Process monitoring methods

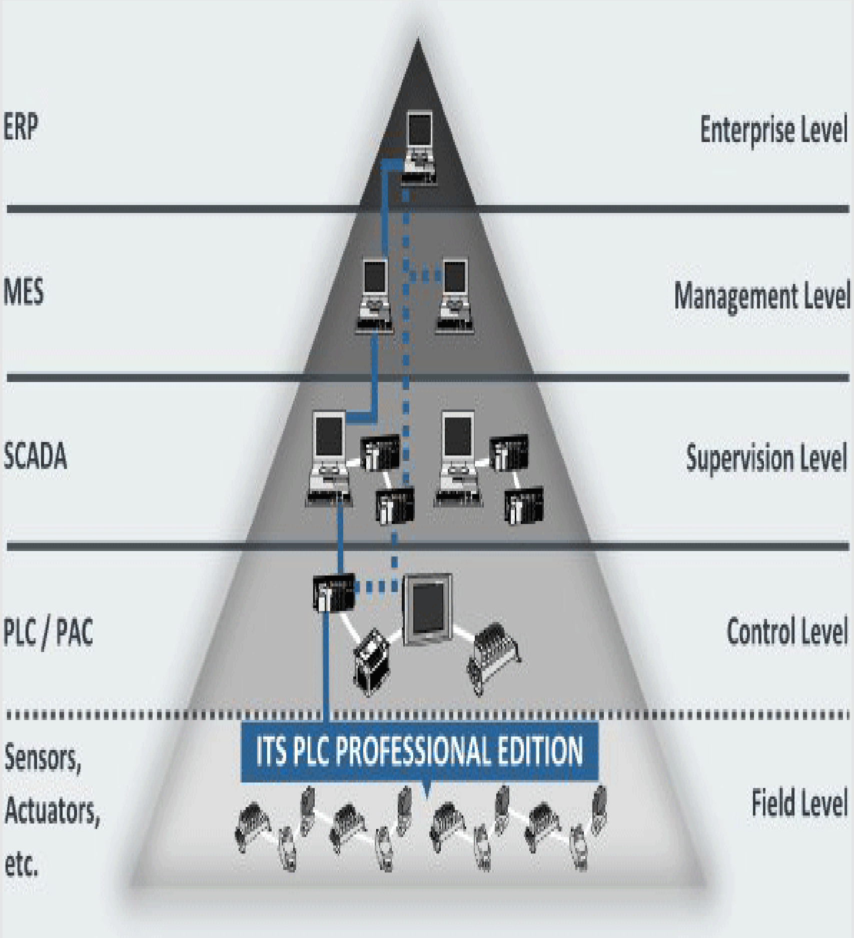
Octavio Pozo Garcia

2ÅF-Automaatika OÜ

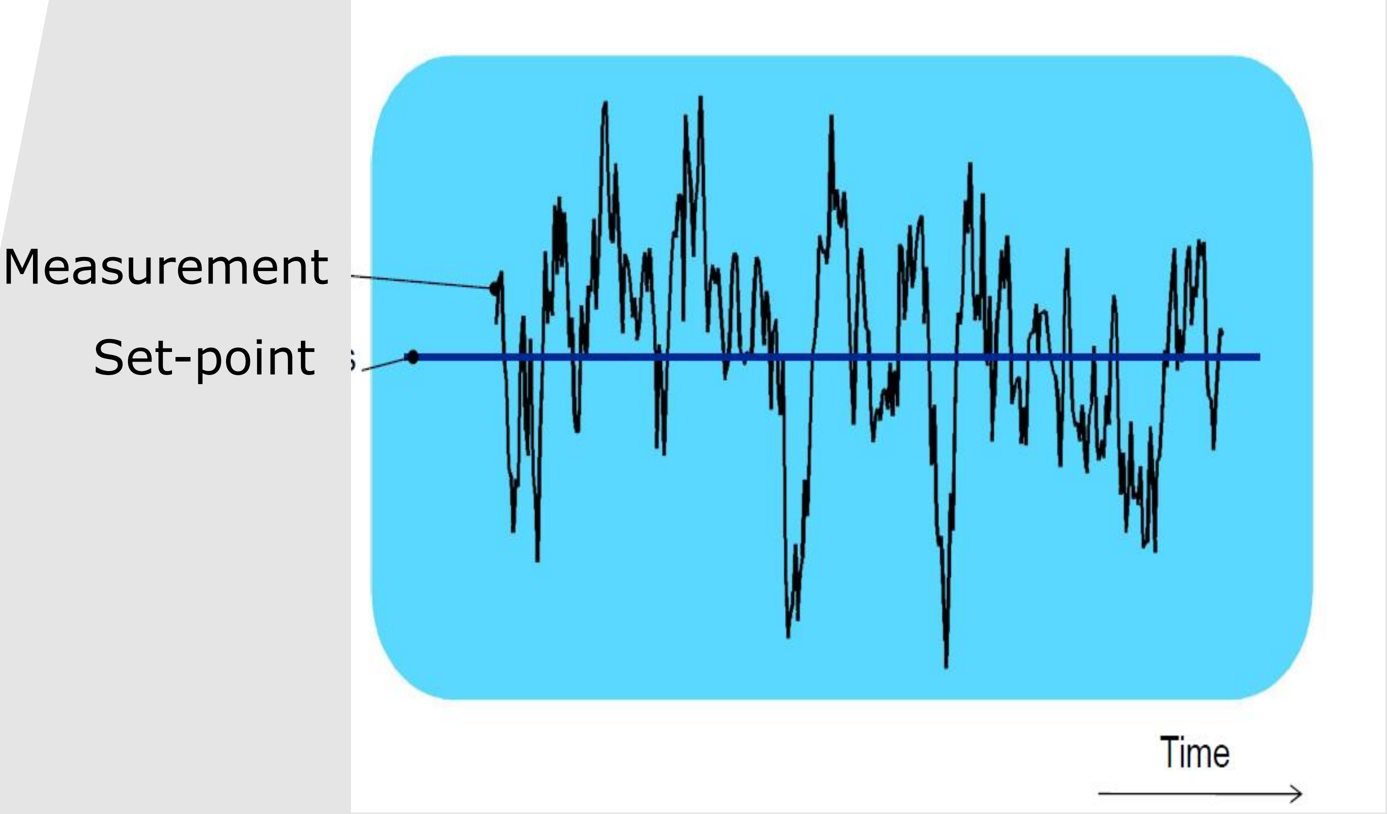
2015

Control systems

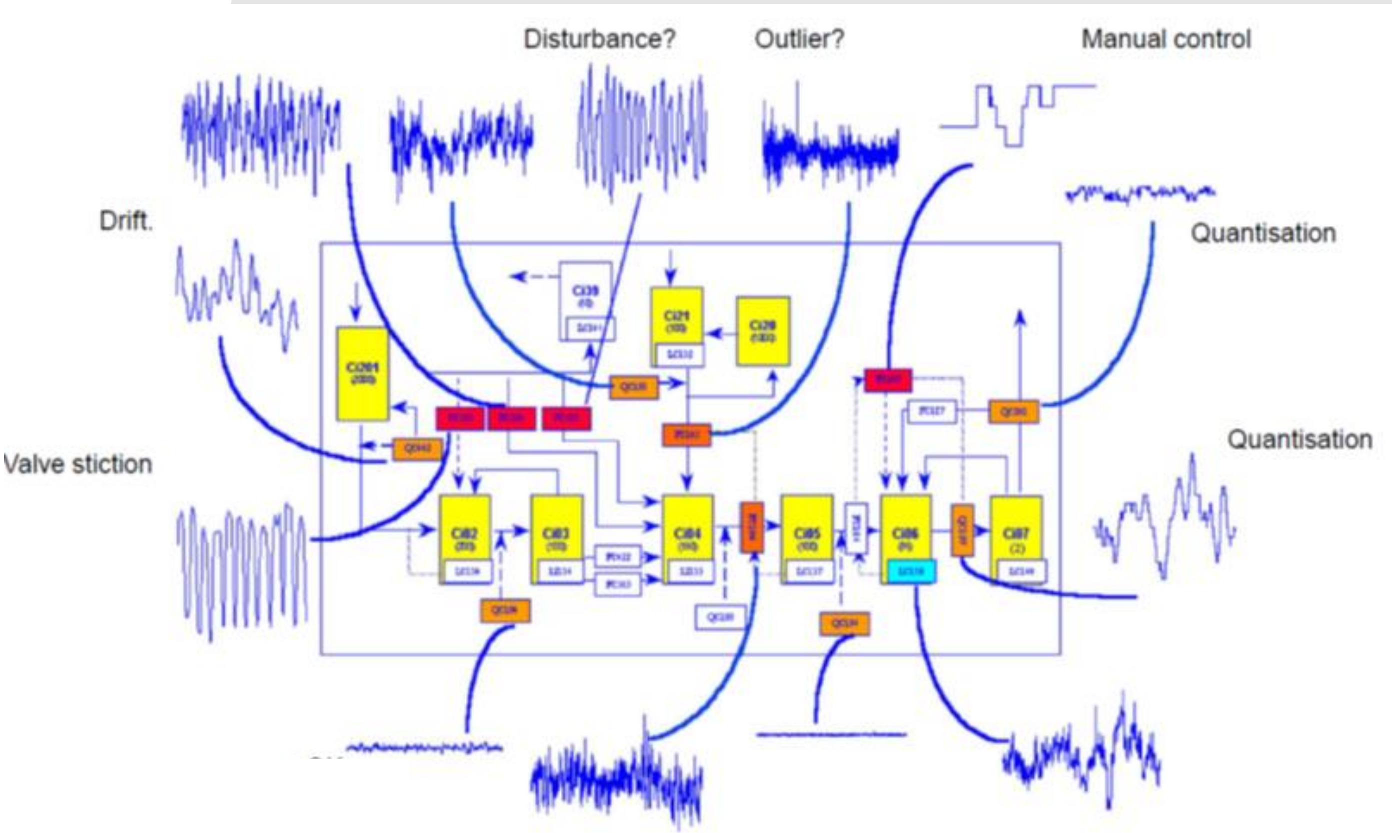
- Regulatory control:
  - Automatic
  - Field devices and sensors
  - Mostly SISO systems
- Supervisory control:
  - Heavy human interaction
  - PLCs, Data acquisition systems, HMI
  - Information is essential



Normal or abnormal behavior?



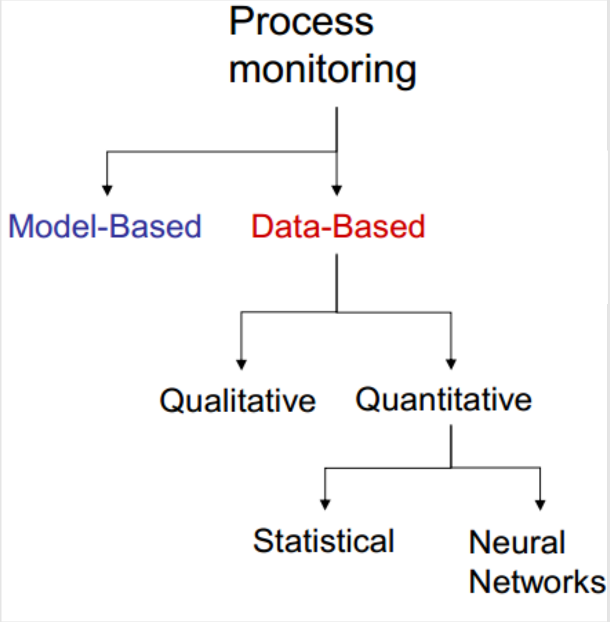
Normal or abnormal behavior?



- Information
  - Quality
  - Quantity
- Process knowledge
  - Training
  - Experience

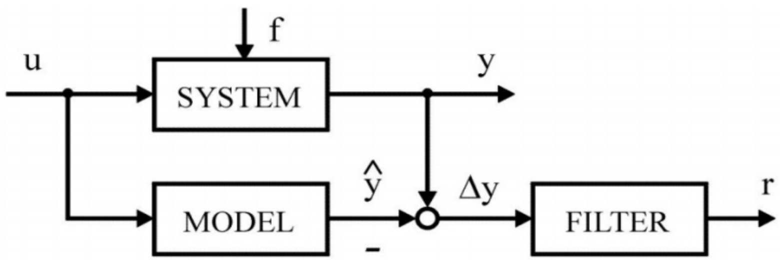
Process monitoring

- Techniques or algorithms that identify and detect changes in the **critical variables** of the process
- Early identification of process abnormal behavior

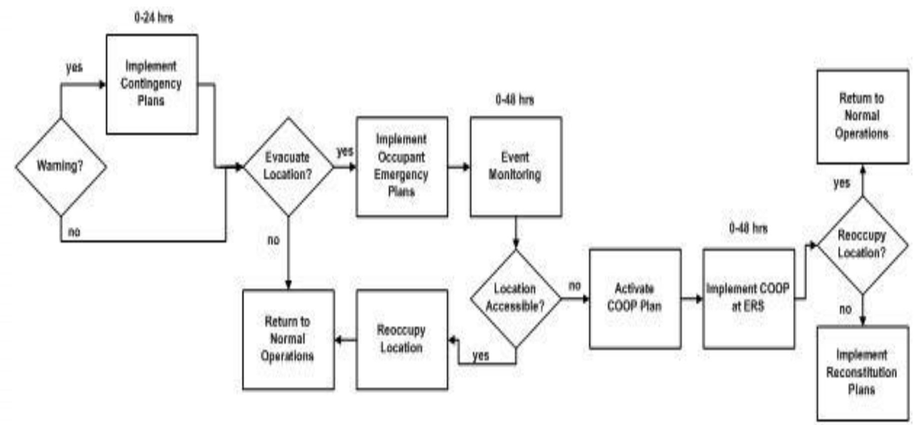


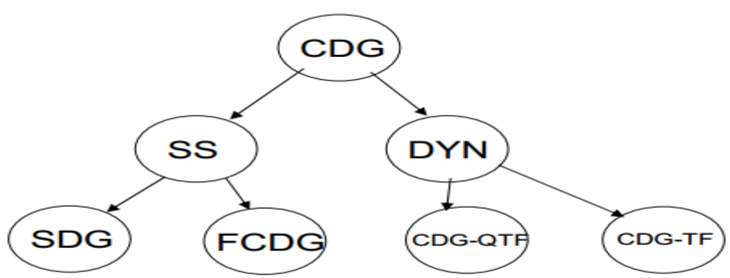
- Model-based:
  - Provide very presice results
  - Requires understanding of physico-chemical relations in the process
  - If the process changes the model is useless
- Data-based:
  - Does not requiere process knowledge
  - Results can depend on the quality of the information
  - Easy to adapt to new processes
  - Requires larges amounts of data

Process monitoring: model based techniques



a)





c)

a) First principles model
b) Expert system (decision trees)
c) Causal model

Process monitoring: Data based techniques

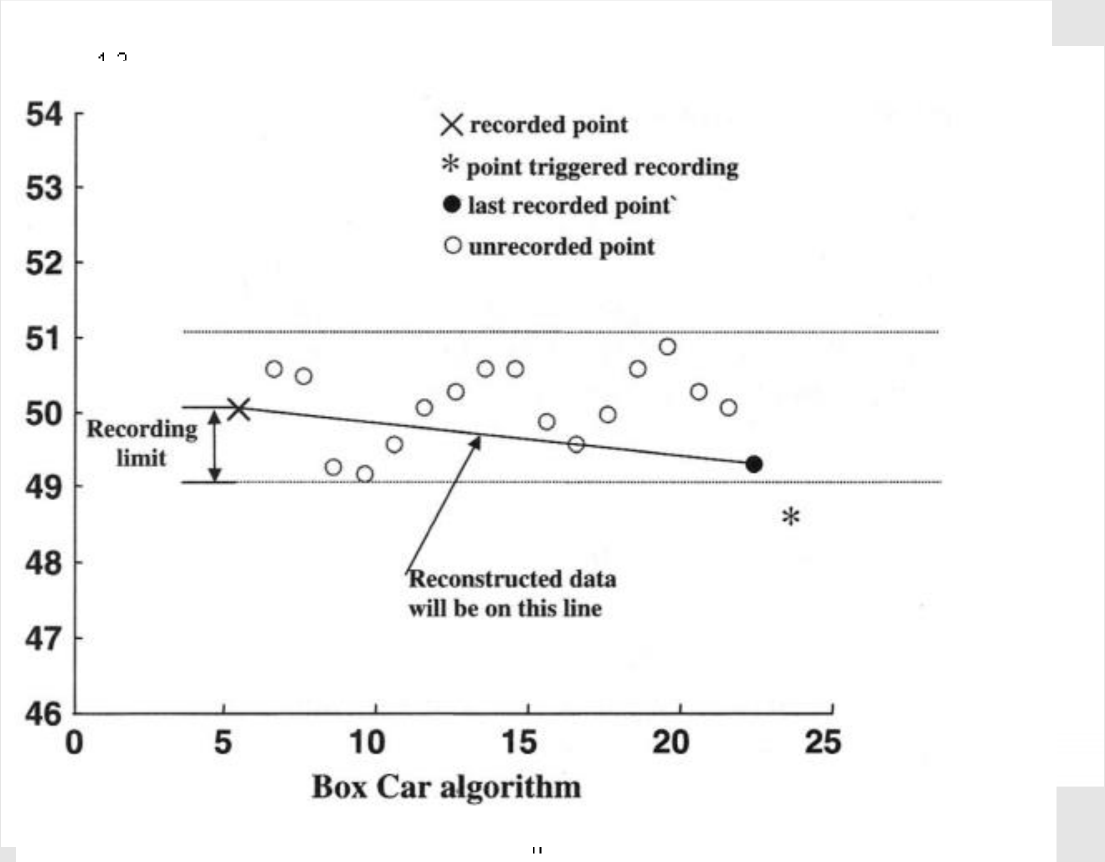
- Statistical
  - Most simple type of monitoring
  - Individual thresholds determined for each variable
  - Correlation between different variables can considered to create more complex systems
- Artificial intelitence
  - Automated reasoning systems
  - Use data to make inferences about the process
  - Captures diverce process behavior

Data preprocessing

- Industrial data is never suitable due to:
  - Noise
  - Uncertainties
  - Disturbances
- Data preparation requires:
  - Selection
  - Varied operation conditions
- Application determine the procedure

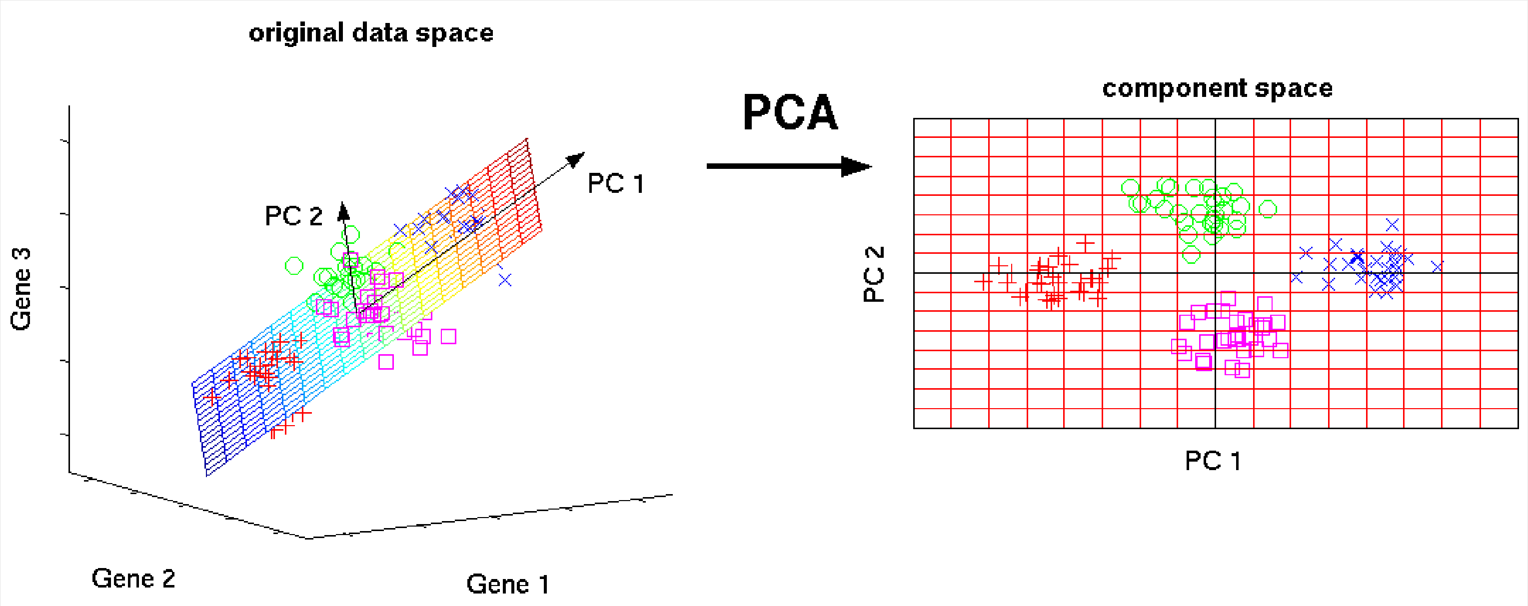
Data preprocessing general recomendations

- Remove Outliers
- Handle missing values
- Resampling
- Scaling
- Filtering
- Compression



Process monitoring methods: Principal component analysis (PCA)

- Statistical procedure
- Describes the variation inside the data
- Reduce complex data sets to a lower dimension
- Symplified dynamics of the system



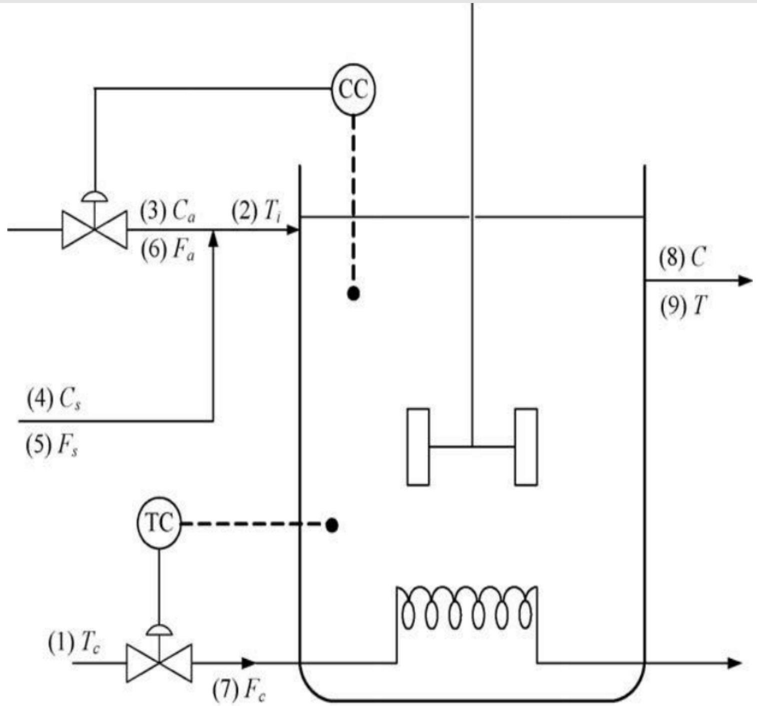
PCA steps

1. Normalize the original data set.
2. Compute the covariance matrix
3. Compute the eigenvalues and eigenvectors.
4. Determine the principal components.
5. Calculate score matrix
6. Determine the monitoring indexes(square predition error or other statistics)

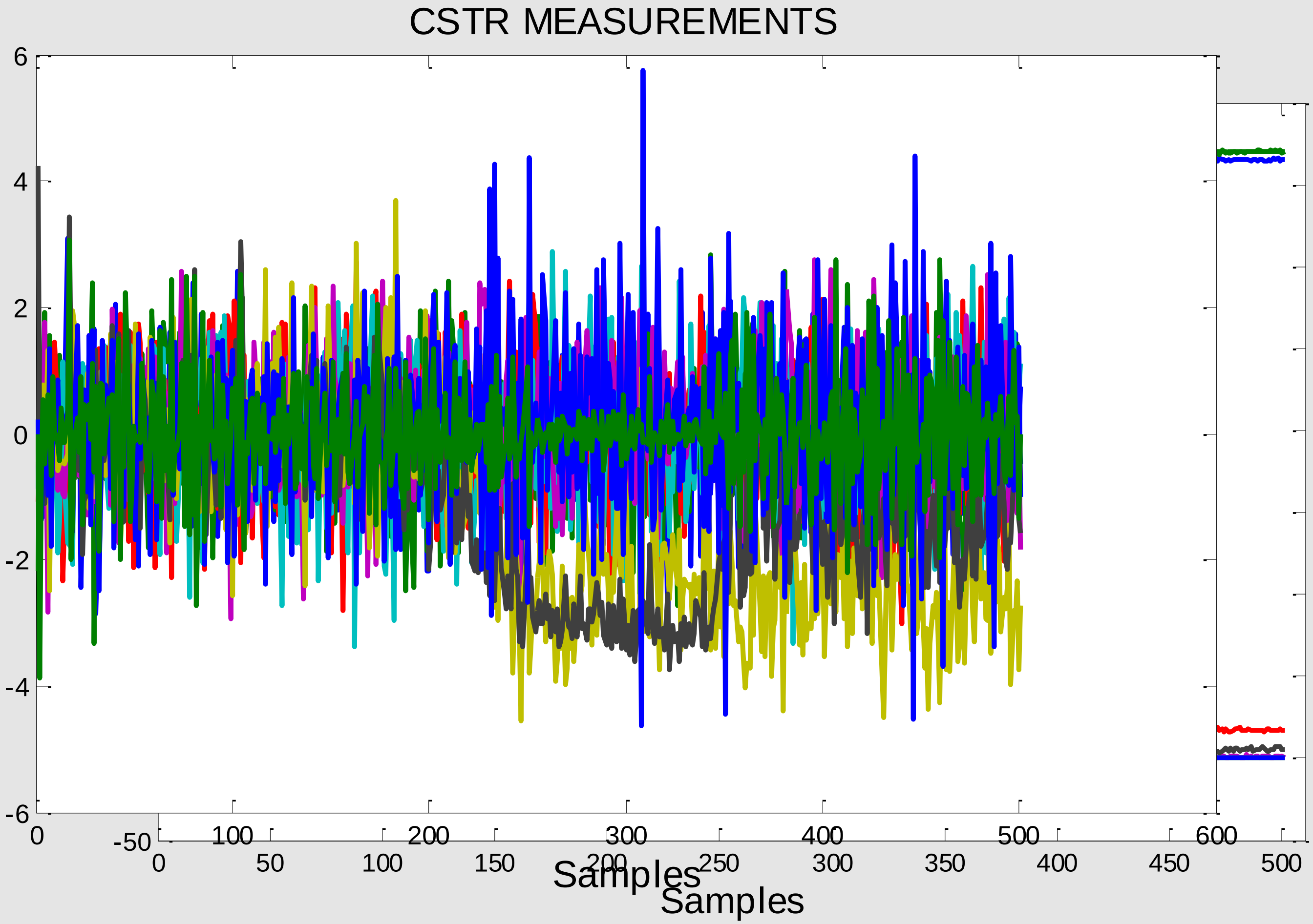
PCA: CSTR example

Measurements:

- Coolant Temperature
- Reactant mixture temperature
- Reactant A concentration in feed
- Reactant A concentration in solvent
- Solvent flow rate
- Feed flow rate
- Coolant flow rate
- Outlet concentration
- Outlet temperature



Data



CSTR MEASUREMENTS
6
4
2
0
-2
-4
-6
0
100
200
300
400
500
600
-50
0
50
100
150
200
250
300
350
400
450
500
Samples
Samples

Covariance matrix and decomposition

Covariance matrix C $C = \frac{1}{N-1} X^T X$

The covariance matrix can be decomposed:

$$C = V \Lambda V^T$$

$$V = \begin{bmatrix} e_1 & \cdots & e_M \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_M \end{bmatrix}$$

The decomposition can be done by solving_

$$\det(C - \lambda I) = 0$$

$$Ce_i = \lambda_i e_i$$

Determine principal components

- The principal components are the eigenvectors with the largest eigenvalues which correspond to the dimensions that have the strongest correlation in the dataset.

Lambda	Variance %	VarianceTotal
2,07	23	23
1,41	16	39
1,31	15	53
0,99	11	64
0,95	11	75
0,83	9	84
0,79	9	93
0,38	4	97
0,26	3	100

Calculate score matrix (PCA model)

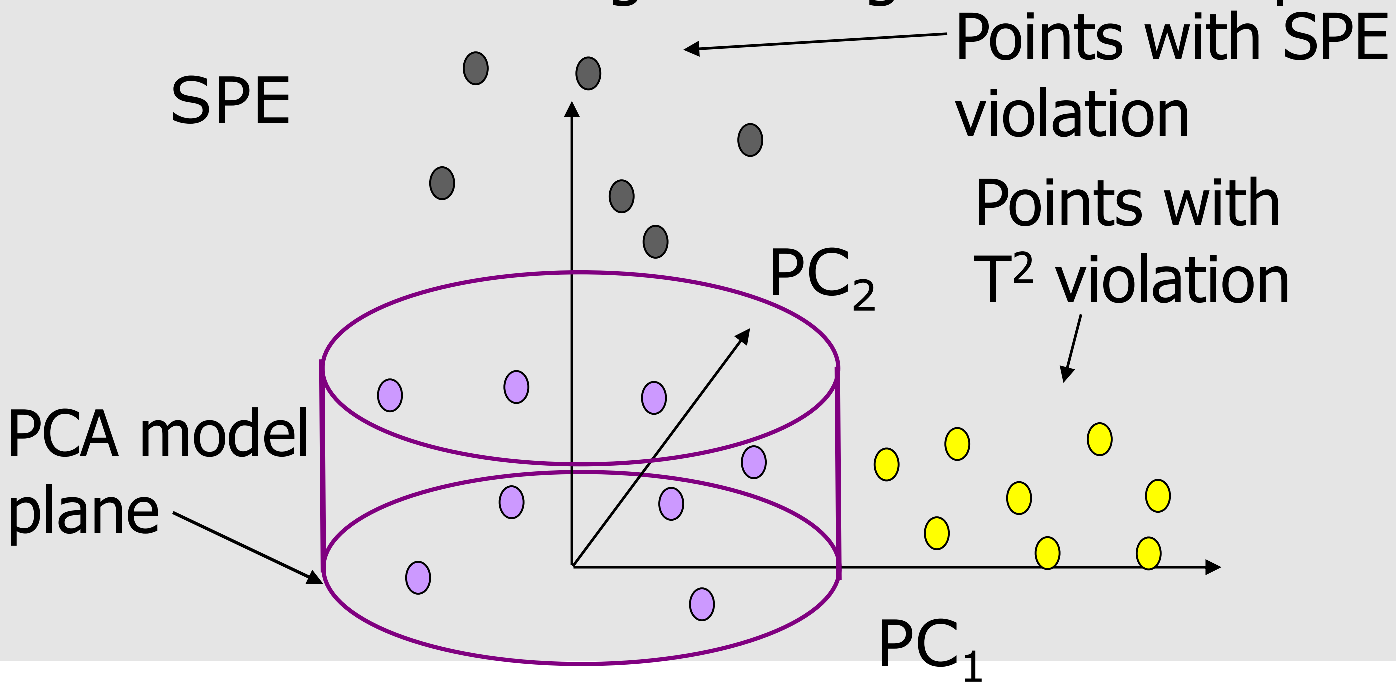
- A score-matrix T can be calculated:

$$\mathbf{T} = \mathbf{X}\mathbf{V}_\mathbf{k}$$

- V_k is a transformation matrix containing the the eigenvectors corresponding to the *k* largest eigenvalues. *k* (principal components)
- T is a ”compressed” version of X

Monitoring indexes

- SPE – variation outside the model, distance off the model plane
- Hotelling T^2 – variation inside the model, distance from the origin along the model plane



Monitoring indexes

$$T^2_{lim} = \frac{K(N-1)}{N-K} F(K, N-K, \alpha)$$

where the $F(K,N\text{-}K,\alpha)$ corresponds to the probability point on the F-distribution with (K,N-K) degrees of freedom and confidence level α. N = # of data samples, K = # of PCs

$$SPE_\alpha = \Theta_1 \left[\frac{c_\alpha h_0 \sqrt{2\Theta_2}}{\Theta_1} + 1 + \frac{\Theta_2 h_0 (h_0 - 1)}{\Theta_1^2} \right]^{\frac{1}{h_0}}$$

"Unused" eigenvalues

$$\Theta_i = \sum_{j=k+1}^{m} \lambda_j^i \quad h_0 = 1 - \frac{2\Theta_1\Theta_3}{3\Theta_2^2}$$

where m= number of original variables, k=number of principal components in the model, c_α= upper limit from normal distribution with conf. level α

Monitor new data

Scale the new data set with training data scaling values

$$\tilde{x}_{new} = \frac{\left(x_{new} - \bar{x}_{train}\right)}{\sigma_{train}}$$

Compute PCA transformation (i.e. scores for all the chosen principal components) using V_k.

$$t = V_k^T \tilde{x}_{new}$$

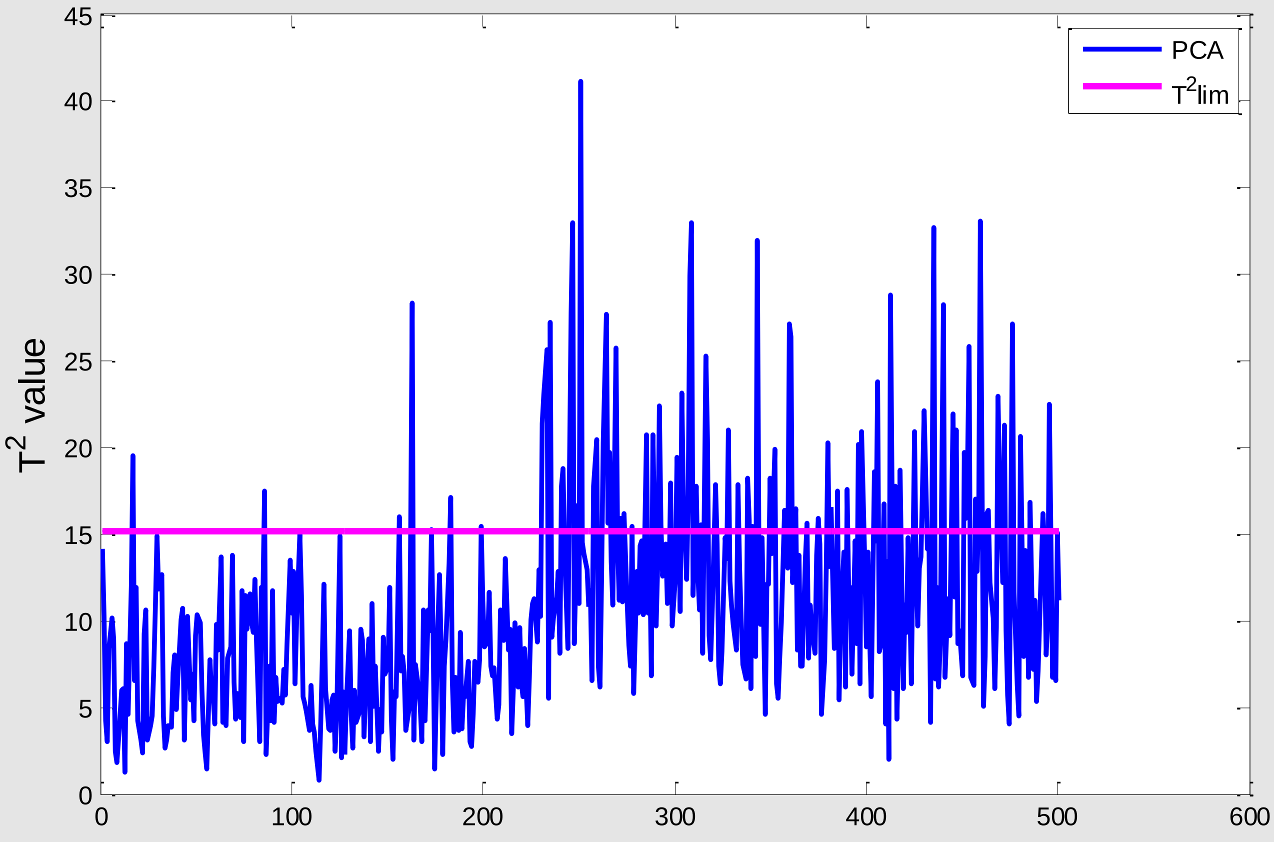
Compare Hottelling and SPE values

- T^2: Measures systematic variations of the process. For individual observation:

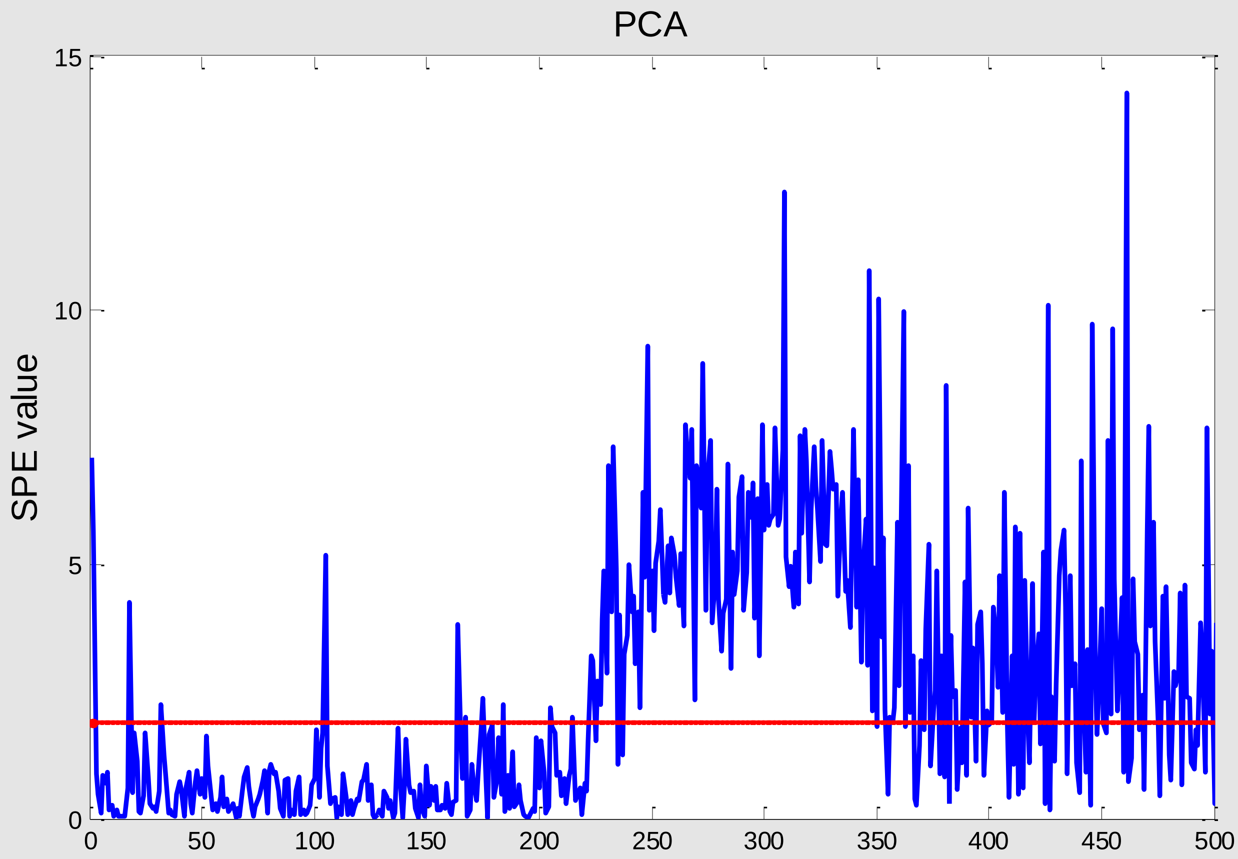
$$T^2(x_{new}) = \tilde{x}_{new}^T V_k \Lambda_K^{-1} V_k^T \tilde{x}_{new} = t^T \Lambda_K^{-1} t$$

- SPE: Measures the random variations of the process

$$SPE = r^T r \quad \text{where}\; r = (I - V_k V_k^T)\tilde{x}_{new}$$



PCA
T^2lim
T^2 value
0
5
10
15
20
25
30
35
40
45
0
100
200
300
400
500
600


PCA
SPE value
15
10
5
0
0
50
100
150
200
250
300
350
400
450
500