Process Monitoring with Canonical Variate Analysis

Steven Xu

Agenda

- Canonical variate analysis
  1. Concept review
  2. How it works
- Kernel canonical variate analysis
  1. Kernel Method
  2. How it works
  3. Pros and Cons
- Application
- Other variants

Process Monitoring with Canonical Variate Analysis

Classical canonical correlation

- Describe to what extend two random vectors are linearly correlated.

- Let $\mathbf{X} = (X_1 \quad X_2 \quad \dots \quad X_p)^T$, $\mathbf{Y} = (Y_1 \quad Y_2 \quad \dots \quad Y_q)^T$ be two centered random vectors with finite second moments.

1 $$\max_{\mathbf{a}_1, \mathbf{b}_1 \neq 0} \frac{\widehat{\mathrm{Cov}}[\mathbf{a}_1^T\mathbf{X}, \mathbf{b}_1^T\mathbf{Y}]}{\widehat{\mathrm{Var}}[\mathbf{a}_1^T\mathbf{X}]^{1/2}\widehat{\mathrm{Var}}[\mathbf{b}_1^T\mathbf{Y}]^{1/2}} \Longleftrightarrow \max_{\mathbf{a}_1, \mathbf{b}_1 \neq 0} \frac{\mathbf{a}_1^T\hat{\Gamma}_{\mathbf{xy}}\mathbf{b}_1}{(\mathbf{a}_1^T\hat{\Gamma}_{\mathbf{xx}}\mathbf{a}_1)^{1/2}(\mathbf{b}_1^T\hat{\Gamma}_{\mathbf{yy}}\mathbf{b}_1)^{1/2}}$$

2 $$\max_{\mathbf{a}_2 \perp \mathbf{a}_1, \mathbf{b}_2 \perp \mathbf{b}_1} \frac{\widehat{\mathrm{Cov}}[\mathbf{a}_2^T\mathbf{X}, \mathbf{b}_2^T\mathbf{Y}]}{\widehat{\mathrm{Var}}[\mathbf{a}_2^T\mathbf{X}]^{1/2}\widehat{\mathrm{Var}}[\mathbf{b}_2^T\mathbf{Y}]^{1/2}} \qquad \max_{\mathbf{a}_1 \perp \mathbf{a}_2, \mathbf{b}_1 \perp \mathbf{b}_2} \frac{\mathbf{a}_2^T\hat{\Gamma}_{\mathbf{xy}}\mathbf{b}_2}{(\mathbf{a}_2^T\hat{\Gamma}_{\mathbf{xx}}\mathbf{a}_2)^{1/2}(\mathbf{b}_2^T\hat{\Gamma}_{\mathbf{yy}}\mathbf{b}_2)^{1/2}}$$

- Let $\mathbf{u} = \hat{\Gamma}_{\mathbf{xx}}^{1/2}\mathbf{a}$, $\mathbf{v} = \hat{\Gamma}_{\mathbf{yy}}^{1/2}\mathbf{b}$. It suffices to solve $\max\limits_{\|\mathbf{u}\|_2^2 = \|\mathbf{v}\|_2^2 = 1} \mathbf{u}^T(\hat{\Gamma}_{\mathbf{xx}}^{-1/2}\hat{\Gamma}_{\mathbf{xy}}\hat{\Gamma}_{\mathbf{yy}}^{-1/2})\mathbf{v}$

Canonical Variate Analysis

- Originates from predictive modeling using state-space structure

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_k + \epsilon_k$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \eta_k$$

- It has been proved that canonical correlation analysis can find linear combination of $\mathbf{y}_{k-1}$ that best predicts $\mathbf{y}_k$.

Process Monitoring with Canonical Variate Analysis

Canonical Variate Analysis

- Let $\mathbf{Y} = (Y_1 \quad Y_2 \quad \ldots \quad Y_q)^T$ be a multivariate process to be monitored

$$\mathbf{y}_{p,k} = \begin{bmatrix}\mathbf{y}_{k-1}^T & \mathbf{y}_{k-2}^T & \cdots & \mathbf{y}_{k-q}^T\end{bmatrix}^T$$
$$\mathbf{y}_{f,k} = \begin{bmatrix}\mathbf{y}_{k}^T & \mathbf{y}_{k+1}^T & \cdots & \mathbf{y}_{k+q-1}^T\end{bmatrix}^T$$

$$\mathbf{Y}_p = \begin{bmatrix}\mathbf{y}_{p,q+1} & \mathbf{y}_{p,q+2} & \cdots & \mathbf{y}_{p,N-q+1}\end{bmatrix}$$
$$\mathbf{Y}_f = \begin{bmatrix}\mathbf{y}_{f,q+1} & \mathbf{y}_{f,q+2} & \cdots & \mathbf{y}_{f,N-q+1}\end{bmatrix}$$

- We want to solve $\max_{\|\mathbf{u}\|_2^2=\|\mathbf{v}\|_2^2=1} \mathbf{u}^T(\hat{\Gamma}_{ff}^{-1/2}\hat{\Gamma}_{fp}\hat{\Gamma}_{pp}^{-1/2})\mathbf{v}$

- $\hat{\Gamma}_{ff}^{-1/2}\hat{\Gamma}_{fp}\hat{\Gamma}_{pp}^{-1/2} = \mathbf{UDV}^T$

Functional Canonical Correlation

Monitoring Statistics

- The state and residual space canonical variates are calculated by

$$\hat{\mathbf{x}}_k = \mathbf{V}_n^T \Gamma_{pp}^{-1/2} \mathbf{y}_{p,k}, \ \hat{\mathbf{e}}_k = (\mathbf{I} - \mathbf{V}_n \mathbf{V}_n^T) \Gamma_{pp}^{-1/2} \mathbf{y}_{p,k}$$

  where n denotes the order of system. Often determined by dominant singular values or AIC.

- $T_k^2 = \hat{\mathbf{x}}_k^T \hat{\mathbf{x}}_k, \ Q_k = \hat{\mathbf{e}}_k^T \hat{\mathbf{e}}_k$ measures the deviation in state and residual space.

- The upper control limit $T_{1-\alpha}^2, \ Q_{1-\alpha}$ are calculate using KDE.

Process Monitoring with Canonical Variate Analysis

Training under normal operation

fix $q, \ \alpha, \ \text{(and n)}$

Construct $\mathbf{Y}_p, \ \mathbf{Y}_f$

Calculate $\hat{\Gamma}_{pp}, \ \hat{\Gamma}_{fp}, \ \hat{\Gamma}_{ff}$

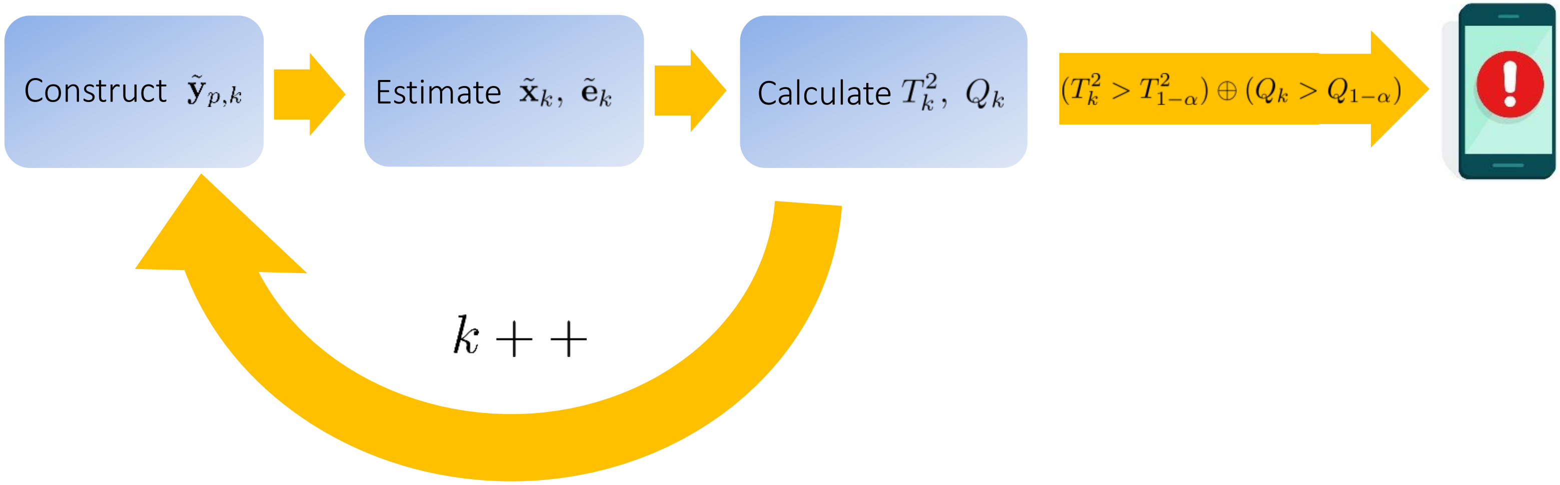
$\hat{\Gamma}_{ff}^{-1/2}\hat{\Gamma}_{fp}\hat{\Gamma}_{pp}^{-1/2} = \mathbf{UDV}^T$

Calculate $T_k^2, \ Q_k$

$T_{1-\alpha}^2, \ Q_{1-\alpha}$ using KDE

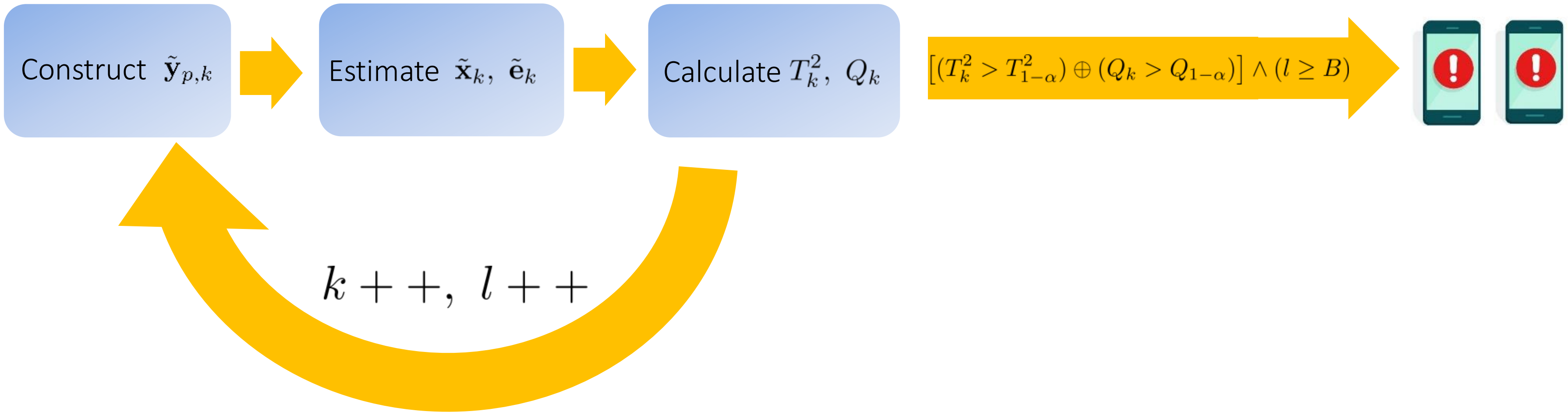
Process Monitoring with Canonical Variate Analysis

Monitoring using real-time data



Process Monitoring with Canonical Variate Analysis

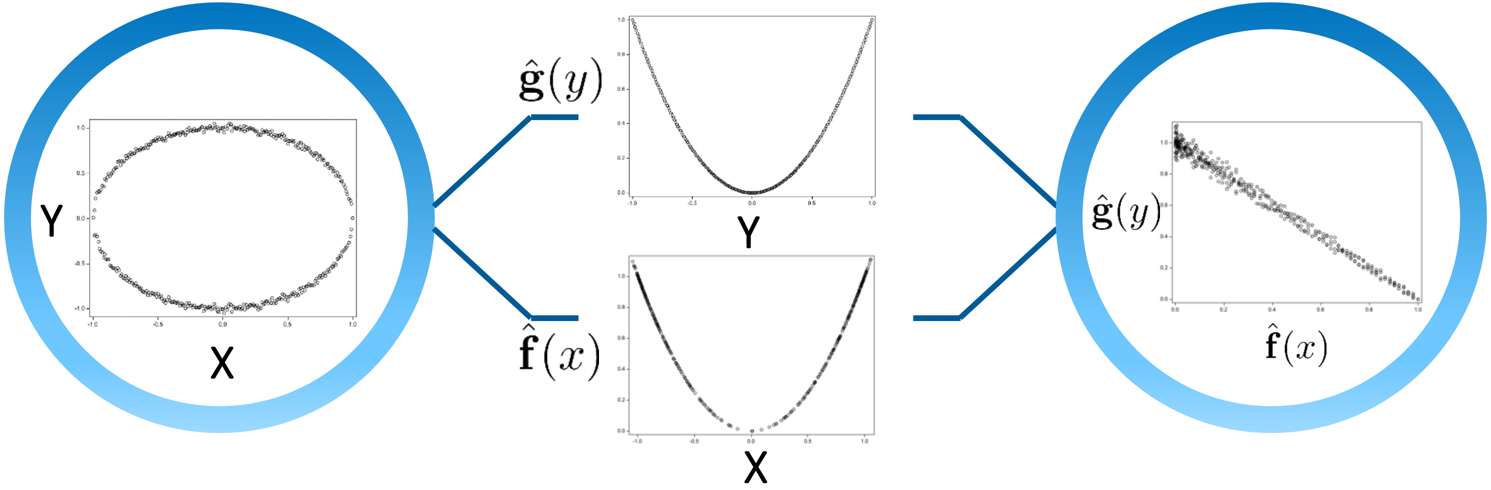
Monitoring using real-time data



Process Monitoring with Canonical Variate Analysis

Kernel Method

- Strongly correlated under nonlinear mapping but not under linear mapping.



Process Monitoring with Canonical Variate Analysis

Kernel canonical correlation

- Describe to what extend two random vectors are nonlinearly correlated.

$$\max_{f\in\mathcal{H}_\mathcal{X},g\in\mathcal{H}_\mathcal{Y}} \frac{\widehat{\mathrm{Cov}}[f(\mathbf{X}),g(\mathbf{Y})]}{(\widehat{\mathrm{Var}}[f(\mathbf{X})]+\epsilon_n||f||^2_{\mathcal{H}_\mathcal{X}})^{1/2}(\widehat{\mathrm{Var}}[g(\mathbf{Y})]+\epsilon_n||g||^2_{\mathcal{H}_\mathcal{Y}})^{1/2}}$$

$$\max_{\mathbf{a}_1,\mathbf{b}_1\neq 0} \frac{\mathbf{a}_1^T\bar{\mathbf{K}}_x\bar{\mathbf{K}}_y^T\mathbf{b}_1}{[\mathbf{a}_1^T(\bar{\mathbf{K}}_x\bar{\mathbf{K}}_x^T+\delta\mathbf{I})\mathbf{a}_1]^{1/2}[\mathbf{b}_1^T(\bar{\mathbf{K}}_y\bar{\mathbf{K}}_y^T+\delta\mathbf{I})\mathbf{b}_1]^{1/2}}$$

- Where $\bar{\mathbf{K}}$ is the centered Gram matrix.
- Gram matrix $[\mathbf{K}]_{ij} := \mathbf{K}(\mathbf{x}_i,\mathbf{x}_j)$ where $\mathbf{K}(\cdot,\cdot)$ is a kernel function.
- Usually $\mathbf{K}(\mathbf{x}_i,\mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i-\mathbf{x}_j\|^2}{2c^2}\right)$, where **c** is the kernel bandwidth

Process Monitoring with Canonical Variate Analysis

Training under normal operation

fix q, α, c (and n)

Construct $\mathbf{Y}_p,\ \mathbf{Y}_f$

Calculate $\bar{\mathbf{K}}_p,\ \bar{\mathbf{K}}_f$

$\hat{\Gamma}_{ff} \propto \bar{\mathbf{K}}_p \bar{\mathbf{K}}_p^T, \ldots$

$(\hat{\Gamma}_{ff} + \delta \mathbf{I})^{-1/2} \hat{\Gamma}_{fp} (\hat{\Gamma}_{pp} + \delta \mathbf{I})^{-1/2} = \mathbf{U}\mathbf{D}\mathbf{V}^T$

Calculate $T_k^2,\ Q_k$

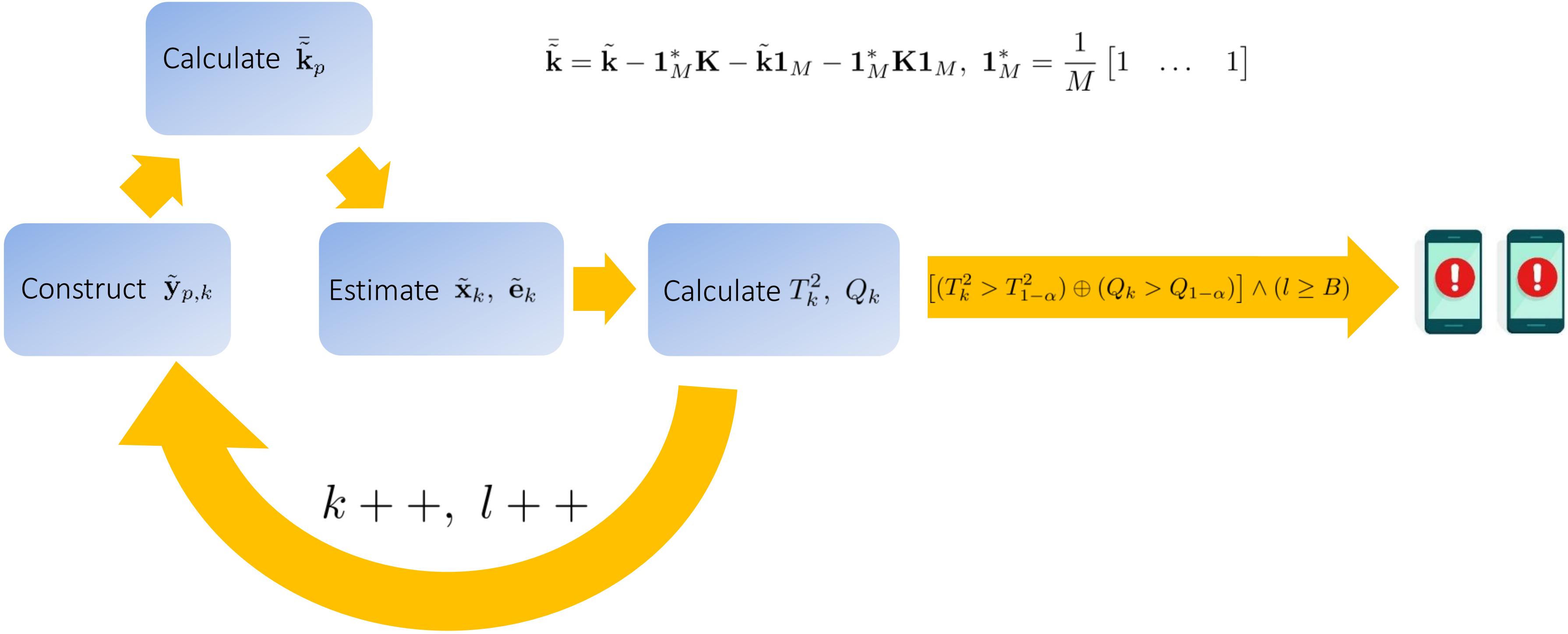
$T_{1-\alpha}^2,\ Q_{1-\alpha}$ using KDE

$$\bar{\mathbf{K}} = \mathbf{K} - \mathbf{1}_M \mathbf{K} - \mathbf{K}\mathbf{1}_M - \mathbf{1}_M \mathbf{K} \mathbf{1}_M,\ \mathbf{1}_M = \frac{1}{M} \begin{bmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{bmatrix}$$

Process Monitoring with Canonical Variate Analysis

Monitoring using real-time data

$$\bar{\bar{\tilde{\mathbf{k}}}} = \tilde{\mathbf{k}} - \mathbf{1}_M^*\mathbf{K} - \tilde{\mathbf{k}}\mathbf{1}_M - \mathbf{1}_M^*\mathbf{K}\mathbf{1}_M, \ \mathbf{1}_M^* = \frac{1}{M}\begin{bmatrix}1 & \dots & 1\end{bmatrix}$$



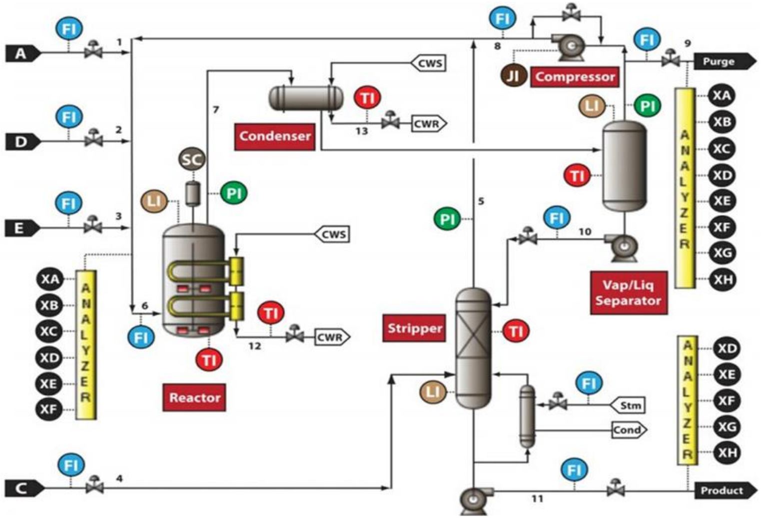
Process Monitoring with Canonical Variate Analysis

Kernel canonical correlation

- Pros:
  - Accounts for correlation under non-linear mapping
  - Flexibility
- Cons:
  - Introduces more parameters
  - Requires regularization
  - Computationally expensive

Process Monitoring with Canonical Variate Analysis

Tennessee Eastman Process



- Benchmark dataset for chemical process control.
- Contains 53 variables and 20 seeded faults.
- 25-hour training and 48-hour testing data.
- 500 simulation runs.

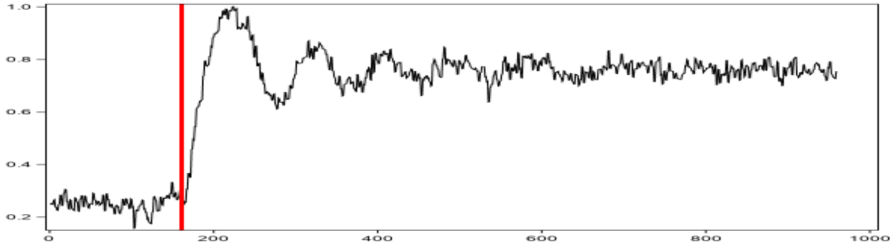
Process Monitoring with Canonical Variate Analysis

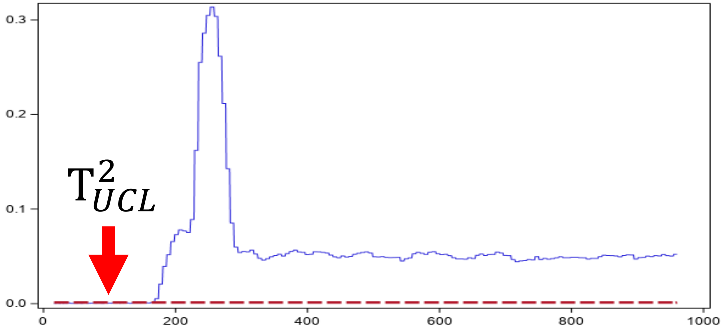
Tennessee Eastman Process

Fault	1	2	4	5	6	7	8	10	11
Accuracy	99.5%	98.8%	99.9%	99.8%	99.9%	99.9%	98.2%	97.5%	98.7%

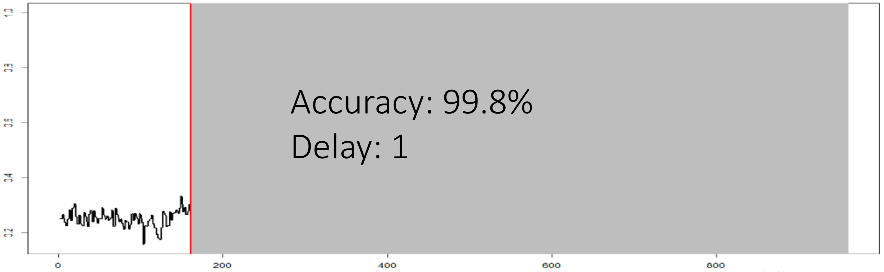
Normal

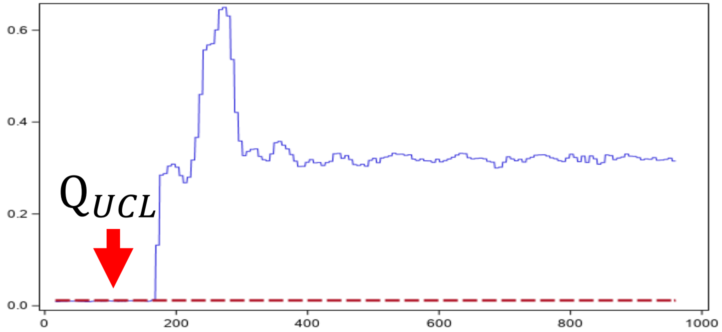
Abnormal





Detect as anomalies

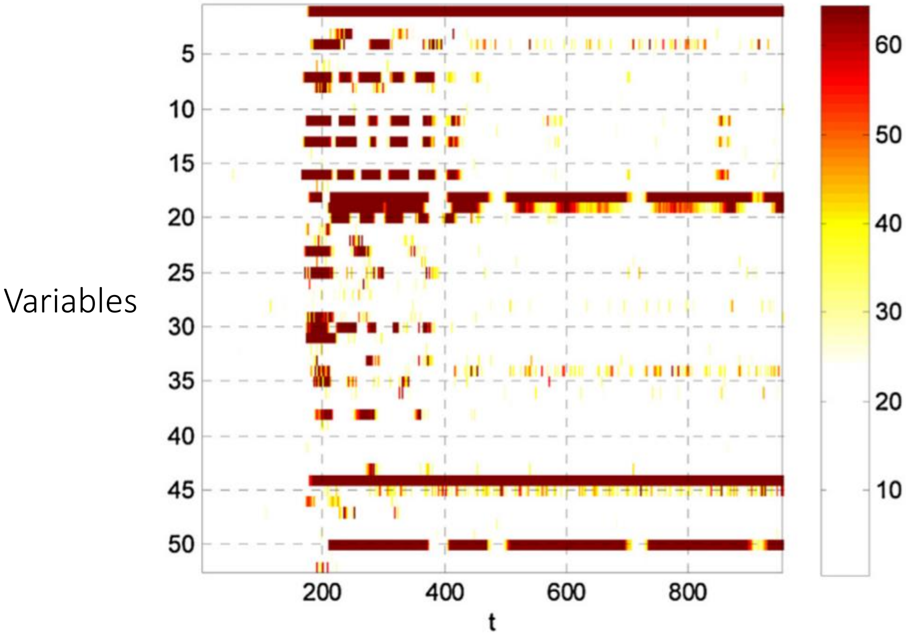




Process Monitoring with Canonical Variate Analysis

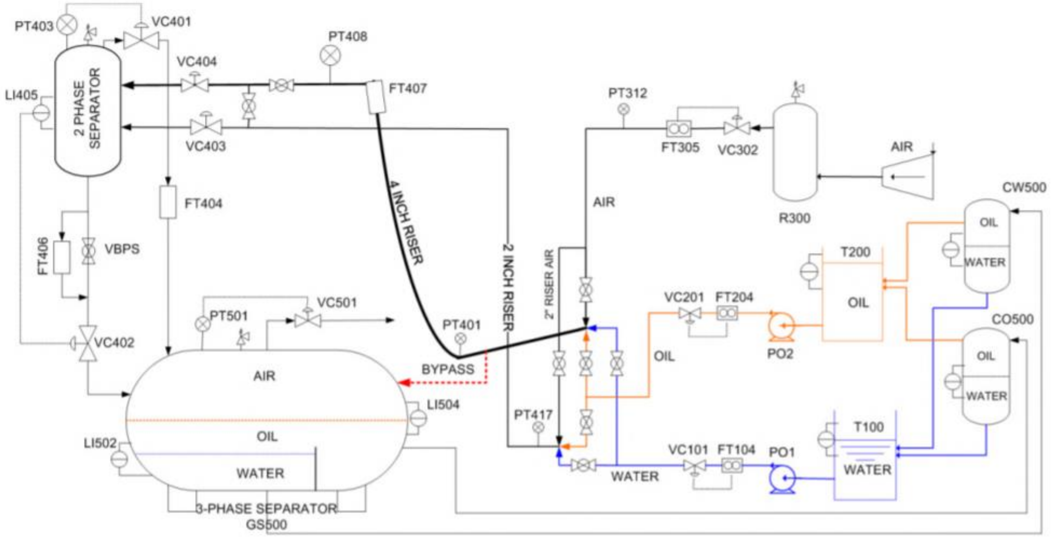
Tennessee Eastman Process

Contribution of each variable to Fault 1



Process Monitoring with Canonical Variate Analysis

Cranfield Three-phase Flow Facility

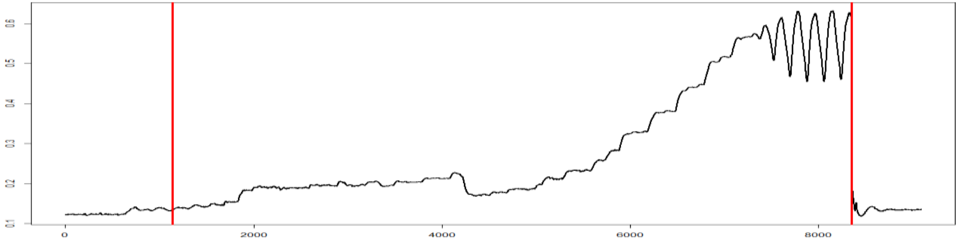


- Various operating conditions
- Contains 24 variables and 6 seeded faults.

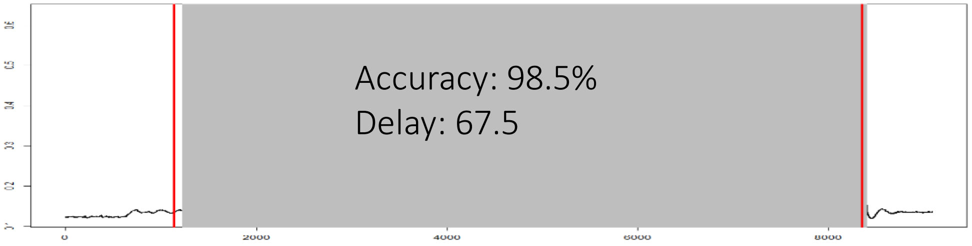
Process Monitoring with Canonical Variate Analysis

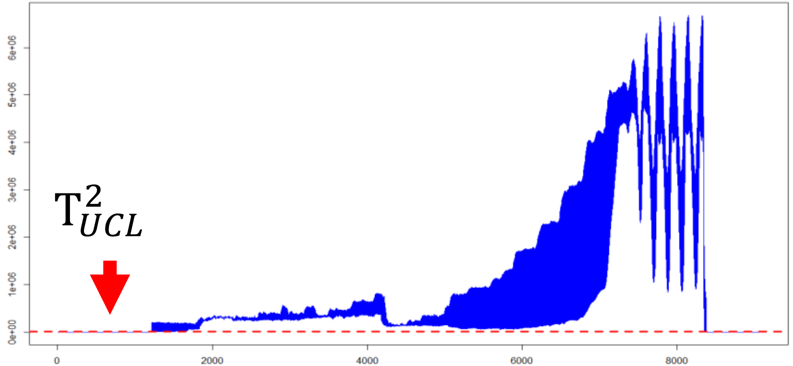
Cranfield Multiphase Facility

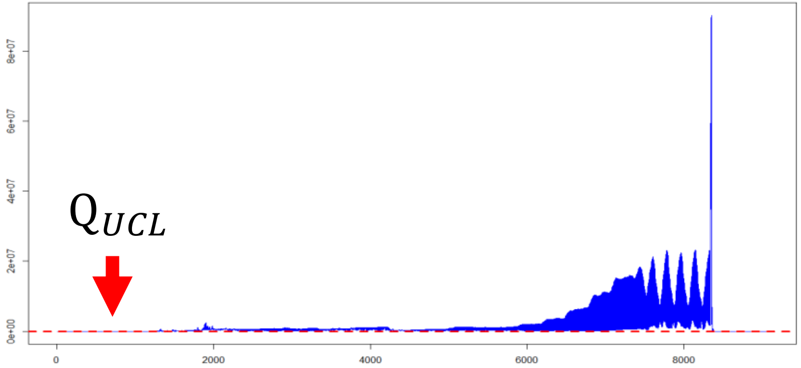
Normal Abnormal Normal



Detect as anomalies







Process Monitoring with Canonical Variate Analysis

Other variant (Extension)

- Sparse Canonical Variate Analysis

$$\max_{\|\mathbf{u}\|_2^2 \leq 1, \|\mathbf{v}\|_2^2 \leq 1, \|\mathbf{u}\|_1 \leq c, \|\mathbf{v}\|_1 \leq c} \mathbf{u}^T (\hat{\Gamma}_{ff}^{-1/2} \hat{\Gamma}_{fp} \hat{\Gamma}_{pp}^{-1/2}) \mathbf{v}$$

  - Can be solved using penalized matrix decomposition
  - Results in sparse contribution: easier to locate variables that cause the fault

Thank you !