

Prédiction de la qualité du produit en fonction de la qualité des équipements:

— Application à la fabrication de semi-conducteurs

Réunion du GT S3 (Sûreté - Surveillance - Supervision)

Arts et Métiers ParisTech, 15 septembre 2016

Presentée par : Mariam MELHEM

Bouchra ANANOU, Mustapha OULADSINE, Jacques PINATON

Plan

1

Introduction

2

Approche I

3

Approche II

4

Conclusion

Plan

1

Introduction

2

Approche I

3

Approche II

4

Conclusion

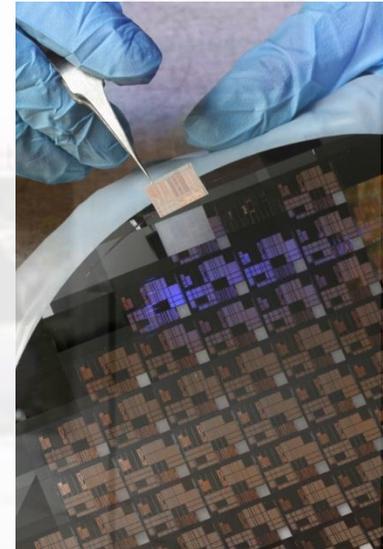
Description du problème industriel

❑ Industrie de semi-conducteurs:

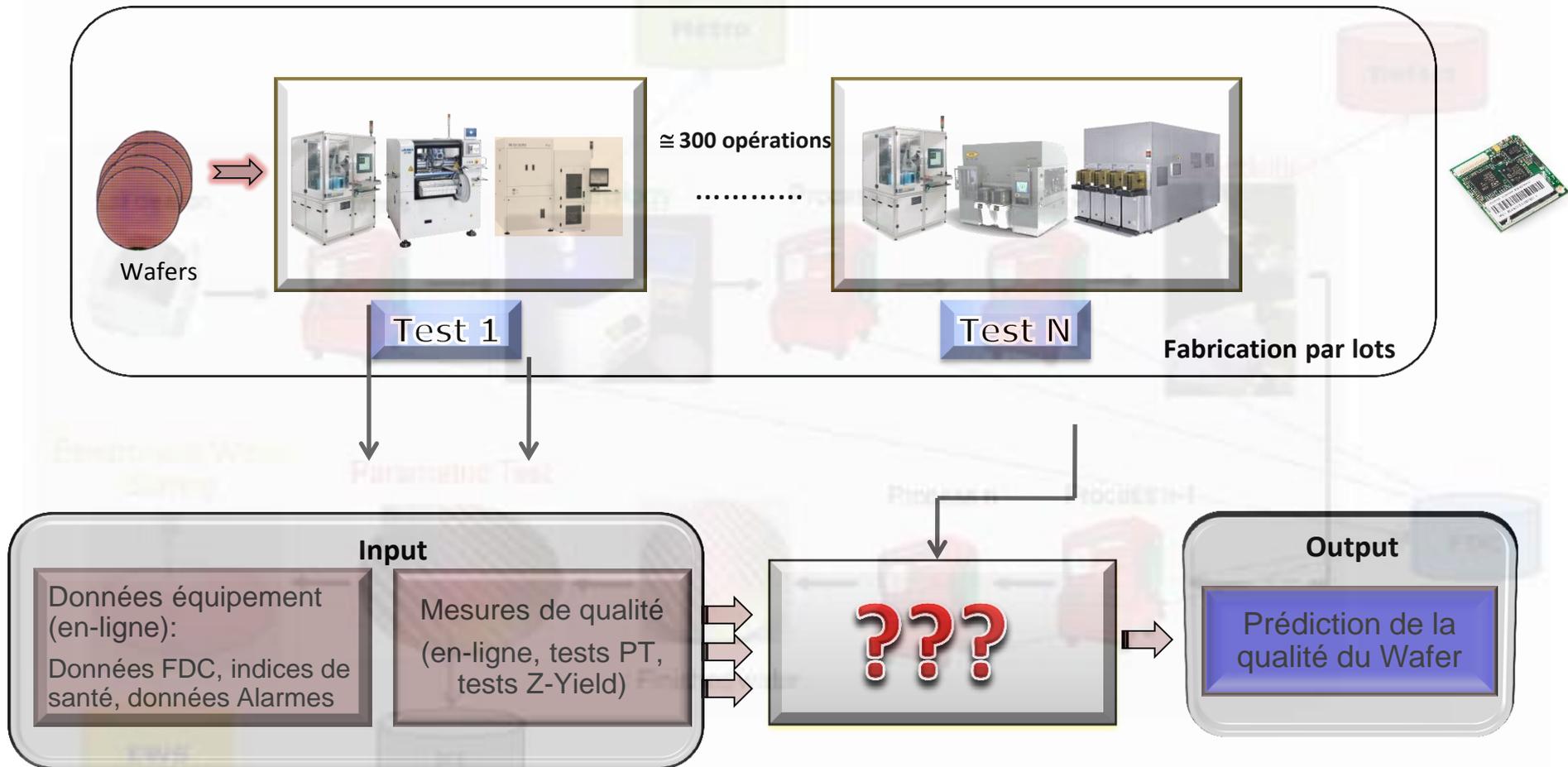
- Industrie qui s'intéresse à la fabrication de composants électroniques qui forment ensemble des circuits intégrés sur des « wafers »
- Processus de fabrication long, complexe et coûteux
- 300 opérations environs pendant 3 mois

❑ Défis

- Grand nombre des équipements contenant chacun plusieurs chambres
- Fabrication par lots
- Nombre important de variables
- Données très corrélées
- Données de test incomplètes (Tests discontinues: échantillonnage aux niveaux lots/wafers)
- La complexité de semi-conducteurs et de la performance de puces (nombre de transistors) double chaque année et demi d'après la loi de Moore



Description du problème industriel



Objectifs

❑ Diagnostic

- **Détection** : le défaut peut-être déterminée par perte de qualité ou une chute de rendement
- **Isolation** : Déterminer l'équipement responsable de défaillance
- **Identification de défauts** : Déterminer les causes du chute de rendement

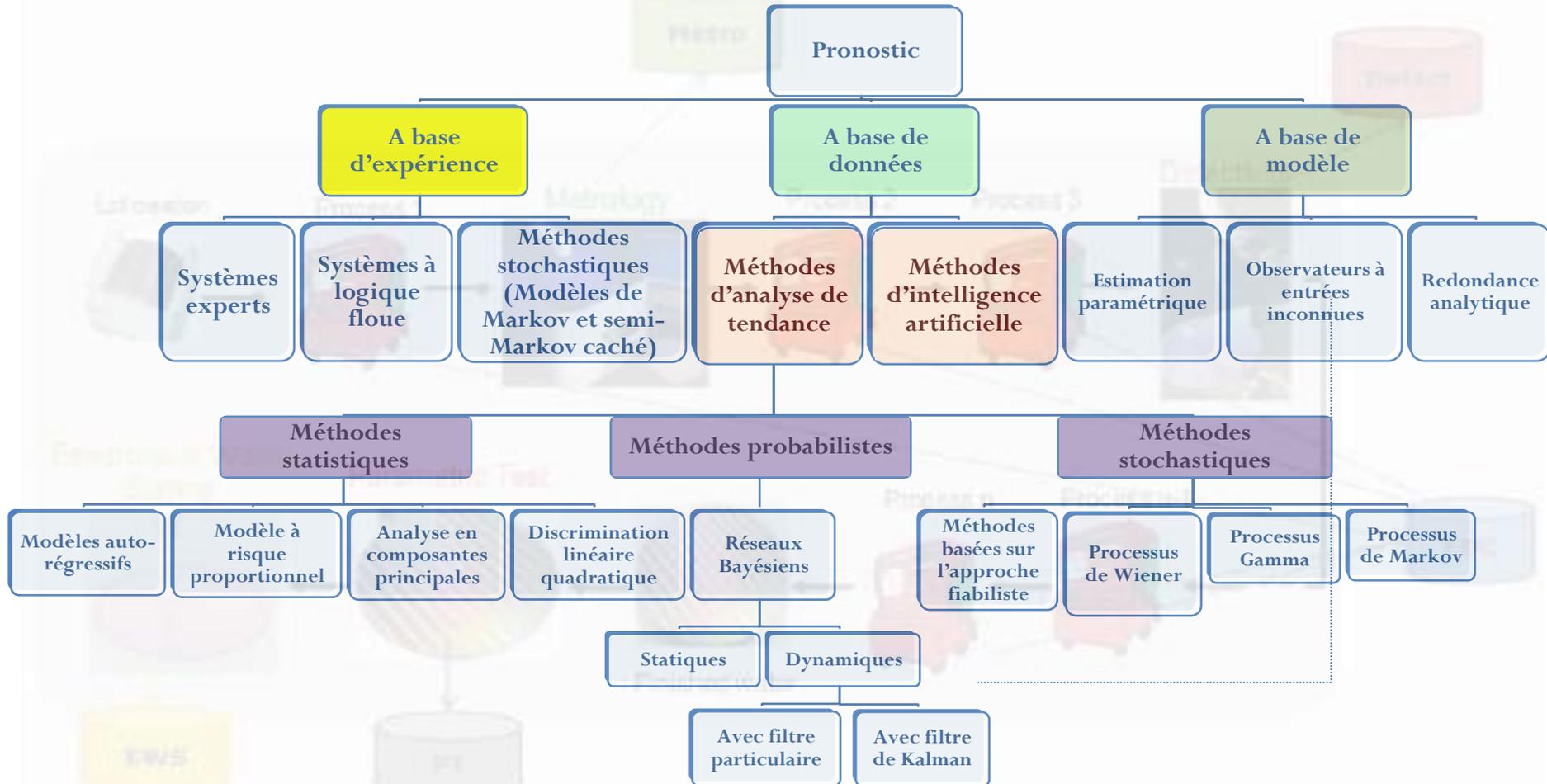
❑ Pronostic :

- **Prédiction de la qualité du produit** :
 - Mesures SPC
 - Scrap
 - Rendement de manufacturing (« Yield »)
- **Prédiction de RUL (Remaining Useful Life)** :
 - Nombre de wafers restant avant perte de rendement

❑ Approches

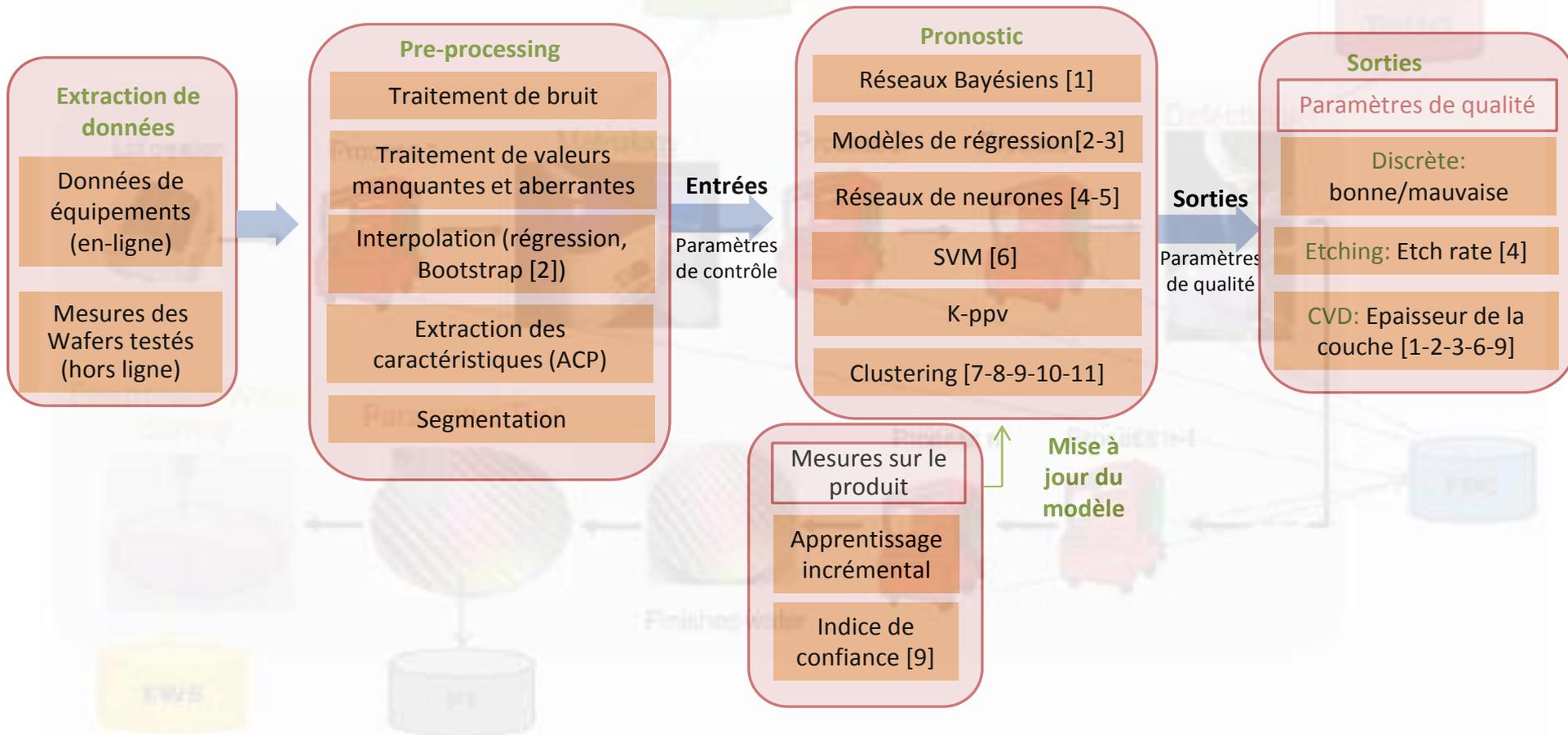
- **Approche I** : Calcul du nombre de wafers restant avant perte de spécifications des wafers (Remaining Useful Wafers RUW)
- **Approche II** : Prédiction de la qualité du Wafer en fonction de données FDC et de la qualité des étapes précédentes de production

Etat de l'art



Etat de l'art

Etat de l'art sur la prédiction de la qualité d'un produit pour un seul équipement :



Plan

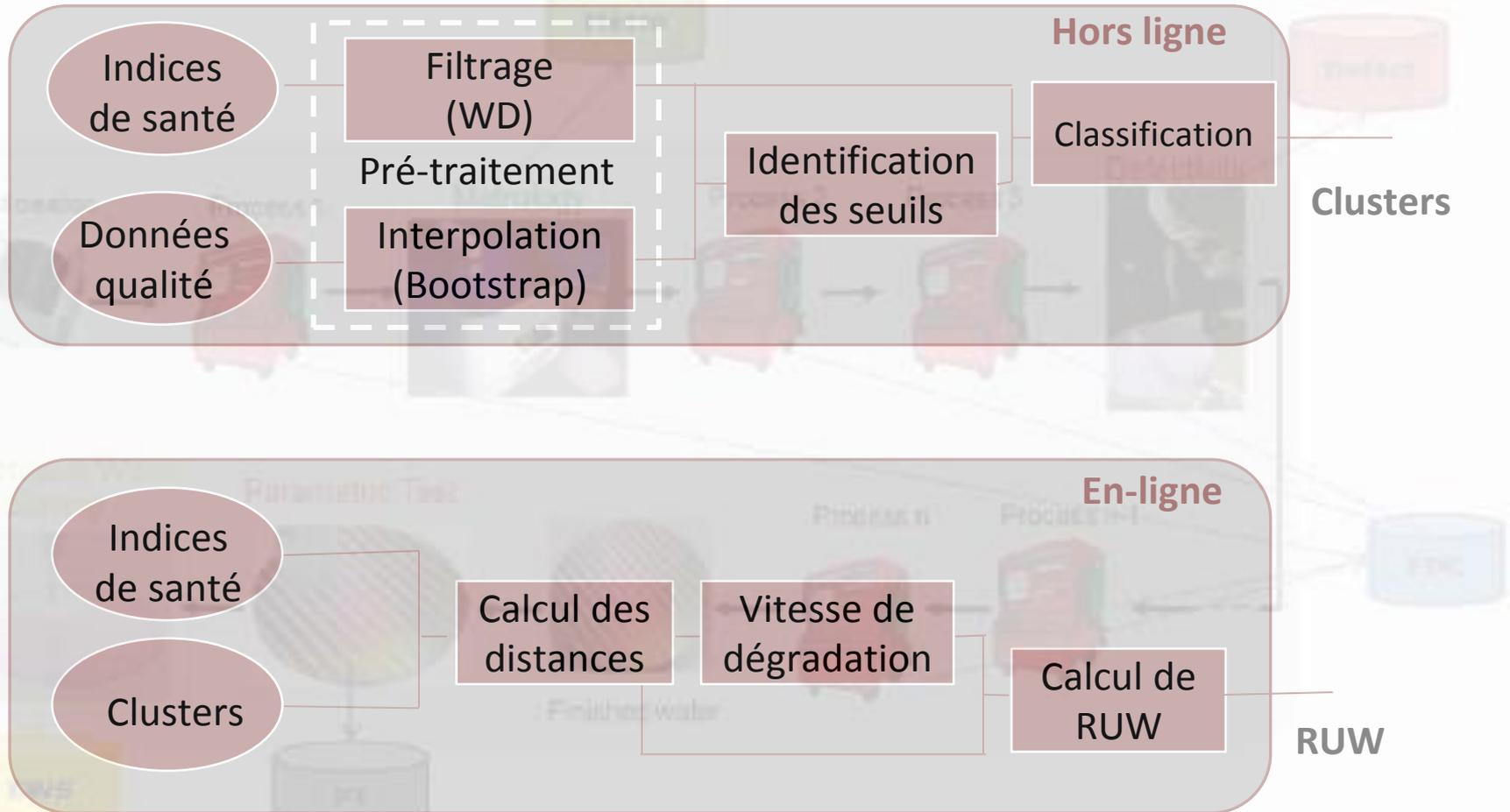
1 Introduction

2 Approche I

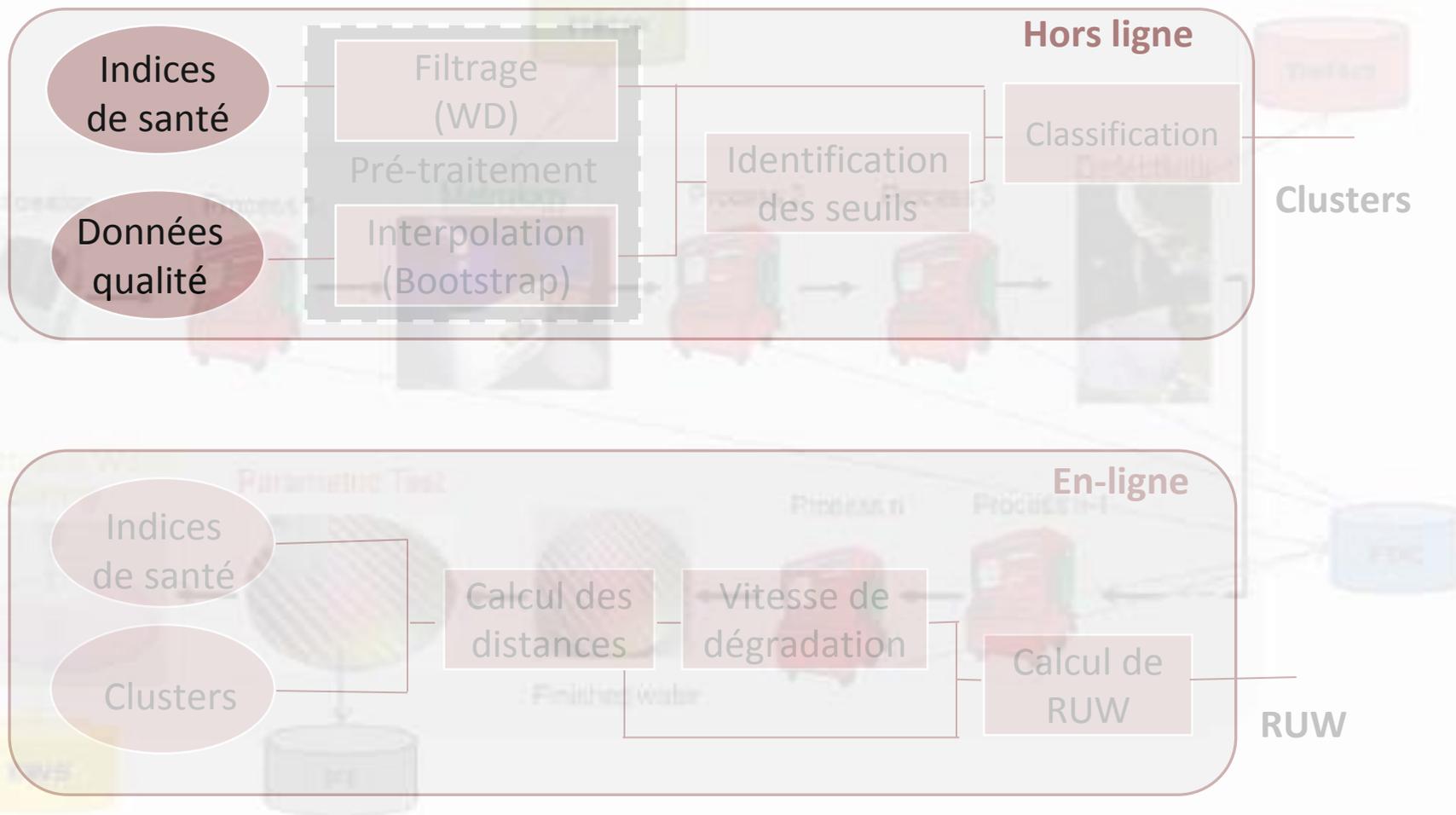
3 Approche II

4 Conclusion

Méthode proposée I



Description de données



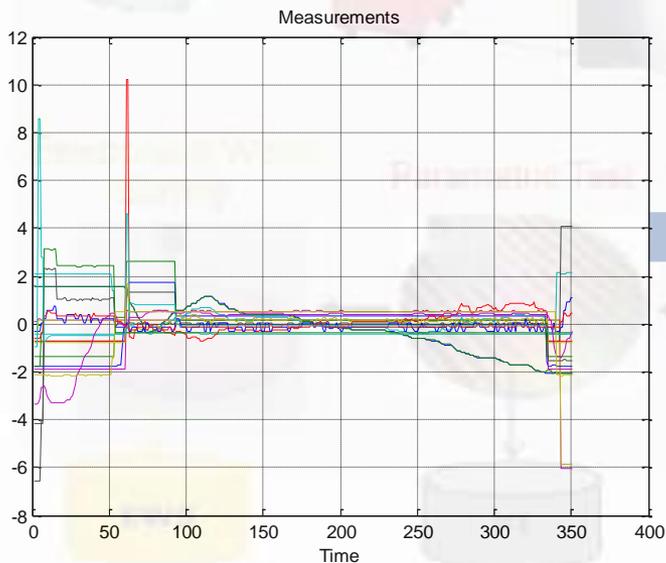
Description de données

❑ Description du système

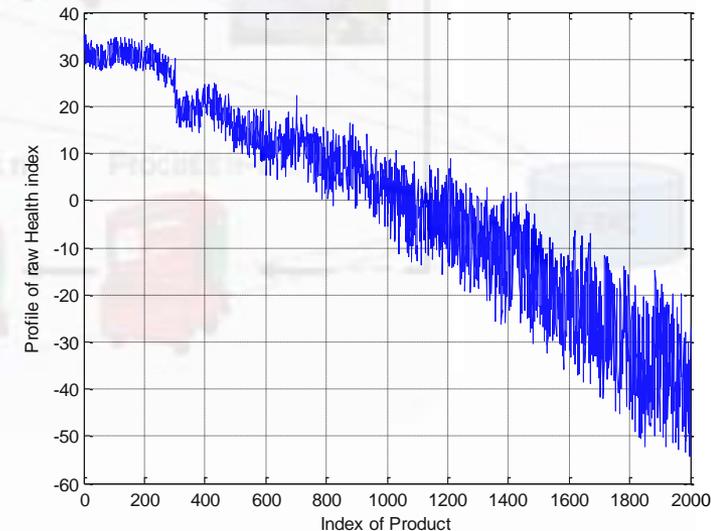
- Un ensemble de 3 équipements de dépôt (A,B,C) dont la fonction est de déposer 3 couches de substances de natures différentes, mais ayant la même Viscosité et la même densité volumique.
- Le wafer est maintenu en rotation constante identique sur les 3 équipements lors du dépôt.

❑ Instrumentation

- Indices de santé pour chaque machine (Significant points combined PCA (SP-PCA))



[SP-PCA]
[Nguyen et al.]



Description de données

❑ Données disponibles

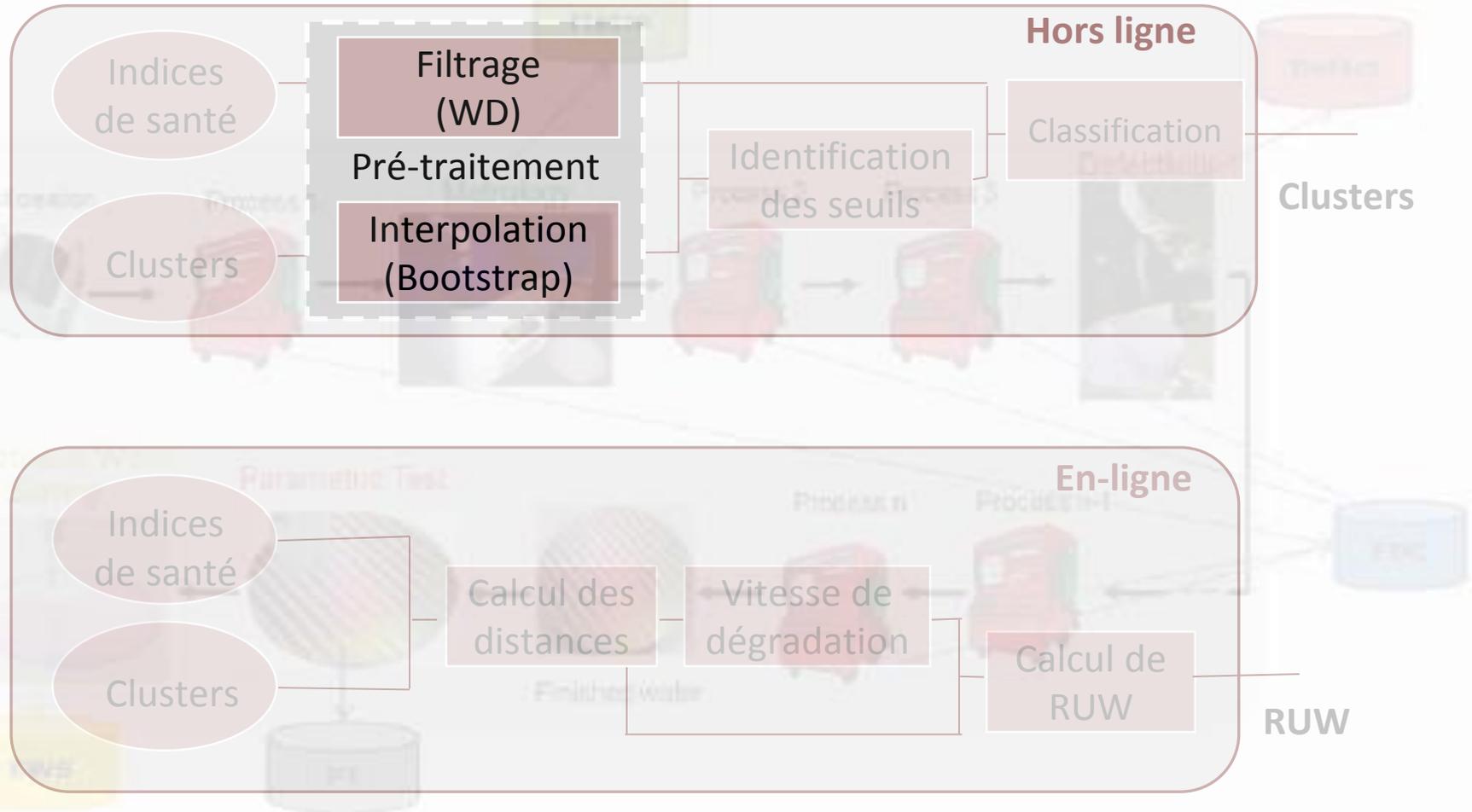
- 1 cycle complet de fonctionnement (du fonctionnement normal jusqu'à défailtant) pour chacun des 3 équipements

❑ Paramètres de qualité

- Les paramètres de qualité peuvent être des mesures électriques, chimiques et géométriques
- Paramètre géométrique choisi (épaisseur de la couche déposée pour chaque machine)

Equipement	Epaisseur de la couche	Tolérance
A	6mm	+ ou -10%
B	3mm	+ ou -5%
C	1,5mm	+ ou -5%
A+B+C	9,5mm	+ou- 10%

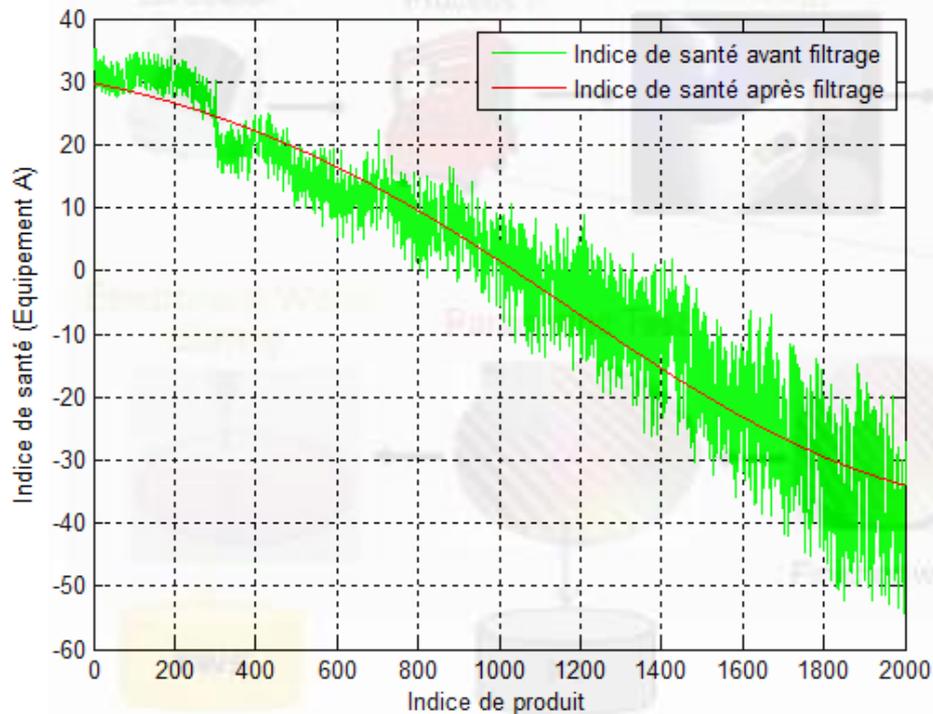
Pré-traitement



Pré-traitement

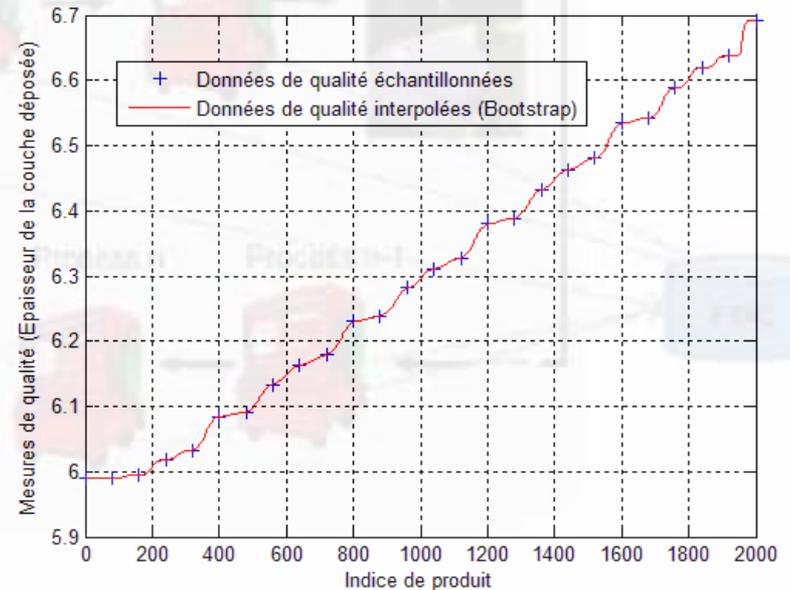
❑ Filtrage des indices de santé

- Décomposition en ondelettes de Daubechies

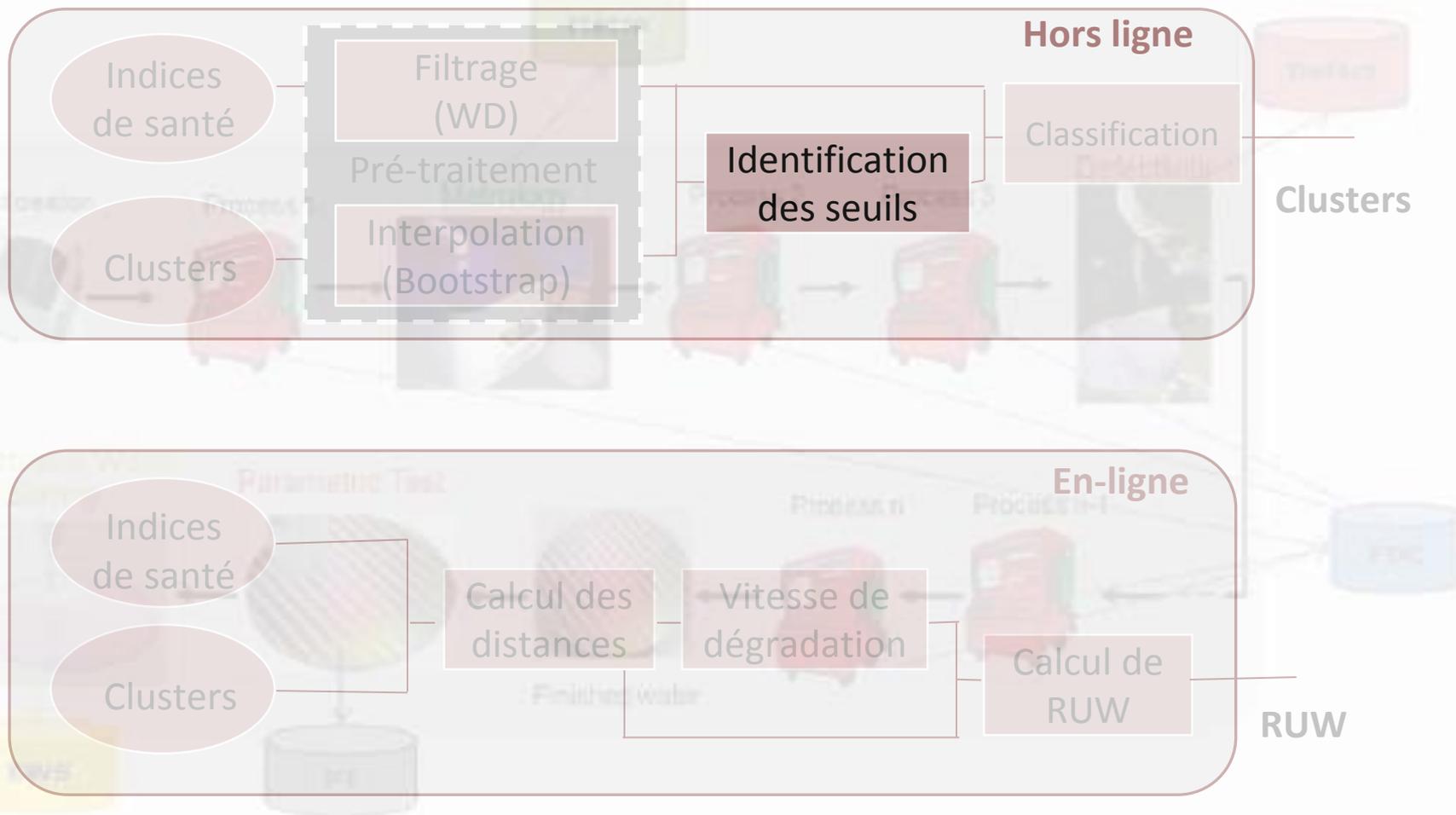


❑ Interpolation de données qualité

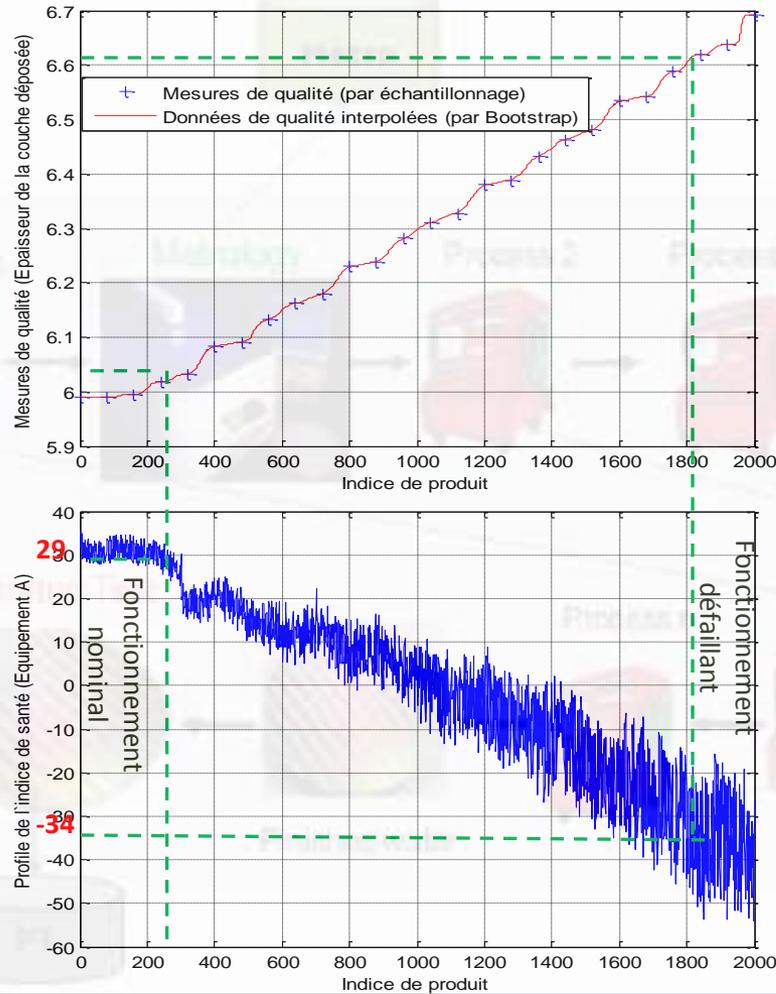
- 25/2000 de mesures
- Bootstrap : ré-échantillonnage
 - ✓ « M » échantillons avec remise
 - ✓ Arrangement en ordre croissant des échantillons
 - ✓ Moyenne à chaque point $Y^i = \frac{1}{M} \sum_{m=1}^M Y_m^i, i = 1, \dots, 2000$



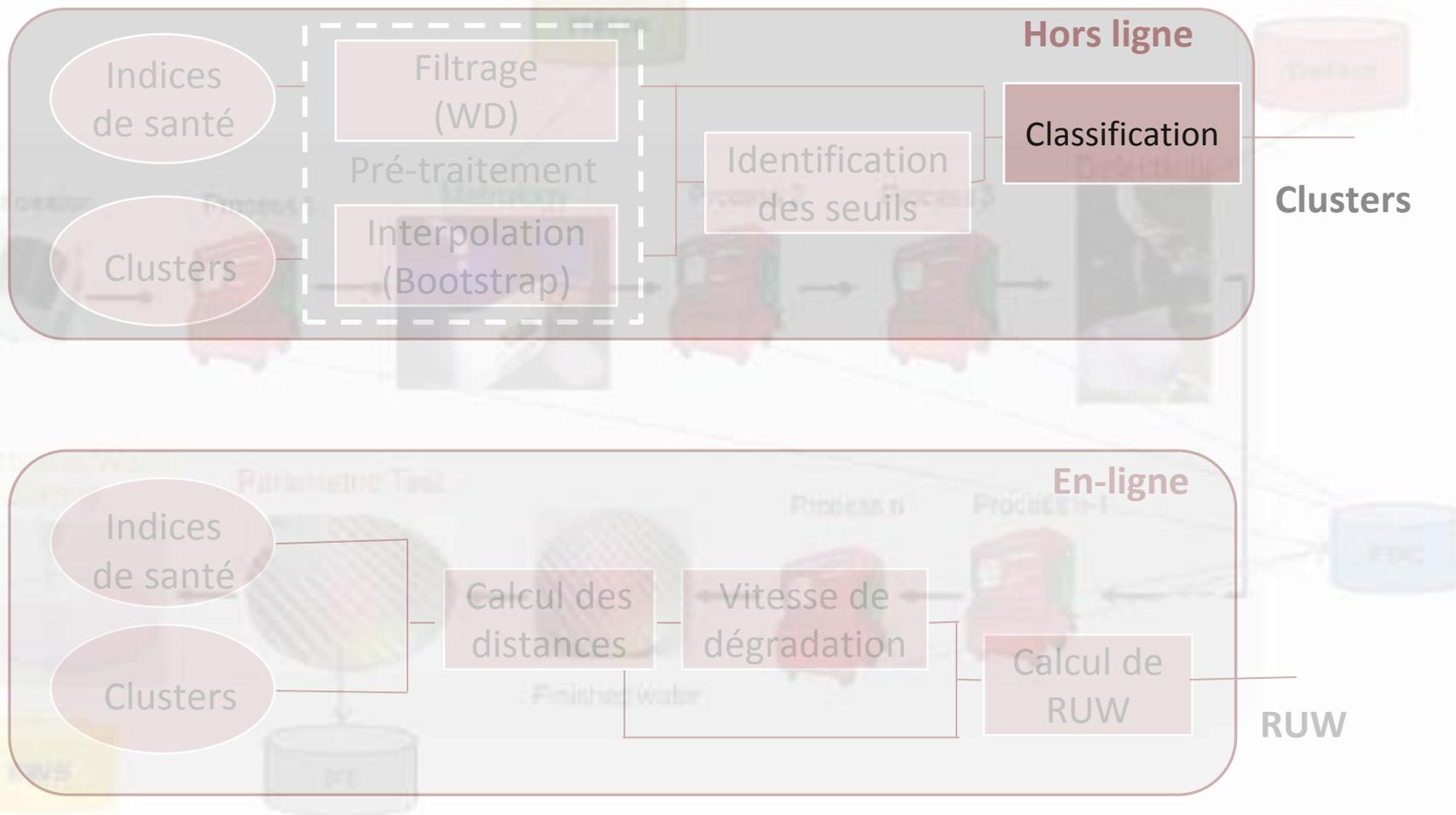
Identification des seuils



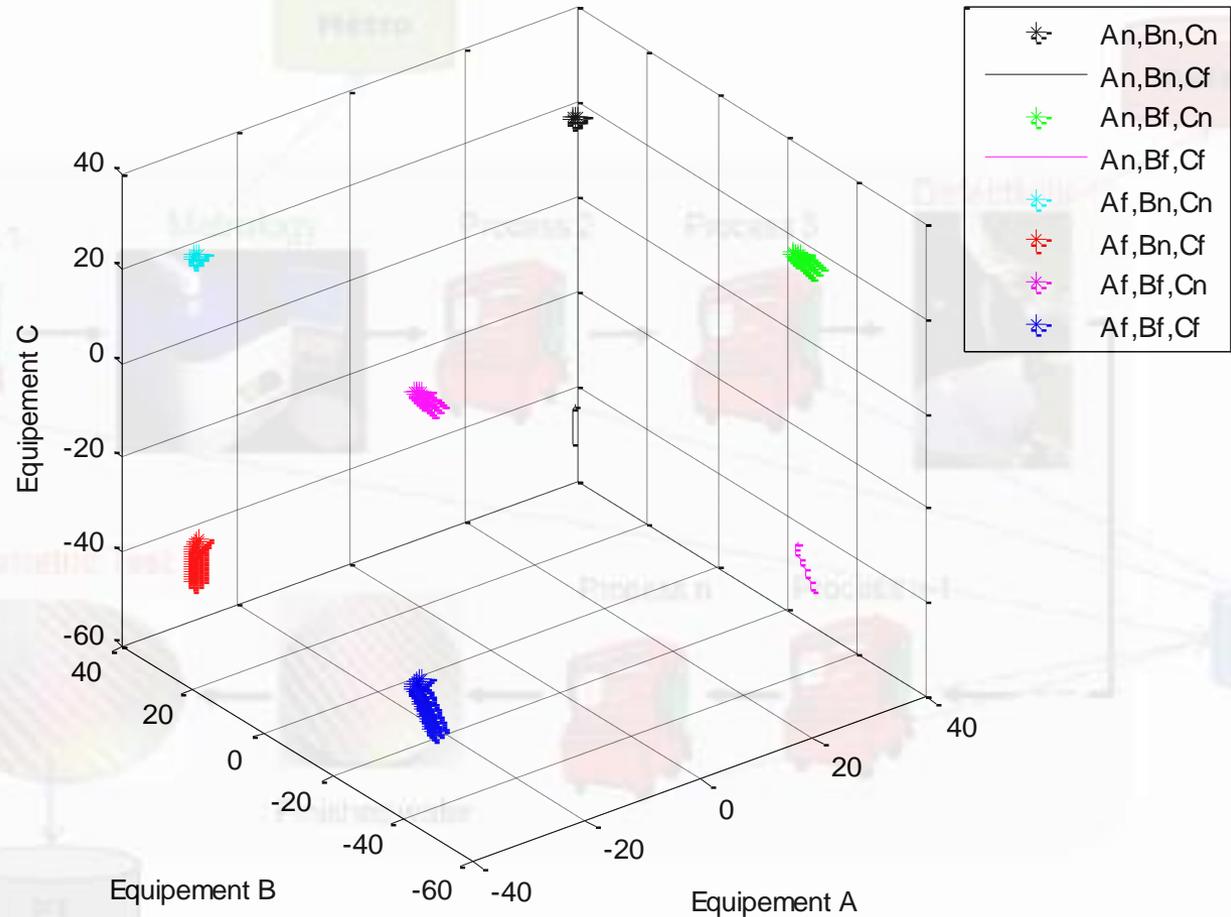
Identification des seuils



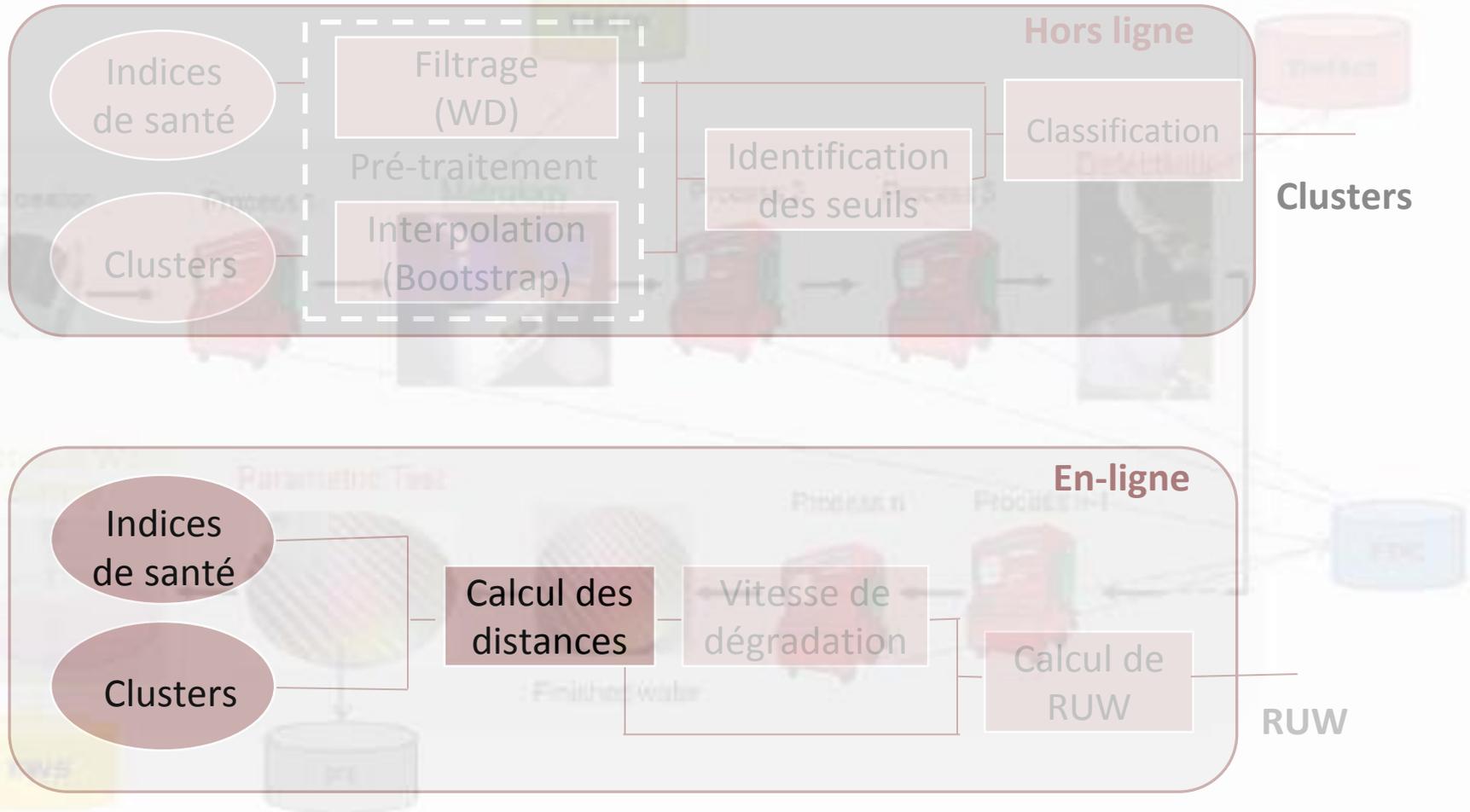
Classification



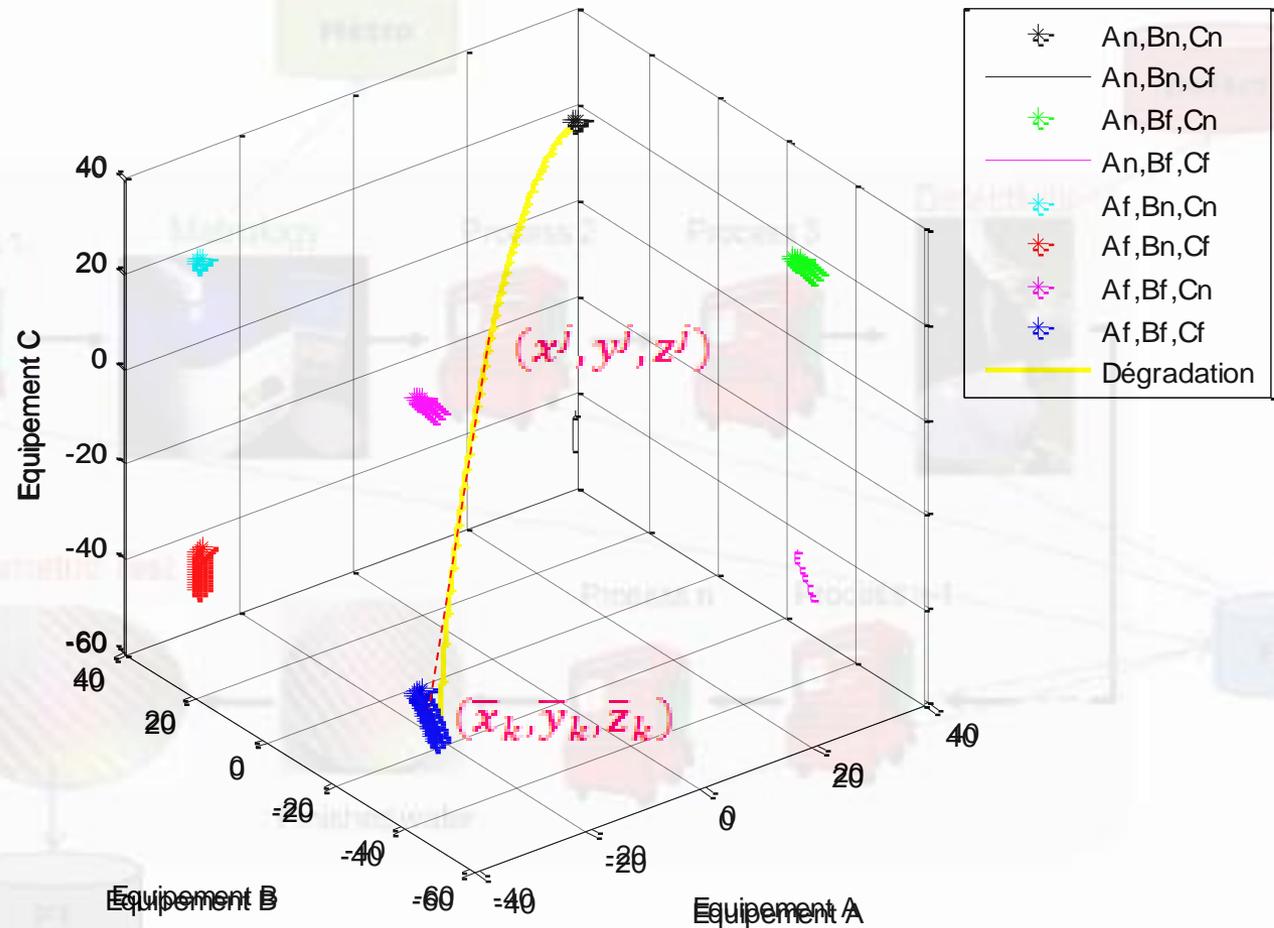
Classification



Calcul des distances



Calcul des distances



Calcul de distances

- Calculer les barycentres des clusters $\{C_k\}_{k=1,\dots,8}$

$$\bar{x}^k = \frac{\sum_{i=1}^{N_k} x_i^k}{N_k}, \bar{y}^k = \frac{\sum_{i=1}^{N_k} y_i^k}{N_k}, \bar{z}^k = \frac{\sum_{i=1}^{N_k} z_i^k}{N_k}$$

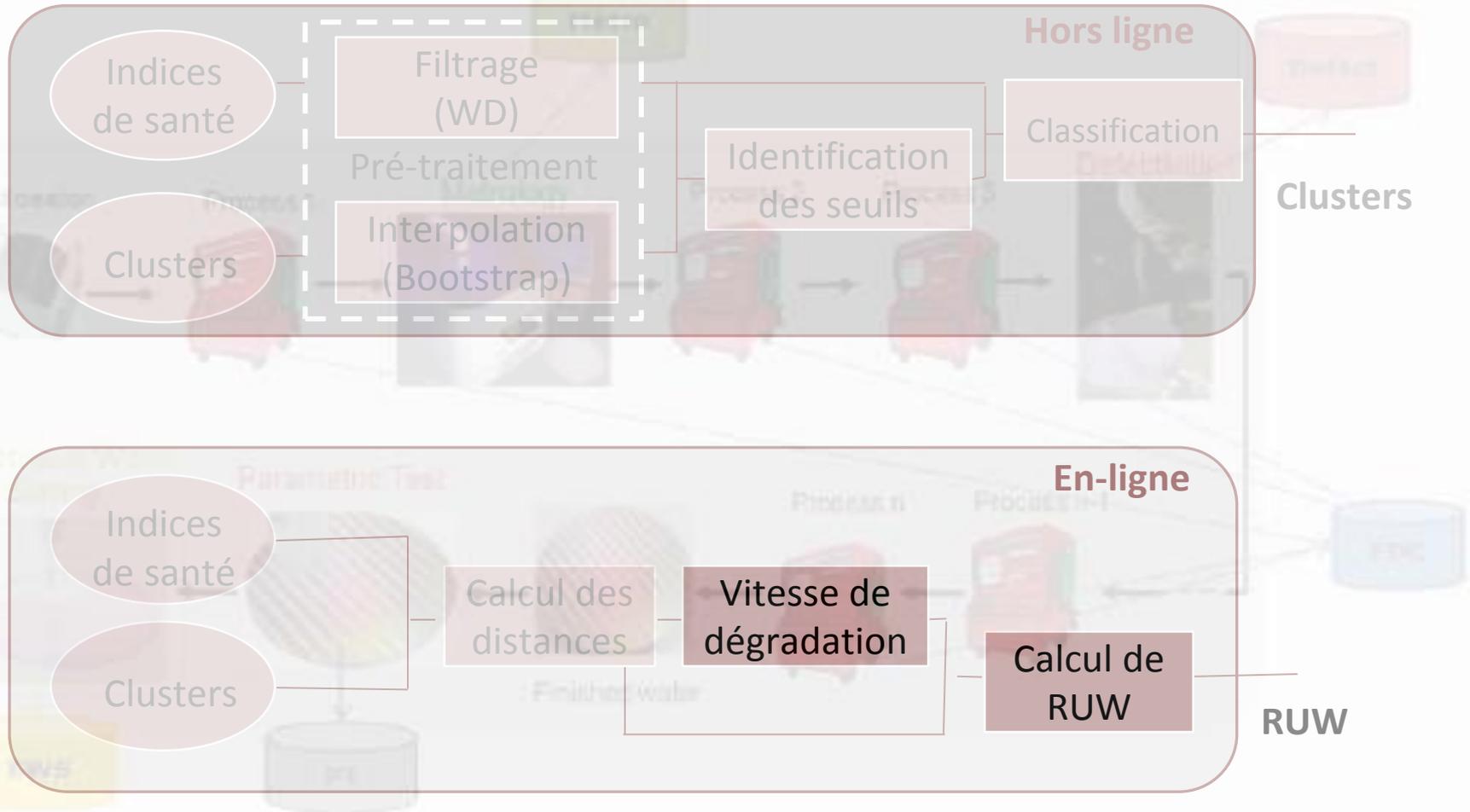
- Calculer la dispersion de chaque cluster

$$Disper_k = \sqrt{\frac{1}{N_k-1} \sum_{i=1}^{N_k} \sqrt{(x_i^k - \bar{x}^k)^2 + (y_i^k - \bar{y}^k)^2 + (z_i^k - \bar{z}^k)^2}}$$

- Calculer la distance entre chaque point de la trajectoire et chaque cluster C_k

$$D_j^k = \frac{\sqrt{(x_j^k - \bar{x}^k)^2 + (y_j^k - \bar{y}^k)^2 + (z_j^k - \bar{z}^k)^2}}{Disper_k}, j = 1, 2, \dots$$

Calcul de RUW



Calcul de RUW

- Calculer la vitesse de déplacement vers chaque cluster C_k

$$v_j^k = \frac{d}{dt} (D_j^k) = \frac{d}{dt} \left(\frac{\sqrt{(x_j^k - \bar{x}^k)^2 + (y_j^k - \bar{y}^k)^2 + (z_j^k - \bar{z}^k)^2}}{\text{Disper}_k} \right), \quad j = 1, 2, \dots$$

- Calculer le temps restant avant défaillance

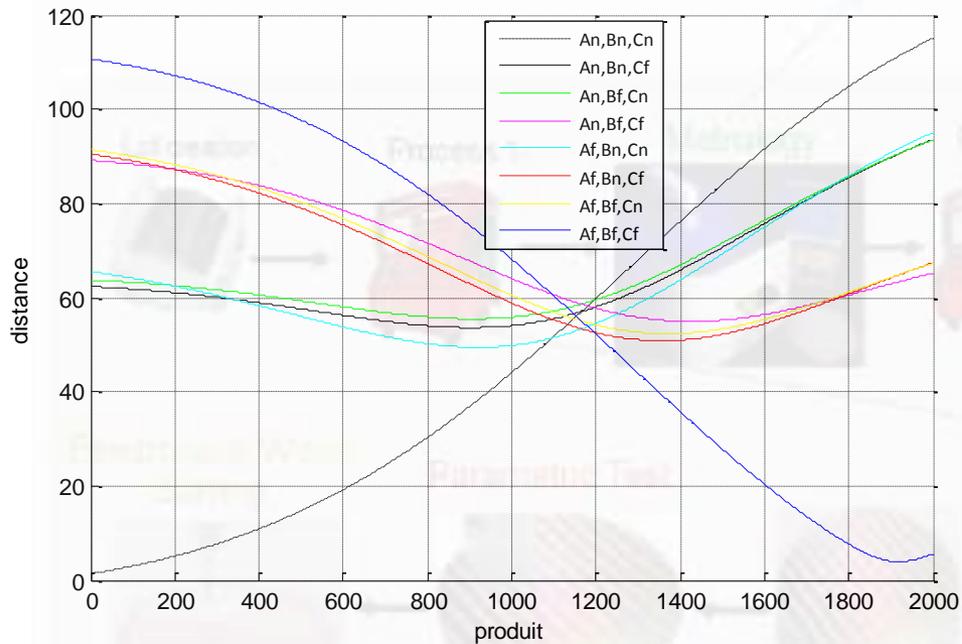
$$RUW_j^k = \frac{D_j^k}{v_j^k}, \quad k = 2, \dots, K, \quad j = 1, 2, \dots$$

$$RUW_j = \min_{k=2 \dots K} \{RUW_j^k\} \quad j=1, 2, \dots$$

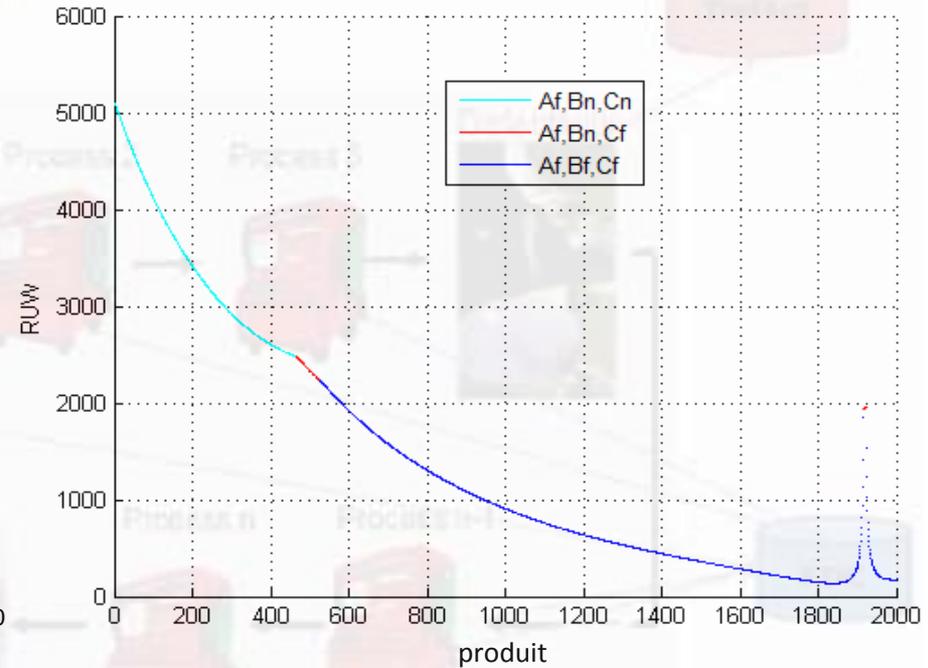
- Déterminer la trajectoire de défaillance

$$C_j = \underset{k=2, \dots, K}{\operatorname{argmin}} RUW_j^k$$

Calcul de RUW



Evolution de la distance avec les différents modes de fonctionnement



Evolution de RUW (Remaining Useful Wafers)

Plan

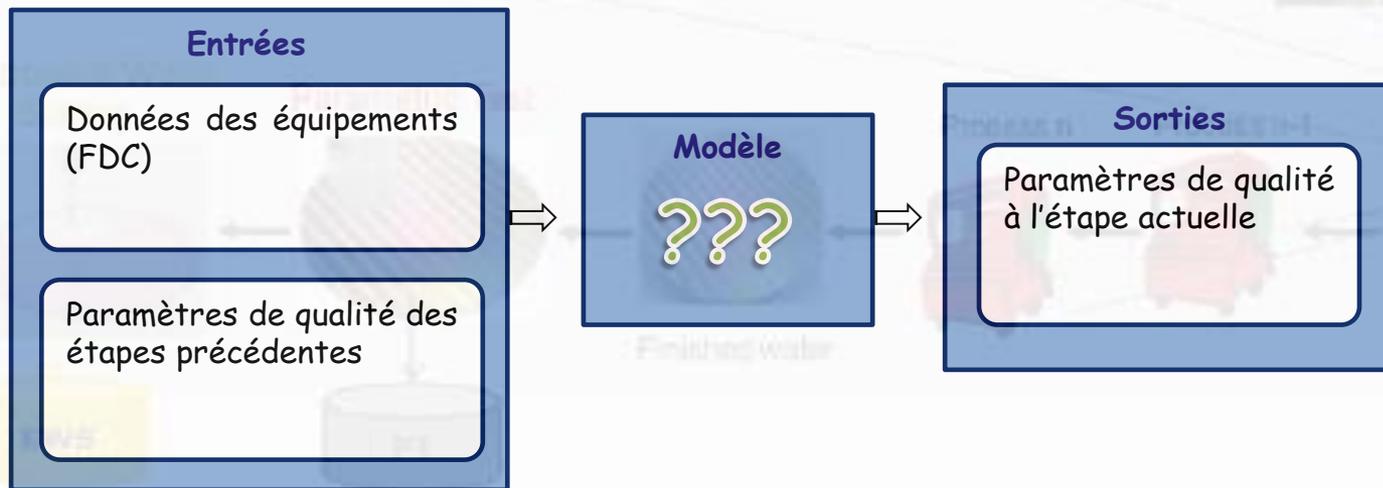
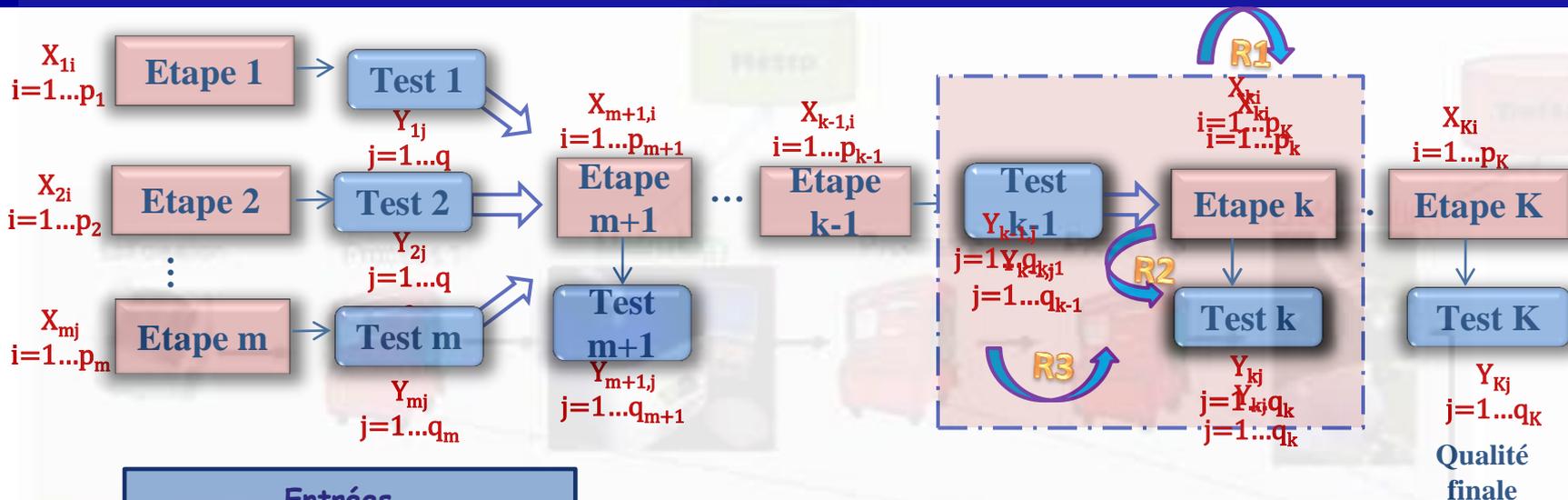
1 Introduction

2 Approche I

3 Approche II

4 Conclusion

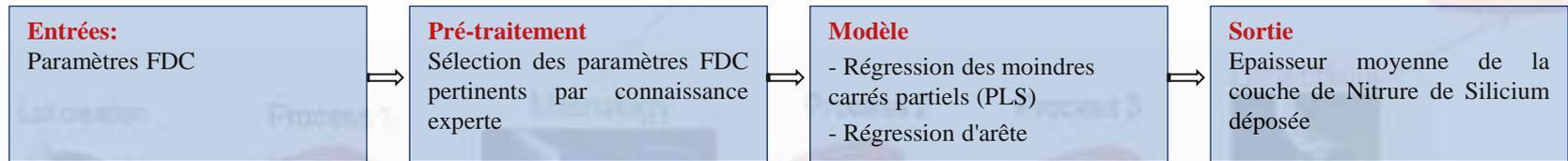
Modèle



Etat de l'art

Prédiction de la qualité dans l'industrie multi-étapes

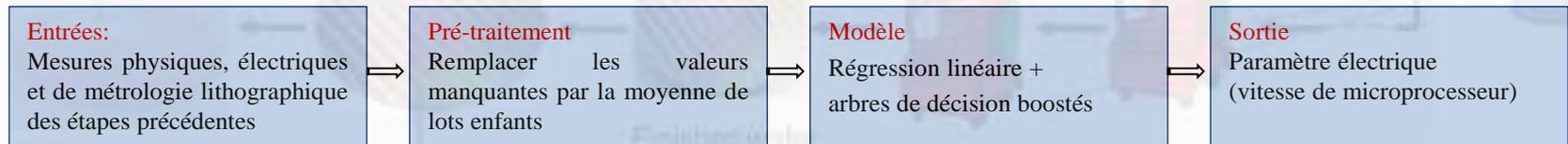
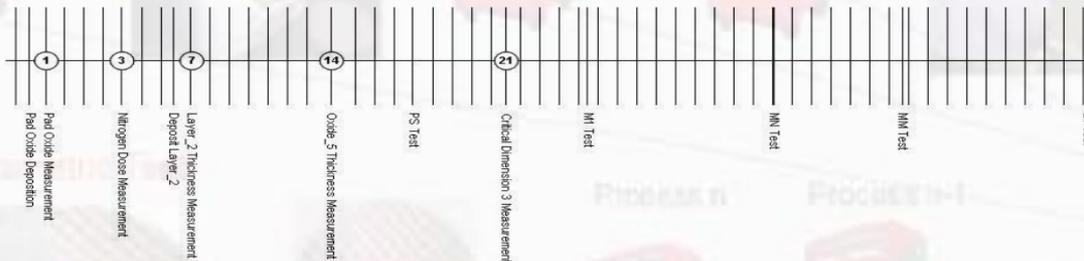
[H. Purwins et al. (2011)]



(+) Le PLS et la régression d'arête sont plus performants que la régression de moindre carré ordinaire

(-) Considère une seule étape de fabrication

[S. M. Weiss et al.]



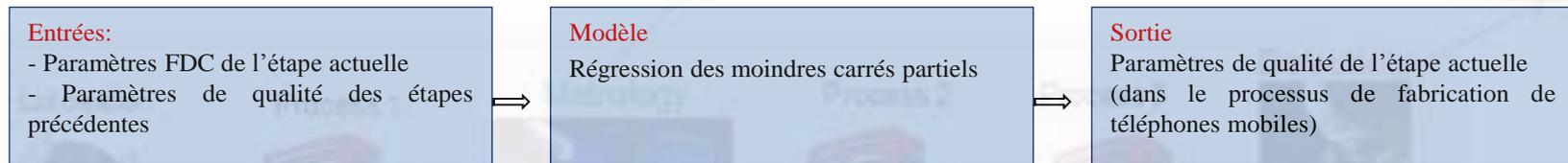
(+) Prend en considération la relation entre étapes de fabrication

(-) Ne prend pas en compte les données des équipements en construisant le modèle de prédiction

Etat de l'art

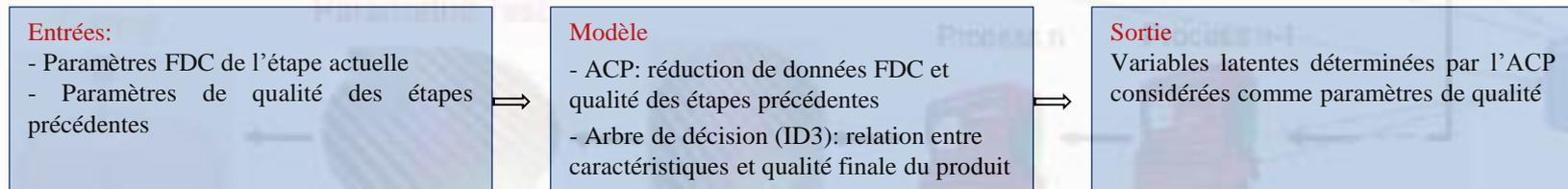
❑ Prédiction de la qualité dans l'industrie multi-étapes

[M. M Movahedi (2011)]



- (+) PLS performant pour la grande quantité de données
- (+) Prise en compte de corrélations entre les étapes
- (-) Mesures de qualité non disponibles dans le processus de fabrication de semi-conducteurs

[F. Arif et al. (2013)]



- (+) Traite le problème de manque de mesures de qualité
- (-) Les variables déterminées par l'ACP n'ont aucune signification physique

Etude théorique

$$X_k = [X_{1k} \dots X_{nk}]^T = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{N1} & \dots & X_{Np} \end{pmatrix}$$

$Y_k = [Y_{1k} \dots Y_{Nk}]^T$ où on dispose seulement de $n \ll N$ mesures

Résoudre l'équation:

$$Y_k = X_k A_k + Y_{k-1} B_{k-1} + \varepsilon_k = W_k \Theta_k + \varepsilon_k$$

où $W_k = [X_k, Y_{k-1}]$ et $\Theta_k = [A_k, B_{k-1}]$

Conditions:

- $p \gg n$
- Grand nombre de variables de processus
- Fortement corrélées
- Petit nombre d'observations (mesures)
- Un seul paramètre de qualité ($X_k \gg Y_{k-1}$)

Moindre carré multiple ordinaire

Minimisation de l'erreur de moindre carré

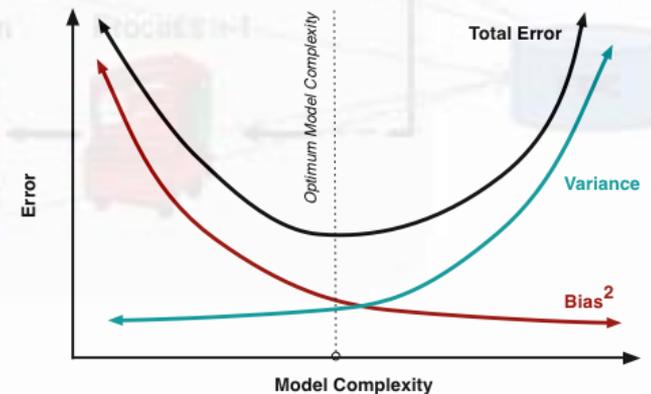
$$E_k = \min_{\Theta_k} \|Y_k - W_k \Theta_k\|^2$$

$$\frac{dE_k}{d\Theta_k} = 0 \Rightarrow \widehat{\Theta}_k = (W_k^T W_k)^{-1} W_k^T Y_k$$

Avantage: Meilleur estimateur non biaisé

Limite: Cas où les variables sont nombreuses et corrélées
 \Rightarrow Sur-apprentissage, $(W^T W)$ devient singulière

Il faut alors chercher un estimateur qui fait un compromis entre biais et variance



Etude théorique

➤ Régression sur composantes principales (*Least Square Error*)

- i. Décomposer W_k en valeurs singulières pour obtenir les composantes principales

$$W_k = T_k P_k^T$$

- ii. Puis faire une régression à partir de composantes principales

$$Y_k = T_k C_k = W_k (P_k C_k) = W_k \Theta_k$$

Avantage: Réduction du nombre de prédicteurs avec le minimum de perte d'information

➤ Régression des moindres carrés partiels (*Partial Least Square Error*)

- i. Décomposer W et Y comme en composantes principales

$$W_k = T_k P_k^T \text{ et } X_k = T_k P_k^T$$

- ii. Puis faire une régression entre T et Y

Avantage: Réduction du nombre de prédicteurs en prenant en compte la corrélation entre prédicteurs et réponse avec le minimum de perte d'information

Etude théorique

Least Absolute Shrinkage and Selection Operator (LASSO)

- Régression linéaire régularisée qui ajoute une pénalisation l'erreur de moindre carré de type L1

$$E_k = \min_{\theta_k} \|Y_K - W_K \theta_K\| + \lambda_k \|\theta_k\|_1$$



$$E_k = \min_{\theta_k} \|Y_K - W_K \theta_K\| \text{ avec } \|\theta_k\|_1 < c$$

Avantage: Sélection de variables par shrinkage de beaucoup de paramètres

Caractéristiques: Renforcer un grand nombre de coefficients à être nuls

Régression d'arête (Ridge Regression)

- Régression linéaire régularisée
- Ajouter une pénalisation l'erreur de moindre carré de type L2

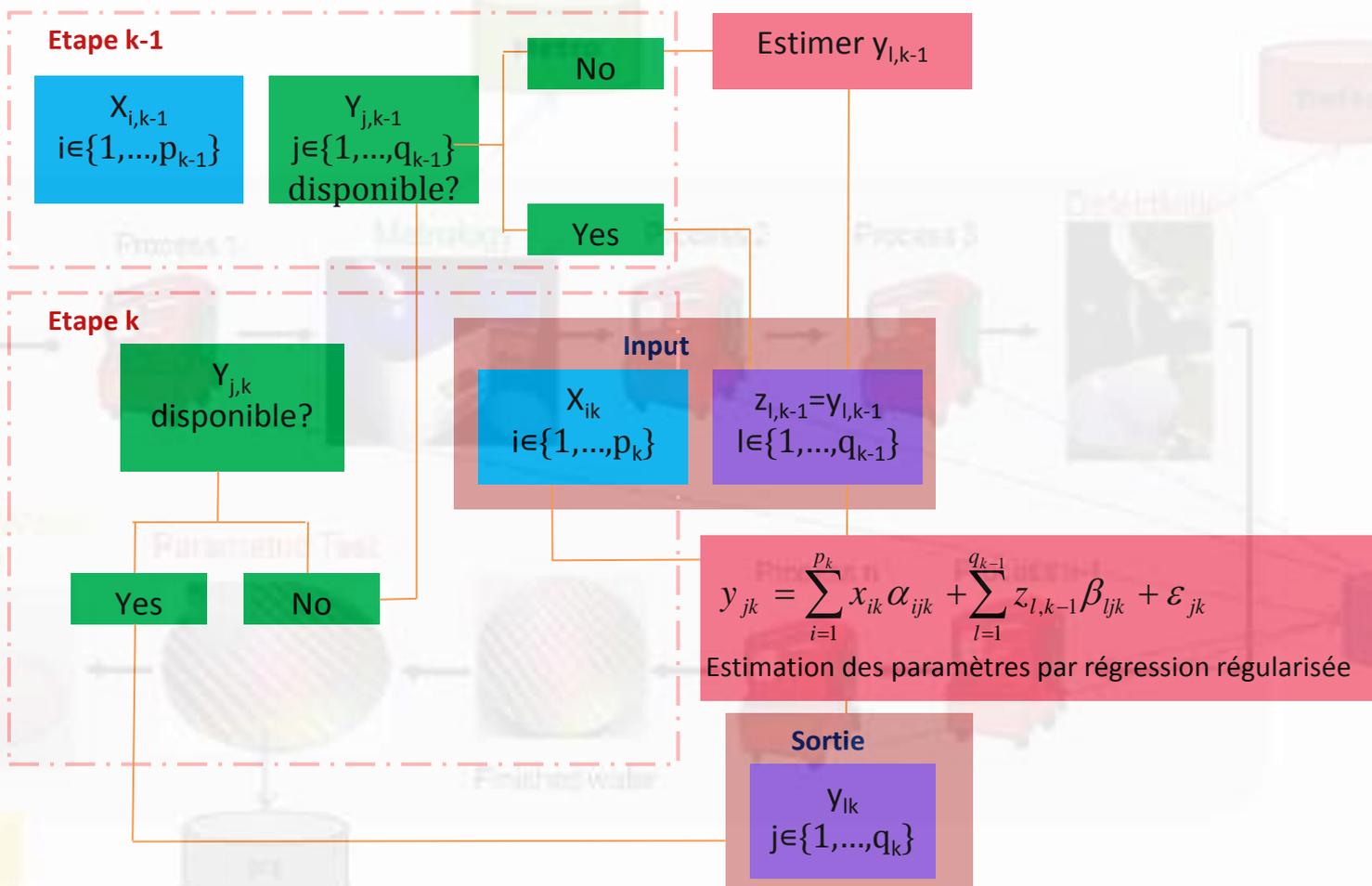
$$E_k = \min_{\theta_k} \|Y_K - W_K \theta_K\| + r_k \|\theta_k\|_2^2$$

$$\frac{dE_k}{d\theta_k} = 0 \Rightarrow \widehat{\theta}_k = (W_k^T W_k + r_k I_{(p+1) \times (p+1)})^{-1} W_k^T Y_k$$

Avantage: Résoudre le problème de corrélation entre variables

Caractéristiques: Diminuer l'effet de certains variables mais sans être nul

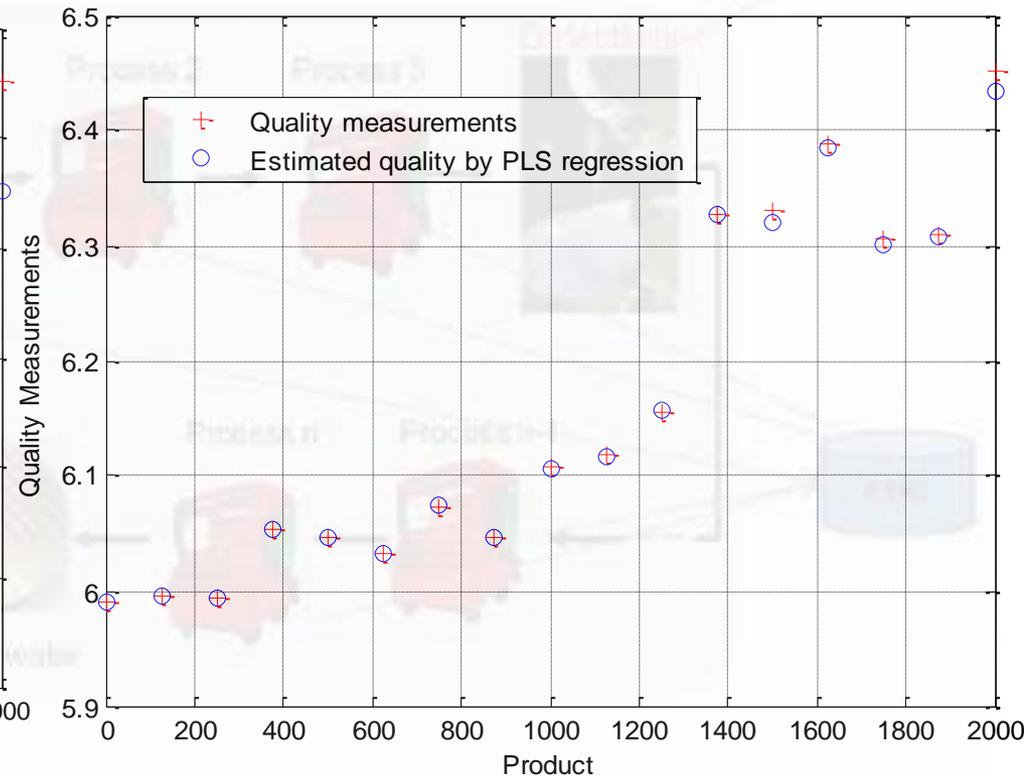
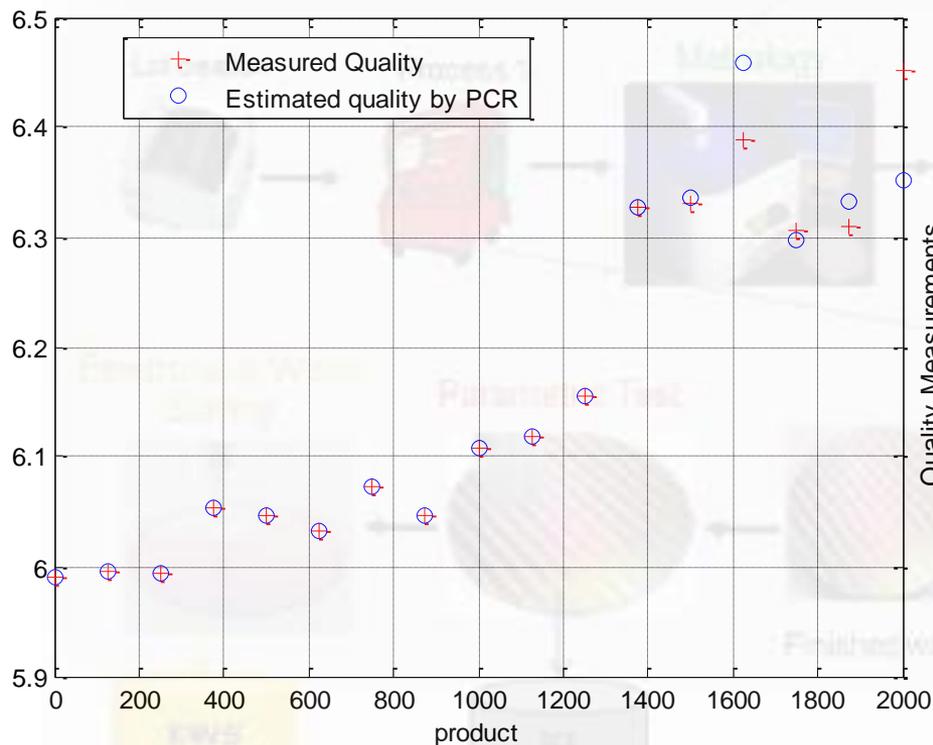
Modèle



Application (sur un seul équipement)

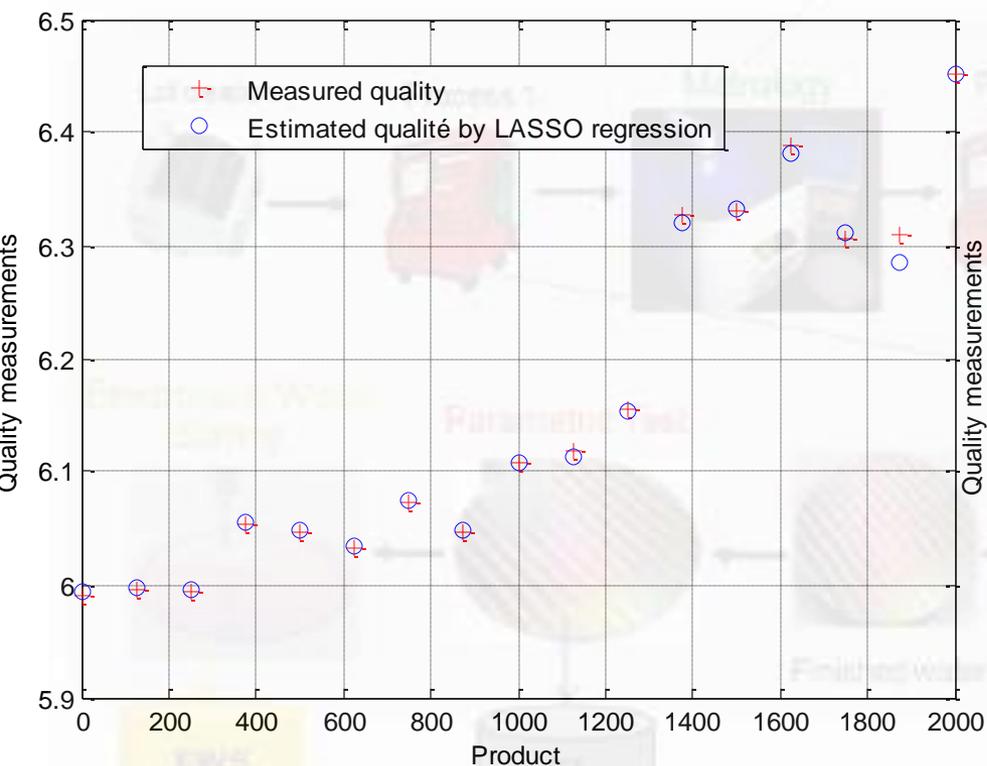
➤ Régression à composantes principales

➤ Régression de moindres carrés ordinaire

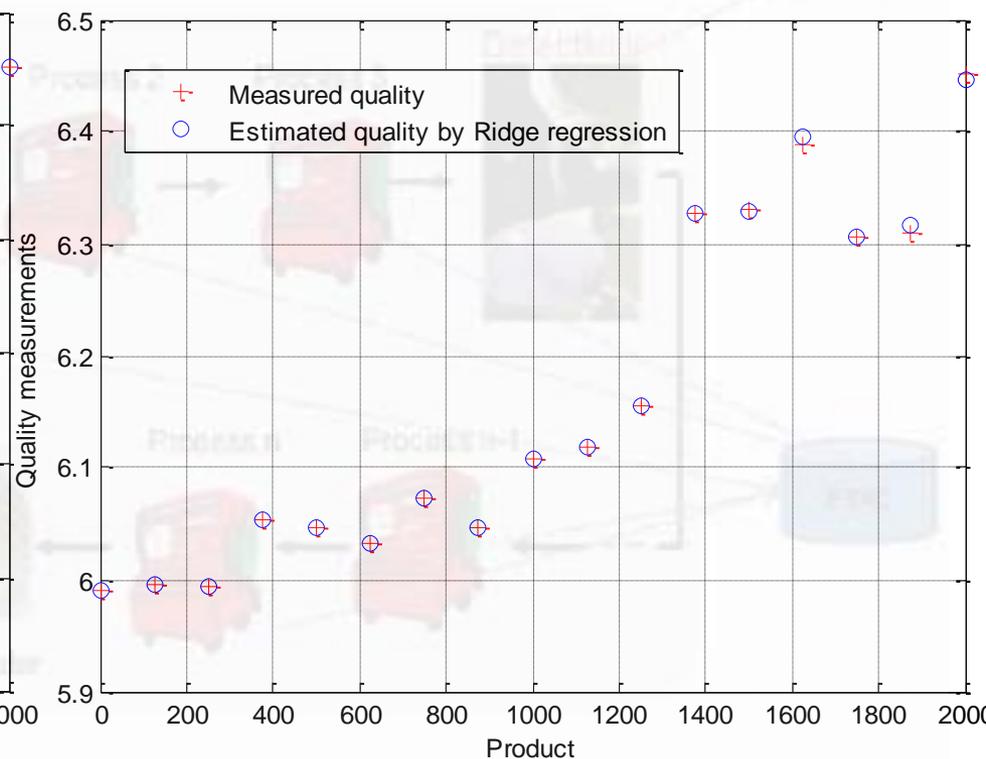


Application (sur un seul équipement)

➤ Régression LASSO



➤ Régression d'arête



Application (sur un seul équipement)

Méthode	Erreur d'apprentissage		Erreur de test	
	Moyenne	Ecart-type	Moyenne	Ecart-type
PCR	2.3124×10^{-15}	6.8195×10^{-16}	0.0158	0.0100
PLS	5.7477×10^{-4}	6.2880×10^{-4}	0.0071	6.9504×10^{-4}
LASSO	0.0043	0.0012	0.0142	0.0043
Régression d'arête	2.7263×10^{-6}	1.3649×10^{-6}	0.0051	0.0022

Moyenne et écart-type de l'erreur quadratique des différentes méthodes de régression en réalisant 5 initialisations aléatoires

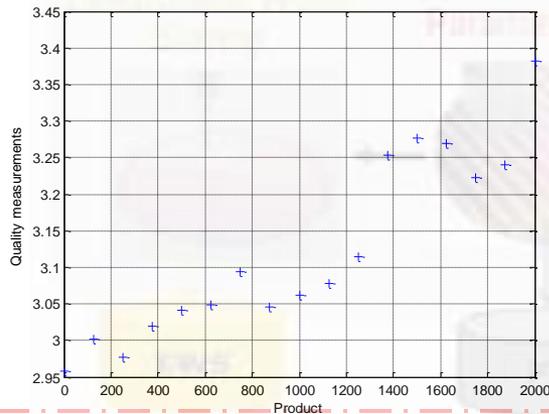
Application (sur 3 équipements)

Description de données

- Données des 3 équipements pour 2000 wafers
- Données FDC (6 variables, 351 observations)
- 1 paramètre de qualité

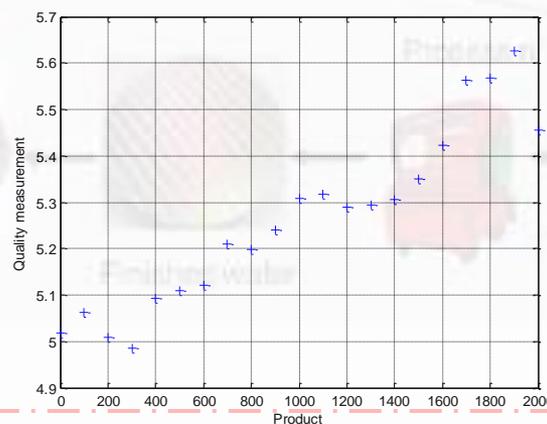
Equipement A

	1	351×6
1			
⋮			
⋮			
2000			



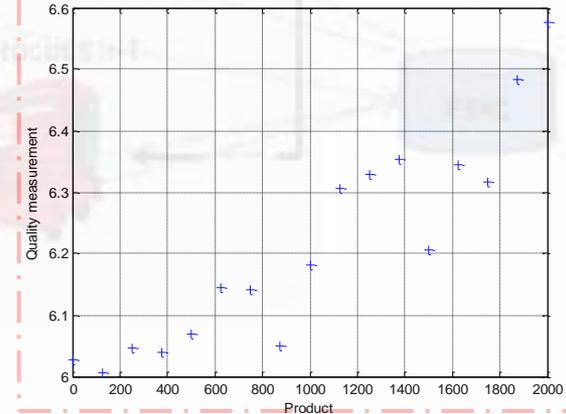
Equipement B

	1	351×6
1			
⋮			
⋮			
2000			



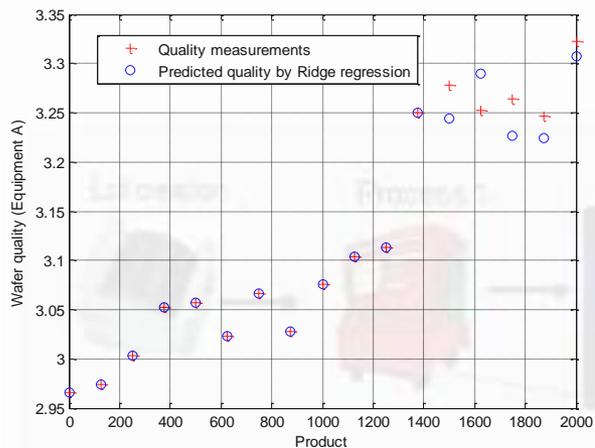
Equipement C

	1	351×6
1			
⋮			
⋮			
2000			

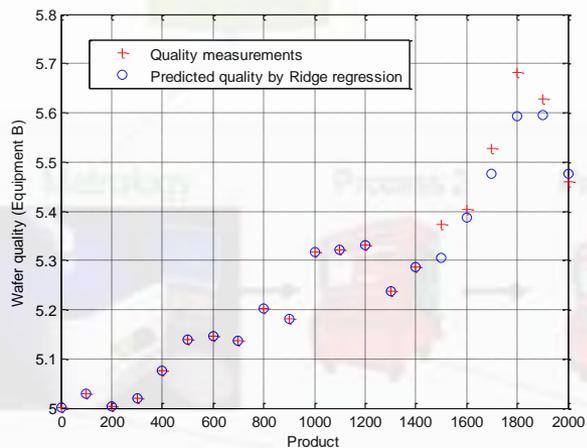


Application (sur 3 équipements)

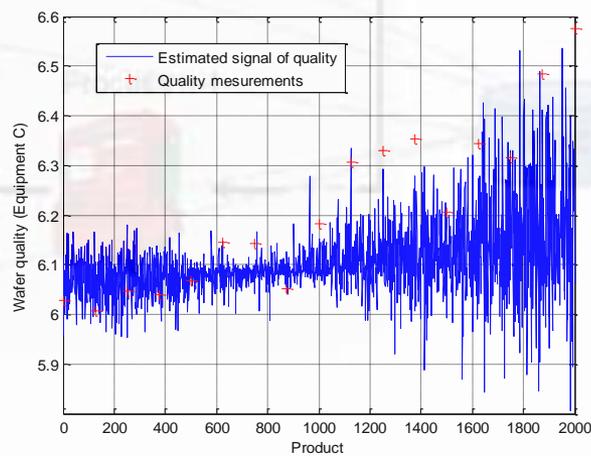
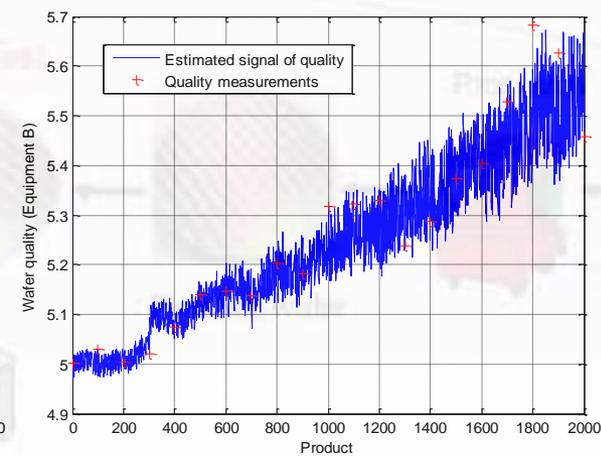
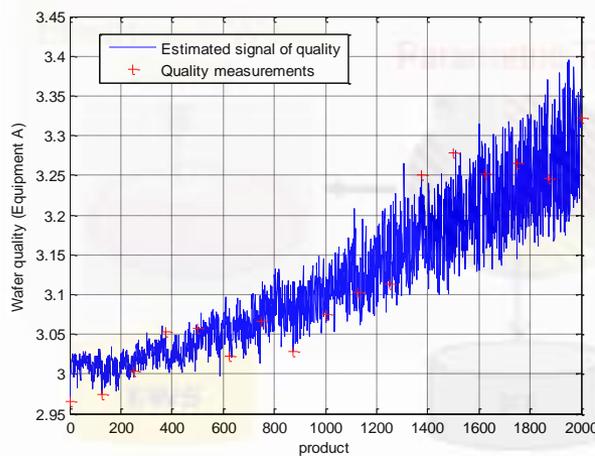
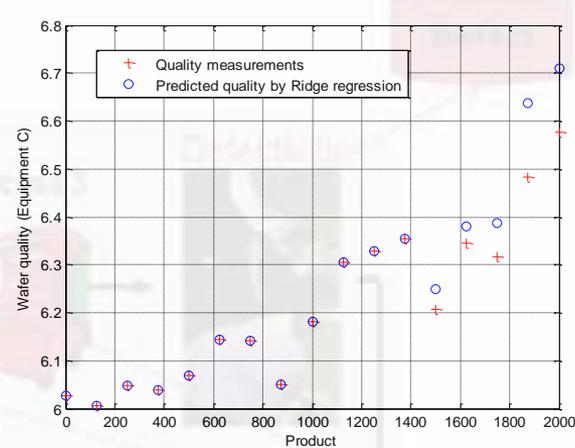
Equipment A



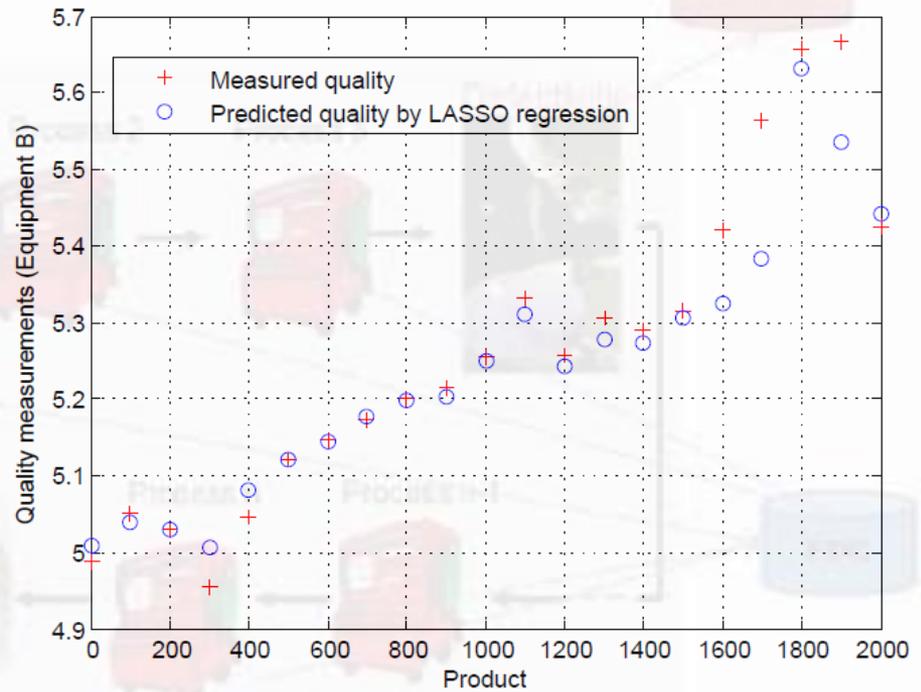
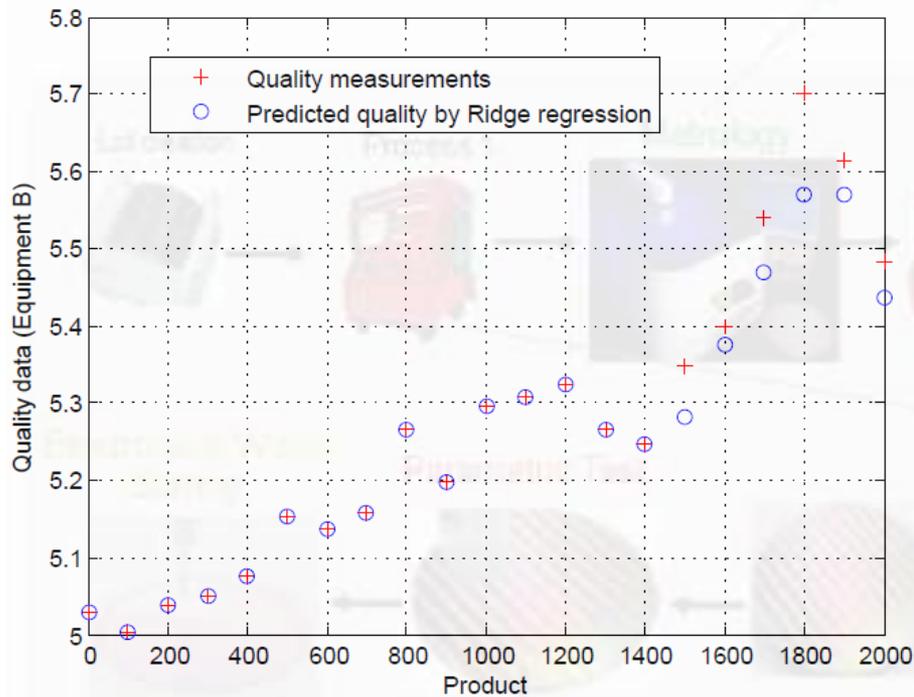
Equipment B



Equipment C



Comparaison



Comparaison

	Equipement A	Equipement B	Equipement C
Méthode à base de régressio d'arête	3.36×10^{-5}	3.35×10^{-5}	2.81×10^{-4}
Méthode à base de LASSO	0.0023	0.0034	0.0152

Erreur quadratique moyenne pour données d'apprentissage

	Equipement A	Equipement B	Equipement C
Méthode à base de régressio d'arête	0.0305	0.0716	0.0152
Méthode à base de LASSO	0.0482	0.1076	0.1022

Erreur quadratique moyenne pour données de test

Plan

1 Introduction

2 Approche I

3 Approche II

4 Conclusion

Conclusion

❖ Méthode I :

- Prédiction de nombre de wafers restant avant de perte de spécifications de produit
- Détection de l'équipement responsable de cette perte de spécifications en suivant l'évolution des indices de santé des équipements

❖ Méthode II :

- Prédiction de la spécification SPC du produit en fonction de données FDC de l'équipement et des spécifications des étapes précédentes
- Régression d'arête est la méthode la plus performante. Elle permet de prendre en compte:
 - La corrélation entre données FDC
 - La relation entre étapes de fabrication
 - L'effet de données FDC sur les données SPC

❖ Méthode III (en cours) :

- Prédiction de Scrap des produits en fonction de données d'alarmes

Références

- [1] - Lei Yang, Jay Lee. *Bayesian Belief Network-based approach for diagnostics and prognostics of semiconductor manufacturing systems*. Robotics and Computer-Integrated Manufacturing 28 (2012) 66-71.
- [2] - Hendrik Purwins, Ahmed Nagi, Bernd Barak, Uwe Höckele, Andreas Kyek, Benjamin Lenz, Günter Pfeifer, Kurt Weinzierl. *Regression Methods for Virtual Metrology of Layer Thickness in Chemical Vapor Deposition*. IEEE/ASME Transactions on Mechatronics, Vol. 19, No. 1, p. 1-8, 2014.
- [3] - Min-Hsiung Hung, Tung-Ho Lin, Fan-Tien Cheng and Rung-Chuan Lin. *A Novel Virtual Metrology Scheme for Predicting CVD Thickness in Semiconductor Manufacturing*. 308 IEEE/ASME Transactions on mechatronics, Vol. 12, No. 3, June 2007.
- [4] - Byungwhan Kim, Dong Won Kim, and Gwi Tae Park. *Prediction of Plasma Etching Using a Polynomial Neural Network*. IEEE TRANSACTIONS ON PLASMA SCIENCE, VOL. 31, NO. 6, DECEMBER 2003.
- [5] - Hyeon Bae, Sungshin Kim, Kwang-Bang Woo, Gary S. May, and Duk-Kwon Lee. *Fault Detection, Diagnosis, and Optimization of Wafer Manufacturing Processes utilizing Knowledge Creation*. International Journal of Control, Automation, and Systems, vol.4, no.3, pp.372-381, June 2006.
- [6] - Pao-Hua Chou, Meng-Jiun Wu, and Kuang-Ku Chen. *Integrating support vector machine and genetic algorithm to implement dynamic wafer quality prediction system*. Expert Systems with Applications 37 (2010) 4413-4424.
- [7] - Hasna Barkia, XAVIER BOUCHER, RODOLPHE LE RICHE, MARIE-AGNES GIRARD, PHILIPPE BEAUNE, DAVID ROZIER. *Extraction de connaissances pour une maîtrise du rendement en industrie du semi-conducteur*. CIGI 2013.
- [8] - Daniel Liebera, Marco Stolpeb, Benedikt Konrada, Jochen Deusea, Katharina Morik. *Quality Prediction in Interlinked Manufacturing Processes based on Supervised & Unsupervised Machine Learning*. Procedia CIRP 00 (2013) 000–000. Forty Sixth CIRP Conference on Manufacturing Systems 2013
- [9] - Yu-Chuan Su, Min-Hsiung Hung, Fan-Tien Cheng, and Yeh-Tung Chen. *A Processing Quality Prognostics Scheme for Plasma Sputtering in TFT-LCD Manufacturing*. IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 19, NO. 2, MAY 2006 183
- [10] - Matthias Becker, Helena Szczerbicka and Mei-Chen Lo. *The Effect of Additional Information on the Prediction Quality of Wafer Fabrication Operation with a Neural Network and Clustering Approach*. International Journal of Information Systems for Logistics and Management. Vol. 6, No. 2 (2011) 15-22
- [11] - Tobias Jung, Luis Herrera and Bernhard Schoelkopf. *Long T erm Prediction of Product Quality in a Glass Manufacturing Process Using a Kernel Based Approach*.
- [12] - Daniel Lieber, Marco Stolpe, Benedikt Konrad, Jochen Deuse, Katharina Morik. *Long Term Prediction of Product Quality in a Glass Manufacturing Process Using a Kernel Based Approach*. Forty Sixth CIRP Conference on Manufacturing Systems 2013. Procedia CIRP 7 (2013) 193 – 198

Références

- [13] - M. Mehdi Movahedi, M.B. Moghadam, H. Saiedi and S. Eftekhari H. *A Solution for Statistical Control of Correlated Processes*. Middle-East Journal of Scientific Research 8 (6): 1034-1045, 2011. ISSN 1990-9233. © IDOSI Publications, 2011.
- [14] - Fahmi Arif, Nanna Suryana, Burairah Hussin. *Cascade Quality Prediction Method Using Multiple PCA+ID3 for Multi-Stage Manufacturing System*. 2013 International Conference on Electronic Engineering and Computer Science.
- [15] - Sholom M. Weiss, Amit Dhurandhar, Robert J. Baseman, Brian F. White, Ronald Logan, Jonathan K. Winslow, and Daniel Poindexter. *Continuous prediction of manufacturing performance throughout the production lifecycle*. J Intell Manuf DOI 10.1007/s10845-014-0911-x.
- [16] - H. barkia, X. Boucher, R. Le Riche, M.-A. Girard, P. Beaune, and D. Rozier. *Extraction de connaissances pour une maîtrise du rendement en industrie du semi-conducteur*.
- [17] - D. Lieber, M. Stolpe, B. Konard, J. Deuse, and K. Morik. *Quality Prediction in Interlinked Manufacturing Process based on Supervised & Unsupervised Machine Learning*. Forty Sixth CIRP Conference on Manufacturing Systems. Procedia CIRP 7, 193 – 198, 2013.
- [18] - R. K. Nurani, R. Akella, and A. J. Strojwas. *In-Line Defect Sampling Methodology in yield Management: An Integrated Framework*. IEEE TRANSACTIONS ON SEMICONDUCTOR MANUFACTURING, VOL. 9, NO. 4, NOVEMBER 1996.
- [19] - T. S. Kim, S. H. Ahn, Y. G. Jang, J. I. Lee, K. J. Lee, B. Y. Kim, and C. H. Cho. *Intelligent yield Prediction Models for High-Speed Microprocessors*. The Ninth International Symposium on Semiconductor Manufacturing.
- [20] - D. C. Krueger. *Semiconductor Yield Modeling Using Generalized Linear Models*. Forty Sixth CIRP Conference on Manufacturing Systems. Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy , March 2011.
- [21] - L. Ludwing, E. Pelz, M. Kessler, W. Sinderhauf, B. Koppenhoefer, and W. Rosenstiel. *Prediction of functional yield of chips in semiconductor industry applications*.
- [22] - Thi bich lien NGUYEN - Mohand DJEZIRI - Bouchra ANANOU - Mustapha OULADSINE - Jacques PINATON. "Fault Prognosis with Stochastic Modelling on Critical Points of Discrete Processes" , *Second European Conference of the Prognostics and Health Management Society 2014, Prognostics and Health Management Society*, jui 2014

Publications

[1] - Mariam Melhem, Bouchra Ananou, Mustapha Ouladsine, Jacques Pinaton. *Regression methods for predicting the product's quality in the semiconductor manufacturing process*, 8th IFAC Conference on Manufacturing Modelling, Management and Control, Troyes, juin 2016

[2] - Mariam Melhem, Bouchra Ananou, Mustapha Ouladsine, Jacques Pinaton. *Regularized regression models to predict the product quality in multistep manufacturing*, International Conference on Systems and Control, Marakech, May 2016.

[3] - Mariam Melhem, Bouchra Ananou, Mohand Djeziri, Mustapha Ouladsine, Jacques Pinaton. *Prediction of the Wafer quality with respect to the production equipments data*, 9th IFAC Symposium on Fault Detection, Supervision (SAFEPROCESS), IFAC, Arts et métiers Polytech, Paris, France, september 2015.

[4] - Mariam Melhem, Bouchra Ananou, Mohand Djeziri, Mustapha Ouladsine, Jacques Pinaton. *Product's Quality Prediction with respect to equipments Data*, 2015 International Conference on Data Mining (DMIN'15), Las Vegas, Nevada, USA, july 2015.

[5] - Mariam Melhem, Bouchra Ananou, Mohand Djeziri, Mustapha Ouladsine, Jacques Pinaton. *Pronostic de la qualité d'un produit dans l'industrie de production : Application à l'industrie de semi-conducteurs*, 6èmes Journées Doctorales / Journées Nationales MACS, INSA Centre - Val de Loire, Campus de Bourges, june 2015.

Prédiction de la qualité du produit en fonction de la qualité des équipements:

Application à la fabrication de semi-conducteurs

Réunion du GT S3 (Sûreté - Surveillance - Supervision)

Arts et Métiers ParisTech, **15 septembre 2016**

Contact : mariam.melhem@isis.org

MERCI POUR VOTRE ATTENTION