Innovative Approaches to Pharmaceutical Development and Manufacturing Seminar Series

The Increasing Need of Multivariate Analysis for Process Understanding and Control (Session 2).

Brad Swarbrick, Product Manager

CAMO Software AS
CAMO Products

The Unscrambler®
A complete Multivariate Analysis and Experimental Design software.

Unscrambler Optimizer
A powerful product formulation and process optimization tool.

Online applications
- Unscrambler On-line
- Unscrambler Classifier
- Unscrambler Predictor

A plug 'n' play product designed to make effective online predictions and classifications, to monitor processes and ensure quality control with spectroscopic measurements.

Unscrambler Connect
A data management strategy

Quali-Sense
the best companion for panel leader detects the personal strengths and weaknesses of each assessor in your panel

Professional Services
Training, Consultancy and Solutions

www.camo.com
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data

2. Principal Component Analysis: The Workhorse of MVA

3. Smart Manufacturing and MAFSS

4. Fundamentals of Multivariate Statistical Process Control

5. Combining Spectral and Process Data

6. Benefits of PAT and Examples
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data
2. Principal Component Analysis: The Workhorse of MVA
3. Smart Manufacturing and MAFSS
4. Fundamentals of Multivariate Statistical Process Control
5. Combining Spectral and Process Data
6. Benefits of PAT and Examples
Historical vs. Designed Data

- **Designed Experiments**: Most Desired and Most Information
  - Maybe difficult to convince Production Manager. Results may not contain much new information

- **Collect From a Process**: May never get an abnormal result from a tight process
  - Sometimes, this is the only data available, but little is known about the state of the process at the time of data collection

- **Historical**: "Excite" the process or use lab scale
Historical vs. Designed Data

• When Historical Data is all that exists, this is not a problem for multivariate methods.

• The aim is to try and establish a Quality Function of the form

\[ \text{Quality} = f(\text{input variables}) \]

• This is the purpose of Designed Experiments, i.e. Screen out what is unimportant and only measure those parameters Critical to Quality (CTQ).
Historical vs. Designed Data

If it’s not important → Don’t Measure It.
The Next Step

• The next question is “how do I know what is important and what is not?”

• The key to finding the answer is........

Multivariate Analysis

• An extremely powerful tool for analysing multivariate data is described in the next section.
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data
2. Principal Component Analysis: The Workhorse of MVA
3. Smart Manufacturing and MAFSS
4. Fundamentals of Multivariate Statistical Process Control
5. Combining Spectral and Process Data
6. Benefits of PAT and Examples
Principal Component Analysis

- Principal Component Analysis (PCA) is considered the “Workhorse” of Multivariate Data Analysis.
- It provides the analyst with a picture of the variability inherent in a process/system.
- It is also an excellent method for detecting outliers and important variables.
Principal Component Analysis

Data + PCA = Information + Noise
Principal Component Analysis

Three important results can be drawn from PCA

1. The variables measured on a system are critical to quality and important variable relationships exist, which should be monitored.
   **Information Rich**

2. The data contain information, but either not enough, or there is too much inherent variability that must be addressed.
   **Balance between Information and Noise**

3. The data contain no information related to quality.
   **Noise Rich**
Principal Component Analysis

• PCA provides (among other things) two very useful plots:
  – The Scores Plot: Which provides a map of the samples
  – The Loadings Plot: Which provides a map of the variables.

• Both plots must be interpreted together for an overall understanding of the system

• An example of a stability study is provided on the next slide
Why is Multivariate Analysis Not Common Place in the Pharmaceutical Industry?
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data
2. Principal Component Analysis: The Workhorse of MVA
3. Smart Manufacturing and MAFSS
4. Fundamentals of Multivariate Statistical Process Control
5. Combining Spectral and Process Data
6. Benefits of PAT and Examples
Processes Are Multivariate in Nature

- Initiatives like Six Sigma and Lean Manufacturing work really well for Nice Systems.
- The Univariante approach of Six Sigma has helped manufacturers to analyse their measurement systems.
- The key area lacking in Six Sigma is that it measures multiple responses, but not multivariately.
Design For Six Sigma (DFSS)

- DFSS was recently introduced to help manufacturers to take the next step in process improvement.
- Companies were adjusting and incrementally improving their processes, but were not achieving the expected performances.
- DFSS incorporates, among other tools, a more focussed approach to Design of Experiment (DoE).
- This is a first step towards a Multivariate Approach to Process Design and Control.
Smart Manufacturing is defined as

*Improved Business Knowledge for Continuous Growth*
The major tool for implementing Smart Manufacturing is MAFSS

**Multivariate Analysis For Six Sigma**

….Combining MVA and Process Knowledge, which provide new insights that........

….changes the basis of competition!
Example from The Mining Industry

Historical data: Production (average yield/day vs week no. in 1990)

Key Question: Why do we have this unwanted variation in produced yield?
Example from The Mining Industry

Situation in weeks 18 – 44 in 1990

• Average yield MgCl₂ ton/day varies from 220 to 360 tons in this period

• 4.1 tons MgCl₂ give 1 ton Magnesium by electrolysis

• Price 1 ton Magnesium (1990): $3000 USD

• The average daily production in this period was 296 tons

Weekly loss of profit > 360-296 x 7 x 3000/4.1 > $0.328 Mill. USD

Period loss of profit > 0.328 x 27 (weeks) = $8.856 Mill. USD
Example from The Mining Industry

Exploratory Data Analysis Using PCA. Biplot – a "window" into the data

RESULT: X-expl: 43%, 25%

Direction of Optimisation
The presence of two regression line populations is an indication that a source of variability has not been accounted for. This is known as a "Lurking Variable".

Example from The Mining Industry

Regression Analysis Using Partial Least Squares (PLS) Regression
Example from The Mining Industry

First observations from the PLS analysis

1) A typical "lurking variable" pattern was observed

2) After discussions with the process engineers in the plant, they claimed that the lurking variable may be the "quality" of the MgO pellets that were chlorinated. They had observed that sometimes when the "quality" by their judgement was "poor", the chlorgas emission was correspondingly higher.

3) By luck, the chlorgas emission was indirectly recorded by consumption of the sulfate solution used to purify the emission.

4) The sulfate consumption was added to our historical data as a substitute for MgO pellet quality.
Example from The Mining Industry

PLS Regression now incorporating Sulfate Solution Data

Note: the "two parallel lines pattern" indicating the presence of a "lurking variable" has disappeared!
Example from The Mining Industry: Conclusions

1. By combining multivariate analysis to historical data and incorporating process knowledge, the process parameters Critical to Quality were identified.

2. The most important process parameters were a function of the ”quality of pellets”, previous to this study, this information was unknown

3. These results initiated an R&D project which lead to a definition of optimal pellet quality (including the level of the water %). On-line process analysers were incorporated to ensure stable pellet quality

4. Some years later the plant had improved their average daily production from roughly 300 ton close to 500 ton, without other investments than a project-team and implementing minor systematic efforts to stabilize the two most important process parameters.
Overview of Smart Manufacturing and MAFSS

- **Smart Manufacturing** is an initiative aimed at providing all manufacturing sections with a 21st century toolkit for improving and optimising their processes.

- Multivariate Analysis and in particular **MAFSS** provides such a toolkit.

- The CAMO model of a complete **Solution** is provided on the next slide.
The CAMO Solution

- Data Repository/Historian
- PAT’s
- Process Controls
- QMS
- CQA, CPP, CTQ
- Process Level

Training Consultancy Development

• The Unscrambler
• R&D, Process Improvement, Technical Improvement
• Custom Applications
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data
2. Principal Component Analysis: The Workhorse of MVA
3. Smart Manufacturing and MAFSS
4. Fundamentals of Multivariate Statistical Process Control
5. Combining Spectral and Process Data
6. Benefits of PAT and Examples
Multivariate Statistical Process Control

- Multivariate Statistical Process Control (MSPC) is the essential next step to univariate SPC, for monitoring and controlling multivariable processes.

- A common misconception is that the operational range of a variable in a multivariate process is the same as the operational range of a variable in a univariate process.

- This would be true only if the process variables were independent of each other.
The Ellipse is known as Hotellings $T^2$ Ellipse and represents a 95% confidence region.

There are regions in the multivariate control chart that are forbidden in the univariate charts.

There are also regions in the univariate sense that are out of control in a multivariate sense.
Fundamentals of MSPC
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data
2. Principal Component Analysis: The Workhorse of MVA
3. Smart Manufacturing and MAFSS
4. Fundamentals of Multivariate Statistical Process Control
5. Combining Spectral and Process Data
6. Benefits of PAT and Examples
Combining Spectral and Process Data

Modelling strategies:
1. Hierarchical model with scores
2. Block-weighing
3. Passify variables
4. Incremental modelling
Combining Spectral and Process Data

Hierarchical model

Process steps

Scores from each

Combine scores from individual steps

Response variables; quality

\[ D \quad X_1 \quad X_2 \]

“Super scores”

\[ Y \]
Putting It All Together

The Holistic/Heuristic (HHL) Approach to MSPC

DATA HISTORIAN

SiPAT or SynTQ

Spectroscopic Analysers

Process Inputs and Outputs

Environmental Outputs

Process Interface

Raw Materials

Granulation and Drying

Blending and Milling

Compression and Coating

Holistic Level

Heuristic Level
Session 2: The Increasing Need of Multivariate Analysis for Process Understanding and Control

Presentation outline

1. Historical vs. Designed Data
2. Principal Component Analysis: The Workhorse of MVA
3. Smart Manufacturing and MAFSS
4. Fundamentals of Multivariate Statistical Process Control
5. Combining Spectral and Process Data
6. Benefits of PAT and Examples
Opportunities

Many opportunities exist where PAT can be applied:

– Raw Materials Understanding
– API Manufacture
– Granulation and Drying
– Blending and Milling
– Biopharmaceutical purification
– Liquid mixing and suspension manufacture
– Compression and coating
Benefits of PAT

• Increase productivity
  Better process understanding resulting in less process issues.

• Reduce Cost
  Less waste, less laboratory based testing, increased throughput

• Reduce Regulatory Oversight
  Results from better control of processes

• Greater Quality Assurance
  Quality should be the key driver, the other benefits naturally follow.
Fluid Bed Drying

- The unit operation of Fluid Bed Drying is a critical process step in the manufacture of solid dose products.

- Over dried granules may result in poor compression characteristics.

- Under dried product may result in microbiological growth.

- Current methods for measuring moisture require stopping the drier and taking samples.
The Use of NIR for Monitoring FBD Endpoints

- Near Infrared (NIR) spectroscopy is a mature technology that can be used in-line to monitor processes in real time.
- Ideally suited for the measurement of moisture (and a whole lot more).
- Probes measure the samples as they exist in the process.
- Requires multivariate techniques to extract the moisture information contained in the spectra.
The System
The Process

Load Wet Granule

Start Drier and Monitor

Shut down process based on real time NIR analysis
Practical Considerations of PAT

- Decide on a product
  - One that will provide most benefit of the PAT implementation.

- Decide on a Process
  - Which unit process causes the most issues or variability

- Decide on a Technology
  - What do you want to measure? Is this measure relate to product quality and performance

- Organise a feasibility study
  - Either use or get access to pilot equipment or test the actual process.
Practical Considerations of PAT

• Project Development
  • Do the feasibility trials indicate that the technology will work?

• System Build
  • Get involvement from all departments and finalise system requirements. This should also be done in conjunction with equipment vendors.

• System Commissioning and Qualification
  • Use GAMP5®, ICH and ASTM guidance

• Develop product methods
  • Use multivariate data analysis techniques to build process models.
System Qualification

• Qualification activities vary depending on the complexity of the system

• Work in conjunction with vendors, using their documents wherever possible: GAMP5® and ASTM E2500

• Avoid individual computerised system qualifications and aim to qualify the entire integrated system: GAMP5® and ASTM E2500
System Qualification

Planning, User Requirements and Safety

Verifies

System Design

Verifies

System Review and Approval

Design Qualification

Verifies

System and Module Build

FAT and SAT

Installation Qualification, Operational Qualification

Module integration and Development Testing, Installation Check, Operational Check

Performance Qualification, Operation and Maintenance

Retirement

Verifies

www.camo.com
System Validation

• Requires different mindset to traditional validation approaches: ASTM E55

• System is validated each time a new product method is developed

• The use of PAT allows for an environment where continuous validation of processes is made

• The combination of all process input/output measurements and any PAT tools are used to define the Design Space
Real Time Monitoring

- NIR instrument collects data every 30 seconds
- Multivariate nature of spectral data requires the use of multivariate tools
  - Multivariate models can be used to
    - Define the Design Space: PCA, SIMCA etc
    - Quantify components: PLS
    - Provide insights into the mechanism of the process
- Can also combine spectral and non-spectral inputs
PCA for Defining the Design Space

Decrease in moisture with a corresponding increase in $-\text{CH}_x$ groups
Principal Component Analysis

**Scores**

- Mix 1
- Mix 2
- Mix 3

**RESULTS**, X-expr: 95%, 3%
Spectral Loadings Analysis

Overlay of PC1 and PC2, shows two different types of moisture, Bound and Free moisture?

RESULT: PC(X-expl): 1(95%) 2(3%)
The Design Space for Drying

RESULT 6, X-expl: 95%, 3%, 1%
Process Improvement

- PAT will show where current processes need improvement.
- Manufacturers worldwide must be prepared for this!
  - Better to understand the process and fix it before it turns into a recall.
  - Also better to have a mechanism to adjust a process then to waste or re-work
- Lost production = lost time, effort and profitability
**A Process Example**

- NIR in FBD showed that product was dry after only 10 min of a 40 min drying cycle
- A process study indicated that the powder bed was not dry
- Spatial analysis showed the sides of the bed were drying, but the centre of the bed was not fluidising
- Added a feedback to the control system that pulsed the bed at predefined time points
Spatial Analysis

Loadings show that improvements in moisture distribution are the major source of variability.

Mix 2 is now more consistent when a pulse is used to initiate fluidisation.
Use of At-line NIR for Confirmation

Pulse Eliminated
The Stop At This Point

Desired State

PC1
PC2
Scores

NIR Stop 1
NIR Stop 1
NIR Stop 1
NIR Stop 1
NIR Stop 1
NIR Stop 1
NIR Stop 1
NIR Stop 1

Mix
Mix
Mix
Mix
Mix
Mix
Mix
NIR Stop 2
NIR Stop 2
NIR Stop 2
NIR Stop 2
NIR Stop 2
NIR Stop 2
NIR Stop 2

PC1 Mix 1 ALL S..., X-expl: 95%4%
Process Understanding

Pulse to initiate bed fluidisation
Conclusions

• Pharmaceutical manufacturers worldwide are now in a position to start the pragmatic implementation of PAT

• If approached correctly, PAT can provide massive benefits to productivity cost and quality assurance.

• Requires Risk-Based Scientific Approaches that use the principles of Multivariate Data Analysis.
Conclusions

• Guidance is now in place to help manufacturers pragmatically implement PAT

• Quality improvement should be the major focus of any PAT implementation

• If cost reduction alone is the driver, regulators will not approve changes

• If quality is the focus, cost reduction naturally follows, productivity increases and there is less regulatory oversight in the event of a deviation
CAMO Training in MVA

• If the seminar interested you and you would like to find out more, CAMO is holding training in the application of MVA to Pharmaceutical and Process Data in May 2009.

• For more information, please see one of us at the end of this presentation or visit www.camo.com for more information.
Get value out of your data

Thank you for your attention

Brad Swarbrick
BradSwarbrick@camo.no

http://www.camo.com