



HAL
open science

Contribution à la formalisation de bilans / états de santé multi-niveaux d'un système pour aider à la prise de décision en maintenance : agrégation d'indicateurs par l'intégrale de Choquet

Bouthaina Abichou

► To cite this version:

Bouthaina Abichou. Contribution à la formalisation de bilans / états de santé multi-niveaux d'un système pour aider à la prise de décision en maintenance : agrégation d'indicateurs par l'intégrale de Choquet. Sciences de l'ingénieur [physics]. Université de Lorraine, 2013. Français. NNT: . tel-00839731

HAL Id: tel-00839731

<https://theses.hal.science/tel-00839731>

Submitted on 29 Jun 2013

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.

Table des matières

Introduction Générale	5
Chapitre 1 : Bilan de santé : du contexte BMCI vers sa généralisation scientifique	11
1. Introduction	11
2. BMCI, description et besoins (contexte industriel)	12
2.1. Constat initial de BMCI.....	13
2.2. Objectif général du projet BMCI.....	14
2.3. Les tâches principales du projet.....	15
2.4. Premières investigations scientifiques	19
2.5. Les verrous industriels.....	23
3. La santé des systèmes industriels (contexte scientifique)	25
3.1. Le PHM et le suivi de la santé des systèmes	26
3.2. Concepts généraux autour de la notion de santé	30
3.3. Vers la formalisation d'un bilan de santé pour des systèmes multi-composants	34
4. Conclusion.....	48
Chapitre 2 : Bilan de santé : la détection des anomalies et la fusion des informations	49
1. Introduction	49
2. La détection des anomalies.....	50
2.1. La détection dans un processus de surveillance	50
2.2. Critères de sélection de l'approche de détection	51
2.3. Classifications existantes des approches de détection des anomalies	52
2.4. Comparaison des méthodes de détection	59
3. L'agrégation des indicateurs	60
3.1. Concepts généraux.....	61
3.2. Critères de sélection de l'opérateur d'agrégation	63
3.3. Classification des outils d'agrégation	63
3.4. Intégrale de Choquet.....	67
4. Conclusion.....	74

Chapitre 3 : Bilan de santé : formalisation et proposition d'outils support à son élaboration.....	77
1. Introduction	77
2. Formalisation d'un bilan de santé système.....	78
2.1. Indicateurs et index pour un élément.....	79
2.2. Relation hiérarchique entre les indicateurs	81
3. Une nouvelle approche pour la détection conditionnelle	83
3.1. Concepts théoriques utilisés : histogramme et entropie relative	84
3.2. Notion de dérive d'un indicateur	87
3.3. Approche de détection contextuelle.....	97
4. Identification des coefficients des capacités pour un système multi-niveaux hiérarchiques .	100
4.1. Apprentissage des capacités au niveau composant.....	101
4.2. Inférence des capacités pour les niveaux supérieurs.....	104
5. Conclusion.....	120
Chapitre 4 : Application des contributions sur le banc d'essai BMCI.....	123
1. Introduction	123
2. Présentation du cas d'application	123
2.1. Description du moteur diesel.....	125
2.2. Structure hiérarchique du bilan de santé.....	127
3. Mise en œuvre de la méthode de détection.....	134
3.1. Indicateur de type I sans impact du contexte	135
3.2. Indicateur de type I avec impact du contexte	136
3.3. Indicateur de type II sans impact du contexte	138
4. La construction des indexes de santé	140
4.1. Apprentissage des capacités pour le niveau inférieur	140
4.2. Inférence de la capacité pour le niveau supérieur	145
4.3. Calcul des états de santé.....	149
5. Conclusion.....	151
Conclusion générale et Perspectives	155
Liste des figures	157
Liste des tableaux	161
Bibliographie de l'auteur.....	163
Bibliographie	165

Introduction Générale

Par rapport aux contextes actuels économiques mais aussi législatifs et environnementaux en perpétuelle évolution, les systèmes industriels doivent non seulement satisfaire, sur les objets/services qu'ils fournissent, aux exigences/performances conventionnelles de coût, de productivité... mais aussi à de nouvelles exigences/performances de sécurité, d'efficacité énergétique, de gestion des déchets... [Sénéchal, 2004].

La maîtrise de ces performances/exigences représente ainsi un enjeu majeur pour la survie des entreprises. En effet, la contradiction potentielle entre l'obtention de certaines de ces performances (e.g. conflit entre sécurité et productivité) prône l'adoption d'une vision nécessairement holistique de leur obtention et non plus spécifique à chacune. Ceci est d'autant plus vrai quelque soit la phase du cycle de vie de ces systèmes considérée [Kiritsis, 2008] [Takata, 2004]. En ce sens, le processus de « maintien en condition opérationnelle (MCO) » prend toute son importance comme souligné par le DoD (Département de la Défense Américain) [DoD, 2008], surtout dans ses considérations anticipatives de « déviations des performances » tel que prôné par la communauté PHM (Prognostics and Health Management) [Kumar, 2010]. Cette communauté cherche à favoriser le déploiement d'approches intégrées et proactives de l'évaluation de la performance illustrées dans le domaine maintenance par des stratégies de type conditionnelle et prévisionnelle.

L'architecture d'un système PHM est généralement formalisée en cohérence avec le standard OSA-CBM (Open System Architecture for Condition Based Maintenance) [Lebold, 2001]. OSA-CBM comprend six couches parmi lesquelles celle du « State Detection » qui a pour objectif principal d'analyser l'état de santé actuel de l'équipement surveillé afin de signaler l'apparition d'un fonctionnement non-optimal ou dégradé [Ribot, 2009]. Cette phase requiert par conséquent de disposer au départ d'informations reflétant l'état réel du système en temps réel et de façon continue. En ce sens, un bilan de santé (vecteur d'indicateurs relatifs au système) est l'élément central d'une prise de décision vis-à-vis du système puisqu'il permet à son utilisateur de suivre l'évolution de l'état réel de ce dernier et de réagir à temps (e.g. par une action de conduite ou de maintenance) en cas d'apparition de dégradations des composants ou de perte de la performance attendue.

Cependant, l'évaluation de la santé des systèmes dans la communauté PHM, est généralement abordée relativement à un niveau équipement sans chercher réellement à considérer le système dans sa globalité [Wang, 2001] [Martin, 1994] [Sabnavis, 2004]. Par conséquent les outils proposés support à l'évaluation de la santé sont souvent spécifiques au

type de l'équipement surveillé et ne peuvent pas être exploités en dehors de ce domaine d'application. En outre, même si la performance fonctionnelle est mentionnée dans des travaux de la communauté PHM comme un objectif pour le suivi de la santé, concrètement la plupart des contributions se focalisent surtout sur des aspects dysfonctionnels associés à la dégradation physique des composants [Lee, 2006].

Dans le domaine naval, ces constats sont à la base du projet du pôle MER-PACA nommé BMCI (Bilan de santé pour la Maintenance et la Conduite Intelligentes) ayant pour objet de contribuer à la formalisation de la notion de « bilan de santé » pour la maintenance et la conduite de navires. Ce projet considère des équipements intégrés vis-à-vis de leur fonction, sous-système, système voire sur-système (notion de composants, fonctions, navire et flotte de navires) et s'intéresse aux différents besoins exprimés par de multiples acteurs métiers (marins, mainteneurs, super intendants, armateurs, centres d'expertises...). Ces travaux de thèse s'inscrivent totalement dans ce projet en se structurant autour d'un protocole de recherche partant des besoins et verrous exprimés dans BMCI sur le concept de « bilan de santé », puis généralisant ces besoins et verrous afin de développer une contribution plus globale applicable dans d'autres domaines tout en la validant au final a minima dans le contexte BMCI.

En ce sens, nos recherches se focalisent plus spécifiquement sur le fondement du concept de bilan/état de santé en regard des considérations de surveillance et ce en partant des définitions liées aux notions de santé des systèmes/équipements issues des communautés PHM, CBM (Condition Based Maintenance), CBM⁺ (Condition Based Maintenance Plus) ... [Amin, 2005] [Niu, 2010 a]. Par conséquent, tout en considérant le système dans sa globalité et non pas en prenant chaque composant de façon séparée, une architecture fonctionnelle générique d'élaboration d'un bilan de santé pour un système multi-composants est proposée. Cette architecture est surtout structurée autour de trois fonctions clés pour l'élaboration d'un bilan de santé que sont la détection des anomalies, la normalisation et l'agrégation d'indicateurs. Ces trois fonctions sont le cœur de notre contribution scientifique majeure. Leur intégration dans une architecture d'élaboration d'un bilan de santé amène à combiner la fonction de détection des anomalies et celle de la normalisation. En effet, le degré de déviation des valeurs prises par chaque indicateur par rapport à la référence nominale (entre 0 et 1) est en même temps un moyen d'apporter une sémantique commune à l'ensemble des indicateurs du bilan de santé et un moyen pour positionner ces indicateurs sur une même échelle.

Pour outiller la fonction de normalisation, nous proposons une nouvelle approche de détection des anomalies. En effet, les approches conventionnelles de détection des anomalies imposent la plupart du temps une connaissance des paramètres spécifiques à chaque type d'indicateur [Chandola, 2009] et ne permettent pas ainsi de les appliquer indépendamment du cas étudié. Pour remédier à ce constat, nous proposons une approche

de détection non paramétrique à base de l'entropie relative [Lee, 2001 a] [Rabatel, 2011]. Le contexte opérationnel et/ou environnemental du système affectant les valeurs prises par les indicateurs en mode nominal est intégré dans l'outil de détection proposé. A la suite de la phase de normalisation, l'agrégation des indicateurs est réalisée à différents niveaux d'abstraction du système multi composants. Les outils d'agrégation les plus développés (Dempster Shafer, réseaux bayésiens, réseaux de neurones..) demandent un temps de calcul conséquent par rapport à une agrégation en ligne [Wu, 2002] [Yu, 2007] ou encore fonctionnent en boîte noire [Niu, 2010 b], ce qui ne permet pas une visibilité nécessaire à une structure étagée de bilan de santé. En ce sens, nous utilisons, en support de l'agrégation, l'intégrale de Choquet [Murofushi, 1991] qui est un opérateur d'agrégation simple et explicite permettant de considérer les interactions horizontales et verticales entre les indicateurs impliqués dans le bilan de santé. Notre contribution vis-à-vis de cet opérateur porte plus spécifiquement sur l'identification des capacités (paramètres pour cette intégrale floue). Cette identification est à base d'apprentissage non supervisé pour le niveau le plus bas dans l'abstraction du système surveillé, à savoir celui des composants. Pour les autres niveaux d'abstraction, les paramètres sont déduits par inférence ascendante depuis les capacités des niveaux inférieurs.

Vis-à-vis de ces différentes contributions, l'intérêt majeur de ces travaux réside dans la généralité des solutions apportées aux problèmes initiaux de manière à pouvoir les appliquer indépendamment de la nature des systèmes industriels considérés (hors du champ naval). De plus, l'apport principal de nos contributions porte sur :

- Le fondement du concept de bilan/état de santé dans une vision système se matérialisant par une structuration fonctionnelle plus pertinente que la structure organique des composants,
- La proposition d'une architecture générique et modulaire d'activités pour l'élaboration d'un bilan de santé pour un élément indépendamment de sa nature ou de son niveau d'abstraction dans un système,
- La prise en compte du contexte opérationnel et environnemental dans le suivi de la santé et plus spécifiquement pour la fonction de détection/normalisation des déviations,
- L'apprentissage des capacités pour l'intégrale de Choquet et leur inférence à travers les niveaux hiérarchiques du système ne nécessitant aucune connaissance a priori.

Ces contributions sont développées et justifiées au sein de ce manuscrit à travers quatre chapitres.

Le chapitre 1 décrit globalement le contexte industriel et scientifique de cette thèse. Le projet BMCI est ainsi détaillé en cherchant à mettre en évidence les besoins industriels majeurs et verrous se référant à la problématique du bilan de santé au sein de ce projet. En

réponse à certains de ces verrous, une première investigation scientifique est présentée en proposant des principes clés comme celui de vecteur de santé. Ensuite, la généralisation du concept « bilan de santé » dans une vision plus académique est abordée avec comme objectif, d'une part, de légitimer les définitions, servant de référence à cette thèse, comme les notions de santé, d'indicateur, d'index... et d'autre part, de proposer une approche de caractérisation et d'élaboration d'un bilan de santé pour des systèmes multi-composants. En ce sens, un bilan de santé est défini comme composé d'indicateurs de performance et/ou de dégradation quantifiant la déviation par rapport à la nominalité, et ce en intégrant le contexte opérationnel/environnemental dans lequel évolue le système surveillé. En outre, des indicateurs synthétiques, désignées par indexes, sont aussi nécessaires pour le bilan de santé afin de donner une information plus générale en confrontant un ensemble d'informations hétérogènes sur l'état du système. Pour l'élaboration de ces indicateurs/indexes, une architecture générique d'activités est ensuite investiguée sur la base du standard OSA-CBM. Cette architecture s'applique à tout élément dans la structure hiérarchique du système. Parmi les activités de cette architecture, celles de la détection-normalisation-agrégation des indicateurs sont clairement isolées et détaillées puisqu'elles représentent le cœur de notre contribution dans l'objectif de répondre à la formalisation du bilan de santé.

Le chapitre 2, par rapport à l'identification précédente des 3 activités clés, porte par conséquent sur un état de l'art concernant les outils/méthodes en support de la détection conditionnelle des anomalies et la fusion des indicateurs (l'activité de normalisation étant considérée comme intégrée à la détection). Ces outils/méthodes sont globalement évalués et comparés en regard des besoins émergents de BMCI mais aussi de leur généralisation. En ce sens, dans une première partie, le recensement des outils/méthodes dans la communauté de la détection/surveillance des anomalies permet de mettre en évidence l'adéquation des approches de détection à base de la théorie d'information pour une détection des anomalies générique et robuste. Par contre, il reste à adapter ces techniques pour la prise en compte du contexte opérationnel dans lequel le système surveillé est amené à évoluer. De façon duale, la seconde partie de ce chapitre s'intéresse, toujours sur la base d'une évaluation et comparaison d'outils/méthodes, à choisir l'outil d'agrégation des indicateurs le plus adéquat parmi ceux utilisés dans le domaine de la fusion d'informations. Cette comparaison révèle l'intérêt qu'apporte l'intégrale de Choquet par rapport aux besoins identifiés dans le chapitre 1. Cependant, un tel outil d'agrégation a comme difficulté majeure l'identification de ses paramètres. En ce sens, une approche d'apprentissage des capacités à base d'entropie relative proposée par [Jullien, 2008] est exploitée par une adaptation potentielle à notre contexte d'étude.

Par rapport aux deux constats et orientations scientifiques du deuxième chapitre dégagés à la fois sur la détection conditionnelle et l'agrégation, le chapitre 3 a pour objet de fonder nos contributions. En ce sens, une nouvelle approche de détection conditionnelle des anomalies

basée sur la partition de l'espace contextuel et la mesure de la déviation de l'observation actuelle par rapport à la référence de normalité est justifiée. Cette approche utilise l'entropie relative pour quantifier l'écart entre les distributions des valeurs prises par chaque indicateur. Cette approche a pour intérêt non seulement d'alerter en cas de dérive significative mais aussi de normaliser les indicateurs. Vis-à-vis de l'agrégation, une approche d'apprentissage non supervisée est proposée et justifiée pour répondre à l'identification des paramètres de l'intégrale de Choquet pour les éléments du niveau inférieur de la hiérarchie d'abstraction du système, à savoir celui des composants. Ensuite, une approche originale pour leur inférence d'un niveau d'abstraction à l'autre est développée pour les éléments des niveaux supérieurs dans le système.

Les contributions formalisées au chapitre précédent sont appliquées sur un cas concret lié au projet BMCI dans le chapitre 4 permettant en premier lieu de juger de la faisabilité, par passage à l'échelle, de la mise en application de nos contributions. Cela permet ensuite d'évaluer la valeur ajoutée de nos travaux avec au final une discussion sur les résultats obtenus à la fois quantitativement et qualitativement. Ce cas concret est ciblé sur un « ensemble moteur » au banc d'essai représentatif de modes de fonctionnement d'installation de production d'énergie électrique (diesel/alternateur) ou d'installation de propulsion (diesel de propulsion).

Enfin, ce mémoire se termine par une conclusion et une élaboration d'un ensemble de perspectives principalement basées sur les résultats obtenus lors du déploiement de notre contribution sur le cas industriel et en regard des hypothèses posées.

Chapitre 1 : Bilan de santé : du contexte BMCI vers sa généralisation scientifique

1. Introduction

Face aux besoins, pour les systèmes industriels, d'un maintien en condition opérationnelle proactif afin d'anticiper les déviations de performances, les pratiques de maintenance basées conventionnellement sur une vue centrée défaillance de niveau composant ont évolué vers des pratiques basées sur une vision de performance globale (vision système) construite à partir de la dégradation des composants et de ses mécanismes d'évolution/propagation sur les différents niveaux d'abstraction. Par conséquent cette évolution se fonde principalement sur un suivi de la santé globale du système pour un objectif étendu de prise de décision (e.g. conduite, maintenance, soutien logistique). Le décideur a donc besoin de disposer au départ des informations reflétant l'état réel du système à travers des indicateurs de santé [Vichare, 2007]. L'évaluation du bilan/état de santé constitue ainsi une étape primordiale dans l'évolution des pratiques de maintenance (e.g. vers des maintenances prévisionnelles) comme prôné par la communauté scientifique/industrielle PHM (Prognostics and Health Management). Cette vision PHM est celle déployée à ce jour au sein du projet BMCI, contexte de nos travaux de thèse, pour une flotte de navires militaires ou commerciaux. Ce projet vise à élaborer un bilan de santé global de la flotte pour une meilleure conduite et une maintenance intelligentes.

Par conséquent ce chapitre vise, premièrement, à introduire et décrire le projet BMCI en cherchant principalement à identifier les besoins industriels en regard de la problématique de bilan/état de santé d'un navire en structurant ce dernier autour des niveaux composant, sous-système, système et flotte. Ces besoins nous amène ensuite à positionner une investigation scientifique, résultante d'une réflexion initiale dans BMCI, et fondatrice pour notre thèse puisque proposant les principes de vecteur de santé composé d'indicateurs fonctionnels/dysfonctionnels/environnementaux et de structuration multi-niveaux. La considération de ces principes dans une vision scientifique globale « bilan de santé » et non plus spécifique au domaine applicatif de BMCI nous permet ensuite de positionner la problématique bilan/état de santé en regard de la communauté PHM, en cohérence, de définir les notions de santé, indicateur, index, et bilan de santé pour des systèmes multi-composants. Ces définitions sont prises comme référence dans tous les travaux et contributions réalisés dans cette thèse et mettent en évidence qu'un bilan de santé est

composé d'indicateurs/index de performance, de dégradation et/ou de contexte permettant de quantifier la déviation par rapport à la nominalité. Il se pose alors la question des processus/activités nécessaires pour élaborer ces indicateurs/index. La réponse à cette question fait l'objet de la fin de ce chapitre où sont discutées les notions d'activités du bilan de santé, d'architecture en mettant plus spécifiquement en évidence le rôle majeur joué par les activités de détection conditionnelle, de normalisation et d'agrégation.

2. BMCI, description et besoins (contexte industriel)

Ce travail de thèse se rattache au projet BMCI (Bilan de santé pour la Maintenance et la Conduite Intelligentes) initié par le pôle de compétitivité Mer Paca. Ce projet a pour objectif global de fournir un outil efficace et à faible coût, d'élaboration d'un bilan de santé pour une flotte de navires. Ce bilan de santé doit permettre à ses utilisateurs de faciliter et d'améliorer la prise de décision en termes de conduite et de maintenance navale.

BMCI a débuté en 2009 et s'étend sur quatre années regroupant autour du même objectif onze partenaires dont six industriels et cinq universités :

- Université Nice Sophia Antipolis LEAT : expertise en antennes miniatures,
- Université de Lorraine CRAN Nancy : expertise en maintenance intelligente,
- Université Paul Cézanne LSIS Marseille : expertise en automatique et détection de défauts,
- Ecole Centrale Marseille CMRT : expertise dans le domaine de la thermique et de l'énergétique,
- Université de Franche Comté FEMTO-ST : expertise en système électronique d'interrogation à distance des capteurs SAW – prototypages de capteurs SAW,
- DCNS : activité en maintenance des navires et construction navale militaire,
- PREDICT : activité en conception et développement de solutions de surveillance, diagnostic, pronostic et maintenance à distance des installations industrielles,
- CMR : activité en conception, fabrication et SAV d'appareils de mesures et contrôle principalement dans le secteur naval,
- ECT : activité en développement de boîtiers embarqués et leur environnement à terre pour le suivi des équipements à bord,
- Moteurs BAUDOIN : concepteur et fournisseur de moteurs navals,
- RIVIERA LINES : société d'excursion en mer et affrètement.

Ces partenaires ont un savoir-faire adapté et complémentaire et partagent les tâches du projet, chacun selon son domaine d'expertise. La gestion du projet BMCI est assurée par le

chef de projet DCNS qui coordonne le travail collaboratif entre l'ensemble des partenaires afin de mener à bien les différentes tâches.

Avec un financement de 2,3 M€, ce projet doit apporter de réelles innovations dans le domaine de la conduite et de la maintenance des installations à bord des navires. Le chiffre d'affaire envisageable par rapport aux solutions résultants du projet est d'environ 90 M€ annuels en régime établi et les emplois générés sont estimés de l'ordre de 350 postes [BMCI, 2009].

2.1. Constat initial de BMCI

Dans le contexte d'activités navales caractérisées par des missions de longue durée en mer, il est de plus en plus nécessaire de s'assurer de la disponibilité des navires ainsi que de la sécurité des équipages tout en optimisant la conduite (réduire les dépenses énergétiques, augmenter l'efficacité logistique de la flotte, maîtriser le niveau de pollution...) tout en diminuant les coûts liés au maintien en condition opérationnelle des installations dans le navire. L'enjeu, d'atteindre un niveau de performance globale pour la flotte au cours de son exploitation, est plus contraignant avec une évolution permanente des conditions opérationnelles des missions navales (contraintes économiques du secteur, exigences de qualité de services, changements du contexte environnemental...). En outre, l'état actuel de la surveillance dans le domaine maritime se base sur l'utilisation des AMS (Alert Maintenance Systems) fournissant des alarmes en cas d'occurrence de défaillances. Ces alarmes ne permettent pas d'éviter les situations critiques. De plus, la stratégie de maintenance est essentiellement systématique. Il existe donc un réel besoin de disposer d'un système de surveillance permettant de détecter les dérives par avance afin d'éviter les dégâts causés par les défaillances et d'optimiser les opérations de maintenance systématique jugées trop coûteuses. D'un point de vue de conduite, un tel système de surveillance permettrait aussi de faciliter la validation des bateaux par les autorités d'accréditation.

Il s'avère donc avantageux d'avoir un outil permettant de surveiller l'état réel des navires ainsi que les évolutions du contexte environnant afin d'évaluer leur performance avant toute prise de décision en conduite ou encore en maintenance. Face à ce contexte, le projet BMCI a pour challenge essentiel de développer une boîte à outils logicielle et matérielle (capteurs, boîtiers d'acquisition, algorithmes de traitement de données, moyens de communication et de stockage informatique...) permettant de mettre en œuvre le concept innovant de bilan de santé multi-niveaux d'une flotte de navires. Cette boîte à outil, appelée aussi application BMCI, est un support à l'élaboration du bilan de santé équipement/navire/flotte et devrait fournir aux utilisateurs tous les éléments nécessaires à l'analyse et au pré-diagnostic temps réel de la santé des équipements facilitant ensuite une prise de décision efficace en termes de maintenance et de conduite.

Le résultat attendu est un « prototype » de l'application BMCI permettant de recueillir les signaux sur les différents équipements, de traiter les données brutes, d'élaborer des indicateurs du bilan de santé et de supporter les fonctions de surveillance et de diagnostic depuis l'équipement jusqu'au niveau de la flotte. Ceci est possible grâce à une structuration de l'application permettant l'échange des données et le partage de connaissances entre les différents niveaux d'abstraction de la flotte.

A moyen terme, cette « boîte à outil » pour le bilan de santé devrait être implantée d'une manière progressive sur les installations à coût de maintenance très élevé. Ces installations sont entre autres les auxiliaires de puissance : les moteurs diesel, turbines à gaz... qui peuvent représenter jusqu'à 50% du coût de maintenance d'un navire [BMCI, 2009]. Cependant, il a été envisagé de développer une solution suffisamment ouverte autorisant leur déploiement sur d'autres équipements du navire tels que les moteurs électriques, auxiliaires de production d'énergie, usine frigo... Par conséquent, la solution apportée doit garantir la portabilité de l'ingénierie du projet d'un type de système à un autre (même en dehors du domaine naval pour être plus générique et acquérir, pour les industriels du projet, d'autres marchés potentiels).

2.2. Objectif général du projet BMCI

Par rapport au challenge présenté ci-dessus, l'objectif principal du projet BMCI porte sur la formalisation et l'élaboration d'un bilan de santé multi-niveaux pour une flotte de navires comme outil facilitant l'aide à la conduite et à la maintenance navales. Il est important de préciser ici que le terme flotte est une notion fondamentale dans le projet [Léger, 2012].

Une flotte doit être considérée comme étant un ensemble de systèmes (par exemple navires), de sous-systèmes (par exemple la propulsion ou la production d'électricité) ou d'équipements (moteur diesel, par exemple, un arbre ...) regroupés en fonction de la nature de l'étude visée [Monnin, 2011]. Ainsi, au sein du projet, le terme flotte est utilisé pour désigner un ensemble de navires ou équipements identiques voire similaires répondant au même critère de classement pour un objectif d'investigation. C'est un moyen innovant de capitalisation de la connaissance et de sa réutilisation pour des situations comparables. Le bilan de santé pour une flotte doit regrouper par conséquent les informations liées aux bilans de santé des navires/équipements qui sont rattachés à cette flotte.

Les utilisateurs de ce bilan de santé (armateurs, affréteurs, exploitants, équipages de navires), chacun à leur niveau de responsabilité, auront un accès à un moyen de surveillance et de veille de la santé globale de leur flotte ou de leur navires, à travers un package clé en main composé de capteurs, de systèmes d'acquisition, de traitement et de visualisation.

Ce package doit permettre :

- Au niveau équipement :

D'établir un bilan de santé pour l'équipement (moteur diesel par exemple) proche de celui de la mesure physique captée. Comme exemple, pour une alimentation électrique, le bilan de santé peut fournir un moyen de suivi dans le temps de l'écart des valeurs de la tension et de l'intensité par rapport à celles correspondantes au mode de fonctionnement nominal.

A ce niveau, le bilan de santé est obtenu en transformant les données issues des capteurs en informations concernant la santé des équipements embarqués. Il a pour objectif de proposer aux exploitants et aux mainteneurs un moyen de prétraitement et de pré-diagnostic temps réel afin de réaliser une première analyse sur la santé des équipements.

- Au niveau navire :

D'établir un bilan de santé global pour tout le système (navire) afin d'évaluer son aptitude et sa capacité opérationnelle. Ce bilan est accompagné d'algorithmes de traitements ciblés pour l'exploitation et l'analyse qui sont destinés au personnel à bord.

Les algorithmes de traitements à bord du navire permettent une aide à la conduite et à la maintenance pour l'équipage et le mainteneur du navire via un outil de management du bilan de santé générique.

- Au niveau flotte :

D'établir un bilan de santé flotte pour le suivi simultané des navires et des équipements et s'assurer de leur disponibilité pour la planification des missions au sein de cette flotte. En outre, la centralisation des informations à ce niveau doit répondre au besoin d'effectuer des expertises approfondies par le retour d'expérience et la capitalisation sur les connaissances. En ce sens, les algorithmes de post-traitements débarqués vont faciliter l'aide à la décision à terre pour les parties prenantes différentes d'un niveau flotte (superintendant, armateur, expert, centre support à terre...).

2.3. Les tâches principales du projet

Par rapport aux objectifs précédemment définis, une première phase du projet consiste à identifier les besoins des parties prenantes, les outils de tests, les modes de défaillances à surveiller pour le test, etc. Cette phase prépare aux trois principales tâches autour desquelles le projet BMCI se structure, à savoir :

- Conception et ingénierie de l'application BMCI,
- Réalisation de la maquette (prototype de l'application BMCI),
- Expérimentation du prototype sur un banc d'essai et sur un navire en mission.

L'innovation du projet porte donc sur la conception/réalisation de l'application BMCI et son expérimentation industrielle. Les tâches multidisciplinaires du projet supports à cette innovation ont été définies et affectées aux membres du consortium en accord avec leurs domaines de compétences.

2.3.1. Ingénierie et conception

Cette phase consiste à définir l'architecture matérielle et logicielle de l'application BMCI à développer selon les spécifications définies lors de la première phase du projet à savoir celle de l'identification des besoins. Cette application permet de supporter la procédure d'élaboration d'un bilan de santé, de surveillance et de diagnostic à tous les niveaux d'abstraction dans la flotte (équipement/navire/flotte) (Figure 1).

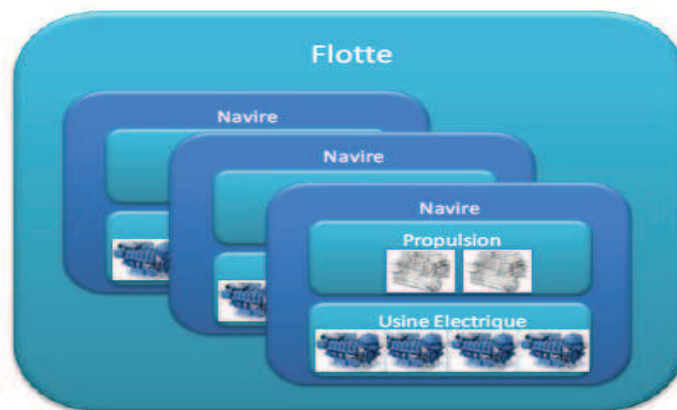


Figure 1 : Principe de structuration multi-niveaux d'abstraction appliqué à la flotte.

Le traitement multi-niveaux ainsi développé consiste à considérer une flotte constituée d'un ensemble de navires, chaque navire étant constitué d'un ensemble d'équipements. A chaque niveau d'abstraction correspond un ensemble de fonctionnalités de traitement des données. En ce sens, un bilan de santé est élaboré à chaque niveau d'abstraction dans la flotte accompagné des traitements pour la surveillance et le diagnostic adéquats aux besoins des utilisateurs à ce niveau d'abstraction. Cette architecture matérielle et logicielle de l'application BMCI permet ainsi des fonctionnalités allant de la caractérisation des équipements à bord dans les salles machines jusqu'à la vision globale de la flotte via l'échange de données et le dialogue entre les niveaux équipement /navire /flotte. La conception de cette application multi-niveaux d'abstractions repose sur trois niveaux de traitements (Figure 2):

- Pré-traitements embarqués (niveau « BMCI Equipement »)
- Post-traitements embarqués (niveau « BMCI Navire »)
- Post-traitements débarqués (niveau « BMCI Flotte »)

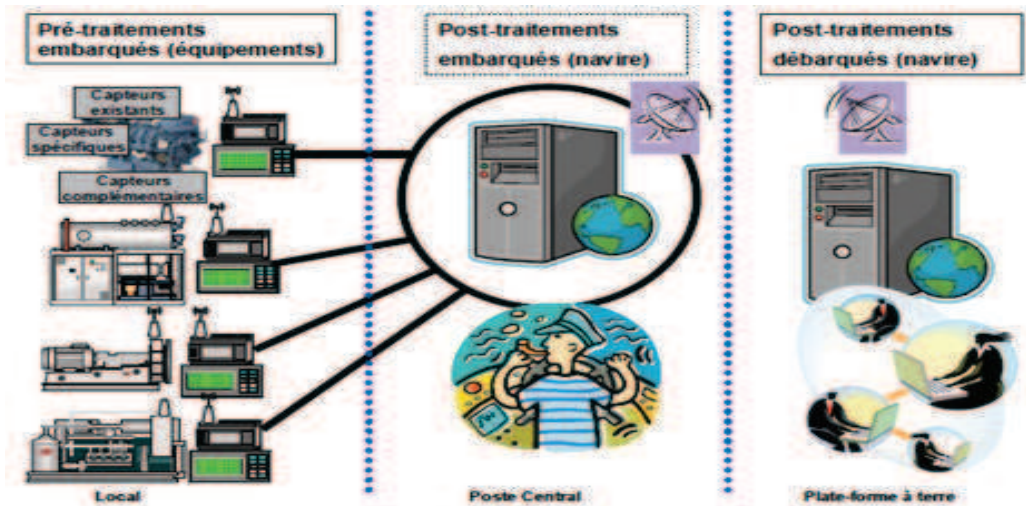


Figure 2 : Architecture logicielle de l'application BMCI.

Dans la Figure 2, les traits discontinus verticaux représentent les séparations entre ce qui est nécessairement au plus près des équipements physiques et qui doit respecter les contraintes industrielles et les conditions environnementales, de ce qui peut être conçu sur des matériels plus conventionnels et qui peut être organisé (à bord, à quai) selon les besoins. Par exemple, une architecture pour un navire de type « transrade » (équipage très réduit et traversée limitée à moins d'une heure) n'est pas la même que celle pour un porte « containers » (équipage double compétences intégrant l'aspect maintenance et missions de plusieurs jours).

Les traitements de base de l'application BMCI se structurent autour de la norme ISO 13374 qui décrit les fonctionnalités standards nécessaires pour la surveillance. BMCI adapte ces fonctionnalités comme le montre la Figure 3.

Cette approche normative est liée à un équipement. Le projet se propose d'utiliser cette structuration au niveau navire et flotte.

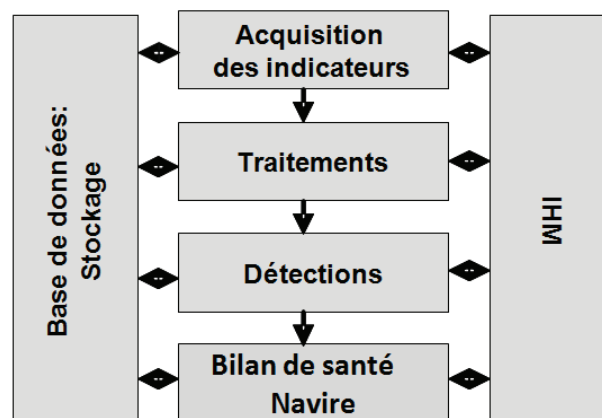


Figure 3 : Fonctionnalités liées au processus de surveillance adaptées au projet [ISO13374, 2002].

2.3.2. Développement du prototype application BMCI

Après une première phase de conception, l'étape suivante consiste à développer les éléments de l'application BMCI aux niveaux : équipement, navire, flotte comme suit :

- Réalisation de deux maquettes de capteurs SAW (capteur de couple, capteurs sans fil en zigbee...) pour l'acquisition des signaux bruts,
- Réalisation des logiciels de base de l'application BMCI,
- Réalisation des bases de données/ontologies supports d'échanges entre les différents niveaux (équipements, navire, flotte),
- Intégration des composants de l'application BMCI,
- Test d'ensemble et validation des différents composants de l'application développée aux différents niveaux (Équipement, navire, flotte).

La dernière sous tâche cherche à valider le bon fonctionnement global des aspects communications, cohérence des informations stockées... Cette phase doit permettre de valider les exigences en lien avec les besoins initiaux.

2.3.3. Expérimentation (banc d'essai puis navire en exploitation)

Les essais d'expérimentation de la boîte à outil (matérielle et logicielle) sont effectués dans un premier temps sur un banc d'essais moteurs. Ensuite, des tests sont effectués en vraie grandeur sur des moteurs dans un navire en exploitation, i.e. l'Adroit. Ainsi le banc moteur doit servir de complément aux essais à la mer pour rejouer, analyser, expertiser des comportements anormaux rencontrés et ainsi confirmer ou infirmer les résultats des essais navire.

Par rapport aux trois tâches précédemment développées, ce travail de thèse se positionne principalement dans la tâche d'ingénierie et conception de la boîte à outil bilan de santé (ou application BMCI). Notre contribution de thèse porte notamment sur la formalisation de la notion de bilan de santé et son élaboration à différents niveaux d'abstraction d'un système. La phase d'évaluation de nos travaux, plus spécifiquement sur les données d'un moteur diesel marin, fait partie intégrante de la phase de test (troisième tâche du projet). L'objectif à moyen terme est d'intégrer les algorithmes développés dans le cadre de ce travail de thèse dans la boîte à outil logicielle comme support à la fonctionnalité d'agrégation des indicateurs dans le bilan de santé.

2.4. Premières investigations scientifiques

A partir du contexte BMCI, des premières investigations scientifiques, auxquelles nous avons participées au début de la thèse, sur le concept de bilan de santé ont été réalisées par le CRAN et son partenaire industriel PREDICT afin de définir un cadre pour ce sujet. Ces premières investigations concernent plus spécifiquement la définition d'une structure hiérarchique d'un bilan de santé pour une flotte de navires et sont synthétisées dans les paragraphes suivants. Ces investigations scientifiques ainsi que le contexte du projet BMCI nous amènent ensuite à isoler plusieurs verrous sur lesquels nous nous positionnons (paragraphe 2.5).

2.4.1. Bilan de santé et état de santé

L'élaboration d'un bilan de santé pour un système doit tenir compte de son aspect fonctionnel (abordé sous l'aspect de sa performance fortement liée à sa finalité), dysfonctionnel (abordé sous l'aspect des modes de dégradation qui altèrent les performances) et opérationnel (abordé sous l'aspect des conditions qui régissent le comportement des performances et des dégradations) [Monnin, 2009].

Le bilan de santé peut être représenté sous la forme d'un vecteur de dimension finie regroupant par juxtaposition les indicateurs élémentaires (Figure 4). La décomposition proposée permet de distinguer les pertes de performance, des modes de dégradation physiques en ayant la possibilité de se référer au contexte opératoire et/ou environnemental dans lequel l'équipement évolue. Les indicateurs opérationnels et environnementaux permettent de distinguer les comportements nominaux relatifs à différentes conditions environnementales ou opérationnelles, des comportements déviants.

Par exemple, pour un moteur, les indicateurs fonctionnels peuvent être de type consommation/rendement énergétique, vitesse de rotation, durée de fonctionnement à différents taux de charge pour la mise en évidence des rendements mécanique et thermique. Parmi les indicateurs dysfonctionnels, les symptômes de dégradation tels que l'état de l'huile ou le niveau de vibration sont des exemples dans le cas d'un moteur d'un navire. Les indicateurs opérationnels décrivent par exemple le mode de fonctionnement, le type de mission, sa durée, le climat ...

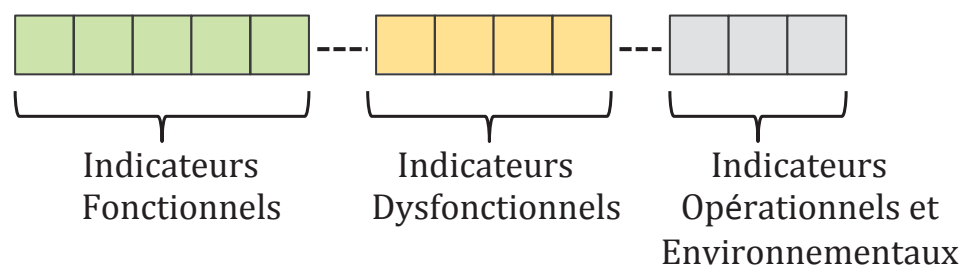


Figure 4 : Représentation d'un vecteur du bilan de santé.

Les indicateurs du vecteur de santé peuvent également être synthétisés/agrégés en une grandeur plus globale : l'état de santé (Figure 5). Cette grandeur permet de faciliter l'interprétation des informations hétérogènes portées par différents indicateurs. L'utilisateur (équipage du navire, mainteneur...) est ainsi alerté rapidement en cas de déviation significative de l'état fonctionnel et/ou dysfonctionnel du système surveillé.

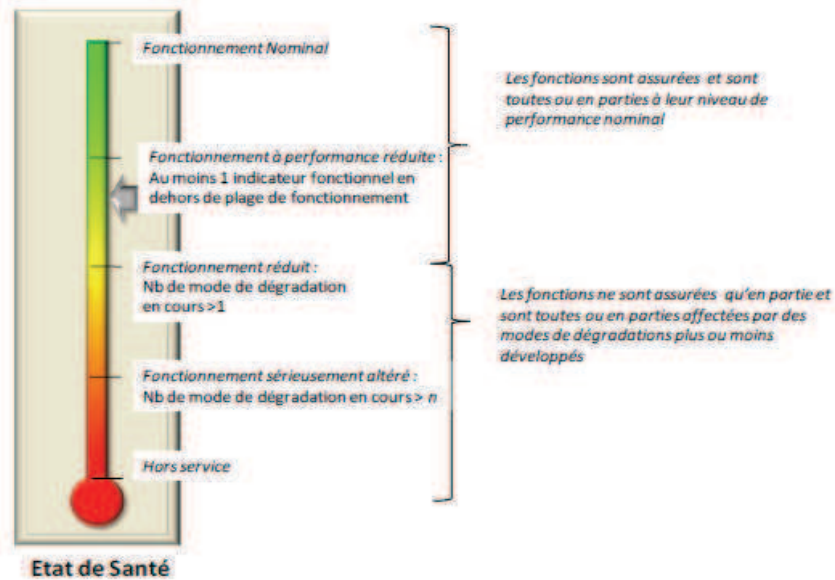


Figure 5 : Représentation de l'état de santé global [Monnin, 2009].

2.4.2. Structure multi-niveaux d'un bilan de santé

Le projet BMCI fonde globalement la structure multi-niveaux sur les trois niveaux déjà expliqués (équipements, navire et flotte) auxquels peut s'ajouter un quatrième niveau intermédiaire entre le niveau équipement et le niveau navire : le niveau installation. Il permet de regrouper sous une même fonction les équipements ayant le même métier au sein du navire. Le principe de construction d'un bilan de santé composé d'un vecteur d'indicateurs rattaché à un état de santé pour un équipement peut être appliqué par conséquent aux différents niveaux d'abstraction de la flotte (Figure 6). En effet, à partir des indicateurs des équipements (ou des mesures physiques captées sur ces équipements), un vecteur d'indicateurs pour les différentes fonctions du navire supportées par ces équipements est construit. En ce sens, un vecteur d'indicateurs global pour le navire peut donc être déterminé de façon hiérarchique en suivant une décomposition fonctionnelle de ce navire [Monnin, 2009].

Concernant le niveau équipement, les indicateurs sont proches des données représentant les grandeurs physiques issues de capteurs. En prenant l'exemple du moteur comme équipement, le bilan de santé peut être constitué des indicateurs suivants :

- Des indicateurs fonctionnels : couple, vitesse, consommation d'huile, rendement

- Des indicateurs dysfonctionnels : état de l'huile, pression différentielle des filtres, températures du moteur...
- Des indicateurs opérationnels/environnementaux : température d'air de la salle moteur, température d'eau de refroidissement moteur, régime moteur...

Pour le niveau installation, il s'agit de regrouper les caractéristiques d'un ensemble d'équipements supportant une fonction particulière pour le navire telle que « la propulsion », dans un bilan de santé donnant une vision globale de cet ensemble (associé à des intervalles de temps plus larges). Des indicateurs pour l'installation de propulsion peuvent être sous forme :

- Des indicateurs fonctionnels : Puissance délivrée, consommation spécifique, durée de fonctionnement...
- Des indicateurs dysfonctionnels : pertes énergétiques, durée totale de dysfonctionnement...
- Des indicateurs opérationnels/environnementaux : durée de la mission, vitesse du bateau, la température de l'eau de mer...

A un niveau supérieur, celui du navire, le bilan de santé a pour objectif de concaténer et consolider les informations issues des indicateurs des niveaux inférieurs. Une vision sur l'ensemble du navire est susceptible d'aider à sa conduite et à identifier la fonction la plus affectée par les éventuels dysfonctionnements. Des indicateurs d'ordre plus global telle que la vitesse réelle, la consommation par mission ou le rendement énergétique global sont alors utilisés comme indicateurs fonctionnels. Les indisponibilités liées à différentes défaillances peuvent être considérées comme exemple d'indicateurs dysfonctionnels. Concernant les indicateurs environnementaux/opérationnels du navire, la nature de la mission ainsi que les conditions climatiques (température, vent...) peuvent être considérées comme des exemples.

Pour le niveau flotte, à partir des indicateurs développés pour chacun des navires/installations/équipements, des indicateurs de niveau flotte sont élaborés en agrégeant ceux du niveau inférieur afin de constituer le bilan de santé de la flotte. Dans la continuité des exemples précédents (basés notamment sur les consommations énergétiques des moteurs), le bilan de santé flotte peut par exemple être constitué par :

- Des indicateurs fonctionnels
 - Efficacité/efficience : Ratio de productivité énergétique = consommation de carburant/Valeur ajoutée brute
 - Performance technique : Cumul heure moteur vs. consommation globale ou cumul énergie vs consommation globale
 - Performance qualité : Poids de capture total (dans le cadre de flotte de pêche par exemple).
- Des indicateurs dysfonctionnels
 - Indisponibilité des navires

- Coûts de maintenance
- Des indicateurs opérationnels/environnementaux
 - Nature de la mission de la flotte
 - Zone de navigation

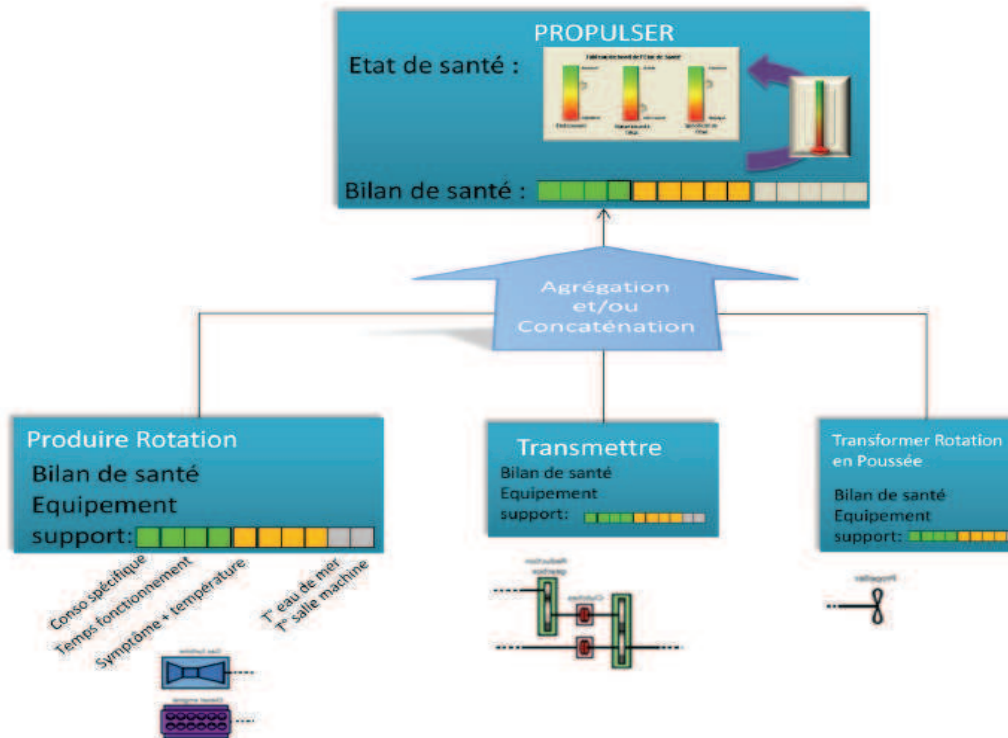


Figure 6 : Exemple de construction du bilan de santé et de détermination de l'état de santé de la fonction "PROPULSER" pour le navire [Monnin, 2009].

2.4.3. Apport de la dimension flotte

La considération de la flotte, d'un point de vue de l'exploitation, apporte une vision globale en agrégeant les différents navires/équipements pour de meilleures décisions stratégiques. Les « individus » qui composent la flotte / sous-flotte peuvent être au besoin, les systèmes eux-mêmes [Bonissone, 2005], [Patrick, 2010], sous-systèmes ou équipements [Umiliacchi, 2011]. Ainsi, la flotte n'est qu'un point de vue abstrait pour considérer un ensemble d'objets dans un but précis (par exemple, une unité de planification de la maintenance), pour un temps donné (par exemple, avant la fin de la mission en cours) [Medina Oliva, 2012].

D'un point de vue de la maintenance, la dimension flotte permet de comparer les navires/équipement entre eux. Elle permet de placer les données ou informations relatives à toutes les unités de la flotte dans le même repère afin de mettre en évidence des résultats pertinents pour la surveillance, le diagnostic, le pronostic ou la décision de maintenance. La dimension flotte fournit par conséquent le moyen de partager des ressources pour les systèmes similaires, de constituer un retour d'expérience, et de capitaliser des connaissances [Monnin, 2011].

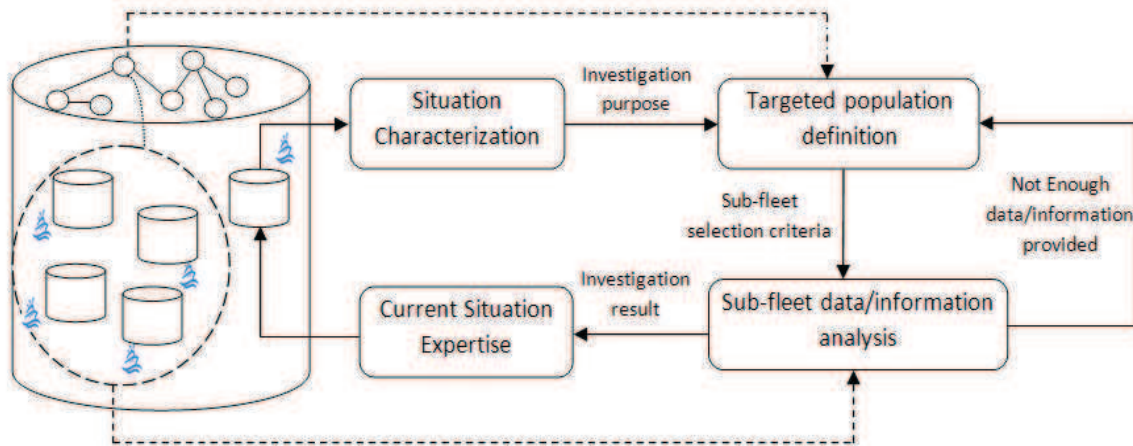


Figure 7 : Apport de la dimension flotte pour l'extraction des connaissances à des fins de diagnostic et d'aide à la décision [Monnin, 2011].

La Figure 7 présente les principales étapes pour réaliser un processus d'extraction de la connaissance/capitalisation au sein de la dimension flotte. Quand un événement ou un comportement anormal d'un équipement dans un navire est détecté (ne pouvant pas être expliqué directement) la situation peut être reportée à une analyse par comparaison à des situations similaires au sein de la flotte afin de compléter la connaissance manquante. Dans ce but, la situation observée doit d'abord être caractérisée pour déterminer l'objectif. La caractérisation de la situation consiste à décrire l'équipement étudié (par exemple le type, l'âge, l'utilisation, l'environnement d'exploitation) ainsi que les données sur lesquelles l'analyse est à effectuer. Selon cette caractérisation, la population cible est définie au sein de la flotte. Les événements similaires sont ainsi examinés dans le but de compléter la connaissance manquante pour la compréhension et la résolution de la situation actuelle. La flotte ciblée peut être affinée ou élargie (démarche itérative) si les résultats potentiels (événements similaires) ne sont pas suffisants ou à l'inverse si les situations similaires sont en trop grand nombre. Ainsi, les résultats des analyses antérieures peuvent être exploités. Dans ce but, les connaissances correspondantes au domaine de la flotte doivent être bien formalisées et structurées afin de faciliter la manipulation de l'aspect multidimensionnel de la flotte et les données hétérogènes entre les bases de données de toutes ses unités. De ces investigations, le sujet de thèse a pris comme référence la notion de bilan de santé multi-niveaux pour les grands systèmes comme les navires afin de la fonder et la généraliser.

2.5. Les verrous industriels

Par rapport aux besoins relatifs à l'élaboration d'un bilan de santé pour la flotte de navires avec une vision multi niveaux d'abstraction tels que souligné dans les paragraphes précédents, les verrous industriels majeurs (liste non exhaustive) mis en évidence dans le projet BMCI sont les suivants :

- Quels indicateurs utiliser pour caractériser les aspects de santé aux différents niveaux depuis les équipements jusqu'à la flotte ?

- Quelle architecture matérielle et logicielle doit être mise en place pour permettre la communication entre les différents niveaux de l'application ?
- Comment exploiter les informations issues du bilan de santé pour des fonctions d'investigation et de diagnostic ?
- Comment exploiter la dimension flotte pour profiter du retour d'expérience et de la connaissance capitalisée ?
- Comment développer des solutions génériques ne nécessitant pas un redéveloppement complet mais uniquement un paramétrage lors du déploiement sur un navire ?

Parmi ces verrous industriels « généraux » et compte tenu du domaine d'implication du CRAN dans le projet, les travaux de thèse se focalisent plus spécifiquement sur les items relatifs à la formalisation et l'élaboration du bilan/état de santé multi-niveaux d'abstraction pour les navires. Ces derniers se traduisent par les verrous suivants :

- Comment définir un bilan de santé de niveau équipement/installation/navire d'une manière générique composé d'indicateurs hétérogènes auxquels il convient de donner du sens, une structure et un rôle vis-à-vis de chaque utilisateur ?
- Comment proposer une vision synthétique et pertinente à partir des informations hétérogènes dans le bilan de santé ?
 - Quelles techniques génériques à mettre en œuvre et permettant la détection de dérives pour les indicateurs utilisés ?
 - Quel(s) opérateur(s) d'agrégation utiliser pour élaborer l'état de santé, synthèse du bilan de santé ?

Ces verrous, initialement spécifiques à BMCI sont tout à fait généralisables par rapport à un enjeu global de bilan de santé. La cible de cette généralisation comporte la variété des systèmes étudiés ainsi que l'hétérogénéité de leurs composants et la divergence entre les besoins en indicateurs entre les différentes parties prenantes. Nos travaux de thèse ont donc pour objet de rationaliser ces verrous afin de travailler sur une généralisation scientifique du bilan de santé. Ceci doit passer nécessairement par approbation au global des notions utilisées pour le suivi de la santé des systèmes industriels. Par cette phase de rationalisation/généralisation, notre thèse est donc bien un travail de recherche (ayant pour objectif de fonder une méthodologie exploitable pour diverses classes d'applications), et non uniquement un travail d'ingénierie devant apporter une réponse précise à un problème particulier.

3. La santé des systèmes industriels (contexte scientifique)

Cette section du chapitre a pour but principal de positionner la problématique industrielle liée au projet BMCI dans un cadre scientifique plus général. Les travaux de recherche sur l'évaluation de la santé des systèmes industriels sont abordés pour en déduire une caractérisation scientifique de la notion de bilan de santé pour des systèmes techniques multi-composants (bien évidemment en cohérence avec les premières investigations synthétisées au paragraphe 2.4 de ce chapitre).

L'évaluation de la santé des systèmes industriels d'une manière générale fait l'objet de notre sujet, et ce pour un double objectif, d'aide à la conduite et à la maintenance. En effet, le suivi de la santé vise non seulement à renseigner sur l'apparition d'une dégradation physique nécessitant une intervention de maintenance, mais aussi à orienter les choix du « pilote » du système par la mise à disposition de diverses informations liées à l'état de fonctionnement dans un contexte opérationnel particulier. Par exemple, une consommation excessive de carburant pour un navire, même si elle ne peut être en lien direct avec un symptôme de dégradation, peut être exploitée afin de réduire la vitesse ou encore changer la planification des missions du navire. Cependant, quelque soit l'objectif visé par l'utilisation des indicateurs de santé (conduite ou maintenance), le processus de détection des dérives pour les valeurs prises par les indicateurs fonctionnels ou de dégradation (dysfonctionnels) reste le même, l'objectif dans les deux cas étant de signaler toute non-nominalité et quantifier la dérive par rapport à une référence (la nominalité). Ceci a pour conséquence de focaliser notre étude sur les approches de maintenance à base du suivi de santé, et les résultats atteints seront valides quelque soit la nature de l'objectif visé (conduite ou maintenance).

En ce sens, le suivi continu de la santé est une composante commune aux approches de maintenance de type conditionnelle/prévisionnelle telle que défendue à travers les « philosophies » du Condition Based Maintenance (CBM) [Thurston, 2001] ou du Prognostics and Health Management (PHM) [Kalgren, 2006] qui visent à utiliser ce moyen afin de mieux gérer les systèmes industriels et à aider la prise de décision en maintenance. Le suivi de la santé de tels systèmes sert à augmenter leur fiabilité globale et à optimiser leurs fonctionnements ainsi que les pratiques de maintenance via une application judicieuse des technologies de surveillance intelligente [Kacprzyński, 2002]. La fonction de surveillance permet de détecter le passage du système dans un fonctionnement anormal. Elle récupère les informations issues des capteurs et les transforme en indicateurs de défaillance à partir d'une référence illustrant le fonctionnement normal ou anormal (dégradé ou inconnu) du système [Ribot, 2009].

3.1. Le PHM et le suivi de la santé des systèmes

Le PHM est une « philosophie » émergente qui étend la notion de maintenance prévisionnelle par l'optimisation de la maintenance et la logistique de soutien afin d'accroître la disponibilité opérationnelle et réduire le « Life-Cycle Cost » tout en augmentant potentiellement la fiabilité et l'espérance de vie de systèmes mécaniques, structurels et électroniques [Kalgren, 2006]. Le PHM s'oriente vers l'évaluation du niveau global de performance des systèmes [Ortiz, 2008] [Amin, 2005]. Le PHM, généralement confondu avec le CBM, est défini comme étant une « philosophie » de gestion de la santé des systèmes techniques basée principalement sur les fonctions de surveillance, diagnostic, pronostic et prise de décision en maintenance. Cette « philosophie » est relativement jeune puisque la communauté PHM a émergé au début des années 2000 pour donner naissance au PHM Society et à l'IEEE PHM.

Par rapport à ces enjeux, la « philosophie » PHM est adoptée dans plusieurs domaines d'application (les avions [Muldoon, 1999] [Byington, 2004], les structures [Ko, 2005], les moteurs [Jaw, 2005]...). Cependant, quelque soit son domaine d'application ou le système concerné, la « philosophie » PHM revendique très clairement un impact en termes d'amélioration de l'aide à la décision en maintenance.

3.1.1. Pour une meilleure aide de décision

Les objectifs derrière la mise en œuvre d'une telle philosophie de maintenance sont d'après [Niu, 2010 a], [Jaloretto Alves, 2009], [Appleby, 2010], [Yu, 2007] [Kumar, 2009] :

- Une meilleure disponibilité et une réduction conséquente des coûts opératoires et de maintenance grâce à une politique de maintenance et de logistique basée sur un suivi de l'état des composants/système en temps réel,
- Une détection plus rapide des dégradations ou des pertes de performance pour une efficacité de fonctionnement accrue,
- Une amélioration de la fiabilité et la sécurité des missions critiques,
- Une diminution des encombrements logistiques et des coûts liés aux investissements sur les matériels supports à la maintenance, leur transport, leur stock, et moins de personnel de maintenance,
- Une réduction de pannes induites par la maintenance,
- La détection et la quantification de la dégradation à long terme pour fournir un moyen de quantifier la vie restante des systèmes. Ceci permet d'identifier ceux qui sont près de leur fin de vie et qui nécessitent des dépenses en immobilisations importantes pour les différencier de ceux qui ne nécessitent qu'une simple remise en état.

Le PHM permet ainsi à la maintenance d'être réalisée de manière plus efficace en intégrant les données PHM (e.g. le RUL, « Remaining Useful Life », issu du pronostic et les résultats du diagnostic), les données de maintenance (ressources et inventaires), les contraintes logistiques (pour les pièces commandées par exemple) et les informations relatives à la mission prévue (Figure 8) [Camci, 2007].

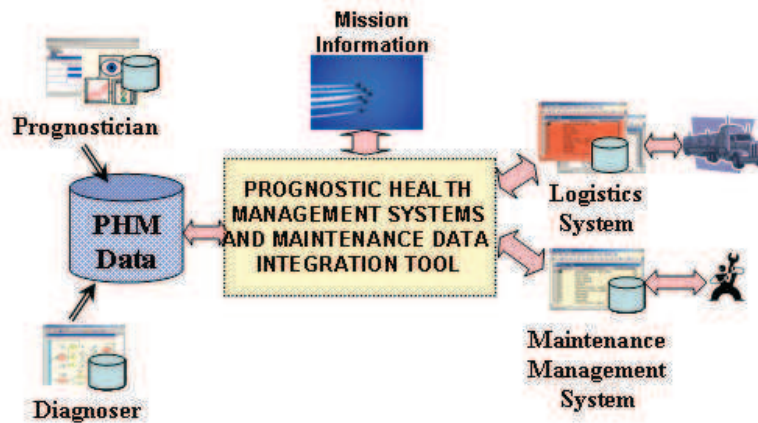


Figure 8 : Rôle du PHM pour une aide à la décision optimale [Camci, 2007].

Le PHM repose par conséquent principalement sur le suivi des variations de l'état du système en analysant les informations pertinentes reflétant des dérives de son comportement. Ces informations sont utilisées pour donner l'alerte sur la détérioration des systèmes ou de leurs composants suffisamment à l'avance de sorte qu'il reste du temps pour prendre la bonne décision de maintenance préventive. Ainsi, en PHM, une succession d'étapes est nécessaire pour l'extraction de telles informations avant la prise de toute décision.

3.1.2. Les étapes clés pour l'aide à la décision

Une procédure d'implantation d'une philosophie de PHM/CBM est considérée comme une suite de phases d'extraction d'informations et de leur traitement, et ce en transformant les données brutes en des informations aidant à la prise de décision [Callan, 2006]. Ces principales phases correspondent aux fonctionnalités suivantes [Thurston, 2001] [Frankle, 2008] :

- Acquisition des données brutes de diverses sources (signaux des capteurs, données d'inspections manuelles...) qui permettent de décrire la santé de l'équipement surveillé,
- Traitement de ces données afin d'en extraire les caractéristiques de santé, nécessaires pour détecter des dérives potentielles,
- Détection des dérives de l'état de santé actuel afin de signaler l'apparition de tout comportement anormal indiquant un mode opératoire non-optimal ou dégradé,
- Analyse des dérives et diagnostic afin de localiser et identifier la cause de l'anomalie détectée,

- Pronostic et étude de l'impact de la situation actuelle sur l'évolution du comportement futur de l'équipement,
- Gestion de la santé en utilisant les informations issues des étapes précédentes pour effectuer une décision concernant l'action à mettre en œuvre.

Cependant, mettre en œuvre les étapes du PHM au niveau du système nécessite aussi la prise en compte des modes de défaillance des différents composants qui sont basés sur des technologies multiples et variées (mécaniques, électriques, électroniques ou logicielles) [Verma, 2010] et dont le comportement peut varier tout au long de leur exploitation [Bonissone, 2005]. La complexité de la structure interne de ces systèmes ainsi que celle de leur environnement et les mécanismes de dégradation qui les régissent [Niu, 2010 b] rend la tâche du suivi de la santé de ces systèmes plus difficile. Un tel suivi constitue l'étape cruciale de surveillance, qui implique l'acquisition de données, leur traitement et l'extraction d'informations utiles représentant le niveau de santé.

3.1.2.1. La surveillance de la santé des systèmes

Au sens du PHM, la fonction de surveillance permet donc de suivre en continu l'évolution de l'équipement à travers des données qualifiables et quantifiables permettant ainsi de prévenir un dysfonctionnement avant qu'il n'arrive [Basseville, 1996]. La surveillance est un dispositif informationnel qui analyse l'état du système et fournit une évaluation de sa condition/santé [Racoceanu, 2006]. Cette condition peut être surveillée à intervalles réguliers, selon un calendrier ou continuellement. Les mesures sur l'état sont analysées et comparées à des mesures similaires concernant un fonctionnement nominal connu afin de détecter le passage du système en fonctionnement anormal.

Néanmoins, l'énorme quantité de données impliquées dans cette fonction conduit généralement le personnel de maintenance à analyser un grand nombre d'informations quotidiennement, ce qui entraîne des limitations dans la considération des alertes [Moore, 2006]. Ce problème est encore plus important pour des systèmes multi-composants.

3.1.2.2. La fusion des informations sur la santé

Avec une telle richesse de l'information, une phase de fusion est souvent intégrée dans une procédure d'implantation PHM [Volponi, 2003]. La fusion des informations vise à améliorer la qualité et l'intelligibilité de l'information [Yu, 2007] à des stades différents du processus PHM. Dans un système de gestion de la santé, il existe trois principaux niveaux où les technologies de fusion jouent un rôle important [Roemer, 2001] [Yu, 2007]. Ces niveaux d'agrégation sont (Figure 9) :

- la fusion au niveau des données brutes : à partir d'un ensemble de données multi-capteurs (Zone 1), la fusion à ce niveau peut servir à valider les signaux et à créer des

caractéristiques de dégradation. Un exemple de fusion de données est la combinaison d'un signal de vitesse et d'un signal de vibration pour obtenir des caractéristiques de vibration synchrones.

- la fusion au niveau des caractéristiques de dégradation : à un niveau supérieur (zone 2), la fusion peut être utilisée pour combiner les caractéristiques résultantes de méthodes de traitements indépendantes de façon intelligente afin d'obtenir la meilleure information possible pour l'analyse et le diagnostic. Par exemple si une caractéristique liée au nombre et à la taille de particules dans l'huile de lubrification d'un roulement est fusionnée avec une caractéristique de vibration, le résultat combiné donne une meilleure information sur la santé du roulement.
- La fusion au niveau de la décision, où les connaissances (zone 3) basées sur l'expérience (taux de défaillance antérieurs par exemple), les prédictions des modèles de pronostic et l'information issue des signaux peuvent être combinées.

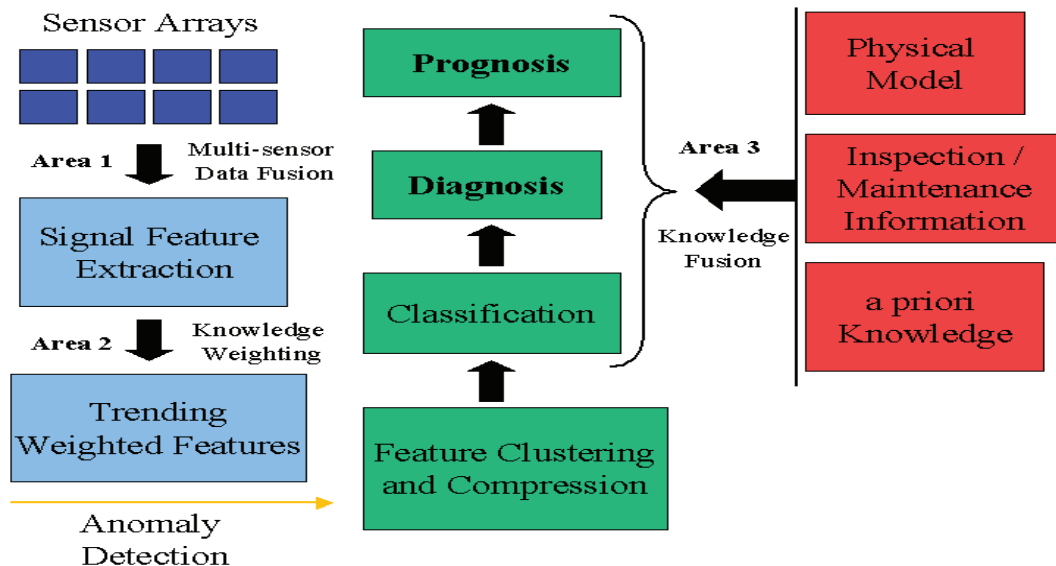


Figure 9 : Les zones d'agrégation dans un processus PHM [Roemer, 2001].

La fusion permet de réduire la quantité d'informations à interpréter à partir des données variées et hétérogènes tout en ayant une vision globale des paramètres indicatifs de la santé du système surveillé. L'information ainsi extraite aide son utilisateur potentiel à prendre des décisions avec un très grand nombre de données à surveiller simultanément.

Les principes de surveillance de la santé et de fusion des informations sont donc fondamentaux pour aborder les verrous scientifiques genèse de cette thèse et portant sur la formalisation d'un bilan de santé pour les systèmes multi-composants.

Nous nous proposons donc de fonder la notion de santé pour de tels systèmes. En ce sens, les concepts généraux autour de la notion de santé ainsi que les moyens pour l'exprimer

vont être dégagés. Cet ensemble d'éléments est primordial pour orienter nos propositions afin de fonder un bilan de santé global et pertinent.

3.2. Concepts généraux autour de la notion de santé

L'évaluation de la santé des systèmes industriels est souvent diffuse dans les travaux de PHM. En effet, la santé est réduite à des alertes de dégradation (résultat de la fonction de détection des anomalies) ou encore renvoyée directement à un résultat de diagnostic sans de réelles explications sur la tendance de déviation par rapport à la normalité [Amin, 2005] [Callan, 2006]. En outre, la description de la santé d'un équipement est souvent liée à un domaine particulier ou encore à une technique relative à la nature des données mesurées, et ce pour le suivi des défauts d'un équipement voire un composant particulier comme dans [Chimentin, 2006], où le suivi de la santé des roulements à billes est dépendant de la technique d'analyse vibratoire utilisée. D'autres travaux comme dans [Abdulmajeed, 2010] proposent des méthodes d'évaluation de la santé spécifiques au cas des câbles des réseaux de distributions électriques. Dans [Zhang, 2007], la santé des machines tournantes est évaluée à travers l'état de dégradation du composant central du système. Ainsi l'évaluation de la santé des systèmes dans la communauté PHM, est généralement abordée relativement à un niveau équipement sans chercher réellement à considérer le système dans sa globalité [Wang, 2001] [Martin, 1994] [Sabnavis, 2004]. Par conséquent les outils proposés support à l'évaluation de la santé sont souvent spécifiques au type de l'équipement surveillé et ne peuvent pas être exploités en dehors de ce domaine d'application. En outre, même si la performance fonctionnelle est mentionnée dans des travaux de la communauté PHM comme un objectif pour le suivi de la santé, concrètement la plupart des contributions se focalisent surtout sur des aspects dysfonctionnels de la dégradation physique des composants [Lee, 2006].

Suite à ces constats, il devient indispensable de rationaliser la notion de bilan de santé de niveau « systèmes industriels » en cherchant à clarifier les concepts généraux en lien avec la notion de bilan de santé tels que le mode de fonctionnement et le concept central de santé.

3.2.1. Modes de fonctionnement

La fonction de surveillance permet de détecter le passage du système du fonctionnement nominal en fonctionnement non nominal. Suivant le mode de fonctionnement sous lequel le système évolue, la mission pour laquelle le système est conçu peut être totalement remplie, partiellement remplie ou non remplie [Rayhane, 2004]. Le mode de fonctionnement non nominal d'un système peut être modélisé par un ensemble de comportements qui ne correspondent pas à une notion bien définie d'un comportement nominal à cause de la présence d'anomalies [Racoceanu, 2006]. Le passage du mode nominal à la défaillance est

souvent évolutif et se caractérise par les étapes suivantes [Villemeur, 1988] [Basseville, 1996] [Iserman, 1997] [Blanke, 2001] :

- Un mode de fonctionnement nominal : le fonctionnement d'un système est dit nominal lorsque ce dernier est capable de fournir les fonctions pour lesquelles il a été conçu et ce avec un certain degré de performance ou de qualité (tel que requis),
- Un mode de fonctionnement dégradé : le mode de fonctionnement dégradé se traduit par une perte de performance d'une des fonctions assurées par un système. Des symptômes de dégradation apparaissent à travers des grandeurs observables qui s'écartent de celles correspondant au fonctionnement nominal. Le passage d'un système d'un mode nominal au mode dégradé est essentiellement lié à l'évolution des mécanismes de dégradation qui peuvent l'affecter,
- Mode de faute : si les performances sont en dessous du seuil d'acceptabilité défini dans les spécifications fonctionnelles pour ce système, la dégradation évolue vers une faute. Cette dernière est définie comme étant une déviation non acceptable d'au moins une propriété caractéristique ou d'un paramètre du système. L'apparition d'une faute peut générer une défaillance,
- Mode de défaillance : suite à l'évolution conjointe ou non de différents mécanismes de dégradations, le système cesse d'accomplir une ou plusieurs de ses fonctions requises.

Les données caractéristiques du comportement nominal peuvent être segmentées en différentes classes, suivant le contexte dans lequel évolue le système, si ce contexte a un impact sur son mode de fonctionnement nominal [Rabatel, 2011]. De telle manière, la prise en compte des critères contextuels pour décrire très précisément le comportement nominal pour un contexte particulier permet d'éviter de confondre comportement nominal contextuel avec un comportement en présence d'anomalie(s).

3.2.2. La santé des systèmes techniques

[Byington, 2004] indique que l'évaluation de la santé est la représentation du niveau réel des déviations dans les systèmes surveillés. [Li, 2007] désigne par la santé d'un système/équipement sa capacité à effectuer la fonction définie et attendue et peut donc être considéré comme le degré d'une certaine performance requise. [Shin, 2009] souligne que la performance d'un produit/composant est évaluée tout au long de sa période d'utilisation par rapport aux spécifications de conception en caractérisant la dégradation de cette performance dans le temps par sa criticité, sa gravité et son anormalité. [Kumar, 2010] renvoie également au fait que le contexte devrait être intégré dans les paramètres de suivi de la santé.

A partir de ces quelques définitions, nous posons comme définition de la santé dans le contexte de nos travaux la définition suivante :

La santé d'un équipement ou d'un système technique peut être définie par : 1) sa performance fonctionnelle en lien étroit avec la finalité du système, 2) les paramètres caractéristiques de ses dégradations, 3) la prise en compte du contexte environnemental et d'usage dans lequel évolue ce système.

Une telle définition de la santé amène à répondre aux deux questions suivantes : 1) comment peut-on représenter la santé, c.à.d. quels sont les moyens de supporter cette information ? 2) comment de telles informations sont elles structurées pour le cas des systèmes multi composants ?

En ce sens, nous introduisons la notion de bilan de santé, comme support de l'information concernant la santé d'un système, et qui est constituée d'indicateurs et d'index de santé (pour représenter des informations multi-niveaux).

3.2.2.1. Indicateur de santé

Un indicateur est défini dans [Brachet, 2000] comme une mesure objective qui reflète un résultat ou un état. Il permet de suivre les changements et les évolutions d'une information quelconque représentant ce résultat/état. En lien avec la notion de santé, pour [Ahren, 2005], un indicateur est exprimé en un seul nombre, délivrant une information sur l'état physique d'un système ou encore reflétant les objectifs et les moyens pour un processus. Il est issu de la transformation des données relatives aux caractéristiques à évaluer afin d'en extraire une quantification facilement interprétable par l'utilisateur [Vichare, 2006]. L'indicateur est alors le principal moyen pour exprimer l'information sur la santé des systèmes techniques. Un indicateur permet de détecter un comportement anormal pendant le fonctionnement d'un système et à mesurer l'étendue de la déviation d'une ou plusieurs caractéristiques d'un fonctionnement nominal [Kumar, 2009].

Plus précisément, un indicateur de défaillance est une quantité significative et pertinente à partir de laquelle il est possible de détecter une défaillance ou une dégradation [Ribot, 2009]. C'est une mesure utilisée afin de quantifier un ou plusieurs symptômes ou caractéristiques de dégradation à partir des données mesurées pour un système variant dans le temps. Ces indicateurs forment ainsi une grandeur fondamentale pour le diagnostic.

Par rapport au contexte BMCI, les indicateurs de performance correspondent aux indicateurs fonctionnels (rendement énergétique du moteur diesel, vitesse du navire, consommation en fuel par exemple). Les indicateurs de dégradation correspondent aux indicateurs dysfonctionnels définis dans le projet (analyse d'huile, niveau vibratoire du moteur diesel par exemple).

Compte tenu de notre contexte et en cohérence avec les définitions précédentes, nous retenons pour nos travaux la définition suivante d'un indicateur d'un bilan de santé :

Un indicateur est une grandeur reflétant une caractéristique de la santé, qu'elle soit de performance ou de dégradation. Un indicateur est issu d'un processus de traitement des données brutes. Il permet de quantifier la déviation par rapport au mode de fonctionnement nominal. Un indicateur porte une information unique sur une partie limitée et particulière de la santé.

Finalement, pour intégrer l'ensemble des dimensions, dans une vision systémique, le bilan de santé intègre des indicateurs de performance, de dégradation et du contexte du système. Les informations représentatives du fonctionnement et de la santé propre de l'élément sont portées par les indicateurs de performance et de dégradation en regard des indicateurs de contexte.

Ainsi, en complément de ces indicateurs élémentaires, un bilan de santé doit contenir aussi une information synthétique de la situation globale du système surveillé. Le support d'une telle information est l'index de santé.

3.2.2.2. Index de santé

Dans [Kalgren, 2006], un index de santé est défini comme étant une variable continue qui varie entre 1 et 0 indiquant le changement de l'état d'un composant à partir d'un état neuf et pleinement opérable jusqu'à une défaillance fonctionnelle.

Les systèmes techniques sont constitués d'un grand nombre de composants présentant des caractéristiques et de multiples modes de dégradation. Il est difficile donc de formuler un seul indicateur pour détecter tous les modes de défaillance d'un composant donné [Bechhoefer, 2007]. Ceci suggère d'apporter à tous les indicateurs hétérogènes une échelle commune pour la santé afin d'améliorer leur interprétation. Ainsi, un index de santé est un outil très utile pour combiner tous ces facteurs de santé. Il permet d'associer des informations variées sur l'état d'un système pour donner une valeur numérique unique comme une information comparative de l'état général dans le temps [Hughes, 2003]. Des exemples d'index composites pour l'évaluation de la santé sont plus explicites dans [Naderian, 2008], [Young, 2010] où un index qui reflète les différents aspects de santé est explicitement calculé pour les systèmes étudiés.

Un bilan de santé est ainsi composé d'indicateurs auquel un index de santé est adjoint, permettant le suivi en continu, de l'état de dégradation et la performance par rapport à une mission ou un contexte environnant.

Cet index de santé est utilisé pour aider à la décision en maintenance à travers les valeurs qu'il prend et les tendances selon lesquelles ses valeurs évoluent (croissance ou décroissance). Ainsi, le décideur, suivant le contexte dans lequel il se situe en termes de charges de travail (par exemple planning de maintenance chargé ou allégé) pourra décider ou non « d'explorer » les causes derrière le changement des valeurs prises par l'index de santé pour aller jusqu'au déclenchement d'une opération. Cependant, l'interprétation de la valeur que l'on obtient pour un index de santé devra être adaptée à la fonction du système étudiée.

Cette adaptation passe par :

- Un apprentissage des paramètres du traitement conduisant à l'élaboration de l'index de santé,
- Une phase d'appropriation, par les opérateurs utilisant cet index de santé, des valeurs prises et de la nature de leur évolution.

Au final, un bilan de santé permet de refléter l'information issue de données brutes mesurées sur les composants du système qui sont en interaction. En outre, la vision globale de la santé du système multi composants requiert une structuration de tels bilans à tous les niveaux du système afin de remonter l'information pertinente à l'utilisateur concerné par un niveau d'abstraction bien déterminé. Cela nécessite de poser quelques caractéristiques d'une « vision système ».

3.3. Vers la formalisation d'un bilan de santé pour des systèmes multi-composants

Un système industriel est appréhendé dans [Von Bertalanffy, 1968] comme un ensemble d'éléments en interrelations entre eux et avec l'environnement. [Aven, 1999] met l'accent sur la finalité (fonction à réaliser) pour laquelle cet ensemble de composants (éléments) est en interaction locale et/ou simultanée. De la même manière, et d'un point de vue de l'ingénierie système [AFIS, 2009], un système a une finalité précise qui répond aux attentes de natures opérationnelles, techniques, économiques... et qui se concrétise à travers des produits ou services réalisés. En outre, le système est en interaction avec l'environnement dans lequel il est placé. Le système échange avec son environnement à travers la transformation des flux entrants (de matière, d'énergie ou d'information) en des flux sortants (les résultats attendus, produits non requis) (Figure 10).

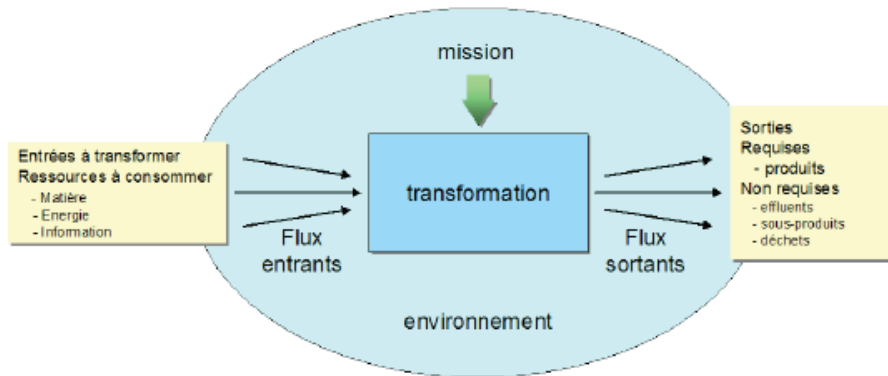


Figure 10 : Le système transforme les flux pour répondre à sa finalité [AFIS, 2009].

Un système multi composants peut être modélisé alors comme un ensemble d'éléments qui sont eux-mêmes en interaction entre eux et avec l'environnement afin de répondre à la finalité requise. Ces éléments évoluent au cours du temps subissant ainsi des changements comme par exemple le vieillissement. Ils présentent alors différents modes et états de fonctionnement qu'il faut maîtriser au sein d'une vision holistique du système surveillé car la perte de performance d'un système provient généralement du mauvais fonctionnement de plusieurs de ses éléments et de leurs interactions [Kumar, 2010].

3.3.1. Structure organique d'un système multi-composants

Sur l'aspect structurel d'un système, deux points de vue complémentaires sont adoptés [AFIS, 2009] :

- le système est vu de l'extérieur comme « une boîte noire » plongée dans un environnement avec lequel il interagit pour lui fournir les services qui répondent à sa finalité, sous forme de transformation de flux entrants en flux sortants.
- le système est vu de l'intérieur comme « une boîte blanche » composé d'éléments en interaction entre eux, la résultante de ces interactions fournissant les services attendus par l'environnement.

En considérant le système de l'intérieur, deux types de modélisation sont communément utilisés : la modélisation par décomposition arborescente (Figure 11-a) et la modélisation par composition/architecture (Figure 11-b). La modélisation par décomposition arborescente (Figure 11-a) considère un système d'un point de vue de niveaux d'abstraction différents qui sont plus adaptés à des prises de décision effectuées également à différents niveaux. Le niveau système représente le niveau d'abstraction le plus haut. Il est décomposé en sous-processus supportés par des sous-systèmes et assurant chacun leur propre fonction. De manière récursive, chaque processus de niveau sous-système peut également être décomposé en un ensemble de processus de niveau inférieur. Cette décomposition est possible jusqu'aux processus de niveaux élémentaires supportés par des composants

physiques du système. La décomposition arborescente d'un système en composants est fonction du degré d'abstraction voulu. Il suffit d'ajouter un niveau d'abstraction dans cette décomposition système/composant pour développer un degré d'abstraction supérieur.

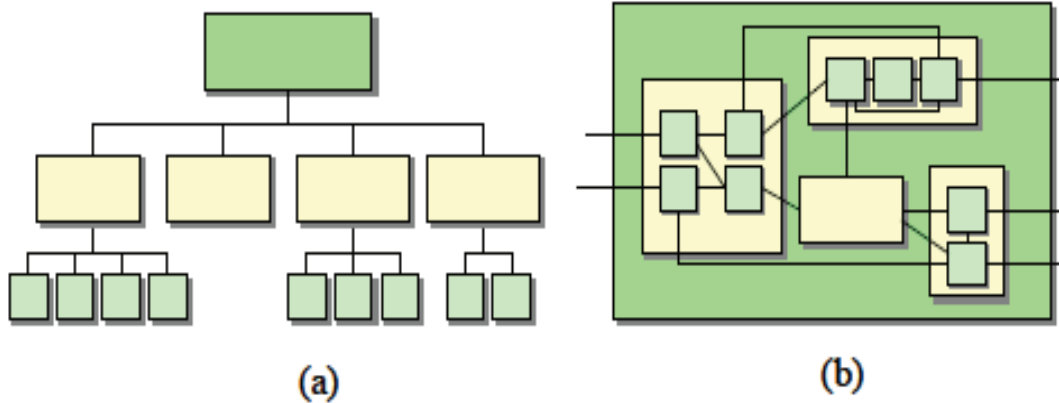


Figure 11 : Aspects structurels internes du système, (a) Décomposition arborescente et (b) Composition ou architecture, [AFIS, 2009].

La modélisation de type composition ou architecture (Figure 11-b) permet de visualiser le système comme étant une succession de processus de niveau sous-système consommant des flux provenant de processus amont et créant d'autres flux consommés par des processus aval. Cette modélisation permet de montrer la chaîne d'interaction créée par les flux échangés. La transparence jusqu'aux composants aboutit au réseau de tous les flux dans la succession de tous les processus élémentaires qui composent le système.

Les deux structures sont importantes pour supporter le bilan de santé. En fait, le premier type de structure permet à l'utilisateur d'avoir une vision d'une grande flexibilité sur un ensemble d'éléments groupés sous une même fonction. La deuxième structure montre les interactions possibles qui peuvent exister entre les composants de fonctions différentes à travers les flux (de matière, d'énergie...) échangés. Cette structure organique des systèmes multi-composants est celle adoptée dans nos travaux pour un bilan de santé multi niveaux hiérarchiques. Afin de fonder un bilan de santé pour un système multi composants, nous proposons donc :

- D'identifier une structure simple et modulaire d'un bilan de santé applicable à tout système multi-composants indépendamment du domaine d'application concerné, et ce en adoptant une décomposition arborescente (Figure 11-a). Une telle forme du bilan donne accès à des informations détaillées ou synthétiques.
- De déterminer la nature des indicateurs d'un bilan de santé, et ce en se basant sur les informations issues de l'architecture interne du système (Figure 11-b).
- D'explicitier les activités menant aux informations délivrées par un tel bilan de santé.

Sur cette base de structure générique, les éléments constitutifs d'un bilan de santé doivent pouvoir être adaptés à tous types de systèmes ou encore à chaque niveau d'abstraction.

3.3.2. Structure hiérarchique d'un bilan de santé

Afin de conserver une vision globale, il est nécessaire d'avoir une représentation générique d'un bilan de santé tout en indiquant les paramètres d'entrée/sortie pour l'élaboration d'un tel bilan.

La détermination de ces indicateurs nécessaires au suivi de la santé a été abordée dans les travaux de ([Xie, 2004], [Appleby, 2010], [Cheng, 2010]...) qui regroupent des techniques de mesures ou capteurs pour divers domaines d'applications de surveillance de la santé. L'approche présentée dans [Kumar, 2009] se base sur une étude AMDEC du système surveillé. Une approche similaire et plus développée pour l'identification des données à surveiller est présentée dans [Kacprzyński, 2002] en se basant sur une analyse fonctionnelle du système à surveiller avant la mise en œuvre de l'AMDEC. Le diagramme fonctionnel contient une représentation de la façon dont les composants, sous-systèmes et systèmes interagissent les uns avec les autres présentant une carte hiérarchique de flux physiques d'énergie (couple, courant, pression) qui sert de base pour la conception des paramètres de suivi de la santé. En effet, ces flux transmettent les symptômes et les effets des modes de défaillance entre les composants et captent leurs évolutions dans le temps. En parallèle, un tableau AMDEC est créé qui contient les symptômes des modes de dégradation, ainsi que les capteurs et les technologies de diagnostic / pronostic correspondants.

A partir de ces approches de détermination des paramètres à surveiller, mais qui restent cependant basés sur une vision « à plat » du système, nous proposons de baser la structure hiérarchique de notre bilan de santé système sur la modélisation fonctionnelle de ce système. Ainsi, un bilan de santé modulaire est associé à chaque fonction de structure fonctionnelle du système, jusqu'au niveau composant (le plus élémentaire) (Figure 12).

De cette façon, le suivi de la santé est possible à partir de plusieurs niveaux selon les besoins de l'équipe de la maintenance ou la conduite. Cette structure permet de regrouper, sous un même bilan de santé d'une fonction, les sous fonctions reliées par les flux. Au niveau du système, le vecteur d'indicateurs comprend principalement des informations sur son fonctionnement global (avec des indicateurs essentiellement de performance) qui résume en quelque sorte la santé de tous ses éléments. Cependant, le vecteur des indicateurs pour les éléments à faible niveau hiérarchique portent surtout des informations reflétant l'état physique et qui sont orientées maintenance (principalement pour des analyses ciblées). Cette structure forme donc une représentation hiérarchique et détaillée du bilan de santé pour un système multi-composants.

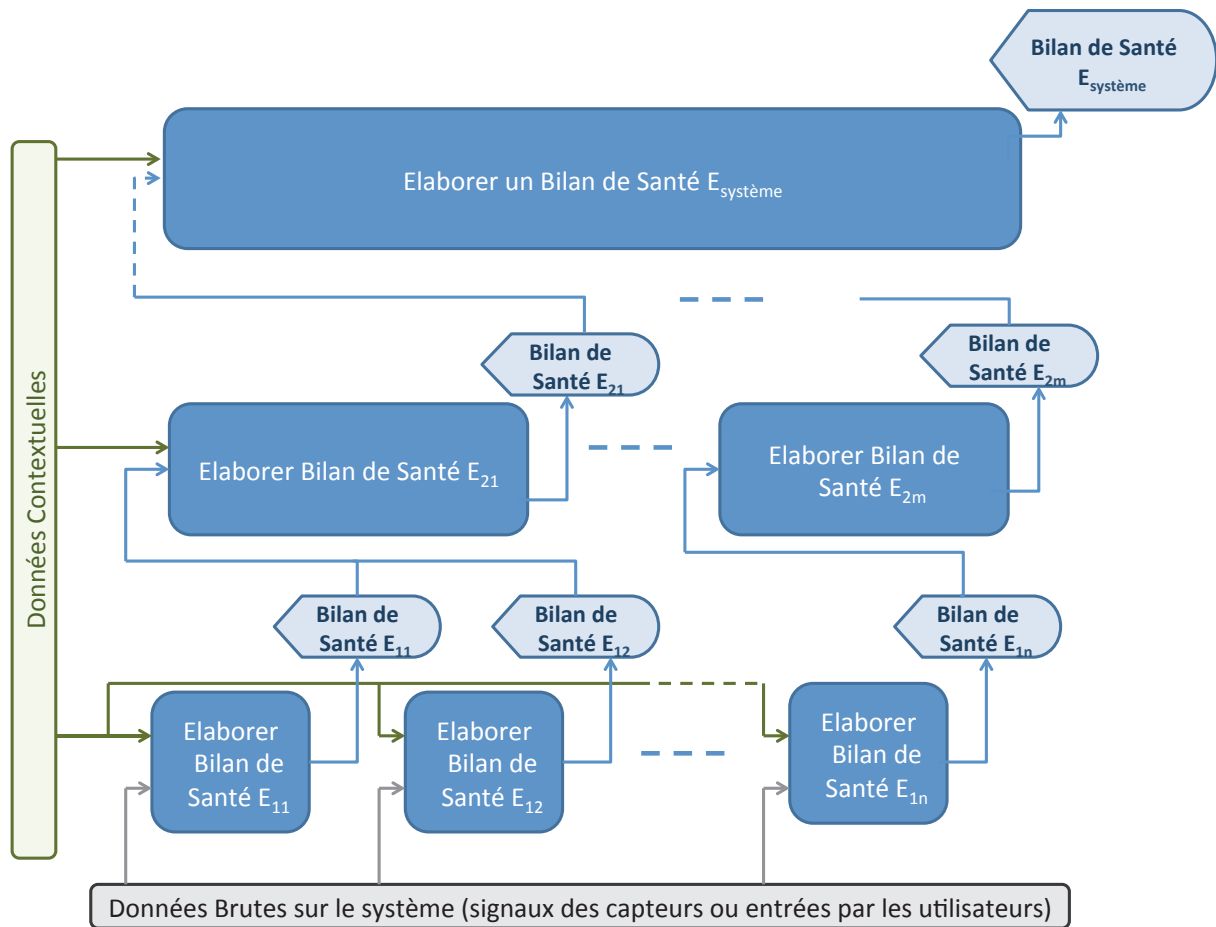


Figure 12 : Structure hiérarchique d'un bilan de santé

En accord avec la définition de la notion de santé construite dans le paragraphe 3.2.2., le bilan de santé doit comprendre les informations liées à la fois à la performance de la fonction via les propriétés de ses flux d'entrée/sortie et aux caractéristiques de dégradation physique liée à son support, à savoir le(s) composant(s) tout en situant ces informations par rapport au contexte de fonctionnement du système.

Dans nos travaux, nous avons décidé, en nous basant sur les contributions de [Léger, 1999], [Muller, 2005], [Cocheteux, 2010] de constituer les indicateurs de performances à partir des propriétés liées aux flux d'entrée et de sortie pour une fonction supportée par un ou plusieurs composants. En effet, la performance de la fonction est décrite par les propriétés des flux de sortie ou des flux d'entrée/sortie pour toute fonction décrite « transformant les flux entrants en flux sortants », et supportée par un équipement. Les indicateurs de dégradation sont liés, par contre, au support de la fonction, à savoir un composant ou un ensemble de composants (ex. sous-système, système). Les indicateurs de dégradation sont de deux types. Ils peuvent être associés soit à des grandeurs physiques du système surveillé (e.g. allongement de la courroie, longueur d'une fissure...) soit à des symptômes induits par les modes de dégradation (ex. vibrations, échauffement d'un moteur). Cependant, il est important de souligner que ces deux types d'indicateurs (de performances ou de

dégradation) sont en interaction puisque les dégradations peuvent être transmises via les flux échangés entre les composants comme le montre l'équation suivante [Léger, 1999] :

$$\text{Etat du Flux Entrant} \wedge \text{Dégradation du Support} \rightarrow \text{Etat du Flux Sortant}$$

Où, \wedge est l'opérateur logique de conjonction ET.

Cette équation exprime la causalité dysfonctionnelle entre la dégradation et la performance mais également l'impact des flux entrants. En effet, le flux entrant peut directement provoquer l'apparition ou modifier l'évolution de modes de dégradations du composant. Ainsi les propriétés du flux d'entrée doivent aussi être reflétées par les indicateurs du bilan de santé. Enfin, les indicateurs contextuels permettent d'évaluer les indicateurs de performance ou de dégradation sans confondre les déviations anormales de comportement avec les variations normales causées par le changement des conditions opérationnelles ou environnementales.

Afin de construire de telles informations, et pour tenir compte du contexte opérationnel et environnemental qui peut influencer la santé d'un composant, les données en entrée pour l'élaboration d'un bilan de santé doivent a minima se structurer autour des catégories suivantes (Figure 13):

- Des données brutes (issues de capteurs ou de bases de données, des observations notées par l'opérateur,...) et qui sont directement liées à l'objet sujet du bilan de santé. Ces données peuvent concerner les propriétés de ses flux énergétiques d'entrée/sortie (vitesse de rotation, courant électrique, débit du fluide...) de la fonction ou encore des mesures relatives aux symptômes de dégradations (vibration, débris d'huile, bruits, température, longueur de fissure...) de son support.
- Des données traitées (indicateurs fonctionnels ou dysfonctionnels) relatives, soit à un élément d'un niveau hiérarchique inférieur, soit à un élément en amont sur le même niveau hiérarchique.
- Données contextuelles : ces données fournissent les informations nécessaires à l'interprétation de l'état de santé d'un objet en fonction de l'état du milieu (les paramètres environnementaux) dans lequel il évolue ou encore en tenant en compte sa mission et son mode opérationnel/utilisation tel que préconisé dans [Muldoon, 1999].

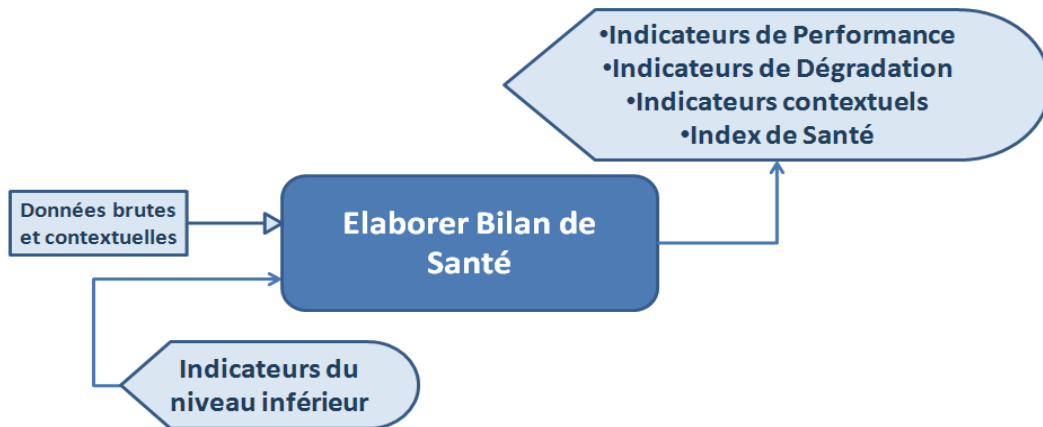


Figure 13 : Les éléments d'entrée/sortie pour l'activité globale d'élaboration d'un bilan de santé.

Sur la base de cette considération du bilan de santé, il est nécessaire d'aborder de manière générique, les fonctions qui permettent d'élaborer un tel bilan avec toutes ses composantes.

3.3.3. Architectures fonctionnelles à base de suivi de la santé

Pour définir l'architecture fonctionnelle, il est nécessaire de passer d'une vision de type «boîte noire» de l'activité d'élaboration d'un bilan de santé à une vision «boîte blanche » des fonctions qui la composent et permettre ainsi d'explicitier les flux entre les données d'entrée et les indicateurs de sortie d'un bilan de santé. Nous proposons ainsi de décrire un ensemble d'architectures fonctionnelles des systèmes à base d'évaluation de la santé, afin d'en extraire les fonctions principales contributrices aux informations exigées par notre bilan de santé. Cet ensemble ne se veut absolument pas exhaustif, mais se rapprochant le plus de la « philosophie » PHM défendue dans nos travaux. Les architectures fonctionnelles qui nous semblent les plus importantes au regard des objectifs posés sont : le standard OSA-CBM [Disenzo, 2001], le Système Intégré de Maintenance Prévisionnelle (SIMP) [Léger, 1999] [Muller, 2005] et le système IVHM (Integrated Vehicle Health Management) [Benedettini, 2009]. Une vision complémentaire sur les architectures de fusion pour le PHM [Roemer, 2001] est aussi considérée pour en déduire la manière d'intégrer la fonction de fusion, nécessaire pour l'élaboration de l'index de santé, dans l'activité mère d'élaboration d'un bilan de santé.

3.3.3.1. OSA CBM

L'architecture d'un système PHM a été formalisée à l'aide du standard OSA-CBM (Open System Architecture for Condition Based Maintenance) [Lebold, 2001]. Le but est de permettre le développement d'une architecture efficace et ouverte qui favorise le déploiement rapide, rentable des technologies connues d'acquisition, de traitement, de suivi, de diagnostic et de pronostic [Jaloretto Alves, 2009].

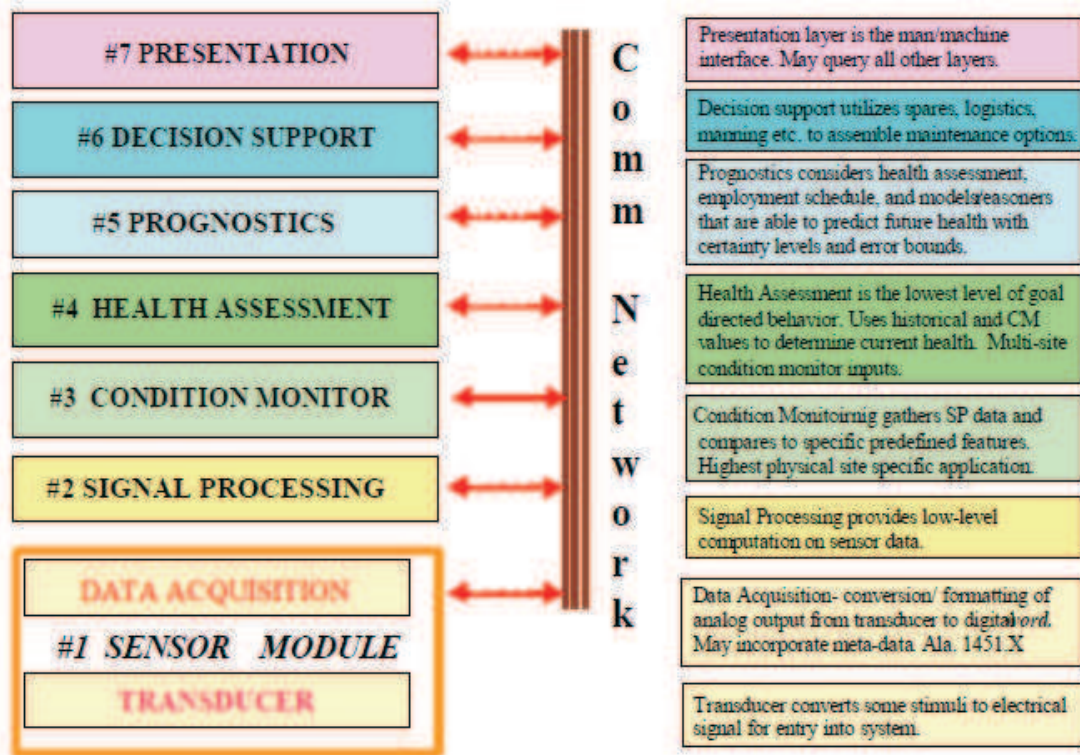


Figure 14 : Architecture du standard OSA-CBM [Discenzo, 2001].

OSA-CBM comprend six couches (Figure 14) correspondantes aux fonctionnalités citées ci-dessus, parmi lesquelles celle du « Condition monitoring ». A ce niveau de traitement, les informations qui décrivent la santé actuelle de l'équipement surveillé sont analysées afin de signaler l'apparition de tout comportement anormal indiquant un mode opératoire non-optimal ou dégradé. Cette phase requiert les informations reflétant l'état réel du système en temps réel et en continu issues des deux couches précédentes à savoir celles de l'acquisition puis du traitement des données. L'information, décrivant le degré de déviation de la santé par rapport au comportement nominal, alimente le processus du « Health Assessment » équivalent à celui du diagnostic. La notion de suivi de santé se rapporte donc à la couche du « condition monitoring » avec les fonctions en amont qui permettent de lui fournir les données nécessaires.

Ce standard a servi de base à la norme ISO 13374 pour la surveillance autour duquel le projet BMCI s'est structuré (Figure 3).

3.3.3.2. SIMP

Le système SIMP a été proposé au sein du CRAN, initié dans les travaux de [Léger, 1999] et formalisé par [Muller, 2005]. Ils proposent un cadre de modélisation du système de maintenance prévisionnelle en l'intégrant avec le système d'information de l'entreprise. Ainsi, le SIMP est un système vis-à-vis de l'ensemble des processus qu'il renferme et qui sont en interaction entre eux et avec l'environnement pour rendre à son environnement les

services correspondants à sa finalité. Il est aussi intégré puisqu'il est voulu coopérant avec les autres processus de l'entreprise: soutien logistique, production... [Cocheteux, 2010]. Pour un objectif de maintenance conditionnelle exécutée en suivant les prévisions extrapolées de l'analyse et de l'évaluation de paramètres significatifs de la dégradation du bien, le SIMP est composé de trois processus : "Processus de Surveillance", "Processus de Pronostic" et "Processus d'Aide à la Décision" intégrés selon le modèle de la Figure 15.

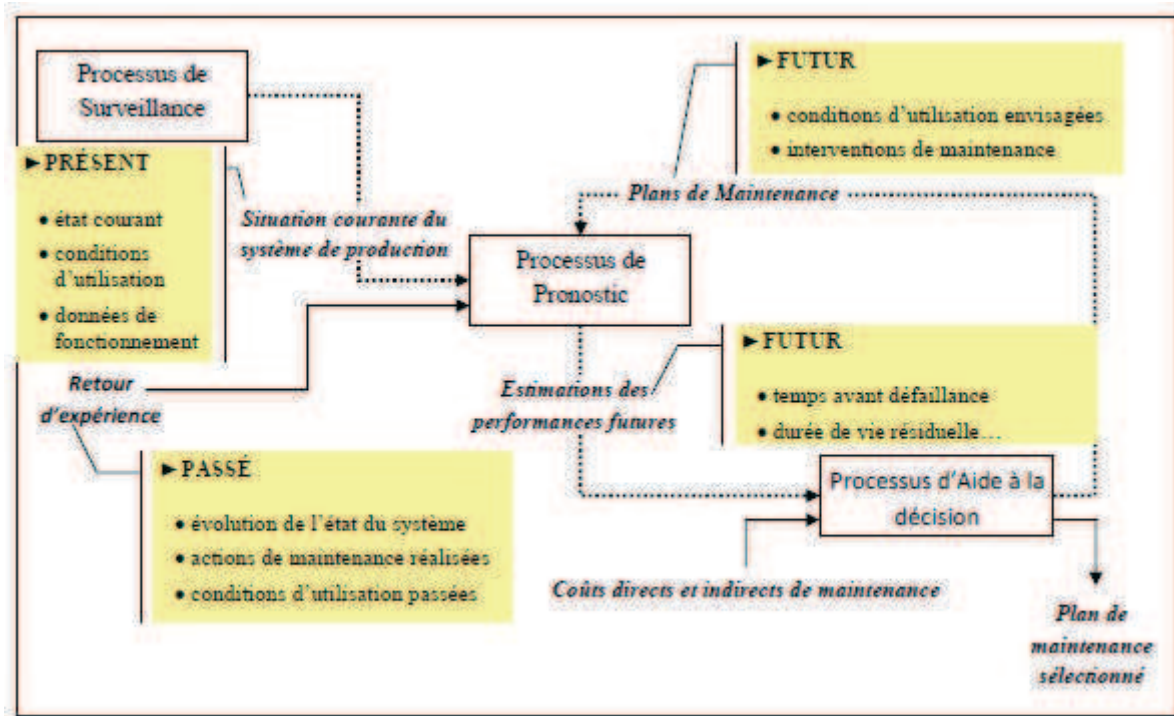


Figure 15 : Architecture du SIMP [Muller, 2005].

Par rapport à cette architecture, la notion de bilan de santé telle que défendue, dans ces travaux, se trouve en sortie de la fonction de surveillance. La fonction de surveillance du modèle SIMP a pour objectif de délivrer l'information concernant la situation courante du processus/système en ayant en amont les données de fonctionnement nécessaires. Cette information intègre aussi le résultat du diagnostic afin de donner exactement la localisation et la cause de l'anomalie une fois détectée (Figure 16).

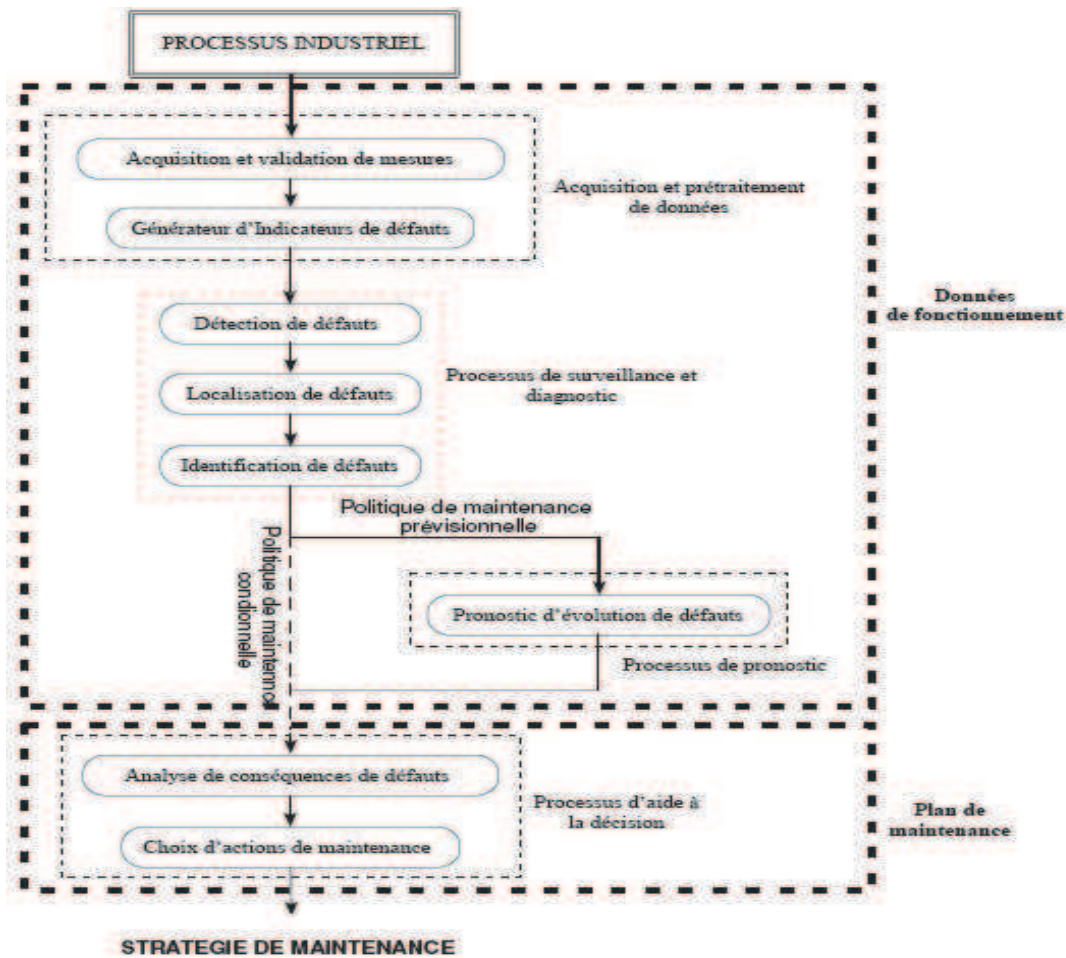


Figure 16 : Architecture de maintenance conditionnelle/prévisionnelle [Ben Salem, 2008].

3.3.3.3. Système IVHM

Le système IVHM (Integrated vehicle health management) a pour objectif de mettre en place une « philosophie » PHM pour permettre un suivi continu et une évaluation en temps réel de la santé fonctionnelle du véhicule, puis prévoir la vie utile restante d'une défaillance proche d'un composant [Williams, 2006]. Ce système utilise cette information pour améliorer les décisions opérationnelles, les actions de maintenance et la performance commerciale ultérieure [Williams, 2006] [Benedettini, 2009].

La Figure 17 illustre l'architecture fonctionnelle de base d'un système IVHM. La première étape consiste à récupérer, des capteurs, les mesures des variables indicatives de modes de dégradation potentiels. Le type de capteurs utilisé est spécifique au domaine d'application. Les mesures issues de ces capteurs sont préalablement traitées pour en retirer les bruits, puis manipulées pour extraire des caractéristiques de santé ou de dégradation. La fonction de diagnostic analyse ces informations afin de détecter, d'identifier et d'isoler les fautes imminentes et naissantes. Les informations issues du diagnostic sont combinées avec les données historiques dans le module de pronostic et sont utilisées pour générer une

estimation de la durée de vie restante des composants et sous-systèmes. Enfin, les informations de diagnostic et de pronostic sont transformées en des actions de maintenance.

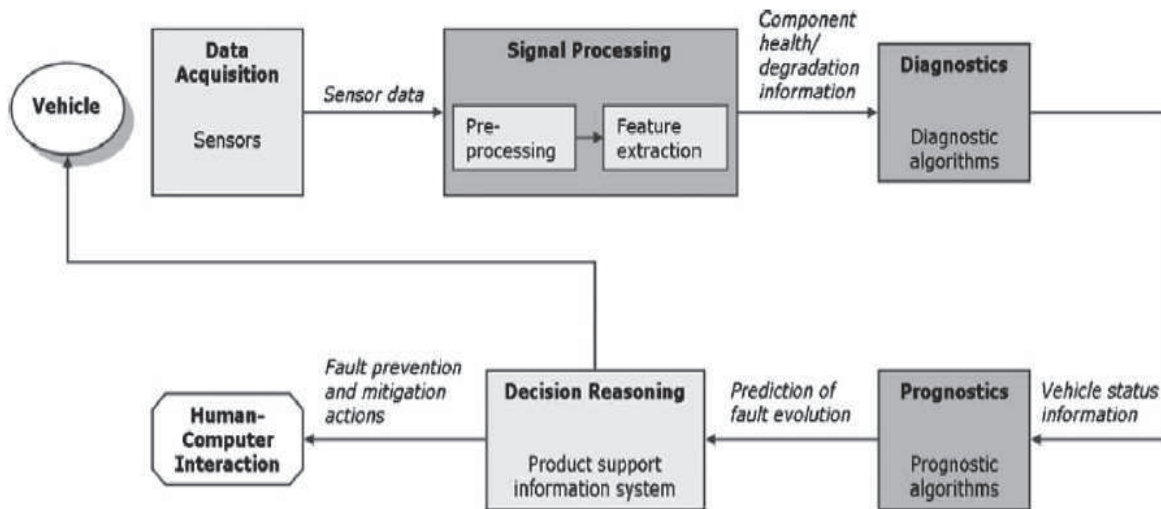


Figure 17 : Architecture du IVHM [Benedettini, 2009].

Ce système met l'accent sur l'importance de la fonction de traitements des données brutes issues des capteurs. Cette fonction correspond à celles de « data acquisition » et « signal processing » dans le standard OSA-CBM. Cette fonction est importante pour son rôle dans l'extraction des caractéristique de santé, appelés indicateur dans notre bilan de santé. La fonction de détection des dérives est cette fois ci intégrée dans le module de diagnostic. Par conséquent la vision bilan de santé de nos travaux se projette entre les deux fonctions d'extraction des caractéristiques de santé (ou indicateurs) et de diagnostic.

3.3.3.4. Architectures de fusion

Par rapport à notre vision de bilan de santé, en complément des architectures fonctionnelles précédemment citées, il est nécessaire aussi de regarder comment la fonction de fusion des informations pour le PHM (comme montré dans la Figure 9) peut être intégrée vis-à-vis des autres fonctions dans une architecture fonctionnelle.

Pour le PHM, il y a trois principaux domaines/niveaux où la fusion peut jouer un rôle contributif. La fusion au niveau des données brutes (souvent celles issues directement des capteurs), est effectuée afin de valider des signaux ou créer de nouvelles caractéristiques de santé. En effet, les outils de la fusion permettent d'augmenter la certitude des données par combinaison de signaux d'un groupe de capteurs afin de fournir un signal qui est généralement de forme identique mais de meilleure qualité [Liu, 2007]. En outre, la fusion à ce stade permet d'extraire des caractéristiques par le filtrage et les tests statistiques (changements de tendance, des paramètres de distribution statistique,...) [Yu, 2007]. A ce niveau, la fusion se représente à travers une structure centralisée de traitement (Figure 18).

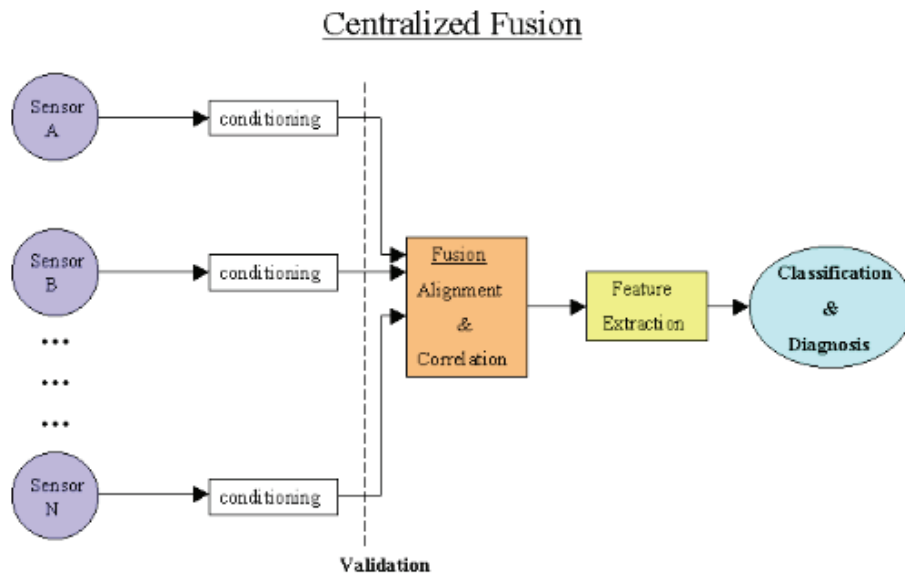


Figure 18 : Architecture de fusion centralisée [Roemer, 2001]

A un niveau d'abstraction plus haut, la fusion des caractéristiques (c'est-à-dire des vecteurs de caractéristiques sont extraits à partir d'observations des capteurs et sont ensuite fusionnés), peut être utilisée pour combiner ces caractéristiques d'une manière intelligente de façon à obtenir la meilleure information possible pour le diagnostic [Dong, 2007]. Ainsi, à un tel niveau, de nouvelles caractéristiques supplémentaires peuvent être utilisées par le système PHM [Luo, 1995]. Elle sert aussi comme une fonction de détection ou de diagnostic par les moyens de classification ou de reconnaissance des formes [Yu, 2007]. L'architecture correspondante à ce niveau correspond à celle de la fusion autonome (Figure 19).

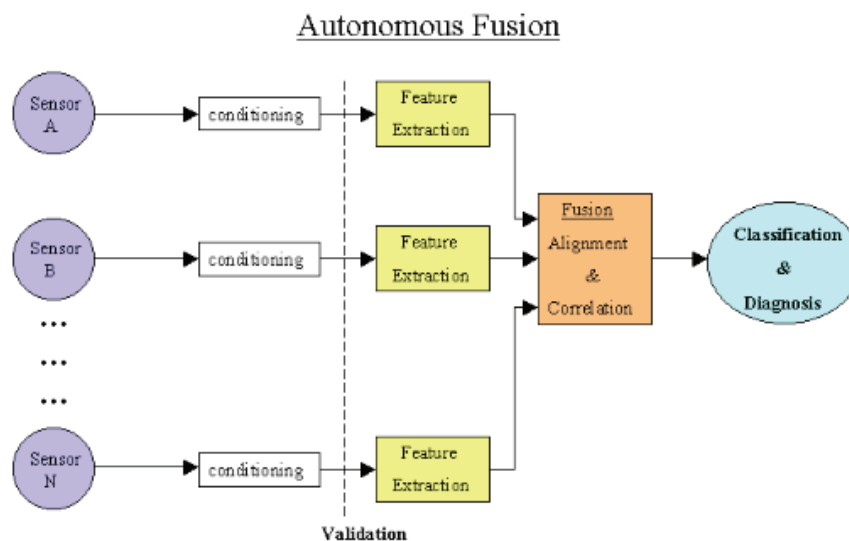


Figure 19 : Architecture de fusion autonome [Roemer, 2001]

Au niveau de décision, la fusion permet aux informations issues à partir de plusieurs décideurs locaux d'être efficacement utilisées ensemble à un plus haut niveau d'abstraction. Les différents acteurs locaux (résultats de classificateurs, information basée sur l'expérience

par exemple), ne sont pas indépendants, les modèles sont alors régulièrement combinés et consolidés pour produire la décision finale [Yu, 2007].

L'architecture de fusion qui paraît la plus appropriée pour notre objectif de thèse est celle de l'architecture de fusion autonome. En effet, les indicateurs sont issus de traitements séparés des données brutes, ensuite, ils sont fusionnés afin de former un support pour le diagnostic, la décision... La fusion au niveau des caractéristiques de santé (indicateurs) succède ainsi à celle de l'extraction de ces indicateurs.

3.3.4. La détection des anomalies et l'agrégation des indicateurs

En nous basant sur la considération du bilan de santé relativement aux architectures présentées précédemment, il est nécessaire à ce stade de synthétiser ces considérations et de les adapter pour que l'architecture résultante des fonctions (architecture simple et générique) supporte l'élaboration du bilan de santé tel que défini dans ce chapitre. La même architecture est appliquée pour chaque élément dans la structure hiérarchique du système indépendamment de sa nature ou de son niveau d'abstraction au sein de cette hiérarchie. Ayant les mêmes entrées et sorties que la «boîte noire» bilan de santé, cette architecture fonctionnelle doit être constituée de trois fonctions clés :

- le traitement de données en entrée pour en extraire les indicateurs tel que préconisé par la fonction « signal processing » de l'OSA-CBM et du IVHM.
- la détection des déviations pour évaluer les dérives des valeurs prises par les indicateurs par rapport au mode nominal. La fonction détection des dérives est la première étape pour toute fonction de surveillance comme dans le SIMP, comme elle est explicitement désignée par « condition monitoring » dans l'OSA-CBM.
- la fusion des indicateurs en un index de santé, en ce sens, l'architecture autonome de fusion est adoptée pour notre bilan étant donné que cette étape de fusion succède à celle d'élaboration des indicateurs.

La Figure 20 propose une vision de cette architecture en montrant l'enchaînement proposé entre ces fonctions pour aboutir à construire les indicateurs et indexes du bilan de santé. Après l'acquisition des données d'entrée, une phase de traitement des données brutes (généralement sous forme de signaux) par filtrage est nécessaire afin d'éliminer les valeurs aberrantes, les bruits... Cette étape est primordiale pour réduire la quantité des données manipulées ou stockées dans les bases de données. Cette même fonction permet aussi d'intégrer les données issues du bilan de santé d'un autre élément (depuis le niveau hiérarchique inférieur). Cette étape n'est pas toujours nécessaire surtout s'il s'agit du niveau hiérarchique le plus bas ou encore si des données du niveau inférieur n'entrent pas dans la détermination des indicateurs du niveau supérieur.

Comme un indicateur sert de support de mesure du degré de déviation de la caractéristique de la santé qu'il représente, la phase de détection est importante. Le résultat d'une telle étape est généralement un score entre 0 et 1 mesurant l'étendue de la déviation par rapport au mode nominal correspondant à cette même caractéristique. La fonction de détection tient compte du contexte environnemental et opérationnel afin que la déviation détectée soit significative par rapport à un tel contexte.

Ensuite, l'ensemble des indicateurs est rassemblé dans un seul index à travers la fonction d'agrégation. Une procédure de fusion des informations nécessite une conversion des données dans un système de coordonnées commun avant la concaténation. Cette étape de normalisation est obligatoire pour faire en sorte que les indicateurs: (1) soient sur une même échelle de valeurs et (2) permettent d'exprimer une sémantique équivalente. Cette étape de normalisation est équivalente à la fonction de détection. En effet, en attribuant pour chaque indicateur une valeur continue entre 0 et 1 ayant un sens bien déterminé s'agissant du degré de la normalité (ou inversement la non-nominalité), une échelle commune à des valeurs hétérogènes pour tous les indicateurs est obtenue.

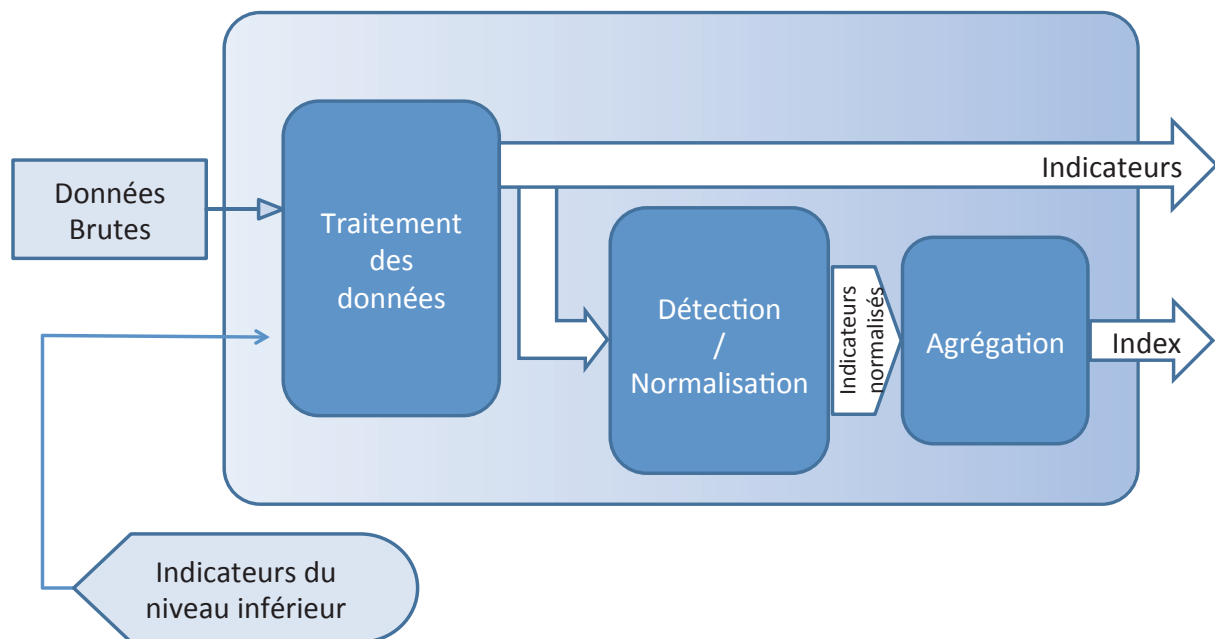


Figure 20 : Boîte générique élémentaire pour les modules d'activités bilan de santé

Les sous activités auxquelles notre thèse contribue plus fortement sont celles de la détection/ normalisation et d'agrégation vu leur importance vis-à-vis du bilan de santé. En effet, la détection/normalisation permet de quantifier la déviation des valeurs prises par les indicateurs par rapport à la normalité tout en tenant compte du contexte opérationnel et environnemental qui peut impacter la référence du mode nominal. En sortie de cette fonction, les indicateurs ont des valeurs situées entre 0 et 1 indiquant le degré de nominalité (1 pour les valeurs des indicateurs correspondant à un fonctionnement nominal). De cette

manière, les indicateurs sont normalisés et prêts à être fusionnés en un index de santé. En ce sens, la fonction de détection/normalisation joue un double rôle : elle permet non seulement de détecter une anomalie locale à partir d'un indicateur mais aussi de normaliser des indicateurs portant des informations hétérogènes. Une deuxième activité sur laquelle porte notre contribution scientifique est celle de l'agrégation. Cette double orientation détection/normalisation et agrégation est totalement cohérente pour répondre aux verrous industriels plus spécifiques mentionnés dans le paragraphe 2.5 en regard du projet BMCI.

4. Conclusion

A partir des besoins exprimés dans le projet industriel BMCI, la problématique d'élaboration de bilan/état de santé d'un navire en structurant ce dernier autour des niveaux composants, sous-système, système a été identifiée. Ensuite, les premières investigations scientifiques, résultante d'une réflexion initiale dans BMCI, ont été présentées pour apporter une première réponse à la problématique initialement posée. Ces investigations proposent une structuration multi niveaux du bilan de santé. Ce dernier se base sur un vecteur d'indicateurs fonctionnels/dysfonctionnels/environnementaux. Afin de rationaliser ce concept et le porter vers des systèmes techniques d'une manière plus générique, la notion de santé est vue comme étant : 1) la performance fonctionnelle en lien étroit avec la finalité du système, 2) les paramètres physiques caractéristiques des dégradations, 3) tout en tenant compte du contexte environnemental et d'usage dans lequel évolue ce système. En outre, le concept d'index permettant de synthétiser la déviation par rapport à la nominalité de ces indicateurs hétérogènes est justifié. Il se pose alors la question des fonctions nécessaires pour élaborer ces indicateurs/indexes. En ce sens, une architecture simplifiée constituée des trois fonctions de traitements des données, de détection d'anomalies et d'agrégation (mettant en œuvre une architecture autonome de fusion) a été introduite à partir d'architectures connues pour des systèmes existants de maintenance basés sur le suivi et l'évaluation de la santé dans un contexte PHM.

Cette architecture simplifiée a permis d'identifier les deux fonctions clés de détection/normalisation et d'agrégation des indicateurs en un index de santé sur lesquelles doit porter notre contribution scientifique. Afin d'identifier les techniques qui peuvent supporter ces deux fonctions en regard des besoins d'élaboration de notre bilan de santé, un état de l'art sur les approches de détection des anomalies ainsi que celles utilisées pour la fusion des informations doit être effectué afin de pouvoir orienter notre contribution vers les méthodes les plus adéquates à notre contexte de recherche. Ceci fait l'objet du deuxième chapitre.

Chapitre 2 : Bilan de santé : la détection des anomalies et la fusion des informations

1. Introduction

Un bilan de santé multi-niveaux d'abstraction est composé de vecteurs d'indicateurs quantifiant la déviation par rapport au mode de fonctionnement nominal lié à un contexte opérationnel/environnemental donné. A chaque vecteur d'indicateurs est associé un index synthétique permettant d'informer son utilisateur en cas d'apparition d'une déviation de l'état du système. Pour l'élaboration de ces indicateurs/indexes, une architecture générique a mis en évidence dans le premier chapitre l'importance des activités/fonctions de détection/normalisation et agrégation qui représentent le cœur de notre contribution. Par rapport à la vision système multi-niveaux, ces fonctions sont à considérer aussi bien sur un même niveau d'abstraction que cohabitant entre niveaux différents. Ce chapitre porte par conséquent sur un état de l'art succinct concernant les outils/méthodes en support de ces deux activités. Cependant, les approches proposées par les deux communautés de fusion de l'information et de la détection des anomalies sont très variées et dépendent du contexte de leur utilisation. Ainsi, après comparaison, les outils les plus proches des besoins émergents de BMCI et de leur généralisation sont isolés pour constituer un sous espace de solutions tout en précisant les adaptations qui seraient nécessaires pour ces outils/méthodes afin de supporter l'ensemble de nos problématiques.

En ce sens, la première partie de ce chapitre s'intéresse au domaine de la détection des anomalies en vue de présenter un état-de-l'art succinct sur les méthodes existantes liées à ce domaine. Par rapport aux besoins liés à la problématique du projet BMCI et sa généralisation, les méthodes les plus susceptibles à répondre à de tels besoins sont mis en évidence. Enfin, les adaptations qui restent à apporter à ces deux outils pour aboutir à notre approche de détection/normalisation sont présentées. D'une façon duale, la deuxième partie de ce chapitre est consacrée au domaine de la fusion des informations et plus précisément à celui de l'agrégation étant donné que l'étape de normalisation est incluse dans l'étape de détection des dérives. Ainsi, une vue synthétique des travaux sur l'agrégation des données nous amène à justifier le choix de l'intégrale de Choquet pour agréger les indicateurs du bilan de santé en un index global. Ensuite, un état-de-l'art sur les méthodes d'identification des paramètres de cet outil, à savoir les coefficients de capacité, permet d'orienter notre démarche d'identification vers une approche non supervisée à base

d'entropie relative. Enfin, les adaptations qui restent à apporter à cette approche pour répondre à nos objectifs sont mentionnées.

2. La détection des anomalies

La détection des anomalies est une activité très importante par la diversité de ses domaines de recherche et d'application. En effet, par exemple, cette activité est appréhendée pour la détection des fraudes sur les cartes de crédit [Aleskerov, 1997], la détection des pathologies en médecine [Spence, 2001] [Tarassenko, 1995], la détection d'intrusion pour la cybersécurité [Pacha, 2007] [Lee, 2001 a], la surveillance militaire des activités des ennemis et la détection d'espionnage [Fujimaki, 2005] [George, 2011], la détection des dégradations dans les machines [Rayhane, 2004] [King, 2002] [Coble, 2010]. C'est par rapport à ce dernier domaine que nos travaux de thèse se situent.

2.1. La détection dans un processus de surveillance

Par rapport à notre contexte de thèse, à savoir celui de la surveillance des systèmes techniques, la détection des anomalies a pour objectif de détecter le passage du système en fonctionnement anormal. Les données issues des capteurs sont transformées en des indicateurs de santé permettant de classer les situations observées comme étant normales ou anormales [Ribot, 2009]. Pour ce domaine, les méthodes de détection sont confondues avec les outils de surveillance d'une manière générale. En effet, la détection est la première fonction dans un processus de surveillance formé par les deux fonctions de détection et de diagnostic. Cette dernière consiste à localiser les anomalies une fois détectées et à identifier leurs causes [Racoceanu, 2006] [Verron, 2007] (Figure 21).

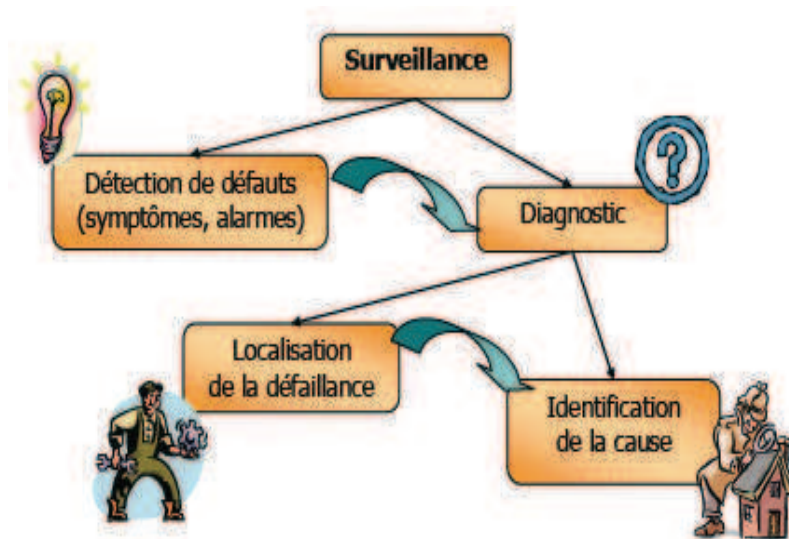


Figure 21 : Positionnement de la détection dans le processus de surveillance [Racoceanu, 2006].

Nous rappelons que par rapport à l'élaboration d'un bilan de santé, la fonction de détection a pour rôle de :

- Quantifier l'écart des valeurs prises par chaque indicateur par rapport à ceux du mode nominal. Les déviations significatives déclencheront l'étape ultérieure de diagnostic (localisation et identification).
- Normaliser des indicateurs hétérogènes en leur apportant une sémantique commune, à savoir celle du degré de nominalité. La fonction de détection prépare ainsi l'agrégation de ces indicateurs en un index global.

Vis-à-vis de ces deux rôles, et en regard de l'objectif de l'élaboration du bilan de santé, il est nécessaire d'identifier les critères caractérisant les attendus de l'approche de détection afin de trouver les sous espaces de méthodes les plus adéquates face à notre objectif.

2.2. Critères de sélection de l'approche de détection

Pour l'identification de ces critères, nous proposons à la fois d'isoler des critères spécifiques à BMCI en lien avec les verrous industriels identifiés dans le premier chapitre mais aussi des critères plus généraux issus d'une généralisation des critères précédents ou des travaux existants.

Parmi les verrous industriels dégagés lors du premier chapitre, un verrou essentiel est relatif à l'intégration du contexte opérationnel et/ou environnemental des missions navales dans l'évaluation des dérives des indicateurs. En effet, dans un contexte changeant, les valeurs prises par des indicateurs peuvent être affectées par ce changement sans que ce comportement ne soit considéré comme étant non-nominal. En outre, les exigences initiales du projet sur la portabilité de la solution apportée à d'autres équipements hors du champ naval nous amènent à considérer le besoin lié à la généricité de l'approche de détection des anomalies à adopter. Cette contrainte de généricité se décline en deux critères de sélection d'un outil de détection adéquat, à savoir, la disponibilité de connaissances sur le système surveillé (connaissances physiques, expertise, données historiques...) pour en constituer un modèle, d'une part, et les hypothèses faites sur les paramètres du comportement nominal des indicateurs, d'autre part.

D'un point de vue de l'intégration de la fonction de la détection/normalisation des indicateurs dans le processus d'élaboration d'un bilan de santé proposé, des contraintes liées à la nature des sorties de cette fonction doivent être aussi considérées dans le choix de l'outil à adopter. En ce sens, l'obtention d'un score en sortie de cette fonction qui varie entre 0 et 1 pour chaque indicateur comme quantification de leur degré de dérive par rapport au mode nominal doit permettre leur normalisation en même temps. Ils sont ainsi prêts à être directement agrégés en un index synthétique.

Par rapport aux besoins et contraintes listés précédemment, il est donc possible de dégager les critères suivants devant être supportés par les approches de détection/normalisation dans notre contexte de thèse :

1. Applicable pour les systèmes sans modèle (modèle physique ou règles de comportement),
2. Délivrance d'un score entre 0 et 1 en sortie de cette fonction,
3. Caractérisation non paramétrique du mode nominal pour un indicateur quelconque,
4. Prise en compte des données contextuelles pour la détection des anomalies.

Ces critères sont utilisés pour comparer les approches de détection afin d'isoler celles les plus proches de nos besoins.

2.3. Classifications existantes des approches de détection des anomalies

De nombreux état-de-l'art ont été déjà réalisés sur les différentes approches et outils pour la résolution des problèmes de la détection des anomalies [Markou, 2003a], [Markou, 2003b], [Chandola, 2009], [Patcha, 2007]. Les approches de détection étudiées pour le domaine de la surveillance ont été aussi classées dans [Venkatasubramanian, 2003a], [Venkatasubramanian, 2003 b], [Venkatasubramanian, 2003c], [Boufaied, 2003], [Katipamula, 2005], [Racoceanu, 2006], [Verron, 2007], ...

Par rapport à ces classifications, nous proposons, en nous appuyant sur le premier critère (applicabilité pour les systèmes sans modèle) de dégager un premier sous espace de solutions. Ensuite, au sein de ce sous espace, nous nous basons sur les trois autres critères afin de comparer les approches et d'isoler celles susceptibles de contribuer à notre objectif.

2.3.1. Classification suivant l'utilisation de modèle

Les deux groupes de méthodes issus de cette classification sont les méthodes à base de modèles et celles sans modèle (Figure 22) [Racoceanu, 2006], [Katipamula, 2005], [Verron, 2007], [Boufaied, 2003], [Venkatasubramanian, 2003 a]. Les méthodes à base de modèles s'appuient sur des modèles comportementaux explicites du système surveillé [Mokhtari, 2007]. Ces modèles peuvent être de type quantitatif ou qualitatif [Katipamula, 2005]. Les modèles quantitatifs sont exprimés sous forme d'équations mathématiques basées sur des principes physiques fondamentaux (appelée aussi les approches FDI, Fault Detection and Isolation) [Theilliol, 2010].

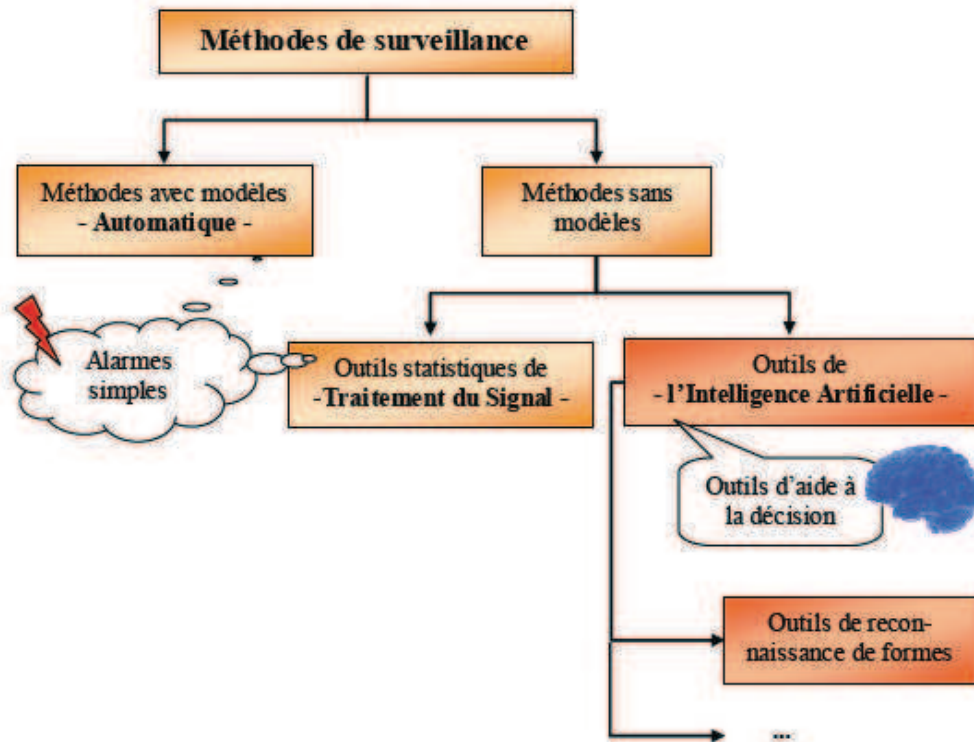


Figure 22 : Classification des méthodes de surveillance [Racoceanu, 2006].

Ces équations sont de type algébro-différentielles permettant de représenter le comportement continu du système avec une certaine précision numérique [Ribot, 2009]. La détection des anomalies est effectuée en comparant les mesures sur le système aux informations prévues par le modèle sur la base des entrées mesurées par des capteurs sur le système réel [Frank, 1990]. Tout écart (appelé souvent résidu) est étudié afin de déterminer ensuite s'il est dû à des aléas normaux, comme par exemple le bruit de mesure, ou s'il traduit une dégradation du système [Verron, 2007] (Figure 23).

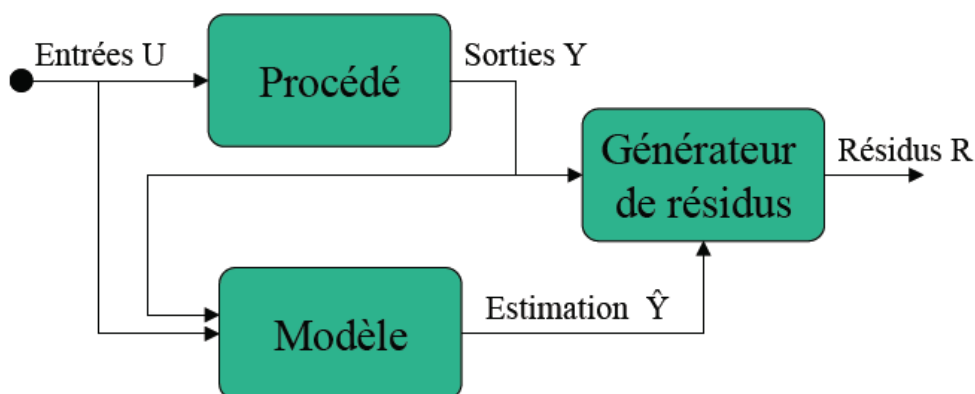


Figure 23 : Procédure de détection à base de modélisation quantitative [Verron, 2007]

Des approches de détection à base de modèles qualitatifs sont aussi utilisées, appelées des méthodes à base de connaissance [Verron, 2007]. Elles exploitent les compétences, le raisonnement et les connaissances des experts sur le processus pour les transformer en

règles logiques de bon comportement, de manière à différencier le mode nominal du mode de comportement avec défauts [Venkatasubramanian, 2003 a].

Les approches de détection à base de modèles (quantitatifs ou qualitatifs) ont certainement des avantages, comme par exemple la précision des modèles quantitatifs et la transparence et la simplicité de l'application des modèles qualitatifs [Katipamula, 2005]. Cependant, de telles méthodes nécessitent une connaissance physique ou comportementale spécifique du système à surveiller. Par rapport au premier critère de « non nécessité de modèle » identifié dans le paragraphe 2.2., ces approches ne sont donc pas réellement adaptées. Il est ainsi plus logique de ne retenir que les méthodes de détection sans modèle pour la suite de notre investigation.

Les méthodes sans modèle cherchent à caractériser le comportement type d'un système en mode nominal, au sein des mesures sur le système étudié. Ainsi la seule information disponible sur le système est sous la forme de données historiques ou de résultats de traitement en ligne des signaux issus des capteurs [Mokhtari, 2007], [Venkatasubramanian, 2003 c].

Au sein de cette catégorie (les méthodes sans modèle), différentes classifications ont été proposées. La classification la plus répandue est celle synthétisée dans [Racoceanu, 2006], [Markou, 2003 a], [Markou, 2003 b], à savoir celle divisant les méthodes de détection sans modèles (ou à base des données) en deux principales catégories : les méthodes statistiques à base de traitement du signal, d'une part, et les méthodes à base de reconnaissance des formes, d'autre part (Figure 22).

2.3.1.1. Détection à base d'outils statistiques

Les outils statistiques de détection des anomalies consistent à supposer que les données étudiées possèdent certaines propriétés statistiques. Ces outils statistiques établissent des tests sur les données acquises afin de détecter les échantillons de mesures qui ne correspondent pas à ces propriétés [Boufaied, 2003]. Une grande variété de tests, applicables sur un échantillon de signaux mesurés tel que le test de franchissement de seuils, le test de moyenne et le test de variance, test de distribution... sont utilisés selon l'allure et/ou les caractéristiques du signal suivi [Basseville, 1988].

2.3.1.2. Détection par reconnaissance des formes

Ces approches à base des données font appel à des techniques d'apprentissage numérique et de classification (Figure 24). L'objectif de ces méthodes est d'associer un ensemble de mesures observées à des états de fonctionnement connus parmi plusieurs classes (comportement nominal, dégradé ou défaillant par exemple). Les classes peuvent être définies de façon neuronale, probabiliste ou encore comme des classes floues [Racoceanu,

2006]. Les méthodes utilisées les plus répandues sont les réseaux de neurones, les Support Vector Machines... [Markou, 2003 b].

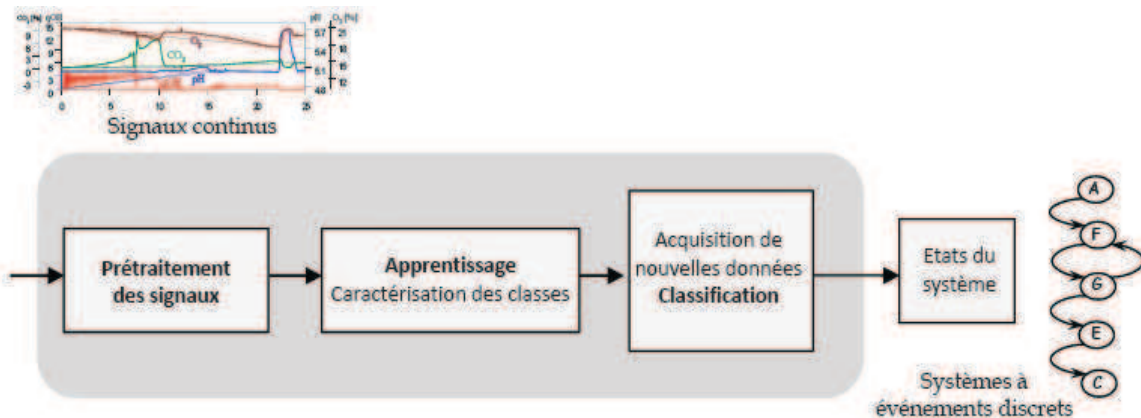


Figure 24 : Surveillance par reconnaissance de formes [Ribot, 2009].

Sur la base de ce positionnement des méthodes sans modèle comme plus cohérentes par rapport à notre premier critère à respecter, nous proposons maintenant d'affiner ce sous-espace par rapport au deuxième critère relatif à la nature de la sortie de la fonction de détection.

2.3.2. Classification suivant la nature de la sortie

Le critère que nous utilisons maintenant est celui de la nature de la sortie de cette fonction de détection. En effet, la sortie peut être sous forme d'un label explicitant qualitativement le mode dans lequel se situe le système (mode nominal ou en présence d'une anomalie) (Figure 25) ou d'un score reflétant le degré avec lequel la situation actuelle est considérée comme non-nominale [Chandola, 2009].

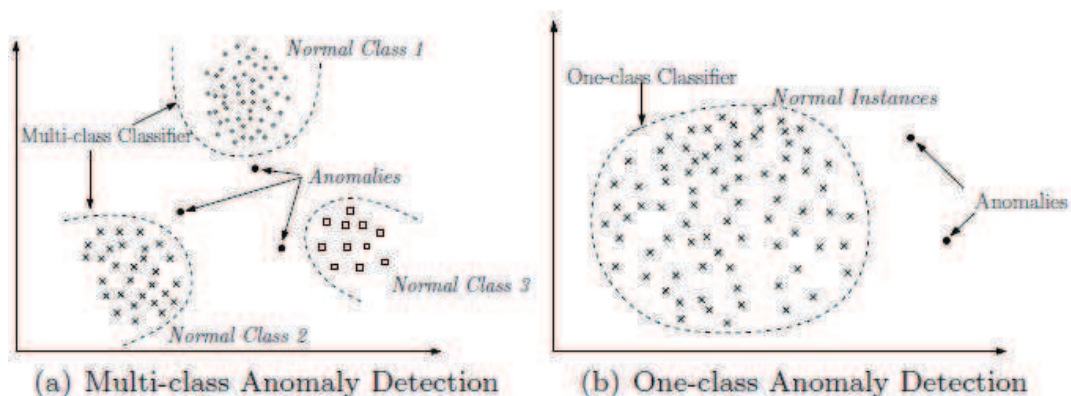


Figure 25 : Méthode de détection avec un label en sortie [Chandola, 2009].

Ainsi, en considérant ce critère, deux groupes de méthodes sans modèle peuvent être dégagés : les méthodes avec label en sortie et celles délivrant un score.

2.3.2.1. Les méthodes avec label en sortie

Ce sous ensemble de méthodes est constitué essentiellement par les outils de classification et des méthodes à base d'outils statistiques du traitement du signal (dépassement de seuils, test de distribution, de variance...). Ces derniers outils permettent par exemple d'alerter dès que le signal surveillé dépasse un seuil prédéfini [Ginart, 2006], [Patrick, 2010], [Zhong Qing, 2010], [Niu, 2010 b]. Pour les techniques de classification, les labels peuvent être uniquement de type nominal et non nominal et visent à alerter en cas de déviation par rapport au mode nominal quelque soit la nature de l'anomalie constatée. D'autres méthodes de classification effectuent la détection avec plus d'informations sur le type de l'anomalie relevée lorsque les labels désignent, à côté du mode nominal, différents modes de défaillance. Dans ce dernier cas, la détection joue aussi le rôle de diagnostic [Byington, 2001] [Sotiris, 2010] [Yang, 2004], [Volponi, 2004].

2.3.2.2. Les méthodes avec un score en sortie

Parmi les méthodes de détection des anomalies délivrant un score reflétant le degré de déviation, nous pouvons citer les méthodes de clustering [Pan, 2010], celles à bases d'histogrammes tel que proposé par [Rizzolo, 2011] [Goldstein, 2012], ou encore celles à base des mesures de la théorie de l'information [Srivastav, 2009] [Yan, 2007].

Les techniques de détection les plus simples sont celles utilisant des histogrammes pour caractériser efficacement le profil des données d'une manière générale. Ces approches de détection à base histogrammes pour des données unidimensionnelle se déroulent en deux étapes [Chandola, 2009]. La première étape consiste à construire un histogramme sur la base des différentes valeurs prises par la variable suivie (un indicateur par exemple) lors de la phase d'apprentissage. Dans la seconde étape, la technique accorde un score à l'instance (valeur) en cours relatif à la hauteur de la barre de l'histogramme dans laquelle cette instance est localisée [Goldstein, 2012]. Dans l'approche proposée dans [Rizzolo, 2011], les scores sont accordés à l'aide d'une fonction d'utilité associée à l'histogramme.

Ce paragraphe a présenté de manière synthétique une classification des méthodes de détection sur la base du critère de la nature de la sortie (score ou label). Sur la base de cette classification, seules les approches délivrant un score sont retenues afin de satisfaire le besoin explicité en termes de normalisation. Une autre classification sur la base de la nécessité ou pas de prendre en compte des paramètres pour caractériser l'évolution de l'indicateur de santé est discutée dans la section suivante.

2.3.3. Classification suivant la nécessité de paramètres pour caractériser le mode nominal

Par rapport à ce critère, à savoir la manière dont le comportement des indicateurs est caractérisé/modélisé, il existe deux types d'approches [Markou, 2003 a] [Chandola, 2009]:

- Les méthodes paramétriques telles que les approches probabilistes de test d'hypothèses, mélange de modèles gaussiens (GMM), les modèles de Markov cachés (HMM), les réseaux de pétri...
- Les méthodes non paramétriques : celles à base d'histogrammes, estimation de densité par noyau (ou par fenêtre de Parzen), les méthodes basées sur les mesures issues de la théorie de l'information et la plus part des méthodes de classification...

Par rapport aux besoins liés au projet, nous nous intéressons aux approches non paramétriques afin d'assurer la généralité de l'approche proposée. Les approches à base d'outils de la théorie de l'information (entropie, entropie conditionnelle, entropie relative...) [Cover, 2006] pour la détection des anomalies peuvent être utilisées sans nécessité de paramètres pour caractériser le mode nominal ou de faute. Ces approches trouvent une plus grande utilisation dans les domaines de détection des anomalies liées aux attaques des réseaux informatiques ou de communication [Kar, 2009] [Gu, 2005] [Lee, 2001 a]. Ces outils ont l'intérêt de pouvoir travailler en mode non supervisé et ne nécessitent aucune hypothèse a priori concernant la distribution statistique des données. Le principe de ces approches se base sur la mesure de la régularité au sein d'un échantillon de données observées et la détection des points représentant des anomalies faisant diminuer cette régularité [Ando, 2007]. De telles approches s'appuient sur des algorithmes d'optimisation (chercher un maximum ou une limite pour une mesure entropique) afin de maximiser la régularité des valeurs prises par la variable suivie [He, 2005]. D'autres approches non paramétriques et simples à base d'histogrammes sont aussi utilisées pour la modélisation du comportement des variables suivies pour la détection [Kind, 2009], [Lakhina, 2005]... Les histogrammes permettent en effet de construire, à partir des différentes valeurs prises par les variables en phase d'apprentissage, une référence [Goldstein, 2012] [Perona, 2010].

Il nous reste maintenant à considérer le dernier critère relatif à la prise en compte du contexte pour la détection des anomalies. En ce sens, [Chandola, 2009] présente dans son survey une classification originale des méthodes de détection des anomalies sur la base de la nature des données en entrée.

2.3.4. Classification suivant la nature des entrées

[Chandola, 2009] adopte la classification des méthodes en trois groupes (Figure 26) :

- Les méthodes de détection des anomalies ponctuelles,

- Les méthodes de détection des anomalies séquentielles (ou collectives),
- Les méthodes de détection des anomalies contextuelles.

En effet, une situation du système à un instant donné peut être considérée comme étant une anomalie dans un contexte bien spécifique et non par ailleurs. Ainsi, elle est qualifiée d’anomalie contextuelle ou conditionnelle. En outre, si une séquence de points représentant un attribut est considérée comme une anomalie sans que les données individuelles ne soient des anomalies en eux-mêmes, on parle dans ce cas d’une anomalie collective. Les méthodes de détection ponctuelle considèrent, pour la détection des instances, des données séparément les unes des autres.

Les problèmes de détection des anomalies contextuelles et collectives peuvent être ramenés à des problèmes de détection d’anomalies ponctuelles. En effet, une approche pour les anomalies collectives peut être résolue en transformant les séquences des données en un espace d’un nombre fini de caractéristiques liées à ces données, et utiliser ensuite une technique parmi celles de la détection d’anomalies ponctuelles dans le nouvel espace afin de détecter les anomalies.

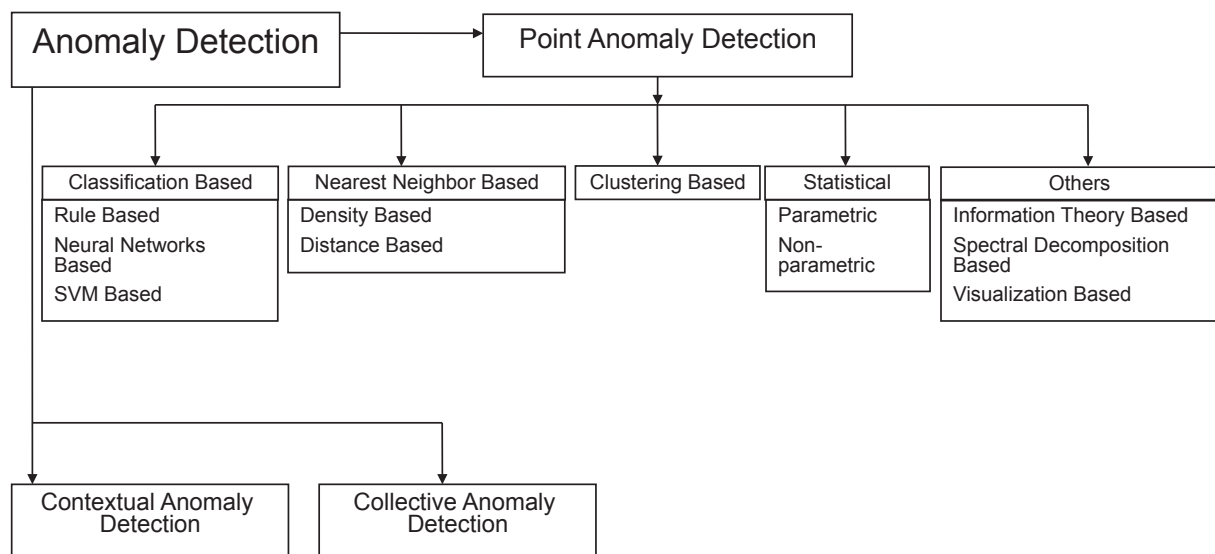


Figure 26 : Classification des méthodes de détection des anomalies [Chandola, 2009].

En outre, les anomalies contextuelles sont en fait des instances individuelles de données (comme les anomalies ponctuelles) mais ne sont considérées comme étant non nominales que par rapport à un contexte particulier. Il suffit alors d’utiliser une méthode de détection d’anomalie ponctuelle dans chaque sous espace du contexte [Rabatel, 2011] [Williem, 2012] [Song, 2007].

2.4. Comparaison des méthodes de détection

Par rapport aux différents positionnements des approches au regard des critères identifiés en paragraphe 2.2.2., nous proposons donc sous la forme d'un tableau une synthèse des approches/critères ainsi identifiés. Comme mentionné dans le paragraphe 2.3.4., un problème de détection contextuelle peut être ramené à un cas de détection ponctuelle avec la répartition de l'espace contextuel. Ainsi, nous allons favoriser les deux critères de « la nature de la sortie : score ou label » et « la nécessité ou pas de paramètres pour caractériser le mode nominal » pour comparer les approches de détection sans modèle.

Groupes de méthodes	Paramétrique ou Non paramétrique	Score ou label
Les méthodes statistiques paramétriques : <ul style="list-style-type: none"> - Dépassement du seuil, - Test de moyenne, - Test de variance, - Test de distribution (GMM) - HMM 	Paramétriques	Label
Les méthodes de détection par classification : <ul style="list-style-type: none"> - Réseaux de neurones, - SVM, - Neuro-floue 	Non paramétrique	Label
Méthodes statistiques non paramétriques : <ul style="list-style-type: none"> - Histogrammes, - Clustering, - kNN 	Non paramétrique	Score
Méthodes statistiques non paramétriques : <ul style="list-style-type: none"> - Estimation par le noyau 	Non paramétrique	Label
Détection à base de la théorie de l'information	Non paramétrique	Score

Tableau 1 : Comparaison entre les différents groupes de méthodes sans modèle selon les deux critères de généralité et de la nature de la sortie

D'après le Tableau 1, les méthodes statistiques non paramétriques délivrant un score (ceux à base d'histogrammes, de « Clustering » ou de « k plus proche voisin ») ainsi que les méthodes à base de mesures de la théorie de l'information ou d'histogrammes se rapprochent le plus de notre objectif de détection/normalisation. Nous nous proposons ainsi de nous positionner dans cet espace de solutions réduit pour la détection et la normalisation aussi.

Cependant, les méthodes à base de clustering, de k-plus proche voisin ou celles issues du domaine de la théorie de l'information sont par nature des approches à base d'algorithmes d'optimisation ce qui entraîne des temps de calcul importants non adaptés à une détection

en ligne (ce qui est notre cas pour une surveillance en continu). En outre, les méthodes à base d'histogrammes ne permettent pas de positionner l'observation courante des valeurs prises par un indicateur par rapport aux deux modes de comportement de référence, à savoir le mode nominal et le mode de faute tel que l'exige la définition de la notion de santé. En effet, cette méthode se base sur un seul modèle sous forme d'histogramme [Goldstein, 2012] [Perona, 2010] caractérisant l'ensemble des valeurs prises par une variable. Ceci ne permet pas de séparer les deux modes de référence (nominal et faute) et de déterminer ainsi l'écart du comportement observé en ligne par rapport à ces deux modes. En outre, ces méthodes à base d'histogramme ne tiennent compte que d'un seul point d'observation. Si ce dernier correspond à une valeur aberrante, il risque d'être signalé par erreur comme étant une déviation due à l'apparition de dégradations.

Notre objectif est ainsi de profiter de l'avantage majeur des histogrammes, à savoir leur pouvoir de modélisation non paramétrique des comportements des signaux, en les intégrant dans une approche originale permettant de délivrer un score positionnant l'observation en cours entre les deux modes de référence et ce de manière robuste (prise en compte de toute une séquence de points). L'approche proposée doit aussi prendre en compte la répartition de l'espace contextuel tel que préconisé dans le paragraphe 2.3.4 pour permettre une détection contextuelle des anomalies.

3. L'agrégation des indicateurs

Une fois les indicateurs normalisés, l'étape suivante est leur agrégation en un index de santé global et synthétique du composant, du sous-système ou du système surveillé. En effet, les défaillances sont souvent liées à de nombreux mécanismes de dégradation corrélés, qui peuvent se traduire par des indicateurs de dégradation extraits de plusieurs signaux de capteurs [Lee, 2006]. En outre, la performance peut être influencée par ces dégradations. En ce sens, un indicateur (de performance ou de dégradation) pris à part n'est efficace que pour révéler une certaine caractéristique de la santé à un certain stade. Par conséquent, une fusion effective de plusieurs indicateurs est susceptible de fournir une information de plus haute qualité (riche et robuste) car issue de différentes sources. L'index de santé ainsi obtenu permet d'apporter aux processus ultérieurs de diagnostic et/ou de pronostic une entrée plus fiable [Niu, 2010 b].

Nous proposons par conséquent de dresser un état de l'art succinct des outils d'agrégation afin de choisir celui qui permet de satisfaire aux besoins issus de la problématique d'agrégation des indicateurs.

3.1. Concepts généraux

L'agrégation fait partie du domaine plus vaste de la fusion des informations. La fusion des informations est une procédure adoptée dans de nombreux domaines (aide à décision multicritères ou la fusion des données multi-capteurs par exemple). Ainsi, des informations de différentes natures (signaux, données, connaissances...) et complémentaires concernant un objet unique sont fusionnées afin de donner une description globale de celui-ci à partir des vues partielles (Figure 27). Cela a pour objet d'éliminer les incohérences de la description partielle [Bouchon-Meunier, 1998].

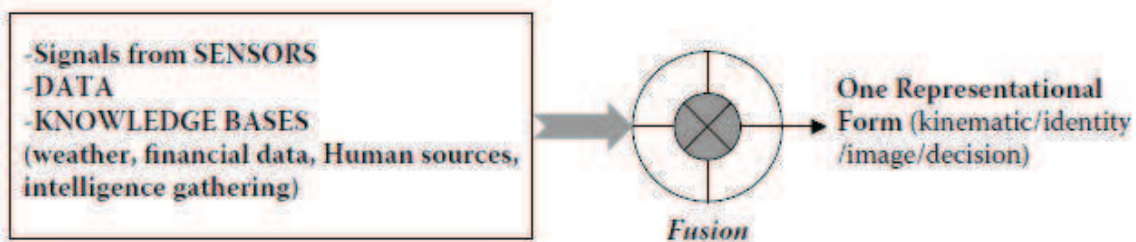


Figure 27 : Schéma conceptuel de la fonction de fusion [Raol, 2010]

Les méthodes de fusion les plus simples sont les opérateurs d'agrégation [Torra, 2007] [Beliakov, 2007], les méthodes de classement et de vote [Byington, 2001], [Roemer, 2001]. Des méthodes plus compliquées sont aussi utilisées comme les réseaux de neurones artificiels [Niu, 2010 a], la logique floue [Arif, 2005], les théories de l'incertain (théorie de l'évidence, théorie des possibilités, théorie de probabilités) donnant naissance à des outils de fusion comme ceux de l'inférence Bayésienne et de Dempster-Shafer [Wu, 2002] [Yang, 2006] [Yu, 2007] [Perrot, 1996] [Raol, 2010].

La fusion des données nécessite une conversion de ces données dans un système de coordonnées communes avant leur combinaison ou concaténation (Figure 28). Cette étape correspond à la fonction d'alignement (appelée aussi conditionnement ou encore parfois, harmonisation) [Arif, 2005] et celle que nous avons appelée la phase de normalisation et assurée par la détection.

L'opération de détection/normalisation va permettre d'assurer la commensurabilité de l'ensemble des variables (i.e. la mise sur une échelle permettant leur agrégation [Berrah & Clivillé, 2007]). Cette étape n'est possible que si l'opération d'agrégation est significative au sens de la théorie du mesurage [Grabisch & Labreuche, 2004]. Ceci passe par la construction d'une échelle commune (échelle ordinale, échelle intervalle, échelle de ratio) dans laquelle des relations d'ordre, de différence ou de ratio entre les données à agréger sont significatives [Grabisch, 2006] [Torra, 2007] [Tversky, 1992]. Cette étape d'alignement ou normalisation est jugée comme étant un problème difficile et pour lequel il n'existe pas de théorie générale [Wald, 1999]. L'agrégation quant à elle consiste à intégrer des données provenant

de sources multiples afin de produire une donnée précise, complète et unifiée sur une entité [Hall, 1997].

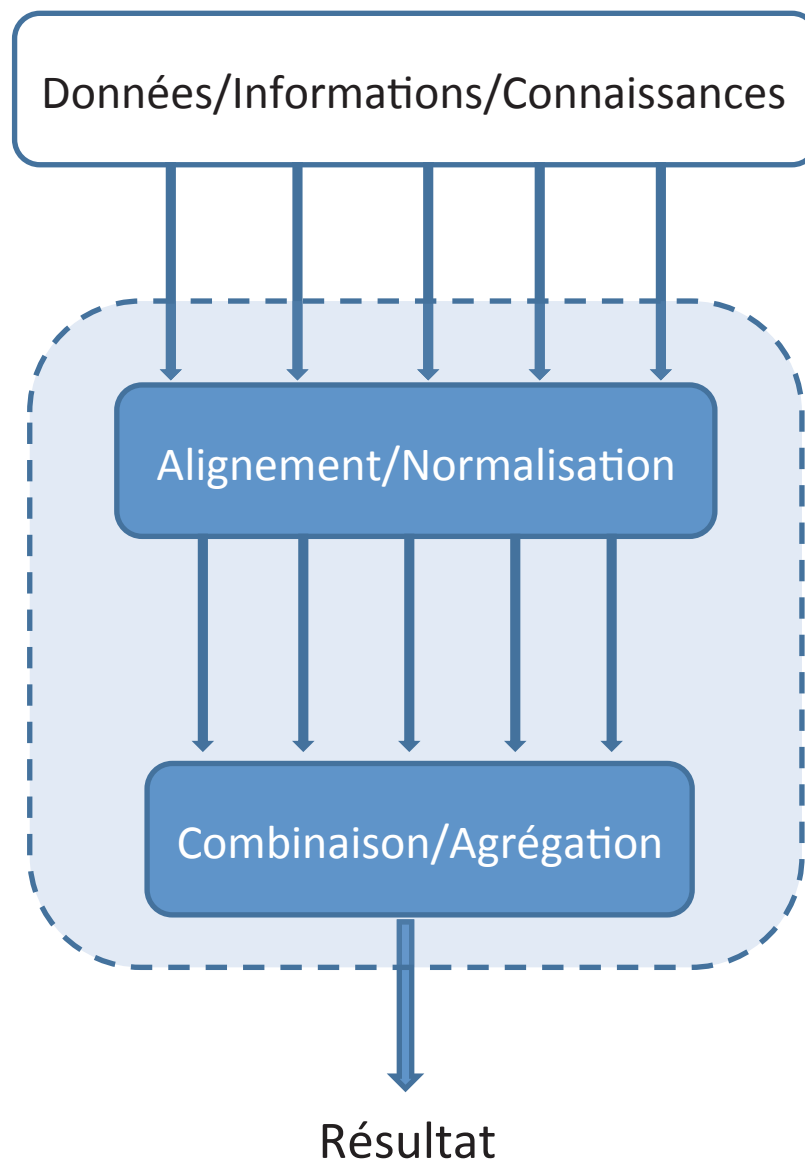


Figure 28 : Les deux étapes essentielles de la fusion.

Généralement le terme agrégation est très souvent utilisé à la place de fusion. Cependant, même si l'objectif dans les deux cas est d'obtenir un résultat final plus pertinent et plus robuste à partir d'un ensemble de données disponibles, le terme « agrégation » est plutôt utilisé pour les données numériques et/ou des objets (critères) ordonnés [Arif, 2005] [Dubois, 2004]. En outre [Torra, 2007] fait aussi la différenciation entre la fusion et l'agrégation en accordant le terme « agrégation » plutôt à des opérateurs mathématiques utilisés dans la fusion des données. Le domaine de la fusion est ainsi plus vaste et nécessite une étape de normalisation alors que l'agrégation peut être ramenée à une fonction de combinaison étant donné que les données en entrée sont prêtes à être combinées par une fonction mathématique.

Dans la suite, nous proposons de présenter les caractéristiques des outils d'agrégation les plus utilisées afin de bien choisir l'opérateur le plus adéquat à nos besoins. Ce choix doit tenir compte de plusieurs critères.

3.2. Critères de sélection de l'opérateur d'agrégation

Ces critères de sélection de l'opérateur d'agrégation doivent se définir à partir de la problématique BMCI et de sa généralisation. En ce sens, l'index doit représenter au mieux l'état de santé du système sur la base d'indicateurs en interaction et de poids différents. Par exemple, en considérant un moteur diesel avec « le dérèglement de l'injecteur » et « la consommation de carburant » comme indicateurs de santé, un « mauvais » résultat sur le premier indicateur (une avance élevée de l'injecteur) suppose un mauvais résultat sur le second (augmentation de la consommation). Une telle interaction doit être prise en compte. En outre, les indicateurs ne représentent pas nécessairement le même degré d'importance vis-à-vis de l'état de santé global. Ceci est relatif à la criticité et l'impact des mécanismes de dégradations que représentent ces indicateurs. De plus, la santé étant représentée par un spectre large d'aspects de performance et de dégradation, il est nécessaire pour l'agrégation de pouvoir refléter un raisonnement de compromis.

Les contraintes précédemment citées permettent d'isoler des critères de classification des opérateurs d'agrégation existants comme suit :

- Prise en compte par l'opérateur de la notion de compromis,
- Considération des poids des indicateurs,
- Considération des interactions entre les indicateurs à agréger.

3.3. Classification des outils d'agrégation

Les opérateurs d'agrégation transforment un nombre fini d'entrées réelles appartenant à un sous-domaine de \mathbb{R} en un résultat unique [Grabisch, 2011 a]. Ils sont appliqués dans des domaines diverses tels que : économie (élaboration d'index synthétiques), éducation (évaluation des élèves), informatique (aide à la décision)... [Torra, 2007]. Des exemples d'utilisation des opérateurs d'agrégation les plus répandus sont liés au domaine de l'aide à la décision multicritères [Beliakov, 2007] [Grabisch, 2003] [Montmain, 2004] [Berrah, 2006]. Cela permet de :

- Choisir une alternative parmi d'autres en se basant sur plusieurs critères (souvent conflictuels) qui servent à évaluer ces alternatives numériquement (par exemple à partir de fonctions d'utilité),

- Aider à la décision du groupe ; en groupant les évaluations d'un ensemble d'experts sur un objet ou plus.

Une fonction ou un opérateur d'agrégation est défini dans [Beliakov, 2007] comme étant une fonction mathématique qui prend en entrée des valeurs appelés des arguments et produit en sortie une autre valeur. Dans la suite, nous considérons $[0,1]$ comme domaine pour la fonction d'agrégation :

$$f: [0,1]^n \rightarrow [0,1]$$

$$(x_1, x_2, \dots, x_n) \rightarrow f(x_1, x_2, \dots, x_n) = y$$

Cette fonction doit vérifier a minima les propriétés [Detyniecki, 2001] [Grabisch, 2003]:

- Conditions aux extrêmes :

$$f(0,0, \dots, 0) = 0$$

$$f(1,1, \dots, 1) = 1$$

- Condition d'idempotence

$$f(x_i, x_i, \dots, x_i) = x_i$$

- Condition de monotonie

$$\forall (x_1, x_2, \dots, x_n), (z_1, z_2, \dots, z_n) \in [0,1]^n$$

$$\text{Si } (x_1, x_2, \dots, x_n) \leq (z_1, z_2, \dots, z_n)$$

$$\text{alors } f(x_1, x_2, \dots, x_n) \leq f(z_1, z_2, \dots, z_n)$$

D'autres propriétés peuvent compléter ces trois premières mais sans que celles-ci ne soient forcément nécessaires telles que la symétrie, la stabilité pour une fonction linéaire, l'existence d'un élément neutre ou absorbant. Plus de détails sur la vérification de ces propriétés sont donnés dans [Grabisch, 2003] [Gomez, 2004].

Les opérateurs d'agrégation sont nombreux et sont généralement classifiés en quatre grands groupes (Figure 29) [Torra, 2007] [Beliakov, 2007] :

- Les opérateurs de compromis dont le résultat est compris entre le minimum et le maximum des valeurs en entrée. La valeur agrégée est alors représentative de l'ensemble des valeurs en entrée.

$$\forall x = (x_1, x_2, \dots, x_n),$$

$$\min(x) \leq Agr(x) \leq \max(x)$$

Parmi ces opérateurs, les plus utilisés sont la moyenne arithmétique, la moyenne pondérée, les moyennes pondérées ordonnées (OWA), l'intégrale de Choquet, l'intégrale de Sugeno...

- Les opérateurs conjonctifs dont le comportement est dit « pessimiste », puisque la conjonction ne permet pas de compenser les valeurs les plus faibles par les autres.

$$Agr(x) \leq \min(x)$$

Parmi les opérateurs conjonctifs, il est possible de citer le minimum et les t-normes.

- Les opérateurs disjonctifs qui ont un comportement inverse à celui des opérateurs conjonctifs (« optimiste ») car la disjonction tient compte de l'entrée dont la valeur est la plus importante.

$$\max(x) \leq Agr(x)$$

Parmi les opérateurs disjonctifs, les fonctions maximum et t-conormes sont les plus utilisées.

- Les opérateurs mixtes

Une fonction d'agrégation est dite mixte si elle n'appartient pas à l'une des catégories ci-dessus, c'est à dire, elle présente différents types de comportements pour différentes parties du domaine.

Cette classification se base sur la nature du comportement décisionnel voulu. Par exemple, un comportement décisionnel pessimiste ou strict favorise le choix d'un opérateur du groupe conjonctif tel que les t-normes ou au contraire un comportement tolérant et optimiste amène vers les opérateurs disjonctifs de type t-conormes.

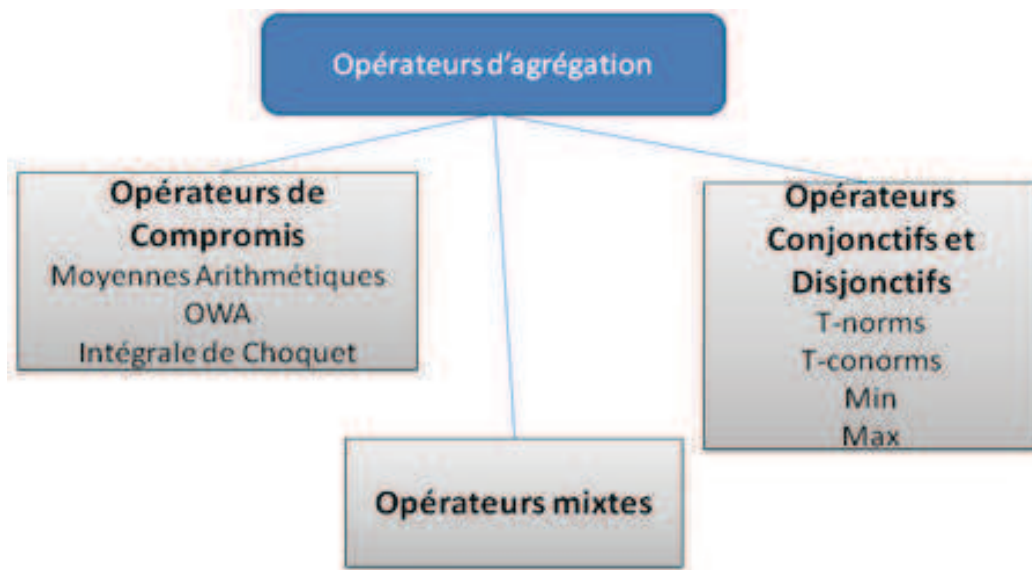


Figure 29 : Différents groupes d'opérateurs d'agrégation

Les opérateurs conjonctifs et disjonctifs sont utilisés afin de donner une grande importance aux cas extrêmes. Cependant, les opérateurs de compromis sont employés lorsqu'un comportement, ayant tendance à tenir compte de l'ensemble des critères et tout en ramenant l'évolution globale entre les évaluations extrêmes, est souhaité. Il découle ainsi de cette classification, que pour satisfaire le premier critère relatif au besoin de prise en compte de la notion de compromis comme décrit lors du paragraphe 3.2., il faut opter pour le groupe des opérateurs de compromis. Parmi les opérateurs les plus connus de ce groupe, une comparaison vis-à-vis des deux autres critères est effectuée afin d'orienter le choix de l'outil adéquat à l'agrégation des indicateurs en un index de santé.

Les opérateurs de compromis les plus simples sont ceux de la moyenne arithmétique ou de la moyenne pondérée. Le premier type considère un poids égal à 1 pour toutes les variables en entrée. Pour la moyenne pondérée un poids différent est accordé à chaque variable en entrée (la somme des poids est égale à 1). Cependant, ces deux opérateurs ne prennent pas en compte des interactions entre les variables à agréger. De la même manière, le groupe des moyennes ordonnées ne permet pas de représenter la dépendance entre les variables en entrée de l'opérateur d'agrégation. Cependant, une autre famille d'opérateurs de compromis constituée des intégrales de Choquet et de Sugeno utilise pour agréger des variables d'entrée, des paramètres reflétant non seulement les poids de ces variables prises individuellement mais aussi le degré d'interaction qui peut exister entre elles. Ces paramètres se présentent sous forme de mesures floues (ou capacités) remplaçant la notion des coefficients de pondérations pour les moyennes classiques. Le Tableau 2 résume l'ensemble de ces considérations.

Groupes de méthodes	Prise en compte des poids des indicateurs	Prise en compte des interactions entre les indicateurs
Moyenne classiques : - Moyenne arithmétique, - Moyenne pondérée	Oui	Non
Moyenne ordonnées : - OWA, - WOWA	Oui	Non
Intégrale de Choquet	Oui	Oui
Intégrale de Sugeno	Oui	Oui

Tableau 2 : Comparaison entre opérateurs d'agrégation par rapport aux deux critères relatifs aux poids et interactions des indicateurs

Il découle de ce tableau que parmi les opérateurs d'agrégation de compromis, l'intégrale de Choquet et l'intégrale de Sugeno sont ceux qui permettent de satisfaire les deux critères de considération des poids et des interactions entre les indicateurs à agréger.

Ces deux intégrales ont d'autres propriétés intéressantes pour notre cas d'utilisation et notamment celle de la continuité [Marichal, 2000 b] [Grabisch, 2011 a] [Grabisch, 2011 b]. De cette manière, une légère variation dans les valeurs des indicateurs en entrée ne produit pas un changement discontinu de la sortie. Cette propriété est une garantie pour une agrégation robuste, cohérente et avec un comportement non chaotique [Gomez, 2004]. Il est intéressant aussi de signaler que l'intégrale de Choquet généralise la moyenne pondérée et l'opérateur OWA, tandis que l'intégrale de Sugeno généralise les opérateurs minimum pondéré et maximum pondéré [Beliakov, 2007].

Cependant, l'intégrale de Sugeno est plus appropriée pour l'agrégation ordinale (où seul l'ordre des éléments est important), tandis que l'intégrale de Choquet est adaptée pour l'agrégation cardinale (où la distance entre les nombres a un sens) [Gomez, 2004]. Etant donné que nous proposons d'agréger des degrés de déviation par rapport aux modes nominaux des différents indicateurs, l'opérateur choisi doit se baser alors sur l'existence d'une échelle numérique où l'addition et la multiplication sont applicables et significatives. C'est pourquoi nous avons opté pour l'intégrale de Choquet.

3.4. Intégrale de Choquet

Le concept de mesures et d'intégrales floues a été proposé en 1974 dans [Sugeno, 1974]. Depuis, la théorie a été développée par la communauté de la logique floue. La mesure floue est une extension de la mesure classique dans le sens où l'additivité de la mesure classique est remplacée par une condition plus faible, la monotonie. La non additivité est la caractéristique principale de la mesure floue, (elle est aussi appelée une mesure non additive). L'intégrale floue est un terme général désignant les intégrales par rapport à la mesure floue. Il existe plusieurs types d'intégrales floues. L'intégrale non additive la plus connue est l'intégrale de Choquet [Murofushi, 1991].

La particularité d'une intégrale floue est son aptitude à représenter certains types d'interactions entre les critères, tel que la redondance (interaction négative) et la synergie (interaction positive). L'intégrale de Choquet généralise en plus des opérateurs d'agrégation de compromis : « Ordered Weighting Mean » (OWA) et la moyenne pondérée.

3.4.1. Mesure floue

La plupart des applications réelles de mesures floues portent sur les problèmes de décision multicritères, où elles sont utilisées en conjonction avec les intégrales floues. Les mesures floues (appelées aussi des capacités) sont définies sur un ensemble (fini) de critères, et modélisent l'importance relative de ces critères ainsi que leurs interactions mutuelles [Grabisch, 1995]. Il est important de mentionner que dans un contexte d'aide à la décision, l'agrégation porte sur des critères, mais ici, les critères se réfèrent aux indicateurs normalisés.

Soit $\{X\}: \{x_1, \dots, x_n\}$ un ensemble d'indicateurs normalisés, une capacité (ou une mesure floue) $\mu : \wp(X) \rightarrow [0,1]$ pour cet ensemble est une application vérifiant:

$$\begin{cases} \mu(\emptyset) = 0 \\ \mu(X) = 1 \\ \mu(A) \leq \mu(B), \forall A \subseteq B \text{ et } \forall B \subseteq X \end{cases} \quad (2.1)$$

Pour l'intégrale de Choquet, les coefficients de la capacité peuvent être interprétés de la même manière que les poids d'une moyenne pondérée. En effet, $\mu(A)$ représente le degré d'importance, ou le poids, de l'ensemble $A \subseteq X$ dans le calcul de la valeur globale de X . Ainsi, lorsque $A = \emptyset$, leur importance est 0 et l'importance maximale est obtenue lorsque tous ces indicateurs sont considérés (et l'importance maximale est égale à 1). En outre, la condition de monotonie dans la définition d'une capacité signifie que plus on considère d'indicateurs, plus grande est leur importance relative.

Afin de mieux expliciter cette notion fondamentale de mesure floue, un exemple souvent utilisé sur le calcul de la moyenne des notes dans différentes matières [Beliakov, 2007] est présenté ci-dessous.

Des élèves sont évalués en fonction de leur niveau en mathématiques (M), physique (P), et en littérature (L) à travers une moyenne finale des notes des trois matières. L'évaluation finale dépend de l'importance donnée à chaque sous ensemble de matières. Pour régler ces importances, une mesure floue est utilisée. Ici, X est l'ensemble de tous les sujets ($X = \{M, P, L\}$), et $\mu(A)$ est l'importance d'un ensemble particulier de matières A . La définition de la mesure prend en compte les éléments suivants.

- Conditions aux limites:

$$\mu(\emptyset) = 0, \mu(\{M, P, L\}) = 1$$

- L'importance relative des matières scientifiques par rapport à la littérature:

$$\mu(\{M\}) = \mu(\{P\}) = 0,35, \mu(\{L\}) = 0,3$$

- L'importance des mathématiques et de la physique est plus importante que celle de la littérature.

- Redondance entre les mathématiques et la physique:

$$\mu(\{M, P\}) = 0,5 < \mu(\{M\}) + \mu(\{P\})$$

Les mathématiques et la physique sont des sujets similaires. L'importance de l'ensemble des deux sujets ne devrait pas être plus grande que l'addition des importances des deux sujets séparément.

- Synergie entre les sujets de la littérature et scientifique:

$$\mu(\{M, L\}) = 0,85 > \mu(\{M\}) + \mu(\{L\}) = 0,35 + 0,3 = 0,65$$

$$\mu(\{P, L\}) = 0,85 > \mu(\{P\}) + \mu(\{L\}) = 0,35 + 0,3 = 0,65$$

Les mathématiques et la littérature sont des sujets complémentaires. Deux notes élevées en ces matières sont souhaitables, ainsi l'importance de l'ensemble $\{M, L\}$ doit être supérieure à la somme des importances des deux sujets. Le même principe s'applique pour la littérature et la physique.

Les coefficients de la capacité (ou mesure floue) sont nécessaires afin de paramétrer l'intégrale de Choquet comme le montre le paragraphe suivant.

3.4.2. Intégrale de Choquet : définition

L'intégrale de Choquet par rapport à cette capacité est définie comme suit :

$$C_{\mu}(x_1, \dots, x_n) = \sum_{i=1}^n (x_{(i)} - x_{(i-1)})\mu(A_{(i)}) \quad (2.2)$$

Où $(.)$ en indice est l'opérateur de permutation tel que:

$$x_{(1)} \leq \dots \leq x_{(n)} \text{ et } A_{(i)}: \{x_{(i)}, \dots, x_{(n)}\} \text{ avec } x_{(0)} = 0$$

L'interprétation géométrique de cette expression a été donnée par [Murofishi, 1991] comme le montre la Figure 30. Il est intéressant de noter ici que dans cette figure $a_i \equiv x_{(i)}$ et $A_{(i)} \equiv \{x|f(x) \geq a_i\}$.

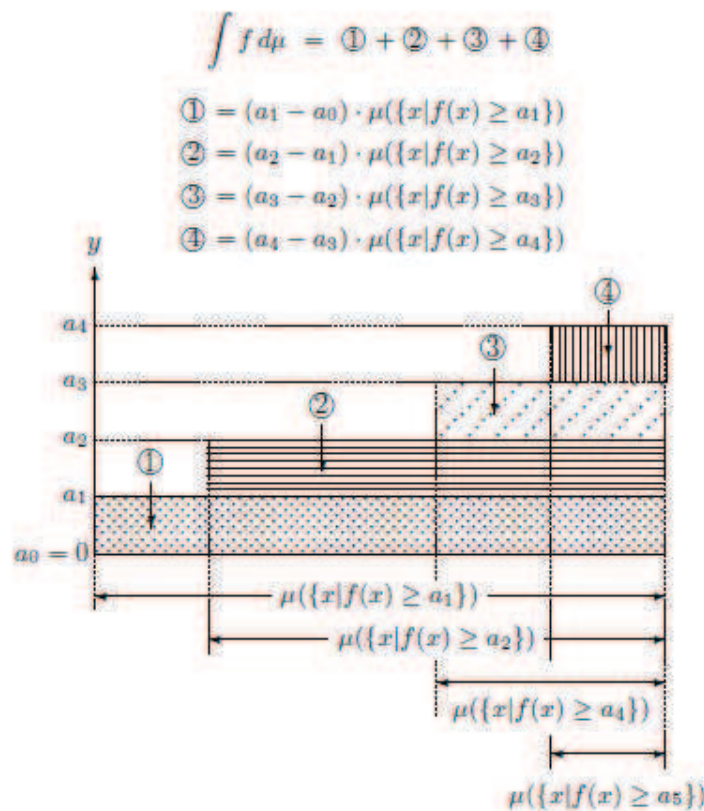


Figure 30 : Interprétation graphique de l'intégrale de Choquet [Murofishi, 1991].

Afin d'expliciter le comportement de cette intégrale, un exemple simple de calcul des moyennes de quatre élèves est présentée. Les coefficients de la capacité utilisés sont ceux de l'exemple du paragraphe précédent (Tableau 3).

$\mu(\{M\}) = 0.35$ $\mu(\{P\}) = 0.35$ $\mu(\{L\}) = 0.3$ $\mu(\{M, P\}) = 0.5$	$\mu(\{M, L\}) = 0.85$ $\mu(\{P, L\}) = 0.85$ $\mu(\{M, P, L\}) = 1$
---	--

Tableau 3 : Liste des coefficients de la capacité pour le calcul du score global des élèves

Une comparaison des résultats de cet opérateur avec ceux donnés par la moyenne pondérée avec pour coefficients de pondération ($\omega_M = \omega_P = 0.35$ et $\omega_L = 0.3$) est effectuée à travers le Tableau 4.

Elève	Mathématique	Physique	Littérature	Intégrale de Choquet	Moyenne Pondérée
1	18	17	15	16.35	16.75
2	18	17	12	14.85	15.85
3	14	13	15	14.15	13.95
4	14	13	12	12.85	13.05

Tableau 4 : Différentes situations pour le calcul des moyennes par Intégrale de Choquet et la moyenne pondérée.

Cet exemple illustratif montre comment l'interaction de redondance entre les deux matières scientifique fait diminuer leur importance vis-à-vis des notes de la littérature et ce par rapport à la moyenne pondérée. Ainsi, les moyennes calculées par l'intégrale de Choquet ont tendance à tenir compte davantage des notes de la littérature étant donné que cette matière est en synergie avec les deux autres matières. Son importance globale dans le calcul de la moyenne se trouve augmentée.

3.4.3. Eléments complémentaires

Par contre, les expressions des degrés d'importance et d'interaction totaux sont obtenues en considérant simultanément tous les coefficients de capacités des sous-ensembles d'indicateurs en lien avec les indicateurs évalués. Ainsi, la représentation Möbius de la capacité $\mu(\cdot)$ est nécessaire [Grabisch, 2007]. Elle est définie comme suit :

$$\forall B \subseteq X, m_\mu(B) = \sum_{A \subseteq B} (-1)^{b-a} \mu(A)$$

Où a et b représentent respectivement les cardinalités des sous ensembles A et B. Avec cette représentation, les coefficients de la capacité pour les couples d'indicateurs reflètent les interactions entre eux de la manière suivante :

- Pas d'interaction entre les deux indicateurs i et j: $m_\mu(\{x_i, x_j\}) = 0$,
- Une interaction de synergie entre i et j: $m_\mu(\{x_i, x_j\}) > 0$,
- Une redondance entre i et j: $m_\mu(\{x_i, x_j\}) < 0$.

L'importance d'un critère ne peut pas être quantifiée que par le coefficient du singleton lui-même. Il faut considérer toutes les interactions entre ce critère et les autres. Il a été proposé à cet effet d'utiliser l'indice d'importance de Shapley (ou valeur de Shapley) qui tient compte de l'importance moyenne du critère par rapport à toutes les interactions de celui-ci avec le reste des critères [Roth, 1988]:

$$\phi_{m_\mu}(\{x_i\}) = \sum_{A \subseteq X \setminus \{x_i\}} \frac{1}{a+1} m_\mu(A \cup \{x_i\}) \quad (2.3)$$

De plus, il est possible d'exprimer le degré d'interaction entre deux critères en considérant toutes leurs interactions avec le reste des critères comme suit [Grabisch, 1997 b]:

$$I_{m_\mu}(\{x_i, x_j\}) = \sum_{A \subseteq X \setminus \{x_i, x_j\}} \frac{1}{a+1} m_\mu(A \cup \{x_i, x_j\}) \quad (2.4)$$

Après cette étape de sélection de l'opérateur d'agrégation, la détermination des paramètres (les coefficients de la capacité pour le cas de l'intégrale de Choquet) est l'étape suivante.

3.4.4. Identification des capacités

Cette étape d'identification est très importante car toutes les connaissances concernant les critères sont intégrées dans la capacité. Le nombre des coefficients de la capacité augmente de façon exponentielle en fonction du nombre de critères ($2^n - 2$ sous-ensembles de critères). Par conséquent, pour un grand nombre de critères, il devient pratiquement impossible pour un décideur de fournir les données nécessaires à l'identification de la mesure floue. Par conséquent, diverses méthodes d'identification des mesures floues ont été proposées :

- Les méthodes basées sur les connaissances de l'expert :

L'expert fournit directement les coefficients requis ou encore des informations pertinentes qui sont ensuite utilisées pour la détermination des paramètres [Grabisch, 2007].

- Les approches basées sur les données :

Ces approches sont aussi appelées les méthodes d'optimisation ou d'apprentissage. La fonction de coût utilisée dans cette étape est de minimiser la différence entre les résultats globaux calculés au moyen de l'intégrale de Choquet et les scores globaux souhaités fournis par le décideur pour les situations observées. Les approches proposées pour résoudre ce type de problème sont basées sur les moindres carrés [Grabisch, 1995]. Une autre alternative consiste à considérer l'ordre total ou partiel dans un classement des situations disponibles. L'approche d'optimisation proposée (par programmation linéaire) consiste à maximiser la différence minimale entre les résultats obtenus à l'aide de l'intégrale de Choquet des situations classées [Marichal, 2000 a].

- Les approches mixtes :

Le troisième groupe de méthodes combine des méthodes basées sur des connaissances de l'expert et des données d'apprentissage. Dans ce troisième groupe de méthodes, des considérations sémantiques peuvent être utilisées pour aider à réduire la complexité dans le problème d'optimisation. [Grabisch, 2007] examine les méthodes de résolution pour ce troisième type de méthodes. Dans ce cas, la connaissance complémentaire fournie est sous forme de classement ou quantification de l'importance des critères, de leurs interactions, ainsi que d'autres contraintes de comportement de certains critères.

- Par simplification de la mesure floue :

Des approches par simplification de la nature de la capacité elle-même ont été aussi proposées. Dans ce cas des catégories de capacités spécifiques définies par moins de paramètres sont adoptées comme les capacités k-additives [Grabisch, 1997 a], les capacités hiérarchiquement décomposables [Domingo, 2002], les mesures Sugeno lambda [Liu, 2006] [Larbani, 2011]...

Mais bien que ces types de capacités fassent réduire la complexité liée au nombre de coefficient à identifier, elles ont été définies pour un certain contexte d'application et ne peuvent pas être facilement généralisables.

- Les approches non supervisées :

Par ailleurs, pour certains problèmes de décision extrêmement complexes, les connaissances requises pour l'identification de mesures floues ne peuvent pas être facilement données par un décideur. Pour résoudre ce genre de problème, Kojadinovic [Kojadinovic, 2002], [Kojadinovic, 2004] a proposé une méthode d'identification non supervisée basée sur l'estimation des coefficients de la capacité à l'aide d'une mesure de quantité d'information, l'entropie. L'approche consiste essentiellement à remplacer la notion assez subjective d'importance d'un sous-ensemble de critères par la notion probabiliste de contenu d'informations d'un sous-ensemble de critères. En contexte d'aide à la décision multi-critères, cette approche permet de définir les poids des sous-ensembles non vides d'attributs (ou critères) par l'intermédiaire d'une mesure d'entropie, permettant d'estimer l'importance du contenu d'information probabiliste d'un sous-ensemble d'attributs à partir d'un ensemble de profils ou de situations d'apprentissage suffisamment représentatif.

Chaque critère x_i ($1 \leq i \leq m$) est associé à une variable aléatoire X_i . On considère aussi que ces variables aléatoires X_1, \dots, X_m sont discrètes et que leurs valeurs sont finies. Soit

$p(X_1, \dots, X_m)$ la distribution de probabilités jointes des valeurs prises par ces critères (différente de la masse de Dirac).

$$\mu(S) := \begin{cases} 0, & \text{if } S = \emptyset, \\ \frac{H(p(X_{i_1}, \dots, X_{i_s}))}{H(p(X_1, \dots, X_m))}, & \text{if } S = \{x_{i_1}, \dots, x_{i_s}\}, \end{cases}$$

$H(p(X_1, \dots, X_m)) \neq 0$ puisque $p(X_1, \dots, X_m)$ est différente de la masse de Dirac, p est approchée par l'estimateur classique de maximum de vraisemblance \hat{p} à partir d'un échantillon de profils. Ainsi,

$$\hat{\mu}(S) = \frac{H(\hat{p}(X_{i_1}, \dots, X_{i_s}))}{H(\hat{p}(X_1, \dots, X_m))}, \forall S \subseteq \{x_1, \dots, x_m\}$$

Un estimateur naturel de l'intégrale de Choquet par rapport à μ est obtenu simplement par substitution [Kojadinovic, 2002] (les vérifications axiomatiques sont fournies dans cette même référence) :

$$\hat{C}_\mu = C_{\hat{\mu}}$$

Le poids d'un sous-ensemble non vide de critères dépend directement de l'uniformité de la distribution de probabilité estimée correspondante. En ce sens, si un sous ensemble de critères montre des variations importantes de ses valeurs prises entre les différentes instances de l'échantillon, la probabilité jointe estimée relative à ce sous ensemble de critères est plus uniforme. Par conséquent, leur poids est plus élevé.

Cette méthode est à la base de plusieurs travaux d'identification des capacités de l'intégrale de Choquet. En effet, la méthode proposée par Kojadinovic à base d'entropie ne permet d'extraire que les poids et les relations de type redondance entre les critères et pas les relations de type « synergie ». Afin de palier à cette limite, [Ramasso, 2011], [Jullien, 2008] proposent de remplacer la fonction d'entropie par celle de l'entropie relative afin de pouvoir extraire les relations de synergie entre les critères à agréger (ce qui n'est pas possible par l'entropie). En ce sens, ils utilisent cette approche pour identifier les coefficients de la capacité pour l'agrégation des caractéristiques dans une problématique de classification.

L'estimation de la capacité via la mesure d'entropie relative passe par l'estimation de la distribution jointe d'un ensemble de critères à partir d'un ensemble représentatif de situations (appelée aussi échantillon, données d'apprentissage...). L'estimateur de maximum de vraisemblance est utilisé.

Pour un problème de classification, chaque coefficient de la capacité $\mu_k(S)$ (où pour chaque classe ω_k une capacité μ_k est à identifier) exprime l'importance relative du sous ensemble de

caractéristiques S dans la distinction de la classe ω_k parmi les autres classes. Ainsi deux distributions sont utilisées :

- P_k^S : la distribution de probabilité du degré d'appartenance à la classe ω_k étant donné que réellement il s'agit de la classe ω_k
- P_k^N : la distribution de probabilité du degré d'appartenance à la classe ω_k étant donné que réellement il s'agit des autres classes

$$\mu_k(S) = \frac{H_R(P_k^S)}{H_R(P_k^N)}$$

$$\text{Avec } H_R(P_k^S) = \sum P_k^S * \log\left(\frac{P_k^S}{P_k^N}\right)$$

Les propriétés de la capacité μ_k issue de la mesure d'entropie relative ont été vérifiées dans [Jullien, 2008]. Cette approche à base d'entropie relative permet non seulement l'identification des interactions de type « redondance », mais aussi celles de type « synergie » entre les critères. En ce sens, en appliquant cette approche, aucune connaissance experte n'est nécessaire pour identifier tous les coefficients de la capacité. Cependant, cette approche a été appliquée pour une problématique de classification mettant en œuvre plusieurs classes à distinguer. Il reste donc à adapter cette technique pour un contexte d'agrégation des indicateurs de santé reflétant le degré d'écart par rapport au mode nominal. En outre, des systèmes multi-composants ont des bilans de santé multi-niveaux et par conséquent un grand nombre de capacités à identifier. Etant donné que cette approche nécessite un lot de données d'apprentissage de taille importante, il devient intéressant de profiter des relations hiérarchiques qui existent entre les indicateurs de niveaux différents pour compléter cette technique d'apprentissage non supervisé par une approche d'inférence des capacités entre les niveaux hiérarchiques. De cette manière, l'effort d'apprentissage est minimisé en l'appliquant uniquement au niveau le plus bas.

L'ensemble de ces adaptations nécessaires à la mise en œuvre de l'intégrale de Choquet pour l'agrégation des indicateurs des bilans de santé pour des systèmes multi-composants est abordé dans le chapitre suivant.

4. Conclusion

Par rapport aux activités clés de notre proposition, ce chapitre porte sur un état de l'art concernant les outils/méthodes en support de la détection conditionnelle des anomalies et la fusion des indicateurs (l'activité de normalisation étant considérée comme intégrée à la détection). Ces outils/méthodes sont globalement évalués et comparés en regard des besoins émergents de BMCI mais aussi de leur généralisation. Un premier constat souligne l'adéquation des approches de détection à base d'histogrammes pour une quantification

génériques des anomalies. Il reste à tirer profit de cet avantage de l'outil histogramme afin de construire une approche adaptée à notre contexte de détection-normalisation.

De façon duale, la seconde partie de ce chapitre s'intéresse, toujours sur la base d'une évaluation et comparaison d'outils/méthodes, à choisir l'outil d'agrégation des indicateurs le plus adéquat parmi ceux utilisés dans le domaine de la fusion d'informations. L'intégrale de Choquet semble être un opérateur adéquat à notre contexte puisque c'est un opérateur de compromis permettant de tenir compte des interactions entre les indicateurs à agréger. Cependant l'apprentissage de ses paramètres reste à adapter pour le cas d'un bilan de santé multi niveaux hiérarchiques.

Ces deux problématiques d'adaptation font l'objet de nos contributions scientifiques qui sont développées dans le chapitre suivant.

Chapitre 3 : Bilan de santé : formalisation et proposition d'outils support à son élaboration

1. Introduction

Par rapport aux deux constats et orientations scientifiques dégagés à la fois sur la détection conditionnelle et l'agrégation, le chapitre 3 a pour objet de présenter la contribution principale de notre travail. En ce sens, une nouvelle approche de détection conditionnelle des anomalies basée sur la partition de l'espace contextuel et la mesure de la déviation de l'observation actuelle par rapport à la référence de normalité est proposée. Cette approche utilise l'entropie relative pour quantifier la déviation de l'état actuel par rapport à des états de référence (nominal et de faute). Les états sont représentés par les distributions des valeurs prises par chaque indicateur sur une fenêtre de temps. La détection a pour intérêt non seulement de quantifier les dérives des indicateurs mais aussi de normaliser l'ensemble des indicateurs dans un vecteur pour permettre leur agrégation dans une étape suivante.

L'agrégation est quant à elle abordée à différents niveaux d'abstraction du système multi-composants. L'opérateur d'agrégation choisi est l'intégrale de Choquet. L'intérêt de cet outil est de prendre en compte des interactions existantes entre des indicateurs à agréger. Cependant, cet opérateur nécessite un nombre important de coefficients de capacité à identifier. Pour palier cette contrainte, nous proposons suivant le niveau hiérarchique considéré:

- Un apprentissage non supervisé de ces coefficients pour le niveau le plus bas dans l'abstraction du système, à savoir celui des composants,
- Une déduction des coefficients des capacités pour les éléments des niveaux supérieurs à partir des coefficients relatifs à leurs constituants (i.e. des éléments appartenant au niveau immédiatement inférieur).

Ainsi, seul le niveau le plus bas nécessite un apprentissage des capacités à partir de données de fonctionnement, pour les autres niveaux, les capacités se déduisent alors de celles du premier. L'approche proposée peut être déployée pour différents cas industriels grâce à sa généralité.

Les notations qui formalisent les différents aspects d'une représentation générique d'un bilan de santé multi-niveaux font l'objet de la première partie de ce chapitre. Ensuite, la deuxième partie est consacrée à la présentation de la nouvelle approche de détection conditionnelle/normalisation proposée. Enfin, la troisième partie décrit la méthodologie d'obtention des capacités pour l'intégrale de Choquet. Les deux phases proposées sont ainsi décrites : l'apprentissage non supervisé de capacités utilisées au niveau des composants, et l'inférence ascendante des capacités pour les éléments des niveaux supérieurs dans la hiérarchie du système.

2. Formalisation d'un bilan de santé système

Après avoir légitimé la notion de bilan de santé à différents niveaux pour un système multi-composants, nous souhaitons apporter une notation formelle pour une expression homogène des informations portées dans les éléments d'un tel bilan ainsi que les relations entre les bilans de santé. Par souci de généralité, la notation décrivant un bilan de santé doit pouvoir être adaptée à tout type de systèmes et également à chaque niveau d'abstraction. Cette formalisation permet de fixer les notations liées au bilan de santé et sert de base pour présenter nos contributions.

En résumant les fondements liés au bilan de santé et apportés dans le premier chapitre, nous adoptons une présentation du bilan de santé basée sur la décomposition fonctionnelle du système, c'est-à-dire, associer un ensemble d'indicateurs et un index de santé à chaque fonction du système (en allant de la fonction la plus élémentaire supportée par un composant jusqu'à la fonction globale réalisée par l'ensemble du système) (Figure 31).

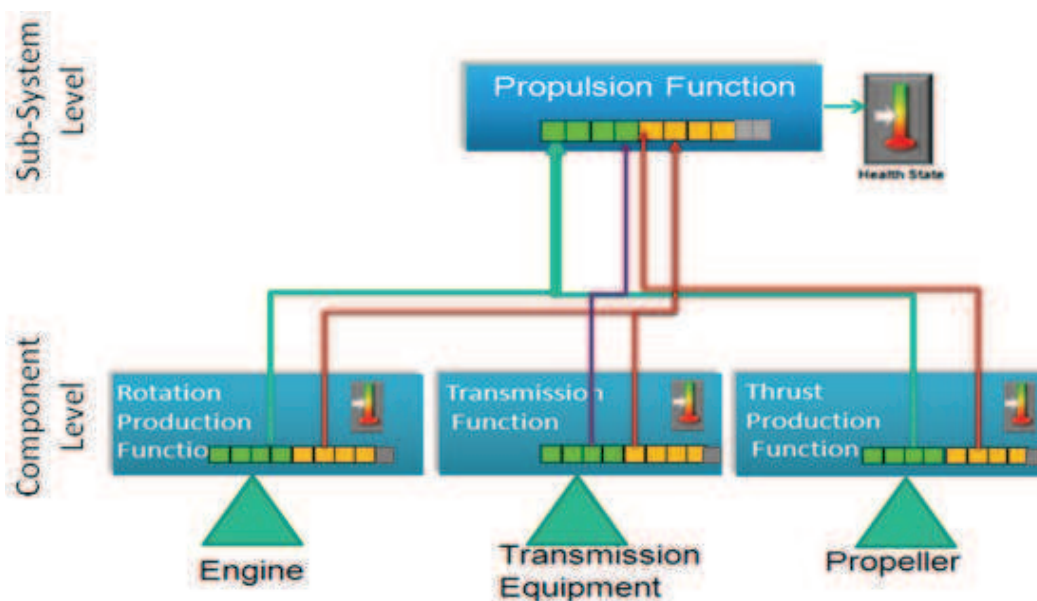


Figure 31 : Architecture fonctionnelle sur laquelle se base la présentation du bilan de santé.

En effet, la structure fonctionnelle des systèmes est un moyen générique pour définir la performance liée à chaque fonction isolée et également toutes les caractéristiques de dégradation qui pourraient apparaître sur le composant, sous-système ou système supportant ces fonctions.

Nous rappelons aussi que les indicateurs de dégradation ou de performance dans un bilan de santé sont issus d'une fonction de traitement des données (filtrage, extraction des caractéristiques de santé...). Le degré de déviation des valeurs prises par ces indicateurs est ensuite évalué par la détection/normalisation vis-à-vis du contexte dans lequel évolue le système surveillé. Les indicateurs après cette étape de détection peuvent servir à une analyse détaillée ou être agrégés en un index de santé global (Figure 32).

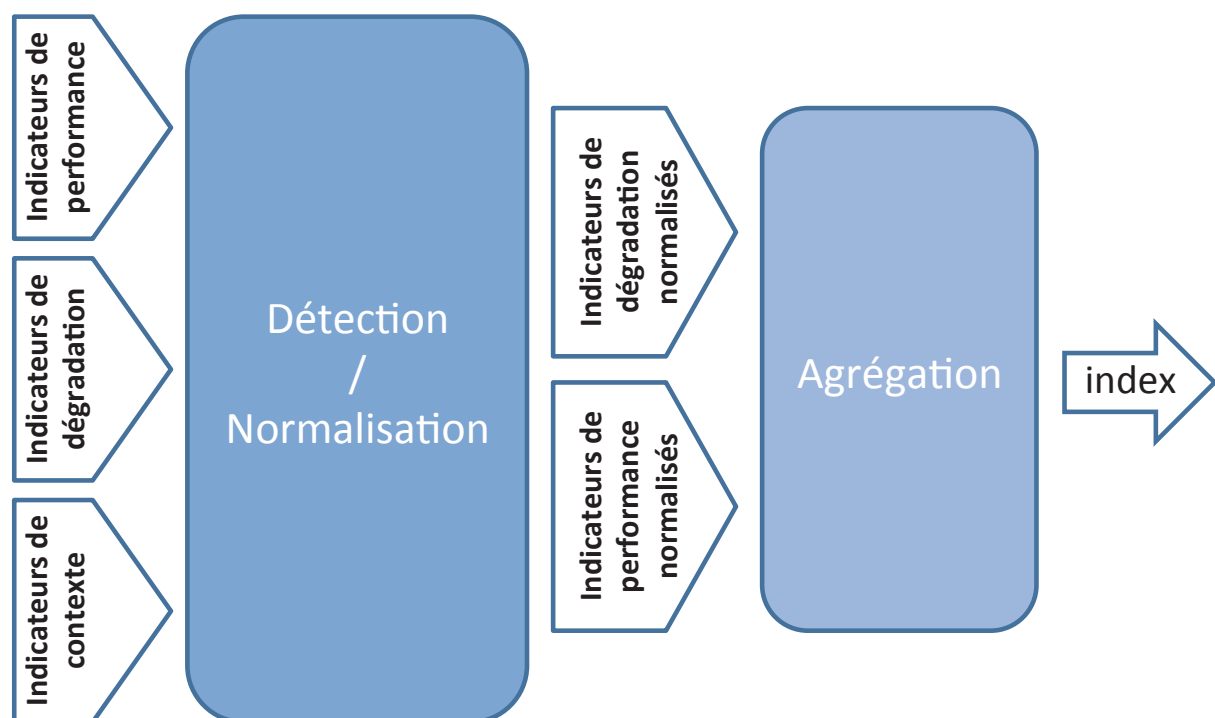


Figure 32 : Etapes de construction d'un index de santé à partir de l'ensemble des indicateurs

2.1. Indicateurs et index pour un élément

Depuis le niveau composant jusqu'au système global, nous appelons un élément E une fonction avec son support dans la structure. Notons l'ensemble des indicateurs de santé (avant détection/normalisation) attaché à un élément E au sein de la structure fonctionnelle d'un système industriel (composant / sous-système / système) par:

$$\mathcal{V}(E) \triangleq \{I_1^P, I_2^P, \dots, I_{n_P}^P\} \cup \{I_1^D, I_2^D, \dots, I_{n_D}^D\} \cup \{I_1^C, I_2^C, \dots, I_{n_C}^C\}$$

Où:

- $\mathcal{V}(E)$ désigne l'ensemble des indicateurs pour un élément quelconque E du système,
- I_j^P est le $j^{\text{ème}}$ indicateur de performance ($j = 1, \dots, np_E$), np_E est le nombre des indicateurs de performance affectés à l'élément en question,
- I_j^D est le $j^{\text{ème}}$ indicateur de dégradation ($j = 1, \dots, nd_E$), nd_E est le nombre des indicateurs de dégradation affectés à cet élément,
- I_j^C est le $j^{\text{ème}}$ indicateur de contexte (opérationnel et/ou environnemental) ($j = 1, \dots, nc_E$), nc_E est le nombre des indicateurs contextuels affectés à cet élément.

Les indicateurs de performance permettent de surveiller la fonction accomplie par l'élément adressé à tous les niveaux de la structure fonctionnelle du système. En effet, un dysfonctionnement peut être révélé par des performances en dessous des attentes conventionnelles (correspondantes aux valeurs nominales) [Zille, 2009]. Ces indicateurs sont principalement axés sur la représentation, d'une part, des flux (d'énergie ou d'information) de la fonction, et d'autre part, de ses performances. Nous trouvons dans cette deuxième catégorie en particulier: les indicateurs représentant l'efficacité (rapport des résultats aux objectifs), et ceux reflétant l'efficacité (rapport entre les résultats et les ressources engagées). Ainsi, les indicateurs de performance sont élaborés en fonction des propriétés des flux d'entrée et de sortie de la fonction considérée et/ou plusieurs données synthétiques (efficacité énergétique, l'écart par rapport aux objectifs).

Les indicateurs de dégradation d'un système sont essentiellement liés à l'évolution des mécanismes de dégradation qui peuvent l'affecter. Ce type d'indicateurs traduit :

- Le mécanisme de dégradation lui-même (usure, fissure, fuite ...). Les indicateurs sont construits à partir des paramètres physiques caractéristiques du composant supportant la fonction suivie,
- Les symptômes observables révélateurs de l'apparition de ces mécanismes (hausse de température, vibrations, caractéristiques acoustiques...),
- Les facteurs externes de dégradation, s'ils existent et peuvent être surveillés (comme les chocs subits par le système, contraintes mécaniques d'usage...).

Les indicateurs contextuels permettent de situer les valeurs prises par les indicateurs de performance ou de dégradation par rapport aux conditions dans lesquelles le fonctionnement du système a lieu. Ceci a pour objectif de considérer toute variation des valeurs de ces indicateurs, résultant d'un changement des conditions opérationnelles ou environnementales, comme étant normale. De cette façon, seules les dérives des valeurs des indicateurs de performance ou de dégradation causées par des anomalies sont effectivement prises en compte.

La phase de détection/normalisation permet de transformer les indicateurs de performance $\{I_1^P, I_2^P, \dots, I_{n_{pE}}^P\}$ et de dégradation $\{I_1^D, I_2^D, \dots, I_{n_{dE}}^D\}$ afin de quantifier la déviation vis-à-vis de la nominalité et ainsi les normaliser sur une même échelle sémantique (déviation par rapport à la nominalité) et numérique dans l'intervalle $[0,1]$. La fonction de détection/normalisation notée $f_{Dét}$ est générique et permet l'intégration des indicateurs contextuels (opérationnels ou environnementaux) $\{I_1^C, I_2^C, \dots, I_{n_{cE}}^C\}$ qui peuvent affecter les valeurs des autres indicateurs.

Notons alors :

$$i_j = f_{Dét}(I_j, \{I_1^C, I_2^C, \dots, I_{n_{cE}}^C\})$$

Où i_j est le $j^{ème}$ indicateur normalisé correspondant à l'indicateur $I_j \in \{I_1^P, I_2^P, \dots, I_{n_{pE}}^P\} \cup \{I_1^D, I_2^D, \dots, I_{n_{dE}}^D\}$ (de performance ou de dégradation)

Après la phase de détection/normalisation des indicateurs de performance et de dégradation, l'index de santé pour l'élément E est calculé par agrégation des indicateurs normalisés $\{i_1^P, i_2^P, \dots, i_{n_{pE}}^P\} \cup \{i_1^D, i_2^D, \dots, i_{n_{dE}}^D\}$. Pour l'agrégation nous utilisons l'intégrale de Choquet. Cet index composite vient compléter cette vision distribuée/hétérogène/complexe du bilan de santé pour en produire une vision synthétique directement/simplement interprétable par le décideur. Notons alors :

$$IS_E = Agr(i_1, \dots, i_{n_{iE}})$$

Où $Agr: [0,1]^{n_E} \rightarrow [0,1]$ est la fonction/opérateur d'agrégation appliqué et n_{iE} le nombre total d'indicateurs normalisés pour l'élément E.

2.2. Relation hiérarchique entre les indicateurs

Comme présenté dans le chapitre 1 (Figure 12), les indicateurs d'un élément E sont calculés à partir de ceux de ses sous-éléments. En effet, les indicateurs du niveau le plus bas, i.e. ceux des composants, sont élaborés à partir de mesures brutes. Nous formalisons dans cette partie la relation existante entre les indicateurs d'un élément E et les indicateurs de ses n_E sous-éléments, notés E_k , par la fonction g_k relative à chaque élément E_k définie comme suit :

$$\begin{aligned} \forall k = 1 \dots n_E \\ g_k: \mathcal{V}(E) \rightarrow \wp(\mathcal{V}(E_k)) \\ \forall I \in \mathcal{V}(E), g_k(I) = \{I_j \in \mathcal{V}(E_k) \mid I_j \text{ est contributeur de } I\} \end{aligned}$$

Les indicateurs de performance à ces niveaux sont exprimés de la manière suivante :

$$I_j^P = \mathcal{H}_j^P \left(\bigcup_{k \in \{1, \dots, n_E\}} g_k(I_j^P), \{I_1^C, I_2^C, \dots, I_{n_{cE}}^C\} \right)$$

Où, I_j^P est l'indicateur fonctionnel de l'élément E, $g_k(I_j^P)$ sont les indicateurs du sous-élément E_k contribuant à I_j^P , $\{I_1^C, I_2^C, \dots, I_{n_{cE}}^C\}$ les indicateurs contextuels de E et \mathcal{H}_j^P la fonction de traitement permettant d'obtenir I_j^P .

La contribution dans \mathcal{H}_j^P des indicateurs contextuels consiste par exemple à valider les traitements dans certaines phases de fonctionnement. Par exemple, le calcul d'un rendement dans une phase de démarrage à froid d'un moteur, même s'il a un sens théorique, dans la pratique, les résultats obtenus ne sont pas significatifs.

Concernant les indicateurs de dégradation, les liens entre les indicateurs de niveaux différents sont identifiés à partir des relations entre les modes de dégradation d'un élément E et ceux des éléments E_k de niveau inférieur. Ainsi, pour chaque élément E, les indicateurs de dégradation sont calculés en intégrant les indicateurs des modes de dégradation des sous éléments les plus significatifs et les plus critiques. De la même manière que pour les indicateurs fonctionnels, nous pouvons écrire :

$$I_j^D = \mathcal{H}_j^D \left(\bigcup_{k \in \{1, \dots, n_E\}} g_k(I_j^D), \{I_1^C, I_2^C, \dots, I_{n_{cE}}^C\} \right)$$

Où, I_j^D est l'indicateur dysfonctionnel de l'élément E, $g_k(I_j^D)$ sont les indicateurs de l'élément E_k contribuant à I_j^D , $\{I_1^C, I_2^C, \dots, I_{n_{cE}}^C\}$ les indicateurs contextuels de E et \mathcal{H}_j^D la fonction de traitement/agrégation permettant d'obtenir I_j^D .

De la même manière que pour \mathcal{H}_j^P , la contribution dans \mathcal{H}_j^D des indicateurs contextuels consiste par exemple à adapter les traitements dans certaines phases de fonctionnement. Par exemple, la température de carter d'un moteur diesel marin n'est pas significative dans une phase de démarrage à froid.

Enfin, quant aux indicateurs contextuels, ils ne sont pas constitués à partir de ceux du niveau inférieur puisque chaque contexte d'un élément E provient de l'environnement dans lequel il évolue.

Dans la suite et plus particulièrement pour le calcul des capacités pour l'intégrale de Choquet, nous avons besoin de généraliser les fonctions g_k à tous les sous-ensembles d'indicateurs de $\mathcal{V}(E)$ à savoir $\wp(\mathcal{V}(E))$.

Ainsi, nous redéfinissons g_k comme : $\forall k = 1 \dots n_E$

$$g_k: \wp(\mathcal{V}(E)) \rightarrow \wp(\mathcal{V}(E_k))$$

$$\forall S \subseteq \mathcal{V}(E), g_k(S) = \{I_j \in \mathcal{V}(E_k) \mid I_j \text{ est contributeur dans } S\}$$

Ainsi, $g_k(S)$ est l'ensemble des indicateurs de $\mathcal{V}(E_k)$ qui ont contribué (via les opérateurs \mathcal{H}_j^P et \mathcal{H}_j^D) au calcul de tous les indicateurs de S . g_k est une application puisque chaque ensemble d'indicateurs du domaine $\wp(\mathcal{V}(E))$ a une et une seule image qui est, un sous-ensemble de $\wp(\mathcal{V}(E_k))$.

La principale contribution de nos travaux de thèse réside dans la fusion des indicateurs de différentes natures en un index unique. Cette fusion nécessite une étape cruciale de normalisation. Cette fonction est définie de manière à assurer aussi la détection des dérives des valeurs de chaque indicateur par rapport au mode de comportement nominal. En outre, cette détection permet de tenir compte de l'effet potentiel du contexte (opérationnel et/ou environnemental) sur les valeurs d'un indicateur.

3. Une nouvelle approche pour la détection conditionnelle

La quantification des dérives des valeurs prises par chaque indicateur permet de lui attribuer un score dans $[0,1]$. En effet, à un instant t , les valeurs des indicateurs normalisés de santé proches de 1 indiquent un mode nominal. Des valeurs nulles indiquent un mode de faute ou de défaillance (jugées non acceptables).

De plus, l'étape suivante d'agrégation nécessite que l'ensemble des valeurs à agréger soit sur une échelle commune tant sémantique que numérique. Pour l'aspect numérique, nous avons défini l'intervalle $[0,1]$. Quand à l'aspect sémantique, c'est la notion de quantification des dérives relativement au comportement nominal qui constitue la base commune.

Cependant, il est nécessaire de faire la distinction entre une variation normale et anormale des valeurs des indicateurs. En effet, les changements des conditions opérationnelles et/ou environnementales peuvent induire des variations parfois significatives dans le comportement sans que cela ne soit considéré comme étant une anomalie. C'est pourquoi les variations dues au changement de contexte doivent être identifiées et séparées des variations indiquant l'apparition de dégradations ou de perte de performance.

Pour une détection des anomalies conditionnelle et pour une généralité, notre approche s'appuie sur l'utilisation d'histogrammes décrivant les comportements des valeurs prises par les indicateurs pour les modes nominaux et de faute. Ensuite, le calcul de l'écart des

observations actuelles (sous forme d'histogramme) par rapport aux histogrammes de référence est effectué en utilisant l'entropie relative.

D'un point de vue détection par apprentissage automatique, il existe trois scénarios différents selon la disponibilité des informations concernant la nature du mode de fonctionnement caractérisant le signal étudié (mode nominal ou de faute) [Michie, 1994]:

- La détection d'anomalies supervisée qui se base sur des données d'apprentissage correspondantes au mode nominal et de faute simultanément,
- La détection d'anomalies semi-supervisée qui utilise un ensemble de données d'apprentissage composé que des données correspondant au mode nominal seulement,
- La détection d'anomalie non supervisée, où les données d'apprentissage sont dépourvues de labels (type de mode associé).

La démarche proposée repose sur une détection supervisée. Ainsi, des histogrammes caractérisant les profils des valeurs prises par chaque indicateur pour les deux modes nominal et de faute sont construits et stockés. Ces histogrammes sont appelés **histogrammes de référence**. Ensuite, l'étape de test en ligne, consiste à comparer un histogramme des observations actuelles par rapport à ces histogrammes de référence. Un score est alors calculé pour l'histogramme actuel selon son degré de similarité ou d'écart. La comparaison d'histogrammes est effectuée par la mesure d'entropie relative qui sera détaillée dans le paragraphe suivant.

Notre approche nécessite une phase d'apprentissage des histogrammes de référence. Dans cette phase d'apprentissage, une étape est la subdivision de l'espace contextuel. Cette étape est nécessaire si le contexte a une influence sur les distributions de référence.

La méthode proposée repose donc sur deux concepts principaux à savoir la construction d'histogrammes et le calcul de la distance entre deux distributions par entropie relative.

3.1. Concepts théoriques utilisés : histogramme et entropie relative

3.1.1. Les histogrammes pour la détection

Un histogramme est une représentation d'une distribution d'une variable sous la forme de l'ensemble des effectifs, des valeurs prises par cette variable, par modalité ou par intervalles selon que la variable soit qualitative ou quantitative. On peut le représenter à l'aide d'un graphique représentant la répartition de la variable. C'est une distribution du nombre

d'occurrences de chaque classe/intervalles de valeurs prise par cette variable sur le nombre total des valeurs [Pillet, 1995].

Dans notre problématique, les indicateurs sont enregistrés sur une fenêtre de temps d'observation et correspondent donc à des variables quantitatives. La technique standard de construction des histogrammes utilise k classes de valeurs représentées par des intervalles sur le support de la variable. Habituellement, les intervalles sont de largeur égale sur toute la plage de valeur. La fréquence (quantité relative) des échantillons dont les valeurs sont dans un intervalle est utilisée comme estimation de la densité (représentée par la hauteur des barres) [Goldstein, 2012].

Il convient dans un premier temps de définir le nombre de classes k . Dans la littérature on peut trouver différentes formules permettant de déterminer le nombre de classes [Pillet, 1995]. Dans notre approche, pour un indicateur donné, les classes des histogrammes utilisés doivent avoir la même largeur. En ce sens, nous allons partir de la règle la plus simple qui consiste à prendre la racine carrée du nombre d'instances N :

$$K = \sqrt{N}$$

Pour nos travaux, la notation d'un histogramme unidimensionnel proposée dans [Kind, 2009] est adoptée. Si un indicateur I a K classes de valeurs, et chaque intervalle $k \in [1, \dots, K]$ a une probabilité d'occurrence $p_k \in [0,1]$ alors l'histogramme peut être présenté par un vecteur :

$$\text{Hist}(I) = (p_1, p_2, \dots, p_K)$$

Où la probabilité p_k d'occurrence à une classe, est approchée par la formule :

$$p_k = \frac{\text{nombre des instances appartenant à cette classe}}{\text{nombre total des instances}}$$

3.1.2. Entropie et entropie relative

Il est nécessaire maintenant de présenter les concepts de mesure de contenu d'information dans les signaux. Pour plus de détails, le lecteur peut se référer à l'ouvrage de Cover et Thomas [Cover, 2006]. Parmi les mesures de contenu d'information définies dans cet ouvrage, les plus utilisées sont celles de l'entropie et l'entropie relative.

L'entropie, appelée aussi entropie de Shannon car introduite par Claude Shannon en 1948 [Shannon, 1948], est un concept très important dans le domaine de la théorie de l'information et plus particulièrement pour la communication. L'entropie mesure l'incertitude dans un ensemble de données ou dans un signal. En effet, l'interprétation classique de l'entropie, est qu'elle spécifie le nombre minimum de bits nécessaire pour coder un signal sans perte d'information. Plus un signal est redondant (i.e. qui comporte un grand

nombre de valeurs semblables) plus son entropie est faible. Pour un nombre de valeurs différentes fixées dans un signal, l'entropie est maximale lorsque toutes les valeurs sont équiprobables, i.e. équiprésentes.

Soit X une variable aléatoire discrète qui prend ses valeurs parmi un ensemble de valeurs C_X suivant une distribution de probabilité $p(x)$. L'entropie de X [Shannon, 1948] est définie comme suit :

$$H(X) = - \sum_{x \in C_X} p(x) \cdot \log(p(x))$$

Où le logarithme est un logarithme de base 2.

L'entropie est exprimée en bits et elle est positive. Par convention, on pose $0 \cdot \log(0) = 0$. Ainsi, H est nulle lorsque p est une masse de Dirac (la distribution de probabilité est nulle sauf pour un singleton où elle est égale à 1). Ainsi, si toutes les valeurs d'un signal appartiennent à une seule classe, alors l'entropie est nulle. D'autre part, H atteint son maximum ($\log(|C_X|)$) si et seulement si p est une distribution uniforme [Cover, 2006].

L'entropie relative est utilisée pour mesurer la similarité entre deux distributions. Par exemple entre la distribution de probabilités d'observations et la distribution de probabilités d'un modèle choisi. Elle permet notamment de mesurer la distance entre deux distributions de données [Lee, 2001 a]. C'est une mesure qui est non-symétrique puisque $H_R(p_1, p_2) \neq H_R(p_2, p_1)$. De plus, elle ne respecte pas l'inégalité triangulaire, mais il est évident que plus p_1 et p_2 sont proches, plus la mesure d'entropie relative est faible.

Ainsi, l'entropie relative (ou encore la divergence de Kullback-Leibler, KL) [Kullback, 1951] entre deux distributions de probabilité $p_1(x)$ et $p_2(x)$ est définie comme suit :

$$H_R(p_1, p_2) = \sum_{x \in C_X} p_1(x) \cdot \log\left(\frac{p_1(x)}{p_2(x)}\right)$$

Où, comme dans le cas de l'entropie, le logarithme est un logarithme de base 2.

L'entropie relative est également exprimée en bits et elle est positive (elle est nulle si $p_1 = p_2$) et par convention, on pose [Cover, 2006] :

$$0 \cdot \log\left(\frac{0}{p_2(x)}\right) = 0 \text{ et } 0 \cdot \log\left(\frac{0}{0}\right) = 0$$

$$p_1(x) \cdot \log\left(\frac{p_1(x)}{0}\right) = \infty$$

Afin d'éviter la divergence de l'entropie relative vers l'infini lorsque $p_2(x) = 0$ et $p_1(x) \neq 0$, [Jullien, 2008] propose d'utiliser la « skew divergence » afin de contourner cette

contrainte. Cette divergence est un cas particulier de l'entropie relative [Lee, 2001 b] et consiste à ajouter un faible pourcentage à la distribution comparée comme suit :

$$SD_{\alpha}(p_1, p_2) = H_R(p_1, \alpha \cdot p_1 + (1 - \alpha) \cdot p_2) \\ = \sum_{x \in C_X} p_1(x) * \log \left(\frac{p_1(x)}{\alpha \cdot p_1(x) + (1 - \alpha) \cdot p_2(x)} \right)$$

Ainsi, lorsque deux distributions, p_1 et p_2 sont disjointes, l'entropie relative approximée par cette divergence devient :

$$H_R(p_1, p_2) = \sum_{x \in C_X} p_1(x) * \log \left(\frac{p_1(x)}{\alpha \cdot p_1(x) + (1 - \alpha) \cdot 0} \right) \\ = \sum_{x \in C_X} p_1(x) * \log \left(\frac{1}{\alpha} \right) \\ = -\log(\alpha)$$

Nous proposons d'utiliser l'entropie relative pour mesurer la déviation par rapport au mode nominal. Les distributions (sous forme d'histogrammes) construites à partir des données d'apprentissage servent de modèles de référence auxquels la distribution de valeurs en cours d'observation est comparée. Cependant, la comparaison des données observées vis-à-vis des modèles de référence est dépendante de la manière dont ces modèles eux mêmes se positionnent l'un par rapport à l'autre. C'est pourquoi, la tendance de la déviation des indicateurs au cours d'une perte de performance ou d'une dégradation façonne la formule par laquelle l'écart (score de déviation) est calculé.

3.2. Notion de dérive d'un indicateur

L'approche proposée pour la quantification des déviations des valeurs prises par les indicateurs nécessite d'identifier, d'une manière générique, les tendances habituelles de ces dérives. Pour cela, à partir des déviations généralement observées, nous identifions l'impact de la déviation sur l'histogramme de la variable. Ainsi, nous avons identifié 2 grands types de dérives sur les histogrammes. Ceci permet d'introduire 2 fonctions de calcul de la quantification de la déviation.

3.2.1. Les deux types de tendance de la dérive

Notre objectif à ce stade est d'identifier les différents types de déviations pour les paramètres suivis qui peuvent apparaître lors d'une dégradation. Pour cela, nous nous basons sur les grandes classes de déviations présentées dans [Léger, 1999] qui fait une présentation explicite et exhaustive des différentes allures types de déviation qui sont recensées sur la base des travaux du domaine [Ragot, 1990] et [Ghetie, 1997]. L'ensemble de

ces déviations permet de mieux appréhender les différentes variations possibles (Tableau 5) sans toutefois en avoir une représentation formelle. Sur chaque allure de déviation, le signal correspondant à un mode de fonctionnement nominal est coloré en vert. La zone de déviation est par contre soulignée par un rectangle de couleur orange.

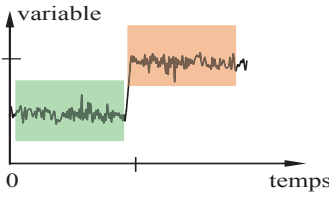
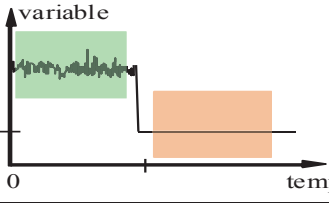
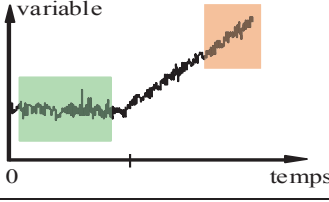
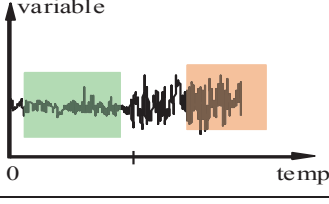
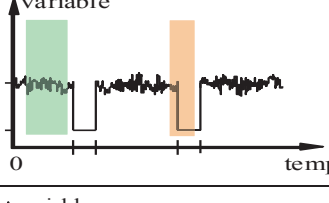
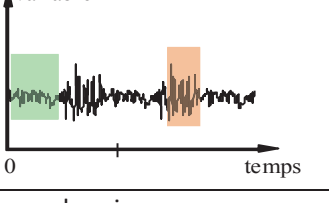
Cas de figure	Type de la dérive	Schématisation
(1)	Déviati on représentant un biais	
(2)	Déviati on représentant un arrêt brusque de fonctionnement	
(3)	Déviati on représentant une dérive	
(4)	Déviati on représentant une amplitude augmentée	
(5)	Déviati on représentant un fonctionnement intermittent	
(6)	Déviati on représentant un fonctionnement irrégulier	

Tableau 5 : Différentes classes de déviations pour les signaux

En relation avec l'objectif de détection et vis-à-vis de l'utilisation de l'entropie relative, deux grand types de variation d'histogramme doivent être distingués car ils vont nécessiter une mesure différente de similarité entre les deux histogrammes de référence (distribution du mode nominal et celle en présence de faute). En effet, la mesure d'entropie relative entre

les deux histogrammes est maximale ($-\log(\alpha)$) dès que ces deux derniers sont disjoints. Il faut donc adapter la formule pour le cas où les histogrammes sont disjoints. Ainsi, nous proposons de séparer les deux types d'histogrammes suivants :

- Type I : Cas où les histogrammes du mode nominal et de faute sont disjoints, c.à.d. lorsque les deux histogrammes ne partagent pas de classes de valeurs communes. Ceci correspond à la situation de déviations dans les trois premiers cas de figures (1-2-3) du Tableau 5. Nous allons adopter la disjonction des histogrammes de référence (Figure 33) pour représenter ce type de déviation. Il convient de noter ici que la déviation de ce type peut être croissante ou décroissante. Pour une déviation croissante l'histogramme correspondant au mode nominal se situe à gauche. Pour le cas contraire, il est à droite par rapport à l'histogramme du mode de faute.

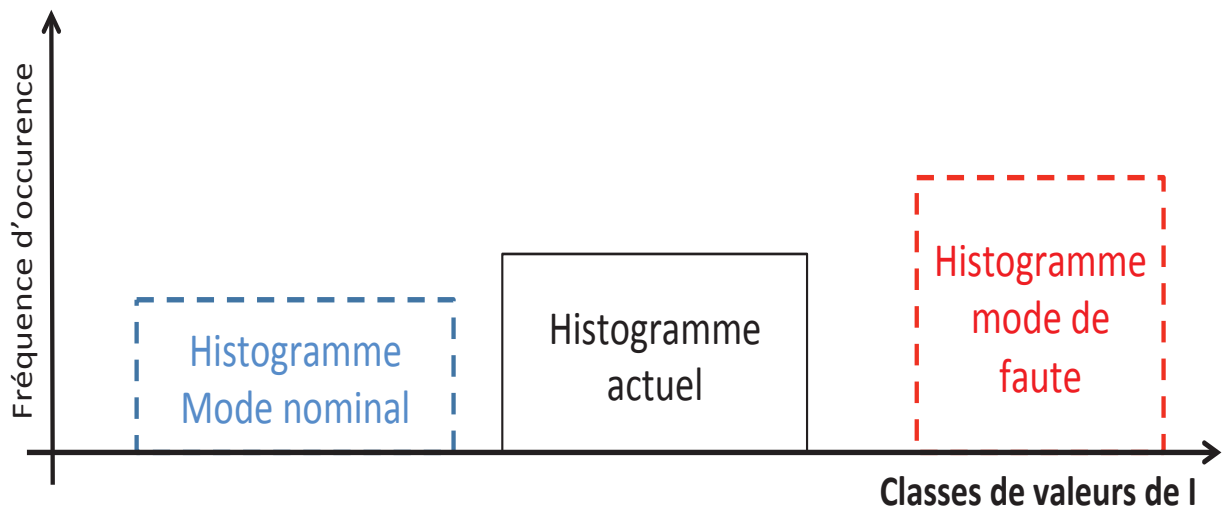


Figure 33 : Disposition des histogrammes pour le type I avec une déviation croissante

- Type II : Cas où les histogrammes du mode nominal et celui en présence de faute sont superposés (i.e. ils peuvent être emboîtés ou partiellement superposés), c.à.d. lorsque ces deux histogrammes ont des classes de valeurs en communs. C'est le cas des déviations représentées par les allures (4-5-6) dans le Tableau 5. La disposition des histogrammes pour ce type de dérive peut être représentée par la Figure 34.

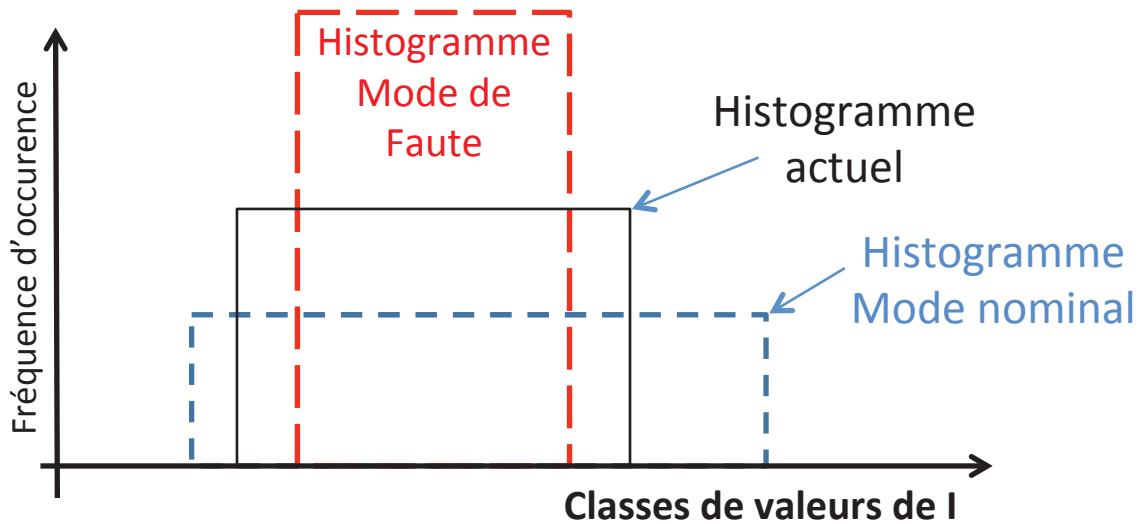


Figure 34 : Disposition des histogrammes pour le type II

Pour les deux allures génériques inhérentes aux déviations, une formule pour le calcul du score de déviation pour chaque indicateur est élaborée. Ces deux formules sont la base de la procédure de détection/normalisation présentée dans ce chapitre.

3.2.2. Quantification de la dérive

L'écart pris par un indicateur I pour l'observation actuelle donne un score $i \in [0,1]$ calculé comme suit :

- Soit $\text{Hist}(I) = (p_1, p_2, \dots, p_k)$ (l'histogramme (distribution de densité) correspondant aux dernières observations en ligne de cet indicateur (sur une fenêtre glissante d'un certain temps correspondant à une séquence de N points). Le nombre de classes est déterminé en premier lieu à l'aide de la longueur de la fenêtre d'observation en ligne à savoir N).
- Soit $\text{Hist}_{\text{nom}}(I) = (p_1^{\text{nom}}, p_2^{\text{nom}}, \dots, p_{k_n}^{\text{nom}})$ (l'histogramme correspondant aux valeurs prises par cet indicateur en mode nominal).
- Soit $\text{Hist}_f(I) = (p_1^f, p_1^f, \dots, p_{k_f}^f)$ (l'histogramme correspondant aux valeurs prises par cet indicateur en présence de faute. Le nombre de classes de l'histogramme du mode de faute est k' . k' est déterminé à partir de la largeur des classes du mode nominal et de l'étendue de $\text{Hist}_f(I)$).

Le calcul du score pour un indicateur du premier type est effectué de la manière suivante :

$$\begin{aligned}
 i &= 1 - \frac{H_R(\text{Hist}(I), \text{Hist}_{\text{nom}}(I))}{H_R(\text{Hist}_f(I), \text{Hist}_{\text{nom}}(I))} \cdot \frac{(l_{\text{act}} - l_{\text{nom}})}{(l_f - l_{\text{nom}})} \\
 &= 1 - \frac{H_R(\text{Hist}(I), \text{Hist}_{\text{nom}}(I))}{-\log(\alpha)} \cdot \frac{(l_{\text{act}} - l_{\text{nom}})}{(l_f - l_{\text{nom}})}
 \end{aligned}$$

Avec :

- l_{act} : limite supérieure de l'histogramme correspondant à l'observation actuelle en ligne pour une déviation croissante. En cas d'une déviation décroissante, cette limite correspond à la limite inférieure de l'histogramme actuel,
- l_{nom} : limite inférieure de l'histogramme correspondant au mode nominal pour une déviation croissante, ou alors la limite supérieure pour une déviation décroissante,
- l_f : valeur moyenne du support de l'histogramme correspondant au mode de faute (pour les deux cas).

Cette formule tient compte du rapport entre la distance de la distribution actuelle par rapport à celle du mode nominal et la distance entre les deux distributions de référence qui sont disjointes. Dans ce cas, on montre que cette dernière distance est égale à $-\log(\alpha)$. Ce rapport d'entropies relatives, multiplié par le rapport des distances séparant les supports des distributions forme l'écart de l'histogramme actuel vis-à-vis du mode nominal (relativement au mode de faute aussi). Le score est obtenu en retranchant l'écart de 1. Cependant le deuxième type de dérive est caractérisé par la superposition des distributions (Figure 34). Dans ce cas, le score pour une distribution courante est obtenu en retranchant de 1 la quantité d'écart de la distribution actuelle par rapport à la distribution nominale rapportée à la somme des distances entre la distribution actuelle et celles des deux distributions de référence (celles du mode nominal et du mode de faute). Ainsi, la formule donnant la valeur i pour un indicateur I de type II est donnée par :

$$i = 1 - \frac{H_R(Hist(I), Hist_{nom}(I))}{H_R(Hist(I), Hist_{nom}(I)) + H_R(Hist(I), Hist_f(I))}$$

Afin de pouvoir appliquer l'entropie relative entre deux histogrammes, l'objectif prioritaire est de fixer la largeur des classes ainsi que d'avoir une référence commune pour les supports des histogrammes. Le nombre de classe k doit permettre de construire une graduation composée de toutes les limites potentielles des barres des histogrammes avec le même écart.

Afin d'illustrer la manière avec laquelle cette méthode permet d'accorder un score de déviation dans les deux cas de figure, nous proposons dans la suite un exemple simple mais démonstratif. Les deux tendances typiques considérées pour quantifier la dérive, à savoir le type 1 et le type 2, sont respectivement présentées par les indicateurs dans (Figure 35 et Figure 36). Sur chaque figure, la zone du signal correspondante au comportement nominal des valeurs de l'indicateur en question est colorée en bleu. Celle correspondante aux valeurs de faute est par contre colorée en rouge. Ensuite, trois fenêtres d'observation (fenêtres glissantes de même durée) sont présentées à différents instants de la dérive (de plus en plus

avancée). Le score correspondant à chaque fenêtre d'observation est alors calculé pour l'indicateur suivi.

3.2.2.1. Exemple d'indicateur de type 1

Afin de pouvoir appliquer les calculs par entropie relative, les classes de tous les histogrammes mis en œuvre doivent avoir la même largeur et le même support (une graduation commune à toutes les classes des histogrammes). La largeur des classes est identifiée à partir du nombre des points dans la fenêtre d'observation.

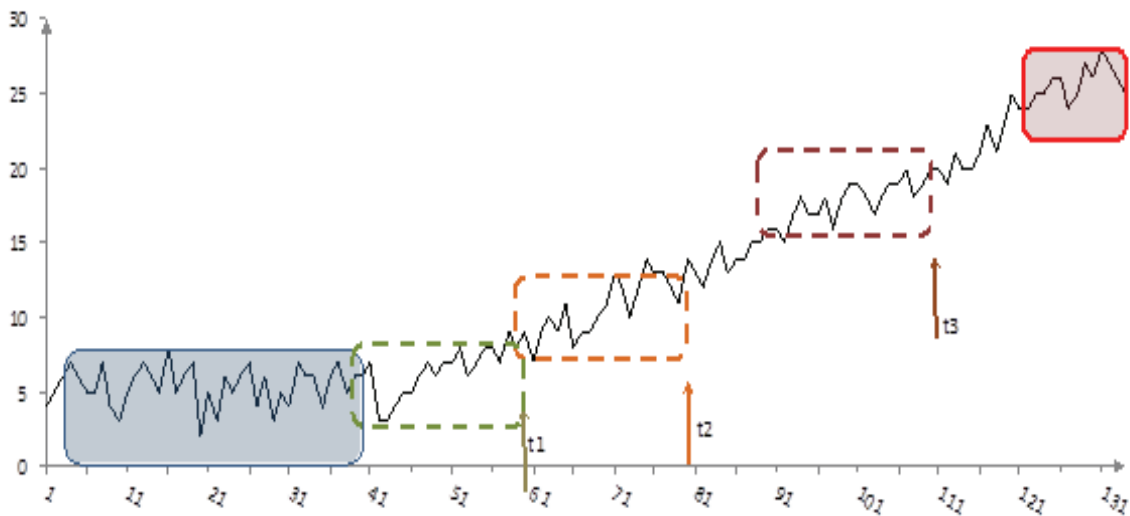


Figure 35 : Zones de valeurs nominales et de faute pour un exemple d'indicateur de type 1.

Pour ce premier exemple, la distribution sous forme d'un histogramme des valeurs prises par cet indicateur en mode nominal ainsi que pour chacune des observations, aux instants t_1 , t_2 et t_3 , est présentée dans le Tableau 6.

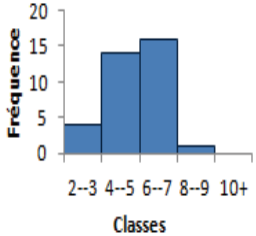
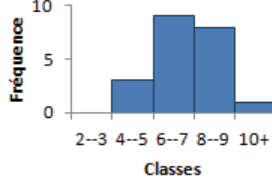
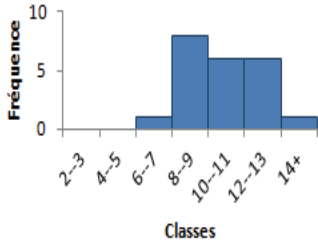
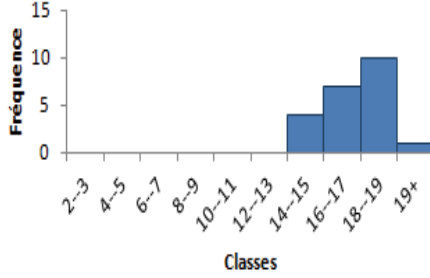
Mode	distribution	histogramme	Score	
Nominal	Classes			
	Probabilité			
	2--3			0,11
	4--5			0,40
	6--7			0,46
8--9	0,03			
Faute	$I_f=25.5$			
Observation à t1	Classes		0.9	
	Probabilité			
	2--3			0,00
	4--5			0,14
	6--7			0,43
	8--9			0,38
10+	0,05			
Observation à t2	Classes		0.59	
	Probabilité			
	6--7			0,05
	8--9			0,36
	10--11			0,27
	12--13			0,27
14+	0,05			
Observation à t3	Classes		0.4	
	Probabilité			
	14--15			0,18
	16--17			0,32
	18--19			0,45
19+	0,05			

Tableau 6: Représentation des distributions correspondantes aux instants d'observation et les scores relatifs pour le premier signal.

Le calcul détaillé du score est donné dans la suite. Notons que pour cet exemple, afin de pouvoir appliquer la formule de calcul (par skew divergence) alpha est fixé à 0.1 comme dans [Jullien, 2008].

- Observation à t1 :

$$i_1(t1) = 1 - \left[0.14 * \log\left(\frac{0.14}{0.1 * 0.14 + 0.9 * 0.40}\right) + 0.43 * \log\left(\frac{0.43}{0.1 * 0.43 + 0.9 * .46}\right) + 0.38 * \log\left(\frac{0.38}{0.1 * 0.38 + 0.9 * 0.03}\right) + 0.05 * \log\left(\frac{0.05}{0.1 * 0.05}\right) \right] * \left[\frac{(9.5 - 1.5)}{-\log(0.1) * (25.5 - 1.5)} \right]$$

$$= 0.9$$

- Observation à t2 :

$$i_1(t2) = 1 - \left[0.05 * \log\left(\frac{0.05}{0.1 * 0.05 + 0.9 * 0.43}\right) + 0.36 * \log\left(\frac{0.36}{0.1 * 0.36 + 0.9 * .03}\right) + 0.27 * \log\left(\frac{0.27}{0.1 * 0.27}\right) * 2 + 0.05 * \log\left(\frac{0.05}{0.1 * 0.05}\right) \right] * \left[\frac{(13.5 - 1.5)}{-\log(0.1) * (25.5 - 1.5)} \right]$$

$$= 0.59$$

- Observation à t3 :

$$i_1(t3) = 1 - \frac{-\log(0.1)}{-\log(0.1)} * \frac{12}{20} = 0.4$$

Cet exemple illustratif montre que les scores ainsi calculés correspondent aux degrés de déviation du signal dans les trois fenêtres d'observation par rapport aux modes nominal et de faute (Figure 35). En effet, la première fenêtre d'observation montre un signal très proche du comportement nominal, le score est alors proche de 1. Lorsque la déviation s'accroît, le score diminue aussi pour refléter la distance du signal observé par rapport au mode de faute.

3.2.2.2. Exemple d'indicateur de type 2 :

Dans ce deuxième exemple, un signal d'un indicateur type 2 est présenté dans la Figure 36.

Le Tableau 7 présente les histogrammes de référence aux instants t1, t2 et t3 ainsi que les scores obtenus. De la même manière, la largeur de classes est la même pour tous les histogrammes (en cours d'observation et ceux de référence). La graduation de l'axe support doit aussi permettre la superposition des classes.

Le détail de calcul du score pour ce deuxième type est donné via l'exemple de l'observation à t1 comme suit :

$$i_2(t1) = 1 - \left[0.11 * \log \left(\frac{0.11}{0.1 * 0.11 + 0.9 * 0.04} \right) + 0.30 * \log \left(\frac{0.30}{0.1 * 0.30 + 0.9 * 0.28} \right) + 0.32 * \log \left(\frac{0.32}{0.1 * 0.32 + 0.9 * 0.39} \right) + 0.27 * \log \left(\frac{0.27}{0.1 * 0.27 + 0.9 * 0.28} \right) \right] / \left[\left(0.11 * \log \left(\frac{0.11}{0.1 * 0.11 + 0.9 * 0.04} \right) + 0.30 * \log \left(\frac{0.30}{0.1 * 0.30 + 0.9 * 0.28} \right) + 0.32 * \log \left(\frac{0.32}{0.1 * 0.32 + 0.9 * 0.39} \right) + 0.27 * \log \left(\frac{0.27}{0.1 * 0.27 + 0.9 * 0.28} \right) \right) + \left(0.11 * \log \left(\frac{0.11}{0.1 * 0.11 + 0.9 * 0.28} \right) + 0.30 * \log \left(\frac{0.30}{0.1 * 0.30 + 0.9 * 0.16} \right) + 0.32 * \log \left(\frac{0.32}{0.1 * 0.32 + 0.9 * 0.12} \right) + 0.27 * \log \left(\frac{0.27}{0.1 * 0.27 + 0.9 * 0.16} \right) \right) \right]$$

De la même manière que pour le premier exemple, les scores ainsi calculés correspondent aux degrés de déviation du signal dans les trois fenêtres d'observation par rapport aux modes nominal et de faute (Figure 36). En effet, le score diminue avec l'approche du mode de faute.

Pour la détection d'anomalie, nous pouvons utiliser l'entropie comme une mesure de la distance entre les données en cours d'observation et celles des données du mode nominal par rapport à cette même différence entre les deux comportements de référence (mode nominal et de faute). Cependant ces comportements (observés et de référence) peuvent être dépendants du contexte dans lequel évolue le système surveillé. Dans la partie qui suit, nous présentons comment nous prenons en compte ce phénomène.

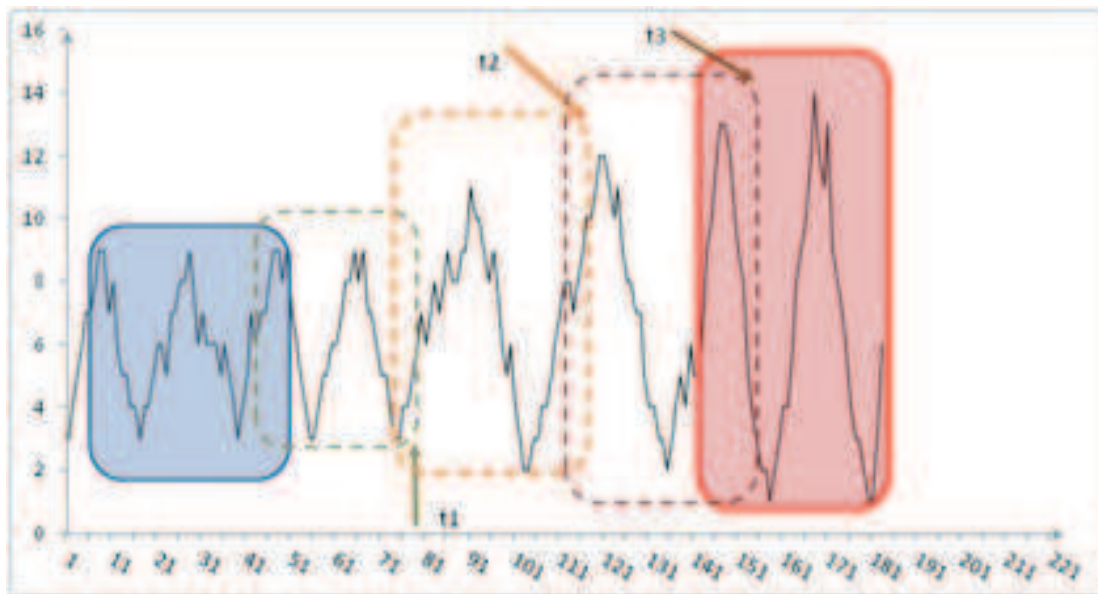


Figure 36: Zones de valeurs nominale et de faute pour un exemple d'indicateur de type 2.

Mode	distribution	histogramme	Score	
Nominal	Classes	Probabilité		
	2--3	0,04		
	4--5	0,28		
	6--7	0,39		
	8--9	0,28		
Faute	Classes	Probabilité		
	2--3	0,28		
	4--5	0,16		
	6--7	0,12		
	8--9	0,16		
	10--11	0,09		
	11--13	0,16		
	14--15	0,02		
Observation t1	Classes	Probabilité		0.93
	2--3	0,11		
	4--5	0,30		
	6--7	0,32		
	8--9	0,27		
Observation t2	Classes	Probabilité		0.58
	2--3	0,12		
	4--5	0,17		
	6--7	0,24		
	8--9	0,37		
	10--11	0,10		
Observation t3	Classes	Probabilité		0.19
	2--3	0,10		
	4--5	0,25		
	6--7	0,15		
	8--9	0,21		
	10--11	0,17		
	12--13	0,13		

Tableau 7 : Représentation des distributions correspondantes aux instants d'observation et les scores relatifs pour le deuxième signal.

3.3. Approche de détection contextuelle

La démarche de détection/normalisation proposée consiste en deux étapes principales (Figure 37):

- Un apprentissage hors ligne : l'objectif de cet apprentissage est de caractériser les comportements nominal et de faute a priori en tenant compte de l'éventuel effet du contexte opérationnel et environnemental.
- La détection en ligne qui consiste à calculer, pour chaque indicateur, un score quantifiant le degré d'écart par rapport au mode nominal.

L'étape apprentissage doit permettre de générer à partir de données historisées, pour chaque indicateur, les histogrammes de référence, i.e. celui du mode nominal et de faute. Le traitement des données brutes permet de cadrer le mode nominal de chaque indicateur. L'évolution vers le mode de faute peut différer suivant le mécanisme de dégradation mis en jeu. Cependant, en choisissant celui le plus proche du mode nominal, cela nous permet de réduire leur nombre en un seul. Suivant la dépendance de l'indicateur au contexte environnemental, l'espace des variables d'influence est subdivisé, i.e. construire des intervalles pour les indicateurs contextuels et apprendre des histogrammes de référence dans chacun d'eux. Cette procédure dépend bien sûr des données disponibles.

Ainsi les données d'apprentissage sont segmentées en différents lots. Chaque lot est défini par le contexte dans lequel les données ont été enregistrées. Cette organisation rassemble toutes les données issues de scénarios menés dans des conditions similaires. Les variables utilisées pour définir les lots de données sont les indicateurs contextuels (par exemple la température environnante, le régime opératoire...).

Ensuite, les histogrammes qui caractérisent les comportements pour chaque lot de valeurs sont construits. Une base contenant les histogrammes de référence permettant de décrire le comportement normal et de faute est ainsi élaborée. Cette base permet de trouver les histogrammes de comportements spécifiques à un contexte défini par sous-espaces associés aux indicateurs contextuels.

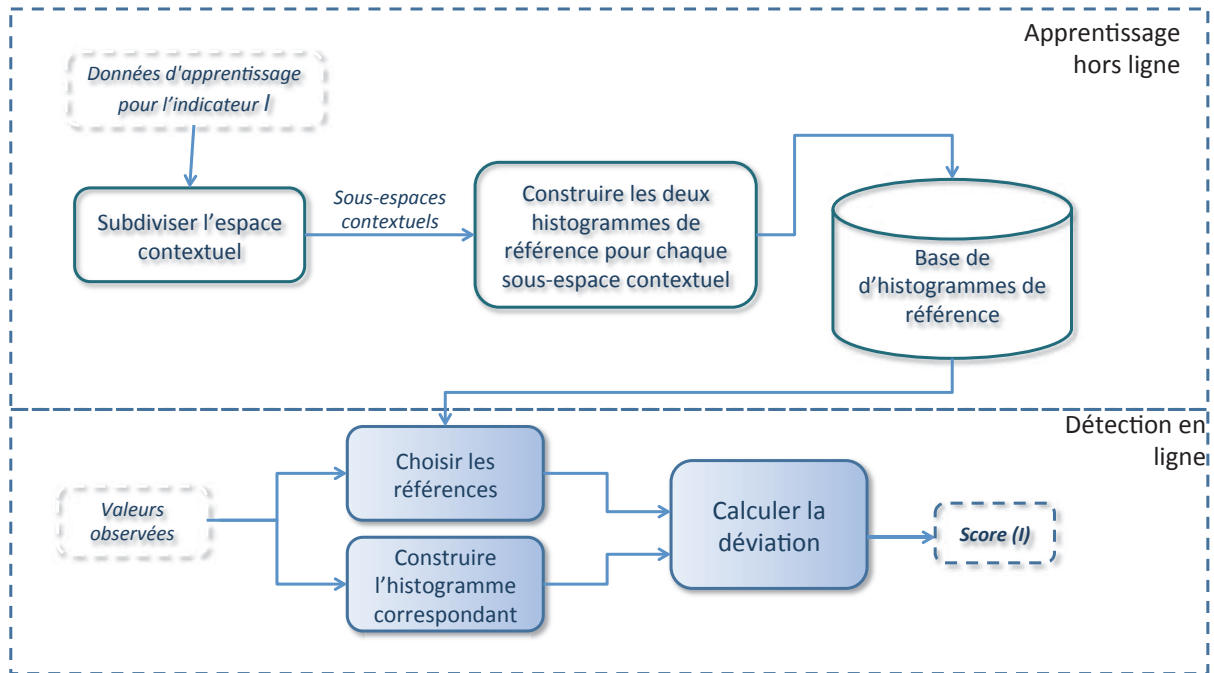


Figure 37 : Structure de l'approche de la normalisation contextuelle des indicateurs

L'apprentissage nécessite de tester pour un indicateur donné si son comportement nominal dépend du changement des conditions environnementales ou opérationnelles. Afin de réduire la quantité de données nécessaires pour l'apprentissage, cette proposition considère par hypothèse, qu'un indicateur peut être affecté au plus par un paramètre contextuel. Cependant, la même approche reste valide pour le cas de plusieurs paramètres contextuels pour un indicateur.

Pour le premier type de dérive ce test consiste à calculer le rapport entre la largeur de l'histogramme du mode nominal (sous les différentes conditions) et la distance séparant les deux histogrammes de référence (Figure 38). Si ce rapport est supérieur à γ , l'indicateur est alors dépendant de l'indicateur contextuel en question. γ pourra être fixé soit par un expert soit relativement aux signaux analysés. On cherche ainsi à vérifier que l'histogramme du mode nominal est suffisamment éloigné de celui de faute.

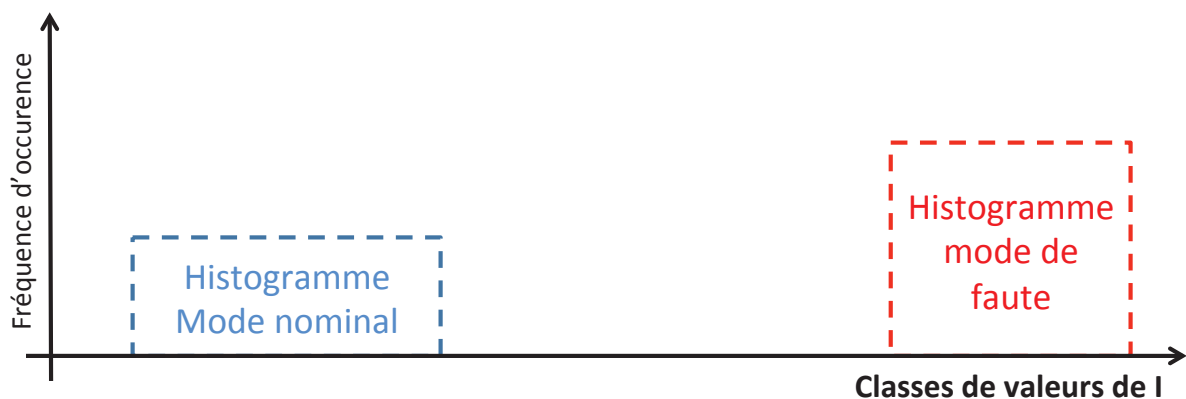


Figure 38: Paramètres pour le test de dépendance au contexte

Si I (un indicateur donné) est dépendant du contexte, alors :

$$\frac{(b - a)}{(c - a)} > \gamma$$

Avec a et b représentent les limites inférieure et supérieure de l'histogramme du mode nominal et c la classe moyenne du support de l'histogramme du mode de faute.

Pour le deuxième type de dérive, le test consiste à vérifier simplement si les deux distributions du mode nominal correspondantes aux deux valeurs extrêmes du paramètre contextuel sont disjointes (pas de classes de valeurs communes).

Une fois l'impact d'un paramètre contextuel identifié, l'étape suivante consiste à diviser l'espace des indicateurs contextuels en sous-espaces dans lesquels nous considérons que l'indicateur est indépendant des indicateurs contextuels. Il s'agit donc de définir les intervalles des indicateurs contextuels pour lesquels des histogrammes de référence doivent être construits. Dans chaque intervalle de l'étendue du paramètre contextuel, les valeurs prises par un indicateur en mode nominal varient légèrement (ou quasiment pas). De cette manière, un tel intervalle est représenté par un couple de distributions (nominal et faute) dans la base d'histogrammes.

En s'inspirant des méthodes par dichotomie, la subdivision de l'étendue de l'indicateur contextuel en sous-intervalles contextuels s'effectue d'une manière itérative. En effet, la valeur moyenne entre deux conditions extrêmes (en termes d'indicateur contextuel) est insérée à chaque nouvelle itération. Ensuite, l'histogramme correspondant à cette nouvelle valeur de l'indicateur contextuel est comparé à ceux correspondant aux deux valeurs extrêmes (Figure 39). Si cette nouvelle distribution est proche de l'une des deux, la valeur moyenne forme avec la valeur de l'indicateur correspondante un intervalle contextuel (Figure 40). L'intervalle restant est ainsi segmenté de la même manière en recherchant de la même manière des intervalles dans lesquels l'indicateur en question est considéré stable.

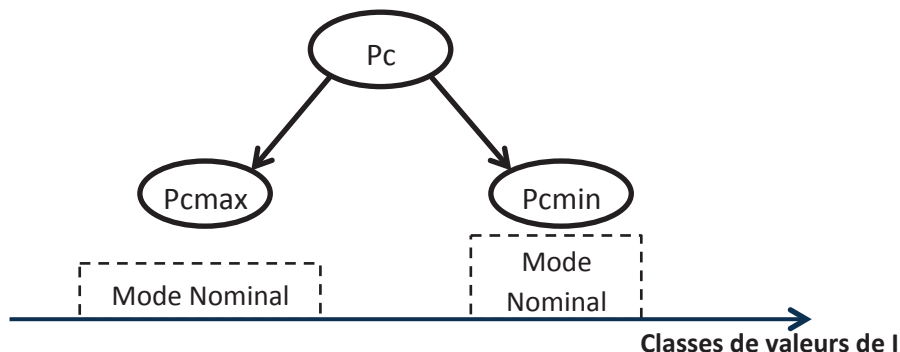


Figure 39 : Première étape de répartition de l'espace contextuel.

Cependant, si la valeur moyenne donne un histogramme à la fois proche des deux distributions extrêmes, le paramètre contextuel est segmenté en deux intervalles [P_{cmin} , P_{cmoy}] et [P_{cmoy} , P_{cmax}].

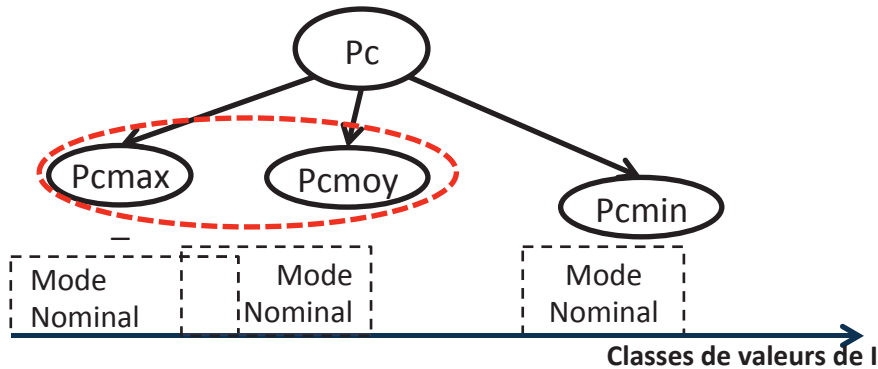


Figure 40 : Définition d'un intervalle du paramètre contextuel.

Les distributions (histogrammes) issues de l'apprentissage sont nécessaires pour la quantification en ligne de toute déviation potentielle pour chaque indicateur. Les fonctions de comparaison des valeurs observées en ligne durant une fenêtre de temps glissante utilisent les formules décrites dans ce chapitre selon le type de la dérive de l'indicateur en question.

La procédure de détection des anomalies ainsi décrite permet de normaliser les indicateurs de différentes natures. L'ensemble de ces indicateurs sont ensuite agrégés en un index de santé synthétique.

4. Identification des coefficients des capacités pour un système multi-niveaux hiérarchiques

Comme nous l'avons montré au chapitre 2, l'intégrale de Choquet est adoptée pour la fonction d'agrégation des indicateurs pour deux raisons principales :

- La prise en compte, en plus du poids des indicateurs, de leurs interactions mutuelles,
- La possibilité d'apprentissage des paramètres en mode non supervisé.

Cette partie présente notre proposition pour la détermination des paramètres de l'intégrale de Choquet, appelés capacités ou mesures floues. Le nombre de paramètres à déterminer pour une capacité étant important, nous proposons de procéder, d'une part, pour le niveau composant par identification à partir de signaux, et d'autre part, pour les niveaux supérieurs par calcul à partir des capacités du niveau immédiatement inférieur. Ainsi, seul le niveau composant nécessite des données pour réaliser une identification des capacités, celles des niveaux supérieurs étant déduites de celles de ce premier niveau.

4.1. Apprentissage des capacités au niveau composant

Lorsqu'aucune connaissance préalable concernant les contraintes sur/entre les variables à agréger ou sur l'importance des indicateurs ou de leurs interactions n'est disponible, l'identification des capacités ne peut pas être un problème d'optimisation comme dans les approches recensées dans [Grabisch, 2007]. Il convient alors d'utiliser une méthode non supervisée d'apprentissage, comme par exemple celle proposée dans [Jullien, 2008] pour le domaine de la classification et de l'adapter à notre contexte d'agrégation des indicateurs de santé en un index synthétique.

A la différence du domaine de la classification où les critères sont évalués par leur capacité à distinguer une classe des autres, dans notre cas, les indicateurs (remplaçant la notion de critères) doivent avoir plus d'importance lorsqu'ils permettent de mettre en évidence l'apparition de dégradation ou de perte de performance. L'idée derrière l'utilisation de l'entropie relative est d'accorder un poids à un indicateur selon son pouvoir "statistique" de discrimination par rapport à un mode nominal. Ce que nous attendons de cet apprentissage est que les indicateurs qui réagissent à l'apparition d'un ou plusieurs mécanisme(s) de dégradation aient un poids plus important. Étant donné qu'un indicateur peut traduire plus qu'une dégradation ou déviation, et qu'un même phénomène peut être représenté par plus d'un indicateur, les interactions de synergie ou de redondance entre indicateurs sont caractérisées par ces relations.

Une relation de redondance entre les indicateurs traduit leur comportement corrélé en présence d'un même phénomène. Ainsi, deux indicateurs sont en redondance dans un contexte de "condition monitoring" lorsqu'il existe une relation de cause à conséquence entre ces deux indicateurs ou encore lorsqu'ils évoluent suite à une cause primaire commune, i.e. lorsque deux indicateurs évoluent en même temps à cause d'une même dégradation.

Deux indicateurs sont en synergie lorsque l'information fournie par ces deux indicateurs correspond à deux modes de dégradation différents. En terme de synergie, la phase de détermination des modes de dégradations correspondant aux causes primaires va orienter les calculs entropiques à pointer "statistiquement" des synergies par le fait que les déviations par rapport au mode nominal ne se produisent pas simultanément ou alors, s'il s'agit de deux sources indépendantes pour un même mode de dégradation, cela va être révélé par leur comportement différent (ou différencié).

4.1.1. Démarche d'apprentissage

Afin de mettre en œuvre cette approche au cas des indicateurs dans un bilan de santé, les données d'apprentissage doivent être organisées suivant les étapes suivantes:

- Etape 1 : à partir d'une étude AMDEC (Analyse des Modes de Dégradations, de leurs Effets et de leur Criticité), identifier tous les modes de dégradations qui peuvent survenir sur l'élément étudié,
- Etape2 : isoler par cette étude les modes de dégradation qui représentent les causes primaires des dégradations ou de déviations des flux de sortie (en appliquant une étude HAZOP : HAZard and OPerability study). Une cause primaire est un mode de dégradation dont aucune cause dans les grilles AMDEC/HAZOP n'est identifiée, elle peut être aussi une déviation d'un flux d'entrée,
- Etape 3 : former un lot d'apprentissage composé à parts égales de données correspondantes au mode dégradé (avec la présence de chacune des causes primaires) et de données représentatives du mode nominal.

Soit $p_{nom(I_1, \dots, I_{niE})}$ et $p_{deg(I_1, \dots, I_{niE})}$, les distributions de probabilités jointes correspondants respectivement au mode nominal et au mode dégradé. Le poids de chaque sous ensemble d'indicateurs $S \in \wp(\mathcal{V}(E))$ est défini par :

$$\mu(S) := \begin{cases} 0, & \text{if } S = \emptyset, \\ \frac{H_R(p_{nom(I_{i_1}, \dots, I_{i_s})}, p_{deg(I_{i_1}, \dots, I_{i_s})})}{H_R(p_{nom(I_1, \dots, I_{niE}), p_{deg(I_1, \dots, I_{niE})})}, & \text{if } S = \{I_{i_1}, \dots, I_{i_s}\}, \end{cases}$$

4.1.2. Exemple

Quatre indicateurs et deux causes primaires de dégradations vont être considérés pour cet exemple illustratif comme suit:

- I1: indicateur lié à la déviation du flux d'entrée identifié comme la cause primaire 1,
- I2: indicateur lié à la dégradation du support représentant la deuxième cause primaire,
- I3, I4 sont des indicateurs représentant une dégradation qui est la conséquence de la cause primaire 1 (donc elles apparaissent simultanément),
- I5 est un indicateur attaché à un mode de dégradation qui est la conséquence de la cause primaire 2.

Les indicateurs sont normalisés et leurs scores en mode nominal et au cours de l'occurrence des deux causes primaires de dégradation ou de déviation de flux sont recensés dans le Tableau 8.

	I1	I2	I3	I4	I5
Mode Nominal	1	1	0,9	1	0,9
	1	1	0,9	1	0,9

	0,8	1	0,8	0,95	0,8
	0,8	1	0,8	0,95	0,8
Présence de la Cause Primaire 1	0,5	1	0,7	0,6	0,9
	0,4	0,9	0,6	0,6	0,9

	0,3	1	0,6	0,4	0,8
	0,2	0,9	0,6	0,5	0,8
Présence de la Cause Primaire 2	1	0,6	0,9	1	0,7
	1	0,5	0,9	1	0,6

	0,9	0,3	0,85	1	0,5
	1	0,2	0,8	1	0,5

Tableau 8 : Extrait des données d'apprentissage et leur structure pour l'identification des coefficients de la capacité

En appliquant l'approche d'apprentissage des capacités ainsi proposée, nous obtenons les résultats présentés sous forme de degrés d'interaction (Eq. 2.4) et de valeurs de Shapley (Eq. 2.3) représentés respectivement dans le Tableau 9 et le Tableau 10.

	I ₁	I ₂	I ₃	I ₄	I ₅
I ₁		0.095	-0.16	-0.17	0.04
I ₂	0.095		0.02	0.05	-0.31
I ₃	-0.16	0.02		-0.11	0.03
I ₄	-0.17	0.05	-0.11		0.02
I ₅	0.04	-0.31	0.03	0.02	

Tableau 9 : Valeurs des degrés d'interaction

Nous retrouvons ainsi les relations de synergie existantes entre les indicateurs 1 et 2. Des interactions de redondance sont aussi trouvées entre les indicateurs qui régissent simultanément une même cause primaire, i.e entre I1, I3 et I4 d'une part et entre I2 et I5 d'autre part. En outre, les indicateurs qui sont en redondance avec I1 ou I2 sont aussi en synergie avec l'autre indicateur (présentant la cause primaire).

Indicateur	I_1	I_2	I_3	I_4	I_5
Valeur de Shapley	0.22	0.32	0.13	0.15	0.17

Tableau 10 : Valeurs de Shapley correspondants à la capacité identifiée

Concernant l'importance de chaque indicateur, comme attendu, elle est proportionnelle au degré de variabilité des scores entre 1 et 0. En effet, pour les indicateurs I_1 et I_2 , le nombre de scores différents entre le mode dégradé et le mode nominal est supérieur à celui des autres indicateurs dont le comportement est plus homogène.

4.2. Inférence des capacités pour les niveaux supérieurs

Nous présentons dans cette partie une nouvelle approche pour déduire la capacité d'un élément E étant donné les capacités de ses sous-éléments E_k . Pour cela, les dépendances physiques entre les sous-éléments portées par les flux échangés doivent être considérées. En outre, trois types de relations structurelles entre les éléments d'une même fonction sont traités : des sous-éléments en série ou en parallèle ; le cas composite série-parallèle se déduisant des deux précédents. Cependant, il existe d'autres cas de montage d'éléments (e.g. en électronique le montage bridge) qui n'ont pas été traités dans cette thèse.

4.2.1. Cas parallèle

Les sous-éléments en parallèle sont caractérisés par l'absence de flux échangés, i.e. le flux de sortie d'une fonction n'entre pas dans l'une des autres fonctions considérées (Figure 41). Toutefois, des fonctions en parallèle peuvent avoir la même finalité, i.e. le même flux de sortie, tout en ayant des composants différents comme support. De même, elles peuvent avoir des flux d'entrée en commun. Nous faisons l'hypothèse que leurs modes de dégradation n'interagissent pas mutuellement.

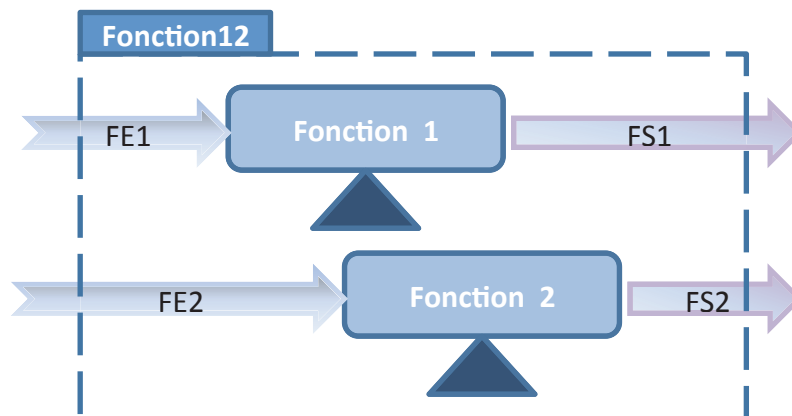


Figure 41 : Fonction composée de deux éléments en parallèle

Dans le cas des sous-fonctions en parallèle, nous nous inspirons de l'inter-additivité de mesures floues définie dans [Torra, 2007]. Si A et B sont deux ensembles de critères indépendants, alors la capacité de l'union de A et B est donnée par :

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

En s'inspirant de cette notion, nous proposons d'inférer la capacité μ_E d'un élément E à partir des capacités μ_{E_k} des éléments E_k qui le composent comme suit:

Pour tout sous-ensemble $S \subseteq \mathcal{V}(E)$,

$$\mu_E(S) = \frac{\sum_{k=1}^{n_E} \mu_{E_k}(g_k(S))}{n_E}$$

Où $g_k(S)$ est l'ensemble des indicateurs contributeurs du sous élément E_k à S.

Il suffit ainsi d'additionner les importances des indicateurs contributeurs des éléments en parallèle (donc qui sont indépendant étant donné qu'ils n'échangent pas des flux entre eux). La somme doit être normalisée (en la divisant sur le nombre de sous éléments en parallèle) afin de garantir la propriété de normalité pour la capacité ainsi inférée.

Maintenant, il faut montrer que μ_E vérifie bien les trois propriétés d'une capacité (Eq. 2.1)

$$- \mu_E(\emptyset) = \frac{\sum_{k=1}^{n_E} \mu_{E_k}(g_k(\emptyset))}{n_E} = 0$$

puisque $g_k(\emptyset) = \emptyset$ et $\mu_{E_k}(\emptyset) = 0, \forall k = 1 \dots n_E$ (μ_{E_k} est une capacité)

$$- \mu_E(\mathcal{V}(E)) = \frac{\sum_{k=1}^{n_E} \mu_{E_k}(g_k(\mathcal{V}(E)))}{n_E} = \frac{\sum_{k=1}^{n_E} \mu_{E_k}(\mathcal{V}(E_k))}{n_E} = \frac{n_E}{n_E} = 1$$

comme $\mu_{E_k}(\mathcal{V}(E_k)) = 1 \forall k = 1 \dots n_E$

- Soit $B_1 \subseteq B_2 \subseteq \mathcal{V}(E)$

alors $g_k(B_1) \subseteq g_k(B_2), \forall k = 1 \dots n_E$, par définition de g_k

d'ou $\mu_{E_k}(g_k(B_1)) \leq \mu_{E_k}(g_k(B_2)), \forall k = 1 \dots n_E$,

car $\forall k = 1 \dots n_E, \mu_{E_k}$ est une capacité,

ainsi $\sum_{k=1}^{n_E} \mu_{E_k}(g_k(B_1)) \leq \sum_{k=1}^{n_E} \mu_{E_k}(g_k(B_2))$

et finalement $\mu_E(B_1) \leq \mu_E(B_2)$

De cette façon, les interactions et l'importance mutuelle des indicateurs liés à une fonction peuvent être déduits à partir des capacités liées aux sous-éléments qui la supportent.

Exemple :

Afin de montrer la faisabilité et l'intérêt de la technique d'inférence des coefficients de capacité proposée dans le cas parallèle, nous proposons de l'appliquer sur un cas d'étude simple mais représentatif. En ce sens, un système de deux pompes fournissant de l'eau à un processus aval a été choisi. La première pompe (P1) est le fournisseur principal d'eau. Par conséquent, elle fonctionne de façon quasi continue. La deuxième pompe (P2) fournit de l'eau quand la première ne peut plus répondre à la demande. Par conséquent, la deuxième pompe fonctionne de façon intermittente. Pour les deux pompes, les indicateurs suivis sont les mêmes et ont été identifiés dans [Shukla, 2002] : l'efficacité (E), la corrosion (C), présence de fuite (F) et Analyse du lubrifiant (A). L'efficacité montre à quel point la pompe est performante. Les indicateurs dysfonctionnels reflètent les problèmes de dégradation potentiels :

- La corrosion conduit à l'usure de l'hélice, à son déséquilibre, à la rupture de l'arbre ... Ce mécanisme de dégradation est plus critique lorsque la pompe fonctionne de manière intermittente que lorsqu'elle fonctionne en continu.
- Une fuite indique un niveau de dégradation élevé du joint entraînant une perte dans le flux de sortie de l'eau.
- L'analyse du lubrifiant, qui en montrant la présence d'eau ou de particules, révèle une dégradation des roulements. Cette dégradation doit être mise en évidence surtout pour la pompe 2 en regard de son mode de fonctionnement.

Cet exemple traitant de l'agrégation, les indicateurs sont des indicateurs normalisés et nous n'avons plus d'indicateurs contextuels, ceux-ci ayant été pris en compte dans l'étape de détection/normalisation. Une valeur de 1 pour ces indicateurs signifie que la pompe est en situation nominale, c'est à dire qu'aucun problème n'existe vis-à-vis des mécanismes de dégradation surveillés. D'autre part, une faible valeur signifie qu'une défaillance suite aux dégradations suivies risque de se produire dans une date proche.

L'index de santé pour chaque pompe est calculé en utilisant des capacités différentes pour l'intégrale de Choquet. En effet, la capacité doit refléter les contraintes à propos de l'importance des indicateurs et de l'interaction entre eux pour chaque cas d'utilisation comme décrit ci-dessus. Le Tableau 11 montre les capacités des deux pompes compte tenu de ces contraintes. Ainsi, comme la pompe 1 fonctionne de manière continue et fournit la quantité d'eau principale, le principal indicateur à prendre en compte est l'efficacité conduisant à favoriser le rendement de la pompe dans l'index de santé. Le deuxième indicateur le plus important est l'indicateur de fuite. La deuxième pompe fonctionne de

manière intermittente afin de supporter la première pompe. Ainsi, l'importance est accordée à sa disponibilité instantanée pour toute demande plutôt qu'à son efficacité.

	{E}	{C}	{L}	{A}	{E,C}	{E,L}	{E,A}	{C,L}	{C,A}	{L,A}	{E,C,L}	{E,C,A}	{E,L,A}	{C,L,A}	{E,C,L,A}
P1	0.50	0.30	0.30	0.20	0.60	0.90	0.70	0.60	0.50	0.50	0.95	0.90	0.90	0.80	1
P2	0.20	0.50	0.10	0.40	0.60	0.30	0.50	0.55	0.95	0.45	0.65	0.97	0.60	0.95	1

Tableau 11 : Les coefficients de capacité pour les deux pompes

Les valeurs Shapley (Eq. 2.3) (Tableau 12) calculées sur les capacités de P1 et P2 permettent de montrer l'importance relative de chaque indicateur compte tenu de sa contribution dans l'ensemble de la capacité. Ainsi on peut retrouver les connaissances intégrées dans les coefficients de la capacité.

	E	C	L	A
P1	0.385	0.204	0.254	0.154
P2	0.118	0.451	0.061	0.368

Tableau 12 : Les valeurs de Shapley pour les deux pompes

En considérant la fonction globale constituée des deux pompes « pomper l'eau », les indicateurs surveillés sont :

- L'efficacité de la fonction de pompage (E) : cet indicateur est calculé à partir des deux indicateurs correspondant à l'efficacité pour les deux pompes de manière à favoriser l'efficacité de la pompe 1,
- La fuite maximale (L) : la valeur minimale entre les deux indicateurs de fuite pour les pompes,
- L'indicateur de dégradation mécanique (M) : cet indicateur reflète les caractéristiques de corrosion et de dégradation liées au lubrifiant. À un niveau supérieur, à l'encontre des indicateurs fonctionnels qui doivent être mis en évidence, les indicateurs de dégradation peuvent être synthétiques. Pour cet indicateur, l'occurrence de corrosion ou de détérioration de la qualité du lubrifiant doit être, pour la pompe 2, plus significative par rapport à ceux de la pompe 1.

Les indicateurs contributeurs pour les indicateurs de la fonction globale (P12) sont présentés dans le Tableau 13.

$S \subseteq \mathcal{V}(P12)$	$g_1(S)$	$g_2(S)$
{E}	{E}	{E}
{L}	{L}	{L}
{M}	{C,A}	{C,A}
{E,L}	{E,L}	{E,L}
{E,M}	{E,C,A}	{E,C,A}
{L,M}	{L,C,A}	{L,C,A}
{E,L,M}	{E,L,C,A}	{E,L,C,A}

Tableau 13 : Les indicateurs contributeurs pour la fonction P12

Le calcul des coefficients de capacité lié à cette fonction est montré dans le Tableau 14.

$\mathcal{V}(P12)$	μ_{P12}
{E}	$= \frac{\mu_1(E)+\mu_2(E)}{2} = \mathbf{0.35}$
{L}	$= \frac{\mu_1(L)+\mu_2(L)}{2} = \mathbf{0.2}$
{M}	$= \frac{\mu_1(C,A)+\mu_2(C,A)}{2} = \mathbf{0.725}$
{E,L}	$= \frac{\mu_1(E,L)+\mu_2(E,L)}{2} = \mathbf{0.6}$
{E,M}	$= \frac{\mu_1(E,C,A)+\mu_2(E,C,A)}{2} = \mathbf{0.935}$
{L,M}	$= \frac{\mu_1(L,C,A)+\mu_2(L,C,A)}{2} = \mathbf{0.875}$
{E,L,M}	$= \frac{\mu_1(E,L,C,A)+\mu_2(E,L,C,A)}{2} = \mathbf{1}$

Tableau 14 : Le calcul des coefficients de la capacité pour la fonction globale.

Ces coefficients montrent que {E,L} sont en synergie comme dans le niveau inférieur. Les couples d'indicateurs {L,M} et {E,M} conservent eux aussi leur interaction initiale, à savoir celle de la redondance. Ainsi, cette inférence permet de conserver le type d'interactions initiales entre les indicateurs du niveau inférieur. Concernant les valeurs de Shapley calculées pour cette fonction globale (Tableau 15), et comme attendu, l'indicateur le plus important est {M}, puis {E} et finalement {L}. Cet ordre d'importance est expliqué par le fait que l'indicateur {M} regroupe 4 indicateurs relatifs aux deux pompes à la différence des indicateurs {E} et {L} qui eux en intègrent uniquement 2. Ainsi, nous pouvons conclure que lorsqu'un indicateur est construit à partir d'un grand nombre de contributeurs, son importance a tendance à augmenter.

	E	L	M
P12	0.260	0.155	0.585

Tableau 15 : Les valeurs de Shapley correspondantes aux indicateurs de la fonction globale

Afin de montrer le comportement de cette déduction hiérarchique des capacités et ainsi que sa cohérence, le Tableau 16 présente quelques situations avec des valeurs d'indicateurs différentes pour P1, P2. Le tableau montre aussi les valeurs prises par les indicateurs de la

fonction globale ainsi que les indexes de santé (IS) calculés sur la base des coefficients de capacité présentés ci-dessus.

Cas	P1 E/C/L/A	IS _{P1}	P2 E/C/L/A	IS _{P2}	P12 E/L/M	IS _{P12}
1	0.9/0.8/0.85/0.9 nominal	0.88	0.9/0.8/0.85/0.9 nominal	0.855	0.9/0.85/0.85	0.867
2	0.9/0.2/0.85/0.9 corrosion	0.81	0.9/0.8/0.85/0.9 nominal	0.855	0.9/0.85/0.76	0.831
3	0.9/0.8/0.85/0.9 nominal	0.88	0.9/0.2/0.85/0.9 corrosion	0.61	0.9/0.85/0.64	0.783
4	0.3/0.8/0.85/0.9 Perte d'efficacité	0.67	0.9/0.8/0.85/0.9 nominal	0.855	0.48/0.85/0.85	0.803
5	0.9/0.8/0.85/0.9 nominal	0.88	0.3/0.8/0.85/0.9 Perte d'efficacité	0.725	0.72/0.85/0.85	0.833
6	0.3/0.2/0.85/0.9 Corrosion et perte d'efficacité	0.54	0.9/0.8/0.85/0.9 nominal	0.855	0.48/0.85/0.76	0.743
7	0.9/0.8/0.85/0.9 nominal	0.88	0.3/0.2/0.85/0.9 Corrosion et perte d'efficacité	0.485	0.72/0.85/0.64	0.714

Tableau 16 : Propagation des phénomènes de santé entre les deux niveaux.

Nous avons utilisé des situations typiques identiques pour les deux pompes: nominales, corrodées, moins efficace et moins efficace et corrodée (cas de 1 à 7). Nous avons cherché à montrer qu'une même situation des deux pompes P1 à P2 ne présente pas le même effet sur la fonction globale. Par exemple, les cas 4 et 5 soulignent que la perte d'efficacité de P1 induit une plus grande diminution de IS_{P12} que la même perte d'efficacité de P2. En fait, lorsque l'efficacité de la pompe1 diminue, elle a plus de conséquence sur l'efficacité de la fonction, car cet indicateur est essentiel pour cette pompe et donc il a un coefficient important qui se reflète par celui de l'indicateur de la fonction globale. De manière similaire, la dérive de l'indicateur de la corrosion a plus d'impact quand elle se produit sur la pompe 2 (cas 2 et 3). Lorsque les deux dégradations sont observées ensemble (cas 6 et 7), la corrosion a un peu plus d'effet sur l'index de santé de la fonction de pompage vu l'importance de l'indicateur de dégradation mécanique {M} pour la fonction globale.

4.2.2. Cas série

Dans le cas série, contrairement au cas précédent, certains indicateurs des sous-fonctions sont dépendants. En effet, les sous-éléments qui se suivent en série s'échangent des flux d'entrée/sortie (Figure 42). Ainsi, quand ils sont rassemblés dans la fonction globale (en fonction de leur contribution à l'objectif final), seule une partie des indicateurs des vecteurs locaux $\mathcal{V}(E_k)$ est reflétée dans $\mathcal{V}(E)$, i.e. les flux intermédiaires d'entrée/sortie n'apparaissent plus.

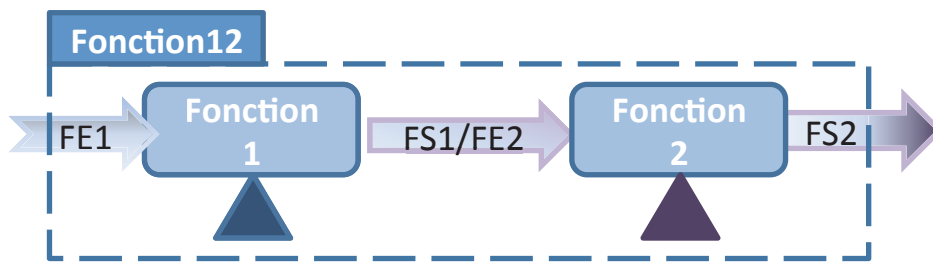


Figure 42 : Fonction composée de deux éléments en série

Les indicateurs contributeurs de chaque sous-élément E_k , i.e. $g_k(\mathcal{V}(E))$, sont de deux types:

- Indicateurs de dégradation qui doivent être signalés à des niveaux supérieurs et ce par rapport à une étude AMDEC de la fonction globale à surveiller.
- Indicateurs en lien avec les flux d'entrée du premier élément de la séquence (IF) et aussi ceux en lien avec les flux sortant de la séquence (OF). Dans cette même catégorie sont classés aussi les indicateurs de type rendement.

Afin de pouvoir visualiser les interactions (de synergie ou de redondance) entre chaque paire indicateurs contributeurs de chaque sous-élément E_k , un graphe est construit pour refléter ces interactions (Figure 43).

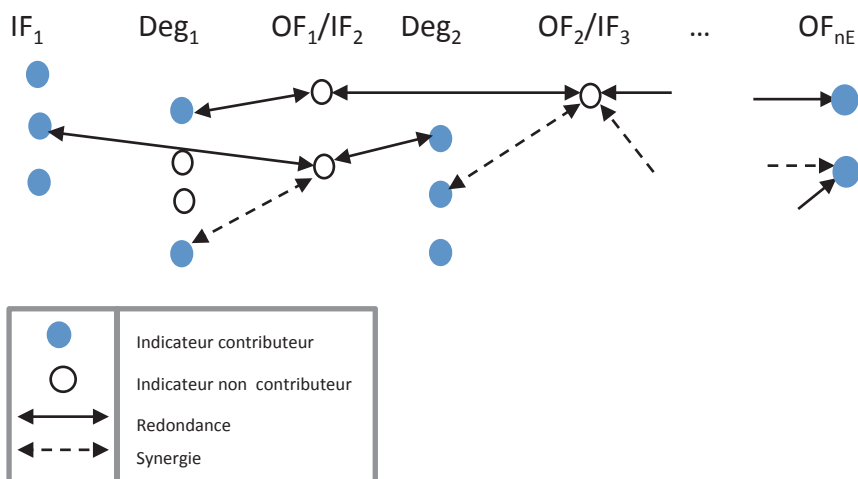


Figure 43 : Graphe de relations entre les indicateurs d'éléments en série

On note par :

- IF_k : les indicateurs de performance liés aux flux entrants du $k^{\text{ème}}$ sous-élément de la séquence en série (dans FE_1 , on trouve tous les indicateurs liés aux flux d'entrée qui ne sont pas un flux sortant d'une sous-fonction précédente dans la série),
- Deg_k : les indicateurs de dégradation du $k^{\text{ème}}$ sous-élément (ou encore les I^D),
- OF_k : les indicateurs de performance liés aux flux de sortie du sous élément à la position k . De même que FE_1 , FS_n regroupe les indicateurs qui représentent des flux de sortie qui n'alimentent pas un sous-élément successeur interne à la fonction en question,
- n_E : est le nombre de sous éléments en série.

Les indicateurs de chaque sous-élément sont répartis dans des colonnes correspondantes à chaque type d'indicateur (flux entrant, dégradation, flux sortant). Ensuite, les interactions (positive ou négative) qui peuvent exister entre un indicateur contributeur et un autre indicateur appartenant à la colonne (OF_k/IF_{k+1}) doivent être identifiées. Le type de relation entre deux indicateurs est déduit à partir des informations contenues dans les coefficients de la capacité déjà calculés pour chaque vecteur d'un sous-élément et plus précisément à partir des degrés d'interaction entre ces deux indicateurs. Un graphe devient le support pour qualifier les relations entre les indicateurs des sous-éléments dans la séquence et par la suite entre ceux du niveau supérieur.

Les coefficients de la capacité liée au vecteur d'indicateurs de la fonction globale sont déduits à partir de ceux des indicateurs contributeurs ainsi que des interactions qui peuvent exister entre les indicateurs contributeurs appartenant à des sous-éléments distincts. Ainsi, l'interaction entre deux indicateurs du niveau supérieur est issue, d'une part, de l'interaction initiale entre les indicateurs contributeurs d'un même sous-élément et, d'autre part, dans la contribution des interactions entre les sous-éléments.

Etant donné que seuls certains indicateurs de chaque sous-élément contribuent à la construction des indicateurs de E , on ne peut pas utiliser directement μ_k . Nous calculons un coefficient de contribution μ'_k , sur la base de μ_k qui ne prend en compte que les indicateurs de $g_k(\mathcal{V}(E))$. Pour cela, μ'_k est déduite de μ_k en normalisant le coefficient de capacité de $S_k \subseteq g_k(\mathcal{V}(E))$ par rapport au poids de $g_k(\mathcal{V}(E))$, i.e. $\mu_k(g_k(\mathcal{V}(E)))$. En outre, et afin de tenir compte du nombre d'éléments de E , n_E , nous multiplions le résultat précédent par la proportion d'indicateurs contributeurs de l'élément E_k sur le nombre total des indicateurs contributeurs à $\mathcal{V}(E)$.

μ'_k est définie alors pour chaque sous élément E_k comme :

$$\forall S_k \subseteq g_k(\mathcal{V}(E)), \mu'_k(S_k) = \frac{\mu_k(S_k)}{\mu_k(g_k(\mathcal{V}(E)))} * \frac{|g_k(\mathcal{V}(E))|}{|\cup_k g_k(\mathcal{V}(E))|}$$

Avec $|\cdot|$ désigne la cardinalité de l'ensemble en question.

Pour les indicateurs appartenant à $\mathcal{V}(E)$, la capacité résultante est inférée de la manière suivante.

$$\forall S \subseteq \mathcal{V}(E),$$

$$\mu(S) = \sum_{k=1}^{n_E} \mu'_k(g_k(S)) - \theta \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |S^-(g_k(S)) \cap g_j(S)| + \sigma \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |S^+(g_k(S)) \cup g_j(S)|$$

Ce calcul requiert les paramètres suivants :

- σ est le coefficient de synergie entre les sous-éléments de la séquence,
- θ est le coefficient de redondance lié à chaque nœud i puisqu'il porte la relation de cause à conséquence entre les indicateurs de deux sous-éléments de la séquence.

$S^-(g_k(S))$ et $S^+(g_k(S))$ sont respectivement les successeurs positifs et négatifs de $g_k(S)$. Ce sont un ou plusieurs indicateurs appartenant aux vecteurs des sous éléments en aval du sous élément E_k avec lesquels une relation de synergie (S^+) ou de redondance (S^-) existe.

Afin de s'assurer des propriétés d'une capacité pour la capacité inférée, σ et θ doivent vérifier ces deux conditions :

- Condition de normalité:

$$\theta \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |S^-(g_k(\mathcal{V}(E))) \cap g_j(\mathcal{V}(E))| = \sigma \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |S^+(g_k(\mathcal{V}(E))) \cup g_j(\mathcal{V}(E))|$$

Nous faisons l'hypothèse que:

$$\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |S^-(g_k(\mathcal{V}(E))) \cap g_j(\mathcal{V}(E))| \neq 0$$

Cette condition est facilement vérifiée pour les cas réels.

- Condition de monotonie :

$$\theta \leq \frac{\min_k \left\{ \min_{T, D \subseteq g_k(\mathcal{V}(E))} \{ \mu'_k(T \cap D) - \mu'_k(T) \} \right\}}{\text{tr}}$$

Où tr est le nombre de relations d'interactions de redondance entre les indicateurs appartenant à différents sous éléments, ce nombre est calculé comme suit :

$$tr = \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(\mathcal{V}(E))) \cup g_j(\mathcal{V}(E))|$$

De cette façon, sur la base des valeurs des coefficients de capacité déjà calculées pour les vecteurs des indicateurs liés à des sous-éléments, la capacité, qui modélise les poids et les interactions entre les indicateurs de la fonction globale est déduite.

Vérification des propriétés de la capacité ainsi inférée

$$- \mu(\emptyset) = \sum_k \mu'_k(g_k(\emptyset)) = \sum_k \mu'_k(\emptyset)$$

$$\text{Or, } \mu'_k(\emptyset) = \frac{\mu_k(\emptyset)}{\mu_k(g_k(\mathcal{V}(E)))} * \frac{|g_k(\mathcal{V}(E))|}{|U_k g_k(\mathcal{V}(E))|} = 0 \text{ car } \mu_k(\emptyset) = 0$$

$$\text{D'où, } \mu(\emptyset) = 0,$$

$$- \mu(\mathcal{V}(E)) = \sum_{k=1}^{n_E} \mu'_k(g_k(\mathcal{V}(E))) - \theta \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(\mathcal{V}(E))) \cup g_j(\mathcal{V}(E))| + \sigma \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^+(g_k(\mathcal{V}(E))) \cup g_j(\mathcal{V}(E))| = \sum_{k=1}^{n_E} \mu'_k(g_k(\mathcal{V}(E))) \text{ étant donné la condition de normalité exprimée ci-dessus,}$$

Et comme :

$$\sum_{k=1}^{n_E} \mu'_k(g_k(\mathcal{V}(E))) = \sum_{k=1}^{n_E} \frac{\mu_k(g_k(\mathcal{V}(E)))}{\mu_k(g_k(\mathcal{V}(E)))} * \frac{|g_k(\mathcal{V}(E))|}{|U_k g_k(\mathcal{V}(E))|} = \frac{1}{|U_k g_k(\mathcal{V}(E))|} \sum_{k=1}^{n_E} |g_k(\mathcal{V}(E))| = 1$$

$$\text{Car } \forall i, j \in \{1, \dots, n_E\}, i \neq j, g_i(\mathcal{V}(E)) \cap g_j(\mathcal{V}(E)) = \emptyset \text{ et donc } |U_k g_k(\mathcal{V}(E))| = \sum_{k=1}^{n_E} |g_k(\mathcal{V}(E))|$$

$$\text{Alors } \mu(\mathcal{V}(E)) = 1$$

- Soit $C1 \subseteq C2 \subseteq \mathcal{V}(E)$ alors :

$$\begin{aligned} \mu(C2) - \mu(C1) &= \sum_{k=1}^{n_E} \mu'_k(g_k(C2)) - \sum_{k=1}^{n_E} \mu'_k(g_k(C1)) \\ &- \theta \left[\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C2)) \cap g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C1)) \cap g_j(C1)| \right] \\ &+ \sigma \left[\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^+(g_k(C2)) \cap g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^+(g_k(C1)) \cup g_j(C1)| \right] \end{aligned}$$

Et comme $C1 \subseteq C2$ alors $g_k(C1) \subseteq g_k(C2)$, d'où le nombre de successeurs positifs de $g_k(C2)$ est supérieur ou égal à celui de $g_k(C1)$.

$$\sigma \left[\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^+(g_k(C2)) \cap g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^+(g_k(C1)) \cup g_j(C1)| \right] \geq 0$$

La condition de monotonie donne aussi :

$$\forall k = 1 \dots n_E, \mu'_k(g_k(C2)) - \mu'_k(g_k(C1)) \geq \theta * \text{tr}$$

$$\sum_{k=1}^{n_E} \mu'_k(g_k(C2)) - \mu'_k(g_k(C1)) \geq \theta * \text{tr}$$

Or,

$$\begin{aligned} & \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(\mathcal{V}(E))) \cap g_j(\mathcal{V}(E))| \\ & \geq \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C2)) \cap g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C1)) \cup g_j(C1)| \\ & \theta \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(\mathcal{V}(E))) \cap g_j(\mathcal{V}(E))| \\ & \geq \theta * \left[\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C2)) \cap g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C1)) \cup g_j(C1)| \right] \\ & \theta * \text{tr} \geq \theta \left[\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C2)) \cup g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C1)) \cup g_j(C1)| \right] \end{aligned}$$

D'où,

$$\begin{aligned} & \sum_{k=1}^{n_E} \mu'_k(g_k(C2)) - \sum_{k=1}^{n_E} \mu'_k(g_k(C1)) \\ & - \theta \left[\sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C2)) \cap g_j(C2)| - \sum_{k=1}^{n_E-1} \sum_{j=k+1}^{n_E} |\mathcal{S}^-(g_k(C1)) \cup g_j(C1)| \right] \geq 0 \end{aligned}$$

Nous concluons ainsi que $\mu(C2) - \mu(C1) \geq 0$

Exemple :

Afin d'illustrer la technique précédemment définie, nous l'appliquons dans un contexte réel, sur la plate-forme TELMA. L'exemple choisi pour la mise en œuvre est simple mais assez démonstratif. La plate-forme TELMA (située au CRAN [Levrat, 2007]) simule un procédé

industriel de déroulement et d'estampage / de coupe de bande de métal pour l'industrie automobile. Le système est composé de quatre sous-systèmes: Changement des bobines, accumulation de la bande, estampage/coupe et avance (Figure 44).

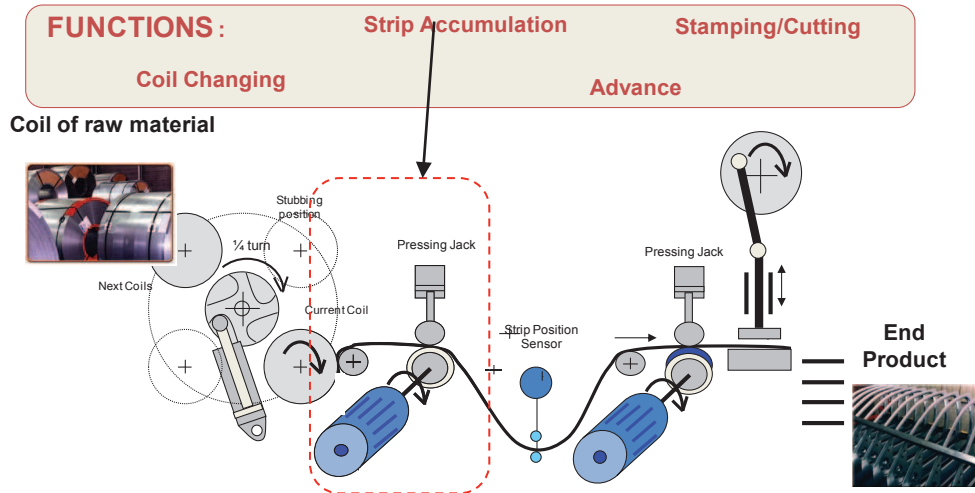


Figure 44 : Fonction d'estampage/coupe dans la plateforme TELMA.

La fonction d'accumulation de bande a été choisie pour illustrer la technique d'inférence verticale des coefficients de la capacité. Une telle fonction est essentiellement composée de trois sous-éléments montés en série: un moteur qui produit une rotation, un rouleau qui tire la bande au moyen de cette rotation transmise par un système de poulies/courroie située entre les deux sous-éléments. Ces trois sous-éléments sont reliés entre eux par des flux échangés comme le montre la Figure 45.

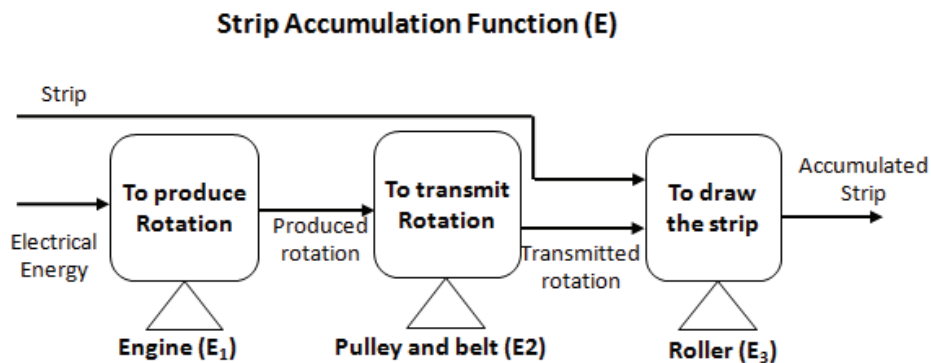


Figure 45 : Décomposition de la fonction « accumulation de la bande ».

A chaque sous élément est accordé un bilan de santé : $\mathcal{V}(E_1)$, $\mathcal{V}(E_2)$ et $\mathcal{V}(E_3)$ ainsi qu'un index de santé. A partir de tels bilans, le bilan de santé de la fonction globale est défini (Tableau 17). Nous illustrons comment la capacité de E est déduite.

$\mathcal{V}(E_1)$	$\mathcal{V}(E_2)$	$\mathcal{V}(E_3)$	$\mathcal{V}(E)$
I11: Energie consommée	I21: Couple en entrée	I31: Couple en entrée	I1: Vitesse d'accumulation (I33, I34)
I12: Couple	I22: Vitesse en entrée	I32: Vitesse en entrée	I2: Energie consommée (I11, I33)
I13: Vitesse	I23: Couple en sortie	I33: Niveau du stock	I3: Glissement total (I25, I35)
I14: Vibration	I24: Vitesse en sortie	I34: Vitesse de la bande	I4: Niveau vibratoire (I14)
	I25: Glissement	I35: Glissement	
	I26: Usure		

Tableau 17 : Les indicateurs pour chaque élément

Une étape importante avant de commencer le calcul de la capacité est de représenter les interactions négatives et positives entre les indicateurs des sous éléments (Figure 46) issues de leurs capacités correspondantes initiales, μ_k .

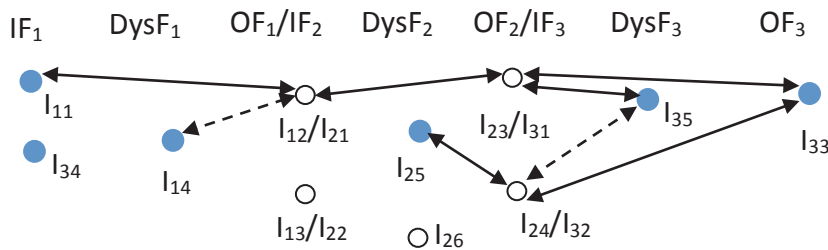


Figure 46 : Les interactions mutuelles entre les indicateurs des sous éléments en série.

Les valeurs des coefficients de μ'_k sont calculés à partir de ceux de μ_k (Tableau 18).

μ'_1	μ'_2	μ'_3
$\mu'_1(I_{11}) = \frac{\mu_1(I_{11})}{\mu_1(I_{11}, I_{14})} * \frac{2}{6} = 0.15$	$\mu'_2(I_{25}) = \frac{\mu_2(I_{25})}{\mu_2(I_{25})} * \frac{1}{6} = 0.16$	$\mu'_3(I_{33}) = \frac{\mu_3(I_{33})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.21$
$\mu'_1(I_{14}) = \frac{\mu_1(I_{14})}{\mu_1(I_{11}, I_{14})} * \frac{2}{6} = 0.17$		$\mu'_3(I_{34}) = \frac{\mu_3(I_{34})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.17$
$\mu'_1(I_{11}, I_{14}) = \frac{\mu_1(I_{11}, I_{14})}{\mu_1(I_{11}, I_{14})} * \frac{2}{6} = 0.33$		$\mu'_3(I_{35}) = \frac{\mu_3(I_{35})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.19$
		$\mu'_3(I_{33}, I_{34}) = \frac{\mu_3(I_{33}, I_{34})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.28$
		$\mu'_3(I_{33}, I_{35}) = \frac{\mu_3(I_{33}, I_{35})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.36$
		$\mu'_3(I_{34}, I_{35}) = \frac{\mu_3(I_{34}, I_{35})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.41$
		$\mu'_3(I_{33}, I_{34}, I_{35}) = \frac{\mu_3(I_{33}, I_{34}, I_{35})}{\mu_3(I_{33}, I_{34}, I_{35})} * \frac{3}{6} = 0.5$

Tableau 18 : Les coefficients de capacité des sous éléments.

Afin de satisfaire les conditions de normalisation et de monotonie, les paramètres σ et θ sont fixés à 0.02.

En effet :

- La condition de normalité donne:

$$\theta \sum_{k=1}^2 \sum_{j=k+1}^3 |\mathcal{S}^-(g_k(\mathcal{V}(E))) \cap g_j(\mathcal{V}(E))| = \sigma \sum_{k=1}^2 \sum_{j=k+1}^3 |\mathcal{S}^+(g_k(\mathcal{V}(E))) \cap g_j(\mathcal{V}(E))|$$

$$\theta (|\mathcal{S}^-(g_1(\mathcal{V}(E))) \cup g_2(\mathcal{V}(E))| + |\mathcal{S}^-(g_1(\mathcal{V}(E))) \cap g_3(\mathcal{V}(E))| + |\mathcal{S}^-(g_2(\mathcal{V}(E))) \cap g_3(\mathcal{V}(E))|)$$

$$= \sigma (|\mathcal{S}^+(g_1(\mathcal{V}(E))) \cup g_2(\mathcal{V}(E))| + |\mathcal{S}^+(g_1(\mathcal{V}(E))) \cap g_3(\mathcal{V}(E))| + |\mathcal{S}^+(g_2(\mathcal{V}(E))) \cap g_3(\mathcal{V}(E))|)$$

$$\theta (|\mathcal{S}^-(\{I_{11}, I_{14}\}) \cap \{I_{25}\}| + |\mathcal{S}^-(\{I_{11}, I_{14}\}) \cap \{I_{33}, I_{34}, I_{35}\}| + |\mathcal{S}^-(\{I_{25}\}) \cap \{I_{33}, I_{34}, I_{35}\}|) =$$

$$\sigma (|\mathcal{S}^+(\{I_{11}, I_{14}\}) \cup \{I_{25}\}| + |\mathcal{S}^+(\{I_{11}, I_{14}\}) \cap \{I_{33}, I_{34}, I_{35}\}| + |\mathcal{S}^+(\{I_{25}\}) \cap \{I_{33}, I_{34}, I_{35}\}|)$$

$$\theta (|\{I_{33}, I_{35}\} \cap \{I_{25}\}| + |\{I_{33}, I_{35}\} \cup \{I_{33}, I_{34}, I_{35}\}| + |\{I_{33}\} \cap \{I_{33}, I_{34}, I_{35}\}|)$$

$$= \sigma (|\{I_{33}, I_{35}\} \cup \{I_{25}\}| + |\{I_{33}, I_{35}\} \cup \{I_{33}, I_{34}, I_{35}\}| + |\{I_{35}\} \cap \{I_{33}, I_{34}, I_{35}\}|)$$

$$\theta(0 + 2 + 1) = \sigma(0 + 2 + 1)$$

$$\theta = \sigma$$

- La condition de monotonie donne :

$$\theta \leq \min \left\{ \min \left\{ \left(\mu'_1(I_{11}, I_{14}) - \mu'_1(I_{14}) \right), \left(\mu'_1(I_{11}, I_{14}) - \mu'_1(I_{11}) \right) \right\}, \min \left\{ \left(\mu'_3(I_{33}, I_{34}, I_{35}) - \mu'_3(I_{34}, I_{35}) \right), \left(\mu'_3(I_{33}, I_{34}, I_{35}) - \mu'_3(I_{33}, I_{34}) \right), \left(\mu'_3(I_{34}, I_{35}) - \mu'_3(I_{35}) \right), \left(\mu'_3(I_{34}, I_{35}) - \mu'_3(I_{33}, I_{35}) \right), \left(\mu'_3(I_{34}, I_{35}) - \mu'_3(I_{33}) \right), \left(\mu'_3(I_{33}, I_{35}) - \mu'_3(I_{33}) \right), \left(\mu'_3(I_{33}, I_{34}) - \mu'_3(I_{33}) \right), \left(\mu'_3(I_{33}, I_{34}) - \mu'_3(I_{34}) \right) \right\} \right\} / 3$$

$$\theta \leq \min \{ \min \{ 0.16, 0.18 \}, \min \{ 0.09, 0.14, 0.22, 0.24, 0.22, 0.17, 0.15, 0.07, 0.11 \} \} / 3$$

$$\theta \leq 0.023$$

Le Tableau 19 montre les valeurs calculées pour les coefficients de la capacité de la fonction d'accumulation.

$\wp(\mathcal{V}(E))$	μ	$\wp(\mathcal{V}(E))$	μ
{I ₁ }	0.28	{I ₂ , I ₄ }	0.54
{I ₂ }	0.34	{I ₃ , I ₄ }	0.56
{I ₃ }	0.37	{I ₁ , I ₂ , I ₃ }	0.77
{I ₄ }	0.17	{I ₁ , I ₂ , I ₄ }	0.61
{I ₁ , I ₂ }	0.41	{I ₁ , I ₃ , I ₄ }	0.87
{I ₁ , I ₃ }	0.66	{I ₂ , I ₃ , I ₄ }	0.85
{I ₁ , I ₄ }	0.47	{I ₁ , I ₂ , I ₃ , I ₄ }	1
{I ₂ , I ₃ }	0.63		

Tableau 19 : Les valeurs de la capacité pour la fonction globale

Ces valeurs modélisent les interactions attendues entre les indicateurs de la fonction « Accumuler la Bande ». En effet, une synergie est trouvée entre les deux indicateurs I1 et I3 car $\mu(\{I_1, I_3\}) > \mu(\{I_1\}) + \mu(\{I_3\})$. Cette synergie résulte de la synergie initiale entre la vitesse de rotation et l'indicateur de glissement du troisième sous élément. De la même manière, les interactions de redondance portées par les flux échangées sont reflétées par cette inférence. Par exemple, sont en redondance I2 et I3 car $\mu(\{I_2, I_3\}) > \mu(\{I_2\}) + \mu(\{I_3\})$.

Cas	I ₁₁ / I ₁₂ / I ₁₃ / I ₁₄	IS _{E_k}	I ₁ / I ₂ / I ₃ / I ₄	IS _E
	I ₂₁ / I ₂₂ / I ₂₃ / I ₂₄ / I ₂₅ / I ₂₆			
	I ₃₁ / I ₃₂ / I ₃₃ / I ₃₄ / I ₃₅			
1	0.9/ 0.9/ 0.9/ 0.9	0.90	0.9/0.9/0.6/0.9	0.78
	0.9/ 0.9/ 0.9/ 0.9/ 0.6/ 0.8	0.85		
	0.9/ 0.9/ 0.9/ 0.9/ 0.6	0.86		
2	0.9/ 0.9/ 0.9/ 0.4	0.70	0.9/0.9/0.9/0.4	0.78
	0.9/ 0.9/ 0.9/ 0.9/ 0.9/ 0.9	0.90		
	0.9/ 0.9/ 0.9/ 0.9/ 0.9	0.90		
3	0.9/ 0.5/ 0.9/ 0.9	0.70	0.7/0.7/0.9/0.9	0.80
	0.5/ 0.9/ 0.5/ 0.9/ 0.9/ 0.9	0.72		
	0.5/ 0.9/ 0.6/ 0.9/ 0.9	0.70		

Tableau 20 : Propagation de la déviation à travers les deux niveaux hiérarchiques.

Le Tableau 20 montre quelques situations (cas de 1 à 3) représentées par différentes valeurs d'indicateurs, pour les trois sous éléments. Un index de santé est calculé pour chaque sous élément (IS_{E1}, IS_{E2}, IS_{E3}). Pour chaque situation, l'index de santé global pour la fonction « accumuler la bande » est aussi calculé en utilisant la capacité inférée.

Dans le premier cas, pour un moteur en bon état, et avec un faible glissement dans le système de poulie/courroie et dans le rouleau, les indexes de santé présentent des déviations localement faibles, mais la fonction prend en compte la synergie entre ces deux phénomènes et le glissement au niveau de la fonction globale a plus d'effet sur son index de santé. Le second cas vise à montrer que les vibrations, même si elles ont une grande importance dans l'évaluation de l'état de santé du moteur, leur projection sur la santé de la fonction globale n'a pas le même impact. Enfin (pour le cas 3), la perte de couple au niveau du moteur est transmise aux deux sous-éléments qui se succèdent. Ceci est signalé localement pour chaque sous-élément. Toutefois, la capacité de la fonction globale tient compte de la redondance entre ces déviations corrélées, ainsi l'index de santé pour la fonction globale est moins dégradé (plus proche de 1).

4.2.3. Cas série- parallèle

Des éléments sont en série-parallèle lorsque la configuration globale ne peut pas être classée dans l'un des deux cas de base, à savoir le cas série ou le cas parallèle. Pour ce cas de figure, il faut procéder à une segmentation des éléments en des blocs d'éléments permettant de retrouver des configurations élémentaires (sous-éléments en série ou en parallèle). Une règle simple de découpage consiste à minimiser le nombre total de blocs virtuels de sous-éléments positionnés en série ou en parallèle. Ensuite, un bilan de santé virtuel est associé à chaque bloc ainsi formé. Ce bilan est composé des indicateurs de ce bloc contribuant à des indicateurs du bilan de santé du niveau supérieur (de la fonction globale). Pour le cas parallèle les coefficients d'indicateurs sont normalisés par rapport à la capacité des indicateurs contributeur de chaque sous-élément du bloc. En effet, pour les grands systèmes, seulement une partie des indicateurs de ces éléments contribue au bilan de santé de l'élément global.

La méthode d'inférence des coefficients de capacité du cas série est appliquée tel que défini dans la section 4.2.2. car la normalisation des coefficients de capacité par rapport à celui des indicateurs contributeurs est déjà considérée.

Ensuite, les blocs de sous-éléments ainsi définis, associés à leurs bilans de santé virtuels sont considérés comme de nouveaux sous-éléments. Si ces blocs sont organisés en parallèle ou en série, l'inférence des coefficients de la capacité liée à l'élément global est effectuée, sinon une nouvelle itération débute et un nouveau regroupement de blocs est effectué. Les itérations s'arrêtent avec l'aboutissement à une structure simple permettant d'appliquer directement l'une des deux méthodes correspondantes au cas série ou parallèle.

En prenant l'exemple de la Figure 47, le regroupement des sous éléments en blocs est effectué en une itération.

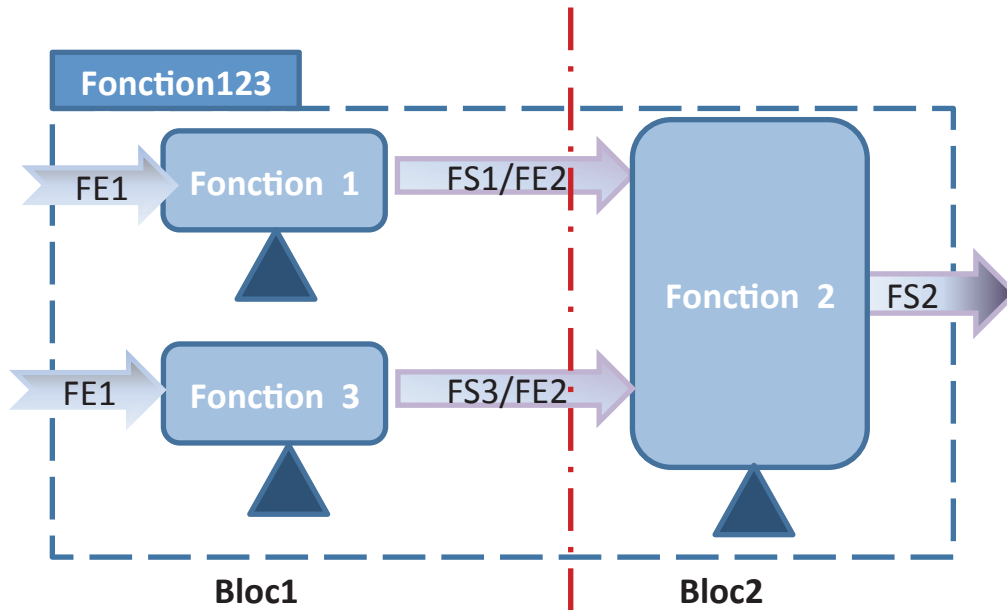


Figure 47 : Répartition de la fonction en deux couches parallèles

En effet les deux fonctions 1 et 3 forment un premier bloc auquel la méthode du cas parallèle peut être appliquée. Ensuite, ce bloc est lié à la fonction 2 en série. Ce qui permet d'appliquer directement la méthode d'inférence des coefficients de la capacité de la fonction globale à partir de ceux du bilan virtuel du bloc 1 et ceux de la fonction 2.

Un exemple d'illustration détaillé pour un cas réel sera présenté dans le chapitre 4 sur notre cas d'application.

5. Conclusion

Afin de formaliser la notion de bilan de santé pour des systèmes techniques multi-composants, une structure simplifiée et générique pour présenter les principaux paramètres qui peuvent être inclus dans le suivi de l'état de santé a été décrite en début de ce chapitre. Cette structure regroupe des indicateurs liés à la performance globale et ceux orientés vers la dégradation dans un vecteur d'indicateurs attaché à tout élément quelque soit sa nature ou le niveau hiérarchique. Un index de santé issu de l'agrégation des indicateurs dans un vecteur doit être calculé. A cette fin, nous avons présenté une approche robuste et générique de normalisation des indicateurs à travers la quantification des dérives des valeurs prises par les indicateurs par rapport aux références de fonctionnement nominal et de faute. Concernant la fonction d'agrégation, la contribution présentée dans ce chapitre s'intéresse à la tâche d'identification des coefficients de la capacité pour tous les vecteurs de la structure du bilan de santé. L'apprentissage non supervisé à base d'entropie relative proposé dans le domaine d'aide à la décision a été adapté au contexte de la fusion des indicateurs. Une approche originale permettant de déduire les capacités des éléments dans

les niveaux hiérarchiques supérieurs à partir de ceux identifiés pour les composants est aussi proposée.

Le chapitre suivant porte sur la mise en œuvre de ces propositions sur un système réel. Ce cas d'application est extrait du projet BMCI. L'objectif est de montrer la faisabilité de la méthode proposée et d'établir un premier bilan en termes de valeur ajoutée et de limites.

Chapitre 4 : Application des contributions sur le banc d'essai BMCI

1. Introduction

Dans un objectif de monter la faisabilité de l'ensemble de nos propositions, ce chapitre présente leur application sur un système réel, à savoir celui d'un moteur diesel marin. Ce système permet un passage à une dimension quasi industrielle caractérisée par la complexité des données vu la multitude des composants et leurs interactions fonctionnelles.

Nous proposons ainsi la mise en œuvre de l'approche d'élaboration d'un bilan de santé multi niveaux hiérarchiques sur un banc d'essai lié au projet BMCI. Le suivi de la santé est effectué « hors ligne » puisque nous avons travaillé sur des données historisées issues du système de surveillance en ligne (les outils utilisés pour la détection/agrégation ne pouvant se connecter à ce jour au système de surveillance en ligne). Cependant, le principe général des tests est de se rapprocher de la réalité du métier naval par des simulations mécaniques des contraintes de fonctionnement (modes de dégradation et conditions opérationnelles et environnementales changeantes). En ce sens, la première section de ce chapitre décrit le banc d'essai sur lequel les tests ont été réalisés. Ensuite, la méthode de détection contextuelle des dérives des indicateurs est appliquée. Afin de procéder à l'étape d'agrégation des indicateurs, la troisième partie de ce chapitre montre la mise en œuvre de l'approche d'identification des coefficients des capacités par apprentissage non supervisé. A partir des coefficients identifiés pour le niveau hiérarchique inférieur, les capacités nécessaires à l'agrégation des indicateurs des niveaux supérieurs sont inférées. Les résultats obtenus sont discutés afin de noter les points forts et les éventuelles limites.

2. Présentation du cas d'application

Dans le cadre du projet BMCI, avant la phase de validation du système complet en mer, une phase de validation sur banc d'essai a été définie. Outre le système BMCI, cette phase nécessite trois éléments clés pour être menée à bien :

- Un moteur diesel représentatif d'un système marin,
- Une cellule d'essai,
- Un protocole de tests.

Le moteur sélectionné pour la campagne d'essai est le moteur 6 M26 SR P1 de la Société Internationale des Moteurs Baudouin (SiMB). Le moteur diesel a été retenu pour les tests du banc BMCI car il est fréquemment mis en service pour plusieurs fonctions d'un navire dont celle de propulsion [Morachioli, 2011]. La SiMB étant partenaire du projet, le choix du moteur s'est porté naturellement sur un des leurs car en premier lieu la SiMB équipe un nombre important de bâtiments opérationnels. De plus, la SiMB, à travers leurs personnels impliqués, a fait bénéficier au projet de son savoir et savoir-faire sur les moteurs diesel mais également a permis de mettre à disposition du projet des pièces sur mesure permettant de simuler mécaniquement des dégradations/défaillances du moteur. Une cellule d'essai a été mise à disposition par l'Ecole Centrale Marseille (ECM). Cette cellule comporte un système complet d'acquisition en redondance avec le système BMCI et permettant ainsi de valider ce dernier. L'architecture du banc d'essais BMCI est présentée à la Figure 48. Le banc comporte, outre le moteur diesel et ses accessoires, un premier système d'acquisition propre à l'ECM, permettant la conduite d'essais (« Système d'Acquisition ECM ») et poste de contrôle commande). Le banc intègre également les éléments propres au projet à savoir :

- Le boîtier BMCI qui assure l'acquisition des signaux issus des capteurs,
- Les modules d'acquisition complémentaires liés au boîtier,
- Le serveur BMCI intégrant la base de données, l'exécution des traitements, et le serveur d'Interface Homme Machine (IHM),
- Le poste Utilisateur (Client BMCI), qui permet d'accéder au travers d'un navigateur web aux interfaces IHM de l'application.

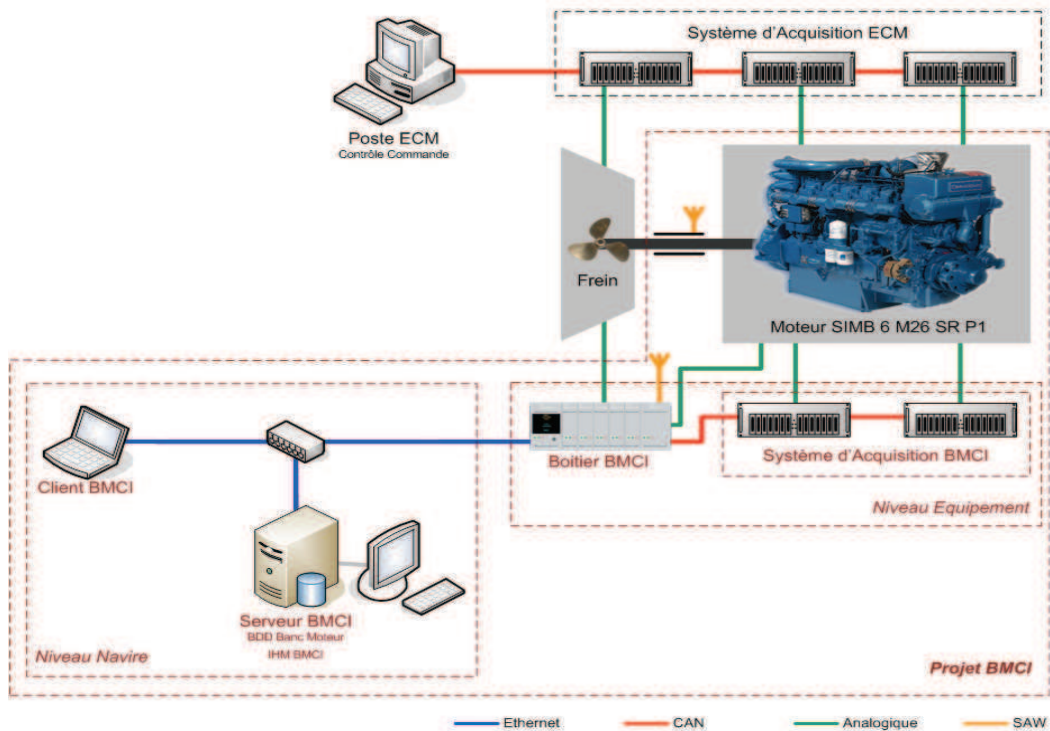


Figure 48 : Architecture du banc d'essai [Peysson, 2011]

Enfin, le protocole de test a été défini par les acteurs du projet ayant besoin de mesures représentatives de modes de fonctionnement conjointement avec les experts des moteurs diesel. Ainsi, les capteurs nécessaires pour acquérir les données brutes sur les différents modes de fonctionnement ont été définis par ces acteurs. Ensuite, les indicateurs indispensables pour signaler l'apparition des principaux modes de dégradation [Khelil, 2010] ont été construits. Les principaux indicateurs sont répartis à travers les bilans de santé présentés dans la suite.

Les données brutes issues des mesures des capteurs sur le banc ou encore des mesures manuelles sont enregistrées sous forme de fichiers Excel sur un serveur accessible aux différents partenaires du projet. Chaque fichier regroupe les données enregistrées tout au long des tests de fonctionnement nominal ou dégradé. En ce sens, les dégradations ont été reproduites expérimentalement (fuite d'air en aval compresseur, encrassement filtre à air, fuite carburant, fuite eau...) pour en étudier les effets sur les mesures du moteur. Afin d'atteindre le mode de fonctionnement dégradé, le moteur a subi au préalable des modifications pour simuler mécaniquement les dysfonctionnements. En outre, différentes conditions environnementales et opérationnelles ont été utilisées lors des tests afin de pouvoir évaluer le comportement du moteur vis-à-vis de tels changements de contexte de fonctionnement (différents régimes, essais réalisés sous différentes températures ambiantes...).

A partir de ces fichiers de données, de très grandes tailles, les mesures sont labellisées suivant le mode de fonctionnement (nominal ou en présence d'une dégradation), le contexte opérationnel (régime moteur) et environnemental (température de la cellule ou de l'eau brute de refroidissement). Chaque lot de données, ainsi labellisé, a été partitionné en sous lots correspondant chacun à un sous-système du moteur. Ensuite les indicateurs ont été construits et enregistrés dans des fichiers csv. Ces fichiers csv alimentent les fonctions de détection et d'agrégation codés sous R (le package Kappalab intégré à ce logiciel nous a facilité la mise en œuvre des calculs relatifs à l'intégrale de Choquet).

2.1. Description du moteur diesel

Pour comprendre le fonctionnement du moteur diesel marin, quelques éléments sont à expliquer, plus particulièrement ceux concernant le fonctionnement du moteur diesel suralimenté à injection directe.

Le moteur du banc est un moteur à quatre temps qui appartient à la classe moyenne puissance (entre 100 et 450 chevaux avec une vitesse de rotation atteignant 1800 tr/min), à six cylindres montés en ligne, suralimenté en air (avec refroidisseur d'air aspiré comme le montre la Figure 49) et dont l'injection est mécanique.

Le moteur diesel est un moteur à combustion interne dont le fonctionnement est basé sur le principe d'auto-inflammation du carburant. Il développe ainsi sa puissance en convertissant en chaleur, l'énergie chimique contenue dans le carburant, puis en transformant cette chaleur en travail mécanique.

En plus de cette fonction principale de combustion, d'autres fonctions supports sont aussi nécessaires pour un moteur diesel comme l'admission d'air, l'alimentation en carburant, le refroidissement et la lubrification.

La suralimentation consiste à augmenter la quantité d'air admise au cylindre. Cela permet d'augmenter la puissance à une même vitesse de rotation, et en même temps, d'améliorer la combustion. Sur le banc test, cette procédure est assurée par un appareil du type centrifuge comportant une turbine entraînée par les gaz d'échappement et accouplée directement à un compresseur (turbocompresseur) (Figure 49).

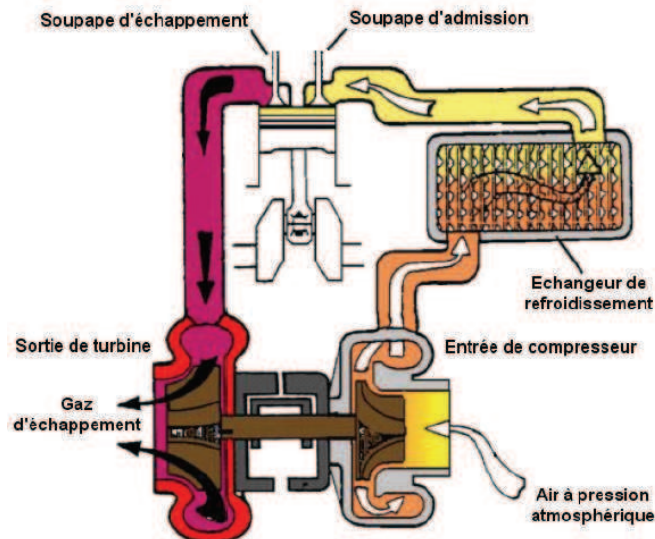


Figure 49 : Schéma de fonctionnement de la suralimentation par turbocompresseur [Sabe, 2006]

La plupart des moteurs Diesel actuels ont des systèmes de refroidissement fonctionnant en boucle fermée (Figure 50). Le rôle de ces systèmes est d'évacuer l'excès de chaleur hors du compartiment moteur. Ils sont généralement composés d'une pompe à eau, d'un ou plusieurs refroidisseurs (radiateur, échangeurs), d'une tuyauterie et d'un thermostat. [Morachioli, 2011].

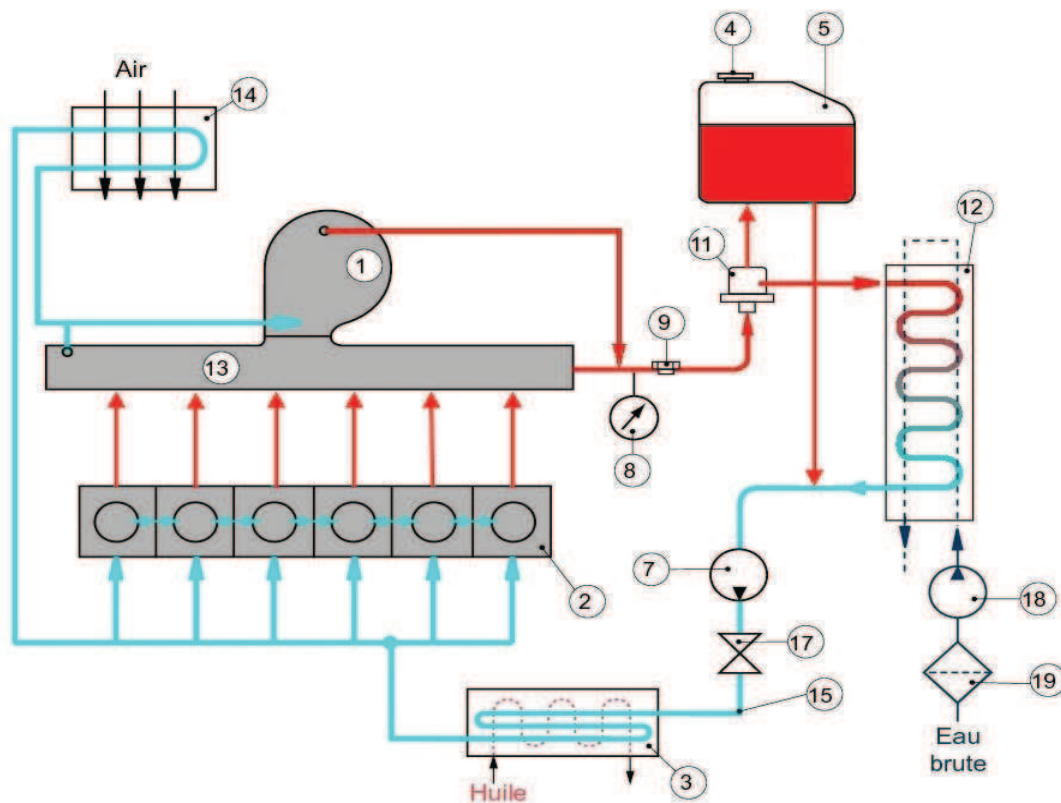


Figure 50 : Schéma du sous système de refroidissement du moteur diesel

La fonction de lubrification dans un moteur est aussi très importante afin d'éviter la fusion des métaux des pièces en mouvement en présence de la chaleur et des frottements. Le rôle du système de lubrification est de lubrifier les surfaces anti-frictionnées des paliers, de refroidir certains organes, d'assurer l'étanchéité du système piston-segments-chemise et de nettoyer, en évacuant les particules issues de l'usure et de la combustion lors de la vidange [Morachioli, 2011].

Afin de synthétiser les éléments décrits dans ce paragraphe et à partir d'une analyse fonctionnelle du moteur, une architecture simplifiée du moteur diesel et de ses fonctions support est présentée dans le paragraphe suivant. Cette architecture fonctionnelle permet de montrer le positionnement mutuel des différents sous systèmes (en série, en parallèle ou en série-parallèle) ainsi que les flux échangés entre eux. C'est sur cette architecture technique que nous avons conduit le déploiement de notre approche d'élaboration d'un bilan de santé hiérarchique.

2.2. Structure hiérarchique du bilan de santé

L'analyse fonctionnelle du moteur diesel du banc d'essai donne lieu aux deux structures : arborescente (Figure 51), et décomposition horizontale (Figure 52).

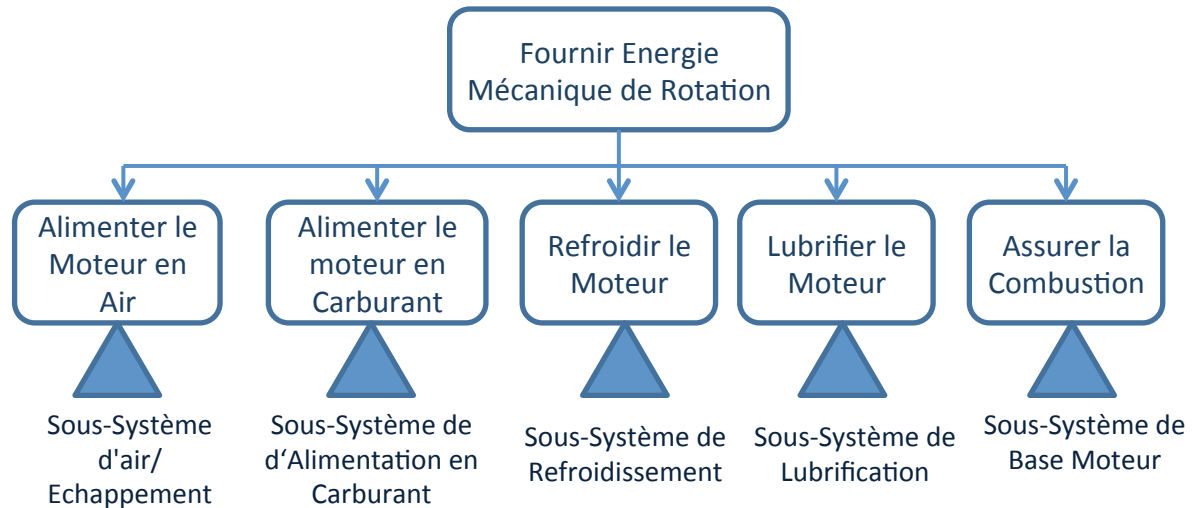


Figure 51 : Structure arborescente du moteur diesel

Cette première structure permet de supporter l'architecture du bilan de santé multi niveaux hiérarchiques. En effet, à chaque fonction (ou élément) de cette structure et pour les deux niveaux, un bilan de santé est à affecter.

D'un autre point de vue, la Figure 52 montre les interactions entre les fonctions élémentaires à travers les flux échangés. Ces interactions définissent ainsi les positionnements de ces fonctions les unes par rapport aux autres (fonctions en série, en parallèle ou en série-parallèle). Une telle structure sert de base pour :

- Identifier les flux entrant et sortant pour chaque fonction. Leurs principales propriétés servent de base pour la définition des indicateurs de santé,
- Le positionnement mutuel entre les fonctions appartenant à un même élément du niveau hiérarchique supérieur permet de choisir la méthode d'inférence des coefficients de la capacité relative à cet élément.

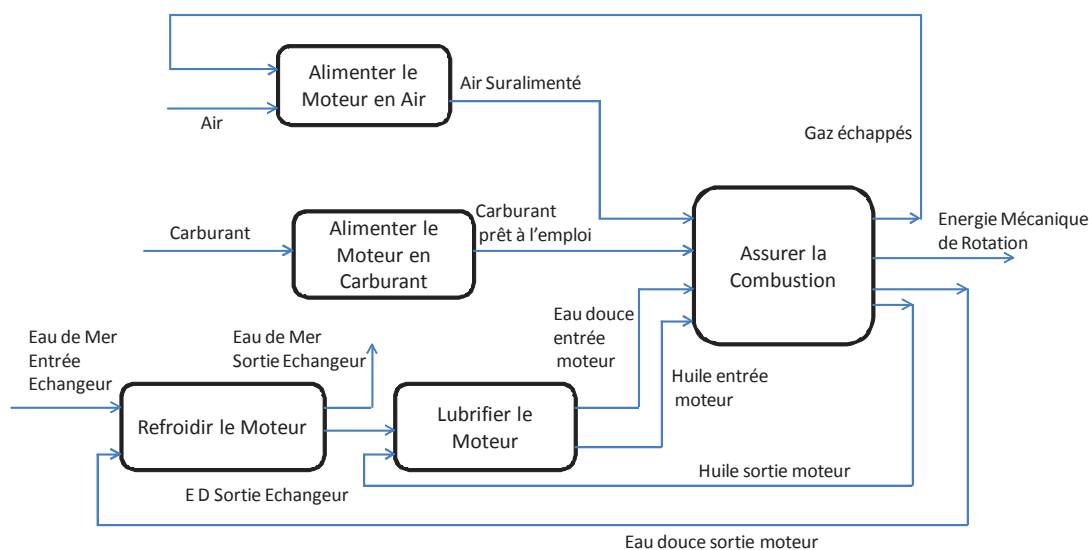


Figure 52 : Organisation des flux échangés entre les fonctions du moteur

Il est important de noter ici que la décomposition des fonctions est limitée d'une part par le choix du modélisateur (concepteur du système BMCI) et d'autre part, par l'instrumentation disponible.

Dans la suite, pour chaque fonction (ou élément) de la structure arborescente (Figure 51), les indicateurs qui lui sont associés sont résumés dans un tableau. Ces indicateurs sont définis par les partenaires du projet experts dans le domaine des moteurs marins. La culasse du moteur a subi quelques transformations afin de pouvoir installer des capteurs supplémentaires à ceux habituellement utilisés pour la surveillance de ce type de moteur diesel sans affecter bien évidemment ses caractéristiques fonctionnelles. Les indicateurs définis pour le moteur dans son ensemble ont été répartis sur les bilans des différents éléments de l'architecture du système. Ces indicateurs sont aussi classés selon les trois types : les indicateurs de performance, de dégradation et contextuels. Chaque tableau d'indicateurs contient les informations suivantes :

- La désignation de la fonction,
- la liste des indicateurs avec désignation et nature (de performance ou de dégradation),
- les indicateurs contextuels susceptibles d'affecter le comportement des indicateurs de performance ou de dégradation,
- les dégradations susceptibles d'être indiquées par la dérive de tels indicateurs.

2.2.1. Alimenter le Moteur en Air

D'après une analyse dysfonctionnelle (AMDEC, circuits d'isolation des dégradations [Khelil, 2010]), les dégradations les plus fréquentes affectant le sous système d'air/échappement sont les suivantes :

- Dégradation du turbocompresseur,
- Encrassement filtre à air,
- Dégradation du refroidisseur air,
- Dégradation du silencieux,
- Fuite d'air ou des gaz d'échappement...

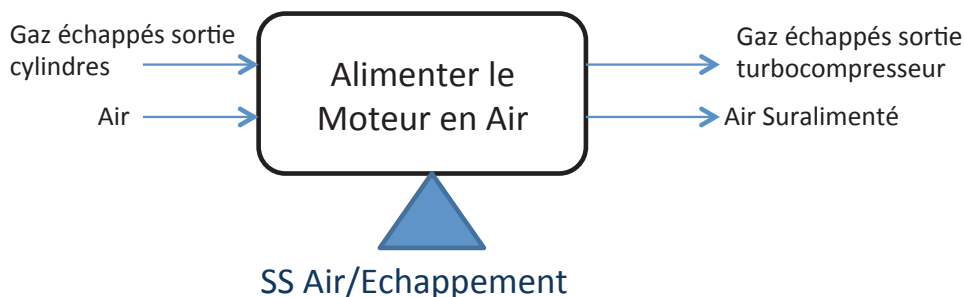


Figure 53 : Sous-système admission en air et échappement

Le Tableau 21 fournit l'ensemble des indicateurs utilisés pour évaluer la santé du sous-système d'admission en air et échappement :

Fonction	Indicateurs	Désignation	Nature	Dégradation
Admission en Air et Echappement	Pression des gaz d'échappement sortie du turbocompresseur	P4	performance	Silencieux/ Turbo / Fuites air ou gaz
	Température maximale des gaz d'échappement sortie des cylindres	Tech_max	performance	Turbo/ Fuites air ou gaz
	Pression air suralimenté	P2	performance	Turbo/ Fuites air ou gaz
	Pression différentielle filtre à air	P1	dégradation	Encrassement filtre à air
	Rendement du Refroidisseur d'air Suralimenté RAS	RAS	dégradation	Echangeur
	Température de la cellule	T1	environnemental	-

Tableau 21 : Les indicateurs de santé pour le sous système d'Air/Echappement

2.2.2. Alimenter le Moteur en Carburant

Concernant le sous-système d'alimentation en carburant, la Figure 54 montre une présentation simplifiée de la fonction accomplie par ce sous-système.

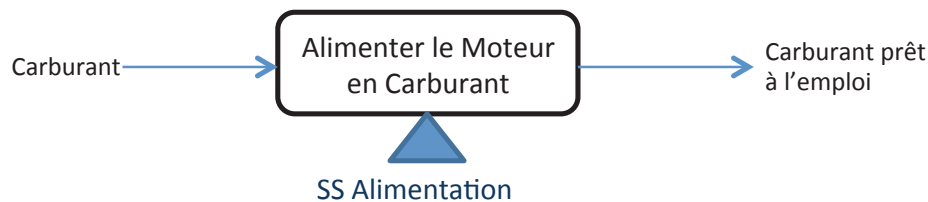


Figure 54 : Sous-système alimentation en carburant

Le bilan de santé élaboré pour ce sous-système doit permettre la détection des dégradations les plus fréquentes [Khelil, 2010] à savoir :

- Avarie pompe alimentation,
- Fuite carburant,
- Encrassement du filtre à carburant,
- Dégradation de la qualité du carburant...

Les indicateurs définis pour ce sous-système sont présentés dans le Tableau 22.

Fonction	Indicateurs	Désignation	Nature	Dégradation
Alimentation	Débit carburant consommé	Dcarb	performance	Fuite carburant/ Injecteur
	Pression entrée pompe d'injection	Pepi	dégradation	Pompe alimentation/ Fuite de carburant

	Pression différentielle filtre carburant	Pfc	dégradation	Encrassement filtre à carburant/ Qualité du carburant
	Température de la cellule	T1	environnemental	-

Tableau 22 : Les indicateurs de santé pour le sous système d'Alimentation en carburant

2.2.3. Lubrifier le Moteur

Pour le sous-système de lubrification (Figure 55) le bilan de santé doit permettre de détecter les dégradations suivantes [Khelil, 2010] :

- Les fuites d'huile,
- L'encrassement filtre à huile,
- Dégradation refroidisseur huile...



Figure 55 : Sous-système de lubrification

Les indicateurs du bilan de santé qui permettent de détecter ces dégradations ou autres pertes de performances sont présentés dans le Tableau 23.

Fonction	Indicateurs	Désignation	Nature	Dégradation
Lubrification	Température eau douce entrée moteur	Teem	performance	Encrassement refroidisseur huile
	Température huile entrée moteur	Them	performance	Encrassement refroidisseur huile
	Pression différentielle filtre huile	Pfh	dégradation	Encrassement filtre huile
	Température eau brute entrée échangeur	Tebte	environnemental	-

Tableau 23 : Les indicateurs de santé pour le sous système de lubrification

2.2.4. Refroidir le Moteur

La Figure 56 illustre les flux d'entrée/sortie pour cette fonction supportée par le sous-système de refroidissement.

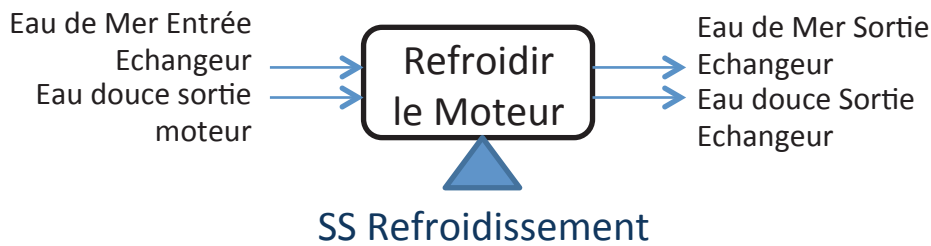


Figure 56 : Sous-système de refroidissement

Le bilan de santé élaboré pour ce sous-système doit permettre la détection des dégradations les plus fréquentes à savoir :

- Avaries échangeur eau brute/eau douce,
- Avarie thermostat,
- Fuite d'eau douce,
- Fuite eau brute,
- Avarie pompe eau brute,
- Encrassement filtre eau brute...

De la même manière, le bilan de santé pour la fonction de refroidissement est présenté dans le Tableau 24.

Fonction	Indicateurs	Désignation	Nature	Dégradation
Refroidissement	Température eau douce sortie moteur	Tesm	performance	Thermostat
	Pression eau brute entrée échangeur	Pebte	performance	Fuite EB/ Pompe
	Pression caisse d'expansion	Pce	dégradation	Fuite eau douce
	Efficacité de l'échangeur ED/EB	Eech	dégradation	Echangeur
	Température eau brute entrée échangeur	Tebte	environnemental	-

Tableau 24 : Les indicateurs de santé pour le sous système de refroidissement

2.2.5. Assurer la combustion

Cette fonction assure la transformation de l'énergie chimique du carburant en énergie mécanique de rotation par la combustion. Cette fonction (Figure 57) nécessite alors en entrée une alimentation en air et en carburant ainsi que d'un liquide de refroidissement ou encore d'huile de lubrification. Les gaz d'échappement et la puissance de rotation sont les principaux flux de sortie issus de cette fonction. En outre, les retours des liquides de refroidissement et de lubrification doivent être aussi considérés comme des flux sortants secondaires.

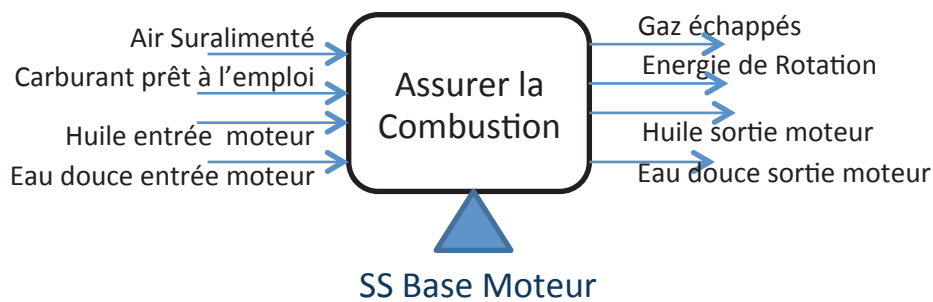


Figure 57 : Flux entrée/sortie pour la base moteur

Le Tableau 25 présente les indicateurs définis pour cette fonction afin de pouvoir détecter les dégradations les plus fréquentes.

Fonction	Indicateurs	Désignation	Nature	Dégradation
Base Moteur	Température eau douce entrée moteur	Peem	performance	-
	Température huile sortie moteur	Thsm	performance	Dégradation des coussinets des paliers, tout blocage bas moteur pouvant entraîner une élévation de température
	Pression air suralimenté entrée cylindres	P2	performance	-
	Pression entrée pompe d'injection	Pepi	performance	-
	Puissance Moteur	Pm	performance	Dégradation des coussinets des paliers, tout blocage bas moteur pouvant entraîner une élévation de température
	Débit dégazage carter	Blowby	dégradation	Dégradation du trio che/seg/pis
	L'asymétrie entre les températures des cylindre (Tech1,...Tech6)	Asy_tech	dégradation	Jeu aux soupapes anormal, usure de siège des soupapes, brûlure, dérèglement injecteurs et pompes à injection...
	Température de la cellule	T1	Environnemental	-

Tableau 25 : Les indicateurs de santé pour la base du moteur

2.2.6. Fournir Energie Mécanique de Rotation

Les indicateurs qui synthétisent la santé de l'ensemble du système « moteur diesel » sont fournis dans le Tableau 26. Les indicateurs à ce niveau d'abstraction sont plutôt de nature de performance (énergétique, de consommation d'huile) dont la dérive peut révéler plusieurs

dégradations. Les indicateurs définis pour le moteur Diesel ont été structurés à travers des bilans de santé pour les différents sous-systèmes. L'étape suivante consiste à les fusionner en utilisant les démarches proposées.

Fonction	Indicateurs	Désignation	Nature	dégradation
Système Moteur Diesel	La consommation spécifique	Cs	performance	Avaries injecteurs/ Fuites de carburant/ Air dans le circuit de carburant/ Avaries pompe d'alimentation
	Pression des gaz d'échappement sortie du turbocompresseur	P4	performance	Silencieux/ Turbo / Fuites
	Pression eau brute entrée échangeur	Pebte	performance	Filtre EB/ Pompe
	Consommation huile/carburant	Rhc	performance	Huile dégradé/ dégradation joint culasse/ usure segmentation
	Température eau brute entrée échangeur	Tebte	environnemental	-
	Température de la cellule	T1	environnemental	-

Tableau 26 : Les indicateurs de santé pour le moteur Diesel

3. Mise en œuvre de la méthode de détection

Dans cette partie, l'approche de détection est mise en œuvre pour les bilans de santé définis dans la partie précédente. En ce sens, dans un premier temps, les différents cas de figures pour la quantification des déviations sont présentés sur des indicateurs individuellement. Ensuite, une vue d'ensemble sur le système est donnée à travers des scénarios montrant l'effet de quelques dégradations sur le comportement des indicateurs de tout le système, et ce sous différentes conditions.

Afin d'appliquer la méthode de quantification de dérive avec la considération du contexte, nous allons considérer des indicateurs répondant aux critères suivants :

- 1^{er} cas : un indicateur de type I dont la dérive n'est pas impactée par le contexte,
- 2^{ème} cas : un indicateur de type I avec un contexte affectant le degré de la déviation par rapport au mode nominal,
- 3^{ème} cas : un indicateur de type II.

Les données exploitées sont celles contenues dans les différents fichiers csv qui viennent alimenter les fonctions de détection codées sous R.

3.1. Indicateur de type I sans impact du contexte

Afin d'illustrer le 1^{er} cas, nous considérons l'indicateur de pression différentielle filtre à air (P1) du bilan de santé de la fonction « Alimenter le Moteur en Air ».

Tout d'abord, le test de dépendance au contexte doit être effectué. La Figure 58 montre l'évolution de cet indicateur en faisant varier la température d'air en entrée de cette fonction (T1) entre les valeurs correspondantes aux conditions extrêmes. Cette température est le paramètre contextuel le plus impactant.

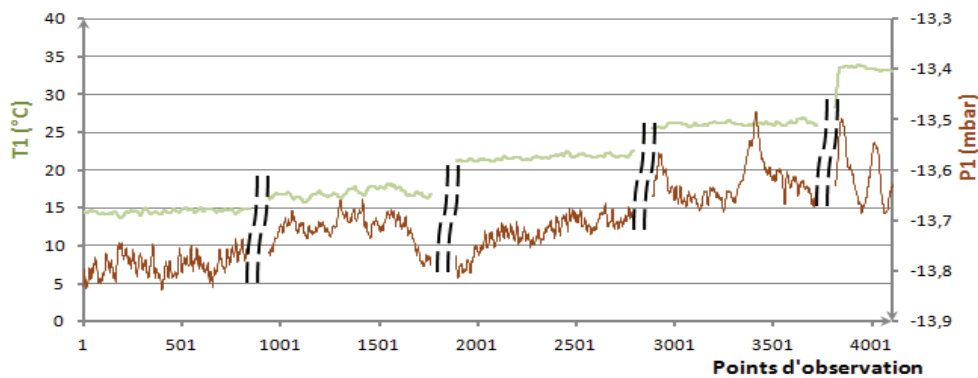


Figure 58 : Effet de la variation de la température pour la pression différentielle filtre à Air (P1)

Comme indiqué dans le chapitre précédent, afin de tester si la température (T1) affecte la dérive de l'indicateur P1, il faut calculer la distance de test de dépendance. T1 est jugé affectant la déviation par rapport au mode nominal si cette distance est supérieure à γ . Nous avons fixé γ à 10%.

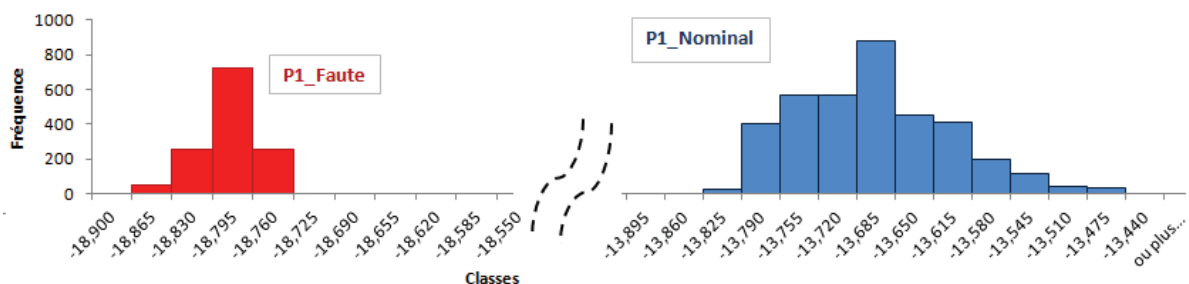


Figure 59 : Histogramme correspondant au mode nominal pour P1 à différentes valeurs de T1

Etant donné l'histogramme représentant les données prises par l'indicateur P1 en mode nominal sous différents contextes, y compris les valeurs de températures extrêmes, ainsi que celui du mode de faute (Figure 59), la distance est calculée comme suit :

$$DistP1 = \frac{-13.82 + 13.47}{-18.79 + 13.47} \cdot 100 = 6\%$$

Ainsi, l'indicateur P1 est indépendant de la température T1 puisque $DistP1 \leq 10\%$

Le calcul de la dérive par cette approche pour l'indicateur P1 est effectué en présence d'un encrassement partiel du filtre à air (le mode de faute étant simulé par un encrassement important de ce filtre). Le résultat est montré dans la Figure 60.

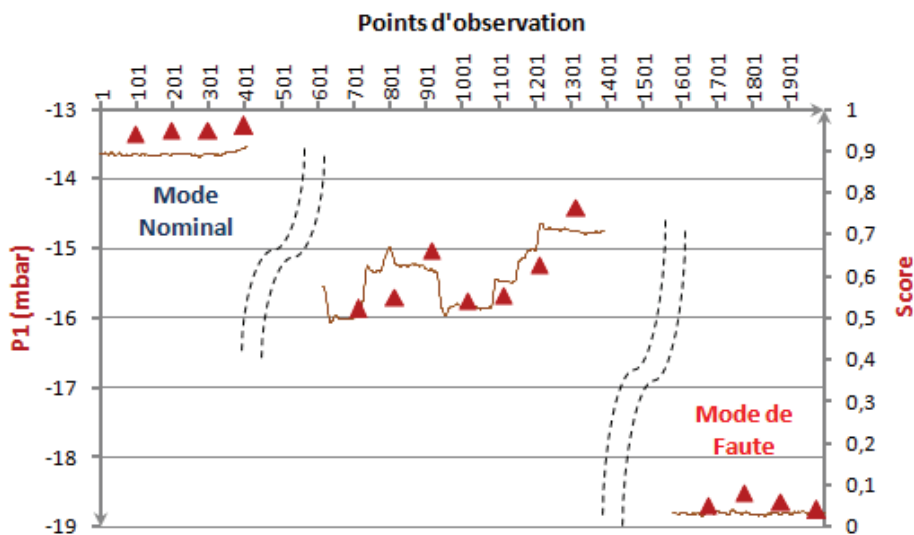


Figure 60 : Scores accordés au signal de P1 à différents instants de la phase d'encrassement partiel du filtre à air.

Sur cette figure, les triangles en rouge correspondent aux valeurs du score. Le signal correspond au mode nominal et le mode de faute est aussi ajouté afin de montrer le positionnement réel du signal en cours d'observation par rapport à ces deux modes. Les scores ainsi affichés sur le graphe représentent la distance du signal observé sur une fenêtre de cents points par rapport aux deux modes de référence (nominal et de faute). La valeur du score est adéquate au positionnement de la fenêtre du signal entre ces deux modes.

3.2. Indicateur de type I avec impact du contexte

Afin d'illustrer le 2^{er} cas, l'indicateur Pression air suralimenté (P2 prime) du bilan de santé de la fonction « Alimenter le moteur en air » est considérée. Il permet de détecter la fuite d'air ou de gaz ou une dégradation du turbocompresseur. Les données du banc d'essai permettent d'étudier le comportement de cet indicateur lors d'une simulation d'une fuite

d'air en aval du compresseur. Le mode de faute pour cet indicateur correspond alors au niveau de fuite le plus élevé possible permis par les expérimentations.

De la même façon, le test de dépendance au contexte doit être effectué. La Figure 61 montre l'évolution de cet indicateur en augmentant la température d'air en entrée de cette fonction (T1) entre les valeurs correspondantes aux conditions extrêmes.

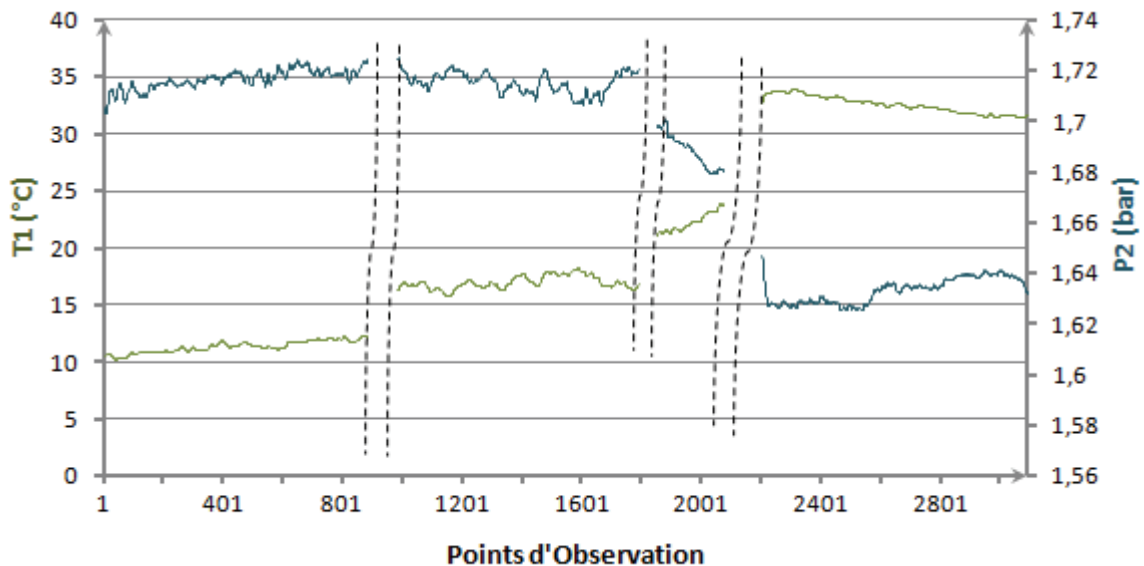


Figure 61 : Comportement de P2 (pression d'air suralimenté) vis-à-vis au changement du paramètre contextuel principal, T1

Le test de dépendance est réalisé en utilisant les paramètres fournis par l'histogramme du mode nominal sous différentes valeurs du paramètre contextuel et celui correspondant au mode de faute (Figure 62).

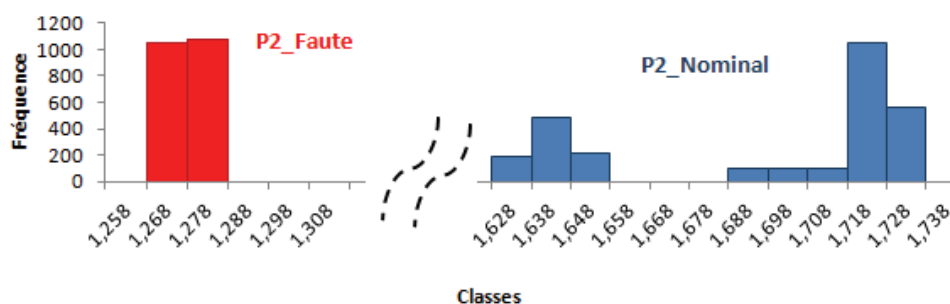


Figure 62 : Histogramme correspondant au mode nominal pour P2 à différentes valeurs de Tcel

La distance pour le test de dépendance de la dérive de cet indicateur au contexte est donnée par l'équation suivante :

$$Dist1 = \frac{1,62 - 1,72}{1,27 - 1,72} \cdot 100 = 21\%$$

L'indicateur P2 est bien affecté par le changement de contexte. La répartition de l'étendu des valeurs de T1 donne les intervalles de température suivants : [10,22.5[; [22.5,29.5[; [29.5,35]. En effet, les températures extrêmes étant 10 et 35, la valeur moyenne est 22.5°. Les distributions correspondantes à cette valeur moyenne et la valeur minimale sont proches, d'où l'intervalle de température [10,22.5[qui forme un premier sous-espace contextuel. Cependant, P2 varie considérablement entre la température moyenne et la température, la répartition contextuelle de cet intervalle de température donne lieu à deux sous espaces contextuels de température à savoir : [22.5,29.5[et [29.5,35].

Le calcul de la dérive pour l'indicateur P2 est effectué en présence d'une fuite moins importante d'air en aval du compresseur à une température appartenant au premier intervalle (Figure 63).

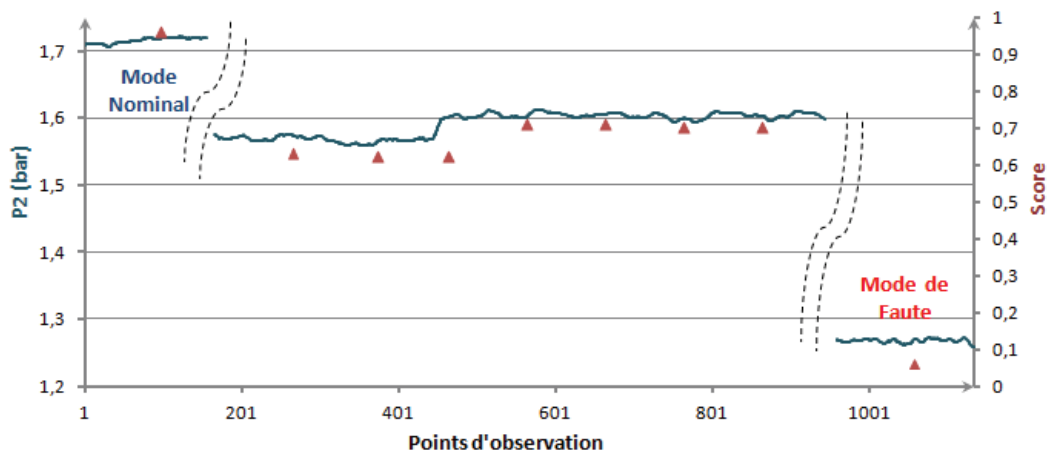


Figure 63 : Scores accordés au signal de P2 à différents instants de la phase de fuite partielle d'air en aval du compresseur

Sur cette figure, les triangles en rouge correspondent aux valeurs du score. Le signal correspond au mode nominal et le mode de faute est aussi ajouté afin de montrer le positionnement réel du signal en cours d'observation par rapport à ces deux modes.

De la même manière, les scores calculés pour P2 reflètent la distance qui sépare le signal observé sur une fenêtre de temps aux deux modes de référence (nominal et de faute) tel que visuellement affiché.

3.3. Indicateur de type II sans impact du contexte

Concernant le troisième type d'indicateur, l'indicateur de dégazage carter (Blowby) a été choisi en présence d'augmentation de température du moteur (simulé sur le banc par blocage du thermostat). Cet indicateur n'est pas dépendant de la température d'eau en

entrée du refroidisseur d'eau douce (bien que ce paramètre soit censé être le plus impactant sur la température globale du moteur) (Figure 64).

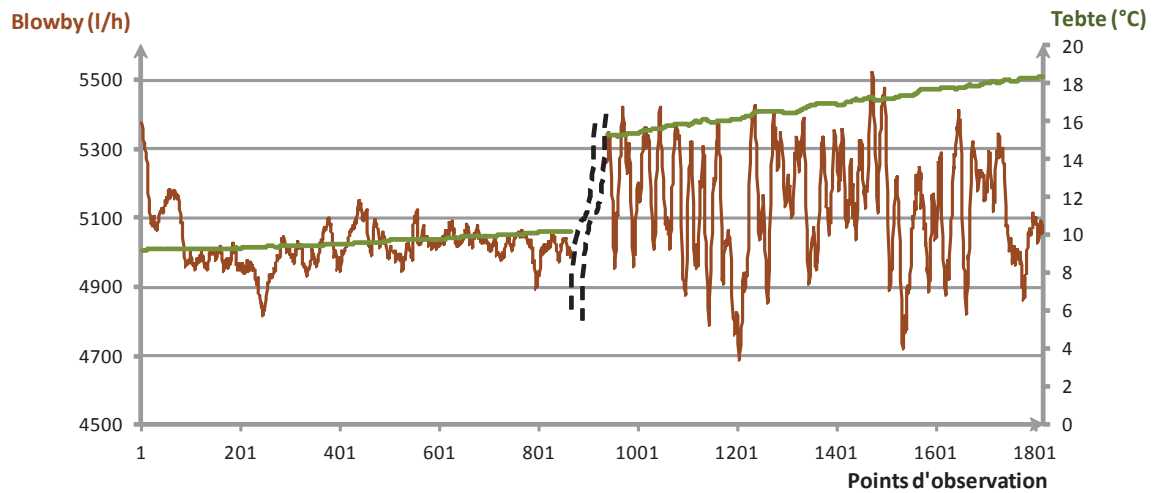


Figure 64 : Comportement de Blowby vis-à-vis au changement du paramètre contextuel principal, Tebte

De la même manière, le calcul de la dérive pour l'indicateur Blowby est effectué en présence d'un blocage du thermostat du circuit d'eau douce. L'histogramme du mode nominal est représenté dans la Figure 65, ainsi que celui du mode de faute.

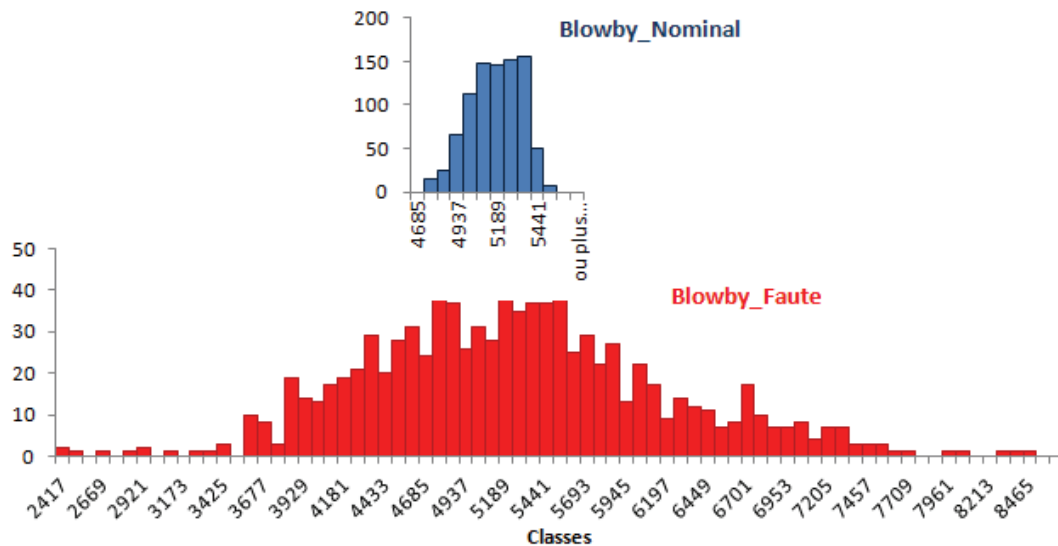


Figure 65 : Histogramme correspondant au mode nominal pour Blowby

La Figure 66 montre les scores accordés à ce signal sur une fenêtre de temps. Notons que ces scores affichés correspondent bien au degré de dégradation du signal observé (correspondant à un blocage partiel du thermostat) par rapport au mode nominal (présenté par le tronçon du signal à gauche) étant donné que le score 0 correspond au signal de faute (thermostat bloqué fermé).

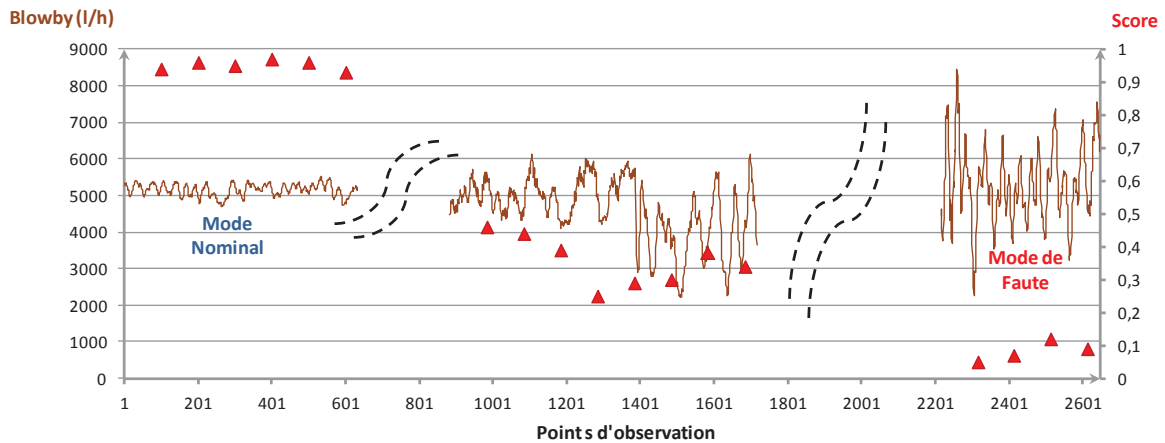


Figure 66 : Scores accordés au signal de Blowby à différents instants de la phase de blocage du thermostat

L'étape suivante correspond à l'agrégation des indicateurs de chaque bilan de santé dans l'architecture du moteur en un index de santé. C'est pourquoi, une identification des capacités pour chaque sous-système par apprentissage est nécessaire. Ensuite, les coefficients du niveau moteur sont inférés en utilisant l'approche correspondante aux sous-éléments en position série-parallèle.

4. La construction des indexes de santé

4.1. Apprentissage des capacités pour le niveau inférieur

Pour chaque sous système du niveau le plus bas dans l'architecture du moteur, nous procédons à un apprentissage des coefficients des capacités. Ces sous-systèmes sont les suivants :

- Sous-système d'air/échappement,
- Sous-système d'alimentation en carburant,
- Sous-système de lubrification,
- Sous-système de refroidissement,
- Sous-système base moteur.

Nous allons détailler la démarche d'apprentissage pour le premier sous système (Air/Echappement). Pour le reste des sous systèmes, les résultats ainsi que leurs interprétations sont directement présentés.

4.1.1. Les coefficients de capacité pour le sous-système d'air/échappement

La Figure 67 montre le lot de données utilisé pour l'apprentissage des coefficients de la capacité propre à l'agrégation des indicateurs du sous-système d'air/échappement. La première partie de ce lot de données est composée de valeurs normalisées correspondantes au mode nominal de fonctionnement du sous système. Les scores des indicateurs pour ce premier ensemble de données sont alors proches de 1. La deuxième partie des données se décompose en deux parties qui reflètent les scores des indicateurs en présence de deux dégradations différentes (causes primaires) qui sont l'encrassement du filtre à air (caractérisé par la chute des scores de l'indicateur P1) et ensuite une dégradation de fuite d'air en aval du compresseur (caractérisée par une déviation de l'indicateur P2).

Sur la base de l'apprentissage par la procédure proposée, les coefficients de la capacité ainsi identifiés sont donnés dans le Tableau 27.

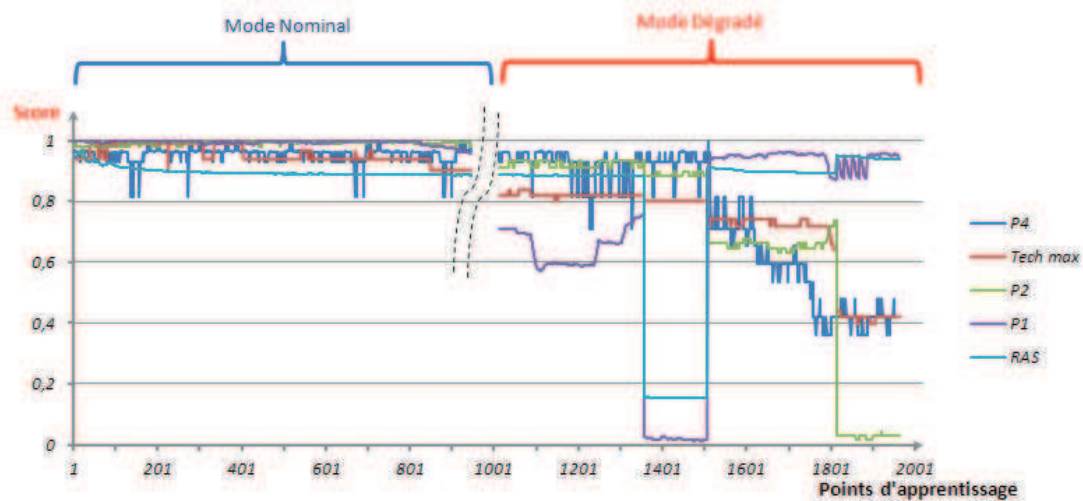


Figure 67 : Données d'apprentissage des coefficients de la capacité pour le sous-système d'air/échappement.

{}=0	{Techmax,P1}=0.99	{Techmax,P2,P1}=0.99
{P4}=0.32	{Techmax,RAS}=0.50	{Techmax,P2,RAS}=0.50
{Techmax}=0.50	{P2,P1}=0.99	{Techmax,P1,RAS}=0.99
{P2}=0.35	{P2,RAS}=0.49	{P2,P1,RAS}=0.99
{P1}=0.37	{P1,RAS}=0.37	{P4,Techmax,P2,P1}=0.99
{RAS}=0.10	{P4,Techmax,P2}=0.50	{P4,Techmax,P2,RAS}=0.50
{P4,Techmax}=0.50	{P4,Techmax,P1}=0.99	{P4,Techmax,P1,RAS}=1.0
{P4,P2}=0.35	{P4,Techmax,RAS}=0.50	{P4,P2,P1,RAS}=1.0
{P4,P1}=0.91	{P4,P2,P1}=0.99	{Techmax,P2,P1,RAS}=0.99
{P4,RAS}=0.46	{P4,P2,RAS}=0.50	{P4,Techmax,P2,P1,RAS}=1.0
{Techmax,P2}=0.50	{P4,P1,RAS}=0.91	

Tableau 27 : Coefficients de la capacité pour le bilan de santé du sous-système d'air/échappement.

Afin de mieux interpréter ces résultats, il est nécessaire de calculer les valeurs de Shapley (Eq. 2.3) (Tableau 28) ainsi que les degrés d'interaction (Eq. 2.4) (Tableau 29) entre ces indicateurs.

Indicateur	P4	Techmax	P2	P1	RAS
Valeur de Shapley	0.13	0.2	0.15	0.47	0.04

Tableau 28 : Valeurs de Shapley pour les indicateurs du sous-système d'air/échappement.

Les valeurs de Shapley montrent que globalement c'est l'indicateur P1 qui est le plus important ensuite les indicateurs Techmax, P2, P4 et RAS. Ces résultats peuvent être interprétés par la variation des scores pris par les indicateurs avec l'apparition des dégradations. Ainsi, les indicateurs qui ont des scores qui diminuent en mode dégradé ont plus d'importance que ceux qui réagissent peu à l'apparition d'une dégradation. En outre, l'indicateur P1 acquiert plus d'importance puisqu'il est en synergie avec plusieurs indicateurs, et notamment avec P2. Le Tableau 29 montre aussi une interaction de redondance entre P2 et P4. En effet, ces deux paramètres évoluent de la même manière à la présence des dégradations mis en œuvre pour l'apprentissage (Figure 67).

	P4	Techmax	P2	P1	RAS
P4		-0.19	-0.20	0.06	0.01
Techmax	-0.19		-0.25	0.01	-0.06
P2	-0.20	-0.25		0.10	0.01
P1	0.06	0.01	0.10		-0.06
RAS	0.01	-0.06	0.01	-0.06	

Tableau 29 : Degrés d'interaction pour les indicateurs du sous-système d'air/échappement.

De la même manière, un apprentissage des capacités pour les sous systèmes d'alimentation en carburant, de lubrification, de refroidissement et ceux pour les indicateurs de la base moteur sont effectués. Les résultats sont présentés dans la suite.

4.1.2. Les coefficients de capacité pour le sous-système d'alimentation en carburant

Lors de la phase d'apprentissage pour ce sous-système, les données correspondant au mode dégradé sont issues de la simulation d'une fuite de carburant à l'amont du filtre gazole. L'indicateur le plus impacté est celui de la pression du carburant à l'entrée de la pompe d'injection (Pepi). Les coefficients de la capacité sont donnés dans le Tableau 30.

{}=0	{Pfc}=0.36	{Pepi,Pfc}=1.0
{Dcarb}=0.83	{Dcarb,Pepi}=1.0	{Dcarb,Pepi,Pfc}=1.0
{Pepi}=1.0	{Dcarb,Pfc}=0.86	

Tableau 30 : Les coefficients de la capacité propres aux indicateurs du sous-système carburant.

Les valeurs de Shapley ainsi que les indices d'interaction sont respectivement présentés dans le Tableau 31 et le Tableau 32.

Indicateur	Dcarb	Pepi	Pfc
Valeur de Shapley	0.36	0.51	0.13

Tableau 31 : Valeurs de Shapley relatives aux indicateurs du sous-système carburant.

Etant donné qu'une seule cause primaire a été simulée pour ce sous système, aucune relation de synergie n'est observée.

	Dcarb	Pepi	Pfc
Dcarb		-0.66	-0.16
Pepi	-0.66		-0.19
Pfc	-0.16	-0.19	

Tableau 32 : Degrés d'interaction relatifs aux indicateurs du sous-système carburant.

4.1.3. Les capacités pour le sous-système de lubrification

Afin de recueillir les données d'apprentissage pour le mode dégradé, le fonctionnement du moteur en présence de l'encrassement du refroidisseur huile a été utilisé. De la même manière, les coefficients de la capacité ainsi élicités sont présentés dans le Tableau 33.

{}=0	{Pfh}=0.64	{Them,Pfh}=0.93
{Teem}=0.90	{Teem,Them}=0.90	{Teem,Them,Pfh}=1.0
{Them}=0.73	{Teem,Pfh}=1.0	

Tableau 33 : Les coefficients de la capacité propres aux indicateurs du sous-système de lubrification.

Une première analyse de ces résultats montre que la température d'eau douce à l'entrée du moteur est l'indicateur le plus important (Valeur de Shapley égale à 0.41) étant donné qu'il réagit le plus à la cause primaire de dégradation mise en œuvre lors de l'apprentissage.

Indicateur	Teem	Them	Pfh
Valeur de Shapley	0.41	0.29	0.29

Tableau 34 : Valeurs de Shapley relatives aux indicateurs du sous-système de lubrification.

Cependant, étant donné qu'une seule cause primaire (encrassement du refroidisseur huile) a pu être simulée pour ce sous système, aucune relation de synergie n'est observée. Ces interprétations sont confortées par les valeurs de Shapley (Tableau 34) et les degrés d'interaction (Tableau 35).

	Teem	Them	Pfh
Teem		-0.51	-0.33
Them	-0.51		-0.22
Pfh	-0.33	-0.22	

Tableau 35 : Degrés d'interaction relatifs aux indicateurs du sous-système de lubrification.

4.1.4. Les capacités pour le sous-système de refroidissement

Les causes primaires de dégradation mis en œuvre pour l'apprentissage des coefficients de capacité pour ce sous-système, sont la fuite d'eau brute en amont de l'échangeur eau brute/eau douce et l'avarie thermostat. Les résultats sont présentés dans le Tableau 36.

{}=0	{Tesm,Pce}=0.46	{Tesm,Pebte,RtEch}=1.00
{Tesm}=0.35	{Tesm,RtEch}=0.46	{Tesm,Pce,RtEch}=0.46
{Pebte}=0.39	{Pebte,Pce}=1.0	{Pebte,Pce,RtEch}=1.0
{Pce}=0.46	{Pebte,RtEch}=0.39	{Tesm,Pebte,Pce,RtEch}=1.00
{RtEch}=0.1	{Pce,RtEch}=0.46	
{Tesm,Pebte}=1.00	{Tesm,Pebte,Pce}=1.00	

Tableau 36 : Les coefficients de la capacité pour les indicateurs du sous-système de refroidissement.

Ces valeurs montrent une synergie entre les deux indicateurs caractéristiques des deux dégradations simulées pour les données d'apprentissage, à savoir le Pebte pour la fuite d'eau et le Tesm pour l'avarie thermostat. Les redondances sont issues de déviation simultanée de deux indicateurs en présence d'une même dégradation (Tableau 38).

Indicateur	Tesm	Pebte	Pce	RtEch
Valeur de Shapley	0.22	0.48	0.25	0.03

Tableau 37 : Valeurs de Shapley relatives aux indicateurs du sous-système de refroidissement.

Les valeurs de Shapley données dans le Tableau 37 montrent une grande importance pour Pebte (étant donné son coefficient de capacité et les interactions de synergie vis-à-vis des deux indicateurs Tesm et Pce). Cependant le rendement de l'échangeur eau brute/eau douce varie très peu en mode dégradé, d'où une valeur de Shapley faible.

	Tesm	Pebte	Pce	RtEch
Tesm		0.12	-0.48	0.00
Pebte	0.12		0.06	-0.05
Pce	-0.48	0.06		-0.05
RtEch	0.00	-0.05	-0.05	

Tableau 38 : Degrés d'interaction relatifs aux indicateurs du sous-système de refroidissement

4.1.5. Les capacités pour le sous-système base moteur

De la même manière, les coefficients de la capacité propre à ce sous-système sont identifiés. Vu le grand nombre de coefficients (128 coefficients pour 7 indicateurs), les résultats sont affichés uniquement sous forme de Valeurs de Shapley (Tableau 39) et degrés d'interaction (Tableau 40). Les dégradations correspondantes aux causes primaires sont liées aux flux entrants à ce sous système, c.à.d. ceux des sous-systèmes en aval.

Indicateur	Teem	Thsm	P2	Pepi	Pm	Blowby	AsyT
------------	------	------	----	------	----	--------	------

Valeur de Shapley	0.18	0.16	0.13	0.19	0.14	0.12	0.05
--------------------------	------	------	------	------	------	------	------

Tableau 39 : Valeurs de Shapley relatifs aux indicateurs du sous-système base moteur

Les importances des indicateurs sont proches les unes des autres étant donné le grand nombre de causes primaires mis en œuvre lors de l'apprentissage (fuite eau, fuite carburant, encrassement refroidisseur huile, avarie thermostat, encrassement filtre à air). Comme attendu, les relations de synergie sont essentiellement entre les indicateurs caractéristiques de dégradations différentes.

	Teem	Thsm	P2	Pepi	Pm	Blowby	AsyT
Teem		-0.12	0.03	0.07	0.03	-0.12	-0.01
Thsm	-0.12		0.02	0.08	-0.05	-0.04	-0.04
P2	0.03	0.02		0.07	-0.12	-0.08	-0.01
Pepi	0.07	0.08	0.07		-0.11	-0.08	-0.03
Pm	0.03	-0.05	-0.12	-0.11		-0.06	-0.02
Blowby	-0.12	-0.04	-0.08	-0.08	-0.06		-0.00
AsyT	-0.01	-0.04	-0.01	-0.03	-0.02	-0.00	

Tableau 40 : Degrés d'interaction relatifs aux indicateurs de la base moteur

A partir des coefficients des capacités identifiés pour les éléments du niveau inférieur de la hiérarchie du moteur diesel étudié dans cette application, ceux du niveau supérieur, à savoir le moteur dans son ensemble, doivent être inférés.

4.2. Inférence de la capacité pour le niveau supérieur

Les sous-systèmes du moteur diesel présentent une structure série-parallèle comme le montre la Figure 68. La répartition en blocs de ces sous-systèmes est effectuée en deux étapes. Une première étape consiste à regrouper les deux sous-systèmes de refroidissement et lubrification, qui sont en série, en un premier bloc. Ensuite, dans un deuxième regroupement, le bloc ainsi formé peut être regroupé avec la fonction d'admission de l'air et celle de l'alimentation en carburant pour former un bloc constitué de trois fonctions en parallèle. Enfin, ce deuxième bloc est positionné en série avec la base moteur (Figure 68).

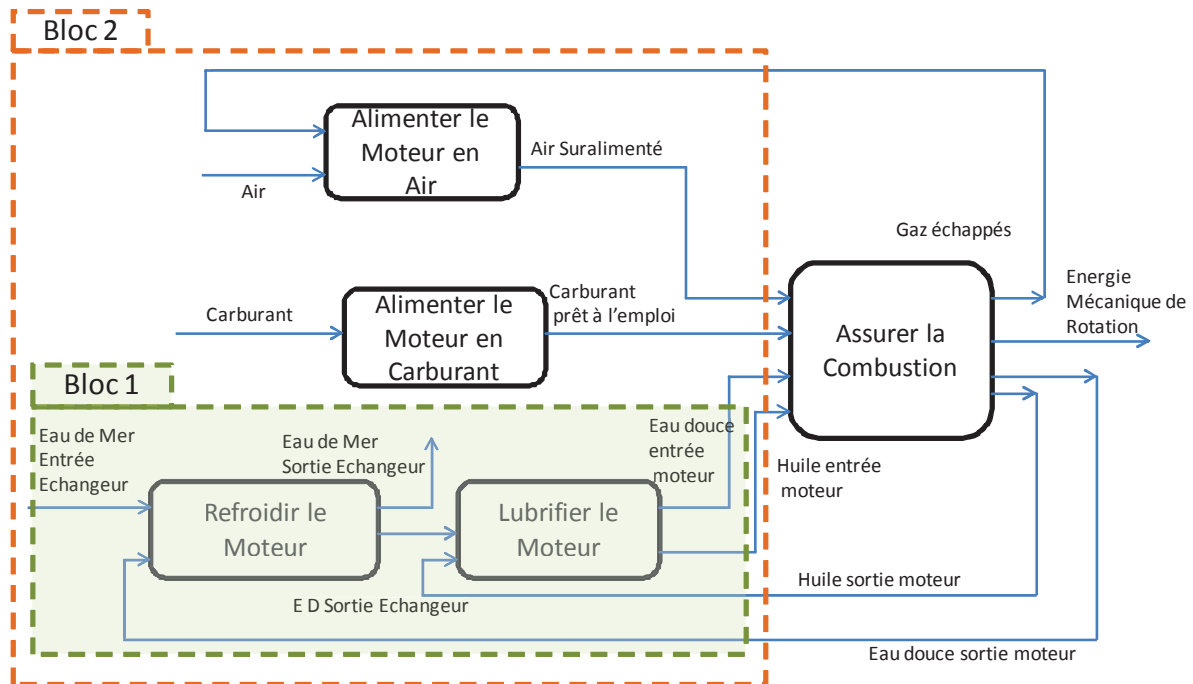


Figure 68 : Décomposition en blocs des sous-systèmes du moteur diesel.

Le tableau montre les indicateurs contributeurs à la construction des indicateurs du moteur diesel.

$\mathcal{V}(\mathbf{E})$	$\mathbf{g}_{b1}(\mathcal{V}(\mathbf{E}))$	$\mathbf{g}_{b2}(\mathcal{V}(\mathbf{E}))$	$\mathbf{g}_{bmot}(\mathcal{V}(\mathbf{E}))$
Cs		Dcarb	Pm
P4		P4	
Pebte	Pebte	Pebte	
Rhc		Dcarb	Thsm

Tableau 41 : Indicateurs contributeurs pour la construction des indicateurs du moteur diesel.

4.2.1. Calcul de la capacité pour le bloc 1

Parmi les indicateurs du premier bloc, l'indicateur Pebte est contributeur pour le niveau supérieur (par choix de construction des indicateurs). Le coefficient de capacité de cet indicateur est alors calculé.

$$D'ou \mu_{b1}(\{Pebte\}) = \mu'_{ref}(\{Pebte\})$$

$$Avec \mu'_{ref}(\{Pebte\}) = \frac{\mu_{ref}(\{Pebte\})}{\mu_{ref}(\{Pebte\})} * \frac{|g_{ref}(\{Pebte\})|}{|g_{lub}(\{Pebte\}) \cap g_{ref}(\{Pebte\})|} = 1$$

4.2.2. Calcul de la capacité pour le bloc 2

Etant donné que le deuxième bloc est composé de trois sous éléments montés en parallèle, le calcul des coefficients de la capacité pour les indicateurs de ce bloc est effectué comme le montre le Tableau 42.

$\mu_{b2}(\{Dcarb\}) = \frac{\mu_{a1}(\{Dcarb\})}{\mu_{a1}(\{Dcarb\}) * 3} = \frac{1}{3} = 0.33$	$\mu_{b2}(\{Dcarb, P4\}) = 0.66$
$\mu_{b2}(\{Pebte\}) = \frac{\mu_{b1}(\{Pebte\})}{\mu_{b1}(\{Pebte\}) * 3} = \frac{1}{3} = 0.33$	$\mu_{b2}(\{Pebte, P4\}) = 0.66$
$\mu_{b2}(\{P4\}) = \frac{\mu_{ad}(\{P4\})}{\mu_{ad}(\{P4\}) * 3} = \frac{1}{3} = 0.33$	$\mu_{b2}(\{Dcarb, Pebte, P4\}) = 1$
$\mu_{b2}(\{Dcarb, Pebte\}) = \frac{\mu_{a1}(\{Dcarb\})}{\mu_{a1}(\{Dcarb\}) * 3} + \frac{\mu_{b1}(\{Pebte\})}{\mu_{b1}(\{Pebte\}) * 3} = 0.66$	

Tableau 42 : Les coefficients de la capacité pour le bloc2.

4.2.3. Calcul de la capacité pour l'ensemble du moteur

Dans une dernière étape, le calcul des coefficients de capacité des indicateurs du niveau supérieur, à savoir celui du niveau moteur vu dans son ensemble sont obtenus en appliquant la méthode d'inférence relative au cas série. Les deux sous éléments sont alors le bloc 2 associé à son bilan virtuel (Dcarb, Pebte, P4) et la fonction de combustion assurée par la « base moteur ». Les indicateurs contributeurs dans le bilan de santé du moteur issus de cette dernière fonction sont Thsm et Pm (Figure 69).

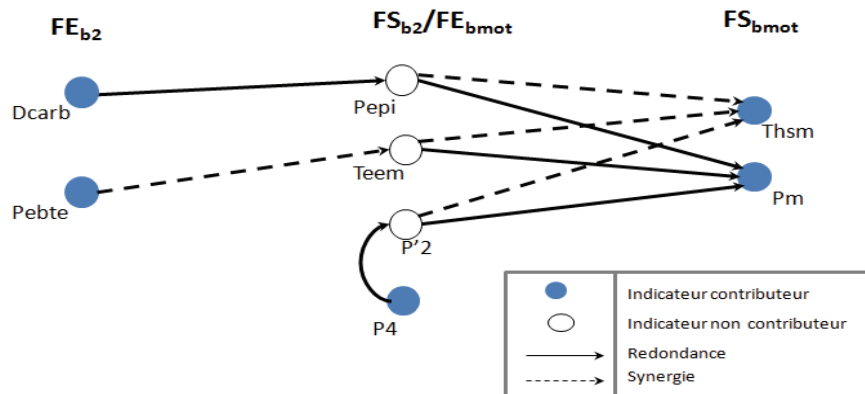


Figure 69 : Représentation des interactions entre les indicateurs du bloc 2 et de la base moteur.

Détermination de θ et σ :

- L'application de la condition de normalité donne :

$$\theta * \left| \mathcal{S}^- \left(g_{b2}(\mathcal{V}(E)) \right) \cup g_{bmot}(\mathcal{V}(E)) \right| = \sigma * \left| \mathcal{S}^+ \left(g_{b2}(\mathcal{V}(E)) \right) \cap g_{bmot}(\mathcal{V}(E)) \right|$$

$$\mathcal{S}^- \left(g_{b2}(\mathcal{V}(E)) \right) = \{Pm\}$$

$$\mathcal{S}^+ \left(g_{b2}(\mathcal{V}(E)) \right) = \{Pm, Thsm\}$$

$$\theta = 2\sigma$$

- L'application de la condition de monotonie :

$$\theta \leq \frac{\min \left\{ 0.33 * \frac{3}{5}, 0.45 * \frac{2}{5} \right\}}{1}$$

$$\theta \leq 0.19$$

$$\theta = 0.10 \text{ et } \sigma = 0.05$$

Enfin, un exemple de calcul des coefficients de la capacité pour le moteur est donné par la suite :

$$\mu_E(\{Cs\}) = \mu'_{b2}(\{Dcarb\}) + \mu'_{bmot}(Pm) - \theta$$

Avec :

$$\mu'_{b2}(\{Dcarb\}) = \frac{\mu_{b2}(\{Dcarb\})}{\mu_{b2}(g_{b2}(\mathcal{V}(E)))} * \frac{|g_{b2}(\mathcal{V}(E))|}{|g_{b2}(\mathcal{V}(E))| + g_{bmot}(\mathcal{V}(E))} = 0.33 * \frac{3}{5} = 0.20$$

$$\mu'_{bmot}(\{Pm\}) = \frac{\mu_{bmot}(\{Pm\})}{\mu_{bmot}(g_{bmot}(\mathcal{V}(E)))} * \frac{|g_{bmot}(\mathcal{V}(E))|}{|g_{b2}(\mathcal{V}(E))| + g_{bmot}(\mathcal{V}(E))} = \frac{0.41}{0.86} * \frac{2}{5} = 0.19$$

$$\mu_E(\{Cs\}) = 0.29$$

De la même manière, le reste des coefficients de la capacité est calculé (Tableau 43).

$\wp(\mathcal{V}(E))$	μ_E	$\wp(\mathcal{V}(E))$	μ_E
{Cs}	0.29	{P4,Rhc}	0.69
{P4}	0.20	{Pebte,Rhc}	0.69
{Pebte}	0.20	{Cs,P4,Pebte}	0.64
{Rhc}	0.43	{Cs,P4,Rhc}	0.70
{Cs,P4}	0.39	{Cs,Pebte,Rhc}	0.85
{Cs,Pebte}	0.54	{P4,Pebte,Rhc}	0.93
{Cs,Rhc}	0.55	{Cs,P4,Pebte,Rhc}	1
{P4,Pebte}	0.40		

Tableau 43 : Coefficients de la capacité relative au niveau moteur diesel

Cette inférence a permis de trouver des coefficients de capacité plus importants pour les indicateurs ayant un plus grand nombre d'indicateurs contributeurs (Cs et Rhc). Ce résultat est intéressant car de tels indicateurs sont plus riches en informations pour le suivi de la santé au niveau supérieur. Cependant, en termes de valeur de Shapley, les deux indicateurs Rhc et Pebte gagnent de l'importance globale (Tableau 44) étant donné les interactions de synergie qui les lient avec les deux autres indicateurs (Tableau 45).

Indicateur	Cs	P4	Pebte	Rhc
Valeur de Shapley	0.18	0.17	0.25	0.0.39

Tableau 44 : Valeurs de Shapley relatives aux indicateurs du moteur diesel

En outre, notons qu'aucune interaction n'existe entre les deux indicateurs (P4 et Pebte) appartenant respectivement aux deux sous-éléments en parallèle (admission en air/échappement et bloc1).

	Cs	P4	Pebte	Rhc
Cs		-0.09	0.05	-0.17
P4	-0.09		-0.00	0.05
Pebte	0.05	-0.00		0.05
Rhc	-0.17	0.05	0.05	

Tableau 45 : Degrés d'interaction relatifs aux indicateurs du moteur

Cependant, pour les indicateurs ayant des indicateurs contributeurs en commun tels que Cs et Rhc, une interaction de redondance entre eux est inférée. De tels résultats sont en concordance avec le comportement attendu en termes d'importance et d'interaction entre les indicateurs à ce niveau hiérarchique.

4.3. Calcul des états de santé

Par rapport aux étapes suivies pour identifier les capacités et paramétrer l'opérateur d'agrégation aux différents niveaux du système, il est logique à ce stade de chercher à

évaluer si les indexes calculés reflètent le degré de déviation réel du comportement du système. Cette phase de validation a été faite pour deux situations de comportement mettant en évidence des degrés différents de propagation des anomalies dans le système. Les différentes situations considérées sont les suivantes:

- Situation 1 : Encrassement filtre à air, cette dégradation n'est observée qu'au niveau du sous-système d'air/échappement (Tableau 46)
- Situation 2 : Fuite eau brute par réduction du débit d'eau brute en entrée échangeur, en fermant la vanne. Cette dégradation a un impact majeur sur le sous-système de refroidissement et par conséquent sur la base moteur (Tableau 47).

Sous-Système	Scores des indicateurs	IS	Scores des indicateurs du Système (Moteur Diesel)	IS
Air-Échappement	0.95/0.92/0.86/0.23/0.89	0.58	0.91/0.95/0.91/0.92	0.92
Carburant	0.94/0.91/0.93	0.93		
Lubrification	0.93/0.95/0.91	0.94		
Refroidissement	0.88/0.91/0.90/0.92	0.90		
Base Moteur	0.93/0.92/0.86/0.91/0.95/0.88/0.95	0.91		

Tableau 46 : Valeurs des états de santé pour le premier cas.

Cette première situation montre un effet local de l'encrassement du filtre à air qui, d'après les mesures des autres indicateurs, ne montre pas de conséquence importante sur le fonctionnement global du moteur. Ainsi, le seul index de santé impacté est celui du sous-système d'air/échappement. Cette structure multi-niveaux hiérarchiques du moteur permet ainsi d'orienter l'attention du pilote du navire vers les anomalies locales qui ne peuvent pas être décelées par une vision globale du moteur.

D'un point de vue d'aide à la prise de décision en maintenance, une question se pose : comment un opérateur pourrait interpréter ces valeurs ?

Un opérateur ayant un comportement optimiste trouve que la valeur 0.92 pour l'index de santé du moteur diesel est proche de la normalité et n'est ainsi pas inquiété (d'où ne rien faire). Par contre, un opérateur ayant un comportement rationnel peut trouver que cette valeur incite à aller contrôler les index de santé du niveau inférieur (ceux des sous-systèmes). Il focalise ainsi sur la valeur 0.58 de l'index de santé du sous-système d'air/échappement (la valeur la plus impactée parmi les autres indexes qui sont eux plutôt proches de la normalité). Cependant, tant que la tendance d'évolution négative de 0.58 ne diminue pas significativement la valeur de l'index de santé du moteur, l'opérateur peut garder en tête le sous-système air/échappement pour une inspection future mais non urgente.

Sous-Système	Scores des indicateurs	IS	Scores des indicateurs du Système (Moteur Diesel)	IS
Air-Echappement	0.92/0.76/0.91/0.93/0.90	0.92	0.91/0.92/0.36/0.73	0.69
Carburant	0.93/0.92/0.90	0.92		
Lubrification	0.94/0.93/0.88	0.93		
Refroidissement	0.51/0.36/0.81/0.27	0.56		
Base Moteur	0.84/0.67/0.91/0.92/0.95/0.89/0.90	0.81		

Tableau 47 : Valeurs des états de santé pour la troisième situation

La fuite d'eau brute a, comme attendu, un impact sur le sous-système de refroidissement ainsi que sur celui de la base moteur (augmentation de la température huile sortie moteur). Ainsi, les deux indicateurs Pebte et Rhc du niveau moteur se dégradent et l'index de santé diminue par conséquent. De cette manière le pilote, averti par une dégradation de l'index de santé global, pourra s'orienter vers plus d'informations sur les causes des anomalies dans le niveau inférieur.

Dans un objectif d'aide à la prise de décision en maintenance, ces résultats peuvent être interprétés de la manière suivante. Un opérateur ayant un comportement rationnel trouve que l'index de santé de valeur 0.69 pour le moteur diesel est représentatif a priori d'une situation dégradée importante, ce qui l'incite à aller contrôler les indexes de santé des différents sous-systèmes. Il regarde ainsi les indexes de niveau inférieur et se focalise sur le sous-système refroidissement avant de planifier les inspections nécessaires. En outre, sur la base de la valeur de l'index de santé de la base moteur, une tâche de maintenance opportune peut être aussi envisagée (regroupement des actions de maintenance des unités de refroidissement et de la base moteur).

5. Conclusion

Ce chapitre a fait l'objet de la mise en œuvre des approches proposées sur un système réel multi-composants lié au projet BMCI, le moteur diesel marin. Ce cas réel a permis de conforter les propositions scientifiques pour l'élaboration d'un bilan de santé multi-niveaux hiérarchiques et l'agrégation des indicateurs de santé après la phase de détection-normalisation. L'applicabilité et la faisabilité de telles propositions ont été démontrées. En effet, ce cas applicatif a permis de révéler les apports suivants des contributions des travaux de cette thèse :

- Une vision étendue et détaillée des indicateurs de santé pour le moteur diesel marin. Une telle vision facilite le suivi des dégradations ainsi que leurs interactions et propagations au sein de ses sous-systèmes.
- Une méthode pour la quantification contextuelle des dérives par rapport au mode nominal et de faute pour des indicateurs ayant des tendances de déviations différentes (type 1 et 2).
- La méthode d'apprentissage des coefficients de la capacité pour les éléments du niveau inférieur du moteur diesel donne lieu à des résultats traduisant les comportements des indicateurs vis-à-vis à l'apparition des causes primaires de dégradation. Les coefficients de chaque sous-ensemble d'indicateurs sont alors proportionnels à leur pouvoir de révéler au plus vite l'apparition des anomalies ainsi que leurs interactions vis-à-vis de telle apparition. Cette approche permet bien de se passer des connaissances fournies par les experts concernant chaque élément du système.
- Une inférence des coefficients de la capacité au niveau moteur à partir de ceux des sous-systèmes montés en série-parallèle a été appliquée de manière à refléter les degrés d'importance des indicateurs contributeurs ainsi que leurs interactions.

Cette application a permis également de mettre en avant les difficultés liées à la confrontation avec un système réel. Les limites les plus importantes sont :

- Des limites dans les conditions d'application
 - Le manque de valeurs pour les indicateurs contextuels (e.g. la température de la cellule dans laquelle se trouve le moteur ainsi que celle de l'eau brute de refroidissement) a limité les données représentatives du comportement nominal de l'indicateur vis-à-vis du changement continu des conditions environnementales.
 - En outre, les dégradations mécaniquement simulées sont parfois insuffisantes pour représenter toutes les causes primaires des dégradations pour l'apprentissage des capacités.
 - La simulation mécanique des dégradations ne permet pas de représenter une évolution continue des déviations car elle est effectuée à des niveaux ponctuels d'anomalies (discrétisation).
 - Des limites dans le déploiement de la méthodologie proposée :
 - Concernant l'approche de détection contextuelle, l'approche proposée se limite par hypothèse à un seul indicateur contextuel, ce qui n'est pas le cas pour différents cas de figures où plusieurs paramètres contextuels influencent d'une manière simultanée le comportement nominal d'un système.
 - En outre, cette application a mis en avant que la construction des histogrammes de référence pour la détection contextuelle requiert une
-

quantité importante de données relatives aux comportements nominaux et de faute à chaque sous espace contextuel et ce pour tous les indicateurs, ce qui peut être difficile pour des cas réels.

Conclusion générale et Perspectives

La thèse a pour objet principal de fonder les éléments essentiels du concept novateur de bilan de santé qui représente l'état d'un système sous la forme d'un vecteur d'indicateurs de différentes natures aidant ensuite à la prise de décision. Ce fondement s'appuie sur les exigences exprimées dans BMCI puis généralisées et rationalisées afin de développer une contribution plus générique déclinable ensuite dans divers domaines d'applications (au-delà du moteur diesel marin).

La thèse se focalise plus spécifiquement sur les fonctions de détection-normalisation-agrégation d'indicateurs pour élaborer un index synthétique représentatif de l'état de santé global d'un élément et ce pour des systèmes multi-composants structurés suivant une architecture fonctionnelle multi-niveaux hiérarchiques.

Nous avons donc, pour le contexte de systèmes industriels, repreciser les définitions de la santé et son support exploitées ensuite comme référence pour ce concept de bilan de santé en cohérence principalement avec les travaux de la communauté PHM (Prognostics and Health Management). Sur cette base de définitions, les contributions scientifiques majeures défendues dans la thèse portent, par rapport au périmètre des activités traitant des indicateurs/indexes, sur:

- Une nouvelle approche de détection conditionnelle des anomalies. Cette approche a pour intérêt (1) de quantifier la déviation pour chaque indicateur par rapport à ses modes de comportement nominal et de faute, et (2) de prendre en compte le contexte opérationnel/environnemental dans lequel se trouve le système. Cette quantification des déviations peut être appliquée pour différents types de tendances de déviations des indicateurs. L'écart par rapport au mode nominal des scores accordés à ces indicateurs tient compte aussi de la distance par rapport au comportement de faute. En outre, la détermination des modes nominaux de référence pour chaque indicateur tient compte du paramètre contextuel susceptible d'influencer le comportement de cet indicateur. La méthode de détection ainsi construite constitue aussi une technique générique pour la normalisation des indicateurs hétérogènes à agréger ;
- Une extension à l'exploitation de l'intégrale de Choquet en tant qu'opérateur d'agrégation des indicateurs. Cette extension concerne, d'une part, un processus d'apprentissage non supervisé des paramètres/capacités pour le niveau le plus

inférieur dans l'abstraction, à savoir celui des composants, et d'autre part, une approche de leur inférence d'un niveau à l'autre.

L'ensemble de ces propositions a été mis en œuvre sur un cas réel. Ce cas applicatif est un moteur diesel marin lié au projet BMCI. Une telle quantification a permis de mettre en avant la faisabilité des approches proposées et leurs intérêts pour une application générique et étendue à des systèmes variés. Cependant, des perspectives à la fois scientifiques et applicatives peuvent être formulées pour ces travaux :

- Confronter la génération des scores de dérives à une démarche d'expertise pour éventuellement reconsidérer la fonction de détection [Peres, 2007],
 - Une amélioration de l'approche de détection contextuelle pourrait être réalisée en explorant la manière de tenir compte de plus d'un indicateur contextuel à la fois [Song, 2007]. En outre, l'influence du contexte sur le comportement de faute peut aussi être considérée,
 - Evaluer les états de santé proposés par rapport aux attentes des experts en vue de proposer une approche conjointe expertise/données pour l'apprentissage des capacités,
 - Traiter le cas des éléments montés en bridge pour l'inférence des capacités,
 - Étendre la validation/expérimentation à une phase opérationnelle sur un système navire en mission (utilisation par le bord *on line*) pour réaliser un véritable passage à l'échelle tel que préconisé par les exigences du projet,
 - Considération du processus d'élaboration du bilan de santé au sein d'un boucle décisionnelle plus large en s'intégrant avec le diagnostic voire le pronostic.
-

Liste des figures

Figure 1 : Principe de structuration multi-niveaux d'abstraction appliqué à la flotte.	16
Figure 2 : Architecture logicielle de l'application BMCI.	17
Figure 3 : Fonctionnalités liées au processus de surveillance adaptées au projet [ISO13374, 2002]. ...	17
Figure 4 : Représentation d'un vecteur du bilan de santé.	19
Figure 5 : Représentation de l'état de santé global [Monnin, 2009].	20
Figure 6 : Exemple de construction du bilan de santé et de détermination de l'état de santé de la fonction "PROPULSER" pour le navire [Monnin, 2009].	22
Figure 7 : Apport de la dimension flotte pour l'extraction des connaissances à des fins de diagnostic et d'aide à la décision [Monnin, 2011].	23
Figure 8 : Rôle du PHM pour une aide à la décision optimale [Camci, 2007].	27
Figure 9 : Les zones d'agrégation dans un processus PHM [Roemer, 2001].	29
Figure 10 : Le système transforme les flux pour répondre à sa finalité [AFIS, 2009].	35
Figure 11 : Aspects structurels internes du système, (a) Décomposition arborescente et (b) Composition ou architecture, [AFIS, 2009].	36
Figure 12 : Structure hiérarchique d'un bilan de santé.	38
Figure 13 : Les éléments d'entrée/sortie pour l'activité globale d'élaboration d'un bilan de santé. ...	40
Figure 14 : Architecture du standard OSA-CBM [Discenzo, 2001].	41
Figure 15 : Architecture du SIMP [Muller, 2005].	42
Figure 16 : Architecture de maintenance conditionnelle/prévisionnelle [Ben Salem, 2008].	43
Figure 17 : Architecture du IVHM [Benedettini, 2009].	44
Figure 18 : Architecture de fusion centralisée [Roemer, 2001]	45
Figure 19 : Architecture de fusion autonome [Roemer, 2001]	45
Figure 20 : Boîte générique élémentaire pour les modules d'activités bilan de santé	47
Figure 21 : Positionnement de la détection dans le processus de surveillance [Racoceanu, 2006].	50
Figure 22 : Classification des méthodes de surveillance [Racoceanu, 2006].	53
Figure 23 : Procédure de détection à base de modélisation quantitative [Verron, 2007].	53
Figure 24 : Surveillance par reconnaissance de formes [Ribot, 2009].	55
Figure 25 : Méthode de détection avec un label en sortie [Chandola, 2009].	55
Figure 26 : Classification des méthodes de détection des anomalies [Chandola, 2009].	58

Figure 27 : Schéma conceptuel de la fonction de fusion [Raol, 2010]	61
Figure 28 : Les deux étapes essentielles de la fusion.....	62
Figure 29 : Différents groupes d'opérateurs d'agrégation.....	65
Figure 30 : Interprétation graphique de l'intégrale de Choquet [Murofishi, 1991].....	69
Figure 31 : Architecture fonctionnelle sur laquelle se base la présentation du bilan de santé.	78
Figure 32 : Etapes de construction d'un index de santé à partir de l'ensemble des indicateurs.....	79
Figure 33 : Disposition des histogrammes pour le type I avec une déviation croissante	89
Figure 34 : Disposition des histogrammes pour le type II.....	90
Figure 35 : Zones de valeurs nominales et de faute pour un exemple d'indicateur de type 1.	92
Figure 36: Zones de valeurs nominale et de faute pour un exemple d'indicateur de type 2.	95
Figure 37 : Structure de l'approche de la normalisation contextuelle des indicateurs	98
Figure 38: Paramètres pour le test de dépendance au contexte.....	98
Figure 39 : Première étape de répartition de l'espace contextuel.	99
Figure 40 : Définition d'un intervalle du paramètre contextuel.	100
Figure 41 : Fonction composée de deux éléments en parallèle.....	104
Figure 42 : Fonction composée de deux éléments en série.....	110
Figure 43 : Graphe de relations entre les indicateurs d'éléments en série	110
Figure 44 : Fonction d'estampage/coupe dans la plateforme TELMA.	115
Figure 45 : Décomposition de la fonction « accumulation de la bande ».	115
Figure 46 : Les interactions mutuelles entre les indicateurs des sous éléments en série.....	116
Figure 47 : Répartition de la fonction en deux couches parallèles	120
Figure 48 : Architecture du banc d'essai [Peysson, 2011].....	124
Figure 49 : Schéma de fonctionnement de la suralimentation par turbocompresseur [Sabeh, 2006]	126
Figure 50 : Schéma du sous système de refroidissement du moteur diesel.....	127
Figure 51 : Structure arborescente du moteur diesel.....	128
Figure 52 : Organisation des flux échangés entre les fonctions du moteur.....	128
Figure 53 : Sous-système admission en air et échappement.....	129
Figure 54 : Sous-système alimentation en carburant	130
Figure 55 : Sous-système de lubrification	131
Figure 56 : Sous-système de refroidissement	132
Figure 57 : Flux entrée/sortie pour la base moteur	133

Figure 58 : Effet de la variation de la température pour la pression différentielle filtre à Air (P1)	135
Figure 59 : Histogramme correspondant au mode nominal pour P1 à différentes valeurs de T1	135
Figure 60 : Scores accordés au signal de P1 à différents instants de la phase d'encrassement partiel du filtre à air.....	136
Figure 61 : Comportement de P2 (pression d'air suralimenté) vis-à-vis au changement du paramètre contextuel principal, T1.....	137
Figure 62 : Histogramme correspondant au mode nominal pour P2 à différentes valeurs de Tcel ...	137
Figure 63 : Scores accordés au signal de P2 à différents instants de la phase de fuite partielle d'air en aval du compresseur	138
Figure 64 : Comportement de Blowby vis-à-vis au changement du paramètre contextuel principal, Tebte	139
Figure 65 : Histogramme correspondant au mode nominal pour Blowby.....	139
Figure 66 : Scores accordés au signal de Blowby à différents instants de la phase de blocage du thermostat	140
Figure 67 : Données d'apprentissage des coefficients de la capacité pour le sous-système d'air/échappement.	141
Figure 68 : Décomposition en blocs des sous-systèmes du moteur diesel.	146
Figure 69 : Représentation des interactions entre les indicateurs du bloc 2 et de la base moteur....	147

Liste des tableaux

Tableau 1 : Comparaison entre les différents groupes de méthodes sans modèle selon les deux critères de généralité et de la nature de la sortie	59
Tableau 2 : Comparaison entre opérateurs d'agrégation par rapport aux deux critères relatifs aux poids et interactions des indicateurs	66
Tableau 3 : Liste des coefficients de la capacité pour le calcul du score global des élèves	70
Tableau 4 : Différentes situations pour le calcul des moyennes par Intégrale de Choquet et la moyenne pondérée.....	70
Tableau 5 : Différentes classes de déviations pour les signaux	88
Tableau 6: Représentation des distributions correspondantes aux instants d'observation et les scores relatifs pour le premier signal.....	93
Tableau 7 : Représentation des distributions correspondantes aux instants d'observation et les scores relatifs pour le deuxième signal.....	96
Tableau 8 : Extrait des données d'apprentissage et leur structure pour l'identification des coefficients de la capacité	103
Tableau 9 : Valeurs des degrés d'interaction	103
Tableau 10 : Valeurs de Shapley correspondants à la capacité identifiée	104
Tableau 11 : Les coefficients de capacité pour les deux pompes	107
Tableau 12 : Les valeurs de Shapley pour les deux pompes	107
Tableau 13 : Les indicateurs contributeurs pour la fonction P12	108
Tableau 14 : Le calcul des coefficients de la capacité pour la fonction globale.....	108
Tableau 15 : Les valeurs de Shapley correspondantes aux indicateurs de la fonction globale.....	108
Tableau 16 : Propagation des phénomènes de santé entre les deux niveaux.....	109
Tableau 17 : Les indicateurs pour chaque élément	116
Tableau 18 : Les coefficients de capacité des sous éléments.....	116
Tableau 19 : Les valeurs de la capacité pour la fonction globale	118
Tableau 20 : Propagation de la déviation à travers les deux niveaux hiérarchiques.....	118
Tableau 21 : Les indicateurs de santé pour le sous système d'Air/Echappement	130
Tableau 22 : Les indicateurs de santé pour le sous système d'Alimentation en carburant	131
Tableau 23 : Les indicateurs de santé pour le sous système de lubrification	131

Tableau 24 : Les indicateurs de santé pour le sous système de refroidissement	132
Tableau 25 : Les indicateurs de santé pour la base du moteur.....	133
Tableau 26 : Les indicateurs de santé pour le moteur Diesel	134
Tableau 27 : Coefficients de la capacité pour le bilan de santé du sous-système d'air/échappement.	141
Tableau 28 : Valeurs de Shapley pour les indicateurs du sous-système d'air/échappement.	142
Tableau 29 : Degrés d'interaction pour les indicateurs du sous-système d'air/échappement.....	142
Tableau 30 : Les coefficients de la capacité propres aux indicateurs du sous-système carburant.	142
Tableau 31 : Valeurs de Shapley relatives aux indicateurs du sous-système carburant.....	143
Tableau 32 : Degrés d'interaction relatifs aux indicateurs du sous-système carburant.	143
Tableau 33 : Les coefficients de la capacité propres aux indicateurs du sous-système de lubrification.	143
Tableau 34 : Valeurs de Shapley relatives aux indicateurs du sous-système de lubrification.	143
Tableau 35 : Degrés d'interaction relatifs aux indicateurs du sous-système de lubrification.....	143
Tableau 36 : Les coefficients de la capacité pour les indicateurs du sous-système de refroidissement.	144
Tableau 37 : Valeurs de Shapley relatives aux indicateurs du sous-système de refroidissement.	144
Tableau 38 : Degrés d'interaction relatifs aux indicateurs du sous-système de refroidissement	144
Tableau 39 : Valeurs de Shapley relatifs aux indicateurs du sous-système base moteur.....	145
Tableau 40 : Degrés d'interaction relatifs aux indicateurs de la base moteur.....	145
Tableau 41 : Indicateurs contributeurs pour la construction des indicateurs du moteur diesel.	146
Tableau 42 : Les coefficients de la capacité pour le bloc2.	147
Tableau 43 : Coefficients de la capacité relative au niveau moteur diesel	149
Tableau 44 : Valeurs de Shapley relatives aux indicateurs du moteur diesel	149
Tableau 45 : Degrés d'interaction relatifs aux indicateurs du moteur.....	149
Tableau 46 : Valeurs des états de santé pour le premier cas.....	150
Tableau 47 : Valeurs des états de santé pour la troisième situation	151

Bibliographie de l'auteur

Conférences internationales

Abichou B., Voisin A., lung B. (2012 c). Choquet integral parameters inference for health indicators fusion within multi-levels industrial systems: application to components in series. 2nd IFAC Workshop on Advanced Maintenance Engineering, Service and Technology, IFAC A-Mest'12, 22-23 November, Seville, Spain.

Cet article est sélectionné pour une publication potentielle (après extension) dans le « Journal of Management Mathematics », pour le special issue on « maintenance modeling and engineering ».

Abichou B., Voisin A., lung B., Kosayyer N. (2012 b). Choquet integral capacities-based data fusion for system health monitoring. 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, IFAC SAFEPROCESS 2012, 29-31 August, Mexico City, Mexico.

Abichou B., Voisin A., lung B. (2012 a). Bottom-up capacities inference for health indicator fusion within multi-level industrial systems. IEEE International Conference on Prognostics and Health Management, IEEE PHM 2012, 18-21 June, Denver, USA.

Monnin M., **Abichou B.**, Voisin A., Mozzati C. (2011). Fleet historical cases for predictive maintenance. International Conference on acoustical and vibratory methods in surveillance and diagnostics, Surveillance 6, 25-26 October, Compiègne, France.

Rizzolo L., **Abichou B.**, Voisin A., Kosayyer N. (2011). Aggregation of Health Assessment Indicators of Industrial Systems. The 7th conference of the European Society for Fuzzy Logic and Technology, EUSFLAT-2011, 18-22 July, Aix-les-Bains, France.

Rapports de projets

Monnin M., Voisin A., **Abichou B.**, Levrat E., Morel D., lung B., Léger J.-B. (2010) BMCI - WP 2.4 Conception - Définition Système Intelligent Navire, Partie III – IHM. Deliverable projet BMCI, Juin 2010.

Monnin M., Voisin A., **Abichou B.**, Medina-Oliva G., Morel D., Levrat E., Léger J.-B., lung B. (2012 a). BMCI - WP 2.5 Conception - Définition Système Intelligent Flotte Partie I – Stockage. Deliverable projet BMCI, Février 2012.

Monnin M., **Abichou B.**, Medina-Oliva G., Voisin A., Morel D., Levrat E., Léger J.-B., lung B. (2012 b). BMCI - WP 2.5 Conception - Définition Système Intelligent Flotte Partie II – Traitement. Deliverable projet BMCI, Février 2012.

Monnin M., Voisin A., **Abichou B.**, Medina-Oliva G., Morel D., Levrat E., Léger J.-B., lung B. (2012 c). BMCI - WP 2.5 Conception - Définition Système Intelligent Flotte Partie III – IHM. Deliverable projet BMCI, Février 2012.

Medina-Oliva G., **Abichou** B., Monnin M., Voisin A., Morel D., lung B., Léger J.-B. (2012 a). BMCI - WP 3.8 Application - Métier Système Intelligent Flotte Descriptif des Architectures Technique et Fonctionnelle. Delivrable projet BMCI, décembre 2012.

Medina-Oliva G., **Abichou** B., Monnin M., Voisin A., Morel D., lung B., Léger J.-B. (2012 b). BMCI - WP 3.9 Tests Système Intelligent Flotte. Delivrable projet BMCI, décembre 2012.

Bibliographie

Abdulmajeed J., Abd Ghani A., Osman H., Ghosh P. (2010). Implementation of condition based maintenance for medium voltage underground system for TNB distribution, CIRED Workshop, Lyon France.

Ahren T. (2005). A study of maintenance performance indicators for the Swedish railroad system, Licentiate Thesis, Division of Operation and Maintenance Engineering, Lulea University of Technology, Sweden.

AFIS (2009). Découvrir et comprendre l'ingénierie Système, Ouvrage collectif AFIS.

Aleskerov E., Freisleben B., Rao B. (1997). Cardwatch: A neural network based database mining system for credit card fraud detection. IEEE Computational Intelligence for Financial Engineering, NY, USA.

Amin S., Byington C., Watson M. (2005). Fuzzy inference and fusion for health state diagnosis of hydraulic pumps and motors. In Proceedings of North American Fuzzy Information Processing Society NAFIPS, Michigan, USA.

Ando S. (2007). Clustering needles in a haystack: An information theoretic analysis of minority and outlier detection. IEEE International Conference on Data Mining, Omaha, USA.

Appleby M. P. (2010). Wear debris detection and oil analysis using ultrasonic and capacitance measurements, A Master of Science Thesis, the Graduate Faculty of the University of Akron.

Arif M. (2005). Fusion de Données : Ultime Etape de Reconnaissance de Formes, Applications à l'Identification et à l'Authentification. Thèse de doctorat, Université de Tours.

Aven T., Jensen U. (1999). Stochastic Models in Reliability. Springer-Verlag, (applications of mathematics), vol. (41), New York.

Basseville M., Cordier M. O. (1996). Surveillance et diagnostic de systèmes dynamiques : approches complémentaires du traitement du signal et de l'intelligence artificielle, Rapport INRIA, N°2861.

Basseville M. (1988). Detecting Changes in Signals and Systems – A Survey. Automatica, vol. (24), pp. 309-326.

Bechhoefer E., Duke A., Mayhew E. (2007). A Case for Health Indicators vs. Condition Indicators in Mechanical Diagnostics, Presented at the American Helicopter Society 63 th Annual Forum, Virginia Beach, VA.

Beliakov G., Pradera A., Calvo T. (2007). *Aggregation Functions: A Guide for Practitioners*. Springer Berlin Heidelberg.

Benedettini O., Baines T.S., Lightfoot H.W., Greenough R.M. (2009). State-of-the-art in integrated vehicle health management. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. (223), pp: 157—170.

Ben Salem A. (2008). *Modèles Probabilistes de Séquences Temporelles et Fusion de Décisions : Application à la Classification de Défauts de Rails et à leur Maintenance*. Thèse de Doctorat de l'Université Henri Poincaré, Nancy.

Berrah L., Clivillé V. (2007) Towards an aggregation performance measurement system model in a supply chain context, *Computers in industry*, vol. 58, pp 709-719

Berrah L., Mauris G., Montmain J. (2006). Monitoring the improvement of an overall industrial performance based on Choquet integral aggregation. *Omega, Special Issue on Multiple Criteria Decision Making for Engineering*, vol. (36), pp: 340–351.

Blanke M., Staroswiecki M., Wu N. E. (2001). Concepts and methods in fault-tolerant control. *American Control Conference*, Washington DC, USA.

BMCI (2009), *Bilan de santé pour la Maintenance et la Conduite Intelligentes, Description Détaillée du projet*.

Bonissone P.P., Varma A. (2005). Predicting the Best Unit within a Fleet: Prognostic Capabilities Enabled by Peer Learning, Fuzzy Similarity, and Evolutionary Design Process. *IEEE International Conference on Fuzzy Systems, FUZZ'05*.

Bouchon-Meunier B. (1998). *Aggregation and fusion of imperfect information*. Physica-Verlag, Heidelberg, Germany.

Boufaied A. (2003). *Contribution à la Surveillance Distribuée Des Systèmes à Evénements Discrets Complexes*. Thèse de doctorat, Université Paul Sabatier de Toulouse.

Brachet D., Thibault F. (2000). *Indicateurs de maintenance, des outils à la mesure d'une fonction*. Publication du Centre Technique des Industries Mécaniques (CETIM).

Byington C. S., Garga A. K. (2001). Data Fusion for Developing Predictive Diagnostics for Electromechanical Systems. Hall D.L., Llinas J. (Eds.), *Handbook of Multisensor Data Fusion*.

Byington C. S., Watson M., Edwards D. (2004). Dynamic signal analysis and neural network modeling for life prediction of flight control actuators. *American Helicopter Society Annual Forum*, Baltimore, USA.

Callan R., Larder B., Sandiford J. (2006). An integrated approach to the development of an intelligent prognostic health management system. In *Proceedings of the IEEE Aerospace Conference*, Montana, USA.

- Camci F., Valentine G., Navarra K. (2007). Methodologies for integration of PHM systems with maintenance data, in IEEE Aerospace Conference Proceedings, Montana, USA.
- Chandola V., Banerjee A., Kumar V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, vol. (14), pp: 3-18.
- Chen S. L., Craig M., Wood R. J. K., Wang L., Callan R., Powrie, H. E. G. (2008). Bearing condition monitoring using multiple sensors and integrated data fusion techniques. *International Conference in Vibrations in Rotating Machinery*, Oxford, UK.
- Cheng S., Azarian M.H., Pecht M.G. (2010). Sensor systems for prognostics and health management. *Sensors 2010*, vol. (10), pp: 5774-5797.
- Chimentin X. (2006). Localisation et quantification des sources vibratoires dans le cadre d'une maintenance préventive conditionnelle en vue de fiabiliser le diagnostic et le suivi de l'endommagement des composants mécaniques tournants : application aux roulements à billes. Thèse de doctorat de l'Université de Reims, Champagne Ardenne.
- Clivillé V. (2004). Approche systémique et méthode multicritère pour la définition d'un système d'indicateurs de performance. Thèse de Doctorat, Université de Savoie.
- Coble J., Humberstone M., Hines J. W. (2010). Adaptive Monitoring, Fault Detection and Diagnostics, and Prognostics System for the IRIS Nuclear Plant. *Annual Conference of the Prognostics and Health Management Society*, Portland, USA.
- Cocheteux P. (2010). Contribution à la maintenance proactive par la formalisation du processus de pronostic des performances des systèmes industriels, thèse de doctorat, Université de Lorraine.
- Cover T. and Thomas J. (2006). *Elements of Information Theory*. Wiley- Interscience, 2006.
- Detyniecki M. (2001). *Fundamentals on aggregation operators*. AGOP, Berkeley, 2001.
- Discenzo F. M., Nickerson W., Mitchell C. E., Keller K. J. (2001). *Open Systems Architecture Enables Health Management for Next Generation System Monitoring and Maintenance*, Development White Paper, OSA-CBM™ Development Group.
- DoD (2008). *Condition Based Maintenance Plus*, Department of Defense Guidebook, May 2008.
- Domingo-Ferrer J., Torra V. (2002). Approximating fuzzy measures by hierarchically decomposable ones. *International Conference on Information Fusion*, Annapolis, USA.
- Dong M., He D. (2007). Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis. *Stochastics and Statistics. European Journal of Operational Research*, vol. (178), pp: 858–878.
-

Dubois D., Prade H. (2004). On the use of aggregation operations in information fusion processes. *Fuzzy Sets and Systems*, vol. (142), pp : 143–161.

Frank P.M. (1990). Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge Based Redundancy – A survey and New Results. *Automatica*, vol. (26), pp: 459-474.

Fujimaki R., Yairi T., and Machida K. (2005). An approach to spacecraft anomaly detection problem using kernel feature space. *ACM SIGKDD international conference on Knowledge discovery in data mining*, New York, USA.

Ghetie M. A. (1997). Vers le concept de redondance algorithmique pour le diagnostic. Thèse de doctorat, Université de Nancy 1, Nancy

George J., Crassidis J.L., Singh T., Fosbury A.M. (2011). Anomaly Detection using Context-Aided Target Tracking. *Journal of Advances in Information Fusion*. Vol. (6), pp : 39—56.

Ginart, A. and Barlas, I. and Goldin, J. and Dorrrity, J.L. (2006). Automated feature selection for embeddable prognostic and health monitoring (PHM) Architectures. *IEEE Autotestcon*, Anaheim, USA.

Goldstein M., Dengel A. (2012). Histogram-based Outlier Score (HBOS): A fast Unsupervised Anomaly Detection Algorithm. *Conference on Artificial Intelligence*, Saarbrücken, Germany.

Gomez D., Montero J. (2004). A discussion on aggregation operators. *Kybernetika*, vol. (40), pp : 107—120.

Grabisch M. (1995). A new algorithm for identifying fuzzy measures and its application to pattern recognition. *IEEE International Conference on Fuzzy Systems*, Yokohama, Japan.

Grabisch M. (1997 a). k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, vol.(92), pp :167–189.

Grabisch M. (1997 b). Alternative representations of discrete fuzzy measures for decision making. *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems*, vol.(5), pp :587–607.

M. Grabisch, C. Labreuche, (2004) Fuzzy measures and integrals, in: J. Figueira, S. Greco, M. Ehrgott (Eds.), *MCDAs. Multiple Criteria Decision Analysis*, Kluwer Academic Publishers, pp. 563–608.

Grabisch M., Perny P. (2003). Agrégation multicritère. in *Logique Floue, principes, aide à la décision*, Bouchon-Meunier B. and Marsala C., Eds. *Hermès-Lavoisier*, 2003, pp. 81–120.

Grabisch M. (2006). Evaluation subjective. In D. Bouyssou, D. Dubois, M. Pirlot, and H. Prade, editors, *Concepts et Méthodes pour l'Aide à la Décision*, IC2, pages 175-232. *Hermès*.

Grabisch, M., Kojadinovic, I., Meyer, P. (2007) A review of methods for capacity identification in Choquet integral based multi-attribute utility theory Applications of the Kappalab R package. *European Journal of Operational Research*, vol. (186), pp. 766—785.

Grabisch M., Marichal J.-L., Mesiar R., Pap E. (2011 a). Aggregation functions: Means. *Information Sciences*, vol. (181), pp: 1–22.

Grabisch M., Marichal J.-L., Mesiar R., Pap E. (2011 b). Aggregation functions: Construction methods, conjunctive, disjunctive and mixed classes. *Information Sciences*, vol. (181), pp: 23–43.

Gu Y., McCallum A., Towsley D. (2005). Detecting anomalies in network traffic using maximum entropy estimation. *Internet Measurement Conference*, Berkeley, USA.

Hall D.L., Llinas J. (1997). An introduction to multisensor data fusion. *IEEE Conference of Micro Electro Mechanical Systems, MEMS '97*, Nagoya , Japan.

He Z., Xu X., Deng S. (2005). An optimization model for outlier detection in categorical data. *Conference on Intelligent Computing*, Hefei, China.

Hjartarson T., Ota S. (2006). Predicting future asset condition based on current Health Index and maintenance level. *Transmission and distribution Construction, Operation and Live Line Maintenance*.

Hughes D. T. (2003). The use of “Health indices” to determine end of life and estimate remnant life for distribution assets, *International conference on electricity distribution*, Barcelona, Spain.

Isermann R., Balle P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control engineering practice*, vol. (5), pp : 709—719.

Iso 13374 - condition monitoring and diagnostics of machines – data processing, communication and presentation, 2002.

Jaloretto Alves M. R., Bizarria C. O., Harrop Galvao R. K. (2009). Trend analysis for prognostics and health monitoring. *Brazilian Symposium on Aerospace Eng. & Applications*, SP, Brazil.

Jaw L. C. (2005), Recent advancements in aircraft engine health management (ehm) technologies and recommendations for the next step. `Turbo Expo 2005, 50th ASME International Gas Turbine & Aero engine Technical Congress', Scientific Monitoring Inc, Scottsdale, U.S.A.

Jullien S., Mauris G., Valet L., Bolon P., Teyssier S. (2008). Identification of Choquet integral's parameters based on relative entropy and applied to classification of tomographic images. *Proceedings of IPMU'08*, Malaga, Espagne.

- Kacprzyński G. J., Roemer M. J., Hess A.J. (2002). Health management system design: Development, simulation and cost/benefit optimization, IEEE Aerospace conference proceeding, Montana, USA.
- Kalgren P.W., Byington C.S., Roemer, M.J. (2006). Defining PHM, a lexical evolution of maintenance and logistics. In proceedings of IEEE Autotestcon, Anaheim, CA, USA.
- Kar S., Sahoo B. (2009). An anomaly detection system for DDoS attack in grid computing. International journal of computer applications in engineering, technology and sciences, vol. (1), pp: 553-558.
- Katipamula S., Brambley M.R. (2005). Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I. HVAC\R Research, vol. (11), pp: 3—25.
- Khelil Y., Djezeri M., Graton G., Ouladsine M. (2010). BMCI WP2.6 Diagnostic bas niveau pour le moteur Diesel marin. Délivrable projet BMCI
- Kind A., Stoecklin M.P., Dimitropoulos X. (2009). Histogram-based traffic anomaly detection. IEEE Transactions on Network and Service Management, vol. (2), pp : 110--121.
- King S.P., King D.M., Anuzis P., Astley K., Tarassenko L., Hayton P., Utete S. (2002). The use of novelty detection techniques for monitoring high-integrity plant, International Conference on Control Applications, Cancun, Mexico.
- Kiritsis D., Nguyen V. K., Stark J. (2008). How closed-loop PLM improves Knowledge Management over the complete product lifecycle and enables the factory of the future. International Journal of Product Lifecycle Management, vol. (3), pp : 54-77.
- Ko J.M., Ni Y.Q. (2005). Technology developments in structural health monitoring of large-scale bridges. Engineering Structures, vol. (27), pp: 1715-1725.
- Kojadinovic I. (2002). Estimation of the weights of interacting criteria from the set of profiles by means of information-theoretic functional. European Journal of Operational Research, vol.(155), pp : 741–751.
- Kojadinovic I. (2004 a). Estimation of the weights of interacting criteria from the set of profiles by means of information-theoretic functionals. In European Journal of Operational Research, vol. (155), pp: 741–751.
- Kojadinovic I. (2004 b). Unsupervised aggregation by the choquet integral based on entropy functionals: Application to the evaluation of students. Modeling Decisions for Artificial Intelligence Conference, Barcelona, Spain.
- Kullback S., Leibler R. A. (1951). On information and sufficiency. The Annals of Mathematical Statistics, vol. (22), pp : 79–86.
-

- Kumar S., Dolev E., Pecht M. (2009). Parameter selection for health monitoring of electronic products. *Microelectronics Reliability*, vol. (50), pp: 161-168.
- Kumar S., Pecht M. (2010). Modelling approaches for prognostics and health management of electronics. *International journal of performability engineering*, vol. (6), pp: 467-476.
- Larbani M., Huang C.Y., Tzeng G.H. (2011). A novel method for fuzzy measure identification. *International Journal of Fuzzy Systems*, vol. (13), pp : 24—34.
- Lakhina A., Crovella M., Diot C. (2005). Mining anomalies using traffic feature distributions. *ACM SIGCOMM Computer Communication Review*, vol. (35), pp: 217-228
- Lebold M., Thurston M. (2001). Open Standards for Condition-Based Maintenance and Prognostic Systems. 5th Annual Maintenance and Reliability Conference (MARCON 2001), Gatlinburg, USA.
- Lee W., Xiang D. (2001 a). Information-theoretic measures for anomaly detection. *IEEE Symposium on Security and Privacy*, Oakland, USA.
- Lee L. (2001 b). On the Effectiveness of the Skew Divergence for Statistical Language Analysis. *Artificial Intelligence and Statistics*.
- Lee J., Ni L., Djurdjanovic D., Qiu H., Liao H. (2006). Intelligent prognostics tools and e-maintenance. *Computers in Industry*, vol. (57), pp: 476-489.
- Léger J. B., lung B., Ferro de Beca A., Pinoteau J. (1999 a). An innovative approach for new Distributed Maintenance System: application to Hydro Power Plants of the REMAFEX project. *Computer in Industry*, vol. (38), pp. 133 – 150.
- Léger J. B. (1999). Contribution méthodologique à la maintenance prévisionnelle des systèmes industriels de production, thèse de niveau doctorat, Université de Nancy 1, Nancy.
- Léger J.B., lung B. (2012). Ships fleet-wide management and naval mission prognostics: Lessons learned and new issues. *IEEE Conference on Prognostics and Health Management (PHM)*, Denver, USA.
- Levrat E., lung B. (2007). TELMA : A full e-maintenance platform. *Proceedings of the Second World Congress on Engineering Asset Management and the Fourth International Conference on Condition Monitoring WCEAM/CM*, Harrogate, Angleterre.
- Liu X., Ma L., Mathew J. (2006). Machinery fault diagnosis based on feature level fuzzy integral data fusion techniques. *IEEE Conference on Industrial Informatics*, Singapore.
- Liu X. (2007). Machinery fault diagnostics based on fuzzy measure and fuzzy integral data fusion techniques, PhD thesis, Queensland University of Technology.
-

Marichal J-L., Roubens M. (2000 a). Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, vol. (124), pp: 641–650.

Marichal J.L. (2000 b). On Choquet and Sugeno integrals as aggregation functions. In *Fuzzy measures and integrals*, vol. (40), pp: 247-272.

Markou M., Singh S. (2003 a). Novelty detection: a review-part 1: statistical approaches. *Signal Processing*, vol. (83), pp: 2481-2497.

Markou, M. and Singh, S. 2003b. Novelty detection: a review-part 2: neural network based approaches. *Signal Processing*, vol. (83), pp : 2499-2521.

Martin K.F. (1994). A review by discussion of condition monitoring and fault-diagnosis in machine-tools. *International Journal of Machine Tools and Manufacture*, vol. (34), pp: 527-551.

Mokhtari A. (2007). Diagnostic des systèmes hybrides : développement d'une méthode associant la détection par classification et la simulation dynamique. Thèse de doctorat, Université de Toulouse.

Monnin M., Voisin A., Morel D., Levrat E., Leger J.B., and lung B. (2009). BMCI WP 1.6 spécification des besoins système intelligent navire. *Delivrable du projet BMCI*.

Montmain J., Denguir A., Tardy J. (2004). An Interactive Decision-Making Support System for Organization Management and Diagnosis. *Modelling, Computation and Optimization in Information Systems and Management Sciences*, MCO'04, Metz, France.

Morachioli F., Delorme T., Piron-Cabaret D. (2011). BMCI WP 2.7 : Définition des campagnes banc et navire. *Délivrable projet BMCI*.

Muller A. (2005) Contribution à la maintenance prévisionnelle des systèmes de production par la formalisation d'un processus de pronostic. *Doctorat de l'Université Henri Poincaré, Nancy 1*.

Murofushi T., Sugeno M. (1991). A theory of fuzzy measures. Representation, the Choquet integral and null sets. In *Journal of Mathematical Analysis and Applications*, vol. (159), pp: 532-549.

Muldoon R.C., Gill J., Brock L.D. (1999). Integrated mechanical diagnostic (IMD) health and usage monitoring system (hums); an open system implementation case study. *Digital Avionics Systems Conference, Saint-Louis, USA*.

Naderian A., Cress S., Piercy R., Wang F., Service J. (2008). An Approach to Determine the Health Index of Power Transformers. *IEEE international symposium on electrical insulation, Vancouver, Canada*.

Niu G., Yang B.S., Pecht M. (2010 a). Development of an optimized condition based maintenance system by data fusion and reliability-centered maintenance. *Reliability Engineering and System Safety*, vol. (95), pp :786–796.

Niu G., Yang B.S. (2010 b). Intelligent condition monitoring and prognostics system based on data-fusion strategy. *Expert Systems with Applications*, vol. (37), No: 12, pp : 8831–8840.

Ortiz E.M., Babbar A., Syrmos V.L., Clark G.J., Vian J.L., Arita M.M. (2008). Multi source data integration for aircraft health management. In proceedings of IEEE Aerospace Conference, Montana, USA.

Pan Y., Chen J., Li X. (2010). Bearing performance degradation assessment based on lifting wavelet packet decomposition and fuzzy c-means. *Mechanical Systems and Signal Processing*, vol. (24), pp: 559-566.

Patcha A., Park J.M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, vol. (51), pp : 3448—3470.

Patrick R., Smith J.S., Byington C.S., Vachtsevanos G.J., Tom K., Ly C. (2010). Integrated Software Platform for Fleet Data Analysis, Enhanced Diagnostics and Safe Transition to Prognostics for Helicopter Component CBM. In Proceedings of the Annual Conference of the prognostics and health management society, Portland, USA.

Peres F., Bouzaïene L., Bocquet J. C., Billy F., Lannoy A., Haïk P. (2007). Anticipating aging failure using feedback data and expert judgment. *Reliability Engineering & System Safety*, vol. (92),pp : 200-210.

Perona I., Albisua I., Arbelaitz O., Gurrutxaga I., Martin J.I., Muguerza J., Pérez J.M. (2010). Histogram based payload processing for unsupervised anomaly detection systems in network intrusion. 14th Portuguese Conference on Artificial Intelligence, Aveiro, Portugal.

Perrot, N., Trystram, G., Le Guennec, D., Guely, F., 1996. Sensor fusion for real time quality evaluation of biscuit during baking, Comparison between Bayesian and fuzzy approaches. *Journal of Food Engineering*, vol. (29), pp: 301–315.

Peysson F., Rizzolo L., Monnin M., Morel D. (2011). BMCI – WP 3.7 Test du Système Intelligent Navire. Délivrable Projet BMCI.

Pillet M. (1995). Appliquer la maîtrise statistique des procédés MSP/SPC. Les Éditions d'Organisation.

Rabatel J., Bringay S., Poncelet P. (2011). Anomaly detection in monitoring sensor data for preventive maintenance. *Expert Systems with Applications*, vol. (38), pp: 7003–7015.

Racoceanu D. (2006). Contribution à la surveillance des Systèmes de Production en utilisant les Techniques de l'Intelligence Artificielle. Synthèse des activités de recherche et d'encadrement, Université de FRANCHE-COMTÉ de Besançon.

Ragot J., Darouach M., Maquin M., Bloch G. (1990). Validation de données et diagnostic. *Traité des nouvelles technologies, série diagnostic et maintenance*, Hermès.

Ramasso E., Jullien S. (2011). Parameter identification in Choquet Integral by the Kullback-Leibler divergence on continuous densities with application to classification fusion. *European Society for Fuzzy Logic and Technology, EUSFLAT - LFA'11.*, Aix-Les-Bains : France.

Raol J. R. (2010). *Multi-Sensor Data Fusion with MATLAB®*. Taylor and Francis Group, LLC.

Rayhane H. (2004). *Surveillance des systèmes de production automatisés : détection et diagnostic*. Thèse de doctorant, Institut National Polytechnique de Grenoble.

Ribot P. (2009). *Vers l'intégration diagnostic/pronostic pour la maintenance des systèmes complexes*, thèse de doctorat, Université de Toulouse, 2009.

Rizzolo L., Abichou B., Voisin A., Kosayyer N. (2011), Aggregation of Health Assessment Indicators of Industrial Systems. In *Proceedings of the 7th conference of the European Society for Fuzzy Logic and Technology, EUSFLAT-2011*, Aix-Les-Bains, France.

Roemer M. J., Kacprzyński G. J., Orsagh R. F. (2001). Assessment of data and knowledge fusion strategies for prognostics and health management. In *proceedings of aerospace conference*, IEEE, Big Sky, USA.

Roth A. (1988). *The Shapley value: Essays in honor of Lloyd S. Shapley*. Cambridge University Press, Cambridge.

Sabeh Z. (2006). *Diagnostic à base de modèle : application à un moteur diesel suralimenté à injection directe*. Thèse de doctorat. Institut National Polytechnique de Lorraine

Sabnavis G., Kirk R. G., Kasarda M., Quinn D. (2004). Cracked shaft detection and diagnostics: A literature review. *The Shock and Vibration Digest*, vol. (36), pp: 287–296.

Sénéchal O. (2004). *Pilotage des systèmes de production vers la performance globale. Habilitation à diriger les recherches*, Université de Valenciennes et du Hainaut Cambrésis (UVHC).

Shannon C. E. (1948). A mathematical theory of communication. *Bell System Technical J.*, vol. (27), pp: 379–423.

Shin J. (2009). *Decision support methods for closed-loop conceptual design*. Thèse de doctorant, Ecole Polytechnique Fédérale de Lausanne.

Shukla S. N., Kshirsagar J. T. (2002). Numerical experiments on a centrifugal pump. *ASME 2002 Joint U.S.-European Fluids Engineering Division Conference*, Montreal, Canada

Song X., Wu M., Jermaine C., Ranka S. (2007). Conditional anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, vol. (19), pp: 631-645.

Sotiris V. A., Tse P. W., Pecht M. G. (2010). Anomaly detection through a bayesian support vector machine. *IEEE Transactions on Reliability*, vol. (59), pp: 277-286.

Spence C., Parra L., Sajda P. (2001). Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, Washington, DC, USA.

Srivastav A., Ray A., Gupta S. (2009). An information-theoretic measure for anomaly detection in complex dynamical systems. *Mechanical Systems and Signal Processing*, vol. (23), pp: 358– 371.

Sugeno M. (1974). *Theory of fuzzy integrals and its applications*, Doctoral Thesis, Tokyo Institute of Technology.

Takata S., F. Kimura, F.J.A.M. van Houten, E. Westkämper, M. Shpitalni, D. Ceglarek et J. Lee (2005). Maintenance: changing role in lice cycle management. *CIRP Annals - Manufacturing Technology*, vol. (53), pp: 643-655.

Tarassenko L. (1995). Novelty detection for the identification of masses in mammograms. *IEEE International Conference on Artificial Neural Networks*, Cambridge, UK.

Theilliol D., Mahfouf M., Ponsart J. C., Sauter D., Gama M. A. (2010). Design of a FDI System based on a Bank of Filter-Observers applied to a Hot Rolling Mill. *Transactions of the Institute of Measurement and Control*, vol. (32), pp : 265-285.

Thurston M., Lebold M. (2001). Standards Developments for Condition-Based Maintenance Systems, Improving Productivity Through Applications of Condition Monitoring. 55th Meeting of the Society for Machinery Failure Prevention Technology, Virginia, USA.

Torra V., Narukawa Y. (2007). *Modeling decisions information fusion and aggregation operators*. Springer-Verlag, Berlin, Germany.

Tversky A., Kahneman D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and uncertainty*, vol. (5), pp : 297—323.

Umiliacchi P., Lane D., Romano F. (2011), Predictive Maintenance of Railway Subsystems Using an Ontology based Modelling Approach. In proceedings of the 9th World Conference on Railway Research, Lille, France.

Venkatasubramanian V., Rengaswamy R., Kavuri S. N., Yin K. (2003 a). A review of process fault detection and diagnosis Part I: Quantitative model based methods. *Computers & chemical engineering*, vol. (27), pp : 293–311.

Venkatasubramanian V., Rengaswamy R., Kavuri S. N. (2003 b). A review of process fault detection and diagnosis Part II: Qualitative models and search strategies. *Computers & chemical engineering*, vol. (27), pp : 313-326.

- Venkatasubramanian V., Rengaswamy R., Kavuri S. N., K. Yin (2003 c). A review of process fault detection and diagnosis Part III: Process history based methods. *Computers & chemical engineering*, vol. (27), pp : 327–346.
- Verma A. K. , Srividya v, Ramesh P. G. (2010). A Systemic Approach to Integrated E-maintenance of Large Engineering Plants. *International Journal of Automation and Computing*, vol. (7), pp.173-179.
- Verron S. (2007). Diagnostic et surveillance des processus complexes par réseaux bayésiens. Thèse de doctorat, Université d'Angers.
- Vichare N., Pecht M. (2006). Prognostics and Health Management of Electronics. *IEEE Transactions on Components and Packaging Technologies*, vol. (29), pp: 222–229.
- Vichare N., Rodgers P., Evely V., Pecht M. (2007). Environment and usage monitoring of electronic products for health assessment and product design. *International Journal of Quality Technology and Quantitative Management*, vol. (4), pp: 235-250.
- Villemeur A. (1988) Sûreté de fonctionnement des systèmes industriels : Fiabilité, Facteurs humains, Informatisation, Eyrolles, Paris.
- Volponi A. J., Brotherton T., Luppold R., Simon D. L. (2003). Development of an information fusion system for engine diagnostics and health management. JANNAF 27th airbreathing propulsion subcommittee meeting, Colorado, USA.
- Volponi A. J., Brotherton T., Luppold R. (2004). Development of an Information Fusion System for Engine Diagnostics and Health Management. AIAA Intelligent Systems Technical Conference, Chicago, USA.
- Von Bertalanffy L. (1968). *General System Theory*. G. Braziller Editions.
- Wald L. (1999). Definitions and terms of reference in data fusion. fusion of sensor data, knowledge sources and algorithms for extraction and classification of topographic objects, Valladolid, Spain.
- Wang W.Q., Ismail F., Farid Golnaraghi M. (2001). Assessment of gear damage monitoring techniques using vibration measurements. *Mechanical Systems and Signal Processing*, vol. (15), pp: 905–922.
- Williams Z. (2006). Benefits of IVHM: an analytical approach. In *Proceedings of the 2006 IEEE Aerospace Conference*, Big Sky, Montana, USA.
- Wiliem A., Madasu V., Boles W., Yarlagadda P. (2012). A suspicious behaviour detection using a context space model for smart surveillance systems. *Computer Vision and Image Understanding*, vol. (116), pp : 194–209.
-

Wu H., Siegel M., Stiefelhagen R., Yang J. (2002). Sensor fusion using Dempster-Shafer theory. IEEE Instrumentation and Measurement Technology Conference, Anchorage, USA.

Xie J., Pecht M. (2004). Application of in situ health monitoring and prognostic sensors. 9th Pan Pacific Microelectronics Symposium Exhibits and Conference, Oahu, USA.

Yan R., Gao R. X. (2007). Approximate Entropy as a diagnostic tool for machine health monitoring. Mechanical Systems and Signal Processing, vol. (21), pp: 824–839.

Yang B.S., Kim K.J. (2006). Application of Dempster–Shafer theory in fault diagnosis of induction motors using vibration and current signals. Mechanical Systems and Signal Processing, vol. (20), pp: 403–420.

Young J., Lin M., Liyuan W., Jian T., Lin H. (2010). A new method to determine condition index for equipment condition assessment. Prognostics and health management conference Macao, China.

Yu L.J., Cleary D., Osborn M., Rajiv V. (2007). Information fusion strategy for aircraft engine health management. ASME Turbo Expo 2007: Power for Land, Sea and Air, Montréal, Canada.

Zhang L., Yan R., Gao R. X., Lee K. (2007). Design of a Real-time Spindle Health Monitoring and Diagnosis System Based on Open Systems Architecture. International Smart Machining Systems Conference, Gaithersburg, USA.

Zille V. (2009). Modélisation et évaluation des stratégies de maintenance complexes sur des systèmes multi-composants. Thèse de Doctorat, Université de technologie de Troyes.

Zhong-Qing W., Zhi-Nong J., Zhong Xin M. B. (2010). The slow-changing alarm system of condition monitoring for rotating machinery. WSEAS Transactions on Systems, vol. (9), pp: 52-61.

Titre :

Contribution à la formalisation de bilans/états de santé multi-niveaux d'un système pour aider à la prise de décision en maintenance : agrégation d'indicateurs hétérogènes par « l'intégrale de Choquet ».

Résumé : Dans cette thèse est défendu l'intérêt d'évaluer la santé d'un système/objet industriel multi-composants à travers un bilan de santé multi-niveaux hiérarchiques. Elle a donc pour objet principal de justifier les éléments essentiels du concept de bilan de santé générique qui représente l'état réel d'un système sous la forme d'un vecteur d'indicateurs de différentes natures. Vis-à-vis de ce fondement, la thèse se focalise plus spécifiquement sur les fonctions de détection des anomalies-normalisation et agrégation d'indicateurs pour élaborer un index synthétique représentatif de l'état de santé global pour chaque élément du système. Il est ainsi proposé, une nouvelle approche de détection conditionnelle des anomalies. Cette approche a pour intérêt de quantifier la déviation pour chaque indicateur par rapport à son mode de comportement nominal tout en prenant en compte le contexte dans lequel évolue le système. Une extension à l'exploitation de l'intégrale de Choquet en tant qu'opérateur d'agrégation des indicateurs est aussi proposée. Cette extension concerne, d'une part, un processus d'apprentissage non supervisé des capacités pour le niveau le plus inférieur dans l'abstraction, à savoir celui des composants, et d'autre part, une approche de mise en œuvre de leur inférence d'un niveau à l'autre. Ces propositions sont appliquées sur un moteur diesel de navire, système essentiel du projet BMCI du pôle MER-PACA dans lequel s'inscrit cette thèse.

Mots-clés : PHM, Projet BMCI, Intégrale de Choquet, Bilan de santé, détection des dérives, agrégation d'indicateurs, identification des capacités.

Title : Contribution to the formalization of health assessment for a multi-layers system to aid maintenance decision making: Choquet integral-based aggregation of heterogeneous indicators.

Abstract : This work is addressing the health assessment of a multi-component system by means of multi-levels health check-up. Thus scientific Ph. D. objective aims to establish items of a generic health check-up concept. It focuses specifically on the functions of anomaly detection, normalization and aggregation of different indicators to develop a synthetic index representing the overall health status for each element within the system. In that way, it is proposed a new approach for detecting conditional anomalies. This approach has the advantage of quantifying the deviation for each indicator compared to its nominal behavior while taking into account the context in which the system operates. An extension of the Choquet integral used as an operator aggregating indicators is also proposed. This extension regards on the one hand, a process of an unsupervised learning of the capacity coefficients for the lowest level of abstraction, namely components level, and on the other hand, an approach to inference them from one level to another. These contributions are implemented on a ship diesel engine which is the most critical system for the BMCI project of the MER-PACA pole to which this thesis is attached.

Keywords: PHM, BMCI Project, Choquet integral, Health check-up, anomalies detection, indicators aggregation, capacities identification.