

# THESE DE DOCTORAT DE

L'UNIVERSITE D'ANGERS

COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 602

*Sciences pour l'Ingénieur*

Spécialité : « *Procédés de Fabrication, Optimisation de Process et de Produits* »

Par

« **Khaoula TIDRIRI** »

« **Fusion de décisions dédiée à la surveillance des systèmes complexes** »

Thèse présentée et soutenue à « Angers », le « 16 Octobre 2018 »

Unité de recherche : Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)

Thèse N° : 139864

## Rapporteurs avant soutenance :

Louise TRAVE-MASSUYES Directrice de Recherche,  
LAAS-CNRS, Université de  
Toulouse

Didier THEILLIOL Professeur,  
CRAN, Université de  
Lorraine

## Composition du Jury :

### Président:

Kamal MEDJAHER

Professeur,  
ENIT, Université de Toulouse

### Examineurs:

Audine SUBIAS

Maître de Conférence-HDR,  
LAAS, Université de Toulouse

Belkacem OULD BOUAMAMA

Professeur,  
CRISTAL, Université de Lorraine

### Directeur de thèse:

Teodor TIPLICA

Maître de Conférence-HDR,  
LARIS, Université d'Angers

### Co-directeurs :

Sylvain VERRON

Maître de conférences,  
LARIS, Université d'Angers

Nizar CHATTI

Maître de conférences,  
LARIS, Université d'Angers



**L'auteur du présent document vous autorise à le partager, reproduire, distribuer et communiquer selon les conditions suivantes :**



- Vous devez le citer en l'attribuant de la manière indiquée par l'auteur (mais pas d'une manière qui suggérerait qu'il approuve votre utilisation de l'œuvre).
- Vous n'avez pas le droit d'utiliser ce document à des fins commerciales.
- Vous n'avez pas le droit de le modifier, de le transformer ou de l'adapter.

**Consulter la licence creative commons complète en français :**  
**<http://creativecommons.org/licences/by-nc-nd/2.0/fr/>**



# REMERCIEMENTS

Ce projet de recherche a été mené à bien grâce à l'aide et au soutien incroyable de nombreuses personnes. Comment ne pas prendre le temps de vous remercier, tous.

Je tiens tout d'abord à remercier Monsieur Teodor TIPLICA, qui a dirigé ce travail avec une grande motivation. Son soutien constant, ses conseils judicieux et ses encouragements m'ont motivé tout au long de ces trois années de thèse. Je lui témoigne toute ma reconnaissance pour sa grande disponibilité, sa générosité intellectuelle et pour la confiance et l'autonomie qu'il a su m'accorder dès le départ.

Je remercie chaleureusement Messieurs Sylvain VERRON et Nizar CHATTI pour leur encadrement et leur aide précieuse. Merci pour toutes ces discussions perspicaces sans fin, qui ont eu un impact considérable sur mon travail de recherche, mais qui ont surtout suscité mon enthousiasme pour la recherche. Merci d'avoir partagé votre temps et vos connaissances.

Merci Sylvain également, pour ta sympathie et ta bonne humeur.

Merci Nizar également, pour tes nombreux conseils. Je les garde tous en mémoire.

Mes remerciements s'adressent ensuite aux membres du jury qui m'ont fait l'honneur de participer à la soutenance. Tout d'abord, je remercie Madame Louise Travé-Massuyès et Monsieur Didier THEILLIOL, pour avoir accepté d'évaluer mes travaux en qualité de rapporteurs. Ensuite, je remercie Madame Audine SUBIAS, Monsieur Belkacem OULD BOUAMAMA et le Président du jury Monsieur Kamal MEDJAHER, pour l'intérêt qu'ils ont porté à mes activités en tant qu'examineurs.

Je voudrais également exprimer toute ma gratitude au personnel du LARIS et de l'ISTIA, et particulièrement à Marie-Françoise GERARD et Kristell TANQUERAY pour leur disponibilité et leur gentillesse.

Je n'oublie pas mes collègues du LARIS, anciens et présents, pour leur soutien et leurs qualités humaines : Lamia, Fally, Nihad, Nadare, Ali, Thérèse, Mirvana, Mohamad, Selma, Pejman, Rodrigue,... Merci pour les moments agréables partagés durant ces trois ans. J'ai une grande pensée également pour mes amies de longue date, Salama et Khaoula, qui ont toujours été à mes côtés.

Un énorme merci à ceux qui m'entourent depuis toujours, mes parents, ma sœur et mon frère. Ils sont mon cercle premier et mon plus grand soutien. Je remercie chaleureusement et sans mesure mes parents pour leur encouragement et leur soutien indéfectible. Je vous serai éternellement et immensément reconnaissante pour TOUT.

Je conclurai en remerciant celui qui me supporte au quotidien, mon mari. Merci pour le soutien que tu m'apportes dans tout ce que j'entreprends, pour ta patience, pour ta confiance et pour tout le reste.



# Table des matières

<b>1</b>	<b>Introduction générale</b>	<b>7</b>
1.1	Contexte . . . . .	7
1.2	Contributions . . . . .	8
1.3	Organisation . . . . .	10
<b>2</b>	<b>Surveillance des systèmes complexes</b>	<b>11</b>
2.1	Introduction . . . . .	11
2.2	Approches pour la surveillance . . . . .	13
2.2.1	Détection et diagnostic à base de modèle . . . . .	13
2.2.2	Détection et diagnostic à base de données historiques . . . . .	17
2.2.3	Pronostic à base de modèle . . . . .	23
2.2.4	Pronostic à base de données . . . . .	25
2.2.5	Discussion . . . . .	27
2.3	Fusion d'approches pour la surveillance . . . . .	29
2.3.1	Fusion de méthodes . . . . .	30
2.3.2	Fusion de décisions . . . . .	36
2.3.3	Discussion . . . . .	46
2.4	Positionnement scientifique . . . . .	49
2.5	Conclusion . . . . .	50
<b>3</b>	<b>Fusion de décisions pour le diagnostic</b>	<b>53</b>
3.1	Introduction . . . . .	54
3.2	Formulation du problème . . . . .	54
3.3	Démarche méthodologique . . . . .	57
3.3.1	Evaluation des performances individuelles . . . . .	58
3.3.2	Stratégie de FDD : Définition des objectifs . . . . .	62
3.3.3	Stratégie de FDD : Cadre générique de fusion . . . . .	65
3.3.4	Stratégie de FDD : Développement des objectifs . . . . .	67
3.3.5	Stratégie de FDD : Prise de décision multi-objectifs sous contraintes . . . . .	72
3.3.6	Fusion des décisions en ligne . . . . .	76
3.4	Discussion . . . . .	78
3.4.1	Choix des objectifs . . . . .	78
3.4.2	Analyse des paramètres obtenus . . . . .	79
3.4.3	Synthèse . . . . .	87
3.5	Application : Diagnostic du Tennessee Eastman Process . . . . .	88
3.5.1	Introduction . . . . .	88
3.5.2	Présentation du TEP . . . . .	89
3.5.3	Diagnostic du TEP par le modèle . . . . .	91
3.5.4	Diagnostic du TEP par les données . . . . .	100
3.5.5	Diagnostic du TEP par la fusion de décisions . . . . .	101

3.6	Conclusion	106
<b>4</b>	<b>Fusion de décisions pour le pronostic</b>	<b>107</b>
4.1	Introduction	108
4.2	Formulation du problème	109
4.3	Démarche méthodologique de pronostic	111
4.3.1	Stratégie d'apprentissage	112
4.3.2	Estimation de l'état de santé et prédiction de la RUL	115
4.4	Démarche méthodologique de fusion	119
4.4.1	Evaluation des performances individuelles	120
4.4.2	Stratégie de pronostic : Définition des objectifs	122
4.4.3	Stratégie de pronostic : Cadre générique de fusion	123
4.4.4	Stratégie de pronostic : Développement des objectifs	123
4.4.5	Stratégie de pronostic : Prise de décision multi-objectifs sous contraintes	127
4.4.6	Fusion des décisions en ligne	128
4.5	Synthèse	131
4.6	Application : Pronostic de turboréacteurs à double flux	132
4.6.1	Introduction	132
4.6.2	Présentation du turboréacteur à double flux	132
4.6.3	Pronostic de turboréacteurs par des méthodes individuelles	134
4.6.4	Pronostic de turboréacteurs par la fusion de décision	139
4.7	Discussion	141
4.7.1	Prédiction de la RUL	141
4.7.2	Analyse des paramètres obtenus	147
4.7.3	Comparaison avec la littérature	153
4.8	Conclusion	155
<b>5</b>	<b>Conclusions et perspectives</b>	<b>157</b>
5.1	Conclusions	157
5.2	Perspectives	159
	<b>Annexes</b>	<b>167</b>

# Acronymes

<b>ACP</b>	Analyse en Composantes Principales
<b>AD</b>	Analyse Discriminante
<b>ARMA</b>	Auto Regressive Moving Average (modèles autorégressifs et moyenne mobile)
<b>BG</b>	Bond Graph
<b>CFN</b>	Conditions de Fonctionnement Normal
<b>CHA</b>	Classification Hiérarchique Ascendante
<b>DM</b>	Distance de Mahalanobis
<b>DWT</b>	Discrete Wavelet Transform (Transformée en ondelettes discrète)
<b>EKF</b>	Extended Kalman Filter (FK étendu)
<b>EOL</b>	End Of Life (Fin de vie)
<b>FAR</b>	False Alarm Rate (Taux de fausses alarmes)
<b>FDD</b>	Fault Detection and Diagnosis (Détection et Diagnostic de Défauts)
<b>FDR</b>	Fault Diagnosis Rate (Taux de diagnostic)
<b>FK</b>	Filtre de Kalman
<b>FP</b>	Filtre à Particules
<b>FS</b>	Fault Signature (Signature du défaut)
<b>FSM</b>	Fault Signature Matrix (Matrice de signature de défauts)
<b>HI</b>	Health Indicator (Indicateur de santé)
<b>HS</b>	Health State (Etat de santé)
<b>MAE</b>	Mean Absolute Error (Erreur absolue moyenne)
<b>MAPE</b>	Mean Absolute Percentage Error (Erreur absolue moyenne en pourcentage )
<b>MSE</b>	Mean Square Error (Erreur quadratique moyenne)
<b>PG</b>	Processus de Gauss
<b>PSL</b>	Projection en Structures Latentes
<b>QTA</b>	Qualitative Trend Analysis (Analyse qualitative de tendances)
<b>RB</b>	Réseau Bayésien
<b>RBD</b>	Réseau Bayésien Dynamique
<b>RN</b>	Réseau de Neurones
<b>RR</b>	Recognition Rate (taux de bonne classification)
<b>RRA</b>	Relation de Redondance Analytique
<b>RtFD</b>	Run to Failure Data (données de passage à l'échec)
<b>RUL</b>	Remaining Useful Life (Durée de vie résiduelle)
<b>SDG</b>	Signed Directed Graph (Graphe orienté signé)
<b>SVR</b>	Support Vector Regression (Machine à vecteurs de support)
<b>TEP</b>	Tennessee Eastman Process





# Introduction générale

## 1.1 Contexte

De nos jours, les systèmes sont devenus de plus en plus complexes. Ceci s'explique par la modernisation croissante, la forte compétitivité ainsi que les exigences de performances associées, notamment les exigences de sûreté de fonctionnement (fiabilité, disponibilité, maintenabilité et sécurité) qui ne cessent de croître, devenant de véritables enjeux. Tout cela induit la nécessité de développer de nouvelles approches de surveillance permettant de rendre compte de l'état du système à chaque instant.

De nombreux travaux [Isermann, 2006, Travé-Massuyès, 2014b, Ding, 2008b, Ding et al., 2011, Chiang et al., 2000a, Basseville et al., 1993, Medjaher, 2014, Si et al., 2011, Zio, 2012, Baraldi et al., 2013a, Lei et al., 2018] ont abordé le domaine de la surveillance, témoignant ainsi de l'intérêt croissant manifesté par la communauté scientifique et les industriels par rapport à cette problématique.

La surveillance est définie comme l'ensemble d'actions permettant de déterminer l'état d'un système, en détectant, en diagnostiquant et en prédisant des défauts qui peuvent survenir lors de son fonctionnement. Ces défauts peuvent non seulement dégrader les performances du système mais aussi avoir de graves conséquences sur la sécurité des personnes ainsi que de lourds impacts sur l'environnement.

Généralement, on distingue deux approches de surveillance : surveillance à base de modèles [Ding, 2008b, Isermann, 2005a, Baraldi et al., 2013a], et surveillance à base de données [Ding et al., 2011, Joe Qin, 2003, Si et al., 2011, Baraldi et al., 2013a]. Chaque approche présente ses propres avantages et inconvénients.

Ces dernières années, on s'est intéressé de plus en plus à l'intégration de méthodes et à la fusion de décisions. En effet, il paraît judicieux de développer des approches hybrides exploitant les méthodes à base de modèle et de données, par l'élaboration d'un cadre commun permettant la fusion des différentes approches de surveillance [Ding et al., 2009, Venkatasubramanian et al., 2003a, Travé-Massuyès, 2014a].

Cependant, malgré les avantages évidents que la fusion d'approches peut offrir pour la détection, le diagnostic et le pronostic des défauts, son potentiel reste encore faiblement exploité en raison de plusieurs obstacles. Tout d'abord, il faut noter le manque de généralité des approches proposées dans la littérature qui se traduit par l'absence d'un cadre unifié. Ensuite, la fusion nécessite de relier différents concepts et suppose donc d'avoir des connaissances solides relatives à plusieurs théories. Le choix des méthodes à fusionner est également très important puisqu'un des plus grands défis dans la conception de la fusion d'approches consiste à profiter des avantages de chaque méthode tout en évitant leurs inconvénients. Enfin, aucune approche de fusion ne permet d'assurer, durant sa phase de conception, une amélioration des performances globales en fonction des performances individuelles des méthodes qu'elle fusionne, ce qui représente un verrou scientifique majeur pour l'intégration de méthodes et la fusion de décisions.

Notre travail s'inscrit dans ce contexte particulier de détection, diagnostic et pronostic de défauts affectant les systèmes complexes. L'objectif principal est de proposer une approche générique d'intégration de diverses méthodes de surveillance, avec une méthodologie basée sur la fusion de décisions, dont la performance serait meilleure que celles des méthodes individuelles la composant.

## 1.2 Contributions

Les principales contributions de ce travail sont résumées comme suit :

- Développement d'une nouvelle démarche méthodologique de fusion de décisions, basée sur la théorie Bayésienne. Cette démarche s'appuie sur un calcul des paramètres d'un Réseau Bayésien (RB) en fonction des objectifs de performance à atteindre. Le développement conduit à un problème multi-objectif sous contraintes, résolu par une approche lexicographique. La démarche proposée se veut générique puisqu'elle ne considère dans son développement aucune contrainte sur les méthodes à intégrer ni sur les applications à considérer.
- Adaptation de cette méthodologie au diagnostic. La première étape est une étape de conception hors-ligne qui permet de définir les objectifs de performance à respecter afin d'améliorer les performances globales du système en diagnostic, puis de déduire de façon théorique les paramètres du RB permettant de respecter ces objectifs. La deuxième étape se déroule en ligne et permet de fusionner les décisions et d'évaluer les performances de cette fusion.
- Application de l'approche proposée à un système complexe de grande taille : le Tennessee Eastman Process (TEP) afin d'évaluer l'efficacité de la méthodologie. Tout d'abord, un modèle Bond Graph (BG) est développé à partir des connaissances physiques du système puis utilisé à des fins de diagnostic. Les résultats fournis par cette méthode à base de modèle sont fusionnés, par le biais de l'approche proposée, avec les résultats d'une méthode à base de données afin de fournir la décision globale.
- Développement d'une approche de pronostic reposant sur le principe de classification. Une première étape conduite hors-ligne permet d'apprendre les liens existants entre les signaux disponibles et les états de dégradations correspondants tandis qu'une seconde étape vise à identifier l'état de santé courant du système et à prédire sa RUL à tout instant.

- Adaptation de la méthodologie de fusion au pronostic. L'étape de conception hors-ligne permet de fixer les objectifs de performance à atteindre dans le cadre du pronostic, puis de proposer un RB paramétré respectant ces objectifs. L'étape en ligne renvoie la durée de fonctionnement avant la défaillance du système et évalue les performances de la fusion de décisions.
- Application de l'approche proposée de pronostic à un ensemble de turboréacteurs à double flux. La première étape consiste à développer deux méthodes de pronostic à partir des données disponibles du système. La fusion des décisions résultant de ces deux méthodes fournit le résultat global correspondant à une estimation de la durée de fonctionnement restante avant la défaillance du système.

**Valorisation du travail de recherche :** Les développements réalisés et les résultats obtenus ont fait l'objet de publications dans des revues internationales ainsi que des présentations dans des conférences internationales et nationales.

#### Articles dans des revues répertoriées (3)

- [R1] : **Tidriri, K.**, Chatti, N., Verron, S., Tiplica, T. (2016). Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring : A review of researches and future challenges. *Annual Reviews in Control*, 42, 63 – 81. (IF = 2.627)
- [R2] : **Tidriri, K.**, Tiplica, T., Chatti, N., Verron, S. (2018). A Generic Framework for Decision Fusion in Fault Detection and Diagnosis. *Journal of Engineering Applications of Artificial Intelligence*, 71, 73 – 86. (IF = 2.894)
- [R3] : **Tidriri, K.**, Chatti, N., Verron, S., Tiplica, T. (2018). Model-based Fault Detection and Diagnosis of Complex Chemical Processes : A case study of the Tennessee Eastman Process. *Proceedings of the Institution of Mechanical Engineers, Part I : Journal of Systems and Control Engineering*. (IF = 1.420)

#### Conférences internationales avec actes et comités de lecture (4)

- [C1] : **Tidriri, K.**, Chatti, N., Verron, S., Tiplica, T. A Bond Graph Modeling for Health Monitoring and Diagnosis of the Tennessee Eastman Process. In : *IEEE International Conference on Control, Decision and Information Technologies*. 05/04/2017, Barcelone p. 6. Spain.
- [C2] : **Tidriri, K.**, Chatti, N., Verron, S., Tiplica, T. A New Hybrid Approach for Fault Detection and Diagnosis. In : *IFAC 2017 World Congress*. 12/07/2017, Toulouse. France
- [C3] : **Tidriri, K.**, Chatti, N., Verron, S., Tiplica, T. Generic Framework for Hybrid Fault Diagnosis and Health Monitoring of the Tennessee Eastman Process. In : *International Conference on Control, Automation and Systems*. 18/10/2017, Jeju. South Korea
- [C4] : **Tidriri, K.**, Tiplica, T. Chatti, N., Verron, S. A New Multi-Objective Decision-Making Approach Applied to the Tennessee Eastman Process. In : *10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS 2018*. 31/08/2018, Warsaw. Poland

### Conférences nationales

- [C5] : **Tidiri, K.**, Chatti, N., Verron, S., Tiplica, T. Intégration de modèles quantitatifs et qualitatifs pour le diagnostic de systèmes dynamiques. *17èmes Journées Doctorales*. 04/05/2017, Nantes, France.
- [C6] : **Tidiri, K.**, Chatti, N., Verron, S., Tiplica, T. Fusion de décisions dédiée à la surveillance des systèmes complexes. *Réunion du GT S3 (Sûreté-Surveillance-Supervision)* . 29/03/2018, Paris, France.

## 1.3 Organisation

Le rapport est structuré principalement autour de trois chapitres :

Dans le chapitre 2, les différentes approches de surveillance proposées dans la littérature jusqu'à ce jour sont passées en revue. Le principe pour chaque approche ainsi que les travaux correspondants sont exposés. Ceci permet de mettre en exergue les avantages, les inconvénients et les limites des approches de surveillance existantes.

Dans le chapitre 3, on propose une nouvelle approche de fusion de décisions pour le diagnostic des systèmes complexes. Tout d'abord, le problème de fusion de décisions est formulé. Ensuite, la démarche méthodologique est détaillée et le cadre générique de fusion est introduit. L'approche est adaptée pour le diagnostic de défaillances. Enfin, l'application et l'évaluation de cette approche sur un système complexe de grande taille, le TEP, permet de montrer son efficacité ainsi que sa capacité à assurer une amélioration significative des performances en diagnostic.

Enfin, dans le chapitre 4, on propose une démarche méthodologique de pronostic basée sur le principe de classification, qui permet de récupérer à tout instant le temps restant avant la défaillance. Ensuite, on adapte notre approche de fusion de décisions pour le pronostic des systèmes complexes. L'efficacité de l'approche proposée est évaluée à travers l'application sur un ensemble de turboréacteurs à double flux. L'objectif est de prédire le temps restant avant la prochaine défaillance du système en utilisant les méthodologies proposées de pronostic et de fusion de décisions.

---

# Surveillance des systèmes complexes- Etat de l'art

## Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>11</b>
<b>2.2</b>	<b>Approches pour la surveillance</b>	<b>13</b>
2.2.1	Détection et diagnostic à base de modèle	13
2.2.2	Détection et diagnostic à base de données historiques	17
2.2.3	Pronostic à base de modèle	23
2.2.4	Pronostic à base de données	25
2.2.5	Discussion	27
<b>2.3</b>	<b>Fusion d'approches pour la surveillance</b>	<b>29</b>
2.3.1	Fusion de méthodes	30
2.3.2	Fusion de décisions	36
2.3.3	Discussion	46
<b>2.4</b>	<b>Positionnement scientifique</b>	<b>49</b>
<b>2.5</b>	<b>Conclusion</b>	<b>50</b>

---

## 2.1 Introduction

Le niveau de complexité croissant des systèmes et les exigences de performances et de sûreté de fonctionnement associées à ces systèmes ont induit la nécessité de développer de nouvelles approches de surveillance plus performantes.

**Définition 1.** *Un système est un ensemble d'éléments, de composants, dispositifs, etc., cohérent vis à vis d'un objectif. Les éléments interagissent entre eux et avec l'environnement.*

Un système est qualifié de complexe quand il est composé d'un grand nombre d'entités qui interagissent de manière non-linéaire, et ayant une dynamique forte.

Ensuite, on expose la notion de surveillance de la manière suivante :

**Définition 2.** *La surveillance est une action visant à observer et analyser l'état d'une entité à chaque instant. Elle repose sur trois volets distincts mais complémentaires : la détection, le diagnostic et le pronostic.*

**Définition 3.** *La détection consiste à déterminer toute anomalie dans le fonctionnement d'un système.*

**Définition 4.** *Le diagnostic consiste à déterminer les causes pouvant être à l'origine d'un dysfonctionnement.*

**Définition 5.** *Le pronostic consiste à prédire la durée de vie résiduelle (appelée RUL : Remaining Useful Life) ou la fin de vie (appelée EOL : End Of Life). La RUL d'un système correspond au temps restant avant la défaillance tandis que le EOL représente l'instant à partir duquel le système n'est plus opérationnel.*

Un système fonctionne «normalement» lorsqu'il est capable de fournir les fonctions pour lesquelles il a été conçu. On dit qu'il est en Conditions de Fonctionnement Normal (CFN). L'apparition d'un défaut peut générer une défaillance.

**Définition 6.** *Un défaut est un écart par rapport à la valeur nominale d'une caractéristique, cet écart dépassant des limites d'acceptabilité.*

En général, les défauts affectant un système physique peuvent être classés en trois catégories, à savoir les défauts d'actionneurs (changement de paramètres par exemple), les défauts de capteurs (biais par exemple) et les changements structurels au niveau du processus (une fuite dans un réservoir par exemple). Ces défauts peuvent conduire à des défaillances critiques s'ils ne sont pas prédits ou du moins détectés et diagnostiqués rapidement. La défaillance a été définie par Isermann [Isermann, 2006] de la manière suivante :

**Définition 7.** *Une défaillance est l'altération ou la cessation de l'aptitude d'un ensemble à accomplir sa ou ses fonctions requises avec les performances définies dans les spécifications techniques.*

Souvent, on distingue la défaillance partielle et la défaillance complète. La première correspond à une dégradation.

**Définition 8.** *Une dégradation représente une perte de performances d'une des fonctions assurées par un équipement.*

Si le système ne peut plus réaliser les fonctions pour lesquelles il a été conçu, on parle de défaillance complète et on dit que le système est défaillant.

Par ailleurs, le choix d'une méthode de surveillance repose essentiellement sur les informations disponibles sur le système : retours d'expérience représentés par les connaissances des experts concernant le système, données historiques enregistrées suite au fonctionnement du système dans différents états, modèles physiques développés à partir de la compréhension fondamentale de la physique du système et exprimés en termes de relations fonctionnelles mathématiques, etc.

Classiquement, on distingue deux catégories d'approches de surveillance :

- les approches à base de **modèles** : observateur [Luenberger, 1964, Luenberger, 1966], Filtre de Kalman (FK) [Kalman, 1960], Filtre à Particules (FP) [Del Moral, 1996], espaces de parité [Chow, 1980], estimation de paramètres [Isermann, 1984], Bond Graph (BG) [Paynter, 1961], Graphe orienté signé (SDG) [Iri et al., 1979], Graphe causal temporel (TCG) [Mosterman and Biswas, 1999], etc.
- les approches à base de **données** : Analyse en Composantes Principales (ACP) [Pearson, 1901, Hotelling, 1933], Réseaux Bayésiens (RB) [Pearl, 1988], Réseaux de Neurones (RN) [Rosenblatt, 1958], Projections en Structures Latentes (PSL) [Wold, 1975], Analyse discriminante (AD) [Friedman, 1989], Analyse qualitative des tendances (QTA) [Cheung and Stephanopoulos, 1990, Janusz and Venkatasubramanian, 1991], k plus proches voisins (kNN) [Cover and Hart, 1967], Arbre de décisions [Breiman et al., 1984], Décomposition en ondelettes [Mallat, 1989], séries temporelles [Alwan and Roberts, 1988], processus de Gauss (PG) [Seeger, 2004], machine à vecteurs de support (SVR) [Boser et al., 1992], logique floue [Zadeh, 1965], etc.

[Venkatasubramanian et al., 2003b, Tidriri et al., 2016, Baraldi et al., 2013a, Venkatasubramanian et al., 2003c, Venkatasubramanian et al., 2003a] ont passé en revue les différentes approches de surveillance en établissant une étude comparative de quelques unes d'entre elles. Ils ont mis en exergue le fait qu'il y ait un intérêt certain à fusionner différentes approches émanant d'une même catégorie mais aussi et surtout plusieurs approches provenant de différentes catégories. Dans ce sens, plusieurs travaux de recherche ont été récemment menés pour établir une stratégie d'aide à la prise de décision en combinant différentes approches de surveillance.

Dans ce chapitre, les caractéristiques des méthodes issues respectivement des approches basées sur les modèles et les données historiques sont étudiées séparément. On mettra en évidence les avantages et les limites de chaque approche grâce à une analyse comparative et on soulignera l'intérêt des approches hybrides. Ensuite, les travaux existants qui ont tenté de combiner différentes méthodes de surveillance seront présentés. Enfin, une discussion permettra d'identifier les verrous scientifiques et de positionner l'apport méthodologique développé dans ce travail par rapport à l'existant.

## 2.2 Approches pour la surveillance

Dans cette partie, on passe en revue les principales approches pour la surveillance des systèmes complexes, en mettant l'accent sur le principe, le cadre méthodologique et l'apport de chaque méthode. On insistera davantage sur les approches qui seront exploitées dans le cadre de nos contributions.

### 2.2.1 Détection et diagnostic à base de modèle

#### Principe

De nombreuses approches de détection et de diagnostic basées sur un modèle font généralement appel à des algorithmes de génération d'indicateurs de défauts, appelés résidus [Isermann, 2005b, Ding, 2008a]. Ces résidus représentent l'évaluation numérique des Relations de Redondances Analytiques (RRAs) qui traduisent les écarts entre le comportement



réel et celui prédit par un modèle. Une règle de décision est ensuite choisie pour évaluer ces résidus afin de détecter et isoler un défaut. Ce principe est illustré dans la Figure 2.1.

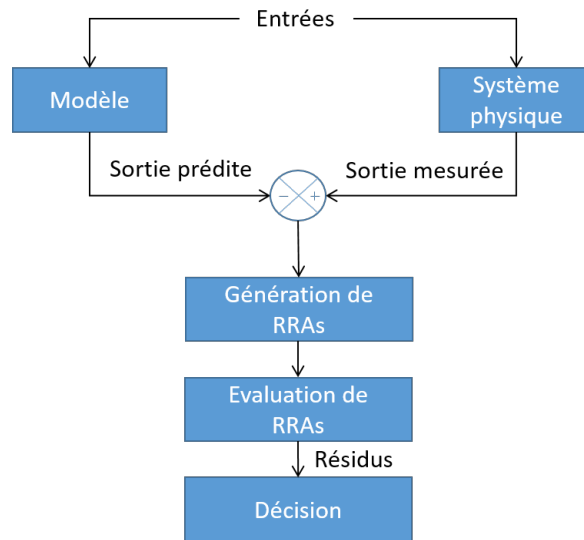


FIGURE 2.1 – Principe de la détection et du diagnostic à base de modèle.

La génération de RRA est une étape fondamentale car elle conditionne la qualité de la détection et du diagnostic et doit ainsi intégrer des considérations sur la robustesse. Elle peut être réalisée par diverses méthodes comme les observateurs, les espaces de parité, l'estimation de paramètres, les BG, etc.

En présence d'un défaut, les résidus diffèrent de zéro, ce qui permet de détecter, voire même d'identifier le défaut. En pratique, les bruits de capteurs, les non-linéarités, les perturbations, les variations de paramètres et les dynamiques non modélisées influent beaucoup sur le résidu et il est donc peu réaliste de considérer que le résidu diffère de zéro uniquement en présence de défaut.

Très souvent, les informations de sensibilité et de robustesse des résidus par rapport aux défauts potentiels sont décrites sous la forme d'une matrice de signatures de défauts (FSM pour Fault Signature Matrix). Les colonnes de cette matrice représentent l'ensemble des résidus tandis que ses lignes représentent les défauts. Cette matrice évalue donc les possibilités du système de détecter et d'identifier des défauts, et ses éléments (en phase d'exploitation) sont déterminés comme suit :

$$S_{ij} = \begin{cases} 1 & \text{si le résidu } i \text{ est sensible au défaut } j \\ 0 & \text{sinon} \end{cases}$$

## Méthodes

**Observateur [Luenberger, 1964]** Un observateur est un filtre dynamique qui permet d'estimer les variables d'état d'un système à partir des grandeurs accessibles telles que les entrées et les mesures. Pour des fins de détection et de diagnostic, l'observateur estime les sorties du système physique, au lieu de son état caché. Les écarts entre ces sorties estimées et les sorties mesurées représentent les résidus.

Suite à de nombreux travaux [Luenberger, 1966, Patton and Chen, 1997, Sotomayor and Odloak, 2005, Liu et al., 2015], les observateurs sont devenus robustes aux incertitudes et aux perturbations, mais en retour, leur sensibilité aux défauts de faible amplitude a diminué. Pour contourner cette limitation, Bastin et Gevers ont proposé d'utiliser des observateurs adaptatifs [Bastin and Gevers, 1988]. Cette solution

est applicable non seulement aux systèmes linéaires, mais aussi à certaines classes de systèmes non-linéaires [Caccavale et al., 2009]. Néanmoins, de nombreux systèmes physiques ne se prêtent pas à une modélisation par le biais de ces classes de systèmes non-linéaires. Ceci représente un inconvénient majeur quant à l'utilisation des approches à base d'observateurs.

**Filtre de Kalman (FK) [Kalman, 1960]** Contrairement à l'observateur déterministe présenté ci-dessus, le FK permet de traiter les problèmes linéaires stochastiques. Il repose sur l'inférence Bayésienne, dans laquelle les données mesurées, qui peuvent être incomplètes ou bruitées, sont utilisées pour estimer et mettre à jour les paramètres inconnus (variables d'états). Comme son nom l'indique, le FK est basé sur le filtrage qui met à jour les paramètres de manière récursive en prenant une mesure à la fois.

Une grande variété de FK a été développée. On peut citer le FK étendu (EKF pour Extended Kalman Filter) [Ljung, 1979] qui traite les problèmes non-linéaires en utilisant les jacobiniennes, et le FK sans parfum (UKF pour Unscented Kalman Filter) [Julier and Uhlmann, 1996] qui permet de pallier le manque de robustesse pressenti du EKF face aux non linéarités par la création de la transformée "unscented".

**Espace de parité [Chow, 1980]** L'espace de parité utilise le modèle du système ainsi que les techniques de l'algèbre linéaire, comme les projections matricielles, pour la génération de RRAs. Ces RRAs lient entre elles uniquement les variables connues ou mesurées et sont obtenues en éliminant les variables d'état non observables sur un horizon fini [Gertler, 1997].

La détection et le diagnostic s'effectuent en suivant l'évolution du vecteur de parité, qui représente une image du désaccord entre les mesures [Chow and Willsky, 1984]. En CFN, le vecteur de parité est de faible amplitude (en raison des bruits aléatoires). En présence d'un défaut, l'amplitude du vecteur augmente et s'oriente suivant la "direction de défaut" associée au capteur défaillant.

L'approche de l'espace de parité est généralement limitée à des systèmes linéaires invariants. De plus, il convient de noter qu'il existe un compromis entre la performance et le coût de calcul dans la conception d'un générateur de résidus basé sur des relations de parité [Hwang et al., 2010, Ye et al., 2004].

**Estimation de paramètres [Isermann, 1984]** Parallèlement au développement des espaces de parité et des observateurs, des efforts ont été déployés pour résoudre le problème de détection et de diagnostic des défauts à partir de la surveillance de l'évolution des paramètres structuraux du système. En effet, les grandeurs non mesurables telles que les paramètres peuvent être estimées à partir de la connaissance des entrées et des sorties mesurables.

Cette approche a été largement étudiée dans le domaine de l'identification des paramètres du système pour le contrôle et la commande, puis appliquée à la détection et au diagnostic des défauts, comme l'a montré Isermann [Isermann, 2006].

Ainsi, un défaut est détecté si les valeurs estimées en ligne sont différentes des valeurs nominales [Isermann, 1993]. L'estimation des paramètres peut être très simple lorsque ces derniers sont directement liés aux coefficients physiques (tels que la résistance, les charges, etc.) [Dai and Gao, 2013]. Cependant, cette technique nécessite d'avoir un système physique excité en permanence afin d'estimer en continu ses paramètres.

**Grphe orienté signé (SDG) [Iri et al., 1979]** Le SDG est un graphe orienté signé basé sur un modèle qualitatif. Les sommets représentent des variables tandis que les arcs

traduisent le signe des influences entre variables. Le SDG a prouvé son efficacité pour le diagnostic des défauts simples et multiples dans les systèmes dynamiques [Iri et al., 1979, Tateno et al., 2006]. Néanmoins, le développement de modèles basés sur le SDG n'est pas une tâche facile en pratique puisqu'il requiert la connaissance précise du système via des schémas détaillés d'instrumentation ou la présence d'un modèle analytique décrivant le comportement du système.

**Grphe causal temporel (TCG) [Mosterman and Biswas, 1999]** Le TCG est un graphe orienté labelé basé, tout comme le SDG, sur un modèle qualitatif. Les sommets représentent les variables du système tandis que les arcs décrivent les relations temporelles et causales entre les différentes variables du système. Le diagnostic à base de TCG part du constat d'occurrence d'un défaut à travers une variable mesurée et tente de déterminer toutes les variables à l'origine de ce défaut. Son utilisation a été étendue aux systèmes hybrides [Daigle, 2008]. Cependant, comme pour le SDG, la construction du TCG repose sur les équations analytiques du système, ce qui représente un inconvénient majeur en pratique. Différents travaux ont essayé de simplifier la construction du TCG, notamment en le déduisant du modèle BG comme dans [Daigle, 2008].

**Bond Graph (BG) [Paynter, 1961]** Le BG est une approche énergétique multidisciplinaire qui fournit une modélisation graphique des systèmes physiques de différents domaines (électrique, mécanique, hydraulique, thermodynamique, etc.). Elle a été introduite par Paynter [Paynter, 1961] afin d'illustrer, sous forme graphique, l'échange de puissance entre les éléments d'un système physique. Le BG est composé de sommets et de demi-flèches appelées "bond", comme illustré dans la Figure 2.2.

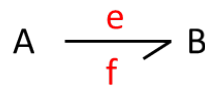


FIGURE 2.2 – Bond Graph.

La direction de la puissance entre les éléments est fixée par les liens de puissance ou bond. Deux variables génériques de puissance sont associées à chaque liaison : un effort  $e$  et un flux  $f$  tel que  $e \times f = Puissance$ . Deux autres variables généralisées peuvent être définies et sont appelées respectivement impulsion généralisée  $p$  et déplacement  $q$ .

Le BG fournit une compréhension physique du système et une description graphique unifiée de son comportement dynamique. En raison de ses propriétés structurelles et causales, il peut être exploité à des fins différentes, telles que le placement des capteurs, l'identification et l'estimation des paramètres du système, mais aussi à des fins de surveillance. En effet, la modélisation BG est particulièrement adaptée à la détection et au diagnostic car elle intègre l'information temporelle et causale requise pour détecter et isoler des défauts. La procédure de génération des résidus repose sur le parcours des chemins causaux [Samantaray and Bouamama, 2008] et est mise en oeuvre dans des logiciels dédiés [Bouamama et al., 2005, Chatti et al., 2014]. Ainsi, le BG est utilisé pour générer des résidus de manière systématique et générique grâce au principe de causalité. Ce concept important de causalité représente l'avantage de l'outil BG par rapport aux autres méthodes graphiques puisqu'il permet le développement de modèles mathématiques dynamiques, en précisant dans quel ordre les variables inconnues doivent être calculées à partir des variables connues, en structurant les équations et en évitant d'éventuelles erreurs

de modélisation.

Un trait causal placé perpendiculairement au bond exprime graphiquement cette causalité et est placé près de l'élément pour lequel l'effort est connu et loin de l'élément pour lequel le flux est connu. Chaque type d'élément présente une propriété causale spécifique. De plus amples détails sur la modélisation BG et des discussions sur la causalité peuvent être trouvés dans les références suivantes [Chatti et al., 2014, Bouamama et al., 2014, Samantaray and Bouamama, 2008].

Récemment, le modèle BG a été étendu en prenant en compte les incertitudes sur les paramètres par la génération de seuils robustes et adaptatifs pour l'étape d'évaluation des résidus. [Djeziri et al., 2007] ont proposé une approche de modélisation BG avec la configuration LFT (Linear Fractional Transformation) permettant de prendre en compte les incertitudes de paramètres. Cependant, cette approche ne considère pas les incertitudes de mesures. Dans [Chatti et al., 2014], un diagnostic qualitatif et quantitatif sur la base de la représentation BG est développé. En effet, les auteurs ont proposé un nouveau formalisme appelé Signed Bond Graph (SBG), construit directement à partir du BG. Ce formalisme incorpore des caractéristiques qualitatives et exploite en plus les propriétés structurelles quantitatives du BG afin de permettre la détection et le diagnostic de défauts multiples.

## Synthèse

Les méthodes de détection-diagnostic à base de modèle fournissent une description du comportement dynamique et une bonne compréhension physique du système, ce qui représente un avantage majeur [Chatti et al., 2014]. Cependant, en pratique, il est très difficile de développer un modèle mathématique précis qui tient compte de différentes sources d'incertitudes en raison de la complexité des systèmes [Tidiri et al., 2016].

En plus du challenge concernant le compromis précision/robustesse, plusieurs autres problématiques restent à résoudre comme l'application des méthodes à base de modèle aux systèmes non-linéaires. Pour pallier cette problématique, une attention particulière a été accordée aux approches de détection de défauts pour les systèmes linéaires à paramètres variants LPV (Linear Parameters Varying) puisqu'elles représentent une alternative attrayante aux dynamiques non linéaires [López-Estrada et al., 2016, Montes de Oca et al., 2014].

De nombreux efforts ont également été faits pour traiter la détection et le diagnostic des défaillances dans les systèmes décentralisés et distribués [Keliris et al., 2015, Indra et al., 2013] et dans les systèmes hybrides [Pons et al., 2015, de Mortain et al., 2015, Chanthery et al., 2015].

Enfin, il faut souligner que l'approche à base de modèles est généralement appliquée sous l'hypothèse de l'apparition de défauts simples uniquement.

Malgré les avantages apportés par les méthodes de détection-diagnostic à base de modèle, il convient de préciser que lorsqu'un modèle n'est pas disponible mais qu'un accès à une quantité importante de données historiques est possible, les approches à base de données représentent une bonne alternative.

## 2.2.2 Détection et diagnostic à base de données historiques

### Principe

De nos jours, les systèmes sont de plus en plus instrumentés, ce qui offre la possibilité de récupérer une quantité importante de données historiques qu'il convient d'exploiter pour

la détection et le diagnostic. Il paraît donc naturel de surveiller le processus en utilisant des méthodes basées sur ces données.

La plupart des méthodes à base de données historiques considèrent la détection et le diagnostic comme des tâches de classification. En effet, la détection vise à identifier si un fonctionnement anormal est survenu, ce qui correspond à une classification en deux catégories : la classe de  $CFN$  et la classe de défaut. Quant au diagnostic, il vise à déterminer le type de défaut, ce qui peut être considéré comme une classification en plusieurs classes : défaut 1, défaut 2, etc. Le principe est illustré dans la Figure 2.3.

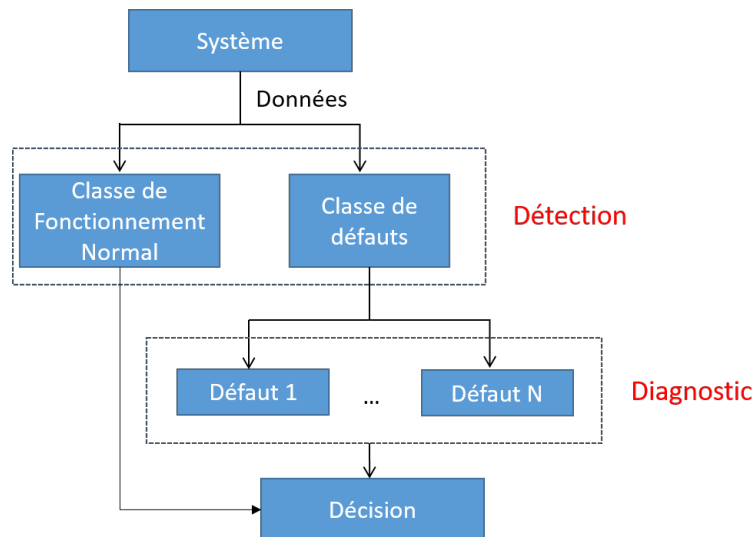


FIGURE 2.3 – Principe de la détection et du diagnostic par données.

Cette classification peut être supervisée ou non-supervisée. Dans la classification supervisée, on détermine les frontières de séparation entre les classes en se basant sur la connaissance de la composition des classes. Cependant, il est difficile d’anticiper a priori toutes les possibilités de défauts, d’où l’intérêt pour la classification non-supervisée. Cette dernière ne requiert aucune information concernant la composition des classes, mais utilise plutôt les données disponibles pour obtenir les frontières de séparation des classes selon des principes comme l’homogénéité des classes ou bien la distance aux différents voisins.

## Méthodes

**Réseau Bayésien (RB) [Pearl, 1988]** Un RB est une représentation graphique acyclique dirigée qui s’appuie sur un modèle probabiliste. Le RB permet de visualiser les relations entre variables. Ces dernières, modélisées par un noeud sur le graphe, représentent les connaissances du réseau et peuvent prendre des valeurs dans un ensemble discret ou continu. Les noeuds sont reliés entre eux par des arcs dirigés qui décrivent un lien de dépendance directe. Un exemple de RB est donné dans la Figure 2.4. Dans cet exemple,  $Y$  (qui est un noeud enfant) dépend directement de  $X$  (qui est un noeud parent), d’où la présence d’un arc allant de la variable  $X$  à la variable  $Y$ .

Le RB permet également de décrire de façon quantitative le fonctionnement du système par le calcul des probabilités des variables [Pearl, 2014]. En effet, chaque noeud du réseau possède une table de probabilités conditionnelles qui permet de quantifier l’effet du (des) noeud(s) parents sur ce noeud : elle décrit les probabilités associées aux noeuds enfants suivant les différentes valeurs des noeuds parents. Pour les noeuds racines (sans parents), on fixe des probabilités a priori concernant les valeurs de la variable.

L’inférence, qui représente le calcul des probabilités a posteriori des variables non observées étant donnée(s) la ou les nouvelle(s) information(s) disponible(s) (appelées

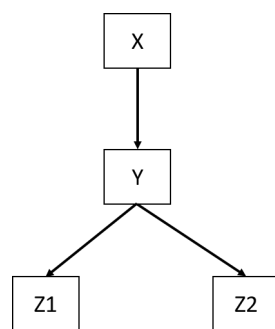


FIGURE 2.4 – Exemple de Réseau Bayésien.

évidences), permet de mettre à jour le réseau. Cette évidence peut être certaine (sûre à 100%) ou bien vraisemblable (par exemple sûre à 80%). Une fois l'information indiquée, celle-ci est propagée dans le réseau par le moteur d'inférence. L'inférence permet également de calculer la distribution d'une ou de plusieurs variables : c'est ce que l'on appelle la marginalisation.

Selon la description des noeuds utilisés (noeud discret ou continu, loi utilisée), on définit plusieurs types de RB : les RB discrets, gaussiens, gaussiens conditionnels, etc.

Une extension intéressante des RB est représentée par la classe de RB *dynamiques* (RBD) qui intègrent la notion de temps (une variable peut influencer sa propre valeur à l'instant suivant) [Murphy, 2002]. Ces RBD ont été largement appliqués dans des applications réelles comme par exemple la détection des défauts transitoires [Jha et al., 2009], l'identification du chemin de propagation de défauts [Yu and Rashid, 2013], etc. Une autre extension intéressante des RB, qui connaît ces dernières années un développement très fort, est représentée par les RB *orientés objet* qui sont de puissants outils de modélisation de la connaissance, très adaptés pour des systèmes complexes. Par analogie à la programmation orientée objet, les RB orientés objet permettent de simplifier la représentation graphique du réseau dans le sens où certaines parties sont regroupées en un seul objet nommé instance. Enfin, une dernière extension des RB que nous mentionnons dans ce mémoire est le diagramme d'influence qui s'obtient en ajoutant des noeuds d'*utilité* et de *décision*. Les noeuds de décisions représentent le choix d'une décision prise par le décideur. Les noeuds d'utilité représentent l'objectif à maximiser.

Dans le contexte des techniques de classification supervisée, les classifieurs Bayésiens ont fait l'objet d'une attention particulière et leurs performances ont été analysées. Les trois structures les plus utilisées ont toutes un point commun : la présence d'un noeud discret racine nommé "noeud de classe". Ce noeud discret possède  $k$  modalités, avec  $k$  le nombre de classes du problème. Voici un bref aperçu de ces différentes structures :

- *RB Naïf*

Le RB naïf repose sur l'hypothèse que les variables observées, représentées par les noeuds enfants du noeud de classe, sont conditionnellement indépendantes étant donné ce noeud racine discret, comme représenté sur la Figure 2.5. Si cette hypothèse est valide et si les probabilités marginales et conditionnelles sont estimées avec assez de précision, le classifieur Bayésien naïf est optimal [Yang and Webb, 2009]. Néanmoins, la présence de corrélations peut réduire l'efficacité du classifieur Bayésien naïf car le poids de l'information redondante tend à minimiser l'influence des autres caractéristiques, conduisant ainsi à des prédictions biaisées [Langley and Sage, 1994]. Cependant, il a été observé que même dans le cas où l'hypothèse d'indépendance est



violée, ce classifieur donne de bons résultats par rapport à d'autres classifieurs plus sophistiqués. De plus, il est peu sensible au bruit, ce qui fait de lui un classifieur très performant pour le diagnostic des processus complexes.

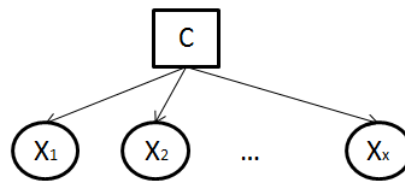


FIGURE 2.5 – RB classifieur Naïf.

- *RB Naïf Augmenté ou TAN (Tree Augmented Naïve Bayes)*

Le RB naïf augmenté repose sur l'hypothèse que les variables observées, représentées par les noeuds enfants, sont liées sous forme d'un arbre, c'est-à-dire qu'on ajoute un arc entre les variables qui partagent une information mutuelle importante. Cette structure a été proposée dans le but d'alléger l'hypothèse très forte d'indépendance des variables vue précédemment, et permet donc d'améliorer les performances du RB naïf. Les relations entre variables sont établies à partir de l'expertise et des données disponibles.

- *RB Naïf Semi-Condensé*

Le RB naïf semi-condensé a été proposé pour prendre en compte la corrélation entre les variables, en introduisant une variable jointe regroupant plusieurs variables corrélées. La matrice de variance-covariance du noeud joint permet de définir les corrélations entre variables.

Le RB a été appliqué avec succès dans divers domaines, comme par exemple le soutien à la décision clinique [Sesen et al., 2013, Antal et al., 2003, Wang et al., 2014], l'analyse des données de génotype [Yan and Cercone, 2010], la prédiction du comportement d'insectes [Pawson et al., 2017], la prévision du flux de trafic [Sun et al., 2006], la détection et le diagnostic de défauts [Zhao et al., 2015, Atoui et al., 2015b, Zhao et al., 2013, Verron et al., 2010c, Verron et al., 2008a, Hu et al., 2018, Lin et al., 2018, Cai et al., 2016]. Dans ce contexte, [Verron et al., 2010c] ont proposé un RB permettant la détection de défauts ainsi que l'identification de l'ensemble des variables à l'origine de ce défaut. Pour cela, les auteurs se sont basés sur le travail de [Li et al., 2008] proposant une décomposition causale de la statistique  $T^2$ , ainsi que sur les résultats de [Verron, 2007], notamment sur la modélisation de cartes de contrôle multivariées sous forme de RB. Une généralisation des statistiques quadratiques sous RB a été ensuite proposée par [Atoui et al., 2015b]. Cette généralisation permet la réalisation simultanée de la détection et du diagnostic sous RB tout en prenant en compte la notion de rejet de distance.

En particulier, le RB est un outil qui permet la représentation des connaissances en prenant en compte différentes sources d'incertitudes [Gaymard and Tiplica, 2014, Weidl et al., 2005, Lu et al., 2009, Correa et al., 2009].

Outre sa capacité à raisonner avec des informations incertaines, le RB peut être complété par des données historiques et des jugements d'experts [Zhao et al., 2013]. Il permet également de modéliser des méthodes comme l'ACP et les cartes de contrôle multivariées afin d'utiliser un seul outil pour la détection et le diagnostic des défauts [Verron et al., 2010a].

Il convient de préciser que la structure du RB dépend de la connaissance a priori du système et nécessite des données pour l'apprentissage et pour la détermination des probabilités a priori et des tables de probabilités conditionnelles. Quand des jugements experts sont disponibles, ils peuvent également être utilisés pour l'apprentissage de la structure et des paramètres [Flores et al., 2011, Masegosa and Moral, 2013].

Par ailleurs, l'efficacité du RB dépend des différentes hypothèses sous lesquelles il est développé. Les RN représentent une bonne alternative car ils peuvent s'adapter aux données sans aucune hypothèse sur la distribution.

**Réseau de Neurones (RN)** [Rosenblatt, 1958] Un RN est défini comme étant un ensemble «*d'automates connectés en réseau, fonctionnant en parallèle, dans lequel les connexions ont la connaissance du domaine*» [Zwingelstein, 1995]. Ces automates sont des neurones artificiels qui permettent de modéliser une quantité considérable de données suivant les poids synaptiques et la fonction d'activation qu'ils renferment.

Le problème de classification dans un RN est réduit à l'estimation des poids synaptiques, appris en exploitant le décalage entre les sorties désirées et réelles. Le mécanisme d'apprentissage peut être supervisé ou non-supervisé.

Au cours des dernières années, les RN ont été utilisés avec succès dans le domaine de la reconnaissance des formes, même en présence de données bruitées ou incomplètes, et pour les problèmes de prédiction et de classification [Zhang, 2000]. Parmi les réseaux les plus communs utilisés pour la classification, on peut citer les réseaux de perceptrons monocouche et multicouche [Rosenblatt, 1958, Sorsa et al., 1991].

Il convient de noter que les performances du RN sont caractérisées par la structure du réseau, l'algorithme d'apprentissage et la fonction d'activation. De plus, le RN est rarement utilisé seul pour la détection et le diagnostic. En effet, une phase de prétraitement des données, qui consiste à réduire la dimension du problème et à sélectionner un ensemble de variables, est souvent utilisée en amont, grâce à des méthodes comme l'ACP.

**Analyse en Composantes Principales (ACP)** [Pearson, 1901, Hotelling, 1933] L'ACP est une méthode d'analyse statistique multivariée qui réduit la dimension des données de sorte que les informations essentielles soient conservées. Cette méthode a été largement utilisée dans la surveillance des processus, et notamment dans le contrôle en temps réel des processus continus [Lu et al., 2003, Zhou et al., 2014]. Elle permet de transformer des variables corrélées en nouvelles variables décorrélées appelées «composantes principales». Ceci se fait par la recherche d'axes géométriques orthogonaux où la variabilité des données est maximale.

Dans le cas des systèmes non-linéaires, il existe des solutions comme l'ACP basée sur les RN [Kramer, 1992], ou l'ACP à noyau [Phillpotts, 2007]. Concernant la surveillance des systèmes à temps variant, des algorithmes récursifs ou adaptatifs ont été développés [Zhang et al., 2012] ainsi que des extensions de l'ACP comme Moving-ACP [Liu et al., 2009], Multi-block et Multi-scale ACP [Lee et al., 2005], WARP (weighted adaptive recursive PCA) [Portnoy et al., 2016], just-in-time learning PCA (JITL-PCA) [Cheng and Chiu, 2005], etc.

D'autre part, plusieurs travaux de recherche ont été menés pour améliorer le diagnostic de défauts en s'appuyant sur l'approche de l'ACP. Parmi ces travaux, on peut citer les cartes de contribution et les approches multi-blocs [Kourti and MacGregor, 1996] ou encore Sparse ACP et shrinking ACP [Xie et al., 2013].

**Projection en Structures Latentes (PSL)** [Wold, 1975] Outre l'ACP, la PSL est une autre approche statistique utilisée à des fins de détection et de diagnostic de défaut.



Cette approche est basée sur la projection de l'information contenue dans les espaces de données de haute dimension sur des espaces de dimension faible, définie par un certain nombre de variables latentes. A l'instar de l'ACP, des extensions de l'approche PSL ont été proposées dans la littérature. On peut citer par exemple Multi-bloc et Multi-way PSL [Kourti et al., 1995] qui ont été développées pour faciliter la procédure de diagnostic.

**Analyse Discriminante (AD) [Friedman, 1989]** L'AD est une technique statistique de classification supervisée qui se base sur la règle de Bayes pour prédire la classe d'une nouvelle observation [McLachlan, 2004]. Cette approche repose le plus souvent sur l'hypothèse d'une distribution normale. Les séparations entre classes réalisées grâce à l'AD peuvent être soit quadratiques, soit linéaires. L'AD quadratique nécessite une grande quantité de données pour estimer ses différents paramètres. Quant à l'AD linéaire, elle est très souvent utilisée pour sa simplicité et sa robustesse face aux fluctuations sur les hypothèses de normalité des classes.

**k plus proches voisins (kNN) [Cover and Hart, 1967]** La méthode des kNN est une technique de classification non-paramétrique, qui repose sur le calcul de distance d'une nouvelle observation à chaque observation présente dans l'ensemble d'apprentissage. Les k voisins les plus proches de cette observation sont sélectionnés et la classe la plus représentée parmi les k voisins est choisie comme classe d'appartenance de la nouvelle observation. Contrairement aux approches paramétriques, le kNN ne nécessite aucune estimation de paramètres. Néanmoins, il convient de fixer la valeur de "k", qui peut avoir une grande influence sur la performance de classification. De plus, le kNN peut s'avérer très coûteux en calcul et en mémoire de stockage puisqu'il entraîne, pour chaque nouvelle observation, le calcul des distances de cette observation à chaque observation présente dans l'ensemble d'apprentissage.

**Analyse qualitative des tendances (QTA) [Cheung and Stephanopoulos, 1990]** La QTA est une technique semi-quantitative pilotée par les données. Elle repose sur la représentation des signaux mesurés comme une séquence de formes (croissante, décroissante, constante, etc.) appelées primitives [Janusz and Venkatasubramanian, 1991]. Chaque séquence de primitives constitue une tendance. L'idée principale est que tout signal raisonnablement lisse peut être représenté comme une séquence de ces primitives. QTA a été utilisée pour la surveillance des processus et la détection et le diagnostic des défaillances [Maurya et al., 2005] grâce au diagnostic rapide et précis qu'elle offre. Cependant, l'augmentation de la complexité de calcul de QTA avec l'augmentation du nombre de capteurs utilisés pour le diagnostic peut interdire son application en temps réel pour les systèmes complexes.

**Arbre de décisions [Breiman et al., 1984]** Un arbre de décisions est un outil de classification et d'aide à la décision représentant un ensemble de choix sous la forme graphique d'une arborescence. Les différentes décisions possibles sont situées aux extrémités des branches, et sont atteintes en fonction de décisions prises à chaque étape. L'arbre de décision a l'avantage d'être lisible et rapide à exécuter, d'autant plus qu'il peut facilement être transformé en règles logiques compréhensibles. Par ailleurs, l'arbre s'agrandit en fonction du nombre de variables et du nombre de modalités pour chaque variable, rendant ainsi sa construction difficile et coûteuse [Cornuéjols and Miclet, 2011].

**Décomposition en ondelettes** [Mallat, 1989] La décomposition en ondelettes est un outil de traitement du signal qui a été utilisé depuis plus d'une décennie pour la surveillance des systèmes [Gao and Yan, 2010, Yan et al., 2014]. Elle repose sur une décomposition dans les bases d'ondelettes, avec l'avantage d'avoir une taille de la fenêtre d'analyse qui est variable. Cette décomposition conserve les informations temporelles et fréquentielles, fournissant ainsi une localisation temps-fréquence des différents défauts qui peuvent survenir. Par conséquent, cette méthode permet de mieux caractériser les défauts par leurs composantes fréquentielles et temporelles.

**Techniques de sous-espaces** [Van Overschee and De Moor, 2012] Le développement de modèles pour des systèmes complexes n'est pas toujours possible. En utilisant des données mesurées d'entrée/sortie, les méthodes d'identification de modèles basées sur les techniques de sous-espace permettent d'estimer le modèle sous l'hypothèse que le système est linéaire, invariant dans le temps, et de dimension finie.

## Synthèse

Les approches de détection-diagnostic basées sur les données sont adaptées pour les systèmes complexes à grande échelle et peuvent réduire le temps et le coût, puisqu'elles ne nécessitent pas le développement de modèles physiques [Alzghoul et al., 2014]. Grâce aux avancées réalisées dans le milieu informatique, ces méthodes se sont dotées d'algorithmes d'inférences de plus en plus rapides, capables de gérer les données manquantes [Pearl, 2014] ou encore les non-linéarités [Zhang, 2000]. Ainsi, elles jouent un rôle important dans les systèmes de surveillance moderne.

Cependant, une étape de prétraitement est très souvent nécessaire pour extraire les informations importantes [Portnoy et al., 2016]. En effet, la plupart des méthodes à base de données voient leurs performances diminuer en présence de variables non-informatives qui contribuent à ajouter du bruit. De plus, leurs performances dépendent de la quantité et de la qualité des données historiques disponibles.

### 2.2.3 Pronostic à base de modèle

Tout comme pour la détection et le diagnostic, le choix d'une méthode de pronostic dépend particulièrement de l'information disponible sur le système considéré. Par la suite, nous allons présenter un état de l'art actuel sur les principales méthodes de pronostic basées sur les modèles et les données. Il faut noter que les méthodes pour le pronostic sont souvent similaires à celles pour la détection-diagnostic. Par conséquent, nous allons particulièrement mettre l'accent sur les méthodes qui n'ont pas été évoquées dans la partie précédente.

#### Principe

Les approches de pronostic basées sur un modèle s'appuient sur une représentation mathématique disponible du phénomène de dégradation. Ainsi, un modèle de comportement est construit à partir de la connaissance des principes et des mécanismes physiques de dégradation affectant le système considéré.

La dégradation est considérée comme une variable dont l'évolution dynamique est modélisée par un processus déterministe ou stochastique si l'on prend en compte les incertitudes sur les conditions initiales et les paramètres.

Plusieurs modèles de dégradation existent selon la discipline et le domaine considérés. On peut citer les modèles de vieillissement, de fatigue, de corrosion, d'érosion, etc.

L'objectif du pronostic à base de modèle consiste donc à prédire l'évolution future du phénomène physique de dégradation. Lorsque ce modèle atteint un seuil prédéfini, on en déduit que le composant ou système surveillé est entré dans un état défaillant. Le schéma illustratif est donné dans la Figure 2.6.

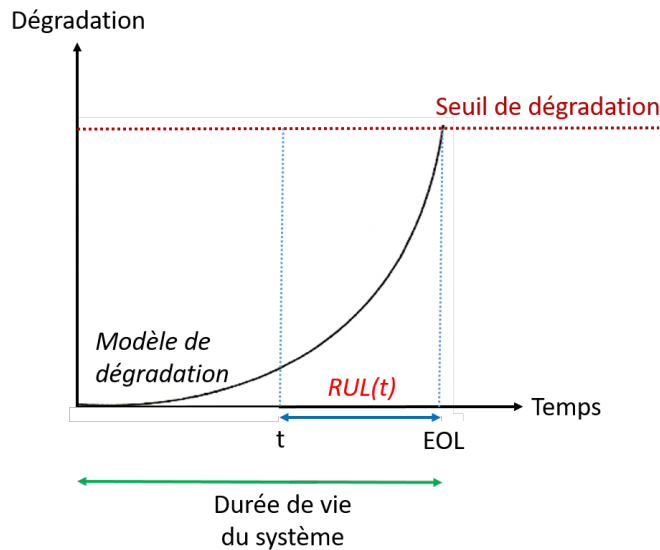


FIGURE 2.6 – Principe du pronostic à base de modèle.

## Méthodes

**Modèle physique** L'approche à base de modèle physique relie par une expression mathématique les dommages accumulés par la dégradation d'un composant aux défaillances qui peuvent survenir. Le modèle utilisé est généralement continu et intègre les contraintes environnementales.

Plusieurs lois ont été utilisées pour déterminer le nombre de cycles qui restent jusqu'à la défaillance d'un composant. On peut citer la loi de Miner [Byington et al., 2002], la loi de Coffin-Manson [Ghelam et al., 2006], ou encore la loi de Paris Erdogan [Qian et al., 2017].

**Filtre à Particules (FP) [Del Moral, 1996]** Le FP est une méthode de simulation séquentielle de type Monte-Carlo. Les particules explorent l'espace d'état de manière indépendante selon le système dynamique sous-jacent (le modèle a priori). Ensuite, lorsqu'une nouvelle mesure est disponible, l'adéquation de chaque particule avec cette mesure est évaluée grâce à la fonction de vraisemblance, afin d'assigner à chaque particule un poids proportionnel à cette valeur. Un mécanisme de sélection permet ensuite d'éliminer ou de multiplier les particules, selon la valeur de leurs poids. Une présentation détaillée du FP peut être trouvée dans [Gordon et al., 2001].

Le FP est l'algorithme d'estimation le plus utilisé pour le pronostic [Mechri et al., 2017]. Il prend en compte la non-linéarité des systèmes et les bruits non gaussiens qui peuvent l'affecter. En effet, le FK puis le EKF ont d'abord été utilisés pour prédire la RUL. Cependant, lorsque le système est fortement non-linéaire ou lorsque les bruits ne sont pas gaussiens, l'EKF est susceptible de diverger, d'où l'intérêt croissant pour les FP [Daigle and Goebel, 2011, Jha et al., 2016].

## Synthèse

L'approche de pronostic à base de modèle offre généralement des résultats précis puisqu'elle se base et intègre dans son raisonnement une connaissance du système à surveiller. Cette connaissance permet également de relier la dégradation observée aux paramètres qui pourraient en être la cause [Byington et al., 2002, Ghelam et al., 2006, Qian et al., 2017].

Cette approche fait l'hypothèse d'existence ou de possibilité de développement d'une représentation mathématique du phénomène de dégradation du système. La connaissance de l'évolution de la dégradation, en plus des données issues du module de diagnostic, permettent de déterminer l'état de dégradation courant et de prédire la durée de vie résiduelle. Un des avantages de cette approche est la flexibilité puisque le modèle peut être réajusté pour prendre en compte tout changement de propriétés du système ou de sa dégradation. De plus, il est possible d'adresser le problème du pronostic des systèmes hybrides [Prakash et al., 2018, Yu et al., 2014] et non-linéaires [Bressel et al., 2016]. Enfin, certaines méthodes comme le FP permettent de gérer efficacement les incertitudes qui résultent des mesures bruitées, du processus de dégradation ou encore des conditions environnementales [Jha et al., 2016].

Par ailleurs, comme pour le cas du diagnostic, il est très difficile d'obtenir le modèle dynamique analytique qui prend en compte le phénomène de dégradation. De plus, les données provenant de capteurs utilisées pour l'identification de paramètres pour des fins de pronostic sont souvent bruitées. Coppe et al. [Coppe et al., 2009] ont montré que l'incertitude paramétrique peut être progressivement réduite, en dépit du bruit et des biais dans les mesures des capteurs. Dans [Hall and Strutt, 2003], les incertitudes de paramètres et de modèles sont prises en compte dans la méthodologie de pronostic développée. En outre, les modèles pour le pronostic sont souvent dédiés à des mécanismes de dégradations spécifiques. Ceci implique, pour un même système complexe, le développement de plusieurs modèles pour les différents composants critiques. C'est ce que l'on qualifie de "pronostic de niveau composant". D'ailleurs, très peu de travaux s'intéressent au "pronostic de niveau système", ce qui représente un challenge à relever pour les prochaines années.

Pour toutes ces raisons, l'intérêt pour les approches pilotées par les données est justifié puisque le cadre applicatif des approches à base de modèle est restreint en pratique.

### 2.2.4 Pronostic à base de données

#### Principe

Les approches de pronostic basées sur les données reposent sur l'utilisation des données disponibles de surveillance. Pour beaucoup d'applications, les données mesurées constituent la principale source d'information pour une meilleure compréhension du comportement de dégradation du système. Les approches axées sur les données considèrent que les caractéristiques statistiques des données sont relativement inchangées à moins qu'un événement dysfonctionnel se produise dans le système.

A partir de ces données, des caractéristiques reflétant le comportement du système et de sa dégradation sont extraites : c'est ce que l'on appelle des indicateurs de santé. Ces derniers permettent d'identifier quantitativement ou qualitativement l'état de dégradation du système surveillé et sont exploités pour l'apprentissage de modèles de prédiction des états actuels et futurs du système. Un schéma résumant le principe de pronostic est décrit dans la Figure 2.7.

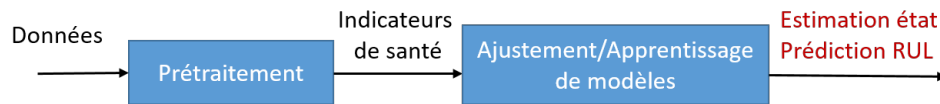


FIGURE 2.7 – Principe du pronostic à base de données.

L'objectif du pronostic à base de données consiste à extrapoler l'évolution des indicateurs de dégradation dans le but de suivre l'état de santé du système et de prédire les éventuelles défaillances qui pourraient survenir. Une estimation de la RUL peut être déduite en calculant le temps qui sépare l'état courant du système de l'état correspondant à sa défaillance.

## Méthodes

**Séries temporelles** [Alwan and Roberts, 1988] Une série temporelle est une suite de valeurs numériques décrivant l'évolution d'une quantité spécifique au cours du temps. L'objectif de l'analyse de séries temporelles est de comprendre l'évolution passée de la quantité étudiée et prédire son comportement futur.

Pour un système physique, la RUL peut être estimé à l'aide d'un modèle ARMA (Auto Regressive Moving Average) [Yan et al., 2004], ou bien par l'utilisation d'un modèle ARIMA (Auto Regressive Integrated Moving Average) [Jardim-Gonçalves et al., 1996]. Dans d'autres cas, un modèle AR peut suffire [Zhang et al., 1994].

Bien que possédant d'excellentes qualités de prédictions [Barraza-Barraza et al., 2017], les modèles ARIMA ne sont pas adaptés aux problèmes multivariables et non-linéaires.

**Processus de Gauss (PG)** [Seeger, 2004] Les techniques de régressions utilisées pour le pronostic des défaillances permettent de prédire la RUL et sa confiance associée sous une forme probabiliste. Le PG est une des méthodes de régression les plus populaire pour le pronostic [Mohanty et al., 2009, Brahim-Belhouari and Bermak, 2004, Liu et al., 2012, Aye and Heyns, 2017]. Le PG est une approche probabiliste puissante de modélisation de l'incertitude sur les données et de la réalisation de l'inférence Bayésienne. Cette approche se base sur la construction d'un modèle qui relie les entrées (les indicateurs de santé appris hors ligne par exemple) et la sortie (les valeurs de fin de vie correspondantes). Ce modèle est ensuite utilisé pour prédire la sortie (la valeur de fin de vie) en fonction d'une nouvelle entrée (indicateur de santé) [Brahim-Belhouari and Bermak, 2004, Liu et al., 2012].

**Séparateur à vaste marge (SVM)** [Boser et al., 1992] Les SVM font partie des techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression. L'objectif principal est de déterminer une ou plusieurs frontières de séparation entre les données.

Très souvent, la méthodologie de pronostic basée sur le SVM est la suivante : tout d'abord, les indicateurs de santé sont extraits des données. Ensuite, les paramètres du SVM sont déterminés et le modèle de régression est appris à partir des indicateurs disponibles. Enfin, le SVM est utilisé en ligne pour déterminer l'état de santé du composant. Les SVM peuvent obtenir des résultats très précis avec des entrées non linéaires [Soualhi et al., 2015, Benkedjough et al., 2013].

Plusieurs variantes de SVM ont été proposées dans la littérature, afin de pouvoir l'adapter à des systèmes dynamiques non-linéaires et non stationnaires. A cet égard, on peut citer le RVM (Relevance Vector Machine) [Hu and Tse, 2013], et le LSSVR (Least Squares Support Vector Regression) [Li et al., 2013].

**Apprentissage profond** [[Schmidhuber, 2015](#)] L'avantage de l'application du RN pour le pronostic réside dans le fait qu'il peut modéliser les systèmes non linéaires, complexes et multidimensionnels sans expertise préalable sur le comportement physique du système [[Li and Ray, 1995](#), [Shao and Nezu, 2000](#), [Liu et al., 2010](#)].

Récemment, la communauté de l'apprentissage s'intéresse de plus en plus à l'apprentissage profond qui s'avère être très efficace pour la reconnaissance de formes. Cet apprentissage est caractérisé par une architecture où plusieurs couches sont empilées dans le réseau.

Les abstractions de haut niveau qui peuvent caractériser les données sont bien modélisées à l'aide des structures profondes complexes, permettant ainsi une extraction de caractéristiques plus efficace par rapport aux RN classiques. Les méthodes d'apprentissage en profondeur ont suscité un grand intérêt et ont permis d'obtenir des résultats significatifs dans de nombreux domaines, et notamment dans l'estimation de la RUL [[Zhang et al., 2017](#), [Ren et al., 2017](#), [Li et al., 2018](#)].

**Logique floue** [[Zadeh, 1965](#)] Contrairement à la logique Booléenne, où une proposition peut prendre soit la valeur "0" soit la valeur "1", la logique floue introduit un degré de vérité que l'on peut choisir dans l'intervalle  $[0, 1]$ , permettant ainsi à la proposition d'être dans un état autre que 0 et 1. La logique floue a été très souvent intégrée dans les RN afin de capturer plus d'informations, d'assurer la convergence de l'apprentissage des paramètres et de gérer les incertitudes et les imprécisions auxquelles n'importe quel système de pronostic doit faire face [[Wang et al., 2004](#)].

## Synthèse

L'approche de pronostic à base de données est plus simple à mettre en oeuvre puisqu'elle ne requiert pas la connaissance physique et le développement mathématique des phénomènes de dégradation affectant le système. Ainsi, le cadre applicatif de cette approche est beaucoup plus large puisqu'elle convient à toutes les applications qui disposent d'une quantité suffisante de données historiques [[Zhang et al., 2017](#), [Specht, 1990](#)]. La principale force de cette approche réside dans sa capacité à apprendre et à capturer les relations importantes entre les données, sans aucune considération physique ou connaissance sur le système. De plus, elle permet de gérer les incertitudes et les non-linéarités [[Khawaja et al., 2005](#)].

Par ailleurs, cette approche est moins précise que celle basée sur les modèles et dépend de la qualité et de la quantité de données que contient l'échantillon d'apprentissage. Ce dernier doit contenir différents scénarios de dégradation.

### 2.2.5 Discussion

L'absence d'une méthode générique de détection, de diagnostic et de pronostic, adaptée à toutes les exigences d'un système de surveillance est la principale conclusion de cette étude bibliographique et ce malgré de nombreuses études publiées dans la littérature. En effet, chaque approche possède ses avantages et inconvénients.

Les approches basées sur les données reposent sur une grande quantité d'informations et se concentrent sur l'analyse des signaux de sortie du système. Elles peuvent gérer des variables corrélées de grande dimension [[Venkatasubramanian et al., 2003b](#)]. Ainsi, elles sont adaptées à des systèmes complexes à grande échelle et ne nécessitent pas l'élaboration de modèles physiques [[Alzghoul et al., 2014](#)].

Par ailleurs, l'étape de prétraitement est généralement nécessaire, que ce soit pour le



pronostic, la détection ou le diagnostic, afin d'extraire les informations utiles avec un temps de calcul parfois élevé. En outre, selon [Venkatasubramanian et al., 2003b], le diagnostic de défauts peut être difficile avec les méthodes basées sur les données puisqu'elles ne possèdent pas de «signature» de défauts. Enfin, la performance de ces approches dépend des données d'apprentissage et peut se dégrader si le système fonctionne dans un état inconnu ou avec des défauts ou des dégradations inconnus.

Contrairement aux approches à base de données, les approches basées sur des modèles nécessitent une quantité moins importante de données historiques et reposent sur un modèle mathématique explicite, développé à partir des lois de la physique ou identifié grâce à des techniques d'estimation. À mesure que la complexité du système augmente, l'estimation devient plus difficile. Ainsi, une étape de validation du modèle est nécessaire pour vérifier la cohérence des hypothèses et des approximations par rapport au comportement du système réel. De plus, ces approches ne peuvent gérer que des systèmes linéaires et des classes très spécifiques de non-linéarités. Selon [Chiang et al., 2000a], les approches à base de modèle ne conviennent pas aux systèmes complexes à grande échelle et leur performance dépend de la qualité et de la précision du modèle. D'ailleurs, la plupart des applications de surveillance dans l'industrie reposent sur des approches pilotées par les données car elles sont capables de détecter plus rapidement les situations défaillantes, sont plus faciles à mettre en oeuvre et nécessitent moins de connaissances a priori. Cependant, quand un modèle précis et complet est disponible, ces approches deviennent plus pertinentes que celles basées sur les données puisqu'elles intègrent la compréhension physique du système et des phénomènes de dégradation qui l'affectent [Chiang et al., 2000a].

Notons cependant que la distinction entre catégories d'approches (à base de données/modèles) n'est pas toujours évidente. En effet, la classification choisie dans ce mémoire, comme dans beaucoup d'autres travaux, doit être considérée avec précaution. A titre d'exemple, le FP est considéré ici comme appartenant à la catégorie d'approche de pronostic basée sur le modèle puisqu'il vise à identifier les paramètres et l'état associés à un modèle de dégradation. Dans [Si et al., 2011], toutefois, il apparaît dans la catégorie d'approches à base de données puisque les méthodes de filtrage sont considérés comme guidés par les données. Plusieurs autres exemples peuvent être donnés dans ce sens. Enfin, il est important de remarquer que l'étape de détection-diagnostic située en amont conditionne généralement les performances du pronostic par sa capacité à fournir une estimation fiable et précise de l'état de santé courant du système.

Il apparaît donc que le principal critère pour choisir une approche est la disponibilité des données historiques ou la possibilité de développement de modèles précis décrivant le système et sa dégradation.

Récemment, il y a eu un intérêt croissant pour l'intégration de différentes stratégies de surveillance de défauts. La principale motivation est qu'aucune méthode individuelle n'est capable de satisfaire toutes les exigences d'un système de surveillance précis. Ainsi, on peut donner comme exemple la méthode d'estimation des paramètres qui a été intégrée aux observateurs et à l'espace de parité [Dai and Gao, 2013]. En effet, les schémas basés sur l'estimation d'état sont efficaces pour la détection de défauts car ils sont intrinsèquement rapides [Patton and Chen, 1997] par rapport à la méthode d'estimation de paramètres, d'où l'intérêt de la combinaison. Des techniques statistiques comme l'ACP et le PSL ont été combinées avec des méthodes supervisées comme les RN [Ghate and Dudul, 2011] ou avec l'analyse temps-fréquence basée sur les ondelettes

[Lu et al., 2003] pour extraire des caractéristiques statistiques. Plusieurs systèmes de classifieurs ont également été combinés pour améliorer la performance de classification [Ghosh et al., 2011], etc.

Dans le prochain paragraphe, la fusion d'approches pour la surveillance est détaillée.

## 2.3 Fusion d'approches pour la surveillance

Selon plusieurs études [Ding et al., 2011, Ding et al., 2009, Tidriri et al., 2016, Venkatasubramanian et al., 2003b], l'intégration de plusieurs approches pour la surveillance des systèmes complexes présente de nombreux intérêts :

- Comblent les faiblesses d'une méthode de surveillance par les avantages d'une autre méthode afin de parvenir à une meilleure performance.
- Utiliser plusieurs sources d'information quand celles-ci sont disponibles, telles que les retours d'expériences, les historiques de maintenance et de réparation, les données historiques, les modèles mathématiques, les connaissances d'experts, etc.
- Bénéficier des progrès réalisés au sein de chaque communauté scientifique.

Classiquement, on distingue deux façons de fusionner des approches de surveillance, à savoir la **Fusion de méthodes** et la **Fusion de décisions**, comme illustré par la Figure 2.8 :

- **Fusion de méthodes** : Cette catégorie, que l'on peut qualifier de "*Fusion en série*", peut être réalisée de deux manières :
  - soit en utilisant différentes méthodes séparément pour les différentes tâches de surveillance, à savoir une méthode pour la détection, une pour le diagnostic et une autre pour le pronostic par exemple (Figure 2.8 a)). Ainsi, les méthodes sont utilisées l'une à la suite de l'autre pour réaliser les différentes étapes de surveillance. Celle-ci est décomposée en un certain nombre de sous-tâches. Chaque méthode est concernée par la recherche d'une solution pour une sous-tâche et contribue ainsi à la réalisation de la surveillance. Cette catégorie fait penser à une "*juxtaposition de méthodes*" et sera donc appelée ainsi. Cette juxtaposition de méthodes peut également se faire au niveau de la même tâche, c'est-à-dire que plusieurs méthodes sont utilisées pour le pronostic par exemple, mais une pour l'estimation de l'état courant et l'autre pour la prédiction de la RUL.
  - soit en combinant les caractéristiques de différentes méthodes afin de fournir une nouvelle méthode plus performante (Figure 2.8 b)). Cette seconde catégorie sera nommée "*intégration de caractéristiques*".
- **Fusion de décisions** : Cette catégorie, qualifiée de "*Fusion en parallèle*", consiste à utiliser conjointement plusieurs méthodes pour la même tâche (détection, diagnostic ou pronostic) puis à combiner leurs sorties pour reconstruire la sortie d'ensemble globale (Figure 2.8 c)). En effet, les décisions générées par les méthodes sont de même nature, ce qui peut se résumer à une fusion de décisions données par les méthodes individuelles. Les méthodes de surveillance, dans ce cas, sont capables de fonctionner de façon indépendante et simultanée, en fournissant une solution à la même tâche.



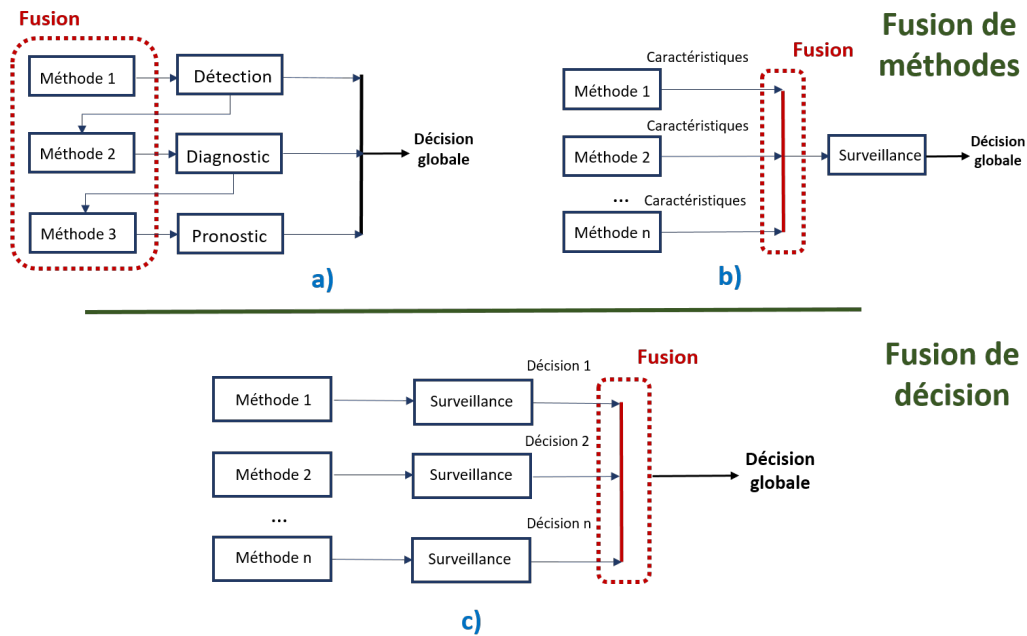


FIGURE 2.8 – Fusion d'approches pour la surveillance.

Bien évidemment, des travaux utilisant conjointement plusieurs méthodes pour l'étape de détection par exemple (ce qui correspond à une fusion de décisions) suivi d'une autre méthode pour le diagnostic (ce qui correspond alors à une fusion de méthodes) peuvent être trouvés dans la littérature. Par conséquent, cette classification doit être considérée également avec précaution.

Dans ce qui suit, un examen des travaux traitant de la fusion de méthodes de surveillance est effectué. L'accent est porté sur les approches dites "hybrides", c'est-à-dire qui fusionnent une approche à base de modèle avec une autre orientée données. Dans ce dernier contexte, il n'y a eu que peu de travaux publiés dans la littérature.

### 2.3.1 Fusion de méthodes

#### Juxtaposition de méthodes

La juxtaposition de méthodes repose sur l'utilisation en série de plusieurs méthodes pour accomplir la tâche de surveillance. Dans [Zhang and Hoo, 2011], un schéma est proposé pour gérer le problème de détection et de diagnostic de défauts par une stratégie hiérarchique, consistant à réduire successivement l'espace de recherche des défauts possibles. Cette stratégie est définie par les quatre étapes qui suivent :

1. Le système est décomposé en fonction des phénomènes physiques impliqués dans la modélisation BG (thermodynamique, chimique, mécanique, etc).
2. L'ACP est appliquée pour réduire la dimension des données et éliminer les valeurs aberrantes.
3. Une transformée en ondelettes discrète (DWT) est appliquée aux scores obtenus par l'ACP pour fournir une décomposition à plusieurs échelles de temps.
4. Enfin, la distance de Mahalanobis (DM) est déterminée et le niveau de confiance est calculé à partir de la distribution de  $\chi^2$ .

Si l'occurrence d'un défaut est détectée (c'est-à-dire s'il y a une violation de la DM à n'importe quelle échelle), le BG et le RB sont utilisés pour localiser le domaine de la cause. En effet, une comparaison est faite entre les échelles de temps auxquelles s'est produite la violation du niveau de confiance nominal et les constantes de temps nominales associées à chaque domaine physique déterminées par le BG. Par la suite, un RBD utilise la dépendance conditionnelle entre les signatures de défauts et les domaines de défaillances pour trouver l'origine possible du défaut.

Cette approche a été appliquée sur un réacteur à cuve à agitation continue et sur une station d'épuration des eaux usées. Cependant, étant donné que l'ACP est utilisée en premier pour réduire la dimension des données, les méthodes qui suivent à l'étape 3 et 4 ne sont appliquées qu'aux résultats obtenus. Ainsi, une erreur qui se produit dans l'espace résiduel peut ne pas être détectée, ce qui représente une limite majeure.

Une autre approche hybride "en série" qui utilise dans un premier temps le SDG pour réduire l'espace de recherche des défauts possibles et ensuite QTA pour identifier le défaut réel est proposé dans [Maurya et al., 2007].

Tout d'abord, le SDG réduit l'ensemble des défauts candidats (appelé défauts du niveau 1 (cf Figure 2.9)), en comparant la réponse mesurée avec celle prédite par le SDG. Ensuite, QTA est appliquée pour améliorer la résolution du diagnostic en classant les défauts du niveau 1. Même si les deux méthodes sont utilisées pour la même tâche, ici le diagnostic, on peut remarquer qu'elles sont appliquées "en série".

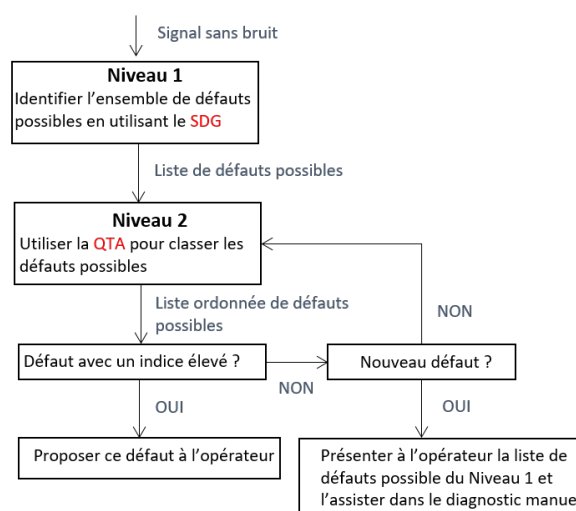


FIGURE 2.9 – Le cadre proposé combinant SDG et QTA pour le diagnostic des défauts.

Les auteurs ont prouvé que la résolution du diagnostic est améliorée tandis que la complexité de calcul est réduite. Cette approche a également abordé le problème de non-détection et de fausses alarmes en raison du bruit et des incertitudes de modélisation ainsi que la question de diagnostic en présence de nouveaux défauts. Cependant, un modèle mathématique est nécessaire pour développer le SDG comme mentionné dans la partie 2.2.1, ce qui limite l'utilisation de cette méthode, surtout dans le cas des systèmes complexes.

Afin d'éviter le développement d'un modèle précis pour les systèmes complexes à grande échelle, plusieurs contributions ont suggéré l'utilisation des techniques d'identification de modèle [Van Overschee and De Moor, 2012, Ding et al., 2009, Sotomayor and Odloak, 2005]. Ces travaux ont proposé un cadre pour surveiller un grand nombre de variables hautement corrélées et pour gérer la tâche de modélisation pour les

systèmes à grande échelle, mais ils n'ont pu traiter que des défauts simples de capteurs et d'actionneurs.

Dans [Schubert et al., 2011], une approche unifiée visant la détection et le diagnostic des défauts de capteurs, d'actionneurs et de processus a été développée. En effet, les auteurs ont proposé une méthodologie qui repose sur les techniques de sous-espaces pour identifier le modèle dynamique puis sur les techniques de reconstruction et d'estimation pour générer des résidus. Une statistique ( $T^2$  ou SPE) est ensuite appliquée pour détecter un comportement anormal. Enfin, les résidus structurels permettent le diagnostic du défaut. Ce cadre est décrit par la Figure 2.10.

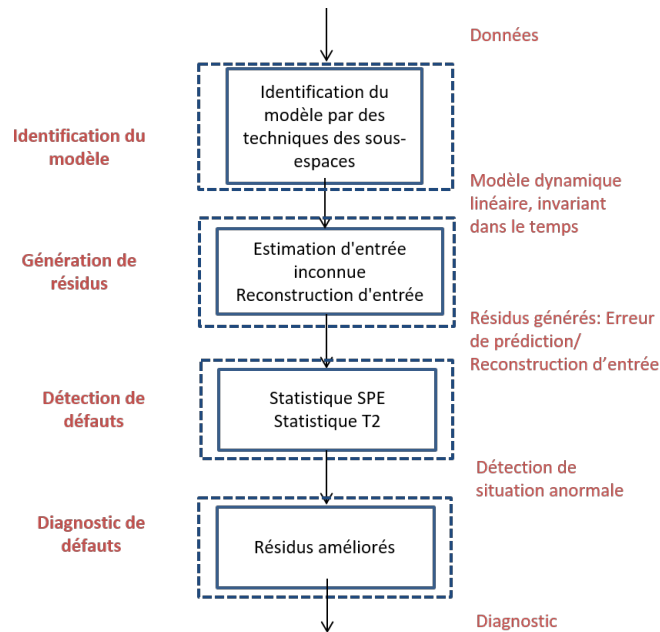


FIGURE 2.10 – Cadre pour l'identification de modèle, la génération de résidus, la détection de défaut et le diagnostic.

Cette approche, conçue pour détecter les défauts simples, a été appliquée à trois procédés industriels. L'utilisation des techniques de sous-espaces et des statistiques permet de simplifier la tâche de surveillance et d'éviter le problème lié aux prérequis de modélisation. Par ailleurs, le système est considéré comme linéaire, invariant dans le temps et de dimension finie. Ceci implique que l'approche ne peut être utilisée que pour certaines applications en pratique.

Or, la plupart des systèmes complexes sont caractérisés par leur non-linéarité et il convient donc de la prendre en compte dans le processus de surveillance. C'est ce qui a été proposé dans [Benkouider et al., 2009] où les auteurs se sont concentrés sur la détection des défauts dans les réacteurs semi-discontinus non-linéaires en utilisant des méthodes statistiques et l'EKF.

Ultérieurement, ils ont proposé une stratégie hybride basée sur l'EKF et le RN [Benkouider et al., 2012]. L'EKF utilise le modèle du réacteur comme prédicteur et permet la détection en ligne des défauts en estimant le coefficient de transfert thermique du réacteur. L'étape de diagnostic est effectuée à l'aide du RN. Ce dernier identifie le type et l'emplacement du défaut détecté. L'approche a été évaluée pour deux types de réactions exothermiques et a montré de bons résultats. Cependant, un modèle précis du réacteur est requis pour l'algorithme EKF.

La prise en compte de la non-linéarité est également abordée pour le pronostic des systèmes. Dans [Daroogheh et al., 2015], une nouvelle structure hybride reposant sur un FP et un RN est proposée. Le mécanisme de dégradation qui affecte le système est considéré connu dans cette approche, et les incertitudes qui l'entachent sont prises en compte à travers les paramètres inconnus qui doivent être estimés en ligne. Par conséquent, deux étapes principales sont mises en œuvre dans cette approche : (a) estimation jointe de l'état / paramètre et (b) prédiction de la distribution état / paramètre. La première étape est assurée grâce à un FP, tandis qu'un schéma de prédiction d'observations est développé à partir du RN pour l'étape de prédiction.

Trois types différents de RN, à savoir le perceptron multicouche, les RNs récurrents et les réseaux à base d'ondelettes sont utilisés à des fins de prévision d'observation. Une méthodologie pour mettre à jour les RN avec de nouvelles observations disponibles est également fournie. Les auteurs déduisent que le cadre hybride développé est robuste par rapport au choix du RN utilisé dans le module de prédiction. Cette approche hybride prend donc en compte les non-linéarités du système mais également les incertitudes.

Ainsi, il apparaît que la juxtaposition de méthodes permet très souvent d'améliorer les performances de surveillance. Par ailleurs, les hypothèses émises généralement (linéarité, stationnarité, etc.) limitent l'application des approches proposées. De plus, comme les méthodes sont utilisées les unes à la suite des autres, les erreurs faites par chaque méthode s'accumulent au fur et à mesure de la réalisation de la tâche de surveillance, contrairement à l'intégration de caractéristiques ou encore à la fusion de décisions.

### **Intégration de caractéristiques**

L'intégration de caractéristiques de plusieurs méthodes a pour objectif d'améliorer les performances de la méthode qui en découle ou de faciliter son application.

Certaines recherches ont porté leur attention sur la combinaison de la facilité de déploiement des méthodes statistiques avec les puissantes capacités de diagnostic des méthodes basées sur la redondance analytique. En effet, [Gertler et al., 1999] ont montré qu'il existe une équivalence entre l'ACP et les relations de parité, et ont prouvé strictement cette dualité. L'idée derrière ce travail est d'identifier une forme primaire de générateur de résidus à partir des données, en transférant à l'ACP certains concepts de redondance analytique comme la capacité de diagnostic des défauts.

Ainsi, l'article présente comment les résidus structurés avec de bonnes propriétés de diagnostic peuvent être obtenus par l'ACP. La performance de cette étude est évaluée sur le TEP pour les systèmes linéaires statiques. L'extension aux systèmes dynamiques et non linéaires n'est pas discutée en détail dans cet article.

Le travail dans [Ding et al., 2009] a également porté sur l'identification d'une forme primaire de résidus à partir de données, en utilisant l'espace de parité ou encore les observateurs. L'objectif principal consiste à prouver que les résidus d'observateur et d'espace de parité peuvent être obtenus directement à partir des données, sans utiliser le modèle du procédé. Cette approche s'est basée sur les travaux de Wang et Qin [Wang and Qin, 2002] pour modéliser des systèmes linéaires invariants dans le temps.

Les approches décrites ci-dessus bénéficient des capacités de diagnostic des approches basées sur les modèles tout en surmontant la tâche difficile du développement de représentations mathématiques. En effet, aucune connaissance particulière sur la théorie des systèmes n'est nécessaire pour générer des résidus structurés, ce qui constitue un avantage pratique significatif. Par ailleurs, les modèles identifiés sont linéaires et généralement stationnaires. De plus, l'identification s'accompagne d'incertitudes sur le

modèle et de dynamiques non modélisées.

Cette même limite relative à la linéarité des systèmes est retrouvée dans [Siswantoro et al., 2016]. En effet, les auteurs ont proposé un modèle linéaire, dont les paramètres sont estimés par un FK, qui fait tendre la sortie prédite par le RN vers la sortie souhaitée. En conséquence, l'erreur du RN est réduite et la précision de la classification est améliorée. Cette approche est illustrée dans la Figure 2.11, et validée empiriquement sur un ensemble de données provenant de l'UCI, qui gère une base de données au service de la communauté de l'apprentissage automatique.

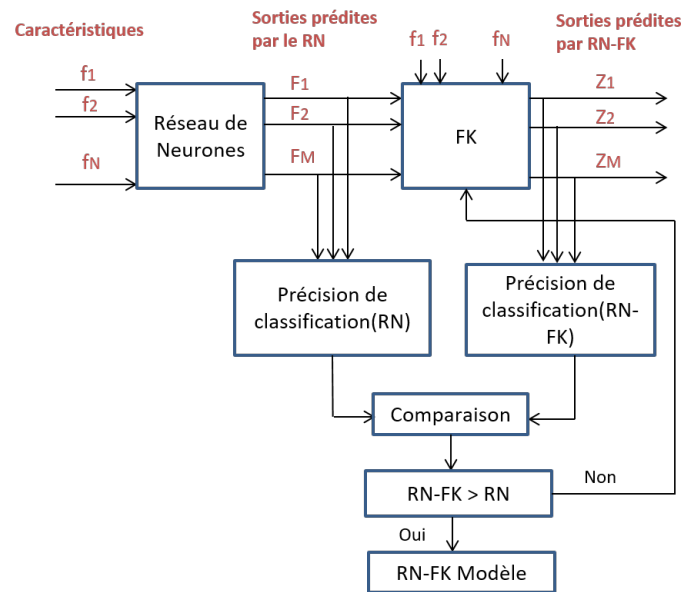


FIGURE 2.11 – Combinaison du RN et du FK.

Même si cette combinaison de méthodes améliore les résultats de classification, le système est supposé linéaire et stationnaire et le modèle dépend des entrées et des sorties prédites. Ainsi, s'il n'existe pas de relation linéaire entre ces variables, cette méthode ne peut pas être envisagée. Mais il faut noter que cette démarche peut être appliquée à d'autres classificateurs et que le FK peut être remplacé par un EKF par exemple pour traiter les systèmes non-linéaires.

En plus du potentiel d'amélioration de performances présenté par les approches hybrides, elles permettent également d'exploiter plusieurs sources d'information quand elles sont disponibles. Ceci a entraîné un intérêt considérable pour le développement de cadres unifiés et génériques permettant l'intégration de différentes méthodes ou de diverses sources de connaissances.

Dans [Weber et al., 2007], les auteurs ont intégré des données de fiabilité dans un RBD. Cependant, cette méthode considère uniquement les composants avec des fonctions de distribution de probabilité exponentielles pour simplifier les tâches de calcul.

Dans ce contexte, Zaidi et al. [Zaidi et al., 2010] ont proposé de combiner la méthode BG pour la génération de résidus avec le RB pour l'intégration de données.

Les informations concernant la fiabilité de chaque composant sont utilisées pour améliorer l'étape de prise de décision, permettant ainsi de gérer des signatures de défauts inconnues ou identiques. Les résidus et la fiabilité des composants sont intégrés dans un RB hybride (à noeuds continus et discrets) qui calcule les probabilités a posteriori des défauts des composants. La partie discrète du RB est constituée des résidus et des défaillances des

composants. Cette structure dépend des signatures des défauts : si un résidu  $r_i$  est sensible au défaut du composant  $C_j$ , alors un arc va relier  $F_j$  à  $r_i$ , comme l'illustre la Figure 2.12. La partie continue représente la fiabilité associée à chaque composant et donne l'information préalable sur la défaillance du composant.

Cette approche a été simulée sur un système à deux réservoirs et les résultats ont été jugés satisfaisants.

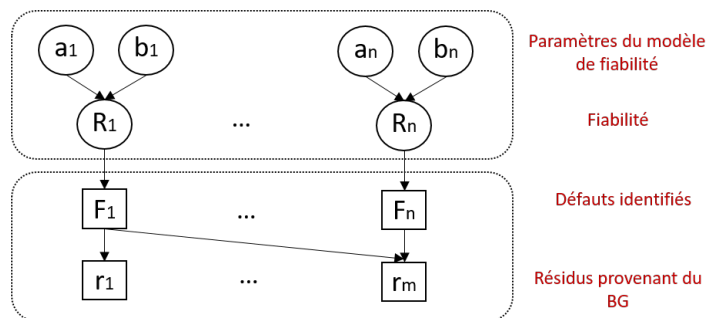


FIGURE 2.12 – RB proposé pour intégrer la fiabilité et les résidus générés par BG.

Toujours dans le contexte d'intégration du BG avec le RB, [Medjaher et al., 2009] ont proposé une procédure de génération d'un modèle de pronostic pour un système hydromécanique. Le BG fournit un modèle dynamique du système, et le RB, grâce à sa capacité d'inférence, permet de prendre en compte l'incertitude et d'effectuer le diagnostic et le pronostic. La procédure proposée consiste en 5 étapes :

- développement du modèle BG ;
- génération des résidus à partir du BG ;
- construction du TCG à partir du BG. Dans cet article, le TCG est utilisé uniquement pour construire le RBD ;
- construction du RBD à partir du TCG et des informations sur les résidus ;
- utilisation du RBD pour estimer l'état futur du système dynamique et la RUL.

Contrairement à cette approche, [Hu et al., 2015] traitent le problème de prédiction de RUL pour lesquels un modèle mathématique décrivant la dégradation des composants est disponible, mais les valeurs des paramètres du modèle sont inconnues et les observations des trajectoires de dégradation dans des composants similaires sont indisponibles. L'approche proposée résout ce problème en utilisant une technique de FP combinée à une méthode de lissage du noyau (KS pour Kernel Smoothing). Cette dernière est utilisée pour améliorer la précision et la robustesse des estimations, surtout dans le cas où il faudrait utiliser un très grand nombre de particules, avec pour conséquence une augmentation de la complexité et du coût de calcul. Le lissage permet également d'utiliser le FP dans des situations où peu de mesures de dégradation sont disponibles.

Ainsi, la méthode FP-KS permet d'estimer simultanément l'état de dégradation et les paramètres inconnus du modèle de dégradation. Sur la base du modèle de dégradation mis à jour (où les paramètres inconnus sont remplacés par ceux estimés), la prédiction de la RUL est ensuite réalisée en simulant les évolutions futures des particules. Une application numérique concernant le pronostic des batteries Lithium-ion est présentée et les résultats obtenus montrent que la méthode FP-KS proposée donne des résultats plus satisfaisants



que les méthodes FP classiques.

Cependant, le FP utilise des modèles analytiques du processus de dégradation et du système d'acquisition des mesures. Cela empêche son applicabilité aux cas, très fréquents dans l'industrie, où il manque des modèles fiables. [Baraldi et al., 2013b] proposent une méthode originale pour étendre le FP au cas où un modèle de mesure analytique n'est pas disponible mais qu'à la place, un ensemble de données contenant des paires «état-mesure» est disponible. Pour ce faire, les données sont utilisées afin de former un ensemble de RN qui est ensuite intégré dans le FP comme modèle de mesure empirique. Le RN fournit donc la distribution des mesures, qui substitue le modèle analytique.

## Synthèse

Ce qui ressort de manière indéniable des différents travaux présentés précédemment est que la fusion de méthodes permet d'améliorer les performances de surveillance [Orchard et al., 2010, Zaidi et al., 2010]. La Table 2.1 résume les améliorations obtenues par fusion de méthodes dans les travaux étudiés précédemment. On constate qu'en plus de l'amélioration de la précision et de la fiabilité des résultats, l'intégration de différentes méthodes permet généralement de simplifier la tâche de surveillance en s'affranchissant de certains prérequis de modélisation propres aux méthodes basées uniquement sur les modèles [Ding et al., 2009, Gertler et al., 1999]. En outre, une représentation précise et une bonne gestion des incertitudes sont généralement possibles, ce qui représente un sujet d'importance capitale, notamment pour le pronostic.

En revanche, les hypothèses concernant la linéarité, stationnarité, etc., des systèmes limitent l'application de ces approches. De plus, elles peuvent devenir très coûteuses en ressources de calcul, et sont généralement conçues spécifiquement pour l'application considérée puisqu'elles ne possèdent pas de potentiel de généralisation. Enfin, dans le cas de juxtaposition de méthodes, les erreurs faites peuvent s'accumuler au fur et à mesure de la réalisation de la tâche de surveillance, d'où l'intérêt croissant pour la fusion de décisions qui s'affranchit de ces limites.

### 2.3.2 Fusion de décisions

La fusion de décisions a suscité un intérêt considérable dans le domaine de la surveillance des systèmes complexes. Cependant, une attention particulière doit être portée aux conflits (décisions contradictoires) qui peuvent compliquer la prise de décision, si un schéma efficace de fusion n'est pas établi.

L'approche la plus simple et la plus populaire de fusion est sans doute le vote majoritaire [Niu et al., 2008, Ng and Srinivasan, 2010, Georgoulas et al., 2013]. L'approche de la moyenne pondérée [Lei et al., 2010, Mylaraswamy and Venkatasubramanian, 1997, Ghosh et al., 2011] quant à elle tente d'évaluer les poids optimaux à assigner aux diverses méthodes utilisées. D'autres schémas de fusion se basent sur le rang. On peut citer le compte de Borda qui est une forme de vote pondéré [Niu et al., 2008], l'approche Bayésienne [Atoui et al., 2015a, Ghosh et al., 2011, Niu et al., 2008, Ng and Srinivasan, 2010], la théorie des croyances de Dempster-Shafer [Ghosh et al., 2011, Zhang and Ge, 2015, Tabassian et al., 2012], et la théorie des possibilités avec la logique floue [Kyriazis and Mathioudakis, 2009, Chen et al., 2017]. Un aperçu détaillé des diverses stratégies de fusion décisionnelle est disponible dans [Kuncheva, 2004].

Méthodes	Cas d'application	Améliorations	Référence
ACP-DWT-DM BG-RB	Réacteur à cuve à agitation continue	Détection de différents types de défauts Amélioration de la fiabilité du diagnostic	[Zhang and Hoo, 2011]
SDG-QTA	TEP	Résolution plus simple du diagnostic Réduction de la complexité du calcul	[Maurya et al., 2007]
Statistiques résidus	Four de chauffage Colonne de distillation	Simplification de la surveillance	[Schubert et al., 2011]
EKF-RN	Réacteurs chimiques	Amélioration de la précision du diagnostic des défauts	[Benkouider et al., 2009] [Benkouider et al., 2012]
FP-RN	Turbine à gaz	Prise en compte des incertitudes	[Daroogheh et al., 2015]
ACP-Espace de parité	TEP	Amélioration du diagnostic des défauts	[Gertler et al., 1999]
ACP-Observateur	TEP	Modélisation non requise pour la génération de résidus	[Ding et al., 2009]
RN-FK	Données UCI	Amélioration de la précision du diagnostic	[Siswanto et al., 2016]
RB-BG	Système à deux réservoirs	Prise de décision plus fiable Prise en compte de la fiabilité des composants	[Weber et al., 2007] [Zaidi et al., 2010]
RB-BG	Système hydromécanique	Amélioration de la précision du pronostic	[Medjaher et al., 2009]
FP-KS	Batteries Lithium-ion	Amélioration de la robustesse des estimations de RULs	[Hu et al., 2015]
FP-RN	Fissure de fatigue	Modélisation non requise de la dégradation	[Baraldi et al., 2013b]

TABLE 2.1 – Fusion de méthodes : récapitulatif.



D'autres approches peuvent être considérées dans ce contexte, bien qu'elles soient davantage utilisées pour l'intégration de données ou la classification, comme les FK, les SVM, les chaînes de Markov cachées, et les RN. Ces approches ne seront pas abordées ici.

Par la suite, nous allons donc présenter les principaux cadres théoriques pour fusionner des décisions que sont les méthodes de vote, l'approche Bayésienne, et la théorie des croyances.

**Méthode de vote** Le vote majoritaire consiste à choisir la décision prise par le maximum de méthodes. Si on dispose de  $m$  méthodes de surveillance  $\{S_1, \dots, S_m\}$  et que la méthode  $S_j$  (avec  $j \in \{1 \dots m\}$ ) attribue la décision  $d_i$  à l'observation  $x$ , alors on note  $S_j(x) = i$  et on associe une fonction indicatrice  $M_i^j$  à chaque méthode :

$$M_i^j(x) = \begin{cases} 1 & \text{si } S_j(x) = i \\ 0 & \text{sinon} \end{cases}$$

La fusion des décisions de  $m$  méthodes est réalisée selon l'équation suivante :

$$M_k(x) = \sum_{j=1}^m M_k^j(x) \quad (2.1)$$

La règle du vote majoritaire consiste donc à choisir le maximum de  $M_k^E$ . La classe correcte est celle qui est le plus souvent choisie par différentes méthodes. Il faut noter qu'un résultat important dans [Lam and Suen, 1997] explique l'intérêt porté pour la méthode de vote majoritaire. En effet, les auteurs prouvent que si le nombre de méthodes est impair, que ces méthodes sont indépendantes et que chaque méthode a une probabilité identique de donner un résultat correct qui est supérieure à 0.5 alors la fusion des décisions par le principe du vote majoritaire donne une meilleure probabilité de succès.

Néanmoins, si toutes les méthodes indiquent des classes différentes, ou dans le cas d'une égalité, alors on est en présence de conflit et la méthode n'admet pas de solution. Ceci représente le principal inconvénient du vote majoritaire.

Par conséquent, cette approche est très simple et intuitive et ne nécessite pas de connaissances a priori, ni de grandes charges de calcul [Georgoulas et al., 2013]. Toutefois, elle ne considère pas les caractéristiques liées à la performance de chaque méthode et n'est intéressante que dans le cas où les méthodes de surveillance sont supposées de performances identiques [Mylaraswamy and Venkatasubramanian, 1997].

Au lieu de combiner les décisions par une somme simple, la somme pondérée [Xu et al., 1992] peut être employée si les performances a priori de chaque méthode pour une décision donnée sont disponibles :

$$M_k^E(x) = \sum_{j=1}^m \alpha_j M_k^j(x) \quad (2.2)$$

où  $\alpha_j$  représente le poids lié à la performance de chaque méthode et où  $\sum_{j=1}^m \alpha_j = 1$ .

On peut noter que contrairement à la méthode du vote classique, une étape d'estimation des poids est nécessaire. Elle peut se faire en se basant sur le taux de bonne classification de chaque méthode par exemple et requiert donc une phase d'apprentissage.

Ainsi, la méthode du vote est adaptée à la fusion de décisions, et particulièrement aux applications de classifications. Dans [Xu et al., 1992], la méthode de vote est comparée

aux approches Bayésiennes et à l'approche de Dempster-Shafer pour de la reconnaissance manuscrite. Les auteurs déduisent que la méthode de vote est moins performante que les deux autres. La même conclusion est faite dans [Kuncheva, 2002]. Néanmoins, dans [Rizvi and Nasrabadi, 2003], malgré sa mise en oeuvre très simple, la méthode du vote majoritaire a présenté des résultats aussi performants que l'approche Bayésienne dans le cas de la reconnaissance de cibles sur des images.

**Approche Bayésienne** L'approche Bayésienne repose sur la théorie des probabilités et permet de prendre en compte l'incertitude de l'information.

La combinaison Bayésienne s'appuie sur la règle de Bayes et consiste à déterminer les probabilités conditionnelles  $P(d_i/S_1, \dots, S_m)$  :

$$P(d_i/S_1, \dots, S_m) = \frac{P(S_1, \dots, S_m/d_i)P(d_i)}{P(S_1, \dots, S_m)} \quad (2.3)$$

Les différentes probabilités dans l'équation 2.3 sont difficiles à estimer, d'où l'introduction généralement d'une hypothèse d'indépendance statistique des méthodes conditionnellement à une décision ainsi que d'une hypothèse d'indépendance statistique. L'équation 2.3 devient alors :

$$P(d_i/S_1, \dots, S_m) = \frac{\prod_{j=1}^m P(S_j/d_i)P(d_i)}{\prod_{j=1}^m P(S_j)} \quad (2.4)$$

La décision finale est souvent choisie selon la règle du maximum a posteriori, c'est-à-dire que la décision  $d_k$  est choisie si  $P(d_k/S_1, \dots, S_m) = \max P(d_i/S_1, \dots, S_m)$  pour  $i \in \{1, \dots, n\}$ .

Il convient de noter que la majeure partie des travaux de recherche qui ont été menés dans le domaine de la fusion décisionnelle repose sur l'approche Bayésienne. Des exemples peuvent être trouvés dans les références [Dromigny and Zhu, 1997, Rodrigues et al., 2000, Lucas, 2001]. La fusion Bayésienne est utilisée pour améliorer la qualité d'un système d'imagerie par rayons X en temps réel en intégrant des informations dans deux conditions d'acquisition différentes [Dromigny and Zhu, 1997]. Un RB est utilisé pour combiner un raisonnement probabiliste avec des paramètres dépendants du temps, pour des fins de contrôle qualité dans les environnements de fabrication [Rodrigues et al., 2000]. [Lucas, 2001] décrit une méthode de diagnostic basée sur la théorie de Bayes, où le raisonnement probabiliste offert par le RB et le raisonnement logique basé sur la cohérence sont intégrés pour minimiser l'incertitude de décision. Les performances du modèle Bayésien de fusion sont étudiées dans [Xu et al., 1992], pour la reconnaissance manuscrite. Il est montré que les performances sont fonction de l'apprentissage. Néanmoins, cette approche est instable puisque les performances sont dégradées si l'apprentissage est mauvais.

Ainsi, la fusion Bayésienne repose sur des bases mathématiques solides et permet de modéliser l'incertitude. Néanmoins, elle ne prend pas en compte l'imprécision et nécessite de disposer d'une grande quantité de données pour estimer les probabilités conditionnelles. En effet, une performance de fusion efficace ne peut être atteinte que si des probabilités conditionnelles et a priori adéquates et appropriées sont disponibles. Bien que, du moins dans certaines situations, des hypothèses puissent être formulées en ce qui concerne les

probabilités a priori et a posteriori, ces hypothèses peuvent devenir déraisonnables dans de nombreuses autres situations.

**Théorie des croyances** La théorie des croyances, appelée aussi théorie de Dempster-Shafer (DS) [Dempster, 1967, Shafer, 1976], permet de représenter à la fois l'incertitude et l'imprécision grâce aux fonctions de croyance et de plausibilité. Elle permet de modéliser des degrés de confiance, appelés masses de croyance, associés à la validité d'une décision. La fusion de décisions dans le cadre de la théorie de DS nécessite de définir un espace de fusion discret fini commun appelé cadre de discernement  $\Theta = \{\theta_1, \dots, \theta_n\}$ . L'ensemble de ses parties,  $\mathcal{P}(\Theta)$  est l'ensemble de tous les sous-ensembles de  $\Theta$ , tel que :  $\mathcal{P}(\Theta) = \{\emptyset, \theta_1, \dots, \theta_n, \{\theta_1, \theta_2\}, \dots, \Theta\}$ .

Dans ce contexte, la masse est définie à travers ces deux propriétés :

- La masse de l'ensemble vide est nulle :  $m(\emptyset) = 0$ .
- La somme des masses des autres sous-ensembles vaut 1 :  $\sum_{X \in \mathcal{P}(\Theta)} m(X) = 1$

La probabilité d'un état appartenant au cadre de discernement est bornée par les fonctions de croyance (*bel*) et de plausibilité (*pl*), qui forment donc son intervalle de confiance :  $bel(X) \leq P(X) \leq pl(X)$ , avec :

$$bel(X) = \sum_{Y \subseteq X} m(Y) \quad (2.5)$$

$$pl(X) = \sum_{Y \cap X \neq \emptyset} m(Y) \quad (2.6)$$

La fonction de croyance *bel* représente la masse minimum qu'il est possible d'affecter à une partie  $X$  de  $\Theta$ , c'est-à-dire toutes les preuves  $Y$  qui supportent l'hypothèse  $X$ , tandis que la fonction de plausibilité *pl* représente la masse maximale, soit toutes les observations qui n'excluent pas l'hypothèse  $X$ .

Chaque méthode de décision génère un jeu de masse noté  $m_i(\cdot)$ . La fusion de deux masses selon la règle de DS se fait de la façon suivante :

$$m_{1,2}(\emptyset) = 0 \quad (2.7)$$

$$m_{1,2}(X) = \frac{1}{1 - K} \sum_{Y \cap Z = X \neq \emptyset} m_1(Y) m_2(Z) \quad (2.8)$$

avec  $K = \sum_{Y \cap Z = \emptyset} m_1(Y) m_2(Z)$  un facteur qui mesure le niveau de conflit entre les deux masses.

Enfin, le choix d'une décision passe par la maximisation d'un critère. Plusieurs possibilités sont envisageables : maximum de plausibilité, maximum de croyance, etc.

La fusion basée sur la théorie des croyances a été appliquée dans plusieurs domaines. On peut citer le traitement d'images, et particulièrement les images médicales [Bloch, 1996, Boudraa et al., 2004, Moraru et al., 2018], et la localisation pour des robots [Ji-Gong et al., 2015].

Par ailleurs, l'application de cette théorie dans la détection et le diagnostic des défauts est rapportée dans plusieurs références [Wang and Wang, 2000, Kaftandjian et al., 2003, Parikh et al., 2001]. Un algorithme de fusion de décisions pour la surveillance de l'état des

moteurs diesel est présenté dans [Wang and Wang, 2000]. [Kaftandjian et al., 2003] décrit comment la théorie des croyances et la théorie des ensembles flous sont combinées pour améliorer la qualité de détection des défauts de soudure. L'utilisation de la théorie des croyances pour combiner les résultats de deux classificateurs primaires afin d'améliorer la performance globale de classification est présentée dans [Parikh et al., 2001]. L'efficacité de cette approche est démontrée à travers la détection de défauts de soupape thermostatique dans un système de refroidissement de moteur diesel.

L'approche DS donne généralement de bons résultats mais implique des calculs plus complexes que les autres approches. Son apport par rapport à la théorie des probabilités tient au fait qu'elle permet d'évaluer non seulement des hypothèses ( $\theta_1$  et  $\theta_2$  par exemple), mais également la partie conjointe qui contient les deux hypothèses. Elle permet également de prendre en compte l'imprécision [Tabassian et al., 2012]. Cependant, sa complexité croît exponentiellement avec la taille du cadre de discernement. De plus, l'estimation des fonctions de masse nécessite une bonne connaissance des données. Ses performances dépendent donc de la qualité et de la quantité des données historiques disponibles et de la phase d'apprentissage. Tout comme l'approche Bayésienne, l'approche DS voit ses performances se dégrader dès lors que l'on s'éloigne des conditions d'apprentissage.

Par la suite, nous allons présenter quelques travaux ayant fusionné les décisions de plusieurs méthodes de surveillance, en s'appuyant sur les cadres de fusion présentés auparavant.

## Travaux

L'un des premiers travaux ayant développé une approche fusionnant les décisions de plusieurs méthodes est celui proposé par [Mylaraswamy and Venkatasubramanian, 1997]. Le cadre hybride, appelé DKit, est illustré dans la Figure 2.13.

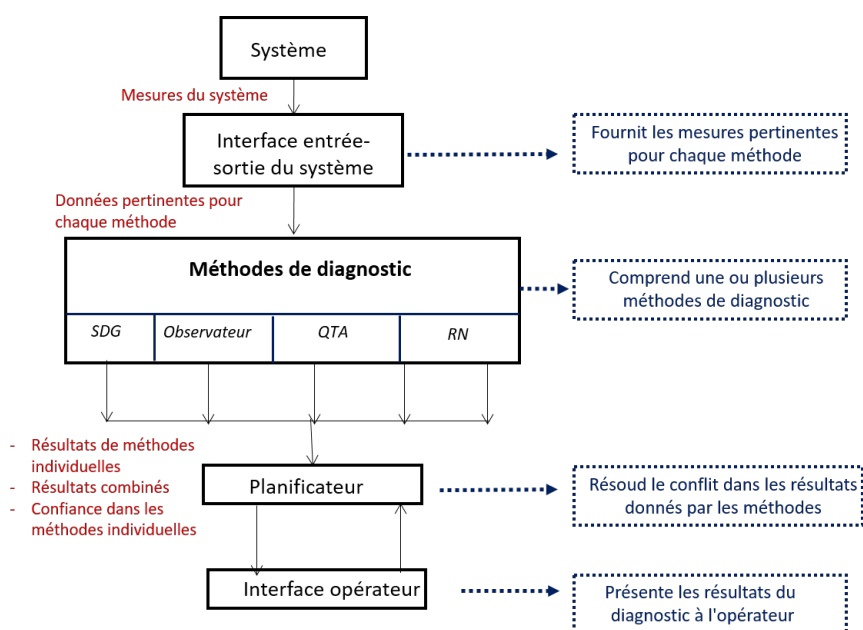


FIGURE 2.13 – Les principaux éléments du Dkit

Il se compose d'une interface d'entrée-sortie qui reçoit les mesures du procédé, d'un ensemble de méthodes de diagnostic, d'un planificateur chargé de résoudre les conflits

dans les résultats et d'une interface d'opérateur qui présente les résultats de diagnostic à l'opérateur. La conception de Dkit dépend de trois paramètres :

- Le choix des méthodes à fusionner (SDG, observateur, QTA et RN)
- Le schéma de déclenchement des méthodes de diagnostic
- L'algorithme de résolution des conflits

Les auteurs ont choisi de combiner la rapidité de réponse des méthodes à base de données (QTA, RN) pour la phase de détection, avec la connaissance des méthodes à base de modèle (SDG, observateur) pour l'identification et l'analyse des causes des défaillances. En effet, selon les auteurs, les méthodes basées sur l'historique des données ne sont pas adaptées pour le diagnostic. Ainsi, une hypothèse de défaut qui ne pourrait pas être expliquée par les méthodes basées sur le modèle sera rejetée. Ceci implique que le SDG et l'observateur sont utilisés pour valider l'hypothèse proposée par la QTA et le RN.

La méthode de fusion utilisée pour combiner les différentes décisions est un schéma de vote pondéré. Les performances de DKit ont été évaluées sur un procédé de craquage catalytique, en simulant différents scénarios de défauts. L'évaluation a pris en compte 4 critères principaux : l'exactitude du diagnostic, la capacité de détection, l'explication de la cause et la rapidité de la réponse. Les résultats ont montré que DKit était capable d'identifier la cause de défaillance et de fournir une explication à la propagation du défaut dans tous les cas simulés.

Le schéma de vote pondéré a également été utilisé dans [Pashazadeh et al., 2018] pour la détection et le diagnostic des défaillances d'éoliennes. Différents classifieurs (RN-Arbre de décisions - kNN) sont implémentés en parallèle et fusionnés. L'analyse des erreurs de classification montre l'efficacité de la fusion des classifieurs par vote pondéré et sa supériorité aux méthodes utilisées séparément.

Ainsi, on constate que de par sa mise en oeuvre très simple, la méthode de vote est largement utilisée. C'est le cas également dans [Vedam et al., 1999]. En effet, la présence de défauts est indiquée à l'aide d'une ACP, tandis que le diagnostic est réalisé avec l'ACP, SDG, et QTA comme indiqué dans la Figure 2.14.

Cette approche fait partie aussi bien de la catégorie de "Fusion de méthodes en série" puisque la détection et le diagnostic sont réalisés par différentes méthodes, que de la catégorie "Fusion de décisions en parallèle" si on prend en compte le fait que deux méthodes sont appliquées pour le diagnostic.

La présence du planificateur vise à éviter le conflit, en utilisant les règles suivantes pour déterminer les pondérations :

- Si ACP-SDG et QTA identifient le défaut, la confiance globale attribuée par le module de diagnostic est la moyenne des probabilités assignées par les différentes méthodes de diagnostic.
- Si ACP-SDG identifient un défaut mais que QTA ne l'identifie pas :
  - Si le défaut fait partie des données historiques utilisées pour développer la base de connaissances de QTA, la confiance attribuée au défaut est de 50% de la confiance attribuée par ACP-SDG.
  - Si le défaut ne fait pas partie des données historiques utilisées pour développer la base de connaissances de QTA, la confiance attribuée au défaut est égale à la confiance attribuée par ACP-SDG.

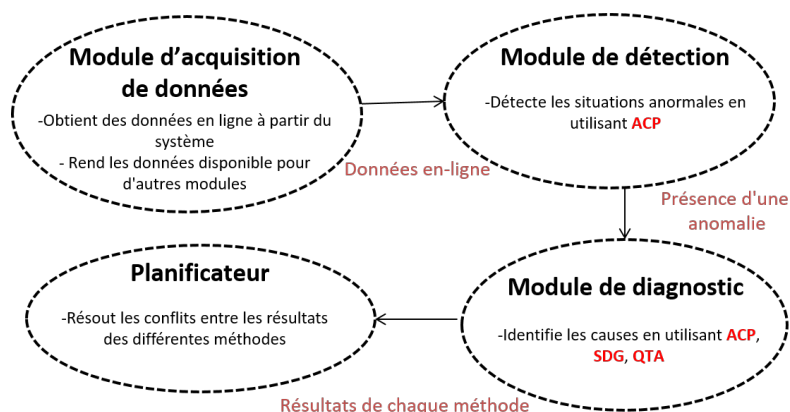


FIGURE 2.14 – ACP, SDG et QTA pour la détection et le diagnostic des défaillances.

- Si ACP-SDG n'identifient pas le défaut et que QTA l'identifie, la confiance globale attribuée par le module de diagnostic est de 50% de la confiance attribuée par QTA.

Ainsi, cette approche a été en mesure d'identifier correctement la nature et l'ampleur des défauts pour un processus de craquage catalytique, mais les questions de robustesse au bruit et de capacité d'adaptation ne sont pas discutées dans cette étude.

Toutefois, comparée aux deux autres approches de fusion, à savoir l'approche Bayésienne et de DS, la méthode du vote présente souvent des performances moins bonnes. L'approche Bayésienne a été utilisée, sous la forme de RB, comme cadre d'intégration de méthodes de détection et de diagnostic dans [Atoui et al., 2015a]. L'objectif du travail était d'associer une méthode guidée par les données, en l'occurrence la statistique  $T^2$ , à un résidu généré par une méthode basée sur un modèle, sous un outil commun : le RB gaussien conditionnel.

Tout d'abord, la statistique  $T^2$  et la matrice d'incidence sont modélisées sous un RB puis combinées grâce à l'approche Bayésienne pour prendre de meilleures décisions, comme décrit dans la Figure 2.15.

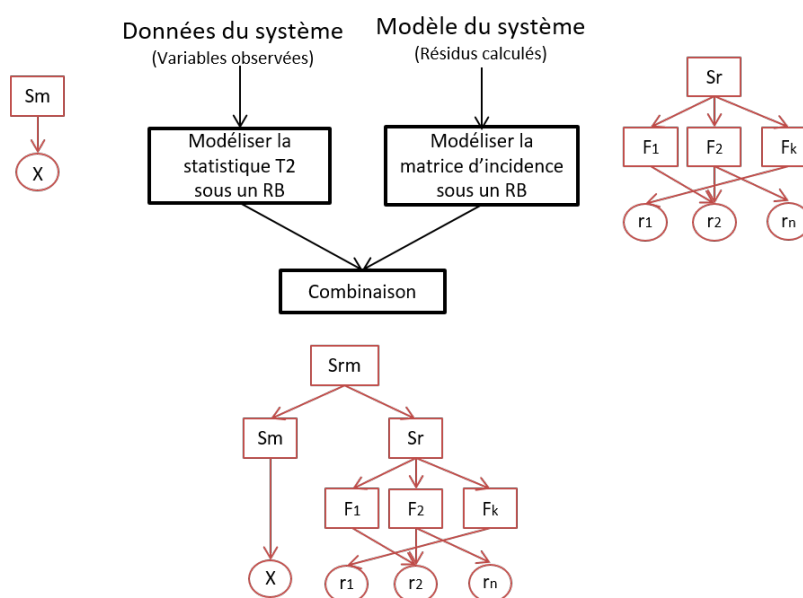


FIGURE 2.15 – Cadre utilisant des mesures et des résidus pour la détection et le diagnostic de défauts.



D'une part, la méthode à base de données fournit une décision à chaque pas de temps, en prenant l'état avec la probabilité a posteriori maximale, compte tenu d'une nouvelle observation. D'autre part, la méthode basée sur un modèle évalue les sorties des résidus et donne à chaque instant la probabilité d'occurrence de chaque défaut. La décision est donnée par le noeud ajouté  $S_{rm}$  qui indique la probabilité d'occurrence pour chaque état. Ce cadre a été simulé sur un procédé de chauffe-eau sous différentes hypothèses (modèle précis, modèle dégradé, grande ou petite quantité de données) et a montré une meilleure fiabilité de décision comparé aux deux méthodes utilisées de façon indépendante. Cependant, il ne gère que les défauts uniques et requiert la modélisation des méthodes à fusionner sous forme de RB, ce qui n'est pas toujours possible. Pour améliorer la prise de décision, d'autres méthodes et informations peuvent être modélisées et intégrées au RB.

Comparé à l'approche Bayésienne, le cadre DS fournit un outil mathématique plus flexible pour traiter des informations imparfaites. Qui plus est, la méthode DS n'impose pas d'hypothèses sur la distribution des données, contrairement à l'approche Bayésienne. Grâce à ces avantages, la méthode basée sur DS a été largement utilisée pour la prise de décision au cours des dernières années. On peut donner comme exemple [Zhang and Ge, 2015]. Cinq méthodes classiquement utilisées (ACP, ACP à noyaux, kNN, AD, RN) sont sélectionnées et leurs décisions sont combinées à travers les règles de DS. En outre, pour améliorer l'efficacité de calcul et la fiabilité du système de fusion, un nouvel indice de mesure de la diversité, appelé coefficient de corrélation, est défini. Les performances de la fusion de décisions sont évaluées par le biais du TEP. Les résultats montrent que les délais de détection et d'identification des défaillances ont diminué par rapport à toute méthode individuelle, que l'efficacité du calcul a été améliorée, ainsi que la fiabilité du système de prise de décision.

Par ailleurs, plusieurs travaux ont tenté de comparer les différentes stratégies de fusion en utilisant la même application et les mêmes conditions d'apprentissage, afin de déterminer laquelle est la plus performante.

Dans ce contexte, [Ghosh et al., 2011] ont abordé la fusion de décisions en mettant en oeuvre l'intégration de méthodes basées sur un modèle (EKF) et sur des données (ACP, RN) pour la surveillance en ligne. L'entrée de chaque méthode représente les données en ligne tandis que la sortie de chaque méthode est une classe assignée (CFN ou classe de défaut). Ces sorties individuelles sont ensuite combinées pour résoudre les conflits et obtenir un résultat consolidé. Dans cet article, les auteurs ont choisi le vote pondéré, la fusion Bayésienne et la fusion DS pour traiter ce problème.

Les auteurs ont réalisé des simulations sur une colonne de distillation, avec deux scénarios différents. Le premier scénario utilisait des méthodes bien conçues présentant de très bonnes performances de classification, qui conduisaient donc à une détection rapide et à l'identification de tous les défauts. Dans ce cas, l'amélioration réalisée par la fusion décisionnelle est marginale et inintéressante. Cependant, en pratique, il est difficile pour une seule méthode de détecter et d'identifier rapidement tous les défauts. Ainsi, dans le deuxième scénario, les méthodes sont repensées afin de détecter et d'identifier seulement certains défauts tout en étant insensibles aux autres, ce qui entraîne un désaccord significatif entre leurs résultats. Dans cette situation, des stratégies de fusion telles que le vote pondéré, la fusion Bayésienne et celle de DS, ont permis d'obtenir une meilleure couverture des défauts, une précision de prédiction plus élevée, et donc une amélioration de la surveillance du système.

Il convient de mentionner que la performance du système de fusion basé sur le vote classique (non pondéré) est plutôt médiocre dans ce travail, et n'est pas meilleure que

n'importe quelle méthode utilisée individuellement. En effet, le vote traite chaque méthode de manière égale sans tenir compte de ses performances spécifiques aux classes. Puisque dans ce travail, chaque méthode est capable de détecter et d'identifier seulement certains défauts, le succès de la fusion de décision dépend principalement de l'utilisation correcte d'informations a priori sur la performance spécifique à la classe. Les méthodes pondérées de vote, Bayésienne et DS, qui utilisent la performance spécifique à la classe à partir de la matrice de confusion, sont donc meilleures.

Si dans le travail précédent, les trois méthodes ont présenté des performances similaires, dans [Niu et al., 2008], il apparaît que l'approche Bayésienne est la plus performante pour le diagnostic des défauts dans un moteur d'ascenseur. L'approche comprend 4 étapes :

- extraction des caractéristiques à partir des données ;
- développement de plusieurs méthodes de classification : SVM, AD, k-NN, RN ;
- sélection de classifieurs basée sur la corrélation ;
- algorithmes de fusion de décision (vote majoritaire, Bayésien, etc.).

Quant au pronostic, très peu de travaux ont tenté de combiner des décisions émanant de différentes méthodes. En effet, alors que des recherches significatives ont été menées sur le pronostic basé sur des modèles et des données, peu de recherches ont été faites sur la prédiction de la RUL en combinant les résultats de prédiction de diverses méthodes. Ceci est justifié par le fait que le principe du pronostic se base généralement sur une estimation ou une prédiction de la RUL plutôt qu'une classification. Néanmoins, plusieurs travaux ont considéré le problème de pronostic comme un problème de discrimination en plusieurs classes, à savoir par exemple : classe *CFN*, classe dégradation et classe défaillance, le rapprochant ainsi du diagnostic.

Ainsi, [Georgoulas et al., 2013] ont considéré le pronostic dans leur travail comme une détection "précoce" d'anomalies. Ceci est réalisé en combinant les décisions de trois méthodes : ACP-RN-PG formées en utilisant différentes entrées et en fusionnant leurs décisions à travers un système de vote majoritaire.

Dans [Li et al., 2017], l'objectif est de proposer une nouvelle approche de pronostic qui prend en compte les effets de la dégradation sur la précision de la prédiction de la RUL. Pour cela, elle attribue à chaque méthode (PF- SVM- RNN) un poids optimisé dépendant de la dégradation de sorte que la somme pondérée des prédictions des différentes RULs soit plus précise. Cette approche est appliquée sur les roulements d'aéronefs et sur des moteurs d'avion, et sa performance surpasse celle des méthodes individuelles.

[Hu et al., 2012] proposent une approche de pronostic basée uniquement sur les données. Ils fusionnent plusieurs algorithmes d'apprentissage (RNN-SVM-RVM) grâce à la somme pondérée. Contrairement aux approches précédentes qui utilisent une seule méthode de pondération pour la fusion, ici trois méthodes de pondération (la pondération basée sur la précision, la pondération basée sur la diversité et celle basée sur l'optimisation), sont proposées pour déterminer les poids de chaque méthode. Les résultats obtenus montrent que n'importe quel schéma de pondération donne des prédictions de RULs plus précises que les méthodes utilisées individuellement. Cependant les méthodes individuelles doivent présenter des performances différentes tout en conservant une précision comparable.

## Synthèse

On peut conclure que la fusion de décisions permet d'améliorer plusieurs critères comme la précision, la fiabilité et la robustesse. La Table 2.2 résume les améliorations ob-



tenues par fusion de décisions, en fonction de l'algorithme utilisé, dans les travaux étudiés précédemment. La diversité des méthodes et la stratégie de fusion sont deux composantes clés du système de fusion de décisions. La diversité peut être obtenue en utilisant des méthodes hétérogènes [Mylaraswamy and Venkatasubramanian, 1997, Pashazadeh et al., 2018, Vedam et al., 1999] ou en utilisant différents ensembles de données d'apprentissage [Zhang and Ge, 2015, Ghosh et al., 2011]. En ce qui concerne la stratégie de fusion, plusieurs types d'approches ont été développés, et il apparaît que les approches les plus performantes sont les approches Bayésiennes et les approches basées sur la théorie de DS.

Contrairement à la fusion de méthodes en série, la fusion de décisions n'entraîne pas une accumulation d'erreurs et ne requiert pas l'intégration de différents concepts et théories.

### 2.3.3 Discussion

Malgré les avantages évidents que la fusion d'approches peut offrir pour la détection, le diagnostic et le pronostic des défauts, son potentiel est encore faiblement exploité. En fait, plusieurs obstacles sont rencontrés lors de l'élaboration d'un système fusionnant plusieurs approches. Les verrous scientifiques révélés par cette étude bibliographique sont les suivants :

**1. Manque de généricité** Tout d'abord, l'absence d'un cadre générique freine le développement des approches de surveillance [Tidiri et al., 2016]. La majorité des travaux de la littérature proposent des solutions dédiées à l'application ou au modèle de détection-diagnostic/pronostic utilisé. Ainsi, le développement d'un système de surveillance dépend de l'application considérée et sa mise en oeuvre nécessite une conception spécifique, ce qui peut se révéler coûteux en ressources financières, matérielles et humaines [Mylaraswamy and Venkatasubramanian, 1997, Pashazadeh et al., 2018, Ghosh et al., 2011, Maurya et al., 2007, Siswanto et al., 2016]. De plus en plus d'études tentent de fournir une solution à ce problème, mais cela reste le principal verrou scientifique pour le développement de la fusion d'approches.

**2. Performances a priori** La plupart des travaux visant la fusion de plusieurs approches pour la surveillance ne considèrent pas les performances individuelles de chaque méthode. Ceci est d'autant plus vrai dans le cas de fusion de méthodes [Maurya et al., 2007, Schubert et al., 2011, Zhang and Hoo, 2011, Daroogheh et al., 2015] ou de caractéristiques [Ding et al., 2009, Siswanto et al., 2016, Medjaher et al., 2009, Hu et al., 2015]. Or, ces informations a priori sont très intéressantes et doivent être intégrées systématiquement dans l'élaboration d'approches de surveillance. Pour cela, il faudrait d'abord évaluer la performance de chaque méthode puis utiliser cette information afin de déterminer le degré de priorité à donner à chaque méthode. C'est le cas de certains travaux fusionnant les décisions avec des approches Bayésiennes, de DS et de vote pondéré [Zhang and Ge, 2015, Ghosh et al., 2011].

**3. Choix des méthodes à fusionner** Un des plus grands défis dans la conception de la fusion d'approches est de trouver des méthodes dont la combinaison permet de profiter des avantages de chaque méthode tout en évitant leurs inconvénients. En effet, l'intégration de différentes méthodes peut s'avérer une tâche difficile en pratique car même si chaque méthode fonctionne individuellement de façon très satisfaisante, rien n'assure que la combinaison présentera des performances comparables ou meilleures. Certains travaux conseillent de choisir des méthodes hétérogènes qui auront tendance à se compléter

Méthodes	Méthode de fusion	Cas d'application	Améliorations	Référence
SDG-QTA-RN Observateur	Vote pondéré	Craquage catalytique	Amélioration de la capacité de détection et de diagnostic	[Mylaraswamy and Venkatasubramanian, 1997]
RN-Arbres de décisions- kNN	Vote majoritaire	Eoliennes	Amélioration du diagnostic	[Pashazadeh et al., 2018]
SDG-ACP	Vote pondéré	Craquage catalytique	Amélioration de la capacité de détection	[Vedam et al., 1999]
Résidus statistique $T^2$	Approche Bayésienne	TEP	Amélioration de la fiabilité de décision	[Atoui et al., 2015a]
ACP-RN kNN-AD	Approche de DS	TEP	Amélioration de la fiabilité de décision	[Zhang and Ge, 2015]
EKF-ACP RN	Vote pondéré Approche Bayésienne Approche de DS	Colonne de distillation TEP	Amélioration de la précision de prédiction et rapidité de la prise de décision	[Ghosh et al., 2011]
SVM-RN AD-kNN	Approche Bayésienne DS- Vote	Moteur d'ascenseurs	Amélioration de la classification	[Niu et al., 2008]
ACP-RN-PG	Vote majoritaire	Palier	Amélioration de la fiabilité de décision	[Georgoulas et al., 2013]
PF-RVM- RNN	Vote pondéré	Moteurs d'avion	Amélioration de la précision du RUL	[Li et al., 2017]
RVM-SVM RNN	Vote pondéré	Turboréacteurs, Ventilateur	Amélioration de la précision du RUL	[Hu et al., 2012]

TABLE 2.2 – Fusion de décisions : récapitulatif.

[Maurya et al., 2007, Schubert et al., 2011, Gertler et al., 1999, Ding et al., 2009] tandis que d'autres proposent d'utiliser différents jeux de données d'apprentissage quand il s'agit de méthodes à base de données [Zhang and Ge, 2015, Ghosh et al., 2011, Niu et al., 2008, Li et al., 2017, Hu et al., 2012].

De plus, la fusion d'approches nécessite généralement de relier différentes théories et concepts, et de modifier certaines caractéristiques de méthodes existantes afin de les rendre compatibles avec d'autres. Ceci suppose d'avoir des connaissances solides relatives à plusieurs théories [Gertler et al., 1999, Ding et al., 2009]. En revanche, la fusion de décisions n'exige pas ce type de connaissances.

La fusion d'approches requiert également la prise en compte de plusieurs hypothèses, comme la possibilité de développer un modèle par exemple, ce qui peut conduire à une réduction de la portée de ces approches, contrairement aux approches qui fusionnent des décisions.

**4. Choix de l'algorithme de fusion** Concernant la fusion des décisions, une fois que les résultats de diverses méthodes sont disponibles, une étape importante consiste à appliquer une stratégie efficace pour les combiner. Cependant, le développement de stratégies de fusion décisionnelle pour combiner les résultats reste largement inexploré.

**5. Assurance *a priori* d'amélioration des performances** D'après notre revue de la littérature, aucune stratégie de fusion d'approches ne peut assurer *a priori*, c'est-à-dire durant sa phase de conception, une amélioration certaine des performances de surveillance. En effet, les différents travaux appliquent leurs stratégies sans savoir si elles apporteront une nette amélioration, ou non, de la performance. Une évaluation de l'amélioration *a priori*, à partir des performances individuelles par exemple, n'est pas considérée.

**6. Prérequis d'apprentissage** La plupart des approches de fusion requièrent une grande quantité de données historiques pour estimer par exemple leurs paramètres. C'est le cas des approches Bayésiennes [Atoui et al., 2015a, Ghosh et al., 2011, Niu et al., 2008, Ng and Srinivasan, 2010], de la théorie des croyances DS [Ghosh et al., 2011, Zhang and Ge, 2015, Tabassian et al., 2012], ou du vote pondéré [Lei et al., 2010, Mylaraswamy and Venkatasubramanian, 1997, Ghosh et al., 2011]. Or, ces données historiques ne sont pas toujours disponibles en pratique, ce qui pourrait fausser, voire limiter l'utilisation de ces approches en réalité.

**7. Intégration du pronostic** Enfin, il apparaît que malgré le lien évident entre la détection, le diagnostic et le pronostic, rares sont les travaux de surveillance qui les ont traités ensemble. Alors que ces tâches peuvent être intrinsèquement différentes les unes des autres (comme vu dans la première partie de l'état de l'art 2.2), elles sont cependant étroitement liées entre elles et ne peuvent pas être considérées comme des tâches isolées ou des problèmes distincts [Tidriri et al., 2016, Venkatasubramanian et al., 2003a]. Cette séparation artificielle devrait disparaître pour que de réels progrès puissent être réalisés dans le domaine de la surveillance. Cela crée d'excellentes opportunités de recherche pour les communautés du diagnostic et du pronostic. Par conséquent, il faudrait s'attacher à développer des approches dans lesquelles la détection, le diagnostic et le pronostic peuvent être intégrés dans un cadre unifié unique, de sorte que la prise de décision soit plus complète et plus efficace.

Ainsi, l'utilité de la fusion d'approches est d'arriver à une solution plus complète afin d'augmenter la confiance globale dans la décision et d'améliorer la performance de la

surveillance. Mais, comment mesurer quantitativement cette amélioration ? Évidemment, il est utile d'avoir un ensemble de critères pour comparer différentes méthodes entre elles. Les systèmes de surveillance basés sur des fusions doivent répondre aux caractéristiques souhaitables des systèmes de surveillance classiques, telles que la détection rapide, le diagnostic avec un taux de fausses alarmes tolérable, le pronostic précis, la robustesse au bruit et aux incertitudes, etc. Des discussions détaillées de ces critères peuvent être trouvées dans [Venkatasubramanian et al., 2003b, Saxena et al., 2008a]. En plus de ces critères, la conception d'une approche qui fusionne plusieurs méthodes devrait inclure d'autres caractéristiques souhaitables telles que :

- Une grande couverture des défauts.
- Peu d'exigences et de prérequis a priori. L'objectif est de fournir une meilleure solution avec moins d'exigences.
- Une stratégie de fusion efficace pour éviter les conflits en amplifiant les bonnes décisions tout en diminuant l'impact des décisions incorrectes.

## 2.4 Positionnement scientifique

Ce chapitre a mis en évidence plusieurs opportunités de recherche qui devraient être abordées pour une conception efficace de stratégies de fusion d'approches.

Dans ce travail de recherche, nous nous intéressons à la fusion de décisions dédiée à la surveillance des systèmes. En effet, suite à l'étude bibliographique, il apparaît que ce sujet reste encore largement inexploré et présente de nombreux défis. Nous avons donc identifié plusieurs verrous scientifiques dans la partie précédente que nous allons tenter de lever à travers l'approche proposée. Cette dernière devra répondre à la problématique de surveillance des systèmes par la fusion de décisions en considérant et en intégrant dans sa démarche les caractéristiques suivantes :

**1. Cadre générique** L'approche proposée doit être construite autour d'un cadre générique et théorique qui s'adapte à plusieurs méthodes de surveillance. L'intégration appropriée de différentes stratégies de surveillance selon un formalisme unifié et un cadre qui n'est pas dédié à une application spécifique semble être une piste intéressante à explorer.

**2. Evaluation et exploitation des performances a priori** Les performances a priori concernant chaque méthode individuelle doivent être évaluées et exploitées dans la nouvelle méthodologie, afin d'assigner à chaque méthode son degré de priorité en fonction des défauts considérés.

**3. Choix non important des méthodes à fusionner** Puisque l'objectif de notre stratégie est de proposer un cadre générique, le choix des méthodes à fusionner n'a pas une grande incidence. Cependant, afin d'apporter la meilleure amélioration de performances possibles, et comme vu dans la littérature, il est recommandé de choisir des méthodes qui présentent une grande diversité afin de tirer profit au maximum de leurs avantages.

**4. Approche Bayésienne pour la fusion** Parmi les approches de fusion présentées, l'approche Bayésienne, et plus particulièrement le RB, a attiré notre attention en raison de ses propriétés en termes de modélisation et d'inférence. En effet, le RB offre un

cadre probabiliste capable de fusionner des informations et des connaissances incertaines provenant de diverses sources. Nous listons ici les principaux avantages du RB :

- Le RB est une représentation graphique compacte et synthétique avec une facilité d'acquisition et d'utilisation de l'information.
- Le RB peut combiner plusieurs aspects à la fois, des statistiques, des probabilités, de l'aide à la décision et de la gestion de l'information.
- La flexibilité de l'inférence permet d'interroger le même modèle graphique pour des objectifs différents, tels que la prédiction (pour le pronostic par exemple) ou la classification (pour la détection-diagnostic par exemple).
- Le RB modélise l'information par une attribution de probabilités permettant ainsi de prendre en compte l'aspect incertain des données et de la tâche de surveillance.
- L'inférence Bayésienne offre un outil puissant de fusion de données, suivant un schéma d'apprentissage.
- Plusieurs méthodes, ou du moins leurs décisions, peuvent être modélisées sous forme de RB [Atoui et al., 2015c, Verron et al., 2010b].

**5. Assurance a priori de l'amélioration** L'approche proposée se base sur une méthodologie qui permet d'assurer à l'utilisateur, durant sa phase de conception, une amélioration des performances en fonction des performances individuelles existantes.

**6. Paramétrage théorique** Afin d'éviter la limite relative à la disponibilité des données d'apprentissage pour l'estimation des paramètres d'un RB, l'approche proposée s'appuie sur un paramétrage théorique des probabilités conditionnelles. Ce paramétrage se fait suivant les performances a priori des méthodes individuelles ainsi que des objectifs de performances à atteindre et fournit alors des paramètres qui respectent les objectifs fixés.

**7. Pronostic** Enfin, le cadre de fusion doit permettre de réaliser aussi bien le diagnostic que le pronostic afin de fournir une solution complète au problème de surveillance des systèmes complexes.

## 2.5 Conclusion

Dans ce chapitre, nous avons présenté un aperçu de différentes méthodes basées sur les modèles et les données pour la détection, le diagnostic, et le pronostic de défauts. Nous avons décrit les principaux concepts sur lesquels reposent les deux communautés de surveillance. Les caractéristiques des différentes approches fondées sur les données et les modèles ont été mises en évidence et les avantages et les limites de chaque approche ont été soulignés par une analyse comparative. La principale conclusion qui découle de cette première partie est que malgré les nombreux travaux qui existent, on ne dispose toujours pas d'un système de surveillance clés en main puisqu'aucune méthode n'est parfaite. En effet, dès lors que l'on s'écarte des conditions académiques idéales où le système est supposé linéaire et invariant, que les bruits sont constants et gaussiens, que les distributions sont connues et gaussiennes, que les données sont nombreuses, disponibles et de bonne qualité, etc., on est confronté au problème de la robustesse et de l'amélioration

des performances.

Un examen des travaux existants ayant tenté de combiner plusieurs méthodes de surveillance a été réalisé dans la seconde partie. Nous avons distingué deux manières de fusionner des méthodes : soit en série lorsque les méthodes sont utilisées l'une à la suite de l'autre pour différentes tâches, soit en parallèle lorsqu'elles sont appliquées pour fournir une solution à la même tâche. La conclusion tirée de cette seconde partie est que la fusion de décisions reste largement inexplorée et ce en raison de plusieurs obstacles qu'elle rencontre : solutions dédiées et manque de généralité, aucune assurance d'amélioration a priori des performances, choix difficile de méthodes à fusionner, choix de la stratégie de fusion, etc.

Enfin, nous avons situé notre travail de recherche par rapport à la problématique liée à la surveillance des systèmes en ayant recours à deux ou plusieurs méthodes. Nous proposons dans ce travail de nous servir des avantages que peut proposer un cadre probabiliste pour la fusion de décisions de manière générique tout en évitant les inconvénients liés à la disponibilité de données d'apprentissage. Ainsi, nous mettons en avant une nouvelle façon de déterminer théoriquement les paramètres du cadre probabiliste, à partir des objectifs de surveillance que l'on veut atteindre et non à partir d'un jeu de données, qui assure a priori une amélioration de performances de surveillance.



# Fusion de décisions pour la détection et le diagnostic

## Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>54</b>
<b>3.2</b>	<b>Formulation du problème</b>	<b>54</b>
<b>3.3</b>	<b>Démarche méthodologique</b>	<b>57</b>
3.3.1	Evaluation des performances individuelles	58
3.3.2	Stratégie de FDD : Définition des objectifs	62
3.3.3	Stratégie de FDD : Cadre générique de fusion	65
3.3.4	Stratégie de FDD : Développement des objectifs	67
3.3.5	Stratégie de FDD : Prise de décision multi-objectifs sous contraintes	72
3.3.6	Fusion des décisions en ligne	76
<b>3.4</b>	<b>Discussion</b>	<b>78</b>
3.4.1	Choix des objectifs	78
3.4.2	Analyse des paramètres obtenus	79
3.4.3	Synthèse	87
<b>3.5</b>	<b>Application : Diagnostic du Tennessee Eastman Process</b>	<b>88</b>
3.5.1	Introduction	88
3.5.2	Présentation du TEP	89
3.5.3	Diagnostic du TEP par le modèle	91
3.5.4	Diagnostic du TEP par les données	100
3.5.5	Diagnostic du TEP par la fusion de décisions	101
<b>3.6</b>	<b>Conclusion</b>	<b>106</b>

---



### 3.1 Introduction

Le développement des approches combinant plusieurs méthodes de surveillance rencontre un intérêt croissant, qui réside dans le fait que diverses méthodes peuvent généralement se compléter.

En effet, par une fusion adaptée de plusieurs méthodes de surveillance, il serait possible de réduire fortement les erreurs de détection et de diagnostic, particulièrement lorsque les méthodes individuelles sont performantes sur des ensembles de défauts disjoints ou se superposant partiellement. Ici, il n'est plus question de sélectionner la méthode la plus performante parmi d'autres pour réaliser la surveillance mais plutôt d'utiliser les méthodes disponibles et de les fusionner de manière appropriée.

Dans ce chapitre, nous proposons un cadre unifié permettant la fusion de décisions de deux méthodes de détection-diagnostic de défauts (FDD). Ce cadre, qui se décline sous la forme d'un RB discret, est générique et adapté pour toute méthode de FDD.

La méthodologie est basée sur un calcul des paramètres du RB, en fonction des objectifs FDD à atteindre. Le développement conduit à un problème multi-objectif sous contraintes, résolu par le biais d'une approche lexicographique. L'efficacité de l'approche de fusion décisionnelle proposée est validée sur le Tennessee Eastman Process (TEP), système devenu une véritable référence au sein de la communauté FDD lorsque l'on teste les approches de détection et de diagnostic. L'application met en exergue la capacité de l'approche à assurer une amélioration significative des performances en FDD, en fournissant un taux élevé de diagnostic de défauts, un faible taux de fausses alarmes et une stratégie efficace pour la résolution des conflits entre différentes méthodes de FDD.

### 3.2 Formulation du problème

Etant donné un système à surveiller, on définit un ensemble de  $n$  défauts  $S = \{s_1, s_2, \dots, s_n\}$  pouvant affecter le système. Ces défauts peuvent présenter différents profils variants dans le temps.

On introduit un ensemble de 2 méthodes  $M = \{m_1, m_2\}$  à fusionner pour le diagnostic du système considéré. Ces méthodes peuvent être à base de modèles ou de données.

Enfin, on présente un ensemble de classes de fonctionnement  $CF = \{CFN, s_1, s_2, \dots, s_n\}$  représentant tous les états possibles du système.

**Définition 9.** Une *classe de fonctionnement*  $cf_t \in CF$  représente l'état dans lequel se trouve le système surveillé à l'instant  $t$ .

Ainsi, on considère qu'un système peut être soit en Conditions de Fonctionnement Normal, représenté par la classe  $CFN$ , soit affecté par un ou plusieurs défauts, représentés par une des classes de fonctionnement  $\{s_1, s_2, \dots, s_n\}$ .

Ceci nous conduit à introduire la notion de *décision* d'une méthode de FDD.

**Définition 10.** Une *décision*  $d_i$  est définie comme la classe ou l'ensemble des classes de fonctionnement  $cf_i \subseteq CF$  choisi(e) par la méthode de FDD  $m_i \subseteq M$  comme état du système.

L'ensemble des décisions  $d_i$  est contenu dans  $D_i$ .

Afin d'évaluer la performance de chaque méthode de FDD, le taux de fausse alarme (FAR pour False Alarm Rate) et le taux de diagnostic de défaut (FDR pour Fault Diagnosis Rate), sont considérés.

**Définition 11.** La **FAR** est le pourcentage d'observations appartenant à la classe  $CFN$  et diagnostiquées à tort comme appartenant à  $S$  par la méthode de FDD. Il est calculé comme suit :

$$FAR = \frac{\text{Nombre d'observations normales identifiées comme défectueuses}}{\text{Nombre total d'observations normales}} \times 100 \quad (3.1)$$

**Définition 12.** La **FDR** est le pourcentage d'observations appartenant à la classe  $S_i$  et correctement diagnostiquées comme appartenant à  $S_i$  par la méthode de FDD. Il est calculé comme suit :

$$FDR = \frac{\text{Nombre d'observations défectueuses correctement diagnostiquées}}{\text{Nombre total d'observations défectueuses}} \times 100 \quad (3.2)$$

Il convient de préciser que si l'on s'intéresse à la performance des méthodes en détection uniquement (et non pas en diagnostic), une autre métrique appelée taux de détection de défaut (FDR pour Fault Detection Rate) pourrait être considérée. Cette dernière ne fait pas la distinction entre les différentes classes de défauts. Dans la suite, la métrique utilisée est bien celle définie dans l'équation 3.2.

En plus de ces deux métriques, on utilise souvent une matrice de confusion afin d'illustrer les performances d'une méthode de FDD. Chaque colonne de cette matrice représente le nombre d'occurrences d'une classe prédite, tandis que chaque ligne représente le nombre d'occurrences d'une classe de fonctionnement réelle, comme illustré dans la Table 3.1. Par exemple,  $x_{11}$  représente le nombre de fois où la méthode de FDD a décidé que le système est en  $CFN$  sachant qu'il était réellement en  $CFN$  tandis que  $x_{1n}$  représente le nombre de fois où la méthode a décidé que le système est en classe de défaut  $s_n$  sachant qu'il était réellement en  $CFN$ . Ces valeurs peuvent également être exprimées en pourcentage.

La matrice de confusion permet de montrer rapidement et visuellement si la méthode de FDD utilisée parvient à prendre les bonnes décisions par rapport à l'état du système, et donc si elle présente de bonnes performances.

On peut déduire les deux métriques  $FAR$  et  $FDR$  directement de la matrice de confusion de la sorte :  $FAR = 1 - x_{11}$  et  $FDR = \frac{x_{22} + x_{33} + \dots + x_{nn}}{n-1}$ .

Matrice de confusion		Décision			
		$CFN$	$s_1$	$\dots$	$s_n$
Réalité	$CFN$	$x_{11}$	$x_{12}$	$\dots$	$x_{1n}$
	$s_1$	$x_{21}$	$x_{22}$	$\dots$	$x_{2n}$
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	$s_n$	$x_{n1}$	$x_{n2}$	$\dots$	$x_{nn}$

TABLE 3.1 – Exemple de matrice de confusion

Ainsi, on peut définir ce que l'on entend par "bonnes performances" comme suit :

**Définition 13.** Une méthode de FDD présente de **bonnes performances** si son  $FAR$  tend vers 0 et que son  $FDR$  tend vers 1 (ou 100% en pourcentage).

En effet, dans l'idéal, le FAR doit être nul tandis que le FDR doit être égal à 1. Quant à la matrice de confusion, tous les éléments en dehors de la diagonale principale doivent être nuls.

Néanmoins, qualifier les performances en FDD de "bonnes" ou "mauvaises" est un concept qui reste très subjectif et qui dépend du système considéré et des attentes de l'utilisateur. En effet, un FAR de 1% peut être considéré comme correct pour une application donnée et inacceptable pour une autre.

Dans la suite, on considère que les performances individuelles de chaque méthode sont évaluées a priori grâce aux métriques et à la matrice de confusion définies ci-dessus. Par conséquent, l'ensemble des défauts diagnosticables avec de bonnes performances par chaque méthode de FDD peut être identifié.

Ainsi, deux types de défauts, résumant tous les scénarios possibles, sont définis.

**Définition 14.** Un défaut  $s_i \subseteq S$  est de **type** 0 (noté  $f_0$ ) si les 2 méthodes de FDD utilisées présentent de bonnes performances pour son diagnostic. L'ensemble de tous les défauts de type 0 est noté  $F_0$ .

**Définition 15.** Un défaut  $s_i \subseteq S$  est de **type**  $k$  (noté  $f_k$ ) ( $k = \{1, 2\}$ ) si uniquement la méthode  $m_k$  présente de bonnes performances pour son diagnostic. L'ensemble de tous les défauts de type  $k$  est noté  $F_k$ .

En effet, pour chaque méthode dont on désire fusionner les décisions, deux possibilités s'offrent pour les défauts qu'elle est capable de diagnostiquer avec de bonnes performances : soit ces défauts sont diagnostiqués tout aussi bien par la seconde méthode et dans ce cas ce sont des défauts communs de type 0, soit ces défauts lui sont spécifiques puisqu'elle est la seule à présenter de bonnes performances pour leur diagnostic. Trois types de défauts sont alors considérés : les défauts de type 0, 1 et 2. Cela implique que les décisions qui peuvent être données par les méthodes  $m_1$  et  $m_2$  sont respectivement :  $D_1 = \{CFN, F_0, F_1\}$  et  $D_2 = \{CFN, F_0, F_2\}$ , avec  $F_0 \subset S$ ;  $F_1 \subset S$ ;  $F_2 \subset S$ .

La formulation du problème de fusion de décisions concerne tout aussi bien le cas de défaut simple que le cas des défauts multiples. Lorsque l'on fait l'hypothèse de défaut simple, l'objectif est d'obtenir une décision globale composée d'une classe de fonctionnement unique qui représente l'état du système à l'instant  $t$ . Cependant, parfois, de multiples défauts peuvent se produire en même temps et par conséquent l'objectif est élargi pour identifier un ensemble de défauts.

Toutefois, nous faisons l'hypothèse suivante (Hypothèse 1) concernant la cardinalité des défauts.

**Hypothèse 1.** Un nombre limité  $l$  de défauts peuvent se produire en même temps dans le système.

Les défauts peuvent différer selon qu'ils affectent les actionneurs, les capteurs ou le procédé lui-même. Nous nous intéressons ici à tout type de défaut.

Par ailleurs, les défauts peuvent également avoir différents profils variant dans le temps. Ils peuvent être persistants si le système reste défaillant après l'occurrence du défaut, ou intermittents, si le système est alternativement en fonctionnement normal et défaillant après l'occurrence du défaut. Comme les défauts intermittents sont difficiles à détecter et à caractériser, et qu'ils ont rarement été traités dans l'état de l'art, nous nous limitons à la détection et au diagnostic des défauts persistants.

**Hypothèse 2.** *Tous les défauts sont persistants.*

Cela nous conduit à la formulation suivante du problème de fusion de décisions.

**Problème-Fusion de décisions.** *Étant donné un système surveillé avec 2 méthodes, et présentant un ensemble de défauts persistants  $S$ , un ensemble de classes de fonctionnement  $CF$  et deux décisions ( $d_1 \subseteq D_1$ ;  $d_2 \subseteq D_2$ ) données à l'instant  $t$  par les 2 méthodes, le problème de fusion de décisions consiste à choisir la décision (appelée décision globale) qui améliore les performances globales du système de surveillance, en terme de FDR et de FAR.*

### 3.3 Démarche méthodologique

La démarche méthodologique proposée peut être découpée en trois étapes principales, illustrées par la Figure 3.1 et décrites plus en détails ci-dessous :

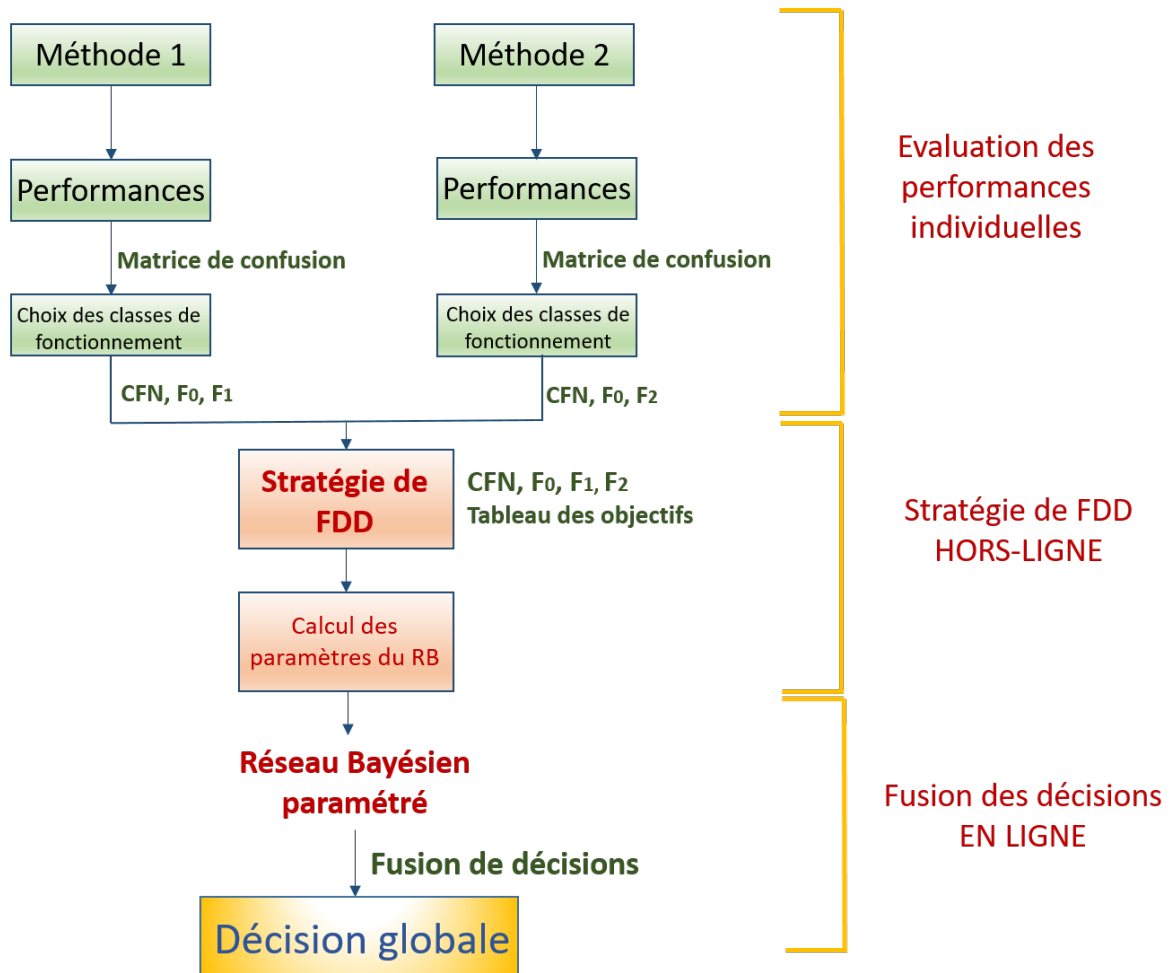


FIGURE 3.1 – La méthodologie proposée

1. **Évaluation des performances individuelles** : Une évaluation des performances de chaque méthode FDD est réalisée. Une matrice de confusion résumant ces performances est construite pour chaque méthode et les FDR et FAR sont déduits. En effet, avant de vouloir fusionner des méthodes FDD, il convient d'abord de les tester individuellement afin de connaître leurs performances. A partir des matrices de confusion, les différents types de défauts ( $F_0, F_1, F_2$ ) sont définis.

2. **Stratégie de FDD (Hors-ligne)** : Par la suite, la stratégie de surveillance est déployée et le tableau des objectifs est défini. Ce dernier permet de construire la décision globale  $d$  à partir des décisions individuelles  $d_1$  et  $d_2$ , et va guider le calcul des paramètres du RB. Une fois les paramètres choisis, le RB est prêt à être utilisé en ligne pour la surveillance du système.
3. **Fusion des décisions (En ligne)** : Les décisions des méthodes individuelles sont introduites dans le RB et fusionnées grâce à la stratégie adoptée. La décision globale est renvoyée par le RB.

Par la suite, nous allons développer les différentes briques de la méthodologie présentée ci-dessus.

### 3.3.1 Evaluation des performances individuelles

Cette première étape permet d'identifier les défauts qui sont diagnostiqués par chacune des méthodes, et ceux qui ne le sont pas, à travers les matrices de confusion et les métriques FDR et FAR (voir Figure 3.2).

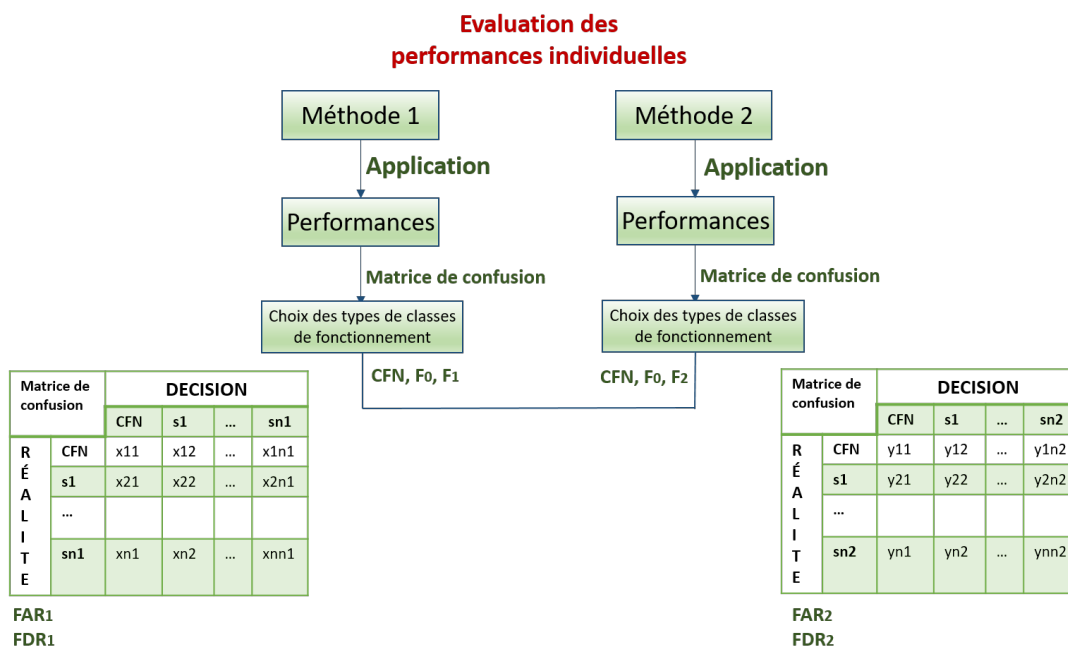


FIGURE 3.2 – Evaluation des performances individuelles

Ainsi, chaque méthode possède sa propre matrice de confusion, qui reflète sa performance. Comme le but est de savoir, pour chaque défaut, quelle méthode est la plus performante, le problème revient à comparer les deux matrices de confusion. Plus particulièrement, il convient de comparer les éléments sur leurs diagonales puisqu'ils reflètent les proportions correctement classées.

L'utilisateur des méthodes de diagnostic peut lui-même fixer les seuils au-delà desquels la performance d'une méthode en FDD est jugée acceptable. Ce choix peut dépendre des expériences passées ou tout simplement de ses connaissances et/ou de ses exigences. Par exemple, on peut considérer qu'une méthode présente de bonnes performances en diagnostic d'un ensemble de défauts si elle est capable de détecter et isoler cet ensemble avec un FDR au moins égal à 80% ( $FDR \geq 80\%$ ) et un FAR qui ne dépasse pas les 0.5%

( $FAR \leq 0.5\%$ ). Comme évoqué précédemment, ce choix dépendra du système considéré.

En pratique, il est important de savoir comment comparer les deux matrices de confusion. Pour cela, nous proposons deux manières de procéder : soit en utilisant les connaissances de l'expert ou à travers les tests statistiques.

1. *Connaissances de l'expert* : La comparaison des matrices de confusion selon les connaissances de l'expert consiste simplement à choisir un seuil au-delà duquel on considère qu'une proportion est plus grande que l'autre. Par exemple, si  $x_{22}$  et  $y_{22}$  sont supérieurs au seuil fixé et que  $x_{22} - y_{22} > 10\%$  alors on considère que la méthode  $m_1$  est experte dans le diagnostic du défaut  $s_1$  et par conséquent  $s_1 \subseteq F_1$ . Par contre, si  $y_{22} - x_{22} > 10\%$  alors  $s_1 \subseteq F_2$ . Enfin, si  $|x_{22} - y_{22}| \leq 10\%$  alors  $s_1 \subseteq F_0$ .

Cependant, les différences observées entre les différentes proportions peuvent être le résultat de simples fluctuations statistiques, d'où l'intérêt d'avoir recours aux tests statistiques.

2. *Tests statistiques* : Les tests statistiques sont bien adaptés à ce genre de problème, et plus spécifiquement les tests de comparaison de proportions. L'objectif est de décider si l'écart observé entre deux proportions est statistiquement significatif. Soit  $p_1^{cfi}$  la proportion donnée par la méthode  $m_1$  d'observations correctement classées pour la classe de fonctionnement  $i$ , à comparer avec  $p_2^{cfi}$  donnée par la méthode  $m_2$ . On va donc tester l'hypothèse  $H_0$  contre l'hypothèse  $H_1$  :

$$\begin{cases} H_0 : p_1^{cfi} = p_2^{cfi} \\ H_1 : p_1^{cfi} > p_2^{cfi} \end{cases} \quad (3.3)$$

ou

$$\begin{cases} H_0 : p_1^{cfi} = p_2^{cfi} \\ H_1 : p_1^{cfi} < p_2^{cfi} \end{cases} \quad (3.4)$$

L'hypothèse nulle  $H_0$  correspond au cas où les deux proportions sont statistiquement égales pour la classe concernée. L'hypothèse alternée  $H_1$  correspond au cas où la proportion donnée par une des méthodes est statistiquement supérieure à l'autre.

La variable de test qui permet de prendre une décision au sujet de l'affirmation faite dans l'hypothèse nulle est la proportion  $P$ , qui représente le meilleur estimateur ponctuel de  $p$ , évaluée sur un échantillon de taille  $n$ . Dans ce cas,  $P$  suit une loi binomiale  $\mathcal{B}(n, p)$ , généralement approximée par une loi normale. La valeur de la statistique test est comparée à un seuil critique qui peut être choisi. Enfin, la décision de rejet éventuel de l'hypothèse nulle est prise en fonction de la valeur  $p$  : si  $p > p_{seuil1}$  ou  $p < p_{seuil2}$  on rejette  $H_0$  sinon on accepte  $H_0$ .

Quelque soit l'approche utilisée (connaissances de l'expert ou tests statistiques), nous pouvons faire correspondre chaque classe de fonctionnement  $CF = \{CFN, s_1, s_2, \dots, s_n\}$  à un type de classe  $\{CFN, F_0, F_1, F_2\}$ . En effet, si les deux proportions comparées ( $p_1^{cfi}; p_2^{cfi}$ ) sont considérées comme proches par les experts, ou comme statistiquement égales par les tests d'hypothèses, alors la classe correspondante est une classe de type 0 :  $cfi \subseteq F_0$ . Par ailleurs, s'il s'avère que  $p_1^{cfi} > p_2^{cfi}$  (respectivement  $p_1^{cfi} < p_2^{cfi}$ ) alors on peut conclure que la méthode  $m_1$  (respectivement la méthode  $m_2$ ) est experte dans cette classe :  $cfi \subseteq F_1$  (respectivement  $cfi \subseteq F_2$ ).

**Exemple** Afin d'illustrer la démarche méthodologique proposée, un exemple numérique est présenté tout au long de ce chapitre.

Dans cet exemple, deux méthodes de classification supervisée (AD linéaire et kNN ( $k = 100$ )) sont utilisées et les scénarios suivants sont traités :

- Échantillons 1 à 100 : Le système est en classe de CFN, caractérisée par une moyenne  $\mu_{CFN} = (2, 3)$  et une matrice de variance-covariance  $\Sigma_{CFN} = \begin{bmatrix} 2 & 3.6 \\ 3.6 & 8 \end{bmatrix}$ .
- Échantillons 101 à 200 : Occurrence du défaut  $s_1$ . Ce défaut représente un saut de moyenne et un changement de variance-covariance par rapport au fonctionnement normal :  $\mu_{s_1} = (-0.2, 6.5)$  et  $\Sigma_{s_1} = \begin{bmatrix} 2 & 0.1 \\ 0.1 & 1 \end{bmatrix}$ . Le défaut disparaît après l'échantillon 200.
- Échantillons 201 à 300 : Occurrence du défaut  $s_2$ . Ce défaut représente un saut de moyenne par rapport au fonctionnement normal :  $\mu_{s_2} = (8, 14)$  et  $\Sigma_{s_2} = \begin{bmatrix} 2 & 3.6 \\ 3.6 & 8 \end{bmatrix}$ . Le défaut disparaît après l'échantillon 300.
- Échantillons 301 à 400 : Occurrence du défaut  $s_3$ . Ce défaut représente un saut de moyenne et un changement de variance-covariance par rapport au fonctionnement normal :  $\mu_{s_3} = (2, -5.5)$  et  $\Sigma_{s_3} = \begin{bmatrix} 2 & 0.1 \\ 0.1 & 1 \end{bmatrix}$ . Le défaut disparaît après l'échantillon 400.

Les données de test sont illustrées dans la Figure 3.3.

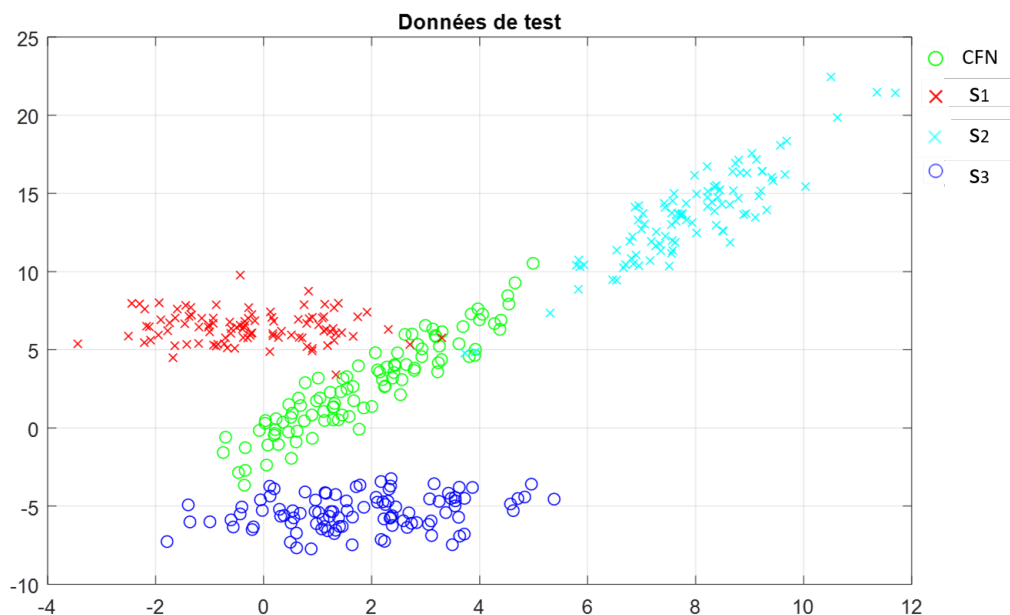


FIGURE 3.3 – Les données de test pour l'exemple numérique

Les résultats de la simulation sont donnés sous la forme de matrices de confusion (voir les tables 3.2a et 3.2b), permettant de visualiser la performance de chaque méthode FDD.

	<i>CFN</i>	$s_1$	$s_2$	$s_3$
<i>CFN</i>	85	15	0	0
$s_1$	0	100	0	0
$s_2$	0	0	100	0
$s_3$	98	2	0	0

(a) Matrice de confusion de la méthode  $m_1$ 

	<i>CFN</i>	$s_1$	$s_2$	$s_3$
<i>CFN</i>	80	20	0	0
$s_1$	0	100	0	0
$s_2$	95	5	0	0
$s_3$	0	0	0	100

(b) Matrice de confusion de la méthode  $m_2$ 

TABLE 3.2 – Evaluation des performances individuelles

- Méthode  $m_1$  : AD linéaire
  - Sur les 100 échantillons en classe de *CFN*, la méthode  $m_1$  en classe correctement 85. Le reste est considéré comme appartenant à la classe  $s_1$ . Le FAR est donc égal à  $FAR_{m_1} = \frac{15}{100} = 15\%$ .
  - Les 100 échantillons appartenant à la classe  $s_1$  sont correctement classés par la méthode  $m_1$ . Le FDR relatif à cette classe est de 100%.
  - La méthode  $m_1$  classe correctement tous les échantillons appartenant à la classe  $s_2$ . Le FDR relatif à cette classe est de 100%.
  - La méthode  $m_1$  présente de mauvaises performances pour la classe  $s_3$  (les données correspondants à cette classe ne lui ont pas été présentées). En effet, sur les 100 échantillons appartenant à cette classe, 98 sont considérés comme en classe de *CFN* et 2 sont considérés comme défaut de type  $s_1$ . Le FDR relatif à cette classe est de 0%.
- Méthode  $m_2$  :  $kNN$ 
  - Sur les 100 échantillons en classe de *CFN*, la méthode  $m_2$  en classe correctement 80. Le reste est considéré comme appartenant à la classe  $s_1$ . Le FAR est donc égal à  $FAR_{m_2} = \frac{20}{100} = 20\%$ .
  - Les 100 échantillons appartenant à la classe  $s_1$  sont correctement classés par la méthode  $m_2$ . Le FDR relatif à cette classe est de 100%.
  - La méthode  $m_2$  présente de mauvaises performances pour la classe  $s_2$  (les données correspondants à cette classe ne lui ont pas été présentées). En effet, sur les 100 échantillons appartenant à cette classe, 95 sont considérés comme en classe de *CFN* et 5 sont considérés comme défaut de type  $s_1$ . Le FDR relatif à cette classe est de 0%.
  - La méthode  $m_2$  classe correctement tous les échantillons appartenant à la classe  $s_3$ . Le FDR relatif à cette classe est de 100%.

En comparant les éléments sur la diagonale, il apparaît que les méthodes  $m_1$  et  $m_2$  sont expertes respectivement dans le diagnostic des défauts  $s_2$  et  $s_3$ . Par conséquent :  $F_1 = \{s_2\}$  et  $F_2 = \{s_3\}$ . Les deux méthodes sont expertes pour le défaut  $s_1$  avec un FDR de 100% donc  $F_0 = \{s_1\}$ .

Il est à noter que la différence entre les deux matrices de confusion est tellement évidente, sur cet exemple académique, que nous n'avons pas besoin d'utiliser les tests statistiques pour les comparer.

Ainsi, lorsque l'ensemble des défauts  $\{F_0, F_1, F_2\}$  affectant le système est pris en compte, les FDRs sont respectivement de :



$$\begin{cases} FDR_{m_{1g}} = \frac{100 + 100 + 0}{300} = 66.67\% \\ FDR_{m_{2g}} = \frac{100 + 0 + 100}{300} = 66.67\% \end{cases} \quad (3.5)$$

Par conséquent, nous sommes bien dans une situation où les méthodes individuelles ont une bonne performance pour un ensemble de défauts mais présentent une mauvaise performance globale, ce qui est le contexte approprié pour notre méthodologie.

### 3.3.2 Stratégie de FDD : Définition des objectifs

Afin d'améliorer la performance en FDD du système, la décision globale, notée  $d$ , devrait respecter certains objectifs. En effet, elle doit être construite en fonction des décisions individuelles  $d_1$  et  $d_2$  données respectivement par les méthodes  $m_1$  et  $m_2$ . L'idée sous-jacente consiste à utiliser les performances a priori de chaque méthode FDD pour lui attribuer un poids dans la décision globale.

Les objectifs à respecter sont nombreux et dépendent de la stratégie adoptée et des performances souhaitées par chaque utilisateur. Nous allons présenter tout d'abord l'ensemble des objectifs possibles puis nous allons faire des choix d'objectifs pour la suite des développements.

#### Les objectifs possibles

Etant donné les décisions individuelles  $d_1$  et  $d_2$  de chacune des méthodes à propos de l'état du système, la décision globale  $d$  doit être choisie de telle sorte que les performances globales du système en FDD soient améliorées. C'est ce qui est appelé l'objectif à respecter par la décision globale  $d$ .

L'ensemble des objectifs possibles est résumé dans la Table 3.3.

$d_2 \backslash d_1$	$CFN$	$F_0$	$F_1$
$CFN$	1. $CFN$	4. $CFN; F_0$	7. $F_1$
$F_0$	2. $CFN; F_0$	5. $F_0$	8. $F_1; F_1 \cap F_0$
$F_2$	3. $F_2$	6. $F_2; F_2 \cap F_0$	9. $F_1 \cap F_2$

TABLE 3.3 – Table des objectifs possibles

La table doit être lue comme suit : les colonnes représentent la décision  $d_1$  donnée par la méthode  $m_1$  tandis que les lignes montrent la décision  $d_2$  obtenue par la méthode  $m_2$ . Les cases du tableau correspondent aux résultats attendus lors de la combinaison des deux décisions, c'est-à-dire la décision globale  $d$ . La couleur rouge indique que l'objectif est unique et qu'il n'y a pas de choix à faire.

Dans ce qui suit, une explication détaillée de chaque objectif est fournie.

**Objectif 1** : Si  $d_1 = \{CFN\}$  et  $d_2 = \{CFN\}$ , alors la décision globale est  $d = \{CFN\}$ .

En effet, si les deux méthodes individuelles décident séparément que le système est en  $CFN$ , alors la décision globale ne peut qu'être  $CFN$ .

**Objectif 2** : Si  $d_1 = \{CFN\}$  et  $d_2 = \{F_0\}$ , alors la décision globale n'est pas unique et les deux possibilités sont  $d = \{CFN\}$  ou  $d = \{F_0\}$ .

Seule la méthode  $m_2$  détecte l'apparition d'un défaut de type 0. Or, la méthode  $m_1$  présente également de bonnes performances dans la détection et le diagnostic de ce type de défauts, puisqu'il s'agit d'un défaut commun aux deux méthodes de surveillance. Par conséquent, le fait qu'une seule méthode détecte un défaut commun conduirait à choisir la classe nominale  $CFN$ . Par ce choix, nous considérons que la décision donnée par la méthode  $m_2$  est certainement une fausse alarme puisque la méthode  $m_1$  n'a pas confirmé l'occurrence du défaut commun. Ce choix pourrait donc diminuer le taux de fausses alarmes global du système.

Par ailleurs, l'utilisateur peut tout aussi bien choisir de déclencher une alarme à chaque fois qu'une méthode détecte un défaut commun, en considérant que la décision globale est :  $d = \{F_0\}$ . Dans ce cas, nous considérons que la décision donnée par la méthode  $m_1$  est certainement une non-détection puisque la méthode  $m_2$  a réussi à détecter l'occurrence du défaut commun. Ce choix pourrait donc augmenter le taux de diagnostic global de défauts du système.

Par conséquent, l'objectif peut être adapté aux besoins de l'utilisateur et à la stratégie de surveillance qu'il adopte puisqu'on peut faire le choix de privilégier le FAR ou alors le FDR.

**Objectif 3** : Si  $d_1 = \{CFN\}$  et  $d_2 = \{F_2\}$ , alors la décision globale est  $d = \{F_2\}$ .

En effet, la méthode  $m_2$  est experte dans le diagnostic des défauts de type 2 contrairement à la méthode  $m_1$ . Cette dernière n'est pas toujours capable de détecter l'occurrence de ce défaut, ce qui se traduit par sa décision à propos de l'état du système qu'elle qualifie comme étant normal. La décision globale privilégie donc la méthode  $m_2$ . Ce choix pourrait augmenter le FDR du système.

**Objectif 4** : Si  $d_1 = \{F_0\}$  et  $d_2 = \{CFN\}$ , alors la décision globale n'est pas unique et les deux possibilités sont  $d = \{CFN\}$  ou  $d = \{F_0\}$ .

Le raisonnement est identique à celui énoncé dans l'Objectif 2.

**Objectif 5** : Si  $d_1 = \{F_0\}$  et  $d_2 = \{F_0\}$ , alors la décision globale est  $d = \{F_0\}$ .

En effet, si les deux méthodes individuelles détectent l'occurrence du défaut commun de type  $F_0$ , alors la décision globale ne peut pas être différente.

**Objectif 6** : Si  $d_1 = \{F_0\}$  et  $d_2 = \{F_2\}$ , alors la décision globale n'est pas unique et les deux possibilités sont  $d = \{F_2\}$  ou  $d = \{F_0 \cap F_2\}$ .

Ici, les deux méthodes sont expertes dans le diagnostic des défauts de type  $F_0$ . Par conséquent, le fait que seule la méthode  $m_1$  détecte ce défaut commun conduit soit à ignorer ce défaut puisqu'il n'a pas été confirmé par la méthode  $m_2$  soit à le choisir, comme dans le cas de l'Objectif 2. Par ailleurs, seule  $m_2$  peut détecter un défaut de type  $F_2$ , ce qui conduit à la décision  $d = \{F_2\}$ . Il est à noter que l'utilisateur peut également choisir de considérer que  $F_0$  et  $F_2$  se sont produits simultanément et donc la décision globale pourrait être :  $d = \{F_0 \cap F_2\}$ . En effet, la méthodologie proposée n'exclut pas l'occurrence de défauts multiples (voir **Hypothèse 2**).

**Objectif 7** : Si  $d_1 = \{F_1\}$  et  $d_2 = \{CFN\}$ , alors la décision globale est  $d = \{F_1\}$ .

Le raisonnement est identique à celui énoncé dans l'Objectif 3.

**Objectif 8** : Si  $d_1 = \{F_1\}$  et  $d_2 = \{F_0\}$ , alors la décision globale n'est pas unique et les deux possibilités sont  $d = \{F_1\}$  ou  $d = \{F_1 \cap F_0\}$ .

Le raisonnement est identique à celui énoncé dans l'Objectif 6.

**Objectif 9** : Si  $d_1 = \{F_1\}$  et  $d_2 = \{F_2\}$ , alors la décision globale est :  $d = \{F_1 \cap F_2\}$ .

Les deux méthodes ont détecté des défauts pour lesquels elles sont expertes. Par conséquent, la décision globale doit tenir compte de cet aspect et signaler la présence des deux défauts.

Il est à noter qu'une symétrie existe dans la Table des objectifs 3.3. En effet, les objectifs (4; 7; 8) suivent le même raisonnement que (2; 3; 6). Par conséquent, le nombre d'objectifs différents à développer se résume à 6 au lieu de 9.

### Les objectifs choisis

Dans la suite, nous allons choisir à titre d'exemple des objectifs stratégiques comme illustré dans la Table 3.4.

$d_2 \backslash d_1$	$CFN$	$F_0$	$F_1$
$CFN$	1. $CFN$	4. $CFN$	7. $F_1$
$F_0$	2. $CFN$	5. $F_0$	8. $F_1$
$F_2$	3. $F_2$	6. $F_2$	9. $F_1 \cap F_2$

TABLE 3.4 – Table des objectifs choisis

Par la suite on montrera que si les 9 objectifs choisis de la Table 3.4 sont respectés, alors la performance globale (en termes de FDR et de FAR) du système de surveillance sera améliorée.

En effet, une décision qui respecte les objectifs 1 et 5 est plus fiable puisque chaque méthode individuelle renforce le résultat de l'autre. Les objectifs 2; 4; 6 et 8 permettent de réduire le FAR car un défaut commun de type  $F_0$  détectable par les deux méthodes est considéré comme la décision globale si et seulement si les deux méthodes le détectent. En ce qui concerne les objectifs 3 et 7, le fait qu'une méthode puisse détecter avec des performances élevées un défaut spécifique contrairement à l'autre méthode, est exploité. Enfin, l'objectif 9 permet d'aborder le cas de défauts multiples, ce qui est très intéressant dans la surveillance des systèmes complexes.

Par conséquent, la décision globale va couvrir plus de défauts ( $D = \{F_0, F_1, F_2\}$ ) que les méthodes utilisées séparément ( $D_1 = \{F_0, F_1\}$  et  $D_2 = \{F_0, F_2\}$ ), assurant ainsi une amélioration du FDR. De plus, un défaut commun aux deux méthodes ( $F_0$ ) est choisi par la décision globale seulement si les deux méthodes l'ont détecté, ce qui peut diminuer le FAR.

### 3.3.3 Stratégie de FDD : Cadre générique de fusion

Dans l'approche proposée, le RB est utilisé d'une part comme cadre de fusion des décisions individuelles données par les méthodes FDD et d'autre part comme outil de prise de décision. En effet, le RB va permettre de décider dans quelle classe de fonctionnement se trouve le système à un instant donné.

Le cadre basé sur le RB permet la prise en considération des incertitudes puisqu'il utilise des probabilités à la place des booléens pour quantifier l'occurrence des défauts. En particulier, il fournit un outil efficace de combinaison de différents types de décisions provenant de diverses méthodes, avec une forte capacité à gérer des décisions incomplètes ou même contradictoires.

Dans la suite, la structure adoptée pour le RB ainsi que ses paramètres sont présentés.

#### Structure

La structure de RB adaptée à notre problématique de fusion de décisions est la structure de classifieur naïf, comme illustré par la Figure 3.4. En effet, le classifieur Bayésien modélise l'appartenance à une classe par un noeud discret à  $k$  modalités, où  $k$  représente le nombre de classes de fonctionnement considéré dans le problème. Les autres noeuds représentent quant à eux les variables descriptives du problème.

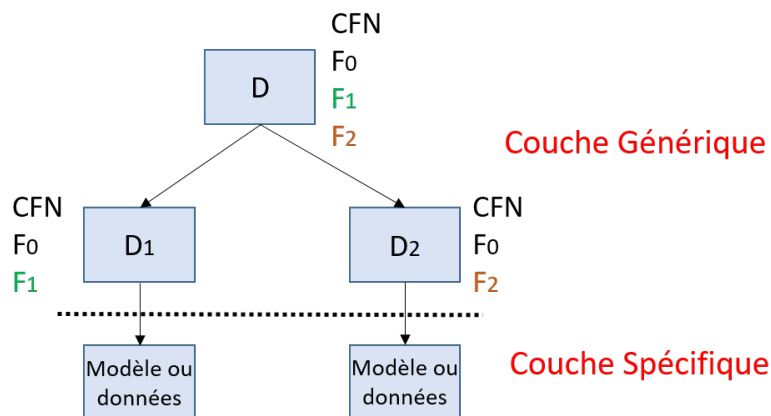


FIGURE 3.4 – Structure du RB

Il s'agit donc d'une structure d'aide à la prise de décision ayant deux couches :

- Une couche générique qui comporte un noeud de décision globale  $D$  et un noeud pour chaque décision individuelle :  $D_1$  et  $D_2$ ,
- Une couche spécifique liée à chaque méthode de surveillance, composée d'indicateurs de défaut pour les méthodes basées sur un modèle, par exemple, et de données pour les méthodes pilotées par les données.

**Propriété 1.** *Le cadre de fusion est générique.*

Il comprend 3 noeuds discrets :

- Le noeud  $D_1$  représentant la décision de  $m_1$ . Ce noeud a 3 types de modalités  $\{CFN; F_0; F_1\}$ .
- Le noeud  $D_2$  représentant la décision de  $m_2$ . Ce noeud a 3 types de modalités  $\{CFN; F_0; F_2\}$ .

- Le noeud  $D$  représentant la décision globale après la fusion. Il peut avoir comme modalités tous les états des deux noeuds de décisions individuelles, soit 4 types de modalités  $\{CFN; F_0; F_1; F_2\}$ .

Chaque noeud de décision individuelle est lié au noeud de décision globale par un arc. Par conséquent, l'inférence basée sur le RB consiste à calculer les probabilités a posteriori de chaque modalité du noeud  $D$  étant donné les probabilités des états des noeuds  $D_1$  et  $D_2$ .

### Paramètres

Pour exploiter le RB de la Figure 3.4, il convient de préciser les probabilités a priori du noeud parent  $D$  (voir la Table 3.5) ainsi que les probabilités conditionnelles des noeuds enfants (voir la Table 3.6).

$D$			
$CFN$	$F_0$	$F_1$	$F_2$
$\alpha_0$	$\alpha_1$	$\alpha_2$	$\alpha_3$

TABLE 3.5 – Table de probabilités a priori du noeud  $D$

$D \backslash D_1$	$CFN$	$F_0$	$F_1$
$CFN$	$a_0$	$a_1$	$a_2$
$F_0$	$b_0$	$b_1$	$b_2$
$F_1$	$c_0$	$c_1$	$c_2$
$F_2$	$d_0$	$d_1$	$d_2$

(a) Table du noeud  $D_1$

$D \backslash D_2$	$CFN$	$F_0$	$F_2$
$CFN$	$e_0$	$e_1$	$e_2$
$F_0$	$f_0$	$f_1$	$f_2$
$F_1$	$g_0$	$g_1$	$g_2$
$F_2$	$h_0$	$h_1$	$h_2$

(b) Table du noeud  $D_2$

TABLE 3.6 – Tables de probabilités conditionnelles

**Définition 16.** La **probabilité a priori** d'un état représente la fréquence normalisée de cet état.

Ici, nous supposons que la connaissance préalable de la fréquence d'occurrence d'un défaut n'est pas disponible, ce qui est généralement le cas en pratique. D'où l'hypothèse 3.

**Hypothèse 3.** Toutes les probabilités a priori peuvent être supposées égales :  $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3$

**Définition 17.** Une table de **probabilités conditionnelles** associée à un noeud permet de quantifier l'effet du noeud parent sur ce noeud. Elle décrit les probabilités associées au noeud enfant suivant les différents états du noeud parent.

Les probabilités conditionnelles sont généralement estimées à partir des données de défauts disponibles. Plusieurs algorithmes d'apprentissage de paramètres existent et sont utilisés selon que les données soient complètes ou manquantes. On peut citer l'apprentissage statistique qui consiste à estimer la probabilité d'un événement par sa fréquence d'apparition dans la base de données ou encore l'apprentissage bayésien qui permet l'obtention des paramètres les plus probables sachant que les données ont été

observées, en utilisant des lois a priori sur les paramètres.

Quelquefois, très peu de données sont disponibles. Les probabilités conditionnelles du RB sont alors estimées de façon empirique en utilisant les connaissances d'experts : ceci est appelé l'élicitation de probabilités [Renooij, 2001].

Dans ce travail, nous allons nous écarter de l'application et de ses données pour conserver un caractère générique. Par conséquent, nous proposons une méthodologie nouvelle et originale qui indique comment déterminer théoriquement les valeurs des paramètres du RB afin d'améliorer les performances de surveillance du système considéré.

### 3.3.4 Stratégie de FDD : Développement des objectifs

La méthodologie proposée a pour but de déterminer l'ensemble des probabilités a priori et conditionnelles possibles (voir Tables 3.5 ; 3.6a et 3.6b), qui font en sorte de remplir les objectifs énoncés dans la Table des objectifs choisis 3.4.

#### Hypothèses et propriétés

Pour cela, nous supposons que, étant donné une décision globale sur le fait que le système se trouve dans l'une des modalités communes aux deux méthodes ( $CFN$  et  $F_0$ ), les probabilités conditionnelles des méthodes individuelles (voir Tables 3.6a et 3.6b) d'être dans ces modalités communes sont les mêmes, comme explicité dans les équations 3.6, 3.7.

$$\begin{cases} P(D_1 = CFN/D = CFN) = P(D_2 = CFN/D = CFN) & \Rightarrow a_0 = e_0 \\ P(D_1 = F_0/D = CFN) = P(D_2 = F_0/D = CFN) & \Rightarrow a_1 = e_1 \end{cases} \quad (3.6)$$

$$\begin{cases} P(D_1 = CFN/D = F_0) = P(D_2 = CFN/D = F_0) & \Rightarrow b_0 = f_0 \\ P(D_1 = F_0/D = F_0) = P(D_2 = F_0/D = F_0) & \Rightarrow b_1 = f_1 \end{cases} \quad (3.7)$$

De plus, si la décision globale est l'une des modalités communes ( $CFN$  et  $F_0$ ), on ne privilégie pas une méthode par rapport à l'autre et on considère donc que les probabilités conditionnelles sont égales, comme décrit dans les équations 3.8.

$$\begin{cases} P(D_1 = F_1/D = CFN) = P(D_2 = F_2/D = CFN) & \Rightarrow a_2 = e_2 \\ P(D_1 = F_1/D = F_0) = P(D_2 = F_2/D = F_0) & \Rightarrow b_2 = f_2 \end{cases} \quad (3.8)$$

Il est à noter qu'en pratique, l'utilisateur peut s'appuyer beaucoup plus sur les décisions d'une méthode particulière et donc choisir  $a_2 \neq e_2$  et  $b_2 \neq f_2$ . Cela montre clairement la souplesse de l'approche proposée, qui offre également à l'utilisateur la possibilité d'ajouter des informations provenant des experts.

Enfin, puisque les méthodes  $m_1$  et  $m_2$  sont respectivement expertes dans la détection de défauts de type  $F_1$  et  $F_2$ , on peut considérer que :

$$\begin{cases} P(D_1 = CFN/D = F_1) = P(D_2 = CFN/D = F_2) & \Rightarrow c_0 = h_0 \\ P(D_1 = F_0/D = F_1) = P(D_2 = F_0/D = F_2) & \Rightarrow c_1 = h_1 \\ P(D_1 = F_1/D = F_1) = P(D_2 = F_2/D = F_2) & \Rightarrow c_2 = h_2 \end{cases} \quad (3.9)$$

$$\begin{cases} P(D_1 = CFN/D = F_2) = P(D_2 = CFN/D = F_1) & \Rightarrow d_0 = g_0 \\ P(D_1 = F_0/D = F_2) = P(D_2 = F_0/D = F_1) & \Rightarrow d_1 = g_1 \\ P(D_1 = F_1/D = F_2) = P(D_2 = F_2/D = F_1) & \Rightarrow d_2 = g_2 \end{cases} \quad (3.10)$$

Ainsi, les probabilités conditionnelles respectent les égalités suivantes :

$$\begin{cases} a_0 = e_0; a_1 = e_1; a_2 = e_2 \\ b_0 = f_0; b_1 = f_1; b_2 = f_2 \\ c_0 = h_0; c_1 = h_1; c_2 = h_2 \\ d_0 = g_0; d_1 = g_1; d_2 = g_2 \end{cases} \quad (3.11)$$

Par conséquent, seulement 12 probabilités conditionnelles ( $\{a_0, a_1, a_2, \dots, d_0, d_1, d_2\}$ ) doivent être déterminées.

En outre, les évènements dont les probabilités sont sur une ligne de la table de probabilités conditionnelles forment une partition de l'évènement  $\{D = d_i\}$  avec  $d_i = \{CFN, F_0, F_1, F_2\}$ . La somme de leurs probabilités vaut donc 1. On déduit de cette propriété les contraintes suivantes :

### Propriété 2.

$$\begin{cases} P(D_1 = CFN/D = CFN) + P(D_1 = F_0/D = CFN) + P(D_1 = F_1/D = CFN) = 1 \\ P(D_1 = CFN/D = F_0) + P(D_1 = F_0/D = F_0) + P(D_1 = F_1/D = F_0) = 1 \\ P(D_1 = CFN/D = F_1) + P(D_1 = F_0/D = F_1) + P(D_1 = F_1/D = F_1) = 1 \\ P(D_1 = CFN/D = F_2) + P(D_1 = F_0/D = F_2) + P(D_1 = F_1/D = F_2) = 1 \end{cases} \quad (3.12)$$

D'où :

$$\begin{cases} a_0 + a_1 + a_2 = 1 \\ b_0 + b_1 + b_2 = 1 \\ c_0 + c_1 + c_2 = 1 \\ d_0 + d_1 + d_2 = 1 \end{cases} \quad (3.13)$$

### Règle de décision

La prise de décisions avec un RB implique le choix d'une règle de décision. Généralement, on distingue deux règles pouvant être utilisées pour décider de l'état du système. Ainsi, une classe de fonctionnement est choisie comme décision  $d$  si :

1. sa probabilité est la plus grande et dépasse un certain seuil représentant généralement la probabilité a priori de cette modalité.
2. sa probabilité est la plus grande et la différence entre cette probabilité et la seconde est supérieure à un certain seuil.

Dans ce travail, toutes les décisions sont prises en utilisant la première règle. On en déduit donc la propriété suivante :

**Propriété 3.** *La probabilité a posteriori de la décision globale  $d$  d'être dans une modalité étant donné les décisions individuelles  $d_1$  et  $d_2$  doit être la plus grande et doit dépasser la probabilité a priori de cette modalité.*

Prenons l'exemple de la classe de fonctionnement  $CFN$ . La propriété se traduit par les équations suivantes 3.14 :

$$D \in CFN \text{ si : } \begin{cases} P(D = CFN/D_1 \cap D_2) > P(D = F_i/D_1 \cap D_2), i \in \{0, 1, 2\} \\ P(D = CFN/D_1 \cap D_2) > P(D = CFN) \end{cases} \quad (3.14)$$

### Objectifs

Dans cette partie, seul l'objectif 1 est développé en détail. Les développements concernant les autres objectifs peuvent être trouvés dans l'Annexe 5.2.

**Objectif 1 :** L'objectif 1 indique que si  $d_1 = \{CFN\}$  et  $d_2 = \{CFN\}$ , alors la décision globale doit toujours être :  $d = \{CFN\}$ . Afin d'assurer cet objectif et selon la règle de décision choisie, la probabilité a posteriori de la modalité  $CFN$  compte tenu des décisions individuelles ( $d_1 = \{CFN\}$  et  $d_2 = \{CFN\}$ ) doit être la probabilité la plus grande et doit dépasser la probabilité a priori de la modalité  $CFN$ .

Par conséquent, les critères suivants (3.15) doivent être respectés :

$$\begin{cases} P(D = CFN/D_1 = CFN \cap D_2 = CFN) > P(D = CFN) \\ P(D = CFN/D_1 = CFN \cap D_2 = CFN) > P(D = F_i/D_1 = CFN \cap D_2 = CFN), \\ i \in \{0, 1, 2\} \end{cases} \quad (3.15)$$

- Critère 1 : La première inégalité est développée comme il suit :

$$\begin{aligned} & P(D = CFN/D_1 = CFN \cap D_2 = CFN) > P(D = CFN) \\ \Leftrightarrow & \frac{P(D_1=CFN/D_2=CFN \cap D=CFN)P(D_2=CFN/D=CFN)P(D=CFN)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} > P(D = CFN) \\ \Leftrightarrow & \frac{P(D_1=CFN/D=CFN)P(D_2=CFN/D=CFN)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} > 1 \\ \Leftrightarrow & \frac{a_0^2}{[\sum P(D_1=CFN/D_2=CFN \cap D=d_i)P(D=d_i)][\sum P(D_2=CFN/D=d_i)P(D=d_i)]} > 1 \\ \Leftrightarrow & \frac{a_0^2}{\alpha_0^2(a_0+b_0+c_0+d_0)^2} > 1 \\ \Leftrightarrow & \frac{a_0}{\alpha_0(a_0+b_0+c_0+d_0)} > 1 \end{aligned} \quad (3.16)$$

La première étape fait appel à la loi de Bayes pour développer les probabilités conditionnelles comme il suit :

$$P(D/D_1 \cap D_2) = \frac{P(D_1/D_2 \cap D)P(D_2/D)P(D)}{P(D_1/D_2)P(D_2)} \quad (3.17)$$

Ensuite, la propriété d'indépendance conditionnelle des noeuds  $D_1$  et  $D_2$  étant donné le noeud racine  $D$  est appliquée pour simplifier les équations :

$$P(D_1/D_2 \cap D) = P(D_1/D) \quad (3.18)$$

Enfin, la marginalisation sur les états du noeud  $D$  est utilisée :



$$P(D_1/D_2)P(D_2) = \sum_d P(D_1/D_2 \cap d_i)P(d_i) \sum_d P(D_2/d_i)P(d_i) \quad (3.19)$$

- Critère 2 : La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$ , on obtient :

$$\begin{aligned} P(D = CFN/D_1 = CFN \cap D_2 = CFN) &> P(D = F_0/D_1 = CFN \cap D_2 = CFN) \\ \Leftrightarrow \frac{P(D_1=CFN/D_2=CFN \cap D=CFN)P(D_2=CFN/D=CFN)P(D=CFN)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} &> \\ \frac{P(D_1=CFN/D_2=CFN \cap D=F_0)P(D_2=CFN/D=F_0)P(D=F_0)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} & \\ \Leftrightarrow \frac{P(D_1=CFN/D=CFN)P(D_2=CFN/D=CFN)P(D=CFN)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} &> \\ \frac{P(D_1=CFN/D=F_0)P(D_2=CFN/D=F_0)P(D=F_0)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} & \\ \Leftrightarrow \frac{a_0^2}{\alpha_0^2(a_0+b_0+c_0+d_0)^2} > \frac{b_0^2}{\alpha_0^2(a_0+b_0+c_0+d_0)^2} & \\ \Leftrightarrow a_0 > b_0 & \end{aligned} \quad (3.20)$$

Pour  $D = F_1$ , on obtient :

$$\begin{aligned} P(D = CFN/D_1 = CFN \cap D_2 = CFN) &> P(D = F_1/D_1 = CFN \cap D_2 = CFN) \\ \Leftrightarrow \frac{P(D_1=CFN/D_2=CFN \cap D=CFN)P(D_2=CFN/D=CFN)P(D=CFN)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} &> \\ \frac{P(D_1=CFN/D_2=CFN \cap D=F_1)P(D_2=CFN/D=F_1)P(D=F_1)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} & \\ \Leftrightarrow \frac{P(D_1=CFN/D=CFN)P(D_2=CFN/D=CFN)P(D=CFN)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} &> \\ \frac{P(D_1=CFN/D=F_1)P(D_2=CFN/D=F_1)P(D=F_1)}{P(D_1=CFN/D_2=CFN)P(D_2=CFN)} & \\ \Leftrightarrow \frac{a_0^2}{\alpha_0^2(a_0+b_0+c_0+d_0)^2} > \frac{c_0d_0}{\alpha_0^2(a_0+b_0+c_0+d_0)^2} & \\ \Leftrightarrow a_0^2 > c_0d_0 & \end{aligned} \quad (3.21)$$

Enfin pour  $D = F_2$ , on obtient la même contrainte :  $a_0^2 > c_0d_0$

Par conséquent, l'objectif 1 peut être atteint si 3 contraintes (équations 3.22, 3.23) impliquant 4 probabilités conditionnelles sont respectées :

$$\frac{a_0}{a_0 + b_0 + c_0 + d_0} > \alpha_0 \quad (3.22)$$

$$\begin{cases} a_0 > b_0 \\ a_0^2 > c_0d_0 \end{cases} \quad (3.23)$$

**Objectif 2 :** L'objectif 2 indique que si  $d_1 = \{CFN\}$  et  $d_2 = \{F_0\}$ , alors la décision globale doit toujours être :  $d = \{CFN\}$ .

Comme pour l'Objectif 1, les critères sont développés (voir Annexe 5.2) et 3 contraintes (équations 3.24, 3.25) sont obtenues.

$$\frac{a_0a_1}{(a_0 + b_0 + c_0 + d_0)(a_1 + b_1 + c_1 + d_1)} > \alpha_0^2 \quad (3.24)$$

$$\begin{cases} a_0a_1 > b_0b_1 \\ a_0a_1 > c_0d_1 \end{cases} \quad (3.25)$$

**Objectif 3 :** L'objectif 3 indique que si  $d_1 = \{CFN\}$  et  $d_2 = \{F_2\}$ , alors la décision globale doit toujours être :  $d = \{F_2\}$ .

Par conséquent, les contraintes suivantes (équations 3.26, 3.27) doivent être respectées :

$$\frac{d_0 c_2}{(a_0 + b_0 + c_0 + d_0)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \quad (3.26)$$

$$\begin{cases} d_0 c_2 > a_0 a_2 \\ d_0 c_2 > b_0 b_2 \\ d_0 c_2 > c_0 d_2 \end{cases} \quad (3.27)$$

**Objectif 4 :** L'objectif 4 indique que si  $d_1 = \{F_0\}$  et  $d_2 = \{CFN\}$ , alors la décision globale doit toujours être :  $d = \{CFN\}$ .

Le raisonnement est identique à celui de l'Objectif 2 et les mêmes contraintes sont déduites.

$$\frac{a_0 a_1}{(a_0 + b_0 + c_0 + d_0)(a_1 + b_1 + c_1 + d_1)} > \alpha_0^2 \quad (3.28)$$

$$\begin{cases} a_0 a_1 > b_0 b_1 \\ a_0 a_1 > c_0 d_1 \end{cases} \quad (3.29)$$

**Objectif 5 :** L'objectif 5 indique que si  $d_1 = \{F_0\}$  et  $d_2 = \{F_0\}$ , alors la décision globale doit toujours être :  $d = \{F_0\}$ .

Par conséquent, les contraintes suivantes (équations 3.30, 3.31) doivent être respectées :

$$\frac{b_1}{a_1 + b_1 + c_1 + d_1} > \alpha_0 \quad (3.30)$$

$$\begin{cases} b_1 > a_1 \\ b_1^2 > c_1 d_1 \end{cases} \quad (3.31)$$

**Objectif 6 :** L'objectif 6 indique que si  $d_1 = \{F_0\}$  et  $d_2 = \{F_2\}$ , alors la décision globale doit toujours être :  $d = \{F_2\}$ .

Par conséquent, les contraintes suivantes (équations 3.32, 3.33) doivent être respectées :

$$\frac{d_1 c_2}{(a_1 + b_1 + c_1 + d_1)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \quad (3.32)$$

$$\begin{cases} d_1 c_2 > a_1 a_2 \\ d_1 c_2 > b_1 b_2 \\ d_1 c_2 > c_1 d_2 \end{cases} \quad (3.33)$$

**Objectif 7 :** L'objectif 7 indique que si  $d_1 = \{F_1\}$  et  $d_2 = \{CFN\}$ , alors la décision globale doit toujours être :  $d = \{F_1\}$ .

Le raisonnement est identique à celui de l'Objectif 3 et les mêmes contraintes sont déduites.

$$\frac{d_0 c_2}{(a_0 + b_0 + c_0 + d_0)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \quad (3.34)$$

$$\begin{cases} d_0 c_2 > a_0 a_2 \\ d_0 c_2 > b_0 b_2 \\ d_0 c_2 > c_0 d_2 \end{cases} \quad (3.35)$$

**Objectif 8 :** L'objectif 8 indique que si  $d_1 = \{F_1\}$  et  $d_2 = \{F_0\}$ , alors la décision globale doit toujours être :  $d = \{F_1\}$ .

Le raisonnement est identique à celui de l'Objectif 6 et les mêmes contraintes sont déduites.

$$\frac{d_1 c_2}{(a_1 + b_1 + c_1 + d_1)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \quad (3.36)$$

$$\begin{cases} d_1 c_2 > a_1 a_2 \\ d_1 c_2 > b_1 b_2 \\ d_1 c_2 > c_1 d_2 \end{cases} \quad (3.37)$$

**Objectif 9 :** L'objectif 9 indique que si  $d_1 = \{F_1\}$  et  $d_2 = \{F_2\}$ , alors la décision globale doit toujours être :  $d = \{F_1 \cap F_2\}$ .

Les contraintes suivantes (équations 3.38, 3.39) doivent donc être respectées :

$$\frac{c_2 d_2}{(a_2 + b_2 + c_2 + d_2)^2} > \alpha_0^2 \quad (3.38)$$

$$\begin{cases} c_2 d_2 > a_2^2 \\ c_2 d_2 > b_2^2 \end{cases} \quad (3.39)$$

Il convient de préciser que le RB tel que présenté dans la partie 3.3.3 ne gère pas les défauts multiples. En effet, une classe de défaut multiple n'apparaît pas dans la modalité du noeud  $D$ . Par ailleurs, notre méthodologie prend en compte ce cas en assurant que le RB va renvoyer des probabilités égales et supérieures aux autres pour les deux modalités concernées par le défaut. Cette égalité de probabilités pour les modalités  $F_1$  et  $F_2$  est interprétée comme une occurrence de défauts multiples.

Par conséquent, la méthodologie proposée part d'un ensemble d'objectifs de diagnostic à atteindre. De ces objectifs découlent des critères qui nous amènent à imposer des contraintes sur les probabilités conditionnelles du RB. Si ces contraintes sont satisfaites, les différents objectifs sont atteints et la performance globale sera améliorée. A ce stade, le but est de résoudre ce problème multi-objectif en trouvant toutes les valeurs possibles des probabilités conditionnelles (c'est-à-dire toutes les solutions du problème).

### 3.3.5 Stratégie de FDD : Prise de décision multi-objectifs sous contraintes

#### Problème multi-objectifs sous contraintes

Un problème multi-objectifs cherche à résoudre simultanément plusieurs objectifs d'un même problème.

De manière formelle, le problème multi-objectif rencontré ici peut s'écrire comme suit :

**Problème multi-objectifs.** Trouver  $x \in \Delta$  tel que  $R(x) = [r_1(x); r_2(x); r_3(x); r_4(x); r_5(x); r_6(x)] > 0$  sous les contraintes  $P(x) = [p_1(x); \dots; p_{14}(x)] > 0$ .

avec  $x = [a_0, b_0, c_0, d_0, a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2]^T$  le vecteur des solutions réalisables (probabilités conditionnelles des noeuds enfants  $D_1$  et  $D_2$ ),  $\Delta$  le domaine de définition des solutions  $x$ ,  $R(x)$  les fonctions objectifs déduites précédemment (correspondant aux équations 3.22, 3.30, 3.24, 3.26, 3.32, 3.38 du premier critère à respecter, transformées en équations aux différences).

$$R(x) \left\{ \begin{array}{l} r_1(x) = \frac{a_0}{a_0+b_0+c_0+d_0} - \alpha_0 \\ r_2(x) = \frac{b_1}{a_1+b_1+c_1+d_1} - \alpha_0 \\ r_3(x) = \frac{a_0 a_1}{(a_0+b_0+c_0+d_0)(a_1+b_1+c_1+d_1)} - \alpha_0^2 \\ r_4(x) = \frac{d_0 c_2}{(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} - \alpha_0^2 \\ r_5(x) = \frac{d_1 c_2}{(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} - \alpha_0^2 \\ r_6(x) = \frac{c_2 d_2}{(a_2+b_2+c_2+d_2)^2} - \alpha_0^2 \end{array} \right. \quad (3.40)$$

et  $P(x)$  l'ensemble de contraintes (correspondant aux équations 3.23, 3.31, 3.25, 3.27, 3.33, 3.39, du second critère à respecter, transformées en équations aux différences)

$$P(x) \left\{ \begin{array}{l} p_1(x) = a_0 - b_0; \quad p_2(x) = a_0^2 - c_0 d_0; \\ p_3(x) = b_1 - a_1; \quad p_4(x) = b_1^2 - c_1 d_1; \\ p_5(x) = a_0 a_1 - b_0 b_1; \quad p_6(x) = a_0 a_1 - c_0 d_1; \\ p_7(x) = d_0 c_2 - a_0 a_2; \quad p_8(x) = d_0 c_2 - b_0 b_2; \quad p_9(x) = d_0 c_2 - c_0 d_2 \\ p_{10}(x) = d_1 c_2 - a_1 a_2; \quad p_{11}(x) = d_1 c_2 - b_1 b_2; \quad p_{12}(x) = d_1 c_2 - c_1 d_2 \\ p_{13}(x) = c_2 d_2 - a_2^2; \quad p_{14}(x) = c_2 d_2 - b_2^2 \end{array} \right. \quad (3.41)$$

**Définition 18.** *L'espace de recherche, représenté par le vecteur  $x$ , désigne l'ensemble des solutions potentielles du problème. En faisant varier ce vecteur, on recherche une solution de la fonction  $r$  qui respecte les contraintes  $p$ .*

L'objectif est de trouver l'ensemble discret de solutions  $x$  respectant les fonctions objectifs  $R(x)$  sous l'ensemble de contraintes  $P(x)$ .

### Résolution du problème multi-objectifs sous contraintes

La première difficulté rencontrée lors de la résolution d'un problème multi-objectif est le choix de la méthode de résolution. En effet, de nombreuses approches ont été développées. Trois grandes catégories peuvent être distinguées [Cohon and Marks, 1975, Jozefowicz, 2013, Hwang and Masud, 2012], chacune avec ses avantages et ses inconvénients, et s'appuient sur les préférences du décideur :

- Préférence a priori [Hwang and Masud, 2012] : Dans cette catégorie, les préférences entre les différentes fonctions objectifs, c'est-à-dire l'ordre d'importance des fonctions objectifs, sont indiquées avant l'exécution de la méthode d'optimisation. Cette approche demande un travail de modélisation du compromis entre les différentes fonctions objectifs afin de les classer. En effet, émettre une préférence sans connaissance sur le comportement des fonctions objectifs les unes envers les autres est une tâche difficile en pratique. Des exemples bien connus de méthodes avec préférence a priori comprennent la méthode de la fonction d'utilité, la méthode lexicographique [Fourman, 1985] et la programmation d'objectifs [Charnes et al., 1955].
- Préférence a posteriori [Hwang and Masud, 2012] : La méthode d'optimisation fournit un ensemble de solutions au décideur. Une solution unique est choisie parmi cet ensemble. Ainsi, il n'est plus nécessaire de modéliser les préférences entre les différentes fonctions objectifs. Par ailleurs, il faut calculer un ensemble de solutions, ce qui peut s'avérer plus coûteux en temps de calcul. Parmi les méthodes avec préférence a posteriori couramment utilisées, on cite : Intersection normale des limites (Normal Boundary Intersection) [Das and Dennis, 1998], Optimisation successive de Pareto (Successive Pareto Optimization) [Mueller-Gritschneider et al., 2009].

- Préférence progressive [Miettinen et al., 2008] : le choix de compromis entre les différentes fonctions objectifs est affiné au fur et à mesure du déroulement de la méthode d'optimisation. Cette approche implique la disponibilité du décideur tout au long de la résolution du problème pour émettre ses choix. On peut citer par exemple la méthode de Zionts-Wallenius [Zionts and Wallenius, 1976]

Il est à noter que certaines méthodes n'entrent pas exclusivement dans une seule catégorie. Par exemple, une méthode à préférence a priori peut également choisir a posteriori une unique solution parmi l'ensemble proposé.

La méthode de résolution adoptée dans ce travail est l'ordonnancement lexicographique proposé par Fourman [Fourman, 1985], qui entre dans la première catégorie. Il s'agit d'une méthode très intuitive qui consiste à résoudre les fonctions objectifs les unes après les autres tout en complétant l'ensemble des contraintes. Elle suppose donc que les fonctions objectifs peuvent être classées selon un ordre d'importance [Coello, 2000, Arunkumar et al., 2006, Marler and Arora, 2004].

**Méthode lexicographique** Supposons par exemple que  $o_1(x)$  soit l'objectif le plus important et  $o_n(x)$  le moins important, classés de telle sorte que si  $i < j$  alors  $o_i$  est prioritaire sur  $o_j$ . Notons  $n_i(x)$  les contraintes à respecter (dans notre cas,  $o_i$  correspond à  $r_i$  et  $n_i(x)$  à  $p_i(x)$ ). La solution optimale peut être obtenue en minimisant (ou en maximisant) les fonctions objectifs, en commençant par les plus importantes :

Minimiser (ou Maximiser)  $o_1(x)$  avec  $n_i(x) > 0, i = 1, \dots, l$ .

On note  $x^*$  la solution obtenue pour ce premier problème.

Ensuite, la première fonction objectif est transformée en contrainte puis la seconde fonction objectif (dans l'ordre d'importance des objectifs) est considérée et le problème suivant est résolu :

Minimiser (ou Maximiser)  $o_2(x)$  avec  $n_i(x) > 0, i = 1, \dots, l$  et  $o_1(x) = o_1(x^*)$

La procédure est itérée jusqu'à ce que tous les objectifs aient été traités.

L'ordre d'importance des objectifs est souvent choisi de façon arbitraire, ce qui représente le principal inconvénient de cette méthode. Par ailleurs, dans ce travail, le but est de trouver toutes les solutions qui satisfont les objectifs et les contraintes. Il est différent de l'optimisation multi-objectif classique qui vise à trouver la solution réalisable (ou l'ensemble de solutions) minimisant ou maximisant simultanément toutes les fonctions objectifs. Ainsi, l'ordre de résolution n'a aucune influence ici. Par conséquent, l'algorithme de résolution multi-objectif [Arunkumar et al., 2006, Marler and Arora, 2004] calcule le vecteur de solutions  $x$  (contenant les 12 probabilités conditionnelles) étant donné le domaine de définition de probabilité  $\Delta$ , les fonctions objectifs  $R(x)$  et les contraintes  $P(x)$ .

L'algorithme correspondant est donné ci-dessous (**Algorithme 1**).

Tout d'abord, l'espace de recherche  $\Delta$ , qui représente le domaine de définition des probabilités, est discrétisé. Nous choisissons  $\Delta = ]0; 1[$  car les valeurs 0 et 1 sont considérées comme triviales. Nous reviendrons sur ces valeurs dans la Partie 3.4.

Dans cet algorithme, les inégalités concernant  $r_1$  et  $r_2$  sont résolues en premier.  $r_1$

---

**Algorithme 1 : Résolution du problème multi-objectif**


---

**Entrée:**  $\Delta, R(x), P(x)$   
 $X = [a_0, b_0, c_0, d_0, a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2]$   
 $X, X_1, X_2, X_3, X_4, X_5, X_6 \leftarrow \emptyset$   
**Tant que**  $R(x) \neq \emptyset$  **faire**  
  **pour**  $x \in \Delta$  **faire**  
    **si**  $r_1(x) > 0$  **alors**  
       $X_1 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_1(x)\}$   
    **fin si**  
    **si**  $r_2(x) > 0$  **alors**  
       $X_2 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_2(x)\}$   
    **fin si**  
  **fin pour**  
  **pour**  $x \in \{X_1 \cup X_2\}$  **faire**  
    **si**  $r_3(x) > 0$  **alors**  
       $X_3 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_3(x)\}$   
    **fin si**  
  **fin pour**  
  **si**  $r_4(X_3) > 0$  **alors**  
     $X_4 \leftarrow X_3$   
     $R(x) \leftarrow R(x) \setminus \{r_4(x)\}$   
  **fin si**  
  **si**  $r_5(X_4) > 0$  **alors**  
     $X_5 \leftarrow X_4$   
     $R(x) \leftarrow R(x) \setminus \{r_5(x)\}$   
  **fin si**  
  **si**  $r_6(X_5) > 0$  **alors**  
     $X_6 \leftarrow X_5$   
     $R(x) \leftarrow R(x) \setminus \{r_6(x)\}$   
  **fin si**  
  **si**  $P(X_6) > 0$  **alors**  
     $X \leftarrow X_6$   
  **fin si**  
**fin Tant que**  
**Retourner**  $X$

---

et  $r_2$  dépendent chacun de 4 probabilités conditionnelles :  $[a_0, b_0, c_0, d_0]$  pour  $r_1$  et  $[a_1, b_1, c_1, d_1]$  pour  $r_2$ . Les solutions relatives à ces deux objectifs sont retenues et sauvegardées respectivement dans  $X_1$  et  $X_2$ . Par la suite, l'inégalité concernant  $r_3$ , qui dépend de  $[a_0, b_0, c_0, d_0, a_1, b_1, c_1, d_1]$ , est résolue. Les valeurs testées pour résoudre cette inégalité proviennent des solutions déjà retenues à l'étape précédente, c'est-à-dire  $X_1$  pour  $[a_0, b_0, c_0, d_0]$  et  $X_2$  pour  $[a_1, b_1, c_1, d_1]$ . Le même raisonnement est suivi par la suite pour la résolution de  $r_4$ ,  $r_5$  et  $r_6$  tout en prenant en compte les contraintes.

Ainsi, l'ensemble des solutions obtenues contient  $q$  combinaisons différentes de probabilités conditionnelles. Chaque combinaison  $[a_0, b_0, \dots, d_2]$  parmi cet ensemble respecte tous les objectifs et toutes les contraintes énoncées précédemment. Par exemple, en prenant un pas de discrétisation de 0.05, 95 solutions sont obtenues.

Par conséquent, cette démarche méthodologique garantit que si l'utilisateur choisit une combinaison de probabilités conditionnelles  $[a_0, b_0, \dots, d_2]$  parmi les solutions proposées, tous les objectifs définis précédemment dans la Table 3.4 seront atteints et la fusion de décision sera améliorée.

En dépit des avancées technologiques et de l'augmentation de la puissance de calcul des ordinateurs, la méthode lexicographique sera nécessairement limitée en temps de calcul lorsque le problème est de grande taille. On parle d'explosion combinatoire. Il est alors possible de se tourner vers les méthodes approchées (comme les méta-heuristiques) qui cherchent en un temps de calcul "raisonnable" des solutions, mais sans garantie sur leur qualité [Talbi, 2009].

### 3.3.6 Fusion des décisions en ligne

Dans l'étape en ligne, les décisions des méthodes individuelles sont introduites dans le RB paramétré, au niveau des noeuds  $D_1$  et  $D_2$ , à chaque instant. Le résultat correspondant à la fusion globale est renvoyé par le noeud  $D$  à chaque instant.

**Exemple** Afin de vérifier l'exactitude des solutions obtenues et ainsi établir l'efficacité de la méthodologie proposée, une phase de test est nécessaire. Le test permet de démontrer que si nous choisissons une combinaison de probabilités conditionnelles parmi l'ensemble obtenu de solutions réalisables (soit un total de 95 solutions), la performance globale est améliorée.

Pour l'évaluation des performances, le cadre générique est développé dans Matlab, en utilisant la BNT (BayesNet Toolbox) proposée par Murphy [Murphy et al., 2001]. La structure du RB, illustrée par la Figure 3.5, correspond à la couche générique montrée auparavant.

Puisque le noeud global  $D$  doit décider entre 4 modalités,  $\{CFN, F_0, F_1, F_2\}$ , la probabilité a priori est égale à  $\frac{1}{4}$ , comme le montre la Table 3.7.

		$D$		
$CFN$	$F_0$	$F_1$	$F_2$	
$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

TABLE 3.7 – Table de probabilités a priori du noeud  $D$

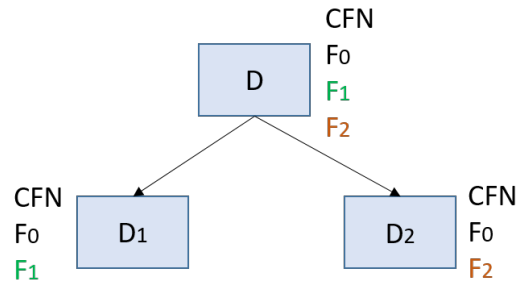


FIGURE 3.5 – Le cadre générique proposé

En ce qui concerne les tables de probabilités conditionnelles des noeuds  $D_1$  et  $D_2$ , n'importe quelle combinaison parmi l'ensemble des solutions déduites précédemment ( $\{a_{0_1}, \dots, d_{2_1}, e_{0_1}, \dots, h_{2_1}\}; \dots; \{a_{0_{95}}, \dots, d_{2_{95}}, e_{0_{95}}, \dots, h_{2_{95}}\}$ ) peut être choisie. Un exemple de tables de probabilités conditionnelles, avec un exemple de combinaison parmi celles obtenues, est donné dans les Tables 3.8a et 3.8b.

$D \backslash D_1$	$CFN$	$F_0$	$F_1$
$CFN$	0.51	0.11	0.38
$F_0$	0.06	0.36	0.58
$F_1$	0.06	0.01	0.93
$F_2$	0.21	0.31	0.48

(a) Table du noeud  $D_1$ 

$D \backslash D_2$	$CFN$	$F_0$	$F_2$
$CFN$	0.51	0.11	0.38
$F_0$	0.06	0.36	0.58
$F_1$	0.21	0.31	0.48
$F_2$	0.06	0.01	0.93

(b) Table du noeud  $D_2$ 

TABLE 3.8 – Tables de probabilités conditionnelles

Les résultats de fusion pour chaque combinaison sont donnés dans la Table 3.9. On constate que la matrice de confusion de la méthode de fusion, relative au noeud  $D$ , est la même pour chaque combinaison de probabilités ( $\{a_{0_1}, \dots, d_{2_1}, e_{0_1}, \dots, h_{2_1}\}; \dots; \{a_{0_{95}}, \dots, d_{2_{95}}, e_{0_{95}}, \dots, h_{2_{95}}\}$ ).

	$CFN$	$F_0$	$F_1$	$F_2$
$CFN$	100	0	0	0
$F_0$	0	100	0	0
$F_1$	0	0	100	0
$F_2$	0	0	0	100

TABLE 3.9 – Matrice de confusion de la méthode de fusion

Ces résultats illustrent les performances excellentes en diagnostic de la méthode de fusion proposée puisque le FDR est de 100% et le FAR est nul pour toutes les combinaisons de probabilités obtenues. Une comparaison est donnée dans la Table 3.10, et reflète l'avantage en termes d'amélioration des performances en diagnostic qui peut être atteint avec la méthodologie proposée.

Par conséquent, si les probabilités conditionnelles sont choisies parmi l'ensemble des solutions proposées par notre méthodologie, nous pouvons nous assurer que la performance de la fusion sera toujours meilleure que celle des méthodes individuelles.



	Méthode 1	Méthode 2	Méthode de Fusion
<b>FDR</b>	66.67	66.67	100
<b>FAR</b>	15	20	0

TABLE 3.10 – Comparaison de performances (%)

### 3.4 Discussion

Dans cette partie, nous allons revenir sur certains points importants de la méthodologie proposée, comme le choix des objectifs à atteindre, la prise en compte d'évidences incertaines ainsi que l'analyse statistique des solutions obtenues.

#### 3.4.1 Choix des objectifs

Le choix des objectifs dépend de la stratégie adoptée par l'utilisateur : soit il décide de privilégier la diminution du FAR ou alors l'augmentation du FDR puisqu'il existe un compromis entre ces deux métriques.

Dans la partie 3.3.2, nous avons privilégié la diminution du FAR en choisissant la modalité *CFN* à la place de  $F_0$  concernant les objectifs 2 et 4 (voir les objectifs en rouge de la Table 3.11a).

Si l'utilisateur souhaite privilégier le FDR, alors la table des objectifs 3.11b est préférable.

$d_2 \backslash d_1$	<i>CFN</i>	$F_0$	$F_1$
<i>CFN</i>	<i>CFN</i>	<i>CFN</i>	$F_1$
$F_0$	<i>CFN</i>	$F_0$	$F_1$
$F_2$	$F_2$	$F_2$	$F_1 \cap F_2$

(a) Objectifs choisis privilégiant le FAR

$d_2 \backslash d_1$	<i>CFN</i>	$F_0$	$F_1$
<i>CFN</i>	<i>CFN</i>	$F_0$	$F_1$
$F_0$	$F_0$	$F_0$	$F_1$
$F_2$	$F_2$	$F_2$	$F_1 \cap F_2$

(b) Objectifs choisis privilégiant le FDR

TABLE 3.11 – Tables des objectifs pour des stratégies différentes

Afin d'assurer les nouveaux objectifs 2 et 4 de la Table 3.11b (qui entraînent les mêmes contraintes puisqu'ils sont symétriques), les critères suivants doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_0/D_1 = CFN \cap D_2 = F_0) > P(D = F_0) \\ P(D = F_0/D_1 = CFN \cap D_2 = F_0) > P(D = (F_i \cup CFN)/D_1 = CFN \cap D_2 = F_0), \\ \quad \quad \quad i \in \{1, 2\} \end{array} \right. \quad (3.42)$$

3 contraintes sont alors obtenues :

$$\frac{b_0 b_1}{(a_0 + b_0 + c_0 + d_0)(a_1 + b_1 + c_1 + d_1)} > \alpha_0^2 \quad (3.43)$$

$$\left\{ \begin{array}{l} b_0 b_1 > a_0 a_1 \\ b_0 b_1 > c_0 d_1 \end{array} \right. \quad (3.44)$$

L'approche lexicographique est appliquée de la même manière afin d'obtenir l'ensemble des solutions respectant les nouveaux objectifs. Pour un pas de discrétisation de 0.05, on déduit alors 80 combinaisons de probabilités conditionnelles respectant l'ensemble de contraintes imposées et qui sont toutes équivalentes.

### 3.4.2 Analyse des paramètres obtenus

Au regard des tables des objectifs privilégiant le FAR (Table 3.11a) ou le FDR (Table 3.11b), on peut identifier les tables de probabilités conditionnelles suivantes (voir Table 3.12 pour une stratégie FAR et Table 3.13 pour une stratégie FDR), en se basant uniquement sur des évidences certaines.

$D \backslash D_1$	$CFN$	$F_0$	$F_1$
$CFN$	$\frac{2}{3}$	$\frac{1}{3}$	0
$F_0$	0	1	0
$F_1$	0	0	1
$F_2$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

(a) Table du noeud  $D_1$ 

$D \backslash D_2$	$CFN$	$F_0$	$F_2$
$CFN$	$\frac{2}{3}$	$\frac{1}{3}$	0
$F_0$	0	1	0
$F_1$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$F_2$	0	0	1

(b) Table du noeud  $D_2$ 

TABLE 3.12 – Tables de probabilités conditionnelles privilégiant le FAR

$D \backslash D_1$	$CFN$	$F_0$	$F_1$
$CFN$	1	0	0
$F_0$	$\frac{1}{3}$	$\frac{2}{3}$	0
$F_1$	0	0	1
$F_2$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

(a) Table du noeud  $D_1$ 

$D \backslash D_2$	$CFN$	$F_0$	$F_2$
$CFN$	1	0	0
$F_0$	$\frac{1}{3}$	$\frac{2}{3}$	0
$F_1$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$F_2$	0	0	1

(b) Table du noeud  $D_2$ 

TABLE 3.13 – Tables de probabilités conditionnelles privilégiant le FDR

En effet, à partir de la Table 3.11a, on constate qu'il y a 3 situations où la décision globale  $d$  doit être égale à  $CFN$ . Sur ces 3 situations, 2 sont réalisées lorsque la méthode  $m_1$  choisit la décision  $d_1 = CFN$  et une lorsque la méthode  $m_1$  choisit la décision  $d_1 = F_0$ . D'où les probabilités conditionnelles suivantes :  $P(D_1 = CFN/D = CFN) = \frac{2}{3}$  et  $P(D_1 = F_0/D = CFN) = \frac{1}{3}$ . Enfin, sur ces 3 situations, la méthode  $m_1$  ne choisit jamais la modalité  $F_1$ . On en déduit :  $P(D_1 = F_1/D = CFN) = 0$ . Le même raisonnement est appliqué pour déduire les autres probabilités conditionnelles.

Cependant, nous allons montrer par la suite que ces solutions ne sont pas toujours équivalentes aux solutions que nous avons déterminées (appelées dans la suite solutions théoriques), notamment en terme de performances (FDR et FAR) avec des évidences incertaines. C'est pour cela que nous avons choisi de comparer, à titre indicatif, les solutions émanant de notre méthodologie par rapport à cette solution, appelée *solution conventionnelle* afin de pouvoir la distinguer des solutions théoriques dans ce qui suit.

#### Prise en compte d'évidences incertaines

Jusqu'à présent, nous nous sommes uniquement concentrés sur la couche générique du RB proposé pour la fusion (voir Figure 3.4). Nous avons considéré les évidences introduites au niveau des noeuds  $D_1$  et  $D_2$  comme certaines (ou binaires), c'est-à-dire que la probabilité de la modalité choisie par la méthode passe à 1 tandis que les autres modalités passent à 0, comme illustré par la Figure 3.6 a). Ces évidences certaines sont obtenues en introduisant

les observations (0 ou 1) directement dans la couche générique ou en appliquant la règle du maximum a posteriori lorsque les méthodes individuelles génèrent des probabilités a posteriori différentes de 0 et 1. Dans ce contexte, toutes les solutions sont équivalentes, c'est-à-dire qu'elles conduisent au même FDR et FAR.

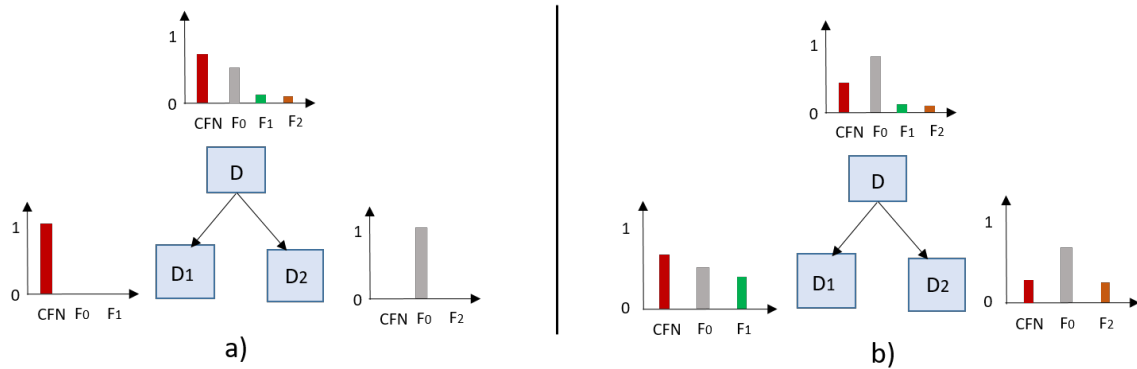


FIGURE 3.6 – a) Evidences certaines VS b) incertaines au niveau des noeuds  $D_1$  et  $D_2$

Cependant, une nouvelle problématique émerge lorsqu'on considère la couche spécifique ou lorsqu'on décide de ne pas appliquer la règle du maximum a posteriori afin de garder toutes les informations contenues dans les décisions individuelles. En effet, les évidences deviennent incertaines, c'est-à-dire qu'elles peuvent prendre toutes les valeurs entre 0 et 1, ce qui est plus réaliste (voir la Figure 3.6 b)). Ainsi, en intégrant cet aspect, il convient d'analyser les différentes solutions obtenues dans ce contexte d'évidences incertaines, et de les comparer suivant les FARs et les FDRs.

Pour cela, nous avons réalisé 580 simulations Monte Carlo, reproduisant des situations où les 4 conditions de fonctionnement  $\{CFN, F_0, F_1, F_2\}$  se produisent, selon les scénarios décrits dans l'exemple 3.3.1.

Pour chaque simulation, les paramètres  $\{\mu_{CFN}, \mu_{s_1}, \mu_{s_2}, \mu_{s_3}, \Sigma_{CFN}, \Sigma_{s_1}, \Sigma_{s_2}, \Sigma_{s_3}\}$  sont choisis aléatoirement afin de générer les données d'apprentissage et de test, tout en veillant à ce que les méthodes  $m_1$  et  $m_2$  soient expertes dans le diagnostic des défauts spécifiques de types  $F_1$  et  $F_2$ . Pour le même scénario de simulation et pour chaque solution (combinaison de probabilités conditionnelles), la fusion basée sur notre méthodologie est réalisée. Les FDRs et les FARs de chaque simulation sont ensuite récupérés pour des fins de comparaison.

Dans la suite, nous allons aborder ces comparaisons pour les deux métriques choisies précédemment : FAR et FDR.

**Comparaison de solutions pour une stratégie privilégiant le FAR** Dans cette partie, la solution conventionnelle est comparée avec les 95 solutions obtenues suivant une stratégie privilégiant le FAR (Table 3.11a).

Tout d'abord, on constate que les distributions de FAR pour les différentes solutions ne peuvent pas être assimilées à la distribution normale selon le test de Shapiro-Wilks (p-value  $< 10^{-6}$ ) [Shapiro and Wilk, 1965]. Par conséquent, afin de les comparer, nous avons opté pour le test de Wilcoxon apparié [Wilcoxon, 1945], qui représente une alternative non paramétrique au test-t pour des échantillons appariés.

La Figure 3.7 illustre les p-values pour le test de Wilcoxon lorsqu'on compare le FAR de la solution conventionnelle avec les FARs des autres solutions théoriques.

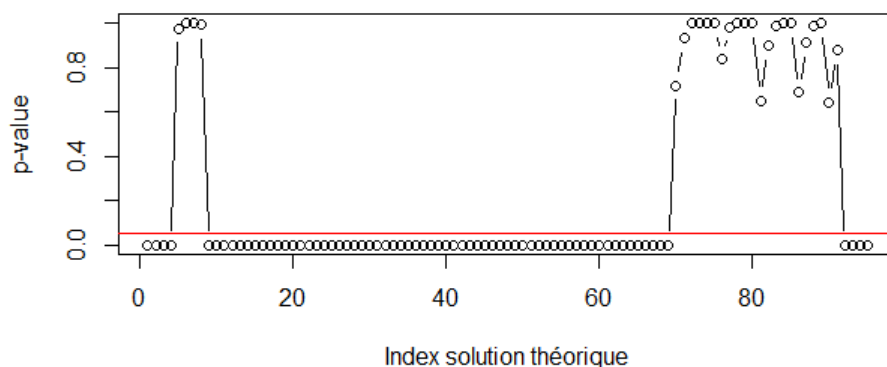


FIGURE 3.7 – P-values du test de Wilcoxon apparié pour les FARs (la ligne continue en rouge représente le seuil de significativité de 5%)

On observe alors que pour un certain nombre de solutions, dont l'index est donné ci-après :  $\{5; 6; 7; 8; 70; 71; 72; 73; 74; 75; 76; 77; 78; 79; 80; 81; 82; 83; 84; 85; 86; 87; 88; 89; 90; 91\}$ , les différences ne sont pas significatives au seuil de 5%. Ce sont donc des solutions qui présentent des FARs statistiquement proches du FAR de la solution conventionnelle.

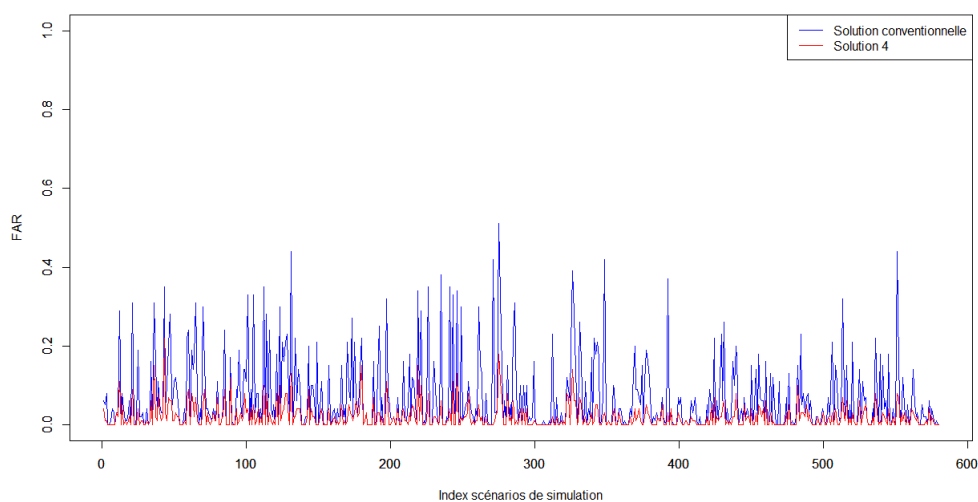


FIGURE 3.8 – Comparaison de FAR : solution conventionnelle VS solution théorique 4

Par ailleurs, toujours selon les p-values du Test de Wilcoxon apparié, les 10 meilleures solutions sont :  $\{4; 3; 9; 10; 33; 1; 11; 92; 2; 16\}$ . La Figure 3.8 illustre la comparaison entre le FAR de la solution conventionnelle (en bleu) et celui de la solution 4 (en rouge). On constate que la solution théorique 4 présente un meilleur FAR sur l'ensemble des simulations.

Afin de comparer les différentes solutions dans l'espace des simulations Monte Carlo (580 scénarios), nous avons réalisé une ACP. Les deux premières dimensions de l'ACP expliquent 76.3% de la variabilité enregistrée dans les données. La projection des différentes solutions dans le plan principal de l'ACP, formé par les deux premières dimensions, est

donnée dans la Figure 3.9.

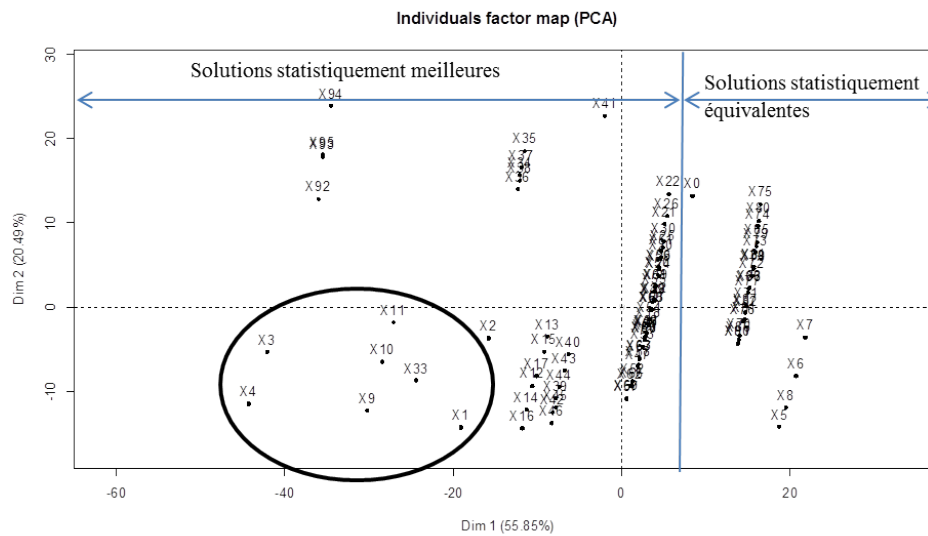


FIGURE 3.9 – Projection des solutions dans le plan principal

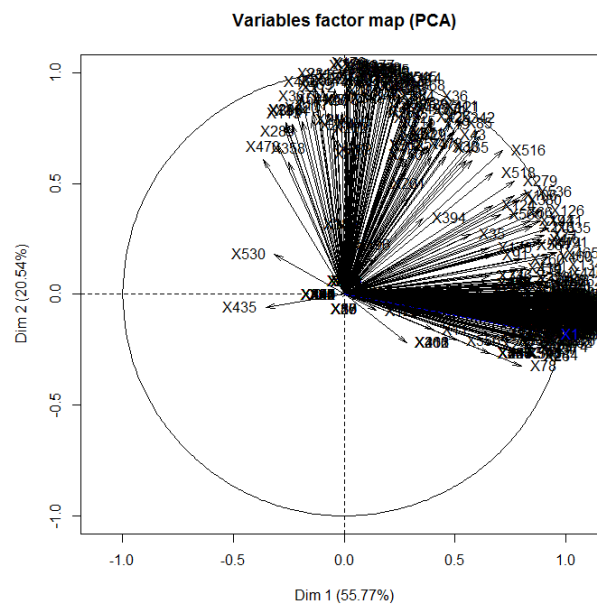


FIGURE 3.10 – Cercle de corrélation dans le plan principal

La Figure 3.10 explique la signification des axes de l'ACP. Les valeurs positives sur l'axe 1 et l'axe 2 représentent des simulations ayant un FAR élevé. Ainsi, les solutions représentées dans la Figure 3.9, et situées dans la direction de la bissectrice du cadran positif présentent des FARs élevés sur l'ensemble des simulations. A l'opposé, les solutions situées dans la direction de la bissectrice du cadran négatif présentent des FARs inférieurs sur l'ensemble des simulations. On peut ainsi repérer deux sous-espaces, comme illustré dans la Figure 3.9 : celui des solutions statistiquement meilleures que la solution conventionnelle (dont les meilleures en termes de FAR se trouvent à l'intérieur de l'ellipse) et respectivement celui des solutions statistiquement équivalentes.

La projection des différentes solutions dans l'espace des probabilités conditionnelles les définissant est illustrée dans la Figure 3.11, avec en bleu la solution conventionnelle, en

rouge les meilleures solutions en termes de FAR selon la p-value du test de Wilcoxon, et en vert les solutions qui ne sont pas significativement différentes de la solution conventionnelle. On constate que les solutions les plus proches de la solution conventionnelle sont celles qui ne sont pas statistiquement différentes.

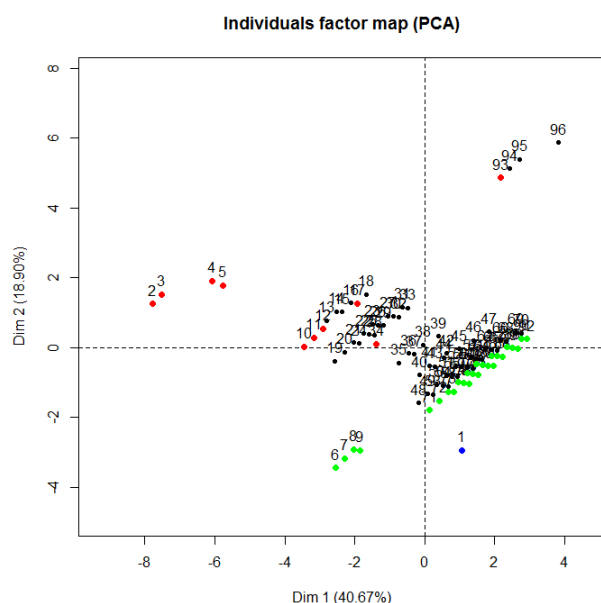


FIGURE 3.11 – Plan principal de projection de l'ACP

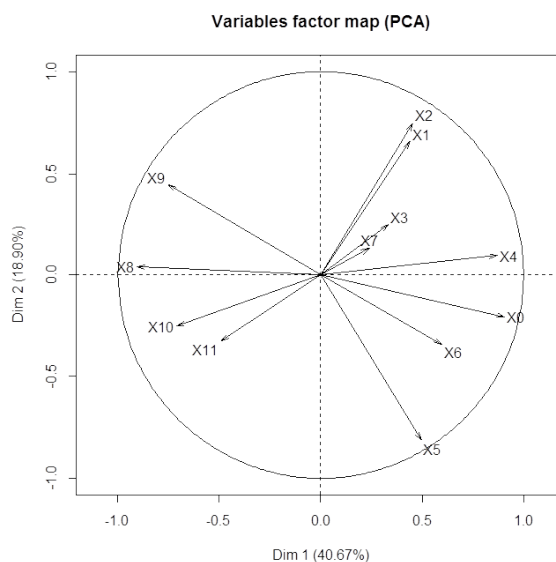


FIGURE 3.12 – Cercle de corrélation

Dans ce cas, les trois premiers axes de l'ACP expliquent 73.84% de la variabilité enregistrée dans les données.

Les variables contribuant le plus à la construction de l'axe 1 de l'ACP sont :  $\{X_0, X_4, X_8, X_9, X_{10}\}$  correspondant respectivement aux probabilités conditionnelles  $\{a_0; a_1; a_2; b_2, c_2\}$ . Ce sont les variables les plus proches de cet axe, et proches du bord du cercle de corrélation. De même, les variables contribuant le plus à la construction de l'axe 2 sont :  $\{X_1, X_2, X_5\}$ , correspondant à  $\{b_0; c_0; b_1\}$ .

On remarque que la solution conventionnelle présente des valeurs plus grandes que les autres solutions au niveau de  $X_5$ , correspondant à la probabilité conditionnelle  $b_1$ .

Les meilleures solutions théoriques  $\{2; 3; 4; 5\}$  présentent des valeurs plus grandes que la solution conventionnelle, surtout au niveau des variables  $X_8$  et  $X_9$ , correspondant à  $\{a_2; b_2\}$ .

Afin de mieux comprendre ce que nous entendons par des valeurs plus "grandes" dans les affirmations ci-dessus, la Table 3.14 illustre les probabilités conditionnelles des 10 meilleures solutions en terme de FAR par rapport à la solution conventionnelle (représentée par la ligne 1 du tableau).

Enfin, les FARs moyens (ainsi que les FDRs correspondants) sur l'ensemble des simulations pour les 10 meilleures solutions sont donnés dans la Table 3.15.

**Comparaison de solutions pour une stratégie privilégiant le FDR** Dans cette partie, la solution conventionnelle est comparée avec les 80 solutions obtenues suivant une stratégie privilégiant le FDR.

	$a_0$	$b_0$	$c_0$	$d_0$	$a_1$	$b_1$	$c_1$	$d_1$	$a_2$	$b_2$	$c_2$	$d_2$
Sol conv.	0.66	0	0	0.33	0.33	1	0	0.33	0	0	1	0.33
Sol 4	0.56	0.11	0.11	0.21	0.16	0.26	0.11	0.21	0.28	0.63	0.78	0.58
Sol 3	0.56	0.11	0.06	0.21	0.16	0.26	0.01	0.36	0.28	0.63	0.93	0.43
Sol 9	0.61	0.11	0.01	0.16	0.21	0.41	0.21	0.26	0.18	0.48	0.78	0.58
Sol 10	0.61	0.11	0.01	0.21	0.21	0.41	0.21	0.31	0.18	0.48	0.78	0.48
Sol 33	0.66	0.01	0.11	0.11	0.26	0.56	0.01	0.61	0.08	0.43	0.88	0.28
Sol 1	0.51	0.06	0.06	0.21	0.11	0.36	0.01	0.31	0.38	0.58	0.93	0.48
Sol 11	0.61	0.11	0.01	0.26	0.21	0.41	0.21	0.36	0.18	0.48	0.78	0.38
Sol 92	0.66	0.46	0.46	0.36	0.31	0.36	0.01	0.31	0.03	0.18	0.53	0.33
Sol 2	0.51	0.06	0.06	0.26	0.11	0.36	0.01	0.36	0.38	0.58	0.93	0.38
Sol 16	0.61	0.16	0.16	0.16	0.21	0.51	0.11	0.56	0.18	0.33	0.73	0.28

TABLE 3.14 – Paramètres de la solution conventionnelle et des 10 meilleures solutions en terme de FAR

	Sol 4	Sol 3	Sol 9	Sol 10	Sol 33	Sol 1	Sol 11	Sol 92	Sol 2	Sol 16	Sol conv. 1
FAR	0.017	0.02	0.03	0.03	0.03	0.04	0.03	0.03	0.04	0.04	0.07
FDR	0.68	0.78	0.78	0.79	0.80	0.78	0.80	0.80	0.80	0.79	0.82

TABLE 3.15 – FARs moyens des 10 meilleures solutions et de la solution conventionnelle

Tout d'abord, les distributions de FDR pour les différentes solutions ne sont pas gaussiennes selon le test de Shapiro-Wilks ( $p$ -value  $< 10^{-6}$ ), d'où l'utilisation du test de Wilcoxon apparié pour les comparer.

La Figure 3.13 illustre les  $p$ -values pour le test de Wilcoxon lorsque le FAR de la solution conventionnelle est comparé avec les FARs des autres solutions théoriques.

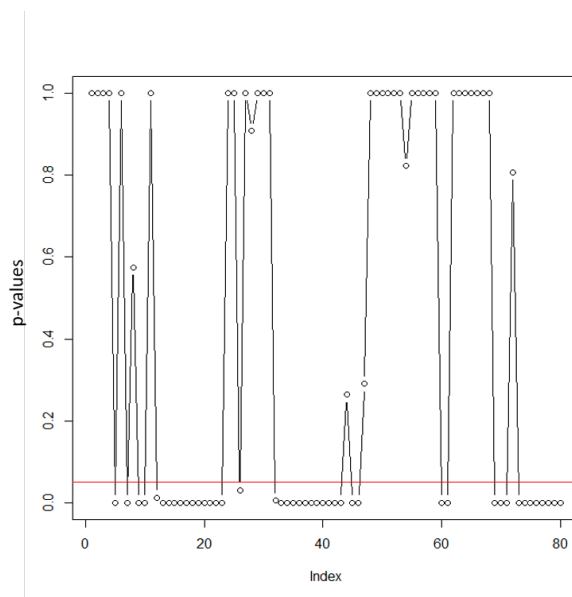


FIGURE 3.13 – Test de Wilcoxon apparié

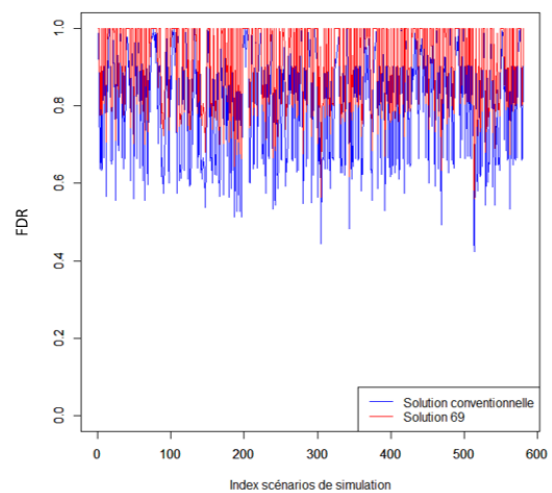


FIGURE 3.14 – Comparaison de FDR

Ainsi, pour 36 solutions, dont l'index est donné ci-après : {1; 2; 3; 4; 6; 8; 11; 24; 25; 27; 28; 29; 30; 31; 44; 47; 48; 49; 50; 51; 52; 53; 54; 55; 56; 57; 58;

59; 62; 63; 64; 65; 66; 67; 68; 72}, les différences ne sont pas significatives au seuil de 5%.

Les 10 meilleures solutions (selon les p-values du Test de Wilcoxon apparié) sont : {69; 75; 45; 46; 73; 70; 9; 74; 71; 76}.

La Figure 3.14 illustre le graphique comparatif du FDR entre la solution conventionnelle (en bleu) et la solution 69 (en rouge). On constate là-encore que la solution théorique présente un meilleur FDR sur l'ensemble des simulations.

Afin de comparer les différentes solutions dans l'espace des simulations Monte Carlo, nous avons réalisé une ACP, comme l'illustre la Figure 3.15. Les deux premières dimensions de l'ACP expliquent 89.56% de la variabilité enregistrée dans les données.

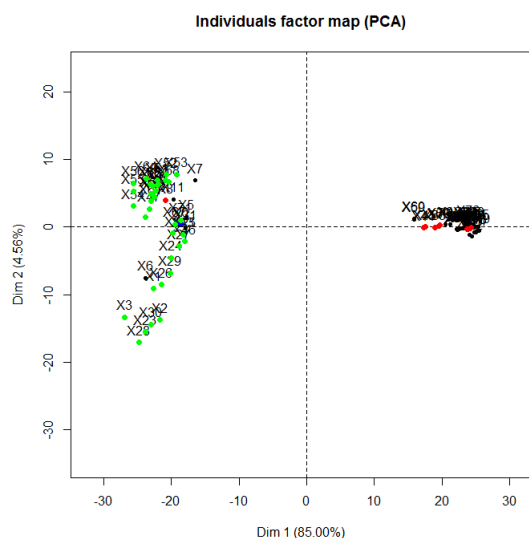


FIGURE 3.15 – Projection des solutions dans le plan principal

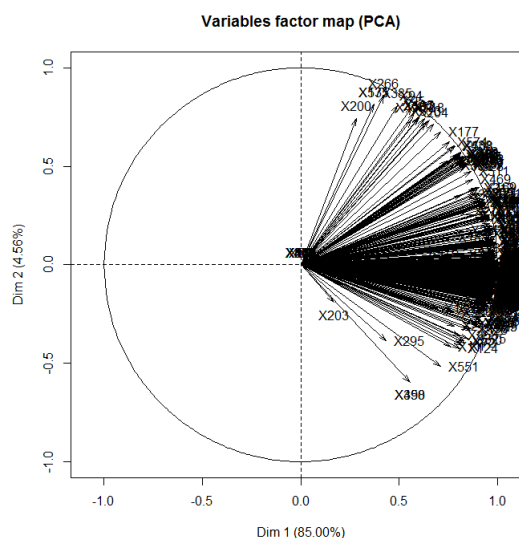


FIGURE 3.16 – Cercle de corrélation

La signification des axes de l'ACP peut être déduite de la Figure 3.16. On peut constater que les valeurs positives sur l'axe 1 traduisent des simulations ayant un FDR élevé. Quelques simulations (moins nombreuses) contribuent à la construction de l'axe 2. Des valeurs positives sur cet axe signifient un FDR grand pour ces simulations.

Par ailleurs, on constate que les observations représentées en rouge dans le plan principal de projection (Figure 3.15) sont les 10 solutions ayant le meilleur FDR. Elles se trouvent majoritairement dans le côté droit de l'axe 1. Celles en vert sont équivalentes à la solution conventionnelle. Ces dernières ont un FDR bien plus petit que les 10 meilleures solutions, et se trouvent toutes dans le côté négatif de l'axe 1 (Figure 3.15).

La projection des différentes solutions dans l'espace des probabilités conditionnelles les définissant est illustrée dans la Figure 3.17. Les deux premiers axes de l'ACP expliquent seulement 56% de la variabilité enregistrée dans les données tandis que les quatre premiers axes expliquent 80.95% de cette variabilité.

Comme constaté sur les Figures 3.17 et 3.18, et à la différence du FAR, on ne dispose plus de caractéristiques dominantes dans l'espace principal de projection nous permettant de distinguer les 10 meilleures solutions suivant le FDR (en rouge) des solutions équivalentes (en vert) à la solution conventionnelle (en bleu). La même conclusion est valable dans les autres espaces de projection.



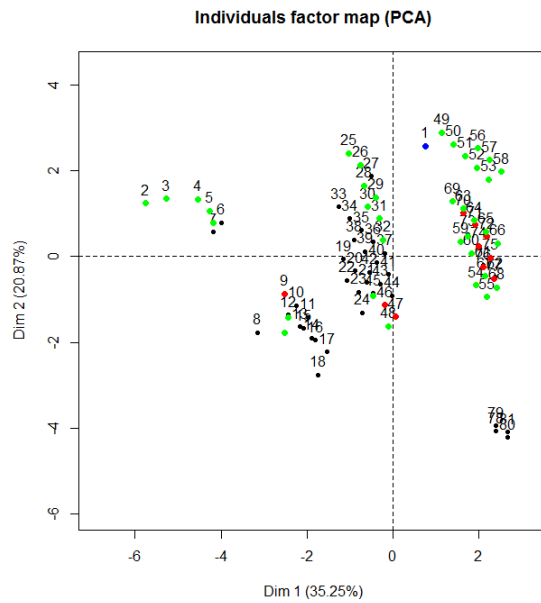


FIGURE 3.17 – Plan principal de projection

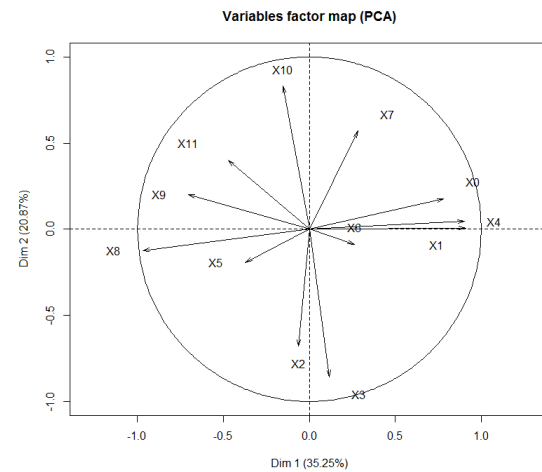


FIGURE 3.18 – Cercle de corrélation

La Table 3.16 illustre les probabilités conditionnelles des 10 meilleures solutions et de la solution conventionnelle (représentée par la ligne 1 de la table).

	$a_0$	$b_0$	$c_0$	$d_0$	$a_1$	$b_1$	$c_1$	$d_1$	$a_2$	$b_2$	$c_2$	$d_2$
Sol conv.	1	0.33	0	0.33	0	0.66	0	0.33	0	0	1	0.33
Sol 69	0.66	0.51	0.31	0.26	0.31	0.46	0.01	0.41	0.03	0.03	0.68	0.33
Sol 75	0.66	0.51	0.41	0.31	0.31	0.46	0.11	0.46	0.03	0.03	0.48	0.23
Sol 45	0.61	0.31	0.16	0.46	0.21	0.66	0.36	0.31	0.18	0.03	0.48	0.23
Sol 46	0.61	0.31	0.16	0.51	0.21	0.66	0.36	0.36	0.18	0.03	0.48	0.13
Sol 73	0.66	0.51	0.36	0.31	0.31	0.46	0.06	0.46	0.03	0.03	0.58	0.23
Sol 70	0.66	0.51	0.31	0.31	0.31	0.46	0.01	0.46	0.03	0.03	0.68	0.23
Sol 9	0.56	0.26	0.41	0.46	0.16	0.56	0.01	0.26	0.28	0.18	0.58	0.28
Sol 74	0.66	0.51	0.36	0.36	0.31	0.46	0.06	0.51	0.03	0.03	0.58	0.13
Sol 71	0.66	0.51	0.31	0.36	0.31	0.46	0.01	0.51	0.03	0.03	0.68	0.13
Sol 76	0.66	0.51	0.41	0.36	0.31	0.46	0.11	0.51	0.03	0.03	0.48	0.13

TABLE 3.16 – Paramètres de la solution conventionnelle et des 10 meilleures solutions en terme de FDR

Enfin, les FDRs moyens (ainsi que les FARs correspondants) sur l'ensemble des simulations des 10 meilleures solutions sont donnés dans la Table 3.17.

Ainsi, nous avons analysé dans cette partie l'ensemble des solutions obtenues, que ce soit pour une stratégie privilégiant le FAR ou le FDR. Nous avons comparé ces solutions avec une solution conventionnelle dans le cadre d'évidences incertaines, et nous avons conclu qu'il existe toujours des solutions théoriques meilleures que la solution conventionnelle. En effet, le FAR obtenu avec la solution conventionnelle (7%) est nettement

	Sol 69	Sol 75	Sol 45	Sol 46	Sol 73	Sol 70	Sol 9	Sol 74	Sol 71	Sol 76	Sol conv. 1
FDR	0.91	0.92	0.90	0.90	0.92	0.91	0.91	0.93	0.91	0.93	0.81
FAR	0.07	0.07	0.08	0.08	0.07	0.07	0.09	0.07	0.07	0.07	0.07

TABLE 3.17 – FDRs moyens des 10 meilleures solutions et de la solution conventionnelle

supérieur au FAR obtenu avec la solution 4 (1.7%), tandis que le FDR de la solution conventionnelle (81%) est inférieur à celui de la solution 74 (93%).

### 3.4.3 Synthèse

En plus des résultats remarquables en FDD, l'originalité et l'intérêt scientifique de la méthodologie proposée sont les suivants :

- Elle fournit un cadre générique où les méthodes, avec différentes connaissances et informations préalables, jeux de données disponibles, exigences de modélisation, etc, peuvent être combinées. De plus, le cadre proposé n'est pas dédié à des applications particulières et répond ainsi au principal obstacle rencontré lors du développement d'approches fusionnant plusieurs méthodes. Généralement, les méthodes de FDD ne possèdent pas de caractère universel. En fonction de la nature du système considéré, il faudra mettre en oeuvre à chaque fois des méthodes spécifiques tenant compte des exigences, prérequis et technologies déployées. Bien qu'il y ait une recherche notable sur les approches fusionnant plusieurs méthodes, très peu de résultats théoriques existent. Ici, nous avons proposé un cadre générique et une méthodologie théorique qui s'affranchissent des contraintes spécifiques aux méthodes de FDD.
- Le cadre n'exclut aucune méthode FDD, qu'elle soit basée sur les données ou sur les modèles. En effet, les méthodes individuelles peuvent fonctionner séparément et leurs décisions peuvent être utilisées directement dans le cadre proposé. Il est également possible d'utiliser le même outil, le RB, pour modéliser les différentes méthodes avant de les combiner. Par exemple, les cartes de contrôle multivariées et d'autres techniques populaires comme l'ACP ont été modélisées par un RB classifieur qui a permis de détecter et de diagnostiquer les défauts avec le même cadre [Verron et al., 2010a], [Atoui et al., 2015b].
- L'approche proposée applique une stratégie efficace pour combiner les résultats individuels une fois disponibles à partir de diverses méthodes. Par conséquent, il surmonte la difficulté de connaître et de combiner différentes théories et concepts spécifiques à chaque méthode FDD, ce qui peut être très compliqué en pratique, comme expliqué dans le Chapitre 2.
- Il fournit également une solution possible pour détecter et diagnostiquer plusieurs défauts simultanés.
- Il permet d'utiliser une variété de sources d'information à travers les méthodes individuelles telles que les données historiques, les modèles mathématiques, les relations causales, etc., conduisant à un diagnostic plus efficace et plus fiable.
- La méthodologie est très flexible et peut s'adapter aux besoins de l'utilisateur et aux stratégies qu'il souhaite adapter pour la surveillance du système. En effet, pour une application spécifique, l'utilisateur peut choisir une stratégie de surveillance

spécifique puisqu'il n'y a pas d'objectifs de décision prédéfinis à respecter dans tous les cas.

- Enfin, il permet d'utiliser, à travers les deux méthodes individuelles, tous les progrès et les avancées réalisés par les différentes communautés de détection-diagnostic.

À notre connaissance, c'est l'un des premiers travaux qui développe une approche théorique et générique assurant la réalisation d'un certain nombre d'objectifs de FDD, et conduisant à une amélioration des performances.

Par la suite, nous allons appliquer la méthodologie proposée sur un système complexe qui est le TEP, afin de montrer la faisabilité et les améliorations apportées par cette approche.

## 3.5 Application : Diagnostic du Tennessee Eastman Process

### 3.5.1 Introduction

Les approches de surveillance peuvent être testées et évaluées sur plusieurs applications comme le système à trois réservoirs [Theilliol et al., 2002], le système de pile à combustible [Chatti et al., 2013a], le véhicule intelligent [Chatti et al., 2013b], le générateur de vapeur [Bouamama et al., 2006] ou le Tennessee Eastman Process (TEP) [Downs and Vogel, 1993]. Ce dernier a été développé par la société Eastman Chemical Company et est considéré comme un benchmark industriel réaliste où les chercheurs peuvent comparer leurs approches de contrôle et de surveillance des systèmes [Chiang et al., 2000a, Ding et al., 2009, Verron et al., 2010a, Atoui et al., 2015a, Ghosh et al., 2011, Lee et al., 2004, Chiang et al., 2000b, Lin et al., 2000, Li et al., 2011, Verron et al., 2006, Maurya et al., 2007, Chiang et al., 2004, Yin et al., 2012, Eslamloueyan, 2011, Jing and Hou, 2015, D'Angelo et al., 2016, Ghosh et al., 2014, Lau et al., 2013, He et al., 2014]. En effet, diverses méthodes ont été appliquées au TEP, telles que l'ACP [Ding et al., 2009], l'ACP dynamique non linéaire [Lin et al., 2000], la PSL [Li et al., 2011], la PSL dynamique [Lee et al., 2004], le RB [Verron et al., 2006], le SDG [Maurya et al., 2007, He et al., 2014], la QTA [Maurya et al., 2007], l'AD [Chiang et al., 2004], les SVMs [Chiang et al., 2004], etc.

Des comparaisons détaillées des méthodes à base de données appliquées au TEP, en particulier les méthodes statistiques telles que l'ACP, l'AD, etc., peuvent être trouvées dans [Chiang et al., 2000a, Yin et al., 2012].

Cependant, il convient de noter que la plupart des méthodes appliquées au TEP appartiennent à l'approche basée sur les données. Ceci s'explique par le fait que les données historiques du TEP sont disponibles, que ce soit pour le fonctionnement normal ou en conditions de défauts, et qu'un modèle pertinent de ce système est difficile à obtenir. Par conséquent, très peu d'attention a été accordée aux méthodes de détection-diagnostic à base de modèles pour le TEP.

Dans ce chapitre, nous allons tout d'abord présenter le procédé du TEP puis nous allons fournir un modèle graphique et analytique qui dérive du formalisme du BG. Ce modèle sera utilisé pour la détection et le diagnostic à base de modèle d'un ensemble de défauts pouvant survenir dans le système. Une autre méthode, cette fois-ci à base de données, sera également appliquée pour le diagnostic de défauts dans le TEP. Ceci nous permettra par la

suite de développer notre approche de prise de décision multi-objectifs proposée pour le diagnostic du TEP.

### 3.5.2 Présentation du TEP

Le TEP est un système basé sur un procédé chimique réel qui se compose de cinq éléments principaux : un réacteur, un compresseur, un décapeur, un séparateur et un condenseur, comme illustré dans la Figure 3.19.

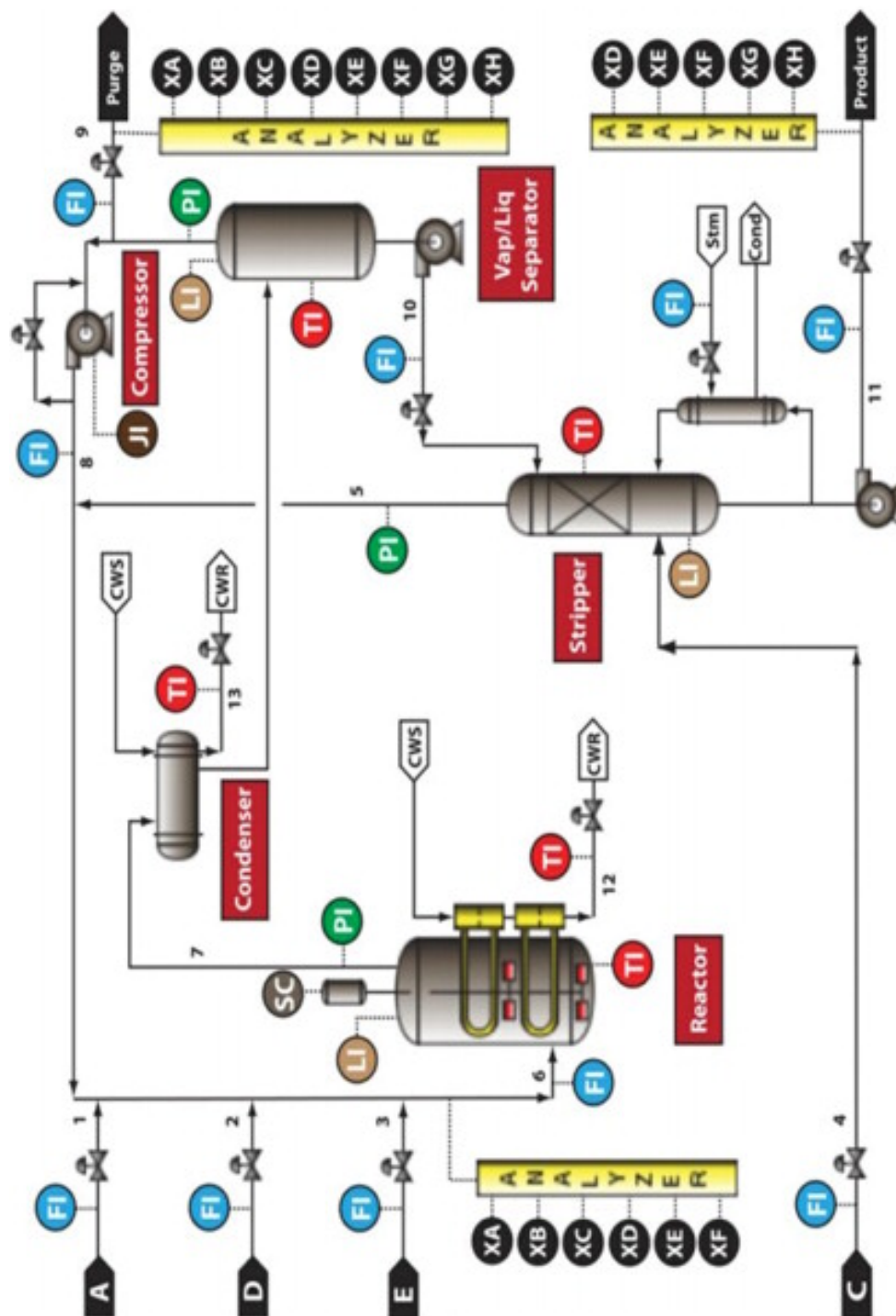
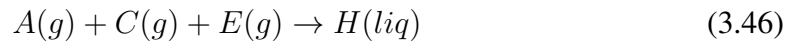
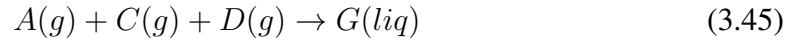


FIGURE 3.19 – Schéma du Tennessee Eastman Process

Quatre gaz réactifs (*A*, *C*, *D* et *E*) alimentent le réacteur où ils réagissent et forment les produits *G* et *H* sous forme gazeuse. Le gaz produit est refroidi au moyen d'un condenseur

et alimente alors le séparateur vapeur-liquide. La vapeur non condensée dans le séparateur est renvoyée vers le réacteur au moyen d'un compresseur. Le gaz inerte  $B$  (non réactif) et le produit dérivé  $F$  sont purgés du procédé dans le séparateur. Le flux condensé dans le séparateur est envoyé au décapeur qui a pour but d'éliminer les dernières traces de réactifs. Enfin, les produits  $G$  et  $H$  sont aspirés à l'extérieur du procédé [Downs and Vogel, 1993]. Les réactions chimiques du procédé sont données par le système d'équation (3.45, 3.46, 3.47, 3.48).



Le TEP comporte 53 variables dont 41 variables de mesure et 12 variables de contrôle. Parmi les variables mesurables, 22 le sont en continu (voir Table 1 en Annexe 5.2), tandis que le reste est échantillonné (voir Table 2 en Annexe 5.2). Les variables de contrôle sont décrites dans la Table 3 en Annexe 5.2. Il faut préciser que la vitesse de l'agitateur (variable  $X_{53}$ ) n'est pas prise en compte puisqu'elle reste toujours constante dû au système d'asservissement mis en place. Ainsi, seules 52 variables sont considérées.

Le TEP peut simuler 15 types de défauts identifiés, ce qui le rend très intéressant pour le test des méthodes de détection et de diagnostic. Les défauts simulés sont de natures diverses : saut en échelon, augmentation de la variabilité, vannes bloquées. Leur description est faite dans la Table 3.18.

Faute	Description	Type
$S_1$	Ratio d'alimentation $A/C$	Saut
$S_2$	Composition en $B$	Saut
$S_3$	Température d'alimentation en $D$	Saut
$S_4$	Température d'entrée du refroidissement liquide au réacteur	Saut
$S_5$	Température d'entrée du refroidissement liquide au condenseur	Saut
$S_6$	Baisse d'alimentation en $A$	Saut
$S_7$	Perte de pression de l'alimentation en $C$	Saut
$S_8$	Composition d'alimentation en $A, B$ et $C$	Variation aléatoire
$S_9$	Température d'alimentation en $D$	Variation aléatoire
$S_{10}$	Température d'alimentation en $C$	Variation aléatoire
$S_{11}$	Température d'entrée du refroidissement liquide au réacteur	Variation aléatoire
$S_{12}$	Température d'entrée du refroidissement liquide au condenseur	Variation aléatoire
$S_{13}$	Cinétiques des réactions	Dérive lente
$S_{14}$	Vanne du refroidissement liquide au réacteur	Bloquée
$S_{15}$	Vanne du refroidissement liquide au condenseur	Bloquée

TABLE 3.18 – Les défauts pouvant affecter le TEP

Deux jeux de données sont générés à partir du TEP : un jeu d'apprentissage et un jeu de test. Chaque jeu contient des données du procédé en fonctionnement normal et des données simulées dans les 15 différentes conditions de défauts, comme décrit dans la Table 3.19. Ainsi, 480 observations d'apprentissage sont disponibles pour le fonctionnement normal et pour chaque type de défaut et 800 observations pour le test.

Condition	Données d'apprentissage	Données de test
Normale $CFN$	480	800
Défaut $S_1$	480	800
Défaut $S_2$	480	800
...	480	800
Défaut $S_{15}$	480	800

TABLE 3.19 – Les jeux de données utilisés

### 3.5.3 Diagnostic du TEP par le modèle

La stratégie de détection et de diagnostic à base de modèle proposée pour le TEP est illustrée dans la Figure 3.20.

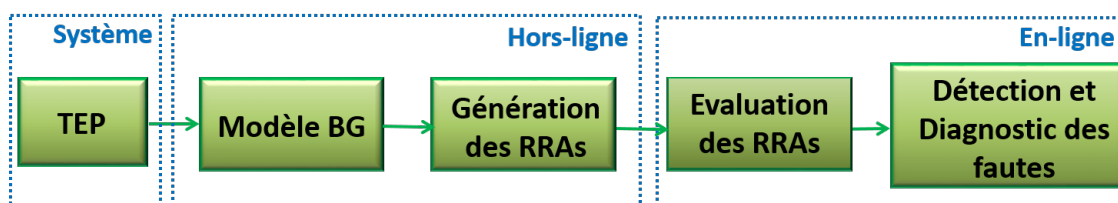


FIGURE 3.20 – Stratégie de détection et de diagnostic de défauts dans le TEP

La première étape consiste à développer un modèle BG pour le TEP. Ce modèle est utilisé pour obtenir les RRAs. L'évaluation numérique de ces RRAs produit des résidus qui sont utilisés pour la détection et le diagnostic des défauts.

#### Modélisation BG du TEP

Dans cette partie, les principales hypothèses prises en compte lors de la modélisation du TEP sont résumées. Ensuite, la première étape pour concevoir un modèle BG, qui consiste à construire un BG à mots, est présentée. Enfin, le BG du bilan massique du TEP est généré. Par souci de clarté, nous présentons ici uniquement le gaz réactif A pour expliquer les détails du modèle proposé. Le modèle mathématique est ensuite déduit du modèle graphique pour compléter la modélisation du procédé.

**Hypothèses** Les principales hypothèses tirées de la littérature [Lee et al., 2004, Downs and Vogel, 1993] et prises en compte pour construire le modèle BG sont les suivantes :

1. Les réactions se déroulent en phase vapeur.
2. L'équilibre est supposé lorsque les phases vapeur et liquide coexistent.
3. Toutes les phases sont bien mélangées.
4. Une zone de mélange contenant 5 flux : {1; 2; 3; 5; 8} (voir Figure 3.19) est définie et représente la zone où les gaz se mélangent avant d'alimenter le réacteur (flux 6). Cette zone est supposée être bien mélangée.
5. Le réacteur est considéré comme un réacteur à cuve agitée en continue.



6. Le compresseur n'est pas modélisé puisque le flux recyclé (flux 8) est considéré comme une variable indépendante.

Dans ce travail, le modèle chimique du TEP (correspondant au bilan massique) est présenté et utilisé à des fins de détection-diagnostic de défauts, tandis que les modèles hydrauliques et thermodynamiques sont uniquement développés et se trouvent dans l'Annexe 5.2.

**BG à mots du TEP** La première étape de la conception du modèle BG consiste en une phase de modélisation conceptuelle appelée BG à mots. Lors de cette phase, le système est décrit comme un ensemble de sous-systèmes interconnectés. Le BG à mots est déduit du schéma du TEP (Figure 3.19) et est présenté dans la Figure 3.21.

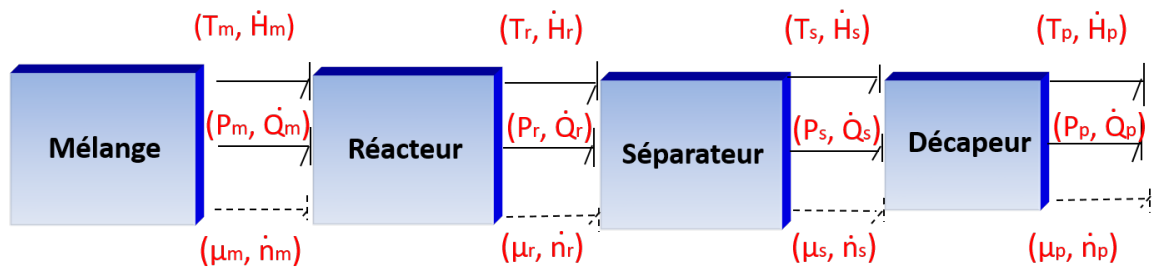


FIGURE 3.21 – BG à mots du TEP

Il visualise l'échange d'énergie ou de puissance entre les unités principales sans spécifier tous les détails concernant chaque partie du système. Cette représentation hiérarchique divise le TEP en quatre sous-systèmes, à savoir une zone de mélange fictive (comme décrit dans l'hypothèse 4), un réacteur, un séparateur et un décapeur. Le compresseur n'est pas modélisé ici comme indiqué dans l'hypothèse 6.

Les indices  $m; r; s; p$  correspondent respectivement aux zones de mélange, du réacteur, du séparateur et du décapeur. Les variables de puissance sont associées respectivement aux domaines chimique, hydraulique et thermodynamique, et représentent l'entrée et la sortie de chaque unité. Ainsi, les variables d'effort et de flux conjuguées sont respectivement le potentiel chimique ( $\mu$ ) et le débit molaire ( $\dot{n}$ ) pour le domaine chimique, la pression ( $P$ ) et le débit volumique ( $\dot{Q}$ ) pour le domaine hydraulique, et le couple (Température, Flux d'enthalpie) =  $(T, \dot{H})$  pour le domaine thermodynamique.

**Modèle BG du TEP** Comme mentionné précédemment, seul le cas particulier du réactif  $A$  est étudié en détail dans cette partie. Les modèles BG relatifs aux autres réactifs et produits sont présentés dans l'Annexe 5.2.

Ainsi, le modèle BG du réactif  $A$  est proposé en causalité dérivée dans la Figure 3.22.

Avant d'alimenter le réacteur, le gaz  $A$  est mélangé avec d'autres réactifs dans la zone de mélange. Ce gaz  $A$  est fourni par plusieurs sources, à savoir :

- La source de flux qui contient le gaz  $A$  (Flux 1 sur la Figure 3.19), qui est représentée par l'élément source de flux ( $S_f : F_1$ ) dans le BG (Figure 3.22). Le gaz traverse une vanne, représentée par la résistance ( $R : R_1$ ). La transformation de puissance lors du passage de la vanne est décrite par l'élément transformateur ( $TF(y_{A1})$ ). Le débit d'alimentation en  $A$  est ensuite mesuré. Cette mesure est représentée par la source de signal modulé ( $SS_f : F_1$ ) dans le BG considéré en causalité dérivée.

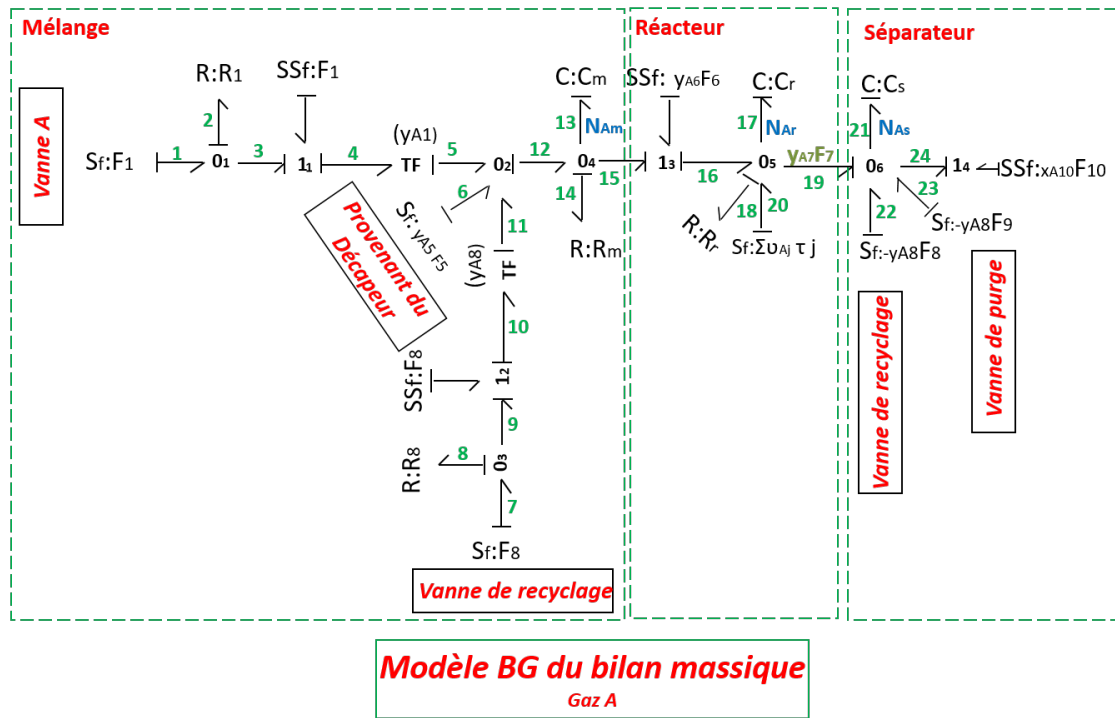


FIGURE 3.22 – Modèle BG du réactif A du TEP, en causalité dérivée

- Les dernières traces de réactifs éliminées par le décapeur (Flux 5 sur la Figure 3.19), représentées par la source ( $S_f : y_{A5}F_5$ ) contiennent une quantité de gaz A.
- Le flux recyclé (Flux 8 sur la Figure 3.19) contient de la vapeur A sortant du séparateur. Il est représenté dans le BG par l'élément source ( $S_f : F_8$ ). La vanne du compresseur est illustrée par l'élément ( $R : R_8$ ). L'élément ( $SS_f : F_8$ ) décrit la mesure du débit recyclé.

Après son passage dans la zone de mélange, le gaz A est ensuite introduit dans le réacteur. Le débit d'alimentation du réacteur (flux 6) est mesuré à travers l'élément ( $SS_f : y_{A6}F_6$ ). Le taux d'avancement de chaque réaction apparaît dans le BG à travers l'élément ( $S_f : \tau_j$ ). La vapeur sortante du séparateur est renvoyée vers la zone de mélange ( $S_f : y_{A8}F_8$ ) tandis qu'une partie est purgée ( $S_f : y_{A8}F_9$ ).

Les éléments  $\{C_m, C_r, C_s\}$  correspondent à l'énergie potentielle stockée dans chaque zone, alors que les éléments  $\{R_m, R_r, R_s\}$  modélisent la puissance dissipée dans les différentes unités. Enfin, les jonctions 1 et 0 sont utilisées pour lier les éléments BG tout en conservant l'énergie.

On peut constater que la partie "décapeur" n'est pas présente sur le modèle BG du gaz A. Ceci est dû au fait que le gaz A est un réactif qui doit être éliminé au niveau du séparateur et que seuls les produits passent normalement dans le décapeur.

Ainsi, le modèle BG représente l'architecture du TEP en tant que sous-systèmes interconnectés échangeant de l'énergie. Il montre la direction et la propagation de l'énergie ainsi que les relations de cause à effet. Cette description énergétique des phénomènes physiques impliqués dans le système permet de décrire son comportement. L'attribution de causalité au sein du modèle BG conduit à la génération des différentes équations constitutives des composants et des contraintes (équations algébriques et différentielles). Contrairement à l'approche mathématique, l'attribution de la causalité permet de résoudre les problèmes algorithmiques en indiquant comment les modèles mathématiques sont



calculés et mis en oeuvre.

**Modèle mathématique du TEP** Le modèle graphique BG permet de générer le modèle mathématique du TEP. En effet, à partir des jonctions 0 de chaque zone (mélange, réacteur, séparateur, décapeur), on peut déduire les caractéristiques essentielles du procédé, y compris les bilans énergétiques des cinq réactifs ( $A, B, C, D, E$ ) et des trois produits ( $F, G, H$ ). Les équations générales décrivant le comportement du système sont données ci-dessous :

$$\frac{dN_{im}}{dt} = \sum_{j=1,2,3,5,8} y_{ij}F_j - y_{i6}F_6 - \frac{\mu_{im}}{R_m} \quad (i = A, B, \dots, H) \quad (3.49)$$

$$\frac{dN_{ir}}{dt} = y_{i6}F_6 - y_{i7}F_7 + \sum_{j=1}^3 \nu_{ij}R_j - \frac{\mu_{ir}}{R_r} \quad (i = A, B, \dots, H) \quad (3.50)$$

$$\frac{dN_{is}}{dt} = y_{i7}F_7 - y_{i8}(F_8 + F_9) - x_{i10}F_{10} \quad (i = A, B, \dots, H) \quad (3.51)$$

$$\frac{dN_{ip}}{dt} = (1 - \phi_i)(x_{i10}F_{10} + y_{i4}F_4) - x_{i11}F_{11} \quad (i = G, H) \quad (3.52)$$

où  $N_{im}$ ,  $N_{ir}$ ,  $N_{is}$  et  $N_{ip}$  [kmol] sont les quantités de matière du gaz  $i$  dans la zone de mélange, du réacteur, du séparateur et du décapeur, respectivement.  $F_j$  [kmol/h] est le débit molaire du flux  $j$ ,  $y_{ij}$  et  $x_{ij}$  représentent respectivement la fraction molaire de la vapeur et la fraction molaire du liquide du composant  $i$  dans le flux  $j$ .  $\mu_{im}$  et  $\mu_{ir}$  [J/mol] sont les potentiels chimiques du gaz  $i$  dans la zone de mélange et du réacteur.  $\nu$  est un coefficient stoechiométrique,  $\phi$  est un facteur de décapage et  $\tau_j$  [kmol/h] est la vitesse de réaction du composant  $j$ .

Concentrons-nous sur le gaz réactif  $A$  (Figure 3.22) pour expliquer les équations différentielles ci-dessus.

Dans la zone de mélange, l'élément de stockage  $C_m$  apparaît à la jonction 0<sub>4</sub>. A partir de cette jonction, on déduit une contrainte exprimant la conservation du potentiel chimique (c'est-à-dire la conservation de l'effort) et une autre indiquant que la somme des débits molaires (c'est-à-dire la somme des flux) est nulle.

$$\begin{cases} e_{12} = e_{13} = e_{14} = e_{15} = \mu_{Am} \\ f_{12} - f_{13} - f_{14} - f_{15} = 0 \end{cases} \quad (3.53)$$

La première contrainte peut être écrite comme suit, en utilisant le déplacement généralisé  $N_{Am}$  représentant la quantité de matière du gaz  $A$  dans la zone de mélange :

$$e_{13} = \frac{1}{C_m} \int f_{13} dt = \frac{N_{Am}}{C_m} \quad (3.54)$$

Afin de développer la deuxième contrainte, on déduit à partir de la jonction 0<sub>2</sub> la contrainte qui lie les débits molaires :

$$f_{12} = f_5 + f_6 + f_{11} \quad (3.55)$$

On remplace alors l'expression de  $f_{12}$  dans la seconde contrainte comme suit :

$$f_{12} - f_{13} - f_{14} - f_{15} = 0 \Rightarrow f_5 + f_6 + f_{11} - f_{13} - f_{14} - f_{15} = 0 \quad (3.56)$$

Enfin, on élimine les variables inconnues en empruntant les chemins causaux comme suit :

$$\begin{cases} f_5 = y_{A1}F_1 \\ f_6 = y_{A5}F_5 \\ f_{11} = y_{A8}F_8 \\ f_{13} = \frac{dN_{Am}}{dt} \\ f_{14} = \frac{\mu_{Am}}{R_m} \\ f_{15} = y_{A6}F_6 \end{cases} \quad (3.57)$$

On en déduit l'équation suivante :

$$y_{A1}F_1 + y_{A5}F_5 + y_{A8}F_8 - \frac{dN_{Am}}{dt} - \frac{\mu_{Am}}{R_m} - y_{A6}F_6 = 0 \quad (3.58)$$

Il convient de noter que les fractions molaires de la vapeur du composant  $A$  dans les flux 2 et 3 sont nulles ( $y_{A2} = y_{A3} = 0$ ), comme mentionné dans [Downs and Vogel, 1993].

Le même raisonnement est appliqué pour le réacteur (à partir de la jonction  $0_5$ ) et la zone de séparation (à partir de la jonction  $0_6$ ), pour obtenir les équations du bilan énergétique du composant  $A$  :

$$\begin{cases} \frac{dN_{Am}}{dt} = y_{A1}F_1 + y_{A5}F_5 + y_{A8}F_8 - y_{A6}F_6 - \frac{\mu_{Am}}{R_m} \\ \frac{dN_{Ar}}{dt} = y_{A6}F_6 - y_{A7}F_7 + \sum_{j=1}^3 \nu_{Aj}R_j - \frac{\mu_{Ar}}{R_r} \\ \frac{dN_{As}}{dt} = y_{A7}F_7 - y_{A8}(F_8 + F_9) - x_{A10}F_{10} \end{cases} \quad (3.59)$$

On retrouve donc les équations 3.49, 3.50, 3.51 présentées ci-dessus. Comme expliqué auparavant, l'équation 3.52 ne concerne que les produits  $G$  et  $H$ .

### Génération des RRAs

La procédure de génération de RRAs est décrite comme suit : d'abord le BG est mis en causalité dérivée afin d'éviter le problème des conditions initiales inconnues. Il est à noter que contrairement au BG en causalité intégrale, le modèle BG en causalité dérivée est seulement utilisé pour la génération de RRAs et non pour la simulation physique. Ceci est réalisé en assignant une causalité différentielle préférée à chaque élément de stockage et en transformant les détecteurs ( $D_e$ ,  $D_f$ ) en sources de signal modulées  $SSf$ . Cela permet de les considérer dans le modèle BG comme des signaux imposés et donc comme un point de départ pour éliminer les variables inconnues, en les substituant par des variables mesurées et connues. Des détails sur cette procédure peuvent être trouvés dans [Chatti et al., 2014], [Chatti et al., 2013a].

Dans cette partie, nous discutons seulement de la génération des RRAs relatives au réactif  $A$ . Les autres RRAs sont facilement déduites en utilisant la même approche et sont données dans l'Annexe 5.2.

La première RRA pour le réactif  $A$  est déduite de la jonction  $0_2$  (voir Figure 3.22) :

$$f_5 + f_6 + f_{11} - f_{12} = 0. \quad (3.60)$$

Les variables inconnues  $f_5$ ,  $f_6$ ,  $f_{11}$  et  $f_{12}$  sont éliminées en parcourant les chemins causaux jusqu'aux variables connues. Par exemple  $f_5$  est éliminée en utilisant le chemin :  $f_5 \rightarrow TF : y_{A1} \rightarrow SSf : F_1$ .

$f_{12}$  peut être réécrite en utilisant la propriété de la jonction  $0_4$ , comme  $f_{12} = f_{13} + f_{14} + f_{15}$  avec  $f_{15} \rightarrow SSf : y_{A6}F_6$ . Cela conduit à la RRA suivante :

$$RRA_1 : y_{A1}F_1 + y_{A5}F_5 + y_{A8}F_8 - C_m \frac{d\mu_{Am}}{dt} - \frac{\mu_{Am}}{R_m} - y_{A6}F_6 = 0 \quad (3.61)$$

Les jonctions  $0_5$  et  $0_6$  dans la Figure 3.22 donnent lieu à une seconde RRA :

$$RRA_2 : y_{A6}F_6 - C_r \frac{d\mu_{Ar}}{dt} - \frac{\mu_{Ar}}{R_r} + \sum_{j=1}^3 \nu_{Aj} \tau_j - C_s \frac{d\mu_{As}}{dt} - y_{A8}(F_8 + F_9) - x_{A10}F_{10} = 0 \quad (3.62)$$

Par conséquent, les résidus du réactif  $A$  en régime permanent sont les suivants :

$$r_1 = y_{A1}F_1 + y_{A5}F_5 + y_{A8}F_8 - C_m \frac{d\mu_{Am}}{dt} - \frac{\mu_{Am}}{R_m} - y_{A6}F_6 \quad (3.63)$$

$$r_2 = y_{A6}F_6 - C_r \frac{d\mu_{Ar}}{dt} - \frac{\mu_{Ar}}{R_r} + \sum_{j=1}^3 \nu_{Aj} \tau_j - C_s \frac{d\mu_{As}}{dt} - y_{A8}(F_8 + F_9) - x_{A10}F_{10} \quad (3.64)$$

Il est important de noter qu'en raison de la complexité du TEP, la modélisation BG a présenté un compromis entre la précision et la complexité du modèle. En effet, le modèle BG proposé a été développé pour assurer un comportement réaliste du système tout en évitant une complexité trop importante du modèle. C'est la raison pour laquelle les incertitudes ne sont pas prises en compte par exemple. L'intégration des incertitudes, que ce soit sous forme LFT comme proposé dans [Djeziri et al., 2007] ou à travers les méthodes par intervalles comme développé dans [Chatti et al., 2016] fait partie des perspectives possibles concernant cette modélisation.

En outre, comme on ne dispose pas des données initiales concernant le TEP, les modèles BG développés sont en causalité dérivée et sont donc spécialement conçus pour des fins de détection et de diagnostic. La validation de ces modèles passe donc par l'évaluation des résidus avec des données expérimentales issues du fonctionnement normal du système. Si ces résidus évoluent autour de zéro, alors on peut considérer que le modèle est valide et reflète bien le fonctionnement nominal du système.

### Evaluation des RRAs

Afin d'évaluer les RRAs, les variables connues du TEP sont utilisées à la fois dans les conditions de fonctionnement normales (CFN) et dans les conditions de défauts.

**Conditions de fonctionnement normal (CFN)** En l'absence de défauts, tous les résidus sont théoriquement nuls comme l'illustre la Figure 3.23. Cette dernière montre les résidus  $r_1$  et  $r_2$  relatifs au réactif  $A$  en régime permanent. Les résidus en CFN des autres réactifs et produits sont donnés dans l'Annexe 5.2.

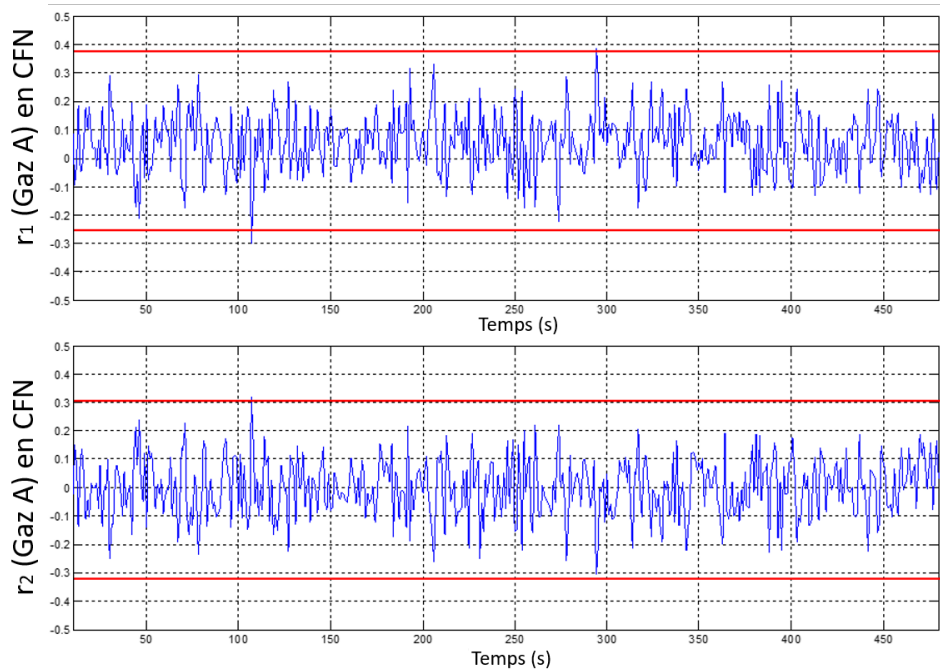


FIGURE 3.23 – Évaluation des résidus  $r_1$  et  $r_2$  relatifs au réactif  $A$ , dans les conditions de fonctionnement normal

L'évaluation des RRA avec des données expérimentales issues du fonctionnement sans défauts permet de valider les résidus obtenus à partir du modèle développé. En effet, un résidu en fonctionnement normal est cohérent avec le modèle du système s'il est nul ou entre certains seuils fixes (voir l'Annexe 5.2 pour l'évaluation des autres résidus).

Les seuils fixes sont obtenus en suivant la procédure de décision simple qui consiste à prendre un seuil égal à  $\mu \pm 3\sigma$  (où  $\mu$  est la moyenne et  $\sigma$  est l'écart-type) du résidu dans les CFNs. Ainsi, pour le résidu  $r_1$ , les limites supérieures et inférieures sont respectivement égales à 0.3644 et  $-0.2600$  et pour le résidu  $r_2$ , elles sont égales à 0.3073 et  $-0.3108$ . Ces seuils garantissent un taux théorique de fausse alarme (FAR) de 0,27%.

**Conditions de défauts** Afin de déterminer les défauts qui peuvent être détectés et isolés par le BG, une Matrice de Signature de Défauts (FSM pour Fault Signature Matrix) est construite. Cette dernière relie les défauts aux résidus.

L'approche classique pour construire une FSM considère que tous les éléments apparaissant dans l'expression du résidu sont associés à des défauts. Par conséquent, l'ensemble des défauts pouvant être détecté et isolé est défini par les résidus générés. Par exemple, si on considère les résidus  $r_1$  et  $r_2$  liés au gaz  $A$ , le système de surveillance peut détecter des défauts impliquant les potentiels chimiques ( $\mu_{Am}$ ,  $\mu_{Ar}$ ), les capteurs qui mesurent plusieurs flux (débit d'alimentation  $F_1$ , débit à la sortie du décapeur  $F_5$ , débit de recyclage  $F_8$ , débit d'alimentation du réacteur  $F_6$ , taux de purge  $F_9$ , flux à la sortie du séparateur  $F_{10}$ ), la cinétique de la réaction à travers la vitesse de réaction  $\tau_j$ , les vannes ( $R_m$ ,  $R_r$ ) et les fuites qui peuvent se produire au niveau de la zone de mélange, du réacteur et du séparateur ( $C_m$ ,  $C_r$ ,  $C_s$ ).

Cependant, ces défauts ne font pas partie des défauts détectables par les approches à base de données. En effet, les données de défauts disponibles sont détaillées dans la Table

3.18. Comme l'objectif de ce chapitre est de développer et fusionner les approches basées sur le modèle et les données, il convient de considérer le même type de défauts.

Par conséquent, une autre manière de construire la FSM a été adoptée ici. Une campagne de tests a été réalisée sur la base des données d'apprentissages disponibles et représentatives des différents scénarios de défauts (voir Table 3.18), afin d'analyser la manière dont les résidus évoluent en présence de ces défauts.

Prenons l'exemple du résidu  $r_1$ . Ce dernier a été évalué en utilisant des mesures impliquées dans son expression, soit  $(F_1; F_5; F_8; F_6)$  pour chaque scénario de défaut  $(S_1; S_2; \dots; S_{15})$ . Il est apparu que  $r_1$  réagit en présence de données provenant de 6 scénarios de défauts :  $(S_1; S_6; S_7; S_8; S_{12}; S_{13})$ . Le même raisonnement est réalisé pour les autres résidus générés.

Ainsi, les données ont été utilisées pour mettre en évidence le lien entre les défauts et les résidus. Les résultats sont donnés dans la FSM, représentée dans la Table 3.20.

	$r_1$	$r_2$	$r_3$	$r_4$	$r_5$	$r_6$	$r_7$	$r_8$	$r_9$	$r_{10}$	$r_{11}$	$r_{13}$	$r_{14}$	$r_{16}$	$r_{17}$
$S_1$	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
$S_2$	0	1	0	1	0	1	0	1	1	0	1	1	1	1	1
$S_5$	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0
$S_6$	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0
$S_7$	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0
$S_8$	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0
$S_{12}$	1	1	0	0	1	1	0	1	1	0	1	0	0	0	0
$S_{13}$	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0

TABLE 3.20 – Matrice de Signature de Défauts (FSM) reliant les défauts aux résidus

Si le résidu est sensible à un défaut (c'est-à-dire si le résidu réagit et dépasse ses seuils fixes), alors la valeur de la colonne correspondant à ce résidu est égale à 1. Il est à noter que 3 résidus ( $r_{12}; r_{15}, r_{18}$ ) n'ont réagi à aucun défaut considéré, ils ne sont donc pas représentés dans la FSM et ne sont pas utilisés pour surveiller le TEP. Un vecteur de 15 résidus ( $r_1; r_2; r_3; r_4; r_5; r_6; r_7; r_8; r_9; r_{10}; r_{11}; r_{13}; r_{14}; r_{16}; r_{17}$ ) représente donc la signature de défaut (FS pour Fault Signature).

Ainsi, parmi les 15 défauts que le TEP peut simuler (Table 3.18), 8 défauts ( $S_1, S_2, S_5, S_6, S_7, S_8, S_{12}, S_{13}$ ) ont été identifiés comme potentiellement sensibles aux résidus générés et peuvent donc être détectés et isolés sur la base du modèle BG, puisque leurs FS sont uniques. Par conséquent, nous allons nous intéresser à la détection et au diagnostic de cet ensemble de 8 défaillances par la suite.

### Détection et diagnostic des défauts en-ligne

La Table 3.20 déduite précédemment est utilisée dans la stratégie de détection et de diagnostic des défauts basée sur le BG, comme montré dans la Figure 3.24..

Afin de détecter un défaut dans le TEP, au moins 1 des 15 résidus générés doit réagir. Cependant, le diagnostic d'un défaut nécessite que la signature du défaut FS ( $r_1; r_2; r_3; r_4; r_5; r_6; r_7; r_8; r_9; r_{10}; r_{11}; r_{13}; r_{14}; r_{16}; r_{17}$ ) corresponde à l'une des signatures de la FSM (Table 3.20). Cela implique que, pour isoler le défaut  $S_1$ , par exemple, les résidus  $r_1$  et  $r_3$  (et uniquement eux) doivent réagir simultanément.

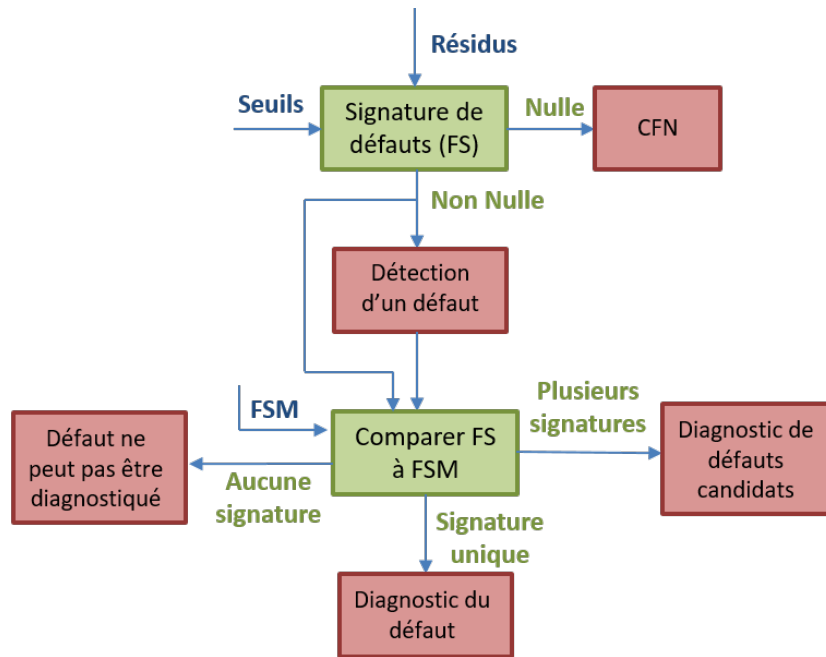


FIGURE 3.24 – Stratégie de FDD basée sur le BG en ligne

Les résultats de la simulation sont donnés dans la Table 3.21. La première ligne correspond au  $FAR$ , qui représente le pourcentage d’observations de résidus en  $CFN$  ayant dépassé le seuil prédéfini. Les autres lignes correspondent au  $FDR$  relatif à chaque classe de défaut. Par exemple, le  $FDR$  relatif au défaut  $S_1$  (voir Table 3.20) est calculé comme le pourcentage d’observations de résidus  $r_1$  et  $r_3$  (et uniquement eux) ayant réagi simultanément.

<b>FAR</b>	<b>1.53</b>
$S_1$	89.04
$S_2$	72.10
$S_5$	83.18
$S_6$	90.32
$S_7$	82.17
$S_8$	83.57
$S_{12}$	73.76
$S_{13}$	70.70
<b>FDR</b>	<b>80.61</b>

TABLE 3.21 – Résultats FDD avec la stratégie BG (en %)

On constate que l’approche BG présente de bonnes performances en FDD avec un faible  $FAR$  de 1.53% et un  $FDR$  de 80.61% [Tidhiri et al., 2018]. En outre, il convient de noter que 15 résidus différents ont été utilisés pour diagnostiquer chaque défaut, ce qui peut augmenter les erreurs de classification. Par exemple, pour isoler le défaut  $S_2$ , 10 résidus ( $r_2, r_4, r_6, r_8, r_9, r_{11}, r_{13}, r_{14}, r_{16}, r_{17}$ ) doivent réagir simultanément. Si l’un des 10 résidus ne réagit pas en même temps que les autres, l’observation ne sera pas considérée comme appartenant à  $S_2$ .

Afin d’améliorer cette prise de décision, deux propositions peuvent être considérées. D’un côté, il est possible de diagnostiquer les 8 défauts en utilisant un sous-ensemble de

résidus ayant une signature de défaut unique. Pour déterminer ce sous-ensemble optimal de résidus, plusieurs algorithmes peuvent être utilisés comme celui basé sur l'information mutuelle entre les résidus, développé dans [Verron et al., 2008b]. D'un autre côté, une prise de décision basée sur les classifieurs (AD, kNN, arbres de décisions etc.) peut être envisagée. Dans ce cas, le classifieur utilise les résidus évalués avec les données d'apprentissage du TEP pour construire son modèle d'apprentissage. Ce modèle est ensuite utilisé pour la surveillance des résidus évalués avec les nouvelles données de test. C'est cette dernière proposition que nous choisissons d'utiliser pour la suite.

Par conséquent, pour présenter la décision de la méthode basée sur le modèle BG à chaque pas de temps, l'AD linéaire (voir Partie 2.2.2) est appliquée aux résidus.

Les résultats de la simulation sont donnés sous la forme d'une matrice de confusion (voir Table 3.22), permettant la visualisation de la performance FDD.

	$CFN$	$S_1$	$S_2$	$S_5$	$S_6$	$S_7$	$S_8$	$S_{12}$	$S_{13}$
$CFN$	99.74	0	0	0	0.13	0	0	0.13	0
$S_1$	10.83	89.17	0	0	0	0	0	0	0
$S_2$	4.20	0	0	0	0	1.4	0.38	1.02	93
$S_5$	2.93	0	0	96.56	0	0	0.38	0.13	0
$S_6$	1.66	0	0	0	98.09	0	0	0.25	0
$S_7$	0	0	0	0	0	100	0	0	0
$S_8$	0.13	0	0	0.13	0	0	99.74	0	0
$S_{12}$	0.25	0	0	0	0	0.25	0	99.5	0
$S_{13}$	0	0	0	0	0	0	0	0	100

TABLE 3.22 – Matrice de confusion de la méthode  $m_1$ , basée sur le BG et l'AD (en %)

Selon la matrice de confusion, les nouvelles performances en FDD sont améliorées puisque le FDR et le FAR sont respectivement de 85.3825% et 0.26% (contre 80.61% et 1.53% pour le BG uniquement).

### 3.5.4 Diagnostic du TEP par les données

La méthode basée sur les données choisie pour surveiller le TEP est une AD linéaire qui utilise des données historiques pour construire un modèle d'apprentissage. Ce dernier est utilisé pour la surveillance des nouvelles données.

La matrice de confusion est donnée dans la Table 3.23.

	$CFN$	$S_1$	$S_2$	$S_5$	$S_6$	$S_7$	$S_8$	$S_{12}$	$S_{13}$
$CFN$	78.98	0	0	3.57	0	0	5.22	12.23	0
$S_1$	0.25	99.75	0	0	0	0	0	0	0
$S_2$	1.91	0	98.09	0	0	0	0	0	0
$S_5$	2.42	0	0	96.69	0	0	0.25	0.64	0
$S_6$	0	0	0	0	100	0	0	0	0
$S_7$	0	0	0	0	0	100	0	0	0
$S_8$	33.12	0.51	0	1.40	0	0	30.70	32.23	2.04
$S_{12}$	31.34	0	0	9.04	0	0	12.99	42.29	4.33
$S_{13}$	14.65	0	0	1.53	0	0	22.42	30.03	31.34

TABLE 3.23 – Matrice de confusion de la méthode 2, basée sur l'AD (en %)



Selon la matrice de confusion, le FDR et le FAR sont respectivement 74.8575% et 21.02%.

### 3.5.5 Diagnostic du TEP par la fusion de décisions

Dans cette partie, nous allons suivre les 3 étapes de la méthodologie proposée auparavant, afin de fusionner les décisions des deux méthodes développées pour le diagnostic du TEP.

1. **Evaluation des performances individuelles** La première étape de la méthodologie de fusion proposée est l'évaluation des performances individuelles.

D1	CFN	S <sub>1</sub>	S <sub>2</sub>	S <sub>5</sub>	S <sub>6</sub>	S <sub>7</sub>	S <sub>8</sub>	S <sub>12</sub>	S <sub>13</sub>	D2	CFN	S <sub>1</sub>	S <sub>2</sub>	S <sub>5</sub>	S <sub>6</sub>	S <sub>7</sub>	S <sub>8</sub>	S <sub>12</sub>	S <sub>13</sub>	
CFN	99.74	0	0	0	0.13	0	0	0.13	0	CFN	78.98	0	0	3.57	0	0	5.22	12.23	0	
S <sub>1</sub>	10.83	89.17	0	0	0	0	0	0	0	S <sub>1</sub>	0.25	99.75	0	0	0	0	0	0	0	0
S <sub>2</sub>	4.20	0	0	0	0	1.4	0.38	1.02	93	S <sub>2</sub>	1.91	0	98.09	0	0	0	0	0	0	0
S <sub>5</sub>	2.93	0	0	96.56	0	0	0.38	0.13	0	S <sub>5</sub>	2.42	0	0	96.69	0	0	0.25	0.64	0	0
S <sub>6</sub>	1.66	0	0	0	98.09	0	0	0.25	0	S <sub>6</sub>	0	0	0	0	100	0	0	0	0	0
S <sub>7</sub>	0	0	0	0	0	100	0	0	0	S <sub>7</sub>	0	0	0	0	0	100	0	0	0	0
S <sub>8</sub>	0.13	0	0	0.13	0	0	99.74	0	0	S <sub>8</sub>	33.12	0.51	0	1.40	0	0	30.70	32.23	2.04	0
S <sub>12</sub>	0.25	0	0	0	0	0.25	0	99.5	0	S <sub>12</sub>	31.34	0	0	9.04	0	0	12.99	42.29	4.33	0
S <sub>13</sub>	0	0	0	0	0	0	0	0	100	S <sub>13</sub>	14.65	0	0	1.53	0	0	22.42	30.03	31.34	0

FIGURE 3.25 – Matrices de confusion des deux méthodes FDD

A partir des matrices de confusion (Figure 3.25), le seuil de performance est fixé à 96%. Ceci implique que les classes dont le FDR relatif est inférieur à 96% ne seront pas considérées comme détectables avec de "bonnes" performances par la méthode en question. C'est le cas des classes (CFN, S<sub>8</sub>, S<sub>12</sub>, S<sub>13</sub>) pour la méthode à base de données ou encore (S<sub>1</sub>, S<sub>2</sub>) pour la méthode basée sur le BG.

Ainsi, le choix des types de classe de fonctionnement se fait de la façon suivante :

- La méthode basée sur le modèle a de meilleures performances pour les classes (CFN, S<sub>8</sub>, S<sub>12</sub>, S<sub>13</sub>), donc nous choisissons  $F_1 = \{CFN, S_8, S_{12}, S_{13}\}$  (en vert sur la Figure 3.25).
- La méthode basée sur les données présente de meilleures performances que la méthode 1 pour les classes (S<sub>1</sub>, S<sub>2</sub>, S<sub>6</sub>), ce qui conduit à  $F_2 = \{S_1, S_2, S_6\}$  (en bleu sur la Figure 3.25).
- Les deux méthodes peuvent détecter et isoler S<sub>5</sub> et S<sub>7</sub> avec des performances élevées et comparables. Par conséquent :  $F_0 = \{S_5, S_7\}$  (en rouge sur la Figure 3.25).

2. **Stratégie de FDD** Le cadre générique est présenté dans la Figure 3.26.

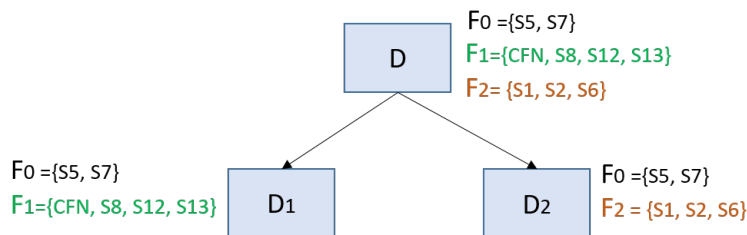


FIGURE 3.26 – Cadre générique de fusion

Le choix des paramètres du RB (probabilités conditionnelles des nœuds D<sub>1</sub> et D<sub>2</sub>) dépend de la stratégie adoptée puisque les évidences incertaines sont prises en compte.



Pour une stratégie privilégiant le FAR (respectivement le FDR) les 96 solutions déduites (voir partie 3.4.2) (respectivement les 81 solutions déduites (voir partie 3.4.2)) sont testées dans l'étape 3, c'est-à-dire l'étape de fusion de décisions.

**3. Fusion des décisions** La dernière étape de la méthodologie proposée consiste à utiliser le RB paramétré en ligne pour la fusion des deux décisions (Figure 3.27). Ces dernières sont données à chaque instant par les deux méthodes et sont introduites au niveau des noeuds  $D_1$  et  $D_2$ .



FIGURE 3.27 – Fusion de décisions en ligne

**Stratégie privilégiant le FAR** Dans cette partie, la stratégie adoptée pour le diagnostic du TEP est celle qui privilégie le FAR. Le tableau des objectifs est celui de la Figure 3.11a et par conséquent les 96 combinaisons de paramètres du RB sont testées. Les FARs déduits pour chacune de ces solutions sont donnés dans la Figure 3.28.

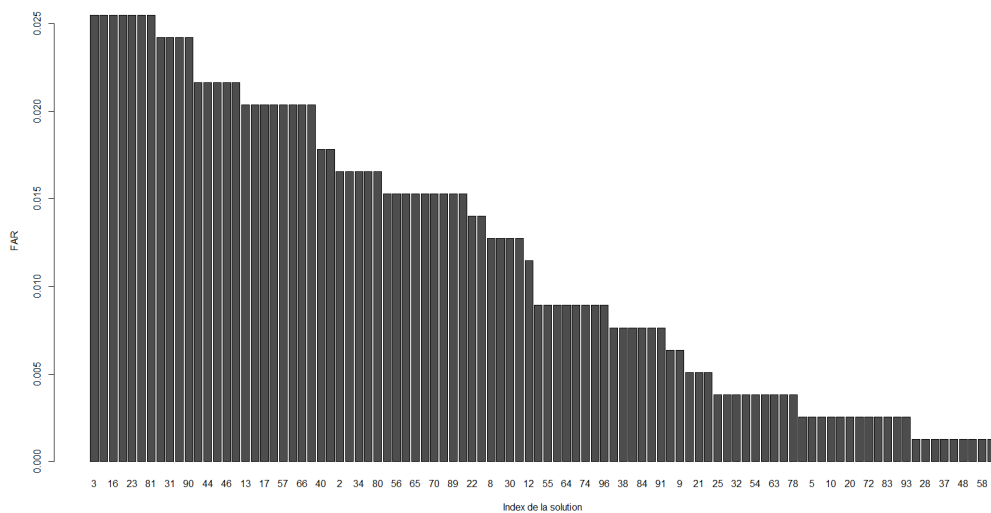


FIGURE 3.28 – FARs obtenus pour chaque solution

On constate que le FAR le plus petit est de 0.13%. Plusieurs solutions (24; 28; 35; 37; 42; 48; 53; 58; 82) mènent à ce FAR. Par ailleurs, celle qui présente le meilleur FDR, soit 98.11%, est la solution 82. Par conséquent, la matrice de confusion de la décision globale  $D$  pour cette solution est montrée dans la Table 3.24.

	$CFN$	$S_1$	$S_2$	$S_5$	$S_6$	$S_7$	$S_8$	$S_{12}$	$S_{13}$
$CFN$	99.87	0	0	0	0	0	0	0.13	0
$S_1$	0.38	95.29	0	0	0	0	4.33	0	0
$S_2$	0	0	94.40	0	0	0	3.69	0	1.91
$S_5$	3.57	0	0	95.92	0	0	0.38	0.13	0
$S_6$	0	0	0	0	100	0	0	0	0
$S_7$	0	0	0	0	0	100	0	0	0
$S_8$	0.25	0	0	0	0	0	99.75	0	0
$S_{12}$	0.50	0	0	0	0	0	0	99.5	0
$S_{13}$	0	0	0	0	0	0	0	0	100

TABLE 3.24 – Matrice de confusion de l’approche de fusion proposée (en %), pour une stratégie privilégiant le FAR

**Stratégie privilégiant le FDR** Dans cette partie, la stratégie adoptée pour le diagnostic du TEP est celle qui privilégie le FDR. Le tableau des objectifs est celui de la Figure 3.11b et par conséquent les 81 combinaisons de paramètres du RB sont testées. Les FDRs déduits pour chacune de ces solutions sont donnés dans la Figure 3.29.

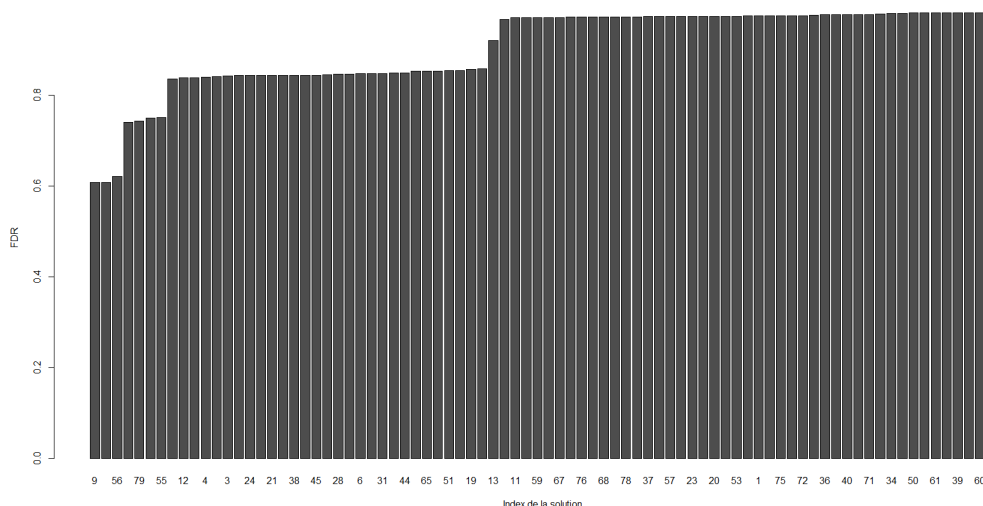


FIGURE 3.29 – FDRs obtenus pour chaque solution

On remarque que le FDR le plus grand est de 98.23%. La seule solution qui mène à ce FDR est la solution 16 et présente un FAR de 0.26%. Ainsi, la matrice de confusion de la décision globale  $D$  est montrée dans la Table 3.25.

	$CFN$	$S_1$	$S_2$	$S_5$	$S_6$	$S_7$	$S_8$	$S_{12}$	$S_{13}$
$CFN$	99.74	0	0	0.13	0	0	0	0.13	0
$S_1$	0.38	95.54	0	0	0	0	4.08	0	0
$S_2$	0	0	94.40	0	0	0	3.69	0	1.91
$S_5$	2.93	0	0	96.69	0	0	0.38	0	0
$S_6$	0	0	0	0	100	0	0	0	0
$S_7$	0	0	0	0	0	100	0	0	0
$S_8$	0.13	0	0	0.13	0	0	99.74	0	0
$S_{12}$	0.25	0	0	0.25	0	0	0	99.5	0
$S_{13}$	0	0	0	0	0	0	0	0	100

TABLE 3.25 – Matrice de confusion de l’approche de fusion proposée (en %), pour une stratégie privilégiant le FDR

Par conséquent, on constate que la méthode de fusion proposée donne de très bonnes performances en diagnostic des défauts considérés dans le TEP, quelle que soit la stratégie adoptée. En effet, le FAR et le FDR sont améliorés dans les deux cas :

- Stratégie FAR :  $FAR = 0.13\%$  et  $FDR = 98.11\%$
- Stratégie FDR :  $FAR = 0.26\%$  et  $FDR = 98.23\%$

Une comparaison est donnée dans la Table 3.26, et reflète l’amélioration des performances en diagnostic qui peut être atteinte avec la méthodologie proposée.

	Method 1	Method 2	Fusion (Stratégie FAR)	Fusion (Stratégie FDR)
FDR	85.38	74.86	98.11	<b>98.23</b>
FAR	0.26	21.02	<b>0.13</b>	0.26

TABLE 3.26 – Comparaison de performances (en %)

**Comparaison avec la littérature** Une étude comparative incluant l’approche de fusion proposée et les résultats rapportés dans la littérature est détaillée dans la Table 3.27. Cette dernière récapitule les performances en diagnostic suivant le FDR pour plusieurs méthodes : RN [Eslamloueyan, 2011], SVM [Jing and Hou, 2015], ACP [Jing and Hou, 2015], RN [Eslamloueyan, 2011], RB [Verron, 2007], BG [Tidri et al., 2018] (voir Partie 3.5.3), les méthodes individuelles développées dans ce travail (voir Parties 3.5.3 pour le BG couplé à l’AD et 3.5.4 pour l’AD) et l’approche de fusion proposée avec les deux stratégies FAR et FDR (voir Partie 3.5.5).

Il faut noter que la performance pour la classe de fonctionnement normal  $CFN$  n’est pas toujours mentionnée dans ces articles car ils se concentrent uniquement sur la performance en diagnostic. Par conséquent, la classe  $CFN$  n’apparaît pas dans la Table 3.27 (voir Table 3.26 pour les FARs liés à notre approche). En outre, il convient de mentionner que les méthodes basées sur les données considérées pour cette comparaison n’ont pas forcément le même contexte expérimental que les approches développées dans notre travail. En effet, certains de ces travaux ont pris en compte d’autres perturbations pouvant affecter le TEP dans leurs résultats de diagnostic. Par conséquent, cette Table 3.27

doit être considérée comme une base de comparaison relative. Enfin, la couleur rouge indique le meilleur FDR pour chaque classe de défaut.

	$RN$	$SVM$	$ACP$	$RN$	$RB$	$BG$	$BG + AD$	$AD$	$Fusion_{FAR}$	$Fusion_{FDR}$
$S_1$	81	87.19	87.19	98	97.5	89.04	89.17	<b>99.75</b>	95.29	95.54
$S_2$	82	85.83	87.5	<b>99</b>	98.13	72.10	0	98.09	94.40	94.40
$S_5$	73	57.60	4.06	<b>98</b>	<b>98</b>	83.18	96.56	96.69	95.92	96.69
$S_6$	84	78.85	90.21	<b>100</b>	<b>100</b>	90.32	98.09	<b>100</b>	<b>100</b>	<b>100</b>
$S_7$	82	88.85	89.69	<b>100</b>	<b>100</b>	82.17	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
$S_8$	45	32.19	85	60	97	83.57	<b>99.75</b>	30.70	<b>99.75</b>	99.74
$S_{12}$	22	50.21	85.83	44	99.25	73.76	99.49	42.29	<b>99.5</b>	<b>99.5</b>
$S_{13}$	16	21.46	69.06	34	76.38	70.70	<b>100</b>	31.34	<b>100</b>	<b>100</b>
<b>FDR</b>	<b>60.63</b>	<b>62.77</b>	<b>74.82</b>	<b>79.13</b>	<b>95.78</b>	<b>80.61</b>	<b>85.38</b>	<b>74.86</b>	<b>98.11</b>	<b>98.23</b>

TABLE 3.27 – Comparaison des performances en diagnostic de défauts (%) pour le TEP

Selon cette comparaison, l'approche proposée de fusion a montré les meilleures performances FDD, avec le meilleur FDR.

## 3.6 Conclusion

Dans le domaine de la fusion de décisions pour le diagnostic des systèmes, de nombreux travaux ont été publiés. La plupart des approches utilisées dans ces travaux sont dédiées à des applications particulières et la base théorique reste toujours manquante.

Ce chapitre a permis de présenter nos principales contributions théoriques apportées au domaine du diagnostic de défauts des systèmes. En effet, une méthodologie théorique et générique permettant la fusion de décisions de deux méthodes de surveillance a été proposée. Cette méthodologie se base sur les performances a priori des méthodes individuelles pour définir un ensemble d'objectifs à respecter. Le RB est utilisé comme cadre générique de fusion et ses paramètres sont déduits théoriquement de sorte à atteindre les objectifs choisis. Cette déduction de paramètres conduit à un problème multi-objectif, résolu avec une approche lexicographique.

Une fois paramétré, le RB est testé à travers un procédé chimique complexe, le TEP. L'application a montré la faisabilité de la méthodologie proposée et sa capacité à résoudre efficacement les conflits entre les différentes méthodes de surveillance. Selon les résultats obtenus, la fusion surpasse les décisions individuelles avec un taux élevé de diagnostic des défauts FDR (98.23%) et le taux de fausse alarme (FAR) le plus bas (0.13%). Par conséquent, notre stratégie de fusion offre une amélioration significative des performances de diagnostic.

# Fusion de décisions pour le pronostic

## Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>108</b>
<b>4.2</b>	<b>Formulation du problème</b>	<b>109</b>
<b>4.3</b>	<b>Démarche méthodologique de pronostic</b>	<b>111</b>
4.3.1	Stratégie d'apprentissage	112
4.3.2	Estimation de l'état de santé et prédiction de la RUL	115
<b>4.4</b>	<b>Démarche méthodologique de fusion</b>	<b>119</b>
4.4.1	Evaluation des performances individuelles	120
4.4.2	Stratégie de pronostic : Définition des objectifs	122
4.4.3	Stratégie de pronostic : Cadre générique de fusion	123
4.4.4	Stratégie de pronostic : Développement des objectifs	123
4.4.5	Stratégie de pronostic : Prise de décision multi-objectifs sous contraintes	127
4.4.6	Fusion des décisions en ligne	128
<b>4.5</b>	<b>Synthèse</b>	<b>131</b>
<b>4.6</b>	<b>Application : Pronostic de turboréacteurs à double flux</b>	<b>132</b>
4.6.1	Introduction	132
4.6.2	Présentation du turboréacteur à double flux	132
4.6.3	Pronostic de turboréacteurs par des méthodes individuelles	134
4.6.4	Pronostic de turboréacteurs par la fusion de décision	139
<b>4.7</b>	<b>Discussion</b>	<b>141</b>
4.7.1	Prédiction de la RUL	141
4.7.2	Analyse des paramètres obtenus	147
4.7.3	Comparaison avec la littérature	153
<b>4.8</b>	<b>Conclusion</b>	<b>155</b>

---

## 4.1 Introduction

Le pronostic est l'une des tâches majeures dans la surveillance des systèmes. Il est étroitement lié au diagnostic et consiste à prédire une certaine caractéristique d'intérêt, qui est généralement la durée de vie restante (RUL), en fonction des informations disponibles.

Au cours de ces dernières années, beaucoup de travaux de recherche s'intéressent au pronostic [Lei et al., 2018, Le Son et al., 2013, Ramasso and Gouriveau, 2010, Yu et al., 2012, Ramasso and Saxena, 2014]. En effet, une stratégie efficace de pronostic peut réduire les opérations inutiles de maintenance et améliorer la fiabilité des systèmes. Comme pour le diagnostic, le choix d'une méthode de pronostic dépend particulièrement de l'information disponible sur le système considéré.

L'approche de pronostic basée sur un modèle développe tout d'abord une représentation mathématique du système, puis prédit son évolution future [Qian et al., 2017]. Lorsque ce modèle atteint un seuil prédéfini, on en déduit que le composant ou système surveillé est entré dans un état défaillant. Cette approche offre généralement des résultats précis puisqu'elle intègre dans son raisonnement une connaissance du comportement du système à surveiller. Cependant, il est très difficile d'obtenir le modèle dynamique analytique qui prend en compte le phénomène de dégradation. De plus, les modèles pour le pronostic sont souvent dédiés à des mécanismes de dégradations spécifiques [Jha et al., 2016]. Pour toutes ces raisons, l'intérêt pour l'approche pilotée par les données est justifié puisque le cadre applicatif des méthodes à base de modèle est restreint en pratique.

L'objectif du pronostic à base de données consiste à extrapoler l'évolution des indicateurs de dégradation dans le but de suivre l'état de santé du système et de prédire les éventuelles défaillances qui pourraient survenir. L'estimation de la RUL peut être déduite en calculant le temps qui sépare l'état courant du système de l'état correspondant à sa défaillance. L'approche à base de données offre de nombreux avantages par rapport à une approche basée sur un modèle. En effet, elle est plus simple à mettre en oeuvre et son cadre applicatif est beaucoup plus large puisqu'elle ne requiert pas le développement mathématique d'un modèle et qu'elle convient à toutes les applications qui disposent d'une quantité suffisante de données [Zhang et al., 2017, Specht, 1990].

Dans ce chapitre, une approche de pronostic basée sur les données est proposée. Par la suite, le même cadre générique ayant permis la fusion de décisions de deux méthodes de diagnostic de défauts est adapté pour le pronostic. La méthodologie de fusion est basée sur un calcul théorique des paramètres du RB, en fonction des objectifs de pronostic à atteindre. Le développement conduit à un problème multi-objectif sous contraintes, résolu par la même approche lexicographique. L'efficacité de l'approche de fusion décisionnelle proposée pour le pronostic est validée sur un problème applicatif très épineux qui a fait l'objet d'un "challenge" proposé par la NASA [Frederick et al., 2007] et qui consiste à estimer les RULs d'un ensemble de turboréacteurs à double flux. L'application souligne la capacité de l'approche à estimer les RULs avec de meilleures performances que la plupart des autres résultats publiés dans la littérature.

Le chapitre est organisé comme suit : tout d'abord, nous présentons une formulation générale du problème de fusion de décisions pour le pronostic et précisons les hypothèses sur lesquelles nous nous basons. Nous décrivons ensuite la méthodologie de pronostic

proposée. Enfin, nous proposons la démarche méthodologique de prise de décision pour le pronostic de défauts ainsi que le cadre générique de fusion en détaillant sa structure et ses paramètres. L'efficacité de l'approche est montrée à travers l'application mentionnée précédemment.

## 4.2 Formulation du problème

On considère ici qu'un système complexe peut comporter un ou plusieurs composants critiques. Ce dernier est défini comme suit :

**Définition 19.** *Un **composant critique** est un composant dont la défaillance entraîne des conséquences graves comme l'indisponibilité du système par exemple. Il peut également s'agir d'un composant dont le taux de défaillance est élevé.*

Le composant critique est particulièrement surveillé dans le cadre du pronostic. D'où l'hypothèse suivante :

**Hypothèse 4.** *La **défaillance** du système est due à la défaillance d'un ou de plusieurs de ses composants critiques.*

On introduit un ensemble de deux méthodes  $M = \{m_1, m_2\}$  à fusionner pour le pronostic du système considéré. Ces méthodes estiment individuellement la RUL du système, en exploitant les données de surveillance disponibles. Les approches basées sur les données peuvent utiliser plusieurs outils afin de prédire la RUL, comme les RNs, les processus Markoviens, ou encore la classification supervisée.

D'où l'hypothèse suivante :

**Hypothèse 5.** *Les **méthodes individuelles** de pronostic exploitent les données et reposent sur le principe de classification supervisée.*

Généralement, on considère que la durée de vie du système ou du composant considéré est segmentée en plusieurs modes, qu'on définit comme suit :

**Définition 20.** *Un **mode** décrit l'évolution du système dans des conditions de santé spécifiques.*

Tout d'abord, le système est dans un mode nominal (correspondant aux conditions de fonctionnement normal CFN), puis dans un mode de dégradation qui ne lui permet plus de revenir dans un mode nominal sans réparation. D'où l'hypothèse suivante :

**Hypothèse 6.** *Les **dégradations** sont progressives et persistantes.*

Enfin, le système converge vers un mode de défaillance, comme illustré par la Figure 4.1.

On définit alors un ensemble de  $n$  états de santé (Health State)  $HS = \{hs_1, hs_2, \dots, hs_n\}$  caractérisant la dégradation du système :

**Définition 21.** *Un **état de santé** représente l'état de dégradation dans lequel se trouve le système surveillé à l'instant considéré  $t$ .*

Ainsi, les états de santé représentent les classes à discriminer par les méthodes individuelles. La Figure 4.1 illustre ces états de santé.

Ceci nous permet d'introduire la notion de *décision* d'une méthode de pronostic.



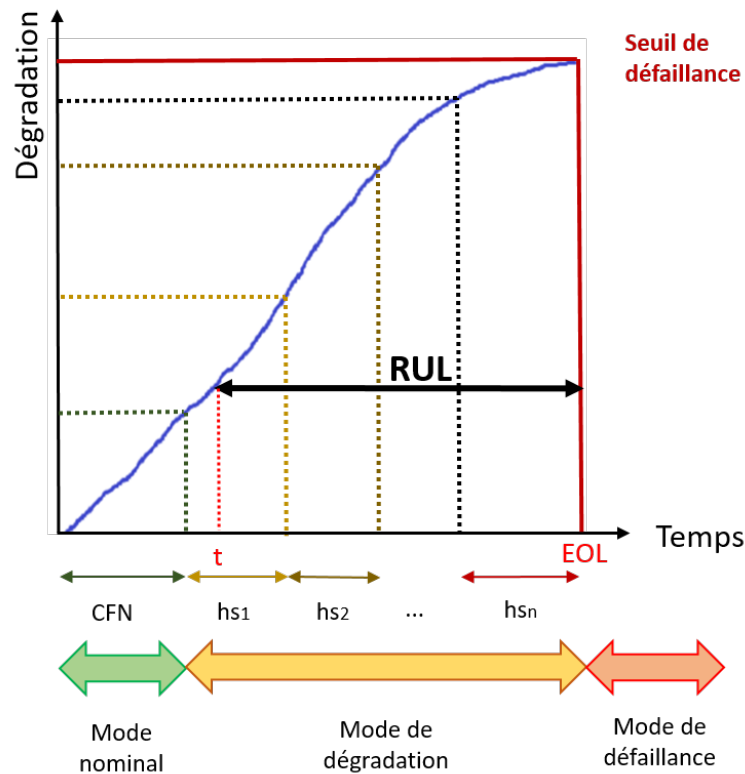


FIGURE 4.1 – Evolution de la durée de vie d'un système

**Définition 22.** Une **décision**  $d$  est définie comme l'état de santé  $hs_i \subseteq HS$  choisi par la méthode de pronostic  $m_i \subseteq M$  comme état du système.

La décision qui résulte de la méthode  $m_i$  est notée  $d_i$ . L'ensemble des décisions  $d_i$  est contenu dans  $D_i$ .

Afin d'évaluer la performance de chaque méthode individuelle en classification, la matrice de confusion ainsi que le taux de bonne classification (RR pour Recognition Rate) sont utilisés.

**Définition 23.** Le **RR** est le pourcentage d'observations correctement classées par la méthode de pronostic. Il est calculé comme suit :

$$RR = \frac{\text{Nombre d'observations correctement classées}}{\text{Nombre total d'observations}} \times 100 \quad (4.1)$$

Dans la suite, on considère que les performances de classification des états de santé peuvent être évaluées, pour chaque méthode. Ainsi, trois types d'états de santé, qui résument tous les scénarios possibles, sont définis.

**Définition 24.** Un état de santé  $hs_i \subseteq HS$  est de **type 0** si les 2 méthodes de pronostic utilisées présentent des performances équivalentes pour sa classification. L'ensemble de tous les états de santé de type 0 est noté  $F_0$ .

On constate que, contrairement au diagnostic dans le Chapitre 3, un état de santé de type 0 traduit le fait que les deux méthodes présentent des performances similaires pour

sa classification. Ceci n'implique pas forcément de "bonnes" performances pour les deux méthodes mais plutôt une même aptitude à classer cet état de santé.

En effet, chaque méthode individuelle prend en considération l'ensemble des états de santé puisqu'ils sont liés entre eux et participent tous à l'évolution de la même dégradation, contrairement aux différents défauts pris en compte dans le problème de diagnostic.

**Définition 25.** *Un état de santé  $hs_i \subseteq HS$  est de type  $k$  ( $k = \{1, 2\}$ ) si la méthode  $m_k$  présente de meilleures performances pour sa classification. L'ensemble de tous les défauts de type  $k$  est noté  $F_k$ .*

En effet, pour chaque méthode, trois possibilités s'offrent pour les états de santé : soit ils sont classés avec la même performance par les deux méthodes et dans ce cas ce sont des états de santé communs de type 0, soit ces états de santé lui sont spécifiques puisqu'elle présente de meilleures performances, soit ces états de santé ne lui sont pas spécifiques. Ainsi, trois types d'états de santé sont considérés : les états de santé de type 0, 1 et 2. Cela implique que les décisions qui peuvent être données par les méthodes  $m_1$  et  $m_2$  sont respectivement :  $D_1 = \{F_0, F_1, F_2\}$  et  $D_2 = \{F_0, F_1, F_2\}$ .

Cela nous conduit à la formulation suivante du problème de fusion de décisions.

**Problème-Fusion de décisions pour le pronostic.** *Étant donné un système surveillé avec 2 méthodes, et présentant un ensemble d'états de santé persistants  $HS$  menant à une défaillance, ainsi que deux décisions ( $d_1 \subseteq D_1$ ;  $d_2 \subseteq D_2$ ) données à l'instant  $t$  par les 2 méthodes, le problème de fusion de décisions consiste à choisir la décision globale qui améliore les performances du système en pronostic.*

L'estimation de la RUL dépend de la décision sur l'état du système. En effet, la RUL représente le temps qui sépare l'état courant de l'état de défaillance.

### 4.3 Démarche méthodologique de pronostic

La démarche méthodologique de pronostic proposée ici est une démarche basée sur les données. Ce choix se justifie par la disponibilité de quantités importantes de données de surveillance, ainsi que par le développement des capteurs et l'augmentation des performances des calculateurs, facilitant la mise en oeuvre de méthodes guidées par les données [Medjaher, 2014]. De plus, le développement d'approches basées sur des modèles physiques est difficile à mettre en oeuvre en raison de la complexité des systèmes industriels.

La méthodologie de pronostic proposée pour les méthodes individuelles est illustrée dans la Figure 4.2. Elle repose sur deux étapes, décrites plus en détails dans ce qui suit.

1. **Apprentissage (Hors-ligne) :** La première étape est conduite hors-ligne et consiste à étudier le lien entre le signal à un instant donné et l'état de santé correspondant. Cet apprentissage se fait à partir des données historiques d'apprentissage, comme les tests de vieillissement accélérés et les données de passage à la défaillance (RtF pour Run to Failure). Ces liens sont enregistrés dans la base d'apprentissage.
2. **Estimation de l'état de santé et de la RUL (En-ligne) :** La deuxième étape se déroule en ligne et consiste à prédire la RUL du composant test. La base d'apprentissage est exploitée pour estimer l'état courant du composant. A partir de cette estimation, la RUL est prédite selon une formule proposée.

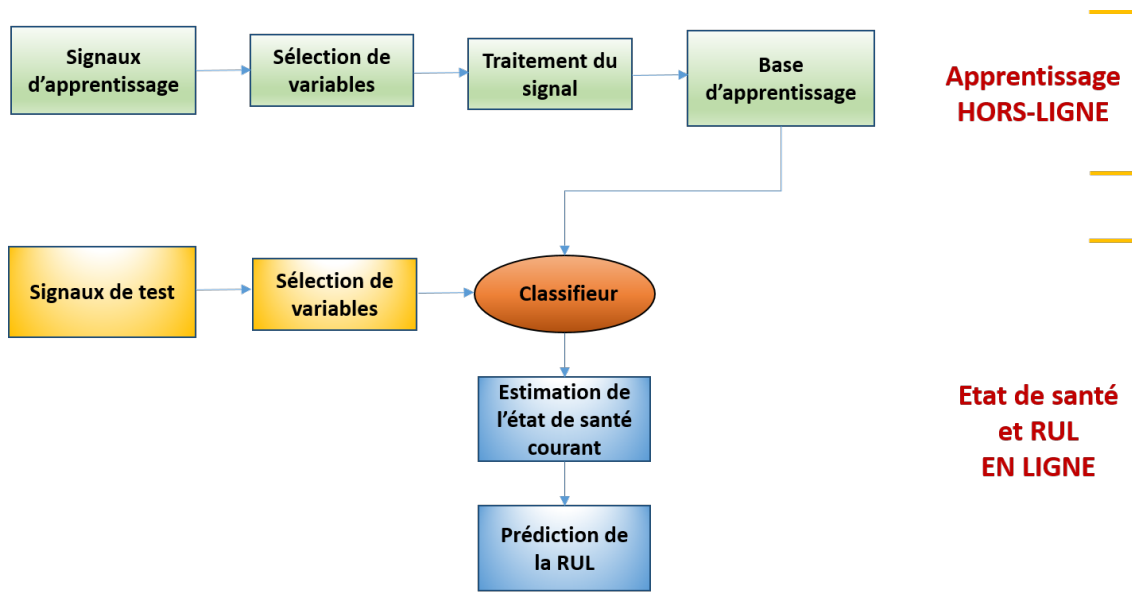


FIGURE 4.2 – La méthodologie de pronostic proposée

Par la suite, nous allons développer les différentes étapes de la méthodologie présentée ci-dessus.

### 4.3.1 Stratégie d'apprentissage

#### Signaux d'apprentissage et sélection de variables

On dispose de  $N$  systèmes semblables  $(S_1, S_2, \dots, S_N)$  dont l'évolution de la dégradation est surveillée.  $p$  capteurs  $(C_1, C_2, \dots, C_p)$  supposés fonctionner sans défaillance sont disposés sur chacun des  $N$  systèmes. Dès lors qu'un système atteint la défaillance, les signaux des  $p$  capteurs sont récupérés. Ces signaux d'apprentissage sont donc représentatifs du fonctionnement du système durant toute sa durée de vie. Ils représentent des données de passage à la défaillance, retraçant le fonctionnement du système de l'état normal jusqu'à l'état défaillant, en passant par plusieurs états de dégradation. L'instant de défaillance, correspondant à la fin de vie du système  $i$ , est noté  $EOL_i$  (EOL pour End of Life). Une sélection de variables est ensuite effectuée afin de ne garder que les signaux pertinents représentant la dégradation du composant. On note  $k$  le nombre de signaux sélectionnés parmi les  $p$ .

La Figure 4.3 illustre ces données d'apprentissage.

Afin d'étudier le lien entre le signal à un instant donné et l'état de santé correspondant, nous avons besoin d'une grande quantité de signaux de même nature, de telle sorte que les données soient représentatives. Pour cela, les signaux de chaque capteur  $j$  ( $j \in \{1, \dots, k\}$ ) sont récupérés et traités pour les  $N$  systèmes disponibles.

On constate alors que les signaux des capteurs n'évoluent pas de la même façon d'un système à l'autre et que chaque système possède sa propre EOL.

Pour exemple, la Figure 4.4 superpose l'évolution du même signal, de l'état normal à la défaillance, pour 2 systèmes de même type.

La plus petite  $EOL_1$  est égale à 179 unités de temps tandis que la plus grande  $EOL_2$  vaut 287 unités de temps.

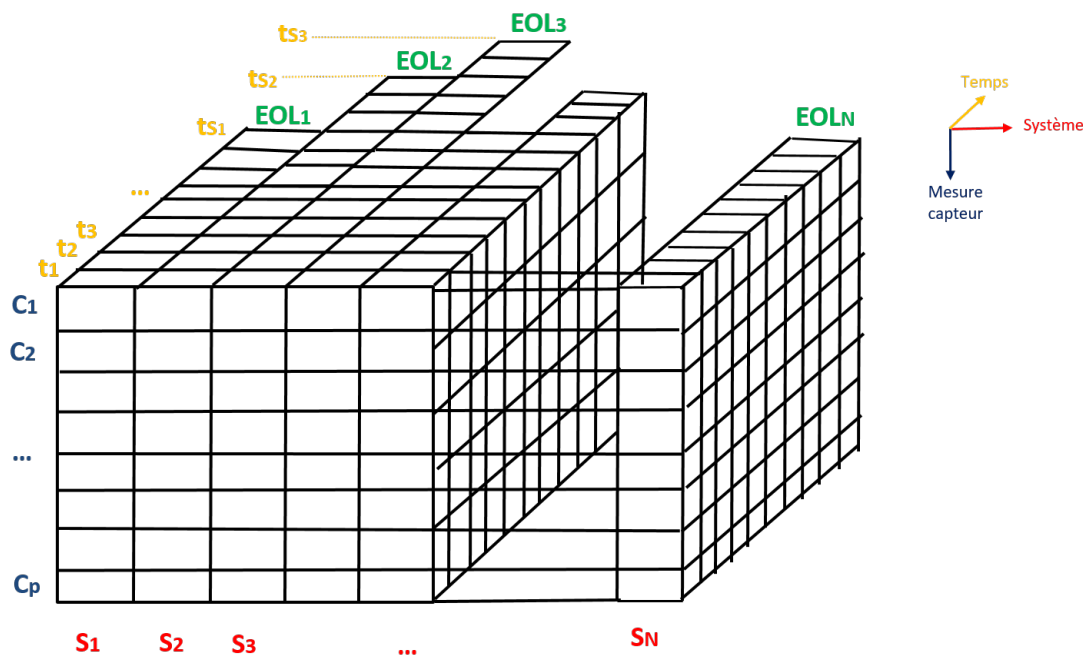


FIGURE 4.3 – Signaux d'apprentissage

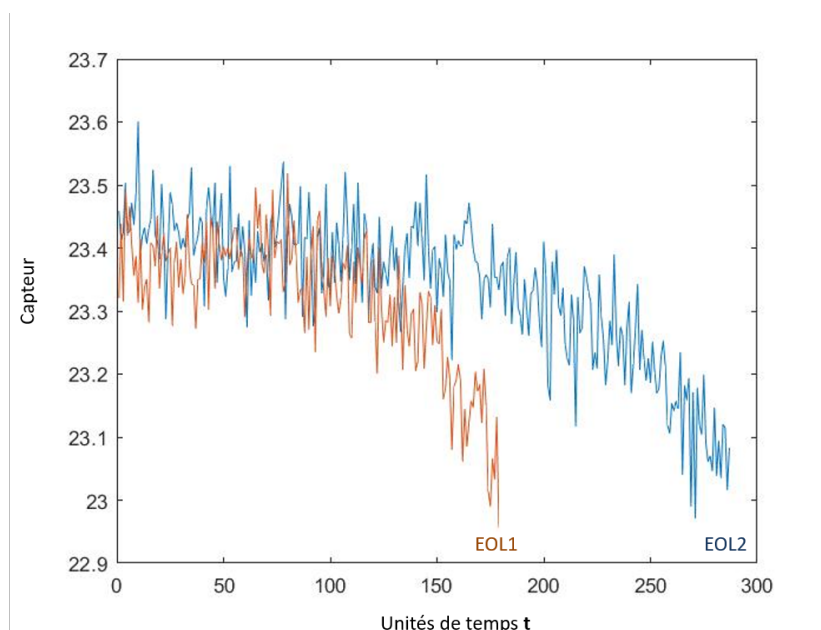


FIGURE 4.4 – Evolution de la même mesure pour les 2 systèmes considérés

Cette comparaison fait émerger une nouvelle problématique pour les méthodes de pronostic utilisant le principe de classification avec une découpe temporelle de la durée de vie du système, comme montré auparavant dans la Figure 4.1. En effet, à un instant donné, les systèmes ne sont pas dans le même état puisque leurs durées de vie sont différentes.

Prenons l'exemple donné dans la Figure 4.4, à l'instant  $t = 50$ , avec une découpe temporelle constante, les deux systèmes sont dans le même état de santé. Or, à chaque état de santé correspond une  $RUL$  prédéterminée, impliquant que les deux systèmes ont la même  $RUL$  à cet instant. Cependant, les deux vraies  $RUL$  ne sont pas égales :  $RUL_1 = 179 - 50 = 129$  contre  $RUL_5 = 287 - 50 = 237$  unités de temps. Le premier

système est proche de la défaillance tandis que le deuxième en est encore très loin, comme on le voit d'ailleurs sur la Figure 4.4. D'où la nécessité de proposer un traitement de ces signaux.

### Traitement des signaux

Le traitement proposé pour pallier ce problème consiste en un recalage des signaux afin de positionner le même état de santé à la fin, c'est-à-dire la défaillance pour tous les systèmes. En effet, le seul état commun et connu avec certitude à tous les systèmes est la fin de vie, d'où le choix de l'*EOL* pour l'ensemble des systèmes comme instant final (point d'arrivée) commun des signaux, comme l'illustre la Figure 4.5.

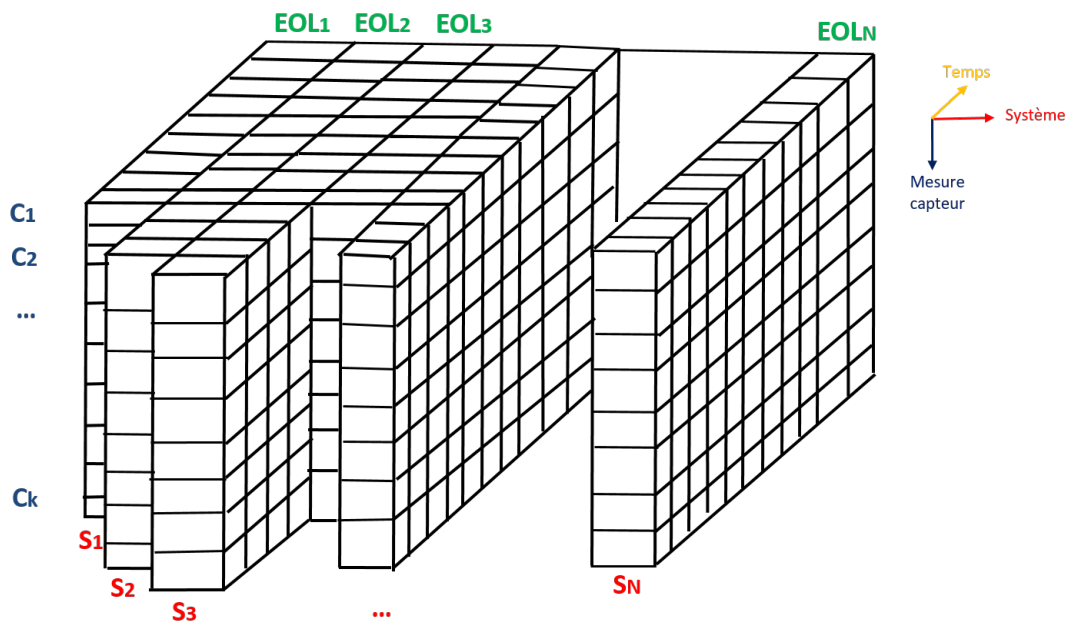


FIGURE 4.5 – Données d'apprentissage, avec l'*EOL* comme point final.

Pour exemple, la Figure 4.6 reprend les deux signaux montrés auparavant, avec l'*EOL* comme instant final.

Le découpage en classes peut donc se faire à partir de l'*EOL* pour la phase d'apprentissage.

Ainsi, il convient de découper le signal en un nombre  $n$  d'états de santé  $hs_i, i \in \{1 \dots n\}$ , de même longueur  $q$ , en partant de l'*EOL*, comme le montre la Figure 4.7, et plus en détails l'exemple de la Figure 4.8.

Dans cet exemple, on découpe le signal en 7 états de santé, de longueur 41. La classe 1, représentant l'état de santé  $hs_1$ , correspond à une dégradation sévère puisqu'elle débute avec l'*EOL*, tandis que la classe 7, représentant l'état de santé  $hs_7$ , correspond à la dégradation la plus légère ou à l'état de fonctionnement normal. A chaque état de santé est associée une RUL.

### Base d'apprentissage

Une méthode de classification (AD, kNN, SVM, arbre de décision, etc.) est ensuite utilisée afin d'apprendre le lien entre les  $k$  signaux à un instant donné pour chaque système et l'état

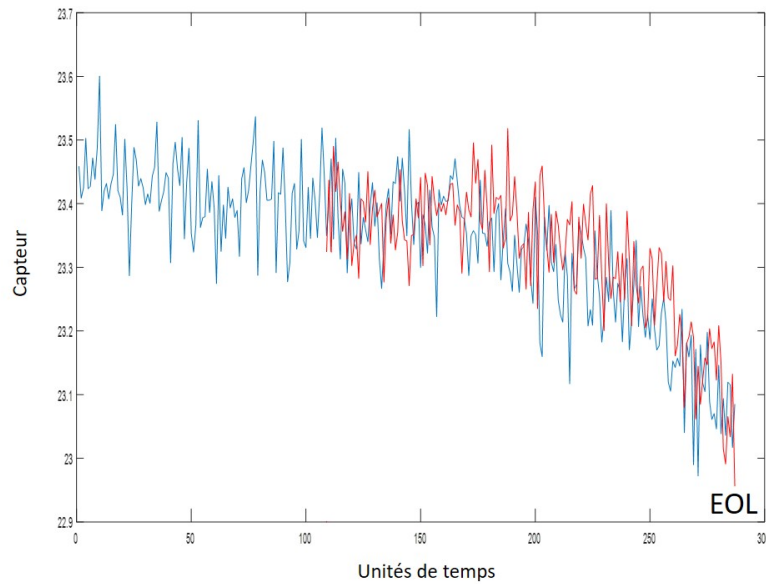


FIGURE 4.6 – Evolution de la même mesure pour les 2 systèmes considérés, avec l'*EOL* comme point de départ.

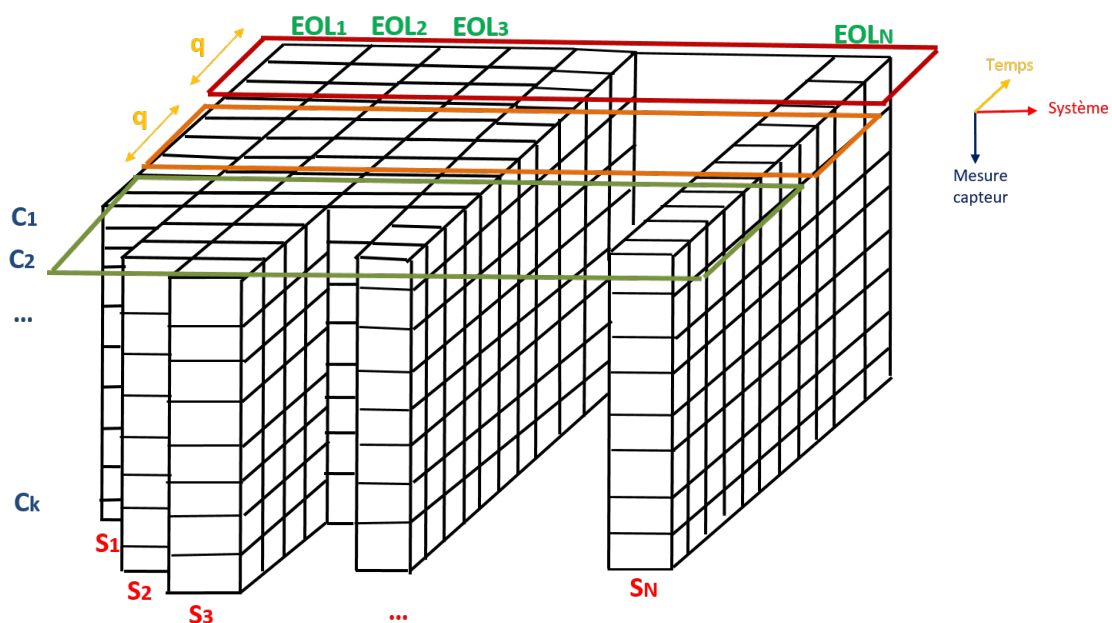


FIGURE 4.7 – Découpage en états de santé

de santé correspondant. Ces liens sont stockés dans une base d'apprentissage et vont servir pour l'estimation de l'état de santé en-ligne. Ainsi, le résultat de classification dépend de la qualité de la base d'apprentissage.

### 4.3.2 Estimation de l'état de santé et prédiction de la RUL

#### Signaux de test et sélection de variables

On dispose de  $T$  nouveaux systèmes de même nature ( $St_1, St_2, \dots, St_T$ ) n'ayant pas été traités en phase d'apprentissage et dont l'évolution de la dégradation est surveillée. Les mêmes capteurs ( $C_1, C_2, \dots, C_p$ ) supposés fonctionner sans défaillance sont disposés

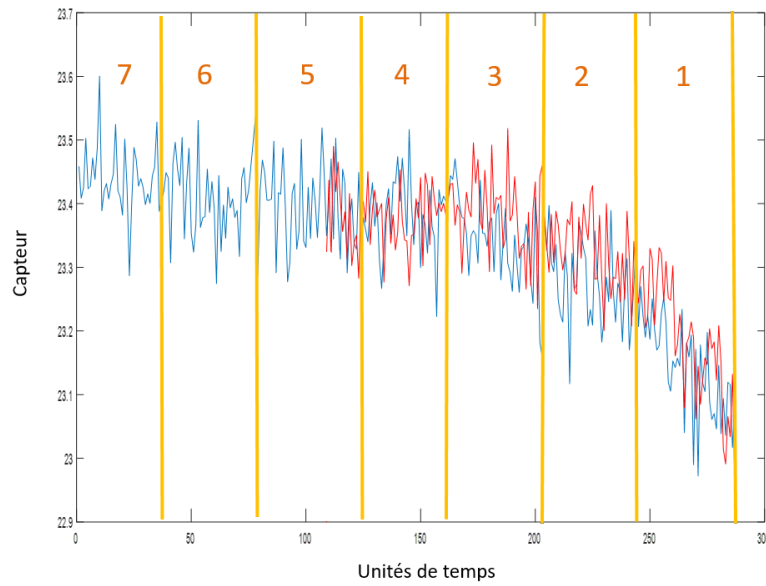


FIGURE 4.8 – Découpage en états de santé pour les deux signaux

sur les  $T$  systèmes. Les  $k$  variables sélectionnées dans la phase d'apprentissage sont considérées dans la phase en-ligne. Contrairement à la phase d'apprentissage, les instants de défaillances de chaque système sont inconnus. L'objectif est de déterminer, pour chaque système, la RUL à un instant donné.

La Figure 4.9 illustre ces données de test.

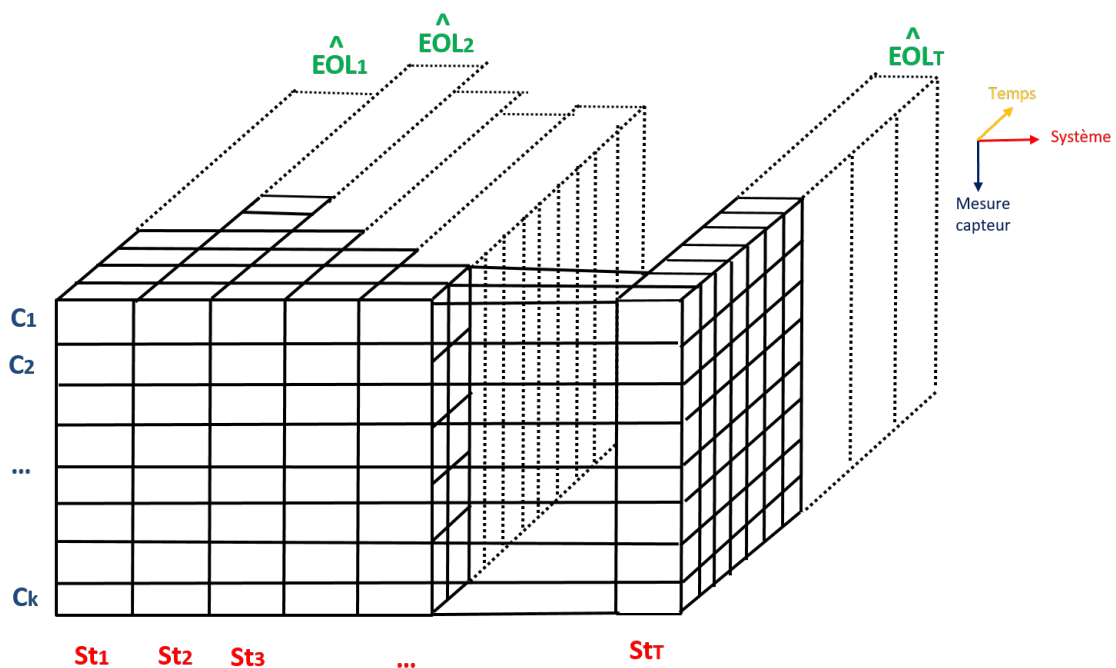


FIGURE 4.9 – Données de test



### Classifieur

Le classifieur utilisé dans la phase d'apprentissage est exploité en-ligne afin d'estimer l'état de santé du système à l'instant  $t$ . Pour cela, le classifieur s'appuie sur la base d'apprentissage comprenant les liens entre les signaux et les états de santé correspondants.

### Estimation de l'état de santé et prédiction de la RUL

A chaque instant, le classifieur choisit un état de santé correspondant aux signaux à l'instant  $t$ . Cet état représente la décision du classifieur.

A chaque état de santé est associée une RUL, calculée de la façon suivante :

$$\widehat{RUL}_i = i \times q - \frac{q}{2} \quad (4.2)$$

avec  $i$  ( $i \in \{1 \dots n\}$ ) le numéro de la classe choisie comme décision par le classifieur. La partie  $i \times q$  permet de situer la tranche de longueur  $q$  étant donné la classe  $i$ . Par exemple, si la décision correspond à la classe 1 (qui comprend le EOL), alors il reste  $q$  unités de temps au système avant son EOL, soit 41 unités de temps pour l'exemple de la Figure 4.8. Par contre, si la décision correspond à la classe 4 (plus éloignée de l'EOL), il lui reste alors  $4q$  unités de temps, soit 164 pour l'exemple présenté. Or, comme on ne sait pas où on se situe exactement dans la classe, on choisit le centre de la classe comme référence, en soustrayant  $\frac{q}{2}$ . Cette dernière représente la valeur de confiance associée à la RUL.

Une fois la RUL estimée, il convient d'évaluer la performance de cette estimation par rapport à la vraie RUL.

### Métriques pour le pronostic

La performance de la prédiction de la RUL est évaluée à travers plusieurs métriques utilisées dans la littérature, comme l'erreur quadratique moyenne (MSE pour Mean Square Error), l'erreur absolue moyenne en pourcentage (MAPE pour Mean absolute percentage error), la précision, le score, etc. Dans la suite, nous définissons les métriques les plus populaires. De plus amples détails peuvent être trouvés dans [Saxena et al., 2008a].

**Définition 26.** *L'erreur quadratique moyenne (MSE pour Mean Square Error) est définie comme le carré moyen des erreurs de prédiction de la RUL. Les erreurs individuelles sont tout d'abord élevées au carré, puis additionnées les unes aux autres. Le résultat obtenu est ensuite divisé par le nombre total d'erreurs individuelles. Sa formule est la suivante :*

$$MSE = \frac{1}{T} \sum_{i=1}^T (RUL_i - \widehat{RUL}_i)^2 \quad (4.3)$$

avec  $T$  le nombre de RULs à estimer,  $RUL_i$  la vraie RUL du système  $i$  et  $\widehat{RUL}_i$  l'estimation de la RUL du système  $i$ .

Lorsqu'on compare deux estimateurs de RUL, on considère que le meilleur est celui qui présente la MSE la plus faible. Par conséquent, la MSE doit être minimisée.

**Définition 27.** *L'erreur absolue moyenne (MAE pour Mean Absolute Error) représente la moyenne arithmétique des valeurs absolues des écarts. Sa formule est la suivante :*



$$MAE = \frac{1}{T} \sum_{i=1}^n |RUL_i - R\hat{U}L_i| \quad (4.4)$$

avec  $T$  le nombre de RULs à estimer,  $RUL_i$  la vraie RUL du système  $i$  et  $R\hat{U}L_i$  l'estimation de la RUL du système  $i$ .  
Tout comme la MSE, la MAE doit être minimisée.

Les erreurs moyennes MSE et MAE résument les performances en prédiction mais sans information évidente à propos de la taille relative de l'erreur, contrairement à l'erreur absolue moyenne en pourcentage.

**Définition 28.** *L'erreur absolue moyenne en pourcentage (MAPE pour Mean absolute percentage error) représente la moyenne des écarts en valeur absolue par rapport aux vraies valeurs de la RUL et s'exprime comme il suit :*

$$MAPE = \frac{100}{T} \sum_{i=1}^T \left| \frac{RUL_i - R\hat{U}L_i}{RUL_i} \right| \quad (4.5)$$

avec  $T$  le nombre de RULs à estimer,  $RUL_i$  la vraie RUL du système  $i$  et  $R\hat{U}L_i$  l'estimation de la RUL du système  $i$ .  
La MAPE doit être minimisée.

Un autre critère est la précision du pronostic, qui doit tendre vers 1 et qui s'écrit comme suit :

**Définition 29.** *Précision  $R$ .*

$$R = 1 - \frac{\sum_{i=1}^T (RUL_i - R\hat{U}L_i)^2}{\sum_{i=1}^T (RUL_i - \overline{RUL})^2} \quad (4.6)$$

avec  $T$  le nombre de RULs à estimer et  $\overline{RUL}$  la moyenne des RULs estimées.

Enfin, la fonction Score a été introduite pour distinguer les différentes pénalités causées par la sous-estimation et la surestimation de la RUL.

**Définition 30.** *La fonction Score, qui doit être minimisée, se calcule de la façon suivante :*

$$Score = \sum_{i=1}^T Sc_i \quad (4.7)$$

$$Sc_i = \begin{cases} e^{-\frac{r_i}{\beta_1}} - 1, & r_i < 0 \\ e^{\frac{r_i}{\beta_2}} - 1, & r_i \geq 0 \end{cases} \quad (4.8)$$

Avec  $T$  le nombre de RULs à estimer et  $r_i = R\hat{U}L_i - RUL_i$ .

Ainsi, elle assigne une pénalité plus élevée aux prédictions de RUL tardives. En effet, une surestimation de la RUL peut entraîner des dommages plus graves qu'une sous-estimation. Un intervalle  $I = [\beta_1, \beta_2]$  est considéré pour évaluer les estimations de RUL comme étant à temps, précoces ou tardives (voir Figure 4.10).

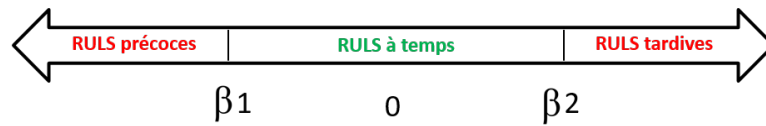


FIGURE 4.10 – Intervalle de prédiction

Il convient de noter que toutes ces métriques sont utilisées uniquement lorsque les vraies RULs sont connues.

Maintenant que l'approche de pronostic individuelle qui va générer une décision a été décrite, nous allons, dans ce qui suit, expliquer la démarche méthodologique de fusion des décisions.

## 4.4 Démarche méthodologique de fusion

La démarche méthodologique proposée peut être découpée en trois étapes principales, illustrées par la Figure 4.11 et décrites plus en détails ci-dessous :

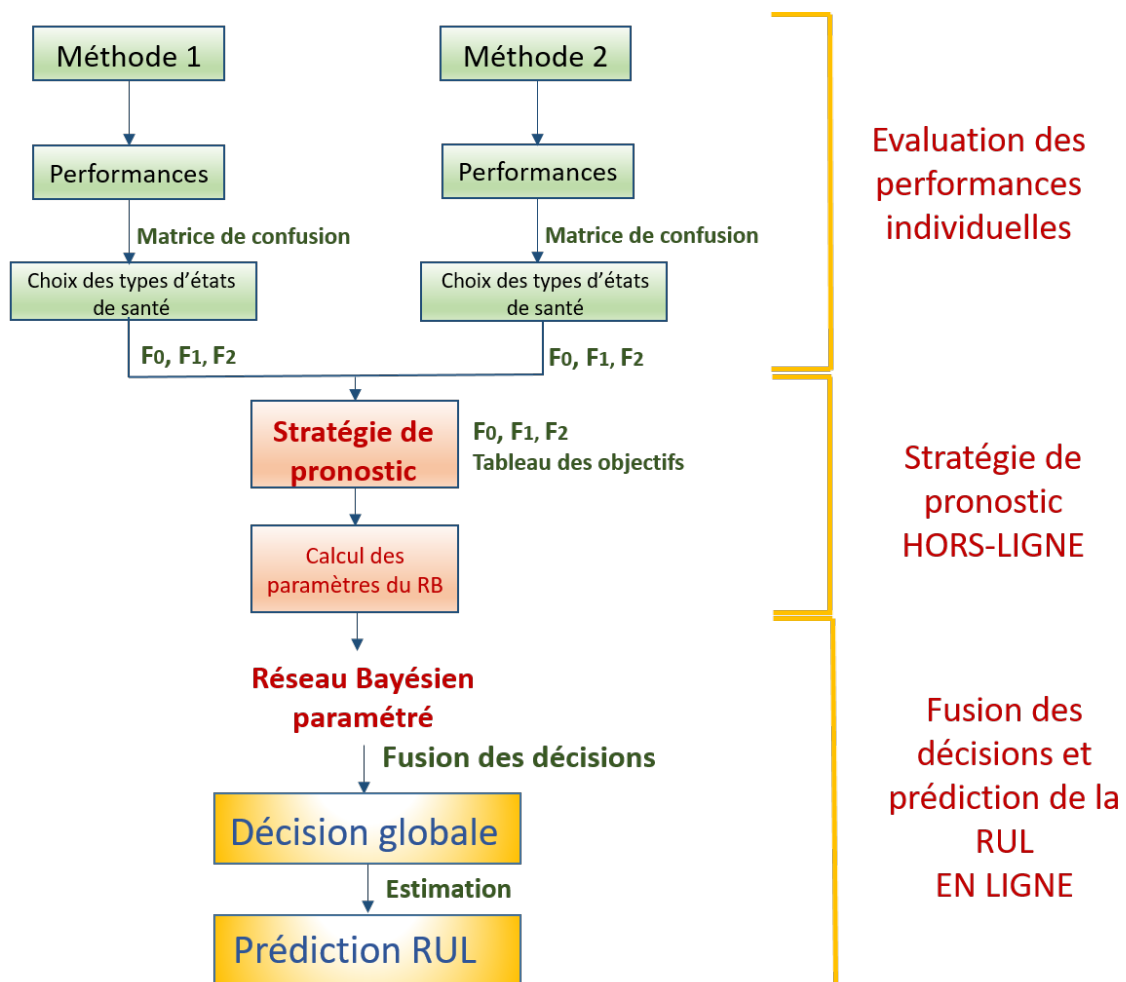


FIGURE 4.11 – Méthodologie proposée de fusion pour le pronostic

- 1. Evaluation des performances individuelles** : Chaque méthode évalue d'abord ses performances individuelles en pronostic du système considéré. Une matrice de

confusion résumant la classification dans les différents états de santé est construite pour chaque méthode. A partir des matrices de confusion, les différents types d'états de santé ( $F_0, F_1, F_2$ ) sont définis et les RULs sont prédites puis évaluées selon les différentes métriques.

2. **Stratégie de pronostic (Hors-ligne)** : Par la suite, la stratégie de pronostic est déployée et le tableau des objectifs est défini. Ce dernier va guider la déduction théorique des paramètres du RB. Une fois les paramètres choisis, le RB est prêt à être utilisé en ligne pour l'estimation de l'état du système et la prédiction de la RUL.
3. **Fusion des décisions et Prédiction de la RUL (En ligne)** : Les décisions des méthodes individuelles (c'est-à-dire les états de santé) sont introduites dans le RB et fusionnées grâce à la stratégie adoptée. La décision globale concernant l'état de santé du système est renvoyée par le RB. La RUL est ensuite prédite.

Par la suite, nous allons développer les différentes briques de la méthodologie présentée ci-dessus.

#### 4.4.1 Evaluation des performances individuelles

Cette première étape permet d'identifier les types d'états de santé qui sont mieux classés par chacune des méthodes, et ceux qui ne le sont pas, à travers les matrices de confusion. A l'égard du diagnostic, ces matrices sont comparées selon le même raisonnement (connaissances de l'expert ou tests statistiques), comme montré dans la partie 3.3.1. Enfin, les différents types d'états de santé ( $F_0, F_1, F_2$ ) sont définis.

**Exemple** Afin d'illustrer la démarche méthodologique proposée, un exemple numérique est présenté tout au long de ce chapitre.

L'évolution de la dégradation enregistrée par un capteur du système considéré est illustrée dans la Figure 4.12.

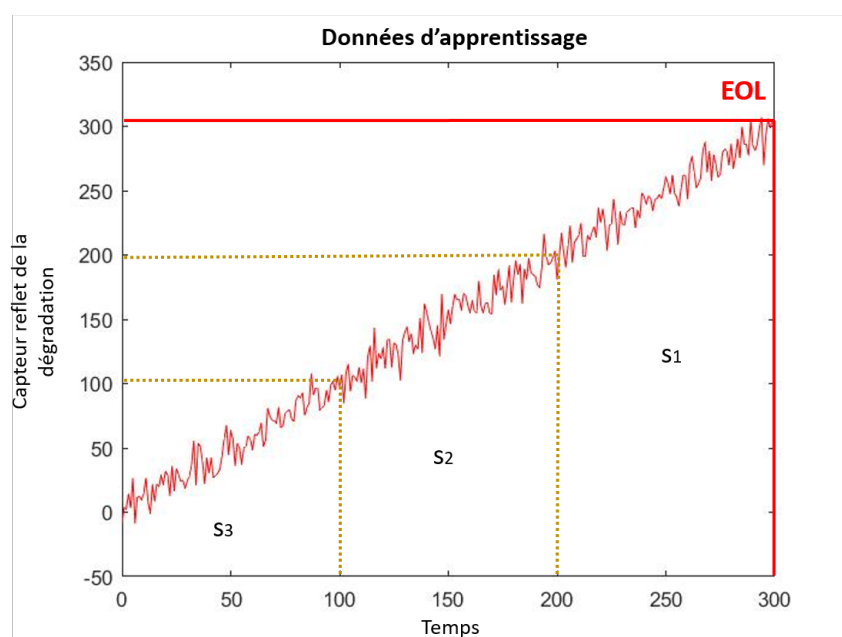


FIGURE 4.12 – Exemple numérique : Evolution de la dégradation

La dégradation débute à l'instant  $t = 0$  et évolue de façon linéaire jusqu'à la défaillance à l'instant  $t = 300$ , correspondant à l'*EOL*. Pour des fins de pronostic, deux méthodes sont développées (AD quadratique et kNN ( $k = 4$ )) afin de déterminer l'état courant du système et prédire la RUL.

Pour cela, la dégradation est découpée en 3 états de santé, selon la méthodologie proposée pour le pronostic (Partie 4.3) :

- Echantillon 1 à 100 : Etat de santé  $hs_3$  correspondant à une dégradation légère.
- Echantillon 101 à 200 : Etat de santé  $hs_2$  correspondant à une dégradation modérée.
- Echantillon 201 à 300 : Etat de santé  $hs_1$  correspondant à une dégradation sévère.

Le but de chaque méthode individuelle est de classer le système dans un de ces 3 états de santé.

Les résultats de la simulation sont donnés sous la forme de matrices de confusion (voir les tables 4.1a et 4.1b), permettant de visualiser la performance en classification de chaque méthode individuelle.

	$hs_1$	$hs_2$	$hs_3$
$hs_1$	98	2	0
$hs_2$	22	78	0
$hs_3$	55	1	44

(a) Matrice de confusion de la méthode  $m_1$ 

	$hs_1$	$hs_2$	$hs_3$
$hs_1$	100	0	0
$hs_2$	52	41	7
$hs_3$	0	5	95

(b) Matrice de confusion de la méthode  $m_2$ 

TABLE 4.1 – Evaluation des performances individuelles

- Méthode  $m_1$  : AD quadratique
  - Sur les 100 échantillons en classe  $hs_1$ , la méthode  $m_1$  en classe correctement 98. Le reste est considéré comme appartenant à la classe  $hs_2$ . Le RR relatif à cette classe est de 98%.
  - Sur les 100 échantillons en classe  $hs_2$ , la méthode  $m_1$  en classe correctement 78. Le reste est considéré comme appartenant à la classe  $hs_1$ . Le RR relatif à cette classe est de 78%.
  - Sur les 100 échantillons en classe  $hs_3$ , la méthode  $m_1$  en classe correctement 44. Le reste est considéré comme appartenant à  $hs_1$  ou  $hs_2$ . Le RR relatif à cette classe est de 44%.
- Méthode  $m_2$  : kNN
  - La méthode  $m_2$  classe correctement tous les échantillons appartenant à la classe  $hs_1$ . Le RR relatif à cette classe est de 100%.
  - Sur les 100 échantillons en classe  $hs_2$ , la méthode  $m_2$  en classe correctement 41. Le reste est considéré comme appartenant à  $hs_1$  ou  $hs_3$ . Le RR relatif à cette classe est de 41%.
  - Sur les 100 échantillons en classe  $hs_3$ , la méthode  $m_2$  en classe correctement 95. Le reste est considéré comme appartenant à la classe  $hs_2$ . Le RR relatif à cette classe est de 95%.

En comparant les éléments sur la diagonale, il apparaît que les méthodes  $m_1$  et  $m_2$  sont expertes respectivement dans la classification des états de santé  $hs_2$  et  $hs_3$ . Par conséquent :  $F_1 = \{hs_2\}$  et  $F_2 = \{hs_3\}$ . Les deux méthodes sont expertes pour la classe  $hs_1$  donc  $F_0 = \{hs_1\}$ .

Par conséquent, les RRs sont respectivement de :

$$\begin{cases} RR_{m_1} = \frac{98 + 78 + 44}{300} = 73.33\% \\ RR_{m_2} = \frac{100 + 41 + 95}{300} = 78.67\% \end{cases} \quad (4.9)$$

Ainsi, les performances en classification des méthodes individuelles sont comparables. Nous allons voir dans la suite comment la fusion va améliorer ces performances.

#### 4.4.2 Stratégie de pronostic : Définition des objectifs

Comme pour le diagnostic (voir partie 3.3.2) la décision globale, notée  $d$  est construite en fonction des décisions individuelles  $d_1$  et  $d_2$  données respectivement par les méthodes  $m_1$  et  $m_2$ . L'ensemble des objectifs possibles est résumé dans la Table 4.2.

$d_2 \backslash d_1$	$F_0$	$F_1$	$F_2$
$F_0$	1. $F_0$	4. $F_1$	7. $F_0$
$F_1$	2. $F_0$	5. $F_1$	8. $F_1; F_2$
$F_2$	3. $F_2$	6. $F_1; F_2$	9. $F_2$

TABLE 4.2 – Table des objectifs

Les objectifs **1**, **5** et **9** (en rouge dans la Table 4.2) sont évidents puisque les deux méthodes choisissent séparément la même décision.

**Objectif 2** : Si  $d_1 = \{F_0\}$  et  $d_2 = \{F_1\}$ , alors la décision globale est :  $d = \{F_0\}$ .

Ici, seule la méthode  $m_2$  a été capable de détecter un état de santé de type 1. Or, la méthode  $m_1$  qui présente de meilleures performances pour sa classification ne le détecte pas, ce qui nous conduit à ignorer l'occurrence de cet état de santé. Par conséquent, la classe de dégradation  $F_0$  est choisie comme décision globale.

Le même raisonnement est appliqué pour le choix de l'Objectif **7** :

Si  $d_1 = \{F_2\}$  et  $d_2 = \{F_0\}$ , alors la décision globale est :  $d = \{F_0\}$ .

**Objectif 3** : Si  $d_1 = \{F_0\}$  et  $d_2 = \{F_2\}$ , alors la décision globale est :  $d = \{F_2\}$ .

La méthode  $m_2$  est experte dans la classification des états de santé de type 2 contrairement à la méthode  $m_1$ . La décision globale suit donc la méthode  $m_2$ .

Le même raisonnement est appliqué pour le choix de l'Objectif **4** :

Si  $d_1 = \{F_1\}$  et  $d_2 = \{F_0\}$ , alors la décision globale est :  $d = \{F_1\}$ .

**Objectif 6** : Si  $d_1 = \{F_1\}$  et  $d_2 = \{F_2\}$ , alors la décision est :  $d = \{F_1\}$  ou  $d = \{F_2\}$ .

Les deux méthodes ont détecté des états de santé pour lesquels elles sont expertes. Contrairement à la stratégie de diagnostic où les deux défauts ont été choisis comme décision globale afin de considérer le cas de défauts multiples, ici les classes représentent des états de dégradation qui mènent au même défaut. Par conséquent, la décision globale ne peut pas être composée de deux états de santé, et un choix parmi ces deux types d'états de santé doit être fait.

Le même raisonnement est appliqué pour le choix de l'Objectif 8 :

Si  $d_1 = \{F_2\}$  et  $d_2 = \{F_1\}$ , alors la décision globale est :  $d = \{F_1\}$  ou  $d = \{F_2\}$ .

On verra par la suite que, quelque soit la décision choisie pour les Objectifs 6 et 8 ( $d = \{F_1\}$  ou  $d = \{F_2\}$ ), les contraintes sont les mêmes.

On en déduit que le nombre d'objectifs différents à développer se résume à 6. Si ces objectifs sont respectés, alors la performance globale en classification (en termes de RR) sera améliorée, et par conséquent, les prédictions de RULs seront plus précises.

#### 4.4.3 Stratégie de pronostic : Cadre générique de fusion

Le cadre générique de fusion utilisé pour le pronostic est le même que celui utilisé pour le diagnostic. Sa structure, qui comprend des modalités différentes, est montrée dans la Figure 4.13.

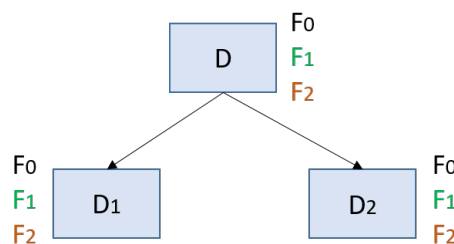


FIGURE 4.13 – Structure du RB pour le pronostic

Par la suite, nous allons déterminer les valeurs des paramètres du RB (les probabilités a priori du nœud parent  $D$  (voir la Table 4.3) et les probabilités conditionnelles des nœuds enfants (voir la Table 4.4)) afin d'améliorer les performances de pronostic du système considéré.

$D$		
$F_0$	$F_1$	$F_2$
$\alpha_0$	$\alpha_1$	$\alpha_2$

TABLE 4.3 – Table de probabilités a priori du nœud  $D$

#### 4.4.4 Stratégie de pronostic : Développement des objectifs

La méthodologie permettant d'atteindre les objectifs fixés est détaillée dans ce qui suit.

$D \backslash D_1$	$F_0$	$F_1$	$F_2$
$F_0$	$a_0$	$a_1$	$a_2$
$F_1$	$b_0$	$b_1$	$b_2$
$F_2$	$c_0$	$c_1$	$c_2$

(a) Table du noeud  $D_1$ 

$D \backslash D_2$	$F_0$	$F_1$	$F_2$
$F_0$	$d_0$	$d_1$	$d_2$
$F_1$	$e_0$	$e_1$	$e_2$
$F_2$	$f_0$	$f_1$	$f_2$

(b) Table du noeud  $D_2$ 

TABLE 4.4 – Tables de probabilités conditionnelles

### Hypothèses et propriétés

Comme pour le diagnostic, toutes les probabilités a priori sont supposées égales :  $\alpha_0 = \alpha_1 = \alpha_2$ .

Ensuite, étant donné une décision globale sur le fait que le système se trouve dans la modalité commune aux deux méthodes ( $F_0$ ), les probabilités conditionnelles des méthodes individuelles d'être dans cette modalité commune sont les mêmes, comme explicité dans l'équation 4.10.

$$P(D_1 = F_0/D = F_0) = P(D_2 = F_0/D = F_0) \Rightarrow a_0 = d_0 \quad (4.10)$$

Enfin, puisque les méthodes  $m_1$  et  $m_2$  sont respectivement expertes dans la classification des états de santé de type  $F_1$  et  $F_2$ , les équations suivantes 4.11, 4.12, 4.13 sont supposées :

$$\begin{cases} P(D_1 = F_1/D = F_0) = P(D_2 = F_2/D = F_0) \Rightarrow a_1 = d_2 \\ P(D_1 = F_2/D = F_0) = P(D_2 = F_1/D = F_0) \Rightarrow a_2 = d_1 \end{cases} \quad (4.11)$$

$$\begin{cases} P(D_1 = F_0/D = F_1) = P(D_2 = F_0/D = F_2) \Rightarrow b_0 = f_0 \\ P(D_1 = F_1/D = F_1) = P(D_2 = F_1/D = F_2) \Rightarrow b_1 = f_2 \\ P(D_1 = F_2/D = F_1) = P(D_2 = F_2/D = F_2) \Rightarrow b_2 = f_1 \end{cases} \quad (4.12)$$

$$\begin{cases} P(D_1 = F_0/D = F_2) = P(D_2 = F_0/D = F_1) \Rightarrow c_0 = e_0 \\ P(D_1 = F_1/D = F_2) = P(D_2 = F_1/D = F_1) \Rightarrow c_1 = e_2 \\ P(D_1 = F_2/D = F_2) = P(D_2 = F_2/D = F_1) \Rightarrow c_2 = e_1 \end{cases} \quad (4.13)$$

Ainsi, les probabilités conditionnelles respectent les égalités suivantes :

$$\begin{cases} a_0 = d_0; & a_1 = d_2; & a_2 = d_1 \\ b_0 = f_0; & b_1 = f_2; & b_2 = f_1 \\ c_0 = e_0; & c_1 = e_2; & c_2 = e_1 \end{cases} \quad (4.14)$$

Par conséquent, seulement 9 probabilités conditionnelles ( $\{a_0, a_1, a_2, \dots, c_0, c_1, c_2\}$ ) doivent être déterminées.

En outre, la propriété sur la somme des probabilités d'une partition qui vaut 1 est appliquée pour obtenir les contraintes supplémentaires suivantes :

$$\begin{cases} a_0 + a_1 + a_2 = 1 \\ b_0 + b_1 + b_2 = 1 \\ c_0 + c_1 + c_2 = 1 \end{cases} \quad (4.15)$$

### Objectifs

Dans cette partie, seul l'Objectif 1 est développé en détail. Les développements concernant les autres objectifs peuvent être trouvés dans l'Annexe 5.2.

La règle de décision choisie est la même que pour le diagnostic, à savoir la règle du maximum a posteriori.

**Objectif 1 :** L'objectif 1 indique que si  $d_1 = \{F_0\}$  et  $d_2 = \{F_0\}$ , alors la décision globale est :  $d = \{F_0\}$ .

Par conséquent, les critères suivants (4.16) doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_0/D_1 = F_0 \cap D_2 = F_0) > P(D = F_0) \\ P(D = F_0/D_1 = F_0 \cap D_2 = F_0) > P(D = F_i/D_1 = F_0 \cap D_2 = F_0), \\ i \in \{1, 2\} \end{array} \right. \quad (4.16)$$

- Critère 1 : La première inégalité est développée comme il suit :

$$\begin{aligned} & P(D = F_0/D_1 = F_0 \cap D_2 = F_0) > P(D = F_0) \\ \Leftrightarrow & \frac{P(D_1=F_0/D_2=F_0 \cap D=F_0)P(D_2=F_0/D=F_0)P(D=F_0)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} > P(D = F_0) \\ \Leftrightarrow & \frac{P(D_1=F_0/D=F_0)P(D_2=F_0/D=F_0)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} > 1 \\ \Leftrightarrow & \frac{a_0^2}{[\sum P(D_1=F_0/D_2=F_0 \cap D=d_i)P(D=d_i)][\sum P(D_2=F_0/D=d_i)P(D=d_i)]} > 1 \\ \Leftrightarrow & \frac{a_0^2}{\alpha_0^2(a_0+b_0+c_0)^2} > 1 \\ \Leftrightarrow & \frac{a_0}{\alpha_0(a_0+b_0+c_0)} > 1 \end{aligned} \quad (4.17)$$

La loi de Bayes, la propriété d'indépendance conditionnelle des noeuds  $D_1$  et  $D_2$  étant donné le noeud racine  $D$  ainsi que la marginalisation sur les états du noeud  $D$  sont utilisés pour développer ce critère.

- Critère 2 : La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_1$ , on obtient :

$$\begin{aligned} & P(D = F_0/D_1 = F_0 \cap D_2 = F_0) > P(D = F_1/D_1 = F_0 \cap D_2 = F_0) \\ \Leftrightarrow & \frac{P(D_1=F_0/D_2=F_0 \cap D=F_0)P(D_2=F_0/D=F_0)P(D=F_0)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} > \\ & \frac{P(D_1=F_0/D_2=F_0 \cap D=F_1)P(D_2=F_0/D=F_1)P(D=F_1)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} \\ \Leftrightarrow & P(D_1 = F_0/D = F_0)P(D_2 = F_0/D = F_0)P(D = F_0) > \\ & P(D_1 = F_0/D = F_1)P(D_2 = F_0/D = F_1)P(D = F_1) \\ \Leftrightarrow & a_0^2 > b_0c_0 \end{aligned} \quad (4.18)$$

Enfin pour  $D = F_2$ , on obtient la même contrainte :  $a_0^2 > b_0c_0$ .

Par conséquent, l'objectif 1 peut être atteint si 2 contraintes (équations 4.19 et 4.20) impliquant 3 probabilités conditionnelles sont respectées :

$$\frac{a_0}{a_0 + b_0 + c_0} > \alpha_0 \quad (4.19)$$

$$a_0^2 > b_0c_0 \quad (4.20)$$



**Objectif 2 :** L'objectif 2 indique que si  $d_1 = \{F_0\}$  et  $d_2 = \{F_1\}$ , alors la décision globale est :  $d = \{F_0\}$  et 3 contraintes (équations 4.21, 4.22) sont obtenues.

$$\frac{a_0 a_2}{(a_0 + b_0 + c_0)(a_2 + b_2 + c_2)} > \alpha_0^2 \quad (4.21)$$

$$\begin{cases} a_0 a_2 > b_0 c_2 \\ a_0 a_2 > b_2 c_0 \end{cases} \quad (4.22)$$

**Objectif 3 :** L'objectif 3 indique que si  $d_1 = \{F_0\}$  et  $d_2 = \{F_2\}$ , alors la décision globale doit toujours être :  $d = \{F_2\}$ .

Par conséquent, les contraintes suivantes (équations 4.23, 4.24) doivent être respectées :

$$\frac{c_0 b_1}{(a_0 + b_0 + c_0)(a_1 + b_1 + c_1)} > \alpha_0^2 \quad (4.23)$$

$$\begin{cases} c_0 b_1 > a_0 a_1 \\ c_0 b_1 > c_1 b_0 \end{cases} \quad (4.24)$$

**Objectif 4 :** L'objectif 4 indique que si  $d_1 = \{F_1\}$  et  $d_2 = \{F_0\}$ , alors la décision globale est :  $d = \{F_1\}$ .

Le raisonnement est identique à celui de l'Objectif 3 et les mêmes contraintes sont déduites (équations 4.23, 4.24).

**Objectif 5 :** L'objectif 5 indique que si  $d_1 = \{F_1\}$  et  $d_2 = \{F_1\}$ , alors la décision globale est :  $d = \{F_1\}$ .

Par conséquent, les contraintes suivantes (équations 4.25, 4.26) doivent être respectées :

$$\frac{b_1 c_2}{(a_1 + b_1 + c_1)(a_2 + b_2 + c_2)} > \alpha_0^2 \quad (4.25)$$

$$\begin{cases} b_1 c_2 > a_1 a_2 \\ b_1 c_2 > b_2 c_1 \end{cases} \quad (4.26)$$

**Objectif 6 :** L'objectif 6 indique que si  $d_1 = \{F_1\}$  et  $d_2 = \{F_2\}$ , alors la décision globale est :  $d = \{F_1\}$  ou  $d = \{F_2\}$ .

Par conséquent, les contraintes suivantes (équations 4.27, 4.28) doivent être respectées :

$$\frac{b_1 c_1}{(a_1 + b_1 + c_1)^2} > \alpha_0^2 \quad (4.27)$$

$$b_1 c_1 > a_1^2 \quad (4.28)$$

**Objectif 7 :** L'objectif 7 indique que si  $d_1 = \{F_2\}$  et  $d_2 = \{F_0\}$ , alors la décision globale est :  $d = \{F_0\}$ .

Le raisonnement est identique à celui de l'Objectif 2 et les mêmes contraintes sont déduites (équations 4.21, 4.22).

**Objectif 8 :** L'objectif 8 indique que si  $d_1 = \{F_2\}$  et  $d_2 = \{F_1\}$ , alors la décision globale doit toujours être :  $d = \{F_1\}$  ou  $d = \{F_2\}$ .

Le raisonnement est identique à celui de l'Objectif 6 mais des contraintes supplémentaires sont déduites (équations 4.29, 4.30).

$$\frac{b_2 c_2}{(a_2 + b_2 + c_2)^2} > \alpha_0^2 \quad (4.29)$$

$$b_2 c_2 > a_2^2 \quad (4.30)$$

**Objectif 9 :** L'objectif 9 indique que si  $d_1 = \{F_2\}$  et  $d_2 = \{F_2\}$ , alors la décision globale est :  $d = \{F_2\}$ .

Le raisonnement est identique à celui de l'Objectif 5 et les mêmes contraintes sont déduites (équations 4.25, 4.26).

Le but maintenant est de résoudre le problème multi-objectif en trouvant toutes les valeurs possibles des probabilités conditionnelles (c'est-à-dire toutes les solutions du problème).

#### 4.4.5 Stratégie de pronostic : Prise de décision multi-objectifs sous contraintes

De manière formelle, le problème multi-objectif rencontré ici peut s'écrire comme suit :

**Problème multi-objectif.** Trouver  $x \in \Delta$  tel que  $R(x) = [r_1(x); r_2(x); r_3(x); r_4(x); r_5(x); r_6(x)] > 0$  sous les contraintes  $P(x) = [p_1(x); \dots; p_9(x)] > 0$ .

avec  $x = [a_0, b_0, c_0, a_1, b_1, c_1, a_2, b_2, c_2]^T$  le vecteur des solutions réalisables (probabilités conditionnelles des noeuds enfants),  $\Delta$  le domaine de définition des probabilités,  $R(x)$  les fonctions objectifs déduites précédemment (correspondant aux équations 4.19, 4.27, 4.29, 4.21, 4.23, 4.25, du premier critère à respecter, transformées en équations aux différences).

$$R(x) \begin{cases} r_1(x) = \frac{a_0}{a_0+b_0+c_0} - \alpha_0 \\ r_2(x) = \frac{b_1 c_1}{(a_1+b_1+c_1)^2} - \alpha_0^2 \\ r_3(x) = \frac{b_2 c_2}{(a_2+b_2+c_2)^2} - \alpha_0^2 \\ r_4(x) = \frac{a_0 a_2}{(a_0+b_0+c_0)(a_2+b_2+c_2)} - \alpha_0^2 \\ r_5(x) = \frac{c_0 b_1}{(a_0+b_0+c_0)(a_1+b_1+c_1)} - \alpha_0^2 \\ r_6(x) = \frac{b_1 c_2}{(a_1+b_1+c_1)(a_2+b_2+c_2)} - \alpha_0^2 \end{cases} \quad (4.31)$$

et  $P(x)$  l'ensemble de contraintes (correspondant aux équations 4.20, 4.22, 4.24, 4.26, 4.28, 4.30, du second critère à respecter, transformées en équations aux différences)

$$P(x) \begin{cases} p_1(x) = a_0^2 - b_0 c_0; \\ p_2(x) = b_1 c_2 - a_1 a_2; & p_3(x) = b_1 c_2 - b_2 c_1; \\ p_4(x) = b_1 c_0 - a_1 a_0; & p_5(x) = b_1 c_0 - c_1 b_0; \\ p_6(x) = a_0 a_2 - b_2 c_0; & p_7(x) = a_0 a_2 - b_0 c_2; \\ & p_8(x) = b_2 c_2 - a_2^2; \\ & p_9(x) = b_1 c_1 - a_1^2; \end{cases} \quad (4.32)$$

L'ordonnancement lexicographique est utilisé comme méthode de résolution pour obtenir l'ensemble des combinaisons de paramètres qui respectent tous les objectifs. L'algorithme correspondant est donné ci-dessous (**Algorithme 2**).

Tout d'abord, l'espace de recherche  $\Delta$ , qui représente le domaine de définition des probabilités, est discrétisé. Nous choisissons  $\Delta = ]0; 1[$  car les valeurs 0 et 1 sont considérées comme triviales.

Dans cet algorithme, les inégalités concernant  $r_1$ ;  $r_2$  et  $r_3$  sont résolues en premier. Ces 3 inégalités dépendent respectivement de 3 probabilités conditionnelles :  $[a_0, b_0, c_0]$  pour  $r_1$ ;  $[a_1, b_1, c_1]$  pour  $r_2$  et  $[a_2, b_2, c_2]$  pour  $r_3$ . Les solutions relatives à ces trois objectifs sont retenues et sauvegardées respectivement dans  $X_1$ ;  $X_2$  et  $X_3$ . Par la suite, l'inégalité concernant  $r_4$ , qui dépend de  $[a_0, b_0, c_0, a_2, b_2, c_2]$ , est résolue. Les valeurs testées pour résoudre cette inégalité proviennent des solutions déjà retenues à l'étape précédente, c'est-à-dire  $X_1$  pour  $[a_0, b_0, c_0]$  et  $X_3$  pour  $[a_2, b_2, c_2]$ . Le même raisonnement est suivi par la suite pour la résolution de  $r_5$  et  $r_6$  tout en prenant en compte les contraintes.

Ainsi, l'ensemble des solutions obtenues contient  $q$  combinaisons différentes de probabilités conditionnelles. Chaque combinaison  $[a_0, b_0, \dots, c_2]$  parmi cet ensemble respecte tous les objectifs et toutes les contraintes énoncées précédemment. Pour un pas de discrétisation de 0.1, 19 solutions sont obtenues.

Par conséquent, cette démarche méthodologique garantit que si l'utilisateur choisit une combinaison de probabilités conditionnelles  $[a_0, b_0, \dots, c_2]$  parmi les solutions proposées, tous les objectifs définis précédemment dans la Table 4.2 seront atteints et la fusion de décision sera améliorée.

#### 4.4.6 Fusion des décisions en ligne

Dans l'étape en ligne, les décisions des méthodes individuelles sont introduites à chaque instant dans le RB au niveau des noeuds  $D_1$  et  $D_2$ . Le résultat correspondant à la fusion globale est renvoyé par le noeud  $D$  à chaque instant.

L'estimation de l'état du système va servir à la prédiction de la RUL. Si le système est dans un état de dégradation légère, la RUL sera grande tandis que s'il se trouve dans un état de dégradation sévère, la RUL sera petite.

**Exemple** Afin de vérifier l'exactitude des solutions obtenues et ainsi établir l'efficacité de la méthodologie proposée, une phase de test est nécessaire. Le test montrera que si nous choisissons une combinaison de probabilités conditionnelles parmi l'ensemble obtenu de solutions réalisables (soit un total de 19 solutions), la performance globale en pronostic sera améliorée.

Les données de test sont illustrées dans la Figure 4.14.

Ces données représentent l'évolution de la dégradation du capteur de même type que celui étudié en phase d'apprentissage. Par ailleurs, contrairement aux données d'apprentissage, les données de test sont arrêtées à un instant donné, correspondant ici à  $t = 110$ . L'objectif est de déterminer la durée de vie résiduelle du système à cet instant, sachant que la vraie RUL est égale à 190.

---

**Algorithme 2** : Résolution du problème multi-objectif pour le pronostic

---

**Entrée:**  $\Delta, R(x), P(x)$   
 $X = [a_0, b_0, c_0, a_1, b_1, c_1, a_2, b_2, c_2]$   
 $X, X_1, X_2, X_3, X_4, X_5, X_6 \leftarrow \emptyset$   
**Tant que**  $R(x) \neq \emptyset$  **faire**  
  **pour**  $x \in \Delta$  **faire**  
    **si**  $r_1(x) > 0$  **alors**  
       $X_1 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_1(x)\}$   
    **fin si**  
    **si**  $r_2(x) > 0$  **alors**  
       $X_2 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_2(x)\}$   
    **fin si**  
    **si**  $r_3(x) > 0$  **alors**  
       $X_3 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_3(x)\}$   
    **fin si**  
  **fin pour**  
  **pour**  $x \in \{X_1 \cup X_3\}$  **faire**  
    **si**  $r_4(x) > 0$  **alors**  
       $X_4 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_4(x)\}$   
    **fin si**  
  **fin pour**  
  **pour**  $x \in \{X_2 \cup X_4\}$  **faire**  
    **si**  $r_5(x) > 0$  **alors**  
       $X_5 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_5(x)\}$   
    **fin si**  
  **fin pour**  
  **pour**  $x \in \{X_4 \cup X_5\}$  **faire**  
    **si**  $r_6(x) > 0$  **alors**  
       $X_6 \leftarrow x$   
       $R(x) \leftarrow R(x) \setminus \{r_6(x)\}$   
    **fin si**  
  **fin pour**  
  **si**  $P(X_6) > 0$  **alors**  
     $X \leftarrow X_6$   
  **fin si**  
**fin Tant que**  
**Retourner**  $X$

---

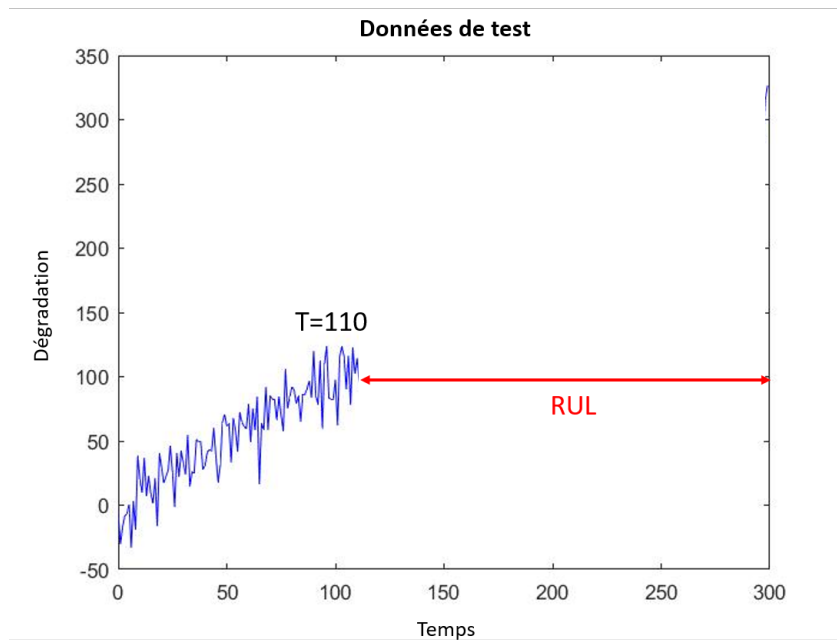


FIGURE 4.14 – Exemple numérique : évolution de la dégradation du capteur test

La structure du RB, illustrée par la Figure 4.13, correspond à la couche générique montrée auparavant.

Puisque le noeud global  $D$  doit décider entre 3 modalités,  $\{F_0, F_1, F_2\}$ , la probabilité a priori de chaque classe est égale à  $\frac{1}{3}$ .

En ce qui concerne les tables de probabilités conditionnelles des noeuds  $D_1$  et  $D_2$ , on peut choisir n'importe quelle combinaison parmi l'ensemble des solutions déduites précédemment.

Les résultats de fusion concernant le signal entier de test sont donnés dans la Table 4.5.

	$hs_1$	$hs_2$	$hs_3$
$hs_1$	98	2	0
$hs_2$	15	78	7
$hs_3$	4	1	95

TABLE 4.5 – Matrice de confusion de la méthode de fusion

Ces résultats illustrent les bonnes performances en classification de la méthode de fusion proposée puisque le taux de bonne classification RR est de 90.33%. Une comparaison est donnée dans la Table 4.6, et reflète l'amélioration des performances en classification qui peut être atteinte avec la méthodologie proposée.

	$m_1$	$m_2$	<b>Fusion</b>
<b>RR</b>	73.33	78.67	90.33

TABLE 4.6 – Comparaison de performances (%)

Enfin, à l'instant  $t = 110$ , la méthode  $m_1$  estime que le système est en état de santé  $hs_2$  tandis que la méthode  $m_2$  estime qu'il se trouve en état de santé  $hs_1$ . D'où les RULs suivantes :

$$RUL_{m_1} = 2 * 100 - \frac{100}{2} = 150 \quad (4.33)$$

$$RUL_{m_2} = 1 * 100 - \frac{100}{2} = 50 \quad (4.34)$$

D'après la stratégie de fusion adoptée, la décision globale suit la méthode experte en  $hs_2$ , soit la méthode  $m_1$ . Ainsi, la décision globale est  $hs_2$ . A partir de cette décision, la RUL est déduite :

$$RUL_{Fusion} = 2 * 100 - \frac{100}{2} = 150 \quad (4.35)$$

Chaque RUL estimée possède une incertitude de  $\pm 50$  unités de temps.

On constate que la fusion et la méthode  $m_1$  présentent les meilleures performances pour la prédiction de la RUL à l'instant  $t = 110$  (par rapport à la RUL réelle de 190).

## 4.5 Synthèse

En plus des résultats remarquables en pronostic, l'originalité et l'intérêt scientifique de la méthodologie proposée sont les suivants :

- Elle fournit un cadre générique où plusieurs méthodes de pronostic basées sur les données peuvent être combinées. La méthodologie à suivre est théorique et répond ainsi au principal obstacle rencontré lors du développement d'approches fusionnant plusieurs méthodes, que ce soit pour le diagnostic ou le pronostic. Par conséquent, l'approche proposée n'est pas dédiée à des applications particulières.
- La méthodologie de pronostic proposée repose sur la classification supervisée et permet donc d'utiliser toute méthode de classification, comme l'AD, le kNN, l'arbre de décision, le SVM, le RB, etc. Les données expérimentales représentatives des phénomènes de dégradation sont exploitées et l'utilisateur s'affranchit du développement de modèles complexes, qui peut être très difficile à mettre en oeuvre.
- L'approche de pronostic proposée n'a pas besoin de définir de seuils de défaillance comme la plupart des méthodes à base de données. Généralement, le calcul de la RUL résulte de l'intersection de la prédiction de l'état futur du composant et de son seuil de défaillance [Medjaher, 2014]. Par conséquent, l'estimation de ces seuils est une tâche très importante puisque des valeurs erronées conduisent à des RULs erronées. Dans l'approche proposée, la RUL est déduite directement de l'estimation de l'état courant de santé, à travers une formule ne faisant pas intervenir de seuil de défaillance.
- L'approche proposée de fusion applique une stratégie efficace pour combiner les résultats individuels une fois disponibles à partir de diverses méthodes. Par conséquent, elle surmonte la difficulté de connaître et de combiner différentes théories et concepts spécifiques à chaque méthode de pronostic. Il convient de noter que l'utilisateur peut appliquer une méthode individuelle différente de celle proposée dans la Partie 4.3, du moment que cette méthode génère des décisions représentant des états de santé.
- Enfin, le cadre permet d'utiliser, à travers les méthodes individuelles, tous les progrès et les avancées réalisés par la communauté de pronostic guidée par les données.

Par la suite, nous allons appliquer la méthodologie proposée sur un système complexe, un ensemble de turboréacteurs, afin de montrer la faisabilité et les améliorations apportées par cette approche.

## 4.6 Application : Pronostic de turboréacteurs à double flux

### 4.6.1 Introduction

Ces dernières années, l'intérêt pour le pronostic des systèmes n'a cessé d'augmenter et on assiste de plus en plus à l'émergence des approches qui exploitent les données fournies par les capteurs équipant les systèmes modernes. Ainsi, au fil du temps, plusieurs ensembles de données ont été spécialement fournis pour le pronostic. On peut citer par exemple les données provenant de la plateforme Pronostia [Nectoux et al., 2012], les données IMS (Intelligent Maintenance Systems (Systèmes de maintenance intelligents)) [Qiu et al., 2006] ou encore les données provenant d'un modèle de simulation de turboréacteurs C-MAPSS (Commercial Modular AeroPropulsion System Simulation) [Frederick et al., 2007]. Ces dernières ont d'abord été utilisées lors de la conférence PHM'08 comme données de challenge, dans le but de tester et comparer plusieurs méthodes de pronostic [Saxena et al., 2008b]. Depuis, ces jeux de données ont été largement utilisés par les chercheurs pour résoudre le problème de pronostic [Ramasso and Saxena, 2014, Khelif et al., 2017, Hu et al., 2012, Xu et al., 2014, Peng et al., 2012, Xi et al., 2013, Wang et al., 2012, Mosallam et al., 2016, Ramasso et al., 2013a]. Ainsi, certains travaux ont profité de la grande quantité de données disponibles concernant le passage à la défaillance (RtFD) pour développer des approches à base d'intelligence artificielle qui apprennent les relations entre les indicateurs de santé et les RULs [Ramasso and Saxena, 2014, Xu et al., 2014, Peng et al., 2012, Xi et al., 2013]. D'autres travaux ont proposé des approches basées sur la similarité, grâce aux différentes données d'apprentissage [Wang et al., 2012, Mosallam et al., 2016, Ramasso et al., 2013a]. Les conditions opérationnelles variables dans le temps de cet ensemble de données ont également été prises en compte dans certaines publications [Lin et al., 2013]. Enfin, plusieurs chercheurs ont transformé le problème de pronostic en un problème de classification, en divisant l'évolution des données en plusieurs états de santé [Javed et al., 2015, Ramasso and Gouriveau, 2014].

Dans ce chapitre, nous allons tout d'abord présenter les données issues des turboréacteurs à double flux [Frederick et al., 2007] puis développer deux approches de pronostic basées sur les données. Les résultats de ces deux approches seront fusionnés selon la méthodologie présentée, puis comparés selon des métriques spécifiques au pronostic.

### 4.6.2 Présentation du turboréacteur à double flux

C-MAPSS est un outil développé à la NASA permettant la simulation du modèle de turboréacteurs à double flux, dont le schéma simplifié est donné dans la Figure 4.15.

Des données provenant de diverses parties du système ont été recueillies pour enregistrer les effets des dégradations sur les mesures des capteurs et fournir des séries temporelles multivariées présentant des comportements de dégradation dans plusieurs unités de turboréacteurs [Ramasso and Saxena, 2014]. Ainsi, cinq ensembles de données

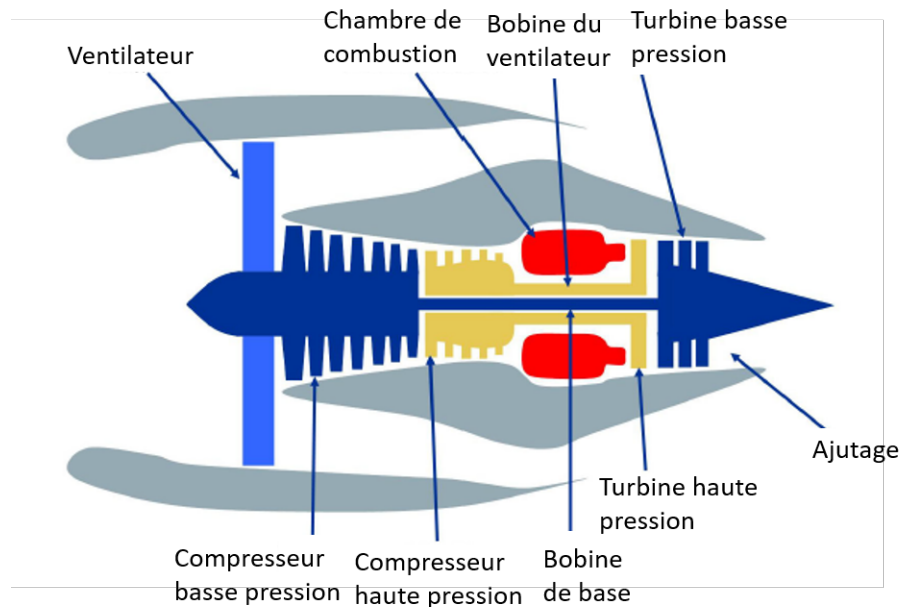


FIGURE 4.15 – Schéma simplifié du turboréacteur simulé dans C-MAPSS [Frederick et al., 2007]

de passage à la défaillance ont été générés dans différentes conditions opérationnelles, comme le résume la Table 4.7.

<b>Ensemble de données</b>	1	2	3	4	PHM'08
<b>Trajectoires d'apprentissage</b>	100	260	100	248	218
<b>Trajectoires de test</b>	100	259	100	249	218
<b>Conditions opérationnelles</b>	1	6	1	6	6
<b>Modes de défaillances</b>	1	1	2	2	2

TABLE 4.7 – Ensembles de données C-MAPSS

Chaque ensemble consiste en un jeu de données d'apprentissage et un jeu de test et comporte 21 paramètres mesurés tels que la température, la pression, la vitesse, etc.

La vraie RUL est fournie pour 4 ensembles de données afin d'évaluer les erreurs de prédiction et calculer les métriques de pronostic à des fins de comparaison. Par ailleurs, la vraie RUL reste inconnue pour l'ensemble qui a été utilisé pour le challenge de la conférence IEEE PHM 2008 [Saxena et al., 2008b].

Les données sont considérées comme provenant d'une flotte de moteurs du même type. Chaque moteur démarre avec différents degrés d'usure initiale et de variations de fabrication, inconnus de l'utilisateur. Cette usure et cette variation sont considérées comme normales. Ainsi, le moteur fonctionne normalement au début de chaque série temporelle et commence à se dégrader à un moment donné. Dans le jeu de données d'apprentissage, la dégradation augmente en amplitude jusqu'à atteindre un seuil prédéfini au-delà duquel il n'est pas préférable de faire fonctionner le moteur. Il s'agit donc de l'EOL du système, comme représenté sur la Figure 4.16. Notons que les utilisateurs n'ont pas d'accès explicite au modèle de simulation et que le seuil de défaillance n'a pas été annoncé a priori.

Dans le jeu de données de test, la série temporelle se termine un certain temps avant la



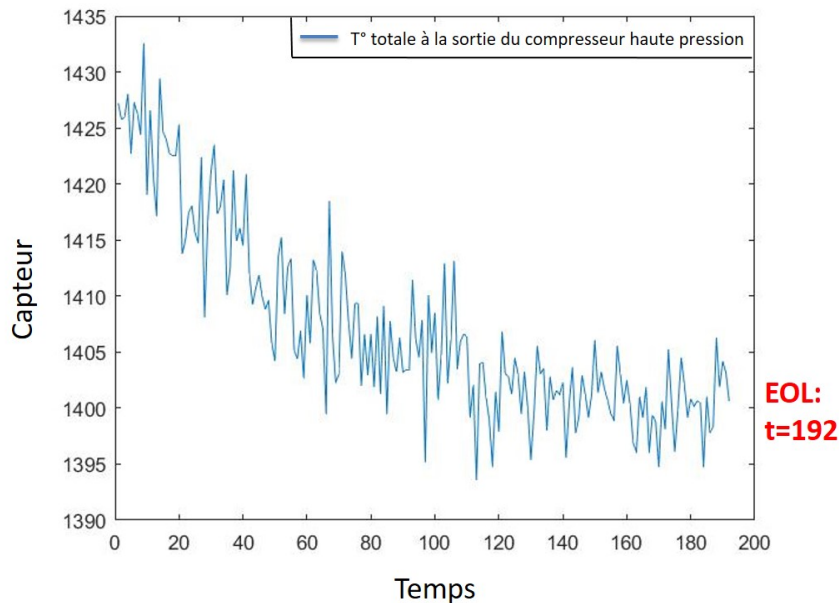


FIGURE 4.16 – Mesure issue d'un capteur

fin de la dégradation. L'objectif est donc de prédire le nombre de cycles de fonctionnement restants avant la défaillance (RUL) dans le jeu de données de test.

L'intérêt pour ces ensembles de données réside dans le fait que C-MAPSS modélise de façon très précise et aussi réaliste que possible le système réel non-linéaire complexe. En outre, le bruit ainsi que de nombreuses sources de variabilité sont pris en compte. Enfin, une grande quantité de données d'apprentissage, provenant d'au moins une centaine de turboréacteurs, est disponible pour chaque ensemble.

Tout ceci rend les données du turboréacteur à double flux adaptées au développement de méthodes de pronostic basées sur les données, où très peu (voire aucune) information sur le système sont disponibles.

Dans ce travail, nous nous intéressons particulièrement au premier ensemble de données, qui a été largement utilisé dans la littérature [Ramasso and Saxena, 2014]. Il consiste en 100 jeux d'apprentissage totalisant 20631 observations, 100 jeux de test totalisant 13096 observations et un fichier enregistrant la vraie RUL des 100 jeux de test.

### 4.6.3 Pronostic de turboréacteurs par des méthodes individuelles

Comme on dispose uniquement de données historiques, la méthodologie de pronostic proposée dans la Partie 4.3 est appliquée. Plusieurs méthodes de classification basées sur les données ont été testées et l'AD ainsi que le kNN ont été choisies pour leurs performances qui conviennent dans notre contexte.

Tout d'abord, une sélection de capteurs est effectuée afin de ne garder que les paramètres informatifs. Ensuite, l'apprentissage est réalisée et les performances individuelles sont évaluées en termes de matrices de confusion pour la phase d'estimation de l'état de santé et en termes de métriques de pronostic pour la prédiction de la RUL.

### Signaux d'apprentissage et sélection de capteurs

L'objectif de cette partie est de développer deux méthodes de pronostic basées sur les données de passage à la défaillance du jeu de données 1, illustré par la Figure 4.17.

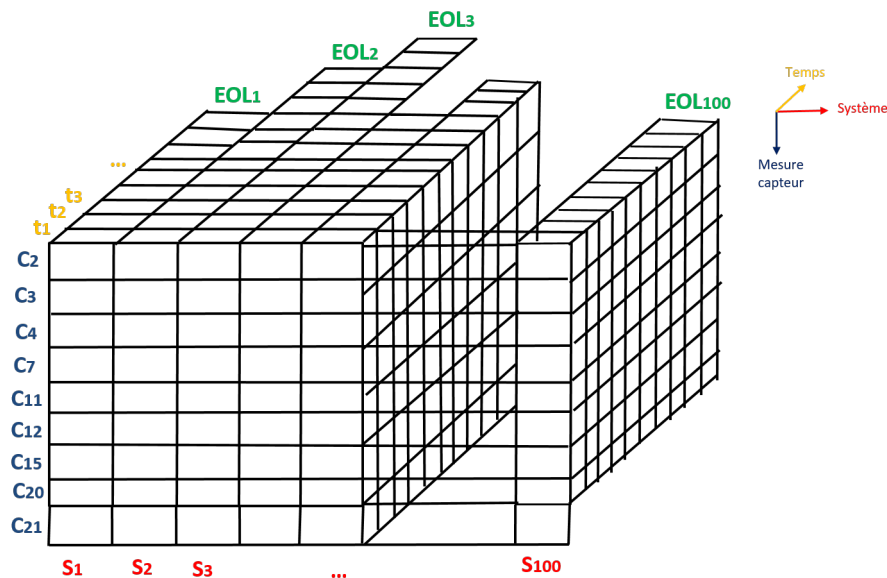


FIGURE 4.17 – Données d'apprentissage issues de 100 turboréacteurs

Ces données représentent l'évolution de la dégradation jusqu'à la défaillance pour 100 turboréacteurs ( $S_1, \dots, S_{100}$ ).

La première étape consiste à sélectionner les capteurs montrant une tendance de dégradation cohérente pour toutes les unités d'apprentissage. En effet, inclure des capteurs dont les mesures ne présentent pas une tendance de dégradation claire, en raison du bruit élevé ou de leur faible sensibilité à la dégradation, peut réduire la précision de la prédiction.

D'après [Wang et al., 2008], l'utilisation du sous-ensemble de capteurs (2, 3, 4, 7, 11, 12, 15, 20 et 21) peut aider à améliorer la précision de l'estimation de la RUL, c'est pourquoi cet ensemble sera sélectionné pour la suite (voir Table 4.8).

Capteur et symbole	Description
2 : $T_{24}$	Température totale à la sortie du compresseur basse pression
3 : $T_{30}$	Température totale à la sortie du compresseur haute pression
4 : $T_{50}$	Température totale à la sortie de la turbine basse pression
7 : $P_{30}$	Pression totale à la sortie du compresseur haute pression
11 : $P_{s30}$	Pression statique à la sortie du compresseur haute pression
12 : $\phi$	Rapport du débit de carburant à $P_{s30}$
15 : $BPR$	Taux de bypass
20 : $W_{31}$	Purge du liquide de refroidissement à turbine haute pression
21 : $W_{32}$	Purge du liquide de refroidissement à turbine basse pression

TABLE 4.8 – Paramètres choisis

**Traitement des données** Chaque turboréacteur se retrouve en état de défaillance à un instant donné représentant son *EOL*. La Figure 4.18 illustre l'évolution de la même mesure, la température totale à la sortie du compresseur basse pression, provenant de 100 turboréacteurs différents.

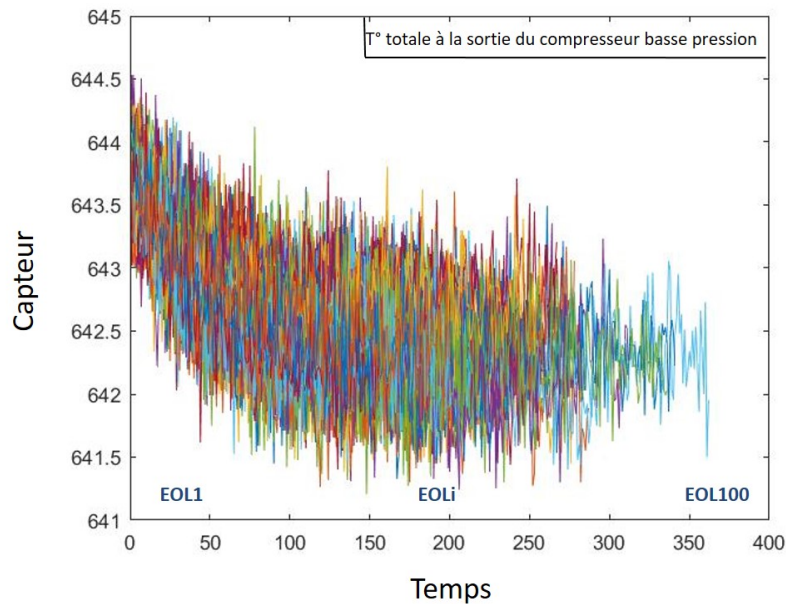


FIGURE 4.18 – Evolution de la température totale à la sortie du compresseur basse pression pour les 100 turboréacteurs considérés

On constate alors que chaque turboréacteur a sa propre *EOL*. La Figure 4.19 représente les 100 différentes *EOLs* dans l'ordre croissant.

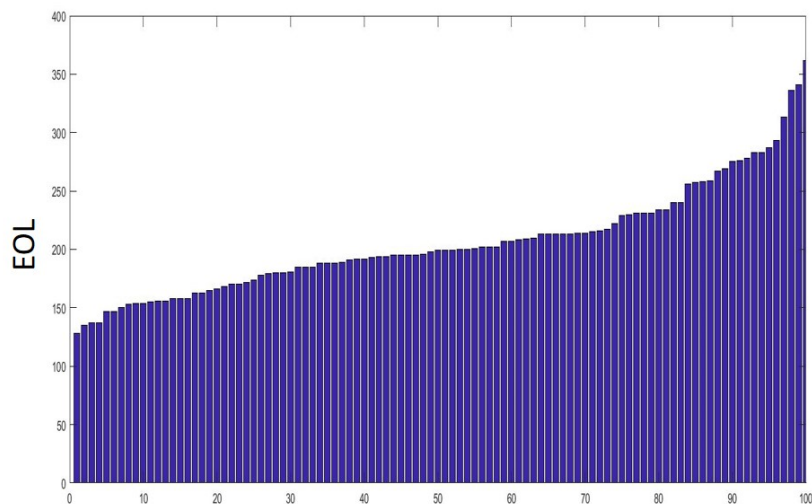


FIGURE 4.19 – *EOL* des 100 turboréacteurs.

On peut remarquer que la plus petite *EOL* est de 128 unités de temps, tandis que la plus grande est de 362. Une comparaison de l'évolution de la même mesure, c'est-à-dire la température totale à la sortie du compresseur basse pression, pour deux turboréacteurs différents : celui ayant la plus petite *EOL* et celui ayant la plus grande, est représentée

dans la Figure 4.20.

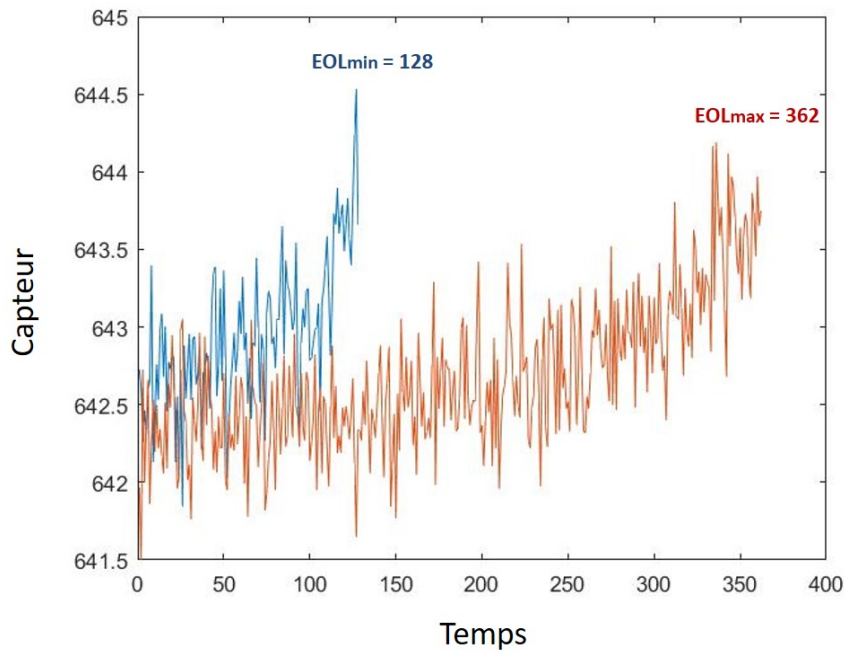


FIGURE 4.20 – Evolution de la même mesure pour les 2 turboréacteurs considérés

Comme décrit précédemment (voir Partie 4.3), nous procédons à un recalage des signaux afin d’avoir le même état de santé à la fin, c’est-à-dire le *EOL* pour tous les turboréacteurs, comme le montre la Figure 4.21.

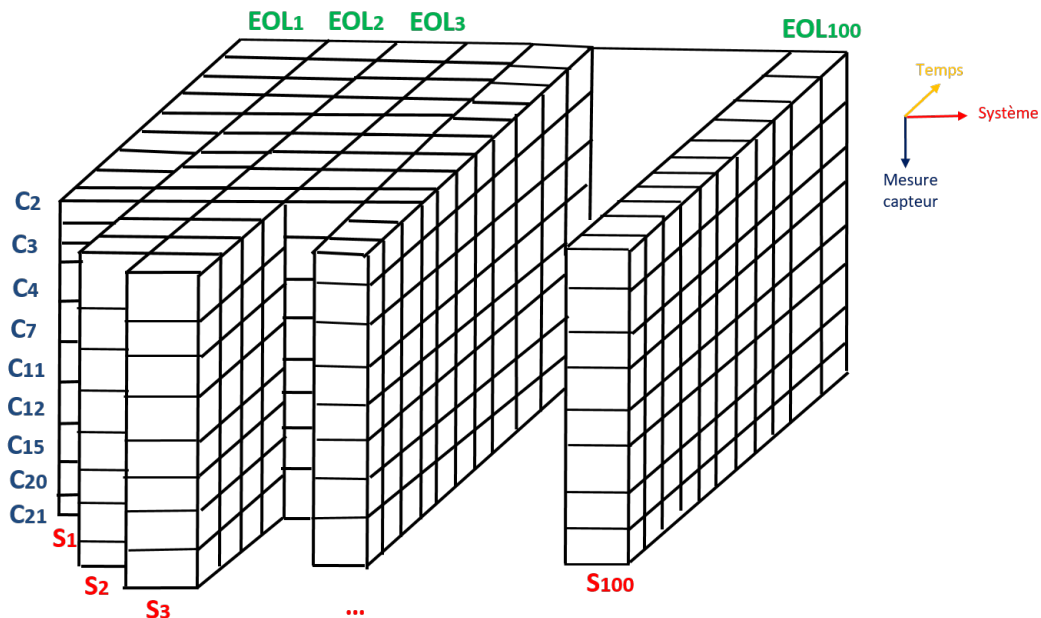


FIGURE 4.21 – Données d’apprentissage issues de 100 turboréacteurs, avec le recalage proposé

**Base d’apprentissage** Le choix du nombre d’états de santé n’est pas prédéterminé. Dans la suite, nous choisissons de détailler le cas où les signaux sont découpés en 4 classes.

Une étude comparative des performances en pronostic de la fusion en fonction du nombre d'états de santé (c'est-à-dire de classe) sera présentée dans la partie Discussion.

Ainsi, afin de construire la base d'apprentissage, 4 états de santé sont considérés ici, à savoir  $hs_1$ ;  $hs_2$ ;  $hs_3$  et  $hs_4$ .

$hs_1$  correspond à l'état de dégradation le plus sévère tandis que  $hs_4$  correspond à l'état de dégradation le plus léger.

Les classifieurs peuvent donc apprendre les liens entre les signaux à un instant donné et l'état de santé correspondant. Les matrices de confusion de l'AD et du kNN sont données dans les Tables 4.9a et 4.9b.

$m_1$	$hs_1$	$hs_2$	$hs_3$	$hs_4$
$hs_1$	83.90	15.93	0.17	0
$hs_2$	11.27	62.00	21.17	5.57
$hs_3$	0.2	28.13	33.97	37.70
$hs_4$	0	9.77	29.77	60.47

(a) Matrice de confusion de  $m_1$  (AD)

$m_2$	$hs_1$	$hs_2$	$hs_3$	$hs_4$
$hs_1$	80.80	16.97	1.33	0.9
$hs_2$	14.13	53.17	18.33	14.37
$hs_3$	1.70	21.80	33.17	43.33
$hs_4$	0.33	7.13	22.27	70.27

(b) Matrice de confusion de  $m_2$  (kNN)

TABLE 4.9 – Evaluation des performances individuelles en classification, pour 4 classes

Selon les matrices de confusion, les taux de bonne classification sur les données d'apprentissage sont respectivement  $RR_{m_1} = 60.09\%$  et  $RR_{m_2} = 59.35\%$ .

**Signaux de test et sélection de variables** Les données de test proviennent de 100 turbo-réacteurs de même type. Les mêmes variables sélectionnées dans la phase d'apprentissage, c'est-à-dire les signaux des capteurs (2, 3, 4, 7, 11, 12, 15, 20 et 21), sont récupérées dans la phase en-ligne. L'objectif est d'estimer la RUL pour chaque système, à un instant donné, comme l'illustre la Figure 4.22.

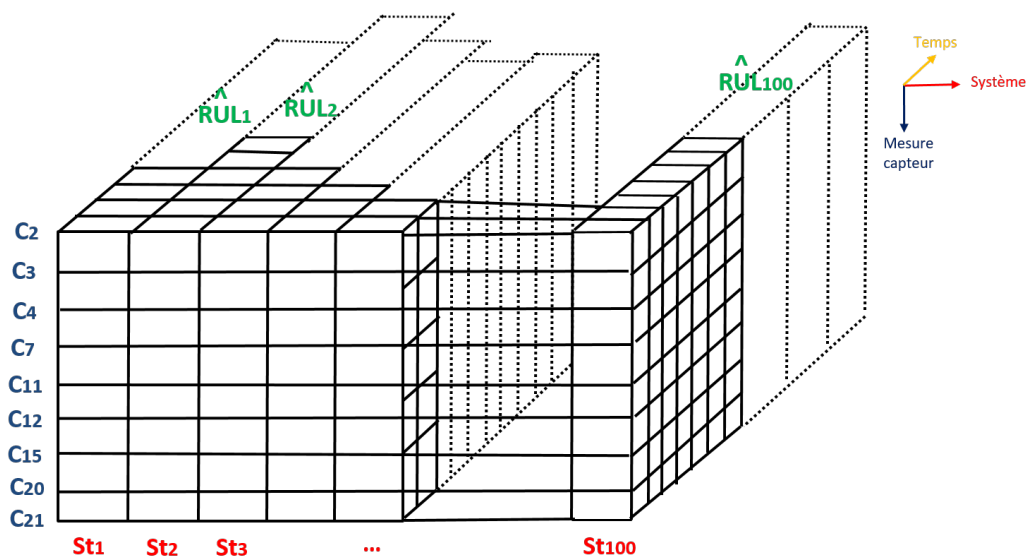


FIGURE 4.22 – Données de test issues de 100 turbo-réacteurs

**Estimation de l'état de santé et prédiction de la RUL** Lorsqu'on applique les deux méthodes individuelles (AD et kNN) sur les données de test, qui ont été arrêtées avant la défaillance, on peut déduire les RULs pour les 100 turboréacteurs. En effet, chaque méthode individuelle exploite sa base d'apprentissage et les données de test pour estimer l'état de santé du système, à l'instant considéré. La décision sur l'état de santé permet de calculer la RUL à cet instant, pour chaque système, selon la formule présentée auparavant (équation 4.2).

Comme la vraie RUL est disponible pour l'ensemble des 100 turboréacteurs, les différentes métriques de pronostic peuvent être calculées, comme l'illustre la Figure 4.23.

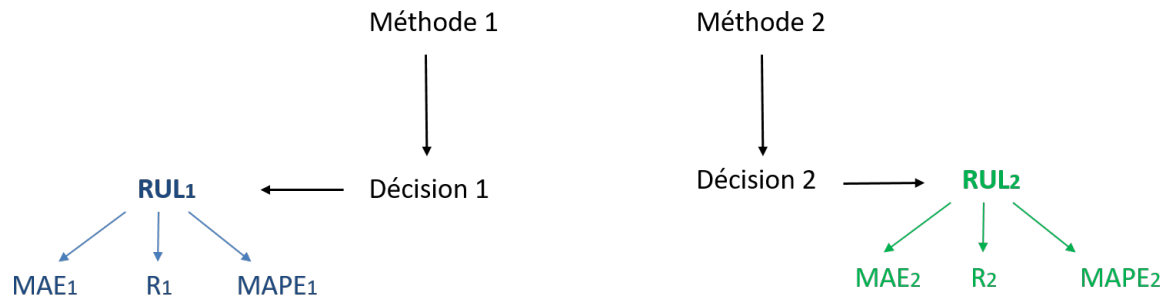


FIGURE 4.23 – Démarche des méthodes individuelles

Une comparaison des performances individuelles en pronostic des deux méthodes est présentée dans la Table 4.10. La couleur rouge indique à chaque fois le meilleur résultat.

	$m_1$ (AD)	$m_2$ (4NN)
Score $S_c$	939.1593	1361.1
Précision $R$	0.7135	0.7610
RULs à l'heure	45	58
Intervalle d'erreur sur la RUL	[-42; 63]	[-66; 48]
MAE	17.64	15.48
MSE	502.18	412.90
MAPE	32.3742	33.9850

TABLE 4.10 – Performances individuelles en pronostic, pour 4 classes

On constate que les performances en prédiction de RULs sont très bonnes dans le cas où on considère 4 classes. En effet, les précisions sont supérieures à 70% et les MAPEs sont inférieures à 34%.

Nous étudierons dans ce qui suit, les performances apportées par la stratégie de fusion de décisions.

#### 4.6.4 Pronostic de turboréacteurs par la fusion de décision

Dans cette partie, nous allons suivre les 3 étapes de la méthodologie proposée auparavant (voir Partie 4.4), afin de fusionner les décisions individuelles des deux méthodes développées pour le pronostic du TEP.

1. **Evaluation des performances individuelles** La première étape de la méthodologie de fusion proposée est l'évaluation des performances individuelles.

Les matrices de confusion des méthodes individuelles sont données dans la Figure 4.24.

m1	hs <sub>1</sub>	hs <sub>2</sub>	hs <sub>3</sub>	hs <sub>4</sub>
hs <sub>1</sub>	83.90	15.93	0.17	0
hs <sub>2</sub>	11.27	62.00	21.17	5.57
hs <sub>3</sub>	0.2	28.13	33.97	37.70
hs <sub>4</sub>	0	9.77	29.77	60.47

m2	hs <sub>1</sub>	hs <sub>2</sub>	hs <sub>3</sub>	hs <sub>4</sub>
hs <sub>1</sub>	80.80	16.97	1.33	0.9
hs <sub>2</sub>	14.13	53.17	18.33	14.37
hs <sub>3</sub>	1.70	21.80	33.17	43.33
hs <sub>4</sub>	0.33	7.13	22.27	70.27

FIGURE 4.24 – Matrices de confusion des deux méthodes de pronostic, pour 4 classes

A partir de ces matrices, le choix des types d'états de santé se fait de la façon suivante :

- La méthode AD ( $m_1$ ) a de meilleures performances pour la classe  $hs_2$ , donc nous choisissons  $F_1 = \{hs_2\}$  (en vert sur la Figure 4.24).
- La méthode kNN ( $m_2$ ) présente de meilleures performances que la méthode 1 pour la classe  $hs_4$ , ce qui conduit à  $F_2 = \{hs_4\}$  (en bleu sur la Figure 4.24).
- Les deux méthodes classent avec des performances similaires  $hs_1$  et  $hs_3$ . Par conséquent :  $F_0 = \{hs_1; hs_3\}$  (en rouge sur la Figure 4.24).

2. **Stratégie de FDD** Le cadre générique utilisé est celui dans la Figure 4.13.

3. **Fusion des décisions et prédiction de la RUL** Cette étape consiste à utiliser le RB paramétré en ligne pour la fusion des deux décisions. Ces dernières sont données à chaque instant par les deux méthodes et sont introduites au niveau des noeuds  $D_1$  et  $D_2$ .

A partir de la décision de la fusion pour chaque turboréacteur, la RUL est estimée selon la même formule proposée précédemment.

La démarche de fusion est illustrée par la Figure 4.25.

Une comparaison entre les métriques de pronostic des méthodes individuelles et celles de la fusion est donnée dans la Table 4.11. La couleur rouge indique le meilleur résultat.

	$m_1$ (AD)	$m_2$ (4NN)	Fusion
Score ( $S_c$ )	939.2	1361.1	695.8
Précision ( $R$ )	0.7135	0.7610	0.7788
RULs à l'heure	45	58	56
Intervalle d'erreur sur la RUL	[-42; 63]	[-66; 48]	[-42; 63]
MAE	17.6	15.5	14.9
MSE	502.18	412.90	383.50
MAPE	32.37	33.99	29.90

TABLE 4.11 – Comparaison des performances en pronostic, pour 4 classes



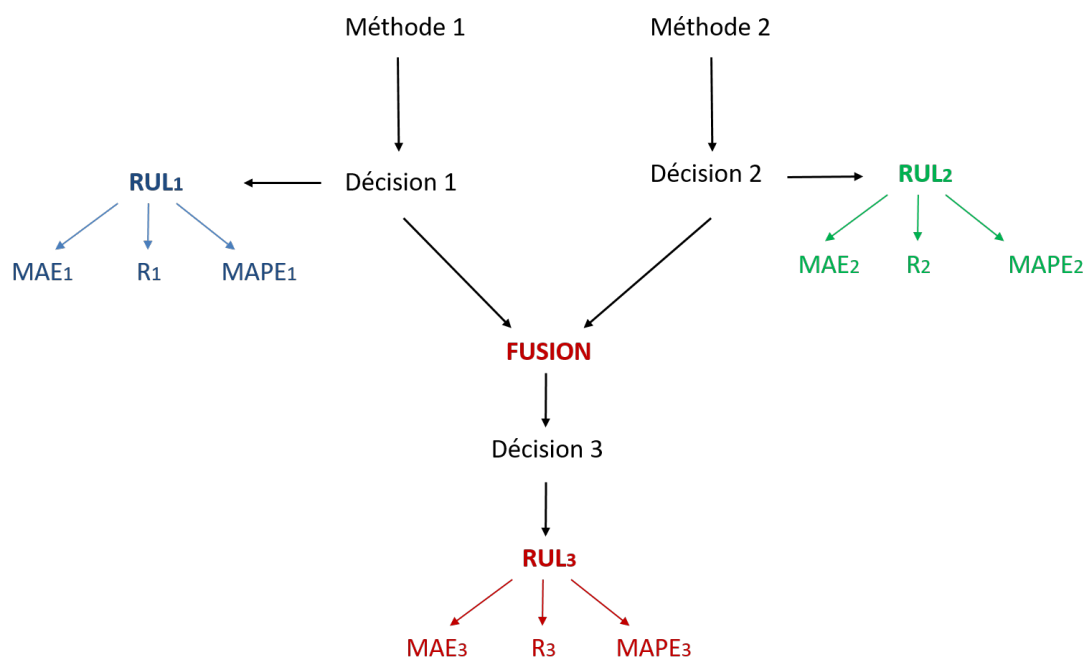


FIGURE 4.25 – Démarche de fusion

On constate que la méthode de fusion proposée avec 4 classes présente de très bonnes performances en pronostic. En effet, le score a été diminué (695.8 contre 939.2 pour l'AD et 1361.1 pour le kNN) et la précision de prédiction de la RUL est augmentée (0.7788 contre 0.7135 pour l'AD et 0.7610 pour le kNN). Parmi les 100 instances de tests, le nombre de RULs à temps, c'est-à-dire se situant dans l'intervalle  $I = [-10, 13]$ , est de 56. Concernant la distribution d'erreur de la RUL, la fusion proposée possède la même étendue que l'AD, soit  $I = [-42; 63]$ . Enfin, les  $MAE$ ,  $MSE$  et  $MAPE$  doivent être les plus petites possibles comme indiqué auparavant, et on peut remarquer que la fusion présente à chaque fois les plus petites erreurs moyennes.

On peut donc en déduire que, pour le cas où on dispose de 4 états de santé, la fusion proposée présente de meilleures performances en pronostic comparée aux méthodes individuelles la composant.

## 4.7 Discussion

Dans cette partie, nous allons revenir sur certains points importants des deux méthodologies proposées, celle du pronostic et celle de la fusion, comme l'influence de certains paramètres sur la prédiction de la RUL, l'analyse statistique des solutions obtenues, ainsi que la prise en compte des évidences incertaines.

### 4.7.1 Prédiction de la RUL

#### Influence des paramètres

Dans la méthodologie de pronostic proposée, la prédiction de la RUL repose sur l'estimation de l'état de santé. Cette estimation exploite le principe de classification supervisée.

Le nombre de classes choisi pour découper le signal n'est pas prédéterminé. Plus le nombre de classes est grand, plus la prédiction de la RUL pourrait s'améliorer. En effet, on



peut calculer autant de RULs différentes que de classes, comme l'illustre la Figure 4.26. Cependant, le taux de bonne classification  $RR$  pourrait diminuer avec l'augmentation du nombre de classes. En effet, plus le nombre de classes augmente, moins les classes comporteront des données. Par conséquent, elles seront moins bien représentées pour la classification (voir Figure 4.26). Il convient donc de tester plusieurs nombres de classes afin de déterminer le meilleur compromis.

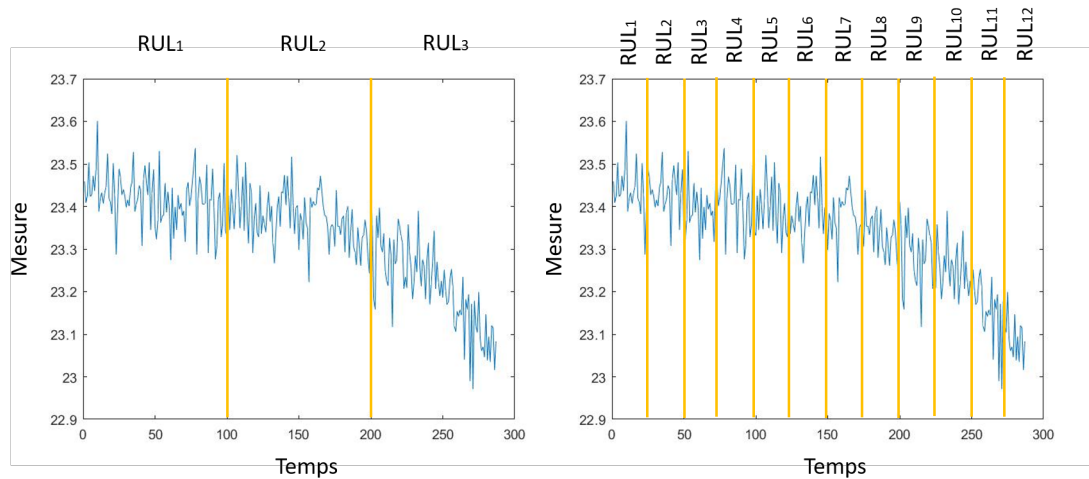


FIGURE 4.26 – Nombre de classes

A partir de l'état de santé estimé par le classifieur, la RUL est prédite selon la formule proposée dans la Partie 4.3 :

$$\widehat{RUL}_i = i \times q - \frac{q}{2} \quad (4.36)$$

Cette prédiction dépend de la classe estimée  $i$  ainsi que de la longueur  $q$  de la tranche correspondant à la classe estimée. Pour plus de simplicité, cette longueur  $q$  a été choisie constante dans notre approche, mais elle peut être variable selon les classes, comme l'illustre la Figure 4.27.

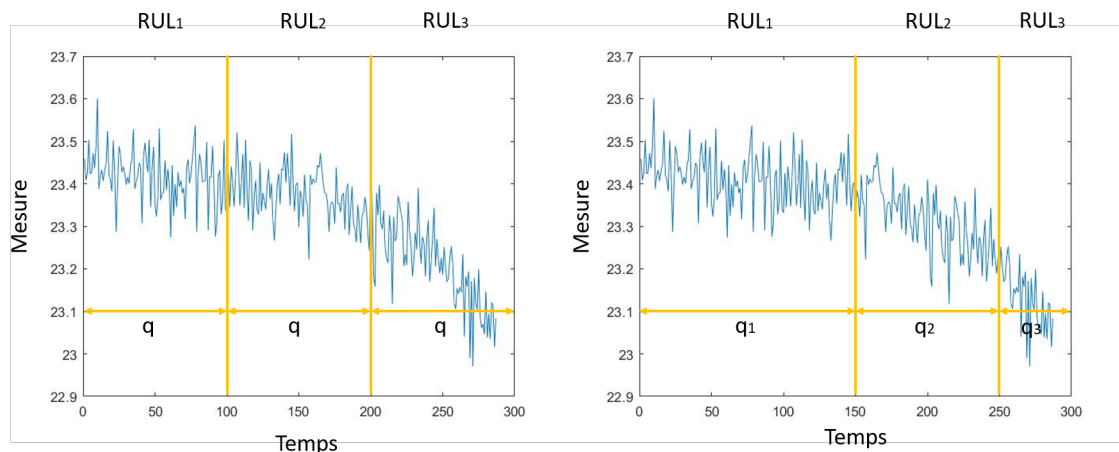


FIGURE 4.27 – Longueur de chaque classe

En outre, la formule 4.36 utilise la classe estimée  $i$  à l'instant  $t$  pour prédire la RUL. Une amélioration possible serait de filtrer les dernières décisions, c'est-à-dire de réaliser une moyenne sur les  $f$  dernières classes estimées. C'est ce qu'on appellera dans la suite la fenêtre de filtrage. La taille de cette fenêtre n'est pas prédéterminée. Par conséquent, il convient de tester plusieurs tailles de fenêtres de filtrage afin de déterminer le meilleur choix.

### Etude comparative

Dans cette partie, nous allons étudier l'influence du nombre de classe et de la taille de la fenêtre de filtrage sur les performances de la fusion en pronostic, pour l'ensemble des turboréacteurs. Pour cela, nous allons tester différentes configurations avec différents nombre de classes  $N_{classes} = \{3; 4; 5; 6; 7; 8; 9; 10\}$  et différentes fenêtres de filtrages  $F = \{1; 2; 3; 4; 5; 6; 7; 8\}$ .

Il convient de noter que le nombre de RULs estimées dépend du nombre de classes choisies. Plus ce nombre augmente, plus la précision de la prédiction de la RUL augmente. Par ailleurs, les classes contiennent de moins en moins de données et sont donc moins bien représentées. C'est pour cela que nous choisissons au minimum 3 classes et au maximum 10 classes. Concernant le filtrage,  $F = 1$  correspond au cas où seule la décision globale à l'instant  $t$  est choisie, tandis que  $F = i$  (avec  $i \in \{2, \dots, 8\}$ ) correspond au cas où la moyenne sur les  $i$  dernières valeurs de la décision globale est considérée.

Les tables suivantes (4.12, 4.13, 4.14, 4.16, 4.15, 4.17) résument les performances en pronostic de l'approche de fusion proposée en termes de métriques de pronostic, selon plusieurs configurations de nombre de classes et de taille de la fenêtre de filtrage, pour les 100 turboréacteurs de test. La couleur rouge indique le meilleur résultat.

Les Figures 4.28 ; 4.29 ; 4.32 ; 4.30 ; 4.31 ; 4.33 représentent ces mêmes métriques de pronostic pour l'approche de fusion en fonction du nombre de classes et de la fenêtre de filtrage.

<b>R</b>	<b>Fenêtre de filtrage</b>							
	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
3	0.64	0.67	0.66	0.68	0.67	0.68	0.68	0.68
4	0.73	0.75	0.76	0.77	<b>0.78</b>	0.77	0.77	0.77
5	0.72	0.75	0.74	0.75	0.76	0.76	0.75	0.75
6	0.69	0.73	0.71	0.73	0.74	0.74	0.74	0.74
7	0.64	0.69	0.67	0.65	0.66	0.67	0.67	0.68
8	0.58	0.66	0.67	0.68	0.70	0.70	0.68	0.70
9	0.71	0.70	0.69	0.67	0.67	0.68	0.67	0.69
10	0.37	0.47	0.47	0.50	0.53	0.57	0.56	0.60

TABLE 4.12 – Comparaison de performances en précision R pour la fusion

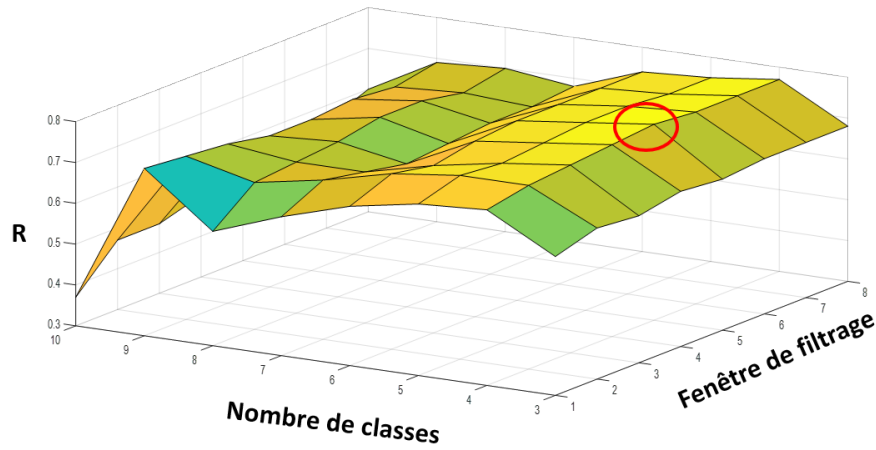


FIGURE 4.28 – Comparaison de la précision de prédiction des RULs

$Sc$	Fenêtre de filtrage							
	1	2	3	4	5	6	7	8
Nombre de classes	1	2	3	4	5	6	7	8
3	1361	1069	1142	1033	1087	1011	994	1006
4	872	854	732	710	<b>696</b>	758	730	758
5	1009	756	812	784	747	795	809	797
6	1184	909	989	905	896	837	869	834
7	1733	1204	1192	1191	1109	1109	1120	1036
8	3977	1277	1079	1128	973	1003	1068	965
9	1092	1141	956	1141	1212	1181	1174	1125
10	21494	5684	5590	4721	3152	2242	2424	1786

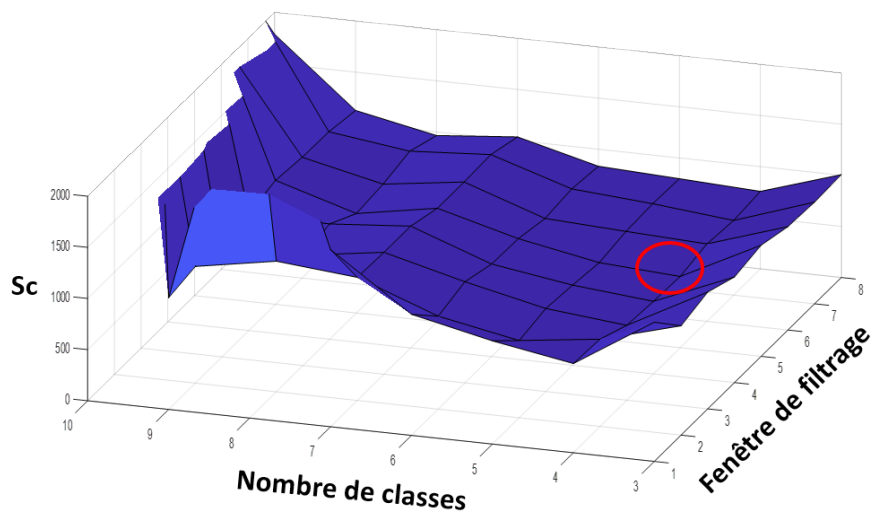
TABLE 4.13 – Comparaison de performances en score  $Sc$  pour la fusion

FIGURE 4.29 – Comparaison du score des RULs

MAE	Fenêtre de filtrage							
	Nombre de classes	1	2	3	4	5	6	7
3	19.38	18.90	19.28	18.74	18.82	18.55	18.79	18.81
4	17.48	16.49	16.16	15.44	14.98	<b>14.86</b>	15.14	15.32
5	16.50	16.02	16.74	16.04	15.70	15.36	15.88	15.92
6	17.28	16.10	16.91	15.96	15.56	15.70	15.99	15.97
7	18.40	17.66	18.34	19.07	18.82	18.59	18.99	18.72
8	19.05	17.94	18.31	17.71	17.16	17.05	17.84	17.43
9	16.66	17.09	17.48	17.82	17.96	17.96	18.33	17.84
10	22.90	22.30	22.28	21.26	20.89	20.14	20.81	19.70

TABLE 4.14 – Comparaison de performances en MAE pour la fusion

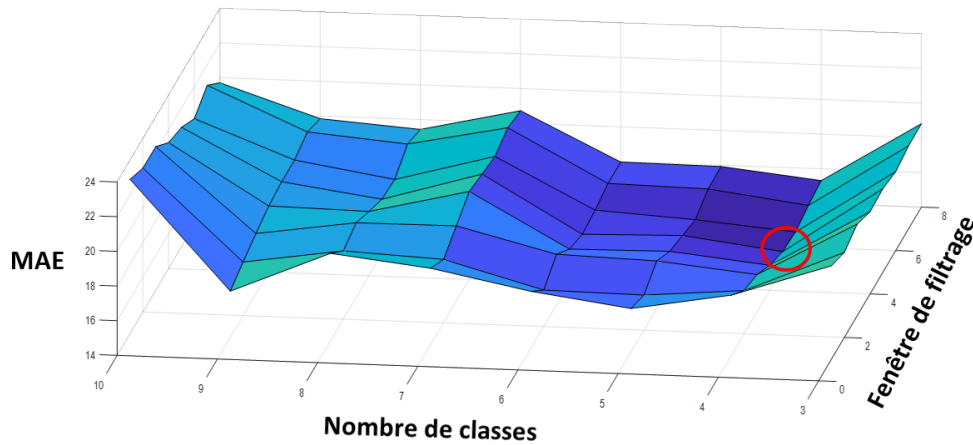


FIGURE 4.30 – Comparaison de l'erreur absolue moyenne des prédictions de RULs

MAPE	Fenêtre de filtrage							
	Nombre de classes	1	2	3	4	5	6	7
3	36.98	36.12	36.77	36.90	37.25	36.87	37.99	38.11
4	31.10	32.22	31.36	30.55	29.91	30.46	31.39	32.91
5	29.48	28.63	30.44	30.48	30.18	30.55	32.17	32.76
6	28.57	<b>27.17</b>	27.89	27.61	27.36	28.82	30.35	31.20
7	28.25	28.10	29.52	31.37	31.53	32.02	33.62	34.18
8	30.83	29.45	29.55	29.54	29.75	30.23	32.16	32.46
9	28.33	29.28	29.12	30.89	31.84	32.32	33.51	33.58
10	37.25	34.33	34.59	32.18	32.24	31.93	33.94	32.64

TABLE 4.15 – Comparaison de performances en MAPE pour la fusion

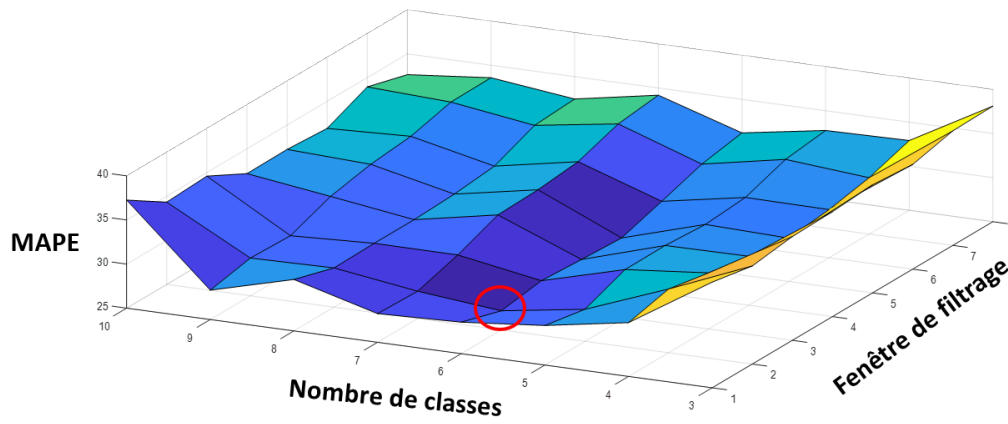


FIGURE 4.31 – Comparaison de l'erreur absolue moyenne en pourcentage des prédictions de RULs

MSE	Fenêtre de filtrage								
	Nombre de classes	1	2	3	4	5	6	7	8
3	635.74	584.14	609.61	577.34	585.66	572.01	570.83	569.59	
4	480.34	431.59	411.94	394.54	<b>383.50</b>	390.44	395.32	395.04	
5	489.34	443.02	463.90	434.14	419.22	414.70	427.94	427.30	
6	546.94	483.74	509.61	467.94	458.22	447.67	461.25	453.04	
7	655.13	562.61	587.62	624.19	596.31	581.50	589.68	567.97	
8	764.59	599.59	582.49	571.25	525.43	529.64	560.02	528.20	
9	522.07	536.05	549.49	579.52	582.24	567.48	584.70	551.37	
10	764.59	599.59	582.49	571.25	525.43	529.64	560.02	528.20	

TABLE 4.16 – Comparaison de performances en MSE pour la fusion

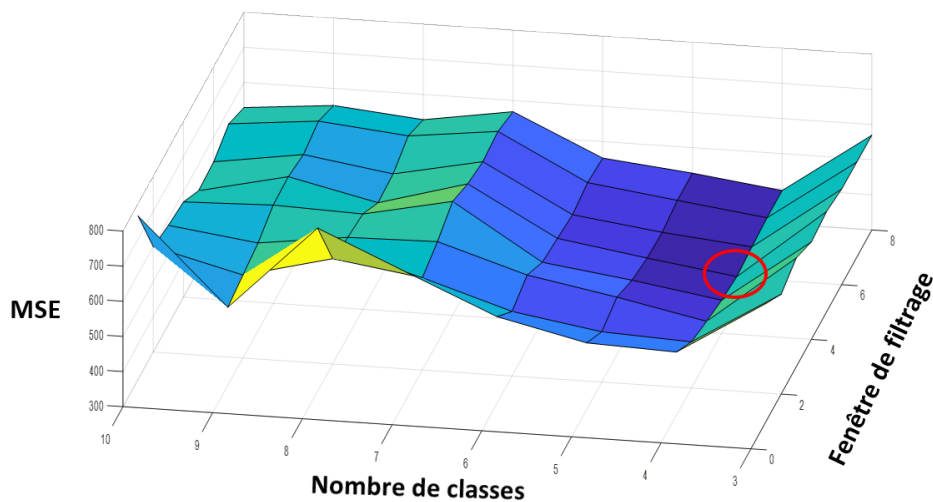


FIGURE 4.32 – Comparaison de l'erreur quadratique moyenne des prédictions de RULs

On-Time	Fenêtre de filtrage							
	1	2	3	4	5	6	7	8
3	46	44	43	43	44	44	44	45
4	43	47	48	51	56	56	51	50
5	51	48	50	53	51	53	51	51
6	50	54	49	<b>58</b>	56	54	53	55
7	44	44	43	45	41	42	42	43
8	48	47	44	49	50	48	45	46
9	50	48	48	46	51	46	44	46
10	47	46	47	46	44	43	39	43

TABLE 4.17 – Comparaison de performances en nombre de RULs à temps pour la fusion

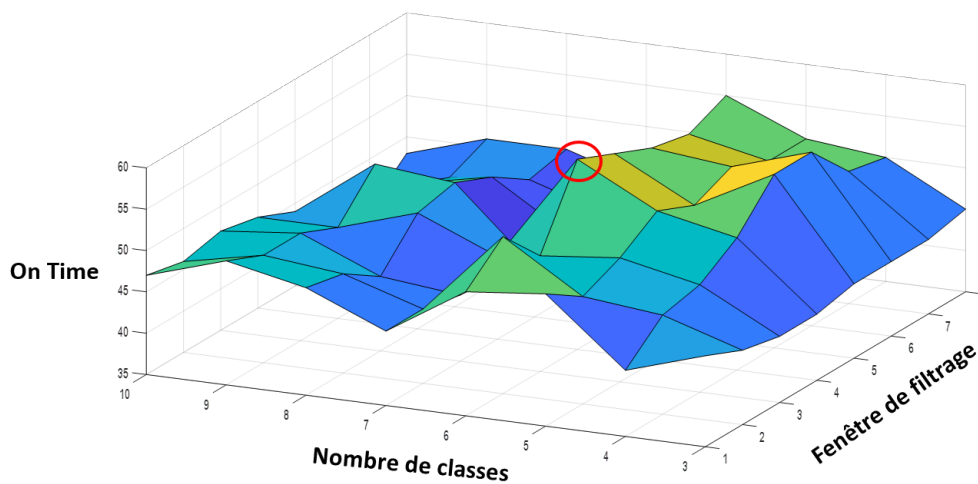


FIGURE 4.33 – Comparaison du nombre de RULs à temps

On constate que la configuration optimale est celle avec un nombre d'états de santé de 4 et une fenêtre de filtrage de 5. En effet, la précision  $R$ , le Score  $Sc$  ainsi que l'erreur quadratique moyenne  $MSE$  sont meilleurs avec cette configuration. Les autres métriques présentent également de très bonnes performances avec cette configuration. Par conséquent, 4 états de santé et une fenêtre de filtrage de 5 sont choisis pour l'analyse qui suit.

#### 4.7.2 Analyse des paramètres obtenus

Au regard de la table des objectifs (Table 4.2), on peut identifier les tables de probabilités conditionnelles suivantes (voir Tables 4.18), en se basant uniquement sur des évidences certaines.

En effet, à partir de la Table 4.2, on constate qu'il y a 3 situations où la décision globale  $d$  doit être égale à  $F_0$ . Sur ces 3 situations, 2 sont réalisées lorsque la méthode  $m_1$  choisit la décision  $d_1 = F_0$  et une lorsque la méthode  $m_1$  choisit la décision  $d_1 = F_2$ . D'où les probabilités conditionnelles suivantes :  $P(D_1 = F_0/D = F_0) = \frac{2}{3}$  et  $P(D_1 = F_2/D = F_0) = \frac{1}{3}$ . Enfin, sur ces 3 situations, la méthode  $m_1$  ne choisit jamais la modalité  $F_1$ . On en déduit :  $P(D_1 = F_1/D = F_0) = 0$ . Le même raisonnement est appliqué pour déduire les autres probabilités conditionnelles.

Cependant, nous allons montrer par la suite que cette solution (appelée *solution*

$D \backslash D_1$	$F_0$	$F_1$	$F_2$
$F_0$	$\frac{2}{3}$	0	$\frac{1}{3}$
$F_1$	0	$\frac{3}{4}$	$\frac{1}{4}$
$F_2$	$\frac{1}{2}$	0	$\frac{1}{2}$

(a) Table du noeud  $D_1$ 

$D \backslash D_2$	$F_0$	$F_1$	$F_2$
$F_0$	$\frac{2}{3}$	$\frac{1}{3}$	0
$F_1$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$
$F_2$	0	0	1

(b) Table du noeud  $D_2$ 

TABLE 4.18 – Tables de probabilités conditionnelles pour le pronostic

*conventionnelle* afin de pouvoir la distinguer des 19 solutions théoriques) n'est pas toujours équivalente aux solutions que nous avons déterminées, notamment avec des évidences incertaines.

### Prise en compte des évidences incertaines

Jusqu'à présent, nous nous sommes uniquement concentrés sur la couche générique du RB proposé pour la fusion (voir Figure 4.13), pour déterminer le nombre de classes et la taille de filtrage optimaux. Nous avons considéré les évidences introduites au niveau des noeuds  $D_1$  et  $D_2$  comme certaines. Ces évidences certaines sont obtenues en introduisant les observations (0 ou 1) directement dans la couche générique ou en appliquant la règle du maximum a posteriori lorsque les méthodes individuelles génèrent des probabilités a posteriori différentes de 0 et 1. Dans ce contexte, toutes les solutions déduites théoriquement ainsi que la solution conventionnelle sont équivalentes, c'est-à-dire qu'elles conduisent à la même classification, et donc au même RUL.

Dans la suite, nous allons considérer le cas où on décide de ne pas appliquer la règle du maximum a posteriori afin de garder toutes les informations contenues dans les décisions individuelles. En effet, les évidences deviennent incertaines, c'est-à-dire qu'elles peuvent prendre toutes les valeurs entre 0 et 1, ce qui est plus réaliste. Ainsi, en intégrant cet aspect, il convient d'analyser les différentes solutions obtenues dans ce contexte d'évidences incertaines, et de les comparer en termes de RULs et de métriques de pronostic, pour les 100 turboréacteurs.

**MSE** Tout d'abord, on va chercher à savoir si des solutions théoriques ont des RULs estimées meilleures que celles de la solution conventionnelle (notée solution 1 dans la suite). Les vraies RULs des 100 turboréacteurs étant connues, on se base dans cette comparaison sur le MSE (Mean Square Error) donné par chaque solution.

Il convient de noter que la même étude peut être réalisée en s'appuyant sur les autres métriques (MAE, MAPE, R, Sc).

La Figure 4.34 illustre les MSEs des différentes solutions. On peut remarquer que la solution n°15 (en bleue dans la figure) présente le MSE le plus réduit ( $MSE = 379.9$ ). La solution conventionnelle ( $MSE = 450$ ) est représentée en rouge dans la Figure 4.34. Une ligne rouge en pointillés permet de comparer les solutions selon le MSE.

On remarque qu'il y a des solutions ayant des MSEs plus grands que celui de la solution conventionnelle (n° 2, 3, 4, 10, 11, 12, 13) mais aussi beaucoup de solutions ayant un MSE meilleur (plus petit) que celui de la solution conventionnelle.

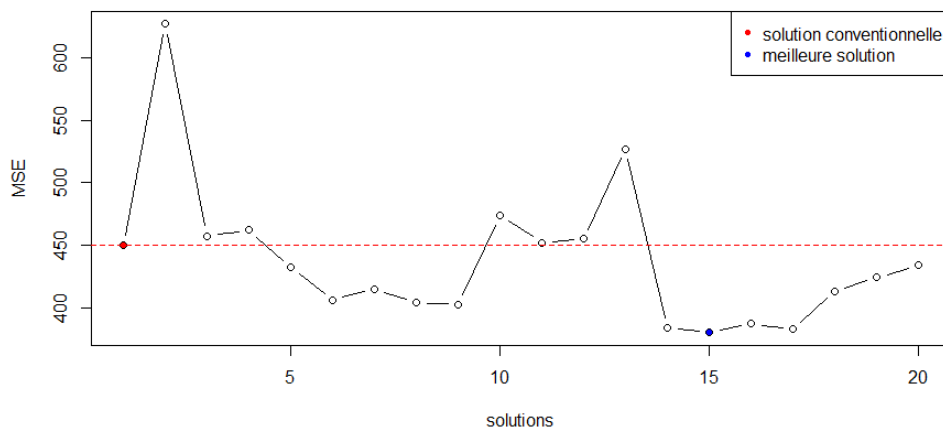


FIGURE 4.34 – Comparaison des solutions selon le MSE

**Exactitude d'estimation des RULs** Une analyse en composantes principales est appliquée dans l'espace des turboréacteurs. Elle montre que les différentes solutions sont assez homogènes par rapport aux vraies RULs, considérées comme une cible. En effet, dans l'espace des 100 turboréacteurs (décrits par leur RULs) (voir la Figure 4.35), les solutions théoriques calculées par l'approche proposée ( $X_1$  à  $X_{19}$ ) ainsi que la solution conventionnelle ( $X_0$ ) sont caractérisées par un certain biais par rapport à la cible  $X_{20}$  (vraies RULs) surtout sur l'axe 1 de l'ACP. Ceci représente l'exactitude (accuracy en anglais) de notre approche d'estimation. Le biais que l'on constate sur toutes les solutions peut être expliquée par la classification et la formule de prédiction de RUL, qui entraîne un intervalle d'incertitude sur toutes les prédictions.

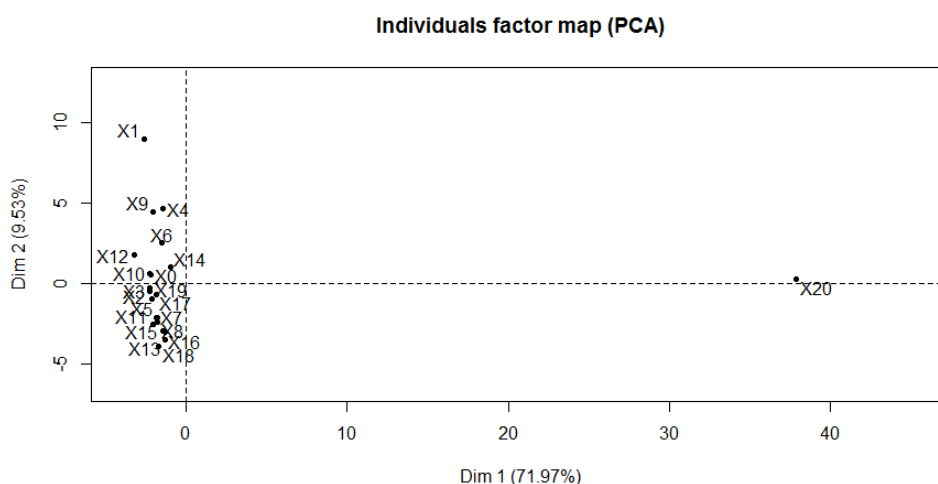


FIGURE 4.35 – Projection des solutions dans l'espace principal de l'ACP (axe 1 et axe 2)

Puisque l'axe 1 de l'ACP explique 71.97% de la variation totale enregistrée dans les données, on peut considérer que cet axe permet d'expliquer en grande partie en quoi se différencient nos solutions de la cible.

**Différenciation des solutions** Pour mieux différencier les solutions entre-elles, il est préférable de regarder leurs projections dans l'espace formé par les axes 2 (9.53% de



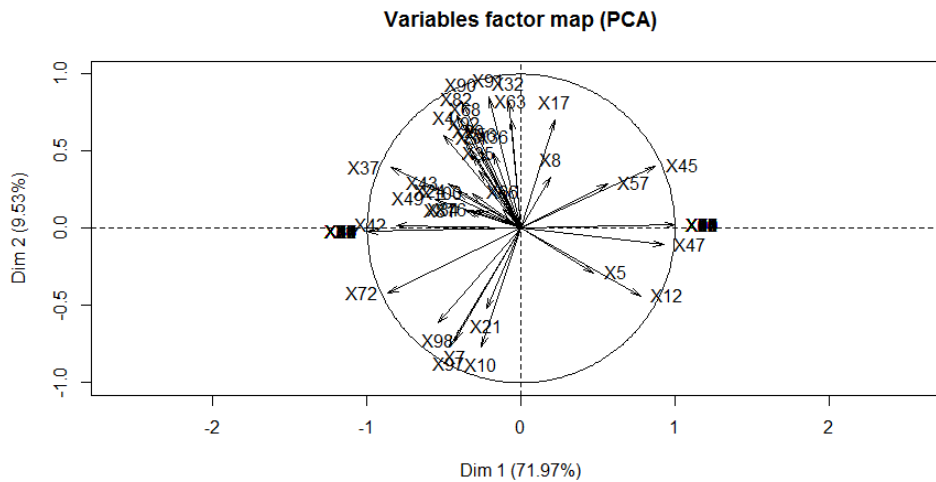


FIGURE 4.36 – Projection des RULs des 100 turboréacteurs dans l'espace principal de l'ACP (axe 1 et axe 2)

variation expliquée) et 3 (7.57% de variation expliquée) (voir les Figures 4.37 et 4.38).

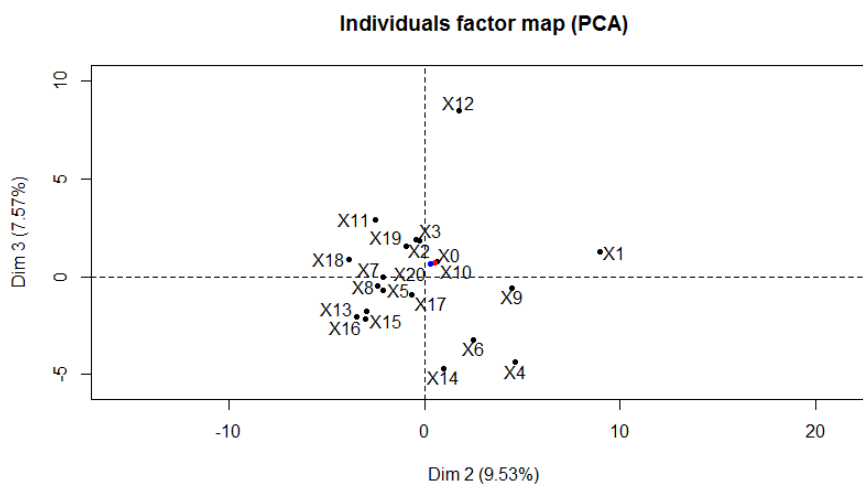


FIGURE 4.37 – Projection des solutions dans l'espace ACP (axe 2 et axe 3)

On remarque que la solution conventionnelle (en rouge dans la Figure 4.37) et la cible (en bleu dans la Figure 4.37) sont proches de l'origine du système de coordonnées et que les différentes solutions sont éparpillées autour d'elles. Les solutions qui se situent à droite dans la Figure 4.37 surestiment les RULs des turboréacteurs 17, 32, 63, 82, 90 et 91 et sous-estiment celles des turboréacteurs 7, 10, 97. C'est exactement l'inverse pour les solutions qui se situent à gauche dans la Figure 4.37. Les solutions qui se situent en haut dans la Figure 4.37 surestiment les RULs des turboréacteurs 24, 31, 34, 35, 36, 49, 56, 66, 76, 81 et 100. Réciproquement, celles qui se situent en bas dans la Figure 4.37 sous-estiment les RULs de ces mêmes turboréacteurs.

Enfin, une classification hiérarchique ascendante (CHA), utilisant la distance Euclidienne et le critère d'agrégation de Ward, conduit à un groupement des solutions comme indiqué dans la Figure 4.39.

Dans ce graphique, la solution conventionnelle est notée  $X_0$  et la cible est notée  $X_{20}$ . Par conséquent, la meilleure solution en termes de MSE (solution n°15 dans la Figure 4.34)

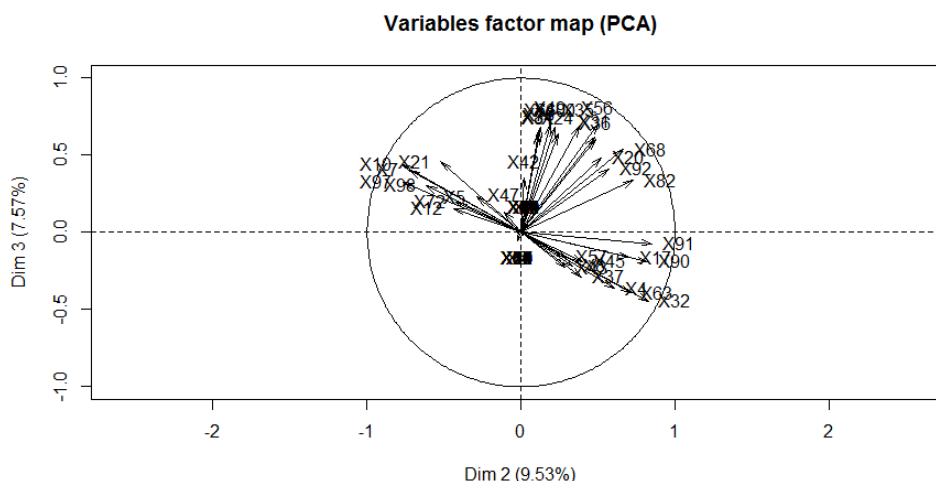


FIGURE 4.38 – Projection des RULs des 100 turboréacteurs dans l'espace ACP (axe 2 et axe 3)

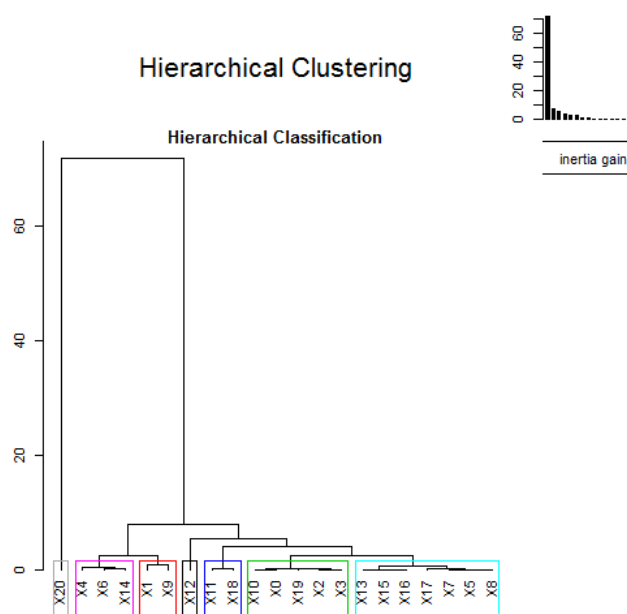


FIGURE 4.39 – CHA des différentes solutions (dendrogramme et indices d'agrégation)

est ici la solution notée  $X_{14}$ . On remarque, encore une fois qu'il n'y a pas de différences très marquées entre les différentes solutions car l'augmentation de l'indice d'agrégation est très lente au fil des regroupements. Pour établir le nombre de classes caractérisant le jeu de données, on devrait couper l'arbre autour de l'indice d'agrégation de 30 (voir Figure 4.39) car c'est l'endroit où la dernière décroissance importante s'est produite. Cela nous ramènerait dans un cas où uniquement deux classes sont présentes (dans une classe la cible et dans l'autre toutes les solutions).

Nous avons fait le choix de couper l'arbre bien plus bas (indice d'agrégation de 2), afin de déterminer quelles sont les solutions qui se ressemblent le plus selon leurs RULs estimées pour les 100 turboréacteurs. On peut remarquer que 6 classes de solutions apparaissent. Ainsi, deux autres solutions ( $X_6$  et  $X_4$ ) ont un comportement très similaire, en termes de RULs estimées pour les 100 turboréacteurs, à celui de la solution ayant le meilleur MSE ( $X_{14}$ ). Elles se caractérisent par des valeurs faibles sur l'axe 3 de l'ACP

(voir Figure 4.37) qui veut dire que toutes sous-estiment les RULs des turboréacteurs 24, 31, 34, 35, 36, 49, 56, 66, 76, 81 et 100.

**Caractérisation des différentes solutions selon les probabilités conditionnelles** Si l'on analyse les différentes solutions dans l'espace des probabilités conditionnelles (voir les Figures 4.40 et 4.41), on constate que la solution ayant le meilleur MSE (en bleu dans la Figure 4.40) est très proche de l'origine du système de coordonnées, ce qui veut dire qu'elle a un profil moyen selon toutes les probabilités conditionnelles.

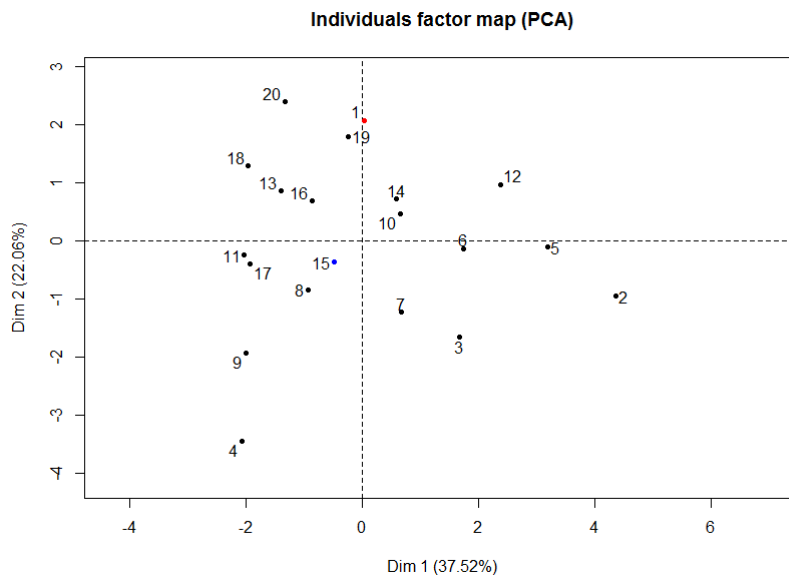


FIGURE 4.40 – Projection des solutions dans l'espace principal de l'ACP

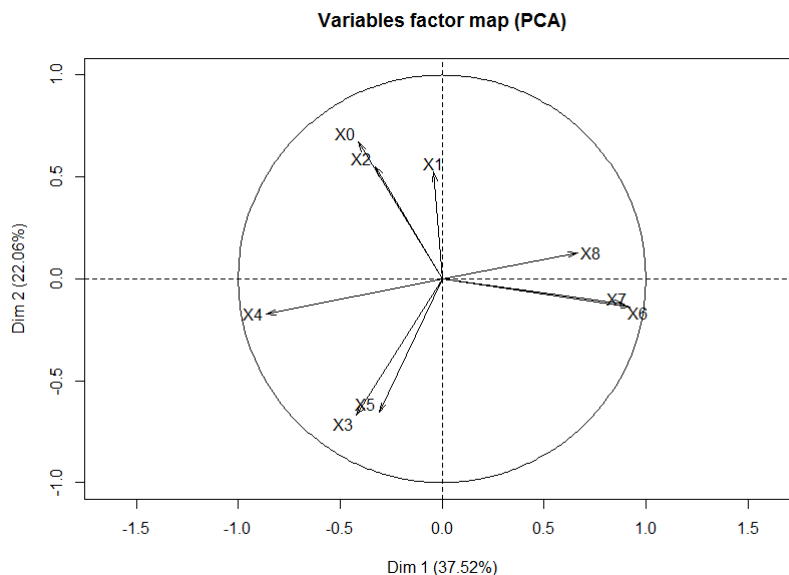


FIGURE 4.41 – Projection des probabilités conditionnelles caractérisant les solutions dans l'espace principal de l'ACP

On constate aussi que la solution conventionnelle présente des valeurs plus grandes (respectivement plus faibles) que la plupart des solutions pour la probabilité conditionnelle

$X_0$  (respectivement  $X_3$  et  $X_5$ ) (voir la Figure 4.41). Globalement, les solutions se trouvant à droite dans la Figure 4.40 ont des probabilités conditionnelles  $X_6$  et  $X_7$  grandes et la probabilité conditionnelle  $X_4$  petite. C'est l'inverse pour ce qui concerne les solutions situées à gauche. Les solutions situées en haut ont une probabilité conditionnelle  $X_0$  grande et des probabilités conditionnelles  $X_3$  et  $X_5$  faibles. C'est l'inverse pour ce qui concerne les solutions situées en bas.

### Résultats

Pour terminer, les 20 combinaisons de paramètres obtenues sont testées avec la configuration optimale déterminée auparavant (4 classes et une taille de filtrage de 5), en utilisant les évidences incertaines, afin de déterminer les RULs des 100 turboréacteurs. Une fois ces RULs déduites avec chaque solution, les métriques d'évaluation de la performance de pronostic sont calculées. La Table 4.19 résume ces performances pour les 20 solutions. La couleur rouge indique le meilleur résultat.

On constate que 4 solutions se différencient des autres, en présentant de très bonnes performances sur la majorité des métriques : ce sont les solutions 14, 15, 16 et 17. En effet, elles présentent toutes la même précision  $R$  de 0.78 et des performances très proches sur les autres métriques. Par ailleurs, la solution 17 se distingue par la meilleure précision, le meilleur score  $Sc = 691.4$ , la meilleure erreur quadratique moyenne  $MAE = 15.24$ , la meilleure erreur quadratique moyenne en pourcentage  $MAPE = 31.22$  et enfin le plus grand nombre de RULs à temps, soit 54. Nous choisissons donc cette solution 17 pour la comparaison avec les autres travaux publiés dans la littérature. Les paramètres lui correspondant sont donnés dans la Table 4.20.

### 4.7.3 Comparaison avec la littérature

Une étude comparative incluant l'approche de fusion proposée et les résultats rapportés dans la littérature est détaillée dans la Table 4.21. Cette dernière récapitule, pour plusieurs travaux, les performances de pronostic en termes de précision  $R$ , score  $Sc$ , nombre de RULs à l'heure  $On - Time$ , ainsi que les erreurs moyennes de prédiction  $MAE$ ,  $MSE$  et  $MAPE$ .

Un tiret dans la table signifie que le travail concerné n'a pas considéré cette métrique pour l'évaluation de ses performances en pronostic tandis que les couleurs rouges et oranges indiquent respectivement le meilleur et le second résultat.

On constate que la plupart des travaux n'évaluent pas leur performance en pronostic avec différentes métriques, ce qui empêche une directe comparaison avec notre travail. Le critère pris en compte par la majorité des méthodes présentées dans cette comparaison est la précision  $R$ , d'où le classement des travaux selon cette métrique, comme on peut le voir dans la Table 4.21.

Selon ce tableau comparatif, l'approche de fusion proposée présente la meilleure précision des estimations RULs ( $R = 0.78$ ), le second meilleur score ( $S = 691$ ) ainsi que les secondes meilleures erreurs moyennes ( $MAE$ ,  $MSE$ ,  $MAPE$ ).

[Shi et al., 2016] présente le plus grand nombre de RULs estimées à temps (on-Time = 60) mais la précision de ses prédictions est plus faible ( $R = 0.69$  contre  $R = 0.78$ ). En outre, les autres métriques ne sont pas précisées.

Enfin, le travail développé dans [Ramasso, 2014] assure le score le moins élevé ( $Sc = 216$ ) ainsi que les erreurs moyennes les plus faibles, mais présente des estimations moins

		Index de la solution									
		1	2	3	4	5	6	7	8	9	10
<b>R</b>		0.74	0.64	0.74	0.73	0.75	0.76	0.76	0.76	0.77	0.73
<b>Sc</b>		861.33	2957.41	904.78	925.59	858.18	740.48	784.01	736.54	734.54	983.27
<b>MSE</b>		449.50	626.86	457.42	461.86	432.22	406.06	414.82	403.90	402.34	473.74
<b>MAE</b>		17.38	19.82	17.56	17.56	16.54	16.18	16.38	16.24	16.06	17.80
<b>MAPE</b>		44.02	62.77	47.00	46.07	39.20	36.05	38.34	36.69	35.89	48.70
<b>On-Time</b>		45	43	44	45	49	51	50	52	53	45
		<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
<b>R</b>		0.74	0.74	0.70	<b>0.78</b>	<b>0.78</b>	<b>0.78</b>	<b>0.78</b>	0.76	0.75	0.75
<b>Sc</b>		866.31	932.05	1035.06	692.73	697.95	701.10	<b>691.40</b>	753.33	877.19	790.07
<b>MSE</b>		452.02	454.90	527.02	383.50	<b>379.90</b>	386.98	382.78	413.14	424.18	433.90
<b>MAE</b>		17.44	17.46	19.36	15.34	<b>15.24</b>	15.36	<b>15.24</b>	16.48	16.44	17.14
<b>MAPE</b>		44.69	44.74	66.14	31.41	32.50	32.10	<b>31.22</b>	37.38	37.30	43.01
<b>On-Time</b>		45	43	38	<b>54</b>	53	<b>54</b>	<b>54</b>	50	50	46

TABLE 4.19 – Comparaison de performances avec les évidences incertaines, pour les différentes solutions

précises que l'approche de fusion (0.67 contre 0.78).

Par conséquent, on peut conclure que l'approche de fusion proposée présente de très bonnes performances en pronostic et surpasse la majorité des travaux publiés dans la littérature.

$D \backslash D_1$	$F_0$	$F_1$	$F_2$
$F_0$	0.81	0.01	0.18
$F_1$	0.01	0.81	0.18
$F_2$	0.21	0.41	0.38

(a) Table du noeud  $D_1$ 

$D \backslash D_2$	$F_0$	$F_1$	$F_2$
$F_0$	0.81	0.18	0.01
$F_1$	0.21	0.38	0.41
$F_2$	0.01	0.18	0.81

(b) Table du noeud  $D_2$ 

TABLE 4.20 – Tables de probabilités conditionnelles pour le pronostic

	<b>R</b>	<b>Sc</b>	<b>On-Time</b>	<b>MAE</b>	<b>MSE</b>	<b>MAPE</b>
[Li et al., 2013]	-	-	-	-	-	43
[Li et al., 2013]	-	-	-	-	-	38
[Peng et al., 2012]	-	-	-	-	3969	-
[Shi et al., 2016]	0.42	—	37	-	-	-
[Ramasso et al., 2013b]	0.53	-	53	-	-	-
[Javed, 2014]	0.55	4463	32	-	-	-
[Shi et al., 2016]	0.58	-	45	-	-	-
[Javed, 2014]	0.61	1046	48	-	-	-
[Ramasso, 2014]	0.67	<b>216</b>	-	<b>10</b>	<b>176</b>	<b>20</b>
[Shi et al., 2016]	<b>0.69</b>	-	<b>60</b>	-	-	-
<b>Fusion</b>	<b>0.78</b>	<b>691</b>	<b>54</b>	<b>15</b>	<b>383</b>	<b>31</b>

TABLE 4.21 – Comparaison de performances en pronostic

## 4.8 Conclusion

Dans le domaine du pronostic des systèmes, rares sont les travaux qui ont fusionné les décisions de plusieurs méthodes. De plus, la plupart de ces travaux sont dédiés à des applications particulières tout comme pour le diagnostic.

Ce chapitre a permis de présenter nos principales contributions théoriques apportées au domaine du pronostic des systèmes. Tout d'abord, une nouvelle démarche de pronostic basée sur les données a été présentée. Ensuite, une méthodologie théorique et générique, similaire à celle développée pour le diagnostic, a été proposée pour permettre la fusion de décisions de deux méthodes de pronostic. Cette méthodologie s'adapte au pronostic en définissant un nouvel ensemble d'objectifs à respecter. Le cadre générique de fusion est un RB dont la déduction de paramètres a conduit à un problème multi-objectif, résolu avec la même approche lexicographique.

Une fois paramétré, le RB a été testé pour le pronostic d'un ensemble de 100 turbo-réacteurs à double flux. L'application a montré la faisabilité de la méthodologie proposée en pronostic et sa capacité à estimer l'état de santé et à prédire la RUL de manière précise. Selon les résultats obtenus, la fusion a surpassé les décisions individuelles ainsi que la majorité des travaux publiés dans la littérature. Par conséquent, notre stratégie de fusion a offert une amélioration significative des performances de pronostic.



# Conclusions et perspectives

## 5.1 Conclusions

Les travaux présentés dans cette thèse ont traité de la surveillance des systèmes complexes avec une méthodologie basée sur la fusion de décisions. La surveillance permet de déterminer l'état du système à travers la détection, le diagnostic et le pronostic des défauts pouvant survenir lors de son fonctionnement.

Dans le chapitre 2, nous avons passé en revue les différentes approches de surveillance proposées dans la littérature jusqu'à ce jour. La principale conclusion tirée de cet état de l'art était que la fusion de décisions reste largement inexplorée, que ce soit pour le diagnostic ou le pronostic, en raison de plusieurs obstacles qu'elle rencontre : solutions dédiées et manque de généricité, aucune assurance d'amélioration a priori des performances, choix difficile de méthodes à fusionner, choix de la stratégie de fusion, etc. Ainsi, l'objectif principal de cette thèse était de proposer une approche générique d'intégration de diverses méthodes de surveillance, assurant de meilleures performances que celles des méthodes individuelles la composant.

La première contribution de cette thèse a été la proposition d'une nouvelle démarche méthodologique de fusion de décisions, basée sur la théorie Bayésienne. En effet, une méthodologie théorique et générique permettant la fusion de décisions de plusieurs méthodes de détection et de diagnostic a été développée dans le Chapitre 3. Cette méthodologie s'est basée sur les performances a priori des méthodes individuelles pour définir un ensemble d'objectifs à respecter afin d'améliorer les performances globales du système en détection et en diagnostic. Le RB a été utilisé comme cadre générique de fusion et ses paramètres ont été calculés de sorte à atteindre les objectifs choisis. Le développement a conduit à un problème multi-objectif sous contraintes, résolu par une approche lexicographique.

La démarche proposée est générique puisqu'elle ne considère dans son développement aucune contrainte sur les méthodes à intégrer ni sur les applications à considérer. De plus, le cadre n'exclut aucune méthode de diagnostic, qu'elle soit basée sur les données ou sur les modèles. Ainsi, différentes méthodes avec diverses connaissances et informations préalables, jeux de données disponibles, exigences de modélisation, etc., peuvent être



combinées, ce qui conduit à un diagnostic plus efficace et plus fiable. En outre, la méthodologie est très flexible et peut s'adapter aux besoins de l'utilisateur et aux stratégies qu'il souhaite adapter pour le diagnostic du système. Enfin, elle permet d'utiliser, à travers les deux méthodes individuelles, tous les progrès et les avancées réalisés par les différentes communautés de détection-diagnostic.

L'approche proposée a été appliquée à un système complexe de grande taille : le Tennessee Eastman Process (TEP). Tout d'abord, un modèle Bond Graph (BG) représentant le fonctionnement du TEP a été développé à partir des connaissances physiques du système puis utilisé à des fins de détection et de diagnostic. Ce développement représente la deuxième contribution de ce travail. Il faut souligner qu'il n'y a eu que très peu de travaux portant sur la modélisation du TEP. D'ailleurs, le modèle BG développé dans cette thèse pourrait servir de "benchmark" pour ceux qui souhaiteraient se pencher sur la modélisation du TEP à des fins de diagnostic et de pronostic.

Les résultats fournis par le modèle BG ont été fusionnés avec les résultats d'une méthode de classification supervisée à base de données par le biais de l'approche développée, permettant ainsi de fournir une décision globale menant à des performances meilleures que celles des méthodes individuelles.

L'application sur le TEP a montré tout d'abord la faisabilité de l'approche proposée et a souligné sa capacité à améliorer significativement les performances en diagnostic, en fournissant un taux élevé de diagnostic de défauts, un faible taux de fausses alarmes et une stratégie efficace pour la résolution des conflits entre différentes méthodes de diagnostic.

La troisième contribution de cette thèse a concerné le développement d'une approche de pronostic reposant sur le principe de classification supervisée. Dans ce cadre, nous avons proposé une démarche méthodologique de pronostic qui s'est basée sur deux étapes principales : une première étape d'apprentissage hors-ligne et une seconde étape de test en ligne. Le but de la première étape a été d'établir les liens existants entre les signaux disponibles et les états de dégradations correspondants tandis que la seconde étape a visé l'identification de l'état de santé du système et la prédiction de sa RUL à tout instant.

Enfin, l'application de cette démarche sur un ensemble de turboréacteurs a montré que malgré sa simplicité, cette approche peut être très efficace.

La quatrième contribution réside dans l'adaptation pour le pronostic du même cadre générique ayant permis la fusion de décisions de plusieurs méthodes de détection-diagnostic de défauts. Tout d'abord, les objectifs de performance à atteindre afin d'améliorer les performances globales du système en pronostic ont été fixés. Pour cela, la méthodologie théorique s'est basée sur les performances a priori des méthodes individuelles, comme pour le diagnostic. Toutefois, les objectifs sont différents dans le cadre du pronostic. Ensuite, le RB, dont les paramètres ont été déduits théoriquement de sorte à respecter tous les objectifs énoncés, a été proposé. Le développement a conduit à un problème multi-objectif sous contraintes, résolu par la même approche lexicographique.

A l'égard du diagnostic, l'approche proposée n'est pas dédiée à des applications particulières et s'affranchit de toute contrainte spécifique au cas d'étude. Elle permet de combiner, à travers son cadre générique, toute approche de pronostic basée sur les données.

Enfin, l'approche de pronostic proposée et la démarche de fusion sont appliquées sur un ensemble de turboréacteurs à double flux. La première étape a consisté à développer deux méthodes de pronostic à partir des données disponibles du système, selon la méthodologie proposée. La fusion des décisions résultant de ces deux méthodes a fourni le résultat

global correspondant à une estimation de l'état de santé du système et une prédiction de la durée de fonctionnement restante avant la défaillance du système. L'application a montré la faisabilité de la méthodologie proposée en pronostic et sa capacité à estimer l'état de santé et à prédire la RUL de manière précise.

Par conséquent, notre stratégie de fusion de décisions a offert une amélioration significative des performances de surveillance, aussi bien en diagnostic qu'en pronostic.

## 5.2 Perspectives

Tout au long de ce travail, nous avons formulé des hypothèses afin de simplifier les développements proposés. Plusieurs pistes de recherche restent donc à explorer dans nos futurs travaux.

- **Egalité des probabilités :**

Tout d'abord, la méthodologie de fusion de décisions repose sur l'hypothèse d'égalité des probabilités a priori des différentes modalités. Cette hypothèse a été adoptée afin de simplifier les calculs des objectifs d'une part, et leur résolution d'autre part. Toutefois, il serait intéressant d'inclure, par exemple, des informations supplémentaires sur ces probabilités a priori et ainsi les considérer comme différentes s'il le faut. Dans ce cas, l'expression des objectifs change mais pas leur nombre.

Les égalités concernant les probabilités conditionnelles peuvent également être remises en cause. En effet, on peut décider de privilégier une méthode par rapport aux autres. Ceci implique que le nombre de probabilités à déduire est plus grand. Dans ce cas, une méthode de résolution, autre que l'approche lexicographique, pourrait être envisagée afin d'éviter l'explosion combinatoire.

- **Modèle BG :**

Le modèle BG développé pour le TEP a pour objectif de fournir une représentation simplifiée du fonctionnement normal du système. Dans nos futurs travaux, nous pourrions envisager de le compléter en prenant en compte les incertitudes qui peuvent l'affecter et d'autres phénomènes physiques inhérents au fonctionnement du système.

- **Pronostic local :**

La démarche de pronostic proposée considère que la RUL du système correspond à la RUL de ses composants critiques. Cette démarche de pronostic local est réalisée au niveau de chaque composant. Cependant, cette hypothèse n'est pas toujours vérifiée en pratique et il convient donc de proposer une démarche de pronostic global, qui prend en compte le système dans son ensemble. L'approche globale impliquerait la prise en compte des liens et des interactions existants entre les différents composants.

- **Extension à  $N$  méthodes :**

Au lieu de combiner un grand nombre de méthodes homogènes, notre approche tente de combiner deux méthodes hétérogènes qui se complètent. D'un point de vue pratique, il est peu probable de développer plusieurs méthodes pour la surveillance d'un même système puisqu'elles reposent sur différentes hypothèses et exigences. Ceci explique en partie notre choix de combiner uniquement deux méthodes de surveillance.

Cependant l'approche proposée possède un potentiel de généralisation à plusieurs

méthodes qu'il convient de mentionner et d'exploiter si nécessaire. En effet, d'un point de vue pratique, rien n'empêche l'extension de la méthodologie et du cadre à un ensemble de  $N$  méthodes.

Dans ce contexte, comme on dispose de  $N$  décisions et que chaque décision a trois possibilités ( $CFN, F_0, F_i, i \in \{1, \dots, N\}$ ), l'ensemble des objectifs à développer est de taille  $3^N$ . Par conséquent, la complexité combinatoire relative au nombre d'objectifs à développer représente la principale limite de l'extension à  $N$  méthodes.

Une proposition, permettant de diminuer le nombre d'objectifs à résoudre, est de ramener le problème à une situation qui correspond au cadre et à la méthodologie déjà développés auparavant, tout en considérant les  $N$  méthodes. Pour cela, on part des  $N$  décisions que l'on fusionne deux à deux en se basant sur la méthodologie proposée de fusion de décisions jusqu'à atteindre le noeud de décision globale.

- **Outil informatique :**

Enfin, il semble important de souligner qu'un outil informatique calculant directement les paramètres du RB à partir de l'évaluation de performances des méthodes individuelles est en cours de développement. Il permet d'automatiser la procédure de définition des objectifs et de calcul des contraintes afin de rendre l'approche accessible à différents utilisateurs, sans connaissances théoriques préalables.

# Liste des tableaux

2.1	Fusion de méthodes : récapitulatif . . . . .	37
2.2	Fusion de décisions : récapitulatif . . . . .	47
3.1	Exemple de matrice de confusion . . . . .	55
3.2	Evaluation des performances individuelles . . . . .	61
3.3	Table des objectifs possibles . . . . .	62
3.4	Table des objectifs choisis . . . . .	64
3.5	Table de probabilités a priori du noeud $D$ . . . . .	66
3.6	Tables de probabilités conditionnelles . . . . .	66
3.7	Table de probabilités a priori du noeud $D$ . . . . .	76
3.8	Tables de probabilités conditionnelles . . . . .	77
3.9	Matrice de confusion de la méthode de fusion . . . . .	77
3.10	Comparaison de performances (%) . . . . .	78
3.11	Tables des objectifs pour des stratégies différentes . . . . .	78
3.12	Tables de probabilités conditionnelles privilégiant le FAR . . . . .	79
3.13	Tables de probabilités conditionnelles privilégiant le FDR . . . . .	79
3.14	Paramètres de la solution conventionnelle et des 10 meilleures solutions en terme de FAR . . . . .	84
3.15	FARs moyens des 10 meilleures solutions et de la solution conventionnelle	84
3.16	Paramètres de la solution conventionnelle et des 10 meilleures solutions en terme de FDR . . . . .	86
3.17	FDRs moyens des 10 meilleures solutions et de la solution conventionnelle	87
3.18	Les défauts pouvant affecter le TEP . . . . .	90
3.19	Les jeux de données utilisés . . . . .	91
3.20	Matrice de Signature de Défauts (FSM) reliant les défauts aux résidus . .	98
3.21	Résultats FDD avec la stratégie BG (en %) . . . . .	99
3.22	Matrice de confusion de la méthode $m_1$ , basée sur le BG et l'AD (en %) .	100
3.23	Matrice de confusion de la méthode 2, basée sur l'AD (en %) . . . . .	100
3.24	Matrice de confusion de l'approche de fusion proposée (en %), pour une stratégie privilégiant le FAR . . . . .	103
3.25	Matrice de confusion de l'approche de fusion proposée (en %), pour une stratégie privilégiant le FDR . . . . .	104
3.26	Comparaison de performances (en %) . . . . .	104
3.27	Comparaison des performances en diagnostic de défauts (%) pour le TEP	105
4.1	Evaluation des performances individuelles . . . . .	121
4.2	Table des objectifs . . . . .	122
4.3	Table de probabilités a priori du noeud $D$ . . . . .	123
4.4	Tables de probabilités conditionnelles . . . . .	124
4.5	Matrice de confusion de la méthode de fusion . . . . .	130
4.6	Comparaison de performances (%) . . . . .	130

4.7	Ensembles de données C-MAPSS . . . . .	133
4.8	Paramètres choisis . . . . .	135
4.9	Evaluation des performances individuelles en classification, pour 4 classes	138
4.10	Performances individuelles en pronostic, pour 4 classes . . . . .	139
4.11	Comparaison de performances en pronostic, pour 4 classes . . . . .	140
4.12	Comparaison de performances en précision R pour la fusion . . . . .	143
4.13	Comparaison de performances en score $S_c$ pour la fusion . . . . .	144
4.14	Comparaison de performances en MAE pour la fusion . . . . .	145
4.15	Comparaison de performances en MAPE pour la fusion . . . . .	145
4.16	Comparaison de performances en MSE pour la fusion . . . . .	146
4.17	Comparaison de performances en nombre de RULs à temps pour la fusion	147
4.18	Tables de probabilités conditionnelles pour le pronostic . . . . .	148
4.19	Comparaison de performances avec les évidences incertaines, pour les différentes solutions . . . . .	154
4.20	Tables de probabilités conditionnelles pour le pronostic . . . . .	155
4.21	Comparaison de performances en pronostic . . . . .	155
1	Variables de mesure en continu . . . . .	173
2	Variables de mesure échantillonnées . . . . .	174
3	Variables de contrôle . . . . .	174

# Table des figures

2.1	Principe de la détection et du diagnostic à base de modèle. . . . .	14
2.2	Bond Graph. . . . .	16
2.3	Principe de la détection et du diagnostic par données. . . . .	18
2.4	Exemple de Réseau Bayésien. . . . .	19
2.5	RB classifieur Naïf. . . . .	20
2.6	Principe du pronostic à base de modèle. . . . .	24
2.7	Principe du pronostic à base de données. . . . .	26
2.8	Fusion d'approches pour la surveillance. . . . .	30
2.9	Le cadre proposé combinant SDG et QTA pour le diagnostic des défauts. . . . .	31
2.10	Cadre pour l'identification de modèle, la génération de résidus, la détection de défaut et le diagnostic. . . . .	32
2.11	Combinaison du RN et du FK. . . . .	34
2.12	RB proposé pour intégrer la fiabilité et les résidus générés par BG. . . . .	35
2.13	Les principaux éléments du Dkit . . . . .	41
2.14	ACP, SDG et QTA pour la détection et le diagnostic des défaillances. . . . .	43
2.15	Cadre utilisant des mesures et des résidus pour la détection et le diagnostic de défauts. . . . .	43
3.1	La méthodologie proposée . . . . .	57
3.2	Evaluation des performances individuelles . . . . .	58
3.3	Les données de test pour l'exemple numérique . . . . .	60
3.4	Structure du RB . . . . .	65
3.5	Le cadre générique proposé . . . . .	77
3.6	a) Evidences certaines VS b) incertaines au niveau des noeuds $D_1$ et $D_2$ . . . . .	80
3.7	P-values du test de Wilcoxon apparié pour les FARs (la ligne continue en rouge représente le seuil de significativité de 5%) . . . . .	81
3.8	Comparaison de FAR : solution conventionnelle VS solution théorique 4 . . . . .	81
3.9	Projection des solutions dans le plan principal . . . . .	82
3.10	Cercle de corrélation dans le plan principal . . . . .	82
3.11	Plan principal de projection de l'ACP . . . . .	83
3.12	Cercle de corrélation . . . . .	83
3.13	Test de Wilcoxon apparié . . . . .	84
3.14	Comparaison de FDR . . . . .	84
3.15	Projection des solutions dans le plan principal . . . . .	85
3.16	Cercle de corrélation . . . . .	85
3.17	Plan principal de projection . . . . .	86
3.18	Cercle de corrélation . . . . .	86
3.19	Schéma du Tennessee Eastman Process . . . . .	89
3.20	Stratégie de détection et de diagnostic de défauts dans le TEP . . . . .	91
3.21	BG à mots du TEP . . . . .	92

3.22	Modèle BG du réactif A du TEP, en causalité dérivée . . . . .	93
3.23	Évaluation des résidus $r_1$ et $r_2$ relatifs au réactif A, dans les conditions de fonctionnement normal . . . . .	97
3.24	Stratégie de FDD basée sur le BG en ligne . . . . .	99
3.25	Matrices de confusion des deux méthodes FDD . . . . .	101
3.26	Cadre générique de fusion . . . . .	101
3.27	Fusion de décisions en ligne . . . . .	102
3.28	FARs obtenus pour chaque solution . . . . .	102
3.29	FDRs obtenus pour chaque solution . . . . .	103
4.1	Evolution de la durée de vie d'un système . . . . .	110
4.2	La méthodologie de pronostic proposée . . . . .	112
4.3	Signaux d'apprentissage . . . . .	113
4.4	Evolution de la même mesure pour les 2 systèmes considérés . . . . .	113
4.5	Données d'apprentissage, avec l' <i>EOL</i> comme point final. . . . .	114
4.6	Evolution de la même mesure pour les 2 systèmes considérés, avec l' <i>EOL</i> comme point de départ. . . . .	115
4.7	Découpage en états de santé . . . . .	115
4.8	Découpage en états de santé pour les deux signaux . . . . .	116
4.9	Données de test . . . . .	116
4.10	Intervalle de prédiction . . . . .	119
4.11	Méthodologie proposée de fusion pour le pronostic . . . . .	119
4.12	Exemple numérique : Evolution de la dégradation . . . . .	120
4.13	Structure du RB pour le pronostic . . . . .	123
4.14	Exemple numérique : évolution de la dégradation du capteur test . . . . .	130
4.15	Schéma simplifié du turboréacteur simulé dans C-MAPSS [Frederick et al., 2007] . . . . .	133
4.16	Mesure issue d'un capteur . . . . .	134
4.17	Données d'apprentissage issues de 100 turboréacteurs . . . . .	135
4.18	Evolution de la température totale à la sortie du compresseur basse pression pour les 100 turboréacteurs considérés . . . . .	136
4.19	EOL des 100 turboréacteurs. . . . .	136
4.20	Evolution de la même mesure pour les 2 turboréacteurs considérés . . . . .	137
4.21	Données d'apprentissage issues de 100 turboréacteurs, avec le recalage proposé . . . . .	137
4.22	Données de test issues de 100 turboréacteurs . . . . .	138
4.23	Démarche des méthodes individuelles . . . . .	139
4.24	Matrices de confusion des deux méthodes de pronostic, pour 4 classes . . . . .	140
4.25	Démarche de fusion . . . . .	141
4.26	Nombre de classes . . . . .	142
4.27	Longueur de chaque classe . . . . .	142
4.28	Comparaison de la précision de prédiction des RULs . . . . .	144
4.29	Comparaison du score des RULs . . . . .	144
4.30	Comparaison de l'erreur absolue moyenne des prédictions de RULs . . . . .	145
4.31	Comparaison de l'erreur absolue moyenne en pourcentage des prédictions de RULs . . . . .	146
4.32	Comparaison de l'erreur quadratique moyenne des prédictions de RULs . . . . .	146
4.33	Comparaison du nombre de RULs à temps . . . . .	147
4.34	Comparaison des solutions selon le MSE . . . . .	149
4.35	Projection des solutions dans l'espace principal de l'ACP (axe 1 et axe 2) . . . . .	149

4.36	Projection des RULs des 100 turboréacteurs dans l'espace principal de l'ACP (axe 1 et axe 2) . . . . .	150
4.37	Projection des solutions dans l'espace ACP (axe 2 et axe 3) . . . . .	150
4.38	Projection des RULs des 100 turboréacteurs dans l'espace ACP (axe 2 et axe 3) . . . . .	151
4.39	CHA des différentes solutions (dendrogramme et indices d'agrégation) . .	151
4.40	Projection des solutions dans l'espace principal de l'ACP . . . . .	152
4.41	Projection des probabilités conditionnelles caractérisant les solutions dans l'espace principal de l'ACP . . . . .	152
1	Modèle BG du réactif $B$ dans le TEP, avec causalité dérivée . . . . .	175
2	Modèle BG du réactif $C$ dans le TEP, avec causalité dérivée . . . . .	175
3	Modèle BG du réactif $D$ dans le TEP, avec causalité dérivée . . . . .	176
4	Modèle BG du réactif $E$ dans le TEP, avec causalité dérivée . . . . .	176
5	Modèle BG du produit $G$ dans le TEP, avec causalité dérivée . . . . .	177
6	Modèle BG du produit $H$ dans le TEP, avec causalité dérivée . . . . .	177
7	Modèle BG du produit dérivé $F$ dans le TEP, avec causalité dérivée . . .	177
8	Modèle BG hydraulique . . . . .	178
9	Modèle BG thermodynamique . . . . .	179
10	Évaluation des résidus $r_3 - r_{18}$ liés aux gaz $B, C, D, E, F, G, H$ , en $CFN$	181





# Annexes

## Annexe A : Développement des objectifs pour le diagnostic

### Objectif 2

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\left\{ \begin{array}{l} P(D = CFN/D_1 = CFN \cap D_2 = F_0) > P(D = CFN) \\ P(D = CFN/D_1 = CFN \cap D_2 = F_0) > P(D = F_i/D_1 = CFN \cap D_2 = F_0), \\ i \in \{0, 1, 2\} \end{array} \right. \quad (1)$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned} P(D = CFN/D_1 = CFN \cap D_2 = F_0) > P(D = CFN) \\ \Leftrightarrow \frac{P(D_1=CFN/D=CFN)P(D_2=F_0/D=CFN)}{P(D_1=CFN/D_2=F_0)P(D_2=F_0)} > 1 \\ \Leftrightarrow \frac{a_0 a_1}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_1+b_1+c_1+d_1)} > 1 \end{aligned} \quad (2)$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$  :

$$\begin{aligned} P(D = CFN/D_1 = CFN \cap D_2 = F_0) > P(D = F_0/D_1 = CFN \cap D_2 = F_0) \\ \Leftrightarrow \frac{P(D_1=CFN/D=CFN)P(D_2=F_0/D=CFN)}{P(D_1=CFN/D_2=F_0)P(D_2=F_0)} > \frac{P(D_1=CFN/D=F_0)P(D_2=F_0/D=F_0)}{P(D_1=CFN/D_2=F_0)P(D_2=F_0)} \\ \Leftrightarrow \frac{a_0 a_1}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_1+b_1+c_1+d_1)} > \frac{b_0 b_1}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_1+b_1+c_1+d_1)} \\ \Leftrightarrow a_0 a_1 > b_0 b_1 \end{aligned} \quad (3)$$

Pour  $D = F_1$ , on obtient :

$$\begin{aligned}
P(D = CFN/D_1 = CFN \cap D_2 = F_0) &> P(D = F_1/D_1 = CFN \cap D_2 = F_0) \\
\Leftrightarrow \frac{P(D_1=CFN/D=CFN)P(D_2=F_0/D=CFN)}{P(D_1=CFN/D_2=F_0)P(D_2=F_0)} &> \frac{P(D_1=CFN/D=F_1)P(D_2=F_0/D=F_1)}{P(D_1=CFN/D_2=F_0)P(D_2=F_0)} \\
\Leftrightarrow \frac{a_0 a_1}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_1+b_1+c_1+d_1)} &> \frac{c_0 d_1}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_1+b_1+c_1+d_1)} \\
&\Leftrightarrow a_0 a_1 > c_0 d_1
\end{aligned} \tag{4}$$

Pour  $D = F_2$ , on obtient la même contrainte :  $a_0 a_1 > c_0 d_1$ .

**Résumé :** Par conséquent, 3 contraintes sont obtenues :

$$\frac{a_0 a_1}{(a_0 + b_0 + c_0 + d_0)(a_1 + b_1 + c_1 + d_1)} > \alpha_0^2 \tag{5}$$

$$\begin{cases} a_0 a_1 > b_0 b_1 \\ a_0 a_1 > c_0 d_1 \end{cases} \tag{6}$$

### Objectif 3

Afin d'assurer cet objectif, les critères suivants (Eq. 7) doivent être respectés :

$$\begin{cases} P(D = F_2/D_1 = CFN \cap D_2 = F_2) > P(D = F_2) \\ P(D = F_2/D_1 = CFN \cap D_2 = F_2) > P(D = (F_i \cup CFN)/D_1 = CFN \cap D_2 = F_2), \\ i \in \{0, 1\} \end{cases} \tag{7}$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned}
P(D = F_2/D_1 = CFN \cap D_2 = F_2) &> P(D = F_2) \\
\Leftrightarrow \frac{P(D_1=CFN/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} &> 1 \\
\Leftrightarrow \frac{d_0 c_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} &> 1
\end{aligned} \tag{8}$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$  :

$$\begin{aligned}
P(D = F_2/D_1 = CFN \cap D_2 = F_2) &> P(D = F_0/D_1 = CFN \cap D_2 = F_2) \\
\Leftrightarrow \frac{P(D_1=CFN/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=CFN/D=F_0)P(D_2=F_2/D=F_0)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} \\
\Leftrightarrow \frac{d_0 c_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} &> \frac{b_0 b_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} \\
&\Leftrightarrow d_0 c_2 > b_0 b_2
\end{aligned} \tag{9}$$

Pour  $D = F_1$ , on obtient :

$$\begin{aligned}
P(D = F_2/D_1 = CFN \cap D_2 = F_2) &> P(D = F_1/D_1 = CFN \cap D_2 = F_2) \\
\Leftrightarrow \frac{P(D_1=CFN/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=CFN/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} \\
\Leftrightarrow \frac{d_0 c_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} &> \frac{c_0 d_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} \\
&\Leftrightarrow d_0 c_2 > c_0 d_2
\end{aligned} \tag{10}$$

Pour  $D = CFN$ , on obtient :

$$\begin{aligned}
P(D = F_2/D_1 = CFN \cap D_2 = F_2) &> P(D = CFN/D_1 = CFN \cap D_2 = F_2) \\
\Leftrightarrow \frac{P(D_1=CFN/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=CFN/D=CFN)P(D_2=F_2/D=CFN)}{P(D_1=CFN/D_2=F_2)P(D_2=F_2)} \\
\Leftrightarrow \frac{d_0c_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} &> \frac{a_0a_2}{\alpha_0^2(a_0+b_0+c_0+d_0)(a_2+b_2+c_2+d_2)} \\
\Leftrightarrow d_0c_2 &> a_0a_2
\end{aligned} \tag{11}$$

**Résumé :** Par conséquent, 4 contraintes sont obtenues :

$$\frac{d_0c_2}{(a_0 + b_0 + c_0 + d_0)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \tag{12}$$

$$\begin{cases} d_0c_2 > a_0a_2 \\ d_0c_2 > b_0b_2 \\ d_0c_2 > c_0d_2 \end{cases} \tag{13}$$

### Objectif 4

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\begin{cases} P(D = CFN/D_1 = F_0 \cap D_2 = CFN) > P(D = CFN) \\ P(D = CFN/D_1 = F_0 \cap D_2 = CFN) > P(D = F_i/D_1 = F_0 \cap D_2 = CFN), \\ i \in \{0, 1, 2\} \end{cases} \tag{14}$$

Le raisonnement est identique à celui de l'Objectif 2 et les mêmes contraintes sont déduites.

$$\frac{a_0a_1}{(a_0 + b_0 + c_0 + d_0)(a_1 + b_1 + c_1 + d_1)} > \alpha_0^2 \tag{15}$$

$$\begin{cases} a_0a_1 > b_0b_1 \\ a_0a_1 > c_0d_1 \end{cases} \tag{16}$$

### Objectif 5

Afin d'assurer cet objectif, les critères suivants (Eq. 77) doivent être respectés :

$$\begin{cases} P(D = F_0/D_1 = F_0 \cap D_2 = F_0) > P(D = F_0) \\ P(D = F_0/D_1 = F_0 \cap D_2 = F_0) > P(D = (F_i \cup CFN)/D_1 = F_0 \cap D_2 = F_0), \\ i \in \{1; 2\} \end{cases} \tag{17}$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned}
P(D = F_0/D_1 = F_0 \cap D_2 = F_0) &> P(D = F_0) \\
\Leftrightarrow \frac{P(D_1=F_0/D=F_0)P(D_2=F_0/D=F_0)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} &> 1 \\
\Leftrightarrow \frac{b_1^2}{\alpha_0^2(a_1+b_1+c_1+d_1)^2} &> 1 \\
\Leftrightarrow \frac{b_1}{a_1+b_1+c_1+d_1} &> \alpha_0
\end{aligned} \tag{18}$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_1$  :

$$\begin{aligned}
 P(D = F_0/D_1 = F_0 \cap D_2 = F_0) &> P(D = F_1/D_1 = F_0 \cap D_2 = F_0) \\
 \Leftrightarrow \frac{P(D_1=F_0/D=F_0)P(D_2=F_0/D=F_0)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} &> \frac{P(D_1=F_0/D=F_1)P(D_2=F_0/D=F_1)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} \\
 \Leftrightarrow \frac{b_1^2}{\alpha_0^2(a_1+b_1+c_1+d_1)^2} &> \frac{c_1d_1}{\alpha_0^2(a_1+b_1+c_1+d_1)^2} \\
 \Leftrightarrow b_1^2 &> c_1d_1
 \end{aligned} \tag{19}$$

Pour  $D = F_2$ , on obtient la même contrainte :  $b_1^2 > c_1d_1$ .

Pour  $D = CFN$  :

$$\begin{aligned}
 P(D = F_0/D_1 = F_0 \cap D_2 = F_0) &> P(D = CFN/D_1 = F_0 \cap D_2 = F_0) \\
 \Leftrightarrow \frac{P(D_1=F_0/D=F_0)P(D_2=F_0/D=F_0)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} &> \frac{P(D_1=F_0/D=CFN)P(D_2=F_0/D=CFN)}{P(D_1=F_0/D_2=F_0)P(D_2=F_0)} \\
 \Leftrightarrow \frac{b_1^2}{\alpha_0^2(a_1+b_1+c_1+d_1)^2} &> \frac{a_1^2}{\alpha_0^2(a_1+b_1+c_1+d_1)^2} \\
 \Leftrightarrow b_1 &> a_1
 \end{aligned} \tag{20}$$

**Résumé :** Par conséquent, 3 contraintes sont obtenues :

$$\frac{b_1}{a_1 + b_1 + c_1 + d_1} > \alpha_0 \tag{21}$$

$$\begin{cases} b_1 > a_1 \\ b_1^2 > c_1d_1 \end{cases} \tag{22}$$

## Objectif 6

Afin d'assurer cet objectif, les critères suivants (Eq. 23) doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_2/D_1 = F_0 \cap D_2 = F_2) > P(D = F_2) \\ P(D = F_2/D_1 = F_0 \cap D_2 = F_2) > P(D = (F_i \cup CFN)/D_1 = F_0 \cap D_2 = F_2), \\ \quad i \in \{0; 1\} \end{array} \right. \tag{23}$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned}
 P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = F_2) \\
 \Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> 1 \\
 \Leftrightarrow \frac{d_1c_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} &> 1
 \end{aligned} \tag{24}$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$  :

$$\begin{aligned}
P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = F_0/D_1 = F_0 \cap D_2 = F_2) \\
\Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_0/D=F_0)P(D_2=F_2/D=F_0)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} \\
\Leftrightarrow \frac{d_1c_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} &> \frac{b_1b_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} \\
&\Leftrightarrow d_1c_2 > b_1b_2
\end{aligned} \tag{25}$$

Pour  $D = F_1$  :

$$\begin{aligned}
P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = F_1/D_1 = F_0 \cap D_2 = F_2) \\
\Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_0/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} \\
\Leftrightarrow \frac{d_1c_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} &> \frac{c_1d_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} \\
&\Leftrightarrow d_1c_2 > c_1d_2
\end{aligned} \tag{26}$$

Pour  $D = CFN$  :

$$\begin{aligned}
P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = CFN/D_1 = F_0 \cap D_2 = F_2) \\
\Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_0/D=CFN)P(D_2=F_2/D=CFN)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} \\
\Leftrightarrow \frac{d_1c_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} &> \frac{a_1a_2}{\alpha_0^2(a_1+b_1+c_1+d_1)(a_2+b_2+c_2+d_2)} \\
&\Leftrightarrow d_1c_2 > a_1a_2
\end{aligned} \tag{27}$$

**Résumé** Par conséquent, 4 contraintes sont déduites :

$$\frac{d_1c_2}{(a_1 + b_1 + c_1 + d_1)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \tag{28}$$

$$\begin{cases} d_1c_2 > a_1a_2 \\ d_1c_2 > b_1b_2 \\ d_1c_2 > c_1d_2 \end{cases} \tag{29}$$

## Objectif 7

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\begin{cases} P(D = F_1/D_1 = F_1 \cap D_2 = CFN) > P(D = F_1) \\ P(D = F_i/D_1 = F_1 \cap D_2 = CFN) > P(D = (F_i \cup CFN)/D_1 = F_1 \cap D_2 = CFN), \\ i \in \{0, 2\} \end{cases} \tag{30}$$

Le raisonnement est identique à celui de l'Objectif 3 et les mêmes contraintes sont déduites.

$$\frac{d_0c_2}{(a_0 + b_0 + c_0 + d_0)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \tag{31}$$

$$\begin{cases} d_0c_2 > a_0a_2 \\ d_0c_2 > b_0b_2 \\ d_0c_2 > c_0d_2 \end{cases} \tag{32}$$

### Objectif 8

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_1/D_1 = F_1 \cap D_2 = F_0) > P(D = F_1) \\ P(D = F_1/D_1 = F_1 \cap D_2 = F_0) > P(D = (F_i \cup CFN)/D_1 = F_1 \cap D_2 = F_0), \\ i \in \{0; 2\} \end{array} \right. \quad (33)$$

Le raisonnement est identique à celui de l'Objectif 6 et les mêmes contraintes sont déduites.

$$\frac{d_1 c_2}{(a_1 + b_1 + c_1 + d_1)(a_2 + b_2 + c_2 + d_2)} > \alpha_0^2 \quad (34)$$

$$\left\{ \begin{array}{l} d_1 c_2 > a_1 a_2 \\ d_1 c_2 > b_1 b_2 \\ d_1 c_2 > c_1 d_2 \end{array} \right. \quad (35)$$

### Objectif 9

Afin d'assurer cet objectif, les critères suivants (Eq. 36) doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_1/D_1 = F_1 \cap D_2 = F_2) > P(D = F_1) \\ P(D = F_2/D_1 = F_1 \cap D_2 = F_2) > P(D = F_2) \\ P(D = F_1/D_1 = F_1 \cap D_2 = F_2) > P(D = (F_0 \cup CFN)/D_1 = F_1 \cap D_2 = F_2) \\ P(D = F_2/D_1 = F_1 \cap D_2 = F_2) > P(D = (F_0 \cup CFN)/D_1 = F_1 \cap D_2 = F_2) \end{array} \right. \quad (36)$$

**Critères 1 et 2 :** La première inégalité est développée comme il suit :

$$\begin{aligned} P(D = F_1/D_1 = F_1 \cap D_2 = F_2) &> P(D = F_1) \\ \Leftrightarrow \frac{P(D_1=F_1/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=F_1/D_2=F_2)P(D_2=F_2)} &> 1 \\ \Leftrightarrow \frac{c_2 d_2}{\alpha_0^2 (a_2 + b_2 + c_2 + d_2)^2} &> 1 \end{aligned} \quad (37)$$

La seconde inégalité conduit à la même contrainte.

**Critères 3 et 4 :** Pour  $D = F_0$ , la troisième inégalité est développée comme il suit :

$$\begin{aligned} P(D = F_1/D_1 = F_1 \cap D_2 = F_2) &> P(D = F_0/D_1 = F_1 \cap D_2 = F_2) \\ \Leftrightarrow \frac{P(D_1=F_1/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=F_1/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_1/D=F_0)P(D_2=F_2/D=F_0)}{P(D_1=F_1/D_2=F_2)P(D_2=F_2)} \\ \Leftrightarrow \frac{c_2 d_2}{\alpha_0^2 (a_2 + b_2 + c_2 + d_2)^2} &> \frac{b_2^2}{\alpha_0^2 (a_2 + b_2 + c_2 + d_2)^2} \\ \Leftrightarrow c_2 d_2 &> b_2^2 \end{aligned} \quad (38)$$

Pour  $D = CFN$ , la troisième inégalité est développée comme il suit :

$$\begin{aligned} P(D = F_1/D_1 = F_1 \cap D_2 = F_2) &> P(D = CFN/D_1 = F_1 \cap D_2 = F_2) \\ \Leftrightarrow \frac{P(D_1=F_1/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=F_1/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_1/D=CFN)P(D_2=F_2/D=CFN)}{P(D_1=F_1/D_2=F_2)P(D_2=F_2)} \\ \Leftrightarrow \frac{c_2 d_2}{\alpha_0^2 (a_2 + b_2 + c_2 + d_2)^2} &> \frac{a_2^2}{\alpha_0^2 (a_2 + b_2 + c_2 + d_2)^2} \\ \Leftrightarrow c_2 d_2 &> a_2^2 \end{aligned} \quad (39)$$

La quatrième inégalité conduit à la même contrainte.

**Résumé** Par conséquent, 3 différentes contraintes sont déduites :

$$\frac{c_2 d_2}{(a_2 + b_2 + c_2 + d_2)^2} > \alpha_0^2 \quad (40)$$

$$\begin{cases} c_2 d_2 > a_2^2 \\ c_2 d_2 > b_2^2 \end{cases} \quad (41)$$

## Annexe B : Description du TEP

Les variables mesurées en continu dans le TEP sont données dans la Table 1 tandis que les variables échantillonnées sont dans la Table 2. Les variables de contrôle sont décrites dans la Table 3.

Description	Variable	Unité
Débit d'alimentation en <i>A</i>	$X_1$	<i>kscmh</i>
Débit d'alimentation en <i>D</i>	$X_2$	<i>kg/h</i>
Débit d'alimentation en <i>E</i>	$X_3$	<i>kg/h</i>
Débit d'alimentation en <i>A</i> et <i>C</i>	$X_4$	<i>kscmh</i>
Débit de recyclage	$X_5$	<i>kscmh</i>
Débit d'alimentation du réacteur	$X_6$	<i>kscmh</i>
Pression du réacteur	$X_7$	<i>kPa</i>
Niveau du réacteur	$X_8$	%
Température du réacteur	$X_9$	°C
Débit de purge	$X_{10}$	<i>kscmh</i>
Température du séparateur	$X_{11}$	°C
Niveau du séparateur	$X_{12}$	%
Pression du séparateur	$X_{13}$	<i>kPa</i>
Débit du séparateur	$X_{14}$	<i>m<sup>3</sup>/h</i>
Niveau du décapeur	$X_{15}$	%
Pression du décapeur	$X_{16}$	<i>kPa</i>
Débit du décapeur	$X_{17}$	<i>m<sup>3</sup>/h</i>
Température du séparateur	$X_{18}$	°C
Débit de gaz au séparateur	$X_{19}$	<i>kg/h</i>
Puissance du compresseur	$X_{20}$	<i>kW</i>
Température de ref. liq. en sortie de réacteur	$X_{21}$	°C
Température de ref. liq. en sortie de séparateur	$X_{22}$	°C

TABLE 1 – Variables de mesure en continu

## Annexe C : Modèle BG du TEP

Les modèles BG (relatifs au domaine chimique) des réactifs (*B*, *C*, *D*, *E*) et des produits (*F*, *G*, *H*) sont donnés dans cette section.

**Réactifs :** Le réactif inerte *B* est fourni par des sources différentes :

- le flux d'alimentation en *A*, représenté dans la Figure 1 par  $F_1$ ,



Composant	Variable	Période d'échantillonnage (min)
A	$X_{23}$	6
B	$X_{24}$	6
C	$X_{25}$	6
D	$X_{26}$	6
E	$X_{27}$	6
F	$X_{28}$	6
A	$X_{29}$	6
B	$X_{30}$	6
C	$X_{31}$	6
D	$X_{32}$	6
E	$X_{33}$	6
F	$X_{34}$	6
G	$X_{35}$	6
H	$X_{36}$	6
D	$X_{37}$	15
E	$X_{38}$	15
F	$X_{39}$	15
G	$X_{40}$	15
H	$X_{41}$	15

TABLE 2 – Variables de mesure échantillonnées

Description	Variable	Unité
Débit d'alimentation en $D$	$X_{42}$	$kg/h$
Débit d'alimentation en $E$	$X_{43}$	$kg/h$
Débit d'alimentation en $A$	$X_{44}$	$kscmh$
Débit d'alimentation en $A$ et $C$	$X_{45}$	$kscmh$
Vanne de recyclage du compresseur	$X_{46}$	%
Vanne de purge	$X_{47}$	%
Débit d'alimentation du séparateur	$X_{48}$	$m^3/h$
Débit d'alimentation du décapeur	$X_{49}$	$m^3/h$
Vanne du décapeur	$X_{50}$	%
Débit du refroidissement liquide au réacteur	$X_{51}$	$m^3/h$
Débit du refroidissement liquide au condenseur	$X_{52}$	$m^3/h$
Vitesse de l'agitateur	$X_{53}$	$rpm$

TABLE 3 – Variables de contrôle

- le flux d'alimentation en  $D$ , décrit par  $F_2$ ,
- le flux sortant du décapeur vers le réacteur, illustré par  $y_{B5}F_5$ ,
- le flux de recyclage  $F_8$ .

Le cheminement du gaz  $B$  dans le procédé est similaire au gaz  $A$ .

Concernant le réactif  $C$  de la Figure 2, il est présent dans le flux sortant du décapeur vers le réacteur, illustré par  $y_{C5}F_5$  et dans le flux de recyclage  $F_8$ . Les autres parties

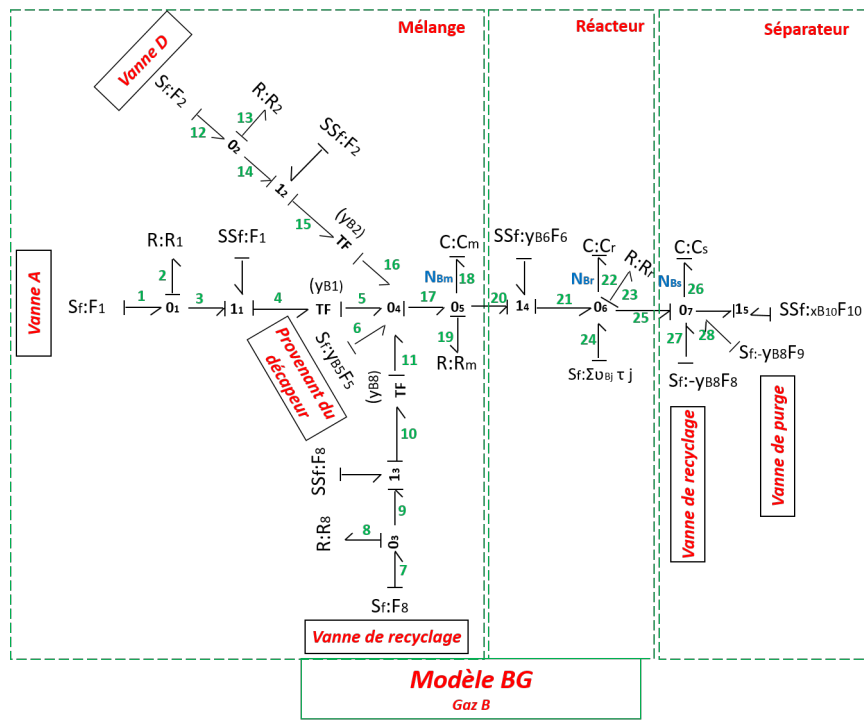


FIGURE 1 – Modèle BG du réactif *B* dans le TEP, avec causalité dérivée

(réacteur et séparateur) ont le même modèle que le gaz *A*.

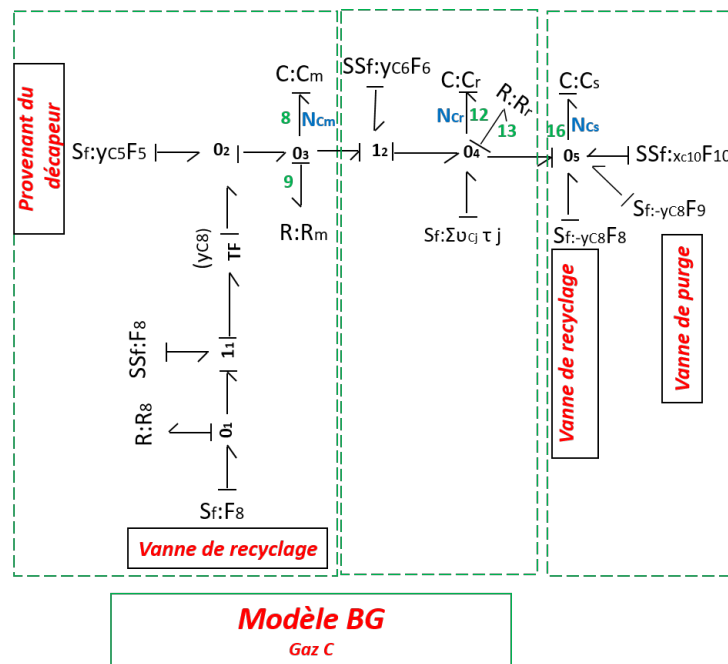
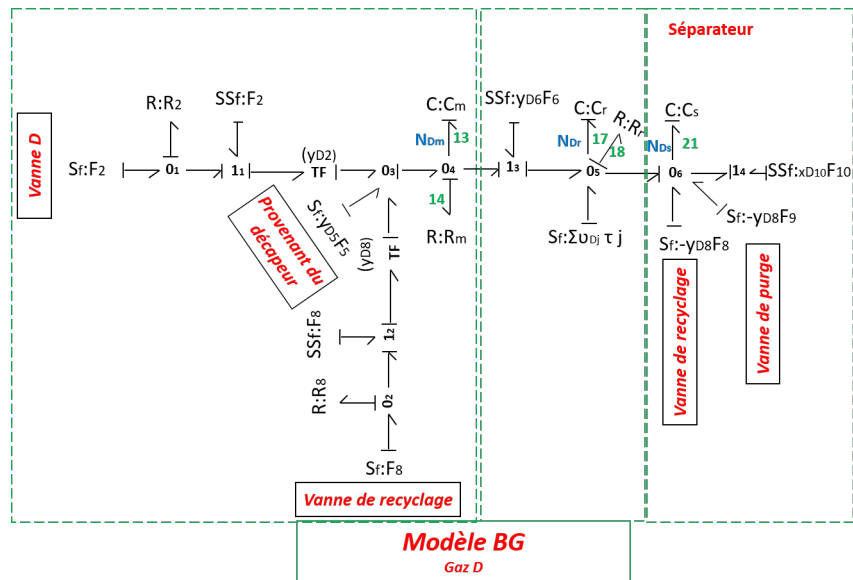
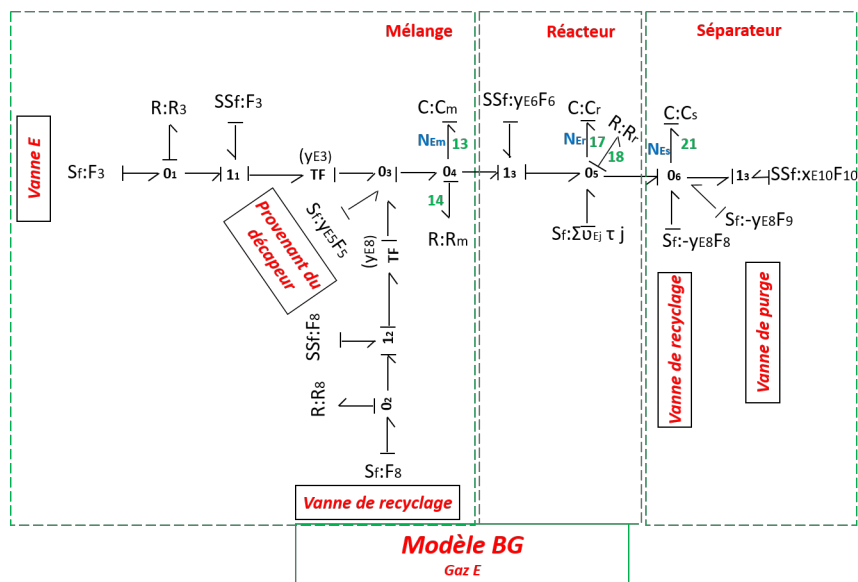


FIGURE 2 – Modèle BG du réactif *C* dans le TEP, avec causalité dérivée

Enfin, les réactifs *D* et *E* sont respectivement présents dans le flux d'alimentation *D*, représenté dans la figure 3 par  $F_2$ , et le flux d'alimentation *E*, représenté dans la Figure 4 par  $F_3$ , en plus du flux sortant du décapeur vers le réacteur et le flux de recyclage. Les modèles de réacteur et de séparateur sont similaires au gaz *A*.

**Produits :** La zone de mélange, comme l'illustre la Figure 5, montre que le produit *G* est présent dans le flux sortant du décapeur vers le réacteur, illustré par  $y_{G5}F_5$  et dans le

FIGURE 3 – Modèle BG du réactif *D* dans le TEP, avec causalité dérivéeFIGURE 4 – Modèle BG du réactif *E* dans le TEP, avec causalité dérivée

flux de recyclage  $F_8$ . Après avoir traversé le réacteur et le séparateur, le gaz *G* se retrouve dans le flux liquide sortant du séparateur (flux 10). Ce dernier est mesuré par le capteur représenté par le signal modulé  $SSf : x_{G10}F_{10}$ . Le gaz *G* passe ensuite dans le décapeur où il est collecté et mesuré ( $SSf : x_{G11}F_{11}$ ).

Le même raisonnement est appliqué au deuxième produit *H* (Figure 6).

Enfin, le produit dérivé *F* (Figure 7) a le même modèle BG que le réactif *E*.

## Annexe D : Modèle hydraulique et thermodynamique

**Modèle hydraulique du TEP** Le modèle BG hydraulique est donné dans la Figure 8.

Dans la zone de mélange, les 3 sources de pression ( $P_1, P_2, P_3$ ) proviennent des réservoirs contenant les gaz *A*, *D* et *E*. Chaque gaz traverse une vanne, représentée par une résistance ( $R : R_{VA}, R : R_{VD}, R : R_{VE}$ ). L'expression de la résistance est comme

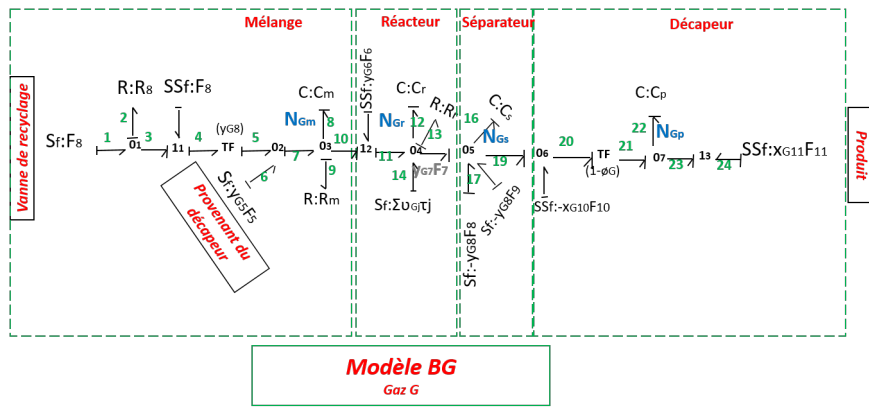


FIGURE 5 – Modèle BG du produit G dans le TEP, avec causalité dérivée

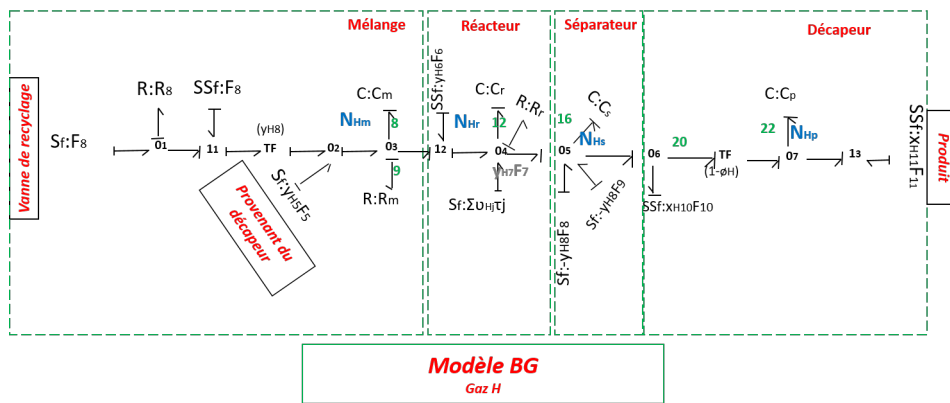


FIGURE 6 – Modèle BG du produit H dans le TEP, avec causalité dérivée

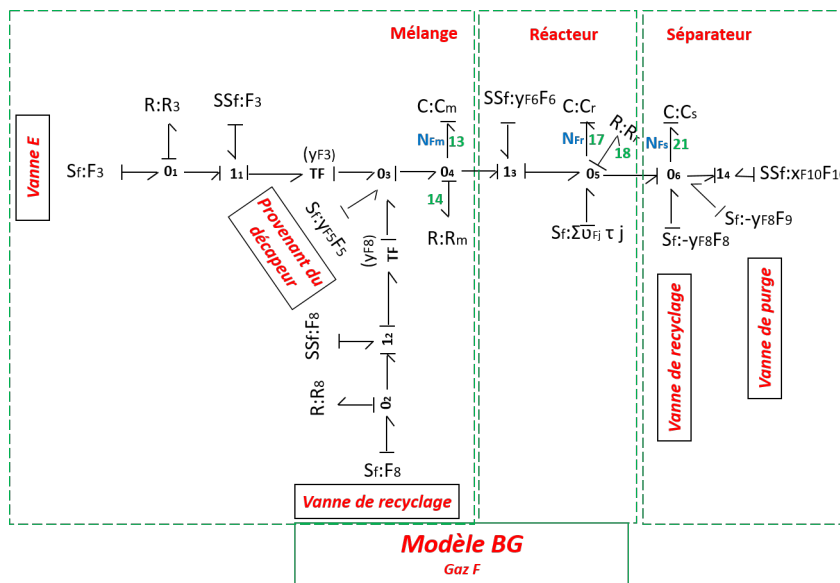


FIGURE 7 – Modèle BG du produit dérivé F dans le TEP, avec causalité dérivée

suit :

$$R_V = \frac{8\chi L}{\pi R^4} \tag{42}$$

avec  $\chi$  la viscosité du fluide,  $L$  la longueur de la vanne et  $R$  son rayon.

Lorsque le compresseur retourne au réacteur une certaine quantité de vapeur qui sort

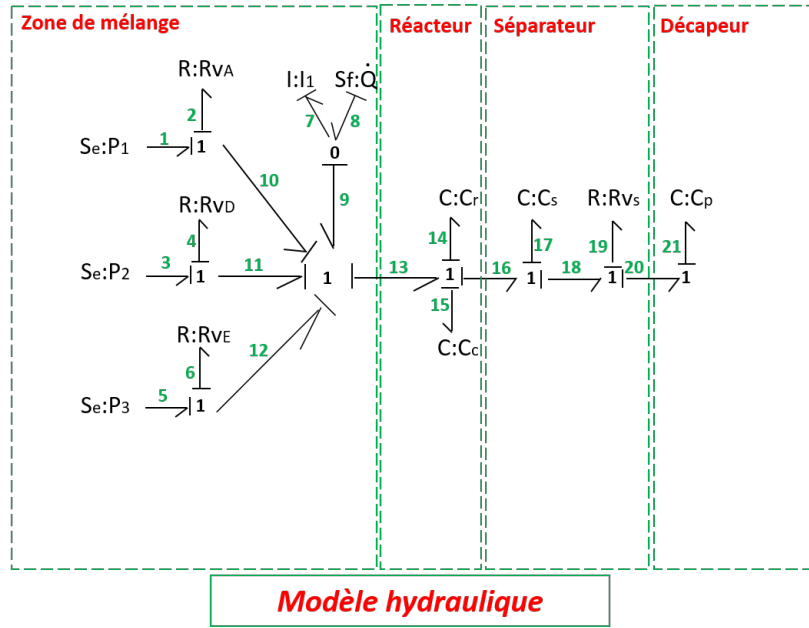


FIGURE 8 – Modèle BG hydraulique

du séparateur, une source de flux  $S_f : \dot{Q}$  représentant le débit volumique de ce gaz apparaît dans le modèle BG. Ce gaz traverse un tuyau qui possède une inertie, représentée dans le modèle BG par l'élément  $I : I_1$ .

A l'entrée du réacteur, les pressions sont ajoutées alors que le débit volumique reste le même, ce qui explique l'utilisation de la jonction 1. Les éléments  $\{C : C_r, C : C_c, C : C_s, C : C_p\}$  correspondent respectivement à la capacité du réacteur, du condenseur, du séparateur et du décapeur. Ces éléments passifs stockent la puissance sous la forme d'énergie potentielle. L'élément  $C$  peut être défini comme suit :

$$C = \frac{A}{\rho g} \quad (43)$$

où  $A$  est la section du réservoir,  $\rho$  la densité et  $g$  la gravité.

Enfin, l'élément  $R : R_{VS}$  modélise la puissance dissipée lorsque le liquide passe du séparateur au décapeur.

**Modèle thermodynamique du TEP** Le modèle BG thermodynamique est donné dans la Figure 9.

Concentrons-nous sur le modèle du réacteur où  $S_{f3} = -\sum_{j=1}^3 \Delta H \tau_j$ ,  $S_{f4} = -\dot{Q}_r$ ,  $S_{f5} = F_6 \left( \sum_{i=A}^H y_{i6} C_{p,vap,i} T_6 \right)$ ,  $S_{f6} = F_6 \left( \sum_{i=A}^H y_{i6} C_{p,vap,i} T_r \right)$  et  $C_r = \sum_{i=A}^H N_{ir} C_{p,i}$ . On peut déduire, à partir de la jonction 0, l'équation suivante du bilan énergétique :

$$\left( \sum_{i=A}^H N_{ir} C_{pi} \right) \frac{dT_r}{dt} = F_6 \left( \sum_{i=A}^H y_{i6} C_{p,vap,i} \right) (T_6 - T_r) - \dot{Q}_r - \sum_{j=1}^3 \Delta H \tau_j \quad (44)$$

où  $T_r$  représente la température du réacteur [ $^{\circ}C$ ],  $Q_r$  est l'énergie qui en est retirée [ $kJ$ ],  $C_{pi}$  et  $C_{p,vap,i}$  représentent respectivement la capacité calorifique spécifique du liquide et de la vapeur du composant  $i$  [ $kJ/kg.C$ ],  $\Delta H$  est la chaleur exothermique [ $kJ/kmol$ ]. Le même raisonnement conduit à trois autres équations du bilan énergétique :

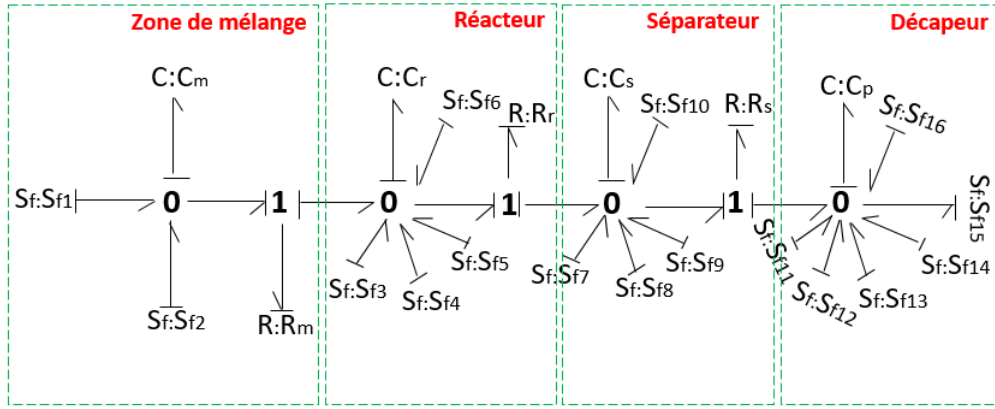


FIGURE 9 – Modèle BG thermodynamique

$$\left( \sum_{i=A}^H N_{im} C_{p,vap,i} \right) \frac{dT_m}{dt} = \sum_{j=1,2,3,5,8} F_j \left( \sum_{i=A}^H y_{ij} C_{p,vap,i} \right) (T_j - T_m) \quad (45)$$

$$\left( \sum_{i=A}^H N_{is} C_{pi} \right) \frac{dT_s}{dt} = F_7 \left( \sum_{i=A}^H y_{i7} C_{p,vap,i} \right) (T_r - T_s) - \dot{Q}_s - H_0 V_s \quad (46)$$

$$\left( \sum_{i=G}^H N_{ip} C_{pi} \right) \frac{dT_p}{dt} = F_{10} \left( \sum_{i=A}^H x_{i10} C_{pi} \right) (T_s - T_p) + F_4 \left( \sum_{i=A}^H y_{i4} C_{p,vap,i} \right) (T_4 - T_p) - \dot{Q}_p - H_0 V_p \quad (47)$$

où  $H_0$  est l'enthalpie de référence [ $kJ$ ],  $V_s$  et  $V_p$  sont les volumes respectivement du séparateur et du décapeur [ $m^3$ ].

## Annexe E : Génération des RRAs

Les résidus liés aux réactifs ( $B, C, D, E$ ) et aux produits ( $F, G, H$ ) sont les suivants :

### Gas B

$$r_3 = y_{B1} F_1 + y_{B5} F_5 + y_{B8} F_8 + y_{B2} F_2 - C_m \frac{d\mu_{Bm}}{dt} - \frac{\mu_{Bm}}{R_m} - y_{B6} F_6 \quad (48)$$

$$r_4 = y_{B6} F_6 - C_r \frac{d\mu_{Br}}{dt} - \frac{\mu_{Br}}{R_r} + \sum_{j=1}^3 \nu_{bj} \tau_j - C_s \frac{d\mu_{Bs}}{dt} - y_{B8} (F_8 + F_9) - x_{B10} F_{10} \quad (49)$$

### Gas C

$$r_5 = y_{C5} F_5 + y_{C8} F_8 - C_m \frac{d\mu_{Cm}}{dt} - \frac{\mu_{Cm}}{R_m} - y_{C6} F_6 \quad (50)$$

$$r_6 = y_{C6} F_6 - C_r \frac{d\mu_{Cr}}{dt} - \frac{\mu_{Cr}}{R_r} + \sum_{j=1}^3 \nu_{Cj} \tau_j - C_s \frac{d\mu_{Cs}}{dt} - y_{C8} (F_8 + F_9) - x_{C10} F_{10} \quad (51)$$

**Gas D**

$$r_7 = y_{D2}F_2 + y_{D5}F_5 + y_{D8}F_8 - C_m \frac{d\mu_{Dm}}{dt} - \frac{\mu_{Dm}}{R_m} - y_{D6}F_6 \quad (52)$$

$$r_8 = y_{D6}F_6 - C_r \frac{d\mu_{Dr}}{dt} - \frac{\mu_{Dr}}{R_r} + \sum_{j=1}^3 \nu_{Dj}\tau_j - C_s \frac{d\mu_{Ds}}{dt} - y_{D8}(F_8 + F_9) - x_{D10}F_{10} \quad (53)$$

**Gas E**

$$r_9 = y_{E3}F_3 + y_{E5}F_5 + y_{E8}F_8 - C_m \frac{d\mu_{Em}}{dt} - \frac{\mu_{Em}}{R_m} - y_{E6}F_6 \quad (54)$$

$$r_{10} = y_{E6}F_6 - C_r \frac{d\mu_{Er}}{dt} - \frac{\mu_{Er}}{R_r} + \sum_{j=1}^3 \nu_{Ej}\tau_j - C_s \frac{d\mu_{Es}}{dt} - y_{E8}(F_8 + F_9) - x_{E10}F_{10} \quad (55)$$

**Gas F**

$$r_{11} = y_{F3}F_3 + y_{F5}F_5 + y_{F8}F_8 - C_m \frac{d\mu_{Fm}}{dt} - \frac{\mu_{Fm}}{R_m} - y_{F6}F_6 \quad (56)$$

$$r_{12} = y_{F6}F_6 - C_r \frac{d\mu_{Fr}}{dt} - \frac{\mu_{Fr}}{R_r} + \sum_{j=1}^3 \nu_{Fj}\tau_j - C_s \frac{d\mu_{Fs}}{dt} - y_{F8}(F_8 + F_9) - x_{F10}F_{10} \quad (57)$$

**Gas G**

$$r_{13} = y_{G8}F_8 + y_{G5}F_5 - C_m \frac{d\mu_{Gm}}{dt} - \frac{\mu_{Gm}}{R_m} - y_{G6}F_6 \quad (58)$$

$$r_{14} = y_{G6}F_6 - C_r \frac{d\mu_{Gr}}{dt} - \frac{\mu_{Gr}}{R_r} + \sum_{j=1}^3 \nu_{Gj}\tau_j - C_s \frac{d\mu_{Gs}}{dt} - y_{G8}(F_8 + F_9) - x_{G10}F_{10} \quad (59)$$

$$r_{15} = (1 - \phi_G)x_{G10}F_{10} - C_p \frac{d\mu_{Gp}}{dt} - x_{G11}F_{11} \quad (60)$$

**Gas H**

$$r_{16} = y_{H8}F_8 + y_{H5}F_5 - C_m \frac{d\mu_{Hm}}{dt} - \frac{\mu_{Hm}}{R_m} - y_{H6}F_6 \quad (61)$$

$$r_{17} = y_{H6}F_6 - C_r \frac{d\mu_{Hr}}{dt} - \frac{\mu_{Hr}}{R_r} + \sum_{j=1}^3 \nu_{Hj}\tau_j - C_s \frac{d\mu_{Hs}}{dt} - y_{H8}(F_8 + F_9) - x_{H10}F_{10} \quad (62)$$

$$r_{18} = (1 - \phi_H)x_{H10}F_{10} - C_p \frac{d\mu_{Hp}}{dt} - x_{H11}F_{11} \quad (63)$$

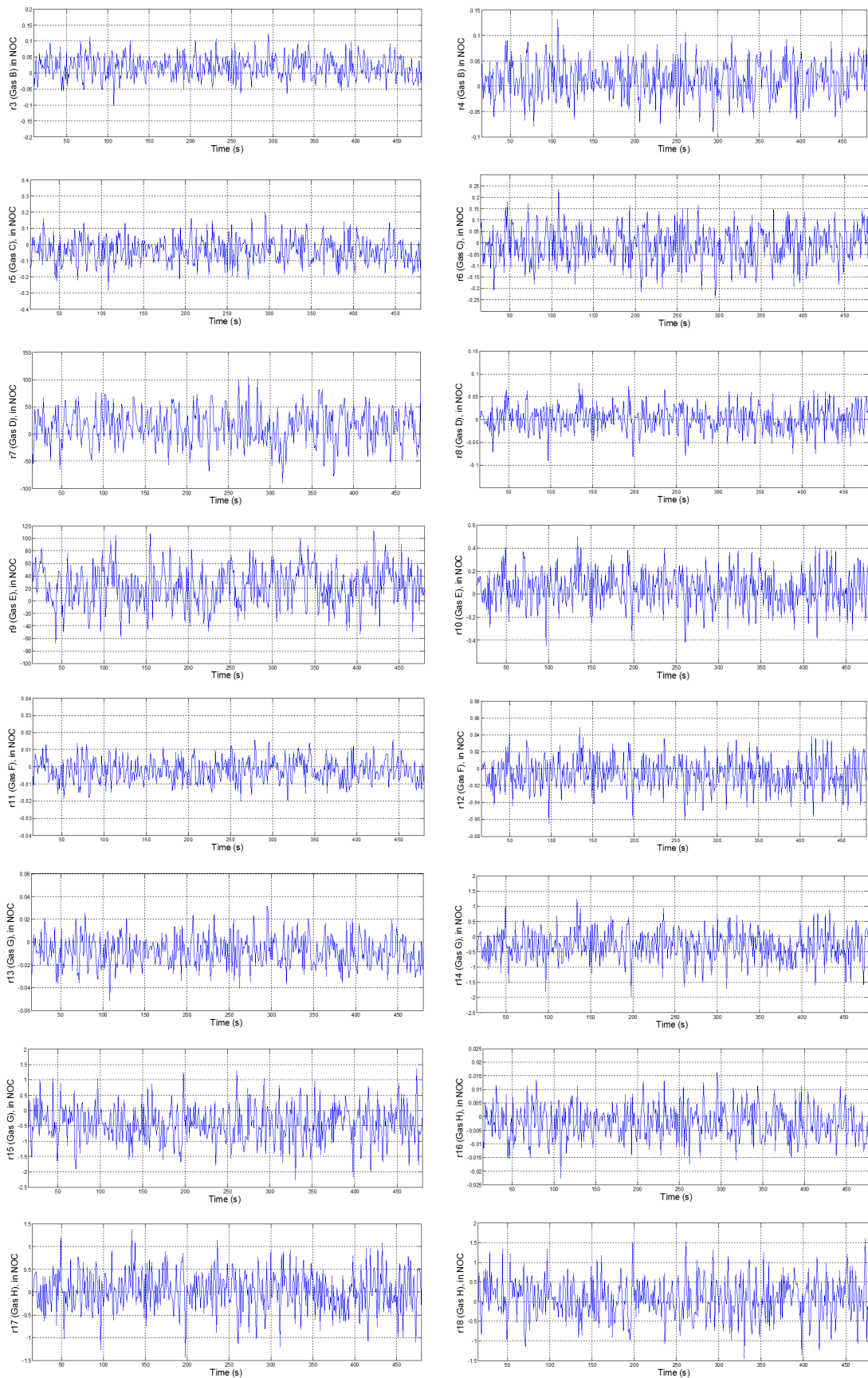


FIGURE 10 – Évaluation des résidus  $r_3 - r_{18}$  liés aux gaz  $B, C, D, E, F, G, H$ , en  $CFN$

## Annexe F : Evaluation des RRA

La Figure 10 montre l'évaluation des résidus avec des variables connues en  $CFN$ , pour les réactifs ( $B, C, D, E$ ) et les produits ( $F, G, H$ ). Tous les résidus sont proches de zéro,



ce qui permet de valider le modèle *BG* du TEP.

## Annexe G : Développement des objectifs pour le pronostic

### Objectif 2

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_0/D_1 = F_0 \cap D_2 = F_1) > P(D = F_0) \\ P(D = F_0/D_1 = F_0 \cap D_2 = F_1) > P(D = F_i/D_1 = F_0 \cap D_2 = F_1), \\ i \in \{1, 2\} \end{array} \right. \quad (64)$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned} P(D = F_0/D_1 = F_0 \cap D_2 = F_1) &> P(D = F_0) \\ \Leftrightarrow \frac{P(D_1=F_0/D=F_0)P(D_2=F_1/D=F_0)}{P(D_1=F_0/D_2=F_1)P(D_2=F_1)} &> 1 \\ \Leftrightarrow \frac{a_0 a_2}{\alpha_0^2 (a_0 + b_0 + c_0)(a_2 + b_2 + c_2)} &> 1 \end{aligned} \quad (65)$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_1$  :

$$\begin{aligned} P(D = F_0/D_1 = F_0 \cap D_2 = F_1) &> P(D = F_1/D_1 = F_0 \cap D_2 = F_1) \\ \Leftrightarrow \frac{P(D_1=F_0/D=F_0)P(D_2=F_1/D=F_0)}{P(D_1=F_0/D_2=F_1)P(D_2=F_1)} &> \frac{P(D_1=F_0/D=F_1)P(D_2=F_1/D=F_1)}{P(D_1=F_0/D_2=F_1)P(D_2=F_1)} \\ \Leftrightarrow a_0 a_2 &> b_0 c_2 \end{aligned} \quad (66)$$

Pour  $D = F_2$ , on obtient :

$$\begin{aligned} P(D = F_0/D_1 = F_0 \cap D_2 = F_1) &> P(D = F_2/D_1 = F_0 \cap D_2 = F_1) \\ \Leftrightarrow \frac{P(D_1=F_0/D=F_0)P(D_2=F_1/D=F_0)}{P(D_1=F_0/D_2=F_1)P(D_2=F_1)} &> \frac{P(D_1=F_0/D=F_2)P(D_2=F_1/D=F_2)}{P(D_1=F_0/D_2=F_1)P(D_2=F_1)} \\ \Leftrightarrow a_0 a_2 &> b_2 c_0 \end{aligned} \quad (67)$$

**Résumé :** Par conséquent, 3 contraintes sont obtenues :

$$\frac{a_0 a_2}{(a_0 + b_0 + c_0)(a_2 + b_2 + c_2)} > 1 > \alpha_0^2 \quad (68)$$

$$\left\{ \begin{array}{l} a_0 a_2 > b_0 c_2 \\ a_0 a_2 > b_2 c_0 \end{array} \right. \quad (69)$$

### Objectif 3

Afin d'assurer cet objectif, les critères suivants (Eq. 70) doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_2/D_1 = F_0 \cap D_2 = F_2) > P(D = F_2) \\ P(D = F_2/D_1 = F_0 \cap D_2 = F_2) > P(D = F_i/D_1 = F_0 \cap D_2 = F_2), \\ i \in \{0, 1\} \end{array} \right. \quad (70)$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned}
 P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = F_2) \\
 \Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> 1 \\
 \Leftrightarrow \frac{c_0 b_1}{\alpha_0^2(a_0+b_0+c_0)(a_2+b_2+c_2)} &> 1
 \end{aligned} \tag{71}$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$  :

$$\begin{aligned}
 P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = F_0/D_1 = F_0 \cap D_2 = F_2) \\
 \Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_0/D=F_0)P(D_2=F_2/D=F_0)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} \\
 \Leftrightarrow c_0 b_1 &> a_0 a_1
 \end{aligned} \tag{72}$$

Pour  $D = F_1$ , on obtient :

$$\begin{aligned}
 P(D = F_2/D_1 = F_0 \cap D_2 = F_2) &> P(D = F_1/D_1 = F_0 \cap D_2 = F_2) \\
 \Leftrightarrow \frac{P(D_1=F_0/D=F_2)P(D_2=F_2/D=F_2)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} &> \frac{P(D_1=F_0/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=F_0/D_2=F_2)P(D_2=F_2)} \\
 \Leftrightarrow c_0 b_1 &> c_1 b_0
 \end{aligned} \tag{73}$$

**Résumé :** Par conséquent, 3 contraintes sont obtenues :

$$\frac{c_0 b_1}{(a_0 + b_0 + c_0)(a_2 + b_2 + c_2)} > 1 > \alpha_0^2 \tag{74}$$

$$\begin{cases} c_0 b_1 > a_0 a_1 \\ c_0 b_1 > c_1 b_0 \end{cases} \tag{75}$$

## Objectif 4

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\begin{cases} P(D = F_1/D_1 = F_1 \cap D_2 = F_0) > P(D = F_1) \\ P(D = F_i/D_1 = F_1 \cap D_2 = F_0) > P(D = F_i/D_1 = F_1 \cap D_2 = F_0), \\ i \in \{0, 2\} \end{cases} \tag{76}$$

Le raisonnement est identique à celui de l'Objectif 3 et les mêmes contraintes sont déduites.

## Objectif 5

Afin d'assurer cet objectif, les critères suivants (Eq. 77) doivent être respectés :

$$\begin{cases} P(D = F_1/D_1 = F_1 \cap D_2 = F_1) > P(D = F_1) \\ P(D = F_i/D_1 = F_1 \cap D_2 = F_1) > P(D = F_i/D_1 = F_1 \cap D_2 = F_1), \\ i \in \{0; 2\} \end{cases} \tag{77}$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned}
 P(D = F_1/D_1 = F_1 \cap D_2 = F_1) &> P(D = F_1) \\
 \Leftrightarrow \frac{P(D_1=F_1/D=F_1)P(D_2=F_1/D=F_1)}{P(D_1=F_1/D_2=F_1)P(D_2=F_1)} &> 1 \\
 \Leftrightarrow \frac{b_1c_2}{\alpha_0^2(a_1+b_1+c_1)(a_2+b_2+c_2)} &> 1
 \end{aligned} \tag{78}$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$  :

$$\begin{aligned}
 P(D = F_1/D_1 = F_1 \cap D_2 = F_1) &> P(D = F_0/D_1 = F_1 \cap D_2 = F_1) \\
 \Leftrightarrow \frac{P(D_1=F_1/D=F_1)P(D_2=F_1/D=F_1)}{P(D_1=F_1/D_2=F_1)P(D_2=F_1)} &> \frac{P(D_1=F_1/D=F_0)P(D_2=F_1/D=F_0)}{P(D_1=F_1/D_2=F_1)P(D_2=F_1)} \\
 \Leftrightarrow b_1c_2 &> a_1a_2
 \end{aligned} \tag{79}$$

Pour  $D = F_2$ , on obtient la contrainte :  $b_1c_2 > b_2c_1$ .

**Résumé :** Par conséquent, 3 contraintes sont obtenues :

$$\frac{b_1c_2}{(a_1 + b_1 + c_1)(a_2 + b_2 + c_2)} > \alpha_0^2 \tag{80}$$

$$\begin{cases} b_1c_2 > a_1a_2 \\ b_1c_2 > b_2c_1 \end{cases} \tag{81}$$

## Objectif 6

Afin d'assurer cet objectif, les critères suivants (Eq. 82) doivent être respectés :

$$\begin{cases} P(D = F_1/D_1 = F_1 \cap D_2 = F_2) > P(D = F_1) \\ P(D = F_i/D_1 = F_1 \cap D_2 = F_2) > P(D = F_i/D_1 = F_1 \cap D_2 = F_2), \\ i \in \{0\} \end{cases} \tag{82}$$

**Critère 1 :** La première inégalité est développée comme il suit :

$$\begin{aligned}
 P(D = F_1/D_1 = F_1 \cap D_2 = F_2) &> P(D = F_1) \\
 \Leftrightarrow \frac{P(D_1=F_1/D=F_1)P(D_2=F_2/D=F_1)}{P(D_1=F_1/D_2=F_2)P(D_2=F_2)} &> 1 \\
 \Leftrightarrow \frac{b_1c_1}{\alpha_0^2(a_1+b_1+c_1)^2} &> 1
 \end{aligned} \tag{83}$$

**Critère 2 :** La seconde inégalité est développée suivant le même raisonnement.

Pour  $D = F_0$  :

$$\begin{aligned}
 P(D = F_1/D_1 = F_1 \cap D_2 = F_2) &> P(D = F_0/D_1 = F_1 \cap D_2 = F_2) \\
 \Leftrightarrow P(D_1 = F_1/D = F_1)P(D_2 = F_2/D = F_1) &> P(D_1 = F_1/D = F_0)P(D_2 = F_2/D = F_0) \\
 \Leftrightarrow b_1c_1 &> a_1^2
 \end{aligned} \tag{84}$$

Les mêmes contraintes sont obtenues pour le deuxième choix de D, soit  $d = F_2$ .

**Résumé** Par conséquent, 2 contraintes sont déduites :

$$\frac{b_1 c_1}{(a_1 + b_1 + c_1)^2} > \alpha_0^2 \quad (85)$$

$$\{ b_1 c_1 > a_1^2 \} \quad (86)$$

### Objectif 7

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_0/D_1 = F_2 \cap D_2 = F_0) > P(D = F_0) \\ P(D = F_0/D_1 = F_2 \cap D_2 = F_0) > P(D = F_i/D_1 = F_2 \cap D_2 = F_0), \\ i \in \{1, 2\} \end{array} \right. \quad (87)$$

Le raisonnement est identique à celui de l'Objectif 2 et les mêmes contraintes sont déduites.

### Objectif 8

Afin d'assurer cet objectif, les critères suivants doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_1/D_1 = F_2 \cap D_2 = F_1) > P(D = F_1) \\ P(D = F_1/D_1 = F_2 \cap D_2 = F_1) > P(D = F_i/D_1 = F_2 \cap D_2 = F_1), \\ i \in \{0\} \end{array} \right. \quad (88)$$

Le raisonnement est identique à celui de l'Objectif 6 et les contraintes suivantes sont déduites.

$$\frac{b_2 c_2}{(a_2 + b_2 + c_2)^2} > \alpha_0^2 \quad (89)$$

$$\{ b_2 c_2 > a_2^2 \} \quad (90)$$

### Objectif 9

Afin d'assurer cet objectif, les critères suivants (Eq. 91) doivent être respectés :

$$\left\{ \begin{array}{l} P(D = F_2/D_1 = F_2 \cap D_2 = F_2) > P(D = F_2) \\ P(D = F_2/D_1 = F_2 \cap D_2 = F_2) > P(D = F_i/D_1 = F_2 \cap D_2 = F_2), \\ i \in \{0; 1\} \end{array} \right. \quad (91)$$

Le raisonnement est identique à celui de l'Objectif 5 et les contraintes suivantes sont déduites.

$$\frac{b_1 c_2}{(a_1 + b_1 + c_1)(a_2 + b_2 + c_2)} > \alpha_0^2 \quad (92)$$

$$\left\{ \begin{array}{l} b_1 c_2 > a_1 a_2 \\ b_1 c_2 > b_2 c_1 \end{array} \right. \quad (93)$$



# Bibliographie

- [Alwan and Roberts, 1988] Alwan, L. C. and Roberts, H. V. (1988). Time-series modeling for statistical process control. *Journal of Business & Economic Statistics*, 6(1) :87–95. [13](#), [26](#)
- [Alzghoul et al., 2014] Alzghoul, A., Backe, B., Löfstrand, M., Byström, A., and Liljedahl, B. (2014). Comparing a knowledge-based and a data-driven method in querying data streams for system fault detection : A hydraulic drive system application. *Computers in Industry*, 65(8) :1126–1135. [23](#), [27](#)
- [Antal et al., 2003] Antal, P., Fannes, G., Timmerman, D., Moreau, Y., and De Moor, B. (2003). Bayesian applications of belief networks and multilayer perceptrons for ovarian tumor classification with rejection. *Artificial Intelligence in Medicine*, 29(1) :39–60. [20](#)
- [Arunkumar et al., 2006] Arunkumar, N., Karunamoorthy, L., Anand, S., and Babu, T. R. (2006). Linear approach for solving a piecewise linear vendor selection problem of quantity discounts using lexicographic method. *The International Journal of Advanced Manufacturing Technology*, 28(11-12) :1254–1260. [74](#)
- [Atoui et al., 2015a] Atoui, M. A., Verron, S., and Kobi, A. (2015a). A Bayesian network dealing with measurements and residuals for system monitoring. *Transactions of the Institute of Measurement and Control*, page 0142331215581446. [36](#), [43](#), [47](#), [48](#), [88](#)
- [Atoui et al., 2015b] Atoui, M. A., Verron, S., and Kobi, A. (2015b). Fault Detection and Diagnosis in a Bayesian Network classifier incorporating probabilistic boundary 1. *IFAC-PapersOnLine*, 48(21) :670–675. [20](#), [87](#)
- [Atoui et al., 2015c] Atoui, M. A., Verron, S., and Kobi, A. (2015c). Fault detection with Conditional Gaussian Network. *Engineering Applications of Artificial Intelligence*, 45 :473–481. [50](#)
- [Aye and Heyns, 2017] Aye, S. and Heyns, P. (2017). An integrated gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission. *Mechanical Systems and Signal Processing*, 84 :485–498. [26](#)
- [Baraldi et al., 2013a] Baraldi, P., Cadini, F., Mangili, F., and Zio, E. (2013a). Model-based and data-driven prognostics under different available information. *Probabilistic Engineering Mechanics*, 32 :66–79. [7](#), [13](#)
- [Baraldi et al., 2013b] Baraldi, P., Compare, M., Saucó, S., and Zio, E. (2013b). Ensemble neural network-based particle filtering for prognostics. *Mechanical Systems and Signal Processing*, 41(1-2) :288–300. [36](#), [37](#)
- [Barraza-Barraza et al., 2017] Barraza-Barraza, D., Tercero-Gómez, V. G., Beruvides, M. G., and Limón-Robles, J. (2017). An adaptive arx model to estimate the rul of

- aluminum plates based on its crack growth. *Mechanical Systems and Signal Processing*, 82 :519–536. 26
- [Basseville et al., 1993] Basseville, M., Nikiforov, I. V., et al. (1993). *Detection of abrupt changes : theory and application*, volume 104. Prentice Hall Englewood Cliffs. 7
- [Bastin and Gevers, 1988] Bastin, G. and Gevers, M. R. (1988). Stable adaptive observers for nonlinear time-varying systems. *Automatic Control, IEEE Transactions on*, 33(7) :650–658. 14
- [Benkedjouh et al., 2013] Benkedjouh, T., Medjaher, K., Zerhouni, N., and Rechak, S. (2013). Remaining useful life estimation based on nonlinear feature reduction and support vector regression. *Engineering Applications of Artificial Intelligence*, 26(7) :1751–1760. 26
- [Benkouider et al., 2009] Benkouider, A. M., Buvat, J. C., Cosmao, J. M., and Saboni, A. (2009). Fault detection in semi-batch reactor using the EKF and statistical method. *Journal of Loss Prevention in the Process Industries*, 22(2) :153–161. 32, 37
- [Benkouider et al., 2012] Benkouider, A. M., Kessas, R., Yahiaoui, A., Buvat, J. C., and Guella, S. (2012). A hybrid approach to faults detection and diagnosis in batch and semi-batch reactors by using EKF and neural network classifier. *Journal of Loss Prevention in the Process Industries*, 25(4) :694–702. 32, 37
- [Bloch, 1996] Bloch, I. (1996). Some aspects of dempster-shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recognition Letters*, 17(8) :905–919. 40
- [Boser et al., 1992] Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM. 13, 26
- [Bouamama et al., 2014] Bouamama, B. O., Biswas, G., Loureiro, R., and Merzouki, R. (2014). Graphical methods for diagnosis of dynamic systems : Review. *Annual Reviews in Control*, 38(2) :199–219. 17
- [Bouamama et al., 2006] Bouamama, B. O., Medjaher, K., Samantaray, A., and Staroswiecki, M. (2006). Supervision of an industrial steam generator. part i : Bond graph modelling. *Control Engineering Practice*, 14(1) :71–83. 88
- [Bouamama et al., 2005] Bouamama, B. O., Samantaray, A., Medjaher, K., Staroswiecki, M., and Dauphin-Tanguy, G. (2005). Model builder using functional and bond graph tools for fdi design. *Control Engineering Practice*, 13(7) :875–891. 16
- [Boudraa et al., 2004] Boudraa, A. O., Bentabet, L., and Salzenstein, F. (2004). Dempster-shafer’s basic probability assignment based on fuzzy membership functions. *ELCVIA : electronic letters on computer vision and image analysis*, 4(1) :1–10. 40
- [Brahim-Belhouari and Bermak, 2004] Brahim-Belhouari, S. and Bermak, A. (2004). Gaussian process for nonstationary time series prediction. *Computational Statistics & Data Analysis*, 47(4) :705–712. 26
- [Breiman et al., 1984] Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and regression trees*. CRC press. 13, 22

- [Bressel et al., 2016] Bressel, M., Hilairet, M., Hissel, D., and Bouamama, B. O. (2016). Extended kalman filter for prognostic of proton exchange membrane fuel cell. *Applied Energy*, 164 :220–227. [25](#)
- [Byington et al., 2002] Byington, C. S., Roemer, M. J., and Galie, T. (2002). Prognostic enhancements to diagnostic systems for improved condition-based maintenance [military aircraft]. In *Aerospace Conference Proceedings, 2002. IEEE*, volume 6, pages 6–6. IEEE. [24](#), [25](#)
- [Caccavale et al., 2009] Caccavale, F., Pierri, F., Iamarino, M., and Tufano, V. (2009). An integrated approach to fault diagnosis for a class of chemical batch processes. *Journal of process control*, 19(5) :827–841. [15](#)
- [Cai et al., 2016] Cai, B., Liu, H., and Xie, M. (2016). A real-time fault diagnosis methodology of complex systems using object-oriented bayesian networks. *Mechanical Systems and Signal Processing*, 80 :31–44. [20](#)
- [Chanthery et al., 2015] Chanthery, E., Pencolé, Y., Ribot, P., and Travé-Massuyès, L. (2015). Hydiag : extended diagnosis and prognosis for hybrid systems. In *The 26th International Workshop on Principles of Diagnosis (DX-2015)*. [17](#)
- [Charnes et al., 1955] Charnes, A., Cooper, W. W., and Ferguson, R. O. (1955). Optimal estimation of executive compensation by linear programming. *Management science*, 1(2) :138–151. [73](#)
- [Chatti et al., 2013a] Chatti, N., Bouamama, B. O., Gehin, A.-L., and Merzouki, R. (2013a). Signed bond graph for health monitoring of pem fuel cell. In *5th International Conference of Development of Fuel Cells*, pages 6–pages. [88](#), [95](#)
- [Chatti et al., 2013b] Chatti, N., Gehin, A., Ould-Bouamama, B., and Merzouki, R. (2013b). Functional and behavior models for the supervision of an intelligent and autonomous system. *Automation Science and Engineering, IEEE Transactions on*, 10(2) :431–445. [88](#)
- [Chatti et al., 2016] Chatti, N., Guyonneau, R., Hardouin, L., Verron, S., and Lagrange, S. (2016). Model-based approach for fault diagnosis using set-membership formulation. *Engineering Applications of Artificial Intelligence*, 55 :307–319. [96](#)
- [Chatti et al., 2014] Chatti, N., Ould-Bouamama, B., Gehin, A.-L., and Merzouki, R. (2014). Signed bond graph for multiple faults diagnosis. *Engineering Applications of Artificial Intelligence*, 36 :134–147. [16](#), [17](#), [95](#)
- [Chen et al., 2017] Chen, Y. M., Hsueh, C.-S., Wang, C.-K., and Wu, T.-Y. (2017). Decision fusion using fuzzy threshold scheme for target detection in sensor networks. *Journal of Computational Science*. [36](#)
- [Cheng and Chiu, 2005] Cheng, C. and Chiu, M.-S. (2005). Nonlinear process monitoring using JITL-PCA. *Chemometrics and Intelligent Laboratory Systems*, 76(1) :1–13. [21](#)
- [Cheung and Stephanopoulos, 1990] Cheung, J.-Y. and Stephanopoulos, G. (1990). Representation of process trends—part i. a formal representation framework. *Computers & Chemical Engineering*, 14(4-5) :495–510. [13](#), [22](#)



- [Chiang et al., 2004] Chiang, L. H., Kotanchek, M. E., and Kordon, A. K. (2004). Fault diagnosis based on fisher discriminant analysis and support vector machines. *Computers & chemical engineering*, 28(8) :1389–1401. [88](#)
- [Chiang et al., 2000a] Chiang, L. H., Russell, E. L., and Braatz, R. D. (2000a). *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media. [7](#), [28](#), [88](#)
- [Chiang et al., 2000b] Chiang, L. H., Russell, E. L., and Braatz, R. D. (2000b). Fault diagnosis in chemical processes using fisher discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemometrics and intelligent laboratory systems*, 50(2) :243–252. [88](#)
- [Chow, 1980] Chow, E. (1980). Failure detection system design methodology. ph. d. thesis. [13](#), [15](#)
- [Chow and Willsky, 1984] Chow, E. Y. and Willsky, A. S. (1984). Analytical redundancy and the design of robust failure detection systems. *Automatic Control, IEEE Transactions on*, 29(7) :603–614. [15](#)
- [Coello, 2000] Coello, C. A. (2000). An updated survey of ga-based multiobjective optimization techniques. *ACM Computing Surveys (CSUR)*, 32(2) :109–143. [74](#)
- [Cohon and Marks, 1975] Cohon, J. L. and Marks, D. H. (1975). A review and evaluation of multiobjective programming techniques. *Water Resources Research*, 11(2) :208–220. [73](#)
- [Coppe et al., 2009] Coppe, A., Haftka, R. T., Kim, N.-H., and Yuan, F.-G. (2009). Reducing uncertainty in damage growth properties by structural health monitoring. In *Annual Conference of the Prognostics and Health Management Society*, volume 27. [25](#)
- [Cornuéjols and Miclet, 2011] Cornuéjols, A. and Miclet, L. (2011). *Apprentissage artificiel : concepts et algorithmes*. Editions Eyrolles. [22](#)
- [Correa et al., 2009] Correa, M., Bielza, C., and Pamies-Teixeira, J. (2009). Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process. *Expert systems with applications*, 36(3) :7270–7279. [20](#)
- [Cover and Hart, 1967] Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1) :21–27. [13](#), [22](#)
- [Dai and Gao, 2013] Dai, X. and Gao, Z. (2013). From model, signal to knowledge : A data-driven perspective of fault detection and diagnosis. *Industrial Informatics, IEEE Transactions on*, 9(4) :2226–2238. [15](#), [28](#)
- [Daigle and Goebel, 2011] Daigle, M. and Goebel, K. (2011). Multiple damage progression paths in model-based prognostics. In *Aerospace conference, 2011 IEEE*, pages 1–10. IEEE. [24](#)
- [Daigle, 2008] Daigle, M. J. (2008). *A qualitative event-based approach to fault diagnosis of hybrid systems*. PhD thesis. [16](#)
- [D’Angelo et al., 2016] D’Angelo, M. F., Palhares, R. M., Camargos Filho, M. C., Maia, R. D., Mendes, J. B., and Ekel, P. Y. (2016). A new fault classification approach applied to tennessee eastman benchmark process. *Applied Soft Computing*, 49 :676–686. [88](#)

- [Daroogheh et al., 2015] Daroogheh, N., Baniamerian, A., Meskin, N., and Khorasani, K. (2015). A hybrid prognosis and health monitoring strategy by integrating particle filters and neural networks for gas turbine engines. In *Prognostics and Health Management (PHM), 2015 IEEE Conference on*, pages 1–8. IEEE. [33](#), [37](#), [46](#)
- [Das and Dennis, 1998] Das, I. and Dennis, J. E. (1998). Normal-boundary intersection : A new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization*, 8(3) :631–657. [73](#)
- [de Mortain et al., 2015] de Mortain, F., Subias, A., Travé-Massuyès, L., and de Flaugergues, V. (2015). Towards active diagnosis of hybrid systems leveraging multimodel identification and a markov decision process. *IFAC-PapersOnLine*, 48(21) :171–176. [17](#)
- [Del Moral, 1996] Del Moral, P. (1996). Non-linear filtering : interacting particle resolution. *Markov processes and related fields*, 2(4) :555–581. [13](#), [24](#)
- [Dempster, 1967] Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339. [40](#)
- [Ding, 2008a] Ding, S. (2008a). *Model-based fault diagnosis techniques : design schemes, algorithms, and tools*. Springer Science & Business Media. [13](#)
- [Ding, 2008b] Ding, S. X. (2008b). *Model-based fault diagnosis techniques : design schemes, algorithms, and tools*. Springer Science & Business Media. [7](#)
- [Ding et al., 2011] Ding, S. X., Zhang, P., Jeinsch, T., Ding, E., Engel, P., and Gui, W. (2011). A survey of the application of basic data-driven and model-based methods in process monitoring and fault diagnosis. *IFAC Proceedings Volumes*, 44(1) :12380–12388. [7](#), [29](#)
- [Ding et al., 2009] Ding, S. X., Zhang, P., Naik, A., Ding, E. L., and Huang, B. (2009). Subspace method aided data-driven design of fault detection and isolation systems. *Journal of process control*, 19(9) :1496–1510. [7](#), [29](#), [31](#), [33](#), [36](#), [37](#), [46](#), [48](#), [88](#)
- [Djeziri et al., 2007] Djeziri, M. A., Merzouki, R., Bouamama, B. O., and Dauphin-Tanguy, G. (2007). Robust fault diagnosis by using bond graph approach. *Mechatronics, IEEE/ASME Transactions on*, 12(6) :599–611. [17](#), [96](#)
- [Downs and Vogel, 1993] Downs, J. J. and Vogel, E. F. (1993). A plant-wide industrial process control problem. *Computers & chemical engineering*, 17(3) :245–255. [88](#), [90](#), [91](#), [95](#)
- [Dromigny and Zhu, 1997] Dromigny, A. and Zhu, Y. M. (1997). Improving the dynamic range of real-time x-ray imaging systems via bayesian fusion. *Journal of Nondestructive Evaluation*, 16(3) :147–160. [39](#)
- [Eslamloueyan, 2011] Eslamloueyan, R. (2011). Designing a hierarchical neural network based on fuzzy clustering for fault diagnosis of the tennessee–eastman process. *Applied soft computing*, 11(1) :1407–1415. [88](#), [104](#)
- [Flores et al., 2011] Flores, M. J., Nicholson, A. E., Brunskill, A., Korb, K. B., and Mascaro, S. (2011). Incorporating expert knowledge when learning bayesian network structure : a medical case study. *Artificial intelligence in medicine*, 53(3) :181–204. [21](#)

- [Fourman, 1985] Fourman, M. P. (1985). Compaction of symbolic layout using genetic algorithms. In *Proceedings of the 1st International Conference on Genetic Algorithms*, pages 141–153. L. Erlbaum Associates Inc. [73](#), [74](#)
- [Frederick et al., 2007] Frederick, D. K., DeCastro, J. A., and Litt, J. S. (2007). User’s guide for the commercial modular aero-propulsion system simulation (c-mapss). [108](#), [132](#), [133](#), [164](#)
- [Friedman, 1989] Friedman, J. H. (1989). Regularized discriminant analysis. *Journal of the American statistical association*, 84(405) :165–175. [13](#), [22](#)
- [Gao and Yan, 2010] Gao, R. X. and Yan, R. (2010). *Wavelets : Theory and applications for manufacturing*. Springer Science & Business Media. [23](#)
- [Gaymard and Tiplica, 2014] Gaymard, S. and Tiplica, T. (2014). Conditional respect towards the pedestrian : difference between men and women and risk modeling by the Bayesian approach. *Quality & Quantity*, 48(1) :91–110. [20](#)
- [Georgoulas et al., 2013] Georgoulas, G., Loutas, T., Stylios, C. D., and Kostopoulos, V. (2013). Bearing fault detection based on hybrid ensemble detector and empirical mode decomposition. *Mechanical Systems and Signal Processing*, 41(1-2) :510–525. [36](#), [38](#), [45](#), [47](#)
- [Gertler, 1997] Gertler, J. (1997). Fault detection and isolation using parity relations. *Control engineering practice*, 5(5) :653–661. [15](#)
- [Gertler et al., 1999] Gertler, J., Li, W., Huang, Y., and McAvoy, T. (1999). Isolation enhanced principal component analysis. *AIChE Journal*, 45(2) :323–334. [33](#), [36](#), [37](#), [48](#)
- [Ghate and Dudul, 2011] Ghate, V. N. and Dudul, S. V. (2011). Cascade neural-network-based fault classifier for three-phase induction motor. *Industrial Electronics, IEEE Transactions on*, 58(5) :1555–1563. [28](#)
- [Ghelam et al., 2006] Ghelam, S., Simeu-Abazi, Z., Derain, J.-P., Feuillebois, C., Vallet, S., and Glade, M. (2006). Integration of health monitoring in the avionics maintenance system. *IFAC Proceedings Volumes*, 39(13) :1449–1454. [24](#), [25](#)
- [Ghosh et al., 2011] Ghosh, K., Ng, Y. S., and Srinivasan, R. (2011). Evaluation of decision fusion strategies for effective collaboration among heterogeneous fault diagnostic methods. *Computers & chemical engineering*, 35(2) :342–355. [29](#), [36](#), [44](#), [46](#), [47](#), [48](#), [88](#)
- [Ghosh et al., 2014] Ghosh, K., Ramteke, M., and Srinivasan, R. (2014). Optimal variable selection for effective statistical process monitoring. *Computers & Chemical Engineering*, 60 :260–276. [88](#)
- [Gordon et al., 2001] Gordon, N., Doucet, A., and Freitas, J. (2001). Sequential monte carlo methods in practice. [24](#)
- [Hall and Strutt, 2003] Hall, P. and Strutt, J. (2003). Probabilistic physics-of-failure models for component reliabilities using monte carlo simulation and weibull analysis : a parametric study. *Reliability Engineering & System Safety*, 80(3) :233–242. [25](#)

- [He et al., 2014] He, B., Chen, T., and Yang, X. (2014). Root cause analysis in multivariate statistical process monitoring : Integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs. *Computers & Chemical Engineering*, 64 :167–177. [88](#)
- [Hotelling, 1933] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6) :417. [13](#), [21](#)
- [Hu et al., 2012] Hu, C., Youn, B. D., Wang, P., and Yoon, J. T. (2012). Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering & System Safety*, 103 :120–135. [45](#), [47](#), [48](#), [132](#)
- [Hu and Tse, 2013] Hu, J. and Tse, P. W. (2013). A relevance vector machine-based approach with application to oil sand pump prognostics. *Sensors*, 13(9) :12663–12686. [26](#)
- [Hu et al., 2018] Hu, M., Chen, H., Shen, L., Li, G., Guo, Y., Li, H., Li, J., and Hu, W. (2018). A machine learning bayesian network for refrigerant charge faults of variable refrigerant flow air conditioning system. *Energy and Buildings*, 158 :668–676. [20](#)
- [Hu et al., 2015] Hu, Y., Baraldi, P., Di Maio, F., and Zio, E. (2015). A particle filtering and kernel smoothing-based approach for new design component prognostics. *Reliability Engineering & System Safety*, 134 :19–31. [35](#), [37](#), [46](#)
- [Hwang and Masud, 2012] Hwang, C.-L. and Masud, A. S. M. (2012). *Multiple objective decision making—methods and applications : a state-of-the-art survey*, volume 164. Springer Science & Business Media. [73](#)
- [Hwang et al., 2010] Hwang, I., Kim, S., Kim, Y., and Seah, C. E. (2010). A survey of fault detection, isolation, and reconfiguration methods. *Control Systems Technology, IEEE Transactions on*, 18(3) :636–653. [15](#)
- [Indra et al., 2013] Indra, S., Travé-Massuyès, L., and Chanthery, E. (2013). A decentralized fault detection and isolation scheme for spacecraft : bridging the gap between model-based fault detection and isolation research and practice. *Progress in Flight Dynamics, Guidance, Navigation, Control, Fault Detection, and Avionics*, 6 :281–298. [17](#)
- [Iri et al., 1979] Iri, M., Aoki, K., O’Shima, E., and Matsuyama, H. (1979). An algorithm for diagnosis of system failures in the chemical process. *Computers & Chemical Engineering*, 3(1-4) :489–493. [13](#), [15](#), [16](#)
- [Isermann, 1984] Isermann, R. (1984). Process fault detection based on modeling and estimation methods—a survey. *Automatica*, 20(4) :387–404. [13](#), [15](#)
- [Isermann, 1993] Isermann, R. (1993). Fault diagnosis of machines via parameter estimation and knowledge processing—tutorial paper. *Automatica*, 29(4) :815–835. [15](#)
- [Isermann, 2005a] Isermann, R. (2005a). Model-based fault-detection and diagnosis—status and applications. *Annual Reviews in control*, 29(1) :71–85. [7](#)
- [Isermann, 2005b] Isermann, R. (2005b). Model-based fault-detection and diagnosis—status and applications. *Annual Reviews in control*, 29(1) :71–85. [13](#)

- [Isermann, 2006] Isermann, R. (2006). *Fault-diagnosis systems : an introduction from fault detection to fault tolerance*. Springer Science & Business Media. 7, 12, 15
- [Janusz and Venkatasubramanian, 1991] Janusz, M. E. and Venkatasubramanian, V. (1991). Automatic generation of qualitative descriptions of process trends for fault detection and diagnosis. *Engineering Applications of Artificial Intelligence*, 4(5) :329–339. 13, 22
- [Jardim-Gonçalves et al., 1996] Jardim-Gonçalves, R., Martins-Barata, M., Assis-Lopes, J. A., and Steiger-Garcão, A. (1996). Application of stochastic modelling to support predictive maintenance for industrial environments. In *Systems, Man, and Cybernetics, 1996., IEEE International Conference on*, volume 1, pages 117–122. IEEE. 26
- [Javed, 2014] Javed, K. (2014). *A robust & reliable Data-driven prognostics approach based on extreme learning machine and fuzzy clustering*. PhD thesis, Université de Franche-Comté. 155
- [Javed et al., 2015] Javed, K., Gouriveau, R., and Zerhouni, N. (2015). A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. *IEEE transactions on cybernetics*, 45(12) :2626–2639. 132
- [Jha et al., 2016] Jha, M. S., Dauphin-Tanguy, G., and Ould-Bouamama, B. (2016). Particle filter based hybrid prognostics for health monitoring of uncertain systems in bond graph framework. *Mechanical Systems and Signal Processing*, 75 :301–329. 24, 25, 108
- [Jha et al., 2009] Jha, S., Li, W., and Seshia, S. A. (2009). Localizing transient faults using dynamic bayesian networks. In *High Level Design Validation and Test Workshop, 2009. HLDVT 2009. IEEE International*, pages 82–87. IEEE. 19
- [Ji-Gong et al., 2015] Ji-Gong, L., Jie-Yong, Z., Jing, Y., Jia, L., Fan-Lin, Z., and Li, Y. (2015). Localizing multiple odor sources with a mobile robot in time-varying airflow environments using dempster-shafer inference. In *Control Conference (CCC), 2015 34th Chinese*, pages 6072–6077. IEEE. 40
- [Jing and Hou, 2015] Jing, C. and Hou, J. (2015). Svm and pca based fault classification approaches for complicated industrial process. *Neurocomputing*, 167 :636–642. 88, 104
- [Joe Qin, 2003] Joe Qin, S. (2003). Statistical process monitoring : basics and beyond. *Journal of chemometrics*, 17(8-9) :480–502. 7
- [Jozefowicz, 2013] Jozefowicz, N. (2013). *Optimisation combinatoire multi-objectif : des méthodes aux problèmes, de la Terre à (presque) la Lune*. PhD thesis, Institut National Polytechnique de Toulouse (INP Toulouse). 73
- [Julier and Uhlmann, 1996] Julier, S. J. and Uhlmann, J. K. (1996). A general method for approximating nonlinear transformations of probability distributions. Technical report, Technical report, Robotics Research Group, Department of Engineering Science, University of Oxford. 15
- [Kaftandjian et al., 2003] Kaftandjian, V., Dupuis, O., Babot, D., and Zhu, Y. M. (2003). Uncertainty modelling using dempster-shafer theory for improving detection of weld defects. *Pattern recognition letters*, 24(1-3) :547–564. 40, 41



- [Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1) :35–45. [13](#), [15](#)
- [Keliris et al., 2015] Keliris, C., Polycarpou, M. M., and Parisini, T. (2015). A robust nonlinear observer-based approach for distributed fault detection of input–output interconnected systems. *Automatica*, 53 :408–415. [17](#)
- [Khawaja et al., 2005] Khawaja, T., Vachtsevanos, G., and Wu, B. (2005). Reasoning about uncertainty in prognosis : a confidence prediction neural network approach. In *Fuzzy Information Processing Society, 2005. NAFIPS 2005. Annual Meeting of the North American*, pages 7–12. IEEE. [27](#)
- [Khelif et al., 2017] Khelif, R., Chebel-Morello, B., Malinowski, S., Laajili, E., Fnaiech, F., and Zerhouni, N. (2017). Direct remaining useful life estimation based on support vector regression. *IEEE Transactions on Industrial Electronics*, 64(3) :2276–2285. [132](#)
- [Kourti and MacGregor, 1996] Kourti, T. and MacGregor, J. F. (1996). Multivariate SPC methods for process and product monitoring. *Journal of quality technology*, 28(4) :409–428. [21](#)
- [Kourti et al., 1995] Kourti, T., Nomikos, P., and MacGregor, J. F. (1995). Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. *Journal of process control*, 5(4) :277–284. [22](#)
- [Kramer, 1992] Kramer, M. A. (1992). Autoassociative neural networks. *Computers & chemical engineering*, 16(4) :313–328. [21](#)
- [Kuncheva, 2002] Kuncheva, L. I. (2002). Switching between selection and fusion in combining classifiers : An experiment. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 32(2) :146–156. [39](#)
- [Kuncheva, 2004] Kuncheva, L. I. (2004). *Combining pattern classifiers : methods and algorithms*. John Wiley & Sons. [36](#)
- [Kyriazis and Mathioudakis, 2009] Kyriazis, A. and Mathioudakis, K. (2009). Gas turbine fault diagnosis using fuzzy-based decision fusion. *Journal of propulsion and power*, 25(2) :335–343. [36](#)
- [Lam and Suen, 1997] Lam, L. and Suen, S. (1997). Application of majority voting to pattern recognition : an analysis of its behavior and performance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 27(5) :553–568. [38](#)
- [Langley and Sage, 1994] Langley, P. and Sage, S. (1994). Induction of selective bayesian classifiers. In *Uncertainty Proceedings 1994*, pages 399–406. Elsevier. [19](#)
- [Lau et al., 2013] Lau, C., Ghosh, K., Hussain, M., and Hassan, C. C. (2013). Fault diagnosis of tennessee eastman process with multi-scale pca and anfis. *Chemometrics and Intelligent Laboratory Systems*, 120 :1–14. [88](#)
- [Le Son et al., 2013] Le Son, K., Fouladirad, M., Barros, A., Levrat, E., and Iung, B. (2013). Remaining useful life estimation based on stochastic deterioration models : A comparative study. *Reliability Engineering & System Safety*, 112 :165–175. [108](#)
- [Lee et al., 2005] Lee, D. S., Park, J. M., and Vanrolleghem, P. A. (2005). Adaptive multiscale principal component analysis for on-line monitoring of a sequencing batch reactor. *Journal of Biotechnology*, 116(2) :195–210. [21](#)

- [Lee et al., 2004] Lee, G., Han, C., and Yoon, E. S. (2004). Multiple-fault diagnosis of the tennessee eastman process based on system decomposition and dynamic pls. *Industrial & engineering chemistry research*, 43(25) :8037–8048. [88](#), [91](#)
- [Lei et al., 2018] Lei, Y., Li, N., Guo, L., Li, N., Yan, T., and Lin, J. (2018). Machinery health prognostics : A systematic review from data acquisition to rul prediction. *Mechanical Systems and Signal Processing*, 104 :799–834. [7](#), [108](#)
- [Lei et al., 2010] Lei, Y., Zuo, M. J., He, Z., and Zi, Y. (2010). A multidimensional hybrid intelligent method for gear fault diagnosis. *Expert Systems with Applications*, 37(2) :1419–1430. [36](#), [48](#)
- [Li and Ray, 1995] Li, C.-J. and Ray, A. (1995). Neural network representation of fatigue damage dynamics. *Smart Materials and Structures*, 4(2) :126. [27](#)
- [Li et al., 2011] Li, G., Alcalá, C. F., Qin, S. J., and Zhou, D. (2011). Generalized reconstruction-based contributions for output-relevant fault diagnosis with application to the tennessee eastman process. *IEEE transactions on control systems technology*, 19(5) :1114–1127. [88](#)
- [Li et al., 2008] Li, J., Jin, J., and Shi, J. (2008). Causation-based t 2 decomposition for multivariate process monitoring and diagnosis. *Journal of Quality Technology*, 40(1) :46–58. [20](#)
- [Li et al., 2018] Li, X., Ding, Q., and Sun, J.-Q. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering & System Safety*, 172 :1–11. [27](#)
- [Li et al., 2013] Li, X., Qian, J., and Wang, G.-g. (2013). Fault prognostic based on hybrid method of state judgment and regression. *Advances in Mechanical Engineering*, 5 :149562. [26](#), [155](#)
- [Li et al., 2017] Li, Z., Wu, D., Hu, C., and Terpenney, J. (2017). An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction. *Reliability Engineering & System Safety*. [45](#), [47](#), [48](#)
- [Lin et al., 2018] Lin, S., Chen, X., and Wang, Q. (2018). Fault diagnosis model based on bayesian network considering information uncertainty and its application in traction power supply system. *IEEJ Transactions on Electrical and Electronic Engineering*. [20](#)
- [Lin et al., 2000] Lin, W., Qian, Y., and Li, X. (2000). Nonlinear dynamic principal component analysis for on-line process monitoring and diagnosis. *Computers & Chemical Engineering*, 24(2) :423–429. [88](#)
- [Lin et al., 2013] Lin, Y., Chen, M., and Zhou, D. (2013). Online probabilistic operational safety assessment of multi-mode engineering systems using bayesian methods. *Reliability Engineering & System Safety*, 119 :150–157. [132](#)
- [Liu et al., 2012] Liu, D., Pang, J., Zhou, J., and Peng, Y. (2012). Data-driven prognostics for lithium-ion battery based on gaussian process regression. In *Prognostics and System Health Management (PHM), 2012 IEEE Conference on*, pages 1–5. IEEE. [26](#)
- [Liu et al., 2010] Liu, J., Saxena, A., Goebel, K., Saha, B., and Wang, W. (2010). An adaptive recurrent neural network for remaining useful life prediction of lithium-ion batteries. Technical report, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION MOFFETT FIELD CA AMES RESEARCH CENTER. [27](#)

- [Liu et al., 2015] Liu, X., Gao, X., and Han, J. (2015). Robust unknown input observer based fault detection for high-order multi-agent systems with disturbances. *ISA transactions*. 14
- [Liu et al., 2009] Liu, X., Kruger, U., Littler, T., Xie, L., and Wang, S. (2009). Moving window kernel PCA for adaptive monitoring of nonlinear processes. *Chemometrics and Intelligent Laboratory Systems*, 96(2) :132–143. 21
- [Ljung, 1979] Ljung, L. (1979). Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1) :36–50. 15
- [López-Estrada et al., 2016] López-Estrada, F. R., Ponsart, J.-C., Theilliol, D., Zhang, Y., and Astorga-Zaragoza, C.-M. (2016). LPV Model-Based Tracking Control and Robust Sensor Fault Diagnosis for a Quadrotor UAV. *Journal of Intelligent & Robotic Systems*, pages 1–15. 17
- [Lu et al., 2009] Lu, J., Bai, C., and Zhang, G. (2009). Cost-benefit factor analysis in e-services using bayesian networks. *Expert Systems with Applications*, 36(3) :4617–4625. 20
- [Lu et al., 2003] Lu, N., Wang, F., and Gao, F. (2003). Combination method of principal component and wavelet analysis for multivariate process monitoring and fault diagnosis. *Industrial & engineering chemistry research*, 42(18) :4198–4207. 21, 29
- [Lucas, 2001] Lucas, P. J. (2001). Bayesian model-based diagnosis. *International Journal of Approximate Reasoning*, 27(2) :99–119. 39
- [Luenberger, 1964] Luenberger, D. G. (1964). Observing the state of a linear system. *IEEE transactions on military electronics*, 8(2) :74–80. 13, 14
- [Luenberger, 1966] Luenberger, D. G. (1966). Observers for multivariable systems. *Automatic Control, IEEE Transactions on*, 11(2) :190–197. 13, 14
- [Mallat, 1989] Mallat, S. G. (1989). A theory for multiresolution signal decomposition : the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7) :674–693. 13, 23
- [Marler and Arora, 2004] Marler, R. T. and Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6) :369–395. 74
- [Masegosa and Moral, 2013] Masegosa, A. R. and Moral, S. (2013). An interactive approach for bayesian network learning using domain/expert knowledge. *International Journal of Approximate Reasoning*, 54(8) :1168–1181. 21
- [Maurya et al., 2005] Maurya, M. R., Rengaswamy, R., and Venkatasubramanian, V. (2005). Fault diagnosis by qualitative trend analysis of the principal components. *Chemical Engineering Research and Design*, 83(9) :1122–1132. 22
- [Maurya et al., 2007] Maurya, M. R., Rengaswamy, R., and Venkatasubramanian, V. (2007). A signed directed graph and qualitative trend analysis-based framework for incipient fault diagnosis. *Chemical Engineering Research and Design*, 85(10) :1407–1422. 31, 37, 46, 48, 88



- [McLachlan, 2004] McLachlan, G. (2004). *Discriminant analysis and statistical pattern recognition*, volume 544. John Wiley & Sons. [22](#)
- [Mechri et al., 2017] Mechri, W., Vu, H.-C., Do, P., Klingelschmidt, T., Peysson, F., and Theilliol, D. (2017). A study on health diagnosis and prognosis of an industrial diesel motor : Hidden markov models and particle filter approach. In *International Conference on Diagnostics of Processes and Systems*, pages 380–389. Springer. [24](#)
- [Medjaher, 2014] Medjaher, K. (2014). *Contribution au pronostic de défaillances guidé par des données*. PhD thesis, Université de Franche-Comté. [7](#), [111](#), [131](#)
- [Medjaher et al., 2009] Medjaher, K., Gouriveau, R., and Zerhouni, N. (2009). A procedure for failure prognostic in dynamic systems. *IFAC Proceedings Volumes*, 42(4) :1597–1602. [35](#), [37](#), [46](#)
- [Miettinen et al., 2008] Miettinen, K., Ruiz, F., and Wierzbicki, A. P. (2008). Introduction to multiobjective optimization : interactive approaches. In *Multiobjective Optimization*, pages 27–57. Springer. [74](#)
- [Mohanty et al., 2009] Mohanty, S., Das, S., Chattopadhyay, A., and Peralta, P. (2009). Gaussian process time series model for life prognosis of metallic structures. *Journal of Intelligent Material Systems and Structures*, 20(8) :887–896. [26](#)
- [Montes de Oca et al., 2014] Montes de Oca, S., Tornil-Sin, S., Puig, V., and Theilliol, D. (2014). Fault-tolerant control design using the linear parameter varying approach. *International Journal of Robust and Nonlinear Control*, 24(14) :1969–1988. [17](#)
- [Moraru et al., 2018] Moraru, L., Obreja, C. D., Dey, N., and Ashour, A. S. (2018). Dempster-shafer fusion for effective retinal vessels’ diameter measurement. In *Soft Computing Based Medical Image Analysis*, pages 149–160. Elsevier. [40](#)
- [Mosallam et al., 2016] Mosallam, A., Medjaher, K., and Zerhouni, N. (2016). Data-driven prognostic method based on bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, 27(5) :1037–1048. [132](#)
- [Mosterman and Biswas, 1999] Mosterman, P. J. and Biswas, G. (1999). Diagnosis of continuous valued systems in transient operating regions. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 29(6) :554–565. [13](#), [16](#)
- [Mueller-Gritschneider et al., 2009] Mueller-Gritschneider, D., Graeb, H., and Schlichtmann, U. (2009). A successive approach to compute the bounded pareto front of practical multiobjective optimization problems. *SIAM Journal on Optimization*, 20(2) :915–934. [73](#)
- [Murphy et al., 2001] Murphy, K. et al. (2001). The bayes net toolbox for matlab. *Computing science and statistics*, 33(2) :1024–1034. [76](#)
- [Murphy, 2002] Murphy, K. P. (2002). *Dynamic bayesian networks : representation, inference and learning*. PhD thesis, University of California, Berkeley. [19](#)
- [Mylaraswamy and Venkatasubramanian, 1997] Mylaraswamy, D. and Venkatasubramanian, V. (1997). A hybrid framework for large scale process fault diagnosis. *Computers & Chemical Engineering*, 21 :S935–S940. [36](#), [38](#), [41](#), [46](#), [47](#), [48](#)

- [Nectoux et al., 2012] Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N., and Varnier, C. (2012). Pronostia : An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM'12.*, pages 1–8. IEEE Catalog Number : CPF12PHM-CDR. [132](#)
- [Ng and Srinivasan, 2010] Ng, Y. S. and Srinivasan, R. (2010). Multi-agent based collaborative fault detection and identification in chemical processes. *Engineering Applications of Artificial Intelligence*, 23(6) :934–949. [36](#), [48](#)
- [Niu et al., 2008] Niu, G., Lee, S.-S., Yang, B.-S., and Lee, S.-J. (2008). Decision fusion system for fault diagnosis of elevator traction machine. *Journal of Mechanical Science and Technology*, 22(1) :85–95. [36](#), [45](#), [47](#), [48](#)
- [Orchard et al., 2010] Orchard, M., Tang, L., Saha, B., Goebel, K., and Vachtsevanos, G. (2010). Risk-sensitive particle-filtering-based prognosis framework for estimation of remaining useful life in energy storage devices. *Studies in Informatics and Control*, 19(3) :209–218. [36](#)
- [Parikh et al., 2001] Parikh, C. R., Pont, M. J., and Jones, N. B. (2001). Application of dempster–shafer theory in condition monitoring applications : a case study. *Pattern Recognition Letters*, 22(6-7) :777–785. [40](#), [41](#)
- [Pashazadeh et al., 2018] Pashazadeh, V., Salmasi, F. R., and Araabi, B. N. (2018). Data driven sensor and actuator fault detection and isolation in wind turbine using classifier fusion. *Renewable Energy*, 116 :99–106. [42](#), [46](#), [47](#)
- [Patton and Chen, 1997] Patton, R. J. and Chen, J. (1997). Observer-based fault detection and isolation : robustness and applications. *Control Engineering Practice*, 5(5) :671–682. [14](#), [28](#)
- [Pawson et al., 2017] Pawson, S. M., Marcot, B. G., and Woodberry, O. G. (2017). Predicting forest insect flight activity : A bayesian network approach. *PloS one*, 12(9) :e0183464. [20](#)
- [Paynter, 1961] Paynter, H. M. (1961). *Analysis and design of engineering systems*. MIT press. [13](#), [16](#)
- [Pearl, 1988] Pearl, J. (1988). Morgan kaufmann series in representation and reasoning. probabilistic reasoning in intelligent systems : Networks of plausible inference. [13](#), [18](#)
- [Pearl, 2014] Pearl, J. (2014). *Probabilistic reasoning in intelligent systems : networks of plausible inference*. Morgan Kaufmann. [18](#), [23](#)
- [Pearson, 1901] Pearson, K. (1901). Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11) :559–572. [13](#), [21](#)
- [Peng et al., 2012] Peng, Y., Wang, H., Wang, J., Liu, D., and Peng, X. (2012). A modified echo state network based remaining useful life estimation approach. In *Prognostics and Health Management (PHM), 2012 IEEE Conference on*, pages 1–7. IEEE. [132](#), [155](#)
- [Phillpotts, 2007] Phillpotts, D. (2007). Nonlinear Fault Detection and Diagnosis Using Kernel PCA applied to a Pilot Distillation Column. *Master's thesis, university of Pretoria*. [21](#)

- [Pons et al., 2015] Pons, R., Subias, A., and Travé-Massuyès, L. (2015). Iterative hybrid causal model based diagnosis : Application to automotive embedded functions. *Engineering Applications of Artificial Intelligence*, 37 :319–335. [17](#)
- [Portnoy et al., 2016] Portnoy, I., Melendez, K., Pinzon, H., and Sanjuan, M. (2016). An improved weighted recursive PCA algorithm for adaptive fault detection. *Control Engineering Practice*, 50 :69–83. [21](#), [23](#)
- [Prakash et al., 2018] Prakash, O., Samantaray, A. K., and Bhattacharyya, R. (2018). Model-based multi-component adaptive prognosis for hybrid dynamical systems. *Control Engineering Practice*, 72 :1–18. [25](#)
- [Qian et al., 2017] Qian, Y., Yan, R., and Gao, R. X. (2017). A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mechanical Systems and Signal Processing*, 83 :549–567. [24](#), [25](#), [108](#)
- [Qiu et al., 2006] Qiu, H., Lee, J., Lin, J., and Yu, G. (2006). Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of sound and vibration*, 289(4-5) :1066–1090. [132](#)
- [Ramasso, 2014] Ramasso, E. (2014). Investigating computational geometry for failure prognostics. *International Journal of Prognostics and Health Management*, 5(1) :005. [153](#), [155](#)
- [Ramasso and Gouriveau, 2010] Ramasso, E. and Gouriveau, R. (2010). Prognostics in switching systems : Evidential markovian classification of real-time neuro-fuzzy predictions. In *Prognostics and Health Management Conference, 2010. PHM'10.*, pages 1–10. IEEE. [108](#)
- [Ramasso and Gouriveau, 2014] Ramasso, E. and Gouriveau, R. (2014). Remaining useful life estimation by classification of predictions based on a neuro-fuzzy system and theory of belief functions. *IEEE Transactions on Reliability*, 63(2) :555–566. [132](#)
- [Ramasso et al., 2013a] Ramasso, E., Rombaut, M., and Zerhouni, N. (2013a). Joint prediction of continuous and discrete states in time-series based on belief functions. *IEEE transactions on cybernetics*, 43(1) :37–50. [132](#)
- [Ramasso et al., 2013b] Ramasso, E., Rombaut, M., and Zerhouni, N. (2013b). Joint prediction of observations and states in time-series : a partially supervised prognostics approach based on belief functions and knn. *networks*, 4 :5. [155](#)
- [Ramasso and Saxena, 2014] Ramasso, E. and Saxena, A. (2014). Performance benchmarking and analysis of prognostic methods for cmapps datasets. *International Journal of Prognostics and Health Management*, 5(2) :1–15. [108](#), [132](#), [134](#)
- [Ren et al., 2017] Ren, L., Cui, J., Sun, Y., and Cheng, X. (2017). Multi-bearing remaining useful life collaborative prediction : A deep learning approach. *Journal of Manufacturing Systems*, 43 :248–256. [27](#)
- [Renooij, 2001] Renooij, S. (2001). Probability elicitation for belief networks : issues to consider. *The Knowledge Engineering Review*, 16(3) :255–269. [67](#)
- [Rizvi and Nasrabadi, 2003] Rizvi, S. A. and Nasrabadi, N. M. (2003). Fusion of flir automatic target recognition algorithms. *Information Fusion*, 4(4) :247–258. [39](#)

- [Rodrigues et al., 2000] Rodrigues, M. A., Liu, Y., Bottaci, L., and Rigas, D. I. (2000). Learning and diagnosis in manufacturing processes through an executable bayesian network. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pages 390–396. Springer. 39
- [Rosenblatt, 1958] Rosenblatt, F. (1958). The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6) :386. 13, 21
- [Samantaray and Bouamama, 2008] Samantaray, A. K. and Bouamama, B. O. (2008). *Model-based process supervision : a bond graph approach*. Springer Science & Business Media. 16, 17
- [Saxena et al., 2008a] Saxena, A., Celaya, J., Balaban, E., Goebel, K., Saha, B., Saha, S., and Schwabacher, M. (2008a). Metrics for evaluating performance of prognostic techniques. In *Prognostics and health management, 2008. phm 2008. international conference on*, pages 1–17. IEEE. 49, 117
- [Saxena et al., 2008b] Saxena, A., Goebel, K., Simon, D., and Eklund, N. (2008b). Damage propagation modeling for aircraft engine run-to-failure simulation. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pages 1–9. IEEE. 132, 133
- [Schmidhuber, 2015] Schmidhuber, J. (2015). Deep learning in neural networks : An overview. *Neural networks*, 61 :85–117. 27
- [Schubert et al., 2011] Schubert, U., Kruger, U., Arellano-Garcia, H., de Sá Feital, T., and Wozny, G. (2011). Unified model-based fault diagnosis for three industrial application studies. *Control Engineering Practice*, 19(5) :479–490. 32, 37, 46, 48
- [Seeger, 2004] Seeger, M. (2004). Gaussian processes for machine learning. *International journal of neural systems*, 14(02) :69–106. 13, 26
- [Sesen et al., 2013] Sesen, M. B., Nicholson, A. E., Banares-Alcantara, R., Kadir, T., and Brady, M. (2013). Bayesian networks for clinical decision support in lung cancer care. *PloS one*, 8(12) :e82349. 20
- [Shafer, 1976] Shafer, G. (1976). *A mathematical theory of evidence*, volume 42. Princeton university press. 40
- [Shao and Nezu, 2000] Shao, Y. and Nezu, K. (2000). Prognosis of remaining bearing life using neural networks. *Proceedings of the Institution of Mechanical Engineers, Part I : Journal of Systems and Control Engineering*, 214(3) :217–230. 27
- [Shapiro and Wilk, 1965] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4) :591–611. 80
- [Shi et al., 2016] Shi, J., Li, Y., WAnG, G., and Li, X. (2016). Health index synthetization and remaining useful life estimation for turbofan engines based on run-to-failure datasets synteza wskaźników stanu technicznego oraz ocena pozostałego okresu użytkowania silników turbowentylatorowych z wykorzystaniem zbiorów danych o pracy do czasu uszkodzenia. *Eksploatacja i Niezawodność*, 18(4) :621. 153, 155

- [Si et al., 2011] Si, X.-S., Wang, W., Hu, C.-H., and Zhou, D.-H. (2011). Remaining useful life estimation—a review on the statistical data driven approaches. *European journal of operational research*, 213(1) :1–14. [7](#), [28](#)
- [Siswanto et al., 2016] Siswanto, J., Prabuwo, A. S., Abdullah, A., and Idrus, B. (2016). A linear model based on Kalman filter for improving neural network classification performance. *Expert Systems with Applications*, 49 :112–122. [34](#), [37](#), [46](#)
- [Sorsa et al., 1991] Sorsa, T., Koivo, H. N., and Koivisto, H. (1991). Neural networks in process fault diagnosis. *Systems, Man and Cybernetics, IEEE Transactions on*, 21(4) :815–825. [21](#)
- [Sotomayor and Odloak, 2005] Sotomayor, O. A. and Odloak, D. (2005). Observer-based fault diagnosis in chemical plants. *Chemical Engineering Journal*, 112(1) :93–108. [14](#), [31](#)
- [Soualhi et al., 2015] Soualhi, A., Medjaher, K., and Zerhouni, N. (2015). Bearing health monitoring based on hilbert–huang transform, support vector machine, and regression. *IEEE Transactions on Instrumentation and Measurement*, 64(1) :52–62. [26](#)
- [Specht, 1990] Specht, D. F. (1990). Probabilistic neural networks. *Neural networks*, 3(1) :109–118. [27](#), [108](#)
- [Sun et al., 2006] Sun, S., Zhang, C., and Yu, G. (2006). A bayesian network approach to traffic flow forecasting. *IEEE Transactions on intelligent transportation systems*, 7(1) :124–132. [20](#)
- [Tabassian et al., 2012] Tabassian, M., Ghaderi, R., and Ebrahimpour, R. (2012). Combination of multiple diverse classifiers using belief functions for handling data with imperfect labels. *Expert systems with applications*, 39(2) :1698–1707. [36](#), [41](#), [48](#)
- [Talbi, 2009] Talbi, E.-G. (2009). *Metaheuristics : from design to implementation*, volume 74. John Wiley & Sons. [76](#)
- [Tateno et al., 2006] Tateno, S., Matsuyama, H., and Tsuge, Y. (2006). Fault diagnosis method using a signed digraph for multiple origins of failures-evaluation of the diagnosis accuracy. In *Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006 IEEE*, pages 3271–3276. IEEE. [16](#)
- [Theilliol et al., 2002] Theilliol, D., Noura, H., and Ponsart, J.-C. (2002). Fault diagnosis and accommodation of a three-tank system based on analytical redundancy. *ISA transactions*, 41(3) :365–382. [88](#)
- [Tidriri et al., 2016] Tidriri, K., Chatti, N., Verron, S., and Tiplica, T. (2016). Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring : A review of researches and future challenges. *Annual Reviews in Control*, 42 :63–81. [13](#), [17](#), [29](#), [46](#), [48](#)
- [Tidriri et al., 2018] Tidriri, K., Chatti, N., Verron, S., and Tiplica, T. (2018). Model-based fault detection and diagnosis of complex chemical processes : A case study of the tennessee eastman process. *Proceedings of the Institution of Mechanical Engineers, Part I : Journal of Systems and Control Engineering (Under Publication)*. [99](#), [104](#)



- [Travé-Massuyès, 2014a] Travé-Massuyès, L. (2014a). Bridges between diagnosis theories from control and ai perspectives. In *Intelligent Systems in Technical and Medical Diagnostics*, pages 3–28. Springer. 7
- [Travé-Massuyès, 2014b] Travé-Massuyès, L. (2014b). Bridging control and artificial intelligence theories for diagnosis : A survey. *Engineering Applications of Artificial Intelligence*, 27 :1–16. 7
- [Van Overschee and De Moor, 2012] Van Overschee, P. and De Moor, B. (2012). *Subspace identification for linear systems : Theory—Implementation—Applications*. Springer Science & Business Media. 23, 31
- [Vedam et al., 1999] Vedam, H., Dash, S., and Venkatasubramanian, V. (1999). An intelligent operator decision support system for abnormal situation management. *Computers & Chemical Engineering*, 23 :S577–S580. 42, 46, 47
- [Venkatasubramanian et al., 2003a] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., and Yin, K. (2003a). A review of process fault detection and diagnosis : Part iii : Process history based methods. *Computers & chemical engineering*, 27(3) :327–346. 7, 13, 48
- [Venkatasubramanian et al., 2003b] Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S. N. (2003b). A review of process fault detection and diagnosis : Part i : Quantitative model-based methods. *Computers & chemical engineering*, 27(3) :293–311. 13, 27, 28, 29, 49
- [Venkatasubramanian et al., 2003c] Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S. N. (2003c). A review of process fault detection and diagnosis : Part ii : Qualitative models and search strategies. *Computers & chemical engineering*, 27(3) :313–326. 13
- [Verron, 2007] Verron, S. (2007). *Diagnostic et surveillance des processus complexes par réseaux bayésiens*. PhD thesis, Université d’Angers. 20, 104
- [Verron et al., 2010a] Verron, S., Li, J., and Tiplica, T. (2010a). Fault detection and isolation of faults in a multivariate process with Bayesian network. *Journal of Process Control*, 20(8) :902–911. 20, 87, 88
- [Verron et al., 2010b] Verron, S., Li, J., and Tiplica, T. (2010b). Fault detection and isolation of faults in a multivariate process with Bayesian network. *Journal of Process Control*, 20(8) :902–911. 50
- [Verron et al., 2006] Verron, S., Tiplica, T., and Kobi, A. (2006). Fault diagnosis with bayesian networks : Application to the tennessee eastman process. In *Industrial Technology, 2006. ICIT 2006. IEEE International Conference on*, pages 98–103. IEEE. 88
- [Verron et al., 2008a] Verron, S., Tiplica, T., and Kobi, A. (2008a). Fault detection and identification with a new feature selection based on mutual information. *Journal of Process Control*, 18(5) :479–490. 20
- [Verron et al., 2008b] Verron, S., Tiplica, T., and Kobi, A. (2008b). Fault detection and identification with a new feature selection based on mutual information. *Journal of Process Control*, 18(5) :479–490. 100

- [Verron et al., 2010c] Verron, S., Tiplica, T., and Kobi, A. (2010c). Fault diagnosis of industrial systems by conditional Gaussian network including a distance rejection criterion. *Engineering applications of artificial intelligence*, 23(7) :1229–1235. [20](#)
- [Wang and Wang, 2000] Wang, H. F. and Wang, J. P. (2000). Fault diagnosis theory : Method and application based on multisensor data fusion. *Journal of Testing and Evaluation*, 28(6) :513–518. [40](#), [41](#)
- [Wang and Qin, 2002] Wang, J. and Qin, S. J. (2002). A new subspace identification approach based on principal component analysis. *Journal of Process Control*, 12(8) :841–855. [33](#)
- [Wang et al., 2014] Wang, K.-J., Makond, B., and Wang, K.-M. (2014). Modeling and predicting the occurrence of brain metastasis from lung cancer by Bayesian network : a case study of Taiwan. *Computers in biology and medicine*, 47 :147–160. [20](#)
- [Wang et al., 2012] Wang, P., Youn, B. D., and Hu, C. (2012). A generic probabilistic framework for structural health prognostics and uncertainty management. *Mechanical Systems and Signal Processing*, 28 :622–637. [132](#)
- [Wang et al., 2008] Wang, T., Yu, J., Siegel, D., and Lee, J. (2008). A similarity-based prognostics approach for remaining useful life estimation of engineered systems. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pages 1–6. IEEE. [135](#)
- [Wang et al., 2004] Wang, W. Q., Golnaraghi, M. F., and Ismail, F. (2004). Prognosis of machine health condition using neuro-fuzzy systems. *Mechanical Systems and Signal Processing*, 18(4) :813–831. [27](#)
- [Weber et al., 2007] Weber, P., Theilliol, D., Aubrun, C., and Evsukoff, A. (2007). Increasing effectiveness of model-based fault diagnosis : A dynamic bayesian network design for decision making. In *In Proceedings of the Conference on Fault Detection, Supervision and Safety of Technical Processes*, pages 90–95. [34](#), [37](#)
- [Weidl et al., 2005] Weidl, G., Madsen, A. L., and Israelson, S. (2005). Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. *Computers & chemical engineering*, 29(9) :1996–2009. [20](#)
- [Wilcoxon, 1945] Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6) :80–83. [80](#)
- [Wold, 1975] Wold, H. (1975). Soft modelling by latent variables : the non-linear iterative partial least squares (nipals) approach. *Journal of Applied Probability*, 12(S1) :117–142. [13](#), [21](#)
- [Xi et al., 2013] Xi, Z., Jing, R., Wang, P., and Hu, C. (2013). A copula-based sampling method for data-driven prognostics and health management. In *Asme 2013 international design engineering technical conferences and computers and information in engineering conference*, pages V03AT03A039–V03AT03A039. American Society of Mechanical Engineers. [132](#)
- [Xie et al., 2013] Xie, L., Lin, X., and Zeng, J. (2013). Shrinking principal component analysis for enhanced process monitoring and fault isolation. *Industrial & Engineering Chemistry Research*, 52(49) :17475–17486. [21](#)

- [Xu et al., 2014] Xu, J., Wang, Y., and Xu, L. (2014). Phm-oriented integrated fusion prognostics for aircraft engines based on sensor data. *IEEE Sensors Journal*, 14(4) :1124–1132. [132](#)
- [Xu et al., 1992] Xu, L., Krzyzak, A., and Suen, C. Y. (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE transactions on systems, man, and cybernetics*, 22(3) :418–435. [38](#), [39](#)
- [Yan et al., 2004] Yan, J., Koc, M., and Lee, J. (2004). A prognostic algorithm for machine performance assessment and its application. *Production Planning & Control*, 15(8) :796–801. [26](#)
- [Yan and Cercone, 2010] Yan, L. J. and Cercone, N. (2010). Bayesian network modeling for evolutionary genetic structures. *Computers & Mathematics with Applications*, 59(8) :2541–2551. [20](#)
- [Yan et al., 2014] Yan, R., Gao, R. X., and Chen, X. (2014). Wavelets for fault diagnosis of rotary machines : A review with applications. *Signal processing*, 96 :1–15. [23](#)
- [Yang and Webb, 2009] Yang, Y. and Webb, G. I. (2009). Discretization for naive-bayes learning : managing discretization bias and variance. *Machine learning*, 74(1) :39–74. [19](#)
- [Ye et al., 2004] Ye, H., Wang, G., and Ding, S. X. (2004). A new parity space approach for fault detection based on stationary wavelet transform. *Automatic Control, IEEE Transactions on*, 49(2) :281–287. [15](#)
- [Yin et al., 2012] Yin, S., Ding, S. X., Haghani, A., Hao, H., and Zhang, P. (2012). A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process. *Journal of Process Control*, 22(9) :1567–1581. [88](#)
- [Yu and Rashid, 2013] Yu, J. and Rashid, M. M. (2013). A novel dynamic bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis. *AIChE Journal*, 59(7) :2348–2365. [19](#)
- [Yu et al., 2014] Yu, M., Wang, D., and Luo, M. (2014). Model-based prognosis for hybrid systems with mode-dependent degradation behaviors. *IEEE Transactions on Industrial Electronics*, 61(1) :546–554. [25](#)
- [Yu et al., 2012] Yu, P., Yong, X., Datong, L., and Xiyuan, P. (2012). Sensor selection with grey correlation analysis for remaining useful life evaluation. In *PHM Society Conference*. [108](#)
- [Zadeh, 1965] Zadeh, L. A. (1965). Information and control. *Fuzzy sets*, 8(3) :338–353. [13](#), [27](#)
- [Zaidi et al., 2010] Zaidi, A., Tagina, M., and Bouamama, B. O. (2010). Reliability data for improvement of decision-making in Analytical Redundancy Relations Bond Graph based diagnosis. In *Advanced Intelligent Mechatronics (AIM), 2010 IEEE/ASME International Conference on*, pages 790–795. IEEE. [34](#), [36](#), [37](#)
- [Zhang et al., 2017] Zhang, C., Lim, P., Qin, A., and Tan, K. C. (2017). Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. *IEEE transactions on neural networks and learning systems*. [27](#), [108](#)



- [Zhang and Ge, 2015] Zhang, F. and Ge, Z. (2015). Decision fusion systems for fault detection and identification in industrial processes. *Journal of Process Control*, 31 :45–54. [36](#), [44](#), [46](#), [47](#), [48](#)
- [Zhang, 2000] Zhang, G. P. (2000). Neural networks for classification : a survey. *Systems, Man, and Cybernetics, Part C : Applications and Reviews, IEEE Transactions on*, 30(4) :451–462. [21](#), [23](#)
- [Zhang et al., 1994] Zhang, Q., Basseville, M., and Benveniste, A. (1994). Early warning of slight changes in systems. *Automatica*, 30(1) :95–113. [26](#)
- [Zhang and Hoo, 2011] Zhang, X. and Hoo, K. A. (2011). Effective fault detection and isolation using bond graph-based domain decomposition. *Computers & chemical engineering*, 35(1) :132–148. [30](#), [37](#), [46](#)
- [Zhang et al., 2012] Zhang, Y., Li, S., and Teng, Y. (2012). Dynamic processes monitoring using recursive kernel principal component analysis. *Chemical Engineering Science*, 72 :78–86. [21](#)
- [Zhao et al., 2015] Zhao, Y., Wen, J., and Wang, S. (2015). Diagnostic Bayesian networks for diagnosing air handling units faults—Part II : Faults in coils and sensors. *Applied Thermal Engineering*, 90 :145–157. [20](#)
- [Zhao et al., 2013] Zhao, Y., Xiao, F., and Wang, S. (2013). An intelligent chiller fault detection and diagnosis methodology using Bayesian belief network. *Energy and Buildings*, 57 :278–288. [20](#)
- [Zhou et al., 2014] Zhou, J., Guo, A., Celler, B., and Su, S. (2014). Fault detection and identification spanning multiple processes by integrating PCA with neural network. *Applied Soft Computing*, 14 :4–11. [21](#)
- [Zio, 2012] Zio, E. (2012). Prognostics and health management of industrial equipment. *Diagnostics and prognostics of engineering systems : methods and techniques*, pages 333–356. [7](#)
- [Zionts and Wallenius, 1976] Zionts, S. and Wallenius, J. (1976). An interactive programming method for solving the multiple criteria problem. *Management science*, 22(6) :652–663. [74](#)
- [Zwingelstein, 1995] Zwingelstein, G. (1995). *Diagnostic des défaillances : théorie et pratique pour les systèmes industriels*. Hermès. [21](#)

## RÉSUMÉ

Le niveau de complexité croissant des systèmes et les exigences de performances et de sûreté de fonctionnement qui leur sont associées ont induit la nécessité de développer de nouvelles approches de surveillance.

Les travaux de cette thèse portent sur la surveillance des systèmes complexes, notamment la détection, le diagnostic et le pronostic de défauts, avec une méthodologie basée sur la fusion de décisions. L'objectif principal est de proposer une approche générique de fusion de diverses méthodes de surveillance, dont la performance serait meilleure que celles des méthodes individuelles la composant.

Pour cela, nous avons proposé une nouvelle démarche de fusion de décisions, basée sur la théorie Bayésienne. Cette démarche s'appuie sur une déduction théorique des paramètres du Réseau Bayésien en fonction des objectifs de performance à atteindre en surveillance. Le développement conduit à un problème multi-objectif sous contraintes, résolu par une approche lexicographique. La première étape se déroule hors-ligne et consiste à définir les objectifs de performance à respecter afin d'améliorer les performances globales du système. Les paramètres du réseau Bayésien permettant de respecter ces objectifs sont ensuite déduits de façon théorique. Enfin, le réseau Bayésien paramétré est utilisé en ligne afin de tester les performances de la fusion de décisions.

Cette méthodologie est adaptée et appliquée d'une part à la détection et au diagnostic, et d'autre part au pronostic. Les performances sont évaluées en termes de taux de diagnostic de défauts (FDR) et taux de fausses alarmes (FAR) pour l'étape de détection et de diagnostic, et en durée de fonctionnement avant la défaillance du système (RUL) pour le pronostic.

**mots-clés :** Détection, Diagnostic, Pronostic, Fusion de décisions, Systèmes complexes, Réseaux Bayésiens

## ABSTRACT

Nowadays, systems are becoming more and more complex and require new effective methods for their supervision. This latter comprises a monitoring phase that aims to improve the system's performances and ensure a safety production for humans and materials.

This thesis work deals with fault detection, diagnosis and prognosis, with a methodology based on decisions fusion. The main issue concerns the integration of different decisions emanating from individual monitoring methods in order to obtain more reliable results.

The methodology is based on a theoretical learning of the Bayesian network parameters, according to monitoring objectives to be reached. The development leads to a multi-objective problem under constraints, which is solved with a lexicographic approach. The first step is offline and consists of defining the objectives to be achieved in order to improve the overall performance of the system. The Bayesian network parameters respecting these objectives are then deduced theoretically. Finally, the parametrized Bayesian network is used online to test the decision fusion performances.

These performances are evaluated in terms of Fault Diagnostic Rate (FDR) and False Alarm Rate (FAR) for the detection and diagnosis stage, and in terms of Remaining Useful Life (RUL) for the prognosis.

**keywords :** Detection, Diagnosis, Prognostic, Decision fusion, Complex systems, Bayesian networks.

# ENGAGEMENT DE NON PLAGIAT

Je, soussigné(e) Khaoula TIDRIRI  
déclare être pleinement conscient(e) que le plagiat de documents ou d'une  
partie d'un document publiée sur toutes formes de support, y compris l'internet,  
constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée.  
En conséquence, je m'engage à citer toutes les sources que j'ai utilisées  
pour écrire ce rapport ou mémoire.

signé par l'étudiant(e) le **06 / 07 / 2018**



**Cet engagement de non plagiat doit être signé et joint  
à tous les rapports, dossiers, mémoires.**

Présidence de l'université  
40 rue de rennes - BP 73532  
49035 Angers cedex

Tél. 02 41 96 23 23 | Fax 02 41 96 23 00



**Titre :** Fusion de décisions dédiée à la surveillance des systèmes complexes.

**Mots clés :** Détection, Diagnostic, Pronostic, Fusion de décisions, Systèmes complexes, Réseaux Bayésiens.

**Résumé :** Le niveau de complexité croissant des systèmes et les exigences de performances et de sûreté de fonctionnement qui leur sont associées ont induit la nécessité de développer de nouvelles approches de surveillance.

Les travaux de cette thèse portent sur la surveillance des systèmes complexes, notamment la détection, le diagnostic et le pronostic de défauts, avec une méthodologie basée sur la fusion de décisions. L'objectif principal est de proposer une approche générique de fusion de diverses méthodes de surveillance, dont la performance serait meilleure que celles des méthodes individuelles la composant.

Pour cela, nous avons proposé une nouvelle démarche de fusion de décisions, basée sur la théorie Bayésienne. Cette démarche s'appuie sur une déduction théorique des paramètres du Réseau Bayésien en fonction des objectifs de performance à atteindre en surveillance.

Le développement conduit à un problème multi-objectif sous contraintes, résolu par une approche lexicographique. La première étape se déroule hors-ligne et consiste à définir les objectifs de performance à respecter afin d'améliorer les performances globales du système. Les paramètres du réseau Bayésien permettant de respecter ces objectifs sont ensuite déduits de façon théorique. Enfin, le réseau Bayésien paramétré est utilisé en ligne afin de tester les performances de la fusion de décisions.

Cette méthodologie est adaptée et appliquée d'une part à la détection et au diagnostic, et d'autre part au pronostic. Les performances sont évaluées en termes de taux de diagnostic de défauts (FDR) et taux de fausses alarmes (FAR) pour l'étape de détection et de diagnostic, et en durée de fonctionnement avant la défaillance du système (RUL) pour le pronostic.

**Title :** Decision fusion dedicated to the monitoring of complex systems.

**Keywords :** Detection, Diagnosis, Prognostic, Decision fusion, Complex systems, Bayesian networks.

**Abstract :** Nowadays, systems are becoming more and more complex and require new effective methods for their supervision. This latter comprises a monitoring phase that aims to improve the system's performances and ensure a safety production for humans and materials.

This thesis work deals with fault detection, diagnosis and prognosis, with a methodology based on decisions fusion. The main issue concerns the integration of different decisions emanating from individual monitoring methods in order to obtain more reliable results.

The methodology is based on a theoretical learning of the Bayesian network parameters, according to monitoring objectives to be reached.

The development leads to a multi-objective problem under constraints, which is solved with a lexicographic approach. The first step is offline and consists of defining the objectives to be achieved in order to improve the overall performance of the system. The Bayesian network parameters respecting these objectives are then deduced theoretically. Finally, the parametrized Bayesian network is used online to test the decision fusion performances.

These performances are evaluated in terms of Fault Diagnostic Rate (FDR) and False Alarm Rate (FAR) for the detection and diagnosis stage, and in terms of Remaining Useful Life (RUL) for the prognosis.