Statistical Process Monitoring with Independent Component Analysis

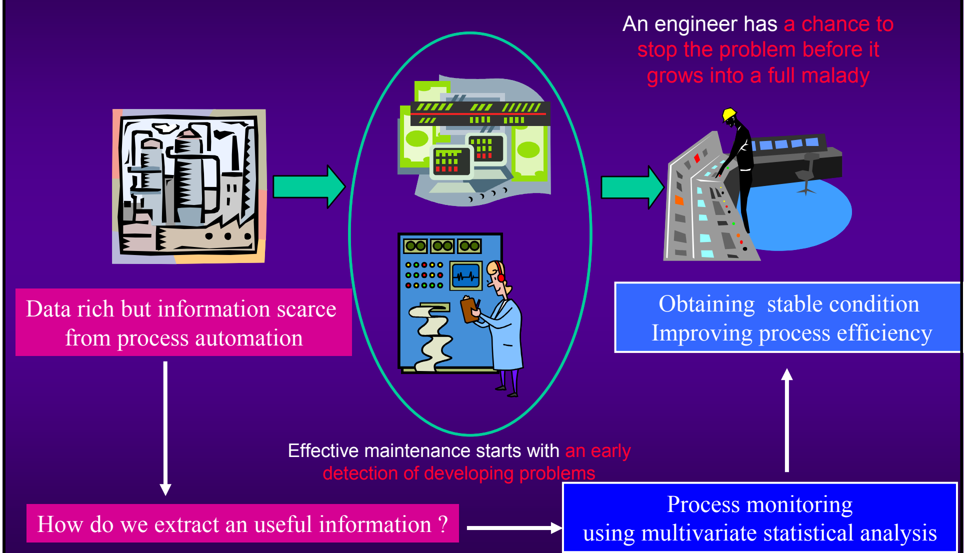
ChangKyoo Yoo
BIOMATH, University Gent, Belgium

Jong-Min Lee, and In-Beum Lee
Department of Chemical Engineering
Pohang University of Science and Technology

Presentation Outline

- Conventional Monitoring Method
- Motivation
- Theory and algorithm of ICA
- Statistical process monitoring procedure with ICA
- Case Study: wastewater treatment plant data
- Conclusion

Concept of Process Monitoring



Conventional Monitoring Methods

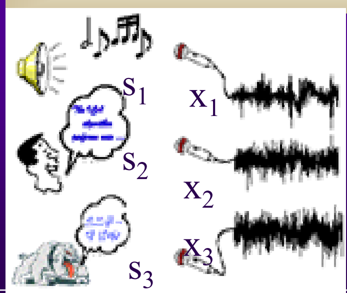
- Multivariate Statistical Methods using PCA/PLS
  - Can handle high dimensional, noisy and correlated data by projecting the original data onto a lower dimensional subspace
  - Use of in-control (normal) runs in the historical database
  - Development of the statistical model that characterizes normal operation (NOC)
  - Computation of control chart limits for use in monitoring future samples
    - ✓ Hotelling T^2 chart, Squared prediction error(SPE) chart
  - Not appropriate for non-stationary, dynamic, or non-Gaussian data

→ **ICA**

The Concept of Independent Component Analysis

• **Independent component analysis (ICA)** is a statistical method, the goal of which is to decompose the multivariate data **x** into a linear sum of statistically independent components, i.e.

$$\mathbf{x} = \mathbf{A}\,\mathbf{s}$$



$x_1(t) = a_{11}s_1(t)+a_{12}s_2(t)+a_{13}s_3(t)$

$x_2(t) = a_{21}s_1(t)+a_{22}s_2(t)+a_{23}s_3(t)$

$x_3(t) = a_{31}s_1(t)+a_{32}s_2(t)+a_{33}s_3(t)$

We should find a_{ij} , $s_1(t)$, $s_2(t)$ and $s_3(t)$ from only $x_1(t)$, $x_2(t)$ and $x_3(t)$

How can we find source signals from only x?

$$\mathbf{x} = \mathbf{A}\,\mathbf{s}$$

s_i : statistically independent and var(s_i)=1

A: unknown mixing matrix (square matrix)

• If $\mathbf{W}=\mathbf{A}^{-1}$, we can exactly recover source signal from $\mathbf{s} = \mathbf{W}\,\mathbf{x}$.

• How can we find **W** from only **x?**

➡ **W** is initialized and updated to maximize the non-Gaussianity of **s**

• More non-Gaussian, more independent!

❑ **Measure of Nongaussianity**

- Kurtosis: $kurt(y) = E(y^4) - 3\,(E(y^2))^2$

  kurtosis is zero for a Gaussian random variable

  very sensitive to outliers

- Negentropy:

  Hyvärinen(1998) develpoed a robust approximation equation of negentropy

Brief procedures of ICA

1. **Centering** (mixed and independent source is zero-mean)
2. **Whitening**

   Transform the observed vector **x** linearly so that $E\{\mathbf{z}(k)\mathbf{z}^T(k)\} = \mathbf{I}$

   $$\mathbf{z}(k) = \mathbf{Q}\mathbf{x}(k) = \mathbf{Q}\mathbf{A}\mathbf{s}(k) = \mathbf{B}\mathbf{s}(k)$$
   $$E\{\mathbf{z}(k)\mathbf{z}^T(k)\} = \mathbf{B}E\{\mathbf{s}(k)\mathbf{s}^T(k)\}\mathbf{B}^T = \mathbf{B}\mathbf{B}^T = \mathbf{I}$$
   $$\mathbf{s}(k) = \mathbf{B}^T\mathbf{z}(k)$$

3. $\mathbf{B}^T$ is initialized and updated to maximize the negentropy of $\mathbf{s}(k)$
4. Since $\mathbf{s}(k) = \mathbf{W}\mathbf{x}(k)$ and $\mathbf{s}(k) = \mathbf{B}^T\mathbf{z}(k) = \mathbf{B}^T\mathbf{Q}\mathbf{x}(k)$, **W** can be obtained by $\mathbf{W} = \mathbf{B}^T\mathbf{Q}$

Comparison between PCA and ICA

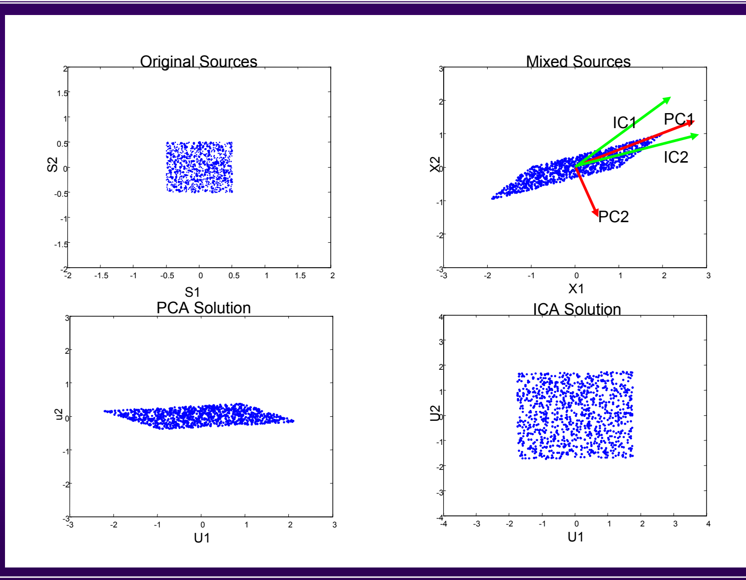
PCA

1. Second-order method (mean, variance)
2. Using only the information contained in the covariance matrix of the data vector x
3. Assume Gaussianity of x
4. Computationally simple

ICA

1. Higher-order method (mean,variance, skewness,kurtosis, etc)
2. Use information on the distribution of x that is not contained in the covariance matrix
3. Assume non-Gaussianity of x
4. More sophisticated techniques

Comparison ICA solution with PCA solution



New monitoring procedures (Modeling Part I)

1. Mean centering and variance saling
2. Whitening procedure: $\mathbf{Z}_{normal} = \mathbf{QX}_{normal}$
3. ICA procedure

   Obtain **W**, **B**, and $\mathbf{S}_{\text{normal}}$ from $\mathbf{S}_{normal} = \mathbf{WX}_{normal} = \mathbf{B}^T\mathbf{Z}_{normal}$

3. Calculate the norm of the row vectors of **W** and separate **W** into the deterministic part and the excluded part based on the magnitude of norms. **B** and $\mathbf{S}_{\text{normal}}$ can be separated with the same criterion.

$\mathbf{W} \rightarrow \mathbf{W}_d, \mathbf{W}_e$ $\quad$ $\mathbf{B} \rightarrow \mathbf{B}_d, \mathbf{B}_e$ $\quad$ $\mathbf{S}_{normal}$ → $\mathbf{S}_d = \mathbf{W}_d\mathbf{X}_{normal}$, $\mathbf{S}_e = \mathbf{W}_e\mathbf{X}_{normal}$

New monitoring procedures (Modeling Part II)

4. Calculate I^2, I_e^2 and *SPE* metrics

$$I^2(n) = \mathbf{s}_d(n)^T \mathbf{s}_d(n)$$ $$I_e^2(n) = \mathbf{s}_e(n)^T \mathbf{s}_e(n)$$ $$SPE(n) = \sum_{j=1}^{d} (x_j(n) - \hat{x}_j(n))^2$$

where $\hat{\mathbf{X}} = \mathbf{Q}^{-1}\mathbf{B}_d\mathbf{S}_d = \mathbf{Q}^{-1}\mathbf{B}_d\mathbf{W}_d\mathbf{X}_{normal}$

5. Obtain control limits of I^2, I_e^2 and *SPE* metrics at each time using kernel density estimation

I^2, I_e^2 and SPE are not normally distributed

Kernel density estimation

Density estimation is the construction of an estimate of the density function from the observed data.

Contribution for fault identification

- Variable contribution to $I_{new}^2(k)$

$$\mathbf{x}_{cd}(k) = \frac{\mathbf{Q}^{-1}\mathbf{B}_d\mathbf{s}_{newd}(k)}{\left\|\mathbf{Q}^{-1}\mathbf{B}_d\mathbf{s}_{newd}(k)\right\|}\left\|\mathbf{s}_{newd}(k)\right\|$$

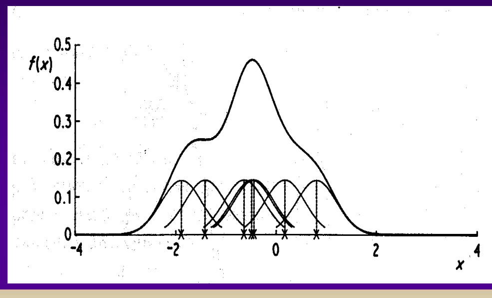
- Variable contribution to $I_{newe}^2(k)$

$$\mathbf{x}_{ce}(k) = \frac{\mathbf{Q}^{-1}\mathbf{B}_e\mathbf{s}_{newe}(k)}{\left\|\mathbf{Q}^{-1}\mathbf{B}_e\mathbf{s}_{newe}(k)\right\|}\left\|\mathbf{s}_{newe}(k)\right\|$$

- Variable contribution to $SPE(k)$

$$\mathbf{x}_{cspe}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k)$$

Kernel density estimation (KDE)



- **Kernel estimator**

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

where h is the window width (smoothing parameter or bandwidth)

K: the shape of the bumps, *h*: width of bumps

$\hat{f}$ will inherit all the continuity and differentiability properties of the kernel K

Process monitoring of ICA

Modeling

Acquire time series data under normal operation

↓

Normalization and and whitening

↓

Obtain ICA model (W,S) and *SPE* from NOC

↓

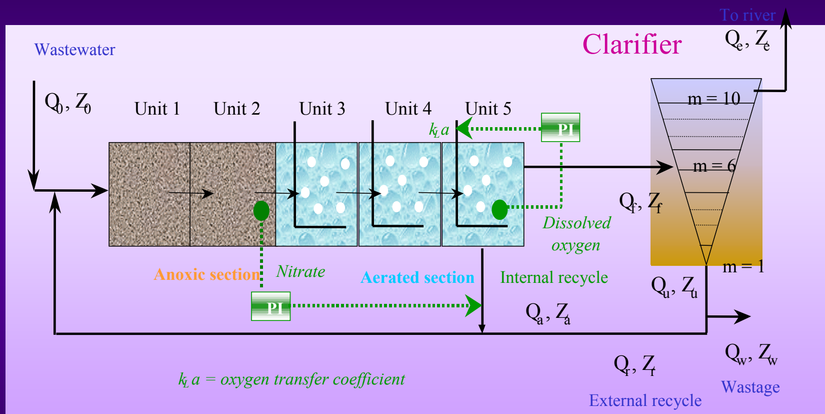
Determine the control chart limits of I^2 and I_e^2 and *SPE* statistics using KDE

↓ ***On-line monitoring***

For new data, calculate I^2 and I_e^2 and *SPE* values

Monitor if I_{new}^2, I_{newe}^2, and *SPE* values exceed the control limits

Case Study: simulation benchmark

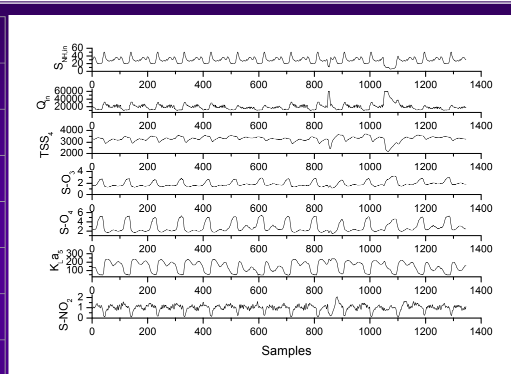


For detail information about benchmark, refer COST working group homepage,

http://www.ensic.u-nancy.fr/COSTWWTP

Variables used in the monitoring and disturbance type

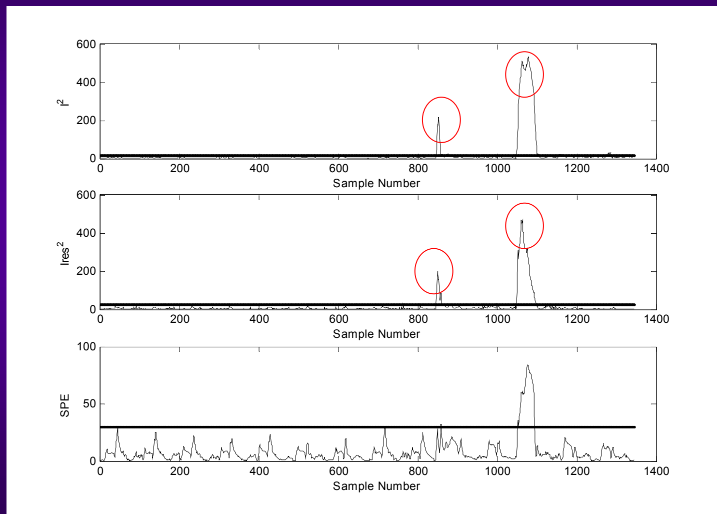
No.	Symbol	Meaning
1	S_{NHin}	**Influent ammoniac concentration**
2	Q_{in}	**Influent flow rate**
3	TSS_4	**Total suspended solid (reactor 4)**
4	$S\text{-}O_3$	**Dissolved oxygen concentration (reactor 3)**
5	$S\text{-}O_4$	**Dissolved oxygen concentration (reactor 4)**
6	$K_L a_5$	**Oxygen transfer coefficient (reactor 5)**
7	$S\text{-}NO_2$	**Nitrate concentration (reactor 2)**



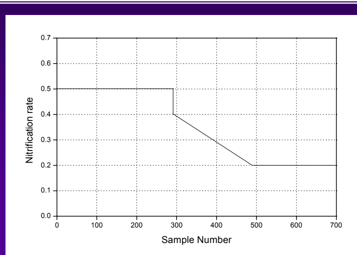
Disturbance type

- **External process disturbance (storm event)**
  - High flow rate, low ammonia conc. in influent load
- **Internal disturbance**
  - Nitrification rate decrease in the aeration basin

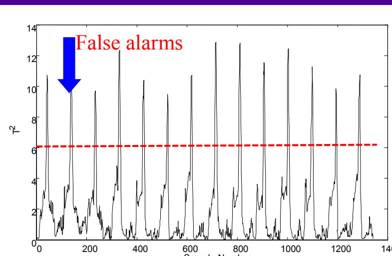
ICA monitoring charts (external disturbance: storm events)

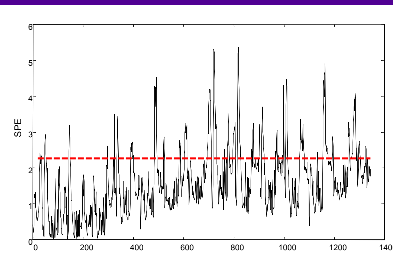


PCA monitoring charts (internal disturbance)

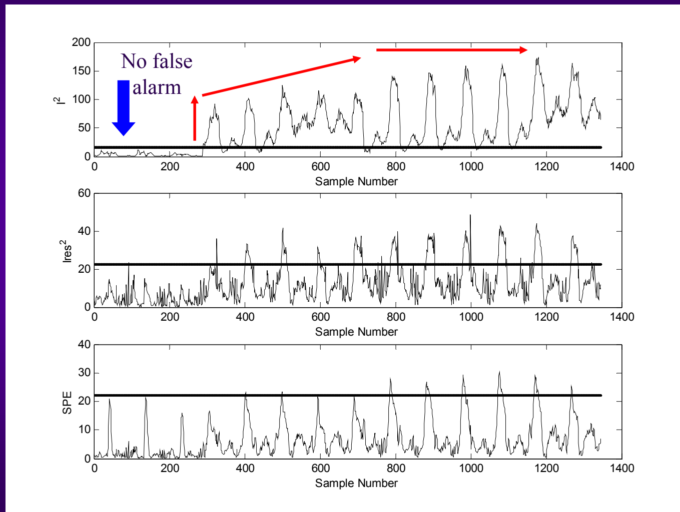


PCA monitoring charts can not detect internal disturbance and just shows many false alarms since the periodic features of WWTP are dominant.

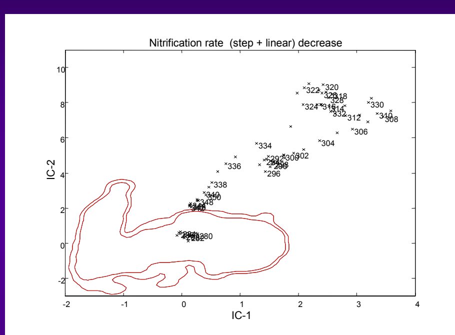


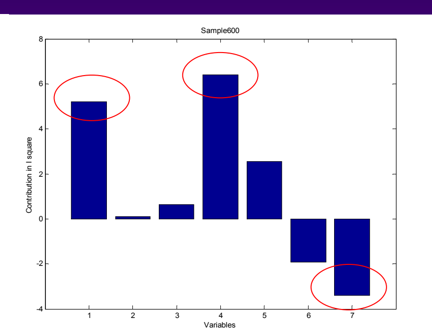


ICA monitoring charts (internal disturbance)



ICs plot and Contribution plot





• ICs plot shows that samples where internal disturbance occur are detected easily.

• Variable 1 ($S\text{-}NH_{in}$), variable 4 ($S\text{-}O_3$), and variable 7($S\text{-}NO_2$) are primarily contributed to the I^2 statistic

Conclusion

1. ICA

   Higher order decorrelation method

   Can find underlying factors from multivariate statistical data

2. ICA monitoring can give better performances rather than PCA monitoring to detect internal disturbance

3. Extension of ICA monitoring