

Applications of Multivariate Data Analysis

Presented by **Petter Möree & Jonas Elfving**



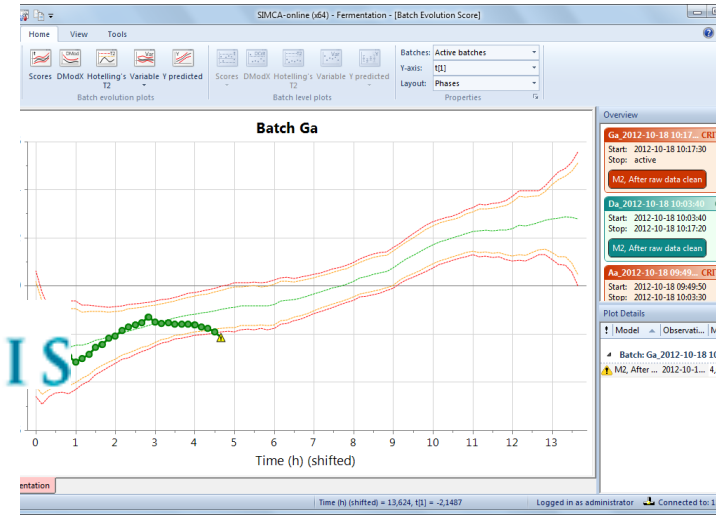
Novartis values from MVDA for PAT & QbD

Increase of **process understanding**

- Identification of influential process parameters
- Identification of correlation pattern among the process parameters
- Generation of process signatures
- Relationship between process parameters and quality attributes

Increase of **process control**

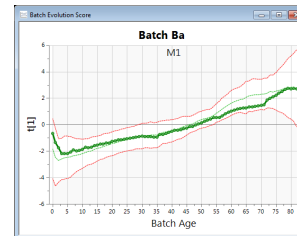
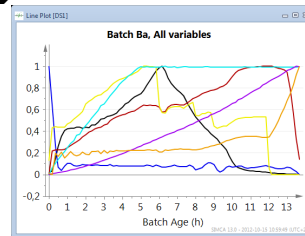
- Enabling on-line early fault detection
- Support for time resolved design space verification (real time quality assurance)
- Predicting quality attributes based on process data
- Excellent tool for root cause, trending analysis and visualization



Business Challenge

- Reduction of dimensionality →
- Conformity check
- Real time release testing
- Trend analysis
- Root cause analysis

Solution



Results and Benefits

- Ability to be proactive rather than reactive to variation or poor quality.
- Saving batches
- Reducing OOS
- Helping to optimize the process

Topics

- Introduction: MVDA in the context of pharmaceutical production
- Case studies for MVDA I
 - **Process monitoring** of a granulation process in pharmaceutical production
- Case studies for MVDA II
 - Statistical Process Control Biopharmaceutical Production **for optimization**
- Short real-time demonstration

Using PI Server, PI Batch and PI Event Frame & PI Interface for SIMCA-online

Umetrics

- Part of ~**1Billion** conglomerate
- The **market leader** in software for multivariate analysis (MVDA) & Design of Experiments (DOE)
- **25+ years** in the market
- Off line analysis tools
- On-Line process monitoring and fault detection
- 700+ companies, 7,000+ users
- Pharmaceutical, Biotech, Chemical, Food, Semiconductors and more
- Worldwide Presence with MKS
- Close collaboration with universities in USA, Sweden, UK and Canada



SIMCA-online

With SIMCA-online , you have the power to monitor manufacturing evolution in real time providing quality information before the product is finished. SIMCA-online makes this possible using multivariate techniques combined with conventional SPC (Statistical Process Control), underpinned by a seamless graphical interface. Finally, you have the ability to react to quality issues as they happen.

SIMCA-OnLine

BY UMETRICS

Average Overall Rating

★★★★★ (1 Reviews)

[Write a Review](#)

[Contact Partner](#)


Share on Twitter



Share on Facebook



Share on LinkedIn



Send as Email

Features & Benefits

- Increase manufacturing efficiency & quality using proven statistical methods
- Optimize process to reach desired quality target while reducing risk
- Very fast return on investment, often from the first major production deviation alert
- Provides information for engineering to make continuous improvements in the process

Industry

Oil & Gas
Chemical & Petrochemicals
Materials, Mines, Metals & Metallurgy
Pharmaceuticals, Food & Life Sciences
Pulp & Paper
Power & Utilities
Critical Facilities, Data Centers & IT

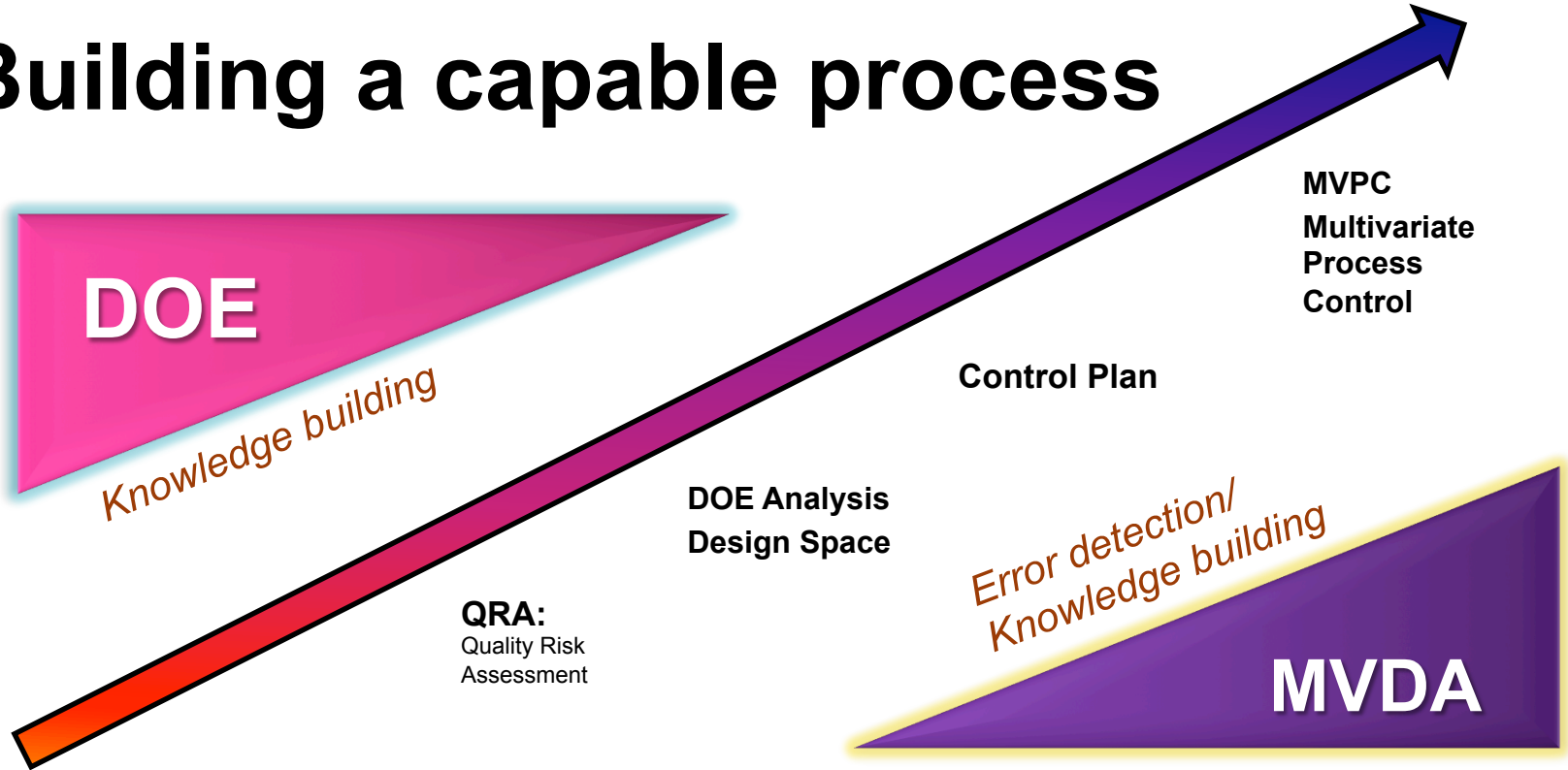
Solution Area

Process Characterization and Analytics
Process Control/Optimization
Visualization

Region Sold

Africa

Building a capable process



QFD
Quality Function
Deployment

QRA:
Quality Risk
Assessment

DOE Analysis
Design Space

Control Plan

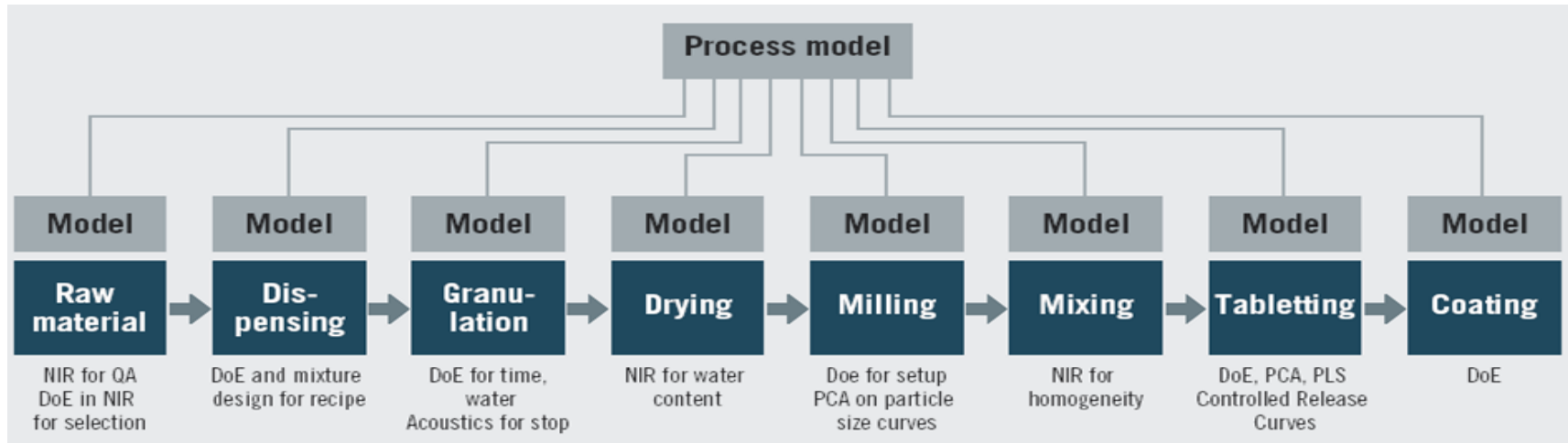
MVPC
Multivariate
Process
Control

MVDA
Error detection/
Knowledge building

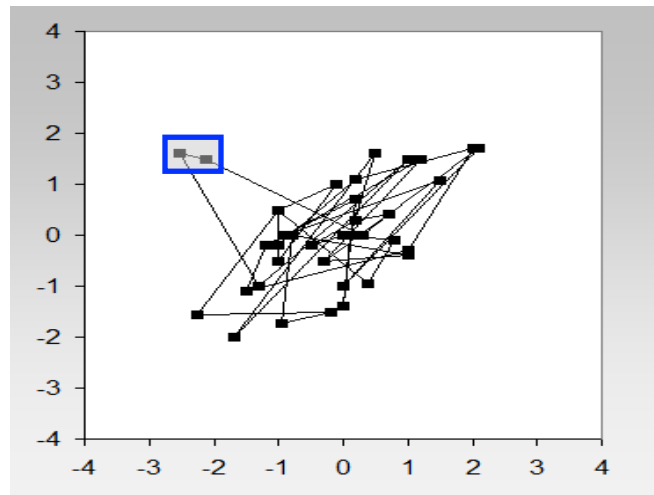
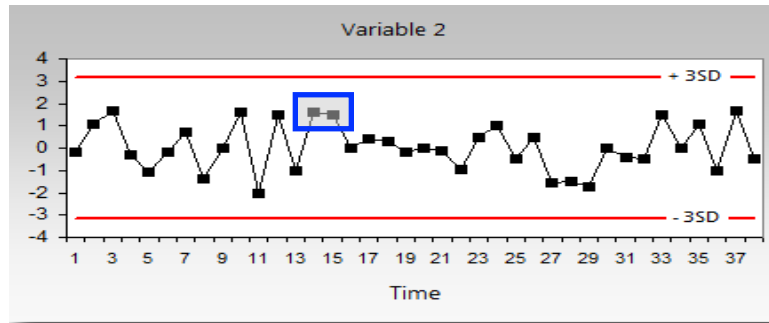
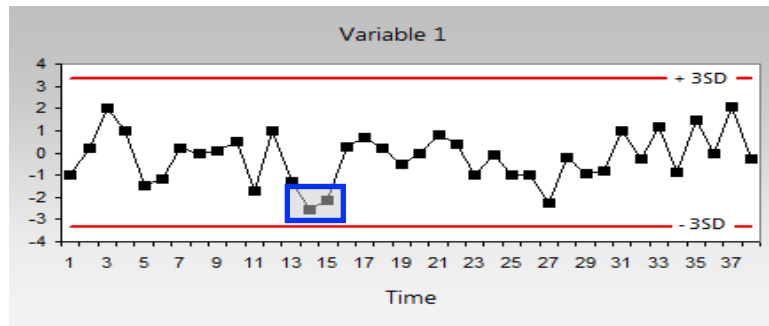
- *DOE is a knowledge building tool for process development*
- *MVDA is used both for process understanding and process monitoring*

Our Customers' Goals in Pharma

- The goal in Pharma production is to help take advantage of data present in the development labs, and the production environments all the way from API to the final product. = **ROI**

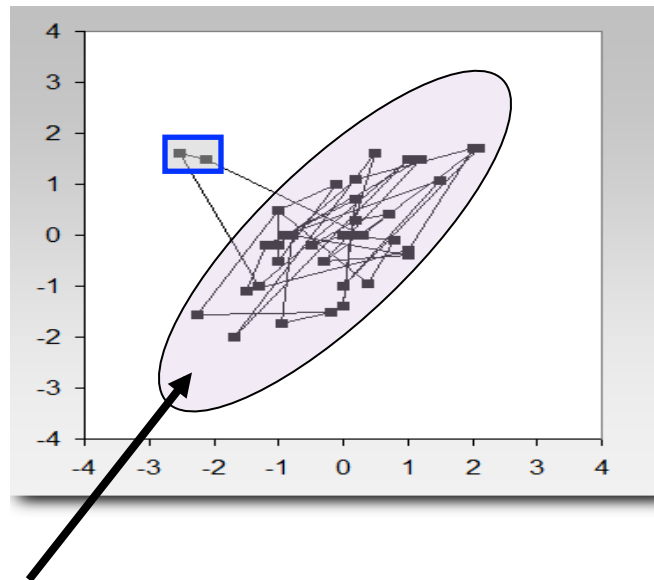
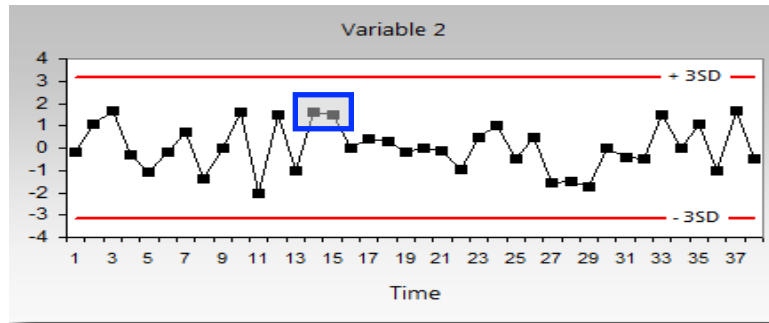
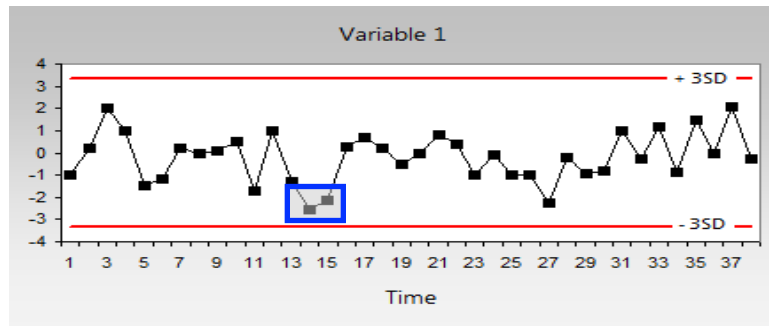


The Need for Multivariate I



The information is found in the correlation pattern - not in the individual variables!

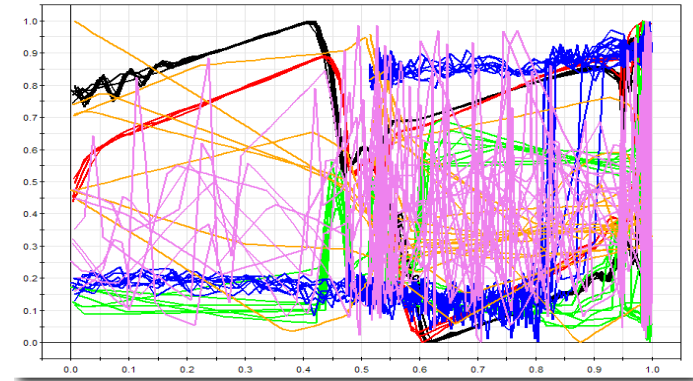
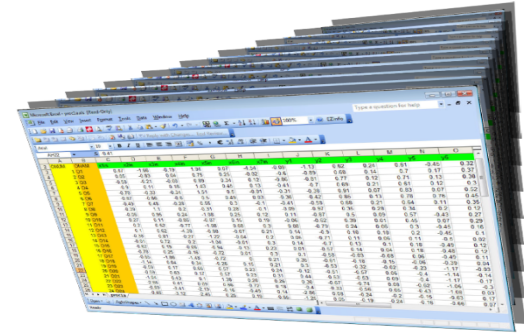
The Need for Multivariate I



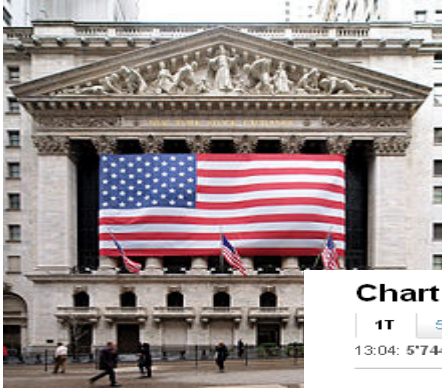
Multivariate Control Limits

The Need for Multivariate II

- Data explosion, more process measurements than ever before, reduce false alarms
- Spectrometers
 - NIR, FTIR, RAMAN, UV, LLSD
 - MS, GC, HPLC
- Process Sensors
 - Acoustic, Video
 - P, T, Flow, pH
 - pO_2 pCO_2
- Require MVDA methods to visualise and extract reliable information from raw data
- MVDA handles noise, missing data, correlation and visualize in graphs



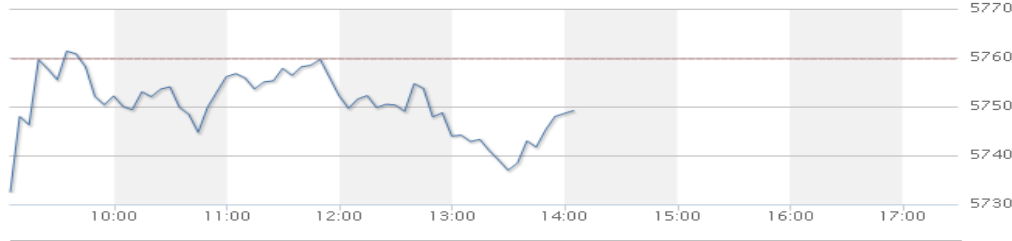
This control chart is familiar to you ?



Chart

1T 5T 1M 6M 1J 5J 10J

13:04: 5'744.2

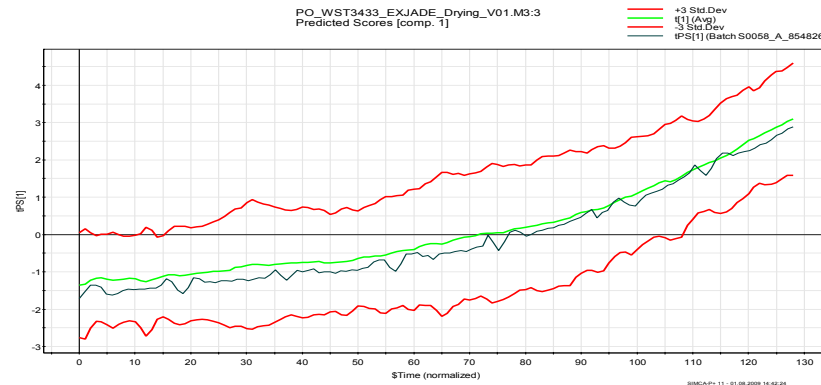


$$\text{SMI} = x1^* \text{Novartis} + x2^* \text{Roche} + x3^* \text{Merck} + x3^* \text{FB} \dots$$

So this control chart is easy to understand....



PO_WST3433_EXJADE_Drying_V01.M3:3
Predicted Scores [comp. 1]



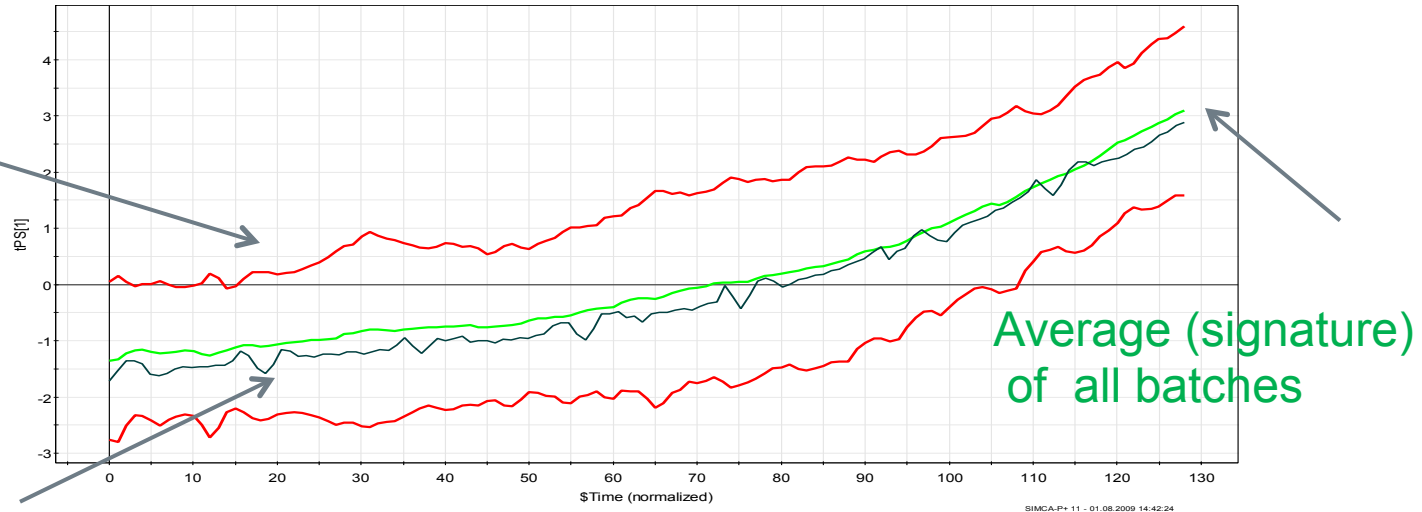
$$t_1 = x1 * \text{Temperature} + x2 * \text{Pressure} + x3 * \text{Agitation speed} \dots$$

MSPC

Observation Level

- Example of a drying step

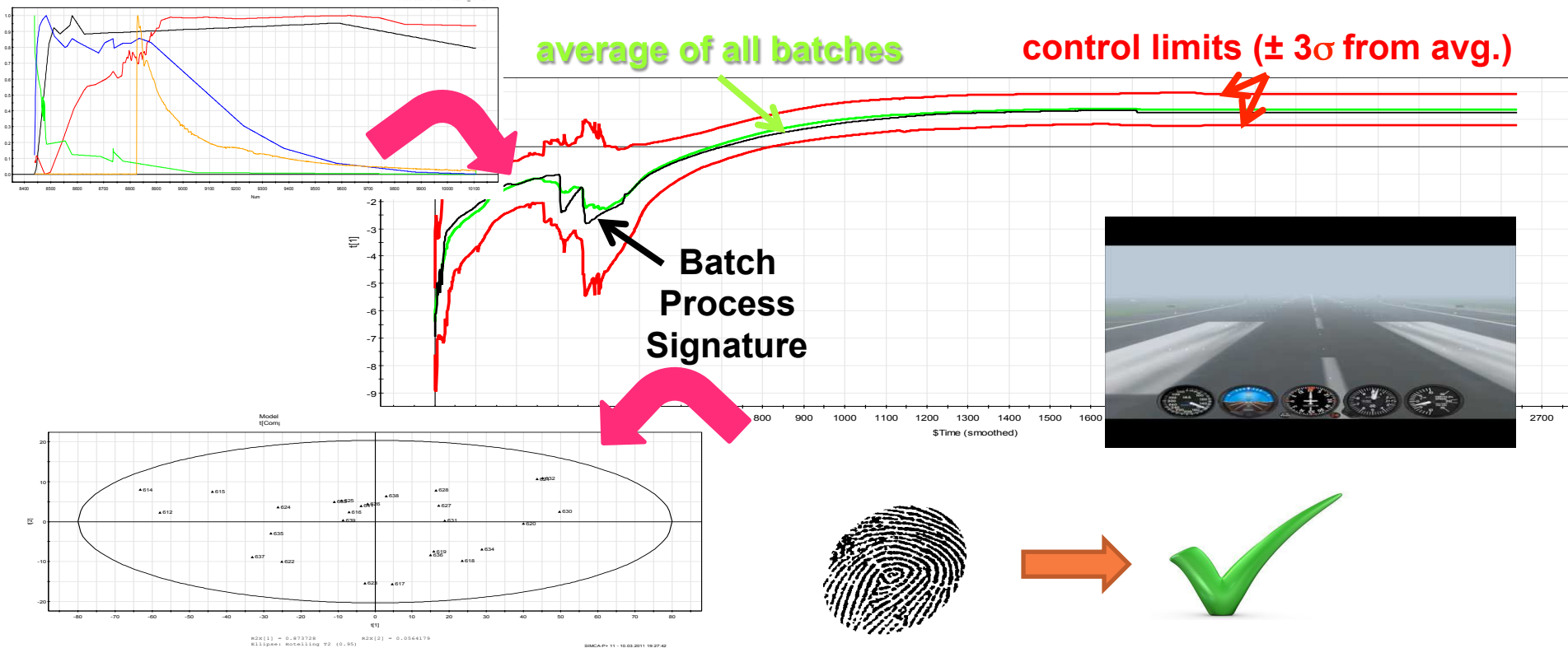
Control limits

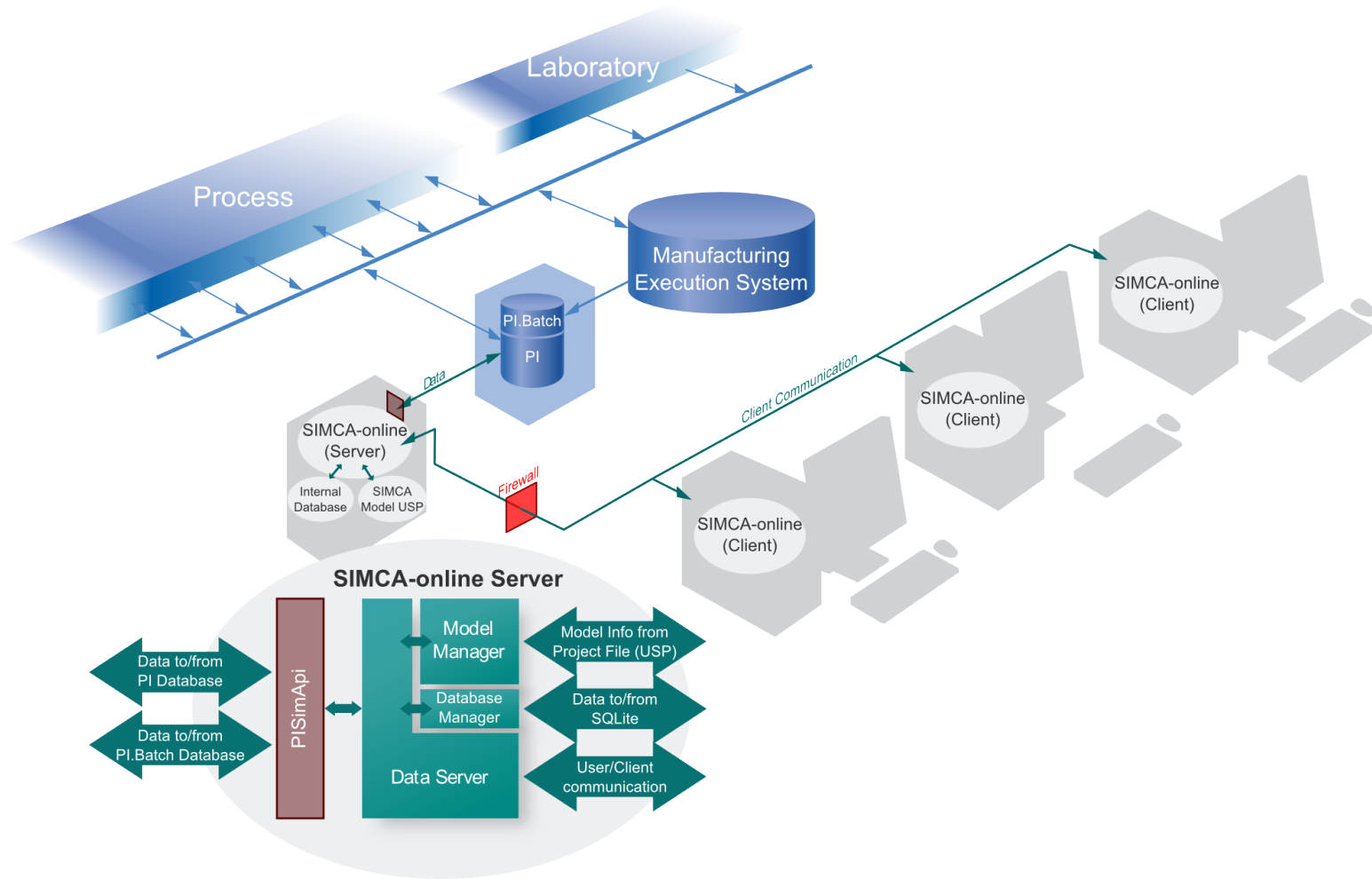


New batch assessed by the model

Statistical Process Control

BATCH CONTROL CHART





Work and Data flow

For Method Development

Reduction of Dimensionality

Final Model

Batch Level

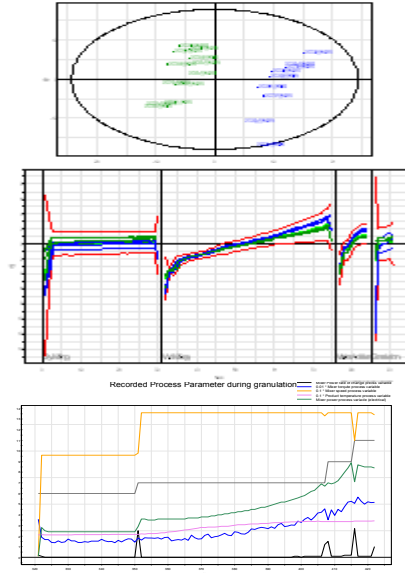
Observation
Level

All Process
Parameters

Individual
Probes

Individual
Probes

...



Aims:

- Creation of batch signature
- Identify correlation patterns

Work and Data flow

For Routine Use in Production

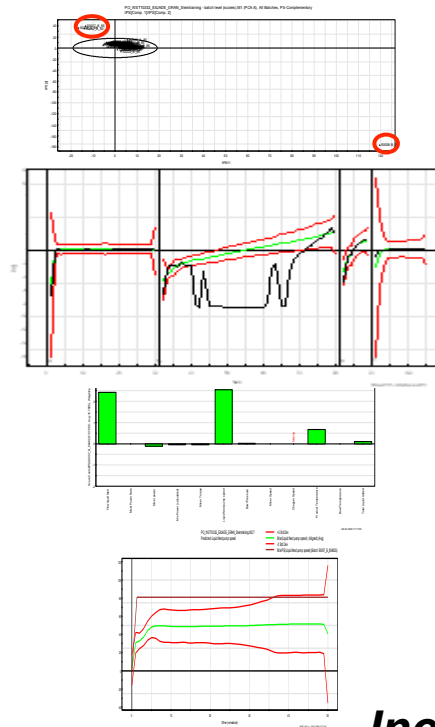
SIMCA-online

Batch Level

Observation
Level

Identification of
responsible Parameter(s)

Investigation on process data



Aims:

- Conformity check
- Real time release testing
- Trend analysis
- Root cause analysis

Increase of level of detail

MVDA applied to granulation

Example for a qualitative model used for MSPC

- **High Shear Granulation (Production Scale)**
- Four phases
 - Dry Mixing
 - Wet mixing
 - Water addition
 - Granulation
- Variables
 - Power consumption and torque
 - Product temperatures
 - Granulator and chopper speed
 - Pump properties and flow parameters

Key consideration for method development

- **Which observations should be included ?**
 - Sufficient number of batches to cover **natural variability**
 - **DoE Data** for special cause variations
 - Exclusion of anomalous, unsteady, discontinuous data (spikes)
- **Which variables should be included ?**
 - Exclusion of variables with no impact and low reliability
 - Weighting and transforming of variables
 - How many scores should be considered
- **Data alignment and synchronization**
 - Definition of start/stop conditions and phases
 - Merger of variables with different acquisition rates
 - Normalization of time based maturity variables vs. absolute time

Granulation Process

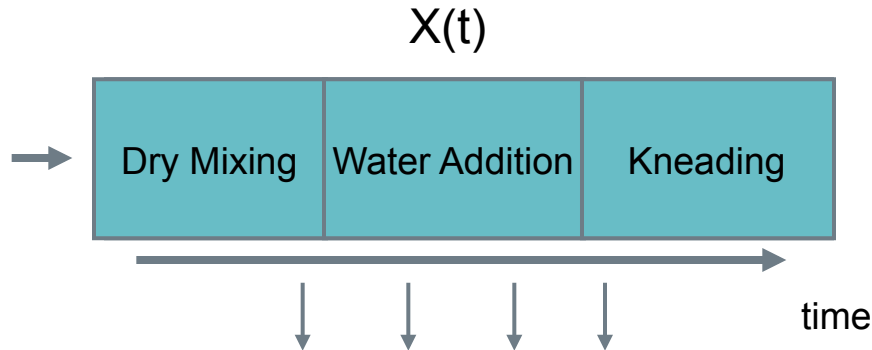
Data Setup

Input:
Initial conditions

Variable initial conditions :

- Amount of granulation liquid
- Granulation time

X_0



1. Power consumption
2. Power consumption rate
3. Torque
4. Product Temperature
5. Mixer Speed
6. Chopper Speed
7. Water addition rate
8. Flow Liquid Speed
9. Liquid Speed Pump

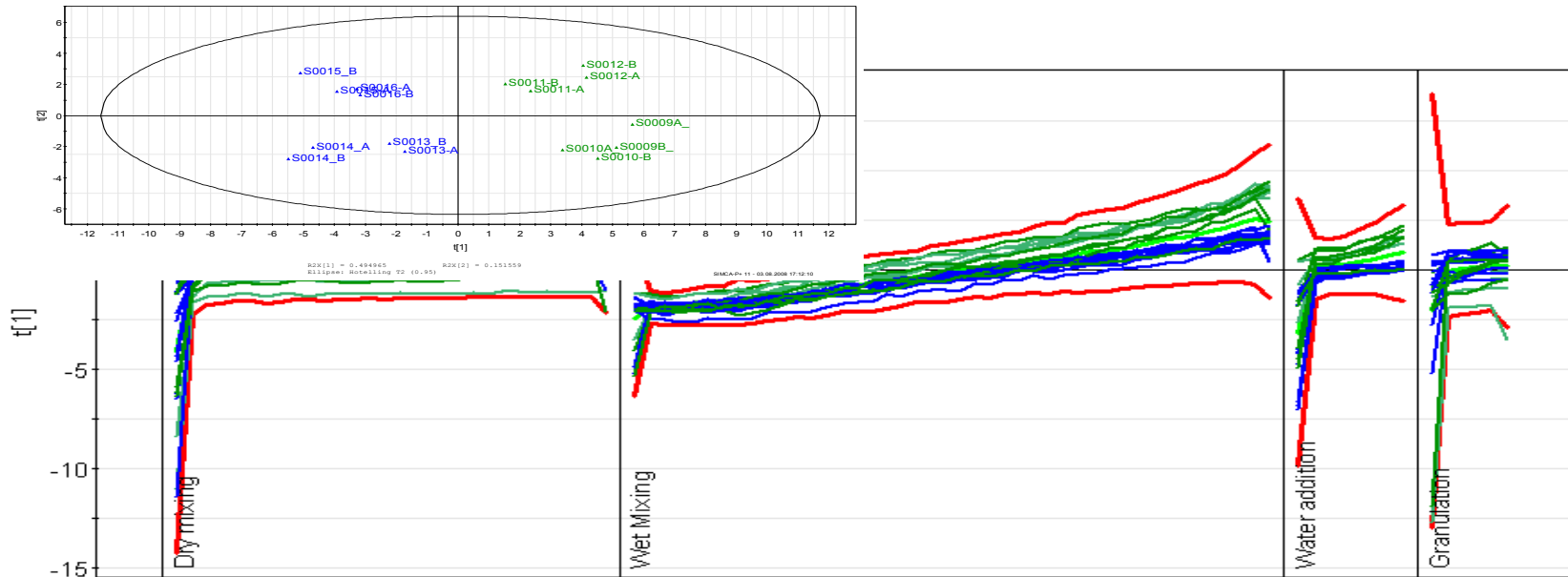
Output:
Quality Attributes

IPC results:
•LOD
•PSD

Y

Example Granulation

Results of DoE investigation

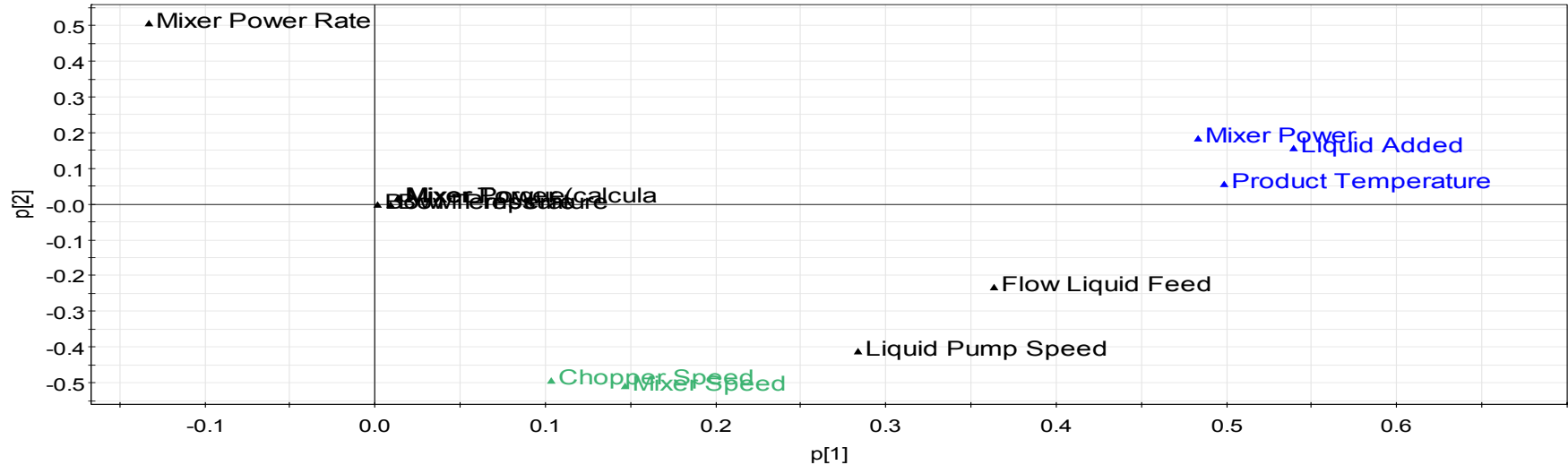


- Different **phases** during granulation are monitored
- Process variability are reflected by the red lines
- Clustering of DoE batches can be visualised
- Common cause vs. special cause variation

Loadings

Identification of process parameters contributing to process variability

Loadings of Granulation Models



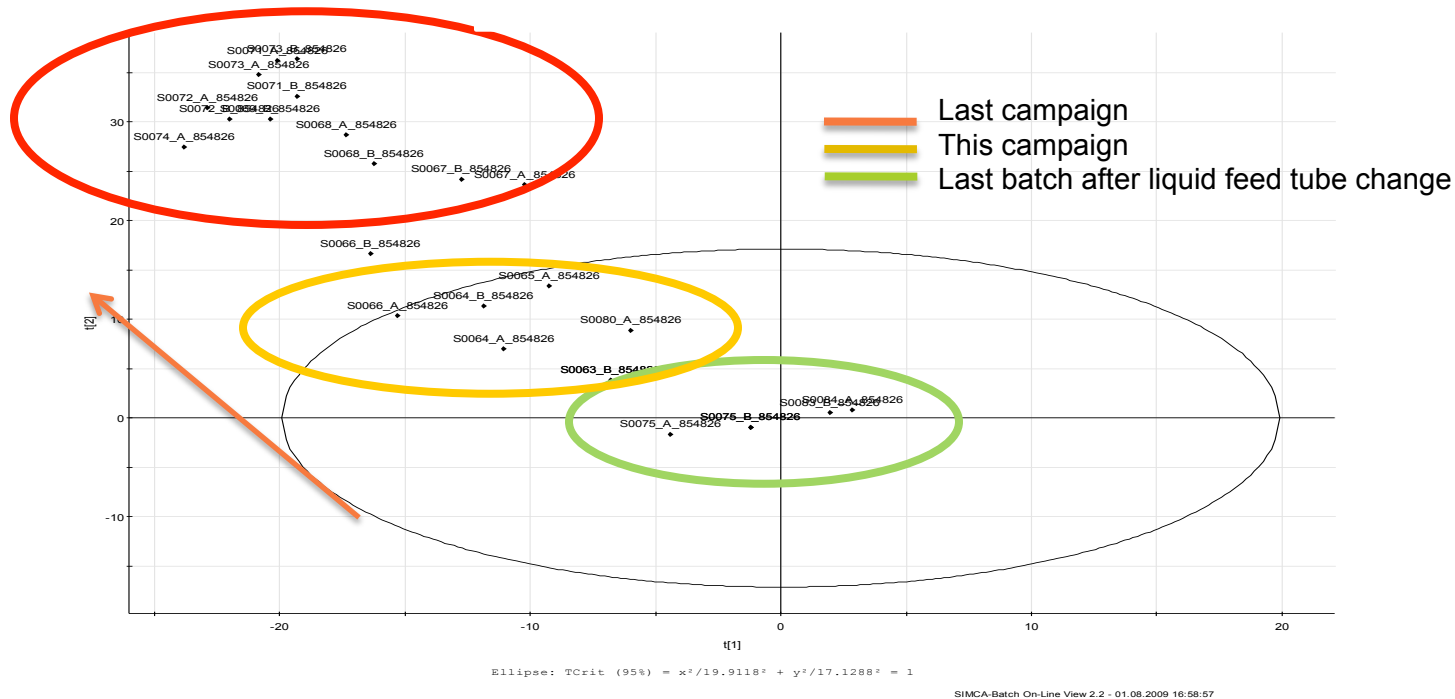
R2X[1] = 0.401621 R2X[2] = 0.318985

SIMCA-P+ 11 - 04.01.2010 18:07:36

- Which parameters are most influential?
- How do the variables correlate to each other?

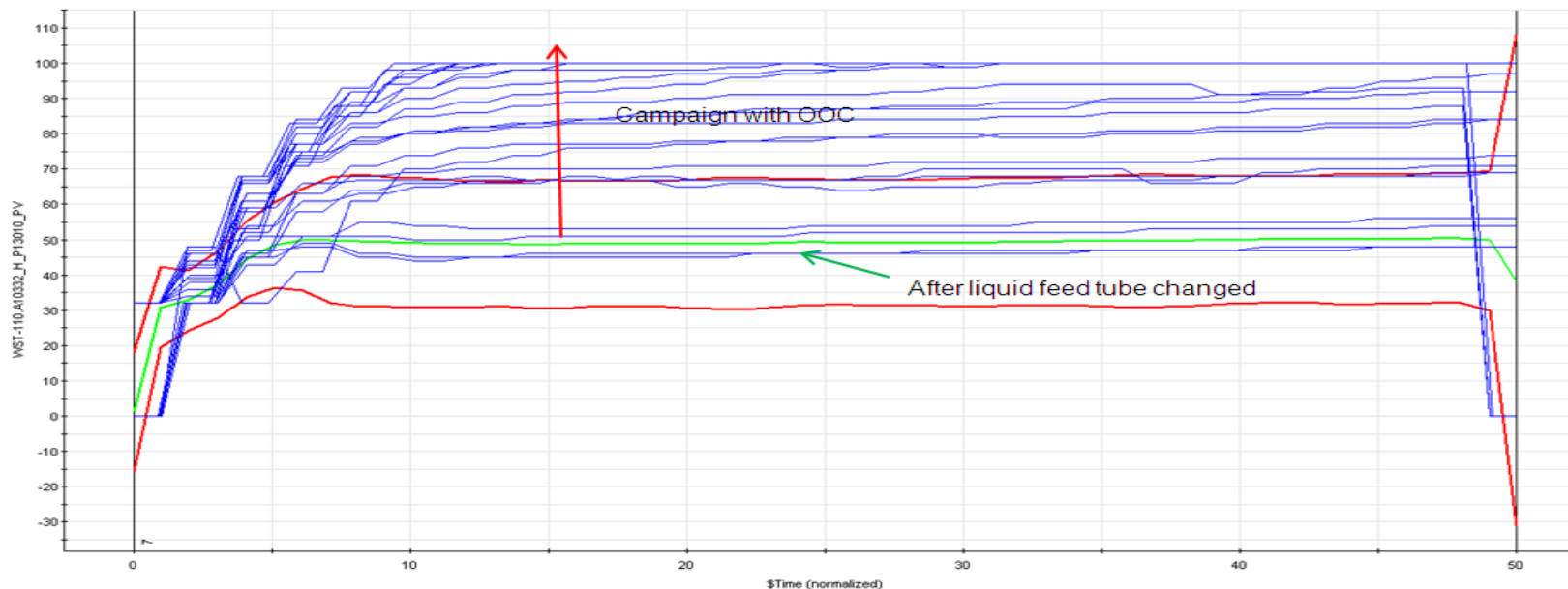
First experience in production

Preventive maintenance

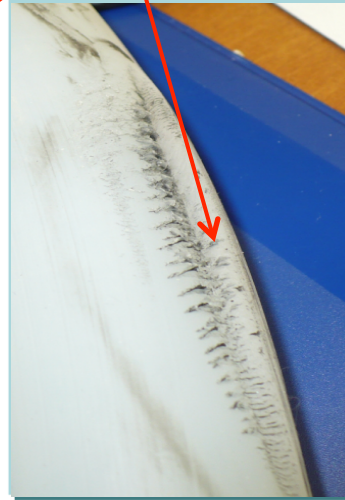
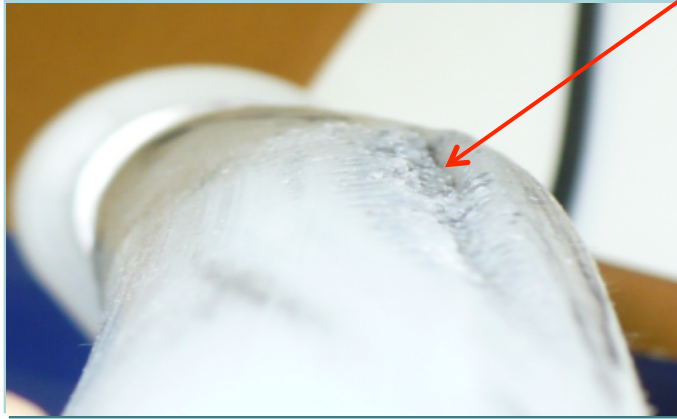
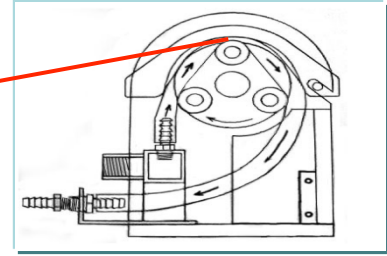
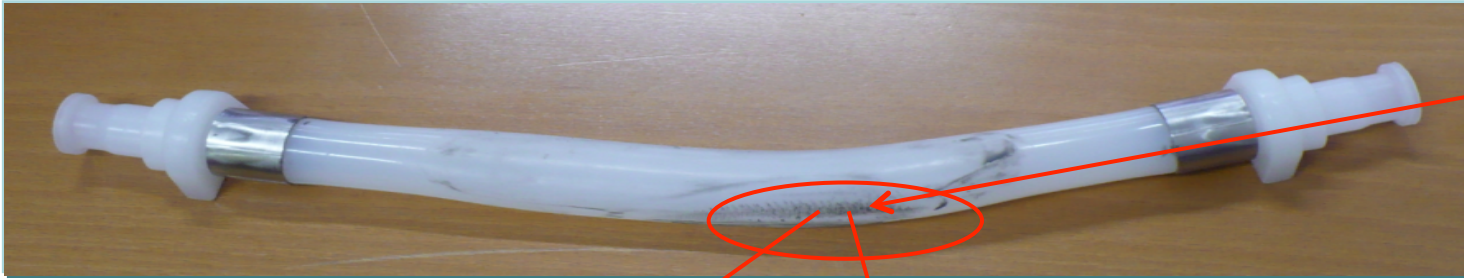


Root cause analysis

Liquid feed pump speed trajectory during the latest 2 campaigns



Root Cause: Worn out liquid feed tube



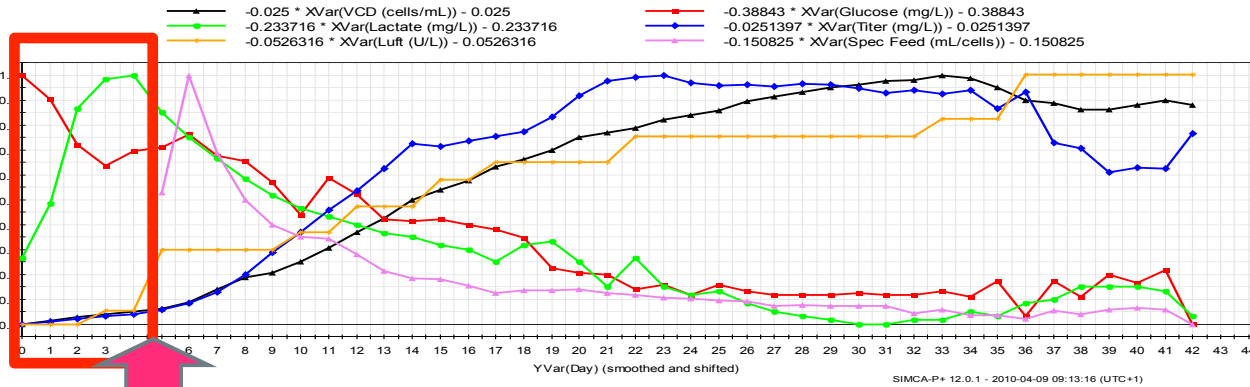
By replacing the tube
batches were saved.

\$ XXXXXXXX saved

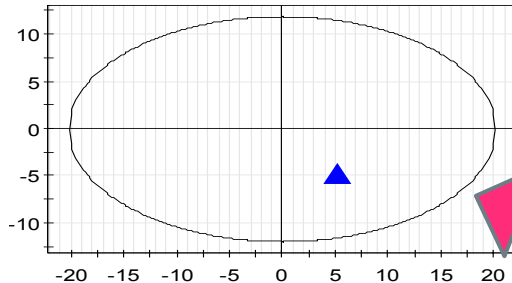
Cell cultivation process

- **Motivation**: established process, not fully characterized, most of process understanding based on experience
- **Modeling**: > 80 DS batches, fully meeting release specifications but some variability in main fermentation yield is observed
- Defined reference (golden) batches, i.e. batches which provided the highest total amount of antibody during cell culture

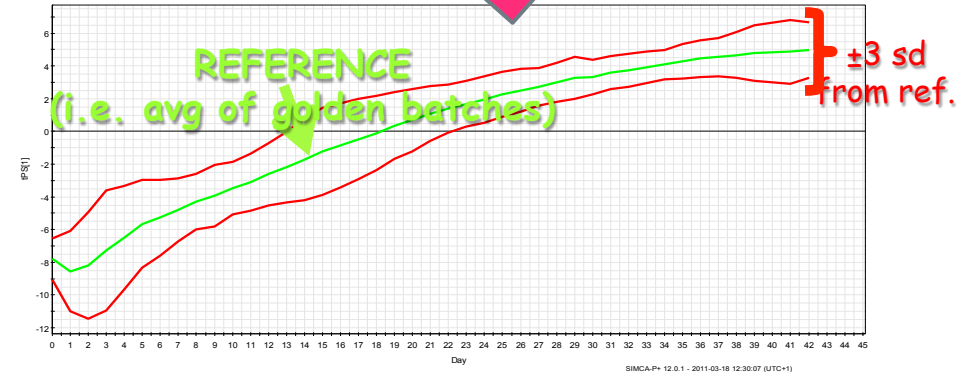
Statistical Process Control for cell cultivation process



Perfusion
Start

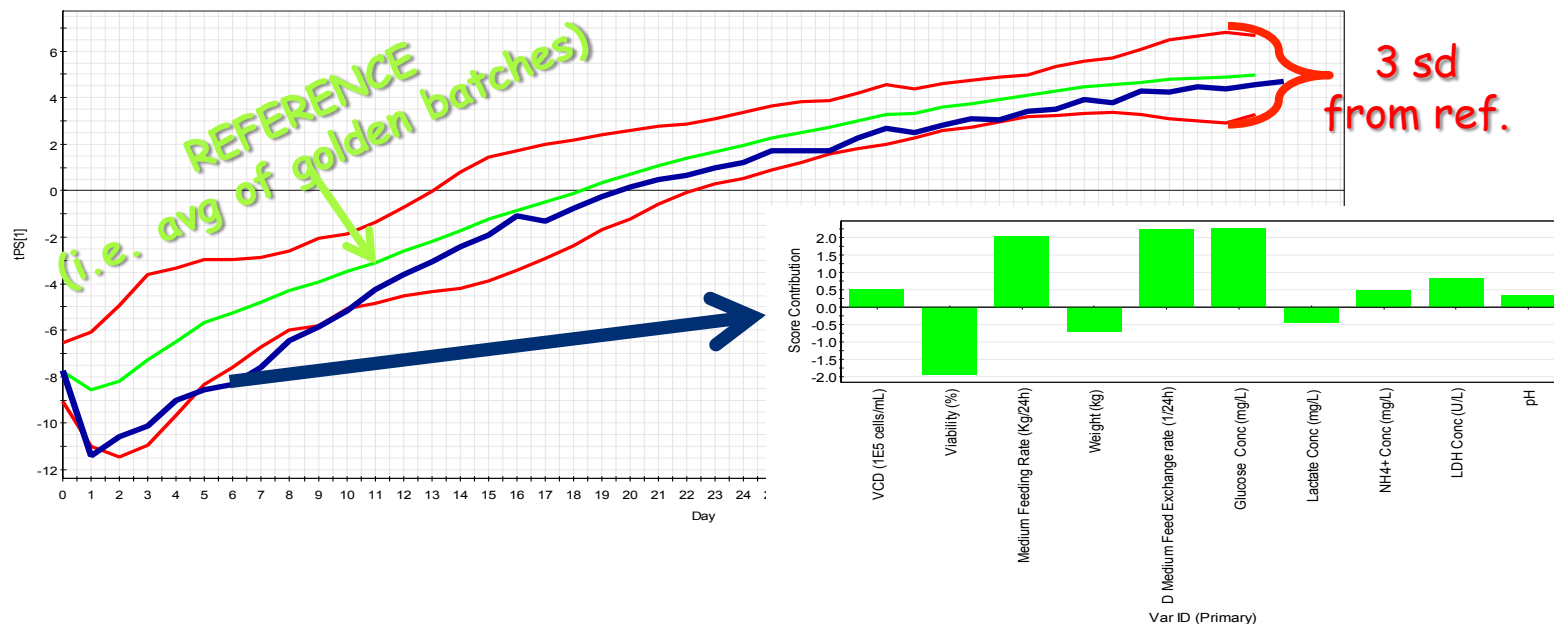


All the variables for each batch are summarized into one quantity (carrier of the information) - process signature



Improve Process Understanding

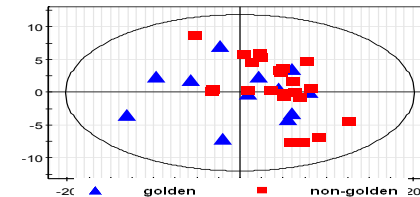
Compared the other batches (“non-golden”) against the “golden” to establish which are the variables responsible for the observed differences



SIMCA-P+ 12.0.1 - 2011-04-29 18:06:16 (UTC+1)

MVDA learning

- Generated useful process knowledge
 - Enhanced process understanding supported by data
 - Improved process consistency
- Established key parameters for cell cultivation
 - medium feeding rate
 - inoculation cell density
 - cells aeration
- No correlation between cell behavior and DS quality attributes



Petter Moree

petter.moree@umetrics.com

Director, Global Product Manager
Umetrics



THANK

YOU

Brought to you by

