Mutlivariate Statistical Process Performance Monitoring

Jie Zhang

School of Chemical Engineering & Advanced Materials

Newcastle University

Newcastle upon Tyne NE1 7RU, UK

jie.zhang@newcastle.ac.uk

Newcastle University

- Located in the northeast of England
- Established in 1834
- More than 20,000 students and 4,500 staff
- A member of the elite Russell Group – an association of 24 major research-intensive universities of the United Kingdom
- Ranked 18th in in The Times and Sunday Times Good University Guide 2014 (UK's top 12 for Research Power in science and engineering, top 8 for Medical Science)
- The university is structured as three faculties:
  - Faculty of Humanities and Social Sciences
  - Faculty of Medical Sciences
  - Faculty of Science, Agriculture and Engineering
- 11 research institutes

Newcastle University

CPACT

School of Chemical Engineering & Advanced Materials

- Chemical Engineering in Newcastle was established in 1952 and the current school was formed in 2002
- 30 academic staff and about 500 students
- Top research ranking (Grade 5 in 2001 RAE, ranked 6th in Chemical Engineering in 2008 RAE, similar ranking in 2014 REF)
- Ranked 4th in UK by Guardian
- Research themes:
  - Chemical Engineering Science
  - Measurement and Analysis
  - Products and Processes
  - Natural Resources

My research activities

- Batch process modelling, control, and optimisation
- Modelling of nonlinear processes (mainly data based modelling)
- Process monitoring and fault diagnosis
- Process control
- Neural networks and neuro-fuzzy systems
- Other computational intelligence techniques

Outline of presentation

- Introduction
- Multivariate statistical process control
- Multivariate statistical process control of batch processes
- Fault localisation through progressive PCA modelling
- Monitoring processes with multiple operating modes
- Some industrial applications of MSPC
- Conclusions

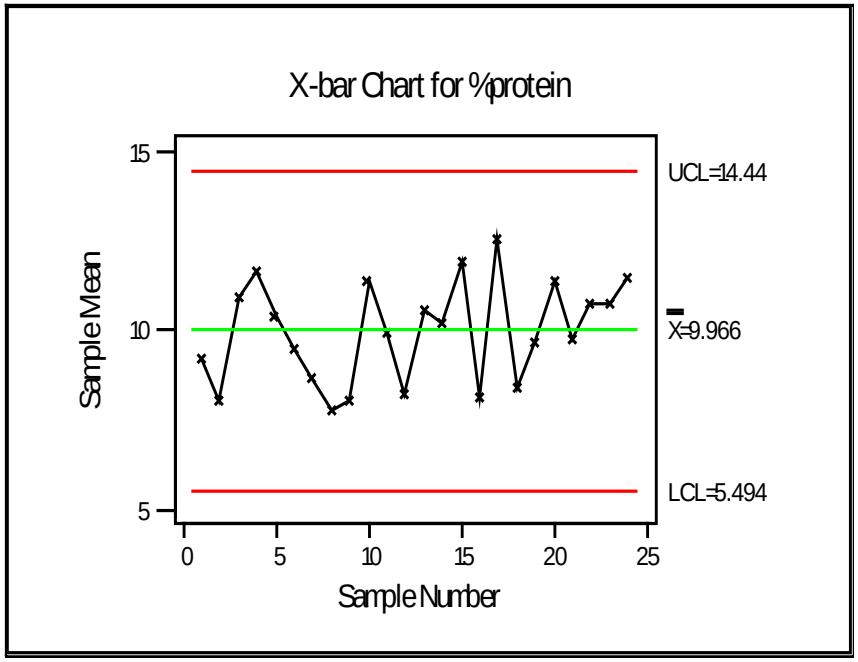
Introduction

- Close monitoring and improved control are increasingly important for the profitable operation of an industrial process.
- A strategically important area of process supervision and control is that of Statistical Process Control (SPC).
- The aim of SPC is to monitor process operation and performance in order to be able to detect the occurrence of:
  - off-specification production
  - important process disturbances
  - process malfunctions
- Process data provides an important basis for the monitoring and control of the quality and consistency of the product and for the statistical control of the process.

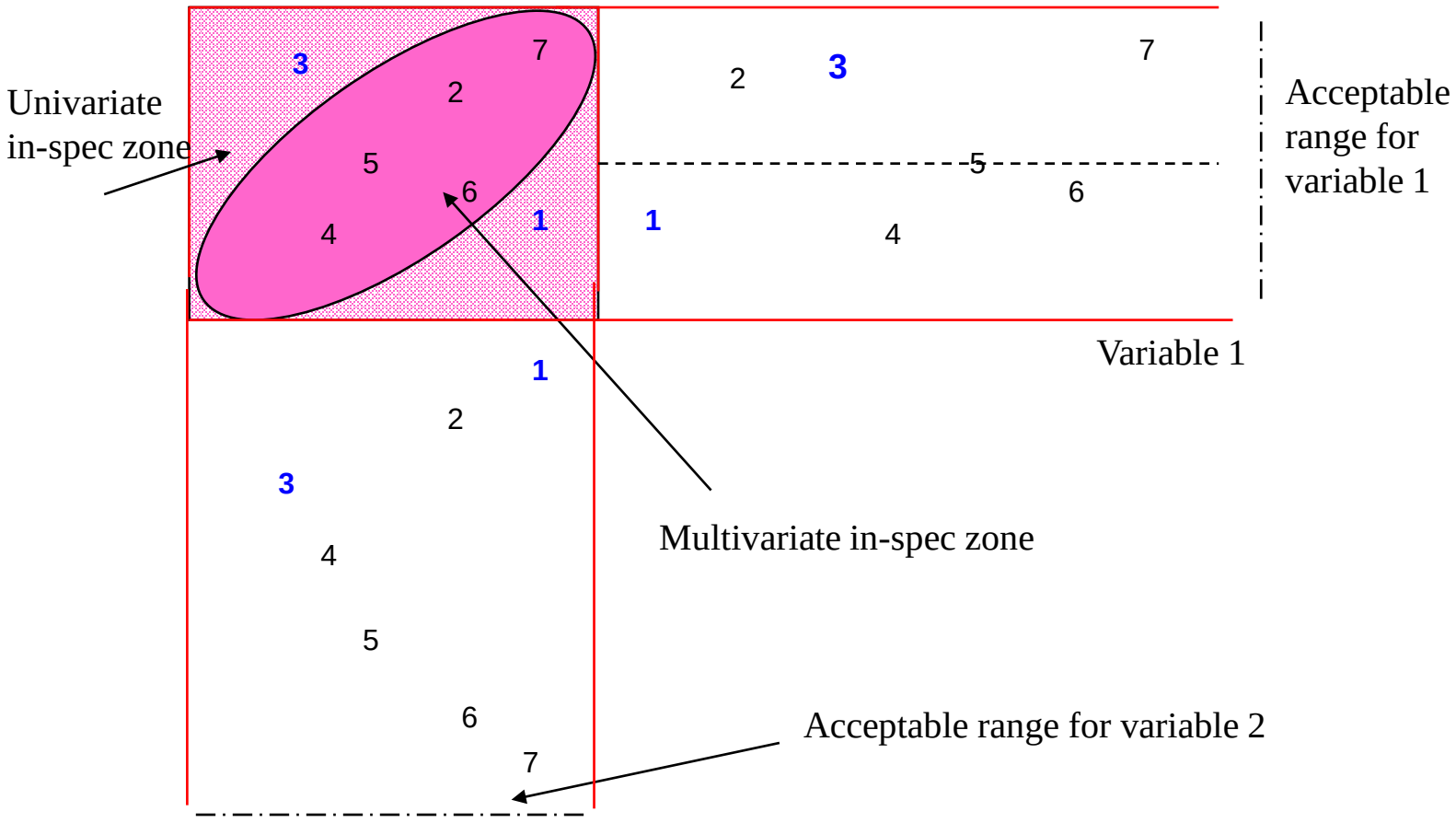
Univariate versus Multivariate SPC

- Univariate SPC looks at the magnitude of the deviations in each variable independent of the other variables.
- Typically, monitored process/quality variables are not independent.
- Only a few underlying events are driving the process at any one time.
- The measurements are simply different reflections of the same underlying events.
- Thus a consequence of examining one variable at-a-time is that the level of control that can be exercised over the quality and consistency of the final product is restricted and plant flexibility impaired.

Traditional univeriate charting approach



Limitation of univeriate SPC



Multivariate SPC

- MSPC or better termed **Process Performance Monitoring**, has the potential to provide operators and plant management with the key process information required to monitor the process effectively, provide early warning of process malfunctions, ensure the manufacture of consistent product and enhance existing plant assets.
- MSPC is based upon the statistical projection techniques of Principal Components Analysis (PCA) and Projection to Latent Structures or Partial Least Squares (PLS).
  - PCA monitors the process through a single block of information - the process and/or quality variables.
  - PLS monitors the process through a model of the quality variables / chemical information developed from the process information.

Principal component analysis (PCA)

- A data matrix X with n samples and m variables (generally assumed to be mean centred and properly scaled) can be decomposed using PCA as

$$X = t_1 p_1^{\mathrm{T}} + t_2 p_2^{\mathrm{T}} + \ldots + t_k p_k^{\mathrm{T}} + \ldots + t_q p_q^{\mathrm{T}}$$

where q=min(m,n), and the $t_i p_i^{\mathrm{T}}$ pairs are ordered by the amount of variance captured.

- Generally, the model is truncated leaving some small amount of variance in a residual matrix:

$$X = t_1 p_1^{\mathrm{T}} + t_2 p_2^{\mathrm{T}} + \ldots + t_k p_k^{\mathrm{T}} + E = T_k T_k^{\mathrm{T}} + E$$

Multivariate statistical process control

- The result of PCA depends on data scales, thus it is a usual practice to scale the nominal data to zero mean and unit variance before applying PCA.
- Once a model has been developed from the nominal data set using a reduced set of principal components, k, $\mathbf{X}=\mathbf{T}\mathbf{P}^{\mathrm{T}}$, the fitted values, $\hat{x}_{new}$, can be calculated for each new multivariate observation:

$$\mathbf{T}_{new} = x_{new}\mathbf{P} \qquad \hat{x}_{new} = \mathbf{T}_{new}\mathbf{P}$$

- The squared prediction error (SPE), also known as the Q statistics, for each new observation is calculated as:

$$Q_{new} = \sum_{i=1}^{k}\left(x_{new} - \hat{x}_{new}\right)^2$$

Multivariate statistical process control

- Approximate control limits for a level of significance α for the SPE are given by (Jackson and Mudholkar, 1979):

$$Q_\alpha = \theta_1 \cdot \left[\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right]^{\frac{1}{h_0}}$$

$$\theta_1 = \sum_{i=r+1}^{a} \lambda_i \qquad \theta_2 = \sum_{i=r+1}^{a} \lambda_i^2 \qquad \theta_3 = \sum_{i=r+1}^{a} \lambda_i^3 \qquad h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$$

where c_α is normal variate with the same sign as h_0.

Multivariate statistical process control

- Another important investigative tool is the T^2 (also known as D-statistic) control chart.
- Assuming multivariate normality of the data, the formula for the T^2 metric for observation i is:

$$T_i^2 = \mathbf{x}_i \mathbf{S}^{-1} \mathbf{x}_i^T$$

where $\mathbf{x}_i$ is a (row) vector representing the process variable measurements for sample i, and $\mathbf{S}=\mathbf{X}^T\mathbf{X}$ is the variance-covariance matrix of the data ($\mathbf{X}$).

- In practice, $\mathbf{S}$ is based on a set of data when the process was in control and $\mathbf{x}_i$ will be some future observation being evaluated.

Multivariate statistical process control

- T^2 can be computed from the PCA scores.
- By selecting only the first k PCA dimensions the T^2 metric for observation i is given by:

$$T_i^2 = \sum_{d=1}^{k} t_{id}^2 / \lambda_d$$

where the PCA score t_{id} for observation i in dimension d has variance λ_d, which is the dth largest eigenvalue of **S**.

- An out of control signal is identified if

$$T^2 > \frac{(n-1)kF_{k,n-k}(\alpha)}{n(n-k)}$$

where α takes 0.05 and 0.01 for the warning and control limits respectively.

Multivariate statistical process control

- The structure of the SPE and T^2 plots reflects two ways in which non-conforming behaviour can be identified:
  - If the process change is caused by a larger than normal shift in one or more of the process variables, but the basic relationship between the quality and/or process variables remains unchanged, then a translation in the scores plane will result (increased T^2 values), with the SPE remaining at an acceptable level.
  - On the other hand, if the abnormality enters through a new event not captured in the reference data set, it will change the nature and possibly the dimension of the relationship between the process and/or quality variables. The SPE will increase.

Fault diagnosis

- Once a fault/abnormal operating condition is detected, its source need to be identified.
- Fault diagnosis can be carried out by the following methods:
  - Contribution plots (Miller et al., 1998)
  - Fault direction characterised by principal components (Zhang et al., 1996)
  - Fault signatures in terms of accumulated principal component scores (Martin et al., 1996; Zhang et al., 1997)
  - Fisher discriminant analysis
  - Reconstruction based approach (Qin, 2003)
  - Neural networks/Fuzzy neural networks
  - Progressive PCA (Hong et al., 2011; 2014)

Contribution Plots

- Fault diagnosis is more easily carried out by reverting back to the original process variables and examining their contribution to the calculated scores and/or the SPE.
- It is then possible to identify the set of original variables whose contribution has changed from that predicted from the nominal model and which may be reflective of the non-conforming behaviour.
- One possible graphical presentation is the variable contribution plot.
- This portrays the change in the new observations relative to their average value calculated from the nominal PCA model.

Fault direction characterised by principal components

- A set of data covering the event of a fault is collected and then scaled using the mean and standard deviations of the normal data.
- The data covering the incidence of a fault are generally polarised such that variances in the data are mainly represented by the first principal component.
- Through PCA, the first loading vector of the data can be calculated and used to represent the direction of a fault in the measurement space.
- The directions of various faults are put in a matrix forming a library of fault directions $F=[D_1\ D_2\ \ldots\ D_n]$

(Zhang, Morris, Martin, *Chem. Eng. Res. Des.*, 1996, 74(1), 89-96)

Fault direction characterised by principal components

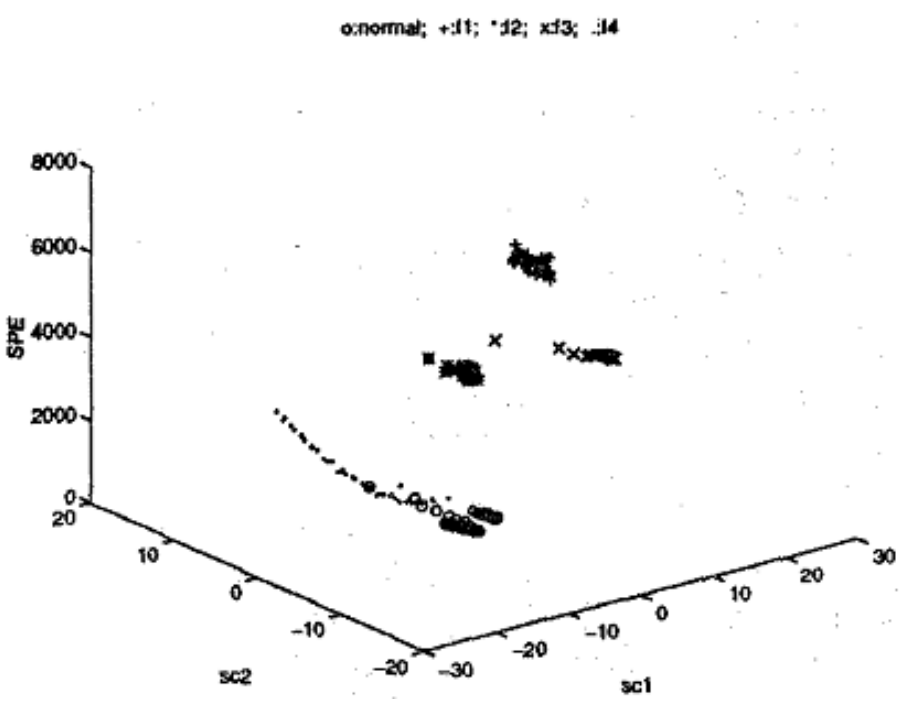
- The currently monitored process measurements can also be analysed through PCA and the first loading vector can be taken as the direction of the current data.
- Denote it by M_D, then the alignment between M_D and the direction of the ith fault, D_i, can be measured by $M_D{}^{\mathrm{T}} D_i$, which is the cosine of the angle between M_D and D_i.
- A diagnostic threshold, τ, is defined such that when

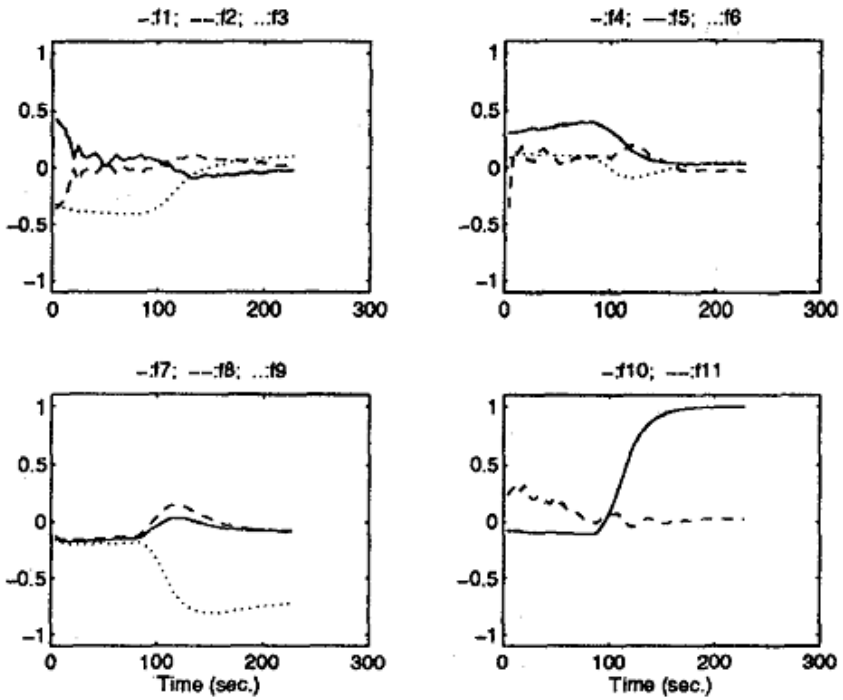
  $$M_D{}^{\mathrm{T}} D_i \geq \tau$$

  it is indicative that the ith fault has occurred.
- Generally, τ is quite close to 1, for example, 0.98.

Fault direction characterised by principal components

- SPE and score plot
- Fault classification





Newcastle University

CPACT

Towards Process / Plant 'Signatures'

- Different faults usually cause process variables to move in different directions.
- The direction of process variable movement can be monitored by studying the projected score movement.
- Accumulated principal component scores:

$$A(n) = \sum_{i=1}^{n} (x(i) - \bar{x})$$

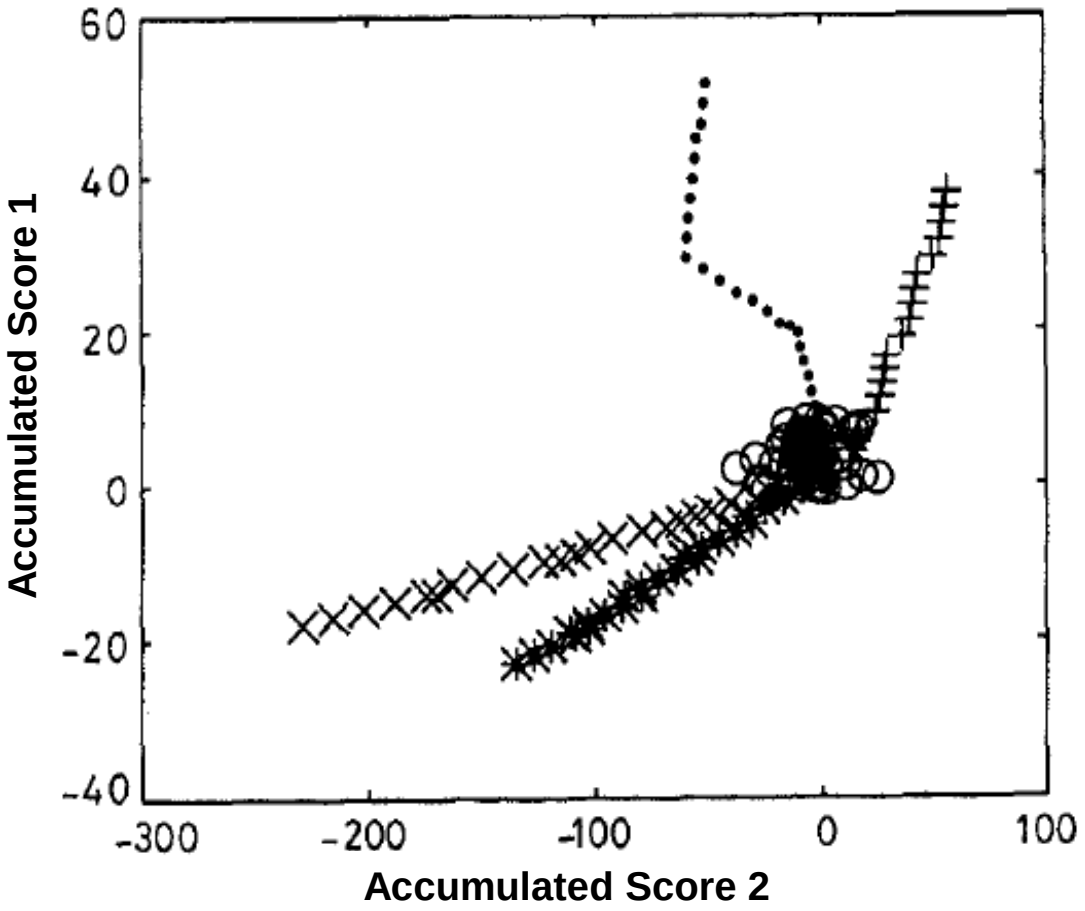
where $x(i)$ is the score for the ith observation, $\bar{x}$ is the mean of the nominal score, and $A(n)$ is the accumulated score till sample n.

- The accumulated scores will remain around the origin when no fault presents.
- Different faults will cause the accumulated score to move in different directions.

Plant Signatures: Cusum Scores Plot (Martin, Morris, Zhang, IEE Proceedings, 1996)

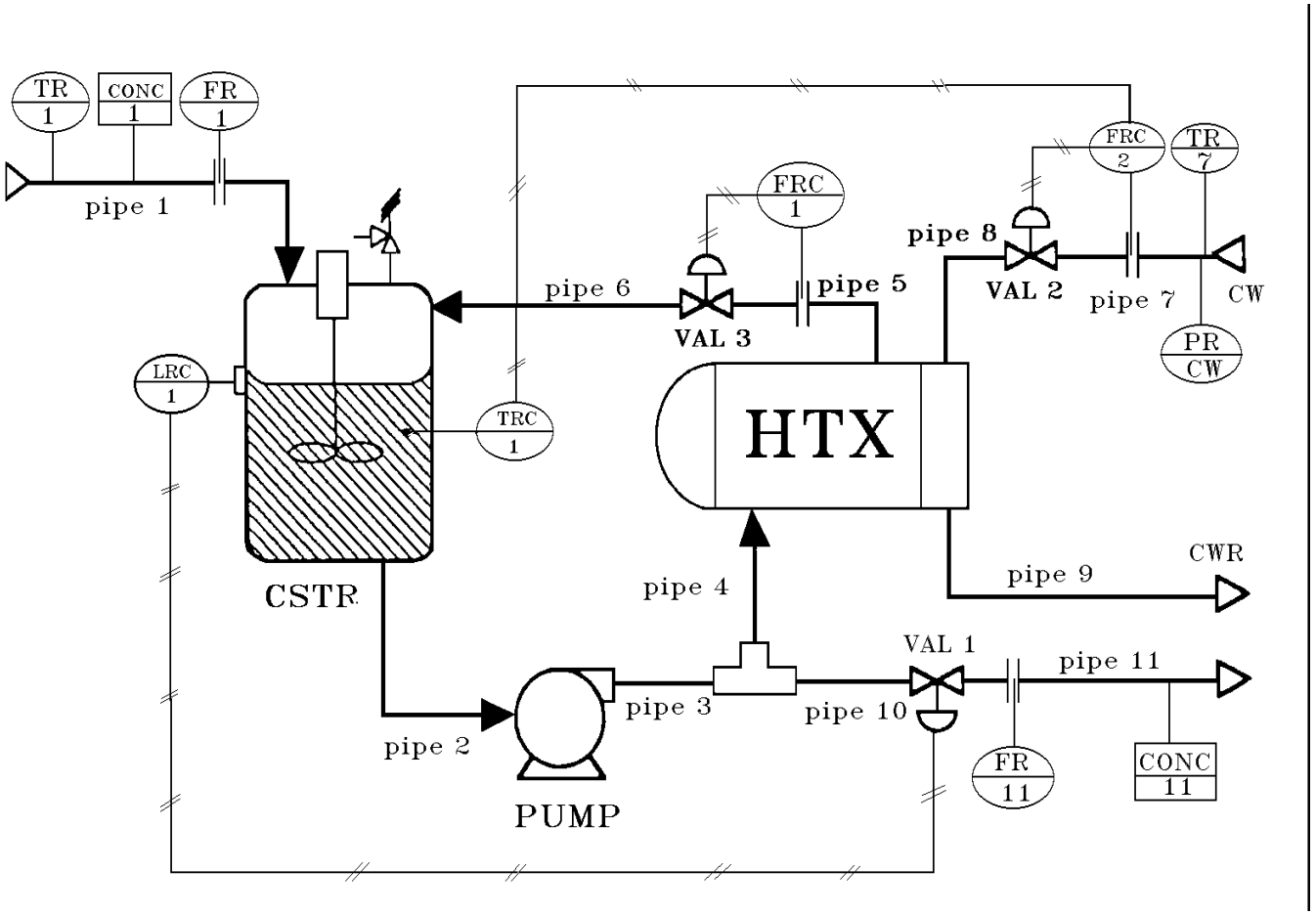
- Fault "directions", "profiles" or "signatures"

  - o Nominal Operations
  - ✶ Solvent Malfunction
  - \+ Reactor Fouling
  - x Impurity Fault
  - • Combined Impurity and Fouling



MSPC Applied to a Complex CSTR Process

- A CSTR with recycle



Application to the CSTR

Fault list:

1). Pipe 1 blockage

2). External feed-reactant flow rate too high

3). Pipe 2 or 3 is blocked or pump fails

4). Pipe 10 or 11 is blocked or control valve 1 fails low

5). External feed-reactant temperature abnormal

6). Control valve 2 fails high

7). Pipe 7, 8, or 9 is blocked or control valve 2 fails low

8). Control valve 1 fails high

9). Pipe 4, 5, or 6 is blocked or control valve 3 fails low

10). Control valve 3 fails high

11). External feed-reactant concentration too low

Measured variables

1). Temperature

2). Tank level

3). T_feed

4). F_inlet

5). F_recycle

6). F_outlet

7). F_cooling

8). Product conc.

9). Feed conc.

10). Pressure

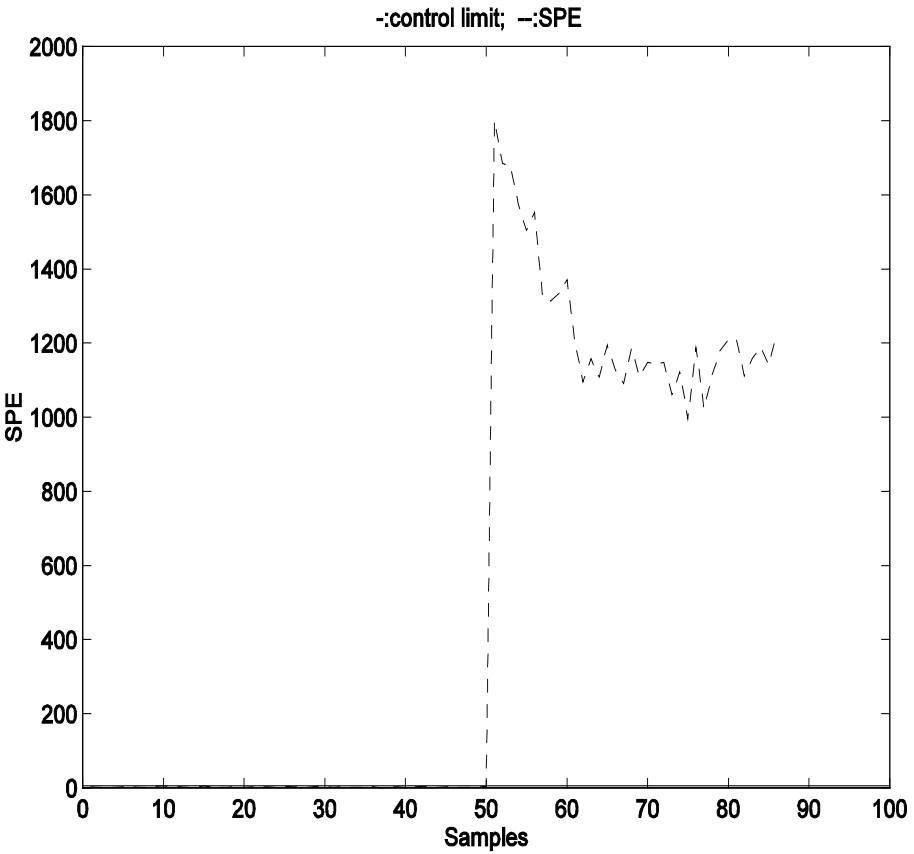
11). Controller 1

12). Controller 2

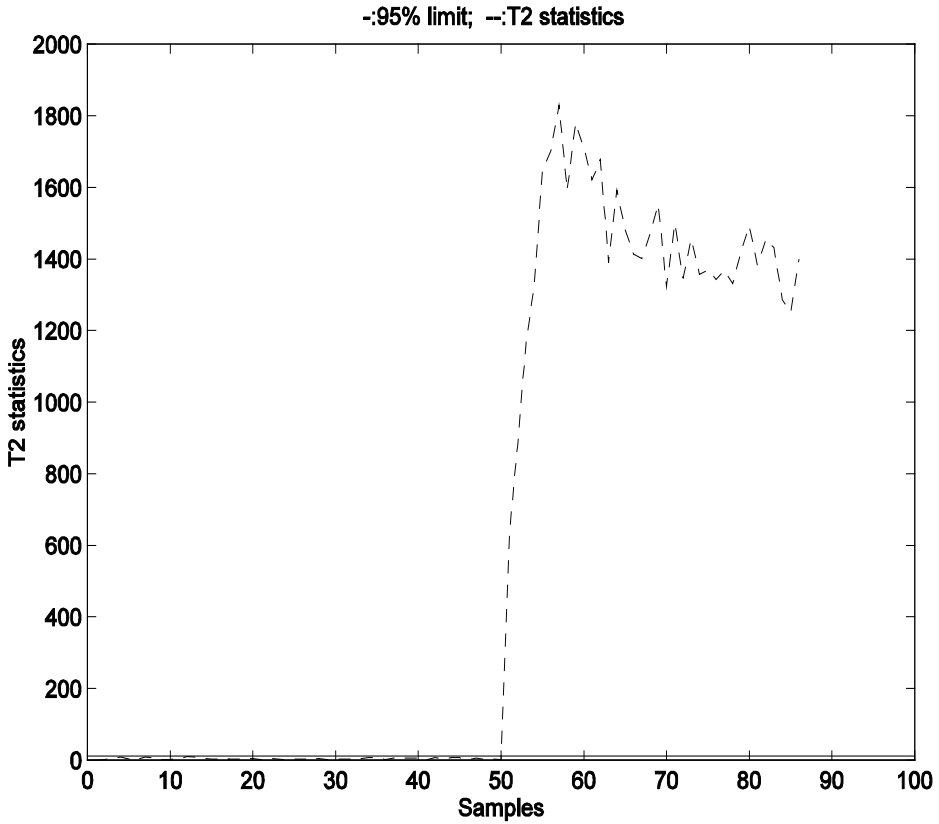
13). Controller 3

MSPC - Fault 1

- SPE plot for Fault 1



- T^2 plot for Fault 1

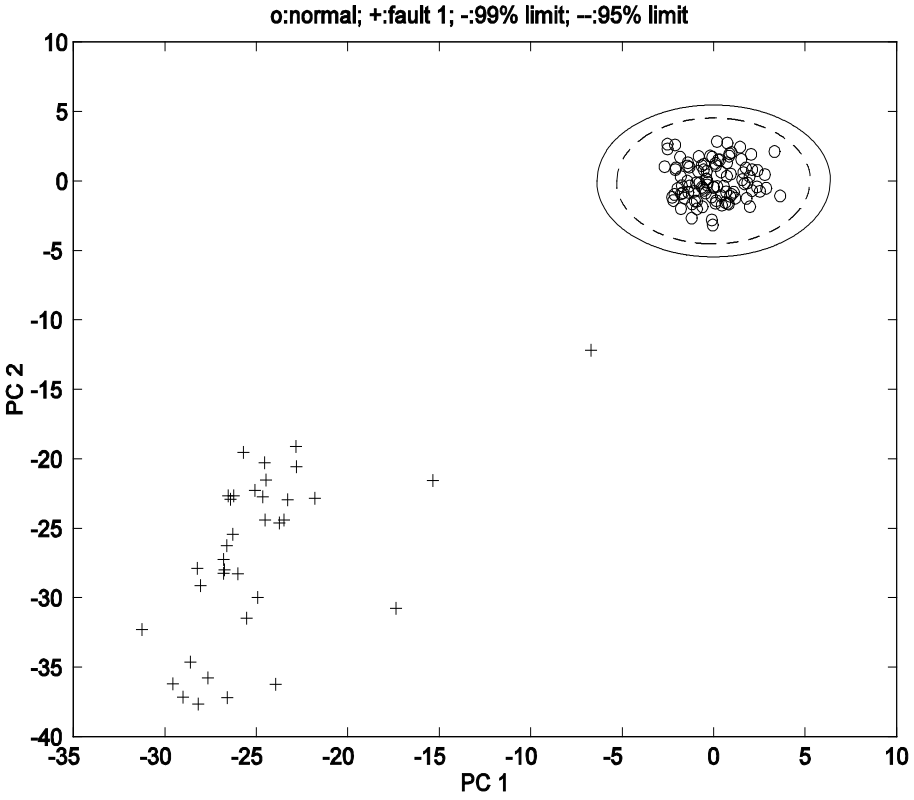


Newcastle University

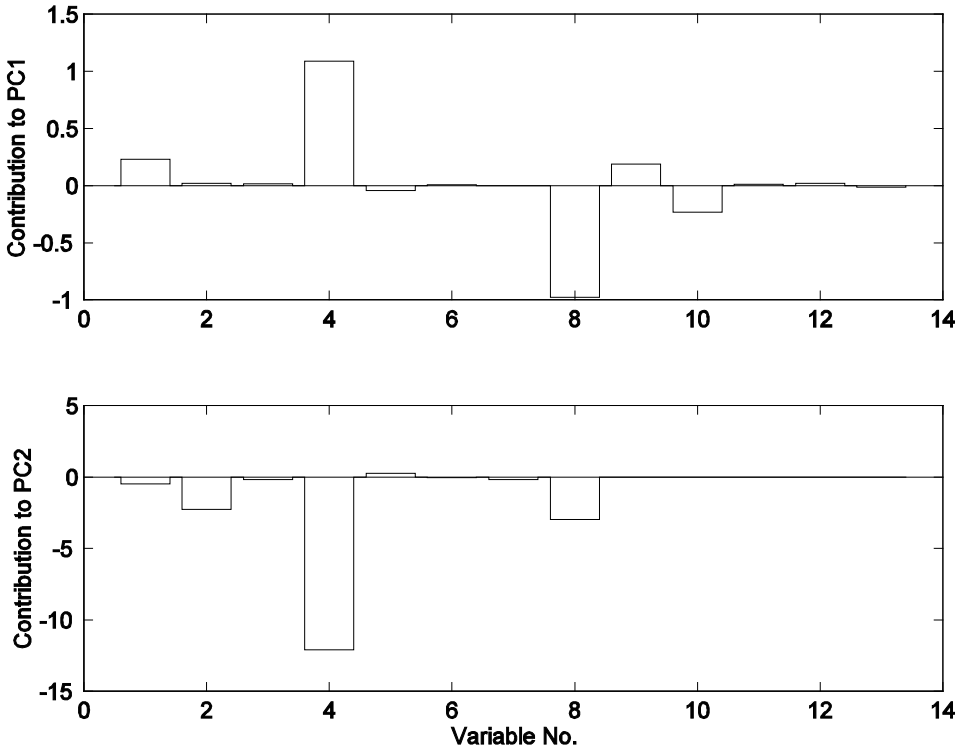
CPACT

MSPC – Fault 1 (cont)

- Plot of the 1st and 2nd PCs under Fault 1



- Contribution plots for the 1st and 2nd PCs under Fault 1

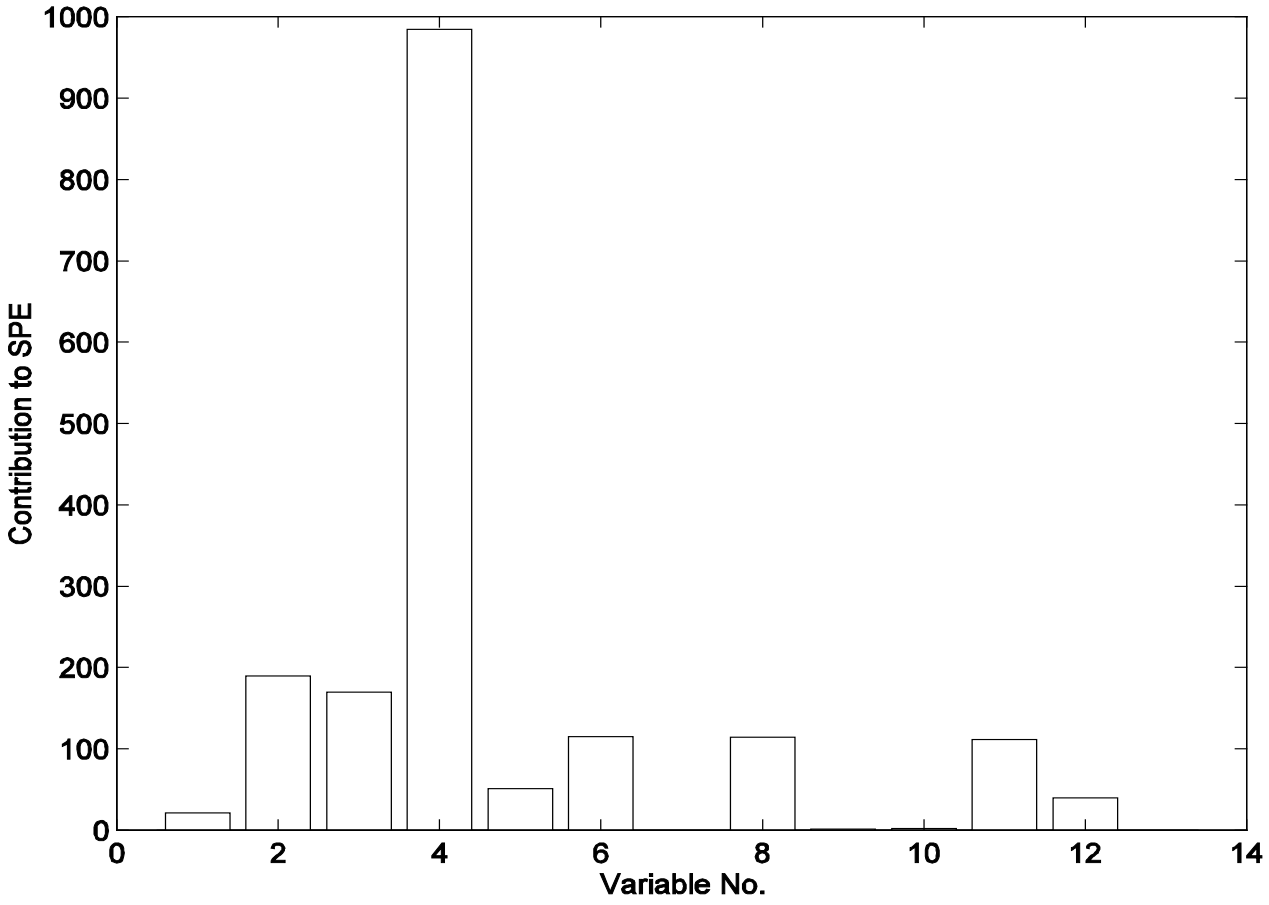


Newcastle University

CPACT

MSPC – Fault 1 (cont)

- Contribution plot for SPE under Fault 1



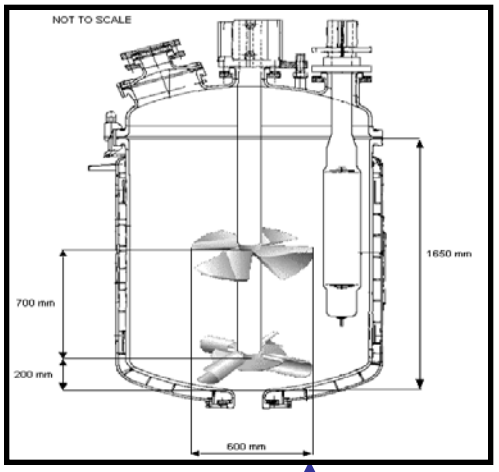
MSPC for Batch Processes

MSPC for Batch Processes

- There is significant interest in the manufacturing of high value added chemicals that are produced in batch reactors that are capable of carrying out multi-processing operations.
- Example processes include crystallisation, brewing, manufacturing of speciality polymers, pharmaceuticals, and biochemicals.
- Global competition requires production to be consistent and of high quality and that the process is operated safely, within environmental guidelines and with minimal energy and raw materials consumption.
- To achieve this, process performance must be monitored in real-time.

Overview of Batch Process MSPC

Reactor



Data Historian
IP-21

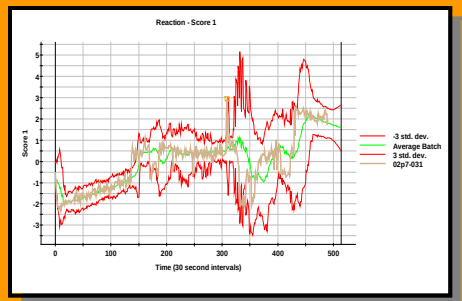
Best Batch Data

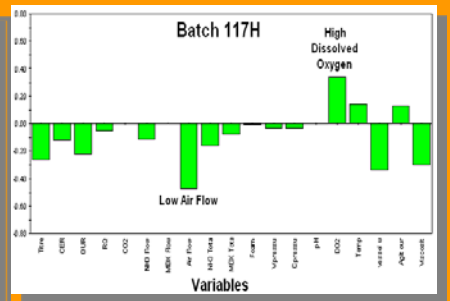
MSPC Model

Process Data
Temp, Flow, Pressure, pH

Current Batch Data

Real Time Performance Monitoring





Feedback Control and Optimisation

Newcastle University

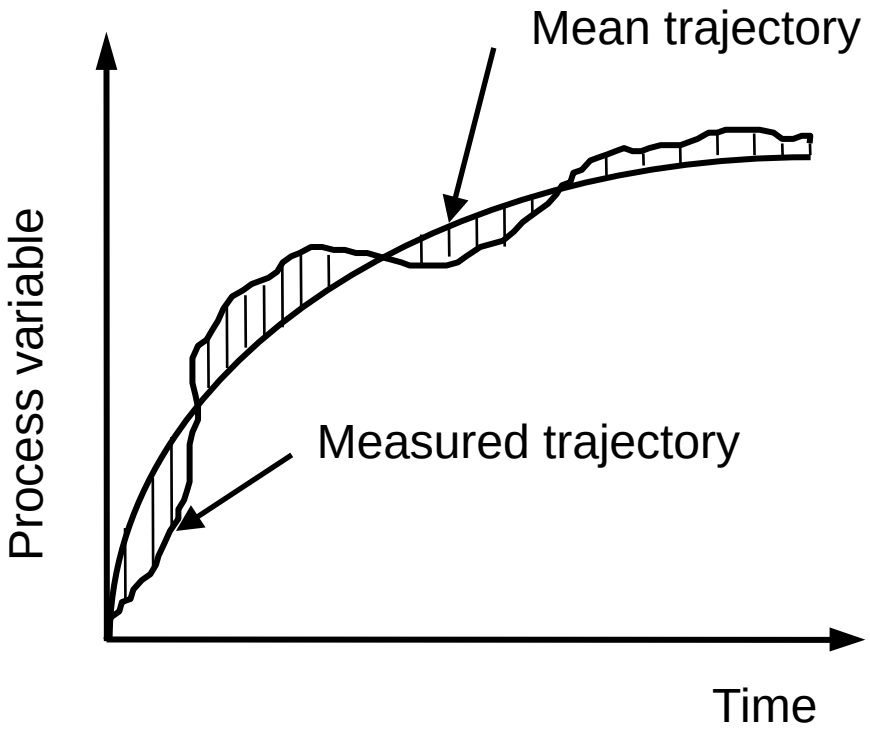
CPACT

MSPC for Batch Processes

- The main characteristics of batch processes are:
  - Finite duration
  - Low volume – small production runs
  - High value products
  - Complex chemical mechanisms
  - Flexibility of production
  - Frequent changing process technology

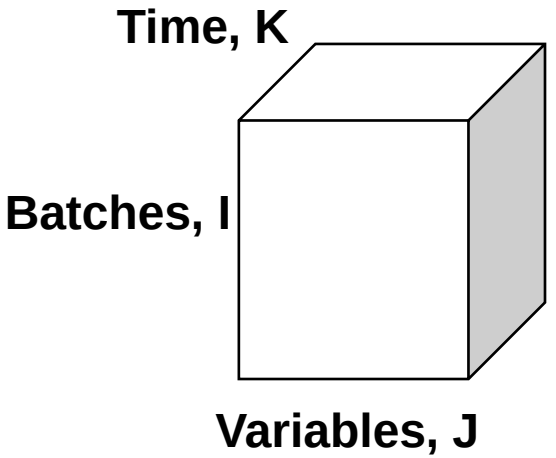
MSPC for Batch Processes

- Batch processes are well known to be nonlinear.
- The nonlinearity can be removed to a great extent through removing the mean batch trajectory from each process variable.
- Process variable variations



Pre-processing of Three-way Data

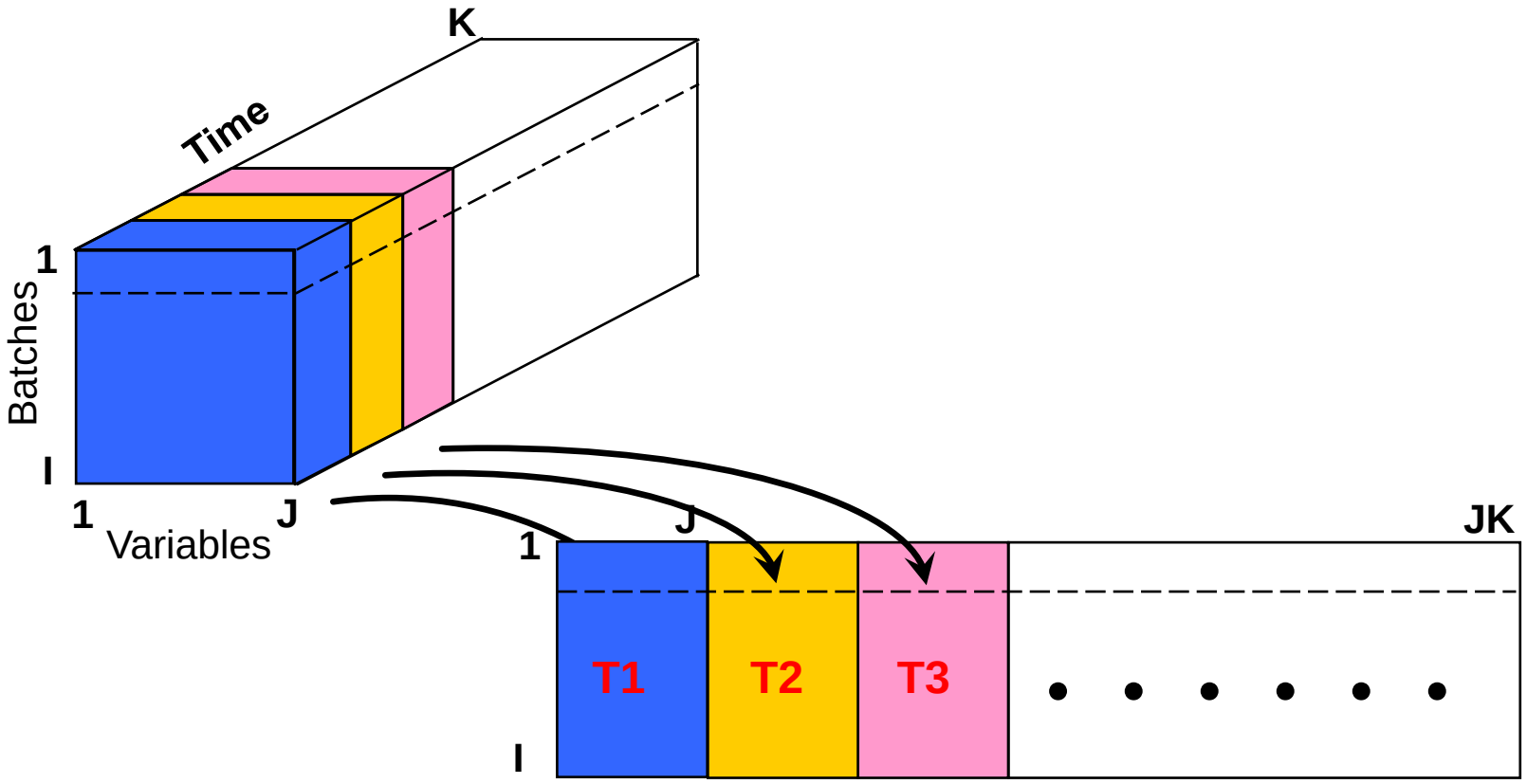
- Batch process data is 3-dimensional
  - Batch (I)
  - Variable (J)
  - Time (K)



- PCA / PLS are bi-linear techniques

- Either
  - Unfold the data into a 2-dimensional array then apply PCA / PLS
- Or
  - Maintain the multi-linear form and apply tri-linear techniques such as multi-linear PLS, PARAFAC etc.

Data Unfolding – Route 1



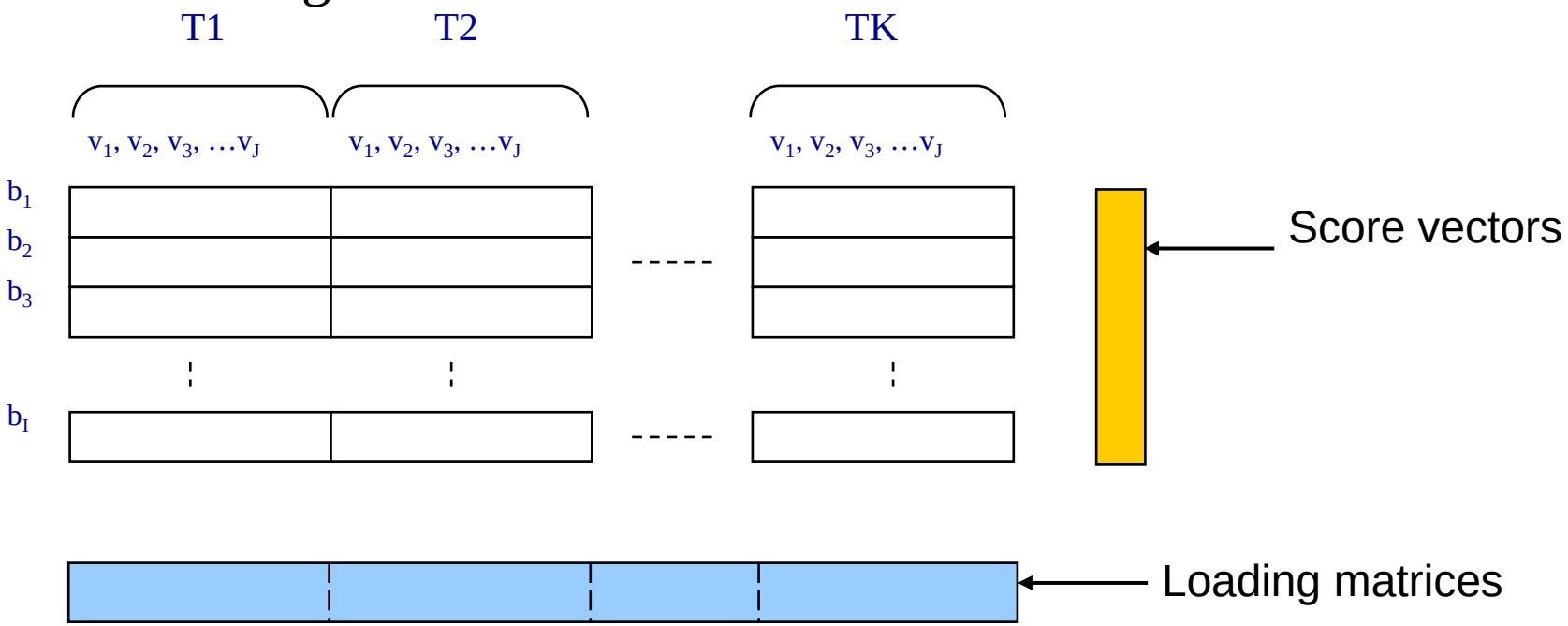
Nomikos and MacGregor (N&M) approach, 1995. “Multivariate SPC charts for monitoring batch processes”. *Technometrics*. **37**, 41 – 59.

Newcastle University

CPACT

Nomikos and MacGregor (N&M) Approach

- Multi-way PCA (MPCA)
- PC score vectors contain information on batch-to-batch variation.
- Loading matrices reflect variable behaviour over time.



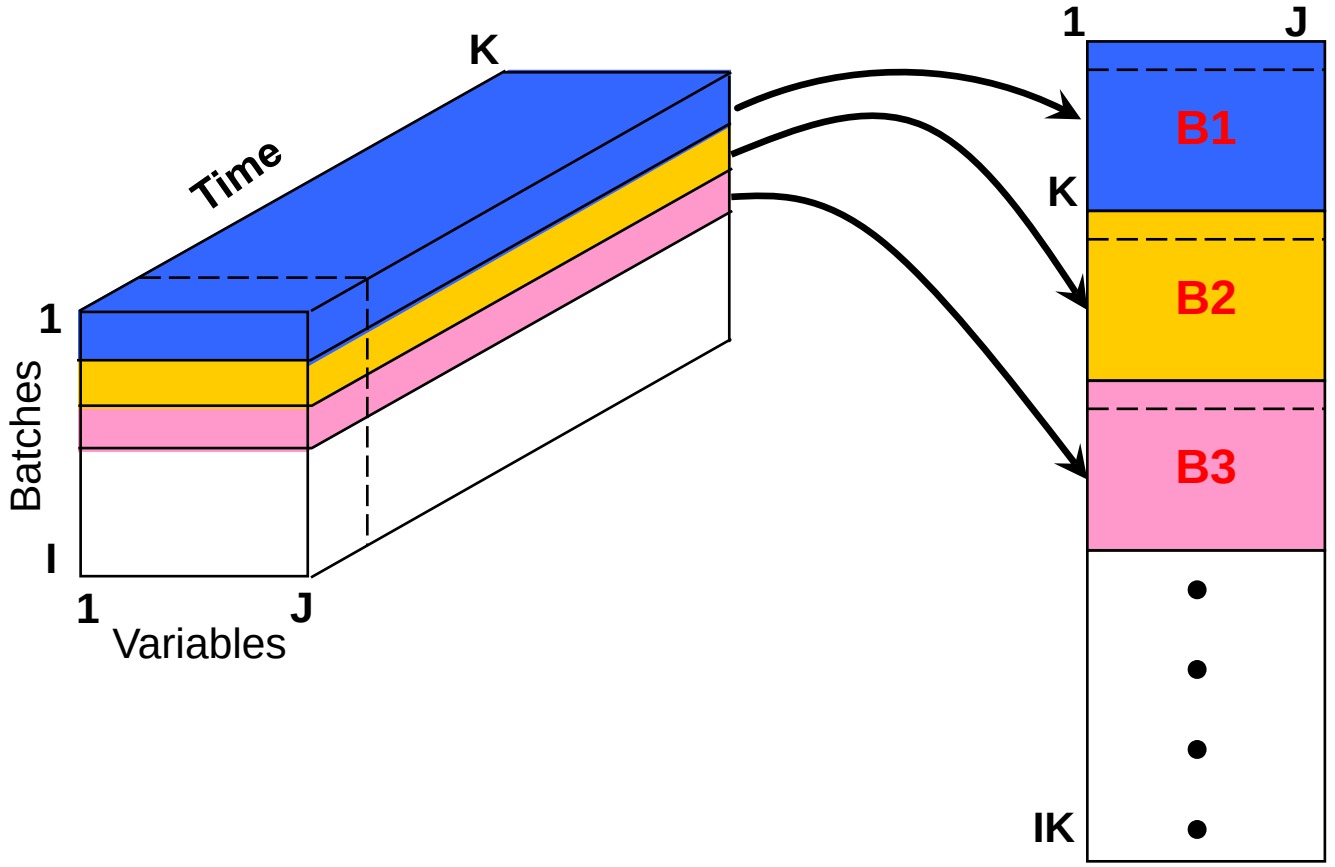
On-line Monitoring

N&M approach

- Monitor the evolution of a new batch using the concept of in-filling future observations.

- Different in-filling methods proposed.
  - Zero deviations ~ assume future measurements to operate along the mean trajectory.
  - Current deviations ~ assume future measurements to continue at the same level as present time.
  - Missing data (Projection) ~ no in-filling applied. Regard the future measurements as missing values.

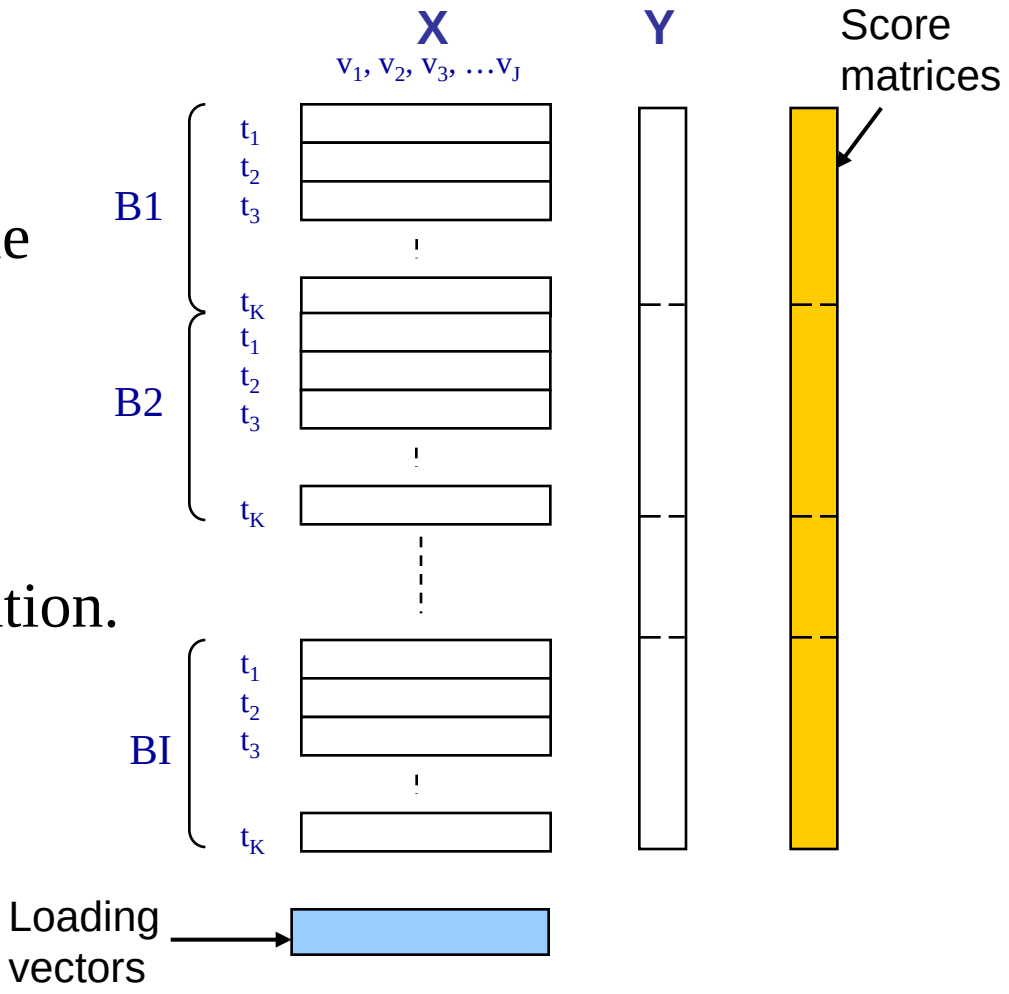
Data Unfolding – Route 2



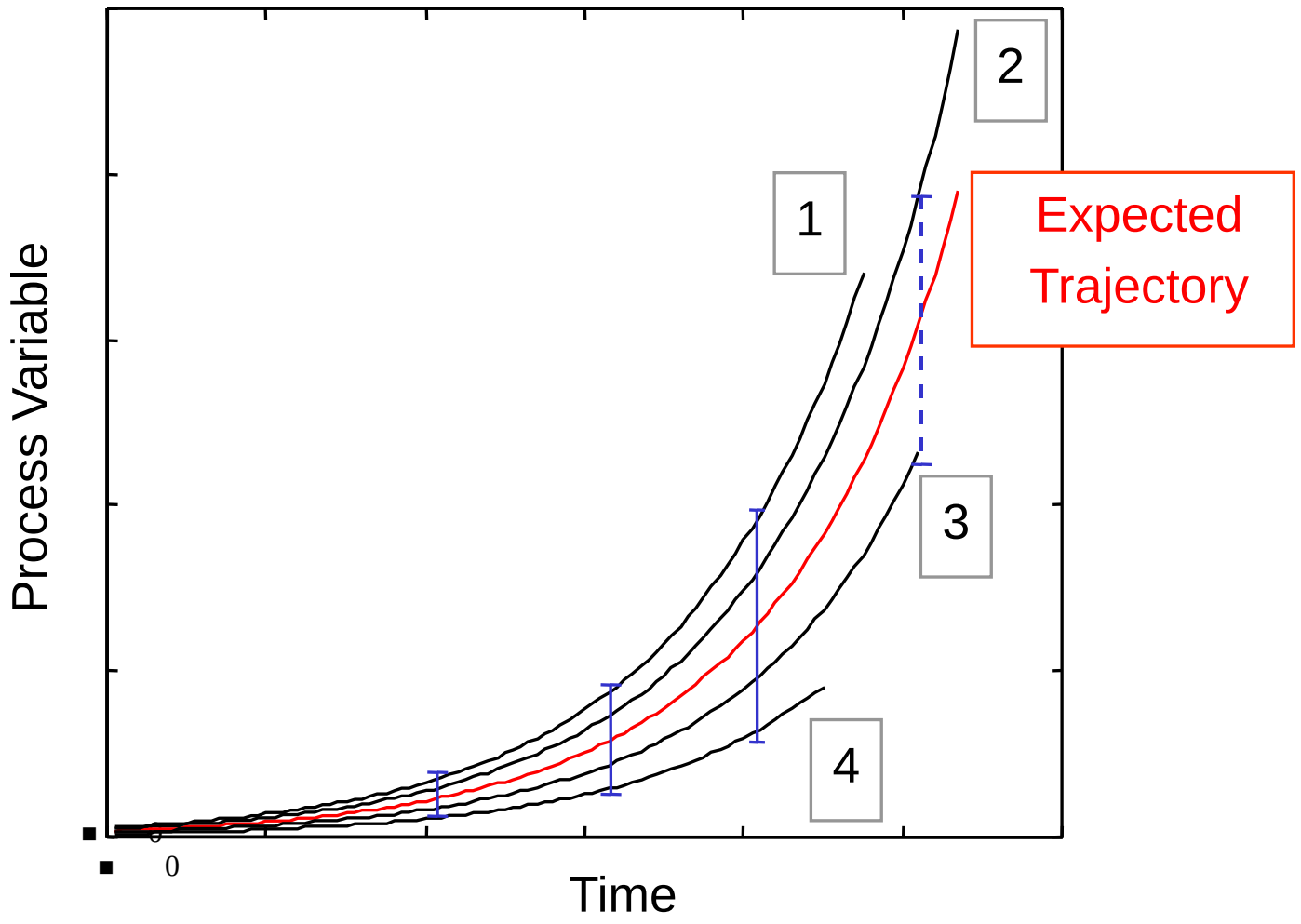
Wold, Kettaneh, Friden and Holmberg approach, 1998. “Modelling and diagnostics of batch processes and analogous kinetic experiments”. *Chemometrics and Intelligent Laboratory Systems*. **44**, 331 – 340.

Wold *et al.* Approach

- Batch observation analysis
- Time maturity as Y variable for PLS analysis.
- Score matrices contain information on batch evolution.
- Loading vectors contain information on variables.



Unequal Length Batches

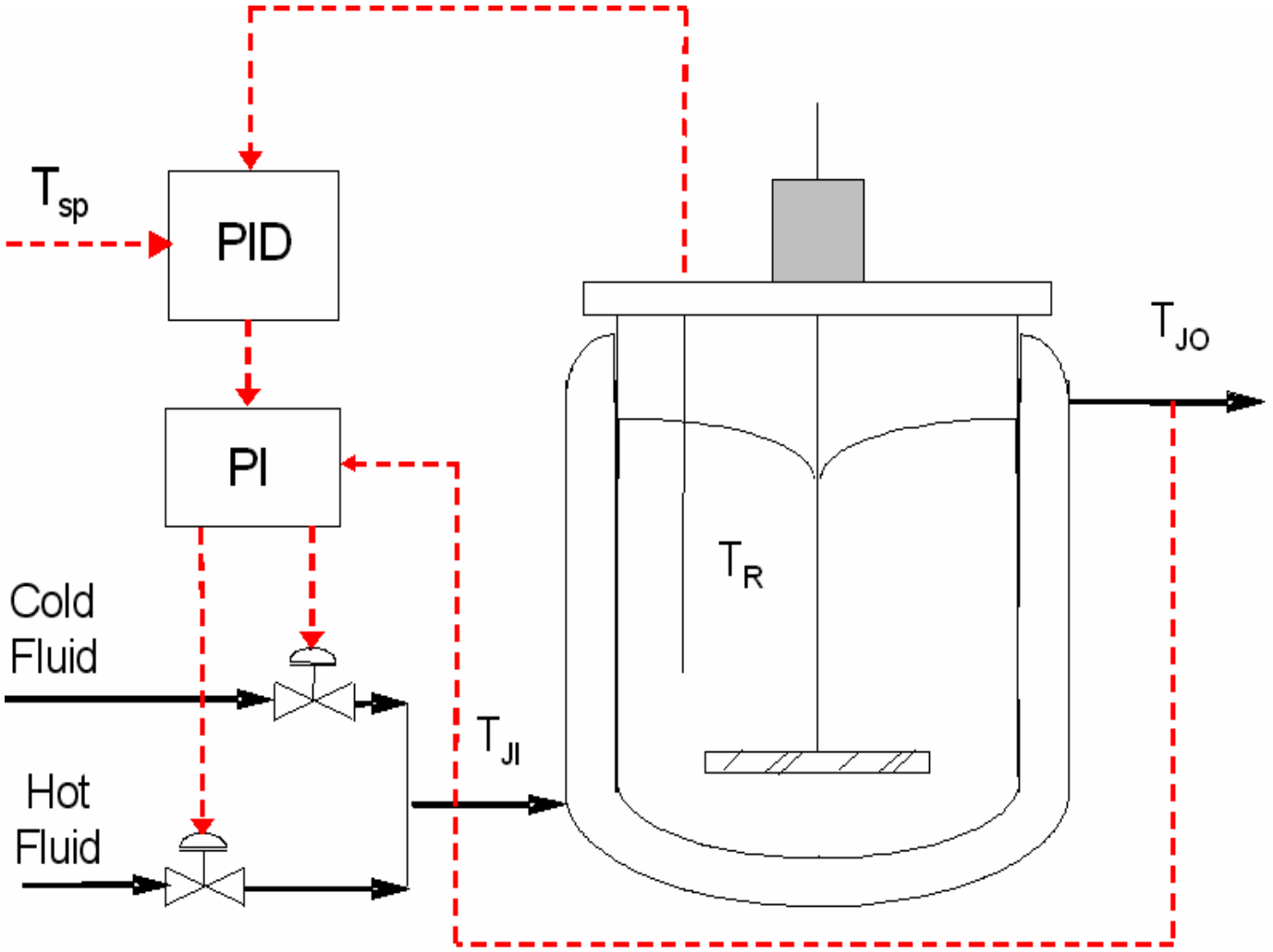


Batch Data Alignment

- The aim of this work is to investigate three approaches for dealing with different length batches and to compare their performance in differentiating between nominal and faulty batch production as part of a multivariate statistical process control scheme.

- Three approaches are:
  - cut to minimum length
  - dynamic time warping
  - use of an indicator (surrogate) variable

Application to a Batch Polymerisation Process

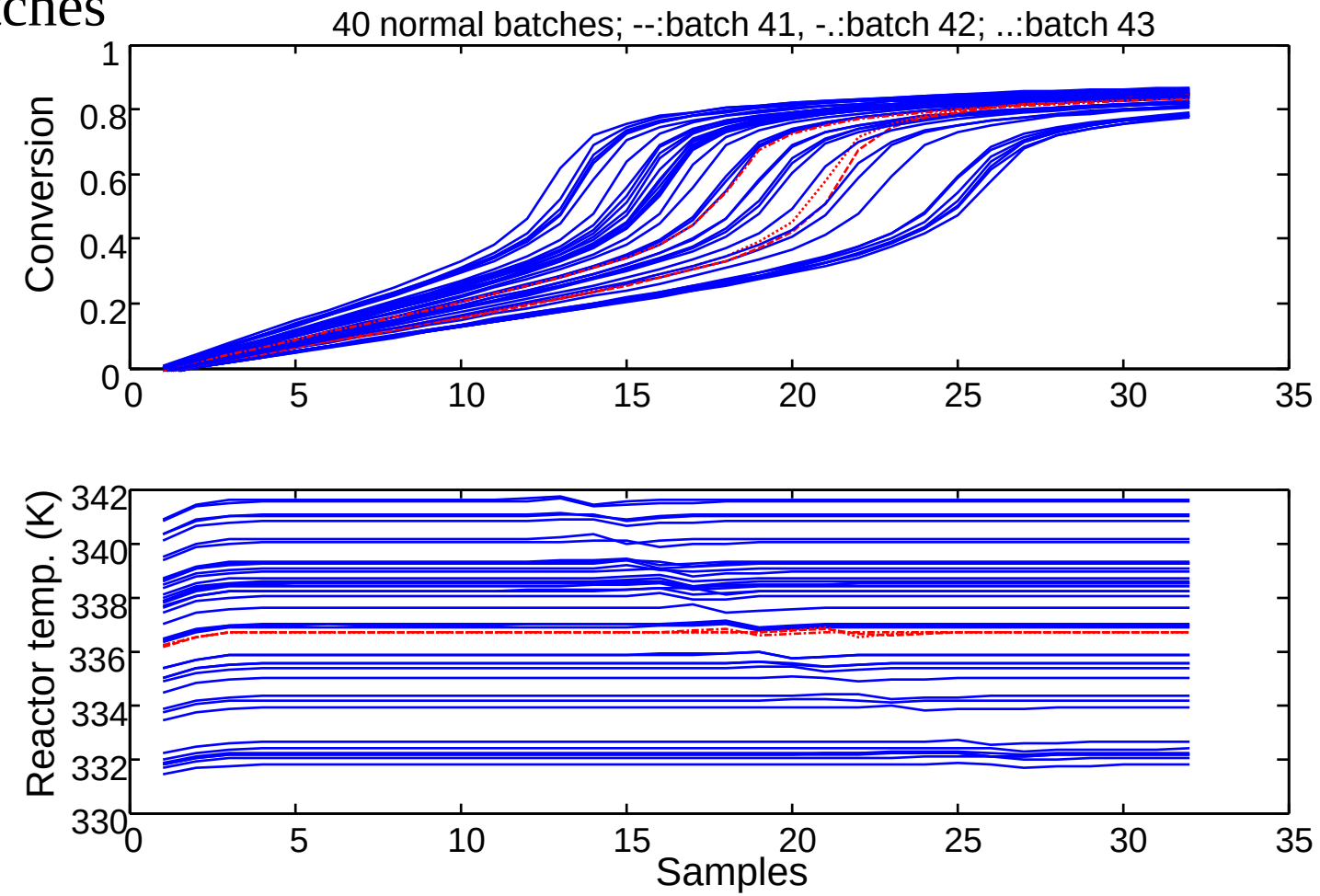


Application to a Batch Polymerisation Process

- 40 normal batch runs and 3 faulty batch runs
- A multi-way PCA model with 38 normal batches were developed
- The 2 unused normal batches and the 3 faulty batches were used to test the multi-way PCA model to see if it can successfully classify the unseen normal batches as being normal and detect the faulty batches
- The three faulty batches:
  - Batch 41: with reactive impurities
  - Batch 42: with reactor fouling
  - Batch 43: with both reactive impurities and reactor fouling

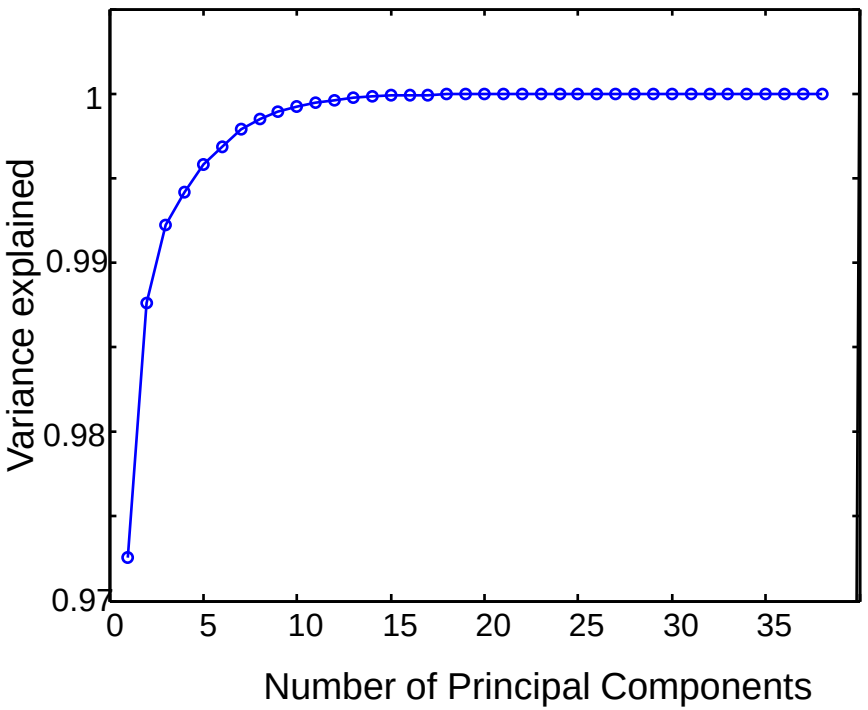
Application to a Batch Polymerisation Process

- Trajectories of temperature and conversion of different batches

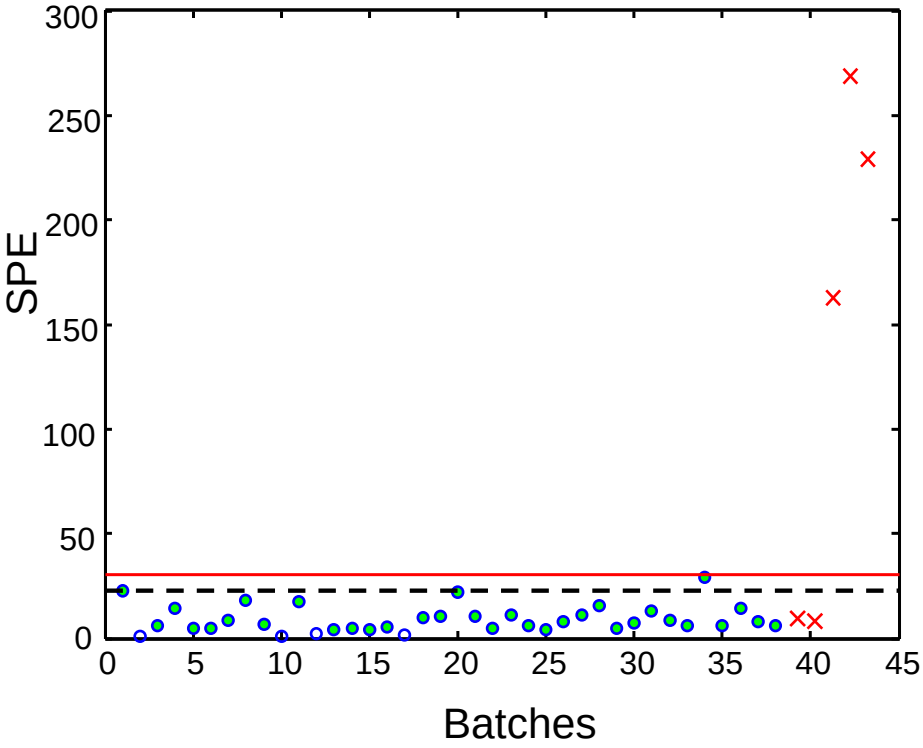


Application to a Batch Polymerisation Process

- Accumulated data variance explanation

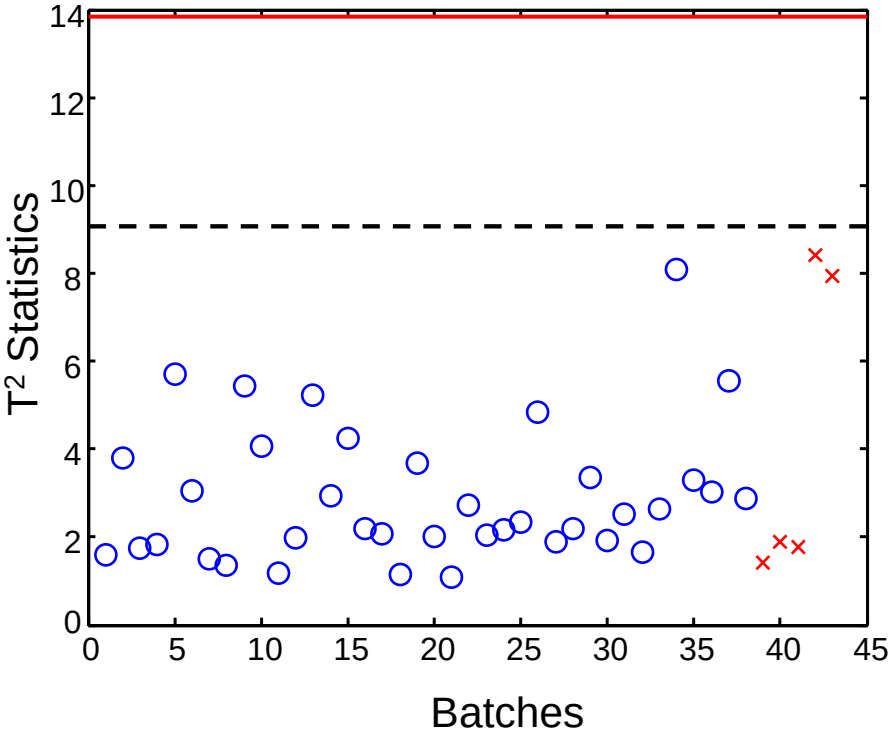


- SPE of the PCA model

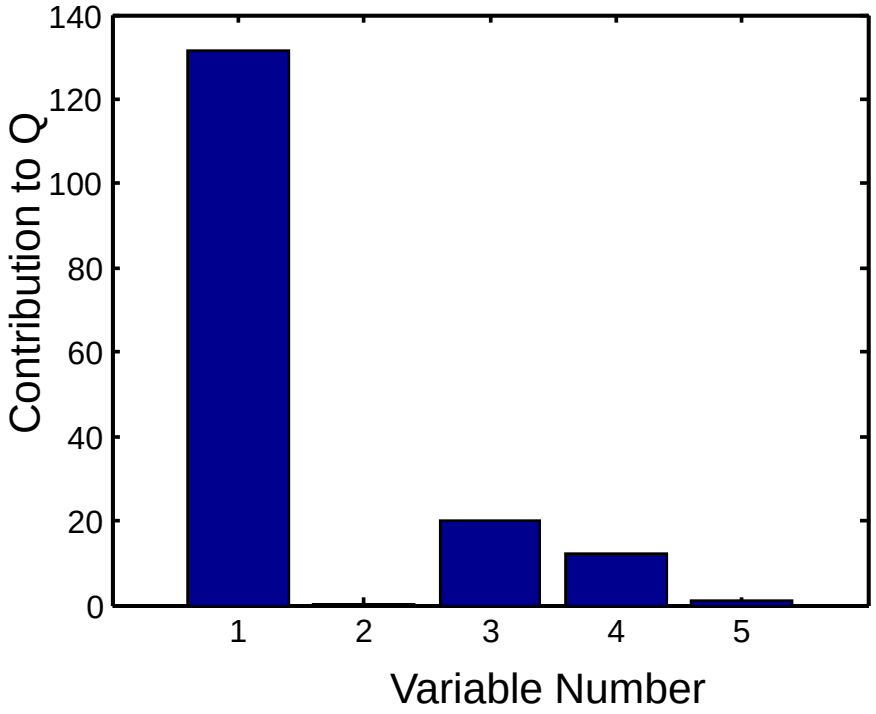


Application to a Batch Polymerisation Process

- T^2 plot

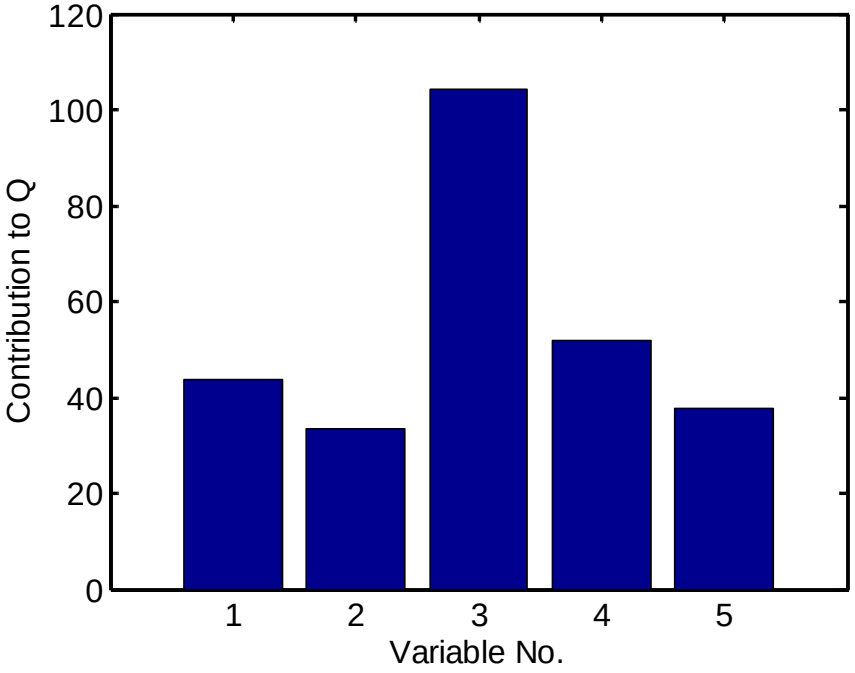


- Contribution plot for SPE of Batch 41

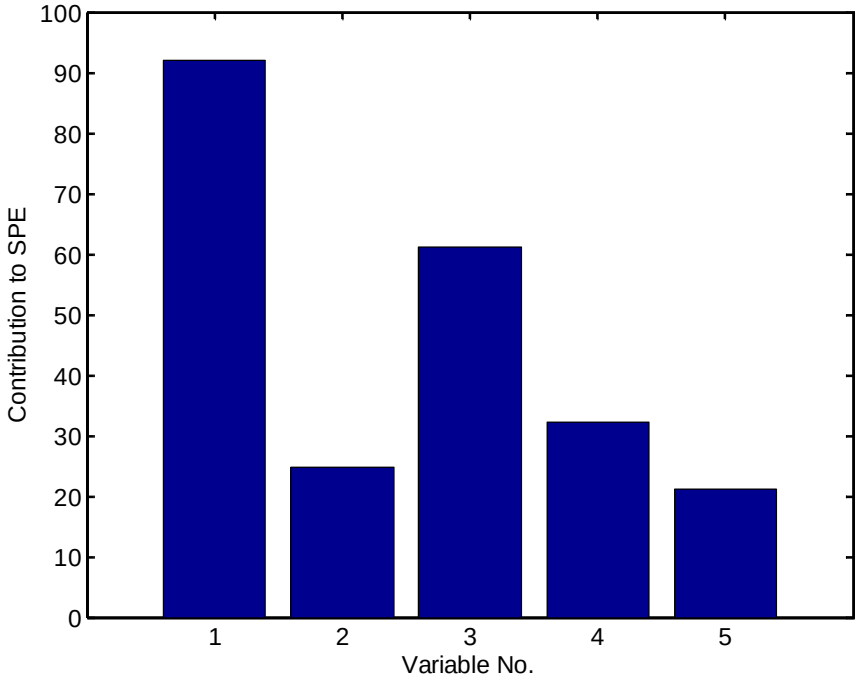


Application to a Batch Polymerisation Process

- Contribution plot for SPE of Batch 42



- Contribution plot for SPE of Batch 43

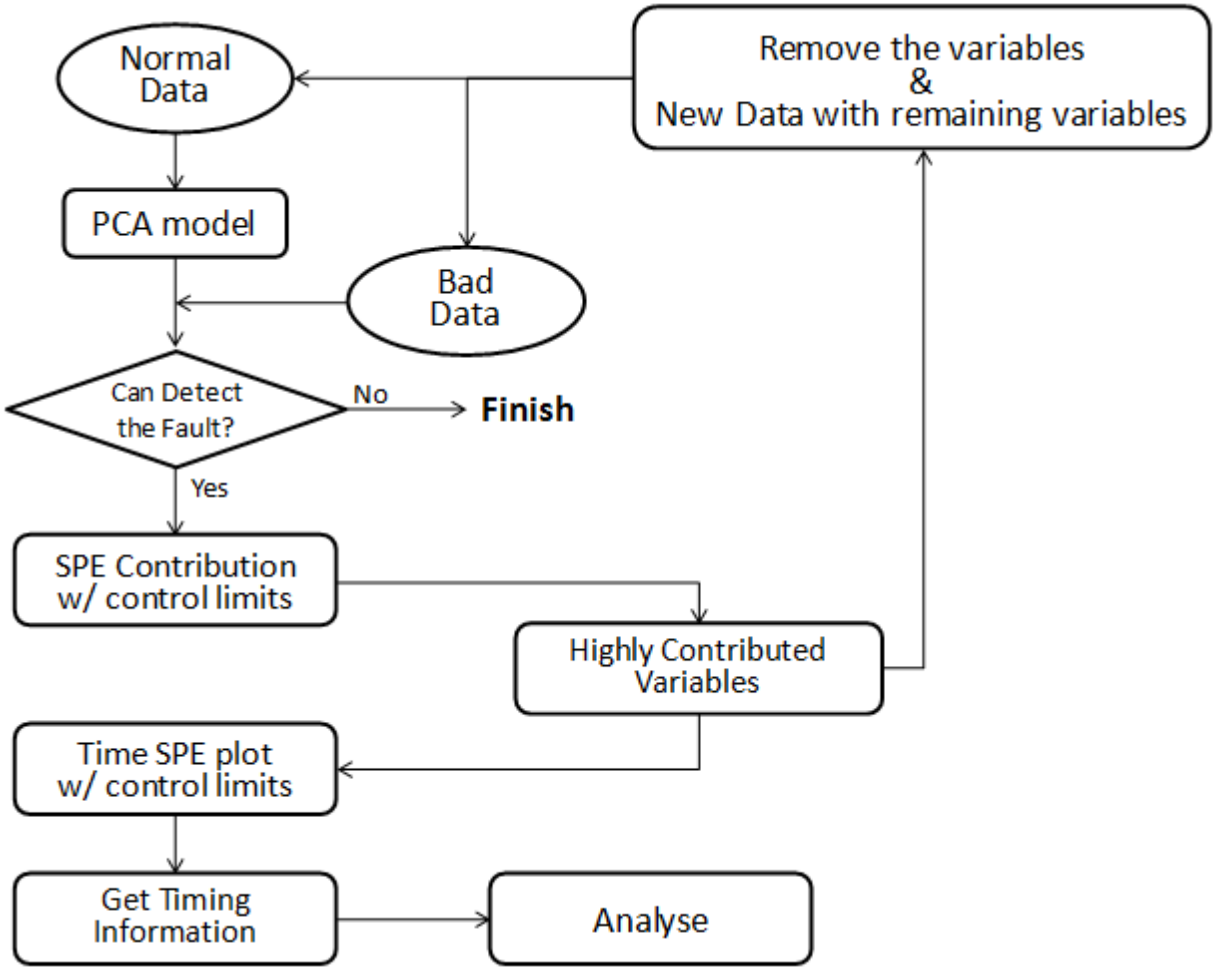


Fault localisation through progressive PCA modelling

(J. J. Hong, J. Zhang, and J. Morris, “Fault localization in batch processes through progressive principal component analysis modeling”, *Ind. Eng. Chem. Res.*, 50(13), 2011, 8153-8162.

J. J. Hong, J. Zhang, and J. Morris, “Progressive multi-block modelling for enhanced fault isolation in batch processes”, *Journal of Process Control*, 2014, 24, 3–26.)

Progressive PCA modelling



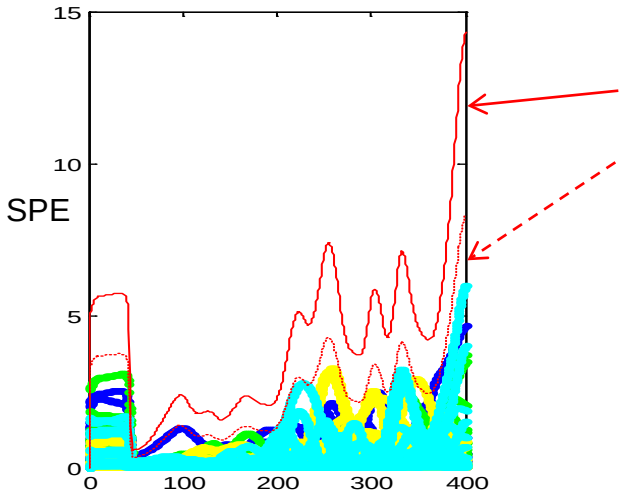
Time series SPE

SPE for variable *j* at time *k* in batch *i*

$$SPE_{ij,k} = (x_{ij,k} - \hat{x}_{ij,k})^2$$

$x_{ij,k}$: data represents variable *j* in the i^{th} batch measured at the k^{th} time point

$\hat{x}_{ij,k}$: PCA model prediction of $x_{ij,k}$



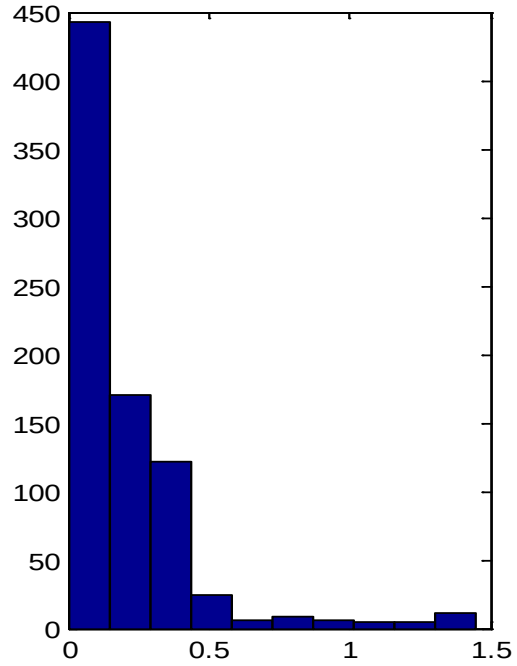
99% Control limits

95% Control limits

Control Limits

Weighted chi-squared distribution can provide a good approximation for the multinormal distribution (Box, 1954).

Control limit



Histogram of SPEs samples

$$\lim_{\alpha} = g\chi^2_{h,\alpha}$$

g : weight
h : degree of freedom
α : Significance level

At a single time point,

Mean and variance of the distribution = Mean and variance of the SPE sample

$$\lim_{\alpha} = (v/2m)\chi^2_{2m^2/v,\alpha}$$

m : mean of sample SPE

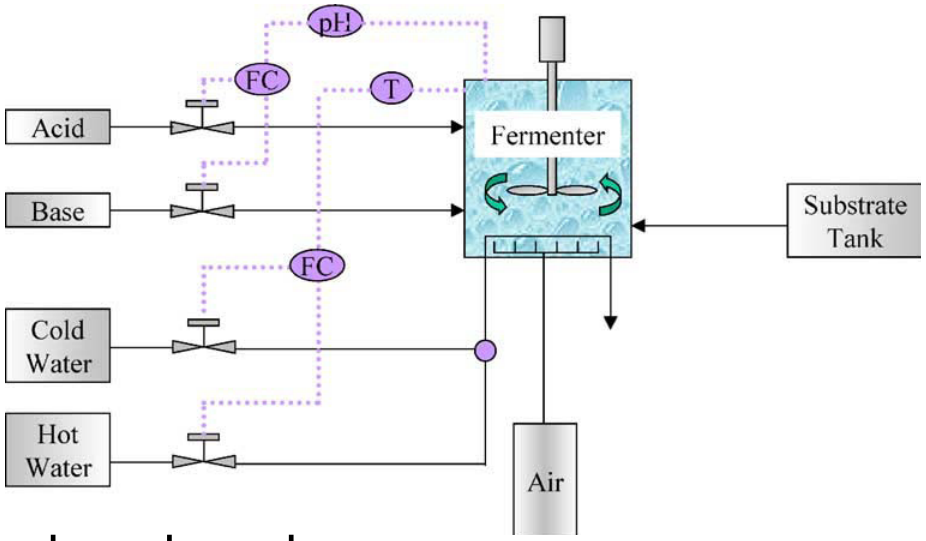
v : variance of sample SPE

99% (when α is 0.01) and 95% (when α is 0.05) control limits are estimated.

Newcastle University

CPACT

Case study



PenSim benchmark process

- Penicillin Production Process running for 400hrs.
- 11 variables are used for process monitoring (Sampling Time: 0.5h).
- 73 normal batches (58 for training and 15 for validation data) and 4 faulty batches

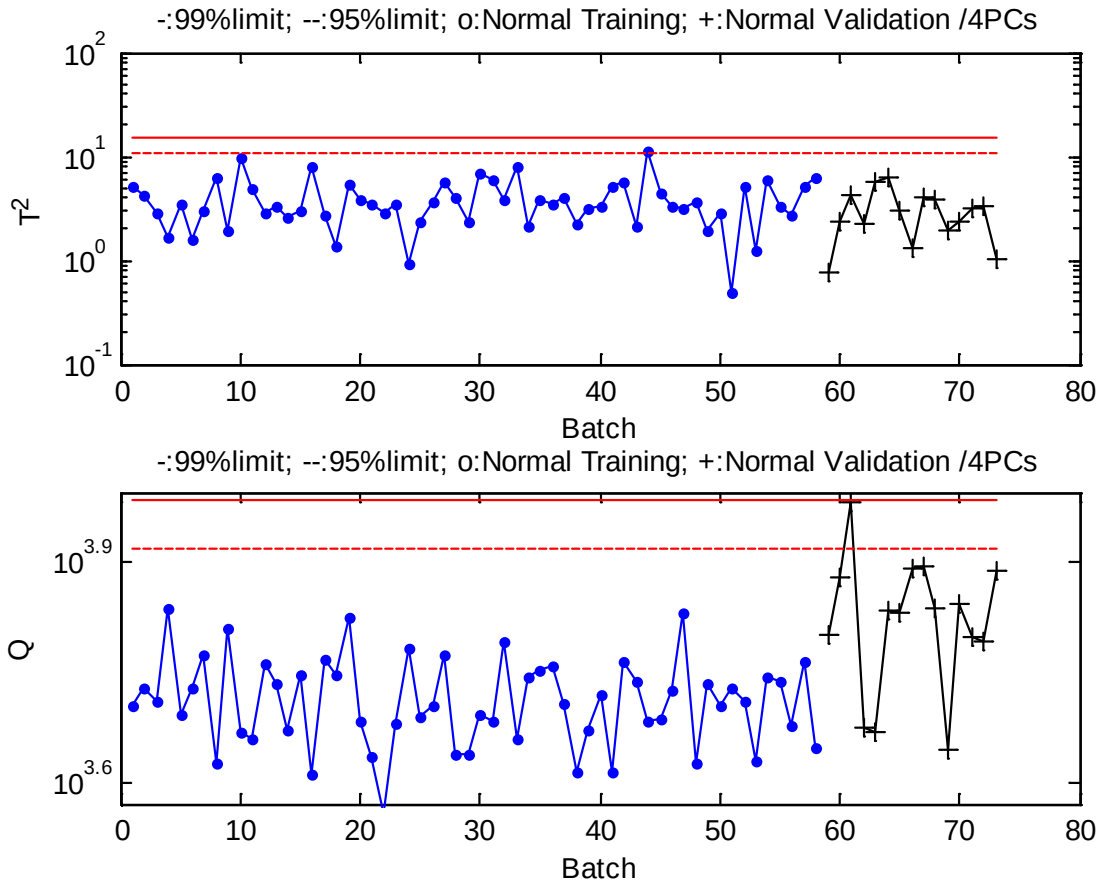
Case study

- Faulty batches:

	Variable that disturbance is introduced
1	Variable #3 (substrate feed rate): -0.03%, ramp, 160h ~ 250h
2	Variable #3 (substrate feed rate): -10%, step, 205h ~ End
3	Variable #8 (pH): pH controller failure, 0h ~ End
4	Initial substrate conc. decreases from 14.9g/L to 11g/L

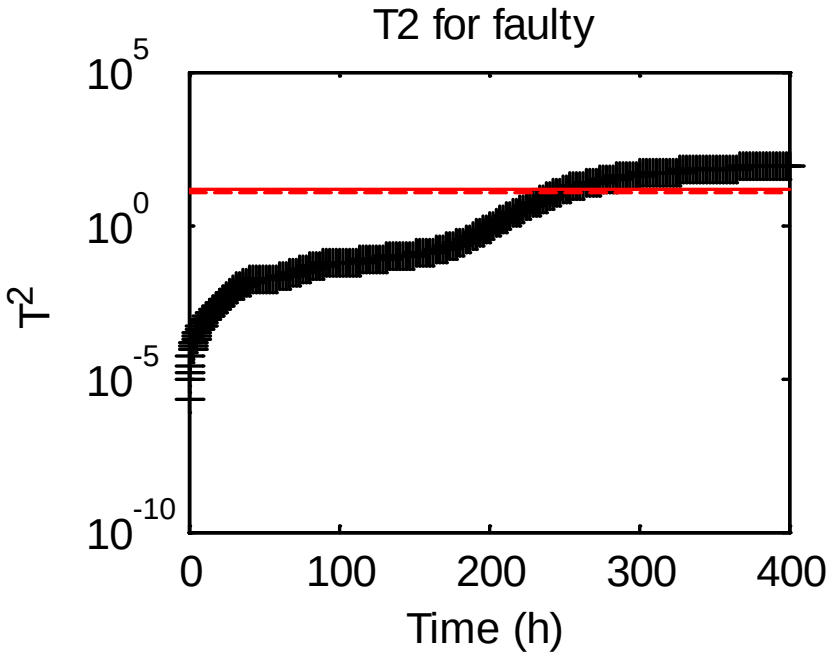
Nominal PCA model

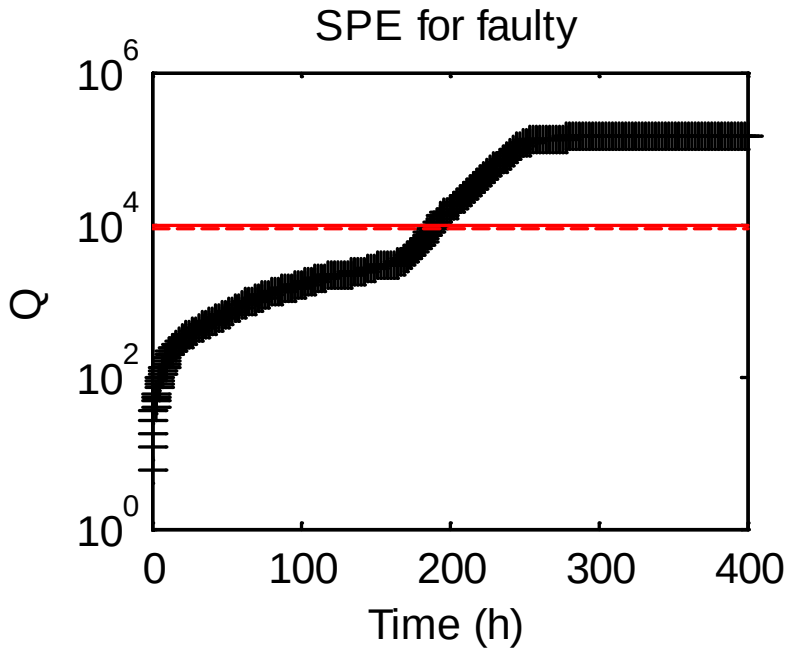
- The nominal PCA model contains 4 PCs



Fault 1

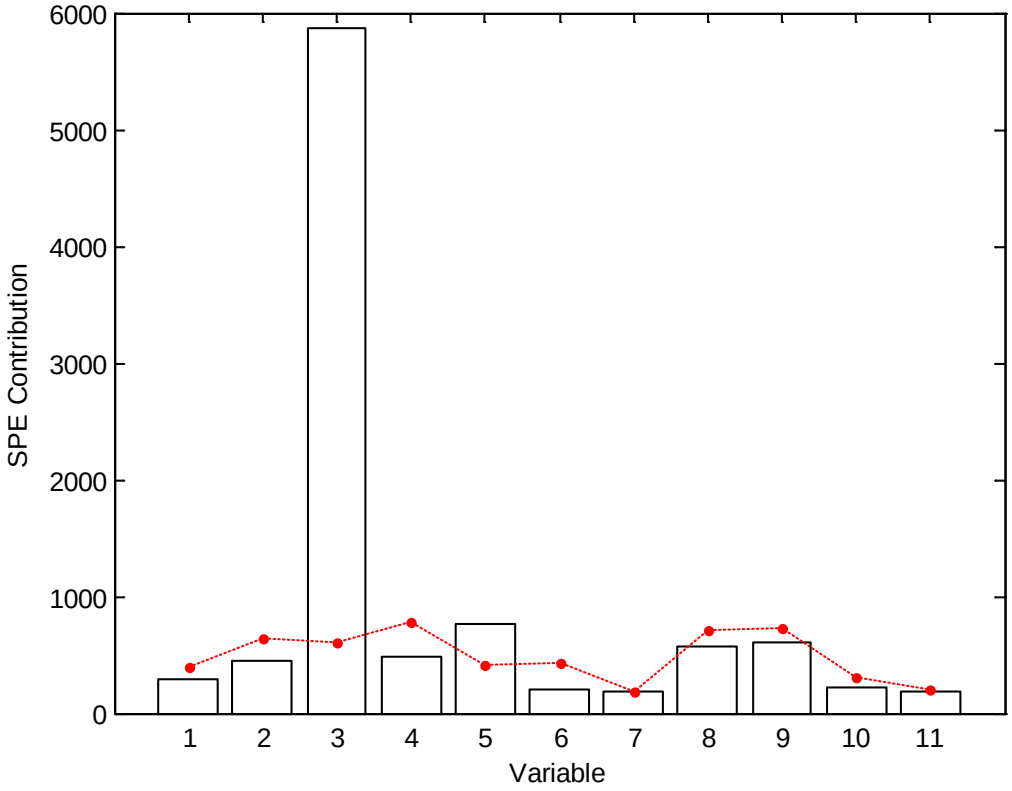
- The PCA model detect the fault first on the Q monitoring chart and then on the T^2 monitoring chart (step 1)





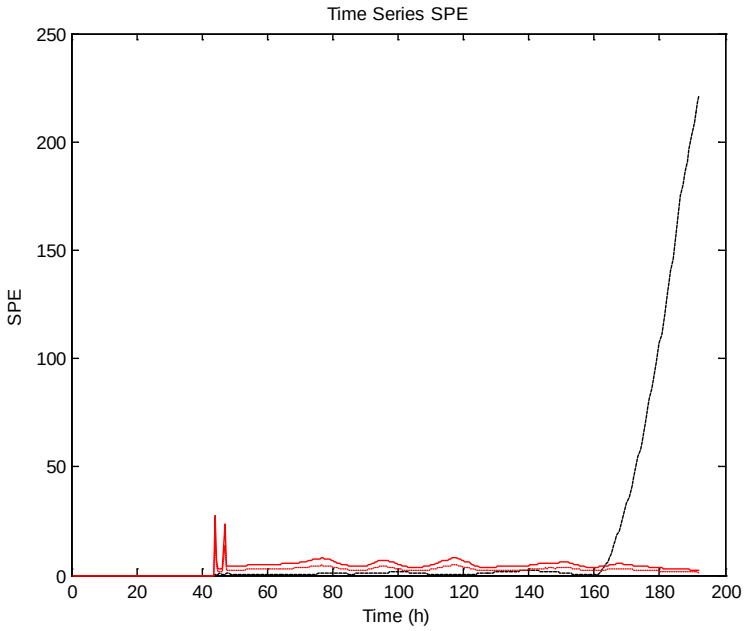
Fault 1

- Contribution plot shows that variables 3 (substrate feed rate) and 5 (dissolved oxygen) contribute to the large SPE.

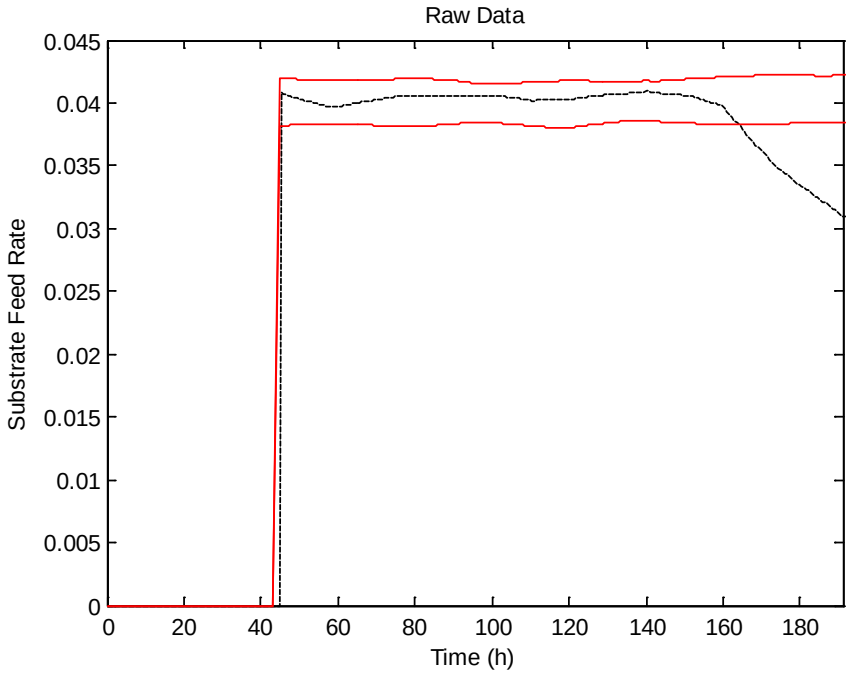


Fault 1

- Time series SPE plot for substrate feed rate (abnormality observed at 163.5h)



- Raw data plot for substrate feed rate (abnormality observed at 164h)

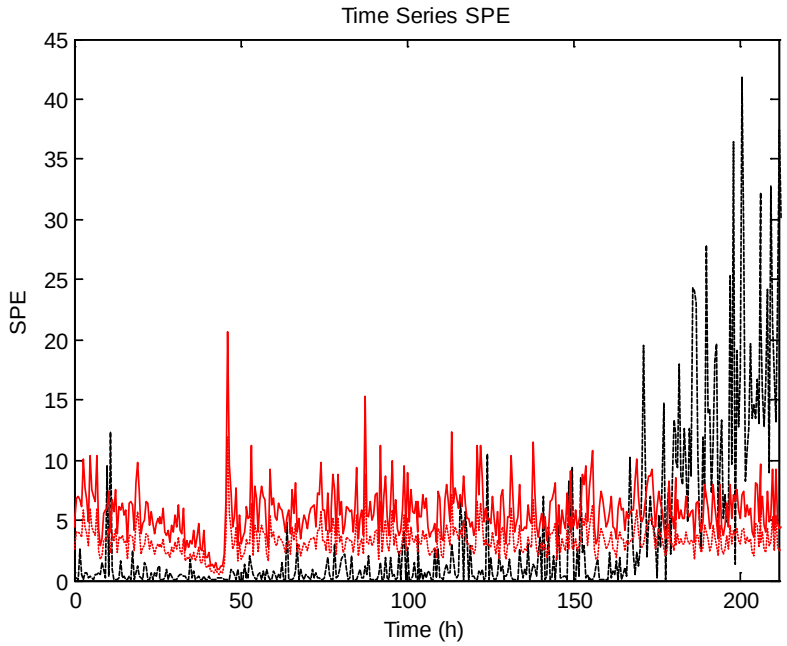


Newcastle University

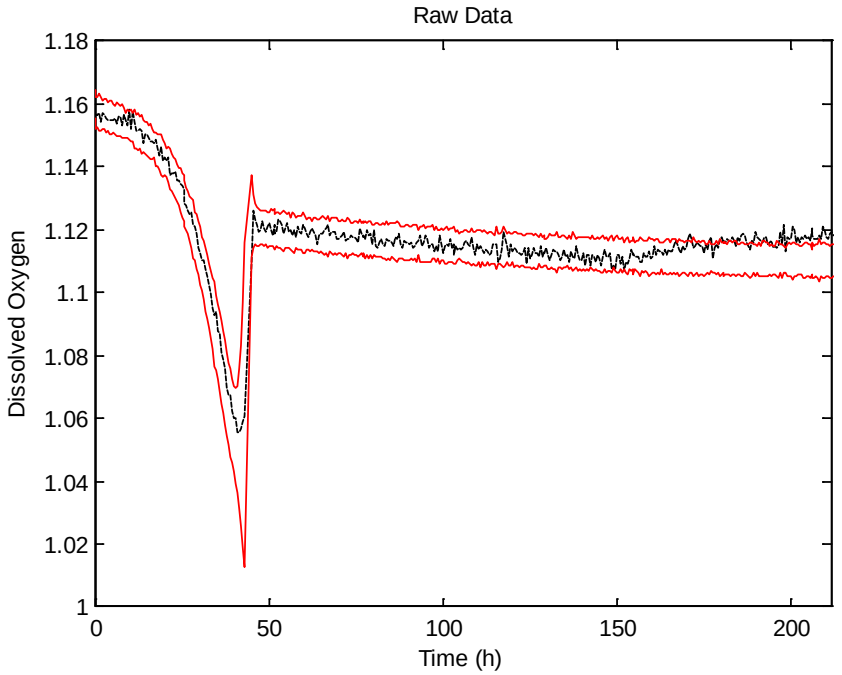
CPACT

Fault 1

- Time series SPE plot for dissolved oxygen rate (abnormality observed at 180h)



- Raw data plot for dissolved oxygen rate (abnormality observed at 185h)

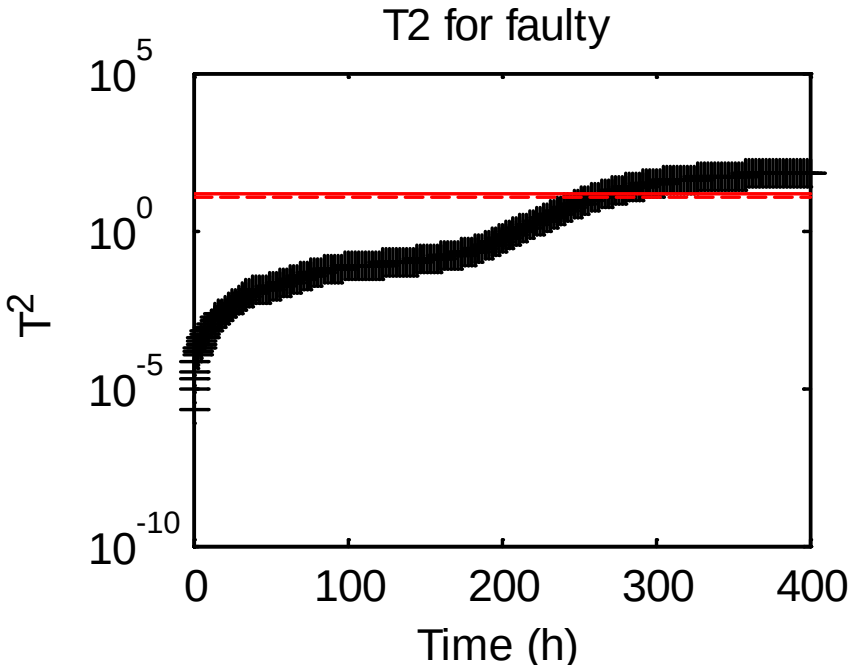


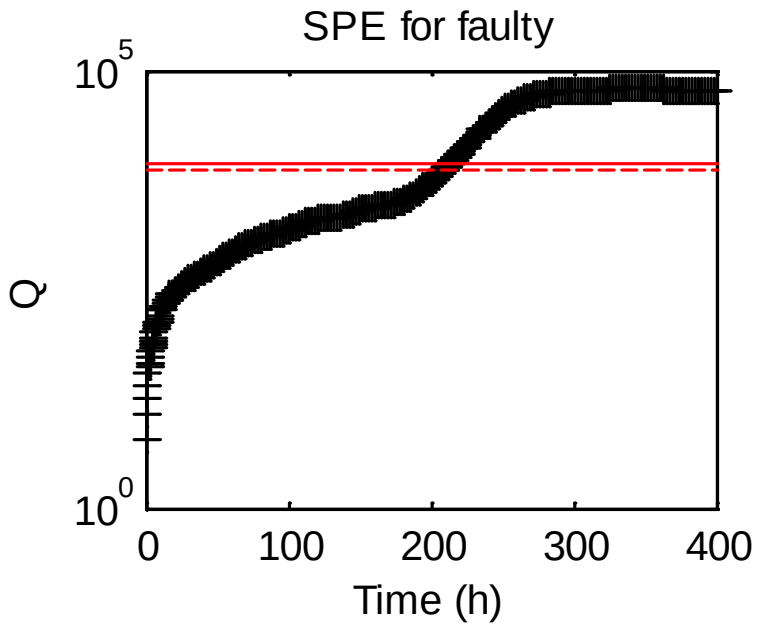
Newcastle University

CPACT

Fault 1

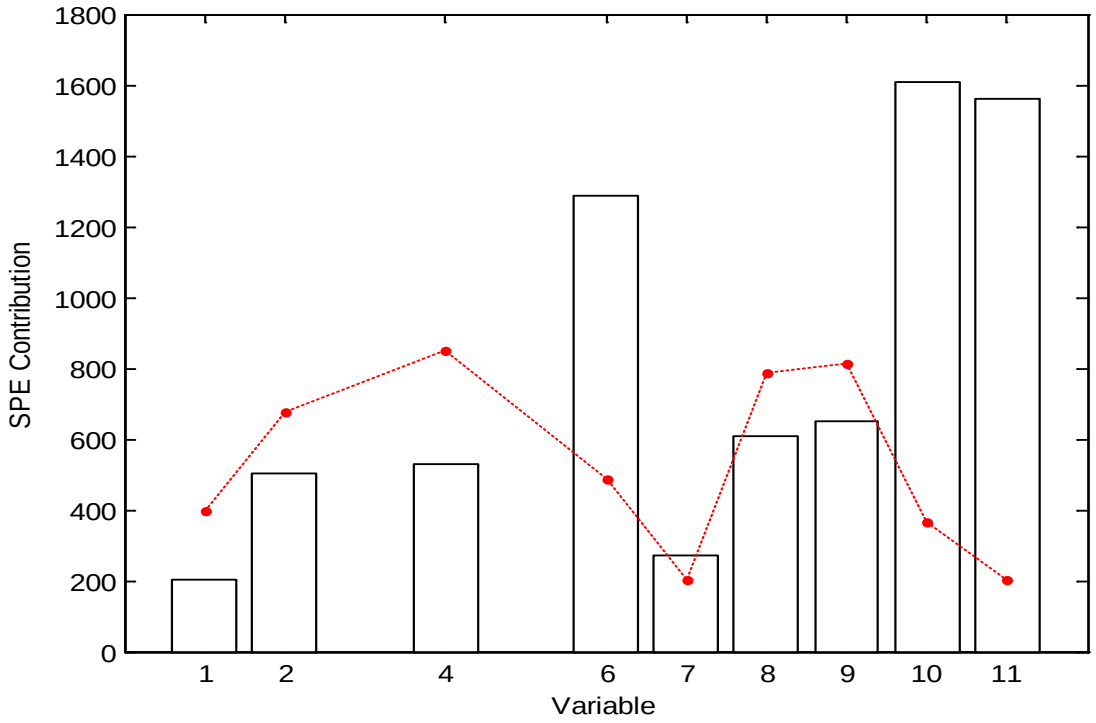
- Monitoring charts for PCA model with variables 3 and 5 removed (Step 2)





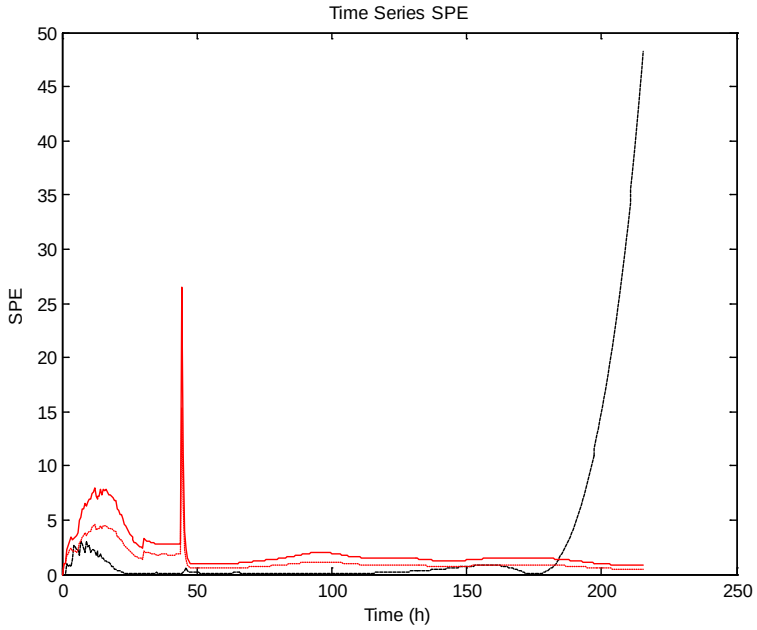
Fault 1

- SPE contribution plot shows that variables 6 (culture volume), 10 (generated heat), and 11 (cold water flow rate) are responsible for large SPE

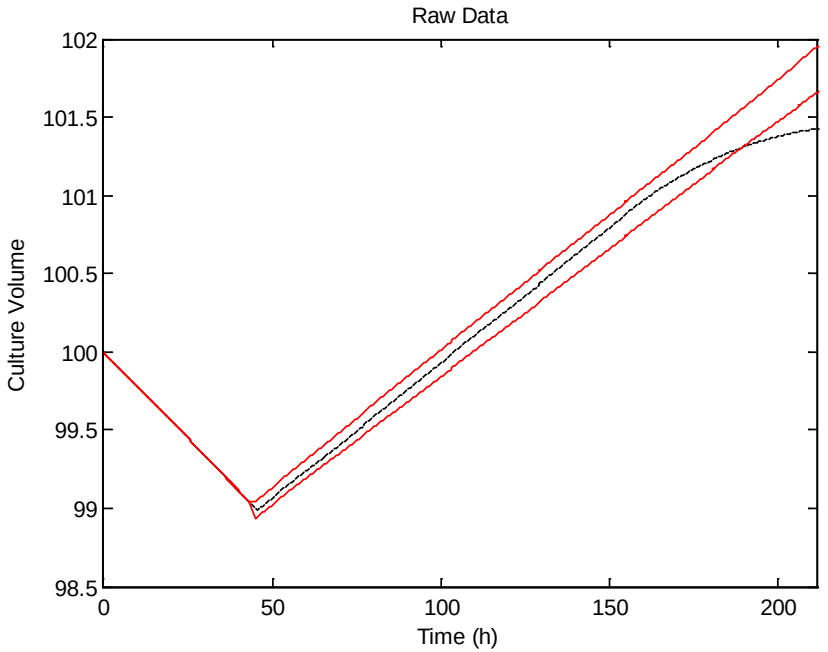


Fault 1

- Time series SPE plot for culture volume

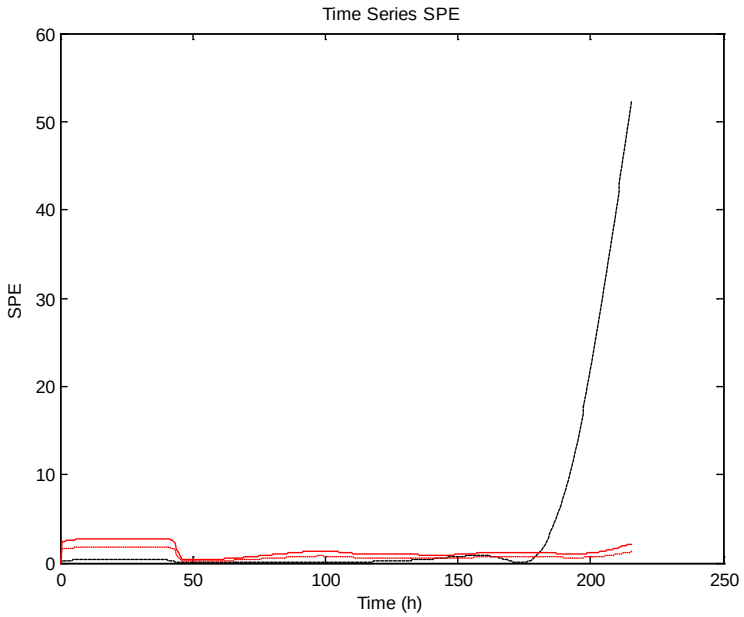


- Raw data plot for culture volume

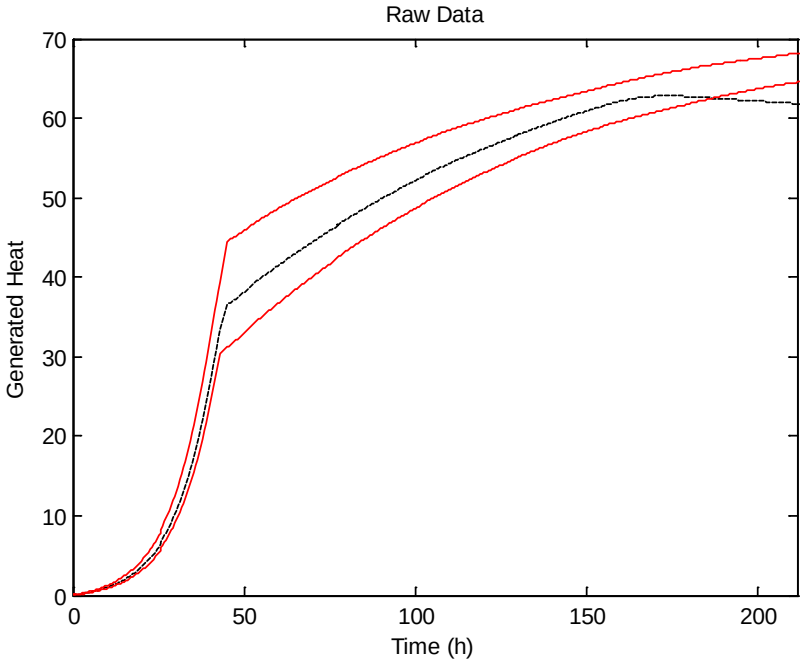


Fault 1

- Time series SPE plot for generated heat



- Raw data plot for generated heat

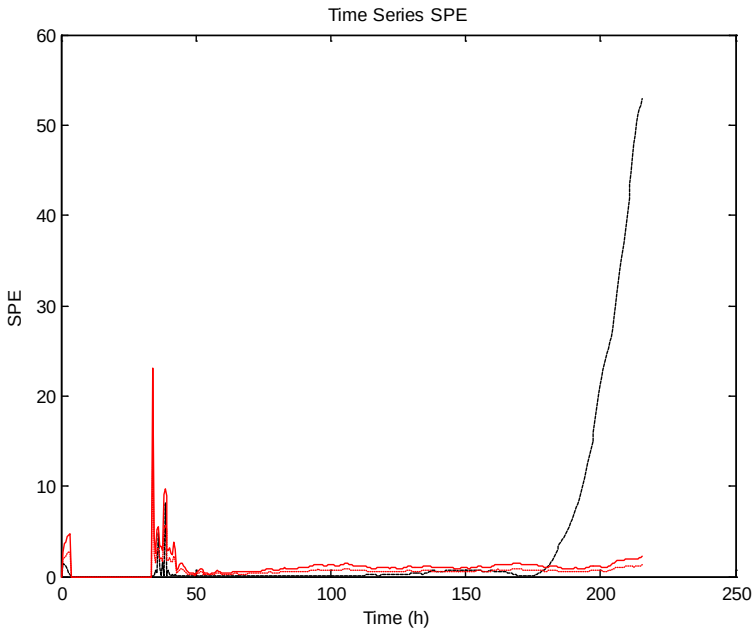


Newcastle University

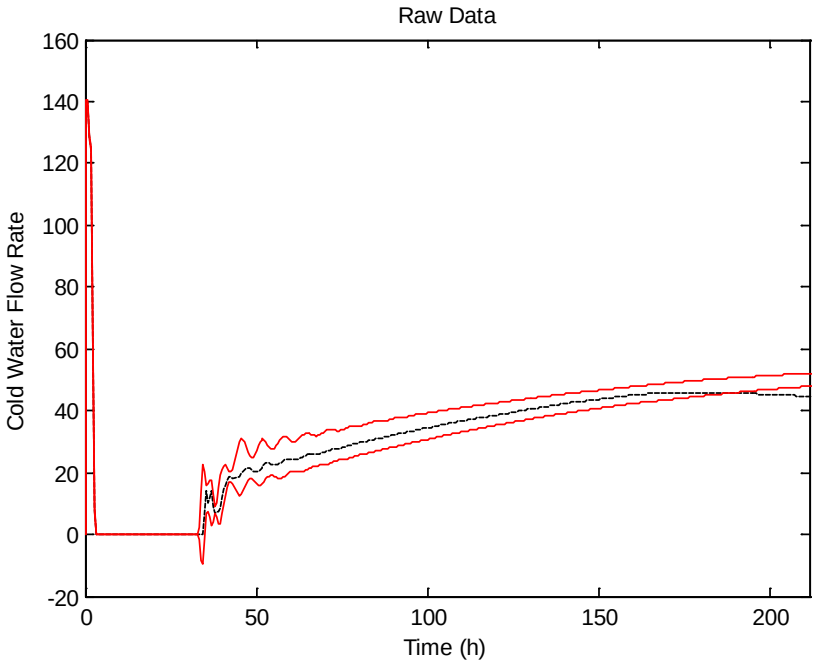
CPACT

Fault 1

- Time series SPE plot for cold water flow rate



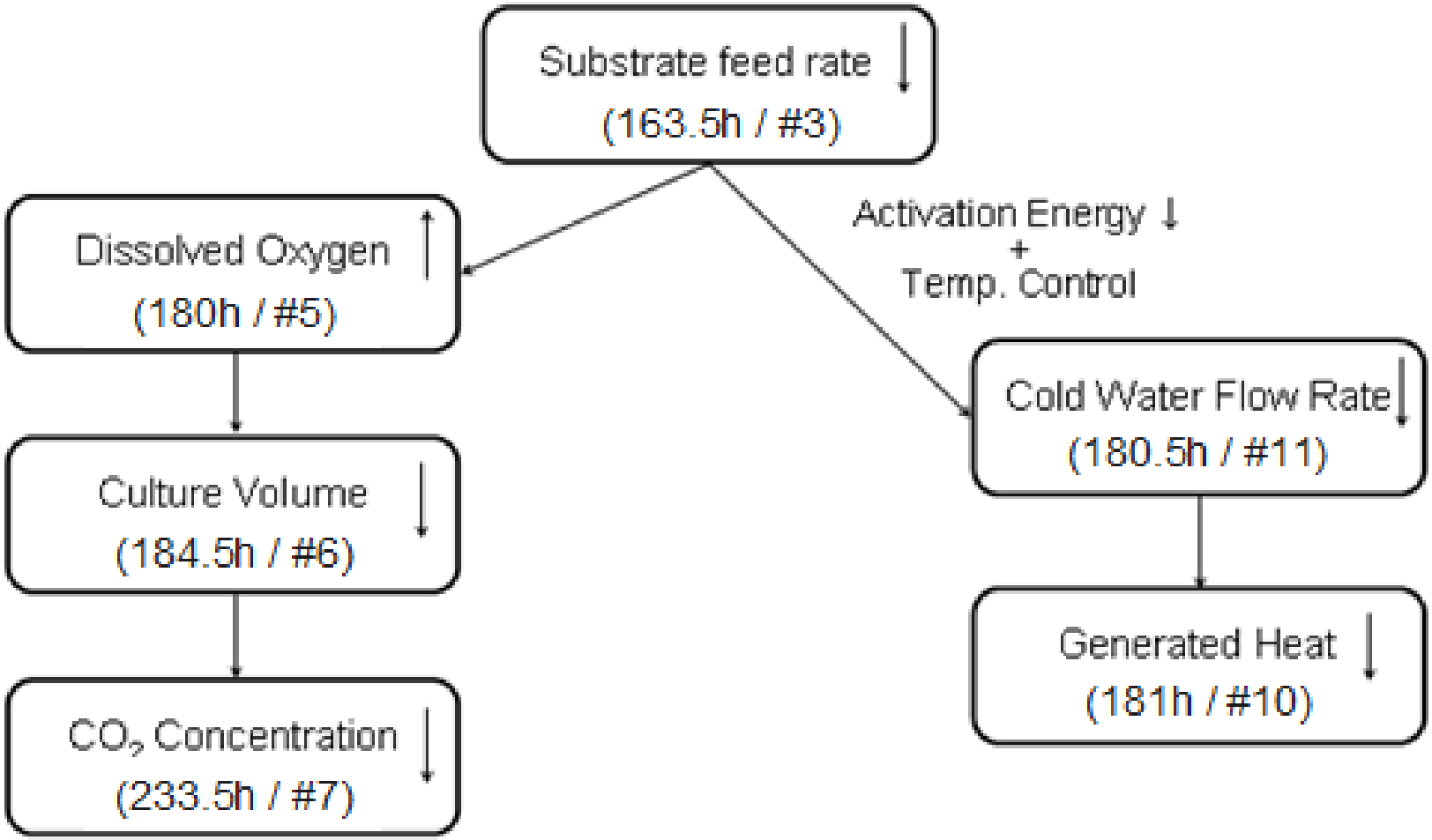
- Raw data plot for cold water flow rate



Result summary for the 1st faulty batch (Online)

Step No.	Identified variables	Time of detection (time series SPE plot)	Time of detection (raw data plot)
1	#3 (Substrate Feed Rate)	163.5h	164h
	#5 (Dissolved Oxygen)	180h	185h
2	#6 (Culture Volume)	184.5h	189.5h
	#10 (Generated Heat)	181h	187.5h
	#11 (Cold Water Flow Rate)	180.5h	188h
3	#7 (CO_2 Concentration)	233.5h	Not Detected

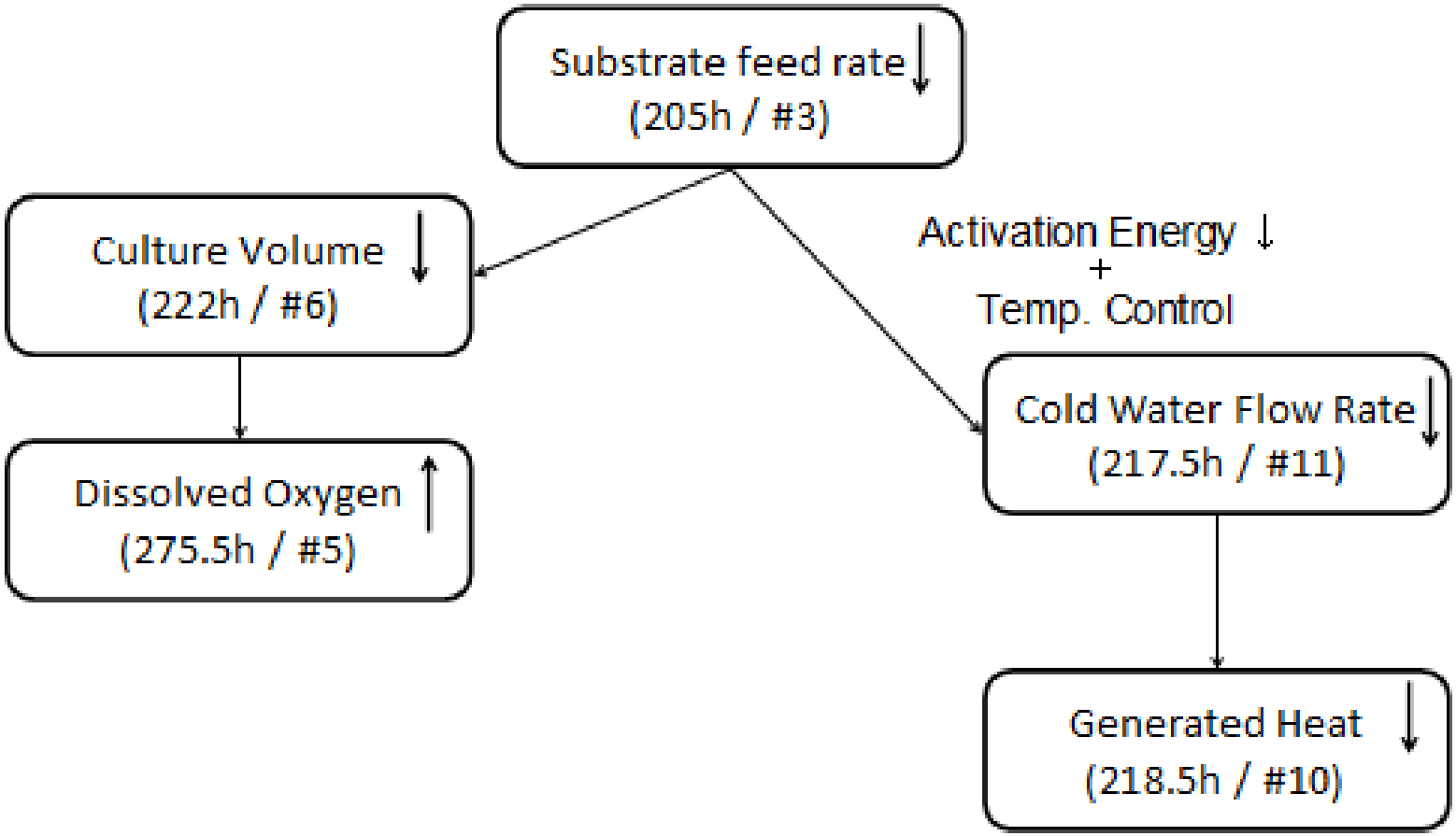
Identified fault propagation path (Fault 1)



Results summary for the 2nd faulty batch (Online)

Step No.	Identified variables	Time of detection (time series SPE plot)	Time of detection (raw data plot)
1	#3 (Substrate feed rate)	206h	205.5h
	#10 (Generated heat)	218.5h	221.5h
	#11 (Cold water flow rate)	217.5h	223h
2	#5 (Dissolved oxygen)	275.5h	Not Detected
	#6 (Culture volume)	222h	233.5h

Identified fault propagation path (Fault 2)



Monitoring of processes with multiple operation modes using principal angle and multiple PCA/PLS models

(S. J. Zhao, J. Zhang, and Y. M. Xu, “Monitoring of Processes with Multiple Operating Modes through Multiple PCA Models”, *Ind. Eng. Chem. Res.*, Vol.43, 2004, 7025-7035.

S. J. Zhao, J. Zhang, and Y. M. Xu, “Performance Monitoring of Processes with Multiple Operating Modes through Multiple PLS Models”, *Journal of Process Control*, 16(7), 2006, 763-772.)

Principal Angles

- The concept of principal angles is originally proposed to measure the distance or angle between two subspaces of higher dimensional linear vector spaces
- Computation of principal angles using singular value decomposition (SVD):

$$\text{For } \mathbf{F} \text{ and } \mathbf{G} \text{ in } \mathbb{R}^m,\ p = \dim(\mathbf{F}) \geq \dim(\mathbf{G}) = q \geq 1$$

$$\cos\theta_1 = \max_{u \in \mathbf{F}} \max_{v \in \mathbf{G}} \left(u^T v\right) = u_1^T v_1,\ \|u\|_2 = \|v\|_2 = 1$$

$$\cos\theta_k = \max_{u \in \mathbf{F}} \max_{v \in \mathbf{G}} \left(u^T v\right) = u_k^T v_k,\ \|u\|_2 = \|v\|_2 = 1$$

$$\text{s.t. } u_i^T u = 0, v_i^T v = 0 \quad i = 1, 2, ..., k-1 \text{ , for } k = 2, ..., q$$

Similarity between PCA models

- For a data matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{L}\ , \mathbf{x}_m] \in \mathbb{R}^{n \times m}$ where *n* denotes the number of samples and *m* the number of process variables or quality information, it can be decomposed using PCA as

$$\mathrm{X} = \mathrm{TP}^T + \mathrm{E}\ with\ \mathrm{P}^T\mathrm{P} = \mathrm{I}$$

- **P** represents an orthonormal basis of the retained principal component subspace

- The definition of principal angles can then be utilized to measure the similarity or distance between two PCA models

Model Construction and Comparison

1). For each group $X_k^{(i)}$, PCA model $P_k^{(i)}$ is *independently* obtained through the traditional PCA method, and the number of retained principal components should be determined through cross-validation with test data set as a part of $X_k^{(i)}$, $k = 1, \ldots, n_i$;

2). Calculate the similarity $S_{j,k}^{(i)}, j = 1, \ldots, n_i, k = 1, \ldots, j-1$ between PCA model $P_j^{(i)}$ and $P_k^{(i)}$

3). If any $S_{j,k}^{(i)} < \theta_t$ where θ_t is the pre-selected threshold of similarity, merge the corresponding data group $X_j^{(i)}$ and $X_k^{(i)}$. Denote these newly generated data groups as $X_1^{(i+1)}, \ldots, X_{n_{i+1}}^{(i+1)}$. Let $i = i+1$ and go to step 1). Otherwise, stop.

Process Monitoring and Fault Detection

- Widely used multivariate statistics such as Squared Prediction Error (SPE) and T^2 can be incorporated straightforwardly.
- Every newly collected observation is first scaled by the corresponding mean and standard deviation of each PCA model and then approximated by it.
- The one yielding the minimum SPE is finally accepted and utilized to determine whether a warning should be triggered.
- Control limits of the control charts generally vary according to the adopted PCA models.

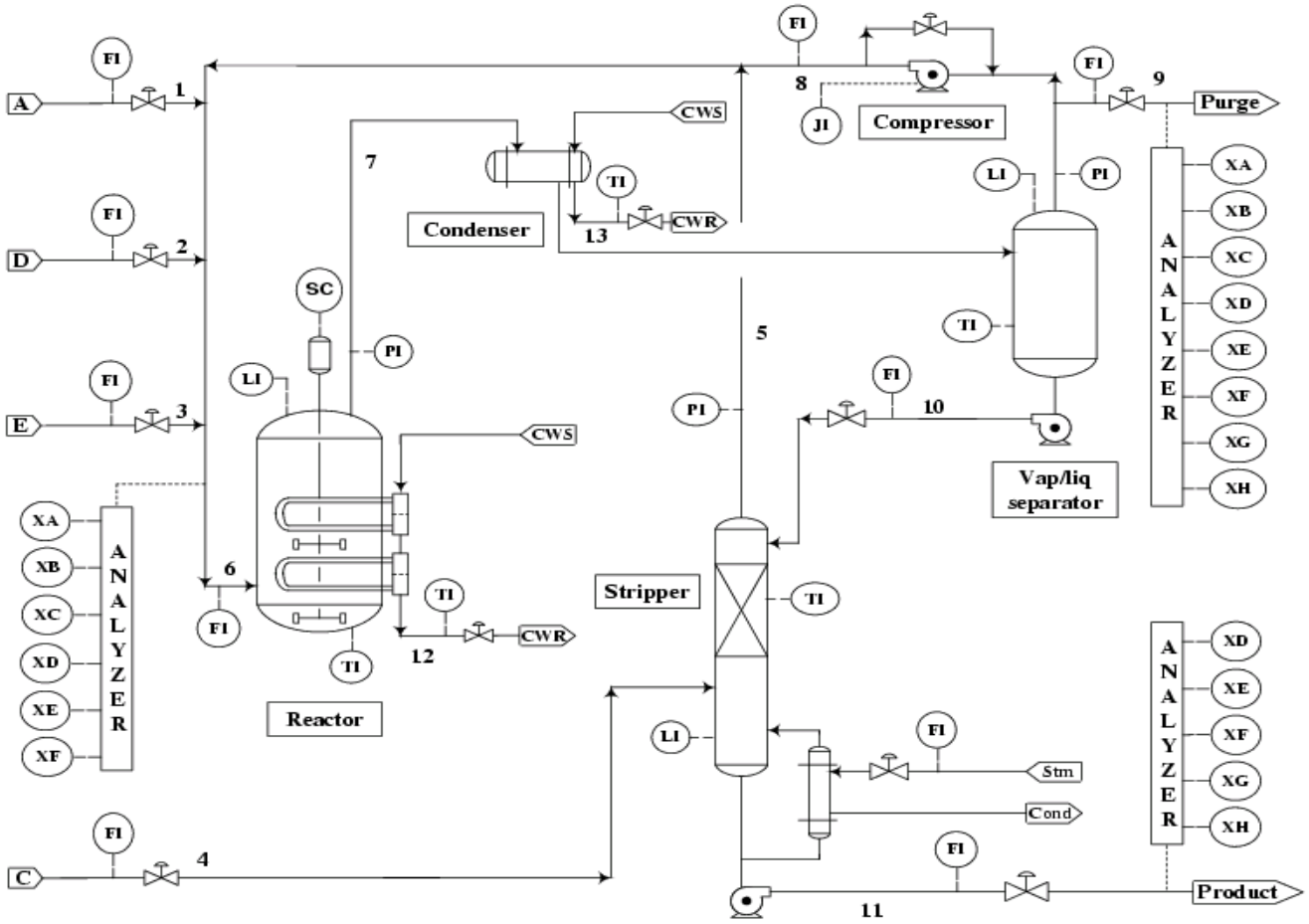
Model Updating

- Industrial processes often experience time-varying changes and are thus desirable to update the model recursively
- With multiple PCA models, the newly collected data should be processed sequentially as follows:
  - determine the model (*adopted model*) it belongs to
  - if this new data does not trigger a warning, it will then be incorporated into the *adopted model*, which can thus be updated recursively analogous to the single-model case
- Model updating should not violate the models similarities constraints

New Model Incorporation

- A new model will be incorporated provided that adequate samples have been collected whilst its distance to every existing one exceeds the predefined threshold.
- Can be performed at runtime if necessary and will not cause a break of the monitoring.
- The data for constructing the new model can be prepared automatically based on certain strategy. Currently, it is achieved manually at appropriate time by the operators.

Case Study 1: The Tennessee Eastman Challenge Process



Newcastle University

CPACT

Results

Six operating modes of TECP

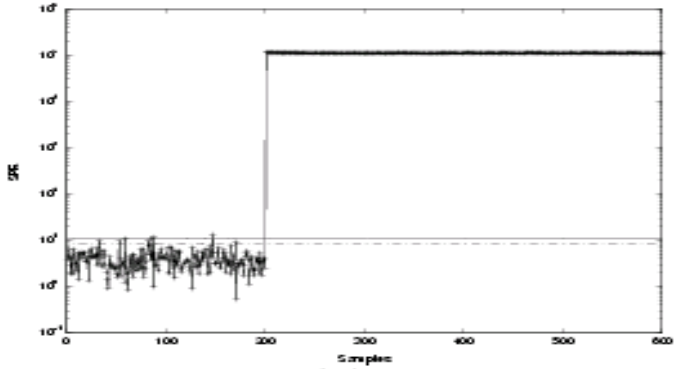
Mode	Desired G/H mass ratio	Desired production (kg/h)
1	50/50	14076
2	10/90	14077
3	90/10	11111
4	50/50	Maximum
5	10/90	Maximum
6	90/10	Maximum

Similarities between the five PCA models corresponding to modes 2 to 6

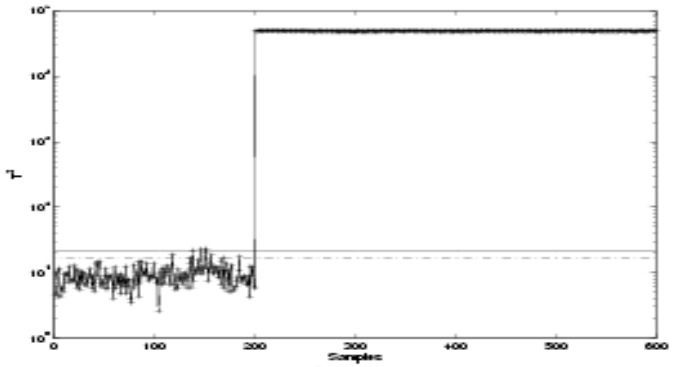
Model No.	mode 2	mode 3	mode 4	mode 5	mode 6
mode 2	1.0000	0.8514	0.9011	0.9346	0.8656
mode 3	0.8514	1.0000	0.8884	0.8467	0.8853
mode 4	0.9011	0.8884	1.0000	0.9117	0.8358
mode 5	0.9346	0.8467	0.9117	1.0000	0.8749
mode 6	0.8656	0.8853	0.8358	0.8749	1.0000

Similarities between the newly developed model and the existing ones

Model	mode 2	mode 3	mode 4	mode 5	mode 6	overall
new model	0.9380	0.8768	0.9033	0.9184	0.8698	4.8511

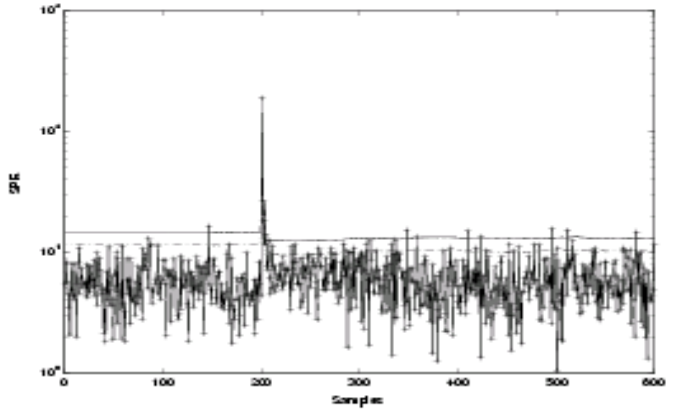


(a)

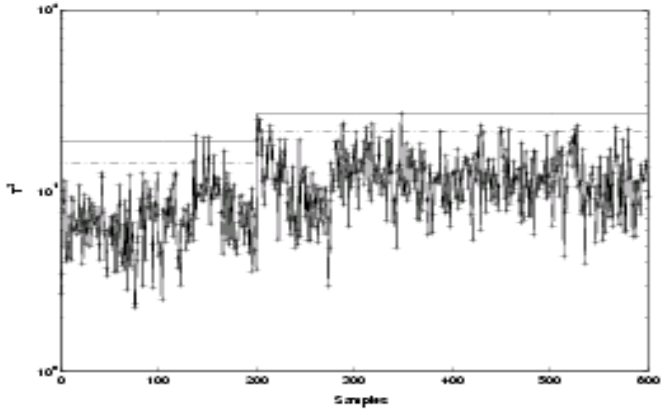


(b)

Fig. 3. The SPE (a) and T^2 (b) control charts by RPCA with 95% (dash-dot line) and 99% (solid line) control limits. The process switches from mode 3 to mode 1 at sampling point 200.



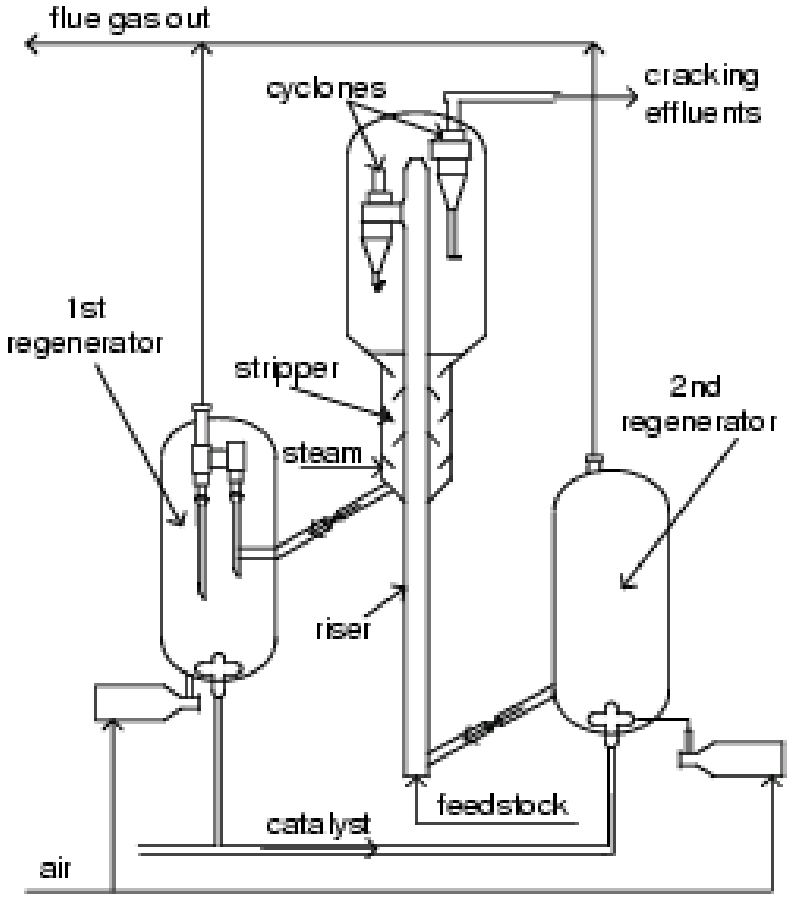
(a)



(b)

Fig. 4. The SPE (a) and T^2 (b) control charts with 95% (dash-dot line) and 99% (solid line) control limits with the new model incorporated. The process switches from mode 3 to mode 1 at sampling point 200.

Case Study 2: Industrial Fluidized Catalytic Cracking Unit



Results

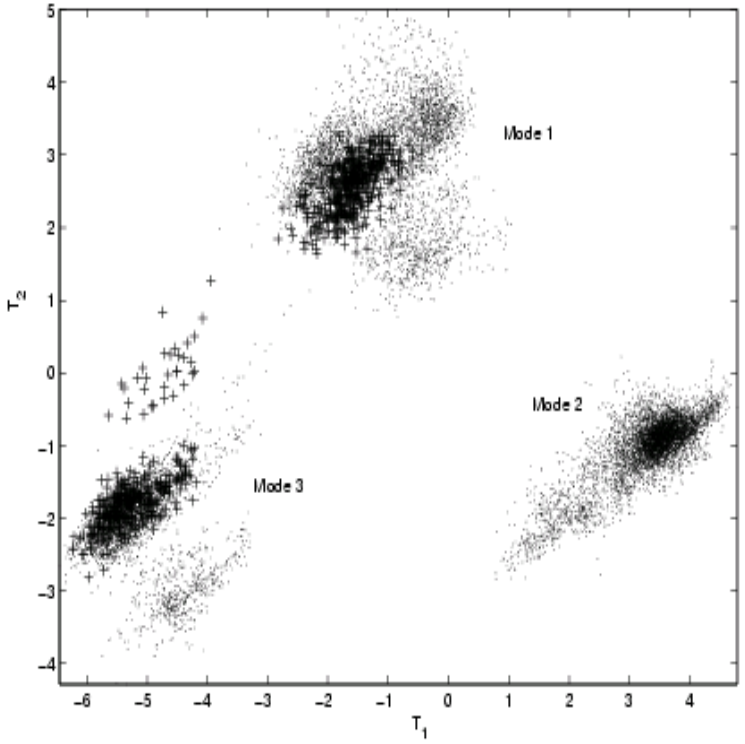


Fig. 11. $\mathbf{T}_1$ vs. $\mathbf{T}_2$ of the PCA model developed from the original data where + denotes the on-line collected data while the others the modelling data

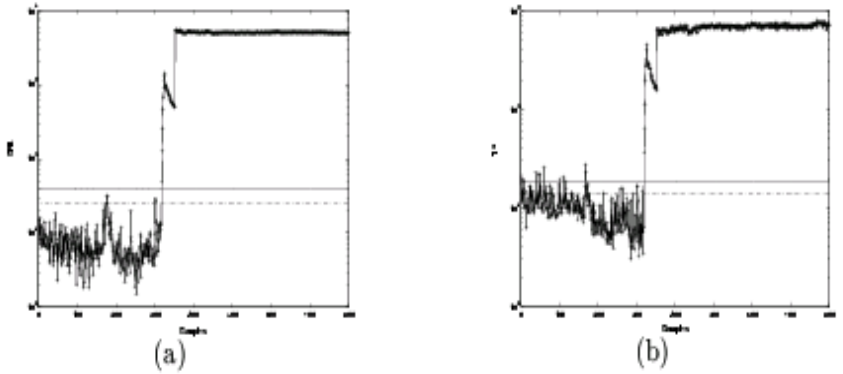


Fig. 12. Real-time SPE (a) and T^2 (b) monitoring charts by RPCA with 95% (dash-dot line) and 99% (solid line) control limits

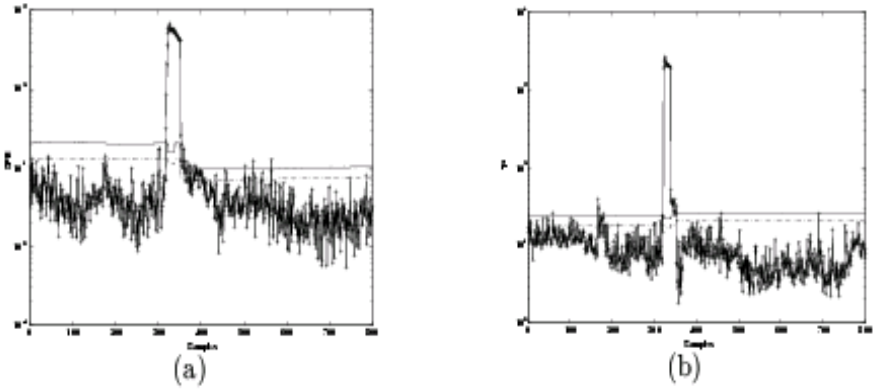


Fig. 13. Real-time SPE (a) and T^2 (b) monitoring charts by the proposed algorithm with 95% (dash-dot line) and 99% (solid line) control limits

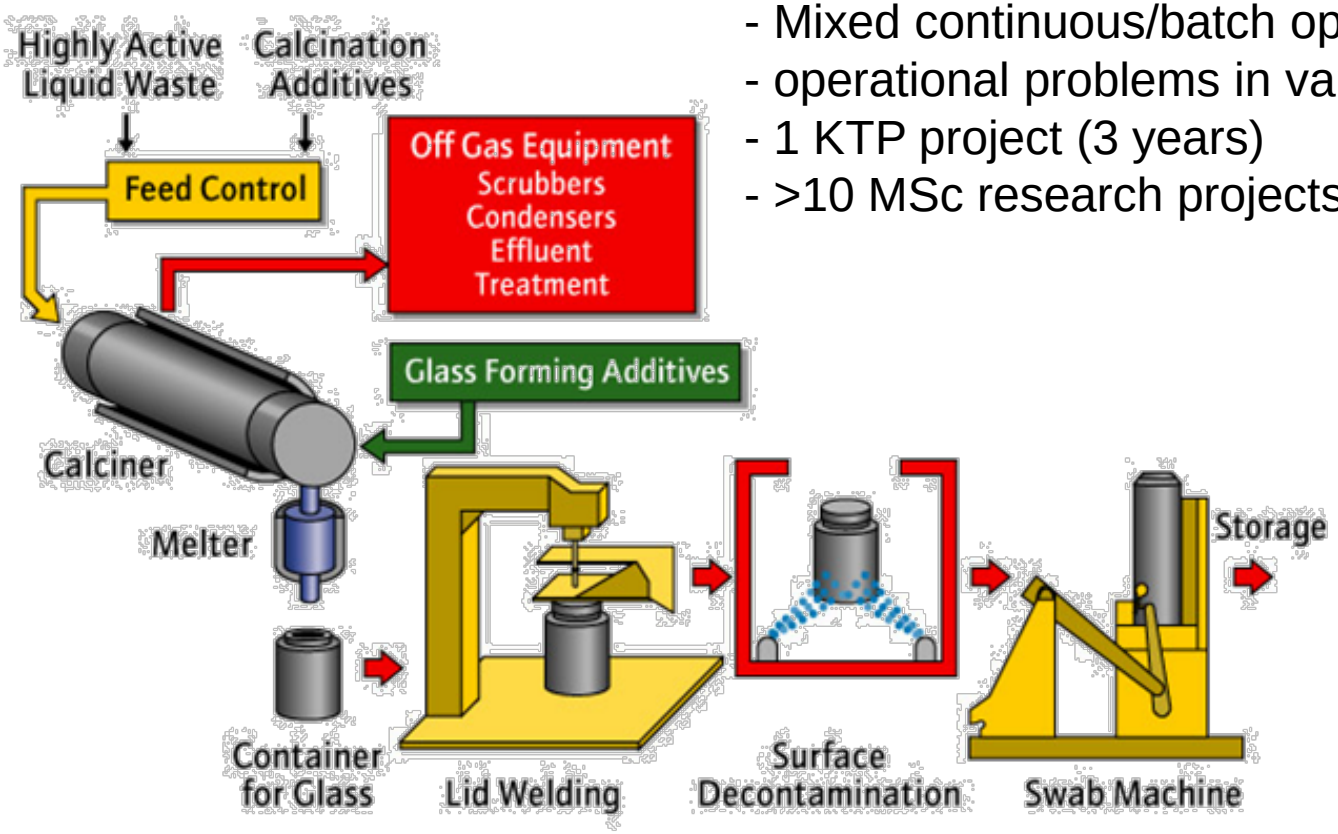
Newcastle University

CPACT

Some industrial applications of MSPC

Monitoring of nuclear waste re-processing plant

- Nuclear waste vitrification plant



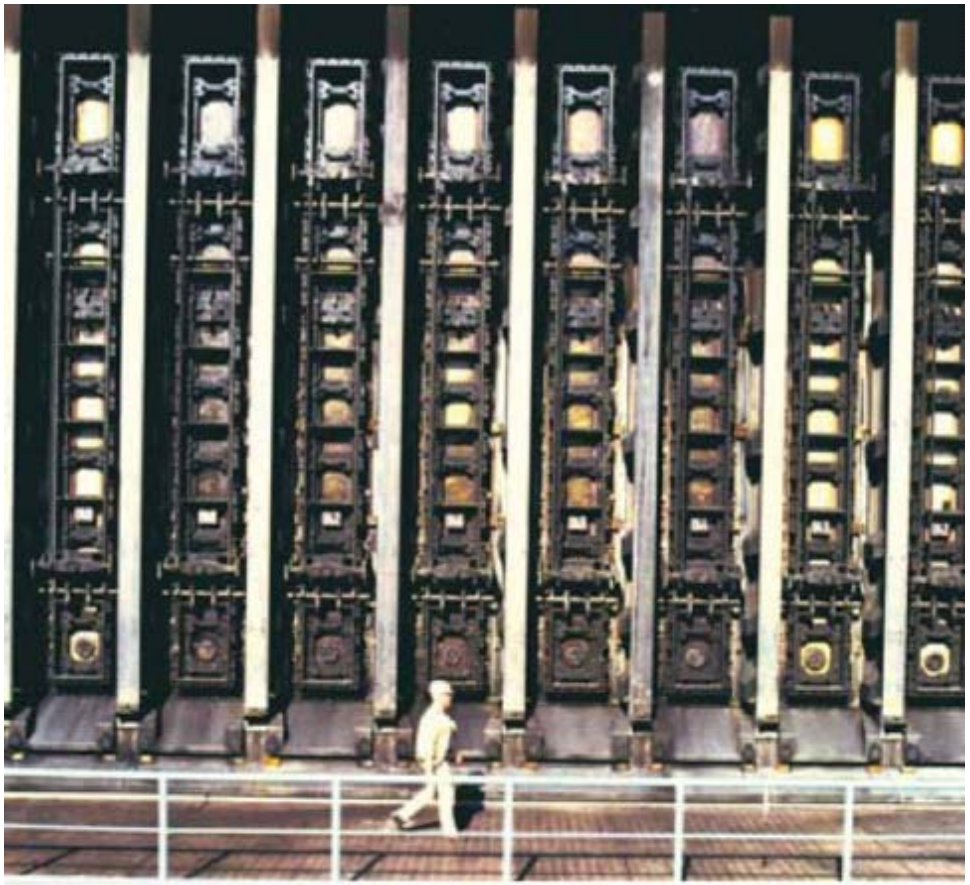
- Mixed continuous/batch operations
- operational problems in various stages
- 1 KTP project (3 years)
- >10 MSc research projects

Newcastle University

CPACT

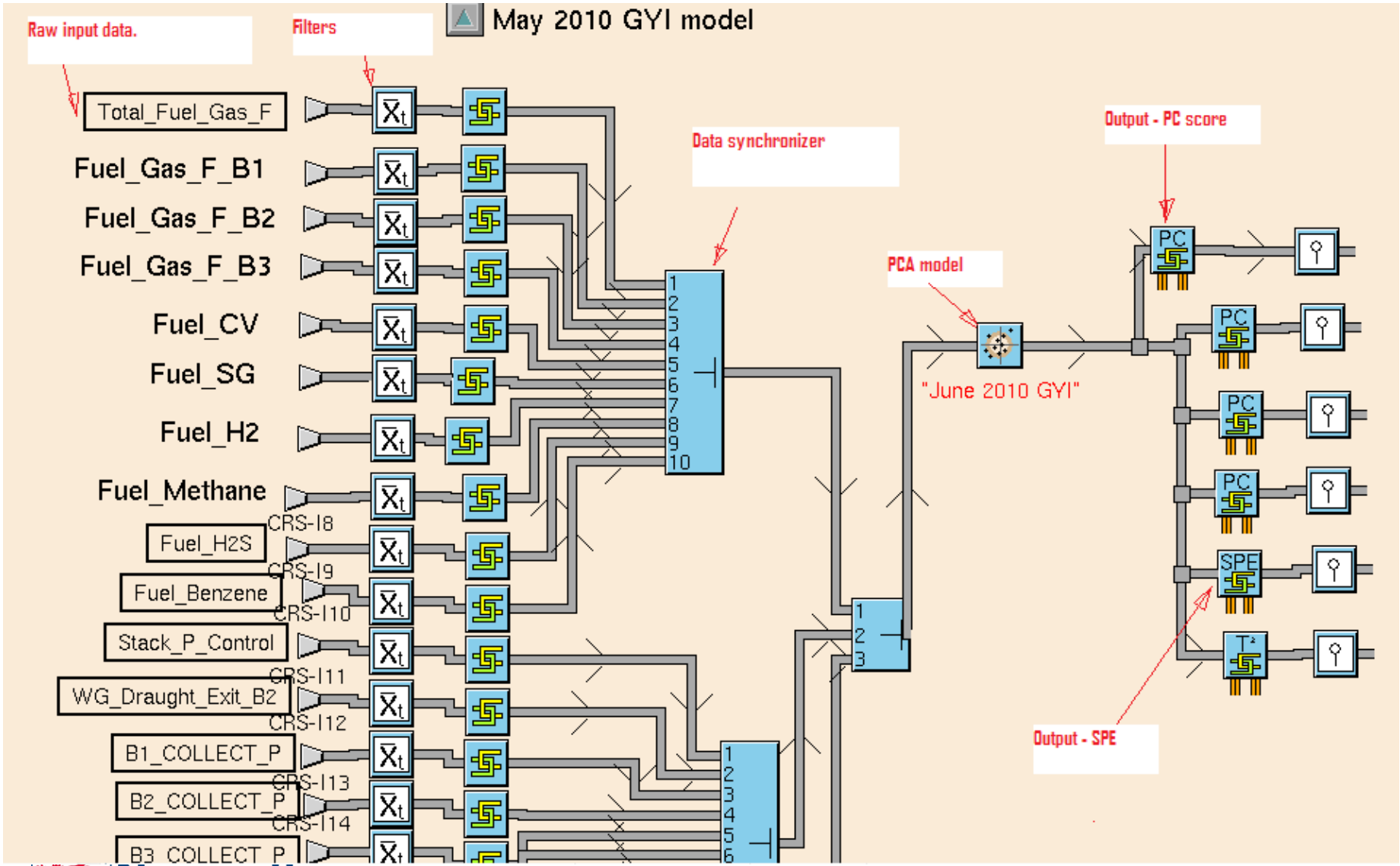
Monitoring of coke ovens at TATA Steel

- The coke-making process: carbonisation of coal to high temperatures in an oxygen deficient atmosphere - pyrolysis.
- Coal is charged into an oven, carbonised for 18 hours and then quenched prior to being used in a blast furnace to make iron for steel producing steel.
- The off-gas (raw coke oven gas) is sent to a by-products plant.
- Main operation problem: through wall leakage.
- EU ECOCARB project



75 coke oven chambers operating up to 1600^0C

A Gensym G2 *iMSPC* Model

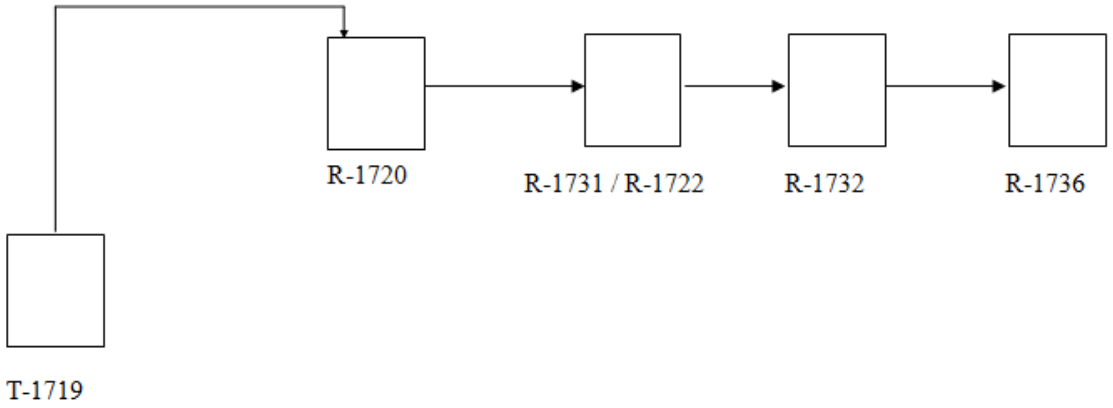


Newcastle University

CPACT

Monitoring a multi-stage batch process at Syngenta

- Multi-block progressive PCA is used to monitor the process and to establish fault propagation paths

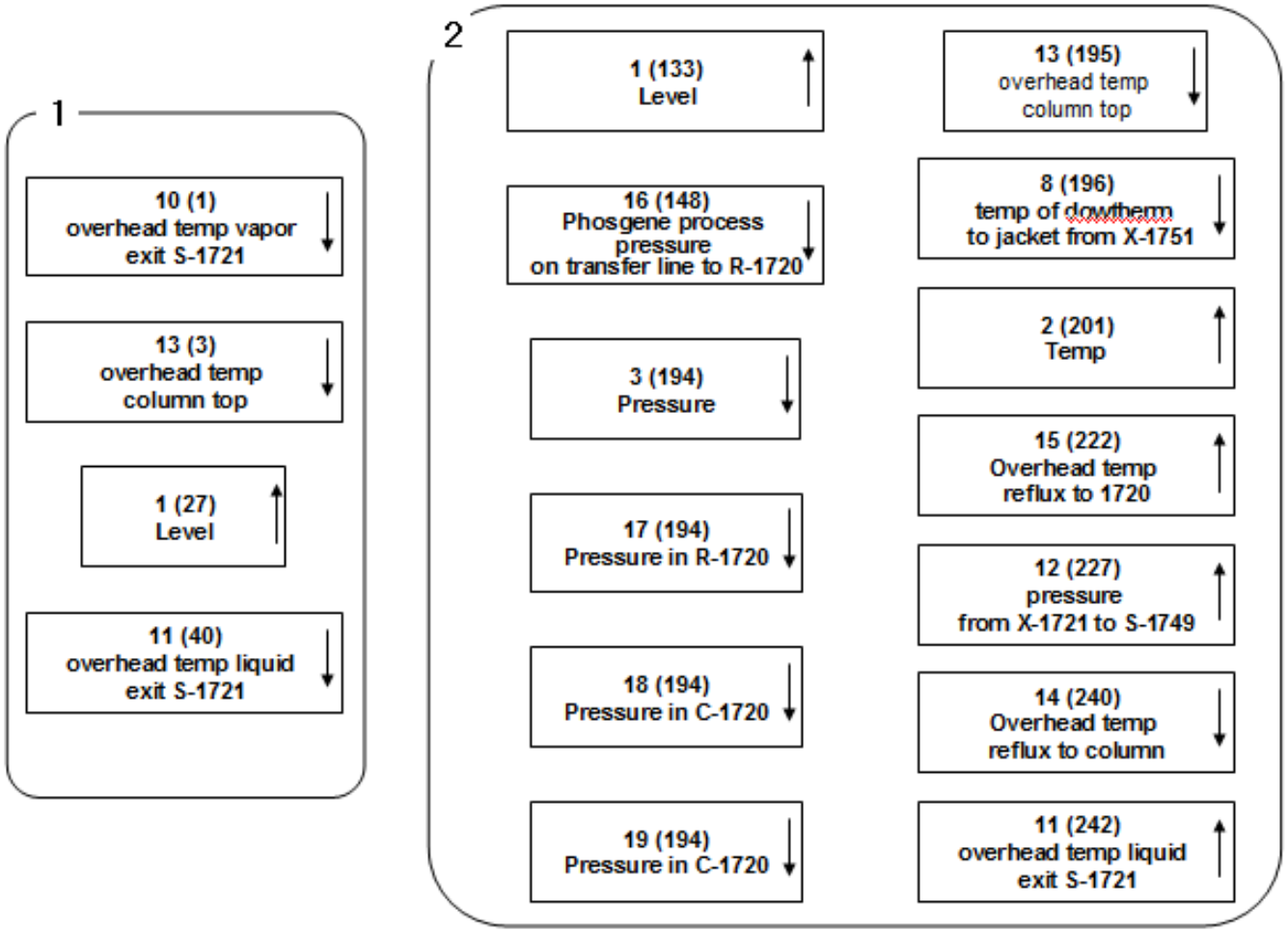


syngenta

Newcastle University

CPACT

Monitoring a multi-stage batch process at Syngenta



Newcastle University

CPACT

Conclusions

- MSPC is very effective in monitoring processes with large number of correlated measured process variables.
- The key in MSPC is through data dimension reduction and monitoring the correlation among process variables in the reduced dimensional space.
- Batch process monitoring can be handled through data un-folding and multi-way version of PCA and PLS.
- Various extensions of MSPC techniques have been proposed to cope with nonlinear, non-Gaussian, dynamic, multi-model, multi-phase processes.
- A hot research area judging from large number of publications, but with many scopes for further research.

Acknowledgements

- The EPSRC, DTI, OST and the EU for Research Funding.

- The collaborating companies for access to their plant data

- My former PhD/MPhil students and RAs: Dr Shijian Zhao, Dr Jeong Jin Hong, Dr Shallon Stubbs, Dr Yusri Yunus, Dr Bo He, Ms Katy Ferguson, Dr Gang Yi

Thank you for your attention.

jie.zhang@newcastle.ac.uk