

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

Université Kasdi Merbah Ouargla



Faculté des nouvelles technologies de l'informations et de télécommunications

Département d'électronique et de télécommunications

**MEMOIRE**

**Pour obtenir le Diplôme de Master**

**Option :** Electronique des systèmes embarqués

Présenté Par :

Bettayeb Mohamed El habib et Belhelib Hossine

-THEME-

---

# **Analyse de données pour le diagnostic des défaillances**

---

Soutenu le :20/06/2023

**Les membres de jury :**

**Président :** Mme Kara Fouzia UKM Ouaregla

**Examineur :** Mr Benathmane Khaled UKM Ouaregla

**Encadreurs :** Mr Chakour Chouaib UKM Ouaregla

**Année Universitaire : 2022/2023**

## ملخص :

تتناول هذه المذكرة العمل في نمذجة وتشخيص الأخطاء في الأنظمة الصناعية ذات الطابع غير الموكد والطريقة المستخدمة هي تحليل العناصر الرئيسية يهدف هذا العمل الى اتقان مختلف النماذج وتحليل وتمثيل البيانات الواردة من المستشعرات ومن ثم تطوير استراتيجية تشخيص تمن من اتخاذ قرار موثوق بشأن القياسات المعيبة

الكلمات المفتاحية : تشخيص، خلل، كشف، تحديد الموقع

## Résumé :

L'objectif de ce travail est de maitriser différents paradigmes de représentation, d'analyse et de modélisation des données « imprécises » venant des capteurs, puis de développer une stratégie de diagnostic permettant une prise de décision fiable sur les mesures en défaut. L'approche par intervalle de l'ACP est utilisée pour répondre à cette question.

**Mots-clés : Diagnostic, défaut, détection, localisation, ACP, ACP par intervalle.**

## Abstract :

The work presented in this thesis is dedicated to the modeling and diagnosis of errors in uncertain industrial systems. The method used is the principal component analysis (PCA). The objective of this work is to master different paradigms of representation, analysis, and modeling, of "imprecise" data coming from sensors, and then develop a diagnostic strategy that enables reliable decision-making an faulty measurement. The interval approach of PCA is used to address this question.

**Keywords: Diagnosis, fault, detection, isolation, PCA, Interval PCA.**



## remerciements


En tout premier lieu, je remercie ALLAH, tout puissant qui nous a donné la santé et le courage pour accomplir ce travail. En achevant ce modeste travail, nous remercions vivement notre promoteur chargé du suivi de notre travail :

### **Dr. Chakour Chouaib**

Pour ses encouragements, ses conseils et sa disponibilité tout au long de l'élaboration de ce travail et afin de nous faciliter le travail et pour son aide précieuse et ses conseils.

Je tiens à exprimer ma gratitude et ma haute considération aux membres de jury, de m'avoir honorée par leurs présences dans mon jury et pour l'intérêt qu'elles ont portées à ce travail.

Nous remercions également tous les professeurs du département d'électronique pour leurs patience et aide tout au long de cette période.



## Liste de figures

Figure I.1: Les méthodes de diagnostic.....	4
Figure II.1 :Evolution des différentes variables de simulation .....	11
Figure II.2 : Evolution des composantes principales .....	13
Figure II.3 : Evolution des différentes mesures et leurs estimations.....	14
Figure II.4: calcule SPE sans défaut.....	15
Figure II.5: calcule SPE dans le cas de défaut .....	16
Figure II.6 : Evolution des différents SPE obtenus par la méthode d'élimination sans défaut.....	17
Figure II.7 : Evolution des différents SPE obtenus par la méthode d'élimination avec un défaut affectant la variable x2 à partir de l'instant 700.....	17
Figure III.1 : Evolution des différentes variables en type intervalle .....	25
Figure III.2 :Evolution des différentes mesures et leurs estimations .....	25
Figure III.3 :Evolution des différentes variables reconstruit.....	25
Figure III.4 :Evolution des différents résidu sans défaut .....	25
Figure III.5 :Evolution des différents résidu dans le cas de défaut .....	28
Figure III.6 :Evolution en temp de l'index [SPE] sans défaut .....	29
Figure III.7 :Evolution en temp de l'index [SPE] avec défaut.....	29
Figure III.8 :Le temps d'évolution de SPE calculé après la reconstruction de différentes variables sans défaut.....	31
Figure III.9 :Le temps d'évolution de SPE calculé après la reconstruction de différentes variables avec défaut --- par la méthode d'élimination avec un défaut affectant la variable x1 à partir de l'instant 700.....	31
Figure IV.1 :la station automatique météorologique.....	32
Figure IV.2 : WindSonic .....	33
Figure IV.3: PT-100 Thermomètre.....	33
Figure IV.4 : Capteur de l'humidité relative et de température RHT175 .....	33
Figure IV.5 : Capture de pression MSB780.....	34
Figure IV.6: MR3 et MR3H Pluviomètre .....	35
Figure IV.7 : CMP10 (CMP11) Pyranomètre .....	35
Figure IV.8: Arceau d'Ombrage CM121B/C .....	35
Figure IV.9 :La description des variables du station.....	37
Figure IV.10 : – L'évolution des différentes mesures .....	37
Figure IV.11 :Evolution des différentes mesures et leurs estimations.....	38
Figure IV.12: L'évolution de l'indices SPE en absence de défaut.....	38

Figure IV.13 : L'évolution de l'indices SPE dans le cas de défaut.....	39
Figure IV.14 : Calcul des reconstruction des variables en absence de défaut.....	39
Figure IV.15 :Calcul des reconstruction des variables en cas de défaut .....	40
Figure IV.16 : Evolution des différents SPE obtenus par la méthode d'élimination avec un défaut affectant la variable $x_2$ à partir de l'instant 7000 .....	40
Figure IV.17 : Evolution des différentes mesures et leurs estimations .....	41
Figure IV.18 : L'évolution de l'indices SPE en type intervalle en absence de défaut.....	42
Figure IV.19 : L'évolution de l'indices SPE en type intervalle dans le cas de défaut .....	42
Figure IV.20 :Calcul des reconstruction des variables à l'indice SPE en type intervalle en absence de défaut .....	42
Figure IV.21 :Calcul des reconstruction des variables à l'indice SPE en type intervalle en cas de défaut.....	43
Figure IV.22 :Evolution des différents SP E obtenus par la méthode d'élimination avec un défaut affectant la variable $x_2$ à partir de l'instant 7000 .....	43

# Sommaire

Résumé.....	I
Liste de figures.....	II
Introduction générale.....	1

## Chapitre I : généralité sur le diagnostic

Introduction.....	2
I.1 Définitions de base :.....	2
I.2 Caractéristiques d'un système de diagnostic :.....	3
I.3 méthodes de diagnostic :.....	3
I.4 Étapes de diagnostic :.....	4
I.5 L'Analyse en Composantes Principales (ACP) :.....	4
Conclusion:.....	5

## Chapitre II : L'analyse en composantes principales

Introduction.....	6
II.1.Historique.....	6
II.2.L'analyse en composantes principales.....	6
II.2.1.Principe.....	6
II.2.2.Avantages et inconvénients de l'ACP.....	6
II.2.2.1.Avantages.....	8
II.2.2.2.Inconvénient.....	8
II.2.4 Détermination de la structure du modèle ACP :.....	10
II.2.5.Pourcentage cumulé de la variance totale (PCV) :.....	10
II.2.6 Exemple de simulation :.....	11
II.3.Détection et localisation des défauts.....	14
II.3.1.Indice <i>SPE</i> .....	14
II.3.2.Reconstruction des variables.....	15
II.3.3.Simulation :.....	15
Conclusion :.....	18

## Chapitre III : Analyse en composantes principales par intervalle

Introduction.....	19
III.1. Analyse en composantes principales par intervalle.....	19
III.1.1. données évaluées à intervalles.....	19

III.2. La normalisation des données évaluées à intervalles : .....	20
III.2.1. normalisation à l'aide de la dispersion des intervalles centraux et de la plage .....	20
III.2.2 normalisation par dispersion des centres d'intervalle : .....	21
III.2.3. normalisation en utilisant la dispersion des intervalles de limites: .....	21
III.2.4. normalisation en utilisant la gamme mondiale : .....	22
III.3. méthode des sommes (VPCA) : .....	22
III.4 Reconstruction des variables pour le modèle (VPCA): .....	25
III.5 PCA à valeur d'intervalle pour la détection des défauts: .....	26
III.5.1 Graphique univarié: .....	26
III.5.2 SPE à valeur intervalle : .....	28
III.5.3. Localisation et isolation des failles à l'aide de PCA pour des données à intervalles de valeur : .....	29
Conclusion : .....	31

## **Chapitre IV : Application sur la station automatique météorologique**

Introduction : .....	32
IV.1 la station automatique météorologique: .....	32
IV.2 Système météorologique intégré: .....	33
IV.2.1 WindSonic: .....	33
IV.2.1.1 Application .....	33
IV.2.2 PT-100 Thermomètre.....	33
.....	33
IV.2.2.1 Application .....	33
IV.2.3 Capteur de l'humidité relative et de température RHT175.....	33
IV.2.4 Capteur de pression MSB780 .....	34
IV.2.4.1 Application .....	34
IV.2.5 MR3 et MR3H Pluviomètre.....	34
IV.2.5.1 Application .....	34
IV.2.6 CMP10 (CMP11) Pyranomètre.....	35
IV.2.6.1 Application .....	35
IV.2.7 Arceau d'Ombrage CM121B/C .....	35
IV.2.7.1 Application .....	35
IV.3 le diagnostic Dans le domaine de la météorologie:.....	36

IV.4 Modélisation par ACP :.....37

    IV.4.1 Détection et localisation de défaut : .....38

    IV.4.2 Modélisation par ACP en type intervalles.....40

Conclusion :.....43

Conclusion générale.....45



## Introduction Générale

L'évolution des défis économiques a conduit à une production sans cesse croissante. Dans un environnement où les performances sont critiques, la moindre défaillance d'un processus peut être préjudiciable. Par conséquent, il est nécessaire de s'assurer en permanence que le processus se déroule de manière optimale. Les informations susceptibles de transformer le comportement du système peuvent être données par des mesures de variables du processus

La qualité des mesures est un facteur important dans la surveillance et l'évaluation des performances du procédé. La qualité des informations peut être améliorée en améliorant la précision des instruments et en ajoutant des capteurs. Pour des raisons techniques ou financières, cette solution de mesure d'une même grandeur par plusieurs capteurs est réservée à l'industrie de pointe. De plus, cette redondance matérielle ne protège pas contre la défaillance de certains éléments communs de la chaîne de mesure : plusieurs capteurs mesurant la même grandeur sont souvent géographiquement proches les uns des autres et alimentés par le même réseau ; une coupure de courant peut arrêter tout le système de mesure. L'utilisation de modèles exacts a priori liant différentes grandeurs mesurées offrent un autre moyen de vérifier la fiabilité des mesures. L'avantage de cette redondance analytique est qu'elle n'augmente pas les coûts d'installation et n'est pas limitée par les matériaux.[13]

Dans le domaine du diagnostic, des méthodes basées sur le concept de redondance d'informations ont été développées. Leur raisonnement est généralement basé sur un test d'accord entre le comportement observé d'un processus fourni par des capteurs et le comportement prédit fourni par une représentation mathématique du processus. Par conséquent, les méthodes de redondance analytique nécessitent des modèles de système de surveillance. La comparaison entre le comportement réel du système et le comportement attendu donné par le modèle fournit une quantité appelée résidu qui sera utilisée pour déterminer si le système est dans un état défaillant.[13]

Le contenu principal de notre mémoire est organisé comme suit :

Le chapitre 01 offre une vue d'ensemble sur la thématique du diagnostic des défauts. Il présente les principes fondamentaux de l'approche diagnostic ainsi que les méthodes les plus couramment utilisées dans ce domaine.

Le chapitre 02 présente la méthode d'analyse en composantes principale pour la détection et l'isolation des défauts.

Le chapitre 03 concerne l'extension de la méthode ACP pour la modélisation et le diagnostic de défauts des données incertaines de type intervalle.

Le quatrième chapitre présente une application des différentes techniques étudiées pour la prédiction des défauts de mesures dans une stations météorologique automatique.

# **Chapitre I : généralité sur le diagnostic**

**Introduction :**

La détection et la localisation de défauts sont des tâches critiques dans de nombreux domaines tels que l'ingénierie, l'industrie, la santé et la sécurité.

**I.1 Définitions de base :**

- ❖ **Diagnostic :** Le diagnostic est le processus de détermination de la cause d'un problème dans un système électronique en utilisant des outils de diagnostic spécialisés tels que des scanners de diagnostic, des oscilloscopes, des multimètres et des analyseurs de spectre.
- ❖ **Capteur :** Un capteur est un dispositif électronique qui mesure une grandeur physique telle que la température, la pression, la vitesse, l'angle, etc. et convertit cette mesure en une forme de signal électrique pour le traitement ultérieur.
- ❖ **Actuateur :** Un actuateur est un dispositif électronique qui convertit un signal électrique en une action mécanique, telle que l'ouverture ou la fermeture d'une vanne, la régulation de la vitesse d'un moteur, etc.
- ❖ **Code de diagnostic :** Un code de diagnostic est un code numérique qui est généré par un système de diagnostic électronique en réponse à un problème détecté dans le système. Les codes de diagnostic sont utilisés pour aider les techniciens à identifier la cause du problème et à effectuer les réparations nécessaires.
- ❖ **Données en temps réel :** Les données en temps réel sont les données en continu qui sont collectées à partir des capteurs et des autres dispositifs électroniques du système en temps réel. Ces données sont utilisées pour surveiller le fonctionnement du système et pour détecter les anomalies et les problèmes.
- ❖ **Système de commande :** Un système de commande est un système électronique qui contrôle le fonctionnement d'un système mécanique, électrique ou électronique. Les systèmes de commande utilisent des capteurs pour surveiller les conditions de fonctionnement et des actionneurs pour réguler le fonctionnement du système.

**I.2 Caractéristiques d'un système de diagnostic :**

Les caractéristiques d'un système de diagnostic peuvent varier en fonction du domaine d'application, mais voici quelques caractéristiques générales que l'on peut retrouver dans un système de diagnostic :

- ✓ **Fiabilité :** Un système de diagnostic doit être fiable et précis dans la détection des problèmes. Cela signifie qu'il doit fournir des résultats cohérents et précis même dans des conditions variables.
- ✓ **Rapidité :** Un système de diagnostic doit être capable de fournir des résultats rapidement pour minimiser le temps d'arrêt des équipements et des systèmes. Cela est particulièrement important dans les domaines industriels où chaque minute d'arrêt peut coûter cher.
- ✓ **Facilité d'utilisation :** Un système de diagnostic doit être facile à utiliser et à comprendre pour les techniciens et les utilisateurs, même s'ils n'ont pas une formation approfondie en électronique ou en mécanique.
- ✓ **Polyvalence :** Un système de diagnostic doit être capable de diagnostiquer différents types de problèmes et de systèmes dans différentes industries et applications.

- ✓ **Capacité de stockage** : Un système de diagnostic doit être capable de stocker les données de diagnostic pour permettre une analyse plus approfondie et pour fournir un historique de diagnostic.
- ✓ **Mises à jour** : Un système de diagnostic doit être capable de recevoir des mises à jour régulières pour améliorer la précision et la fiabilité des résultats de diagnostic.
- ✓ **Support technique** : Un système de diagnostic doit être accompagné d'un support technique pour aider les utilisateurs à comprendre et à résoudre les problèmes rencontrés lors de l'utilisation du système.[1]

### I.3 Méthodes de diagnostic :

On peut globalement distinguer deux grandes familles dans les méthodes de diagnostic (comme il est illustré dans la figure II.1) Les méthodes basées sur une modélisation des systèmes ou des signaux sans modèle, que nous dénommerons diagnostic quantitatif. Les méthodes basées sur l'intelligence artificielle, avec modèle, que nous appellerons diagnostic qualitatif.[1]

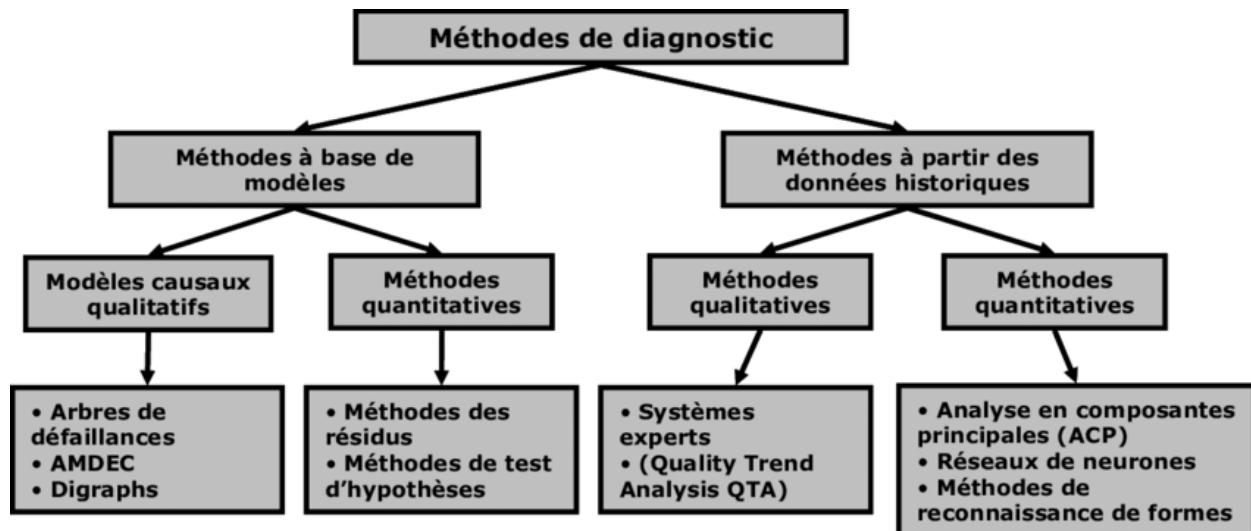


Figure I.1: Les méthodes de diagnostic

### I.4 Étapes de diagnostic:

- **Acquisition de données** : La procédure de diagnostic nécessite de disposer de l'information sur le fonctionnement du système à surveiller, les fonctions suivantes doivent être réalisées : Conditionnement et pré-traitement du signal.
- **Élaboration d'indicateurs de défauts** : à partir des mesures réalisées et des observations issues des opérateurs en charge de l'installation, il s'agit de construire des indicateurs permettant de mettre en évidence les éventuels défauts pouvant apparaître au sein du système. Dans le domaine de diagnostic, les indicateurs de défauts sont couramment dénommés les résidus aux symptômes. Un résidu représente un écart entre une grandeur estimées et mesurées. Cet écart de comportement doit donc être idéalement nul en l'absence de défaut et différent de zéro dans le cas contraire.

- **Détection** : C'est l'opération qui permet de décider si le système est en fonctionnement normal ou non. On pourrait penser qu'il suffit de tester la non nullité des résidus pour décider de l'apparition d'un défaut.
- **Localisation** : la localisation suit l'étape de détection, elle attribue le défaut à un sous-système particulier (capteur, actionneur, organe de commande, processus...).
- **Prise de décisions** : le fonctionnement incorrect du système étant constaté, il s'agit de décider de la marche à suivre afin de conserver les performances souhaitées du système sous surveillance. Cette prise de décision doit permettre de générer, éventuellement sous le contrôle d'un opérateur humain, les actions correctrices nécessaires à un retour à la normale du fonctionnement de l'installation.[1]

### Conclusion :

Dans ce chapitre nous avons vu que l'objectif du diagnostic est de détecter rapidement les divers défauts existant sur un procédé pour éviter la dégradation de ses performances et augmenter la sécurité des opérateurs, de l'environnement, de l'instrumentation et des biens.

**Chapitre II :**  
**l'analyse en**  
**composantes**  
**principales ACP**

## Introduction

L'analyse en composantes principales est une technique numérique bien éprouvée dans le domaine du traitement de données pour réduire la dimension de l'espace de représentation d'un système. L'ACP linéaire est un outil de modélisation des relations linéaires entre les différentes variables représentant le comportement d'un processus quelconque. Ainsi, l'ACP permet de mettre en évidence les relations qui existent entre les variables, sans tenir compte à priori, d'une quelconque structure. En effet, l'ACP élabore implicitement un modèle du système à partir de données prélevées sur ce dernier en fonctionnement normal. Cette technique descriptive peut être considérée comme une méthode d'identification de systèmes.

L'identification du modèle repose sur deux étapes : Estimer ses paramètres et Déterminer sa structure. Une fois le modèle ACP est identifié, des résidus peuvent être générés en comparant le comportement observé à celui donné par le modèle ACP de référence. Ces résidus permettent de détecter puis de localiser l'ensemble des variables en défaut.

## II.2. L'analyse en composantes principales

### II.2.1. Principe

L'analyse en composantes principales permet d'identifier les relations linéaires existantes entre les différentes variables du système. Les N mesures des m variables du processus représentant tous les modes de fonctionnement normal de ce dernier, sont collectées pour former une matrice des données X. Un prétraitement indispensable consiste à centrer et réduire les variables. Soit  $X \in R^{M \times N}$  la matrice de données prenant la forme :[3]

$$X = \begin{matrix} x1(1) & x2(1) & \dots & xm(1) \\ x1(2) & x2(2) & \dots & xm(2) \\ x1(N) & x2(N) & \dots & xm(N) \end{matrix}$$

Où,  $x1(1)$  représente la mesure de la première variable du premier instant de mesure.

Au préalable, afin de rendre le résultat indépendant des unités utilisées pour chaque variable, un prétraitement indispensable consiste à centrer et réduire les variables.

$$X = [X_1 \dots \dots \dots X_m] \quad (2.1)$$

La matrice de corrélation est donnée par :

$$\Sigma = \frac{1}{N-1} X X^T \quad (2.2)$$



L'estimation des paramètres du modèle ACP est très simple et revient en fait à un calcul des valeurs et vecteurs propres de la matrice de corrélation. Cependant le choix de la structure est plus délicat comme nous le verrons par la suite.[4]

### II.2.3 Modélisation des systèmes à partir de l'ACP

L'identification du modèle ACP consiste à estimer ses paramètres par une décomposition en valeurs/vecteurs propres de la matrice des corrélations  $\Sigma$ . Sa structure est fixée par un choix adéquat du nombre de composantes principales à retenir dans le modèle.

En analyse en composantes principales, nous désirons aboutir à un nouvel espace de dimensions plus faibles. Pour cela, nous cherchons un nouveau vecteur, notée  $\mathbf{t}$ , telle que  $\mathbf{t} \in \mathcal{R}^\ell$  où  $\ell$  représente la dimension de ce nouvel espace. Cette projection des variables sur le nouvel espace est conditionnée par l'hypothèse que l'erreur d'estimation soit minimale. [4]

Une relation linéaire doit donc être déterminée entre le vecteur des mesures  $\mathbf{x}$  et les composantes  $\mathbf{t}$ . Cette relation s'écrit ainsi :

$$\mathbf{t} = \mathbf{P}^T \mathbf{x} \text{ ou } \mathbf{P} \in \mathcal{R}^{m \times n}.$$

Grâce à cette transformation, les données sont projetées sur un espace de dimensions  $\ell$  de dimension plus faible ( $\ell < m$ ). Le premier vecteur propre  $\mathbf{P}_1$  de la matrice de projection  $\mathbf{P}$  est celui représentant la direction de l'espace des données qui interprète la plus forte concentration d'informations et est par la même occasion est la plus associé à la plus grande valeur propre  $\lambda_1$ . Réciproquement, le dernier vecteur propre  $\mathbf{P}_m$  de la matrice de projection  $\mathbf{P}$  indique la direction du vecteur propre montrant véhiculant de très faibles informations sur les données de la matrice  $\Sigma$ , ce vecteur propre sera associé à la plus petite valeur propre  $\lambda_m$ .

Par cette approche, on est en mesure de diminuer la dimension de la représentation des données en s'arrêtant au niveau du  $\ell^{\text{me}}$  terme  $\mathbf{t}_j \mathbf{p}_j$  avec ( $j = 1, \dots, \ell$ ).

Associés aux  $\lambda_j$  plus grandes valeurs propres. On aboutit donc à un nouveau vecteur qui n'est que l'estimation de  $\mathbf{x}$  par l'analyse en composantes principales obtenue par la formule ci- dessous:[4]

$$\hat{\mathbf{x}} = \sum_{j=1}^{\ell} \mathbf{t}_j \mathbf{p}_j = \sum_{j=1}^{\ell} (\mathbf{p}_j^T \mathbf{x}) \mathbf{p}_j \quad (2.3)$$

Nous aurons donc transformé le format des données initiales par le biais de  $\ell$  composantes principales  $t_1, t_2, \dots, t_\ell$  nécessaire à la description de ces données. Une partie de l'information est perdue dans les autres composantes principales. Cet "perte" d'informations est calculée par l'erreur  $e$  qui a pour valeur l'expression suivante :

$$e = x - \hat{x} = \sum_{i=\ell+1}^m t_i p_i \quad (2.4)$$

L'erreur d'estimation  $e$ , est évaluée à partir des dernières composantes principales où  $i = \ell + 1, \dots, m$  les dernières valeurs propres, correspondant aux faibles valeurs c'est-à-dire :  $\lambda_i$  avec  $i = \ell + 1, \dots, m$  sont associées à ces composantes principales.[6]

La somme de ces valeurs propres qui n'ont pas été retenue représente la perte d'informations sous la forme :[4]

$$\varepsilon\{ee^T\} = \sum_{i=\ell+1}^m \text{var}\{t_i\} \quad (2.5)$$

La reconstitution du vecteur initial de données  $x$  est donnée par l'expression :

$$x = \hat{x} + e \quad (2.6)$$

Les  $\ell$  premiers vecteurs propres seront représentés dans une matrice notée  $\hat{P}$ . Cette notation nous permet d'écrire :[7]

$$\hat{t} = \hat{P}^T x \quad (2.7)$$

Le vecteur test le vecteur des  $\ell$  premières composantes principales. En tenant compte que:

$$\hat{C} = \hat{P}\hat{P}^T \quad (2.8)$$

Nous déduisons l'estimation du vecteur  $x$  par :[8]

$$\hat{x} = \hat{C}x \quad (2.9)$$

#### II.2.4 Détermination de la structure du modèle ACP :

L'analyse en composantes principales recherche une approximation de la matrice initiale des données  $X$  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 principales.[9]

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 à 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. De plus il y a risque de non détection des défauts, si

certaines variables sont projetées dans le sous- espace des composantes principales.[10]

Il existe plusieurs critères permettant le calcul du nombre de composantes à retenir, la moyenne des valeurs propres, la validation croisée (PRESS), la variance non reconstruite (VNR), et le pourcentage cumulé de la variance (PCV).[10]

### II.2.5. Pourcentage cumulé de la variance totale (PCV) :

La base de cette méthode, notons que chaque composante principale est représentative d'une portion de la variance des mesures du processus étudié. Les valeurs propres de la matrice de corrélation sont les mesures de cette variance et peuvent donc être utilisées dans la sélection du nombre de composantes principales. Pour le choix de  $l$ , il faut choisir le pourcentage de la variance totale qu'on veut conserver. Le nombre de composantes est alors le plus petit nombre pris de telle sorte que ce pourcentage soit atteint ou dépassé ; les composantes sont choisies successivement dans l'ordre des variances décroissantes. Le pourcentage de variance expliquée par les premières composantes est donné par :[7]

$$PVC(l) = \left( \frac{\sum_{j=1}^l \lambda_j}{\sum_{j=1}^m \lambda_j} \right) \% \quad (2.10)$$

### II.2.6 Exemple de simulation :

Pour illustrer ce qui a été dit jusqu'à présent sur l'ACP linéaire, nous présentons ici un premier exemple de simulation. Cet exemple sera également utilisé pour illustrer les différentes démarches présentées dans ce premier chapitre.

Nous disposons de 07 variables représentant un système statique et qui sont décrites par les équations suivantes :

$$X1 = \cos(t)$$

$$X2 = \sin(t)$$

$$X3 = X1 + \epsilon_1$$

$$X4 = X1 + \epsilon_2$$

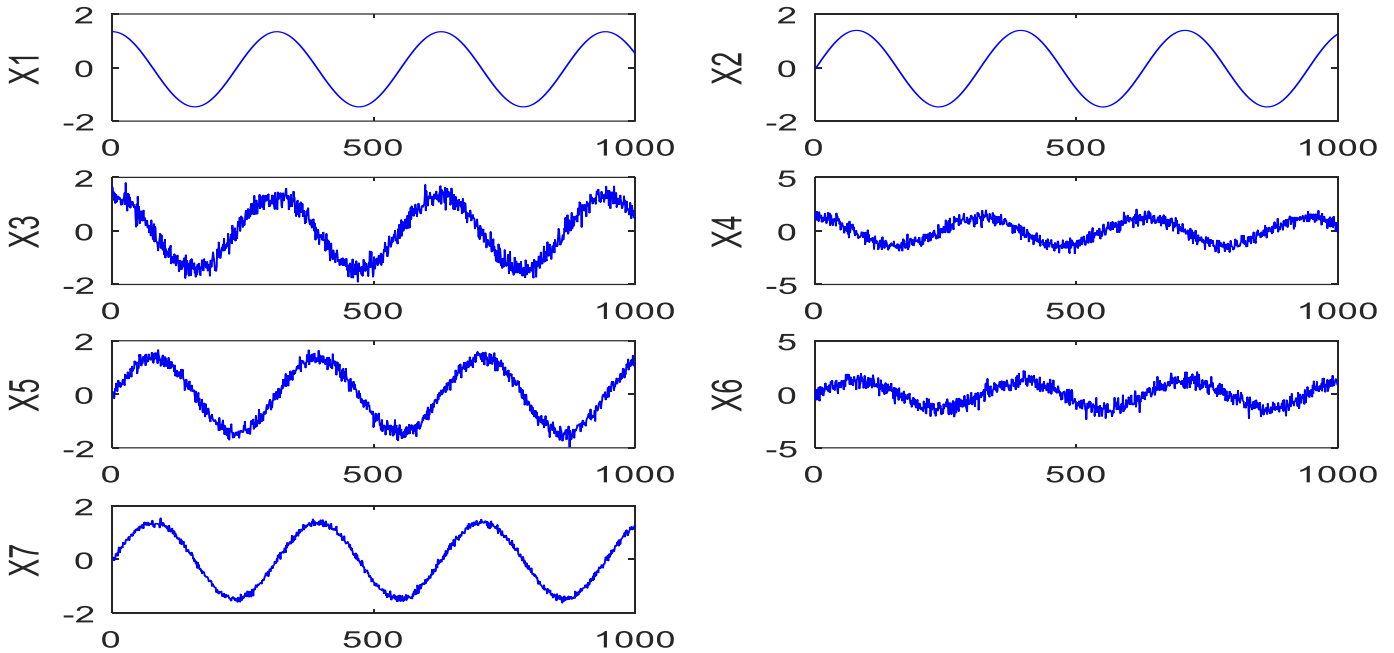
$$X5 = X2 + \epsilon_3$$

$$X6 = X2 + \epsilon_4$$

$$X7 = X2 + \epsilon_5$$

Où les  $\epsilon_i$  sont des bruits aléatoires gaussiens centrés et réduits. Les signaux  $x_1$  et  $x_2$  sont respectivement des signaux sinus et cosinus mesurés dans l'intervalle  $[0, 1000]$ . La matrice des mesures  $X = [x_1, x_2, \dots, x_7]$  est constituée de telle sorte que l'espace de données originelles

est de dimension  $m=7$ , dont chaque variable présente 1000 observations. L'évolution des différentes variables est représentée par la Figure (II.1).



**Figure II.1 :** Evolution des différentes variables de simulation

La matrice de covariance des données normalisées sera donc :

COV=

1	0.00212605	0.97638948	0.94547134	0.00645839	-0.0029132	0.00040677
0.00212605	1	-0.0003276	-0.0064405	0.98972007	0.91630542	0.99760645
0.97638948	-0.0003276	1	0.91997195	0.00299106	-0.0034334	-0.00239
0.94547134	-0.0064405	0.91997195	1	-0.0042727	-0.0054507	-0.0065858
0.00645839	0.98972007	0.00299106	-0.0042727	1	0.90903359	0.98774814
-0.0029132	0.91630542	-0.0034334	-0.0054507	0.90903359	1	0.91579362
0.00040677	0.99760645	-0.00239	-0.0065858	0.98774814	0.91579362	1

L'analyse en composantes principales cherche à identifier les vecteurs propres ortho-normaux et les valeurs propres des variables originelles. Ces vecteurs propres sont utilisés pour construire les composantes principales, et les valeurs propres sont les variances des composantes principales correspondantes.

Les matrices des valeurs et vecteurs propres sont données respectivement par :

$\Delta=$	3.8593973						
	2.89473027						
	0.12440043						
	0.08502997						
	0.0200766						
	0.01409086						
	0.00227457						
$P=$	-0.001695	0.58285156	0.03225885	0.22576669	0.77210981	0.11005878	-0.0018624
	0.50597115	0.00300077	0.26808834	-0.0211575	-0.0531777	0.34235621	0.74271768
	-0.0028721	0.57776634	0.02899537	0.5581052	-0.5910945	-0.0664763	-0.0066247
	-0.0054646	0.57134045	-0.0652819	-0.7945214	-0.1902966	-0.0416947	0.0079393
	0.50375376	0.00495947	0.30149933	-0.0076146	0.10077426	-0.7995262	-0.0764862
	0.48425258	0.00131207	-0.8719658	0.06547841	0.0204269	-0.0215654	-0.0018893
	0.50565185	0.00221001	0.2660062	-0.0386101	-0.0695728	0.47414613	-0.6651363

Maintenant, une fois que les directions du nouvel sous-espace orthonormé sont définies, on effectue un changement de base qui permet de projeter les variables inter-corrélées ( $x_1, x_2, \dots, x_7$ ) sur les directions P. Un autre nombre de variables réduites expliquent au mieux la variabilité des données originales est obtenu. Ces nouvelles variables sont appelées composantes principales  $t_i$ . Leur évolution est alors tracée dans la figure II.2.

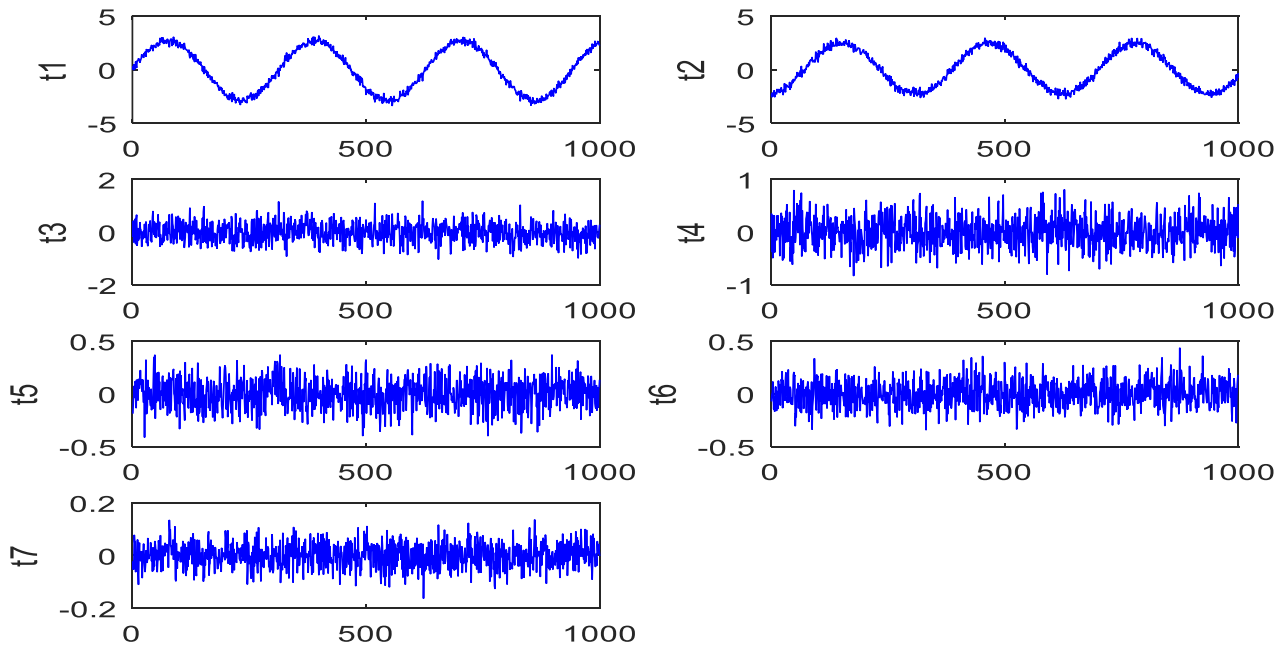


Figure II.2 : Evolution des composantes principales

Après avoir présenté la phase de diagonalisation de la matrice de corrélation, la détermination du nombre de composantes principales est une étape cruciale dans la méthode ACP, puisqu'elle permet d'identifier la structure optimale du modèle. La matrice qui caractérise ce modèle est donnée comme suit :

$$C = \begin{bmatrix} 0.33971881 & 0.00089136 & 0.33675688 & 0.33301593 & 0.00203675 & -5.61E-05 & 0.000431 \\ 0.00089136 & 0.25601581 & 0.00028055 & -0.0010505 & 0.25489976 & 0.24502178 & 0.25585188 \\ 0.33675688 & 0.00028055 & 0.3338222 & 0.33011698 & 0.00141859 & -0.0006327 & -0.0001754 \\ 0.33301593 & -0.0010505 & 0.33011698 & 0.32645977 & 8.07E-05 & -0.0018966 & -0.0015005 \\ 0.00203675 & 0.25489976 & 0.00141859 & 8.07E-05 & 0.25379245 & 0.24395057 & 0.25473498 \\ -5.61E-05 & 0.24502178 & -0.0006327 & -0.0018966 & 0.24395057 & 0.23450228 & 0.24486611 \\ 0.000431 & 0.25585188 & -0.0001754 & -0.0015005 & 0.25473498 & 0.24486611 & 0.25568867 \end{bmatrix}$$

Pour tester le modèle identifié, un jeu de données de validation de 1000 échantillons est utilisé. En fait, la matrice des observations  $X$  peut être donc estimée à partir des  $\ell$  composantes principales sélectionnées ( $\ell = 2$ ) correspondantes aux  $\ell$  plus grandes valeurs propres de la matrice de covariance. L'évolution des données et leur estimation est illustré par la Figure II.3.

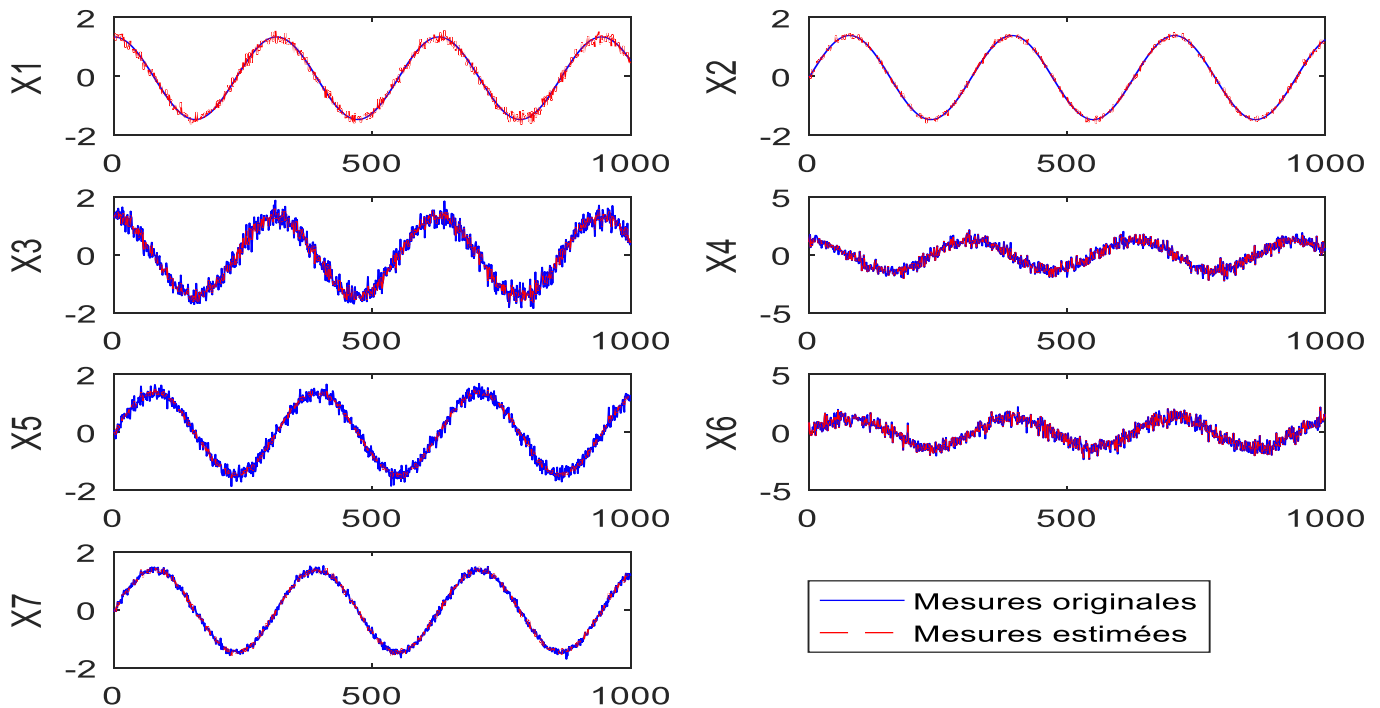


Figure II.3 : Evolution des différentes mesures et leurs estimations

### II.3. Détection et localisation des défauts

#### II.3.1. Indice *SPE*

La statistique *SPE*, également connue sous l'appellation Q (Squared Prediction Error) est une statistique obtenue à partir de l'erreur d'estimation des variables  $e(k)$ , elle nous renseigne donc sur l'évolution des mesures et de leur variation. Dès l'apparition d'une anomalie, la valeur de l'erreur d'estimation croît, entraînant par la même occasion une augmentation de l'indice *SPE*. Dans la détection de défauts par des tests statistiques, il faut avoir recours à des seuils de confiance. Les seuils de confiance représentent la tolérance de la statistique à une valeur du seuil suivant la qualité de la détection souhaitée.

Le critère *SPE* n'est autre que l'erreur quadratique de prédiction qui est égale à la somme des carrés des résidus.

A l'instant  $k$  nous avons :

$$SPE(K) = \sum_{i=1}^m (e_i(k))^2 \quad (2.11)$$

où  $e_i(k)$  est le  $i^{ieme}$  résidus qui donné par :

$$e_i(k) = x_i(k) - \widehat{x_i(k)} \quad (2.12)$$

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

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

Où  $\delta_\alpha^2$  est le seuil de détection du  $SPE(k)$ .

#### II.3.2. Reconstruction des variables

La méthode basée sur le principe de reconstruction repose sur l'hypothèse d'une possible défaillance d'un capteur, et propose de reconstruire la valeur de sa mesure en utilisant le modèle ACP préalablement calculé ainsi que les mesures des autres capteurs.

Quand une erreur est détectée, il est crucial de localiser la variable fautive et de mettre en œuvre les mesures correctives nécessaires pour éliminer les données inhabituelles. Nous pouvons trouver l'approche de reconstruction variable suggérée parmi les différentes techniques d'isolation basées sur la PCA. Cette méthode est efficace et est fréquemment utilisée pour ce sujet dans de nombreuses publications liées à l'sujet (Harkat et al., 2006) (Alcala and Qin, 2009).

Une comparaison des valeurs des variables avant et après reconstruction est effectuée dans la PCA conventionnelle pour isoler les failles. La méthode de reconstruction des variables suppose que chaque variable peut être incorrecte et suggère de reconstruire la variable supposée incorrecte à l'aide du modèle PCA de la variable restante

Cette variable reconstruite est ensuite utilisée pour séparer les défauts détectés. En d'autres termes, la reconstruction de la variable  $i$  utilise toutes les autres variables de données à l'exception de la variable  $i$ . Par conséquent, si seule cette variable présente un défaut, sa reconstruction élimine le défaut qui l'affecte.

$$G_j^T = [\varepsilon_1 \quad \varepsilon_2 \quad \dots \quad g_j \quad \dots \quad \varepsilon_m] \quad (2.13)$$

Avec

$$g_j^T = \frac{1}{1-c_{jj}} [c_{-j}^T \quad \mathbf{0} \quad c_{+j}^T] \quad (2.14)$$

$G_j$  permet de calculer le vecteur de mesures reconstruit  $x_j$  en utilisant le vecteur d'entrée  $x$  :

$$X_j^T = x^T G_j^T = [x_1 \quad \dots \quad z_j \quad \dots \quad x_m] \quad (2.15)$$

On voit ici, qu'on a reconstruit la valeur du capteur  $j$  en utilisant les valeurs des mesures des autres capteurs, et le modèle ACP déjà calculé. Alors le défaut dans ce capteur est éliminé lorsqu'on utilise l'indicateur SPE pour la détection de défaut.

### II.3.3. Simulation :

En reprenant le même exemple de simulation présenté dans la partie de modélisation. Et, afin d'observer l'impact de défaut sur la génération des résidus, ainsi que les composantes principales, un défaut est simulé aux données et précisément à la variable  $x_2$ . Ce défaut est représenté par un biais d'amplitude égale à 20% de la plage de variation de cette variable, à partir de l'instant 700. La figure suivante montre l'évolution de l'indice de détection. Ce dernier dépasse sa valeur de seuil, ce qui reflète la détection du défaut. Dans ce cas anormal du processus, le défaut est détecté par l'indicateur SPE représentant les variations des données dans l'espace résiduel.



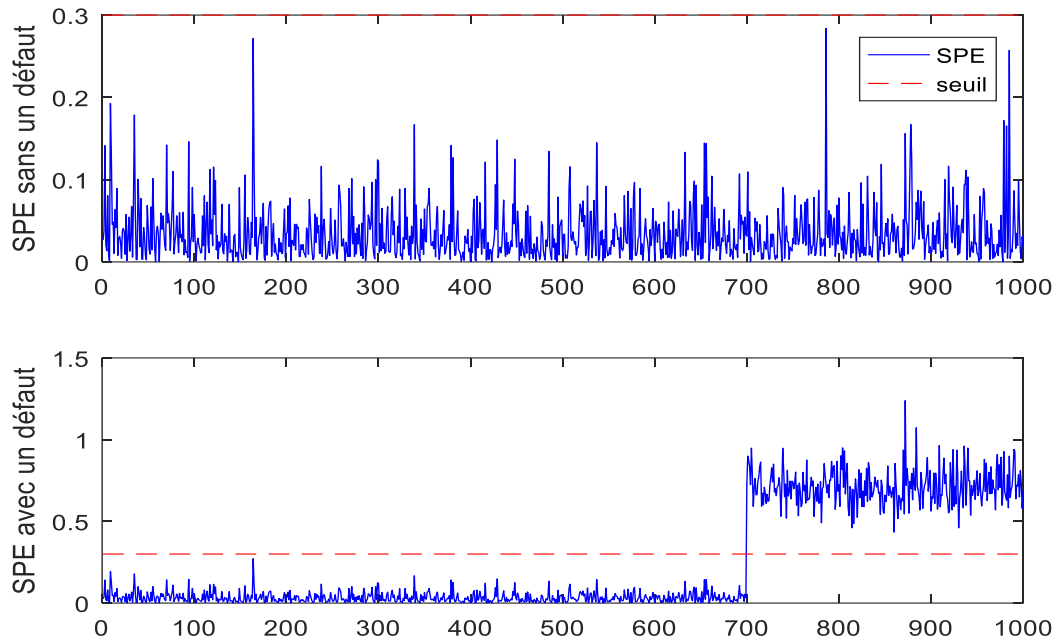


Figure II.4 : Evolution de SPE sans défaut et dans le cas de défaut

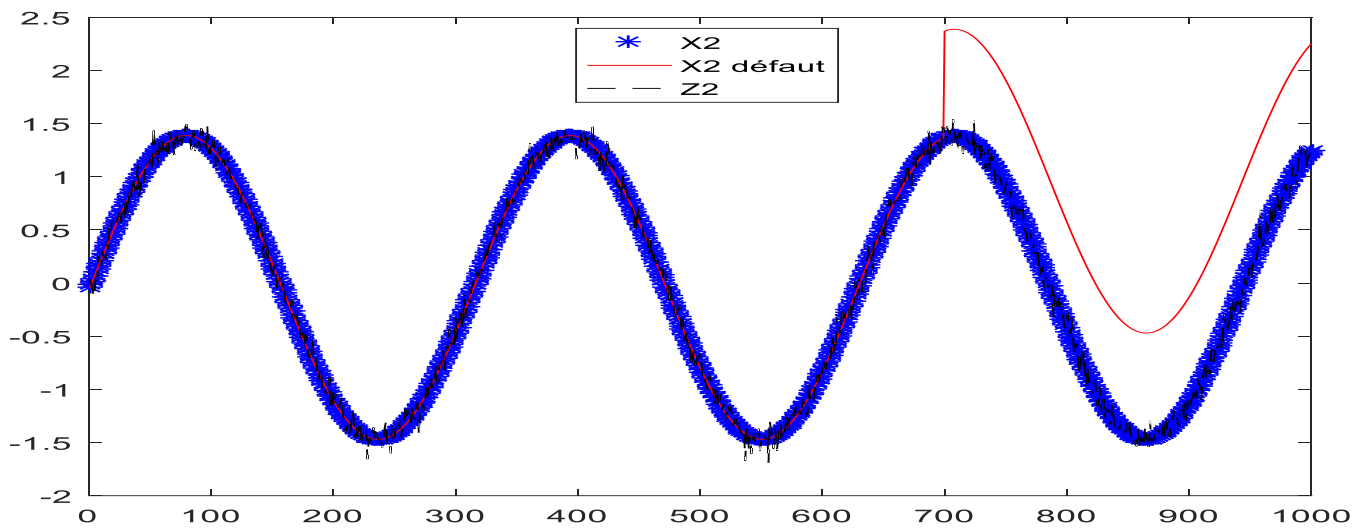
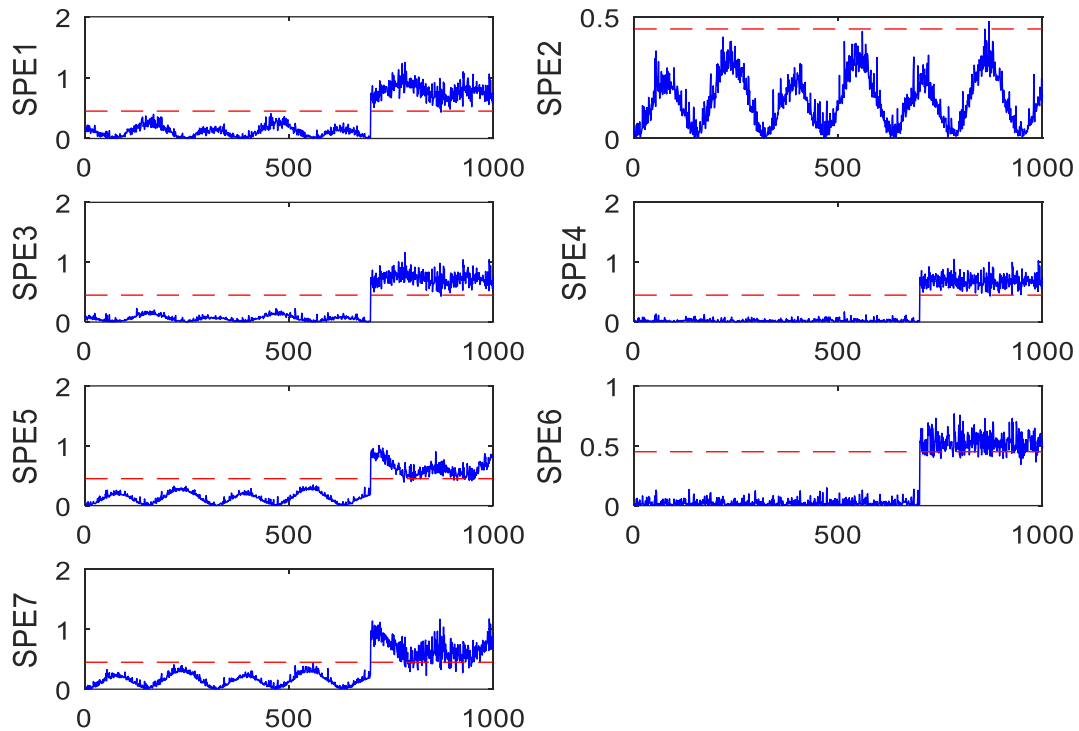


Figure II.5 la reconstruction Z2 de variables X2

Afin de localiser la variable en défaut l'approche de Reconstruction des variables à l'instant de détection ( $k = 700$ ) pour l'indice SPE est appliquée dont le résultat est illustré par la Figure (FigII.6)



**Figure II.6 :** Evolution des différents SPE obtenus par la méthode de Reconstruction avec un défaut affectant la variable  $x_2$  à partir de l'instant 700

### Conclusion :

Dans ce chapitre nous avons présenté les démarches de l'approche ACP pour modélisation des processus industriels en utilisant un historique de données d'entrées/sorties. Ensuite l'application du modèle identifié pour des raisons de diagnostic, i.e., détection, localisation et estimation des défauts.

**Chapitre III :**  
**l'analyse en**  
**composantes**  
**principales par**  
**intervalle**

## Introduction

Les résultats ACP, comme toute autre méthode d'analyse des données, ne sont précis qu'en tenant compte des incertitudes du modèle. Ces incertitudes permettent de prendre en compte les idées d'imprécision des variations des mesures causées par le bruit. Alors, le but de ce chapitre est de résoudre ce problème en proposant une stratégie de modélisation et de détection de défauts basée sur l'ACP intervalle, en particulier la méthode des sommets (Vertus PCA), qui permet la prise en considération des incertitudes de mesures du système pour plus de robustesse dans la détection de défauts.

L'analyse des composantes principales (ACP) a été largement utilisée pour superviser les processus. Il est nécessaire d'utiliser des techniques de surveillance des processus afin d'être sensible à tous les défauts potentiels du processus ou robuste face aux changements effectués dans le processus.

Cependant, car la plupart des véritables procédés industriels sont variables dans le temps et les données normales de fonctionnement peuvent être indiquées comme défectueuses par le modèle conventionnel de la méthode ACP, les méthodes classiques de l'ACP ne sont pas toujours fiables et les résultats fournis ne sont précis qu'en tenant compte des incertitudes du modèle. Le nombre excessif de fausses alarmes limite considérablement l'utilisation de l'ACP traditionnelle.[1]

### III.1. Données de type intervalle

La véritable valeur de la quantité est une notion. En pratique, la valeur véritable ne peut pas être mesurée et les informations recueillies sur un processus sont simplement des estimations données par les capteurs et sont donc inexactes. Cela est principalement dû aux incertitudes causées par des erreurs de mesure ou déterminées par des conditions expérimentales spécifiques.

Soit la matrice de data conventionnelle  $X \in R^{n \times m}$  contenant  $n$  échantillons de  $m$  variables de processus, où  $x_j(k)$ , est la  $k$ -ième observation de la  $j$ -ième variable. Afin de préserver l'information de la variable, il est plus approprié de représenter cette mesure par une valeur incertaine ou d'intervalle plutôt que par une valeur unique. L'incertitude étant inconnue, nous supposons que sa variation est limitée et peut être représentée par un intervalle de la forme  $[x_{j-}(k), \bar{x}_j(k)]$  où  $x_{j-}(k)$  et  $\bar{x}_j(k)$  désignent représenter les valeurs minimales et maximales enregistrées pour la variable à valeurs d'intervalle  $J$ -ième En ce qui concerne la  $k$ -ième unité d'observation. Pour prendre en compte les incertitudes du paramètre et fournir des limites strictes des mesures estimées, seuls les intervalles réels sont pris en compte. Par conséquent, l'observation d'intervalle réel du  $k$  pour la variable  $j$  est représenté par  $x_j(k)$  et la nouvelle matrix d'intervalles globaux  $[x_j]$  est formalisée comme une collection d'observations d'intervalles évaluées pour les différentes variables de

description selon ce qui suit :

$$[X] = \begin{pmatrix} x_{1\_}(1), \overline{x_{1\_}(1)} & \cdots & x_{m\_}(1), \overline{x_{m\_}(1)} \\ \vdots & \ddots & \vdots \\ x_{1\_}(n), \overline{x_{1\_}(n)} & \cdots & x_{m\_}(n), \overline{x_{m\_}(n)} \end{pmatrix}$$

Rendez  $\delta x_j(k)$  l'erreur d'une mesure, où  $x_j(k)$  est la quantité d'intérêt.  $\delta x_j(k)$ , est la différence entre la valeur réelle  $x_j(k)$  et la valeur rapportée en raison d'une mesure

$x_j^c(k)$  :

$$\delta x_j(k) = x_j^c(k) - x_j^r(k) \tag{3.1}$$

En conséquence, lorsqu'on définit un intervalle de données mesurées, l'erreur de mesure  $\delta x_j(k)$ , également connue sous le nom de  $x_j^r(k)$ , donne le radius de l'intervalle observation et le centre de l'intervalle est donné par l'erreur de mesure.  $x_j^c(k)$  Par conséquent, la formule conventionnelle de construction d'intervartion est donnée par :

$$[x_j(k)] = [x_j^c(k) - x_j^r(k), x_j^c(k) + x_j^r(k)] \tag{3.2}$$

Le couple (midpoint, radius) peut également exprimer l'intervalle  $[x_j(k)] = \{x_j^c(k), x_j^r(k)\}$ , dans la forme connue sous le nom de forme de point moyen.

**III.2. La normalisation des données intervalles :**

**III.2.1. Normalisation à l'aide de la dispersion des intervalles centraux et de la plage**

Les intervalles finis  $\{[x_j(1)], [x_j(2)], \dots, [x_j(k)]\}$  doivent être ensemble, de sorte que  $[x_j(k)] \subset R \forall k \in \{1, \dots, n\}$  et  $[m_j]$  est leur intervalle moyen correspondant. Ainsi  $\sum_{k=1}^n [(x_j^c(k) - m_j^c) + (x_j^r(k) - m_j^r)] = 0$  que  $\sum_{k=1}^n d^2([x_j(k)], [m_j])$  est réduit. Où les moyennes et le rayon de  $x_j(k)$  sont  $x_c(k)$  et  $x_r(k)$  et  $m_c, m_r$  sont leurs méthodes respectives La variance est définie comme la somme des distances carrées par rapport au moyen intervalle. En conséquence, la variance  $\sigma_j^2$  Voici les définitions pour l'intervalle de données valorisées :

$$\sigma^2 = \frac{1}{n} \sum_{k=1}^n d^2([x_j(k)], [m_j]) \tag{3.3}$$

On peut également écrire la définition de la variance selon la formule suivante :

$$\sigma^2 = \frac{1}{n} \sum_{k=1}^n |x_j^c(k) - m_j^c| + |(x_j^r(k) - m_j^r)|^2 \tag{3.4}$$

Et :

$$[m_j] = \left[ \frac{1}{n} \sum_{k=1}^n x_j(k) \quad \frac{1}{n} \sum_{k=1}^n \bar{x}_j(k) \right], m_j^c = \frac{1}{n} \sum_{k=1}^n x_j^c(k) \text{ and } m_j^r = \frac{1}{n} \sum_{k=1}^n x_j^r(k) \tag{3.5}$$

Alor :

$$\sigma^2 = \frac{1}{n} [\sum_{k=1}^n (x_j^c(k) - m_j^c)^2 + \sum_{k=1}^n (x_j^r(k) - m_j^r)^2 + 2 \sum_{k=1}^n |x_j^c(k) - m_j^c| |(x_j^r(k) - m_j^r)|] \quad (3.6)$$

La variance des données d'intervalle estimées peut être divisée en trois parties : la variance entre les points moyens, la variance entre les zones et la relation deux fois entre les points moyens et les zones, fournie par :

$$\sum_{k=1}^n |x_j^c(k) - m_j^c| |(x_j^r(k) - m_j^r)| \geq 0 \quad (3.7)$$

Il est possible de généraliser la distance entre les intervalles à la distance elliptique dans l'espace  $R^m$  Un intervalle normalisé est :

$$[\frac{1}{\sigma} (x_j^c(k) - m_j^c - |(x_j^r(k) - m_j^r)|), \frac{1}{\sigma} (x_j^c(k) - m_j^c + |(x_j^r(k) - m_j^r)|)] \quad (3.8)$$

### III.2.2 Normalisations par dispersion des centres d'intervalle :

La première méthode prend en compte la moyenne et la dispersion des intervalles centraux  $(x_{j-}(k) + \bar{x}_j(k)) \div 2$  et normalise de sorte que les points centraux transformés ont une moyenne et une variance uniforme pour chaque variable.

La valeur moyenne et la dispersion de tous les intervalles mi-point sont données par :

$$m_j = \frac{1}{n} \sum_{k=1}^n \frac{(x_{j-}(k) + \bar{x}_j(k))}{2} \text{ et } \sigma_j^2 = \frac{1}{n} \sum_{k=1}^n (\frac{(x_{j-}(k) + \bar{x}_j(k))}{2} - m_j)^2 \quad (3.9)$$

Avec cette notation, l'intervalle normalisé est défini avec des limites :

$$\frac{x_{j-}(k) - m_j}{\sigma_j}, \frac{\bar{x}_j(k) - m_j}{\sigma_j} \quad (3.10)$$

### III.2.3. Normalisation en utilisant la dispersion des intervalles de limites :

Pour chaque variable  $[X_j]$ , la deuxième méthode de normalisation transforme les n intervalles  $x_j(k)$  de manière à ce que la moyenne et la dispersion jointe des intervalles scalés sont respectivement 0 et 1. La dispersion jointe d'une variable  $[X_j]$  est définie par :

$$\sigma_j^2 = \frac{1}{n} \sum_{k=1}^n \frac{(x_{j-}(k) - m_j)^2 + (\bar{x}_j(k) - m_j)^2}{2} \quad (3.11)$$

Pour  $k = 1, \dots, n$ , les intervalles  $[x_j(k)] = [x_{j-}(k), \bar{x}_j(k)]$  sont modifiés en :

$$\frac{x_j(k) - m_j}{\sigma_j} \quad (3.12)$$

### III.2.4. Normalisation en utilisant la gamme mondiale :

La méthode de normalisation troisième transforme les intervalles  $[x_j(k)] = [x_{j-}(k), \bar{x}_j(k)]$ , ( $k = 1, \dots, n$ ) de manière à ce que la plage des intervalles recalés pour n soit l'intervalle unitaire [0, 1]. Les valeurs limites inférieures et supérieures doivent être  $Min_j = \min \{x_{j-}(k), \dots, x_{j-}(k)\}$  et

$Max_j = \{\bar{x}_j(k), \dots, \bar{x}_j(k)\}$ . Avec cette notation, l'intervalle est transformé en un intervalle normalisé avec des limite

$$\frac{\underline{x}_j(k) - Min_j}{Max_j - Min_j} \quad \frac{\bar{x}_j(k) - Min_j}{Max_j - Min_j}$$

**III.3. Analyse en composantes principales par intervalle**

**III.3.1. Méthode des sommets (Vertices PCA, VPCA) :**

La méthode des sommets (VPCA) est l'une des premières méthodes suggérées pour mettre en œuvre la PCA pour des données à valeurs d'intervalle (Cazes et al., 1997), (Chouakria, 1998). VPCA se déroule en deux étapes : Dans la première étape, l'ACP traditionnelle est utilisée pour la matrice sommet codée correspondant aux données intervalles, et dans la seconde étape, une gamme de chaque symbolique objet sur les notes est reconstruite afin de représenter la configuration des intervalles de données.

Selon VPCA, chaque objet  $H(k)$ , décrit par un intervalle évaluer  $[x(k)] = [x_{1-}(k), \bar{x}_1(k)] \dots [x_{m-}(k), \bar{x}_m(k)]$  dans un espace de dimension  $m$ , peut être représenté par une matrix de données numériques  $V(k)$  composée de  $2^m$  lignes et  $m$  colonnes, contenant les vertices des hypercubes associés  $H_k$ . Les verticaux sont fournis par toutes les combinaisons possibles entre les limites du vecteurs  $[x(k)]$ . Pour illustrer, si un intervalle vectoriel à deux dimensions  $[x(k)] = [[x_{1-}(k), \bar{x}_1(k)], [x_{2-}(k), \bar{x}_2(k)]]$  les vertices de la matrix  $V(k)$  sont construits de la manière suivante :[12]

$$V(k) = \begin{pmatrix} x_{1-}(k) & x_{2-}(k) \\ x_{1-}(k) & \bar{x}_2(k) \\ \bar{x}_1(k) & x_{2-}(k) \\ \bar{x}_1(k) & \bar{x}_2(k) \end{pmatrix}$$

Toute limite des observations évaluées par intervalle est une extrémité de l'hypercube correspondante. Par conséquent, une observation classique, c'est-à-dire qu'un point an un point, un segment a deux points, un rectangle a quatre points, etc. Les vertices matrices  $V(k)$  de la matrice VPCA globale peuvent ensuite être combinés pour tous les objets  $H(k)$  de la manière suivante :

$$V = \begin{pmatrix} V(1) \\ \vdots \\ V(k) \\ \vdots \\ V(n) \end{pmatrix} = \begin{pmatrix} \begin{bmatrix} \underline{x}_1(1) & \dots & \underline{x}_m(1) \\ \vdots & & \vdots \\ \bar{x}_1(1) & \dots & \bar{x}_m(1) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \underline{x}_1(k) & \dots & \underline{x}_m(k) \\ \vdots & & \vdots \\ \bar{x}_1(k) & \dots & \bar{x}_m(k) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \underline{x}_1(n) & \dots & \underline{x}_m(n) \\ \vdots & & \vdots \\ \bar{x}_1(n) & \dots & \bar{x}_m(n) \end{bmatrix} \end{pmatrix}$$

C'est pourquoi la PCA peut être effectuée sur les vertices de la matrix classique V selon la formule suivante :  $V = P \Lambda P^T$

Pour les vertices de la matrix V, les valeurs propres et les vectors propres de la covariance matrix  $\Sigma V$  sont calculés par  $\Lambda$  et P. Avant de calculer la covariance matrix V, (Cazes et al. (1997)) et (Chouakria et al. (2011)) suggèrent l'application d'un schéma de poids sur les vertices de la matrix V en fonction de la nature de l'application en cours.

Trois systèmes de poids symbolique principaux sont proposés, qui sont présentés ci-dessous. Pour commencer, nous pouvons utiliser  $w_j(k)$  pour représenter le poids d'observation  $[x_j(k)]$ , tandis que  $w_j^l(k)$ ,  $l = 1, \dots, 2^m$  est le poids du point l Le plan de mesure doit répondre à :

$$w_j(k) = \sum_{l=1}^{2^m} w_j^l(k), \sum_{j=1}^m w_j(k) = 1 \tag{3.13}$$

Une sélection de poids initiale et courante  $w_j(k)$  donne une valeur égale à chaque observation, comme dans :  $w_j(k) = 1/m, j = 1, \dots, m$

Les différences internes des hypercubes sont soulignées par une seconde option de poids :

$$w_j(k) = \frac{v_j(k)}{\sum_{j=1}^m v_j(k)} \tag{3.14}$$

où  $v_j(k)$  est le volume de l'hypercube associé au vector d'observation  $[x(k)]$  calculé comme suit

$$w_j(k) = \frac{\frac{1-v_j(k)}{\sum_{j=1}^m v_j(k)}}{\sum_{j=1}^m \frac{1-v_j(k)}{\sum_{j=1}^m v_j(k)}} \tag{3.15}$$



Par conséquent, la matrice de poids  $D$  est la diagonale de poids  $n \times n$   $w_j(k)$  liée aux vertices de la matrice  $V$ . La matrice de covariance pondérée des vertices est ensuite fournie par :  $\Sigma_V = VDV^T$ . Ensuite, une PCA classique est effectuée sur les vertices de la matrice  $V$  en utilisant la matrice de covariance  $\Sigma_V$ . Les vertices de la matrice  $[x(k)]$  doivent être  $V(k)$ , et les principaux composants de la matrice  $T(k) = [t_1(k), t]$  doivent être fournis par :

$$T(k) = V(k)P$$

Les eigen-vectors de la matrice de covariance  $\Sigma_V$  sont  $P = [p_1, \dots, p_m]$ . (Cazes et al., 1997) ont proposé une méthode de calcul pour les composants évalués à intervalle  $[t_1(k)], \dots, [t_m(k)]$  en utilisant l'algorithme Moore, à partir des composants verticaux principaux de la matrice, qui se présentent de la manière suivante :

$$\begin{cases} \underline{t}_j(k) = \min (T_j(k)) \\ \bar{t}_j(k) = \max (T_j(k)) \end{cases} \quad (3.16)$$

la matrice de sommets estimée  $\widehat{V}_j(k)$  donnée par :  $\widehat{V}_j(k) = V(k)P_l P_l^T = V(k)C_l$

Les mesures estimées par intervalle pour les premiers composants principaux sont fournies par :

$$\begin{cases} \underline{\hat{x}}_j(k) = \min (\widehat{V}_j(k)) \\ \bar{\hat{x}}_j(k) = \max (\widehat{V}_j(k)) \end{cases} \quad (3.17)$$

### III.5.2 SPE à valeur intervalle :

Cette statistique est calculée sur des résultats comme dans le cas classique, ce qui donne un intervalle [SPE] avec une limite supérieure et une limite inférieure, correspondant respectivement aux limites supérieures et inférieures des résultats calculés, comme suit :

$$\begin{cases} \underline{SPE}(k) = \|\underline{e}(k)\|^2 = \underline{e}(k)^T \underline{e}(k) \\ \overline{SPE}(k) = \|\bar{e}(k)\|^2 = \bar{e}(k)^T \bar{e}(k) \end{cases} \quad (3.18)$$

### III.4 Reconstruction des variables pour le modèle (VPCA):

Cette section, qui est une extension de cette technique pour le modèle PCA à valeurs d'intervalle, est utilisée pour isoler les défauts et reconstituer les mesures erronées. La matrice de projection  $G(i)$  est utilisée pour obtenir les reconstructions des modèles d'intervalles PCA différents. Cette matrice fournit des projections reconstruites sur les sous-espaces principaux ou résiduels en fonction du nombre de composants choisis.

Une PCA classique est effectuée sur les vertices de la matrice  $V$  pour VPCA. Nous pouvons considérer un intervalle vectoriel  $[x(k)]$  égal à  $[[x_1(k)], [x_2(k)], \dots, [x_m(k)]]$ , et la matrice verte correspondante de deux lignes et de deux colonnes de  $m$ . La matrice reconstruite  $V^{(i)}(k)$  sera exécutée sur les vertices de la matrice  $V(k)$  de la manière traditionnelle :  $V^{(i)}(k) = G^{(i)}V(k)$

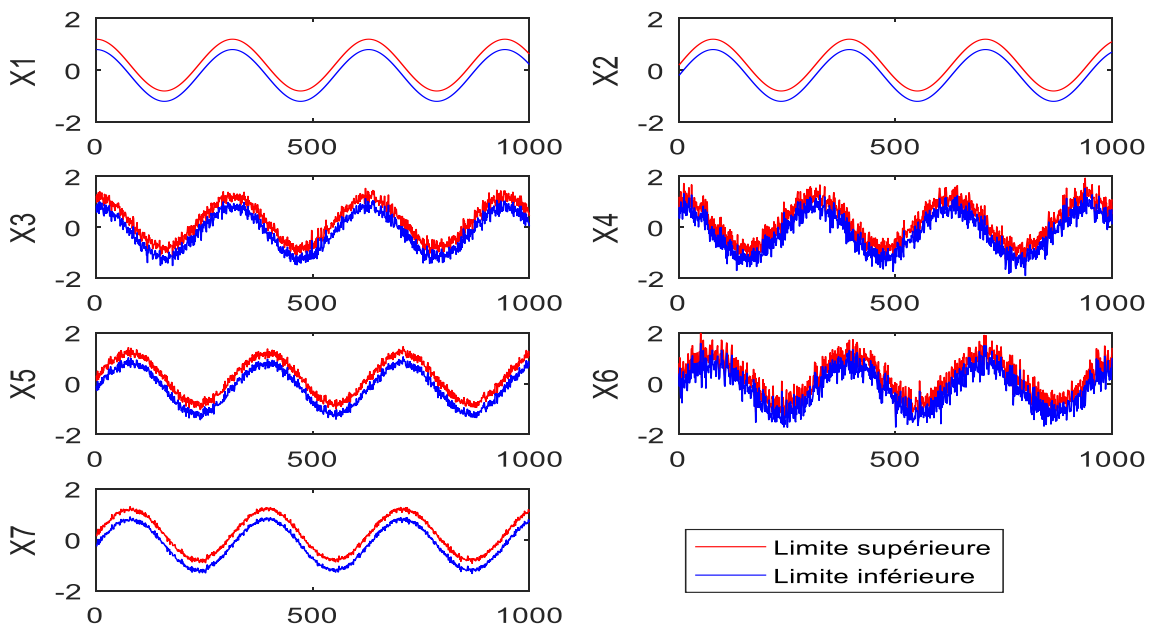
$$\begin{aligned}
 V(k) &= [V_1(k) \dots V_i(k) \dots V_m(k)] \\
 &= \begin{bmatrix} \underline{x}_1(k) & \dots & \underline{x}_i(k) & \dots & \underline{x}_m(k) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \bar{x}_1(k) & \cdot & \bar{x}_i(k) & \cdot & \bar{x}_m(k) \end{bmatrix}
 \end{aligned} \tag{3.19}$$

Puis, l'intervalle de reconstruction  $[z_i(k)] = [z_i(k), \bar{z}_i(k)]$  d'un intervalle de mesure  $[x_i(k)]$  est donné par :

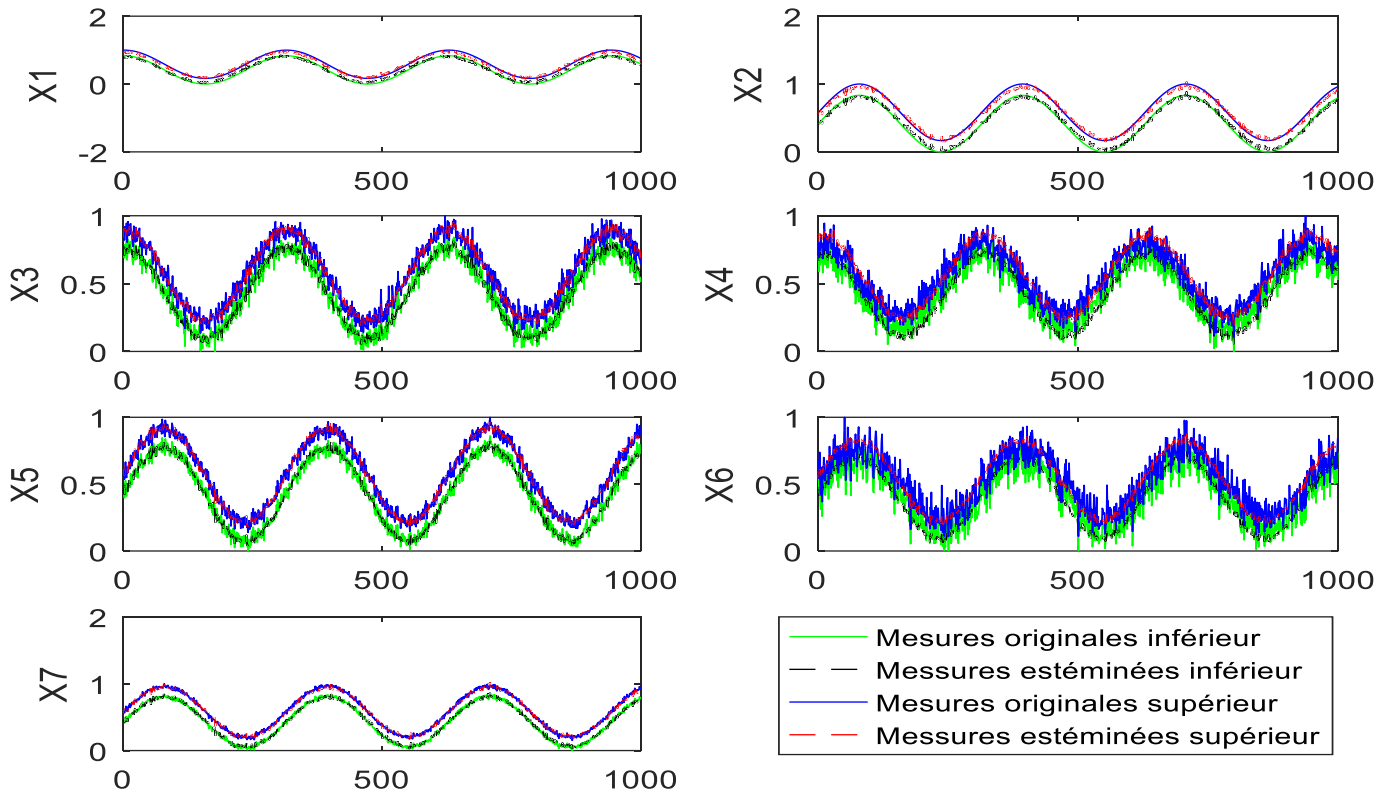
$$\begin{cases} \underline{z}_i = \min V_i^{(i)}(k) \\ \bar{z}_i(k) = \max V_i^{(i)}(k) \end{cases} \tag{3.20}$$

**Simulation :**

Cet exemple simule le système sur 1000 échantillons avec un intervalle de 20% de la plage de variation de chacune des 7 variables :



**Figure III.1 :** Evolution des différentes variables en type intervalle



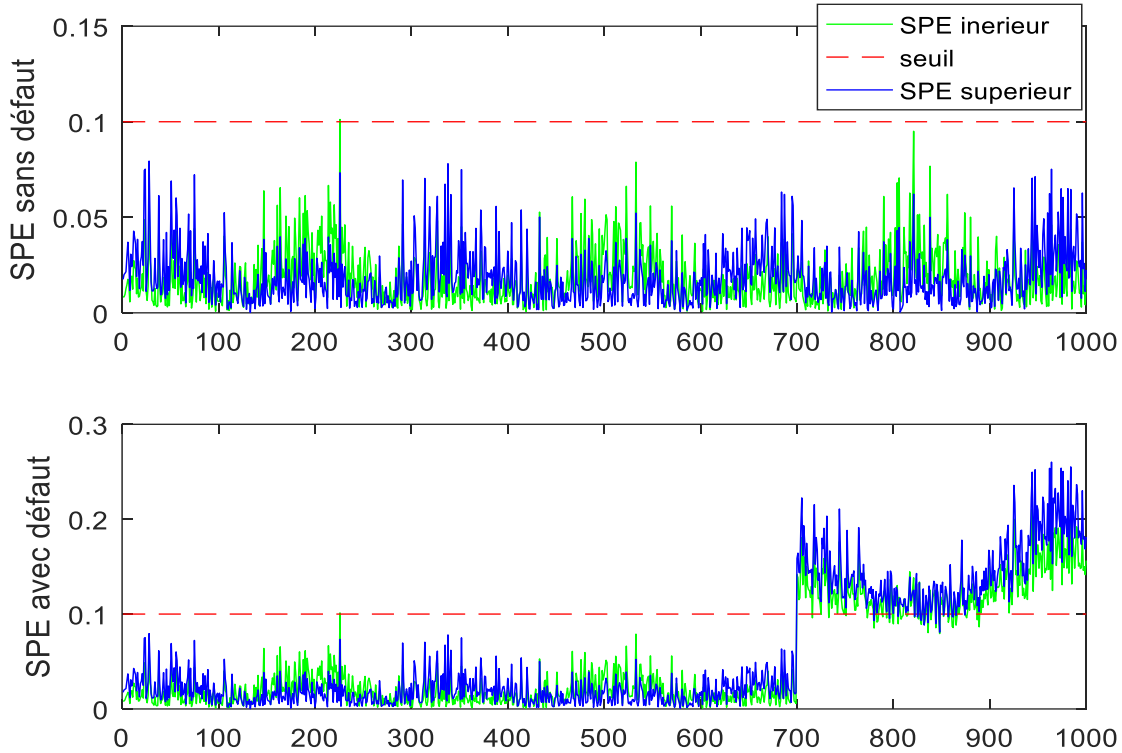
**Figure III.2 :**Evolution des différentes mesures et leurs estimation

L'approche de la détection de défauts basés sur la PCA consiste à analyser les restes de manière distincte. Pour commencer, nous devons définir l'erreur d'estimation par intervalles PCA. Les résidus  $\underline{e}(k)$  et  $\bar{e}(k)$  peuvent être obtenus grâce à la différence entre les estimations d'intervallées initiales des variables, comme :

$$\begin{cases} \underline{e}(k) = \underline{\mathbf{x}}(k) - \bar{\hat{\mathbf{x}}}(k) \\ \bar{e}(k) = \bar{\mathbf{x}}(k) - \underline{\hat{\mathbf{x}}}(k) \end{cases} \quad (3.21)$$

Des événements inhabituels sont projetés sur l'espace résiduel dans l'ACP classique et peuvent être détectés à l'aide de plusieurs statistiques. L'ACP à valeurs d'intervalle suit le même principe. Cependant, une caractéristique intrigante des résidus à valeurs d'intervalle permet une détection précise des défauts dans les cas univariés.

L'offset ou défaut a été ajouté à la variable  $x_2(k)$  dans la plage de temps de 700 à 1000, et il a été simulé comme un biais constant d'ampleur égal à 40% de la plage de variation de la variable  $x_2(k)$ . Les résultats de détection et d'isolation de défauts sont présentés dans les figures suivantes.



**Figure III.3 :** Evolution de SPE sans défaut et dans le cas de défaut pour les mesure en type intervalles

### III.5.3. Localisation et isolation des failles à l'aide de PCA pour des données à intervalles de valeur :

**.Simulation :**

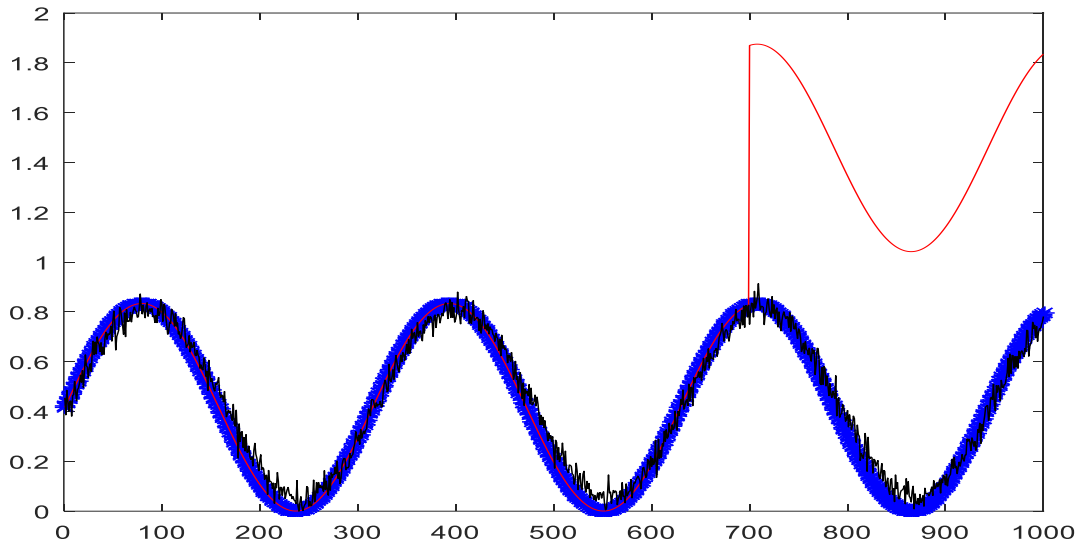


Figure III.4 :Le temps d'évolution de la reconstruction de variables X2

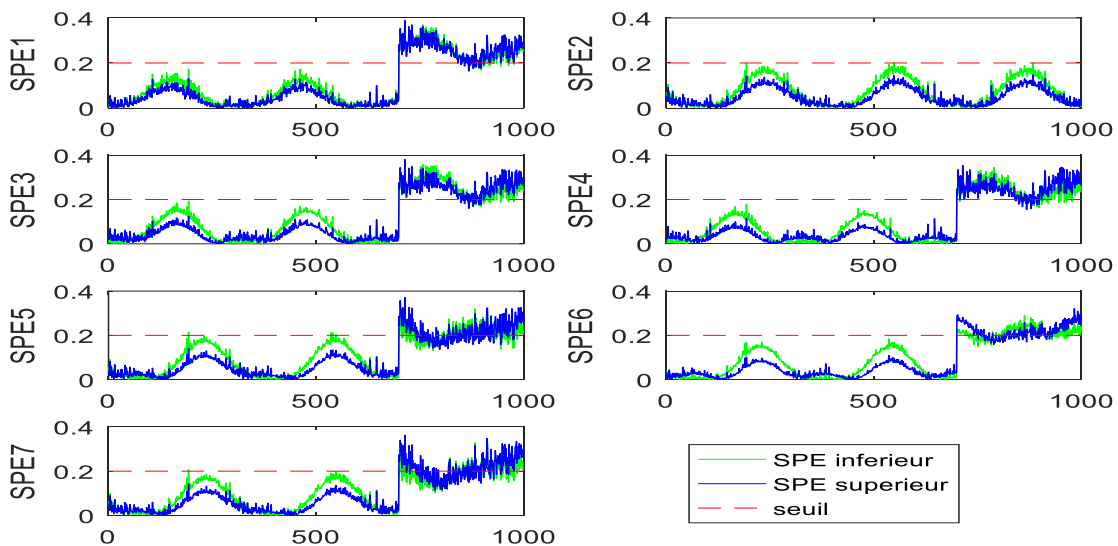


Figure III.5: Evolution de SPE calculé après la reconstruction de différentes variables avec défaut --- par la méthode d'élimination avec un défaut affectant la variable x2 à partir de l'instant 700

**Conclusion :**

Dans ce chapitre, nous présentons les méthodes de modélisation des données imprécises ou à caractère incertain à l'aide de l'analyse en composantes principales (ACP) intervalle. Une fois que

le modèle est identifié, nous pouvons procéder à la détection et à la localisation des défauts en générant des indicateurs spécifiques de défaut. Nous présentons également une extension de l'approche de reconstruction adaptée aux données de type intervalle, que nous avons utilisée avec succès pour localiser les défauts des capteurs incertains.

# **Chapitre IV : Application sur la station météorologique automatique**

**Introduction :**

La météorologie est une science qui a pour objet l'étude des phénomènes atmosphériques tels que les nuages, les précipitations ou le vent dans le but de comprendre comment ils se forment et évoluent en fonction des paramètres mesurés tels que la pression, la température et l'humidité. Le mot vient du grec ancien μετέωρος / metēōros (« qui est au-dessus de la terre »), qui désigne les particules en suspension dans l'atmosphère et -λογία / -logia, « discours » ou « connaissance ».

C'est une discipline qui traite principalement de la mécanique des fluides et de la thermodynamique mais qui fait usage de différentes autres branches de la physique, de la chimie et des mathématiques. Purement descriptive à l'origine, la météorologie est devenue un lieu d'application de ces disciplines. Pour ce faire elle doit s'appuyer sur un réseau cohérent d'observations : le premier du genre — qui concerne un territoire multinational étendu — apparaît en 1854, sous la direction du français Le Verrier qui établit un réseau européen de données atmosphériques et fonctionne de manière opérationnelle dès 1856.

La météorologie moderne permet d'établir des prévisions de l'évolution du temps en s'appuyant sur des modèles mathématiques à court comme à long terme qui assimilent des données de nombreuses sources dont les stations, les satellites et les radars météorologiques. La météorologie a des applications dans des domaines très divers comme les besoins militaires, la production d'énergie, les transports (aériens, maritimes et terrestres), l'agriculture, la médecine, la construction, la photographie aérienne ou le cinéma. Elle est également appliquée pour la prévision de la qualité de l'air ou de plusieurs risques naturels d'origine atmosphérique.

**IV.1 la station automatique météorologique :**

la station automatique météorologique est une ensemble des capteurs destiné de mesurer différentes paramètre météorologique (10 capteur) ,situées à moins de 70 mètres de la station météorologique, constitué d'une tige métallique de 10 mètres positionnée sur son sommet un capteur de vent en direction nord (direction et vitesse), abri à une hauteur de 2 mètres porte trois capteurs de température et humidité et un autre pour le rayonnement direct, capteur de température au sol, deux capteur de température dans le soul , le Premier à 20 cm de profondeur et le deuxième à 50 cm de profondeur, un autre capteur de rayonnement (diffus) au nord et pluviomètre au sud , et un capteur de pression au niveau de la station d'observation, toutes les informations mesurées par les capteur sont collectées au niveau du data logger et ensuite envoyées à l'ordinateur de traitement par un dispositif (moxa) qui les convertit des signaux en informations numériques.





**Figure IV.1 :**la station automatique météorologique

## **IV.2 Système météorologique intégré :**

### **IV.2.1 WindSonic:**



**Figure IV.2 :** WindSonic

Fabriqué par Gill Instruments, le WindSonic est un capteur à ultrasons pour la mesure de la vitesse et de la direction du vent. La construction du capteur est robuste, fiable et il a de faible consommation.

Sans pièce en mouvement, le WindSonic fournit la vitesse et la direction du vent avec de nombreuses options de sorties (numériques, analogiques) couvrant les besoins classiques des utilisateurs.[2]

### **IV.2.2 PT-100 Thermomètre**



**Figure IV.3:** PT-100 Thermomètre

La sonde de température désignée pour mesurer la température de l'air, du sol de surface et du sol. Elle peut être également utilisée pour mesurer la température de la mer avec un câble spécial désigné pour les applications maritimes.

#### IV.2.3 Capteur de l'humidité relative et de température RHT175



**Figure IV.4 :** Capteur de l'humidité relative et de température RHT175

MicroStep-MIS RHT175 Capteur de l'humidité à mesurer l'humidité relative de l'air et la température de l'air dans des applications météorologiques et environnementales. RHT175 est une sonde contrôlée par un microprocesseur à très basse consommation de puissance, convenu aux applications alimentées par énergie solaire ou batterie.

RHT175 fournit trois types de sortie de valeur mesurée :

- SDI-12
- interface en série
- sortie analogique

#### IV.2.4 Capteur de pression MSB780



**Figure IV.5 :** Capture de pression MSB780

MicroStep-MIS MSB780 Baromètre digital est désigné pour mesurer la pression absolue atmosphérique dans des applications météorologiques et aériennes professionnelles. MSB780 fourni une précision excellente, une stabilité à longue durée et la dépendance à la température.

MSB780X est une version élargie de MSB780. Il comporte le transducteur de pression de la même qualité premium que MSB780, mais permet aussi d'utiliser deux et trois transducteurs de pression dans une seule unité.

#### IV.2.5 MR3 et MR3H Pluviomètre

Les pluviomètres MR3 et MR3H sont destinés à la mesure de précipitation liquide (MR3) et de la mesure de précipitation liquide et solide (MR3H).



**Figure IV.6:** MR3 et MR3H Pluviomètre

#### IV.2.6 CMP10 (CMP11) Pyranomètre

La mesure du rayonnement solaire global au-dessus d'une surface plane se fait grâce aux pyranomètres, dans la gamme de longueurs d'onde allant de 285 nanomètres (nm), ou moins, à 2800 nm.

Le CMP10 est le pyranomètre étalon secondaire avec le meilleur rapport qualité-prix-performance du marché. Basé sur la technologie éprouvée du CMP

11, le CMP10 étend cette qualité pour les applications où l'entretien est difficile.[2]



**Figure IV.7 :** CMP10 (CMP11) Pyranomètre

#### **IV.2.7 Arceau d’Ombrage CM121B/C**

L’anneau d’ombrage CM121B/C peut être utilisé avec des pyranomètres, pyrgéomètres et radiomètres UV de Kipp & Zonen pour protéger les instruments contre le rayonnement direct.

L’ombre de l’anneau recouvre complètement le dôme du pyranomètre. L’arceau ne nécessite aucun réglage pendant plusieurs jours. Naturellement, l’arceau intercepte également une petite partie du rayonnement diffus venant du ciel. Une correction est nécessaire pour compenser ce phénomène ; un tableau est donc fourni pour cela.[2]



**Figure IV.8:** Arceau d’Ombrage CM121B/C

#### **IV.3 L’intérêt de l’approche diagnostic dans les station météorologie:**

Dans le domaine de la météorologie, le diagnostic des défauts ou l’identification et la résolution des problèmes dans les systèmes de prévision météorologique sont des tâches essentielles. Des prévisions météorologiques précises sont indispensables dans de nombreux secteurs tels que l’agriculture, l’aviation, la gestion des catastrophes et la planification quotidienne. Le diagnostic

des défauts joue un rôle crucial dans le maintien de la fiabilité et de l'efficacité des systèmes météorologiques.

Les systèmes météorologiques sont composés d'un réseau complexe d'instruments, de capteurs, de satellites et de modèles informatiques qui collectent et analysent des données pour générer des prévisions météorologiques. Cependant, ces systèmes ne sont pas à l'abri de pannes ou d'erreurs pouvant affecter la qualité et la précision des prévisions. Le processus de diagnostic consiste à identifier, analyser et corriger ces défauts afin d'assurer la fourniture d'informations météorologiques fiables et précises.

Un domaine courant de diagnostic des défauts en météorologie concerne le réseau d'observation. Les stations météorologiques, les bouées météorologiques, les radars et les satellites fournissent des données essentielles telles que la température, l'humidité, la pression, le vent et les précipitations. Toute défaillance ou incohérence dans ces observations peut conduire à des prévisions erronées. Des techniques de diagnostic sont utilisées pour détecter les défaillances des capteurs, les problèmes de calibration, les erreurs de transmission des données ou toute anomalie pouvant affecter l'intégrité du réseau d'observation.

Un autre aspect du diagnostic des défauts en météorologie concerne les modèles de prévision numérique du temps (PNT). Les modèles de PNT utilisent des équations mathématiques et des techniques d'assimilation des données pour simuler et prédire les conditions atmosphériques. Cependant, ces modèles peuvent rencontrer des défauts dus à des erreurs dans les données d'entrée, des limitations de calcul ou des lacunes dans le modèle lui-même. Des méthodes de diagnostic sont utilisées pour identifier et corriger ces défauts, garantissant ainsi l'exactitude et la fiabilité des modèles de prévision.

De plus, le diagnostic des défauts en météorologie s'étend aux procédures de traitement et d'assimilation des données. Les données météorologiques brutes collectées à partir de différentes sources sont soumises à un prétraitement et à une assimilation pour créer une image complète de l'atmosphère. Les erreurs ou les défauts introduits lors de ces étapes peuvent avoir un impact significatif sur la qualité des prévisions. Des techniques de diagnostic sont utilisées pour surveiller la qualité des données, identifier les valeurs aberrantes, détecter les incohérences et corriger les algorithmes d'assimilation des données afin d'améliorer la fiabilité globale des résultats des prévisions.

En conclusion, le diagnostic des défauts dans le domaine de la météorologie est un processus essentiel pour maintenir l'exactitude et la fiabilité des systèmes de prévision météorologique. En détectant et en rectifiant les défauts dans le réseau d'observation, les modèles numériques et les procédures de traitement des données, les météorologues peuvent garantir que les prévisions fournies à diverses industries et au grand public sont aussi précises et fiables que possible. Les avancées continues dans les techniques de diagnostic continuent d'améliorer l'efficacité des systèmes météorologiques, contribuant ainsi à de meilleures prévisions météorologiques et à une prise de décision améliorée dans de nombreux secteurs.

#### IV.4 Modélisation par ACP :

Dans cette partie, Nous appliquerons la méthode ACP pour la détection et la localisation de défauts sur la station automatique, seulement 6 variables ont été considérées. La matrice des données contient donc 6 variables, ces variables sont représentées dans le Figure suivant

Variables	Capteurs
X1	Vitesse de vent(max)
X2	Température
X3	Rayonnement global
X4	Humidité
X5	Direction de vent
X6	Vitesse de vent

**Figure IV.9 :** La description des variables de la station

Pour mieux suivre ces variables, Cette figures représente l'évolution des différentes mesures sur une plage de fonctionnement  $t = 10000$  minutes (01 semaine).

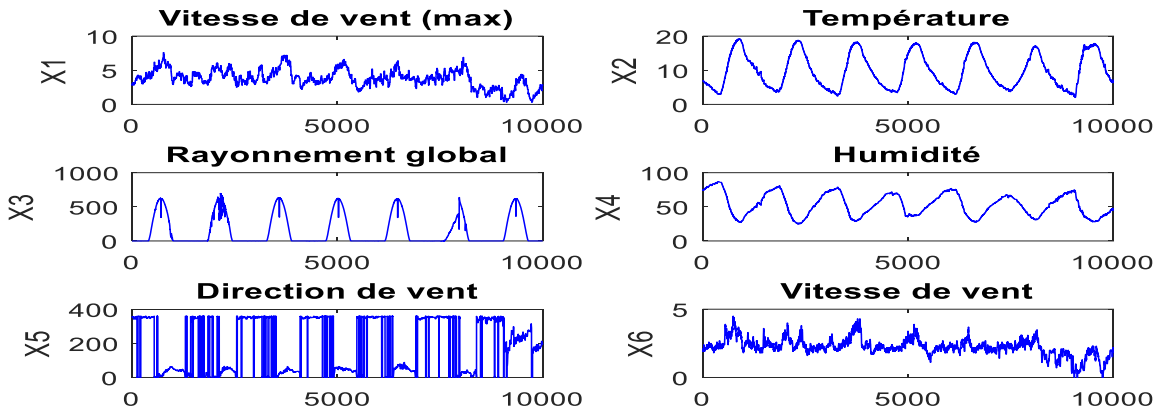


Figure IV.10 : – L'évolution des différentes mesures

Après l'application de modèle ACP en remarque sur la figure que les estimation conforme aux les données initiales ainsi que les erreurs d'estimation de toutes les variables sont presque nulles, Ceci montre que le modèle ACP donne une bonne estimation sur les différentes mesures.

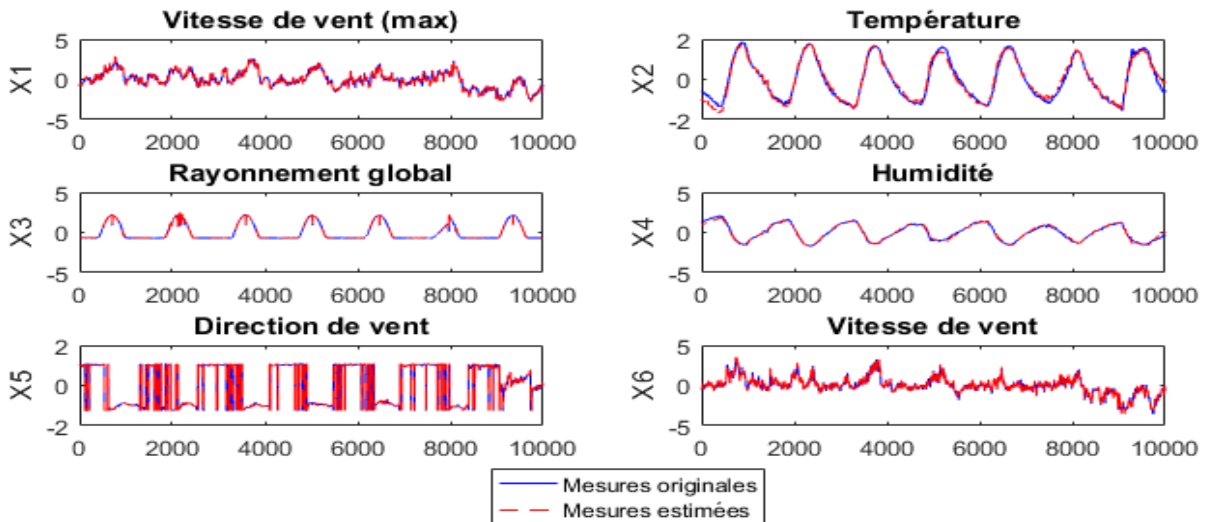


Figure IV.11 : Evolution des différentes mesures et leurs estimations

IV.4.1 Détection et localisation de défaut :

Une fois que le modèle ACP est obtenu, nous simulons un défaut sur la variable x2 avec un biais qui représente 40% de sa plage de variation, à partir de l’instant 7000. Nous calculons l’indice de détection de défauts SPE. La figure (fig,et ) représente l’évolution de cet indice en absence et avec défaut.

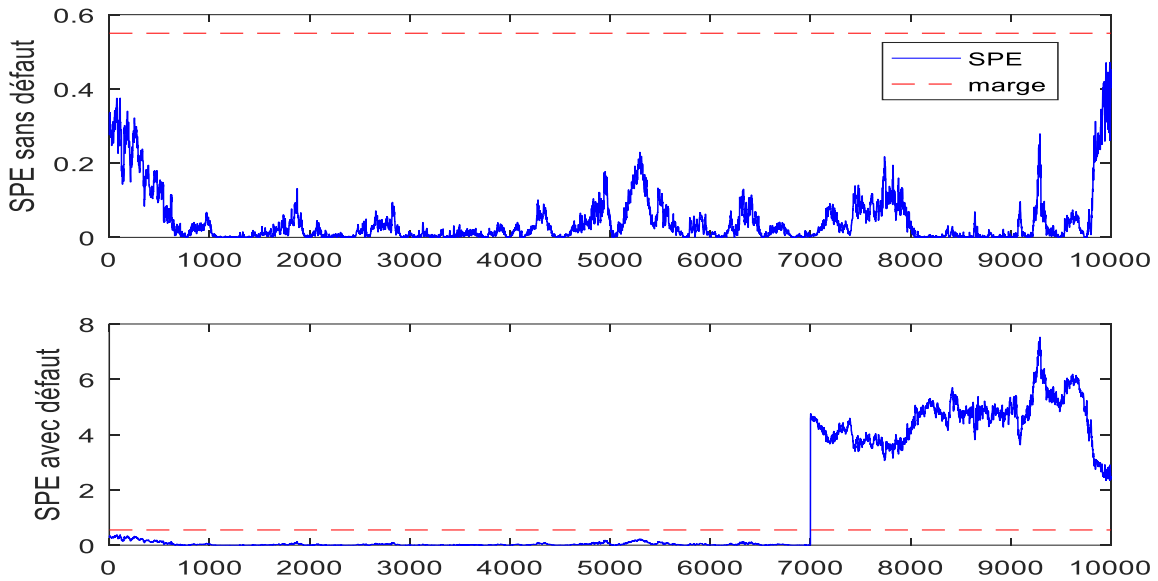


Figure IV.12 : L’évolution de l’indices SPE

Une fois la défaillance est bien détectée par un indice de détections, une étape d’isolation de ce défaut intervient pour savoir la provenance de ce défaut. Dans notre cas nous avons utilisé la méthode de reconstruction des variables pour la localisation de défaut.

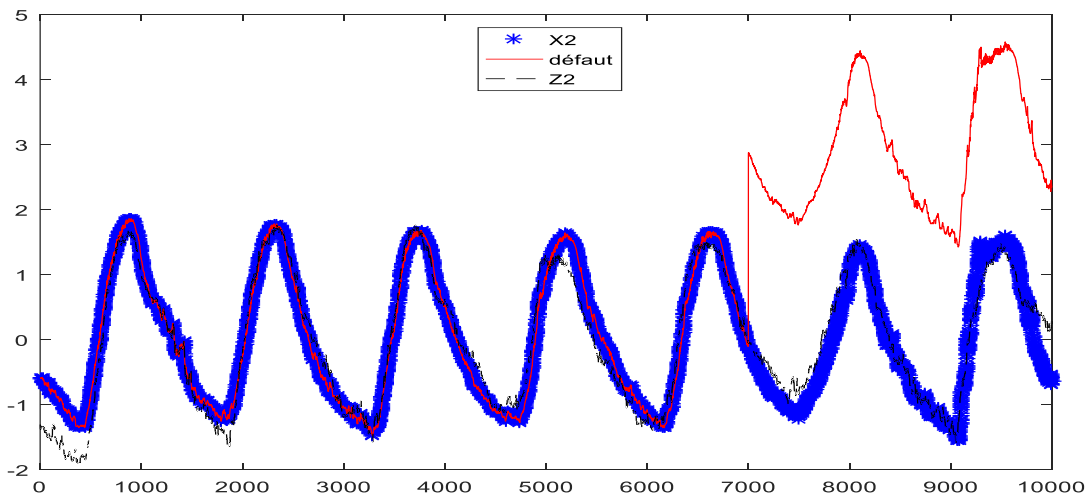
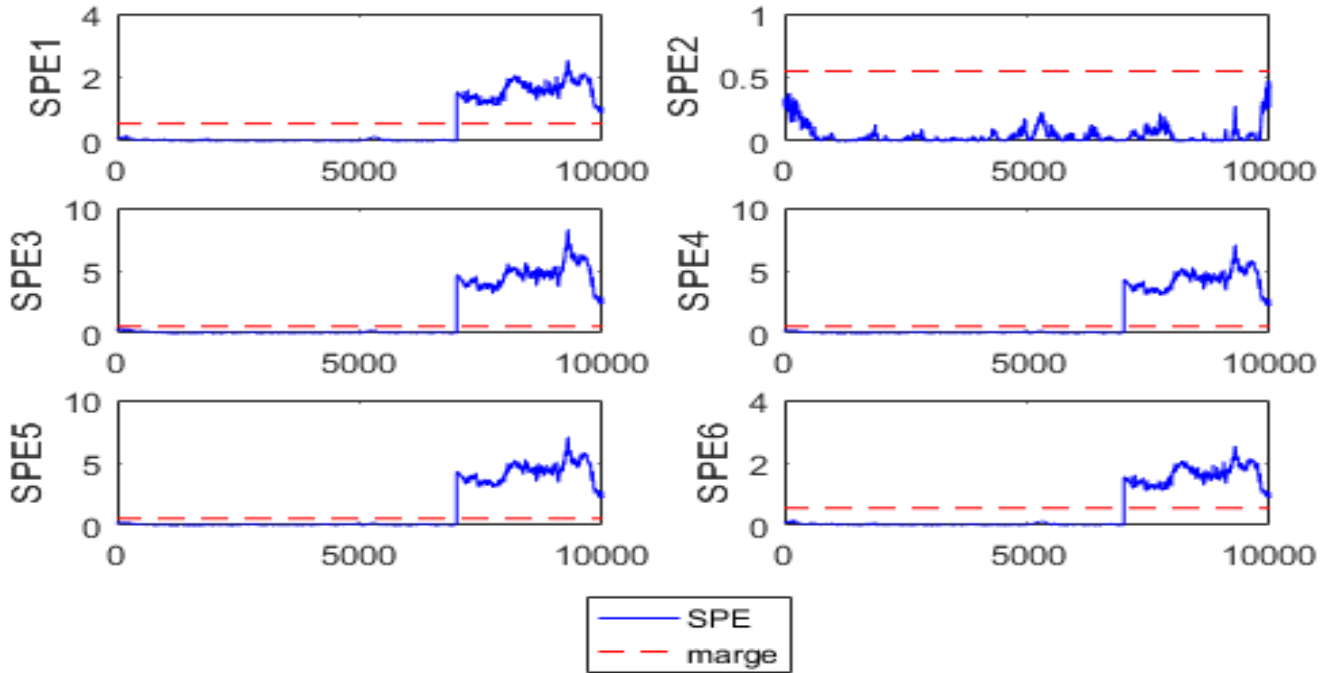


Figure IV.13: Calcul de reconstruction de variables en X2



On remarque que toutes les variables sauf la première variable ont été affectées par l'erreur, ce qui signifie, selon le principe d'erreur de reconstruction, que la première variable est celle qui porte l'erreur.

**Localisation :**



**Figure IV.14:** Evolution des différents SPE obtenus par la méthode de reconstruction avec un défaut affectant la variable  $x_2$  à partir de l'instant 7000

#### IV.4.2 Modélisation par ACP en type intervalles

Dans cette partie, nous appliquerons la méthode ACP pour la détection et la localisation de défauts sur les mêmes données précédentes avec un intervalle qui représente 20% de sa plage de variation, après l'application du modèle ACP, on remarque sur la figure que les estimations sont correctes et les erreurs nulles, Ceci montre que le modèle ACP est valide.

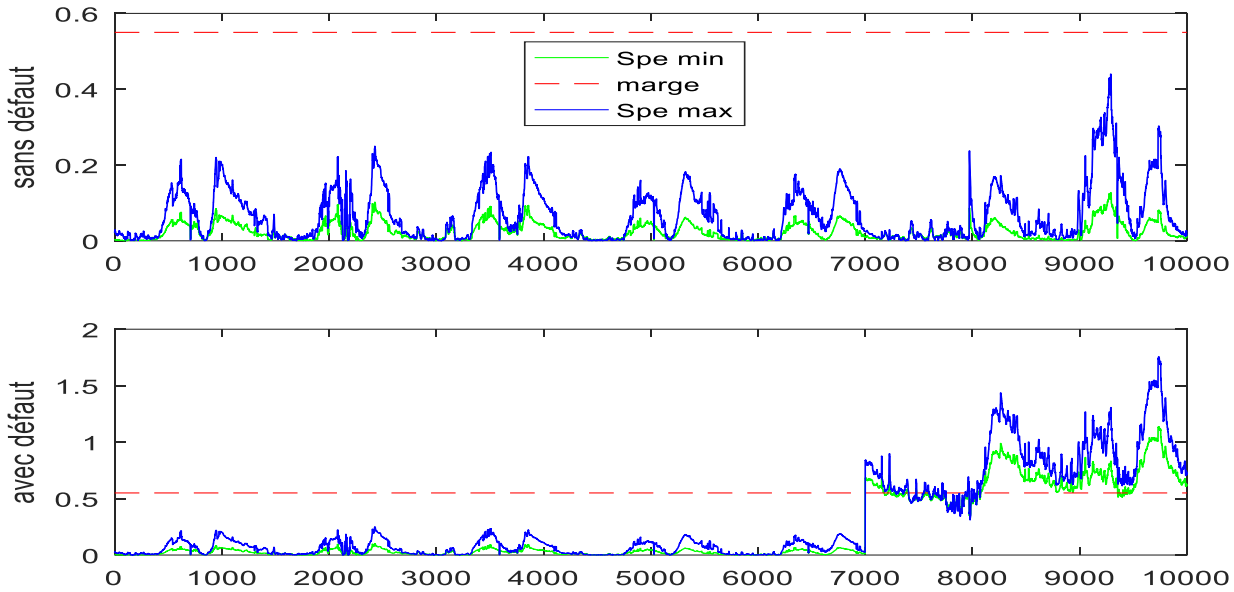


Figure IV.15 : Evolution des différentes mesures et leurs estimations

**Détection et localisation de défaut en type intervalle :**

Pour la détection et localisation nous simulons un défaut sur la variablex2 avec un biais qui représente 40% de sa plage de variation, à partir de l’instant 7000. Nous calculons l’indice de détection de défauts SPE.La figure (fig,et ) représente l’évolution de cet indice en absence et avec défaut.

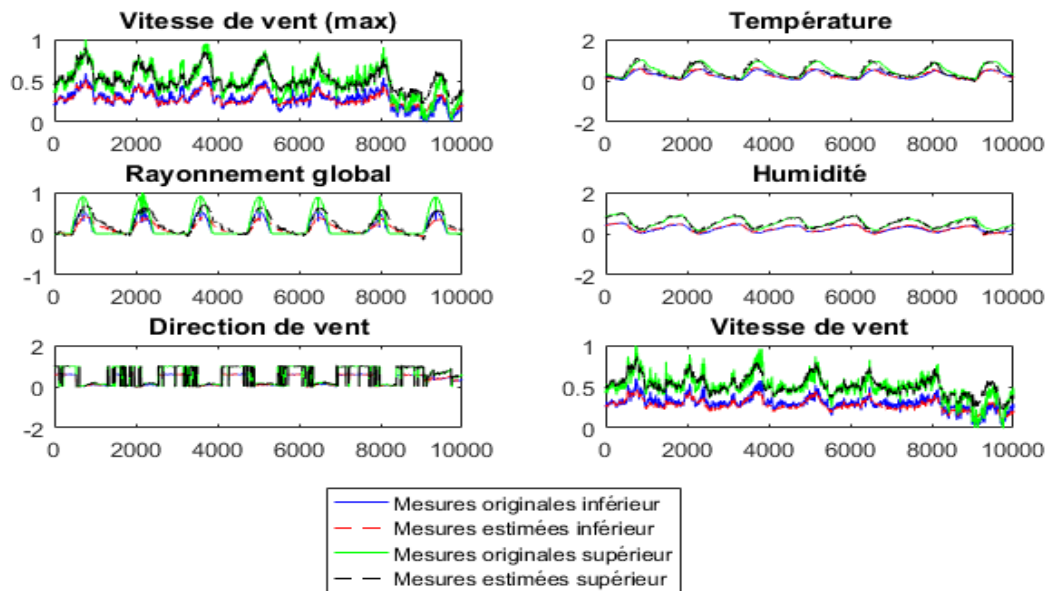
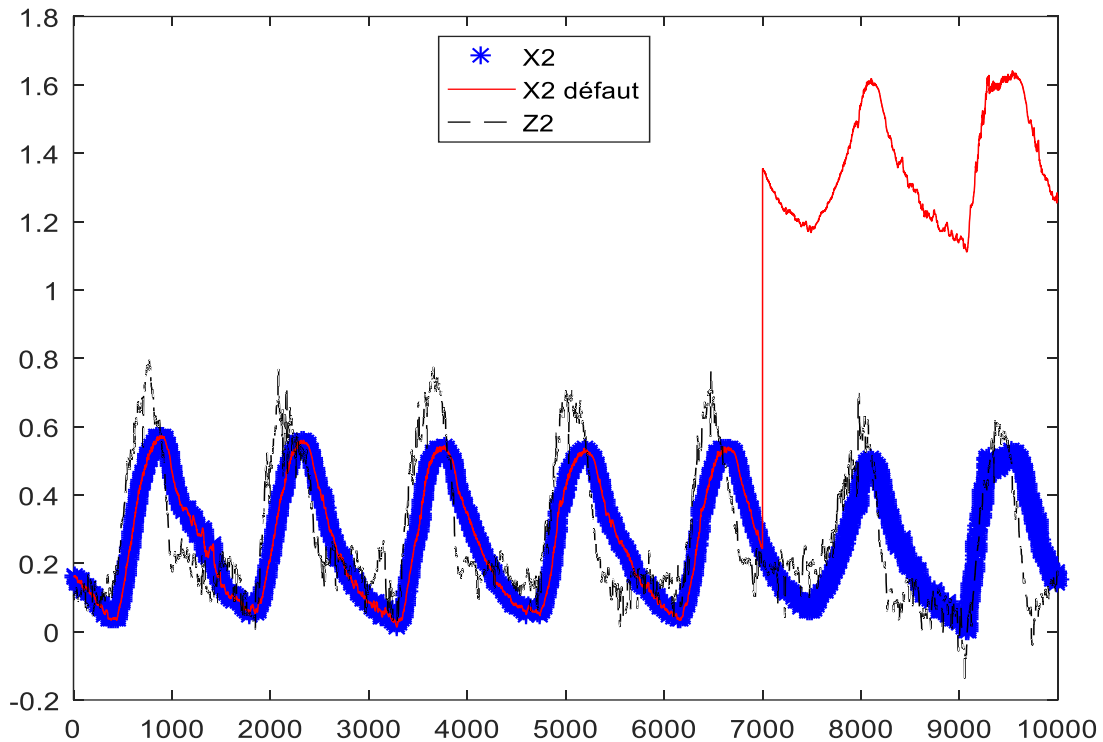


Figure IV.16 : L'évolution de l'indices SPE en type intervalle

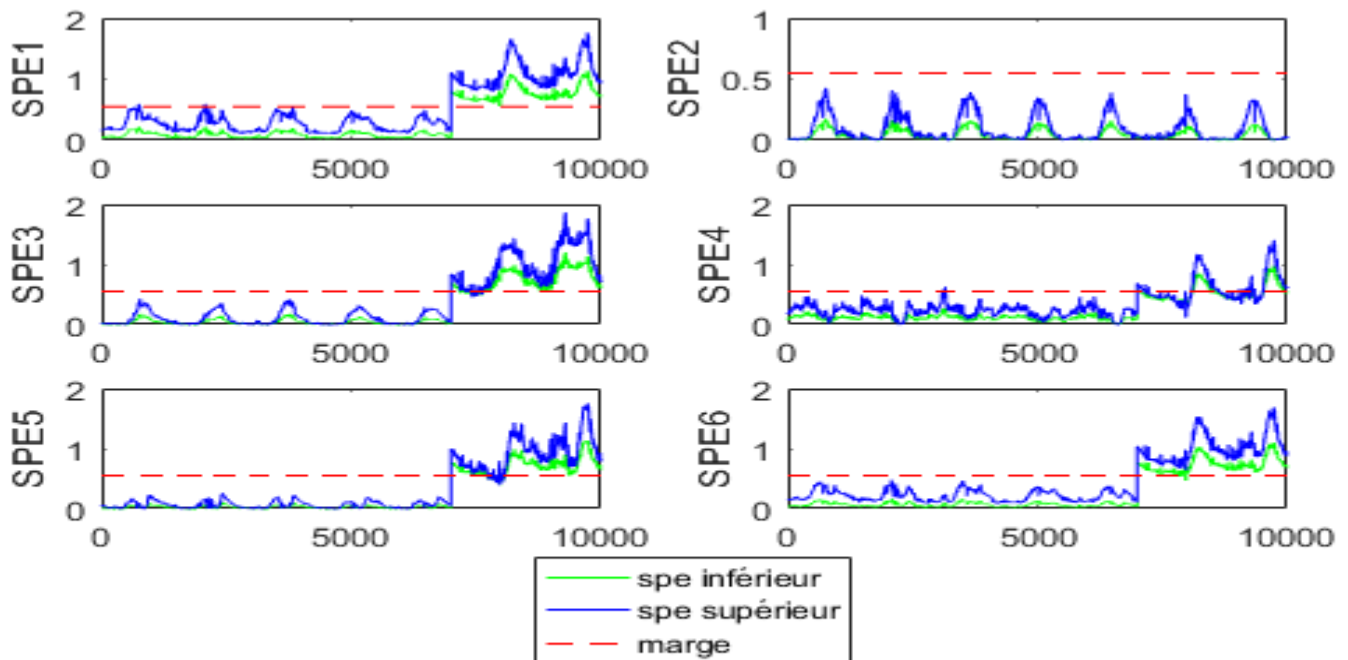
Une fois la défaillance est bien détectée par un indice de détections, une étape d'isolation de ce défaut intervient pour savoir la provenance de ce défaut. Dans notre cas nous avons utilisé la méthode de reconstruction des variables pour la localisation de défaut.



**Figure IV.17 :** Calcul de reconstruction de variable X2 en type intervalle

On remarque que toutes les variables sauf la première variable ont été affectées par l'erreur, ce qui signifie, selon le principe d'erreur de reconstruction, que la première variable est celle qui porte l'erreur.

**localisation :**



**Figure IV.18 :** Evolution des différents SPE obtenus par la méthode de reconstruction avec un défaut affectant la variable  $x_2$  à partir de l'instant 7000

### Conclusion :

Dans le domaine du diagnostic, diverses méthodes basées sur le concept de redondance d'informations ont été développées. Leur principe de fonctionnement repose généralement sur des tests de concordance entre le comportement observé du processus, fourni par les capteurs, et le comportement attendu défini par un modèle mathématique du processus. Ces méthodes d'analyse de redondance nécessitent la présence d'un modèle du système à surveiller. La détection des défauts s'appuie sur l'évaluation de mesures et de statistiques (par exemple, l'erreur carrée de prédiction). La localisation des défauts est réalisée en appliquant des méthodes de calcul de reconstruction, qui ont été identifiées comme étant efficaces. Pour illustrer l'application de ces techniques de modélisation, de détection et d'isolation des défauts, nous avons réalisé avec succès une étude sur une station météorologique automatique.

## Conclusion générale

En conclusion, l'analyse des composantes principales (ACP) est une méthode statistique utile pour déterminer les relations entre les variables et réduire la dimensionnalité des données. Elle permet la transformation d'ensembles de données complexes et multidimensionnels en un espace de dimensions réduites tout en conservant la majorité des informations contenues dans les données d'origine.

L'ACP a de nombreuses utilisations dans divers domaines, notamment la finance, la biologie, la psychologie, la géographie, entre autres. En simplifiant l'ensemble de données et en révélant les structures sous-jacentes, les relations entre les variables et les similitudes entre les observations, elle facilite l'analyse, la modélisation et la visualisation des données.

Cependant, lorsqu'il s'agit de la surveillance des systèmes de mesures à caractère incertaines, les approches traditionnelles de l'ACP peuvent ne pas toujours être fiables. Pour obtenir une analyse robuste et une détection précise des défauts dans de tels cas, il est essentiel de prendre en compte les incertitudes et d'intégrer des représentations par intervalles des données. En fournissant des bornes plus strictes sur les mesures estimées et en tenant compte des incertitudes de mesure. Les méthodes d'ACP basées sur les intervalles, telle que la VPCA, permet de résoudre ces limitations.

En incorporant l'incertitude et les représentations par intervalles dans l'ACP, il est possible de traiter les erreurs, les incertitudes et les variations inhérentes au processus de mesure. Les résultats d'analyse sont plus précis et plus solides, ce qui permet une meilleure détection des défauts, une surveillance améliorée des processus et une prise de décision fiable.

Dans l'ensemble, l'ACP, associée aux techniques basées sur les intervalles, fournit un cadre puissant pour l'analyse des données, la modélisation et la détection de défauts. Elle permet aux chercheurs et aux praticiens d'acquérir une compréhension plus approfondie des systèmes complexes tout en tenant compte des incertitudes et des variations présentes dans les données.

## Références bibliographiques

- [1] Mr.M.F.Harkat-La diagnostic industrielle- université badji mokhtar-annaba
- [2]système météorologique intégré-microStep-avril 2018
- [3] Yvon T , Mohamed F H , Gilles M , Jose R , Détection et localisation de défauts multiples par analyse en composantes principales Bordeaux, France May 2009
- [4] Harkat M-F , Détection et Localisation de défaut par Analyse en Composantes Principales Centre de Recherche en Automatique de Nancy | UMR-CNRS 70392,Avenue de la Foret de Haye 54516 Vandœuvre Lûes-Nancy 30 Juin 2003 .
- [5] Alain R , Gilles G , Analyse en composantes principales INSA Rouen -Département ASI Laboratoire LITISv
- [6] Chakour C , Diagnostic et surveillance des procédés industriel set de le environnement sur la base de l'analyse de données , Université Badji Mokhtar , Annaba 2016
- [7] Arar K , T adjine S , Modélisation et Diagnostic des systèmes par l'Analyse en Composantes Principales Multi-Echelle (MSPCA) Faculté des Sciences de l'Ingénieur ,Département d'Electronique BP. 12, Sidi Amar 23000 Annaba , 2007 .
- [8] Chakour C , Faouzi H , Messaoud D , Neuronal Principal Component Analysis for Nonlinear Time-Varying Processes Monitoring Badji Mokhtar University, Institute of Electronics, Annaba, Algeria. September 2-4, 2015.
- [9] Dimitrios F , Détection et localisation des défauts provenant des capteurs et des actionneurs , application sur un système non linéaire , Université Paul Sabatier – Toulouse III, 2008. Français Submitted on 18 Nov 2008 .
- [10] Yvon T , L'Analyse en Composantes Principales pour le diagnostic , Institut National Polytechnique de Lorraine Centre de Recherche en Automatique de Nancy.
- [11]Amimar Tarek,Akbal Mohamed Amine-Détection et localisation de défauts par l'analyse en composantes principales :application à système hydraulique .
- [12] Tarek Ait-Izema,, M-Faouzi Harkat , Messaoud Djeghaba, Frédéric Kratz- On the application of interval PCA to process monitoring: A robust strategy for sensor FDI with new efficient control statistics - Journal of Process Contro
- [13] Cours Modélisation à base de l'analyse en composantes principales (ACP)-2022