PAT, QbD and Process Performance Monitoring
(The Impact for Process Systems Engineering)

Julian Morris¹, Zeng-ping Chen² and David Lovett³

Centre for Process Analytics and Control Technologies
¹Newcastle University, ²Strathclyde University, ³Perceptive Engineering

CPAC Rome Workshop March 2008
Overview of Presentation

- What is Process Systems Engineering?
- Where are we now, where are we going and how important is variability?
- Look at some novel methodologies for building fit-for-purpose calibration and statistical monitoring models
- Highlight some process analytical technologies through work aiming to improve understanding of monitoring and controlling batch crystallisation processes using advanced Chemometrics
- Show how some new closed loop process PAT control ideas are coming to fruition through innovative process systems engineering - Closing the Analytical Control Loop
- Closure

The EU provides 32% of the world's chemicals manufacturing through some 25,000 enterprises of which 98% are SMEs which account for 45% of the sectors 'added value', and 46% of all employees are in SMEs.
What is Process Systems Engineering?

- Process Systems Engineering (PSE) is a well established scientific area, concerned with the development of methods and computer-based tools for an integrated approach to all aspects of process modelling, simulation, design, operation, control, optimisation and management of complexity in uncertain systems across multiple time and scale lengths.

- These tools enable Process Systems Engineers to systematically develop products and processes across a wide range of systems involving chemical, biological and physical changes from molecular and genetic phenomena to manufacturing processes and related business processes.
Where are we Now and Where are we Going in Process Analytics and Control Technologies?
Capsules of Abbott Laboratories' protease inhibitor Norvir (ritonavir) are likely to become unavailable by the middle of August. The company has a problem with the manufacture of the anti-HIV capsules which it cannot resolve at present.

The problem relates to "undesirable" crystal formation. Abbott says that a series of recent production batches of Norvir capsules failed the approved test for dissolution, and were not released for marketing. Investigation of the reason for the failure showed the presence of a new crystalline form of ritonavir which affects the way it dissolves, and possibly its absorption. Retained samples from a number of marketed batches of capsules were examined and there was no evidence of the unwanted crystalline form.

Mr Mark Haywood (managing director, Abbott Laboratories) said that teams were working round the clock to try to resolve the issue, but at present the company had no idea why the problem was occurring.
Impurities and Polymorphism (2)

- Impurities effect nucleation and growth processes, and hence can stabilise meta-stable polymorphic forms.

- As product purity improves during process chemistry work-up, the “stable” polymorphic form can change!

  e.g. RITONOVIR aids drug which changed from anhydrous to hydrate crystal after launch:

  - with lower solubility and hence bio-availability.
  - product was withdrawn for a year and reformulated.
  - new FDA approval needed – mega cost implication!
During the development and testing of a new drug an issue arose with the ‘smell’ of an API product.

- Although the material was very pure (>99%) there was a residual ‘smell’ of the crystallisation solvent which persisted even after drying.
- This made the ‘placebo’ tablets difficult to disguise (blind).

There was also a problem with the final blending, drying and compressing operations where the formulation team wanted larger crystals which had better flow and compression characteristics for tablet making.
API Crystal Size Distribution

- Origin of the ‘Smell’:
  - Crystal size

Courtesy of Gerry Steele, AstraZeneca. APACT07
Rather than carry out a re-crystallisation, temperature cycling the material causing the difference in the dissolution rates can be used to increase crystal size.

![Temperature Cycling Diagram]

Courtesy of Gerry Steele, AstraZeneca. APACT07
Where Are We Going?

From this ... to this .... to this

Pharmaceutical companies typically operate at around 2 to 2½ Sigma compared with 6 Sigma in world class manufacturing
Variability (or PAT) by Edwards Deming

Cease reliance on mass inspection to achieve quality.

Eliminate the need for mass inspection by building quality into the product in the first place.

Dr W. Edwards Deming
(Circa 1980s)

“Learning is not compulsory, …

…. Neither is survival”
Variability and Coping with Different Product Recipes, Different Processing Units, Different Production Sites, Different Spectroscopic Probes & Probe Locations, ……, etc
Variability with Different Product Recipes, Processing Units, Production Sites, Spectroscopic Probe Locations, …

- The elimination of differences between different modelling or calibration applications is a major issue in flexible process manufacturing
  - Different spectroscopic probes in different vessels or on different sites
  - Different reactors (unit operations), multiple recipes, etc
- Normally this would require the construction of separate models for each individual situation.
- A simple but partial solution is the subtraction of local or global mean levels
- A better solution is to develop a multi-group (sub-space) model or calibration such that the interest can then focus on within process (or product) variability.

Case Study: Between and Within Group Variation

- Multi product manufacturing with Unilever (Lever Fabergé).
- Approximately 50 different products and five main production units, complicated by multiple recipes.
- Monitoring the process using standard MSPC would mean the building and maintenance of approximately 250 statistical monitoring models.
- Too complicated for ‘quality’ and operational personnel.
- High model-maintenance costs
To demonstrate the application of multiple group PLS the production of two recipes are considered. Each process mixer is also considered as a separate ‘Recipe’ and thus the process model contains four distinct groups.

<table>
<thead>
<tr>
<th>Recipe</th>
<th>Group</th>
<th>Mixer</th>
<th>Number of Batches</th>
<th>Raw Materials</th>
<th>Quality Variables</th>
<th>Process Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 (♀)</td>
<td>1</td>
<td>19</td>
<td>23</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td>2</td>
<td>2 (+)</td>
<td>2</td>
<td>21</td>
<td>23</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td>2</td>
<td>3 (♂)</td>
<td>1</td>
<td>20</td>
<td>17</td>
<td>1</td>
<td>90</td>
</tr>
<tr>
<td>4 (♀)</td>
<td>2</td>
<td>29</td>
<td>17</td>
<td>1</td>
<td>90</td>
<td></td>
</tr>
</tbody>
</table>

Note that Recipe 2 has fewer raw materials and fewer process variables than Recipe 1.
This model contains both within group and between group variation.

Subtle process events cannot be detected; the greater the number of distinct groups in the data set, the greater the impact of between group variation.
Impact of Spectroscopic Probe Location in a Reaction Vessel

Spectroscopic Probe Location

PCA Plot of Measured Spectra
The pooled variance-covariance matrix is constructed as a weighted sum of the individual elements of the individual variance-covariance matrices.
The monitoring chart shows the scores jumping outside the action and warning limits as the process malfunction impacts.
Variables (62 - 65) were identified as being related to raw material addition.

A faulty dosing-control valve was located as being the source of the problem.
Variability – Modelling and Calibration Challenges

- **Process Issues:**
  - Equipment characteristics; site-to-site process differences, etc
  - Fluctuations in both control and external process variables
  - Cell improvement; cell line changes; media changes
  - Small data sets are an issue but can be enhanced through Bootstrap Aggregation

- **Analytical Issues:**
  - Separating absorbance from multiplicative light scattering effects caused by the variations in optical path length
  - Inter probe variability: impact of component variance on PLS calibration – can probe differences be accommodated or eliminated?
  - Can calibration models be made generic for different production unit operations / production lines?
Variability - Coping with Minimal Experimental Data

- The need for fit-for-purpose statistical models can be limited due to intrinsic data sparsity resulting from the small number of objects available, e.g. experiments, batches, runs, etc, compared with the large number of process variables wavelengths measured.

- This has been successfully addressed in PLS calibrations and in building statistical monitoring and predictive models through the addition of Gaussian noise to the original data and combining multiple models using Bootstrap Aggregated Regression.

- This allows the development of robust calibration or process models and has been shown to lead to a decrease in prediction errors as a consequence of the increase in the data density.

Advanced Chemometric Methods
Chemicals Behaving Badly (CBBII):
Integrated In-Process Analytics and Closed Loop Control

CBB#2 Project: collaboration with Leeds, Heriot-Watt and Newcastle University

Partners
- AEA Technology
- AstraZeneca
- Bede Scientific Instruments
- BNFL
- Clairet Scientific
- DTI
- EPSRC
- GlaxoSmithKline
- HEL
- Malvern Instruments
- Pfizer
- Syngenta

Chemometrics

Supersaturation

Growth kinetics

Size

Nucleation kinetics

MSZW

UVvis

Video microscopy

Shape

Reactant rheology

Polymorphic form

Process conditions

Mixing & scale-up

Heat transfer

Reaction Calorimetry

Batch process monitoring & control

FTIR

LDA/PIV

CFD

Heat transfer

Chemicals Behaving Badly (CBBII): Integrated In-Process Analytics and Closed Loop Control

CBB#2 Project: collaboration with Leeds, Heriot-Watt and Newcastle University

Partners
- AEA Technology
- AstraZeneca
- Bede Scientific Instruments
- BNFL
- Clairet Scientific
- DTI
- EPSRC
- GlaxoSmithKline
- HEL
- Malvern Instruments
- Pfizer
- Syngenta
Smoothed Principal Component Analysis (SPCA)

- Enhancing signal-to-noise ratio

\[
X^T X r_i = \lambda_i \times (I + k \times Q^T Q) r_i \quad i = 1, 2, \ldots, m
\]

\[
F = [r_1, r_2, \cdots, r_c] \times [r_1, r_2, \cdots, r_c]^+ \quad x_F = Fx = Fx_s + Fx_n
\]

Raw (a) and Processed (b) XRD profiles (by SPCA) for 6 XRD data sets of mannitol-methanol suspensions with the contents of mannitol equalling to 0.0%, 0.178%, 0.389%, 0.533%, 0.8% and 1.0% g/ml, respectively.
The raw and pre-processed XRD profiles of the β-form with concentration varying from 0.02% to 8.00%; relationships between concentrations and peak heights at peaks B3 of the raw and processed spectra.
System provides capability to monitor polymorphic form “in-process” unaffected by product separation prior to analysis.

Typically circa 1 wt % detectable via in-process XRD, much lower with advanced chemometric analysis (Smoothed PCA)
Loading Space Standardisation (LSS)

- Correcting non-linear shift and broadening in spectral bands caused by temperature fluctuations

\[ x_i(t_i) = \sum_{k=1}^{K} c_{i,k}s_k(t_i) + e_i(t_i) \]

\[
\begin{align*}
X(t_1) & \xrightarrow{PCA} TP'(t_1) \\
\vdots & \quad \vdots \\
X(t_m) & \xrightarrow{PCA} TP'(t_m)
\end{align*}
\]

\[
\begin{align*}
P(t_i) & \sim t_i \Rightarrow LSS \\
x(t_{\text{test}}) & \xrightarrow{LSS} x(t_{\text{ref}})
\end{align*}
\]
Optical Path-Length Estimation and Correction (OPLEC)

- Separating absorbance from multiplicative light scattering effects caused by the variations in optical path length

\[ x_i = b_i \sum_{j=1}^{J} c_{i,j}s_j + \varepsilon_i \]

- \( b_i \) - multiplicative parameter

\[ b_i = OPLEC(X, c_j), \ X = [x_1, x_2, \ldots, x_m], \ c_j = [c_{1j}, c_{2j}, \ldots, c_{mj}] \]

- OPLEC has recently been applied for Raman Scattering applications in complex crystallisation processes.
250L Pilot Plant Batch Agitated Vessel

- Temp. & pH interface Box
- Enablr Box & PC
- Turbidity Box
- Agitator speed meter
- Outlet (8mm) port
- Inlet (12mm) port
- α-sizer
- Bath for cooling α-sizer’s transducer
- Mono pump for α-sizer
- Manhole
- 250 L vessel
- α-sizer
Supersaturation Control System Upgrade to PI Capability

Crystallisation Vessel

FTIR Spectrometer

WINISO SOFTWARE

FTIR PC

Spectra

Cal. model

C(t)

ENABLIR SOFTWARE

4 – 20 mA

FTIR Spectrometer

Crystallisation Vessel

FTIR PC

Spectra

Cal. model

C(t)

WINISO SOFTWARE

HEL PC

Data Processing Block

C(t)

Macro

Solubility Model

S(t)

Control Block

IMC Based PI Controller

User Define Model

User Define “S” set point
Supersaturation Control of L-Glutamic Acid
250 litre Plant Crystalliser

Temperature, Concentration, Solubility & Turbidity

Supersaturation Control

5% seeds added

$S = 1.1$

$S_{\text{max}} = 1.125$

$S_{\text{min}} = 1.075$

Started Supersaturation Control

$S_{\text{lim}}$
L-Glutamic Acid Crystals (a) Seeds (b) Product

α

β

100 μm
Multivariate Statistical Process Control (MSPC)
or
Process Performance Monitoring
Is the Data Unfolded and Scaled Properly?
Batch Data Unfolding (Nomikos & MacGregor) - MWPCA_{N&M}
Batch Data Unfolding (Wold et al) - PCA_w
Nomikos and MacGregor (N&M). Batch-wise unfolding.

Wold, Kettaneh, Friden and Holmberg. Variable-wise unfolding
  - Unfolding in this direction followed by mean centering and scaling does not remove the mean trajectories from the data and hence only captures the covariance among the mean trajectories of the variables, which is not of interest in process performance monitoring.
  - Rather, it is the covariance structure of the deviations of the variables about these mean trajectories that is of interest.
  - This unfolding approach also only provides a ‘static’ model that captures only the local covariance structure amongst the variables and does not account for the dynamic behaviour of batch processes.
Non-linearity and Process Dynamics (MWPCA<sub>NM</sub>)

- Scaling can remove major non-linear component in the data.

- Scaling may remove major dynamic component in the data.
Non-linearity and Process Dynamics (PCA_W)

- Scaling cannot remove major non-linear component in the data.

- Scaling cannot remove major dynamic component in the data.
Model Based Closed Loop Process Control
In the presence of auto-correlation, the action and warning limits calculated for a process performance monitoring scheme based on the usual assumptions that the observations are Independent and Identically Distributed (IID), are inappropriate making the monitoring charts unreliable, if not useless.

Developing an MSPC representation for the monitoring of processes, exhibiting time varying characteristics, will result in an increase in the number of false alarms and operational changes not being detected.

So, how do we deal with auto-correlated data?
At the heart of a model-based control or monitoring system is a DYNAMIC model.
Impact of Non-linearity – Normal Probability Plots
Impact of Non-linearity – Normal Probability Plots

- Process Data
- Residuals
- Plant-Model Mismatch
- Model Residuals
- Dynamic PLS
- Model Residuals
- Time Series Model
Closing the Analytical Control Loop
Challenges for Real Time Closed Loop Control and Real Time Release

- At least 15 years worth of academic control theory, simulations and occasionally pilot studies have reported on the benefits and challenges of improving operational performance through improved control, more recently in the Pharmaceutical Industry.

- Since then the FDA has opened the way to a technology-led quality assurance regime there has been hectic activity directed towards defining an implementation methodology for measurement, modelling and data analysis.

- Many of the larger automation companies have focussed on the overall infrastructure to ensure that the integration of data from many sources is recorded securely and a traceable batch record is generated.

- Hopefully one of the key outcomes is that the integration of the data will also integrate the users of the data such as the chemists, biologist, engineers, to provide a process systems engineering approach to design, build, operate, monitor and improve production facilities.
Challenges for Real Time Closed Loop Control

History:
- > 15 years of academic control theory, simulations & pilot studies in Pharma
- 2004 - FDA opened the way to a technology-led quality assurance regime
- Hectic activity to define methodology for measurement, Design of Experiments, modelling, data analysis, etc.
- Large automation companies focus on overall infrastructure for data management.
- Aim to ensure integration of data from many sources is recorded securely and a traceable batch record is generated.

Recent:
- Aim to also integrate the users of data (chemists, engineers, Lab analysts, process development scientists)
- Provide a process systems engineering approach to design, build, operate, monitor and improve production facilities – aligning with Lean and Six sigma initiatives
- Understanding the challenges with data compatibility, data fusion and data validity.
- Exploration of the benefits and challenges of improving operation through improved control.
- Beginning to investigate the use of Advanced Control techniques on unit operations and production lines.
Challenges of Real Time Closed Loop Pharma Control

- Why is Advanced Process Control required?
- Why and How to verify the integrity of data used in PAT applications
- Using Continuous, Discrete and Spectral data in a co-ordinated way
- Batch Control – Advisory or Closed Loop?
Why is Advanced Control Required?
PAT and Advanced Process Control

Why we’re here

- PAT can be used as part of a tool box to optimise the way pharmaceuticals are manufactured
- Provide greater understanding of what to control
- Potentially provide a means to control “Critical Attributes” by monitoring and adjusting “Critical Parameters” in real time
- Provides some of the ability, to reduce the risk of process variability, effecting process capability and product quality

Current Model

Variable Raw Material
Input to process

Fixed Process

Variable Output

Variable Process Model

Variable Raw Material
Input to process

Variable Process

Consistent Output

Continuous Quality Verification
- Understand process constraints and complex process interactions.
- Build multivariate correlation model between variables and actuators, causes and effects.
- Predict impact of known disturbances on operations.
- Predict, Advise, Make co-ordinated moves on multiple actuators.
- Exploit opportunities to push quality / throughput close to constraint limits.
How Do You Verify the Integrity of Data for PAT Real Time Control & Real Time Release
Univariate Data Quality Monitor:

Individual Signal Validation
- Logical Checks
- Statistical checks

Multivariate: Using Robust PCA
- Outlier detection
- Outlier Identification
- Data Quality Monitoring

Data Quality record underpins the validity of the system – critical in a validated environment

21 CFR part 11 Records
All signals used to infer (model) the Critical To Quality Parameters are included and monitored individually within a Data Quality Monitor whose output is automatically logged.
Real Time Quality Control
(Using Spectral Data)

Imported from
Model Development File

Clean Data Set

Spectral Data

Data Quality Monitor

Process Data

Discrete Data

Real Time Pre-Processing

21 CFR part 11 Records

Imported from Model Development File
Real Time Quality Control (Integrated Data Management)

CtQ Parameters

- Continuously measured OR
- A-periodically measured OR
- Real time value inferred from calibration model OR
- End-point value inferred from calibration model OR
- Scores of calibration model are CTQ parameters

Real time pre-processed data

Spectral Data

Process Data

Discrete Data

PLS/PCA Calibration Model

Dynamic PCA Controller

Control Space

Design Space
Key Challenges for Pharma Batch Control
Key Challenges for Pharma Batch Control

Availability of Critical to Quality (CtQ) Measurements
- Are CtQ parameters continuously measured?
- Are CtQ parameters measured by lab assay at end or during batch?
- Can CtQ parameters be inferred from PAT sensors with calibration model?

Process Non-linearity
- As a function of process operating point (e.g. Arrhenius reaction rate, etc)
- Time varying dynamics as batch evolves
- Changes in sign of process gain can be especially challenging

Trajectory tracking
- Many control systems excellent at regulating to fixed set-point
- Many control systems excellent at rejecting disturbances
- Few are good at tracking a time varying set-point profile
Typically the CtQ parameter isn’t directly measured

- Controlled indirectly by making process variable track a pre-defined trajectory

**Successful if:**

- process variable trajectory guarantees CTQ parameter will hit target end-point within acceptable quality limits
- AND variability in initial conditions/raw material is low or does not impact end-point
- AND process dynamics are ‘sufficiently linear’
CtQ parameter is controlled indirectly by controlling a process variable

- Dynamic process model captures dynamic multivariable process interactions.
- Model predicts trajectory of controlled variable(s) over future horizon.
- Controller computes trajectory of MV moves over future horizon to minimise current and future Set-Point error.
- Only the first move is applied, process repeats at next control interval
- This is a ‘moving window’ approach.
- End-point value of CtQ parameter continuously estimated throughout the batch
- MV trajectories modified by controller to make estimated end-point hit target
- Unfolded PLS model - predicts end-point value based on MV and process variable trajectories over the entire batch
- At each point in the batch, computes trajectory of MV moves over entire batch to minimise the error between the predicted end-point value and end-point target
Delivering Real Time Release

Parametric or Real Time Release is a consequence of the process being:

- Well understood,
- Engineered out and residually measured,
- Characterised and controlled,
- VARIABILITY understood and monitored and controlled.

Given that an appropriate measurement platform supports understanding and control, achieving this demands:

- Robust data quality and control of whole system suitability – not just analytical but conventional measured systems must be quality monitored to ensure that behaviour outside the bounds of understanding is identified.
- Appropriate understanding of system sensitivity builds confidence and early warning systems must therefore themselves be sensitive.
- Periodic calibration an observation in operation alone are unlikely to pick up small errors introduced by sensor performance in process.
- Control to multivariate trajectories and end points demands a level of precision and understanding through the system that supports confident statements about quality and uncertainty – nothing else will be acceptable to the FDA.
Practicalities of Real Time Release

- Contrast testing between unit operations and final Quality Assay:
  - Where is the bulk of the testing for a complex product which undergoes (say) 10 processing operations?
  - What is the overall lead time associated with stage-to-stage and final product testing?
  - Is the quality / cost / productivity burden most affected by the ability to reduce variability and get predictable stage to stage testing or by targeting final product tests?
  - With more potent compounds in play, the ability to obviate some stage-to-stage testing can have far reaching benefits beyond simple fixed cost reduction including reduced occupational exposure.

- Final product assay testing is going to be with us for some time yet …..

- Multi stage modelling is the challenge to be met in deriving analogues for final product quality test parameters, these are not yet well understood.

- Data collected must be robust and with only one chance to produce quality at minimum cost, a focus on characterising product and controlling variability in unit operations is logical and builds a platform for Real Time Release.
PAT Sensors in Closed Loop Process Control – Some Challenges

- Real-Time Management of Process Data - maintaining a traceable record of Data Quality.

- Real time pre-processing – needs to be consistent with and traceable to the unique calibration model.

- No control system is going to control a spectrum of several hundred simultaneous values; so what is important?
  - Is there a fit-for-purpose calibration model to infer specific product properties?
  - Are there particular features / segments of the spectrum of interest?
  - Should the scores of the PCA/PLS calibration model be controlled directly using Latent Variable controllers?
In this talk, I have …

- Reviewed variability and coping with different product recipes / formulations, different processing units, different production sites, different spectroscopic probes & probe locations.

- Reviewed some novel methodologies for building robust calibration and statistical monitoring models.

- Highlighted some process analytical technologies through work aiming to improve understanding of batch crystallisation processes and its scale-up, and the provision of enhanced process understanding through advanced chemometrics.

- And shown how some new closed loop process PAT control ideas are coming to fruition through innovative process systems engineering.
High Throughput Experimentation and Process Intensification in Product and Process Development
Many thanks to CPAC for their kind invitation and of course you for your kind attention. I will be happy to attempt to answer your questions.