

Défis pour la Maîtrise Statistique des Processus à l'aube du XXIème siècle

Teodor TIPLICA

LARIS, Université d'Angers



QUALITA'2015

*11ème Congrès International Pluridisciplinaire en
Qualité, Sûreté de Fonctionnement et Développement
Durable*

17-19 mars 2015 Nancy (France)



Contexte



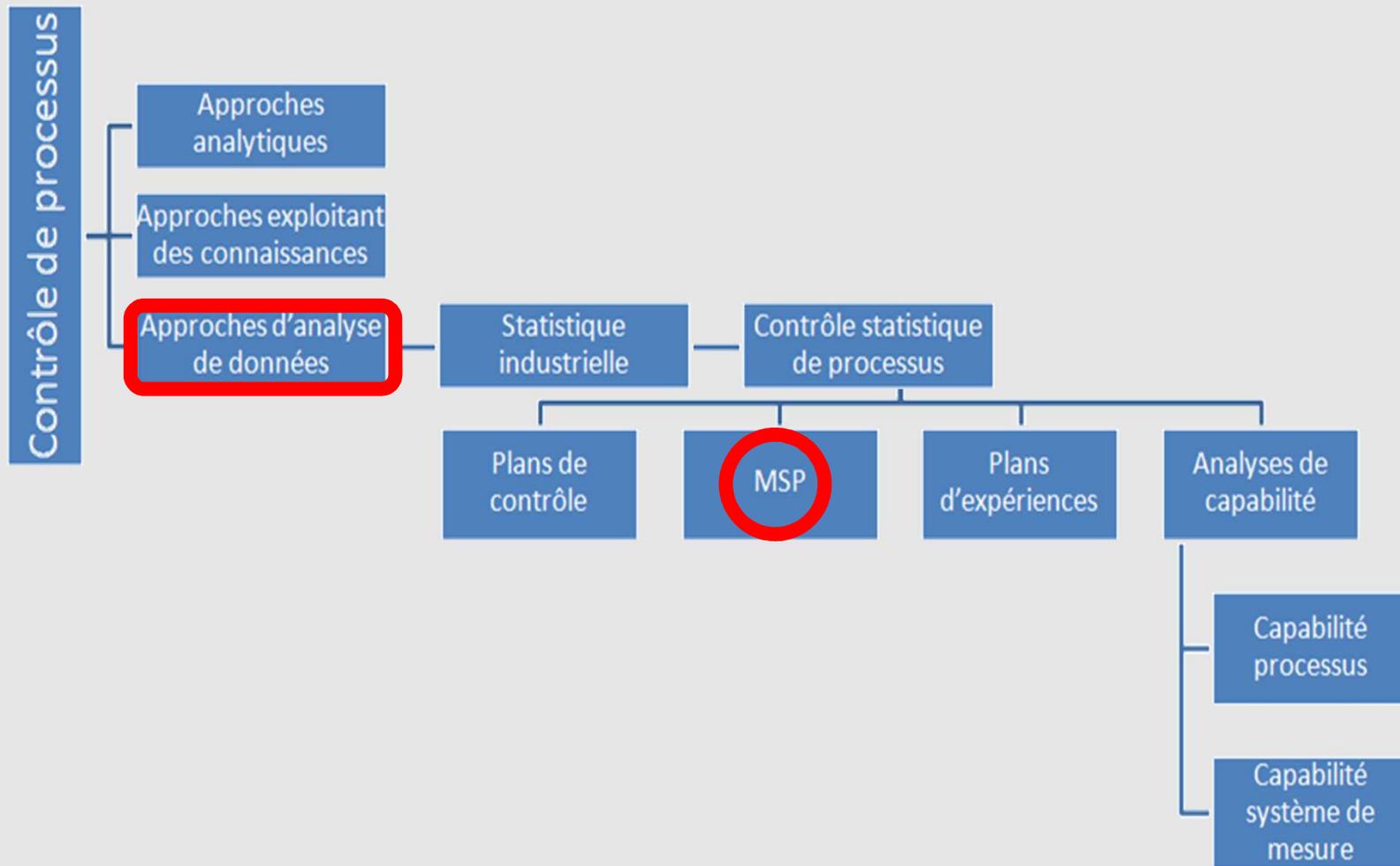
- Les approches statistiques ont joué un rôle très important dans l'accélération du rythme de développement industriel lors du dernier siècle.



- Les applications industrielles ont offert aux statisticiens des grandes opportunités pour le développement d'une recherche méthodologique



Contexte



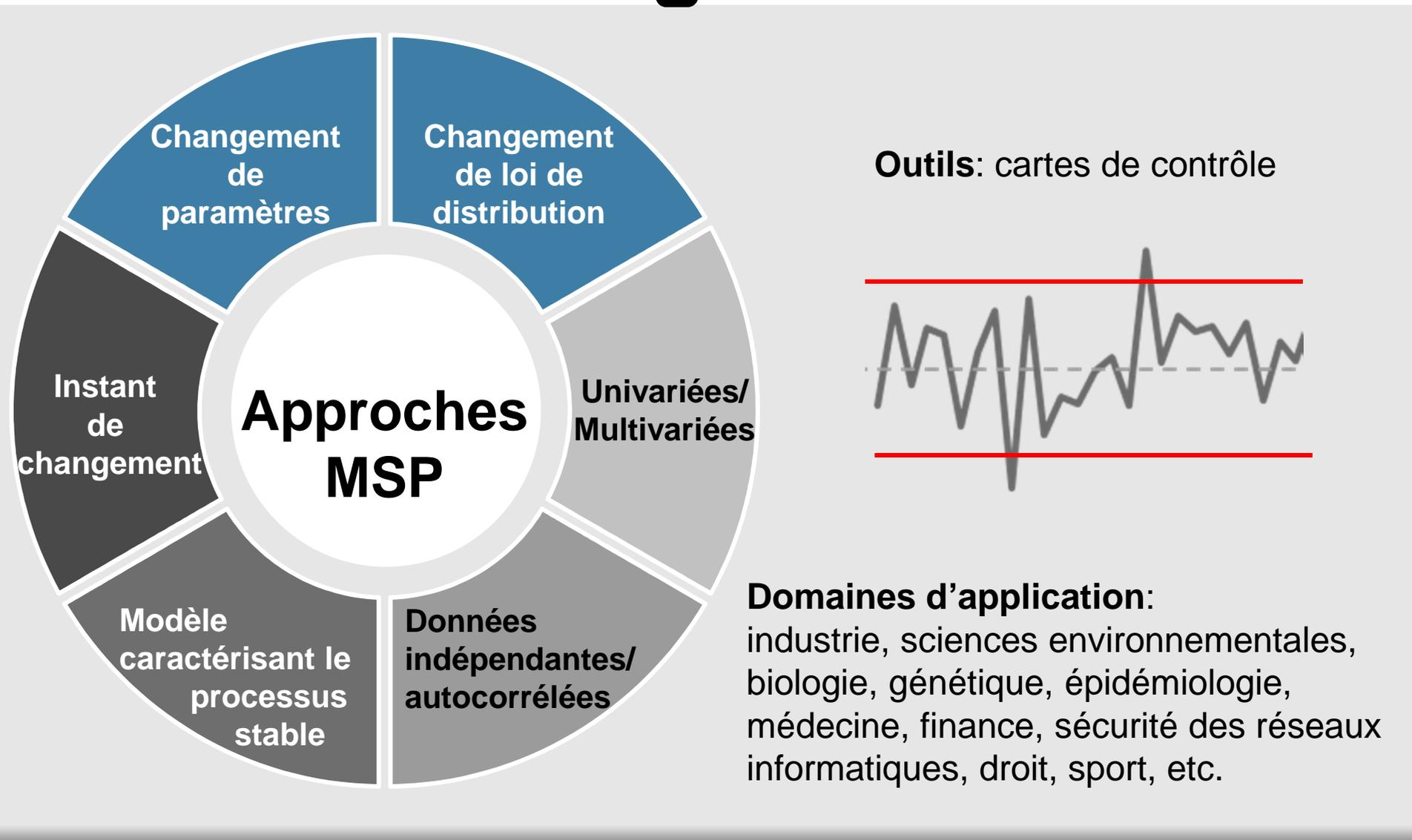
« Anything that can go wrong, will go wrong »¹ (Edward A. Murphy Jr.).

Objectif de la MSP:

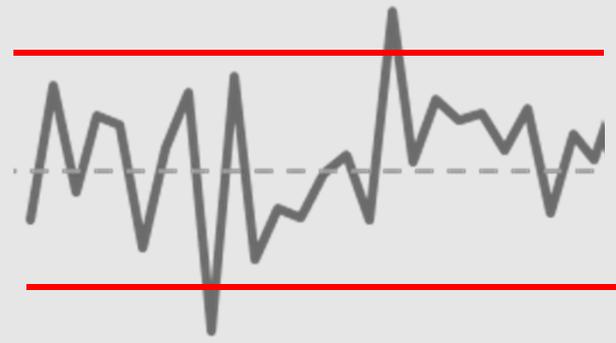
**CONCEPTION DES OUTILS PERFORMANTS
DE DÉTECTION AYANT UN NOMBRE RÉDUIT
DE FAUSSES ALERTES**

1. « *Tout ce qui peut mal tourner va mal tourner* »

Détection $\begin{cases} H_0: \{X_i\}_{i=1,2,\dots,t} \sim F_0 \\ H_1: \{X_i\}_{i=1,2,\dots,\tau} \sim F_0 \text{ et } \{X_i\}_{i=\tau+1,\dots,t} \sim F_1 \end{cases}$



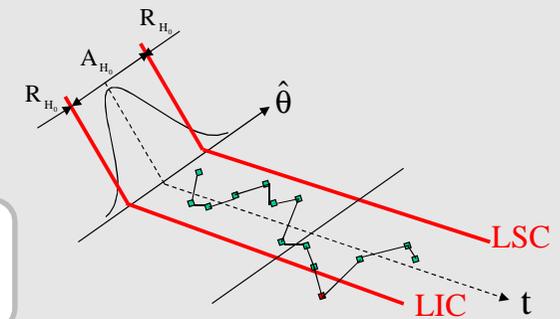
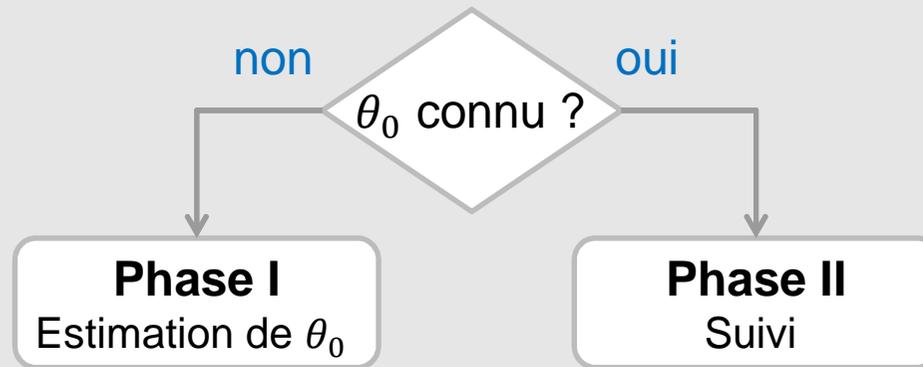
Outils: cartes de contrôle



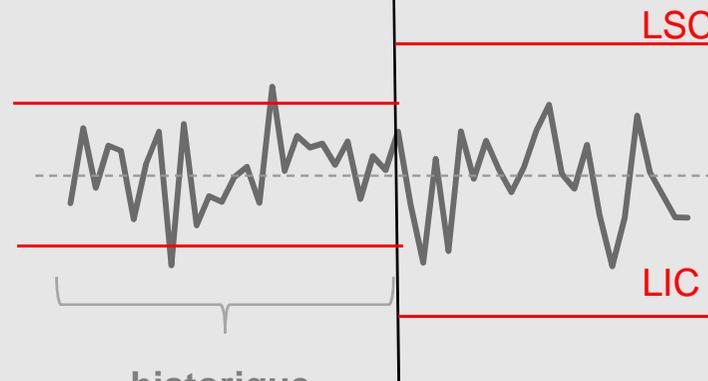
Domaines d'application:
industrie, sciences environnementales, biologie, génétique, épidémiologie, médecine, finance, sécurité des réseaux informatiques, droit, sport, etc.

Construction d'une carte de contrôle

$$\{X_i\} \sim f_\theta(x) \quad \begin{cases} H_0 & \theta = \theta_0 \\ H_1 & \theta \neq \theta_0 \end{cases}$$



Analyse exploratoire de données



Application répétée d'un test d'hypothèse

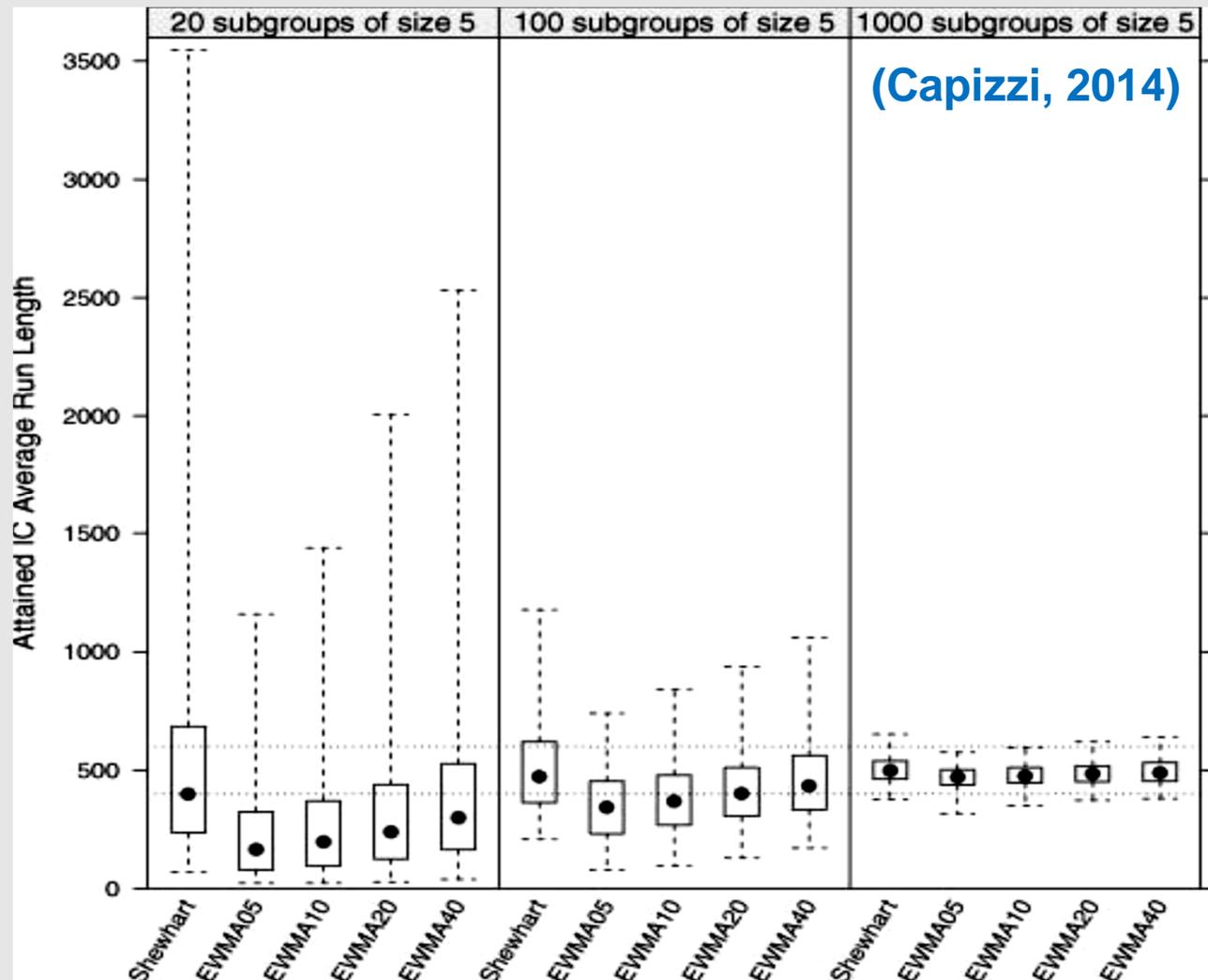
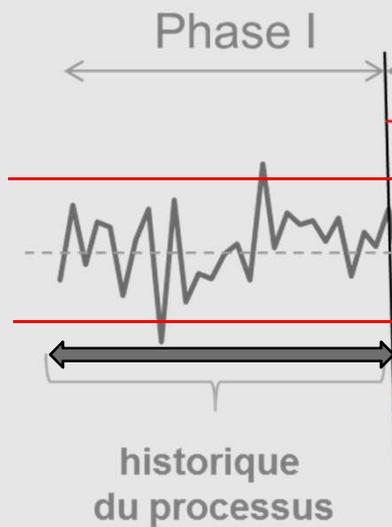
Indicateurs de performance : FAP, FAR, etc.

historique du processus

Indicateurs de performance : ARL, ATS, ANSS, etc.

Phase I – problématique

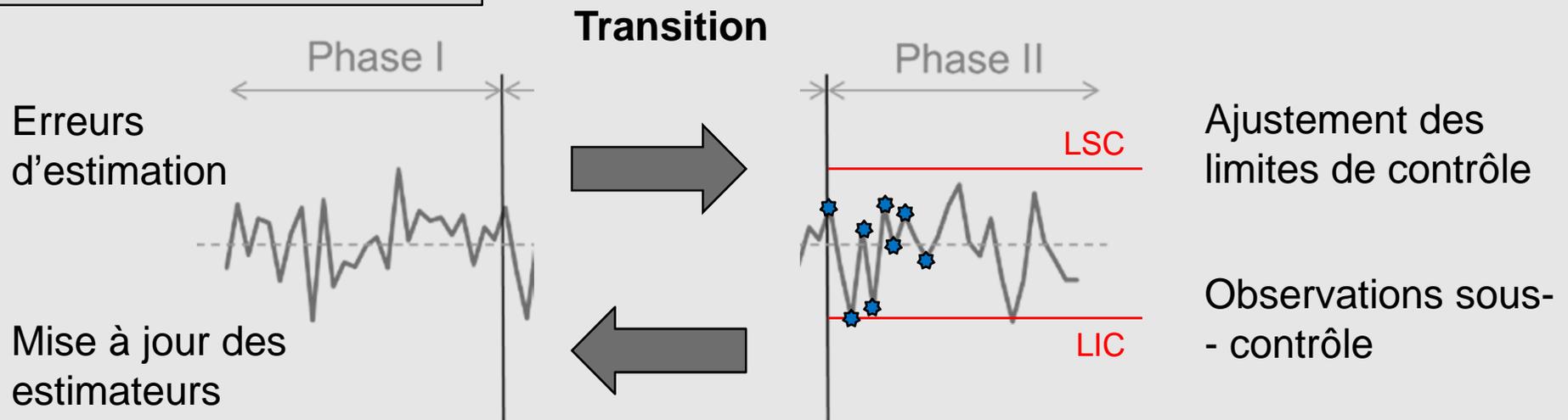
Besoin données ++
temps ++



Phase I – problématique



Contraintes données - -
 temps - -

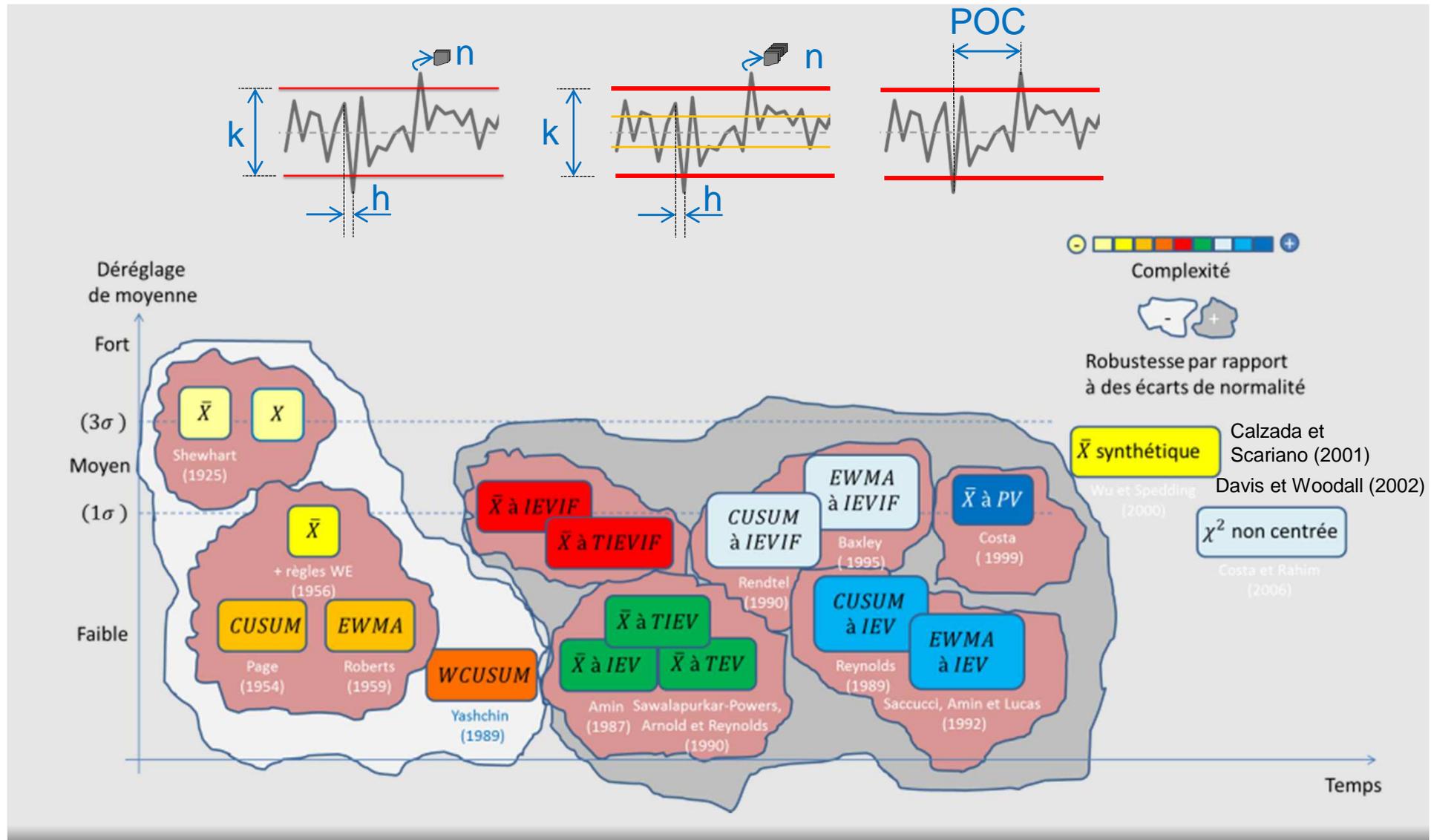


Vrai « challenge » pour la recherche !!!

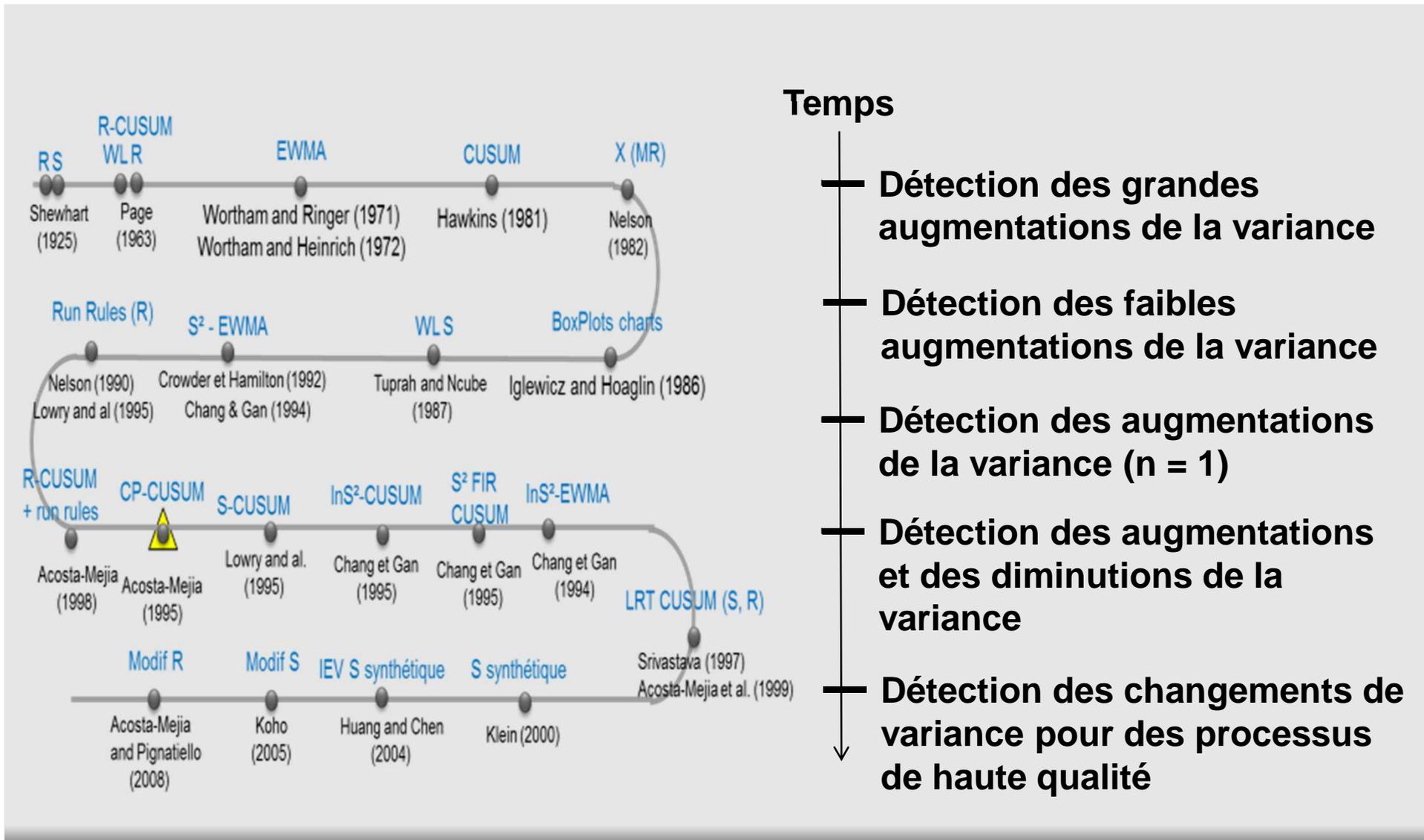


CARTES DE CONTRÔLE UNIVARIÉES

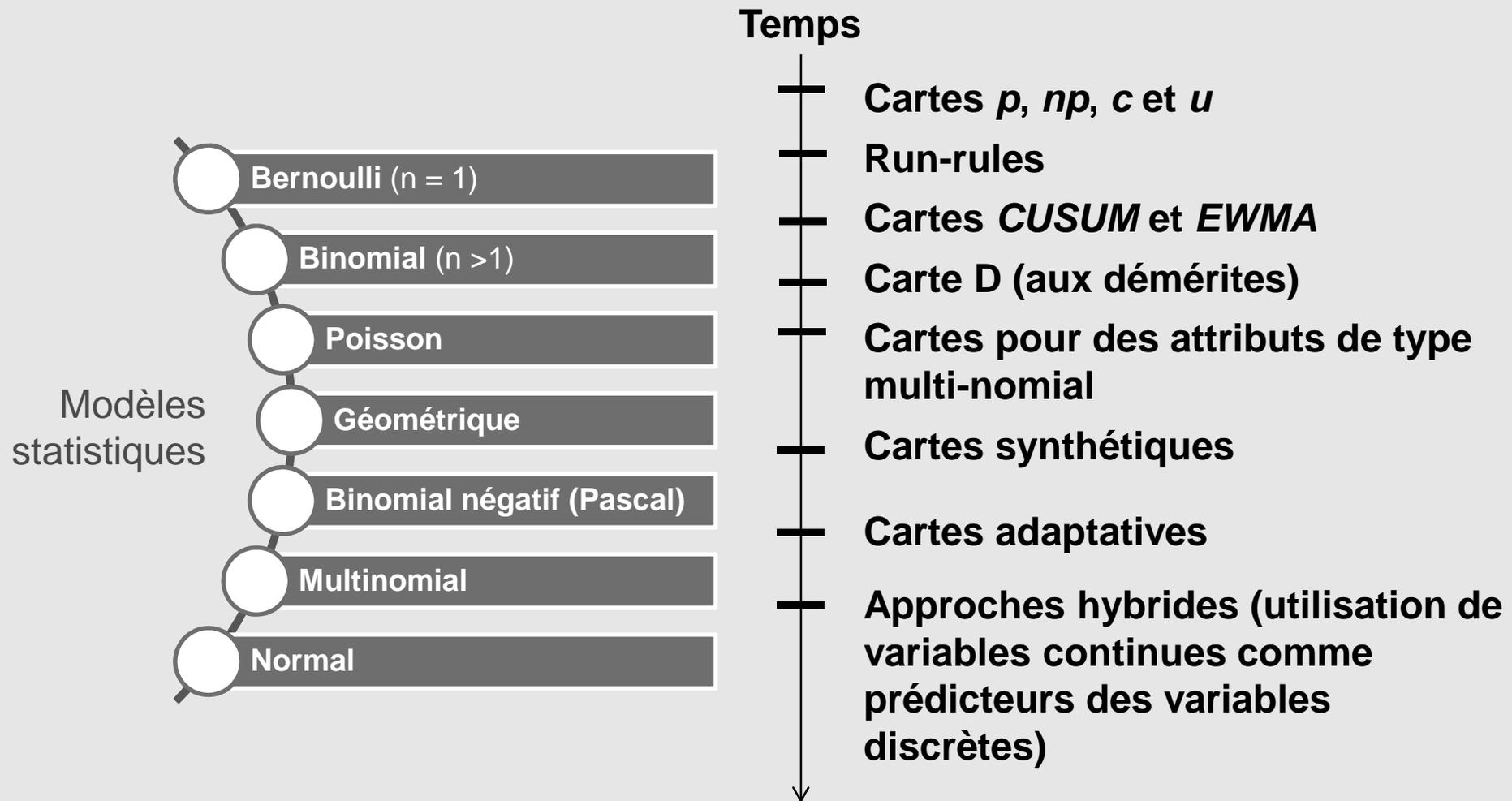
Cartes de contrôle pour la moyenne



Cartes de contrôle pour la variance



Cartes de contrôle pour les attributs





DÉTECTION DE L'INSTANT DE CHANGEMENT

Approches univariées

$$X_i \sim \begin{cases} N(\mu_0, \sigma_0) & \text{si } i \leq \tau \\ N(\mu_1, \sigma_0) & \text{si } i > \tau \end{cases} \quad i = 1, 2, \dots, T \quad \mu_1 = \mu_0 + \delta \frac{\sigma_0}{\sqrt{n}}$$

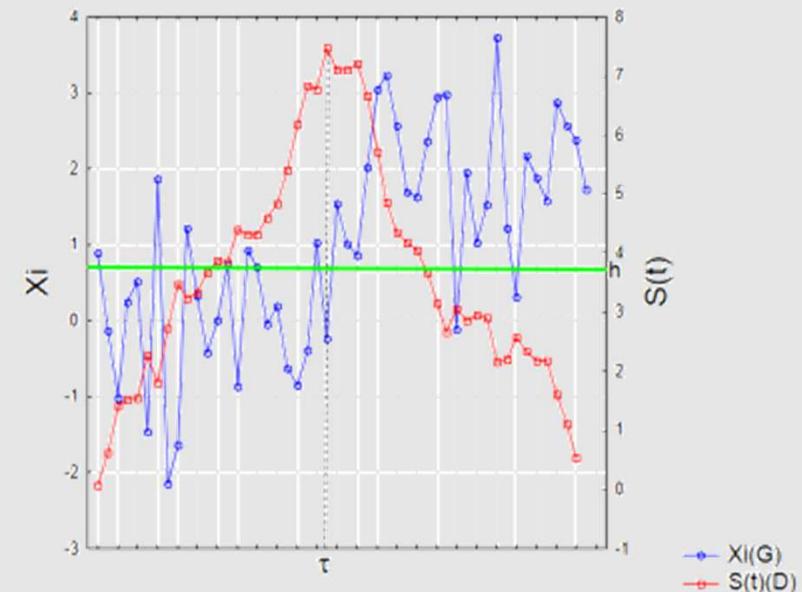
$$\hat{t} = \left\{ t, S_{max} = \max_{1 \leq t < T} \{S(t)\} > h \right\} \quad S(t) = \sqrt{\frac{t(T-t)}{T}} \frac{\bar{x}_{0,t} - \bar{x}_{t,T}}{S_c}$$

$$h = t \frac{\alpha}{T-1} T^{-2}$$

(Hawkins, Qiu et Kang, 2003)

Méthodes reposant sur l'approche du rapport de maximum de vraisemblance.

Basseville et Nikiforov (1993) – Detection of Abrupt Changes: Theory and Applications.



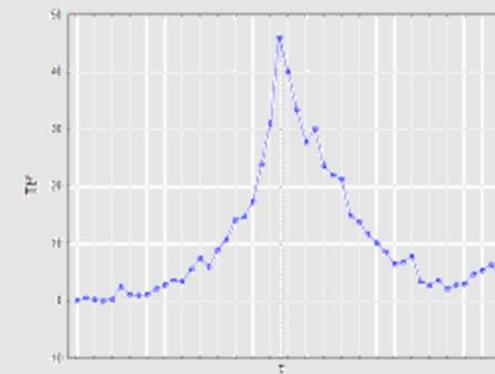
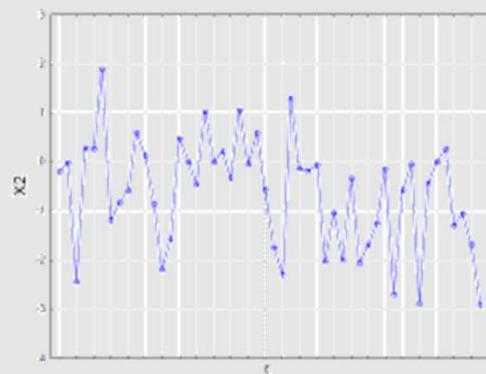
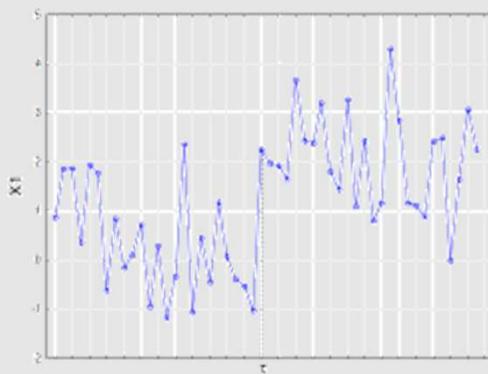
Approches multivariées

$$X_i \sim \begin{cases} N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) & \text{si } i \leq \tau \\ N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_0) & \text{si } i > \tau \end{cases} \quad i = 1, 2, \dots, T$$

$$\hat{t} = \arg \max_{0 \leq t < T} \{T_t^2\} \quad Y_t = \sqrt{\frac{t(T-t)}{T}} (\bar{X}_{0,t} - \bar{X}_{t,T})$$

$$T_t^2 = Y_t^T W_t^{-1} Y_t \quad (\text{Srivastava et Worsley, 1986})$$

Illustration: $\boldsymbol{\mu}_0 = [0; 0]$ $\boldsymbol{\mu}_1 = [2; -1]$





APPROCHES POUR DES DONNÉES NON-NORMALES ET/OU AUTO-CORRÉLÉES

Approches de détection pour des lois non-gaussiennes



Solution 1: Utilisation des cartes « classiques »

- Cartes X-bar de Shewhart (n grand), EWMA, etc.
- Cartes adaptatives avec des limites asymétriques

Solution 2: Utilisation des cartes distributionnelles

- Cartes pour la loi Exponentielle, Poisson, Rayleigh, etc.

Solution 3: Utilisation des transformations

- Box-Cox, Johnson, Manly, John-Draper, etc.

Solution 4: Utilisation de cartes non-paramétriques

- Représentent sur l'utilisation des tests non-paramétriques (rang, signe, rang signé, etc.)

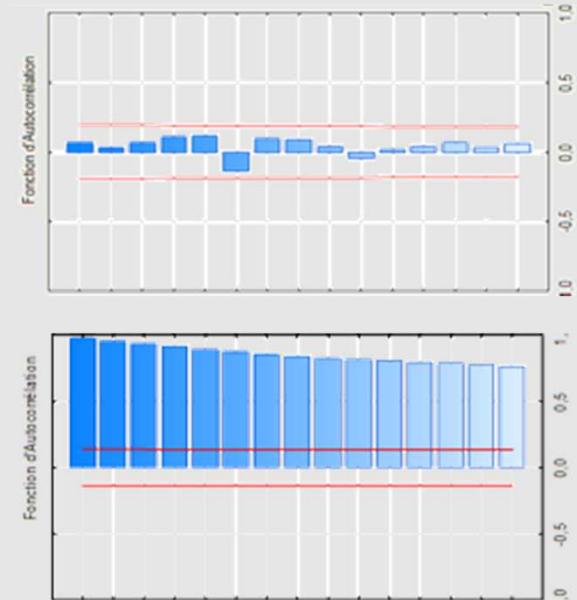
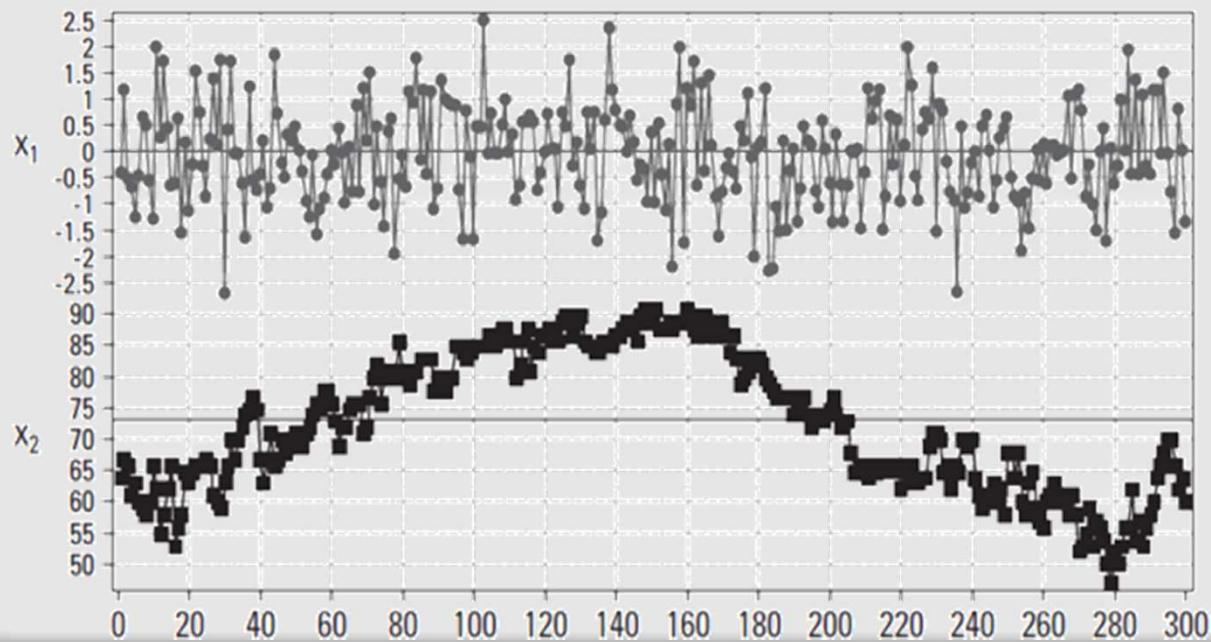
Approches de détection pour des données autocorrélées



Inconvénient de l'autocorrélation

Sous-estimation de la variance

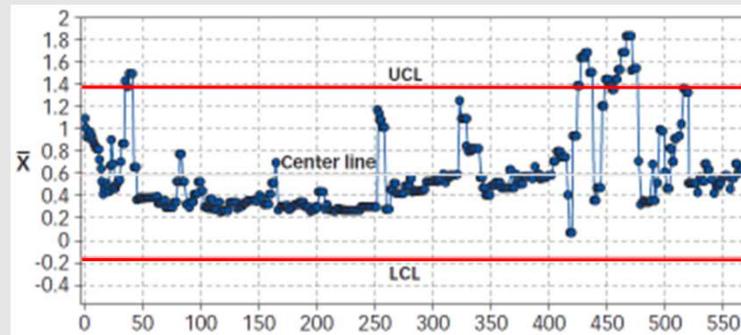
- Augmentation du taux de fausses alertes



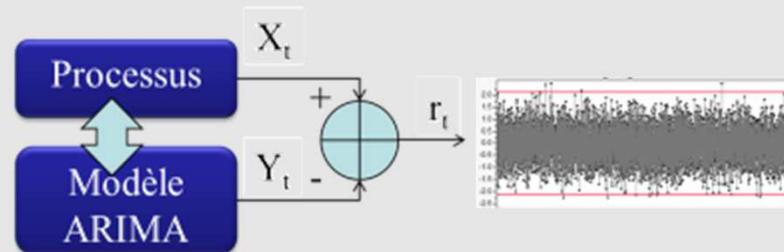
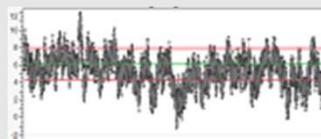
Approches de détection pour des données autocorrélées



Solution 1: Elargir l'intervalle entre les limites de contrôle



Solution 2: Suivi des résidus des modèles ARIMA



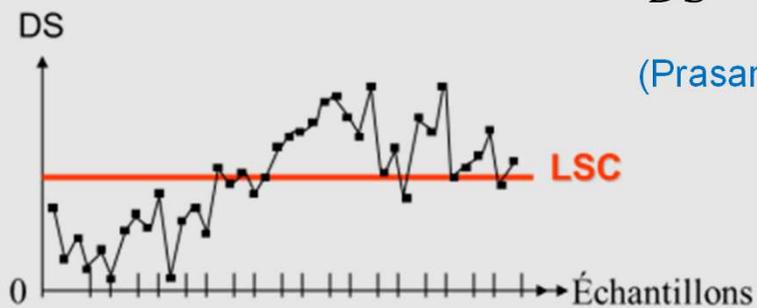
- Estimation
- Interprétation 
- Fenêtre d'opportunité

Solution 3: Utilisation de certaines transformations



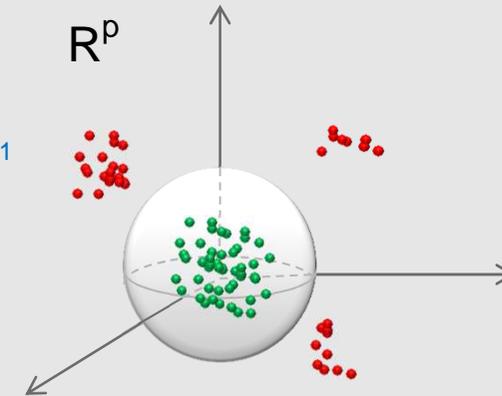
CARTES DE CONTRÔLE MULTIVARIÉES

Principe



$$DS^2 = (X - \mu)^T \Sigma^{-1} (X - \mu)$$

(Prasanta Chandra Mahalanobis, 1936)¹



Avantages

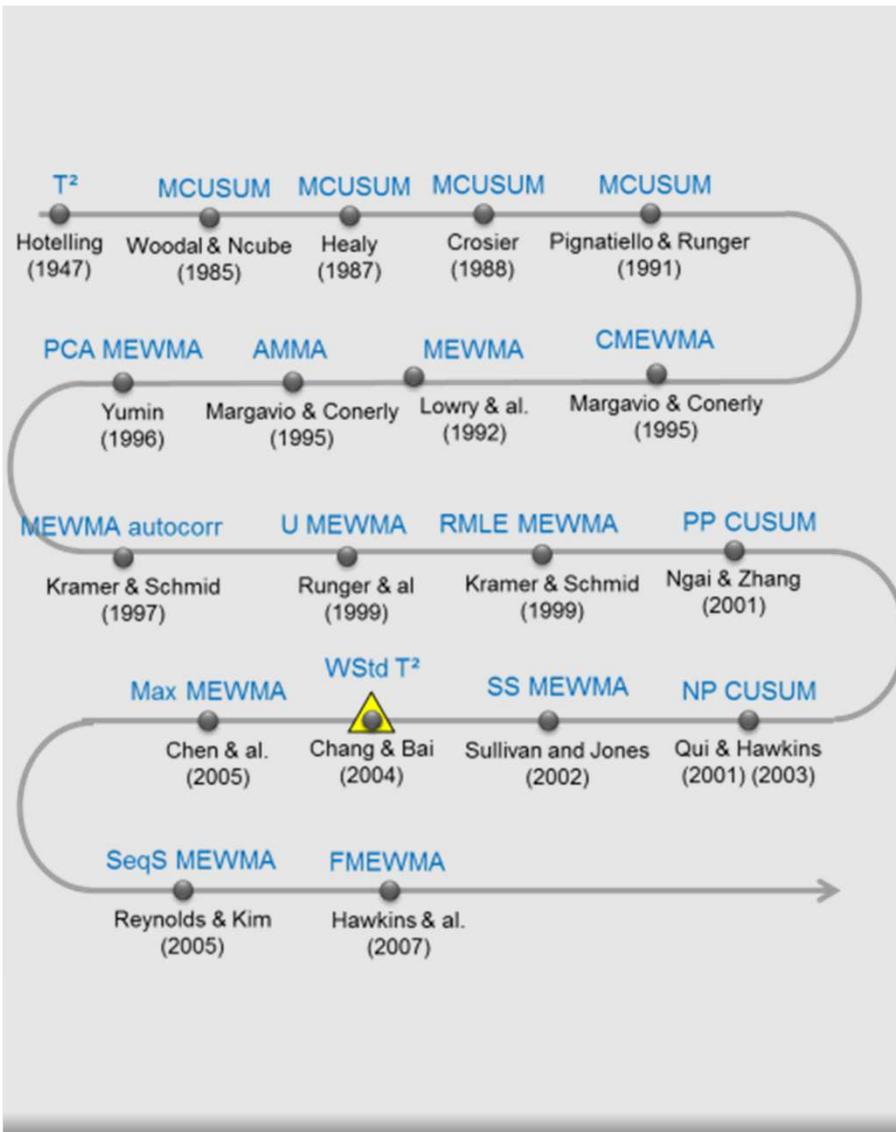
- [Redacted]
- Taux de fausses alertes maîtrisé
- Corrélation entre les paramètres considérée
- Suivi simplifié

Inconvénients

- [Redacted]
- Diagnostic plus difficile
- Complexité plus forte

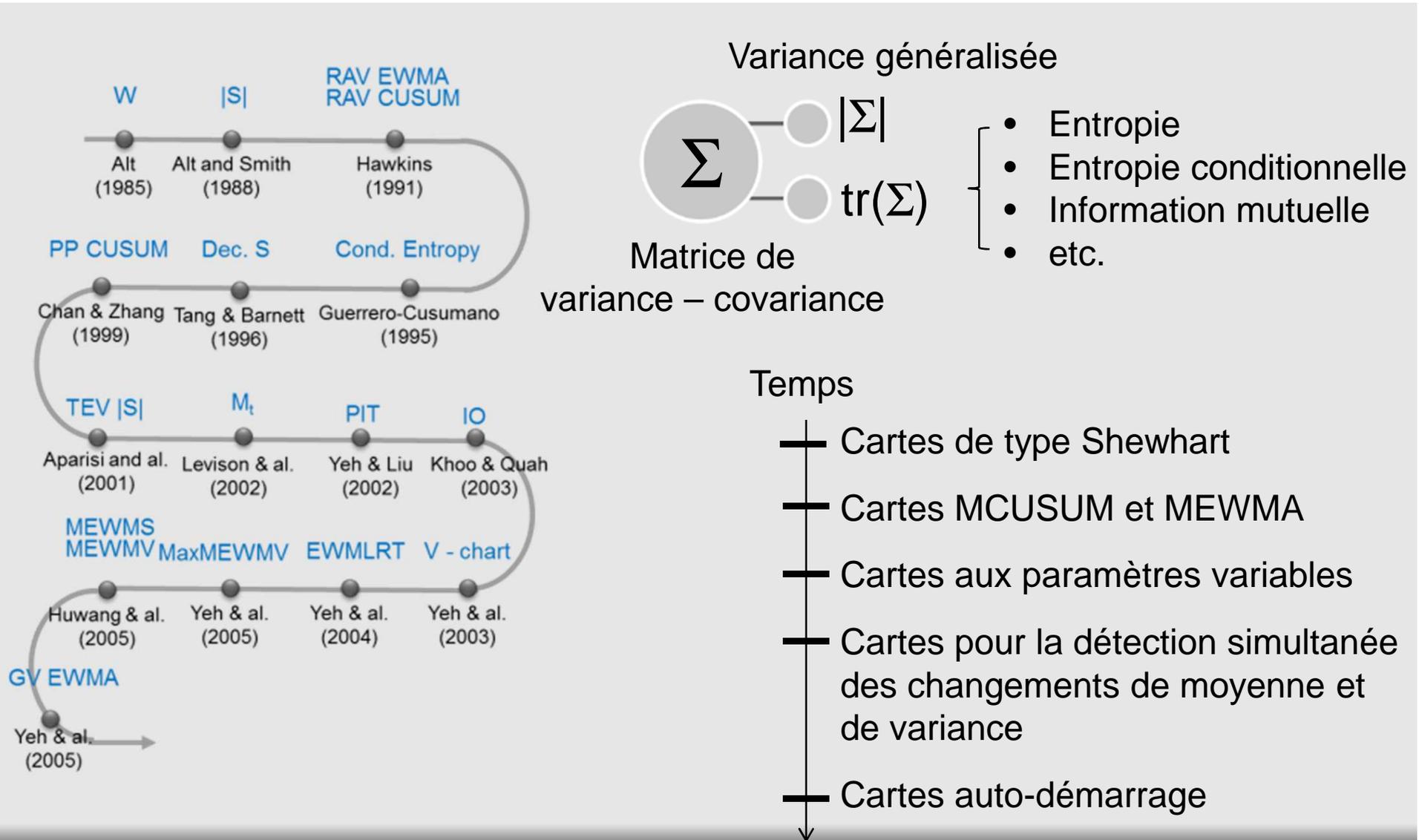
1. P. C. Mahalanobis, « On the generalised distance in statistics », Proceedings of the National Institute of Sciences of India, vol. 2, no 1, 1936, p. 49–55

Détection de changement de **moyenne**



- Efforts importants pour la détection plus efficace de **faibles changements** dans le vecteur des moyennes (MCUSUM et MEWMA)
- Efforts importants pour transiter des **schémas directionnels** vers des **schémas invariants**
- **Peu de travaux en dehors du cadre** $i.i.d. X \sim N_p(\mu, \Sigma)$
- Utilisation des approches « **change-point** » ($n = 1$)
- Utilisation des **estimateurs robustes** ($n = 1$)

Détection de changement de dispersion



Détection de changement de dispersion



- Cartes existantes
- pour $n > 1$ et $n = 1$
 - statistiques de test ($|S|$; rapport de vraisemblance)
 - généralement affectées par des changements de moyenne
 - diagnostic difficile

 Beaucoup moins de travaux que dans le cas de la détection des changements dans le vecteur des moyennes !

 Hypothèse: i.i.d. $X \sim N_p(\mu, \Sigma)$

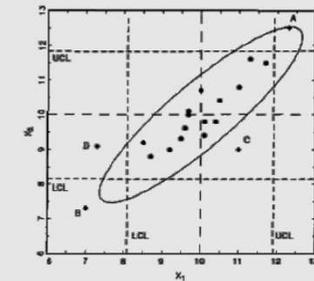
 Peu de travaux portant sur la Phase I



APPROCHES STATISTIQUES DE DIAGNOSTIC

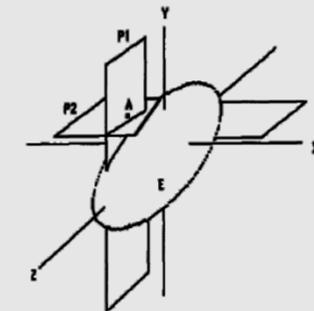
■ Tests univariés

- cartes de Shewhart avec des limites de type Bonferroni (Alt, 1985)
- intervalles de confiance joints pour les moyennes (Hayter & Tsui, 1994)
- ellipse de contrôle (Alt, 1985)
- méthode de l'hyperplan (Chua & Montgomery, 1992)
- carte de contrôle minmax (Sepulveda & Nachlas, 1997)



■ Décomposition de la statistique T^2

- statistique de Student (Doganaksoy & al, 1991)
- décomposition de Murphy (Murphy, 1987)
- méthode « step-down » (Wierda, 1994)
- méthode FIT (Finite Intersection Tests) (Timm, 1996)
- méthode RAM (Runger & al., 1996)
- décomposition MYT (Mason & al., 1995)
- double décomposition MYT (Mason & al., 1996)



Diagnostic par des approches statistiques



A
P
P
R
O
C
H
E
S

■ Régression

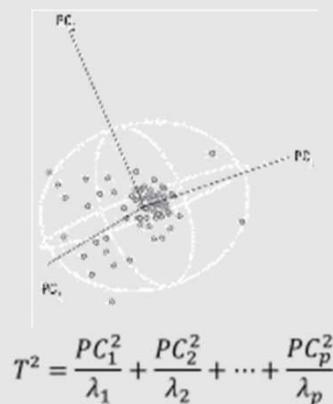
- carte sélection de la cause (Zhang, 1984)
- ajustement régressif de variables (Hawkins, 1993)
- DVR (Dummy Variable Regression) (Kalagonda & Kulkarni, 2003)

■ ACP

- diagrammes de contribution (Kourti & MacGregor, 1996)
- diagrammes de contribution généralisées (Wasterhuis & al, 2000)
- SBNI (Standard Bivariate Normal Integral) (Maravelakis & al, 2002)
- ML-PCA mixture (Choi & al, 2005)

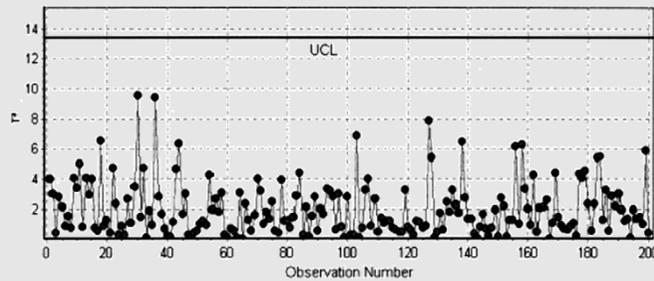
■ Techniques graphiques

- MP chart (Multivariate Profiles) (Fuchs & Benjamini, 1994)
- Gabriel biplot (Sparks & al, 1997)
- diagramme radar (Nottingham & al., 2001)

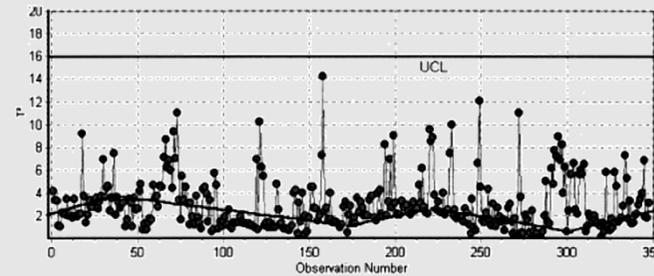


Techniques graphiques

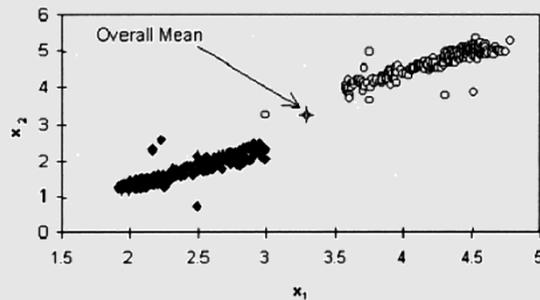
- identification des « profils » dans la statistique T^2 (Mason & al, 2003)



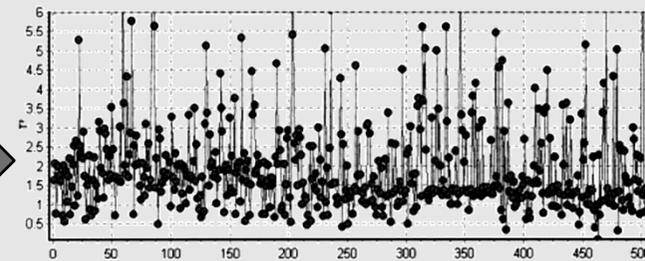
(T^2 normal)



(T^2 avec variations saisonnières)



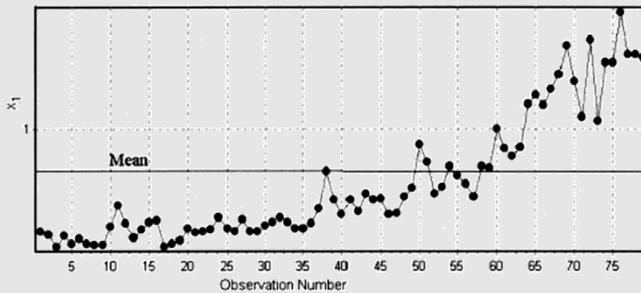
(mixtures)



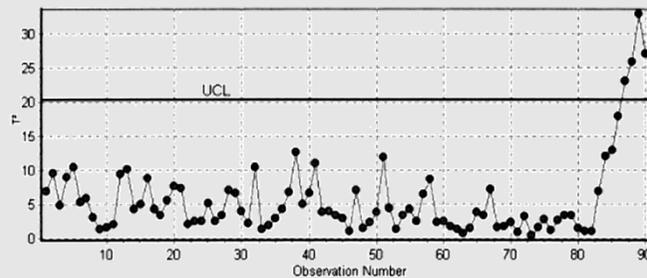
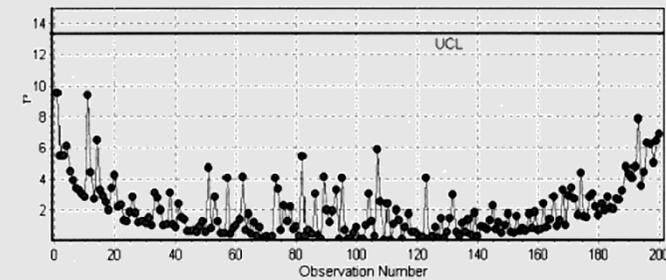
(T^2 pour mixtures)

■ Techniques graphiques

- identification des « profils » dans la statistique T^2 (Mason & al, 2003)



(Autocorrélation + dérive)

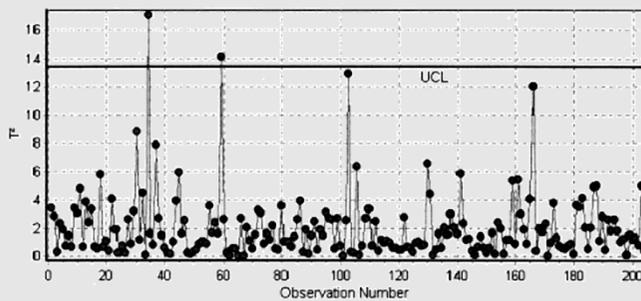


(T^2 saut en échelon)

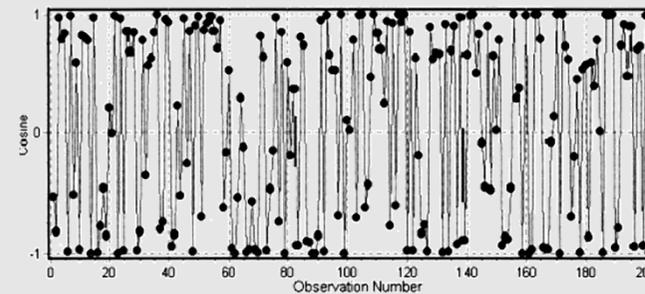
Techniques graphiques

- identification des « profils » dans la statistique T^2 (Mason & al, 2003)

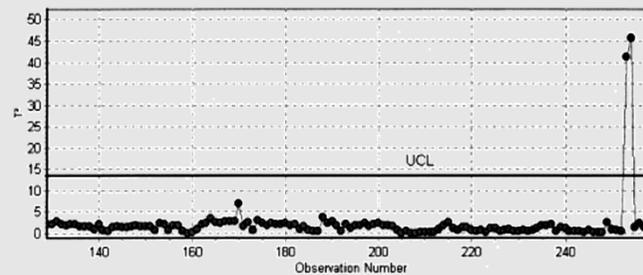
APPROCHES



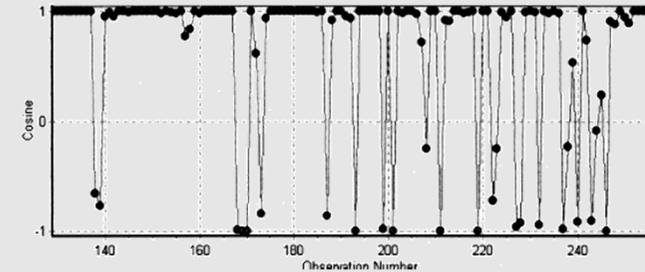
(T^2 normal)



(diagramme cos)



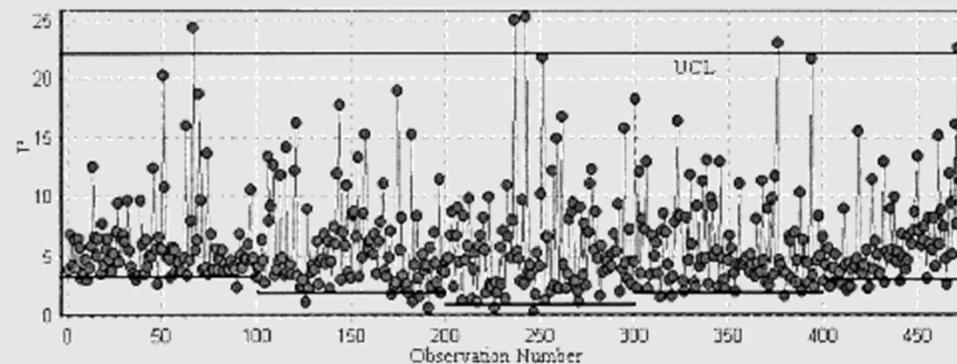
(T^2 + cause assignable)



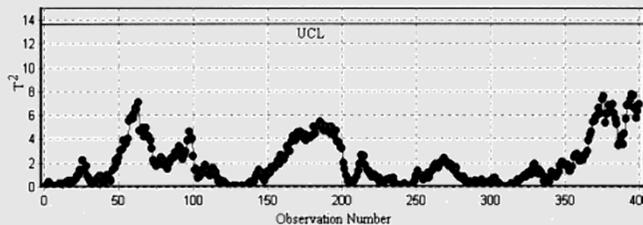
(diagramme cos)

Techniques graphiques

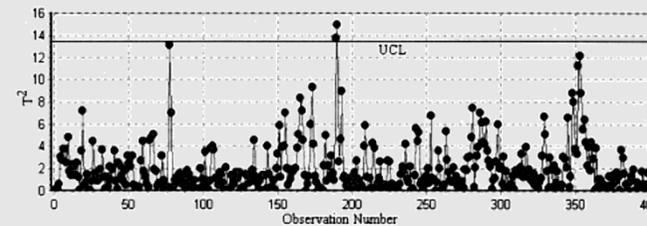
- identification des « profils » dans la statistique T^2 (Mason & al, 2003)



(T^2 en présence de sauts multiples de type échelon)



(T^2 autocorrélation $\rho_1 = 0.99$)



(T^2 autocorrélation $\rho_1 = 0.7$)



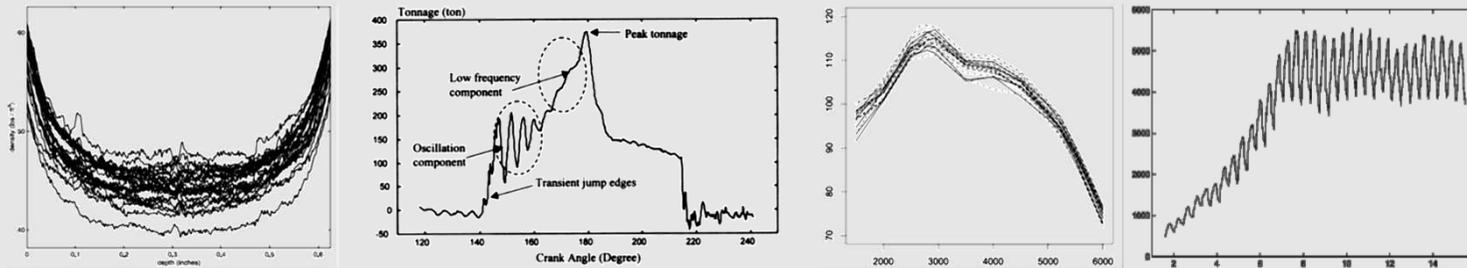
DÉVELOPPEMENT D'APPROCHES PLUS EN « PHASE » AVEC LES SYSTÈMES MODERNES DE PRODUCTION

Avancées importantes dans la de fabrication moderne



- **Exigences** qualité de plus en plus importantes
- Nouveaux systèmes de fabrication (**modulaires**, **flexibles** et **reconfigurables**, etc.) pour des opérations de fabrication complexes
- Opérations de fabrication réalisées sur **plusieurs machines** ou **étapes de fabrication**
- De moins en moins d'expérimentation et tests physiques. Recours aux modèles virtuels (**CAD/CAM**)
- Produits de complexité plus importante
- Données abondantes et de structures variées (**images**; **données fonctionnelles**, **séries chronologiques multidimensionnelles**, etc.)
- Capacités de calcul de plus en plus performantes

Surveillance de profils



$$Y = f(X)$$

■ Solutions

- suivi de certains points clés dans le profile (cartes de contrôle MSP pour ces points)
- suivi de l'intégralité du profile (modèle + approches paramétriques/non-paramétriques)

$$\begin{cases} H_0 & f = f_0 \\ H_1 & f \neq f_0 \end{cases} \iff \begin{array}{l} \text{processus sous-contrôle} \\ \text{processus hors-contrôle} \end{array}$$

Surveillance de profiles

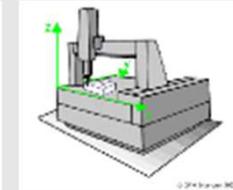
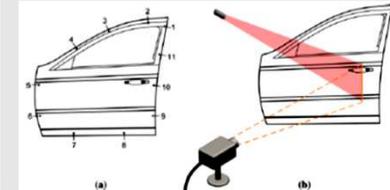
- **Suivi de paramètres des modèles régressifs** (linéaires, non-linéaires, simples et multiples, logistique, etc.)
 - cartes de Shewhart univariées ([Chang et Gan, 2006](#))
 - change point ([Mahmoud et al., 2007](#))
 - cartes self-starting ([Zou et al, 2007](#))
 - PCA and ICA ([Ding et al., 2006](#))
- **Suivi de l'ajustement de profiles**
 - SVM ([Moguerza et al., 2007](#))
 - clustering ([Chen and al., 2015](#))
- **Transformations**
 - Transformation de Fourier Discrète ([Chen et Nembhard, 2007](#))
 - Ondelettes ([Chicken and Pignatiello, 2007](#))

Surveillance spatio-temporelle

- Compter les événements se produisant dans le temps mais en tenant compte de leur location!

- Domaines d'application

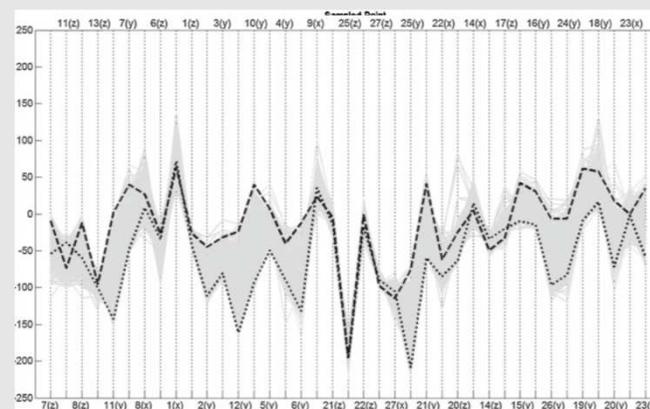
- santé publique
- industrie (scanners laser, machines à mesurer tridimensionnelle (MMT), etc.)
- médecine



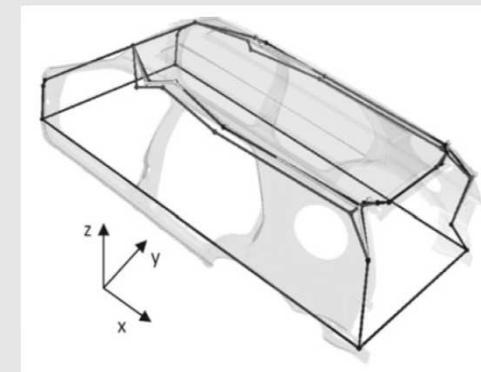
- Exemple (surveillance de profils 3D)



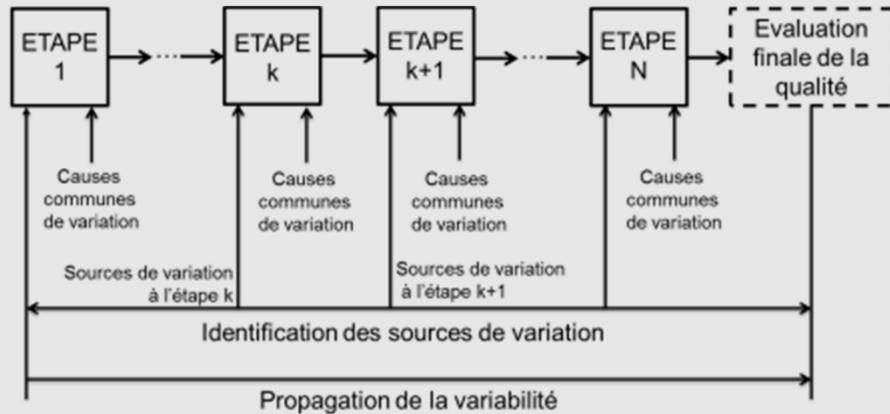
PCA
→



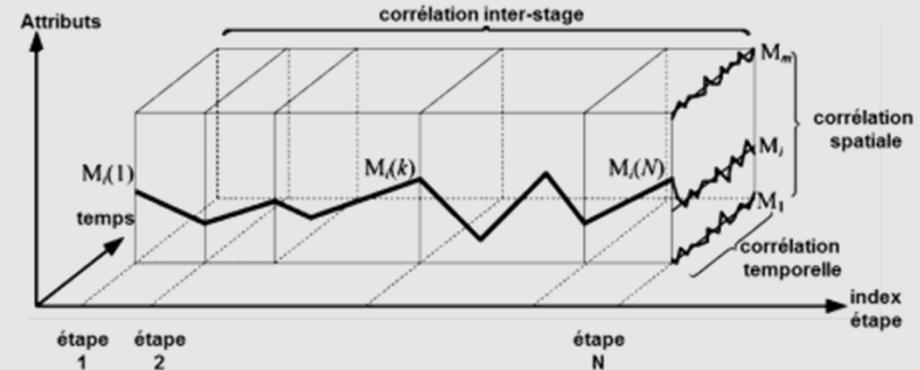
PCA⁻¹
→
+
CAD



Suivi de processus multi-étapes



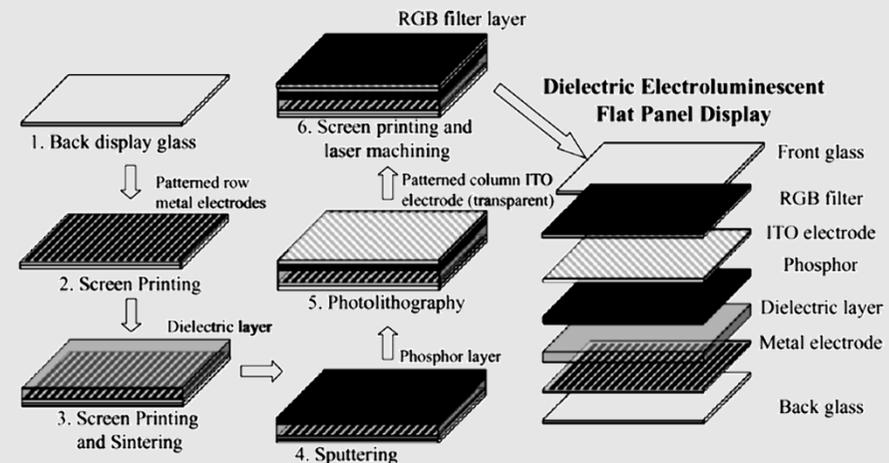
(schéma d'un processus multi-étapes)



(structure de données dans un processus multi-étapes)

■ Difficultés

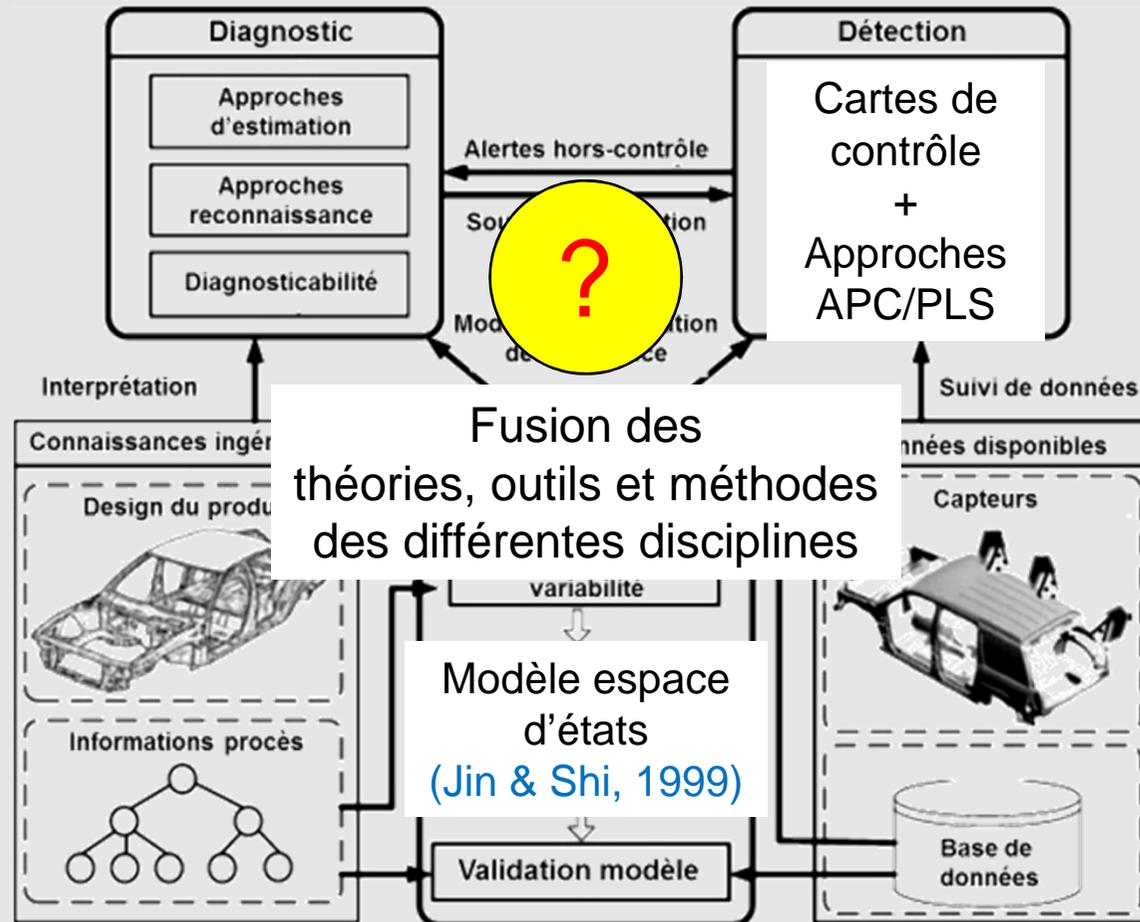
- grand nombre d'étapes
- structures hybrides (séquentielles, parallèles, etc.)
- données mixtes et à des échelles différentes
- performance collective et stochastique



Etapes majeures dans la fabrication d'un affichage diélectrique à écran plat électroluminescent

Suivi de processus multi-étapes

- Utilisation jointe des approches d'ingénierie et statistiques



Suivi de processus multi-étapes – Approches SPC

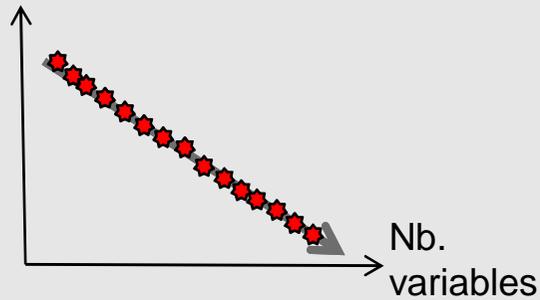


- Cartes de contrôle univariées/multivariées au niveau de l'étape finale
- Suivi séparé des différentes étapes par des cartes de contrôle univariées/multivariées
- Suivi à l'aide de cartes de contrôle des paramètres des modèles régressifs
- Approche d'ajustement des variables par régression (Hawkins, 1991)
- Carte de contrôle « sélection de la cause » (Zhang, 1992)
- Approches APC/PLS (Nomikos & MacGregor, 1995; Kourti & MacGregor, 1996)
- Schémas SPC exploitant la structure détaillée d'un système multi-étapes: « Process-Oriented Basis Representation » - PROREB (Xiang & Tsung, 2008; Li & Tsung, 2009)
- Décomposition causale de la statistique T^2 en utilisant des réseaux Bayésiens (Li et al., 2008)

Suivi de processus dans des espaces de grandes dimensions



Performances de CCM



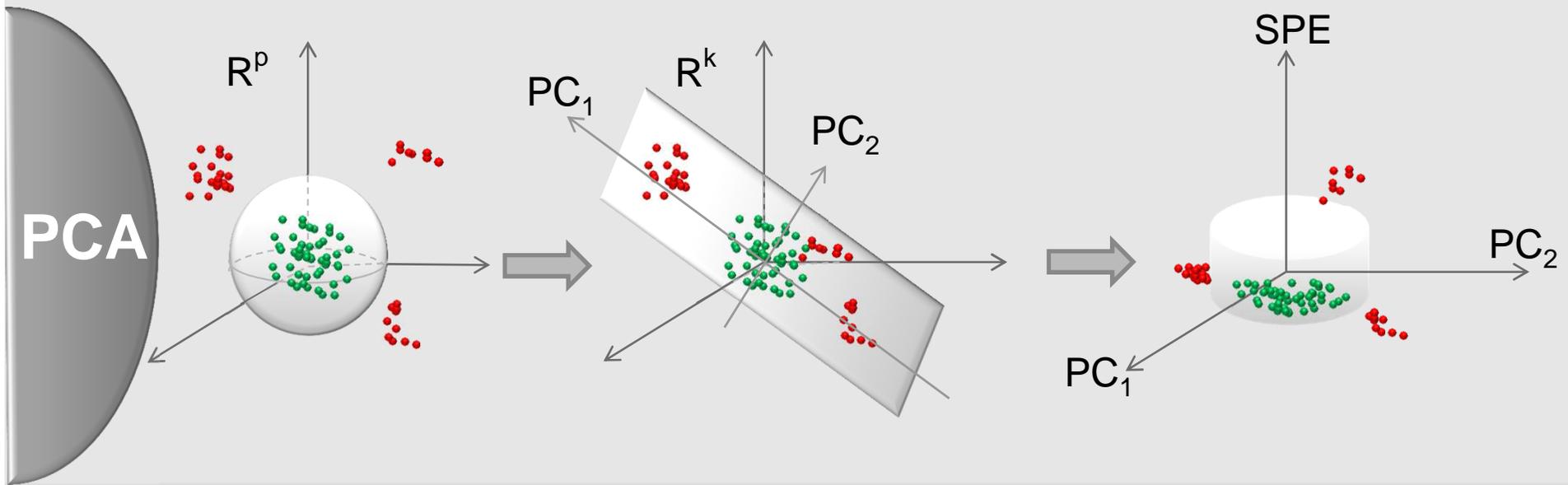
Approches de projection

ACP, PLS

$$T^2 = \frac{PC_1^2}{\lambda_1} + \frac{PC_2^2}{\lambda_2} + \dots + \frac{PC_p^2}{\lambda_p}$$

(Jackson, 1991)

$$T^2 = \underbrace{\sum_{i=1}^k \frac{PC_i^2}{\lambda_i}}_{\text{modèle}} + \underbrace{\sum_{i=k+1}^p \frac{PC_i^2}{\lambda_i}}_{\text{résidus}}$$

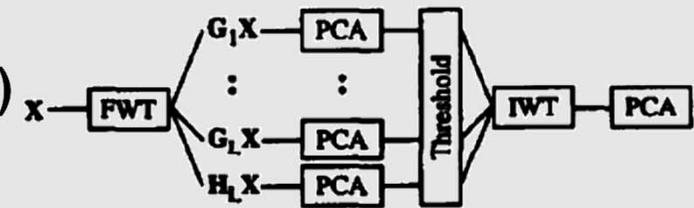
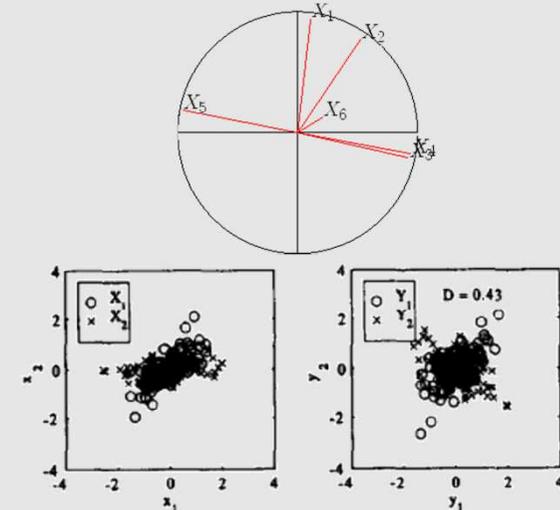


Détection et diagnostic par des approches de projection dans des structures latentes

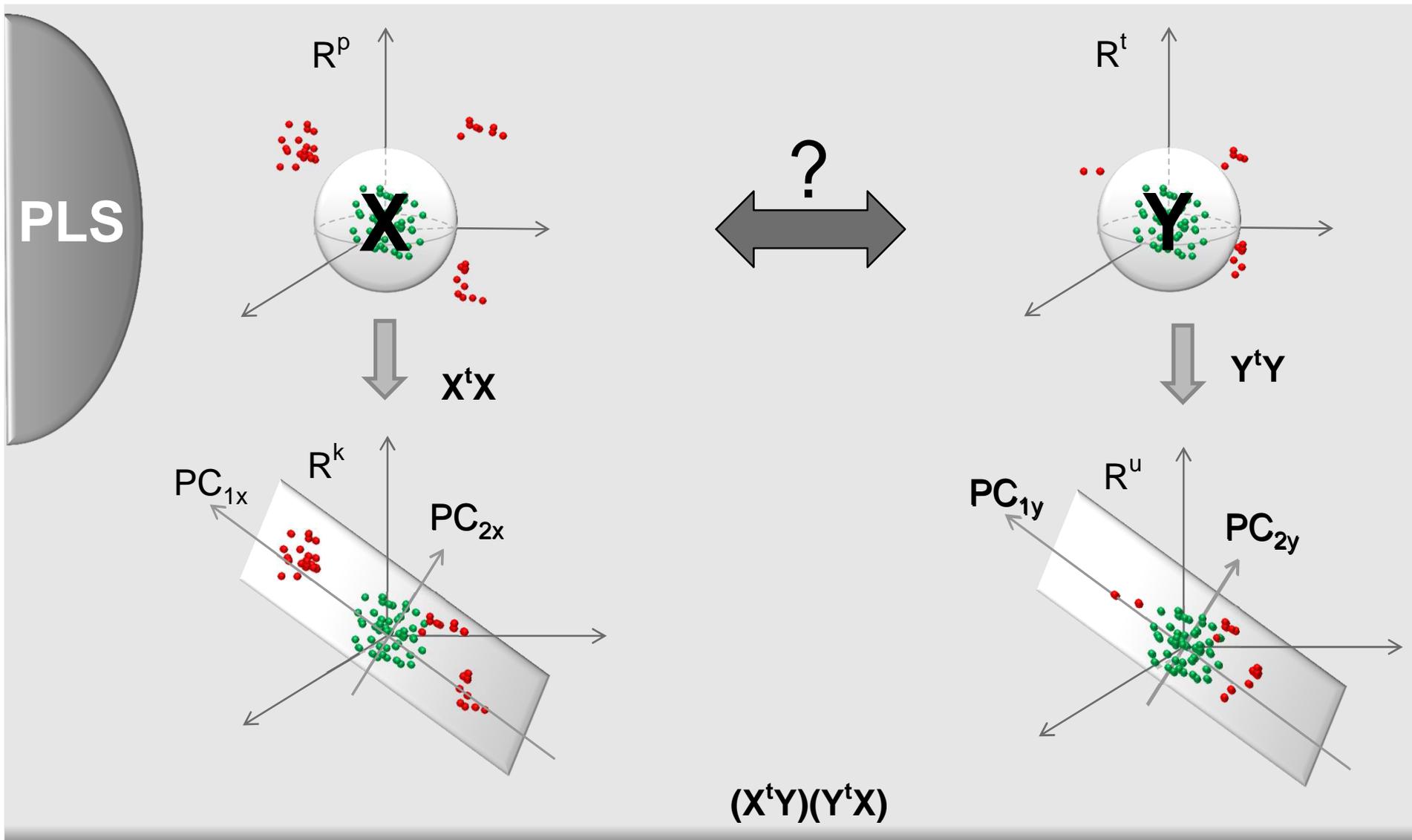


PCA
approches
dérivées

- MPCA (Moving PCA)
- DISSIM
- MS-PCA (Multi Scale PCA)
- MS – MPCA
- MS – DISSIM
- Multi-way PCA (Mw-PCA)



Détection et diagnostic par des approches de projection dans des structures latentes

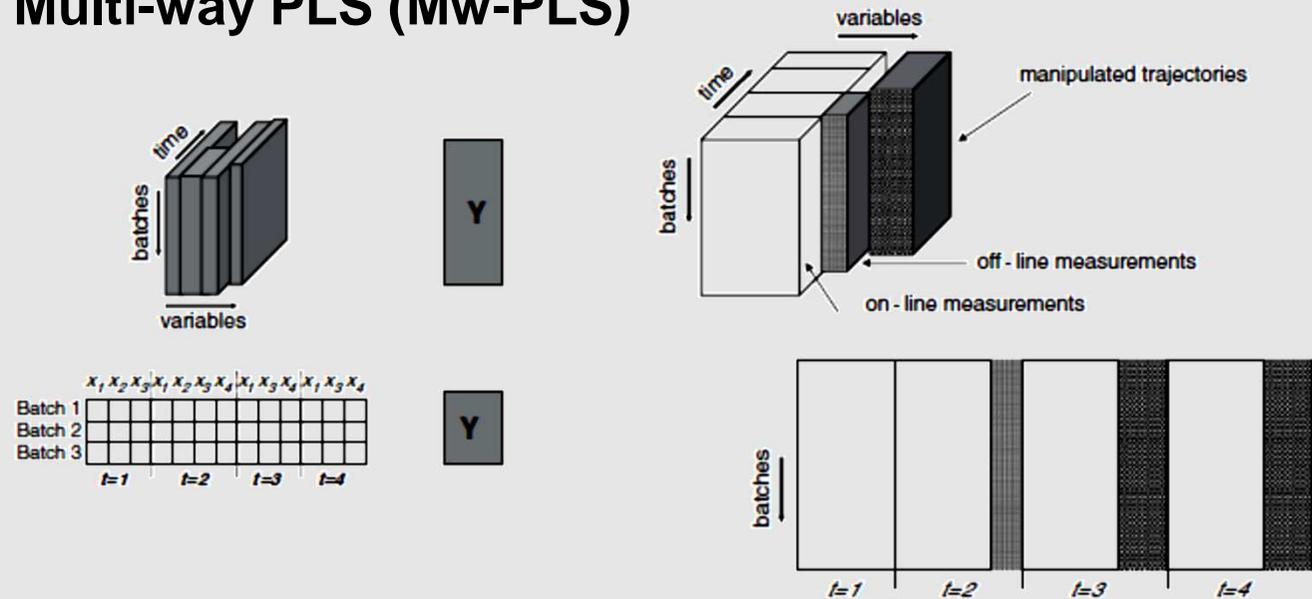


Détection et diagnostic par des approches de projection dans des structures latentes

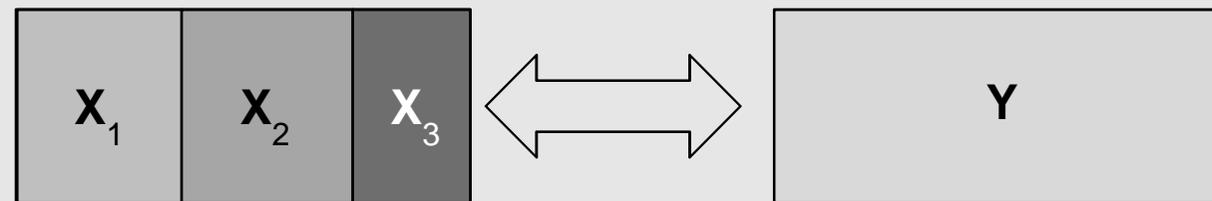


PLS
approches
dérivées

- Multi-way PLS (Mw-PLS)



- Multi-bloc PLS (MB-PLS)

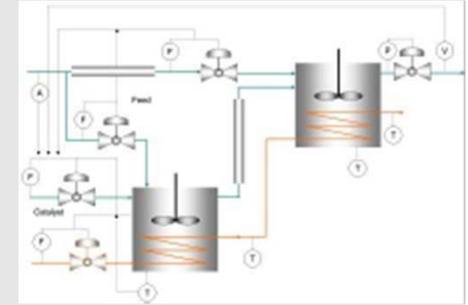


Processus par lot



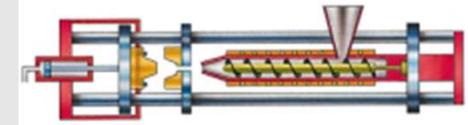
■ Exemples

- fabrication des polymères, de produits pharmaceutiques ou biochimiques
- la séparation et la transformation de matières par distillation et cristallisation par lots
- moulage par injection



■ Caractéristiques

- volume important de données collectées sur un grand nombre de variables (températures, pressions, débits, etc.)



■ Difficultés

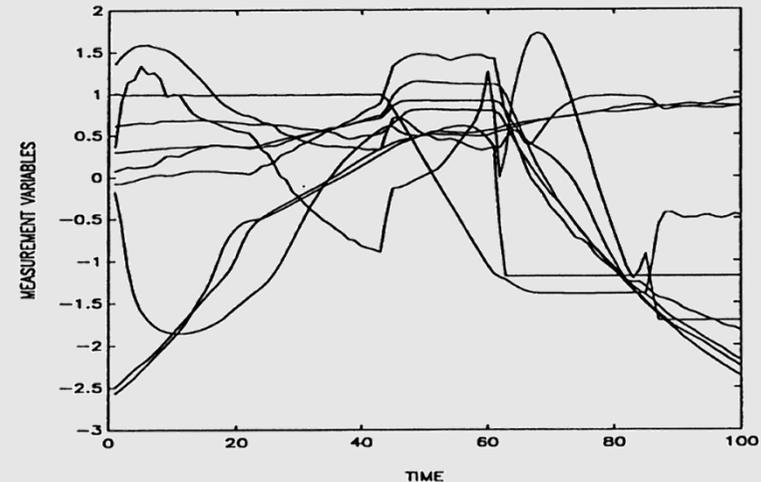
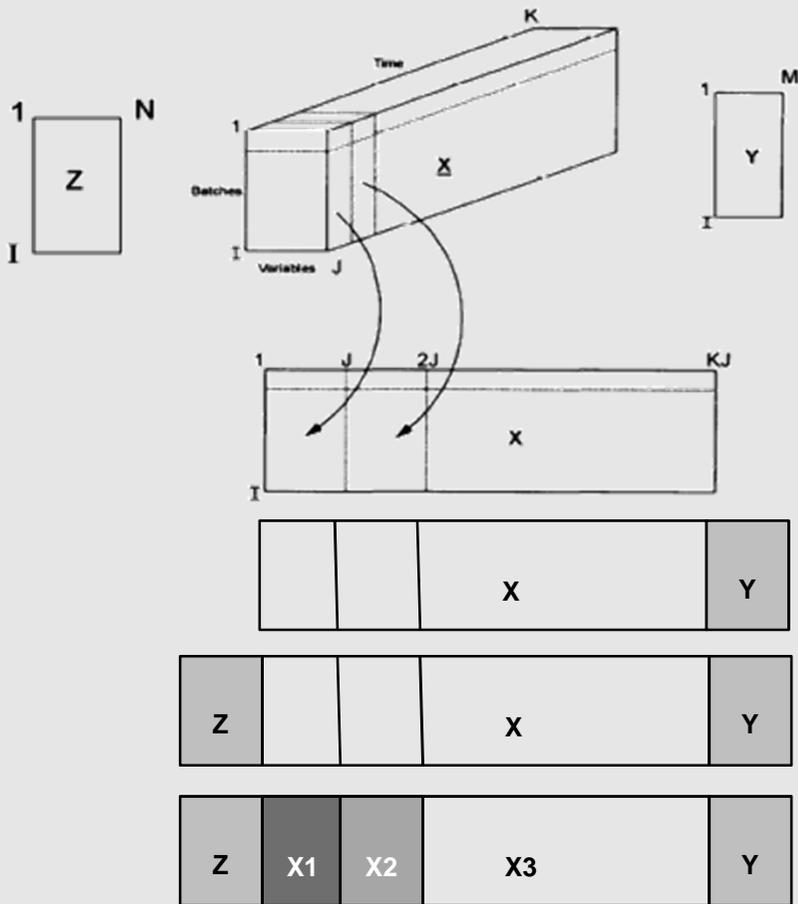
- manipulation d'un grand nombre de variables non-linéaires qui varient dans le temps et qui sont fortement corrélées
- non-stationnarité de variables



Processus de type « lot »



Organisation de données

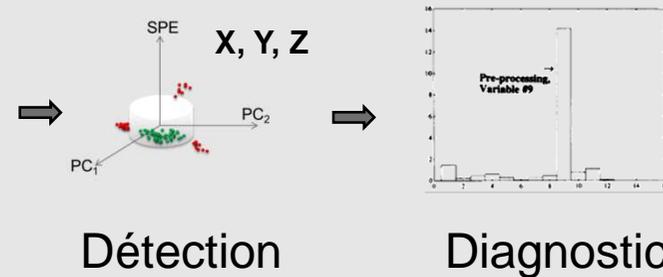


⇒ Mw-PCA

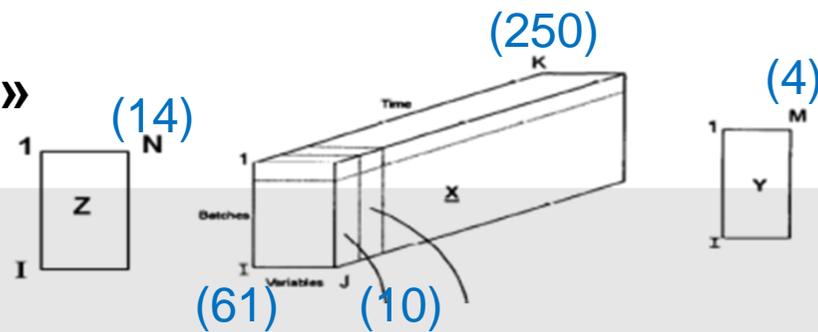
⇒ Mw-PLS

⇒ MB-PLS

⇒ MB-PLS



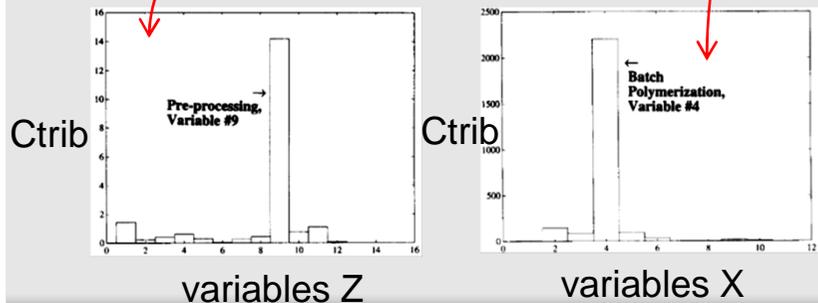
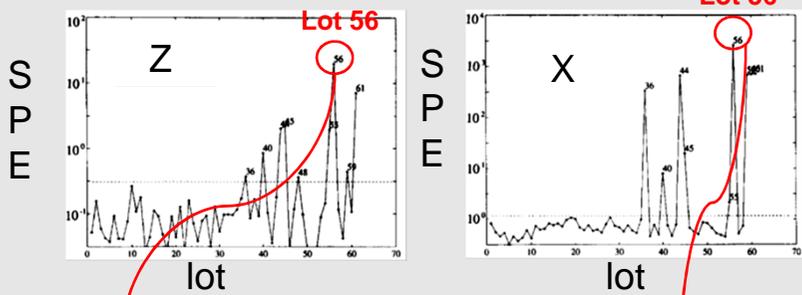
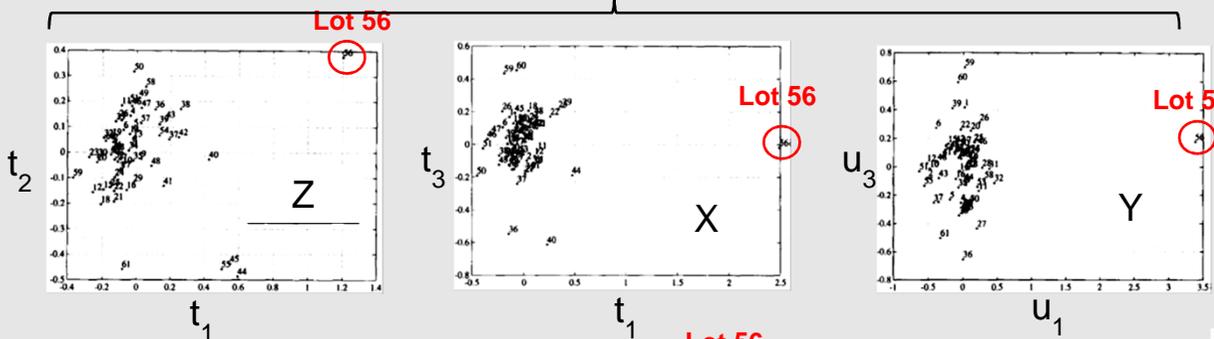
Processus de type « lot »



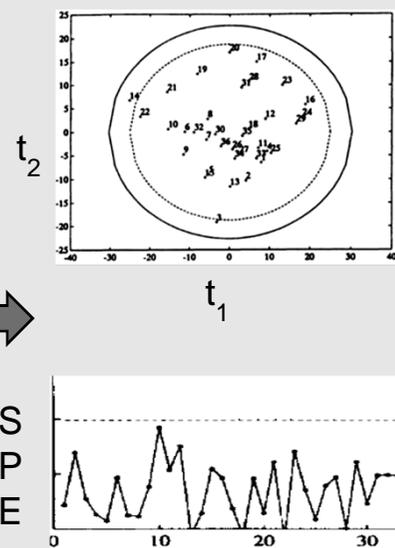
■ Illustration

MB-PLS

Phase 1 – Analyse préliminaire des données
(déterminer un jeu de données SC pour la Phase II)



Décision
écarter le lot 56

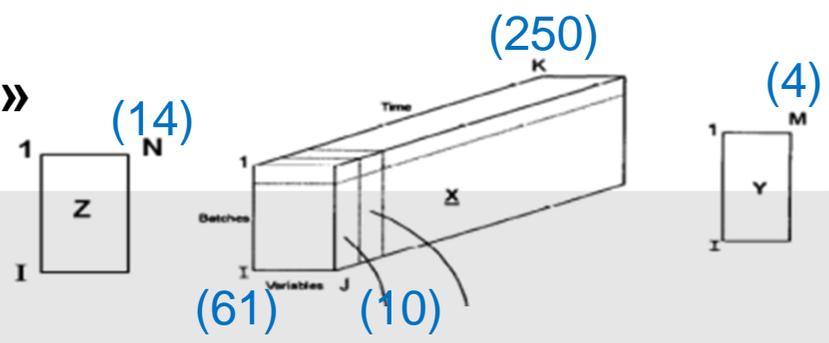


Modèle MB-PLS validé

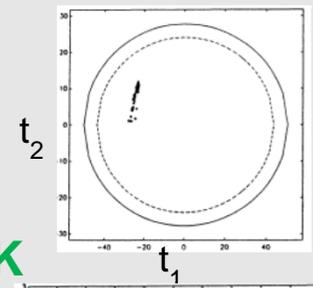
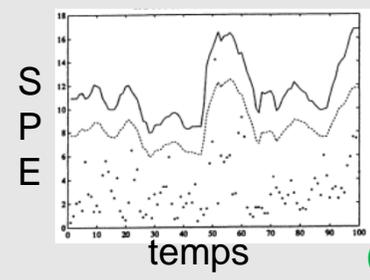
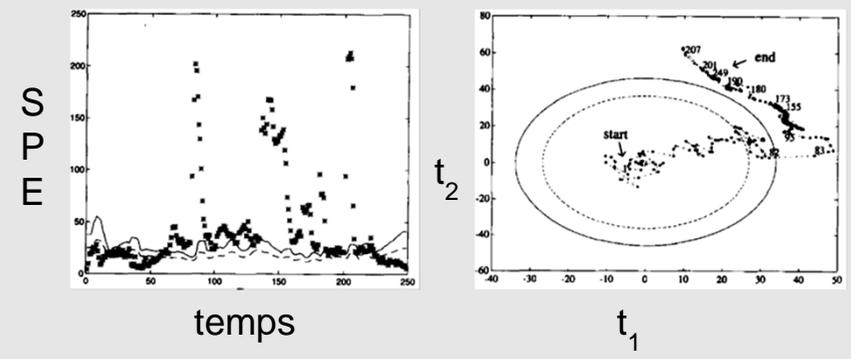
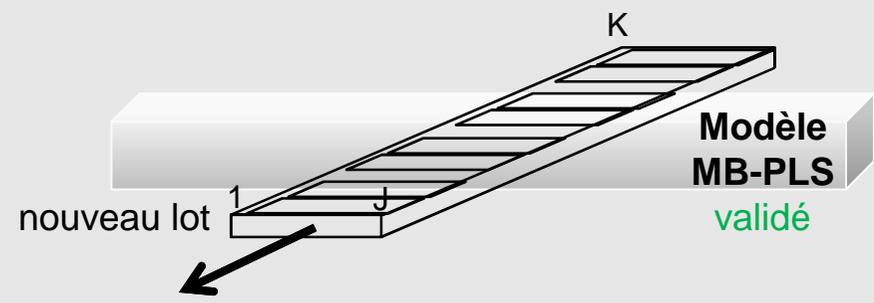
Processus de type « lot »



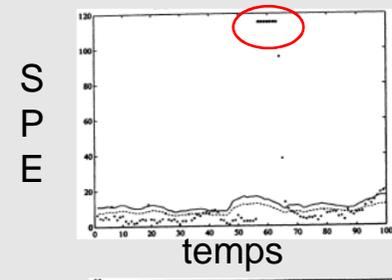
Illustration



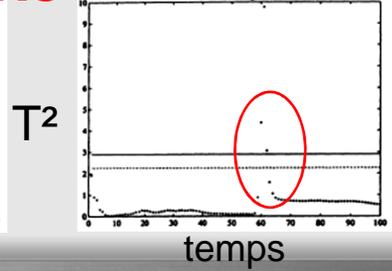
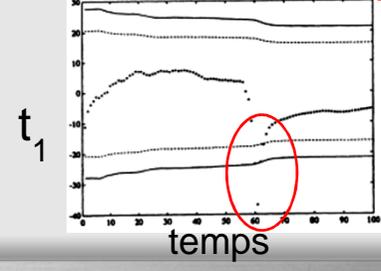
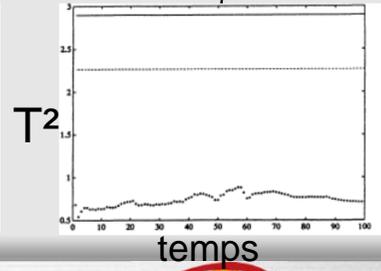
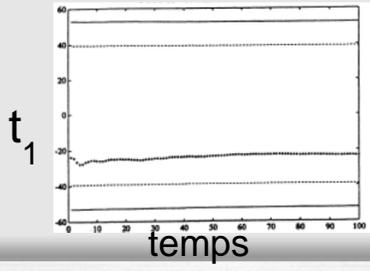
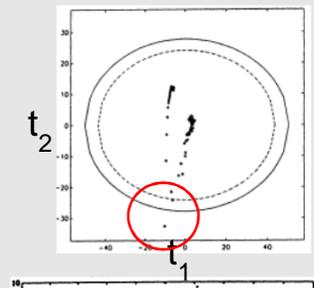
Phase II – Suivi on-line



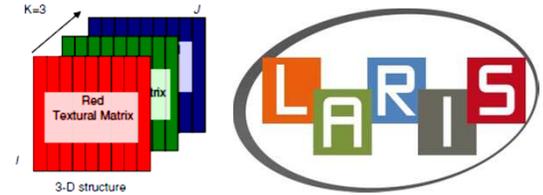
OK



KO



Surveillance par le biais des systèmes de vision artificielle

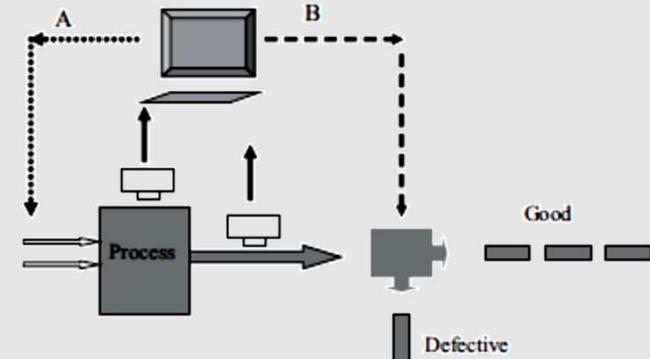


■ Applications industrielles des Systèmes de Vision Artificielle (SVA)

- détection des défauts dimensionnels
- détection des défauts structurels (manque de composants dans les assemblages)
- détection de défauts de surface (rayures, fissures, rugosité, mauvaise finition, etc.)
- défauts opérationnels

■ Domaines industriels d'application

- industries textile, du papier, du plastique, etc.
- industrie alimentaire



Surveillance par le biais des systèmes de vision artificielle



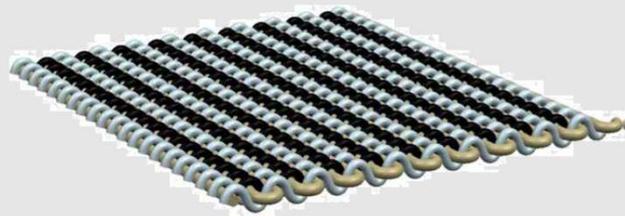
SVA
+
SPC

- **Cartes de contrôle univariées**
 - pour des contrôles dimensionnels ([Horst & Negin, 1992](#); [Tan and al., 1996](#))
 - changement de couleur ([Nembhard and al., 2003](#))
 - usure de forets ([Liang & Chiou, 2008](#))
- **Cartes de contrôle multivariées**
 - suivi de défauts dans les circuits intégrés ([Tong et al., 2005](#))
 - suivi de l'apparence et de l'esthétique de comptoirs en pierre reconstituée (détection de rayures, remous et ondulations) – [Liu & McGregor, 2006](#).
 - défauts dans les composants électroniques ([Lin, 2007](#))
 - système expert d'inspection visuelle automatique ([Lyu & Chen, 2009](#))
- **Utilisation des approches pour les profils**
 - détection de défauts dans des afficheurs LCD (diagramme Q-Q) – [Wang & Tsung, 2005](#)
 - détection de défauts au niveau des surfaces métalliques ([Colosimo et al., 2010](#))

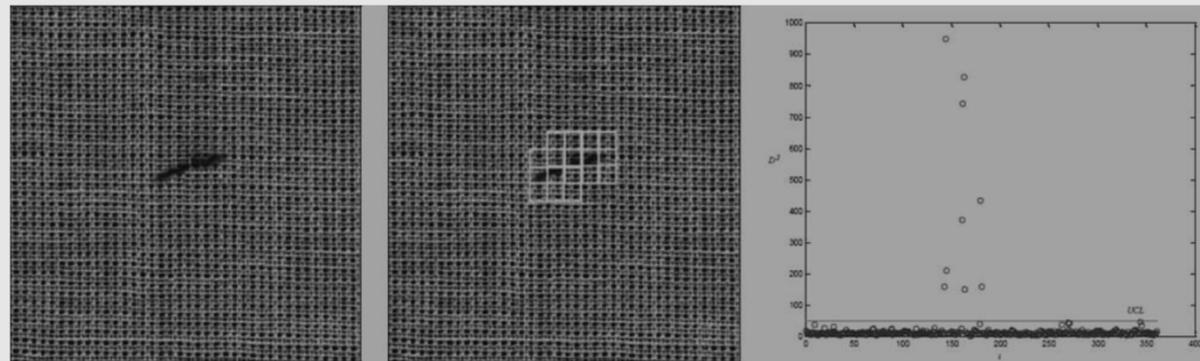
Surveillance par le biais des systèmes de vision artificielle



- Utilisation de cartes de contrôle dans le contexte de la surveillance spatiale
 - l'axe horizontal de la carte représente la position dans l'image
 - utilisation spatiale par fenêtre glissante



Tunàk and Linka (2008)



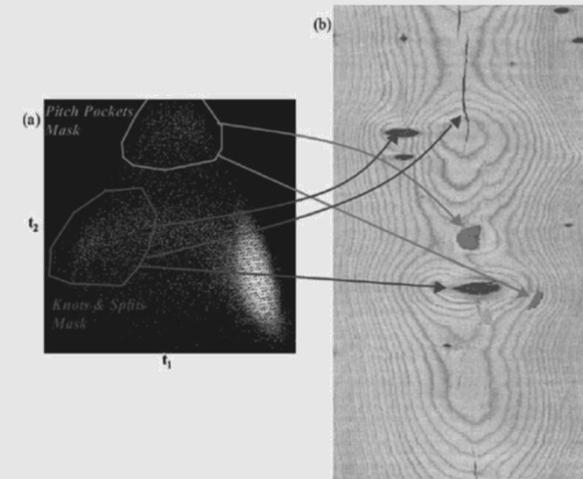
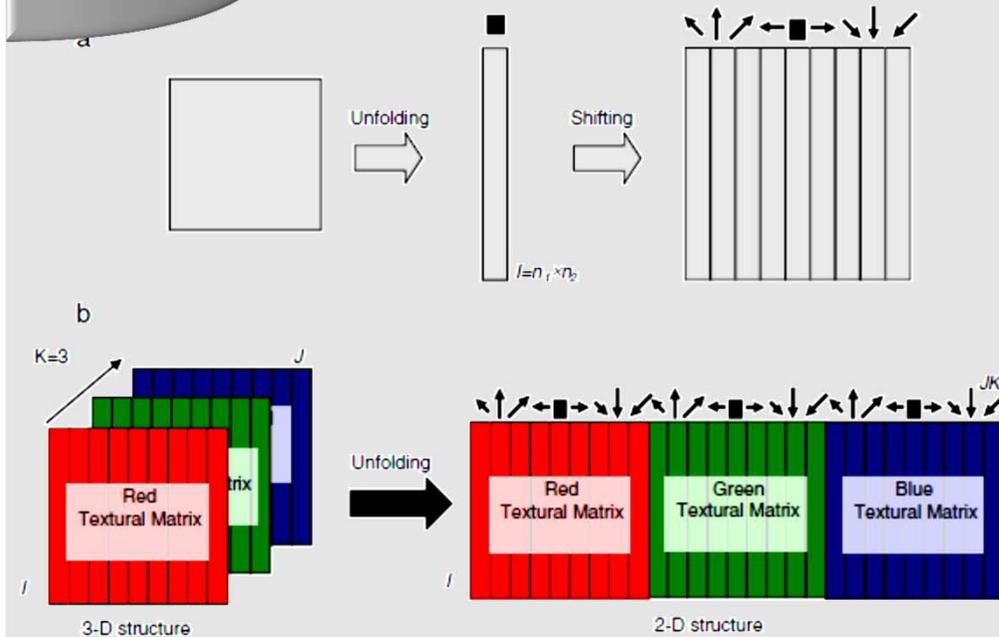
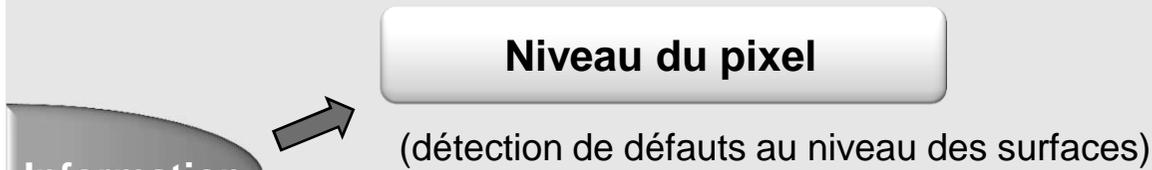
(détection de défaut dans le textile)

Carte T^2

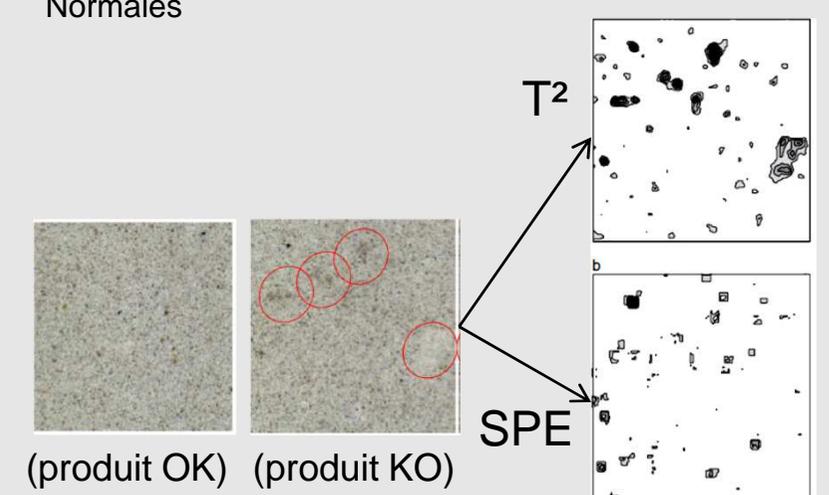
Surveillance par le biais des systèmes de vision artificielle



- Analyse Multivariée d'Imagés (AMI)



Conditions → (PCA, PLS, etc.) → (T², SPE)
Normales



Surveillance par le biais des systèmes de vision artificielle

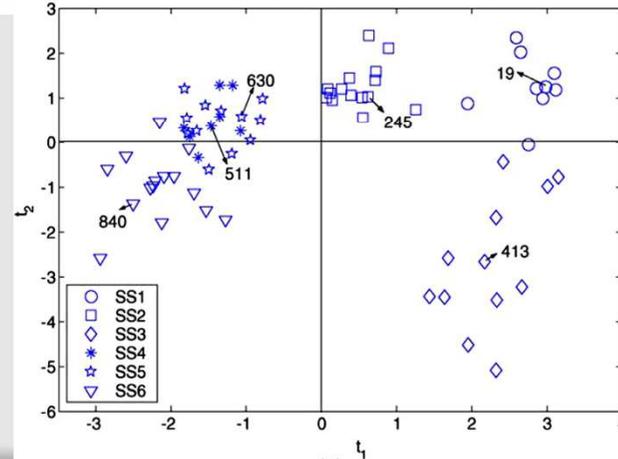
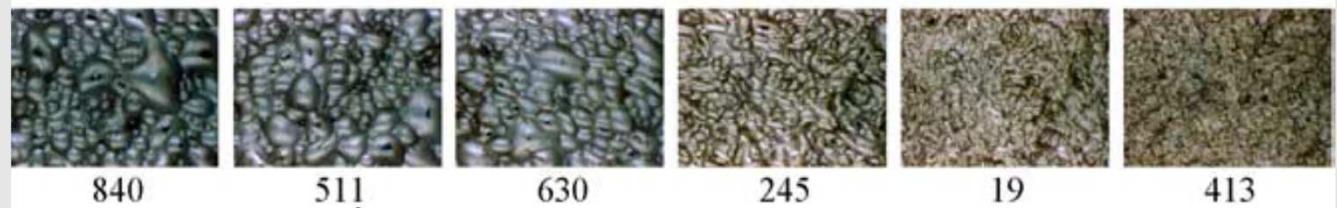
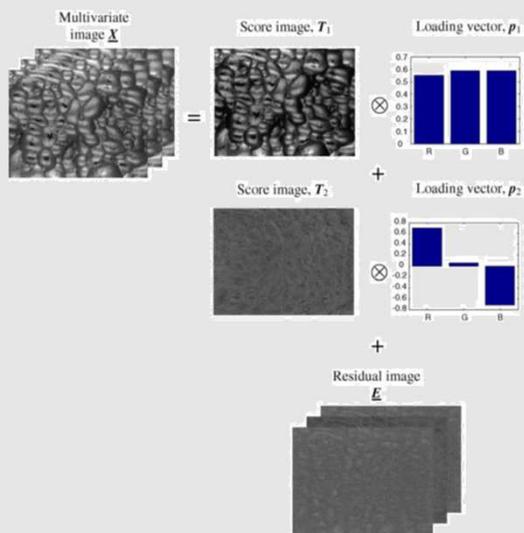


- Analyse Multivariée d'Imagés (AMI)

Niveau de l'image

Information

Aide à la décision
 (détection de changements d'une phase à l'autre)
 (problèmes de classement; de tri; etc.)



(Liu and al., 2005)



CONCLUSION

L'apport de la Maîtrise Statistique de Processus



- Compréhension de la **variabilité** dans le processus
- Décision au sujet de l'**état du processus** (sous-contrôle, hors-contrôle)
- Décision au sujet de la **stabilité** du processus (avant toute étude de capabilité)
- Amélioration de la Qualité et de la Productivité des processus industriels et de services



“There has rarely been a technological invention like statistical quality control, which is so wide in its application yet so simple in theory, which is so effective in its results yet so easy to adopt and which yields so high a return yet needs so low an investment”.

C. R. Rao (1989)

Le paradoxe de la MSP



"There are few areas of statistical application with a wider gap between methodological development and application than is seen in SPC."

Crowder, Hawkins, Reynolds, and Yashchin (1997)

L'augmentation de la complexité des problèmes rencontrés représente une opportunité pour réduire l'écart entre les applications et la recherche théorique en MSP.

Quelques défis à relever

- Méthodes rapides (**temps réel**) et flexibles d'analyse adaptées au volume et à la complexité de données (**structure, big data**)
- Développement des logiciels et progiciels (**environnements adaptés pour le calcul et la visualisation des résultats**)
- Intégration des approches statistiques et des connaissances d'ingénierie (**approches hybrides de détection et de diagnostic**)
- Jonction entre les approches « change point », MSP et diagnostic
- Approches non-paramétriques **multivariées**
- Approches d'estimation robustes (**en Phase I**)
- Intégration des approches mixtes **APC/SPC**