

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE**

**Université de Batna
Faculté des Sciences de l'Ingénieur
Département d'Electronique**

MEMOIRE

Présenté en vue de l'obtention du diplôme de

Magister en Electronique

Par

Saber DIR

Thème

**Utilisation de l'analyse en composantes principales pour
l'identification, détection et localisation de défauts pour les systèmes
statiques. Extension de l'ACP pour les systèmes dynamiques.**

Devant le jury

<u>Président :</u>	Mr. Djamel BENATIA	Prof.	Univ. Batna
<u>Rapporteur :</u>	Mr. Abderraouf MESSAI	M. C.	Univ. Constantine
<u>Examineurs :</u>	Mr. Abdelhafid CHAABI	Prof.	Univ. Constantine
	Mr. Tarek FORTAKI	M. C.	Univ. Batna
	Mr. Yacine ABDESSEMED	M. C.	Univ. Batna

Remerciements

Le travail présenté dans ce manuscrit a été réalisé au Laboratoire d'électronique avancé de l'Université de Batna

Je tiens tout d'abord à remercier Monsieur Noureddine BOUGUECHAL pour m'avoir accueilli au sein de son laboratoire.

J'exprime ma profonde reconnaissance à Messieurs Abderraouf MESSAI qui a su encadrer et diriger ce travail pour m'aider à le mener à bien.

Je remercie l'ensemble des membres du jury qui m'ont fait l'honneur d'accepter de lire et de juger ce mémoire : Messieurs Djamel BENATIA, Tarek FORTAKI et Yacine ABDESSEMED de l'université de Batna et Monsieur Abdelhafid CHAABI de l'université de Constantine.

Je ne saurais terminer sans remercier l'ensemble des collègues et amis qui m'ont aidé durant ces dernières années. Un grand merci à vous tous.

Dédicaces

A :

Mes chers parents

Mes frères et mes soeurs

Toute ma famille

Tous mes amis

Je dédie ce mémoire

Sommaire

<i>Chapitre 1 : Introduction générale</i>	7
<i>Chapitre 2 : Analyse en composantes principales</i>	11
1. Introduction	11
2. Principe de l'analyse en composantes principale	12
3. Détermination de la structure du modèle	16
<i>Chapitre 3: Détection de défauts par analyse en composantes principale</i>	18
1. Introduction	18
2. Principes fondamentaux du diagnostic	19
2.1 Diagnostic.....	19
2.2 Défaillance.....	19
2.3 Défaut	20
2.4 Redondance Analytique.....	20
2.4.1 Exemple.....	21
2.5 Détection et localisation	21
3. Détection de défauts par analyse en composantes principale	22
3.1 Expression des résidus en présence d'un défaut.....	22
3.2 Statistique <i>SPE</i>	23
<i>Chapitre 4 : ACP dynamique</i>	26
1. Introduction	26
2. Identification	28
2.1 Analyse parallèle	29
2.1.1 Exemple illustratif.....	30
2.2 Critère d' <i>AIC</i>	35
<i>Chapitre 5 : Résultats expérimentaux</i>	37
1. Introduction	37

2. Exemple d'application à la détection de défauts	37
2.1. Identification du modèle ACP	38
<i>Conclusion générale et perspectives</i>	48
<i>Annexe</i>	49
Méthodes à base d'observateurs ou de filtres.....	49
Remarques sur les reconstituteurs.....	54
Filtre de Kalman	56
<i>BIBLIOGRAPHIE</i>	59

Table des figures

<i>Figure 1.1 : Principe de génération des résidus</i>	10
<i>Figure 2.1 – Déroulement d’une analyse en composantes principales</i>	15
<i>(a) Distribution d’entrée.</i>	15
<i>(b) Centrage et réduction de cette distribution.</i>	15
<i>(c) Les deux axes principaux, correspondant aux vecteurs propres de la matrice de covariance de la distribution.</i>	15
 <i>Fig. 3.1 : Principe général du diagnostic</i>	19
 <i>Fig. 3.2 : Illustration de procédure de détection de saut de moyenne</i>	24
 <i>Fig. 4.1 : Intercorrélation entre les composantes principales pour un décalage l égal à 0.</i>	31
 <i>Fig. 4.2: Intercorrélation entre les composantes principales pour un décalage l égale à 1.</i>	33
 <i>Fig. 4.3: Intercorrélation entre les composantes principales pour un décalage l égal à 2.</i>	35
 <i>Fig. 5.1 : Mesures simulées de $z_1(k)$, $z_2(k)$, $z_3(k)$ et $z_4(k)$ en absence de défaut.</i>	41
 <i>Fig. 5.2 : Evolution des mesures et des estimations des deux variables $z_1(k)$ et $z_2(k)$ en présence d'un défaut affectant la variable $z_1(k)$ de l'instant 50 à l'instant 120.</i>	43
 <i>Fig. 5.3 : Evolution des 8 résidus $e_1(k)$, $e_2(k)$, $e_1(k)$, $e_3(k)$, $e_4(k)$, $e_1(k+1)$, $e_2(k+1)$, $e_3(k+1)$ et $e_4(k+1)$</i>	44
 <i>Fig. 5.4 : Evolution du SPE avec un défaut affectant la variable $z_1(k)$ de l'instant 50 à l'instant 120.</i>	45

Table des tableaux

<i>Tab. 5.1 : Evolution du taux de fausses alarmes et du taux de bonne détection en fonction de l'écart type de $\varepsilon_3(k)$ et $\varepsilon_4(k)$.</i>	47
<i>Tab. 5.2 : Evolution du taux de fausses alarmes et du taux de bonne détection en fonction de l'amplitude de défaut d affectant la variable $z_2(k)$.</i>	47

Notation

$\mathbf{X} \in \mathfrak{R}^{N \times m}$	Matrice de données représentant le fonctionnement normal du système,
$\hat{\mathbf{X}}$	Estimation de \mathbf{X} par le modèle ACP ,
\mathbf{E}	Matrice des résidus d'estimation de \mathbf{X} ,
$\Sigma \in \mathfrak{R}^{m \times m}$	Matrice de covariance de \mathbf{X} ,
N	Nombre d'échantillons mesurés,
m	Nombre de variables (dimension de l'espace des données mesurées),
l	Nombre de composantes retenues dans le modèle ACP (dimension du sous-espace des composantes principales),
k	Indice du temps,
$\mathbf{x} \in \mathfrak{R}^{N \times m}$	Nouveau vecteur de mesure,
$\hat{\mathbf{x}}$	Estimation du vecteur \mathbf{x} par le modèle ACP ,
x_i	La i ème composante du vecteur \mathbf{x} ,
$\mathbf{P} \in \mathfrak{R}^{m \times m}$	Matrice des vecteurs propres de \mathfrak{S} ,
$\hat{\mathbf{P}} \in \mathfrak{R}^{m \times l}$	Matrice des premiers vecteurs propres de Σ ,
$\tilde{\mathbf{P}} \in \mathfrak{R}^{m \times (m-l)}$	Matrice des $(m-l)$ ` derniers vecteurs propres de Σ ,

$\hat{\mathbf{t}}$	Vecteurs des l premières composantes principales,
$\tilde{\mathbf{t}}$	Vecteur des $(m-l)$ dernières composantes principales,
$\hat{\mathbf{C}} = \hat{\mathbf{P}}\hat{\mathbf{P}}^T$	Matrice représentant le modèle ACP ,
λ_i	i ème valeur propre de Σ ,
p_i	i ème vecteur propre de Σ correspondant à λ_i ,
E	Espérance mathématique,
p_i	Erreur d'estimation sur la i ème variable,
\mathbf{e}	Vecteur des erreurs d'estimation,
SPE	Erreur quadratique d'estimation (squared prediction error),

Chapitre 1 : Introduction générale

La réalisation des systèmes sûrs passe par la prise en compte de la moindre défaillance sur les processus et leurs aptitudes à résister aux éventuelles défaillances matérielles (capteurs, actionneurs, etc.), il est donc nécessaire de développer des nouvelles procédures de surveillance qui permettent la détection, la localisation et le diagnostic des éventuels défauts.

Ces dernières années, diverses méthodologies ont été développées pour répondre aux objectifs de détection et de localisation de défauts. On recense, dans la littérature scientifique différentes méthodes. Parmi celles-ci, on peut citer les méthodes s'appuyant sur l'estimation d'état comme les observateurs et le filtre de Kalman (voir annexe). D'autres méthodes basées sur le concept de redondance de l'information font appel à une connaissance du système constituée par la formulation explicite d'un modèle analytique du système à surveiller.

Le principe de base est fondé sur la prise d'informations, par le biais de capteurs, sur le processus à surveiller. La comparaison entre le comportement réel du processus et le comportement prédit par le modèle, fournit des informations contenues dans un ensemble de signaux indicateurs de défauts ou résidus qui caractérisent un écart par rapport aux conditions de fonctionnement. Pour générer ces résidus, on utilise des informations issues d'un modèle analytique du système, en suite les comparer à celles fournies par les instruments de mesure. L'analyse (temporelle ou fréquentielle) de ces signaux permet de détecter et éventuellement d'interpréter tout comportement anormal du système afin de localiser sa provenance.

Leur principe de base est fondé généralement sur un test de cohérence entre un comportement observé du processus fourni par le biais des capteurs et un comportement prévu fourni par un modèle ou une représentation mathématique du processus.

L'architecture générale du système de diagnostic à base de redondance analytique pour la détection et la localisation de défaillances de capteurs, sachant qu'on ne s'intéresse qu'à des défaillances d'instrumentation (capteurs ou actionneurs) est représentée sur la figure (fig 1.1).

Les méthodes de redondance analytique sont basées sur l'utilisation des redondances présentes dans les mesures des variables. Le prix à payer est toutefois l'élaboration d'un modèle mathématique aussi complet que possible dont la qualité est primordiale pour l'obtention d'un système de détection performant. Cela est tout à fait envisageable lorsqu'il s'agit des systèmes de petite dimension. En revanche, pour les systèmes de taille plus importante, établir de telles relations devient très difficile à obtenir.

L'obtention de toutes relations entre les variables quelque soit la taille quelque soit la dimension du système est précisément le but de l'analyse en composantes principales.

L'analyse en composantes principales (ACP) est une technique descriptive permettant d'étudier les relations qui existent entre les variables, sans tenir compte, a priori, d'une quelconque structure [26-27]. Le but de l'ACP est d'identifier la structure de dépendance entre des observations multivariées afin d'obtenir une description ou une représentation compacte de ces dernières.

Les méthodes de projection statistiques multivariées (MSPM) (Multivariate Statistical Process Monitoring) telles que l'analyse en composantes principales statique (ACP), ont été largement appliquées dans les processus industriels comme méthodes statistiques multivariées pour le diagnostic de défaut [8-3-16-2]. En général, les méthodes MSPM basées sur l'ACP statique sont formulées autour de l'hypothèse qui considère que les données analysées suivent une loi normale et qu'il n'y a pas de corrélation temporelle (autocorrélation) entre les différentes variables, tandis que les données utilisées dans la surveillance de nombreux processus industriels sont caractérisées non seulement par des intercorrélations mais également par des autocorrélations. Ainsi, si les processus ont un comportement dynamique, des problèmes associés à l'inefficacité des méthodes de surveillance basées sur l'ACP statique apparaissent, puisque l'hypothèse de base sur laquelle repose la surveillance par l'ACP statique n'est pas respectée.

Pour tenter de lever cette imperfection de l'ACP statique, plusieurs extensions du formalisme de base ont été développées pour la surveillance des processus dynamiques [1-4-9-10]. Ku et al [1] ont proposé une extension de l'ACP statique intégrant des retards entre les mesures.

Ce travail s'inscrit dans ce contexte de détection et de localisation de défauts de systèmes dynamique. Le but de ce travail est d'employer des méthodes pour inclure le comportement dynamique dans le modèle de l'ACP. Les outils statistiques de l'ACP statique sont étendus pour s'appliquer aux systèmes dynamiques pour prendre en compte les autocorrélations. Nous nous référerons à cette méthode dynamique en tant que **ACP** dynamique ou **ACPD**. Par la suite on présentera une approche d'identification qui a été proposé par Ku et al [1] où nous avons modifié l'algorithme utilisé, pour qu'on puisse chercher l'ordre du modèle et toutes les relations dynamique même si on n'a pas des relations statiques. Un exemple d'application sur à la détection de défauts sur un système stable sera propos

Ce mémoire est décomposé en cinq chapitres et organisé de la façon suivante :

Chapitre 1

Introduction générale

Chapitre 2

Le deuxième chapitre présentera le principe de l'analyse en composantes principales (ACP). La phase d'estimation des paramètres du modèle ACP y sera exposée.

Chapitre 3

Ce chapitre sera consacré à la détection de défauts par analyse en composantes principales linéaires.

Chapitre 4

L'extension de l'analyse en composantes principales dans le cas des systèmes dynamiques sera présentée. Une approche qui permet d'obtenir l'ordre du modèle et le nombre de composantes principales sera présentée comme une technique l'ACP dynamique.

Chapitre 5

Le dernier chapitre de ce travail sera consacré à l'application de l'analyse en composantes principales pour la détection et la localisation de défauts de capteurs d'un système stable.

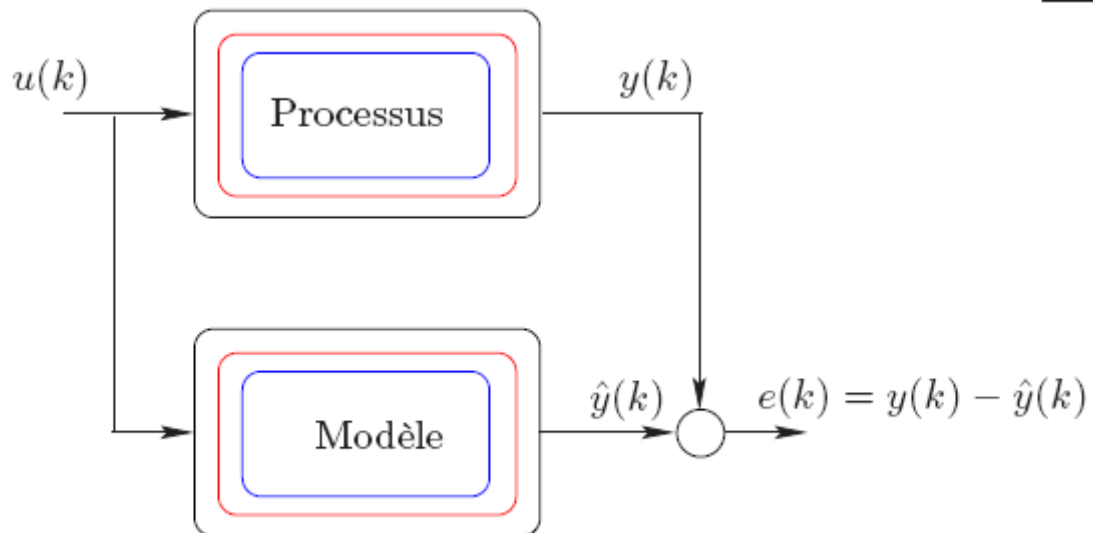


Figure 1.1 : Principe de génération des résidus

Chapitre 2 : Analyse en composantes principales

1. Introduction

Dans le domaine du diagnostic, des méthodes basées sur le concept de redondance de l'information ont été développées. Leur principe repose généralement sur un test de cohérence entre un comportement observé du processus fourni par des capteurs et un comportement prévu fourni par une représentation mathématique du processus. Les méthodes de redondance nécessitent donc un modèle du système à surveiller. Les méthodes de redondance analytique sont fondées sur l'utilisation des redondances présentes dans les mesures des variables. Le prix à payer est toutefois l'élaboration d'un modèle mathématique aussi complet que possible dont la qualité est primordiale pour l'obtention d'un système de détection performant.

Les méthodes statistiques classiques utilisées pour la détection de défauts dans des processus industriels sont connu généralement sous le nom de maîtrise ou contrôle statistique de processus (CSP) (traduction française de SPC, Statistical Process Control). Les premières techniques du CSP remontent aux travaux de Shewhart publiés en 1931 [28]. Parmi ces techniques les plus utilisées et les plus populaires sont des méthodes univariées, c'est-à-dire, surveillant chaque variable ainsi que ses statistiques (moyenne et variance) indépendamment des autres variables.

Cette approche exige un opérateur surveillant sans interruption les différents diagrammes univariés, ce qui réduit sensiblement sa capacité à faire des évaluations fiables au sujet de l'état du processus. En conséquence, l'application de cette approche à des systèmes de grande dimension (grand nombre de variable) devient difficile sinon impossible.

L'intérêt de l'analyse en composantes principales réside dans la réduction de dimension de représentation de données ou plus simplement la compression de données. En effet, cette

technique permet de caractériser les directions orthogonales d'un espace de données porteuses du maximum d'information au sens de la maximisation des variances de projections. L'amplitude des valeurs propres de la matrice de covariance des données quantifie pour chacune de ces directions la quantité d'information encodée.

En effet, l'ACP peut être considéré comme une technique de minimisation de l'erreur quadratique d'estimation ou une technique de maximisation de la variance des projections.

2. Principe de l'analyse en composantes principale

Considérons un vecteur de données aléatoires $\mathbf{x} = [x_1(k), \dots, x_m(k)]^T \in \mathfrak{R}^m$ de moyenne nulle $E\{\mathbf{x}(k)\} = 0$ (E représente l'espérance mathématique).

En analyse en composantes principales, un vecteur caractéristique $\mathbf{t} \in \mathfrak{R}^g$ est associé à chaque vecteur de données \mathbf{x} dont il optimise la représentation au sens de la minimisation de l'erreur d'estimation de \mathbf{x} [8]. Les vecteurs \mathbf{t} et \mathbf{x} sont liés par la transformation linéaire suivante :

$$\mathbf{t} = \mathbf{P}^T \mathbf{x} \tag{1}$$

Où la matrice de transformation \mathbf{P} est composée des vecteurs propres de la matrice de covariance des données $\mathbf{\Sigma} = E\{\mathbf{xx}^T\} \in \mathfrak{R}^{m \times m}$

Le vecteur \mathbf{t} est appelé vecteur de composantes principales.

En pratique, pour la modélisation d'un processus utilisant l'analyse en composantes principales, les mesures des variables du processus représentant tous les modes en fonctionnement normal de ce dernier sont collectés dans une matrice \mathbf{X} . Soit m le nombre de variables et N le nombre d'observations de chaque variable. $\mathbf{X} \in \mathfrak{R}^{N \times m}$ est donnée par :

$$\mathbf{X} = \begin{pmatrix} x_1(1) & x_2(1) & \cdots & x_m(1) \\ x_1(2) & x_2(2) & \cdots & x_m(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(N) & x_2(N) & \cdots & x_m(N) \end{pmatrix} \quad (2)$$

Où $x_1(1)$ représente la mesure de la première variable au premier instant de mesure.

Pour rendre le résultat indépendant des unités utilisées pour chaque variable, les données collectées dans \mathbf{X} , doivent subir un prétraitement qui consiste à centrer et réduire les variable [11-13-8].

La matrice de corrélation est données par :

$$\mathbf{\Sigma} = \mathbf{X}^T \mathbf{X} / N = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T \quad (3)$$

$\mathbf{\Lambda}$ est la matrice diagonale des valeurs propres de $\mathbf{\Sigma}$. On supposera dans cette analyse que les éléments diagonaux de $\mathbf{\Lambda}$ sont ordonnés du plus grand au plus petit. Dans l'ACP seuls les vecteurs propres liés aux g premiers vecteurs propres sont maintenus. Les matrices $\mathbf{\Lambda}$ et \mathbf{P} sont aussi partitionnées comme suit :

$$\mathbf{P} = \begin{bmatrix} \hat{\mathbf{P}} & \tilde{\mathbf{P}} \end{bmatrix}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_{\hat{\mathbf{P}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{\tilde{\mathbf{P}}} \end{bmatrix} \quad (4)$$

Les deux sous-matrices $\hat{\mathbf{P}}$ et $\mathbf{\Lambda}_{\hat{\mathbf{P}}}$ représentent les matrices des g premiers vecteurs propres et des g premières valeurs propres, respectivement. Les deux sous-matrices $\tilde{\mathbf{P}}$ et $\mathbf{\Lambda}_{\tilde{\mathbf{P}}}$ représentent les matrices des $(m-g)$ derniers vecteurs propres et des $(m-g)$ dernières valeurs propres, respectivement.

Rappelons que ces valeurs propres sont non négatives par définition. La direction de l'espace des données matérialisée par le vecteur propre associé à la plus grande valeur propre est la

plus riche d'information. Inversement, la direction du vecteur propre directeur associé à la plus petite valeur propre est celle qui présente le minimum d'information. Il est donc possible de réduire la dimension de la représentation des données en ne retenant que les vecteurs

Les m vecteurs propres unitaires p_i de la matrice de covariance Λ représentent les m directions orthogonales de l'espace des données suivant lesquelles les variances des projections t_i des données sont maximales. La figure 2.1 montre un déroulement d'une analyse en composantes principales.

Ainsi, la matrice des composantes principales $\mathbf{T} \in \mathfrak{R}^{N \times m}$ est donnée par :

$$\mathbf{T} = \mathbf{X}\mathbf{P} = \mathbf{X} \begin{bmatrix} \hat{\mathbf{P}} & \tilde{\mathbf{P}} \end{bmatrix} \quad (5)$$

A partir de cette équation (5), on peut écrire :

$$\hat{\mathbf{T}} = \mathbf{X}\hat{\mathbf{P}} \quad (6)$$

et

$$\tilde{\mathbf{T}} = \mathbf{X}\tilde{\mathbf{P}} \quad (7)$$

$\hat{\mathbf{T}}$ représente la projection de \mathbf{X} sur les g premiers vecteurs propres de la matrice de covariance Σ et $\tilde{\mathbf{T}}$ représente la projection de \mathbf{X} sur les $(m-g)$ derniers vecteurs propres.

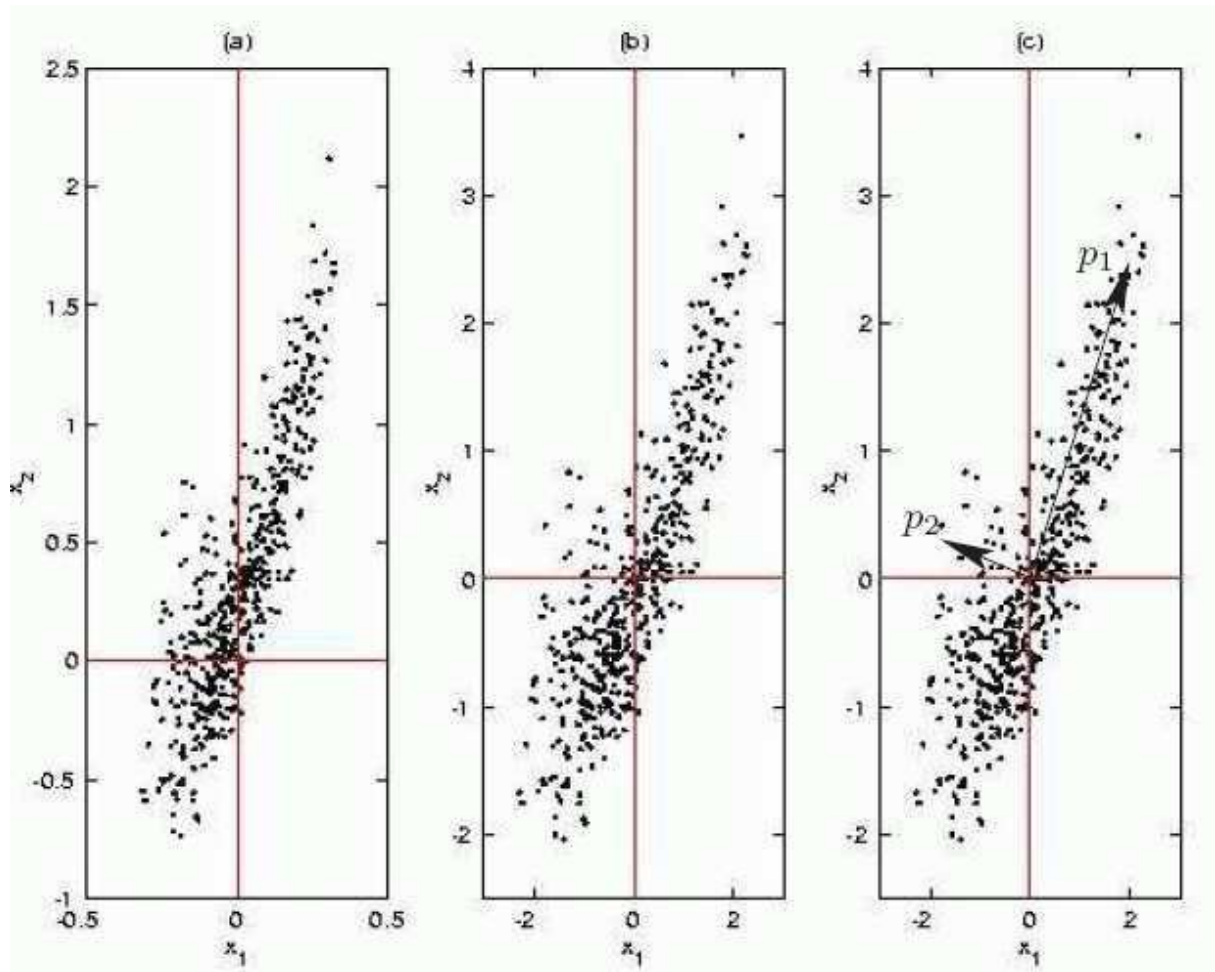


Figure 2.1 – Déroulement d'une analyse en composantes principales.

(a) Distribution d'entrée.

(b) Centrage et réduction de cette distribution.

(c) Les deux axes principaux, correspondant aux vecteurs propres de la matrice de covariance de la distribution.

A partir de l'équation (6) et (7) on peut écrire :

$$\hat{\mathbf{X}} = \hat{\mathbf{T}}\hat{\mathbf{P}}^T \quad (8)$$

et

$$\tilde{\mathbf{X}} = \tilde{\mathbf{T}}\tilde{\mathbf{P}}^T \quad (9)$$

Où $\hat{\mathbf{X}}$ représente l'estimation de \mathbf{X} et $\tilde{\mathbf{X}}$ représente la matrice des résidus notée \mathbf{E} .

On en déduit la décomposition suivante de la matrice \mathbf{X} :

$$\mathbf{X} = \hat{\mathbf{X}} + \tilde{\mathbf{X}} = \hat{\mathbf{X}} + \mathbf{E} \quad (10)$$

Ainsi que l'estimation :

$$\begin{aligned} \hat{\mathbf{X}} &= \hat{\mathbf{T}}\hat{\mathbf{P}}^T \\ \hat{\mathbf{X}} &= (\mathbf{X}\hat{\mathbf{P}})\hat{\mathbf{P}}^T \\ \hat{\mathbf{X}} &= \mathbf{X}\hat{\mathbf{C}} \end{aligned} \quad (11)$$

Et le résidu \mathbf{E} :

$$\begin{aligned} \mathbf{E} &= \mathbf{X} - \hat{\mathbf{X}} \\ &= \mathbf{X} - \mathbf{X}\hat{\mathbf{C}} \\ &= \mathbf{X}(\mathbf{I} - \hat{\mathbf{C}}) \\ &= \mathbf{X}\tilde{\mathbf{C}} \end{aligned} \quad (12)$$

Où $\hat{\mathbf{C}} = \hat{\mathbf{P}}\hat{\mathbf{P}}^T$ et $\tilde{\mathbf{C}} = (\mathbf{I} - \hat{\mathbf{C}})$

3. Détermination de la structure du modèle

L'analyse en composantes principales recherche une approximation de la matrice initiale des données \mathbf{X} par une matrice de rang inférieur issue d'une décomposition en valeurs singulières.

La question qui se pose alors, et qui a été largement débattue dans la littérature, concerne le choix du nombre de composantes principales qui doit être retenu. De nombreuses règles sont proposées dans la littérature pour déterminer le nombre de composantes à retenir [5-6-29-30].

La plupart de ces méthodes sont heuristiques et donnent un nombre de composantes subjectif. Toutefois, dans le cadre de l'application de l'ACP au diagnostic, le nombre de composantes a un impact significatif sur chaque étape de la procédure de détection et de localisation. Si peu

de composantes sont utilisées, on risque de perdre des informations contenues dans les données de départ en projetant certaines variables dans le sous-espace des résidus et donc avoir des erreurs de modélisation qui entachent les résidus ce qui provoque des fausses alarmes. Si par contre beaucoup de composantes sont utilisées, il y a le risque d'avoir des composantes retenues (les composantes correspondantes aux valeurs propres les plus faibles parmi celles retenues dans le modèle) qui sont porteuses de bruit ce qui est indésirable.

Dans la chapitre quatre nous allons présenter une approche pour la détermination du nombre de composantes à retenir et ainsi définir le modèle **ACP**.

La détermination du nombre de composantes du modèle **ACP** n'est en fait qu'une procédure de recherche de la structure du modèle. On retrouve dans la littérature plusieurs critères utilisant différentes approches concernant l'**ACP** statique.

Chapitre 3: Détection de défauts par analyse en composantes principale

1. Introduction

Pour un grand nombre d'applications, il est nécessaire d'implanter un système de surveillance afin de détecter, isoler, voire identifier tout dysfonctionnement.

Le diagnostic est une procédure qui consiste à détecter et localiser un composant ou un élément défectueux dans un système. On désigne par détecter la capacité à mettre en évidence l'apparition d'un ou plusieurs défauts. On parle de localisation quand on est, de plus, capable de préciser la nature du ou des défauts occurrence. La structure générale d'une procédure de diagnostic est représentée sur la figure 3.1 où le module de diagnostic est alimenté par toute la connaissance disponible (observation et modèle) sur le système. Cette connaissance inclut les mesures des variables et toute autre information pouvant être utile pour le diagnostic comme, par exemple, la structure du système. Le module de diagnostic traite les observations et produit un "diagnostic" qui est une liste de défauts possibles pouvant affecter le système au cours du temps.

Un système de surveillance doit permettre de caractériser le mode de fonctionnement d'un système à partir d'informations préalablement collectées, en reconnaissant et en indiquant les anomalies de comportement.

Dans le domaine du diagnostic, des méthodes basées sur le concept de redondance de l'information ont été développées. Leur principe repose généralement sur un test de cohérence entre un comportement observé du processus fourni par des capteurs et un comportement prévu fourni par une représentation mathématique du processus.

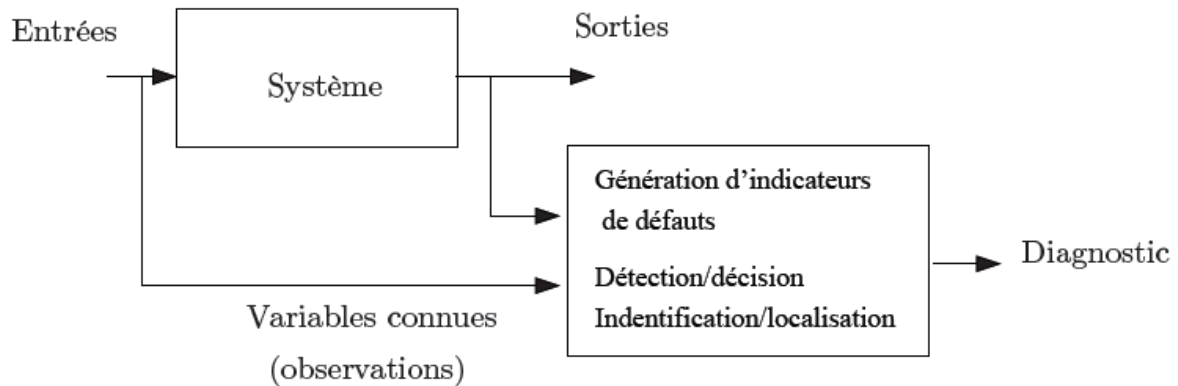


Fig. 3.1 : Principe général du diagnostic

2. Principes fondamentaux du diagnostic

2.1 Diagnostic

Le diagnostic consiste à détecter un fonctionnement anormal au sein du système et à déterminer sa cause en localisant le ou les composants du système présentant une anomalie de fonctionnement et, éventuellement, en caractérisant l'anomalie (sévérité, instant d'apparition, durée, etc.). Typiquement, le diagnostic débute par la comparaison entre le comportement ou le fonctionnement réel du système (dont une image est fournie par les observations) et le comportement ou le fonctionnement théorique attendu fourni par le modèle.

2.2 Défaillance

Une défaillance se rapporte à une anomalie fonctionnelle. L'adjectif défaillant est employé pour qualifier un système physique ou un composant dont une ou plusieurs fonctions sont altérées. Par abus, nous qualifions de défaillant un système physique ou un composant qui présente une anomalie fonctionnelle. En cas de cessation fonctionnelle, le terme panne est utilisé, c'est-à-dire l'inaptitude d'un composant à accomplir sa fonction.

2.3 Défaut

La définition que nous pouvons trouver dans la littérature, présente un défaut comme tout écart entre la caractéristique observée sur le dispositif et la caractéristique de référence [31]. La notion de défaut est donc voisine de celle de défaillance. Toutefois, un défaut n'implique pas nécessairement une défaillance car un défaut, lié au comportement, est plus général qu'une défaillance, qui est quant à elle liée aux fonctions qui peuvent être toujours remplies malgré la présence d'un défaut. De plus la notion de défaut inclut celle de défaillance ; un défaut n'altère pas nécessairement le fonctionnement d'un système physique mais peut présager d'une défaillance à venir.

Dans la pratique, un processus est composé de trois groupes d'organes : les actionneurs, les capteurs et le processus. Le comportement de l'ensemble dépend de chaque composant. Une défaillance peut donc surgir dans un ou plusieurs de ces organes.

Le diagnostic nécessite alors un ensemble de tests permettant de révéler des incohérences entre les observations et le modèle, signes de défauts ou symptômes de défaillances. L'objectif est ensuite de localiser le ou les composants ayant un comportement ou un fonctionnement anormal à partir des incohérences révélées par la procédure de diagnostic. Pour cela, un ensemble d'hypothèses sur l'origine possible des anomalies est généré. Chacune des hypothèses est formulée en termes de changements dans la structure, le comportement ou le fonctionnement du système. La localisation consiste à rechercher les hypothèses permettant d'expliquer l'ensemble des incohérences observées.

2.4 Redondance Analytique

Il existe une redondance analytique s'il existe une ou plusieurs relations ne faisant intervenir que des grandeurs mesurables, vraie(s) en l'absence de défaut. L'infirmité de ces relations met en évidence la présence d'un ou plusieurs défauts.

2.4.1 Exemple

Supposons que nous disposons de deux capteurs y_1 et y_2 qui mesurent la même variable x selon les lois physiques suivantes :

$$\begin{cases} y_1 = \sqrt{x} \\ y_2 = x \end{cases}$$

L'intégrité des deux capteurs peut être alors vérifiée en testant la validité de la relation

$$y_1^2 - y_2 = 0$$

Sur cet exemple, il est facile de voir qu'un dysfonctionnement de l'un des deux capteurs infirmerait la relation et un défaut pourrait être détecté. Dans des cas plus généraux, et pour faciliter non seulement la détection de défauts mais également leurs localisations, il est nécessaire d'avoir plus de relations de redondance analytique (au moins autant que de défauts à localiser) et de vérifier certaines propriétés structurelles.

Par ailleurs, on peut aussi disposer d'informations supplémentaires concernant les défauts : fréquence, incidence sur le système, processus générateurs...

2.5 Détection et localisation

La détection consiste à prendre une décision binaire : soit le système fonctionne correctement, soit une panne s'est produite. Le résultat de la procédure de détection est une alarme signifiant que le fonctionnement réel du système ne concorde plus avec le modèle de fonctionnement sain. Isoler revient à attribuer le défaut au module défectueux du système : capteur, actionneur, processus ou unité de commande. Diagnostiquer consiste à effectuer la classification des défauts selon certains paramètres qui les caractérisent : instant d'apparition, amplitude. Cette étape consiste également à prévoir l'évolution des défauts et à quantifier leur degré de sévérité.

3. Détection de défauts par analyse en composantes principale

La phase de détection, dans le cas du diagnostic à base de modèle analytique, est liée à l'étape de génération de résidus qui a pour but de générer à partir d'un modèle de bon fonctionnement du processus et des mesures disponibles, des signaux révélateurs de la présence de défauts, appelés résidus. A partir de l'analyse de ces résidus, l'étape de prise de décision doit alors indiquer si un défaut est présent ou non.

Dans le cas de l'analyse en composantes principales de trouve ce même principe de l'approche par estimation d'état, les résidus étant obtenus par un test de cohérence entre l'estimation et les mesures.

3.1 Expression des résidus en présence d'un défaut

Avant d'aborder les différents indices de détection, nous expliciterons les expressions des résidus en présence d'un défaut d quelconque sur la $j^{\text{ème}}$ variable du processus à surveiller. Notons $x^*(k)$ le vecteur de mesure à l'instant k en absence de défaut et ξ_j la direction du défaut.

Ainsi on peut écrire que :

$$x(k) = x^*(k) + \xi_j d(k) \quad (13)$$

On a :

$$\begin{aligned} \mathbf{e}(k) &= \tilde{\mathbf{C}}\mathbf{x}(k) \\ &= \begin{bmatrix} \tilde{c}_{11} & \tilde{c}_{11} & \cdots & \tilde{c}_{1m} \\ \tilde{c}_{21} & \tilde{c}_{22} & \cdots & \tilde{c}_{21} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{c}_{m1} & \cdots & \cdots & \tilde{c}_{mm} \end{bmatrix} (x^*(k) + \xi_j d(k)) \\ &= \mathbf{e}^*(k) + \begin{bmatrix} \tilde{c}_{1j} \\ \tilde{c}_{2j} \\ \vdots \\ \tilde{c}_{mj} \end{bmatrix} \end{aligned} \quad (14)$$

Où $\mathbf{e}^*(k)$ représente le vecteur des résidus en absence de défaut, ces résidus étant statistiquement nuls. Ainsi le défaut d se propage dans tous les résidus, et plus précisément un défaut d_j affectant la variable x_j se propage dans le résidu e_j avec une amplitude $\tilde{c}_{ij}d_j$.

3.2 Statistique *SPE*

La procédure de détection vise à déterminer l'apparition et l'instant d'occurrence d'un défaut. Un défaut est détectable si au moins un résidu permet de le détecter [32]. Les bruits affectant les systèmes à surveiller sont souvent modélisés par des variables aléatoires suivant une forme normale $\mathcal{N}(\mu, \Omega)$ où μ est la valeur moyenne et Ω sa matrice de variance covariance. Les résidus étant générés par les défauts et les bruits, l'étude de leurs caractéristiques statistiques peut renseigner sur l'occurrence d'un défaut. Par exemple, dans le cas d'un défaut constant et de bruits suivant une loi normale affectant un système linéaire, l'apparition d'un défaut provoquera un saut de moyenne du résidu.

Le résidu $r(t)$ sera caractérisé par :

$$r(t) \sim \mathcal{N}(\mu_0, \Omega) \text{ dans le cas sans défaut}$$

$$r(t) \sim \mathcal{N}(\mu_1, \Omega) \text{ en présence de défaut}$$

La figure 3.2 présente l'interprétation graphique du principe d'une procédure de décision pour la détection d'un saut de moyenne.

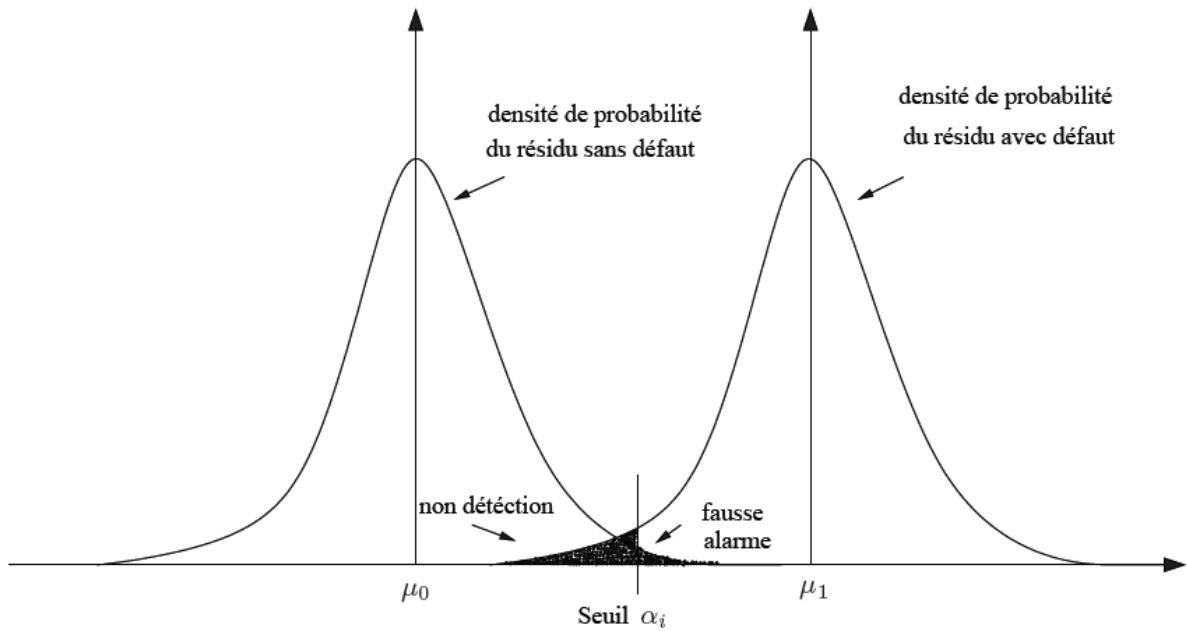


Fig. 3.2 : Illustration de procédure de détection de saut de moyenne

Une des statistiques typique pour la détection des anomalies est la statistique SPE (Squared Prediction Error) [14-16-18], qui s'exprime par :

$$SPE(k) = \mathbf{e}(k)^T \mathbf{e}(k) \quad (15)$$

Le processus est considéré en fonctionnement anormal (présence d'un défaut) à l'instant k si :

$$SPE(k) > \delta_\alpha^2 \quad (16)$$

Où δ^2 est le seuil de confiance du $SPE(k)$, et α est l'intervalle de confiance.

Soit $\sum_{j=g+1}^m \sigma_j^i$ pour $i=1,2,3$ et σ_j est la $j^{\text{ème}}$ valeur propre de la matrice de covariance Σ .

Une approximation du seuil δ^2 a été effectuée par Box [19]. Cette approximation est la suivante :

$$\delta_\alpha^2 = g_\theta \chi_{h,\alpha}^2 \quad (17)$$

Où $g_\theta = \theta_2 / \theta_1$ et $h = \theta_1^2 / \theta_1$

Il faut noter que ce résultat est donné sous la condition que le vecteur des échantillons \mathbf{x} suit une distribution normale.

Chapitre 4 : ACP dynamique

1. Introduction

L'analyse en composantes principales cherche à identifier les vecteurs propres orthonormaux et les valeurs propres de la matrice de covariance des données. On peut montrer que les derniers vecteurs propres représentent en fait les relations linéaires ou quasi linéaires qui existent entre les variables. Identifier les relations linéaires entre les variables revient à trouver l'espace nul de la matrice des données $\mathbf{X} \in \mathfrak{R}^{N \times m}$ (N est le nombre d'observations et m est le nombre de variables), qui est la solution non triviale de l'équation suivante

$$\mathbf{X}\mathbf{b} = 0 \quad (18)$$

Où $\mathbf{b} \in \mathfrak{R}^{m \times 1}$. La condition exigée pour que cette solution existe est que le rang de la matrice \mathbf{X} soit inférieur à m . Si le rang de \mathbf{X} est égal à $(m-r)$, alors il existe r relations linéaires. La solution de l'équation est constituée des vecteurs propres liés aux valeurs propres nulles de la matrice de covariance Σ (ou les vecteurs singuliers liés aux valeurs singulières nulles de la matrice de \mathbf{X}) :

$$\mathbf{b} = \mathbf{p}_k, \quad k = m - r + 1, \dots, m \quad \text{si } \sigma_k = 0 \quad (19)$$

Où \mathbf{p}_k est le $k^{\text{ième}}$ vecteur propre de la matrice de covariance Σ et σ_k est la $k^{\text{ième}}$ valeur propre de Σ .

Définition 1

Si \mathbf{X} est une matrice $N \times m$ ($N > m$), alors il existe deux matrices \mathbf{P} et $\mathbf{\Lambda}$:

$$\begin{aligned} \mathbf{P} &= [\mathbf{p}_1, \dots, \mathbf{p}_m] \in \mathfrak{R}^{m \times m} & (\mathbf{p}_i \in \mathfrak{R}^{m \times 1}) \\ \mathbf{\Lambda} &= \text{diag}(\sigma_1, \dots, \sigma_m) \in \mathfrak{R}^{m \times m} \end{aligned} \quad (20)$$

Telles que :

$$\mathbf{\Sigma} = \mathbf{X}^T \mathbf{X} / N = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T \quad (21)$$

$\mathbf{\Lambda}$ est la matrice diagonale des valeurs propres de $\mathbf{\Lambda}$ et \mathbf{P} est la matrice orthogonale des vecteurs propres de $\mathbf{\Sigma}$. Les valeurs propres sont ordonnées comme suit : $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$.

Cependant, les valeurs propres nulles de la matrice $\mathbf{\Sigma}$ sont rarement obtenues dans la pratique en raison de la présence du bruit. En présence du bruit les valeurs propres nulles deviennent petites. La solution de l'équation (18) sera constituée des vecteurs propres liés aux valeurs propres les plus petites :

$$\mathbf{b} = \mathbf{p}_k, \quad k = r+1, \dots, m \quad \text{si } \sigma_k = \sigma(E_\sigma) \quad (22)$$

Où E_σ est un certain seuil pour de petites valeurs propres. Nous appellerons l'espace qui contient les vecteurs \mathbf{p}_k ($k = m-r+1, \dots, m$) le sous-espace des résidus ; cet espace contient les relations linéaires entre les variables.

Cependant, quand le bruit existe, il est difficile de déterminer le nombre exact de composantes principales ou, d'une manière équivalente, le nombre de relations linéaires entre les variables. Si toutes les relations sont statiques, plusieurs méthodes sont proposées dans la littérature pour déterminer le nombre de composantes à retenir [6-5]. La difficulté principale surgit quand les relations sont dynamiques. Quand les données contiennent l'information dynamique, l'application de l'ACP sur les données n'indiquera pas les relations exactes entre les variables, mais plutôt une approximation statique linéaire. Pour un système dynamique, les valeurs courantes des variables dépendront des valeurs passées ; pour cela nous devons identifier les relations linéaires dynamiques dans $\mathbf{X}(k)$, représentant les données à l'instant (k), et $\mathbf{X}(k+1)$, représentant les données à l'instant ($k+1$) au moins, si toutes les relations dynamiques sont du premier ordre. L'équation (18) devient :

$$[\mathbf{X}(k) \ \mathbf{X}(k-1)]\mathbf{b} = 0 \quad (23)$$

Ou d'une manière générale,

$$\mathbf{X}_d(l)\mathbf{b} = 0 \quad (24)$$

Avec :

$$\mathbf{X}_d(l) = [\mathbf{X}(k) \ \mathbf{X}(k-1) \ \dots \ \mathbf{X}(k-l)] \quad (25)$$

Où $\mathbf{X}_d(l) \in \mathfrak{R}^{m \times N}$ et l désigne le nombre de décalages de la matrice \mathbf{X} dans le temps.

Notons que la matrice $\mathbf{X}_d(l)$ a la forme d'une matrice de Hankel. C'est essentiellement identique à l'approche originale de l'ACP sauf que la matrice de données se compose de vecteurs décalés dans le temps. En choisissant correctement le nombre de décalage l , les relations statiques et dynamiques devraient apparaître dans le sous-espace résiduel avec des valeurs propres de faible amplitude.

Cette méthode, utilisant le décalage de données dans le temps, n'est pas nouvelle et est largement répandue dans le domaine de l'identification [15-12].

2. Identification

Généralement, la procédure d'identification de modèles, consiste, après le choix d'une classe de modèle, à choisir une structure fixe puis à estimer les paramètres du modèle et enfin à valider ce modèle. Dans le cas de l'ACP, l'estimation des paramètres du modèle est très simple et revient en fait à un calcul de valeurs et vecteurs propres de la matrice Σ ou valeurs et vecteurs propres de la matrice \mathbf{X} . Cependant le choix de la structure est plus délicat comme nous verrons par la suite, en présentant deux approches qui nous permettront de déterminer la structure du modèle.

2.1 Analyse parallèle

Cette approche a été proposée par Ku et al [1]. Pour extraire les relations statiques et dynamiques à partir des données par l'utilisation de l'ACP dynamique, on doit déterminer le nombre de décalages l et les nombre de composantes principales à retenir (ou de relations linéaires existant entre les variables). Le nombre l indique aussi l'ordre du système dynamique.

Cette approche utilise l'algorithme suivant pour déterminer le nombre approprié de composantes à retenir et l'ordre du système.

1. Prendre l égale 0, calculer toutes les relations statiques;
2. Mettre $r_n(0)$ égale au nombre de relations statique. r_n représente les nouvelles relations linéaires trouvées à chaque niveau de décalage l ;
3. Prendre l égale 1;
4. Former la matrice de données : $\mathbf{X}_d(l) = [\mathbf{X}(k) \ \mathbf{X}(k-1) \ \dots \ \mathbf{X}(k-l)]$;
5. Calculer toutes les composantes principales;
6. Prendre j égale à $m \times l$ et $r(l)$ à 0. $r(l)$ représente le nombre total de relations linéaires à chaque niveau de décalage l ;
7. Déterminer si, la $j^{\text{ème}}$ composante représente une relation linéaire. Si oui, continuer si non, passer à l'étape 7;
8. Prendre j égale à $j-1$ et $r(l) = r(l) + 1$, retourner à l'étape 5;
9. Calculer le nouveau nombre de relations linéaires, selon l'équation suivantes :

$$r_n(l) = r(l) - \sum_{i=0}^{l-1} (l-i+1) r_n(i) \quad (26)$$

10. Si $r_{new}(l) \leq 0$, aller à l'étape 10, si non continuer;
11. Prendre l égale à $l+1$, retourner à l'étape 2;
12. Fin

Cet algorithme suppose qu'il existe déjà une méthode pour décider si une composante représente une relation linéaire ou non. Dans cette approche on utilise à chaque niveau l l'intercorrélation entre les composantes comme critère de décision. En présence du bruit, une composante principale représente une relation linéaire si elle n'est corrélée avec aucune composante et qu'elle est corrélée avec elle-même qu'à un instant donné c'est à dire que cette composante représente un bruit blanc.

Le nombre de relations linéaires r , sera le nombre total de variables (à chaque niveau l) moins le nombre de composantes principales retenues.

$r_n(l)$ pourrait avoir une valeur négative pour une certaine valeur de l .

2.1.1 Exemple illustratif

Considérons le système dynamique suivant :

$$y(k+1) = 0.7y(k) + 4u(k)$$

Avec $y(k)$ la sortie du système, et $u(k)$ est l'entrée du système. $u(k)$ est un signal en forme de créneaux, dont l'amplitude et la durée changent de manière aléatoire. On dispose de deux variables qui sont décrites par les équations suivantes :

$$z_1(k) = y(k) + e_1(k)$$

$$z_2(k) = u(k) + e_2(k)$$

Avec :

Les bruits affectant sont modélisés par des variables aléatoires suivant une forme normale

$$\varepsilon_1 \sim \mathcal{N}(0, 0.04) \quad \varepsilon_2 \sim \mathcal{N}(0, 0.04)$$

1. Calculer les relations statiques ($l = 0$)

Le vecteur de données $x(k)$ est :

$$\mathbf{x}(k) = [z_1(k) \ z_2(k)]^T$$

Avec $z_1(k) \in \mathfrak{R}^{N \times 1}$ et $z_2(k) \in \mathfrak{R}^{N \times 1}$.

La décomposition en valeurs propres de la matrice Σ donne :

$$\mathbf{\Lambda} = \begin{pmatrix} 1.88 & 0 \\ 0 & 0.11 \end{pmatrix} \quad \mathbf{P} = \begin{pmatrix} 0.70 & 0.70 \\ 0.70 & -0.70 \end{pmatrix}$$

Ce résultat n'indique aucune relation entre les variables, puisqu'il n'y a aucune valeur propre qui est égale à (ou proche de) zéro. La figure (4.1) montre que les deux composantes t_1 et t_2 ne sont pas totalement indépendantes, donc elles ne peuvent pas être des bruits blancs. Les deux composantes principales sont significatives, d'où le nombre à retenir égal à 2.

On a : $r(0) = 0$ et $r_n(0) = 0$.

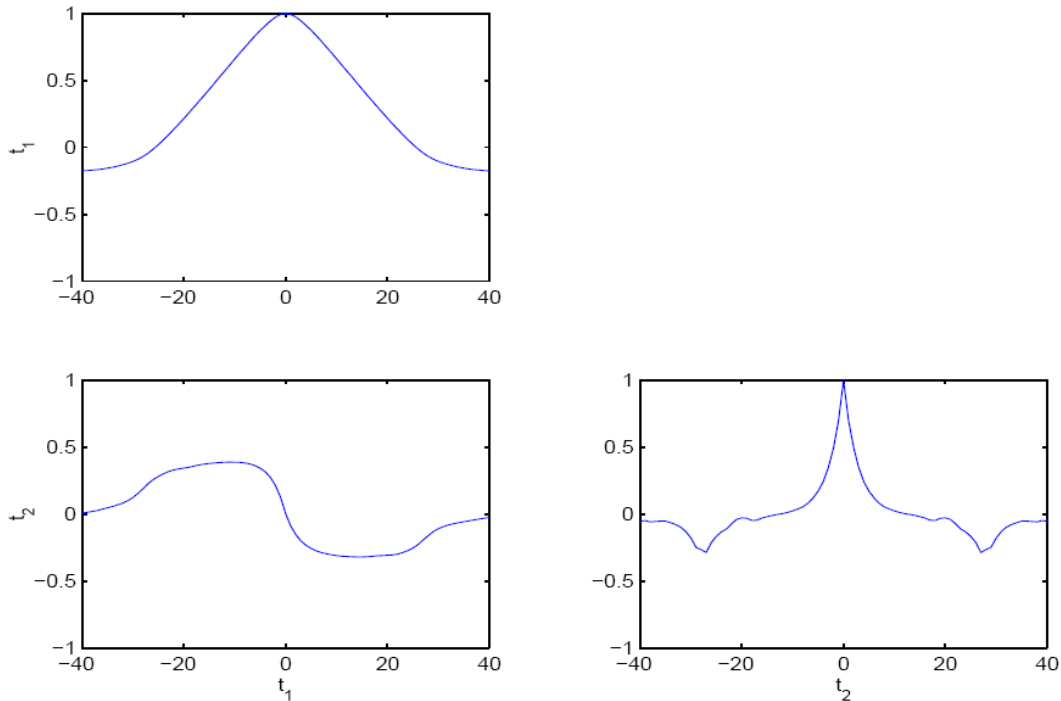


Fig. 4.1 : Intercorrélation entre les composantes principales pour un décalage l égal à 0.

2. Pour un décalage l égal à 1 :

Pour une valeur de décalage l de 1, la matrice de données $\mathbf{X}(k)$ devient :

$$\mathbf{X}_d(1) = [\mathbf{X}(k) \ \mathbf{X}(k-1)]$$

La décomposition en valeurs propres donne les matrices des valeurs propres et des vecteurs propres de la matrice de covariance Σ suivantes :

$$\mathbf{\Lambda} = \begin{pmatrix} 3.74 & 0 & 0 & 0 \\ 0 & 0.20 & 0 & 0 \\ 0 & 0 & 0.07 & 0 \\ 0 & 0 & 0 & 0.00 \end{pmatrix}$$

$$\mathbf{P} = \begin{pmatrix} 0.51 & 0.33 & 0.02 & 0.97 \\ 0.48 & 0.65 & 0.57 & 0.01 \\ 0.49 & 0.60 & 0.25 & 0.56 \\ 0.50 & 0.29 & 0.77 & 0.22 \end{pmatrix}$$

La valeur propre nulle prouve qu'il y a une relation dynamique linéaire entre les quatre variables.

La figure (4.2) montre que la dernière composante t_4 est un bruit blanc car elle est indépendante de toutes les autres composantes (elle n'est corrélée avec aucune), tandis que les trois autres sont corrélées entre elles. Les trois composantes principales sont significatives, d'où le nombre à retenir égal à 3. Le nombre de relations linéaires est donné par :

$$r(1) = m - g = 4 - 3 = 1$$

Avec m et g représentent respectivement le nombre de variables et le nombre de composantes principales à retenir à chaque niveau de décalage l .

Ainsi :

$$r_n(1) = r(1) - \sum_{i=0}^0 (1-i+1) r_n(i) = 1$$

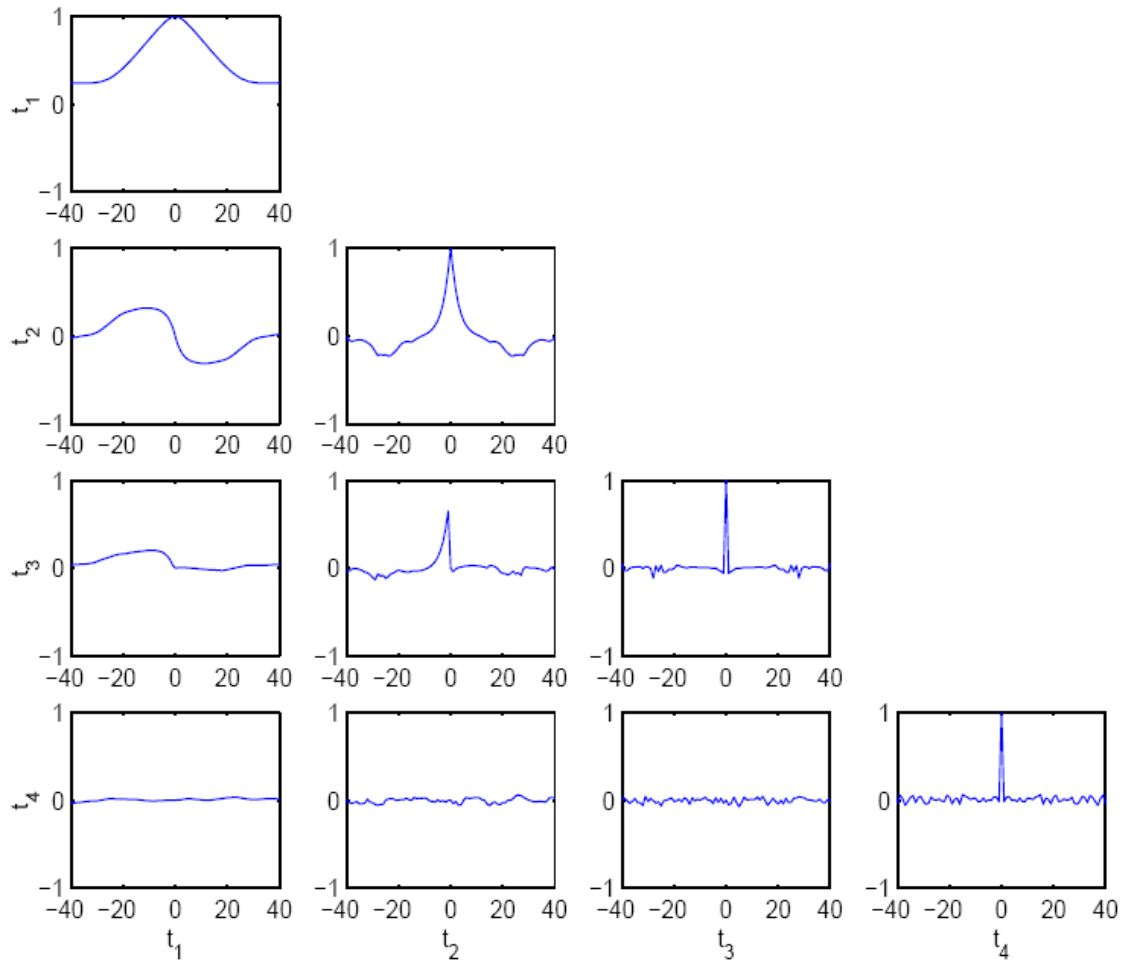


Fig. 4.2: Intercorrélation entre les composantes principales pour un décalage l égale à 1.

r_n n'est pas inférieur ou égal à zéro, donc on décale la matrice de données \mathbf{X} encore une fois.

3. Pour un décalage l égal à 2 :

La matrice de données devient

$$\mathbf{X}_d(2) = [\mathbf{X}(k) \quad \mathbf{X}(k-1) \quad \mathbf{X}(k-2)]$$

Les matrices des valeurs propres et vecteurs propres de la matrice de covariance Σ sont données par :

$$\mathbf{\Lambda} = \begin{pmatrix} 0:41 & 0:13 & 0:02 & 0:11 & 0:74 & 0:48 \\ 0:39 & 0:62 & 0:57 & 0:35 & 0:02 & 0:01 \\ 0:41 & 0:36 & 0:06 & 0:05 & 0:23 & 0:79 \\ 0:40 & 0:39 & 0:25 & 0:74 & 0:21 & 0:10 \\ 0:40 & 0:39 & 0:35 & 0:08 & 0:54 & 0:32 \\ 0:41 & 0:39 & 0:68 & 0:54 & 0:22 & 0:09 \end{pmatrix}$$

$$\mathbf{P} = \begin{pmatrix} 5.46 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.24 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.06 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.04 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.00 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.00 \end{pmatrix}$$

On a deux valeurs propres nulles. Ceci implique qu'il y a deux relations linéaires parmi les six variables. Les deux relations sont la relation dynamique du premier ordre trouvé au niveau $l=1$ répétée deux fois. Cela indique qu'aucune nouvelle information n'est trouvée dans ce niveau d'où $r_n(2)$ est nul. Par conséquent, on devait s'arrêter au niveau $l=1$ et la matrice $[\mathbf{X}(k) \mathbf{X}(k-1)]$ sera suffisante pour extraire les relations qui peuvent exister entre les variables, et le modèle ACPD, d'autre part.

La figure (4.3) montre que les deux dernières composantes t_5 et t_6 sont des bruits blancs car elles sont indépendantes de toutes les autres composantes et chacune d'elles est corrélée avec elle-même qu'à un instant k donné.

Récemment Qin et al [22] ont prouvé que cette forme de l'ACP dynamique est non biaisée seulement si toutes les variables sont entachées par une variance de bruit identique. Si les variables ne sont pas entachées par une variance de bruit identique, le modèle ACP est biaisé.

Qin et al [22-24] avaient proposé une approche d'identification basé sur la méthode des sous-espaces qui trouve un modèle **ACP** non biaisé en utilisant des variables instrumentales.

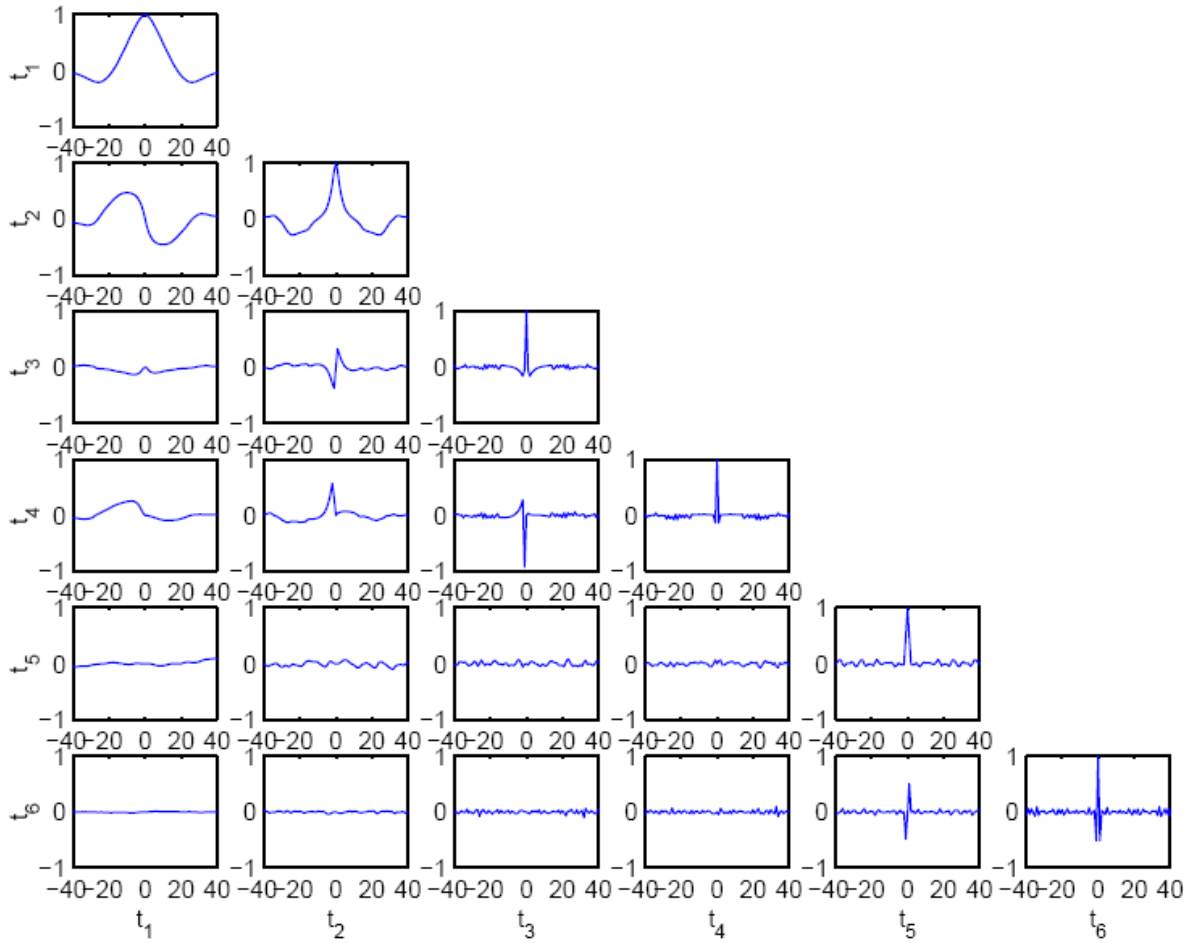


Fig. 4.3: Intercorrélation entre les composantes principales pour un décalage l égal à 2.

2.2 Critère d'AIC

Ici les idées sont brièvement traitées, pour plus de détails voir Larimore [20-21] et Larimore et Mehra [25]. On peut montrer que pour un grand nombre N d'observations, le critère d'AIC est un estimateur optimal et réalise des décisions optimales sur l'ordre du modèle. L'AIC original a été dérivé d'une évaluation de maximum de vraisemblance [23].

Etant donné le nombre N d'observation, et la séquence d'ordre du système n , par exemple $n \in [0 \ 20]$, l'ordre du modèle sera celui qui rendra le critère d'AIC suivant :

$$AIC_n(n) = N(m(s + \log 2\pi) + \frac{1}{N} \log(\det \sum_{k=1}^N e(k)e(k)^T)) + 2\delta_n M_n \quad (27)$$

minimum.

Où m est le nombre de sorties, s est le nombre d'entrées et N est le nombre d'observations.

Avec :

$$\mathbf{e}(k) = (\mathbf{y}(k) - \hat{\mathbf{y}}(k)) \quad (28)$$

Et :

$$M_n = 2n \times m + \frac{m(m+s)}{2} + n \times s + m \times s \quad (29)$$

est le nombre de paramètres dans le modèle de l'espace d'état, et

$$\delta_n = \frac{N}{N - \left(\frac{M_n}{m} + \frac{m+1}{2} \right)} \quad (30)$$

est un facteur pour expliquer le nombre limité d'échantillons.

La clef au calcul d' AIC est le calcul de l'erreur de la prédiction $\mathbf{y} - \hat{\mathbf{y}}$. Qin et al [24] ont utilisé deux algorithmes pour calculer l'erreur de prédiction, le premier est la méthode de polynôme caractéristique (PC) et la deuxième est de l'erreur de reconstruction (ER).

Chapitre 5 : Résultats expérimentaux

1. Introduction

Dans ce chapitre nous allons présenter les expériences que nous avons effectuées. Nous allons présenter l'application de l'ACP à la détection de défauts de capteurs sur un système stable

2. Exemple d'application à la détection de défauts

Considérons le système dynamique suivant :

$$\begin{cases} y_1(k+1) = 0.7y_1(k) + 0.2y_2(k) + 4u_1(k) \\ y_2(k+1) = 0.5y_2(k) + 2u_1(k) + 3u_2(k) \\ y_1(0) = 0 \\ y_2(0) = 0 \end{cases}$$

$y_1(k)$ et $y_2(k)$ sont les sorties du système, $u_1(k)$ et $u_2(k)$ sont les entrées du système.

$u_1(k)$ et $u_2(k)$ sont des signaux en forme de créneaux, dont les amplitudes et les durées changent de manière aléatoire.

Les mesures simulées des variables $y_1(k)$, $y_2(k)$, $u_1(k)$ et $u_2(k)$ sont données, à titre indicatif, sur la figure (5.1).

On dispose de quatre variables décrites par les équations suivantes :

$$\begin{cases} z_1(k) = y_1(k) + \varepsilon_1(k) \\ z_2(k) = y_2(k) + \varepsilon_2(k) \\ z_3(k) = u_1(k) + \varepsilon_3(k) \\ z_4(k) = u_2(k) + \varepsilon_4(k) \end{cases}$$

où les bruits de mesure $\varepsilon_i(k)$ sont des bruits blancs de moyenne nulle et d'écart type égal à 7 pour $\varepsilon_1(k)$ et $\varepsilon_2(k)$ et de 0.5 pour $\varepsilon_3(k)$ et $\varepsilon_4(k)$

Le vecteur de données centrées réduites est donné par :

$$\mathbf{x}(k) = [z_1(k) \ z_2(k) \ z_3(k) \ z_4(k)]^T$$

Ainsi pour N observations la matrice de données centrées réduites est la suivante :

$$\mathbf{X}(k) = [\mathbf{z}_1(k) \ \mathbf{z}_2(k) \ \mathbf{z}_3(k) \ \mathbf{z}_4(k)]^T$$

2.1. Identification du modèle ACP

L'identification du modèle revient à calculer l'ordre du modèle et le nombre de composantes principales à retenir, pour cela on utilise l'approche d'analyse parallèle. On suit le même raisonnement de l'exemple 1 pour identifier le modèle, on trouve :

Le nombre de décalage l est égal à 1, donc le modèle ACP est du premier ordre.

Le nombre total de variable m est égale à 8 et le nombre de composantes principales à retenir g est égal à 6.

Le vecteur $\mathbf{x}_d(1)$ de données est donné par :

$$\mathbf{x}_d(1) = [z_1(k) \ z_2(k) \ z_3(k) \ z_4(k) \ z_1(k-1) \ z_2(k-1) \ z_3(k-1) \ z_4(k-1)]^T$$

La matrice de données $\mathbf{X}_d(1)$ est la suivante :

$$\mathbf{X}_d(1) = [\mathbf{z}_1(k) \ \mathbf{z}_2(k) \ \mathbf{z}_3(k) \ \mathbf{z}_4(k) \ \mathbf{z}_1(k-1) \ \mathbf{z}_2(k-1) \ \mathbf{z}_3(k-1) \ \mathbf{z}_4(k-1)]^T$$

Un modèle **ACP** à été identifié. Alors on peut calculer l'estimation de $\mathbf{x}_d(1)$ comme suit :

$$\hat{\mathbf{x}}_d(1) = \hat{\mathbf{C}}\mathbf{x}_d(1)$$

Avec

$$\hat{\mathbf{x}}_d(1) = [\hat{z}_1(k) \ \hat{z}_2(k) \ \hat{z}_3(k) \ \hat{z}_4(k) \ \hat{z}_1(k-1) \ \hat{z}_2(k-1) \ \hat{z}_3(k-1) \ \hat{z}_4(k-1)]^T$$

La matrice $\hat{\mathbf{C}}$ égale à :

$$\hat{\mathbf{C}} = \hat{\mathbf{P}}\hat{\mathbf{P}}^T = \begin{pmatrix} 0.37 & 0.03 & 0.04 & 0.00 & 0.47 & 0.05 & 0.07 & -0.01 \\ 0.03 & 0.32 & 0.03 & 0.01 & -0.02 & 0.37 & 0.07 & 0.26 \\ 0.04 & 0.03 & 0.99 & 0.00 & -0.03 & -0.02 & -0.01 & -0.01 \\ 0.00 & 0.01 & 0.00 & 0.99 & 0.00 & -0.10 & 0.00 & 0.00 \\ 0.47 & -0.02 & -0.03 & 0.00 & 0.64 & -0.03 & -0.05 & 0.01 \\ 0.05 & 0.37 & -0.02 & -0.01 & -0.03 & 0.78 & -0.05 & -0.14 \\ 0.07 & 0.07 & -0.01 & 0.00 & -0.05 & -0.05 & 0.97 & -0.03 \\ -0.01 & 0.26 & -0.01 & 0.00 & 0.01 & -0.14 & -0.03 & 0.89 \end{pmatrix}$$

Où $\hat{\mathbf{P}}$ est la matrice des g premiers vecteurs propres de la matrice de covariance des données correspondantes aux g premières valeurs propres.

Les variables estimées sont données par :

$$\hat{z}_1(k) = 0.37z_1(k) + 0.03z_2(k) + 0.04z_3(k) + 0.47z_4(k-1) + 0.05z_1(k-1) \\ + 0.07z_2(k-1) - 0.01z_4(k-1)$$

$$\hat{z}_2(k) = 0.03z_1(k) + 0.32z_2(k) + 0.03z_3(k) + 0.01z_4(k) - 0.02z_1(k) + 0.37z_2(k-1) \\ + 0.07z_3(k-1) + 0.26z_4(k-1)$$

$$\hat{z}_3(k) = 0.04z_1(k) + 0.03z_2(k) + 0.99z_3(k) - 0.03z_1(k-1) - 0.02z_2(k-1) - 0.01z_3(k-1) \\ - 0.01z_4(k-1)$$

$$\hat{z}_4(k) = 0.01z_2(k) + 0.99z_4(k) - 0.01z_2(k-1)$$

$$\hat{z}_1(k-1) = 0.47z_1(k) - 0.02z_2(k) - 0.03z_3(k) + 0.64z_1(k-1) - 0.03z_2(k-1) - 0.05z_3(k-1) \\ + 0.01z_4(k-1)$$

$$\hat{z}_2(k-1) = 0.05z_1(k) + 0.37z_2(k) - 0.02z_3(k) - 0.01z_4(k) - 0.03z_1(k-1) + 0.78z_2(k-1) \\ - 0.05z_3(k-1) - 0.14z_4(k-1)$$

$$\hat{z}_3(k-1) = 0.07z_1(k) + 0.07z_2(k) - 0.01z_3(k) - 0.05z_1(k-1) - 0.05z_2(k-1) \\ + 0.97z_3(k-1) - 0.03z_4(k-1)$$

$$\hat{z}_4(k-1) = -0.01z_1(k) + 0.26z_2(k) - 0.01z_3(k) + 0.01z_1(k-1) - 0.14z_2(k-1) \\ - 0.03z_3(k-1) + 0.89z_4(k-1)$$

Ainsi la matrice des résidus est calculée comme suit :

$$\mathbf{e}_d(1) = \tilde{\mathbf{C}}\mathbf{x}_d(1)$$

Avec

$$\mathbf{e}_d(1) = [e_1(k) \ e_2(k) \ e_3(k) \ e_4(k) \ e_1(k-1) \ e_2(k-1) \ e_3(k-1) \ e_4(k-1)]^T$$

Et la matrice $\tilde{\mathbf{C}}$ égale à :

$$\tilde{\mathbf{C}} = (\mathbf{I} - \hat{\mathbf{C}})$$

$$= \begin{bmatrix} 0.629 & -0.036 & -0.046 & 0.001 & -0.470 & -0.058 & -0.071 & 0.010 \\ -0.026 & 0.674 & -0.004 & -0.003 & 0.043 & -0.379 & -0.132 & -0.239 \\ -0.026 & -0.004 & 0.001 & 0 & 0.019 & 0.005 & 0.005 & 0.001 \\ 0.001 & 0.003 & 0 & 0 & 0 & -0.002 & 0 & -0.001 \\ -0.470 & 0.043 & 0.019 & 0 & 0.351 & 0.022 & 0.063 & -0.016 \\ -0.048 & -0.381 & 0.005 & -0.002 & 0.022 & 0.225 & 0.085 & 0.137 \\ -0.091 & -0.132 & 0.005 & 0 & 0.063 & 0.085 & 0.041 & 0.047 \\ 0.010 & -0.239 & 0.001 & -0.001 & -0.016 & 0.137 & 0.047 & 0.086 \end{bmatrix}$$

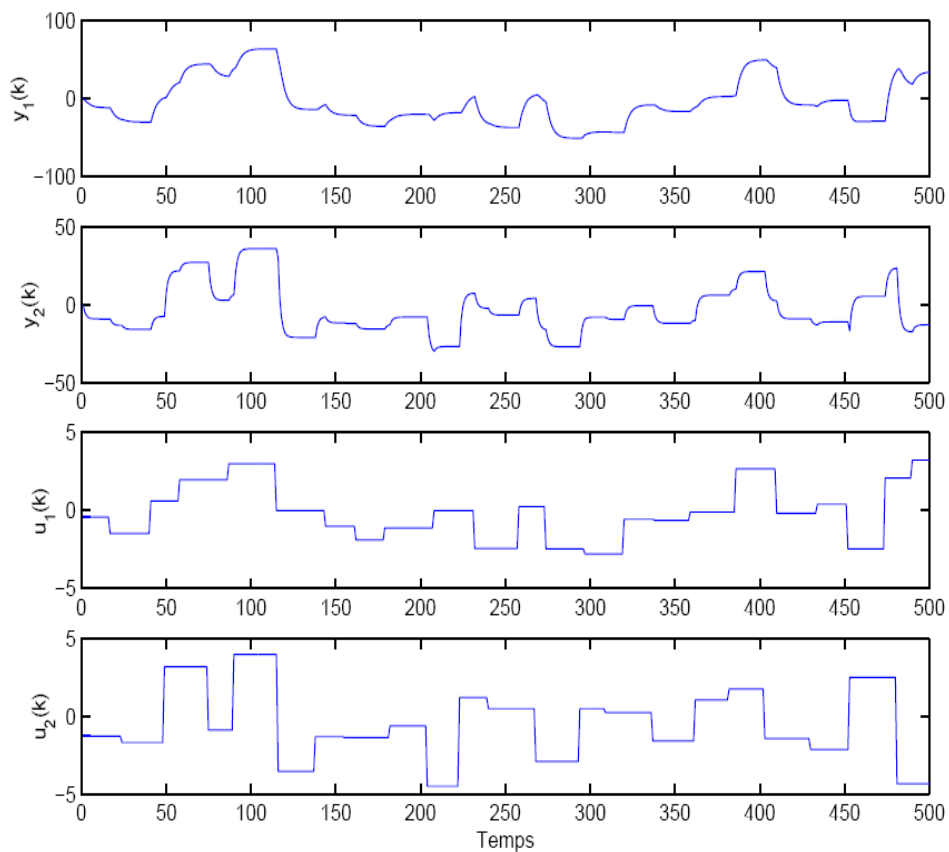


Fig. 5.1 : Mesures simulées de $z_1(k)$, $z_2(k)$, $z_3(k)$ et $z_4(k)$ en absence de défaut.

Les résidus sont donnés par :

$$e_1(k) = 0.629z_1(k) - 0.036z_2(k) - 0.046z_3(k) + 0.001z_4(k) - 0.47z_1(k-1) \\ - 0.058z_2(k-1) - 0.071z_3(k-1) + 0.01z_4(k-1)$$

$$e_2(k) = -0.026z_1(k) + 0.674z_2(k) - 0.071z_3(k) - 0.003z_4(k) + 0.043z_1(k-1) - 0.379z_2(k-1) \\ - 0.032z_3(k-1) - 0.239z_4(k-1)$$

$$e_3(k) = -0.026z_1(k) - 0.004z_2(k) + 0.001z_3(k) + 0.019z_1(k-1) + 0.005z_2(k-1) \\ - 0.005z_3(k-1) + 0.001z_4(k-1)$$

$$e_4(k) = 0.001z_1(k) + 0.003z_2(k) - 0.002z_2(k-1) - 0.001z_4(k-1)$$

$$e_1(k-1) = -0.47z_1(k) + 0.043z_2(k) + 0.019z_3(k) + 0.351z_1(k-1) + 0.022z_2(k-1) \\ + 0.063z_3(k-1) - 0.016z_4(k-1)$$

$$e_2(k-1) = -0.048z_1(k) - 0.381z_2(k) + 0.005z_3(k) - 0.002z_4(k) + 0.022z_1(k-1) + 0.225z_2(k-1) \\ + 0.085z_3(k-1) + 0.137z_4(k-1)$$

$$e_3(k-1) = -0.091z_1(k) - 0.132z_2(k) + 0.005z_3(k) + 0.063z_1(k-1) + 0.085z_2(k-1) \\ + 0.041z_3(k-1) + 0.047z_4(k-1)$$

$$e_4(k-1) = 0.01z_1(k) - 0.239z_2(k) + 0.001z_3(k) - 0.001z_4(k-1) - 0.016z_1(k-1) \\ + 0.137z_2(k-1) + 0.047z_3(k-1) + 0.086z_4(k-1)$$

Pour calculer le seuil δ_α^2 SPE qui nous permet de détecter les conditions anormales, on fixe a priori l'intervalle de confiance $\alpha = 0:05$ et on calcul :

$$\theta_1 = \sum_{j=g+1}^m \sigma_j^1 = 0.028$$

$$\theta_2 = \sum_{j=g+1}^m \sigma_j^2 = 5.44 \times 10^{-4}$$

$$g_\theta = \theta_2 / \theta_1 = 0.018$$

$$h = \theta_1^2 / \theta_2 = 1.527$$

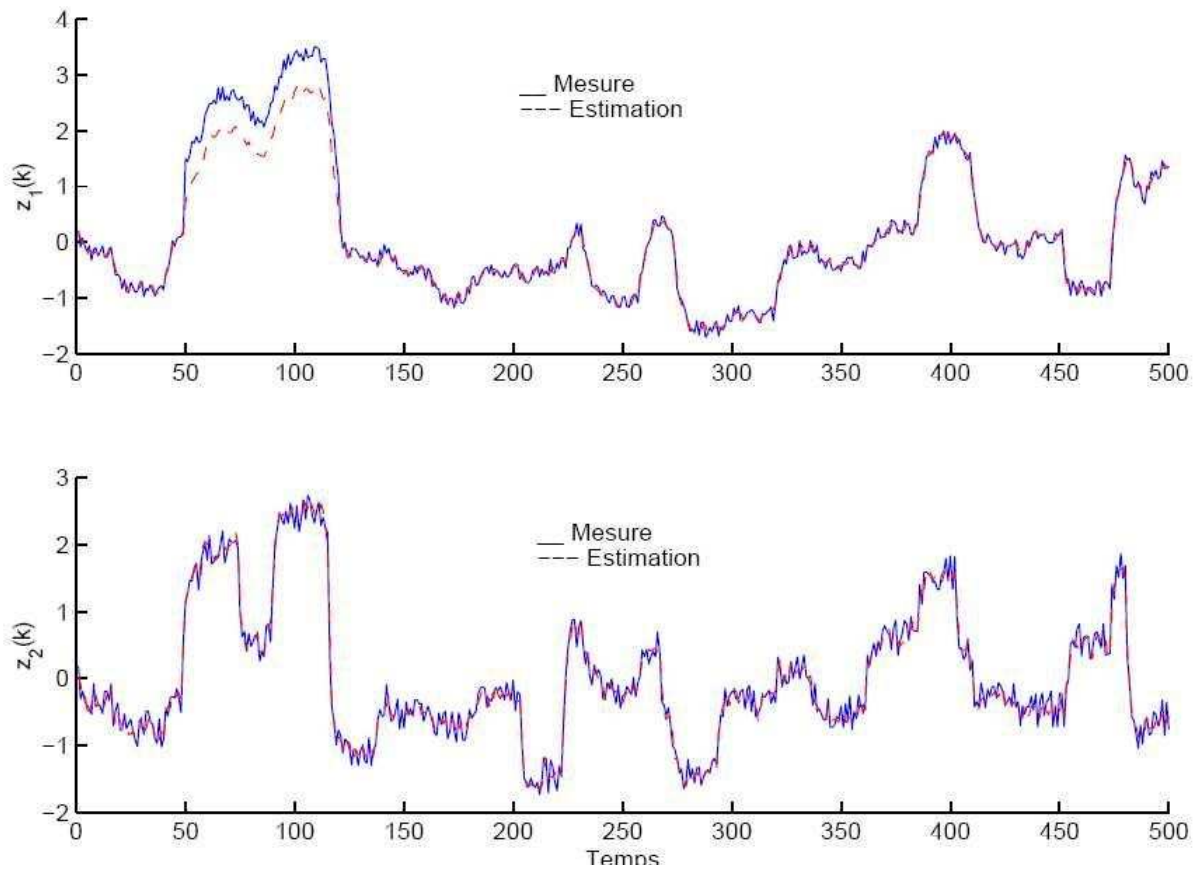


Fig. 5.2 : Evolution des mesures et des estimations des deux variables $z_1(k)$ et $z_2(k)$ en présence d'un défaut affectant la variable $z_1(k)$ de l'instant 50 à l'instant 120.

On peut donc calculer le seuil δ_α^2 qui est égal à :

$$\delta_\alpha^2 = 0.018 \chi_{1.527, 0.05}^2 = 0.095$$

Par la suite on génère un défaut d sur la variable $z_1(k)$ d'amplitude égale à 15% de la plage de variation de $z_1(k)$ de l'instant 50 à l'instant 120.

La figure (5.2) présente les mesures et les estimations des deux variables de sorties $z_1(k)$ et $z_2(k)$ en présence du défaut. On voit bien un écart entre la mesure et l'estimation de $z_1(k)$ de l'instant 50 à l'instant 120, contrairement à $z_2(k)$ où l'estimation suit parfaitement la mesure.

La figure (5.3) montre les 8 résidus $e_1(k)$, $e_2(k)$, $e_3(k)$, $e_4(k)$, $e_1(k+1)$, $e_2(k+1)$, $e_3(k+1)$ et $e_4(k+1)$. On voit bien le défaut sur $e_1(k)$, $e_3(k)$, $e_4(k)$, $e_1(k+1)$ et $e_3(k+1)$ contrairement à $e_2(k)$ et $e_4(k+1)$ où on ne le voit pas, car son amplitude est très faible.

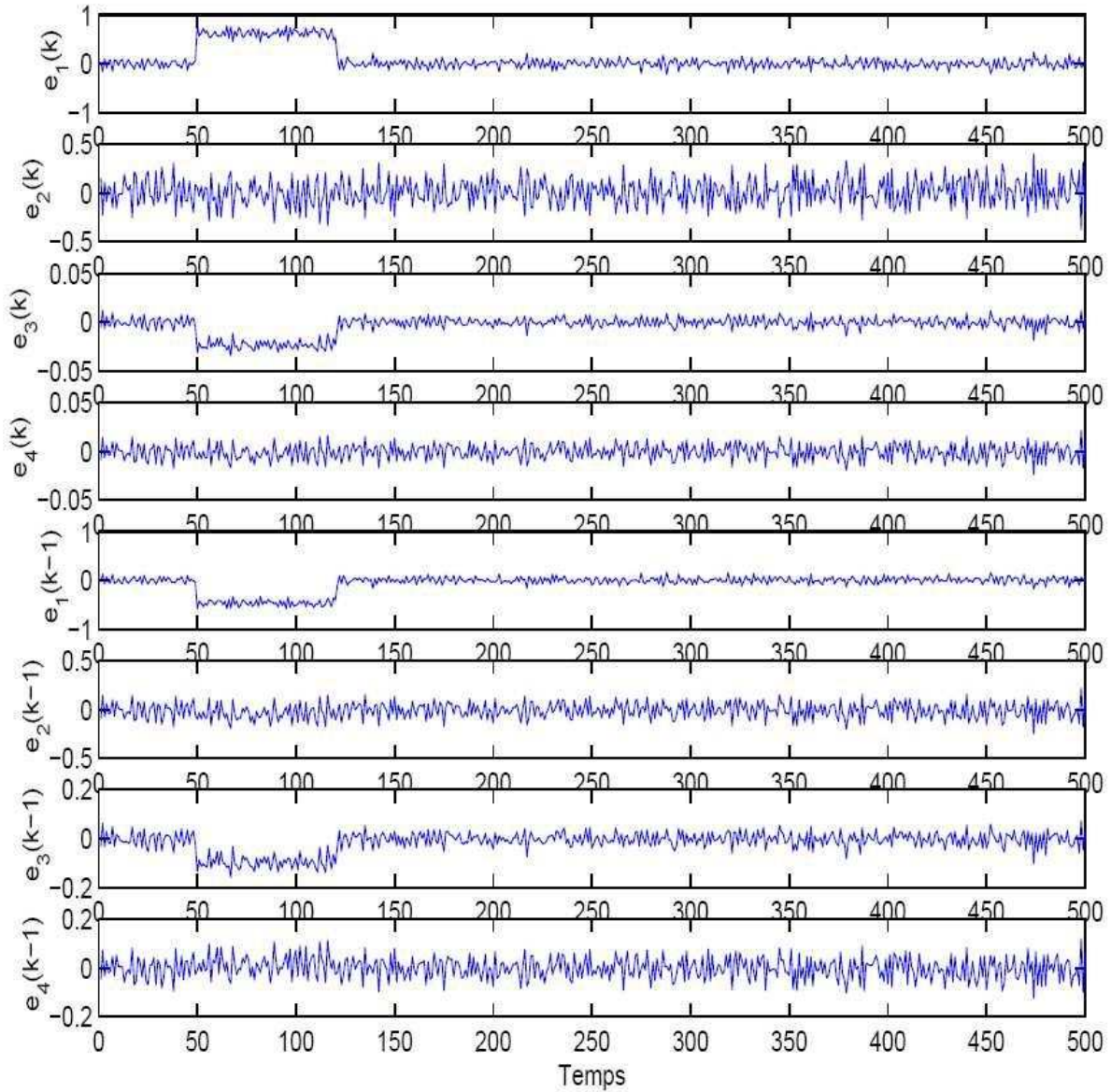


Fig. 5.3 : Evolution des 8 résidus $e_1(k)$, $e_2(k)$, $e_3(k)$, $e_4(k)$, $e_1(k+1)$, $e_2(k+1)$, $e_3(k+1)$ et $e_4(k+1)$

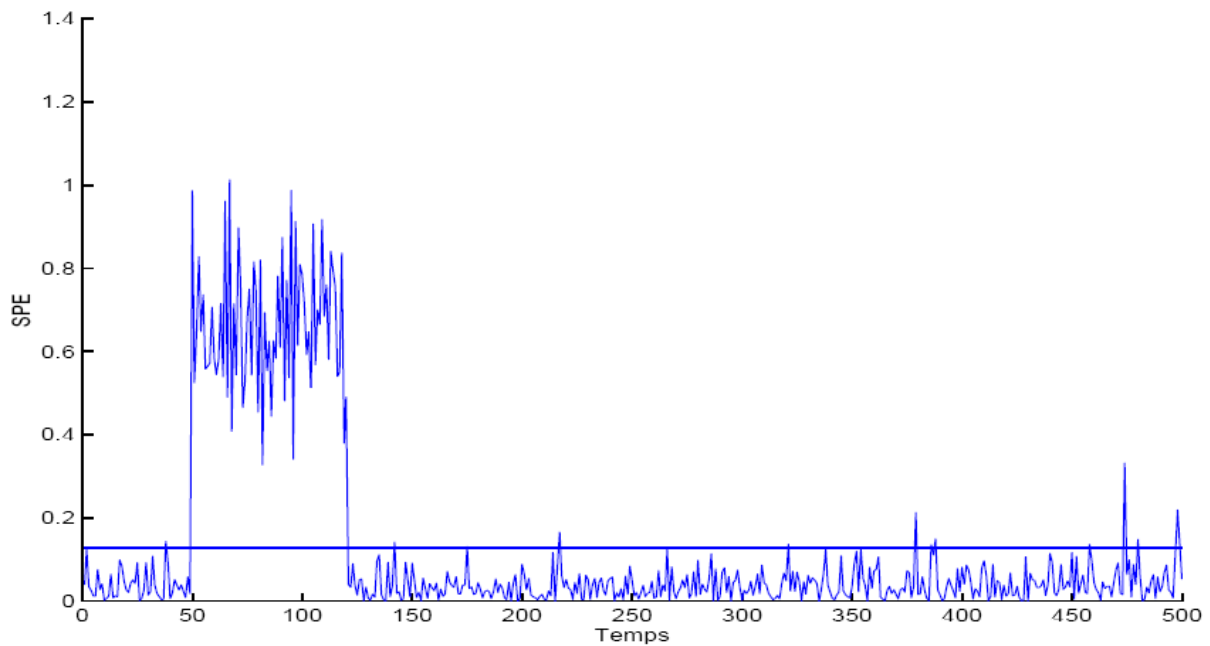


Fig. 5.4 : Evolution du SPE avec un défaut affectant la variable $z_1(k)$ de l'instant 50 à l'instant 120.

La figure (5.4) présente l'évolution du SPE avec un seuil de détection δ_α^2 égal à 0.095 pour une valeur de $\alpha = 0.05$. On voit que la statistique SPE nous permet de détecter le défaut d affectant la variable $z_1(k)$.

Par la suite, pour étudier l'influence des variances des bruits sur la détection des défauts, on génère un défaut d sur la variable $z_1(k)$ d'amplitude égale à 10% de la plage de variation de $z_1(k)$.

On fixe les écarts types des bruits de sorties à 7 et on fait varier les écarts types des bruits d'entrées en % de la plage de variation des entrées ; on calculera après le taux de fausses alarmes (une fausse alarme est la détection d'un défaut qui n'existe pas en réalité) et le taux de bonne détection (une bonne détection veut dire que si un défaut se présente, on le détecte).

A chaque niveau de bruit, on fait 1000 simulations, on calcule le taux de fausses alarmes pour chaque simulation, puis on prend la moyenne des taux de fausses alarmes.

Pour calculer le taux de bonne détection, on s'intéresse qu'à l'intervalle où le défaut intervient; on calcule le taux de bonne détection pour chaque simulation et on prend la moyenne des taux de bonnes détections.

La table 5.1 montre l'évolution du taux de fausses alarmes et de taux de bonne détection en fonction de se l'écart type de $\varepsilon_1(k)$ et $\varepsilon_2(k)$. On voit que si l'écart type $\varepsilon_1(k)$ et $\varepsilon_2(k)$ augmente, le taux de fausses alarmes augmente, et le taux de bonne détection diminue.

Le modèle ACPD utilisé dans cet exemple est un modèle biaisé car les variances des bruits ne sont pas identiques. Le résultat obtenu est assez acceptable puisque même avec un modèle ACPD biaisé et pour un écart type de $\varepsilon_3(k)$ et $\varepsilon_4(k)$ égal à 20% de la plage de variation des entrées ; on a un taux de bonne détection qui est égal à 83% ce qui est assez acceptable.

Maintenant on va voir l'influence de l'amplitude de défaut sur le taux de fausses alarmes et la taux de bonne détection. Pour cela on fixe les écarts types des bruits comme suit :

Un écart type égal à 7 pour $\varepsilon_1(k)$ et $\varepsilon_2(k)$.

Un écart type égal à 2 pour $\varepsilon_3(k)$ et $\varepsilon_4(k)$.

On génère un défaut d sur la variable $z_2(k)$ d'amplitude qui varie en fonction de la plage de variation de $z_2(k)$. A chaque valeur de l'amplitude du défaut, on calcul le taux de fausses alarmes et le taux de bonne détection.

La table 5.2 montre l'évolution du taux de fausses alarmes et le taux de bonne détection en fonction de en fonction de l'amplitude de défaut d affectant la variable $z_2(k)$. On voit que si le défaut diminue, le taux de bonne détection diminue et le taux de fausse alarme sensiblement constant (autour de 4.5 %).

l'écart type de $\varepsilon_3(k)$ et $\varepsilon_4(k)$ en % de la plage de variation des entrées	taux de fausses alarmes	taux de bonne détection
5.0 %	3.8 %	99.8 %
10 %	4.3 %	91.5%
20 %	4.6 %	83.0 %
30 %	5.1 %	73.2%
40 %	5.1 %	66.1%

Tab. 5.1 : Evolution du taux de fausses alarmes et du taux de bonne détection en fonction de l'écart type de $\varepsilon_3(k)$ et $\varepsilon_4(k)$.

l'amplitude du défaut d affectant la la variable $z_2(k)$ en % de la plage de variation de la variable $z_2(k)$	taux de fausses alarmes	taux de bonne détection
40 %	4.5 %	100 %
20 %	4.3 %	100 %
15 %	4.5 %	91.5 %
10 %	4.6 %	44.5 %
5 %	4.5 %	21.1 %

Tab. 5.2 : Evolution du taux de fausses alarmes et du taux de bonne détection en fonction de l'amplitude de défaut d affectant la variable $z_2(k)$.

Conclusion générale et perspectives

Dans le domaine du diagnostic, des méthodes basées sur le concept de redondance de l'information ont été développées. Leur principe repose généralement sur un test de cohérence entre un comportement observé du processus fourni par les capteurs et un comportement prévu fourni par une représentation mathématique du processus. Ces méthodes de redondance analytique nécessitent un modèle du système à surveiller. Ce modèle comprend un certain nombre de paramètres dont les valeurs sont supposés connues lors du fonctionnement normal. Le prix à payer est toutefois l'élaboration d'un modèle mathématique aussi complet que possible dont la qualité est primordiale pour l'obtention d'un système de détection performant. Cela est tout à fait envisageable lorsqu'il s'agit des systèmes de petite dimension (nombre de variable à surveiller). En revanche, pour les systèmes de taille plus importante, établir de telles relations mathématiques entre les variables paraît moins immédiat. De plus, le choix de modèle pour la génération de résidus est arbitraire et il se peut très bien que les corrélations entre certaines variables ne soient pas prises en compte. Comme alternative, les méthodes basées sur l'analyse en composantes principales (ACP), sont très intéressantes pour la mise en évidence des corrélations linéaires significatives entre les variables du processus sans formuler de façon explicite le modèle du système.

Les méthodes de détection de défauts pour des processus dynamiques devraient simultanément considérer les intercorrélations et les autocorrélations entre les données.

Pour satisfaire ce besoin, nous avons présenté un algorithme pour identifier le modèle ACP dynamique qui consiste à calculer l'ordre du modèle dynamique et le nombre de composantes principales à retenir. Une fois le modèle identifié, on peut appliquer les outils de détection développés pour l'ACP statique.

Une perspective de ce travail consiste à utiliser l'approche de l'erreur de reconstruction pour calculer le nombre de composantes principales à retenir afin d'identifier le modèle **ACPD**.

Egalement une autre perspective peut être exploitée concerne de développer une approche pour le cas de l'ACP non linéaire.

Annexe

Dans cette annexe, nous avons regroupé l'étude des observateurs et des filtres en raison de l'analogie de leur formulation. Dans la littérature, on parle exclusivement d'observateurs dans le cas déterministe (en référence à l'observateur classique de Luenberger) et de filtres dans le cas stochastique. Ainsi, pour les systèmes linéaires, la structure de base de ces estimateurs (ou reconstruteurs) est toujours la même.

Méthodes à base d'observateurs ou de filtres

L'estimation en utilisant les observateurs est un processus simulant le fonctionnement du système à partir d'un modèle mathématique où la sortie est corrigée par l'erreur d'estimation de la sortie. De manière générale, ce concept peut être représenté par la figure 1.

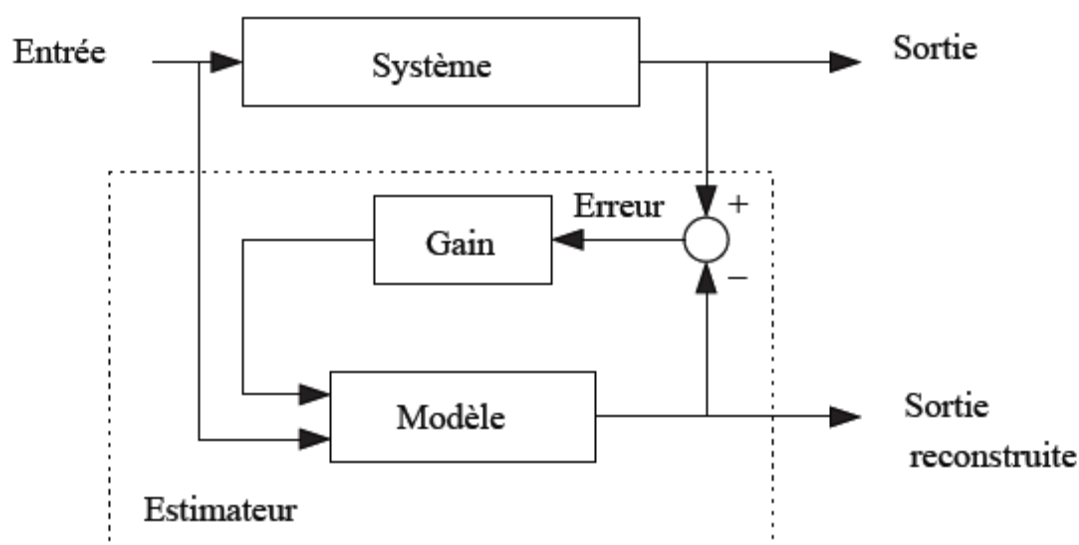


Fig. 1 : Principe général d'un estimateur de la sortie

La différence entre filtres et observateurs provient essentiellement du mode de calcul des paramètres du reconstituteur en fonction du contexte choisi : continu ou discret, déterministe ou stochastique. Cependant, en règle générale, les approches à base d'observateurs Clark et al. [33] consistent à comparer des fonctions des sorties estimées avec les mêmes fonctions de sorties mesurées. L'écart entre ces fonctions est utilisé comme résidu.

Il est intéressant de noter que, pour un schéma de diagnostic à base d'observateurs, seule l'estimation des sorties est nécessaire celle de l'état ne l'étant pas [33].

De plus, l'avantage de l'utilisation d'un observateur est la flexibilité dans le choix de son gain qui conduit à une riche variété de structures pour le diagnostic et la détection.

Dans le but d'obtenir la structure (générale) d'un observateur, considérons le système linéaire temps discret à paramètres invariants suivant

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \end{cases} \quad (1)$$

$u(k) \in \mathfrak{R}^p$ est la commande, $x(k) \in \mathfrak{R}^n$ est l'état et $y(k) \in \mathfrak{R}^m$ la sortie. Les matrices A, B et C sont supposées parfaitement connues et de dimensions adaptées. Un observateur de type proportionnel est utilisé pour reconstruire les variables du système à partir des entrées $u(k)$ et des sorties $y(k)$:

$$\begin{cases} \hat{x}(k+1) = A\hat{x}(k) + Bu(k) + He(k) \\ e(k) = y(k) - C\hat{x}(k) \end{cases} \quad (2)$$

Le schéma de l'observateur décrit par les équations (2) est illustré par la figure 1.

L'évolution de l'erreur d'estimation de l'état est, d'après (2), décrite par :

$$\begin{cases} e_x(k) = x(k) - \hat{x}(k) \\ e_x(k+1) = (A - HC)e_x(k) \end{cases} \quad (3)$$

L'erreur d'estimation de l'état $e_x(k)$ (et l'erreur de l'estimation de sortie $e_y(k)$) tendent asymptotiquement vers zéro

$$\begin{cases} \lim_{k \rightarrow \infty} e_x(k) = 0 \\ \lim_{k \rightarrow \infty} e_y(k) = 0 \end{cases}$$

Si l'observateur est stable, ce qui est garanti par un choix approprié de la matrice de gain H permettant de placer les valeurs propres de $(A - HC)$ dans le disque unité, sous réserve que (A, C) soit détectable.

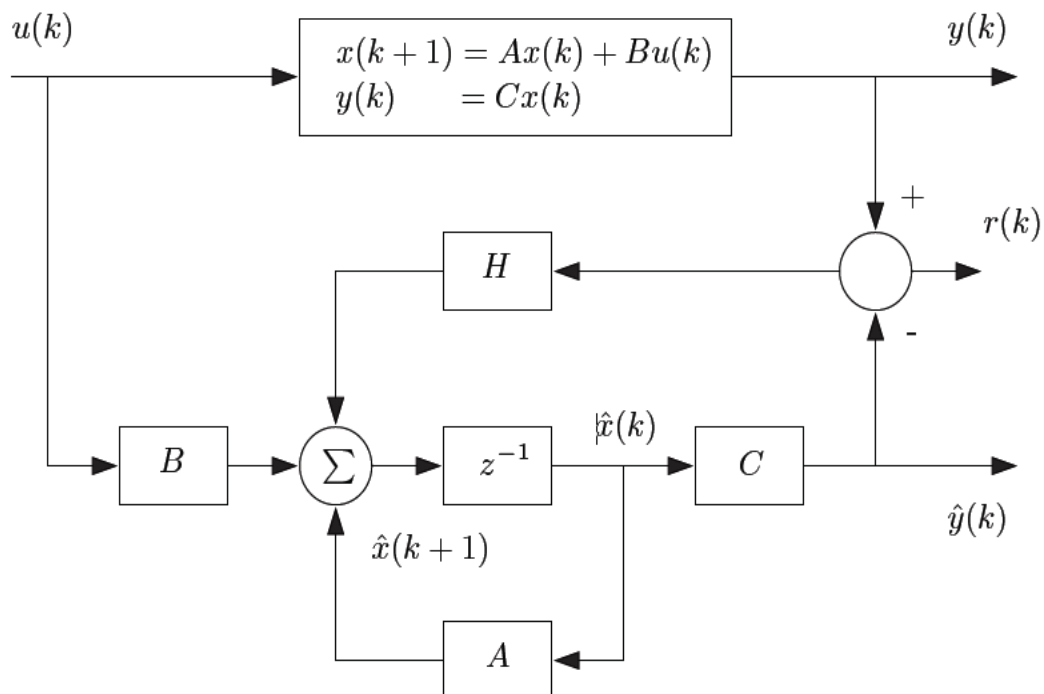


Fig. 2 : Système et observateur

Si le système est influencé par des perturbations et des défauts, le modèle suivant tient compte de ces nouvelles entrées :

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + Qw(k) + L_1 f(k) \\ y(k) = Cx(k) + Rw(k) + L_2 f(k) \end{cases} \quad (4)$$

où $w(k)$ est le vecteur de bruits et $f(k)$ le vecteur de défauts sur les entrées et les sorties agissant à travers des matrices connues L_1 et L_2 respectivement. Ce modèle prend en compte des défauts additifs sur les actionneurs, le système ainsi que les capteurs.

L'erreur d'estimation d'état s'exprime (en absence de bruits $w(k) = 0$) par :

$$e_x(k+1) = (A - HC)e_x(k) + L_1 f(k) - HL_2 f(k) \quad (5)$$

et l'erreur d'estimation de la sortie s'exprime en fonction de l'erreur d'estimation d'état $e_x(k)$.

On a :

$$e(k) = Ce_x(k) + L_2 f(k) \quad (6)$$

Quand un défaut $f(k)$ apparaît, l'erreur d'estimation d'état $e_x(k)$ s'écarte de zéro. De ce fait, l'erreur d'estimation de la sortie $e(k)$ s'écarte également de zéro, donc le signal $e(k)$ est un résidu noté $r(k)$ sur la figure 2.

En présence de bruits et de perturbations les résidus sont optimisés à l'aide d'un critère faisant intervenir la norme H_∞ et/ou H_- des fonctions de transfert matricielles. Ces normes s'avèrent être des outils adaptés pour la représentation des sensibilités des résidus vis-à-vis des perturbations et des défauts dans un contexte du pire des cas. Les normes H_∞ [34] et H_- [35] reflètent respectivement la valeur maximale et minimale des gains entre les signaux. L'objectif est alors de maximiser la sensibilité des résidus aux défauts tout en minimisant leur sensibilité aux perturbations.

A partir de cette structure de base, un autre type d'observateurs a été introduit dans la détection de défauts au milieu des années 80 : l'observateur à entrées inconnues (Unknown Input Observer : UIO). Des travaux concernant les observateurs pour des systèmes dont certaines entrées sont inconnues sont initiés par [36] et les premiers travaux publiés utilisant un observateur à entrées inconnues pour la détection et la localisation de défaut sont dus à [37]. L'observateur à entrées inconnues est aussi utilisé pour le cas des entrées toutes connues mais dont une partie seulement est utilisée afin de ne pas être affectés par d'éventuels défauts sur les entrées ignorées.

Supposons que le système est décrit par la représentation d'état discrète suivante :

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + Ef(k) \\ y(k) = Cx(k) \end{cases} \quad (7)$$

où $x(k) \in \mathfrak{R}^n$ est le vecteur d'état du système, $y(k) \in \mathfrak{R}^m$ le vecteur de sorties, $u(k) \in \mathfrak{R}^p$ le vecteur de commande et $f(k) \in \mathfrak{R}^q$ désigne le vecteur d'entrées inconnues. A , B , C et E sont des matrices de dimensions appropriées. La description générale d'un observateur à entrées inconnues d'ordre plein est donnée par l'expression suivante

$$\begin{cases} z(k+1) = Fz(k) + TBu(k) + Ky(k) \\ \hat{x}(k) = z(k) + Hy(k) \end{cases} \quad (8)$$

où $z(k) \in \mathfrak{R}^n$ est l'état de l'observateur, $\hat{x}(k)$ est l'estimé de l'état du système $x(k)$ et que F , T , H et K sont des matrices à déterminer pour accomplir le découplage de l'erreur d'estimation par rapport aux entrées inconnues.

L'erreur d'estimation d'état obtenue en appliquant l'observateur à entrées inconnues (8) au système (7) est de la forme

$$e_x(k+1) = Me_x(k) \quad (9)$$

Ceci veut dire que si toutes les valeurs propres de M sont à l'intérieur du disque unité, $e_x(k)$ tend de façon asymptotique vers zéro, c'est-à-dire $\lim_{K \rightarrow \infty} e_x(k) = 0$.

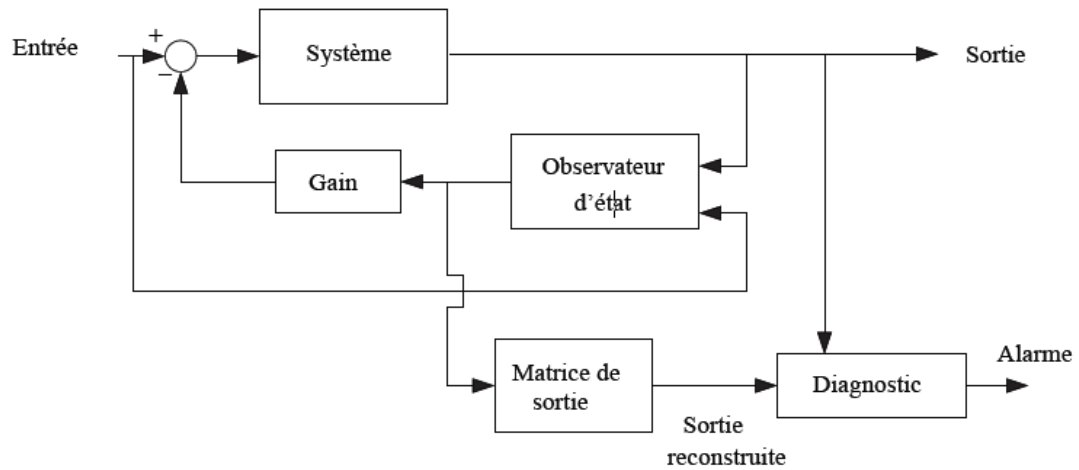
La synthèse d'un observateur à entrées inconnues revient à résoudre le système d'équations matricielles et les conditions de son existence sont données dans [38].

Remarques sur les reconSTRUCTEURS

Généralement, pour le diagnostic de systèmes régulés, on utilise un reconSTRUCTEUR de sortie du système bouclé. Dans ce cas, l'observateur est placé hors de la boucle de régulation.

Dans le cas où le régulateur est à base d'observateur (comme un retour d'état observé) on peut utiliser l'observateur inclus dans la boucle de régulation pour générer l'erreur de sortie. Cependant se pose le problème du réglage du gain de l'observateur pour assurer les performances du diagnostic sans diminuer les performances du système physique. Si on se place dans un contexte de système déterministe, sans incertitude de modèle, le principe de séparation est applicable sous réserve de commandabilité et d'observabilité du système.

On peut alors calculer séparément le régulateur et l'observateur. Ceci sous-entend qu'il est effectivement possible d'utiliser l'observateur de la boucle fermée pour faire du diagnostic comme illustré à la figure 3. On a donc deux possibilités pour le diagnostic des systèmes régulés : soit le module de diagnostic est externe au système régulé, dans ce cas on utilise les techniques de diagnostic en boucle ouverte en considérant le système régulé, soit le module de diagnostic est intégré dans la boucle de régulation. Ce point sera amplement discuté plus loin dans ce mémoire quand le diagnostic en boucle fermée sera exposé



*Fig. 3 : Diagnostic à l'aide d'un observateur d'état dans la boucle de régulation :
reconstruction d'état pour la régulation et la sortie pour la génération de résidus*

Filtre de Kalman

Nous allons faire un rappel non exhaustif de quelques notions sur le filtre de Kalman.

Le filtre de Kalman est l'un des estimateurs les plus utilisés. Élaboré pour un modèle stochastique, il permet de déterminer d'une manière optimale l'état du système lorsque ce dernier a été au préalable modélisé. Dans le cas d'une erreur de modélisation ou d'un changement significatif du mode de fonctionnement du système, le filtre de Kalman dans sa version de base ne garantit plus des résultats optimaux.

Il permet de résoudre de façon élégante le problème du filtrage linéaire. Utilisant la représentation d'état du système, le filtre de Kalman se présente sous la forme d'un ensemble d'équations récurrentes très faciles à résoudre d'un point de vue numérique. Il permet non seulement d'estimation optimale de l'état du système, mais aussi la variance de l'erreur d'estimation.

L'évolution de l'état du système est décrite par le système d'équations récurrentes :

$$\begin{cases} x(k) = Ax(k-1) + Bu(k-1) + w(k-1) \\ z(k) = Cx(k) + v(k) \end{cases} \quad (10)$$

où $x(k) \in \mathfrak{R}^n$ est le vecteur d'état à l'instant k , $A \in \mathfrak{R}^{n \times n}$ est la matrice d'évolution d'état, $u(k-1) \in \mathfrak{R}^p$ est le vecteur de commande à l'instant $k-1$, $B \in \mathfrak{R}^{n \times p}$ est la matrice des gains de l'entrée, $C \in \mathfrak{R}^{q \times n}$ est la matrice des gains de la sortie, $z(k) \in \mathfrak{R}^q$ est la sortie du système à l'instant k , $w(k-1) \in \mathfrak{R}^n$ est le bruit sur l'état à l'instant $k-1$, enfin $v(k) \in \mathfrak{R}^q$ est le bruit mesure w et v sont blancs, gaussiens et centrés tels que :

$$\begin{aligned} E [w(k)w^T(j)] &= Q(k)\delta_{kj} \\ E [v(k)v^T(j)] &= R(k)\delta_{kj} \\ E [v(k)w^T(j)] &= 0 \quad \forall k, j \end{aligned} \quad (11)$$

où $Q(k)$ et $R(k)$ sont les matrices de variance des bruits et δ_{kj} est le symbole de Kronecker

$$\delta_{kj} = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

L'état $x(k)$ et l'observation $z(k)$, qui se déduisent linéairement des bruits $v(k)$, $w(k)$ et des conditions initiales gaussiennes $x(0)$, avec $E[x(k)] = 0$, sont eux aussi gaussiens. L'expression du filtre de Kalman peut s'obtenir de manière itérative. À partir des conditions initiales $\hat{x}(0/0)$ et de la matrice de covariance $P(0/0)$, les équations du filtre de Kalman discret s'écrivent :

$$\begin{aligned} P(k+1|k) &= AP(k|k)A^T + Q(k) \\ \hat{x}(k+1|k) &= A\hat{x}(k|k) + Bu(k) \\ K(k+1) &= P(k+1|k)C^T (CP(k+1|k)C^T + R(k+1))^{-1} \\ \hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + K(k+1)(z(k+1) - C\hat{x}(k+1|k)) \\ P(k+1|k+1) &= (I - K(k+1)C)P(k+1|k) \end{aligned} \tag{12}$$

Le résidu $v(k+1) = z(k+1) - C\hat{x}(k+1|k)$ est la différence entre la mesure actuelle et la meilleure prédiction de la mesure, il sert à la réadaptation de l'estimée du filtre. Si les hypothèses précédentes (11) sur les bruits sont vérifiées, on a :

$$\begin{aligned} E[v(k)] &= 0 \\ E[v(k)v^T(j)] &= 0, \quad k \neq j \\ E[v(k)v^T(k)] &= S(k) = CP(k|k-1)C^T + R(k) \end{aligned} \tag{13}$$

Notons $N(V(k); 0, S(k))$ la distribution gaussienne de $V(k)$ de moyenne nulle et de matrice de variance $S(k)$:

$$N(V(k); 0, S(k)) = |2\pi S(k)|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} V^T(k) S^{-1}(k) V(k) \right\} \tag{14}$$

La densité de probabilité de $\nu(k)$ est alors donnée par :

$$p(\nu(k)) = N(\nu(k); 0, S(k)) \quad (15)$$

Si le résidu n'a pas ces caractéristiques, alors les hypothèses (11) ne sont pas vérifiées, cela peut être dû à une mauvaise modélisation du système ou à un défaut sur le système.

À partir des équations (14), plus la valeur du résidu est proche de 0 avec une variance faible, plus la valeur de la densité de probabilité augmente et plus le modèle du système représente le comportement du système. La valeur de cette densité de probabilité ou fonction de vraisemblance du modèle, sera un élément important dans l'évaluation des modèles dans les méthodes multi-modèle.

Le filtre de Kalman est largement utilisé, mais on peut recourir à d'autres types de filtres, le choix sera motivé par l'application et les objectifs de l'estimation.

BIBLIOGRAPHIE

- [1] W. Ku, R. H. Storer and C. Georgakis, (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chem. Intell. Lab. Syst.* 30, 179-196.

- [2] M. J. Piovoso, K. A. Kosanovich, P. K. Pearson. Monitoring process performance in real time. In : *American Control Conf. Chicago, IL.* 2359-2363.

- [3] J. F. MacGregor, (1989). Multivariate statistical methods for monitoring large data sets from chemical processes. In : *AICHE Meeting.*

- [4] D. M. Himes, R. H. Storer and C. Georgakis, (1994). In : *Proc. of the American Control Conf. IEEE Press, Piscataway, NJ.* 1279-1283.

- [5] L. Ferré, (1995). Selection of components in principal component analysis : A comparaison of méthodes. *Computational Statistics and Data Analysis.* 669-682.

- [6] S. Valle, L. Weihua, and S. J. qin, (1999). Selection of the number of principal components : the variance of the reconstruction error criterion with a comparison to other methods. *Industrial. Engineering Chemistry Research.* 38, 4389-4401.

- [7] J. V. Kresta, J. F. MacGregor and T. E. Marlin, (1991). Multivariate statistical monitoring of process operating performance, *The Canadian Journal of Chem. Engng.* 69, 35-47.

- [8] M. F. Harkat, (2003). Détection et localisation de défauts par analyse en composantes principales. Thèse de l'Institut National Polytechnique de Lorraine.

-
- [9] Y. Huang , J. Gertler and T. J. McAvoy, (2000). Sensor and actuator fault isolation by structured partial PCA with nonlinear extensions. *Journal of Process Control*. 10, 459-469.
- [10] C. Lee, S.W. Choi and I.B. Lee, (2004). Sensor fault identification based on time-lagged PCA in dynamic processes. *Chemometrics and Intelligent Laboratory Systems*. 70, 165-178.
- [11] M. F. Harkat, G. Mourot, and J. Ragot, (2000). Sensor failure detection of air quality monitoring network. *IFAC Symposium on fault detection, Supervision and Safety for Technical Processes, SAFEPROCESS'2000*, Budapest, Hungary, June. 14-16.
- [12] G. C. Tiao and G. E. P. Box, (1981). *J. Am. Statist. Assoc.* 76, 802.
- [13] M. F. Harkat, G. Mourot, and J. Ragot, (2001). Sensor failure detection and isolation of air quality monitoring network. *4th International Conference on Acoustical and Vibratory Surveillance Methodes and Diagnostic Techniques*, Compiègne, France.
- [14] R. Dunia, S. J. Qin, T.F. Edgar and T. J. McAvoy, (1996). Identification of faulty sensors using principal component analysis. *AIChE. J.* 42, 2797-2812.
- [15] L. Ljung, (1987). *System identi_cation : Theory for the user*. Prentice-hall, Englewood Clins, NJ.
- [16] J. V. Kresta, J. F. MacGregor and T. E. Marlin, (1991). Mutivariate statiscal monitoring and diagnosing process and product performance. *Journal of quality technology*. 28, 409-428.
- [17] B. M. Wise and N. B. Gallagher, (1996). The process chemometrics approach to process monitoring and fault détection. *Journal of Process Control*. 6, 329-348.

- [18] T. Kourti and J. F. MacGregor, (1996). Recent developments in multivariate SPC methods for monitoring of process operating performance. *Can. J. Chem. Eng.* 69, 35-47.
- [19] G. E. P. Box, (2000). Some theorems on quadratique forms applied in the study of analysis of variance problems : Efect of inequality of variance in one-way classification. *The Anals of Mathematical Statistics.* 34, 290-302.
- [20] W.E. Larimore, (1983). Predictive inference, suficiency, entropy, and an asymptotic likelihood principle. *Bio-metrika.* 70, 175-81.
- [21] W. E. Larimore, (1983). System identification, reduced-order filtering and modelling via canonical variate analysis, In : H. S. Rao, T. Dorato. Eds. *Proc. American Control Conf. IEEE, New York.*
- [22] W. Li and S.J. Qin, (2001). Consistent dynamic PCA based on errors-invariables subspace identification. *Journal Process Control.* 11, 661-678.
- [23] H. Akaike, (1974). A new look at the statistical model identification. *IEEE Trans. Auto. Cont.* 19, 716-723.
- [24] J. Wang and S. J. Qin, (2002). A new subspace identification approach based on principal component analysis. *Journal of Process Control.* 12, 841-855.
- [25] W. E. Larimore and R.K. Mehra, (1985), The problem of obverting data. *Byte.* 10-167.
- [26] K. I. Diamantaras., S. Y. Kung, (1996). *Principal component neural networks. Theory and applications.* John Wiley & Sons.
- [27] I. T. Jolliffe, (1986). *Principal component analysis.* Springer-Verlag, New York.

-
- [28] W. A. Shewhart, (1931). Economic control of quality of manufactured product. D. VanNostrand, New York, NY.
- [29] Besse P. et Ferré L. (1993). Sur l'usage de la validation croisée en analyse en composantes principales. *Revue de Statistique Appliquée*, XLI (1), 71-76.
- [30] Himes D. M., Storer R. H. and Georgakis C. (1994). Determination of the number of principal component for disturbance detection and isolation. *Proceedings of ACC*, Baltimore.
- [31] Afnor (1994). Maintenance - concepts et définitions des activités de maintenance, norme NF x 60-010. Association Française de Normalisation, 28.
- [32] M. Basseville, (1999). On fault detectability and isolability. In *Proceedings of the 5th European Control Conference*, Karlsruhe, Allemagne.
- [33] R. N. Clark, D. C. Fosth, and W. M. Walton, (1975). Detecting instrument malfunctions in control systems. *IEEE Transactions on Aerospace and Electronic Systems*. 11 (4), 465-473.
- [33] J. Chen, and R. Patton, (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. London : Kluwer Academic Publishers.
- [34] G. H. Golub and C. F. Van Loan, (1991). *Matrix Computations*. Baltimore : The Johns Hopkins University Press.
- [35] F. Rambeaux, (2001). *Génération et évaluation des résidus pour le diagnostic des systèmes incertains : approche fréquentielle*. Thèse, Université Henri Poincaré Nancy 1, France.

- [36] Wang, S., E. J. Davison, and P. Dorato, (1975). Observing the states of systems with unmeasurable disturbances. *IEEE Transactions on Automatic Control* 20(5) 716–717.
- [37] N. Viswanadham and R. Srichander, (1987). Fault detection using unknown input observers. *Control Theory and Advanced Technology*. 3 (2), 91–101.
- [38] J. Chen, and R. Patton, (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. London : Kluwer Academic Publishers.