



# Surveillance et diagnostic de systèmes dynamiques: approches complémentaires du traitement de signal et de l'intelligence artificielle

Michèle Basseville, Marie-Odile Cordier

## ► To cite this version:

Michèle Basseville, Marie-Odile Cordier. Surveillance et diagnostic de systèmes dynamiques: approches complémentaires du traitement de signal et de l'intelligence artificielle. [Rapport de recherche] RR-2861, INRIA. 1996. <inria-00073830>

**HAL Id: inria-00073830**

**<https://hal.inria.fr/inria-00073830>**

Submitted on 24 May 2006

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

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

***Surveillance et diagnostic de systèmes dynamiques:  
approches complémentaires du traitement de signal  
et de l'intelligence artificielle***

Michèle Basseville, Marie-Odile Cordier

**N° 2861**

Avril 1996

\_\_\_\_\_ THÈMES 3 et 4 \_\_\_\_\_



***rapport  
de recherche***



## Surveillance et diagnostic de systèmes dynamiques: approches complémentaires du traitement de signal et de l'intelligence artificielle

Michèle Basseville\*, Marie-Odile Cordier†

Thèmes 3 et 4 — Interaction homme-machine,  
images, données, connaissances — Simulation et optimisation  
de systèmes complexes  
Projets AS, REPCO

Rapport de recherche n° 2861 — Avril 1996 — 40 pages

**Résumé :** La conduite et la maintenance de procédés industriels nécessitent l'analyse de leur comportement tel qu'il est perçu par les divers capteurs. Il s'agit essentiellement de détecter, localiser et diagnostiquer toute déviation par rapport à un comportement de référence. Nous présentons des méthodes statistiques du traitement du signal et des méthodes de l'intelligence artificielle permettant la génération des alarmes et leur interprétation. Nous les illustrons par des exemples tirés d'applications industrielles. Nous nous efforçons de montrer la complémentarité de ces approches.

**Mots-clé :** Surveillance, diagnostic de pannes, systèmes dynamiques, traitement du signal, intelligence artificielle, approches numériques, approches symboliques, applications industrielles.

(Abstract: *pto*)

Cet article a été écrit à l'occasion d'un tutoriel que les auteurs ont présenté lors du Congrès AFCET-INRIA *Reconnaissance des Formes et Intelligence Artificielle*, Rennes, Janvier 1996.

\* IRISA/CNRS – [basseville@irisa.fr](mailto:basseville@irisa.fr).

† IRISA/Université de Rennes I – [cordier@irisa.fr](mailto:cordier@irisa.fr).

Unité de recherche INRIA Rennes

IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)

Téléphone : (33) 99 84 71 00 – Télécopie : (33) 99 84 71 71

# Monitoring and diagnosing dynamical systems: signal processing and artificial intelligence approaches

**Abstract:** The analysis of the behavior of industrial processes as measured with various sensors is necessary for control and maintenance purposes. The basic problem is to detect, localize and diagnose any deviation with respect to a reference behavior. We describe statistical signal processing methods and artificial intelligence methods for alarm generation and interpretation. We give several examples from industrial applications. We attempt to show how each type of approaches complement the other one.

**Key-words:** Monitoring, fault diagnosis, dynamical systems, signal processing, artificial intelligence, numerical methods, symbolic methods, industrial applications.

# 1 Introduction

Nous décrivons tout d'abord le contexte dans lequel nous nous situons, et proposons une définition possible d'un certain nombre de mots-clefs utilisés dans nos communautés avec des connotations souvent différentes; en particulier, nous précisons ce que nous entendons par système dynamique, surveillance, diagnostic. Ensuite, nous décrivons les différents niveaux de traitement relatifs à la surveillance. Enfin, nous introduisons les objectifs d'une bonne surveillance.

## 1.1 Surveillance de systèmes dynamiques

Nous nous intéressons ici à la surveillance et au diagnostic de machines, de structures ou de procédés, que nous désignerons le plus souvent par le terme de *systèmes*, et qui sont essentiellement *dynamiques*.

Nous indiquons maintenant ce que nous entendons par système dynamique et les raisons qui conduisent à le surveiller.

### 1.1.1 Systèmes dynamiques: Définitions

Nous donnons ici des définitions relativement peu formelles; des formules typiques de différentes sortes de modèles sont données au tableau 1 à titre d'illustration. Nous renvoyons le lecteur par exemple à [63] pour des développements.

Nous discutons d'abord des modèles déterministes. Les modèles stochastiques correspondants s'obtiennent en y faisant intervenir des variables aléatoires communément appelées bruits. Ces bruits peuvent intervenir par exemple de manière additive (souvent) ou multiplicative sur les équations du tableau 1. Ils sont souvent supposés Gaussiens, à la fois car cela correspond à une réalité fréquente et parce que cela permet de concevoir des algorithmes simples.

Un système *dynamique* est un système dans lequel les grandeurs qui interviennent (variables d'états  $X$ , entrées  $U$ , sorties  $Y$ ) sont liées par des relations temporelles. Au contraire, dans un système statique, les liaisons entre variables sont exclusivement instantanées. Qu'ils soient statiques ou dynamiques, les modèles sont qualifiés de :

- *linéaires* ou *non-linéaires*, selon que les relations auxquelles on s'intéresse plus particulièrement sont linéaires ou non. Il faut bien prendre garde à ce qu'une grandeur d'un modèle (par exemple une sortie) peut être à la fois linéaire par rapport à une grandeur (par exemple une entrée) et non-linéaire par rapport à une autre (par exemple un état ou un paramètre);
- *continus* ou *échantillonnés* (discrétisés), selon que le temps qui y figure est une variable continue  $t$  ou une variable discrète  $k$  résultant d'un échantillonnage régulier. Il convient de noter que le temps peut aussi être donné par une suite d'événements discrets; on parle alors de *systèmes à événements discrets*. D'autre part, la discrétisation peut aussi faire référence à une discrétisation des variables plutôt que du temps.

Le *régime établi* ("steady-state" pour les Anglo-Saxons) d'un système dynamique est décrit par le modèle statique correspondant à la suppression de l'indexation par le temps des grandeurs en jeu dans le modèle dynamique.

Les modèles dynamiques usuels, c'est-à-dire résultant de l'application des grandes lois fondamentales de la physique, se présentent naturellement sous l'une ou l'autre des deux formes suivantes :

- *représentation entrées-sorties*, sans explicitation d'état; par exemple, équations de bilans de matières en chimie;
- *représentation d'état*, dans laquelle on écrit les relations dynamiques entre les entrées et les sorties en explicitant un état (multidimensionnel) évoluant lui-même de manière dynamique.

Pour ce qui est des modèles dynamiques linéaires (1), on sait passer d'une représentation à une autre [2], les représentations entrées-sorties (2) étant communément appelées ARMA<sup>1</sup>. On sait aussi, pour une large classe de modèles dynamiques non-linéaires, trouver une représentation entrées-sorties (4) correspondant à un modèle d'état (3) [64].

---

1. Auto Regressive Moving Average: autrement dit, la sortie à un instant donné est une combinaison linéaire d'un nombre fini de sorties à des instants précédents, et d'un nombre fini, éventuellement différent, d'entrées à des instants précédents.

TAB. 1 – Exemples de modèles déterministes. On distingue les modèles statiques (relations instantanées) des modèles dynamiques (liaisons temporelles). Pour chacun de ces deux types, on distingue les modèles linéaires des modèles non-linéaires. Pour chacun d'eux, on distingue les modèles avec explicitation d'un état interne  $X$  et les modèles sans explicitation d'état (appelés aussi représentations entrées  $U$ -sorties  $Y$ ). On donne à chaque fois une équation en temps discret  $k$  et une en temps continu  $t$ ; la différence essentielle étant l'introduction de différentiations successives  $Y_t^{(i)}$  à la place des décalages temporels  $Y_{k-i}$ . Les paramètres du modèle sont regroupés génériquement en  $\theta$ , qui, selon les cas, est un vecteur ou une matrice.

Modèles déterministes						
	Statiques			Dynamiques		
Linéaires	Avec	$k$	$X_{k+1} = X_k$	Avec	$k$	$\begin{cases} X_{k+1} = FX_k + GU_k \\ Y_k = HX_k + JU_k \end{cases} \quad (1)$
		état	$t$		$X'_t = 0$	état
	Sans	$k$	$Y_k = \theta U_k$	Sans	$k$	$\sum_{i=0}^p \theta_i Y_{k-i} + \sum_{j=0}^q \theta_{p+j} U_{k-j} = 0 \quad (2)$
		état	$t$		$Y_t = \theta U_t$	état
Non-linéaires	Avec	$k$	$X_{k+1} = X_k$	Avec	$k$	$\begin{cases} X_{k+1} = f(\theta, X_k, U_k) \\ Y_k = h(\theta, X_k, U_k) \end{cases}$
		état	$t$		$X'_t = 0$	état
	Sans	$k$	$g(\theta, Y_k, U_k) = 0$	Sans	$k$	$g(\theta, Y_k, Y_{k-1}, \dots, U_k, U_{k-1}, \dots) = 0$
		état	$t$		$g(\theta, Y_t, U_t) = 0$	état

On dit (informellement) qu'un système dynamique est *stable* si, lorsqu'on lui applique des entrées bornées, les sorties qui en résultent sont elles aussi bornées<sup>2</sup>. Il existe en fait plusieurs définitions formelles de la stabilité correspondant à cette notion intuitive.

On dit qu'un système dynamique est *stationnaire* si son évolution entre deux instants ne dépend que de l'intervalle de temps qui séparent ces deux instants, et pas des instants eux-mêmes.

Les *paramètres*  $\theta$  d'un modèle sont les grandeurs qui caractérisent le système. Ces grandeurs peuvent être :

- constantes, lorsque les caractéristiques du système n'évoluent pas dans le temps;
- variables lentement, par exemple lorsque le système vieillit et s'use;
- variables, parfois de manière fréquente et/ou brutale, selon le mode de fonctionnement du système, par exemple suivant la vitesse de rotation ou suivant la charge thermique d'une machine;
- variables lorsque surgit une perturbation ou une panne; on distingue alors le comportement normal du système des comportements (anormaux) résultant de ces pannes.

Au moment de la modélisation, il convient de bien prendre garde à la répartition des variables entre celles qui doivent être considérées comme variables d'état et celles qui doivent être prises comme paramètres caractéristiques. Pour une application donnée, la tâche peut être délicate : les répartitions intuitives ne sont pas forcément les plus efficaces, au sens où on peut artificiellement augmenter la difficulté d'un problème de surveillance par une mauvaise paramétrisation.

Dans la suite, on désignera par *panne* une modification indésirable du système se traduisant par une variation d'un ou de plusieurs paramètres par rapport à une valeur de référence  $\theta_0$ , c'est-à-dire d'une ou plusieurs composantes du vecteur  $\theta$ . Tout le problème de la surveillance consiste à détecter ces variations de paramètres, à distinguer celles qui résultent de pannes de celles qui résultent de comportements normaux (par exemple, on n'a souvent aucune espèce d'intérêt à détecter les changements de modes de fonctionnement d'une machine), à décider si ces variations sont effectivement significatives par rapport aux incertitudes sur le modèle et sur la référence  $\theta_0$  et aux bruits sur les données mesurées. Le problème du diagnostic est en outre de déterminer les composantes du vecteur paramètre qui sont effectivement modifiées et de trouver les causes, en termes des composants atteints, de ces variations.

### 1.1.2 Systèmes dynamiques: Pourquoi les surveiller?

Les principales raisons qui conduisent à surveiller un système sont :

- La *conduite*, qu'il s'agit d'optimiser et qui est une tâche en-ligne; par exemple, l'aide à la conduite d'un haut-fourneau doit fournir à l'opérateur de conduite les outils nécessaires à la prise de décision d'actions visant à faire fonctionner le procédé au mieux (production maximale; sécurité; non dégradation des équipements). Ceci passe par la surveillance du procédé afin de détecter toutes anomalies de fonctionnement et l'identifier aussi bien que possible.
- Un *service*, qu'il faut assurer; c'est le cas d'un réseau de télécommunications par exemple. Là aussi, la capacité à détecter, localiser et diagnostiquer les événements indésirables, et à y réagir, est un enjeu crucial.
- La *maintenance*, qui a pour l'objet le remplacement ou la réparation d'équipements usés ou défectueux, et que l'on souhaite optimiser. Elle s'effectue le plus souvent hors-ligne : La maintenance corrective intervient après la détection et la localisation d'un défaut; la maintenance préventive est elle le plus souvent systématique. Cependant, la maintenance conditionnelle, qui est une alternative à la maintenance systématique, fait l'objet d'une demande croissante dans un grand nombre d'applications industrielles; par exemple, pour les machines de production d'énergie. Cette maintenance est basée sur la surveillance en continu de l'évolution du système, afin de prévenir un dysfonctionnement avant qu'il n'arrive; elle impose donc des traitements en-ligne, au moins en partie.

Nous n'abordons pas dans la suite les enjeux économiques, technologiques et environnementaux de la surveillance, qui sont liés à des impératifs de sécurité (des hommes et des matériels) et de protection de l'environnement, et à la recherche de gains de productivité.

---

2. On dit inversement que le système est instable lorsque les sorties mesurées peuvent croître indéfiniment.



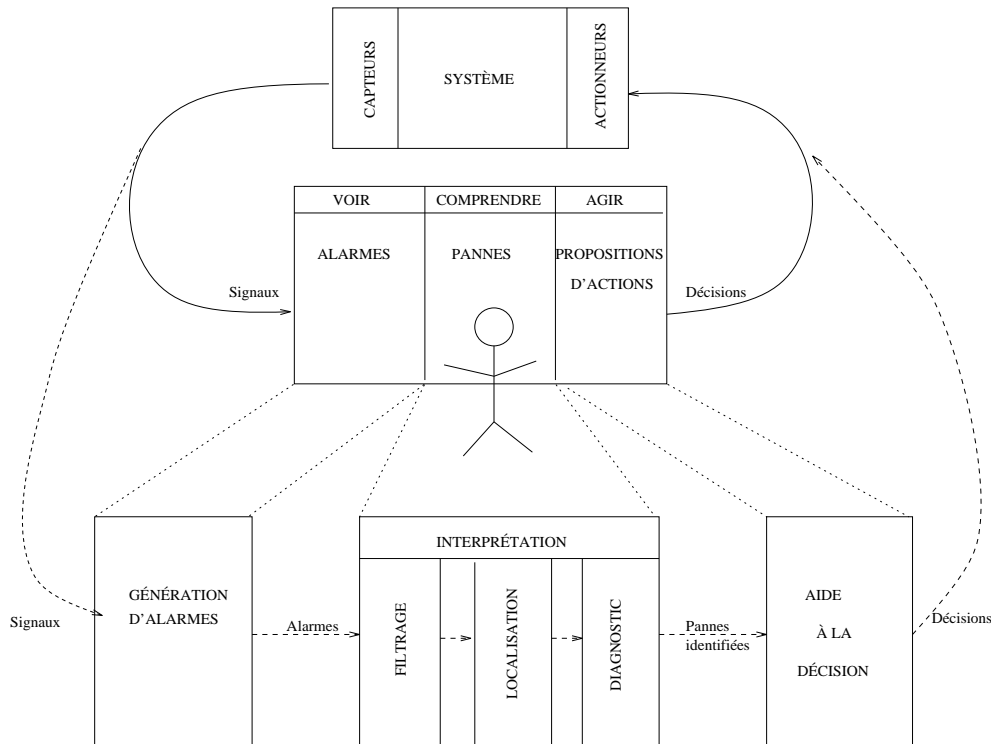


FIG. 1 – Architecture générale d’un système de surveillance en-ligne. On distingue le système muni de ses capteurs et actionneurs et l’opérateur devant son écran de contrôle. Au vu des alarmes, des pannes, des actions qui lui sont proposées, l’opérateur décide d’actions de conduite qui sont rendues effectives par l’intermédiaire d’actionneurs; ici le terme “actionneurs” est pris au sens large et représente tous les moyens permettant d’agir sur le fonctionnement du système.

Il est important de remarquer que, si conduite et maintenance sont des opérations qui s’inscrivent dans le temps de manière différente, la surveillance qu’elles impliquent peut être de la même nature (en-ligne) pour les deux objectifs. En effet, le choix d’une approche en-ligne ou hors-ligne pour la surveillance dépend non seulement de l’objectif – conduite ou maintenance, mais également des ordres de grandeur relatifs des constantes de temps du système, des différents types d’événements à surveiller, et aussi de l’échantillonnage des mesures disponibles. Nous reviendrons plus précisément sur ce point lorsque nous discuterons le problème de la détection de ces événements, et donc de la génération des alarmes.

## 1.2 Architecture générale

La figure 1 présente l’architecture générale d’un poste de conduite en-ligne. On y retrouve les trois grandes fonctions “ Voir ”, “ Comprendre ”, “ Agir ” nécessaires à une bonne surveillance. La fonction de perception s’appuie sur les données acquises par les capteurs du système à surveiller et transmet à l’opérateur des informations plus ou moins élaborées. L’opérateur analyse les informations reçues et décide des actions à entreprendre au travers d’actionneurs qui permettent d’agir sur le système. L’architecture d’un poste de maintenance hors-ligne est proche de celle-ci à la différence près que le flux d’observations ne provient pas directement des capteurs mais d’une base de données où ont été archivées, entre autres, les valeurs caractéristiques des signaux capteurs; d’autre part, les décisions prises sont des actions de remplacement ou de réparation des composants du système.

Nous nous focalisons dans la suite de cet article sur les composantes “ Voir ” et “ Comprendre ”; nous nous intéressons donc à la *détection*, à la *localisation*, et au *diagnostic* d’événements ou d’évolutions imprévus, repérés comme déviations par rapport à un état ou un comportement de référence (normal, souhaitable ou nominal). Nous excluons de notre propos la composante “ Agir ” (“ réaction ”) de la surveillance, qui conduit à la reconfiguration de lois de commande, la réparation ou la mise en oeuvre de stratégies de maintenance.

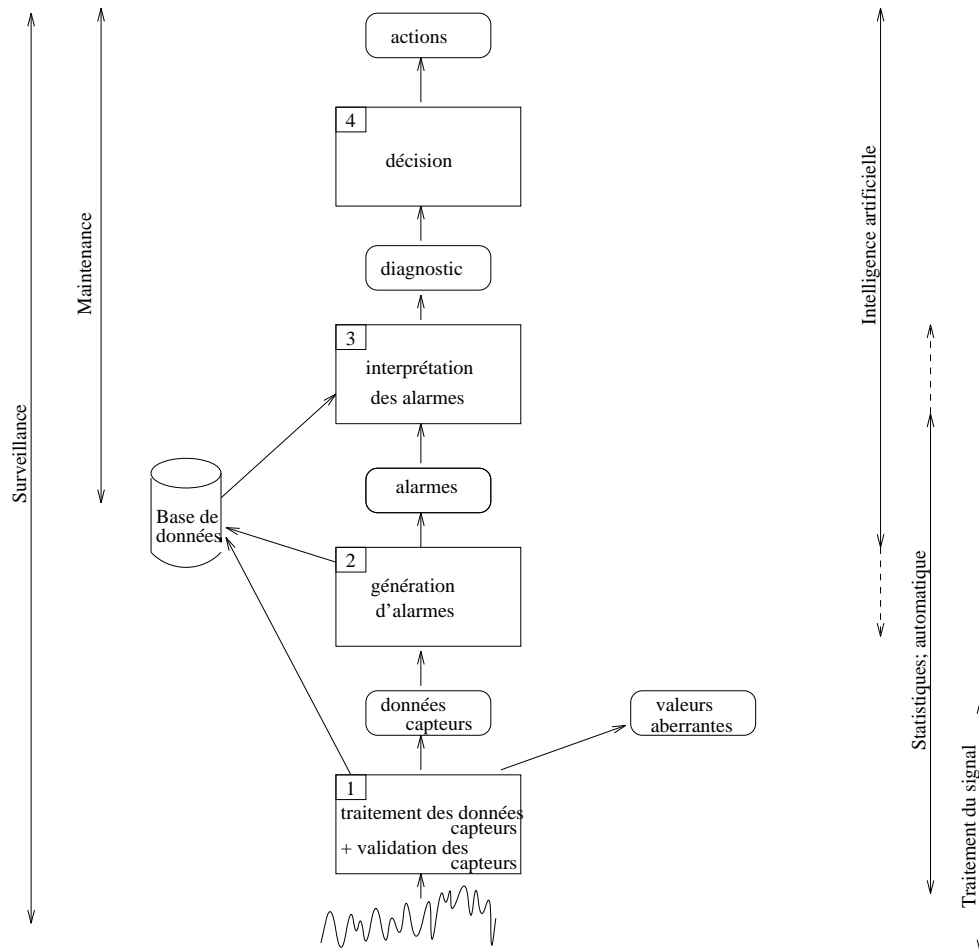


FIG. 2 – Organisation générale des traitements en-ligne. On distingue quatre niveaux de traitement permettant de passer des signaux issus des capteurs à des informations symboliques destinées à l'opérateur. Les flèches indiquent une direction privilégiée de circulation des informations, même si un point important, mais souvent mal traité, est celui du nécessaire aller-retour entre les informations symboliques et numériques qui n'est pas indiqué sur ce schéma. Dans le cas de la surveillance (en-ligne), les actions sont des actions de conduite; dans le cas de la maintenance (hors-ligne), les actions sont des actions de réparation ou de maintenance préventive.

### 1.3 Les différents niveaux de traitement : du signal à la décision

On peut représenter l'organisation générale des traitements en-ligne utiles pour l'aide à la conduite comme indiqué à la figure 2.

On distingue plusieurs niveaux de traitement possibles :

1. Le premier niveau concerne le traitement des données capteurs : il a pour première fonction la détection et l'élimination des valeurs aberrantes. Il prend aussi souvent en charge la validation des capteurs eux-même en étudiant, à l'aide de méthodes numériques, la redondance inter-capteurs.
2. Un second niveau, qui travaille toujours à partir des données capteurs, a pour fonction la *génération des alarmes*.

Les alarmes indiquent l'occurrence d'événements survenus dans le système. Ces indicateurs peuvent être de types variés :

- soit en tout-ou-rien : il s'est passé quelque chose ou pas;
- soit de type numérique discret : elles sont alors étiquetées par un numéro ou par un libellé.

La génération des alarmes peut se faire, par exemple, par segmentation des signaux traités individuellement, ou par un traitement conjoint de plusieurs signaux et utilisant un modèle du système surveillé; nous y reviendrons plus loin.

Il est important de souligner que ces deux premiers niveaux, redondance inter-capteurs et génération d'alarmes, contribuent à un codage numérique/symbolique "intelligent" des données capteurs, point crucial pour la surveillance. En particulier, il est possible de produire directement à ce niveau des *alarmes "intelligentes"* au sens où elles sont porteuses d'informations relatives au diagnostic.

3. Un troisième niveau a pour charge l'*interprétation des alarmes*. Il doit transformer les alarmes (qui signalent des événements survenus sur le système) en une information utile à l'opérateur dans sa prise de décision, que ce soit pour l'inciter à ne rien faire ou à réagir par l'intermédiaire des actionneurs. Les événements à l'origine de l'émission de ces alarmes ne sont pas forcément indésirables. Un des buts de l'interprétation est précisément de le déterminer.

On distingue le *filtrage* (élimination d'alarmes redondantes); la *localisation* d'un dysfonctionnement (en faisant par exemple le lien entre signal ou paramètre perturbé et composant en dysfonctionnement); et le *diagnostic* proprement dit (recherche de la cause profonde expliquant le phénomène observé).

4. Le dernier niveau concerne le système d'aide à la conduite à proprement parler.

Il convient de souligner que la phase d'interprétation des alarmes dépend directement de l'objectif du système de surveillance. La profondeur dans la recherche de causes est par exemple différente selon que la surveillance est effectuée dans un but d'aide à la conduite, comme ici sur cette figure, ou dans un but de maintenance conditionnelle.

En général la surveillance d'un système comprend ces trois niveaux et suppose donc une intégration de traitements de type purement numérique (niveau 1), de type mixte numérique/symbolique (niveau 2) et plutôt symbolique (niveau 3). On peut distinguer les systèmes de gestion d'alarmes qui reçoivent directement des alarmes et ne comprennent donc ni le premier ni le second niveau.

## 1.4 Objectifs d'une "bonne" surveillance

Pour surveiller en continu l'évolution du système et trouver les causes des événements indésirables, il convient d'utiliser au mieux les informations disponibles, qu'elles soient numériques (les données des divers capteurs et actionneurs), ou symboliques (historiques, environnement, événements indésirables survenus antérieurement et réparations effectuées,...). Dans l'optique d'une utilisation optimale des informations disponibles pour la surveillance, deux points nous semblent mériter d'être soulignés :

- Un premier point crucial réside dans le traitement des signaux (données des capteurs et actionneurs) mesurés sur le système, qui doivent être analysés aussi finement que possible afin de limiter au maximum les pertes d'informations et la génération d'alarmes pauvres et par là-même trop nombreuses. Il s'agit alors :
  - d'extraire des signaux des caractéristiques des objets étudiés qui soient pertinentes pour la surveillance;
  - d'être capable ensuite d'analyser ces éléments afin de détecter perturbations et anomalies; ceci en préservant une robustesse suffisante par rapport aux bruits sur le système, aux erreurs sur son modèle, et aux incertitudes sur la référence, mais aussi par rapport aux changements de modes de fonctionnement du système;
  - de marier judicieusement les informations condensées ainsi obtenues avec d'autres informations, de nature plus "discrète" ou symbolique, qui sont disponibles sur l'objet et indispensables à sa surveillance.

En particulier, que ce soit pour la conduite ou la maintenance conditionnelle, la *détection précoce de déviations faibles par rapport à une caractérisation du système en ambiance de travail normale* (sans excitation artificielle, ni ralentissement, ni arrêt) s'avère d'un grand intérêt.

- Le second point crucial est l'adéquation des réponses fournies par le système de surveillance aux besoins des opérateurs chargés de les utiliser. Il est ainsi souvent nécessaire de tenir compte du contexte général de la surveillance et du niveau des réponses attendues. Ceci explique l'intérêt d'une interprétation des

alarmes s'appuyant sur des connaissances symboliques, familières ou au moins partageables par l'opérateur. On est alors en mesure d'expliquer et/ou de justifier les jugements portés sur l'installation ou les décisions proposées. Bien entendu, ceci passe par la définition d'un interface approprié et ergonomique, point important mais en dehors du champ de cet article.

Ces deux points mettent bien en lumière une caractéristique de la surveillance, qui requiert le passage d'une information numérique à une information symbolique, et, point souvent oublié et délicat, le nécessaire aller-et-retour entre ces deux types d'informations. Ils mettent aussi en évidence la nécessité et l'importance du *cahier des charges* de la surveillance : que ce soit au niveau numérique ou au niveau symbolique, on ne peut concevoir un système de surveillance qu'en réponse à une demande<sup>3</sup> de surveillance donnée : quelles pannes doivent être détectées, localisées, diagnostiquées, sont-elles brutales ou au contraire graduelles, de quelles informations l'opérateur a besoin pour réagir ...

Un autre point en aval de la surveillance proprement dite est celui du choix du type des capteurs, de leur nombre et de leur positionnement les plus adéquats pour la surveillance. Des critères permettant de quantifier la qualité d'un jeu de capteurs pour la surveillance peuvent naturellement être associés à des algorithmes de détection. Ces critères quantitatifs permettent de qualifier un positionnement de capteurs pour la surveillance et le diagnostic, et peuvent être utilisés de deux façons : pour un positionnement de capteurs donné (on n'a pas toujours le choix!), déterminer les défauts détectables et diagnosticables; pour un ou des défauts donnés, déterminer les positionnements de capteurs qui permettront de les surveiller. Cette détermination est éventuellement une simple sélection des mesures disponibles. Il peut arriver qu'un défaut donné soit plus facilement diagnostiquable avec un sous-ensemble des mesures qu'avec toutes les mesures traitées de manière conjointe, car rajouter des capteurs permet d'accroître l'information mais aussi le bruit! Par exemple, en surveillance des vibrations, on n'a pas intérêt à traiter de manière conjointe les mesures des capteurs verticaux et horizontaux placés sur les paliers d'un groupe turbo-alternateur.

Nous nous consacrons par la suite aux niveaux 2 et 3 de la figure 2. Pour chacune de ces deux étapes : génération d'alarmes (détection) et interprétation des alarmes (filtrage, localisation et diagnostic), différentes méthodes existent, basées sur des concepts et techniques variés et s'appuyant sur différents modèles du système surveillé. Nous nous proposons de présenter des méthodes statistiques du traitement du signal et des méthodes de l'intelligence artificielle permettant la génération des alarmes et leur interprétation.

## 2 Génération des alarmes

Nous discutons maintenant la génération des alarmes.

### 2.1 Le problème

Comme on l'a déjà indiqué, elle peut intervenir à plusieurs niveaux. Soit au niveau d'un capteur, sur lequel on procède à une segmentation automatique en zones homogènes [12, 13, 16, 4, 11, 19] et à la détection d'événements indésirables, par exemple de type transitoire [80, 49, 50, 1, 3]. Ce traitement ignore le plus souvent le modèle du système surveillé dans son ensemble. Il peut être réalisé à l'aide d'algorithmes basés sur des modèles boîte-noire<sup>4</sup>, comme nous le décrivons ci-dessous. Soit au niveau de plusieurs capteurs que l'on analyse à l'aide d'algorithmes basés sur tout ou partie du modèle physique du système surveillé; même à ce niveau purement numérique de traitement des signaux, on peut alors générer des alarmes intelligentes, au sens où elles contiennent de l'information relative à un diagnostic de plus haut niveau, sous la forme des composants responsables le plus probablement de l'alarme déclenchée.

Les informations dont on dispose sur le système sont :

- des *données mesurées* ( $Y_k$ ), provenant d'un ou plusieurs capteurs ou d'actionneurs, et donc multidimensionnelles;

---

3. Cette demande peut être peu précise au départ, le cahier des charges se précisant au fur et à mesure de l'élaboration du système de surveillance, et ... réciproquement!

4. Par exemple, un algorithme de segmentation spectrale reposant sur une modélisation auto-régressive (AR) peut s'avérer intéressant sur des signaux pour lesquels on n'a pas étudié la pertinence physique d'une telle modélisation. Le modèle est alors jugé en fonction du service rendu par l'algorithme qu'il sous-tend, plus que par sa capacité à représenter les phénomènes en jeu.

- le cas échéant, un ou plusieurs *modèles*<sup>5</sup>, éventuellement de natures différentes : modèles fonctionnels, structuraux, comportementaux;
- le cas échéant aussi, des indications sur la façon dont les événements que l'on souhaite détecter influent sur le modèle; par exemple, de manière additive ou multiplicative sur les équations des modèles donnés dans le tableau 1. Ces indications peuvent avoir des répercussions sur la manière de transformer les données en vue de la génération d'alarmes.

## 2.2 Deux types de modèles: le modèle boîte-noire et le modèle physique

La génération d'alarmes repose le plus souvent sur l'utilisation de modèles numériques. Nous distinguons en général deux types de modélisation : la *modélisation physique* qui, plus ou moins finement permet d'expliquer voire de simuler les phénomènes en jeu dans l'installation, et la *modélisation boîte-noire* qui a pour particularité de pouvoir être utilisée " en aveugle " sur les données, sans considération physique. Il convient de souligner qu'une surveillance peut être effectuée même en l'absence d'un modèle de connaissance (physique) du système surveillé. On peut alors lui substituer un modèle boîte-noire, de type réseau de neurones ou réseau d'ondelettes [106, 92, 58], ce qui revient à comprimer l'information disponible dans les mesures en un vecteur de paramètres de dimension finie (les poids du réseau, et les coefficients caractéristiques de la fonction de base, translation et dilatation par exemple). Nous revenons sur ces questions de modélisation dans la conclusion.

## 2.3 Les étapes de la génération d'alarmes

La génération d'alarmes la plus élémentaire consiste à confronter directement les données mesurées à des seuils (ou gabarits) préétablis; c'est ce que les Anglo-saxons appellent limit-checking. Ce traitement brutal (directement sur le signal) contribue le plus souvent à une perte d'informations importante, et à une génération d'alarmes pauvres et par là-même trop nombreuses. Elles sont d'autre part très sensibles aux seuils qui sont difficiles à régler.

Il existe des méthodes plus élaborées, et pas nécessairement très coûteuses en calculs, qui permettent une génération d'alarmes plus fine. On peut alors distinguer deux composants principaux :

- Une *transformation* des données  $Z = f(Y)$  en ce qui est souvent appelé un *vecteur de résidus*  $Z$ . Celui-ci se doit d'être idéalement nul en l'absence de déviations et significativement non nul en leur présence.
- Une stratégie permettant d'*évaluer les résidus* et donc de *décider du caractère significatif* ou non des déviations qui se produisent, significatif vis-à-vis des bruits, erreurs de modélisation, incertitudes sur la référence, changements de modes de fonctionnement ... Cette décision peut consister en une simple *détection d'anomalie*, mais aussi, et c'est moins courant à ce niveau, comporter un aspect de *diagnostic*<sup>6</sup> quant à la localisation et l'interprétation physique de cette anomalie. L'alarme est alors qualifiée d'*alarme intelligente*, signifiant par là qu'elle est non seulement significative mais aussi porteuse d'une information reliant l'anomalie au composant physique responsable de celle-ci.

La question préalable, que l'on a déjà évoquée et sur laquelle on ne reviendra pas, est bien sûr celle de l'instrumentation optimale du système pour la surveillance [74, 72, 100, 18], autrement dit du choix du positionnement des capteurs, appelé aussi diagnosticabilité<sup>7</sup> [88].

## 2.4 Les méthodes de génération d'alarmes

Nous précisons les méthodes qui permettent d'aborder les deux principales étapes de transformation et de décision (détection et diagnostic). Nous présentons ensuite les méthodes relevant de l'intelligence artificielle et proposant de coupler modèles numériques (ou quantitatifs) et modèles symboliques (ou qualitatifs). Nous terminons en évoquant les critères d'une " bonne " génération d'alarmes.

5. Dans toute la section 2, sauf précision contraire, les modèles utilisés sont des modèles numériques.

6. Attention, ici le mot " diagnostic " a une signification un peu différente de celle que nous lui avons donnée précédemment (voir section 1), et de celui que nous lui donnerons dans la section 3, à savoir recherche de la cause profonde du phénomène détecté. Il est plus proche ici de ce que nous appelons par la suite " localisation " et correspond à une déviation identifiée sur un paramètre le (ou les) composant(s) physique(s) probablement responsable(s) de la déviation.

7. Il convient de ne pas confondre ce terme avec la diagnosticabilité [95] au sens plus statistique de détectabilité (pour un positionnement de capteurs donné), que l'on peut, pour des modèles d'état linéaires tout du moins, relier à la notion d'observabilité des Automaticiens [19].

### 2.4.1 Transformation des données ou génération de résidus

Il s'agit de définir une transformation des données  $Z = f(Y)$  en un vecteur (vecteur de résidus) qui reflète au mieux la présence d'une déviation par rapport à une référence, tout en étant le moins possible sensible aux diverses perturbations et erreurs sur le système et sur la référence. Ce vecteur sera ensuite évalué pour décider de l'occurrence ou non d'événements significatifs (voir 2.4.2).

On distingue deux types de méthodes :

1. des méthodes *géométriques*, qui exploitent les propriétés (algébriques ou géométriques) du modèle dynamique, le plus souvent linéaire, du système considéré; il s'agit essentiellement de transformées des données initiales, linéaires elles-aussi, de type *observateur* ou relations de *redondance analytique* [103, 65, 69, 79, 48, 78, 67, 77];
2. des méthodes *statistiques*, concernant essentiellement ce que l'on appelle les statistiques *exhaustives* [25, 39, 8, 19], autrement dit les transformées des données initiales qui ne perdent pas d'information [59, 38] quant aux caractéristiques propres de ces données.

En fait, les méthodes statistiques résolvent le plus souvent de manière simultanée les deux problèmes de transformation des données (ou génération de résidus) et de décision (évaluation des résidus et génération d'alarmes). On les décrit donc dans la section 2.4.2.

L'avantage des méthodes géométriques est qu'elles exploitent les propriétés de structure du modèle dynamique, en particulier d'observabilité de l'état du système considéré, s'il s'agit d'un système à représentation d'état. L'inconvénient est qu'elles ne donnent pas de mesure du caractère significatif des déviations observées. Alors que l'avantage des méthodes statistiques est d'offrir automatiquement une robustesse par rapport aux problèmes de calibration des capteurs et de bruits. Les deux types de méthodes ne sont cependant pas totalement étrangères [19, 10]: dans les deux cas, la transformation des données repose essentiellement sur des opérateurs de projection (projection sur l'orthogonal d'un sous-espace pour les méthodes géométriques, projection associée à l'estimation de l'état aux moindres carrés pour les méthodes stochastiques).

### 2.4.2 Décision ou génération d'alarme

Il s'agit de détecter (si possible de manière précoce, pour un traitement en-ligne) des déviations (fussent-elles faibles) par rapport à une caractérisation du système en ambiance de travail normale, c'est-à-dire sans excitation artificielle, ni ralentissement, ni arrêt.

Nous appelons *référence* (ou encore signature) une telle caractérisation effectuée au moyen d'une modélisation paramétrique, de vecteur de paramètre  $\theta$  dont la valeur de référence est  $\theta_0$ . Rappelons (voir 2.2) que l'on peut distinguer en général deux types de modélisation: la modélisation physique et la modélisation boîte-noire. Dans certains cas, les deux modélisations peuvent être confondues. Lorsqu'elles sont distinctes, elles sont intéressantes pour la surveillance essentiellement lorsque l'on connaît une relation au moins partielle entre elles. Naturellement, la pertinence et le contenu informatif des éléments de diagnostic que l'on est capable de délivrer alors dépendent de manière cruciale de l'existence ou non d'un lien entre les paramètres du système identifié et tout ou partie d'un modèle physique du système.

Disposant donc d'une référence et de nouvelles mesures, il s'agit de décider si ces mesures sont toujours décrites convenablement par cette référence ou s'il s'est produit un changement significatif (bien qu'éventuellement petit) de comportement. Deux solutions sont envisageables pour ce faire, schématisées par la figure 3.

- *Acquisition d'une nouvelle référence  $\theta_1$ , et comparaison des deux références  $\theta_0$  et  $\theta_1$*  [55, 56] à l'aide d'une distance convenable.

Cette méthode nous semble avoir un certain nombre d'inconvénients :

- L'acquisition de référence peut être coûteuse, par exemple lorsqu'elle ne peut être complètement automatisée et nécessite l'intervention d'un opérateur humain; c'est le cas des installations complexes, et en particulier en mécanique des vibrations [14]: toutes les méthodes d'analyse modale<sup>8</sup> et pas seulement les techniques paramétriques, nécessitent l'intervention humaine pour l'incorporation d'informations externes aux signaux afin de valider la référence finale;
- Le succès d'une telle méthode de surveillance dépend précisément et de manière cruciale du bon choix de la mesure de distance entre deux références [9];

---

<sup>8</sup>. C'est-à-dire d'identification des caractéristiques vibratoires.

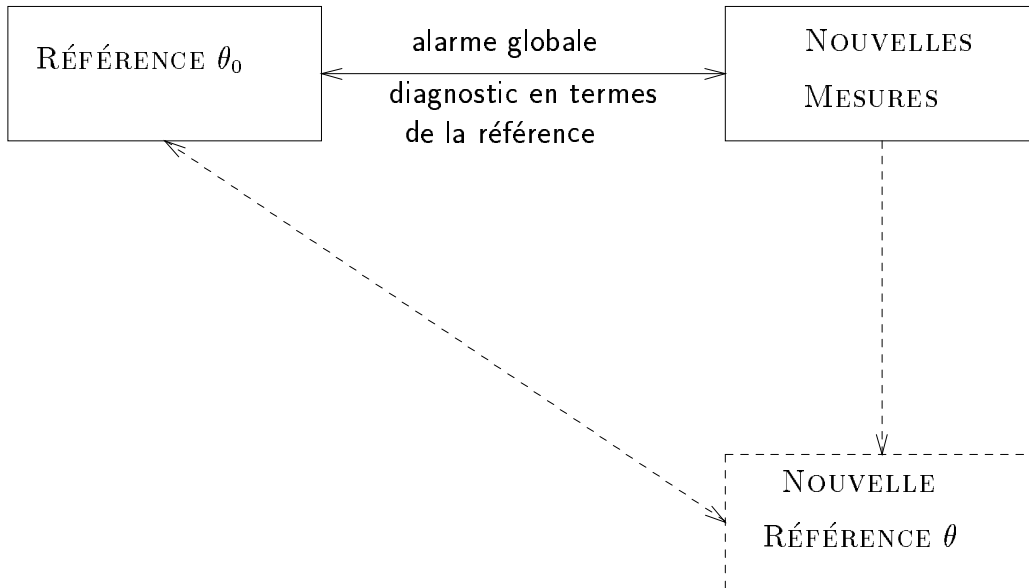


FIG. 3 – Comparer de nouvelles mesures à une référence  $\theta_0$ . On peut soit identifier, sur les nouvelles mesures, une nouvelle référence  $\theta$  et comparer les deux références  $\theta$  et  $\theta_0$ ; soit décider directement de la non-adéquation entre les nouvelles mesures et la référence  $\theta_0$  de départ, ce qui déclenche une alarme “ globale ”. Dans ce dernier cas, on peut en outre effectuer un diagnostic, en termes de la paramétrisation  $\theta$ , de l’alarme globale ainsi déclenchée.

- Dans le cas de systèmes complexes, la comparaison de deux références peut ne pas permettre de diagnostic; par exemple, toujours en mécanique des vibrations, il est bien connu que l’appariement des déformées résultant de deux analyses modales successives est un exercice très difficile lorsque des modes de vibration sont proches les uns des autres.
- *Comparaison directe des nouvelles mesures à la référence disponible.* Il s’agit alors de calculer une transformation convenable des données, en utilisant la référence. C’est d’une certaine façon ce que font les méthodes géométriques de génération de résidus. C’est aussi ce que font les méthodes statistiques. On dispose ainsi d’outils généraux de fabrication d’indicateurs permettant d’une part de mesurer l’adéquation entre une référence et un nouvel enregistrement, et d’autre part de donner des éléments de diagnostic. En particulier en [20, 105] est présentée une méthode générale de conception de l’indicateur approprié. Cet indicateur présente l’intérêt essentiel de transformer le problème de surveillance initial qui, même en termes du modèle boîte-noire, peut être très complexe, en un problème de détection standard : surveillance de la moyenne  $\mu$  d’un vecteur Gaussien (et non pas d’un processus – il y a donc prise en compte du temps de manière intrinsèque), pour lequel on dispose d’un guide pour le réglage des seuils. Cette méthode permet de s’affranchir de tous les inconvénients mentionnés plus haut, et fournit une alarme globale peu coûteuse pouvant être embarquée et un diagnostic en termes de la référence boîte-noire. En outre, elle réalise automatiquement le compromis entre l’amplitude des changements détectés et la précision de l’identification de la référence : les changements déclarés significatifs le sont vis-à-vis de ce domaine de confiance. Il en est de même vis-à-vis du bruit présent sur les mesures.

## 2.5 Génération d’alarmes intelligentes

Nous entendons ici par “ alarme intelligente ” le fait de générer des alarmes affectées d’une information relative au composant physique probablement responsable de la déviation détectée. On peut envisager plusieurs méthodes pour ce faire, procédant soit par sensibilité à une panne ou un groupe de pannes particulier, soit au contraire par réjection de l’ensemble des autres pannes.

Les deux problèmes de diagnostic<sup>9</sup> schématisés par la figure 4 peuvent se présenter, à savoir :

- *Diagnostic en termes de la référence boîte-noire*, par exemple en termes des caractéristiques vibratoires identifiables avec les capteurs dont on dispose, dans le cas de la surveillance des vibrations [17]. Plusieurs

9. Le mot “ diagnostic ” est pris ici dans le sens restreint précisé en 2.3.

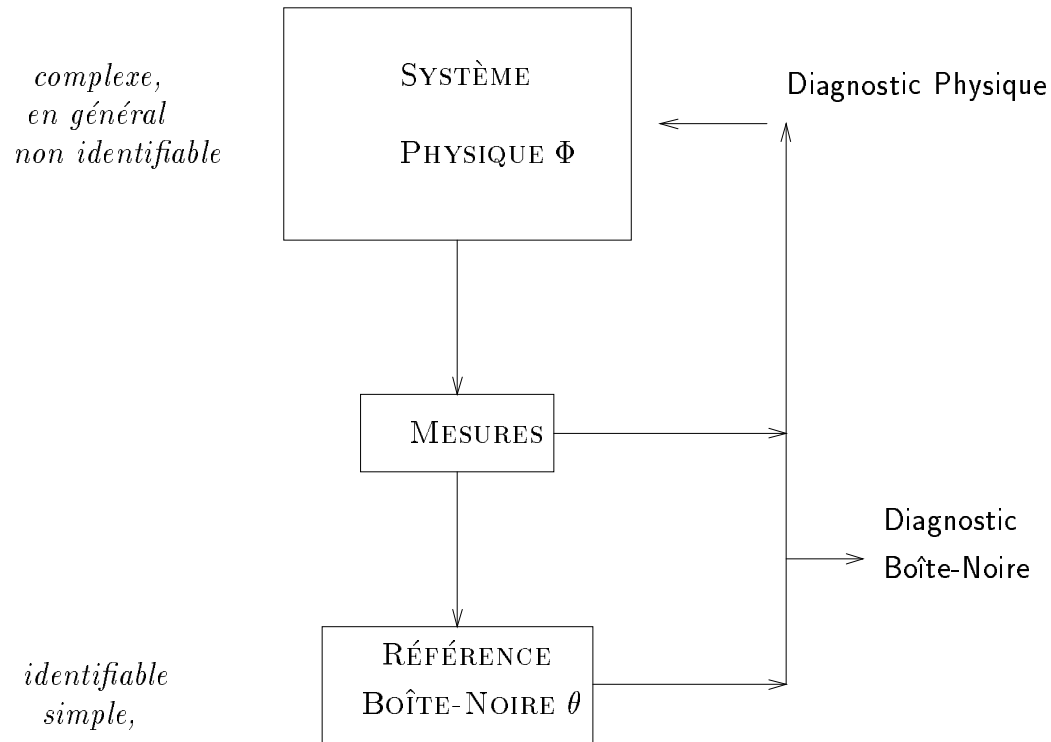


FIG. 4 – Diagnostic en termes de la référence boîte-noire ou du modèle physique: lorsque l’on utilise un modèle boîte-noire pour surveiller un système vu au travers de mesures, il est possible de remonter à un diagnostic en termes de la modélisation physique sous-jacente, et ceci sans résoudre le problème inverse correspondant.

méthodes sont envisageables, et procèdent soit par sensibilité vis-à-vis d’une panne donnée [17], soit par réjection (au sens statistique) des pannes complémentaires [19]. Par panne, on entend ici changement d’une composante ou d’une partie des composantes du vecteur de paramètres. Ce diagnostic apporte des informations qui sont intéressantes mais peuvent être complexes à exploiter : là encore il est bien connu que deux changements d’origines mécaniques très différentes peuvent provoquer des déviations de référence modale voisines.

- *Diagnostic en termes du modèle physique sous-jacent.* Ce problème est moins simple à résoudre, car le modèle physique est en général complexe et non identifiable à partir des signaux mesurés. La solution décrite en [42] contourne la résolution du difficile problème inverse correspondant, et repose sur un calcul de Jacobiens convenables dans les directions de changement qui résultent de la projection des changements physiques élémentaires sur l’espace de la référence boîte-noire – qui sont appelées “ macros-pannes ”<sup>10</sup> – et sur l’utilisation d’une technique de sensibilité. L’apprentissage de ces “ macros-pannes ”, qui sont des sous-espaces de l’espace des paramètres de la référence boîte-noire est réalisé une fois pour toutes à l’étape de conception. Par contre, les tests de sensibilité résultant peuvent être embarqués en vue d’un fonctionnement en temps réel.

Dans le cas où les différentes composantes du paramètre  $\theta$  ont un sens physique – comme c’est le cas pour les turbines à gaz [70], le diagnostic consiste à répondre à la question : quelle(s) est (sont) la (les) composante(s) de  $\theta$  responsable(s) du déclenchement de l’alarme globale? Dans le cas mono-panne, l’idée est de fabriquer

10. Les “ macros-pannes ” sont en fait des sous-espaces de l’espace des paramètres physiques, obtenus par agrégation des changements physiques élémentaires, l’idée étant qu’avec un petit nombre de capteurs on ne peut prétendre discriminer entre tous ces changements élémentaires. Ces sous-espaces sont ensuite projetés sur l’espace de la référence boîte-noire, pour constituer les directions de cet espace dans lesquelles on surveille pour savoir ce qui est à l’origine de la déviation détectée. La cohérence de l’ensemble de la procédure est assurée par le fait que la métrique utilisée pour l’agrégation des changements physiques élémentaires dans le “ gros ” espace, est la même que celle qui est utilisée pour décider du caractère significatif des déviations qui se produisent. De plus, comme, dans ce cas particulier, il y a une adéquation structurelle entre le modèle physique et le modèle boîte-noire, il résulte de cette procédure que les “ macros-pannes ” ont effectivement une cohérence du point de vue physique : ces sous-ensembles sont constitués par exemple de masses et de raideurs correspondant à des éléments du système mécanique effectivement proches les uns des autres.



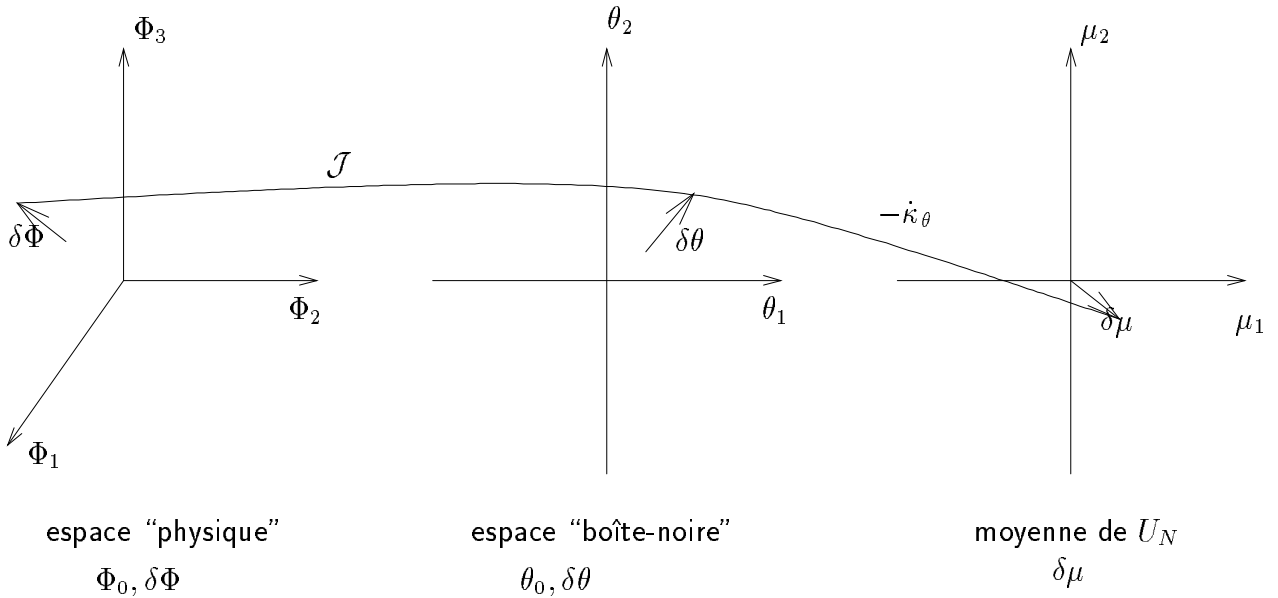


FIG. 5 – Diagnostic en termes de la référence boîte-noire ou du modèle physique avec réduction de modèle. Les déviations dans les différents espaces sont notées à l'aide du symbole  $\delta$ . Le vecteur  $\Delta_N$  est le vecteur Gaussien, indicateur de panne, résultant de la transformation des données initiales et de moyenne  $\mu$ .

autant d'indicateurs que de composantes dans  $\theta$ , chaque indicateur surveillant une composante en étant aveugle aux changements sur les autres composantes. Pour fabriquer ces indicateurs, on peut utiliser soit une technique de sensibilité (vis-à-vis des composantes surveillées), soit une technique de réjection – au sens statistique – des composantes non surveillées. En outre, dans le cas multi-panne, c'est-à-dire lorsque plusieurs composantes du vecteur paramètre peuvent changer simultanément, mais que l'on ne sait ni lesquelles ni même combien d'entre elles, une procédure incrémentale permet de déterminer le nombre de composantes défaillantes [70].

Dans le cas où le paramètre  $\theta$  est relié à un paramètre physique  $\Phi$  comme indiqué à la figure 5 qui résume les relations entre les différents espaces paramétriques en jeu, le diagnostic physique est opéré par une méthode de sensibilité, mais cette fois avec réduction de modèle (effective dans le calcul du Jacobien  $\mathcal{J}$ ), car la dimension du paramètre physique  $\Phi$  est très supérieure à celle de la référence  $\theta$ .

## 2.6 Techniques de l'IA pour la génération d'alarmes

Nous présentons ici des travaux représentatifs d'approches combinant l'utilisation de modèles numériques et de modèles symboliques dans le contexte de la génération d'alarmes. Dans les trois cas, on utilise un graphe d'influence décrivant les dépendances entre les variables du système et qui est donc un modèle qualitatif du système. Trois types d'utilisation des modèles symboliques sont décrits. Dans le premier cas, un graphe d'influence est utilisé pour décomposer le système en sous-systèmes afin d'améliorer l'efficacité de la génération d'alarmes. Dans le second cas, le graphe d'influence permet de focaliser la recherche des paramètres responsables des déviations observées, et les hypothèses retenues sont ensuite validées sur le modèle numérique. Dans les deux cas, il est fait l'hypothèse que le système est en régime établi (et donc décrit par un modèle statique). Dans le troisième cas, les graphes d'influence sont directement utilisés par simulation pour la détection. L'utilisation pour la génération de modèles purement symboliques, tels que celle que l'on peut faire de modèles prédictifs (automates, réseaux de Petri) sera présentée dans la section 3 consacrée aux modèles symboliques. Pour finir, nous présentons rapidement les "Bond graphs" ou graphes de liaison, qui peuvent servir d'outils de base pour obtenir un modèle du système, et en particulier pour obtenir l'ensemble des équations d'état du système (ensemble d'équations différentielles).

### 2.6.1 Graphes d'influence pour la simulation

Les graphes d'influence<sup>11</sup> décrivent les interactions existant entre les variables du procédé et ceci de manière qualitative. Les noeuds de ce graphe sont donc les variables du procédé et les arcs les liens entre les variables. Ces arcs sont en général étiquetés. Ce peut être par un signe positif ou négatif, indiquant si les deux variables liées évoluent dans le même sens ou dans un sens inverse; par des gains exprimant l'ordre de grandeur des variations des variables; ce peut être aussi par des délais indiquant le décalage dans le temps des variations.

Dans les études faites pour la station d'essais de Cadarache dans le cadre du projet Esprit Alliance, les graphes d'influence sont étiquetés par des fonctions de transfert qualitatives [75] précisant les retards, les gains et les temps de stabilisation. Les graphes d'influence sont alors utilisés pour simuler le système et en prédirent les sorties en fonction des entrées. La simulation revient à propager les variables d'entrée à travers le graphe d'influence afin d'obtenir les réponses qualitatives (voir l'algorithme de propagation dans [62]). De même, [28, 27] propose d'utiliser un graphe d'influence (dénommé graphe causal) associé à des contraintes globales numériques pour simuler le système et détecter des différences entre mesures prédites et mesures observées.

Les graphes d'influence peuvent aussi être utilisés pour expliquer les anomalies dans la phase de diagnostic. Ils sont donc de nouveau examinés dans la partie traitant de l'interprétation des alarmes (voir 3.5.1).

### 2.6.2 Graphe d'influence et découpage hiérarchique du système

Il s'agit essentiellement des travaux de Rouquette [87]. S'appuyant sur une modélisation numérique, O. Rouquette propose d'utiliser des techniques de génération d'alarmes assez classiques de type comparaison d'un résidu (ou transformée convenable des entrées et sorties disponibles) à un seuil [87]. Afin de réduire la taille du problème, une décomposition du système en sous-systèmes, indépendants aux boucles de feedback près, est effectuée et permet de traiter le système global de manière hiérarchique [87]. Cette décomposition repose sur une décomposition du graphe de dépendance des variables en sous-graphes fortement connectés reliés entre eux uniquement par des arcs identifiés comme des boucles de feedback. Les dépendances entre variables sont des dépendances de type calculatoire: " pour calculer  $x$ , j'utilise ou j'ai besoin de  $y$ ". Cette technique a été expérimentée pour la détection d'anomalies sur un système de contrôle thermique pour la station spatiale Freedom de Mc Donnell Douglas.

### 2.6.3 Utilisation conjointe d'un modèle quantitatif et qualitatif

L'idée essentielle développée par Gallanti [52] est de générer les candidats au diagnostic en s'appuyant sur un modèle qualitatif et de valider ces candidats en les testant sur un modèle numérique. Il s'agit donc d'une approche de type " Generate and Test".

Le graphe d'influence est construit par abstraction du modèle numérique et reflète les dépendances entre les variations des paramètres du système et les variations des sorties. Les paramètres candidats à être responsables des déviations observées sur les sorties sont générés en partant du principe suivant: une déviation d'une sortie doit être expliquée par au moins un paramètre l'influençant; la non-déviations d'une sortie ne peut s'expliquer que par l'absence de variations des paramètres ou la variation conjuguée de deux paramètres se compensant (le cas de trois ou plus paramètres se compensant n'est pas envisagé).

Les candidats ainsi générés sont ensuite testés par simulation sur le modèle numérique afin d'éliminer les candidats erronés (" spurious") qu'aurait pu produire la recherche à base de modèle qualitatif.

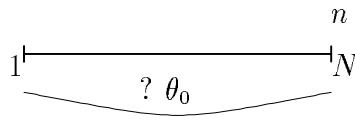
Une approche très semblable est proposée dans [24, 104]; une des particularités de ces travaux est l'emploi de " bond-graphs" ou graphes de liaison comme outil de base pour l'acquisition du modèle (sous forme d'équations d'état) du système (voir 2.6.4).

### 2.6.4 " Bond-graphs" ou Graphes de liaison

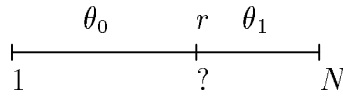
Un problème classique est celui de l'acquisition du modèle, qu'il soit numérique ou qualitatif. Une méthodologie à base de graphes de liaison (ou bond-graphs) a été proposée par [86, 24, 66]. Celle-ci est spécialement adaptée à la modélisation de systèmes physiques dynamiques et repose sur l'identification d'un certain nombre d'éléments de base, indépendants du domaine. Ceux-ci peuvent prendre des formes physiques différentes selon les domaines, mais leurs interactions sont, elles, indépendantes du domaine et bien identifiées. Ainsi, on peut associer à chaque liaison entre deux variables un effort et un flux avec: effort \* flux = puissance. Dans le domaine

11. Nous distinguons ici les graphes d'influence, où les arcs relient des variables entre elles, des graphes causaux où les arcs relient états et phénomènes (qui peuvent être mesurables) entre eux; dans la littérature, le terme graphe causal est souvent utilisé dans les deux cas.

## Validation de Modèle



## Détection de changement hors-ligne



## Détection de changement en-ligne

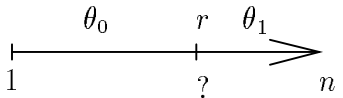


FIG. 6 – Trois problèmes de détection.

de la mécanique, l'effort correspond à la force et le flux à la vitesse; en thermodynamique, l'effort correspond à la température et le flux au flux de chaleur. Nous ne décrivons pas plus en détail cette approche, mais signalons en les points intéressants : la construction du modèle s'appuie sur l'existence d'une bibliothèque de primitives de base et une opération de composition permettant de constituer de manière systématique le modèle global du système; il est possible, à partir de ce modèle de type graphe de liaison, d'engendrer les équations d'état du système, ainsi que les équations d'entrées-sorties, et les liens entre composants physiques et paramètres sont alors explicites. Il fournit donc un modèle numérique du système sur lequel peut travailler un algorithme d'interprétation ou de diagnostic comme le propose [24]. Lucas [66] utilise, lui, ce modèle afin d'engendrer différents types de modèles (numériques, mais aussi modèles qualitatifs, graphes causaux) qui concourent à l'établissement du diagnostic.

## 2.7 Les critères de la génération d'alarmes

Terminons par une remarque concernant les critères qui sous-tendent la conception des algorithmes de génération d'alarmes, au sens où nous l'entendons en 2.4. Comme ces critères sont liés à la nature hors-ligne ou en-ligne des traitements effectués en vue de la détection, nous commençons par revenir sur la discussion en-ligne/hors-ligne que nous avons abordée au début de cet article.

### 2.7.1 Trois problèmes de détection

Il est important de distinguer trois problèmes de détection, et donc trois façons de générer des alarmes à partir de données mesurées. Ces trois types de problèmes sont illustrés par la figure 6 :

- *Validation de modèle* : On dispose d'une part d'une référence  $\theta_0$  et d'autre part d'un enregistrement de taille  $N$  de nouvelles mesures. On cherche à décider si les nouvelles mesures sont bien décrites par la référence. Ce problème peut être posé hors-ligne ( $N$  fixé) ou bien en-ligne (échantillon de taille  $n$  variable).
- *Détection de changement hors-ligne* : On dispose d'un enregistrement de taille  $N$  de nouvelles mesures. On cherche à décider si, à l'intérieur de cet enregistrement, il y a un changement de paramètre, de la valeur  $\theta_0$  à la valeur  $\theta_1$ , se produisant à l'instant inconnu  $r$ .
- *Détection de changement en-ligne* : A chaque instant  $n$ , on cherche à décider si, à l'intérieur de l'enregistrement, il y a un changement de paramètre, de la valeur  $\theta_0$  à la valeur  $\theta_1$ , se produisant à l'instant inconnu  $r$ .

Le plus difficile de ces trois problèmes est évidemment le troisième, puisque l'on cherche à y répondre à la même question en utilisant moins de données, et donc moins d'information. Et c'est d'ailleurs celui pour lequel la quantité de résultats théoriques disponibles est la moindre.

En pratique, pour une installation donnée, on peut utiliser l'un ou plusieurs de ces points de vue, selon les ordres de grandeur relatifs des constantes de temps des différents types d'événements à surveiller et de l'échantillonnage des mesures disponibles. Par exemple, pour la surveillance du système de combustion d'une turbine à gaz en [15, 70] ont été utilisés essentiellement les premier et troisième points de vue, autrement dit la validation de modèles et la détection de changement en-ligne, pour la surveillance de dégradations lentes et la détection de pannes brutales respectivement.

### 2.7.2 Les critères

Les critères qui sont utilisés pour la conception et l'analyse des performances des algorithmes de détection correspondants sont pris parmi les cinq suivants :

1. temps moyen entre fausses alarmes;
2. probabilité de fausse détection;
3. retard moyen à la détection;
4. probabilité de non détection;
5. précision des estimateurs de l'instant d'occurrence et de l'amplitude du changement.

Toutes ces quantités sont très fortement influencées par la distance de Kullback (encore appelée information) entre les distributions avant et après changement. Cette distance peut être utilisée pour définir la détectabilité d'un changement [19].

## 3 Interprétation des alarmes : du filtrage au diagnostic

Nous avons vu dans la section précédente les différentes techniques permettant, à partir des observations recueillies par les capteurs, de détecter des déviations anormales et de générer les alarmes correspondantes. Celles-ci sont en général destinées à un opérateur qui les reçoit en salle de contrôle et doit les gérer. Nous présentons dans cette section les techniques permettant d'interpréter ce flux d'alarmes afin d'aider l'opérateur dans cette tâche. Nous distinguons :

- le *filtrage*, dont le but est de limiter la charge d'information en ne présentant à celui-ci que les alarmes “intéressantes”;
- la *localisation* qui a pour but de caractériser ou d'identifier<sup>12</sup> la situation de dysfonctionnement détectée; ceci consiste souvent à faire le lien entre une alarme ou une suite d'alarmes et le dysfonctionnement d'un(des) composant(s) responsable(s) de celle(s)-ci;
- le *diagnostic* qui a pour but de proposer la cause la plus probable du dysfonctionnement observé. On réserve souvent le terme de *panne* au résultat du diagnostic.

Comme nous l'avons vu dans la section 2, la phase de génération d'alarmes peut fournir des alarmes dites intelligentes qui contiennent une part de localisation des anomalies. Il faut aussi noter que la distinction entre localisation et diagnostic peut être difficile à faire. Le terme “diagnostic” est souvent associé à une recherche plus en profondeur ou plus globale d'une cause de la situation anormale détectée. La recherche des causes d'un dysfonctionnement peut se faire à des niveaux de profondeur variables selon l'objectif visé; la caractérisation d'une situation anormale est souvent suffisante lorsqu'il s'agit de surveillance d'un système en-ligne (et donc de décider d'une action de correction), alors que l'on recherchera souvent la cause profonde à l'origine d'un dysfonctionnement lorsqu'il s'agit de maintenance préventive ou d'analyse a posteriori du système (où l'on doit décider d'actions de réparation ou de remplacement de composants du système).

Après avoir décrit ces différents niveaux d'interprétation, nous présentons les différents types de modèles qui peuvent être utilisés à cet effet. À la différence des modèles utilisés dans la phase de génération d'alarmes, ces modèles sont le plus souvent symboliques<sup>13</sup>; nous différencions les modèles associatifs, les modèles prédictifs et les modèles explicatifs et en décrivons succinctement les caractéristiques. Nous montrons ensuite les techniques exploitant ces modèles et contribuant à l'interprétation des alarmes.

12. On utilise souvent le terme identification au lieu de celui de localisation, mais, par crainte d'ambiguïté avec ce qu'est l'identification des paramètres en automatique et traitement du signal, nous ne l'employons pas ici.

13. Nous préférons utiliser le qualificatif “symbolique” opposé à “numérique”, plutôt que “qualitatif” opposé à “quantitatif”, pour éviter une ambiguïté avec les modèles de la physique qualitative, qui sont des modèles symboliques mais n'en sont pas les seules instances.

### 3.1 Différents niveaux d'interprétation

#### 3.1.1 Filtrage

L'examen des mesures issues des capteurs a permis de générer des alarmes qui sont destinées à l'opérateur chargé de la conduite du système surveillé. Un problème important est celui de la quantité des alarmes reçues qui peut vite transformer le poste de conduite en une centrale d'acquiescement d'alarmes. L'objectif du filtrage est donc de limiter la charge d'information en sélectionnant parmi l'ensemble des alarmes générées un sous-ensemble d'alarmes pertinentes, suffisantes pour la bonne information de l'opérateur. Un premier traitement peut se contenter de distinguer les alarmes selon leur niveau de priorité: les alarmes sont affectées a priori d'un degré de priorité et présentées en respectant ces priorités. On distingue généralement les alarmes dites de sécurité qui doivent être traitées d'urgence et correspondent à des situations très graves; les alarmes de conduite qui sont susceptibles d'entraîner des actions de conduite et les alarmes d'information.

Une des raisons de l'avalanche d'alarmes est la forte corrélation entre les variables observées, lorsque cette corrélation n'est pas prise en compte lors de la génération d'alarmes; ces dépendances provoquent l'affichage d'alarmes redondantes entre elles car fortement corrélées et ceci avec des décalages dans le temps pouvant être importants si l'on a à faire à des procédés lents. Les principaux outils de filtrage sont le graphe causal et le graphe d'influence, qui tous deux expriment les liens de dépendances entre situations et entre variables; ils permettent d'identifier les dépendances entre alarmes et de ne conserver que les alarmes sources. Afin de tenir compte de l'aspect temporel de ces dépendances, ces graphes causaux et d'influence comportent le plus souvent des informations temporelles comme par exemple des décalages (ou retards) entre les variations des variables. Nous reviendrons sur ces graphes lorsque nous parlerons des modèles explicatifs (voir 3.5).

#### 3.1.2 Localisation

L'objectif de la localisation est de faire le lien entre les anomalies détectées dans l'évolution des sorties du système et une (ou plusieurs) situation(s) répertoriée(s) de dysfonctionnement du système. Il s'agit couramment de faire le lien entre les paramètres identifiés comme sortant des normes et le dysfonctionnement d'un des composants du système (ce qui explique le terme "localisation"). Cette phase consiste souvent à repérer un ensemble (ou une séquence) d'alarmes dont les occurrences concomitantes sont caractéristiques d'une situation. Les techniques proposées par l'intelligence artificielle s'appuient sur trois types de modèles bien différents: les modèles associatifs, les modèles prédictifs et les modèles explicatifs que nous décrivons plus largement dans la section 3.2.1.

#### 3.1.3 Diagnostic

La phase de diagnostic a pour but de rechercher les causes premières des phénomènes observés. Il s'agit d'une analyse profonde du procédé, souvent plus utile dans un but de maintenance que de conduite. Les modèles et les méthodes proposés pour le diagnostic sont proches de ceux utilisés dans la phase de localisation, même si les modèles explicatifs sont souvent privilégiés.

### 3.2 Différents types de modèles symboliques

Nous examinons les différents types de modèles utilisés en intelligence artificielle pour l'interprétation d'alarmes dans un contexte de surveillance. Ces modèles se différencient des modèles utilisés en diagnostic statique principalement par l'existence d'une composante temporelle nécessaire pour exprimer les délais entre causes et effets; les durées des phénomènes observés; la persistance ou l'intermittence des dysfonctionnements. Dans les modèles prédictifs, qui permettent de simuler le comportement du système, le caractère dynamique du système induit l'existence (éventuellement sous-jacente) d'un graphe de transitions entre les états du système et ceci dans chacun des modes de fonctionnement considérés.

#### 3.2.1 Trois grandes classes de modèles

Nous distinguons trois grandes classes de modèles symboliques:

- les *modèles associatifs*, tels que ceux utilisés par les systèmes experts ou la reconnaissance de scénarios, associent à des schémas d'alarmes, contraintes temporellement, des situations significatives, en particulier de dysfonctionnements;

- les *modèles prédictifs* permettent de simuler les différents comportements possibles du système dans ses différents modes de comportement; ils peuvent être utilisés pour la génération d’alarmes par comparaison entre les comportements prévus et les comportements observés ainsi que pour l’interprétation;
- les *modèles explicatifs* sont de deux types; les graphes d’influence décrivent les liens de dépendances entre les variables du système et les graphes causaux les liens de causalité entre les états ou situations (le plus souvent de dysfonctionnements) et leurs effets observés; les premiers peuvent être utilisés en simulation pour la détection; les deux sont utilisés de manière abductive pour remonter des alarmes aux variables ou situations les expliquant .

Dans un système de suivi et d’interprétation d’alarmes, il est fréquent de trouver plusieurs modèles de types différents, chacun étant utilisé pour des buts distincts; ils peuvent décrire le système selon différents points de vue, par exemple la structure physique du système, son comportement, ses fonctions. Certains modèles peuvent être orientés composant ou orientés processus; ils peuvent aussi correspondre à des descriptions à différents niveaux de détail.

### 3.2.2 Acquisition des modèles

Ces différents types de modèles se distinguent par leurs modes d’acquisition.

- Les modèles associatifs reposent en général sur une expertise existante; ils résultent d’une phase d’extraction d’expertise opérée auprès des opérateurs (opérateurs de conduite du procédé; opérateurs de supervision d’un réseau ...) ou des ingénieurs de maintenance habitués à analyser en profondeur le comportement du système à partir des observables. Les points délicats sont la validation et la couverture de l’expertise ainsi obtenue (celle-ci est-elle cohérente? couvre-t-elle bien l’ensemble des situations envisageables?). Ces modèles sont utilisables directement aussi bien pour la partie détection, filtrage, interprétation et diagnostic.
- Les modèles prédictifs peuvent s’obtenir soit par abstraction des modèles numériques comme le suggère l’approche de la physique qualitative (encore faut-il disposer des modèles numériques), soit directement à partir des modèles symboliques de conception ou de description du fonctionnement du système s’ils existent ou s’ils peuvent être formalisés par les ingénieurs experts du procédé. Une méthodologie de construction de tels modèles par composition de primitives de base est actuellement souvent préconisée; c’était le cas déjà dans les travaux de [41] où chaque composant était associé à un ensemble de contraintes représentant son fonctionnement (component-connection); cela peut se faire en s’appuyant sur la théorie des graphes de liaison (ou bond-graphs), ou par composition d’automates. Le point-clé de ce type de modélisation est le choix du bon niveau d’abstraction de ce modèle; la coopération entre des modèles à différents niveaux d’abstraction étant par ailleurs souvent souhaitable. Par leur capacité à calculer les comportements prédits, ces modèles sont directement utilisables pour la détection.
- Les modèles explicatifs s’appuient souvent sur une expertise déjà explicitée dans l’entreprise sous forme de catalogue de pannes, d’AMDECs ou d’arbres de défaillances qui expriment pour chaque panne l’ensemble des symptômes observables. Le point faible de ces expertises est l’expression des contraintes temporelles souvent exprimées de manière très imprécise. Un raisonnement abductif est nécessaire pour obtenir à partir des observables les pannes les expliquant.

### 3.2.3 Passage d’un modèle à un autre

Il est souvent envisageable de passer d’un type de modèles à un autre en utilisant des techniques de type compilation. Ceci peut être le cas lorsque l’on veut passer d’un modèle plus facile à acquérir à un autre modèle dont la structure est plus adaptée au problème posé. Il est par exemple possible de transformer un modèle explicatif (causes - effets) en un modèle associatif (effets - causes). Dans un formalisme logique, ceci correspond à faire l’hypothèse de complétion du modèle et à utiliser des mécanismes de type circonscription par exemple. L’idée essentielle est de mettre un ensemble de connaissances sous une forme plus directement et donc plus efficacement utilisable. Au lieu ainsi de rechercher par un raisonnement abductif une explication pour un ensemble de symptômes dans un modèle explicatif, on associera directement l’ensemble de symptômes à son explication. Une autre possibilité est de s’appuyer sur un modèle prédictif pour rechercher, par simulation, les associations observations-pannes, puis d’en dégager les schémas discriminants [22, 89]. L’intérêt d’une telle compilation des connaissances n’est pas propre aux systèmes dynamiques et a été très discutée [57]; nous ne reprendrons pas ce débat ici.

### 3.3 Modèles associatifs

Les modèles associatifs reposent sur l'expression directe des liens entre un ensemble de symptômes ou d'événements et une situation caractéristique (en particulier un dysfonctionnement du système). Nous décrivons deux approches typiques : les systèmes experts et la reconnaissance de scénarios, et leurs applications au problème de la surveillance.

#### 3.3.1 Systèmes experts

Les systèmes experts ont été les premiers outils utilisés pour le diagnostic de systèmes statiques et ils ont été étendus pour le diagnostic de systèmes dynamiques. Ils sont encore à l'heure actuelle les plus utilisés dans les systèmes de surveillance opérationnels. PICON était, dès les années 80, un système destiné à raisonner en temps réel sur des applications de contrôle de procédés et ce système a donné lieu au développement du générateur de systèmes experts G2 (développé par GENSYM) [53] qui est utilisé dans de nombreuses applications industrielles, par exemple par l'IFP [32]. On peut aussi citer RTWorks [5], Chronos, CogSys [5], Ilog Rules ou Kool95 utilisé par Sollac dans le cadre du projet SACHEM. Ces générateurs de systèmes experts dits temps réel se caractérisent par les points suivants :

- la prise en compte de l'aspect temporel : Ceci se traduit par
  - l'existence de primitives temporelles dans le langage d'expression des connaissances permettant d'exprimer des contraintes symboliques et/ou numériques sur les événements observés,
  - la garantie d'une réponse " temps réel " (dépendant du processus surveillé), obtenue par l'existence de mécanismes d'interruption et permettant un comportement réactif, ainsi que par des possibilités de focalisation,
  - la gestion d'une base de faits datés avec des mécanismes d'oubli ou d'archivage.
- une architecture de type tableau noir [54] permettant l'utilisation de bases de connaissances hétérogènes et un contrôle de type réactif par gestion d'un agenda (voir par exemple ATOME développé par le CRIN [61]),
- une architecture permettant une intégration aisée d'outils de traitements des informations issues des capteurs, d'outils de simulation, et fournissant des interfaces utilisateurs conviviales et souples. G2 propose par exemple un simulateur comprenant des algorithmes de résolution d'équations; un langage graphique permettant de décrire des diagrammes; un module permettant d'intégrer des réseaux neuronaux.

Citons quelques réalisations françaises dans le domaine de la supervision utilisant un système expert : le projet Sachel de la Sollac pour le système d'aide à la conduite des hauts fourneaux repose pour la partie détection et interprétation sur une approche de type système expert utilisant le langage orienté objet Kool. L'acquisition de l'expertise s'est faite auprès des experts de la conduite du haut-fourneau de Fos sur Mer. Un premier système de détection et d'interprétation est en cours de mise en place. Sylvie Cauvin [32] a conçu pour l'IFP une architecture de supervision, Alexip, destiné aux procédés de raffinage et de pétrochimie. Il a été implémenté à l'aide du logiciel G2 et expérimenté pour la supervision du procédé Alhabutol. Dans ce système, des modèles de type explicatif (graphes causaux; graphes d'influence) sont utilisés en coopération avec les systèmes à base de connaissances. On peut encore citer les réalisations de Total pour l'unité de craquage de la raffinerie de Provence et l'unité de désasphaltage de la raffinerie de Gondreville, les expériences de Lafarge pour la supervision des fours à ciment ainsi que Elf Aquitaine pour l'optimisation de l'alimentation en hydrogène de la raffinerie de Grandpuits.

#### 3.3.2 Reconnaissance de scénarios

Les logiciels de reconnaissance de scénarios correspondent à un second type de modèles de type associatifs. Contrairement aux précédents, ceux-ci ont été tout spécialement conçus pour le suivi de systèmes dynamiques et donnent une part prépondérante au traitement des contraintes temporelles. Un scénario ou chronique décrit une situation par l'ensemble des événements qui la caractérisent et les contraintes temporelles qui s'y rapportent. La reconnaissance de scénarios a pour but d'analyser un flux d'événements en-ligne et de détecter parmi l'ensemble des scénarios décrits ceux que les événements satisfont. C'est le cas de IxTeT [44] qui a été expérimenté dans le cadre du projet Esprit Tiger sur la surveillance de turbines à gaz. Pour remédier à la difficulté d'acquisition des scénarios, une étude est en cours en vue de l'utilisation des techniques d'apprentissage à cet effet.

### 3.3.3 Avantages/inconvénients

Les points délicats de ce genre d'approche sont l'acquisition, la validation et la couverture de l'expertise obtenue (celle-ci est-elle cohérente? couvre-t-elle bien l'ensemble des situations envisageables?). L'acquisition, comme nous l'avons vu, est faite auprès des experts, ce qui comporte un risque d'inexactitudes, d'incohérences et d'incomplétudes. Concernant la validation, les techniques proposées pour les bases de connaissances [6] paraissent adaptées à ce problème mais se focalisent en général peu sur la vérification des contraintes temporelles. Celle-ci en revanche est prise en charge dans les systèmes de reconnaissance de scénarios [44]. Un autre point est celui de la non-généricité inhérente à ce type de modèles. Une évolution du procédé peut rendre obsolète l'ensemble de l'expertise.

On leur reconnaît en revanche les points forts suivants: leur efficacité résultant de l'association directe entre alarmes et situations; en particulier, les patterns d'alarmes ne contiennent en général que les éléments discriminants nécessaires et suffisants à l'identification d'une situation; leur utilisation possible aussi bien pour la partie détection, filtrage, interprétation et diagnostic.

## 3.4 Modèles prédictifs

Les modèles prédictifs décrivent le comportement du système dans ses différents modes (comportement normal, comportements dégradés, comportements en cas de panne). Ils permettent en général de simuler le système pas à pas et de prédire ainsi les valeurs des observables. Ils sont donc directement utilisables pour la détection de situations anormales par confrontation entre prédictions et observations. En raison du caractère dynamique des systèmes surveillés, un mode de comportement se décrit par un ensemble d'états (stables ou transitoires) et des transitions entre ces états.

### 3.4.1 Modèles de la physique qualitative

Les modèles issus de l'approche de la physique qualitative [47] peuvent être vus comme une abstraction des modèles numériques correspondants. Les motivations mises en avant [73] sont d'éviter la complexité de traitement des équations différentielles numériques; de prendre en compte l'inévitable imprécision des mesures et de se rapprocher du mode de raisonnement d'un ingénieur. Le choix d'un espace quantité pour les variables décrivant le procédé détermine le niveau d'abstraction du modèle qualitatif; à cet espace quantité doit correspondre une algèbre définissant dans ce nouvel univers les résultats des opérations qualitatives retenues, en général les opérateurs arithmétiques, la dérivée ... On définit ainsi une algèbre des signes, des ordres de grandeur, des intervalles. Le modèle qualitatif est constitué par l'espace quantité et les contraintes qualitatives entre les variables. Un état qualitatif est alors décrit par un ensemble de variables d'états et leurs valeurs (éventuellement leurs sens de variation [45]); les états du système sont les états d'équilibre et les états transitoires entre ces états d'équilibre. Les transitions déterminent les différentes évolutions possibles du système.

La simulation qualitative [47, 45, 27] permet, étant donné un état initial, de calculer les états successeurs satisfaisant les contraintes décrites par le modèle. Le suivi et la détection revient à repérer dans quel état est le système en comparant prédictions et observations. L'interprétation consiste à sélectionner, parmi les modes de comportement possibles, celui ou ceux qui permettent de prédire le comportement le plus proche de celui observé. [90] ont proposé un algorithme de simulation qualitative, Fusim, où les variables qualitatives sont associées à un nombre flou, ce qui permet de proposer un classement des comportements qualitatifs calculés. Dans tous les cas, un problème délicat est celui du recalage des modèles. Nous avons déjà en 2.6 présenté les graphes d'influence et l'utilisation qui peut en être faite en tant que modèle prédictif pour la génération d'alarmes, dès que l'on y associe un certain nombre d'informations quantitatives tels que fonctions de transfert ou équations numériques. Une des

principales difficultés rencontrées par cette approche est d'éviter de produire par simulation des comportements erronés ("spurious"), résultant de l'approximation de ces modèles abstraits et qui ont pour effet de diminuer la sensibilité de la détection. Une première solution est de compléter les modèles qualitatifs par des informations numériques, comme cela est proposé dans les approches dites semi-quantitatives [45]; une autre solution consiste à utiliser ces modèles pour générer un ensemble de comportements candidats restreint, avant de les valider sur des modèles numériques. Une approche dans ce sens est celle de [51] présentée en 2.6. Une autre faiblesse de ce genre d'approche est qu'elle ne donne pas de représentation explicite du temps; la seule notion est l'ordre sur les états, mais il n'est pas possible d'exprimer des délais ou la durée des états. La simulation qualitative causale remédie à ce problème en permettant d'effectuer la propagation de signaux incluant une composante temporelle à travers un graphe d'influences. Le système CA-EN qui utilise cette technique complétée d'un modèle qualitatif classique a été expérimenté dans le cadre du projet Esprit Tiger pour la surveillance de turbines à gaz [28].



Un dernier problème est celui du seuillage; en particulier, il serait intéressant d'étudier expérimentalement l'apport des méthodes de segmentation automatique [12, 4, 19] au niveau des seuils implicitement nécessaires en amont de toute utilisation des modèles de physique qualitative actuels, ne serait-ce que pour le passage du numérique au qualitatif, dont on sait – voir par exemple le projet “Alliance” – qu'il est particulièrement délicat en dynamique: le remplacement des seuils directs sur les mesures, souvent utilisés, par les outils de segmentation devrait permettre un gain en robustesse sensible.

Les modèles qualitatifs peuvent être construits en s'appuyant directement sur un modèle numérique, mais un avantage de ce genre d'approche est qu'ils peuvent être développés pour des systèmes complexes en s'appuyant sur les compétences des ingénieurs, même lorsque le développement de modèles numériques est difficilement envisageable.

### 3.4.2 Modèles à événements discrets

Dans les modèles à événements discrets, le temps est discrétisé et les variables le sont aussi. Le formalisme sous-jacent est l'automate d'état fini; les seules dates considérées sont celles correspondant à un changement d'état. Les réseaux de Petri rentrent directement dans ce cadre et ont été enrichis (réseaux de Petri temporisés; temporels; stochastiques; à marquage flou) afin de mieux représenter les évolutions d'un système dynamique. Les réseaux de Petri sont utilisés essentiellement en simulation et permettent en particulier de bien traiter les problèmes de synchronisation d'événements. En revanche, leur utilisation pour le diagnostic est beaucoup plus difficile [99]. Dans [82], Portinale a défini un raisonnement de type chaînage arrière sur les réseaux de Petri afin de les utiliser dans la phase de diagnostic. Leurs applications opérationnelles sont encore restreintes. On peut citer l'utilisation des réseaux de Petri flous pour une unité de raffinage d'Esso au Canada [96] ainsi que la proposition de [32] s'en inspirant pour construire le graphe des situations.

Une autre possibilité est de modéliser directement le fonctionnement du système par des automates. C'est l'approche proposée par [89] ainsi que celle utilisée par [22] dans le cas de la surveillance du réseau Transpac. Dans les deux cas, il est proposé de construire l'automate global du système par composition d'automates élémentaires associés aux composants du système et disponibles en bibliothèque. Cette représentation est adaptée directement à la simulation et la détection; pour la phase d'interprétation et de diagnostic, deux méthodes sont possibles. La première consiste à transformer hors-ligne l'automate en un “diagnostiqueur”; c'est ce que propose [89]; une autre possibilité [23] consiste à simuler cet automate à partir des pannes les plus courantes afin de construire des couples (ensemble d'événements observables, pannes) utilisables en-ligne par un reconnaiseur de scénarios (voir 3.3.2). De même que les modèles qualitatifs, les représentations à base d'automates ne représentent pas en général les contraintes temporelles de manière explicite; une extension possible est celle des automates temporels

### 3.4.3 Autres formalismes

Un certain nombre d'autres formalismes existent qui ne seront pas détaillés ici, en particulier les réseaux tels les réseaux bayésiens ou les réseaux probabilistes temporels ou les approches en logique classique ou linéaire.

## 3.5 Modèles explicatifs

Les modèles explicatifs décrivent partiellement le système en explicitant les liens de dépendance existant entre les variables du procédé dans le cas des graphes d'influence, entre les états du système et leurs effets observés dans le cas des graphes causaux.

### 3.5.1 Graphes d'influence

Les graphes d'influence ont déjà été présenté en 2.6; rappelons qu'ils décrivent les interactions existant entre les variables du procédé et ceci de manière qualitative. Les noeuds de ce graphe sont donc les variables du procédé et les arcs les liens entre les variables. Ces arcs sont en général étiquetés. Ce peut être par un signe positif ou négatif, indiquant si les deux variables liées évoluent dans le même sens ou dans un sens inverse; par des gains exprimant l'ordre de grandeur des variations des variables; ce peut être aussi par des délais indiquant le décalage dans le temps des variations. Comme nous l'avons mentionné en 2.6, ils peuvent être utilisés dans la phase de génération d'alarmes en leur appliquant un algorithme de simulation; c'est le cas du projet Esprit Alliance où les graphes d'influence sont étiquetés par des fonctions de transfert qualitatives [62] ou dans le cas du projet Tiger avec le logiciel CA-EN [28]. Ces graphes sont aussi utilisés de manière abductive: il servent lors la phase de filtrage des alarmes afin de repérer les alarmes qui sont directement dépendantes d'autres et

ne seront pas affichés à l'opérateur [62]. Dans [32], ils permettent, sous le vocable de graphes causaux à court terme, de fournir des explications du comportement du procédé. Une autre utilisation est celle faite par [87] et présentée en 2.6.

Une difficulté de constitution des graphes d'influence, à partir des équations numériques ou qualitatives, et de leur utilisation, est l'existence de boucles (telles que les boucles de feedback).

### 3.5.2 Graphes causaux (et temporels)

Les graphes causaux décrivent les liens causaux existant entre les différents états du système, et en particulier entre les perturbations extérieures et les événements de type panne et leurs effets observables. Les noeuds du graphe représentent donc en général des situations et les feuilles des symptômes ou effets observables. Les arcs sont interprétés comme des liens causaux forts de type "a cause b" ou des liens causaux faibles de type "a peut causer b" (une étiquette peut permettre de distinguer ces deux cas comme dans [37]). Le graphe peut avoir une structure particulière telle un graphe bipartite où les arcs relient exclusivement des causes à des manifestations [85], ou un graphe en couches, mais en général le graphe est un treillis. Les graphes causaux ont été largement utilisés pour le diagnostic de systèmes statiques [37, 85]; ils ont été étendus pour traiter le cas des systèmes dynamiques essentiellement par l'ajout de contraintes temporelles étiquetant les arcs comme on le trouve dans [81] ou dans [30]. Ces graphes sont toujours utilisés de manière abductive et permettent de rechercher, à partir des symptômes, les causes premières de ceux-ci. Ils permettent donc d'établir un diagnostic. L'ajout des contraintes temporelles a conduit à formaliser ce que l'on entend par diagnostic abductif temporel et à en proposer des algorithmes efficaces [30, 29]. Dans le système DIAPO [81], système de diagnostic pour l'aide à la maintenance des pompes de refroidissement des réacteurs nucléaires, un tel graphe causal est utilisé afin de faire le diagnostic a posteriori d'une anomalie de fonctionnement; l'architecture de DIAPO est de type tableau noir et intègre d'autres modèles, en particulier une base de connaissance utilisée par un système expert.

[32] distingue le graphe causal des problèmes, le graphe causal à long terme, et le graphe causal à court terme qui est de type graphe d'influence. Le graphe causal des problèmes décrit les liens causaux entre les dysfonctionnements des composants et relie donc entre elles les alarmes caractéristiques de ces dysfonctionnements; ce graphe est utilisé pour le filtrage des alarmes et des dysfonctionnements associés. Le graphe causal à long terme décrit les liens causaux entre les actions possibles sur le système ou les perturbations et leurs conséquences sur les variables du procédé. Il est utilisé dans une phase de diagnostic pour distinguer les événements dominants, supposés régir l'évolution du système, et les événements masqués.

Cayrac [33, 34] s'appuie sur l'existence d'une connaissance des relations entre les défaillances et leurs effets sous forme d'AMDEC pour construire des graphes causaux bipartites proches de ceux décrits dans [85]; le formalisme qu'il propose permet de prendre en compte l'incomplétude (oubli de certains effets; absence de certaines observations) et l'incertitude (plus ou moins grande certitude quant aux liens entre défaillances et effets) et de classer ainsi les hypothèses de diagnostic proposées. Ce travail a été effectué dans le cadre d'une application de surveillance de satellites chez Matra Marconi Space. Quelques primitives temporelles ont été intégrées au modèle relationnel afin de tenir compte des aspects dynamiques du système surveillé.

## 4 Conclusion

Pour conclure cette présentation générale des méthodes statistiques de traitement du signal et des méthodes de l'intelligence artificielle pour la surveillance et le diagnostic de systèmes dynamiques, nous avons choisi de dégager quatre points qui nous semblent importants et qui nous permettront d'évaluer et de comparer les différentes solutions proposées à ce problème de surveillance, ce que nous faisons dans la conclusion de cet article. Ce sont :

1. Les difficultés liées à l'acquisition des informations nécessaires, en particulier des modèles,
2. La capacité à prendre en compte l'incertain et l'imprécision, problème incontournable lorsque l'on veut traiter des systèmes réels,
3. La généralité des outils proposés et leur capacité à évoluer selon l'évolution du système surveillé,
4. La validation (ou l'évaluation) des résultats obtenus.

Nous donnons ainsi quelques éléments caractérisant les approches numériques et symboliques présentées ci-dessus.

## 4.1 Détection/Interprétation

Une première réflexion très générale s'impose : les méthodes statistiques du traitement de signal sont par nature proches du système surveillé, puisqu'elles travaillent directement à partir des données issues des capteurs; elles sont ainsi principalement utilisées pour la génération d'alarmes; les méthodes de l'intelligence artificielle sont, elles, plus tournées vers la communication avec l'opérateur et se focalisent plus sur la transformation d'un ensemble d'information brutes et non reliées entre elles en une information interprétable directement par l'opérateur chargé de la conduite; elles sont donc utilisées pour l'interprétation des alarmes. Cette distinction est bien entendu simpliste et à moduler : comme nous l'avons vu, la génération des alarmes dites intelligentes comprend une part d'interprétation et, d'autre part, certains modèles symboliques sont utilisés en simulation pour la détection et la génération d'alarmes.

## 4.2 Avantages/inconvénients des approches numériques

Pour en discuter, nous reprenons les quatre points mentionnés précédemment.

### 4.2.1 Acquisition des modèles

Les méthodes numériques reposent, par définition, sur l'utilisation de modèles. Nous ne revenons pas ici sur la distinction entre plusieurs types de modèles faite en 2.2. Nous discutons par contre du choix de modèles pour la surveillance et de l'acquisition (des paramètres  $\theta_0$ ) du modèle qui sert de référence. Même si plusieurs commentaires que nous faisons ici concernent aussi les méthodes géométriques de génération des résidus à base d'observateurs mentionnées en 2.4.1, nous discutons principalement des méthodes numériques statistiques. On renvoie par exemple à [78] pour une discussion de la robustesse des méthodes géométriques vis-à-vis de la (mé)connaissance des modèles.

De manière générale, on peut dire que la pertinence d'un modèle dépend de l'usage que l'on veut en faire. Pour un même système, et selon que l'on veut le simuler, le prédire, le commander, ou le surveiller, le "bon" modèle n'est pas le même, tout simplement parce que les aspects auxquels on a besoin d'être sensible ne sont pas les mêmes. Pire : le critère pour valider le "bon" modèle dépend aussi de l'objectif d'utilisation. En effet, le critère pour accepter et valider le dit modèle de référence doit être cohérent avec le critère qui gouverne l'algorithme de décision : c'est le prix à payer pour l'aptitude à traiter de manière cohérente de grands ensembles de données, question incontournable pour la surveillance en fonctionnement [70]. D'où l'affirmation : *le choix d'un modèle et du critère pour le valider doit être gouverné par l'objectif de traitement*. Et les Automaticiens qui font de la synthèse d'observateurs pour la détection de pannes savent très bien que leurs observateurs ne sont pas les mêmes que ceux qui servent à estimer tout ou partie de l'état du système.

On a déjà indiqué en 2.2 que des modèles de réseaux de neurones ou d'ondelettes [106, 92, 58], souvent considérés comme représentations non paramétriques, peuvent néanmoins être traités comme modèles paramétriques, dans lesquels le vecteur de paramètres est l'ensemble (fini!) des poids et éventuellement des coefficients de translation et dilatation. On peut construire alors, de la même manière que pour les modèles paramétriques plus standard, des détecteurs<sup>14</sup>. Une question intéressante en pratique, lorsque l'on veut surveiller un système sur lequel on n'a que très peu d'informations physiques, est de savoir si l'on a intérêt à se creuser la tête pour imaginer un modèle physique approché, ou bien à utiliser un tel modèle non paramétrique. C'est ce que a été essayé pour la surveillance des turbines à gaz en [70]. La réponse à la question n'est pas évidente du tout, et on rejoint ici la remarque précédente sur le "bon" modèle : un réseau d'ondelettes semble être plus pertinent pour simuler des signaux réalistes; par contre, pour la surveillance, il pose plusieurs problèmes de manque d'interprétation physique (et c'est ennuyeux pour le diagnostic), de manque de robustesse au changement de mode de fonctionnement, et aussi de dimension (typiquement 20 paramètres pour chacun des 18 thermocouples, alors que le modèle physique approché qui traite les 18 signaux simultanément n'a que ... 13 paramètres). Or on sait que plus un modèle est compliqué, plus il est difficile à surveiller.

Si la modélisation de composants (pas forcément très élémentaires, par exemple une machine entière) est souvent réalisable, un problème délicat concerne la modélisation d'un procédé entier (par exemple, un haut-fourneau ou une raffinerie), qui n'est pas toujours disponible, et revêt un caractère mixte numérique/symbolique marqué. On peut dire qu'une des voies de recherche en matière de surveillance de systèmes dynamiques concerne précisément la conception d'algorithmes manipulant de manière conjointe des modèles de natures et de "granulométries" différentes. En particulier, est encore largement ouverte la question de la conception d'algorithmes

14. Il est important de remarquer que ces détecteurs fonctionnent d'une manière essentiellement différente de celle des méthodes de surveillance à l'aide de réseaux de neurones qui travaillent par comparaison entre une sortie mesurée et une sortie simulée du réseau.

de surveillance/diagnostic prenant en compte toutes les connaissances, notamment en termes de modélisation et de relations de cause à effet, contenues dans les analyses de risque et de fiabilité généralement disponibles pour les systèmes industriels complexes (e.g. AMDEC). Nous renvoyons à [91, 97, 98] pour des utilisations de plusieurs types de connaissance.

Nous avons discuté jusqu'ici essentiellement du choix et de l'acquisition du modèle de référence qui représente le système surveillé à l'état sain et en fonctionnement normal. Nous terminons par un commentaire sur la représentation des différents états de pannes entre lesquels on veut être capable de discriminer. Une situation idéale consiste à disposer (des valeurs des paramètres) des modèles représentant le système dans chacun de ces états, et à faire, un peu comme en reconnaissance de formes, de la discrimination entre eux. Cette situation est bien évidemment rarement rencontrée en pratique, parce que la plupart des systèmes à surveiller sont trop coûteux, et/ou trop critiques du point de vue de la sécurité, pour qu'on puisse envisager d'y injecter des pannes. Comme on l'a déjà indiqué, l'information de base qui influe sur la conception de détecteurs concerne la manière dont les pannes se manifestent sur le modèle du système surveillé. Nous revenons plus loin sur ces questions d'informations disponibles concernant les pannes lorsque nous discutons la validation des résultats obtenus par l'algorithme de surveillance.

#### 4.2.2 Prise en compte de l'incertain

Rappelons que nous avons discuté brièvement en 1.1.1 l'introduction des bruits dans les modélisations dynamiques, à l'aide de variables aléatoires dont la loi de probabilité est souvent supposée Gaussienne.

L'un des avantages des méthodes statistiques décrites à la section 2 est qu'elles prennent en compte, de manière intrinsèque, d'une part les problèmes de calibration des capteurs et de bruits, et, d'autre part, l'incertitude qui entache l'acquisition de la référence  $\theta_0$ , quelle que soit d'ailleurs la qualité de la méthode d'identification utilisée [63, 21]. Cet avantage doit cependant être considéré de manière réaliste: il ne signifie pas que l'on peut détecter des déviations faibles sur un système caractérisé de manière imprécise, et instrumenté à l'aide de capteurs délivrant des signaux quantifiés grossièrement et de longueur très courte! Néanmoins, cette prise en compte de l'incertain par le biais de lois de probabilités de bruits est un moyen d'introduire une certaine "souplesse" dans l'algorithmique de surveillance, en particulier parce qu'elle permet de régler des seuils d'une manière souvent peu critique et relativement indépendante des conditions expérimentales.

L'exemple le plus surprenant est sans doute celui de la segmentation automatique de la parole rapporté en [4, 19], où le réglage des paramètres de l'algorithme est indépendant de la fréquence d'échantillonnage, du niveau de bruit (même pour un enregistrement radio-moblie), du locuteur (homme ou femme), et même de l'ordre du modèle auto-régressif utilisé (et ceci renvoie aux commentaires sur la robustesse vis-à-vis des erreurs de modèles); comparer les figures 7 et 8 aux figures 9 et 10.

Un autre exemple intéressant est celui de la surveillance des chambres de combustion d'une turbine à gaz, où l'une des difficultés réside dans le fait que la turbine change fréquemment de mode de fonctionnement, ce qui se traduit, sur les signaux mesurés par les 18 thermocouples, par des instationnarités (voir première partie de la figure 11) plus importantes que celles provoquées par les chutes de température à détecter (voir première partie de la figure 12). Or l'idée générale de la surveillance en fonctionnement discutée ici consiste à acquérir une caractérisation de référence de temps à autre (par exemple pour tenir compte du vieillissement du système surveillé), mais représentative de différents modes de fonctionnement de ce système. Ceci s'obtient par identification en utilisant des données qui mélangent le plus possible d'états de fonctionnement normaux du système (par exemple, différentes charges thermiques pour une turbine à gaz [15, 70], mais aussi différents états de houle pour une structure off-shore [84, 83], différentes vitesses de rotation pour une machine tournante [84, 101]). La deuxième partie de la figure 12 montre qu'il est possible d'effectuer une détection de chute de température qui soit robuste vis-à-vis de ces changements de modes de fonctionnement. Cette robustesse s'obtient par le biais d'une définition appropriée des hypothèses (d'absence et de présence de pannes) entre lesquelles on teste. L'hypothèse d'absence de pannes est non pas restreinte à  $\theta = \theta_0$ , mais élargie à  $\theta \in \Theta_0$ , où  $\Theta_0$  est un "ellipsoïde" de confiance autour de  $\theta_0$  dont la taille est choisie par apprentissage sur les données. L'hypothèse de présence de pannes est elle aussi de la forme  $\theta \in \Theta_1$ , où  $\Theta_1$  est un autre "ellipsoïde" autour de  $\theta_0$ , dont la taille est choisie par apprentissage sur les données en fonction du cahier des charges de la surveillance.

#### 4.2.3 Généricité des outils

Il est clair que le caractère universel de la transformation d'une grande classe de problèmes de surveillance de processus dynamiques en le problème de la détection d'un changement de la moyenne d'un vecteur (statique)

Gaussien<sup>15</sup> confère aux algorithmes de détection qui en résultent une genericité importante par la possibilité de calibration des seuils : les statistiques de test ont une loi connue et tabulée, celle du  $\chi^2$ .

Pour ce qui est de la fabrication des alarmes intelligentes discutées en 2.5, c'est-à-dire des alarmes contenant une information relative au diagnostic en termes du "gros" modèle physique sous-jacent, il convient de souligner que celle qui concerne la surveillance des vibrations en [14, 42] présente un caractère générique important, puisque seule l'interface avec le modèle mécanique est à refaire pour chaque application (machine, structure).

#### 4.2.4 Validation des résultats

Comme on l'a déjà indiqué, l'une des principales difficultés de la validation des résultats obtenus par une algorithmique de surveillance réside dans le fait que, pour beaucoup de systèmes réels à surveiller, on dispose de très peu (voire pas du tout, et c'est tant mieux!) de données correspondant à des pannes effectives. On doit donc souvent se contenter d'une validation par simulations effectuées à l'aide d'un simulateur plus ou moins réaliste de l'application envisagée, à moins que l'on ait la chance de disposer d'une maquette, plate-forme ou site expérimental; encore faut-il que ce dernier puisse supporter lui aussi des pannes non triviales à détecter. Une autre approche expérimentale peut être envisagée dans certains cas, celle qui consiste à utiliser la forme d'un modèle simplifié de l'application pour injecter, sur des données réelles enregistrées en l'absence de pannes, des pannes réalistes et répondant au cahier des charges de la surveillance; c'est ce qui est fait en [15, 70] pour les turbines à gaz.

### 4.3 Avantages/inconvénients des approches symboliques

#### 4.3.1 Acquisition des modèles

Les méthodes symboliques présentées reposent sur des modèles que nous avons présentés en 3.2 sous trois grandes classes : les modèles associatifs, prédictifs et explicatifs. L'acquisition de ces modèles pose des problèmes spécifiques à chacun de ces trois types.

**Les modèles associatifs.** Les modèles associatifs décrivent les liens existant entre symptômes et pannes; ils comportent donc essentiellement des connaissances expertes relatives aux dysfonctionnements du système. Cette expertise est le plus souvent acquise par extraction de connaissances auprès des experts du domaine (opérateurs de conduite; ingénieurs de maintenance). Ainsi dans le projet SACHEM (système d'aide à la conduite de haut-fourneaux en cours de réalisation chez Sollac), le nombre d'experts interviewés est de l'ordre de 15, le travail de recueil d'expertise est estimé à environ 250 journées d'interviews; environ 1000 journées ont été nécessaires aux cognitivistes chargés de conduire les interviews, puis de les retranscrire en une base de connaissances. Celle-ci rassemble les connaissances devant permettre d'estimer à partir des événements signaux les événements haut-fourneaux qui caractérisent l'état global du système et suffisent à l'opérateur de conduite à évaluer la situation et à décider d'éventuelles actions. Dans les approches de type reconnaissance de scénarios, l'acquisition des scénarios peut être envisagée de la même manière; c'est ce qui a été fait par exemple dans le projet TIGER pour la surveillance de turbines à gaz à partir d'IxTeT [44].

La difficulté d'acquisition de l'expertise n'est pas propre aux systèmes dynamiques; les méthodologies d'acquisition de connaissances telles que KADS [102] et les outils existants de validation des bases de connaissances [6] peuvent donc être utilisés. La seule différence notable est l'existence de contraintes temporelles qu'il faut donc pouvoir exprimer dans le formalisme choisi et qui pose des problèmes spécifiques de validation. Dans les techniques de reconnaissance de scénarios, le test de leur cohérence est en général effectué par des algorithmes classiques de gestion de contraintes temporelles. Un autre point délicat est le test de la complétude des connaissances expertes ainsi acquises : couvrent-elles bien tous les cas de dysfonctionnements possibles?

Un autre point déjà mentionné est celui du choix du bon niveau d'abstraction (ou de granularité) du modèle. Ce point est important, quel que soit le type de modèles; dans les modèles associatifs, le niveau de granularité est induit par le langage utilisé par l'expert concerné et correspond donc la plupart du temps au niveau qui est le bon pour l'analyse de la situation et la prise de décision. En revanche, un problème est laissé de côté qui est

---

15. À notre connaissance, il existe à l'heure actuelle (au moins) trois ensembles de travaux permettant d'effectuer une telle transformation, et donc de fabriquer des tests du  $\chi^2$  sur des transformations non-linéaires des signaux observés. Ce sont les travaux [20, 105, 19] dits d'approche asymptotique locale, s'appuyant sur des fonctions souvent alternatives à la vraisemblance, et concernant des processus semi-Markoviens; les travaux autour des notions de quasi-vraisemblance ou pseudo-vraisemblance [71] concernant des semi-martingales, et repris pour la détection par Basawa [7]; les travaux récents [35], s'appuyant sur les processus  $\alpha$ -mélangeants, et généralisant des tests non-paramétriques de conformité classiques en statistique et des tests récents de non-gaussianité de Giannakis. Voir les discussions de [60].

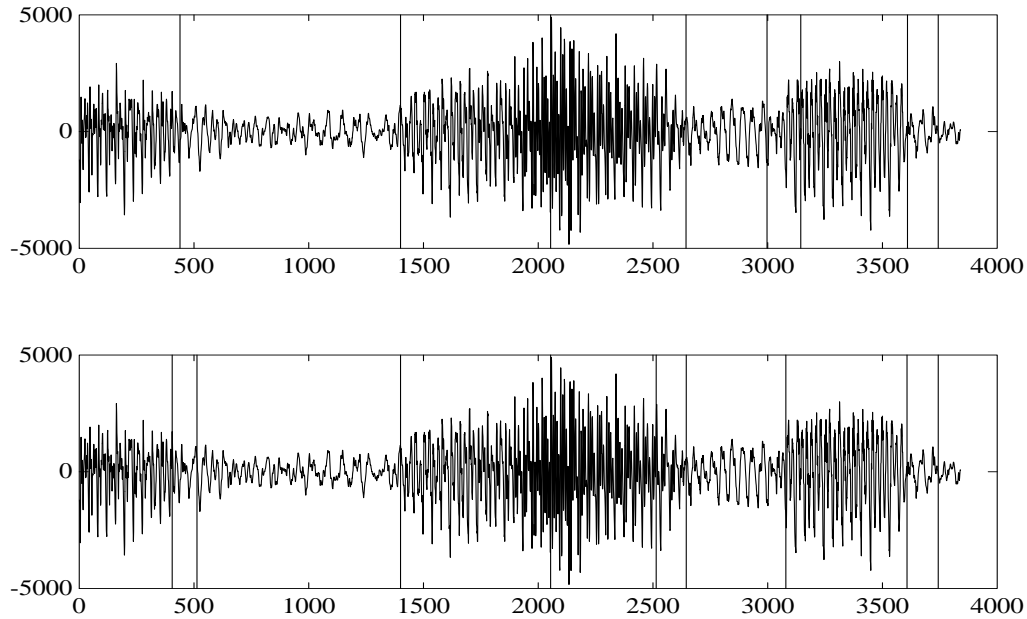


FIG. 7 – Segmentation automatique du signal de parole filtré correspondant au signal bruité de la figure 9, avec ordre AR 16 (signal du haut) et ordre AR 2 (signal du bas). Les traits verticaux indiquent les frontières des segments trouvés par l’algorithme.

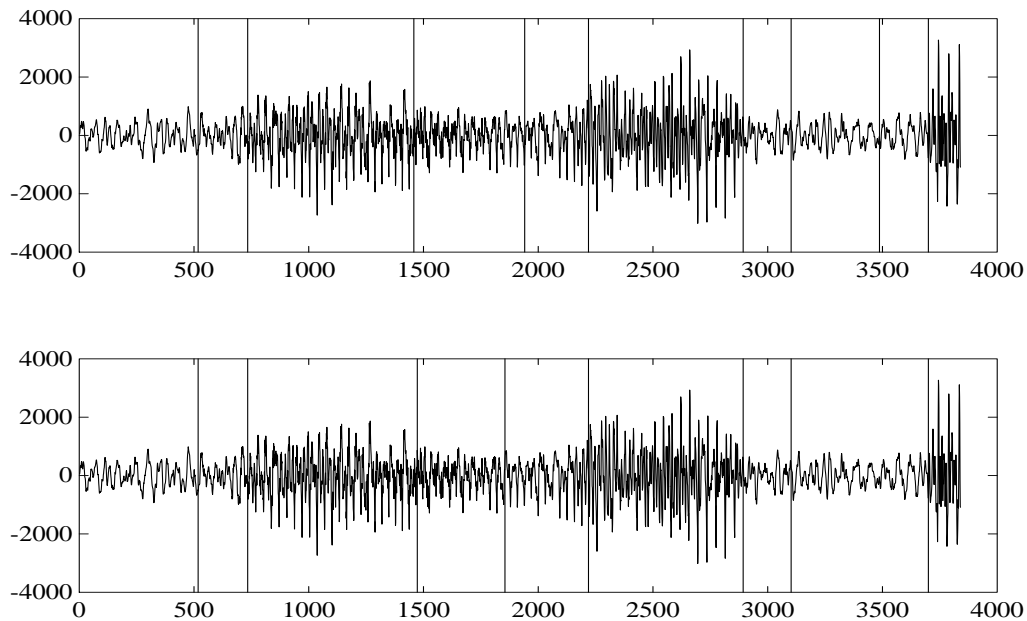


FIG. 8 – Segmentation avec différents ordres (suite.): le signal de parole filtré correspondant au signal bruité de la figure 10.

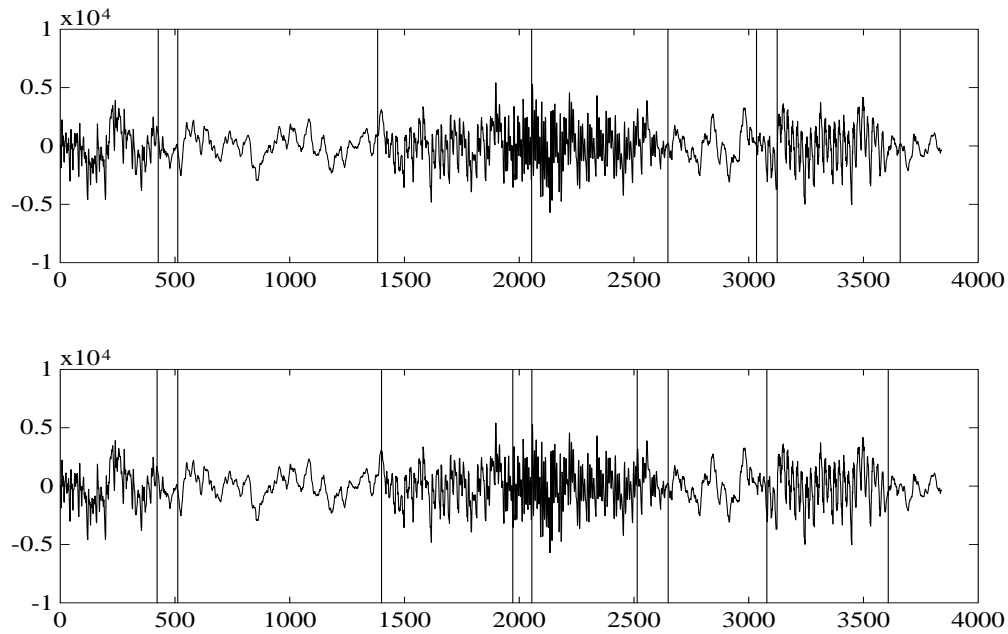


FIG. 9 – Segmentation du signal de parole bruité (enregistrement radio-mobile), avec ordre AR 16 (signal du haut) et ordre AR 2 (signal du bas). Les traits verticaux indiquent les frontières des segments trouvés par l'algorithme.

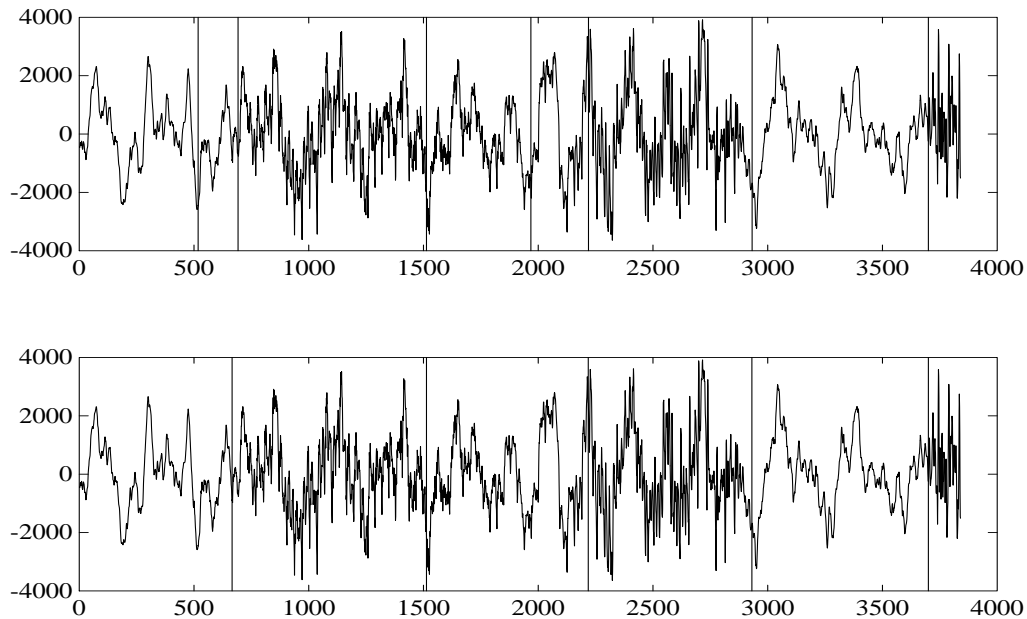


FIG. 10 – Segmentation avec différents ordres : le signal de parole bruité.

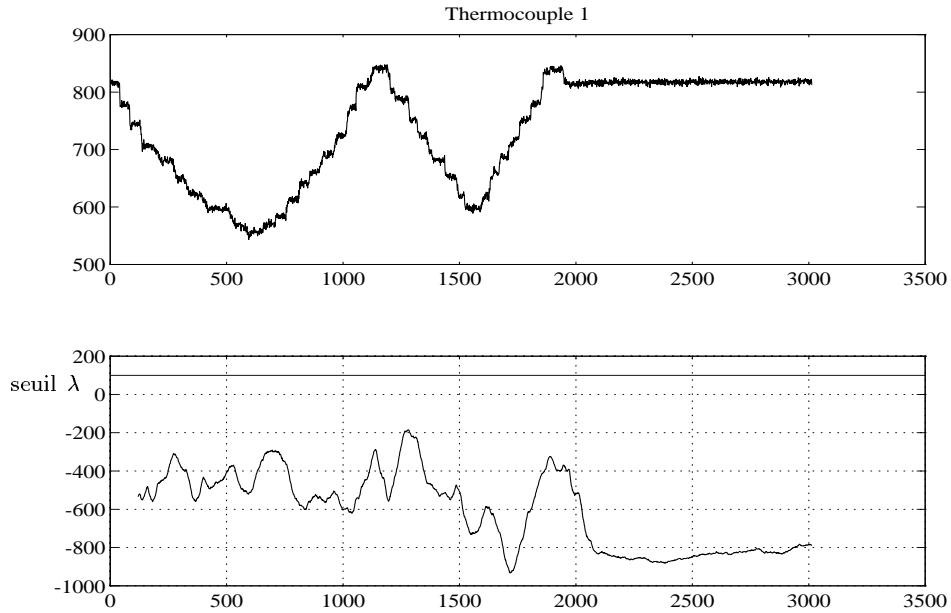


FIG. 11 – *Comportement du test en l'absence de pannes - Noter la robustesse par rapport aux non-stationnarités dues aux changements de modes de fonctionnement: la statistique de test (traitement séquentiel et multi-capteurs) reste négative et ne franchit pas le seuil positif.*

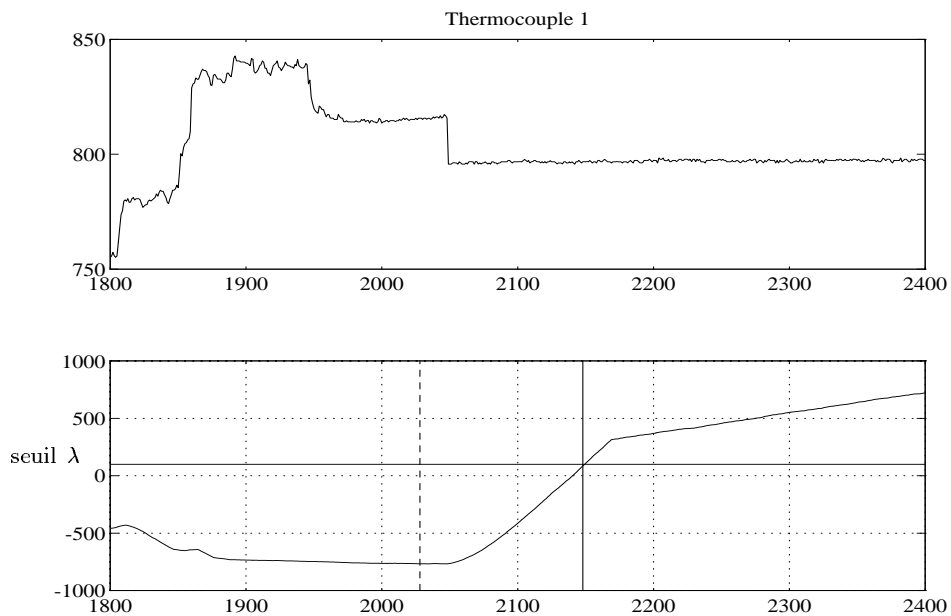


FIG. 12 – *Détection d'une baisse anormale de température de 20 K dans le premier thermocouple à l'instant 2050 - Noter que l'amplitude de ce changement est faible par rapport aux non-stationnarités, et que la statistique de test (traitement séquentiel et multi-capteurs) reste bien négative avant la chute de température.*



celui du passage entre le signal et la description en termes d'événements significatifs, et en particulier la phase d'élaboration des symptômes.

L'acquisition des modèles associatifs repose donc sur l'existence d'expertise et donc la disponibilité d'experts prêts à accoucher de leurs connaissances. Une autre possibilité peut être d'obtenir cette connaissance en transformant une connaissance présente sous une autre forme en connaissance associative ou en les apprenant à partir de données réelles ou simulées; c'est ce qui est proposé dans [23] où les associations sont construites à partir de données obtenues par simulation de pannes sur un modèle prédictif; c'est ce qui est fait aussi dans [68] où les scénarios sont appris à partir d'observations réelles. La phase d'apprentissage pour être utilisable nécessite de travailler à partir d'exemples étiquetés par un expert (apprentissage supervisé), ce qui laisse entier le problème d'existence et de la disponibilité d'un expert. En cas contraire, on obtient des modèles boîte-noire tels que ceux utilisés dans les approches numériques qui peuvent difficilement être considérés comme des modèles symboliques en raison de leur faible capacité à expliquer les phénomènes.

Une autre caractéristique des modèles associatifs est qu'ils reposent sur une bonne connaissance des phénomènes de dysfonctionnements. Ceci est en général le cas lorsqu'il s'agit de déviations courantes que l'opérateur de conduite a l'habitude de repérer. Ce n'est pas du tout le cas lorsqu'il s'agit de dysfonctionnements graves qui sont a fortiori peu courants et même heureusement inexistantes et qu'il s'agit justement de prévenir; les incidents majeurs dans les centrales nucléaires ou thermiques font clairement partie de cette classe.

**Les modèles prédictifs.** Les modèles prédictifs décrivent le comportement du système dans son état normal et éventuellement dans ses états de dysfonctionnement. Ils peuvent être vus comme des "simulateurs symboliques". La construction de tels modèles nécessite donc une bonne connaissance du comportement du système. Les schémas de conception sont une des sources de connaissances privilégiées.

Dans les modèles qualitatifs, l'élaboration des équations qualitatives se fait en général en s'appuyant sur les équations numériques lorsqu'elles sont disponibles. Il est souvent relativement facile de décrire sous forme de graphes d'influence les dépendances entre les variables qualitatives retenues, même lorsqu'il s'agit d'une grosse installation pour laquelle la construction du modèle numérique global est difficilement envisageable [32, 26, 75].

Une des difficultés essentielles est le choix du bon niveau d'abstraction et construire un tel modèle est souvent une étape délicate [40, 94]. Une des directions courantes de recherche consiste à combiner l'utilisation de modèles à différents niveaux d'abstraction ou sous différents points de vue [51, 36, 76].

**Les modèles explicatifs.** Les modèles explicatifs décrivent les relations causes-effets existant entre les pannes, les événements pouvant se produire sur le système et leurs effets observables. L'acquisition d'un tel graphe causal pose le même genre de problèmes que celui de l'expertise nécessaire pour les modèles associatifs. Il nécessite une bonne expérience du système et en particulier de son comportement en cas de dysfonctionnement. Le point spécifique des systèmes dynamiques est le fait que les arcs causaux sont étiquetés par des contraintes temporelles (voir le système DIAPO par exemple [81]) et que l'acquisition et le contrôle de la cohérence de ces contraintes est un point délicat.

#### 4.3.2 Prise en compte de l'incertain

Comme nous l'avons déjà souligné, un des points spécifiques aux systèmes dynamiques dans l'ensemble des méthodes proposés par l'IA est l'expression de contraintes temporelles. Une des facettes de la prise en compte de l'incertain concerne ainsi les contraintes temporelles. En général, l'incertitude sur les dates ou sur les délais (par exemple entre les causes et les effets) est traduite par l'utilisation d'une représentation à base d'intervalles; une date ou un délai est alors représenté par un intervalle dont la longueur correspond directement à l'imprécision sur la valeur. Dans certains travaux tels que [75], les intervalles sont des intervalles flous. Le contrôle de la cohérence des contraintes temporelles ainsi que leur propagation doit donc tenir compte de cette imprécision en contrôlant ou propageant des intervalles et non des valeurs exactes. Ce point se retrouve dans les différents types de modèles que nous avons évoqués.

Une autre source d'incertitude est celle existant directement dans les modèles symboliques utilisés, qui sont par nature des représentations abstraites du système supervisé, et par la même entachées d'incertitude et d'incomplétude. Un des problèmes bien connus est celui de la prise en compte de ce type d'incertitude dans les modèles associatifs. Le formalisme des systèmes experts permet ainsi souvent d'exprimer le degré d'évocation d'un lien effets-causes par une valeur numérique ou symbolique. Ces degrés d'évocation doivent ensuite être combinés entre eux afin de correspondre au mieux à la certitude associée aux conclusions obtenues. Une discussion sur ce type de gestion de l'incertain peut être trouvée dans [93]. De même dans les graphes causaux, il peut être souhaitable d'exprimer la plus ou moins grande certitude entre a et b lorsque l'on sait

que “a cause b”. Il est proposé dans CHECK [37] de différencier les liens forts (“a cause toujours b”) des liens faibles (“a peut causer b”). Une autre possibilité est d’utiliser la théorie des possibilités comme le fait [34] ou encore la théorie des probabilités. Dans les approches de type reconnaissance de scénarios, [43] suggère d’utiliser la théorie des ensembles flous pour affecter un degré de reconnaissance ou une préférence entre les scénarios possiblement reconnus.

Une dernière source d’incertitude concerne les observations elles-mêmes, en particulier lorsqu’elles sont issues de capteurs sujets à erreurs et/ou imprécisions. Ce point est rarement traité par les approches à base de modèles symboliques, à l’exception des approches de type simulation qualitative.

La simulation de modèles prédictifs met particulièrement bien en valeur les problèmes liés à la prise en compte de l’incertain lorsqu’il s’agit de comparer des observations prédites (par un modèle approximatif) et des observations acquises par l’intermédiaire de capteurs imprécis. On retrouve exactement ce problème dans le cas de la simulation qualitative faite sur des graphes d’influence (voir 2.6.1). Dans le système DIAPASON [75], les arcs du graphe d’influence sont étiquetés par des fonctions de transfert qualitatives exprimant les liens existant entre les variables qualitatives. L’incertitude sur la valeur des variables observées ainsi que sur les fonctions de transfert utilisées est exprimée en associant aux variables ainsi qu’aux coefficients des fonctions des intervalles de valeurs plutôt que des valeurs précises, et plus précisément des intervalles flous. La propagation des valeurs des variables à travers le graphe se fonde sur la théorie des possibilités. Dans le système CA-EN [27], le temps est discrétisé et la simulation qualitative s’appuie sur la propagation d’intervalles à chaque instant. Il fournit des enveloppes (bornes supérieures et inférieures des valeurs à chaque instant) pour les variables auxquelles on s’intéresse. Les valeurs observées sont alors comparées par rapport à ces enveloppes proposées par le simulateur.

Pour les systèmes à événements discrets, une extension des réseaux de Petri aux réseaux de Petri flous [31] a été proposée et permet ainsi de prendre en compte l’incertitude.

### 4.3.3 Généricité des outils

La généricité dépend très directement des types d’approches employées. Dans les approches à base de modèles associatifs, il est clair que l’expertise est remise en cause en cas de changement du système surveillé et même en cas de modification de celui-ci. Il est difficile a priori de connaître les répercussions d’une modification sur l’ensemble de l’expertise. C’est ce point qui est souvent reproché à ces approches et qui a eu pour conséquence le développement d’outils d’acquisition de l’expertise (voir 4.3.1) ainsi que d’acquisition automatique utilisant des techniques d’apprentissage; c’est le cas en particulier pour l’apprentissage de scénarios dans les approches de type reconnaissance de scénarios [68]. Le problème de généricité est similaire en ce qui concerne les graphes causaux, puisqu’ils requièrent une bonne connaissance du comportement du système en cas de dysfonctionnement, et que celui-ci peut varier du tout au tout en cas de changement de celui-ci. Dans ces deux cas, les outils (tels que les moteurs d’inférence) exploitant la connaissance contenue dans les modèles restent bien entendu valables.

Un des points couramment mis en avant pour justifier les approches à base de modèles prédictifs (souvent appelées approches à base de modèles) est justement celui de leur généricité. Le modèle donne une description du fonctionnement du système et sa construction peut s’appuyer sur les schémas de conception existants. En cas de modification d’un composant du système, il est relativement facile de modifier la description de celui-ci sans avoir à reconsidérer l’ensemble du modèle. Il est cependant reconnu que la construction du modèle reste une activité difficile [40]. Une solution consiste à constituer une bibliothèque de composants élémentaires et d’utiliser ces descriptions génériques pour construire par composition le modèle global du système. C’est ce qui est mis en avant dans les travaux utilisant les bond-graphs [46, 76, 24]. C’est aussi ce qui est proposé dans [22, 89] qui utilisent la composition synchrone d’automates.

### 4.3.4 Validation des résultats

La validation des résultats pose des problèmes très proches de ceux rencontrés pour les modèles numériques (voir 4.2.4). En l’absence d’un simulateur permettant de simuler le comportement du système en cas de pannes (et en particulier de pannes multiples), il est difficile d’estimer la qualité des diagnostics fournis par les systèmes de surveillance. La confrontation des diagnostics avec ceux émis par les experts et opérateurs de conduite est une des solutions les plus couramment employées.

## 4.4 Complémentarité des deux types d’approches

Dans cet article, nous avons présenté les méthodes et outils proposés par le traitement de signal et l’intelligence artificielle pour la surveillance et le diagnostic de systèmes dynamiques. Comme nous l’avons déjà

mentionné, les méthodes du traitement de signal sont, par nature, proches du système surveillé, et ainsi principalement utilisées pour la génération d'alarmes; alors que les méthodes de l'intelligence artificielle sont plus tournées vers la communication avec l'opérateur, et ainsi plus utilisées pour les phases d'interprétation des alarmes et de diagnostic. Cependant, la description des diverses méthodes et de leurs propriétés fait évidemment apparaître des préoccupations communes, telles la définition d'un cahier des charges de la surveillance; le choix et l'acquisition de modèles pour la surveillance; l'utilisation, pour le diagnostic, des modèles de causes à effets établis à la conception; la conversion des signaux en événements significatifs, . . . , pour ne citer que celles-là.

Il est donc important de favoriser la concertation et les échanges de techniques et d'outils entre les communautés concernées par ces problèmes, qui correspondent à d'importants enjeux industriels, que ce soit en termes de sécurité des personnes et des installations industrielles, de protection de l'environnement, ou de gains de productivité.

## Références

- [1] P. Abry and P. Flandrin. Multiresolution transient detection. In *IEEE-SP Int. Symp. Time-frequency and time-scale analysis*, Philadelphia, VA, October 1994.
- [2] H. Akaike. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. *Annals Institute Statistical Mathematics*, 26:363–387, 1974.
- [3] P.O. Amblard, J.L. Lacoume, and J.M. Brossier. Transient detection, higher-order time-frequency distributions, and the entropy. In *IEEE Workshop on Higher Order Statistics*, Lake Tahoe, June 1993.
- [4] R. André-Obrecht. A new statistical approach for the automatic segmentation of continuous speech signals. *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP-36(1):29–40, January 1988.
- [5] K. E. Arzen. A survey of commercial real time expert system environments. In *International Symposium on Artificial Intelligence in Real Time Control*, pages 611–618, 1992.
- [6] M. Ayel and M.-C. Rousset. *La cohérence dans les bases de connaissances*. Cepadues, 1990.
- [7] M. Basawa. Neyman-le cam tests based on estimating functions. In L.M. Le Cam and R.A. Olshen, editors, *Proc. Berkeley Conf. in honor of Neyman and Kiefer*, volume 2, pages 811–825. Wadsworth Inc., 1985.
- [8] M. Basseville. Detecting changes in signals and systems - a survey. *Automatica*, 24(3):309–326, May 1988.
- [9] M. Basseville. Distance measures for signal processing and pattern recognition. *Signal Processing*, 18(4):349–369, December 1989.
- [10] M. Basseville. Information criteria for FDI. In *IFAC '96 World Conference*, San Francisco, CA, July 1996.
- [11] M. Basseville, A. Benhallam, C. Doncarli, M.F. Lucas, D. De Bruçq, O. Colot, H. Rix, E. Thierry, and F. Kauffmann. Fiches d'algorithmes de segmentation de signaux. *Traitement du Signal*, 9(1), 1992. Supplément au volume 9.
- [12] M. Basseville and A. Benveniste. Design and comparative study of some sequential jump detection algorithms for digital signals. *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-31(3):521–535, June 1983.
- [13] M. Basseville and A. Benveniste. Sequential detection of abrupt changes in spectral characteristics of digital signals. *IEEE Trans. on Information Theory*, IT-29(5):709–724, September 1983.
- [14] M. Basseville, A. Benveniste, Gach-Devauchelle, M. Goursat, D. Bonnecase, P. Dorey, M. Prevosto, and M. Olagnon. Damage monitoring in vibration mechanics : issues in diagnostics and predictive maintenance. *Mechanical Systems and Signal Processing*, 7(5):401–423, September 1993.
- [15] M. Basseville, A. Benveniste, G. Mathis, and Q. Zhang. Monitoring the combustion set of a gas turbine. In *Proc. Safeprocess'94*, Helsinki, Finland, June 1994.
- [16] M. Basseville, A. Benveniste, and G. Moustakides. Detection and diagnosis of abrupt changes in modal characteristics of nonstationary digital signals. *IEEE Trans. on Information Theory*, IT-32(3):412–417, May 1986.
- [17] M. Basseville, A. Benveniste, G. Moustakides, and A. Rougée. Detection and diagnosis of changes in the eigenstructure of nonstationary multivariable systems. *Automatica*, 23(4):479–489, July 1987.
- [18] M. Basseville, A. Benveniste, G. Moustakides, and A. Rougée. Optimal sensor location for detecting changes in dynamical behavior. *IEEE Trans. on Automatic Control*, AC-32(12):1067–1075, December 1987.
- [19] M. Basseville and I.V. Nikiforov. *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall Information and System Sciences Serie, Englewood Cliffs, NJ, 1993.
- [20] A. Benveniste, M. Basseville, and G. Moustakides. The asymptotic local approach to change detection and model validation. *IEEE Trans. on Automatic Control*, AC-32(7):583–592, July 1987.

- [21] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximations*. Springer Series on Applications of Mathematics, New-York, 1990.
- [22] S. Bibas, M.-O. Cordier, P. Dague, F. Lévy, and L. Rozé. Scenario generation for telecommunications networks. In *Proceedings IJCAI Workshop on AI in Distributed Intelligent Networks*, Montréal, 1995.
- [23] S. Bibas and L. Rozé. Modélisation et simulation pour la supervision de réseaux de télécommunications. In *Actes de RFIA '96*, pages 339–346, Rennes, 1996.
- [24] G. Biswas and X. Yu. A formal modeling scheme for continuous systems : Focus on diagnosis. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1474–1479, Chambéry, 1993.
- [25] A.A. Borovkov. *Statistique Mathématique. Estimation et tests d'hypothèses*. Mir, Paris, 1987.
- [26] K. Bousson. *Raisonnement causal pour la supervision de processus basée sur des modèles*. PhD thesis, LAAS-CNRS, Toulouse, 1993.
- [27] K. Bousson and L. Travé. Fuzzy causal simulation in process engineering. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1536–1541, Chambéry, 1993.
- [28] K. Bousson and L. Travé. Putting more numbers in the qualitative simulator CA-EN. In *Proceedings 2nd International Conference on Intelligent Systems Engineering*, Hamburg, 1994.
- [29] V. Brusoni, L. Console, P. Terenziani, and D. Theseider Dupré. An efficient algorithm for computing temporal abductive diagnoses. In *Proceedings of the International Workshop on Principles of Diagnosis(DX'95)*, Goslar(Germany), 1995.
- [30] V. Brusoni, L. Console, P. Terenziani, and D. Theseider Dupre. Characterizing temporal abductive diagnosis. In *Proceedings of the International Workshop on Principles of Diagnosis(DX'95)*, pages 34–40, Goslar(Germany), 1995.
- [31] J. Cardoso, R. Valette, and D. Dubois. Petri nets with uncertain markings. In *Advances in Petri Nets*, volume 483, pages 64–78. Springer Verlag, 1991.
- [32] S. Cauvin. *Un environnement générique à base de connaissances pour la supervision de procédés de raffinage et de pétrochimie*. PhD thesis, Thèse CNAM, 1995.
- [33] D. Cayrac. *Diagnostic opérationnel exploitant des connaissances incomplètes*. PhD thesis, Thèse Université Paul Sabatier, Toulouse, 1995.
- [34] D. Cayrac, D. Dubois, and H. Prade. Une méthode pratique de diagnostic de pannes “basée sur les modèles” avec prise en compte qualitative de l’incertitude. In *Actes de RFIA '96*, pages 320–329, Rennes, 1996.
- [35] K. Choukri and É. Moulines. Time-domain procedures for testing that a stationary time-series is gaussian. *IEEE Trans. Signal Processing*, SP-44, à paraître, 1996.
- [36] B. B. Chu and J. A. Reggia. Modeling diagnosis at multiple levels of abstraction. I. representing causal relations at multiple levels of abstraction. *International Journal of Intelligent Systems*, 6:617–644, 1991.
- [37] L. Console, L. Portinale, D. Theseider Dupré, and P. Torasso. A theory of diagnosis for incomplete causal models. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1311–1317, Milan, 1989.
- [38] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications, New-York, 1991.
- [39] D.R. Cox and D.V. Hinkley. *Theoretical Statistics*. Chapman and Hall, New-York, 1986.
- [40] P. Dague, P. Devès, and O. Raiman. Troubleshooting : when modeling is the trouble. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, pages 590–595, Seattle, 1987.
- [41] J. de Kleer and B. C. Williams. Diagnosis with behavioral modes. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1324–1330, Milan, 1989.

- [42] B. Devauchelle-Gach. *Diagnostic mécanique des fatigues sur les structures soumises à des vibrations en ambiance de travail*. PhD thesis, Université de Paris IX Dauphine, December 1991.
- [43] C. Dousson. *Suivi d'évolutions et reconnaissance de chroniques*. PhD thesis, Thèse Université de Toulouse, 1994.
- [44] C. Dousson, P. Gaborit, and M. Ghallab. Situation recognition: representation and algorithms. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 166–172, Chambéry, 1993.
- [45] D. Dvorak and B. Kuipers. Model-based monitoring of dynamic systems. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1238–1243, Milan, 1989.
- [46] B. Falkenhainer and K. Forbus. Compositional modeling: finding the right model for the job. *Artificial Intelligence*, 51:95–143, 1991.
- [47] K. D. Forbus. Qualitative process theory. In D. G. Bobrow, editor, *Qualitative Reasoning about Physical Systems*, pages 85–168, Cambridge, MA, 1985. MIT Press.
- [48] P.M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge based redundancy - a survey and new results. *Automatica*, 26:459–474, 1990.
- [49] B. Friedlander. Detection of transient signals by the gabor representation. *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP-37(2):169–180, February 1989.
- [50] M. Frisch and H. Messer. Transient signal detection using prior information in the likelihood test. *IEEE Trans. Signal Processing*, ASSP-41(6):2177–2191, June 1993.
- [51] M. Gallanti, M. Roncato, A. Stefanini, and G. Torielli. A diagnostic algorithm based on models at different levels of abstraction. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1350–1355, Milan, 1989.
- [52] M. Gallanti, A. Stefanini, and L. Tomada. ODS: A diagnostic system based on qualitative modeling techniques. In *Proc. of the Fifth Conference on Artificial Intelligence Applications CAIA-89*, pages 141–149, Miami, FL, 1989.
- [53] GENSYM. G2 reference manual version 3.0, 1993.
- [54] B. Hayes-Roth. A blackboard architecture for control. *Artificial Intelligence*, 26:251–321, 1985.
- [55] R. Isermann. Process fault detection based on modeling and estimation methods - a survey. *Automatica*, 20:387–404, 1984.
- [56] R. Isermann. Fault diagnosis of machines via parameter estimation and knowledge processing - tutorial paper. *Automatica*, 29(4):815–836, 1993.
- [57] Special Issue. Knowledge compilation: a symposium. *IEEE Expert*, 6, 2:71–93, 1991.
- [58] A. Juditsky, H. Hjalmärsson, A. Benveniste, B. Delyon, L. Ljung, J. Sjöberg, and Q. Zhang. Nonlinear black-box modelling in system identification: mathematical foundations. *Automatica*, 31(12):1725–1750, December 1995. Special Issue on Trends in System Identification.
- [59] S. Kullback. *Information Theory and Statistics*. Wiley, New-York, 1959. (also Dover, New-York, 1968).
- [60] T.L. Lai. Sequential changepoint detection in quality control and dynamical systems. *Jal Royal Statistical Society B*, 57:613–658, 1995.
- [61] H. Laasri and B. Maître. *Coopération dans un univers multi-agents basé sur le modèle du blackboard : études et réalisations*. PhD thesis, Thèse de l'université de Nancy 1, 1989.
- [62] L. Leyval and S. Gentil. On line event-based simulation through a causal graph. In *IMACS Workshop on Qualitative reasoning and decision support systems*, pages 209–214, Toulouse, 1991.
- [63] L. Ljung. *System Identification - Theory for the User*. Prentice Hall Information and System Sciences Series, Englewood Cliffs, NJ, 1987.

- [64] L. Ljung and T. Glad. On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265–276, March 1994.
- [65] X.C. Lou, A.S. Willsky, and G.C. Verghese. Optimally robust redundancy relations for failure detection in uncertain systems. *Automatica*, 22(3):333–344, 1986.
- [66] B. Lucas. *Méthode d'aide à la modélisation par graphes de liaison et utilisation pour le diagnostic qualitatif de systèmes physiques*. PhD thesis, Thèse Université Paris-Sud, Orsay, 1994.
- [67] J.F. Magni and P. Mouyon. On residual generation by observer and parity space approaches. *IEEE Trans. on Automatic Control*, AC-39(2):441–447, 1994.
- [68] F. Marois. Apprentissage de scénarios. Technical report, LAAS, Toulouse, 1995. Mémoire CNAM.
- [69] M.A. Massoumnia, G.C. Verghese, and A.S. Willsky. Failure detection and identification. *IEEE Trans. Automatic Control*, AC-34(3):316–321, 1989.
- [70] G. Mathis. *Outils de Détection de Rupture et de Diagnostic: Application à la Surveillance de Turbines à Gaz*. PhD thesis, Université de Rennes I, September 1994.
- [71] D.L. McLeish and C.G. Small. *The Theory and Applications of Statistical Inference Functions*, volume 44 of *Lecture Notes in Statistics*. Springer-Verlag, New-York, 1988.
- [72] R.K. Mehra. Optimization of measurement schedules and sensor design for linear dynamic systems. *IEEE Trans. on Automatic Control*, AC-21(1):55–64, 1976.
- [73] R. Milne and L. Travé. On qualitative reasoning. IJCAI Tutorial, 1993.
- [74] P.C. Müller and H.I. Weber. Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems. *Automatica*, 8:237–246, 1972.
- [75] J. Montmain and S. Gentil. Decision-making in fault detection: a fuzzy approach. In *Proceedings of the International Conference on Fault Diagnosis Tooldiag'93*, pages 653–659, Toulouse, 1993.
- [76] P. P. Nayak, L. Joskowicz, and S. Addanki. Automated model selection using context-dependent behaviors. In *Proceedings of the AAAI National Conference on Artificial Intelligence*, pages 710–716, San José, CA, 1992.
- [77] R.N. Nikoukhah. Innovations generation in the presence of unknown inputs: application to robust failure detection. *Automatica*, 30(12):1851–1868, 1994.
- [78] R.J. Patton. Robust model-based fault diagnosis: the state of the art. In *Proc. Safeprocess'94*, Helsinki, Finland, 1994.
- [79] R.J. Patton, P.M. Frank, and R.N. Clark, editors. *Fault Diagnosis in Dynamic Systems: Theory and Application*. Prentice Hall International, UK, 1989.
- [80] B. Porat and B. Friedlander. Adaptive detection of transient signals. *IEEE Trans. Acoustics, Speech and Signal Processing*, ASSP-34(12):1410–1418, December 1986.
- [81] M. Porcheron, B. Ricard, J. L. Busquet, and P. Parent. DIAPO: A case study in applying advanced AI techniques to the diagnosis of a complex system. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, Amsterdam, 1994.
- [82] L. Portinale. Exploiting t-invariant analysis in diagnostic reasoning on a Petri net model. In *LNCS 691, Proceedings of Application and Theory of Petri Nets*, 1993.
- [83] M. Prevosto. *Algorithmes d'Identification des Caractéristiques Vibratoires de Structures Mécaniques Complexes*. PhD thesis, Université de Rennes I, March 1982.
- [84] M. Prevosto, M. Olagnon, A. Benveniste, M. Basseville, and G. Le Vey. State-space formulation, a solution to modal parameter estimation. *Journal of Sound and Vibration*, 148(2):329–342, July 1991.
- [85] J. A. Reggia, D. S. Nau, and P. Y. Wang. Diagnostic expert systems based on a set covering model. *International Journal of Man-Machine Studies*, 19(5):437–460, 1983.

- [86] R. C. Rosenberg and D. C. Karnopp. *Introduction to Physical Systems Dynamics*. Mc Graw Hill, New-York, 1983.
- [87] N. Rouquette. Quantitative diagnosis of continuous-valued, steady-state systems. In *Proceedings of the International Workshop on Principles of Diagnosis(DX'95)*, pages 91–98, Goslar(Germany), 1995.
- [88] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. Diagnosability of discrete event systems. Technical Report CGR-94-02, University of Michigan, 1994.
- [89] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis. A discrete event systems approach to failure diagnosis. In *Proceedings of the International Workshop on Principles of Diagnosis(DX'94)*, pages 269–277, New Paltz(NY), 1994.
- [90] Q. Shen and R. Leitch. Combining qualitative simulation and fuzzy sets. In B. Faltings and P. Struss, editors, *Recent advances in Qualitative Physics*, pages 83–100. MIT Press, 1992.
- [91] M.G. Singh, K.S. Hindi, G. Schmidt, and S.G. Tzafestas, editors. *Fault Detection and Reliability – Knowledge-Based and Other Approaches*, volume 9, International Series on Systems and Control. Pergamon Press, Oxford, UK, 1987.
- [92] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.Y. Glorennec, H. Hjalmärsson, and A. Juditsky. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, December 1995. Special Issue on Trends in System Identification.
- [93] Léa Sombé. Reasoning under uncertain information in artificial intelligence. *International Journal of Intelligent Systems*, 5:4, 1990. Special Issue.
- [94] P. Struss. What's in SD? towards a theory of modeling for diagnoses. In *Proceedings of the International Workshop on Principles of Diagnosis(DX'91)*, pages 41–51, Milan, 1991.
- [95] S. Tanaka. Diagnosability of systems and optimal sensor location. In R.J. Patton, P.M. Frank, and R.N. Clark, editors, *Fault Diagnosis in Dynamic Systems: Theory and Application*. Prentice Hall International, 1989.
- [96] R. Tang and G.K. Pang. Real-time intelligent process control using continuous fuzzy Petri nets. In *Proc. 2nd IFAC Workshop on Computer Software Structures Integrating AI/KBS Systems in Process-control*, pages 98–104, Lund, 1994.
- [97] S.G. Tzafestas, M.G. Singh, and G. Schmidt, editors. *Fault Diagnostics and Reliability*, volume 1 of *System Fault Diagnosis, Reliability and Related Knowledge-based Approaches*. Reidel, Dordrech, Holland, 1987.
- [98] S.G. Tzafestas, M.G. Singh, and G. Schmidt, editors. *Knowledge-based and Fault-tolerant Techniques*, volume 2 of *System Fault Diagnosis, Reliability and Related Knowledge-based Approaches*. Reider, Dordrecht, Holland, 1987.
- [99] R. Valette. Réseaux de Petri pour la détection et le diagnostic. In *Journées d'études S reté, Surveillance et Supervision*. GDR Automatique, 1994.
- [100] W.E. Vander Velde and C.R. Carignan. Number and placement of control system components considering possible failures. *Jal Guidance, Control, Dynamics*, 7(6):703–709, 1984.
- [101] G. Le Vey. *Analyse Modale et Surveillance Vibratoire des Machines Tournantes*. PhD thesis, Université de Rennes I, September 1988.
- [102] B. J. Wielinga and J. A. Breuker. Models of expertise. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, pages 307–318, 1986.
- [103] A.S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, 12:601–611, 1976.
- [104] X. Yu and G. Biswas. A multi-level diagnosis methodology for complex systems. In *Proc. of the Eighth Conference on Artificial Intelligence for Application CAIA-92*, pages 81–87, Monterey, CA, 1992.



- [105] Q. Zhang, M. Basseville, and A. Benveniste. Early warning of slight changes in systems and plants with application to condition based maintenance. *Automatica*, 30(1):95–114, January 1994. Special Issue on Statistical Methods in Signal Processing and Control.
- [106] Q. Zhang and A. Benveniste. Wavelet networks. *IEEE Trans. Neural Networks*, NN-3(6):889–898, November 1992.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Surveillance de systèmes dynamiques . . . . .	3
1.1.1	Systèmes dynamiques: Définitions . . . . .	3
1.1.2	Systèmes dynamiques: Pourquoi les surveiller? . . . . .	5
1.2	Architecture générale . . . . .	6
1.3	Les différents niveaux de traitement : du signal à la décision . . . . .	7
1.4	Objectifs d'une " bonne " surveillance . . . . .	8
<b>2</b>	<b>Génération des alarmes</b>	<b>9</b>
2.1	Le problème . . . . .	9
2.2	Deux types de modèles: le modèle boîte-noire et le modèle physique . . . . .	10
2.3	Les étapes de la génération d'alarmes . . . . .	10
2.4	Les méthodes de génération d'alarmes . . . . .	10
2.4.1	Transformation des données ou génération de résidus . . . . .	11
2.4.2	Décision ou génération d'alarme . . . . .	11
2.5	Génération d'alarmes intelligentes . . . . .	12
2.6	Techniques de l'IA pour la génération d'alarmes . . . . .	14
2.6.1	Graphes d'influence pour la simulation . . . . .	15
2.6.2	Graphe d'influence et découpage hiérarchique du système . . . . .	15
2.6.3	Utilisation conjointe d'un modèle quantitatif et qualitatif . . . . .	15
2.6.4	" Bond-graphs " ou Graphes de liaison . . . . .	15
2.7	Les critères de la génération d'alarmes . . . . .	16
2.7.1	Trois problèmes de détection . . . . .	16
2.7.2	Les critères . . . . .	17
<b>3</b>	<b>Interprétation des alarmes : du filtrage au diagnostic</b>	<b>17</b>
3.1	Différents niveaux d'interprétation . . . . .	18
3.1.1	Filtrage . . . . .	18
3.1.2	Localisation . . . . .	18
3.1.3	Diagnostic . . . . .	18
3.2	Différents types de modèles symboliques . . . . .	18
3.2.1	Trois grandes classes de modèles . . . . .	18
3.2.2	Acquisition des modèles . . . . .	19
3.2.3	Passage d'un modèle à un autre . . . . .	19
3.3	Modèles associatifs . . . . .	20
3.3.1	Systèmes experts . . . . .	20
3.3.2	Reconnaissance de scénarios . . . . .	20
3.3.3	Avantages/inconvénients . . . . .	21
3.4	Modèles prédictifs . . . . .	21
3.4.1	Modèles de la physique qualitative . . . . .	21
3.4.2	Modèles à événements discrets . . . . .	22
3.4.3	Autres formalismes . . . . .	22
3.5	Modèles explicatifs . . . . .	22
3.5.1	Graphes d'influence . . . . .	22
3.5.2	Graphes causaux (et temporels) . . . . .	23
<b>4</b>	<b>Conclusion</b>	<b>23</b>
4.1	Détection/Interprétation . . . . .	24
4.2	Avantages/inconvénients des approches numériques . . . . .	24
4.2.1	Acquisition des modèles . . . . .	24
4.2.2	Prise en compte de l'incertain . . . . .	25
4.2.3	Généricité des outils . . . . .	25
4.2.4	Validation des résultats . . . . .	26
4.3	Avantages/inconvénients des approches symboliques . . . . .	26
4.3.1	Acquisition des modèles . . . . .	26

4.3.2	Prise en compte de l'incertain . . . . .	30
4.3.3	Généricité des outils . . . . .	31
4.3.4	Validation des résultats . . . . .	31
4.4	Complémentarité des deux types d'approches . . . . .	31
<b>Références</b>		<b>33</b>



---

Unit e de recherche INRIA Lorraine, Technop le de Nancy-Brabois, Campus scientifique,  
615 rue du Jardin Botanique, BP 101, 54600 VILLERS L S NANCY  
Unit e de recherche INRIA Rennes, Irista, Campus universitaire de Beaulieu, 35042 RENNES Cedex  
Unit e de recherche INRIA Rh ne-Alpes, 46 avenue F elix Viallet, 38031 GRENoble Cedex 1  
Unit e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex  
Unit e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

---

 diteur  
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)  
ISSN 0249-6399