Diagnostic de fonctionnement par analyse en composantes principales : Application à une station de traitement des eaux usées

Tharrault Yvon

sous la direction de
J. Ragot et G. Mourot

CRAN

cnrs dépasser les frontières

CENTRE DE RECHERCHE PUBLIC HENRI TUDOR
www.tudor.lu

Nancy-Université
INPL

Contexte

La performance et la fiabilité de l'ensemble des moyens de commande et de contrôle des systèmes industriels sont liées à la qualité des systèmes de mesures.

Objectif

Valider les données des capteurs utiles au pilotage d'une station de traitement des eaux usées

Outil de modélisation

Analyse en Composantes Principales

Plan

1. Principe de l'Analyse en Composantes Principales (ACP)
    1. Identification du modèle ACP
    2. Détection de défauts
2. L'analyse en composantes principales robuste
3. Localisation de défauts multiples
4. Application à une station de traitement des eaux usées
5. Conclusion

Principe de l'Analyse en Composantes Principales (ACP)

- Matrice de données $X \in \Re^{N \times n}$ collectées sur le système en fonctionnement normal

ACP

Maximisation de la variance des projections des données sur des axes particuliers $T = XP$

- $T \in \Re^{N \times n}$: Matrice des composantes principales
- $P \in \Re^{n \times n}$: Matrice de projection

Décomposition en valeurs/vecteurs propres de la matrice de covariance

$$\Sigma = \frac{1}{N-1} X^T X = P \Lambda P^T \qquad \text{avec} \qquad PP^T = P^T P = I_n$$

$$\Sigma = \begin{bmatrix} P_\ell & P_{n-\ell} \end{bmatrix} \begin{bmatrix} \Lambda_\ell & 0 \\ 0 & \Lambda_{n-\ell} \end{bmatrix} \begin{bmatrix} P_\ell^T \\ P_{n-\ell}^T \end{bmatrix}$$

Principe de l'Analyse en Composantes Principales (ACP)

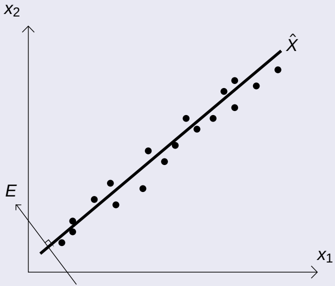
Décomposition

Partie principale

$\hat{X} = X\,C_\ell$
où $C_\ell = P_\ell P_\ell^T$

Partie résiduelle

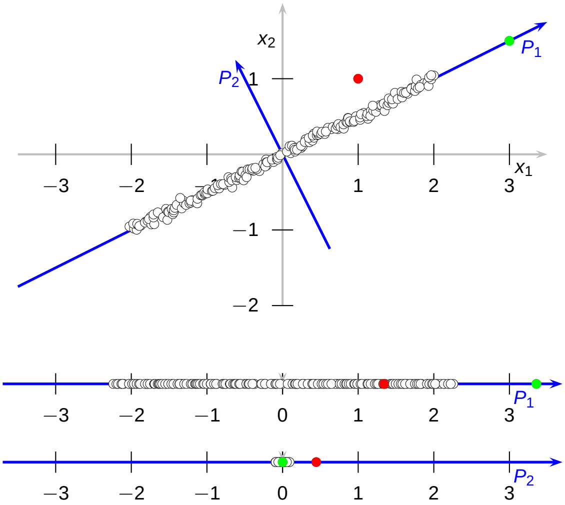
$E = X - \hat{X}$
$E = X\,(I_n - C_\ell)$



Détermination du nombre de composantes principales ℓ

Projection des défauts dans l'espace principal et l'espace résiduel

- $n = 2$ variables x_1 et x_2
- $X = [x_1\ x_2]$
- $\ell = 1$
- défaut 1 :
  observation verte
- défaut 2 :
  observation rouge



Détection de défauts

Indicateurs généralement utilisés

SPE ou T_H^2 pour détecter un défaut dans l'espace résiduel

$$SPE(k) = x(k)^{\mathrm{T}} P_{n-\ell} P_{n-\ell}^{\mathrm{T}} x(k)$$

$$\mathrm{T}_H^2(k) = x(k)^{\mathrm{T}} P_{n-\ell} \Lambda_{n-\ell}^{-1} P_{n-\ell}^{\mathrm{T}} x(k)$$

T^2 pour détecter un défaut dans l'espace principal

$$T^2(k) = x(k)^{\mathrm{T}} P_\ell \Lambda_\ell^{-1} P_\ell^{\mathrm{T}} x(k)$$

Afin d'assurer la détection de l'ensemble des défauts

Utilisation d'un indicateur combiné

$$\eta(k) = \frac{T^2(k)}{\chi_l^2} + \frac{SPE(k)}{\delta^2}$$

avec χ_l^2 seuil de détection de l'indice T^2 et δ^2 le seuil de détection de l'indice SPE.

Utilisation de la distance de Mahalanobis

$$\begin{aligned} D(k) &= T^2(k) + \mathrm{T}_H^2(k) \\ &= x(k)^{\mathrm{T}} P \Lambda^{-1} P^{\mathrm{T}} x(k) \end{aligned}$$

Détection de défauts : exemple

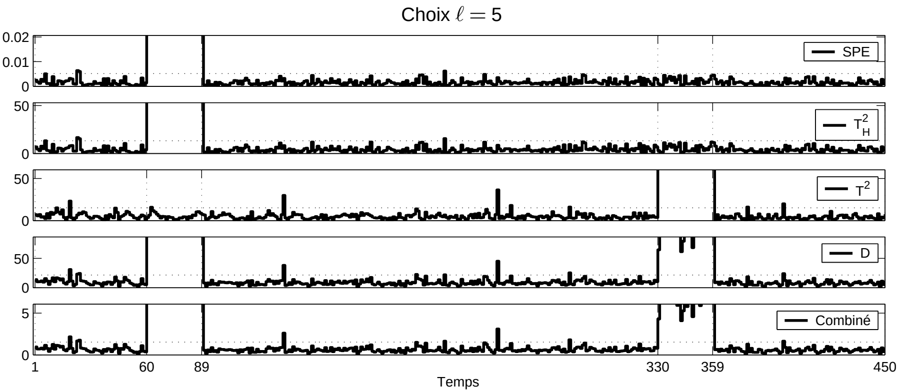
- $x(k) = \begin{bmatrix} x_1(k) & x_2(k) & x_3(k) & x_4(k) & x_5(k) & x_6(k) & x_7(k) & x_8(k) & x_9(k) \end{bmatrix}^T$

$$\begin{aligned}
x_1(k) &= v(k)^2 + \sin(0.1k), \quad v(k) \sim \mathcal{N}(0,1) \\
x_2(k) &= 2\sin(k/6)\cos(k/4)\exp(-k/N) \\
x_3(k) &= \log(x_2(k)^2) \\
x_4(k) &= x_1(k) + x_2(k) \\
x_5(k) &= x_1(k) - x_2(k) \\
x_6(k) &= 2x_1(k) + x_2(k) \\
x_7(k) &= x_1(k) + x_3(k) \\
x_8(k) &\sim \mathcal{N}(0,1) \\
x_9(k) &\sim \mathcal{N}(0,1)
\end{aligned}$$

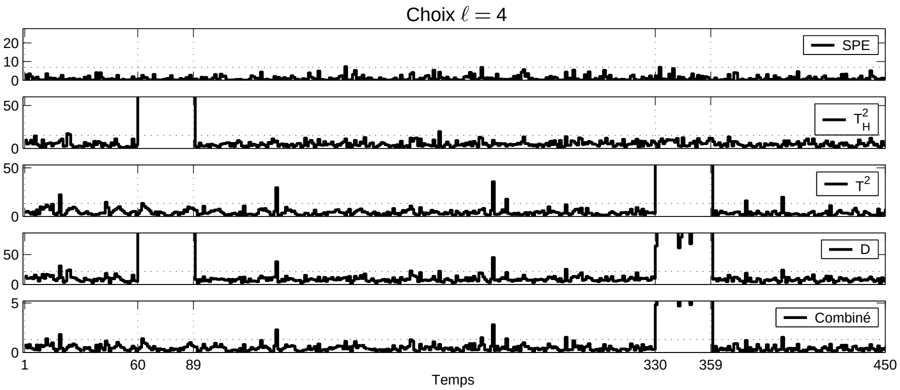
	I_1 {69 : 89}	I_2 {330 :359}
x_1	$\times$	0
x_8	0	$\times$

- La matrice X est composée de $N = 128$ observations du vecteur $x(k)$
- $\ell = 5$ composantes principales

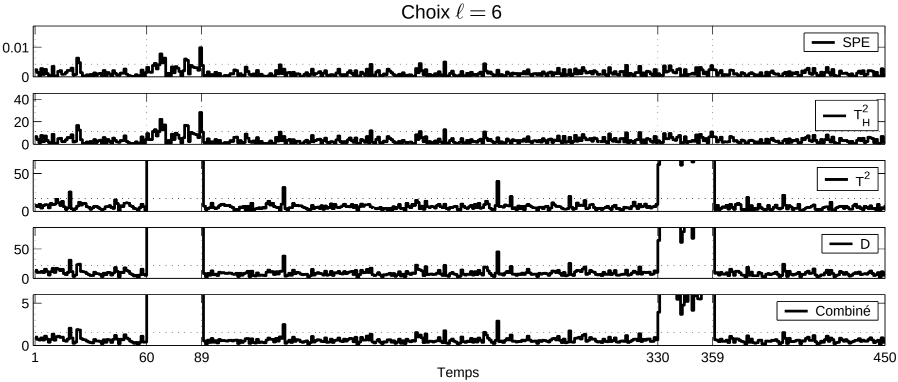
Détection de défauts : robustesse au nombre de composantes principales ℓ



Détection de défauts : robustesse au nombre de composantes principales ℓ



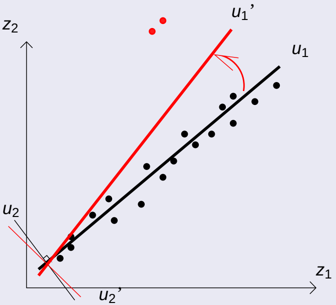
Détection de défauts : robustesse au nombre de composantes principales ℓ



Faiblesse de l'ACP

Sensibilité aux valeurs aberrantes

Valeurs aberrantes : Observations (obtenues durant des périodes de démarrage, d'arrêt, de fonctionnement dégradé, erreurs de mesure, ...) qui ne correspondent pas au fonctionnement normal ou courant du système et qui sont généralement minoritaires.



Modèle ACP robuste aux valeurs aberrantes

MM-estimator Robust Principal Component Analysis (MMRPCA)

MMRPCA

Résidu $r(k)$

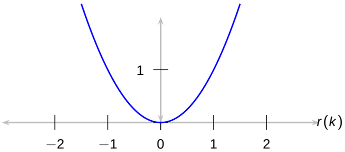
$r(k) = ||P_{n-\ell}^T(x(k) - \mu)||^2$

avec $x(k)$ une observation
μ la moyenne du jeu de données X
$P_{n-\ell}$ est la matrice des vecteurs propres de la matrice de variance-covariance associés aux $n-\ell$ plus petites valeurs propres

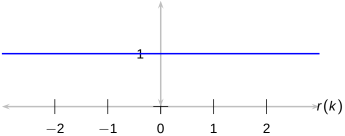
L'ACP minimise le critère suivant sous contrainte $P_{n-\ell}^T P_{n-\ell} = I_{n-\ell}$:

$$\frac{1}{N}\sum_{k=1}^{N}(r(k))$$

Fonction objectif x^2



Fonction poid



MMRPCA

Résidu $r(k)$

$$r(k) = ||P_{n-\ell}^T(x(k)-\mu)||^2$$

avec $x(k)$ une observation
μ la moyenne du jeu de données X
$P_{n-\ell}$ est la matrice des vecteurs propres de la matrice de variance-covariance associés aux $n-\ell$ plus petites valeurs propres

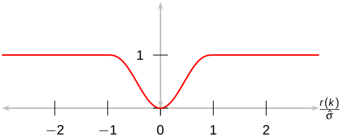
Le MM-estimateur minimise le critère suivant sous contrainte $P_{n-\ell}^T P_{n-\ell} = I_{n-\ell}$:

$$\frac{1}{N}\sum_{k=1}^{N}\rho\left(\frac{r(k)}{\hat{\sigma}}\right)$$

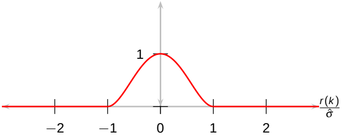
avec ρ la fonction de objectif et $\hat{\sigma}$ la dispersion robuste des résidus r obtenue en minimisant le critère suivant :

$$\frac{1}{N}\sum_{k=1}^{N}\rho\left(\frac{r(k)}{\hat{\sigma}_i}\right) = \delta$$

Fonction objectif ρ



Fonction poid **w**



MMRPCA

Algorithme de Maronna (2005)

1. Initialisation
2. Calcul des résidus r
3. Calcul de la dispertion robuste des résidus $\hat{\sigma}$
4. Calcul des poids associés à chaque observation $\mathbf{w}(k) = \dot{\rho}\left(\frac{r(k)}{\hat{\sigma}}\right)$
5. Calculer μ à partir de $\mu = \frac{\sum_{k=1}^{N} \mathbf{w}(k) x(k)}{\sum_{k=1}^{N} \mathbf{w}(k)}$
6. Calculer S à partir de $S = \sum_{k=1}^{N} \mathbf{w}(k)(x(k) - \mu)(x(k) - \mu)^T$
7. Calculer $P_{n-\ell}$ la matrice des vecteurs propres de la matrice de variance-covariance robuste S correspondant aux $n - \ell$ plus petites valeurs propres.

MMRPCA

Initialisation avec une matrice de variance-covariance robuste

Matrice de variance-covariance robuste (Caussinus et al., 2003)

$$V = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} w(i,j)(x(i)-x(j))(x(i)-x(j))^T}{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} w(i,j)}$$

où les poids $w(i,j)$ sont définis par :

$$w(i,j) = \exp\left(-\frac{\beta}{2}(x(i)-x(j))^T \Sigma^{-1}(x(i)-x(j))\right)$$

avec β un paramètre de réglage

Pour $\beta = 0$, $V = 2\Sigma$

MMRPCA

Robuste seulement aux valeurs aberrantes possédant une projection dans l'espace résiduel

MM-estimateur

Le MM-estimateur maximise le critère suivant sous contrainte $P_\ell^T P_\ell = I_\ell$:

$$\frac{1}{N}\sum_{k=1}^{N}\rho\left(\frac{||P_\ell^T(x(k)-\mu)||^2}{\hat{\sigma}}\right)$$

avec $\hat{\sigma}$ la dispersion robuste des résidus r et ρ est la fonction de pondération.

Algorithme de la méthode MMRPCA

Algorithme

1. Calcul de la matrice de variance-covariance robuste V
2. L'algorithme des MM-estimateurs pour les défauts avec une projection dans l'espace résiduel
   1. Résultats : Poids associés à chaque observation w_r
   2. Résultats : S_r
3. L'algorithme des MM-estimateurs pour les défauts avec une projection dans l'espace principal
   1. $w = min(w_r, w_p)$
4. Détection et élimination des valeurs aberrantes

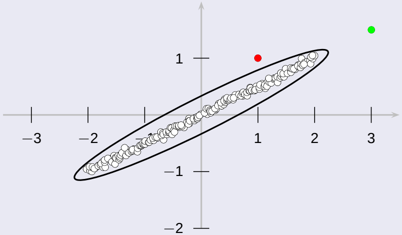
L'analyse en composantes principales robuste aux valeurs aberrantes

Minimum covariance determinant

L'estimateur MCD cherche h observations qui minimisent le déterminant de la matrice de variance-covariance des données.
où h est le nombre d'observations considérées comme "saines"

- Nécessite la connaissance du nombre de valeurs aberrantes $(N - h)$ inconnues a priori

$\Rightarrow$ Pour assurer le rejet des valeurs aberrantes choix de $h = 0.5N$



Exemple de comparaison des méthodes MMRPCA et MCD

- $x(k) = \begin{bmatrix} x_1(k) & x_2(k) & x_3(k) & x_4(k) & x_5(k) & x_6(k) & x_7(k) & x_8(k) & x_9(k) \end{bmatrix}^T$

$$
\begin{aligned}
x_1(k) &= v(k)^2 + \sin(0.1k), \quad v(k) \sim \mathcal{N}(0,1) \\
x_2(k) &= 2\sin(k/6)\cos(k/4)\exp(-k/N) \\
x_3(k) &= \log(x_2(k)^2) \\
x_4(k) &= x_1(k) + x_2(k) \\
x_5(k) &= x_1(k) - x_2(k) \\
x_6(k) &= 2x_1(k) + x_2(k) \\
x_7(k) &= x_1(k) + x_3(k) \\
x_8(k) &\sim \mathcal{N}(0,1) \\
x_9(k) &\sim \mathcal{N}(0,1)
\end{aligned}
$$

	Intervalles	
	I_1	I_2
x_1	$\times$	0
x_8	0	$\times$

- La matrice X est composée de $N = 128$ observations du vecteur $x(k)$

Exemple de comparaison des méthodes MMRPCA et MCD

	le défaut n'est pas détecté
	le défaut est détecté uniquement avec la méthode MMRPCA
	le défaut est détecté uniquement avec la méthode MCD
	le défaut est détecté avec la méthode MCD et avec la méthode MMRPCA

Durée de l'intervalle I_1	Amplitude du défaut sur x_1																
	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3	4	5	6	7	8	9	10
5%																	
15%																	
25%																	
35%																	
45%																	

Durée de l'intervalle I_2	Amplitude du défaut sur x_8											
	6	9	12	15	18	21	30	40	50	60	70	80
5%												
15%												
25%												
35%												
45%												

Méthode MMRPCA domaine d'application plus large que la méthode MCD

Localisation de défauts multiples

Localisation de défauts multiples

Explosion combinatoire du nombre de situations possibles de défauts

$$N_{situations} = \sum_{r=1}^{n} \mathbb{C}_n^r$$

avec $\mathbb{C}_n^r$ représente le nombre de combinaisons possibles de choisir r variables parmi n.

Méthode

Principe de reconstruction des variables (minimisation du défaut)

Localisation de défauts multiples

La reconstruction $\hat{x}_R$

Minimiser l'influence des défauts sur l'indicateur de détection

$$\hat{x}_R(k) = x(k) - \Xi_R f_R$$

avec f_R : amplitude (inconnue) du défaut
Ξ_R : matrice des directions de reconstruction

Par exemple, pour reconstruire 2 variables ($R = 2,4$) parmi 5 variables

$$\Xi_R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T$$

L'estimation de l'amplitude du défaut $\hat{f}_R$:

$$\hat{f}_R = \arg\min_{f_R} \{D_R(k)\}$$

avec $D_R(k) = \hat{x}_R(k)^T P \Lambda^{-1} P^{\mathrm{T}} \hat{x}_R(k)$

Localisation de défauts multiples

Le vecteur de reconstruction $\hat{x}_R(k)$ du vecteur $x(k)$ est donné par :

$$\hat{x}_R(k) = \left(I - \Xi_R(\Xi_R^T P\Lambda^{-1}P^{\mathrm{T}}\Xi_R)^{-1}\Xi_R^T P\Lambda^{-1}P^{\mathrm{T}}\right)\, x(k)$$

Condition de reconstruction structurelle :

Existence de $(\Xi_R^T P\Lambda^{-1}P^{\mathrm{T}}\Xi_R)^{-1} \Rightarrow$ la matrice $\Xi_R^T P\Lambda^{-1}P^{\mathrm{T}}\Xi_R$ de plein rang

Pour qu'un défaut soit reconstructible, il faut qu'il soit au minimum projeté dans l'espace principal ($r \leq \ell$) ou dans l'espace résiduel ($r \leq n-\ell$). Cette condition implique que le nombre de variables reconstruites r doit respecter l'inégalité :

$$r \leq \max(n-\ell, \ell)$$

avec r : Nombre de variables reconstruites
ℓ : Nombre de composantes principales
n : Nombre de variables

Détection et localisation de défauts multiples

Condition de reconstruction : $r \leq \max(n-\ell, \ell)$

Le nombre maximum de reconstructions est le suivant :

$$\sum_{r=1}^{\max(n-\ell,\ell)} \mathbb{C}_n^r$$

avec $\mathbb{C}_n^r$ représente le nombre de combinaisons possibles de choisir r variables parmi n.

Localisation de défauts

Pour les observations en défaut, les ensembles de variables défaillantes $\hat{R}$ sont déterminés comme suit :

$$\hat{R} = \arg_{R \in \Im} D_R(k) < \gamma_\alpha^2$$

où γ_α^2 le seuil de détection de l'indicateur D_R
et $\Im$ est l'ensemble des combinaisons des directions de reconstructions possibles.

Détection et localisation de défauts multiples

Réduction du nombre de reconstruction

Détermination des directions de projections colinéaires.

Un indicateur K est alors construit :

$$K(R_1,R_2)=\max\{(d(R_1,R_2),\tilde{d}(R_1,R_2)\}$$

où $d(R_1,R_2)$ distance entre deux sous espaces dans l'espace principal et $\tilde{d}(R_1,R_2)$ distance entre deux sous espaces dans l'espace résiduel :

$$\begin{aligned}
d(R_1,R_2) &= ||\hat{\Xi}_{R_1}(\hat{\Xi}_{R_1}^T\hat{\Xi}_{R_1})^{-1}\hat{\Xi}_{R_1}^T-\hat{\Xi}_{R_2}(\hat{\Xi}_{R_2}^T\hat{\Xi}_{R_2})^{-1}\hat{\Xi}_{R_2}^T||_2\\
\tilde{d}(R_1,R_2) &= ||\tilde{\Xi}_{R_1}(\tilde{\Xi}_{R_1}^T\tilde{\Xi}_{R_1})^{-1}\tilde{\Xi}_{R_1}^T-\tilde{\Xi}_{R_2}(\tilde{\Xi}_{R_2}^T\tilde{\Xi}_{R_2})^{-1}\tilde{\Xi}_{R_2}^T||_2
\end{aligned}$$

avec $\hat{\Xi}_{R_1}=\Lambda_\ell^{-1/2}P_\ell^T\Xi_{R_1}$, $\tilde{\Xi}_{R_1}=\Lambda_{n-\ell}^{-1/2}P_{n-\ell}^T\Xi_{R_1}$ et R_1 et R_2 correspondent aux ensembles des indices des variables reconstruites.

Par exemple, pour reconstruire 2 variables ($R_1=2,4$) parmi 5 variables

$$\Xi_{R_1}=\left[\begin{array}{ccccc}0&1&0&0&0\\0&0&0&1&0\end{array}\right]^T$$

Localisation des défauts

Algorithme pour la détermination des défauts détectables et localisables

1. $r = 1$
2. Calculer pour l'ensemble des directions possibles l'indicateur $K(R_1, R_2)$.
   Si $K(R_1, R_2)$ est nul :
   - seul l'ensemble des variables potentiellement en défaut peut être déterminé, c'est-à-dire les variables avec les indices R_1, R_2 ou R_1 et R_2. Il est donc nécessaire de considérer une seule direction, par exemple R_1.

   Si $K(R_1, R_2)$ est proche de zéro :
   - l'amplitude du défaut doit être importante pour assurer la localisation du défaut

   Sinon le défaut est localisable
3. $r = r + 1$
4. Tant que $r \leq \max(\ell, n - \ell)$ aller à l'étape 2

Exemple numérique : cas multiple défauts

On considère :

- $x(k) = \left[\begin{array}{ccccccccc} x_1(k) & x_2(k) & x_3(k) & x_4(k) & x_5(k) & x_6(k) & x_7(k) & x_8(k) & x_9(k) \end{array}\right]^T$

$$x_1(k) = v(k)^2 + \sin(0.1k), \quad v(k) \sim \mathcal{N}(0,1)$$
$$x_2(k) = 2\sin(k/6)\cos(k/4)\exp(-k/N)$$
$$x_3(k) = \log(x_2(k)^2)$$
$$x_4(k) = x_1(k) + x_2(k)$$
$$x_5(k) = x_1(k) - x_2(k)$$
$$x_6(k) = 2x_1(k) + x_2(k)$$
$$x_7(k) = x_1(k) + x_3(k)$$
$$x_8(k) \sim \mathcal{N}(0,1)$$
$$x_9(k) \sim \mathcal{N}(0,1)$$

	I_1 {10 : 24}	I_2 {35 : 49}	I_3 {60 : 74}	I_4 {85 : 99}
x_1	$\times$	$\times$	0	0
x_2	0	$\times$	0	0
x_3	$\times$	0	0	0
x_4	0	0	$\times$	0
x_8	0	0	$\times$	$\times$

- La matrice X est composée de $N = 128$ observations du vecteur $x(k)$

Exemple numérique : reconstruction utile

$\max(\ell, n-\ell) = \max(5,4) = 5$

Le nombre maximum de reconstructions est donc de 381.

$r = 1$

K		R_1								
		1	2	3	4	5	6	7	8	9
R_2	1	0	1	0.99	1.00	0.92	0.92	1.00	1.00	0.99
	2		0	1.00	0.96	0.73	0.97	1.00	1.00	1.00
	3			0	1.00	0.96	1.00	0.40	1.00	1.00
	4				0	0.97	0.73	1.00	0.99	1.00
	5					0	0.97	0.99	1.00	1.00
	6						0	1.00	0.99	1.00
	7							0	1.00	0.99
	8								0	1.00

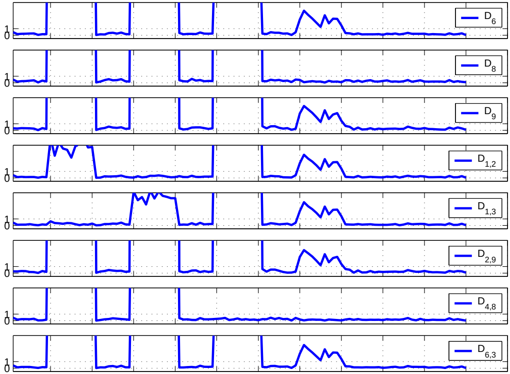
r=2

Pour $R_1 = \{1,3\}$ et $R_2 = \{1,7\}$, $K(R_1, R_2) = 0$

$\Rightarrow$ Les projections des directions 1,3 et 1,7 sont colinéaires

Le nombre de reconstructions nécessaires est de 150 au lieu de 381

Exemple numérique : reconstruction utile



	I_1	I_2	I_3	I_4
$D_{1,3}$	0	$\times$	$\times$	$\times$
$D_{1,2}$	$\times$	0	$\times$	$\times$
$D_{4,8}$	$\times$	$\times$	0	0
D_8	$\times$	$\times$	$\times$	0

Localisations des défauts sur les différents intervalles

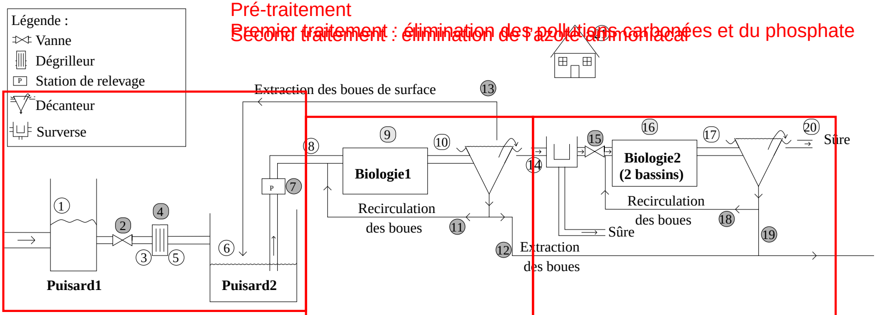
- dans l'intervalle I_1, x_1 et x_3 ou/et x_7 sont en défaut
- dans l'intervalle I_2, x_1 et x_2 sont en défaut
- dans l'intervalle I_3, x_4 et x_8 sont en défaut
- dans l'intervalle I_4, x_8 est en défaut

Description de la station de traitement des eaux usées

Pré-traitement

Premier traitement : élimination des pollutions carbonées et du phosphate

Second traitement : élimination de l'azote ammoniacal



1. Hauteur du puisard1 : $H1$
2. % ouverture vanne : $V2$
3. Hauteur avant le dégrillage : $H3$
4. Cmd râteaux : $C4$
5. Hauteur après le dégrillage : $H5$
6. Hauteur du puisard2 : $H6$
7. Cmd pompes de relevages : $C7_1$ à $C7_6$
8. Débit : $Q8$
   Température : $T8$
   Conductivité : $\sigma 8$
   Absorption UV : $A8$
   Potentiel Hydrogène : $PH8$

9. Matière en suspension : $MES9$
   Oxygène dissous : $OD9$
9. Cmd aérateur : $CA9$
   Cmd floculation phosphate : $Fp9$
10. Ammonium : $NH10$
    Nitrate : $NO10$
    Orthophosphate : $OP10$
11. Cmd recirculation des boues : $C11$
12. Cmd extraction des boues : $C12$
13. Cmd boue de surface : $C13$

14. Hauteur avant la surverse : $H14$
15. % ouverture vanne : $V15$
16. Matière en suspension : $MES16_1$, $MES16_2$
    Oxygène dissous : $OD16_1$ et $OD16_2$
16. Cmd aérateur : $CA16$
17. Ammonium : $NH17$
    Nitrate : $NO17$
18. Cmd recirculation des boues : $C18$
19. Cmd extraction des boues : $C19$
20. Orthophosphate : $OP20$
21. Pluviomètre : $P21$

Application à la station de traitement des eaux usées

Objectifs

Valider les données des capteurs utiles au pilotage d'une station de traitement des eaux usées

Construction de la matrice de données X

Présence de phénomènes dynamiques $\longrightarrow$ Détermination des décalages temporels
Présence de phénomènes non-linéaires $\longrightarrow$ Détermination des transformations non linéaires à prendre en compte

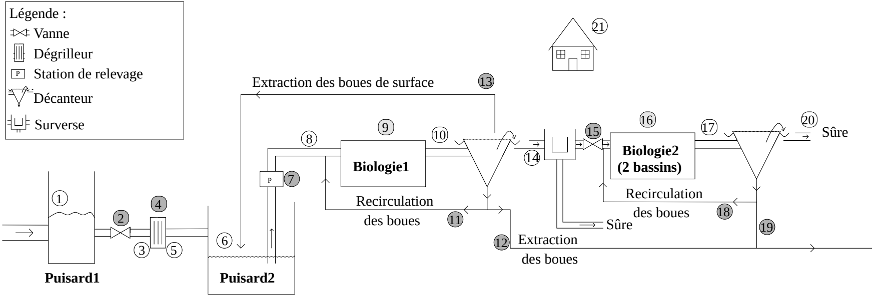
Comprendre les différents phénomènes de la partie biologique Déterminer des transformations non-linéaires

Utilisation du modèle ASM1
Simplification du modèle ASM1 au cas de la station

Technique de modélisation afin de déterminer les décalages temporels et transformations non-linéaires à prendre en compte

Application à la station de traitement des eaux usées

Description de la station



1 Hauteur du puisard1 : $H1$
2 % ouverture vanne : $V2$
3 Hauteur avant le dégrillage : $H3$
4 Cmd râteaux : $C4$
5 Hauteur après le dégrillage : $H5$
6 Hauteur du puisard2 : $H6$
7 Cmd pompes de relevages : $C7_1$ à $C7_6$
8 Débit : $Q8$
Température : $T8$
Conductivité : $\sigma 8$
Absorption UV : $A8$
Potentiel Hydrogène : $PH8$

9 Matière en suspension : $MES9$
Oxygène dissous : $OD9$
9 Cmd aérateur : $CA9$
Cmd floculation phosphate : $Fp9$

10 Ammonium : $NH10$
Nitrate : $NO10$
Orthophosphate : $OP10$

11 Cmd recirculation des boues : $C11$
12 Cmd extraction des boues : $C12$
13 Cmd boue de surface : $C13$

14 Hauteur avant la surverse : $H14$
15 % ouverture vanne : $V15$
16 Matière en suspension : $MES16_1$, $MES16_2$
Oxygène dissous : $OD16_1$ et $OD16_2$
16 Cmd aérateur : $CA16$
17 Ammonium : $NH17$
Nitrate : $NO17$
18 Cmd recirculation des boues : $C18$
19 Cmd extraction des boues : $C19$
20 Orthophosphate : $OP20$
21 Pluviomètre : $P21$

Application sur la partie hydraulique de la station de traitement des eaux usées

Décalages temporels à prendre en compte pour la partie hydraulique

Partie Hydraulique			
Variable	Décalages temporels	Transformation non linéaire	Décalages temporels
H3	q^0, q^{-1}		
H5	q^0		
H6	q^0		
Q8	$q^0,$	$\tanh((Q8-550)./150)$	q^{-1}
H14	q^0, q^{-1}		
$C7$	q^0		

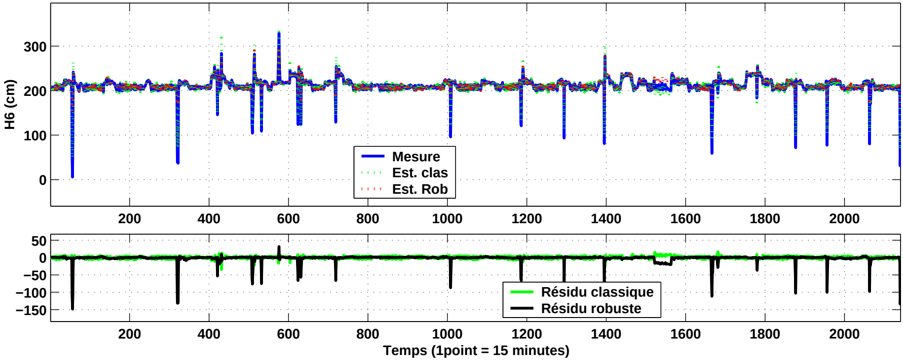
Matrice de données

$$x(k) = \begin{bmatrix} H3(k) & H5(k) & H6(k) & Q8(k) & H14(k) \\ \tanh((Q8(k-1)-550)./150) & H3(k-1) & H14(k-1) & C7(k) \end{bmatrix}^T$$

Application sur la partie hydraulique de la station de traitement des eaux usées

Construction du modèle

4 composantes principales robustes déterminées



Application sur la partie hydraulique de la station de traitement des eaux usées

Analyse des directions de reconstruction utiles dans l'espace global

Le nombre maximum de reconstructions est donc de 381 ($max(n-\ell,\ell)=5$)

K		R_1								
		1	2	3	4	5	6	7	8	9
R_2	1	0	0.85	0.97	1.00	1.00	1.00	0.97	0.99	1.00
	2		0	0.91	1.00	0.97	1.00	0.95	1.00	1.00
	3			0	1.00	0.99	1.00	0.99	1.00	1.00
	4				0	1.00	0.94	0.99	0.98	0.71
	5					0	0.90	0.89	0.90	0.98
	6						0	1.00	1.00	0.91
	7							0	0.80	0.99
	8								0	0.97

TABLE: indicateur K pour $r=1$

Le nombre de reconstructions nécessaires est de 202 au lieu de 381

Application sur la partie hydraulique de la station de traitement des eaux usées

Détection des défauts

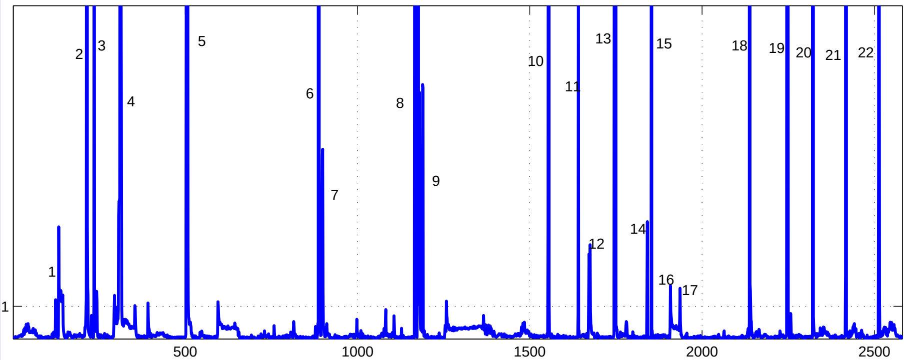


FIGURE: Détection des défauts avec la distance de Mahalanobis

Application sur la partie hydraulique de la station de traitement des eaux usées

Localisation des défauts

Indice des défauts	Direction de reconstruction qui minimise l'influence des défauts
16, 17	D_1
3, 10, 11, 14, 15, 19, 20, 21	D_3
7, 12	$D_{1,6}$
6, 8, 13, 18, 22	$D_{3,9}$
1, 2, 5	$D_{1,3,4}$
4	$D_{3,7,9}$
9	$D_{1,2,3,7}$

Application sur la partie hydraulique de la station de traitement des eaux usées

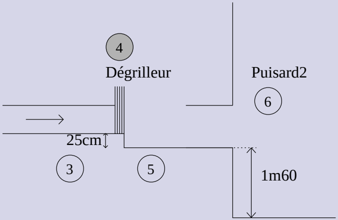
Analyse

Lorsque l'indice D_3 est calculé, les défauts 3, 10, 11, 14, 15, 19, 20 et 21 sont proches de zéro

- sur cette période, les valeurs du capteur $H6$ sont aberrantes.

Interprétation

Raison : Changement des relations entre les différentes mesures lorsque la valeur de $H6$ est inférieure à $1m85$



Détermination de défauts capteurs, ou d'un mode de fonctionnement très particulier

Conclusion & perspectives

Conclusion

- ACP robuste aux valeurs aberrantes sur l'ensemble des variables
- Utilisation de la distance de Mahalanobis pour détecter des défauts sur l'ensemble des variables
- Utilisation du principe de reconstruction afin de localiser les variables défaillantes
- Détermination des défauts multiples isolables
–> évite l'explosion combinatoire des scénarios de défauts
- Application à la partie hydraulique d'une station d'épuration

Perspectives

- Valider la méthode robuste MMRPCA sur un exemple de grandes dimensions
- Application à la partie biologique d'une station d'épuration
- Extension au changement de modes de fonctionnements