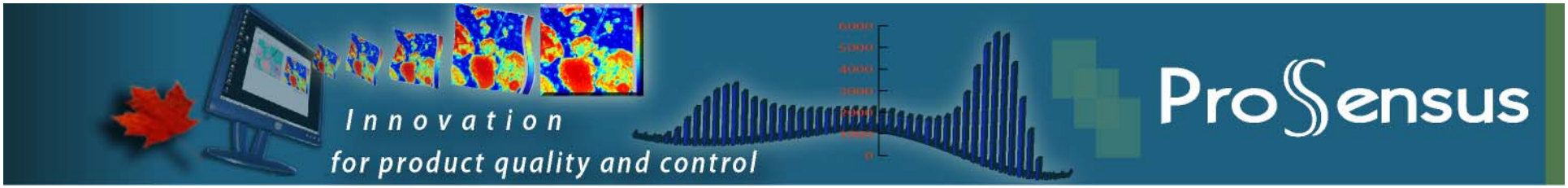


Latent Variable Methods in Process Systems Engineering

John F. MacGregor

ProSensus, Inc. and McMaster University

Hamilton, ON, Canada

www.prosensus.ca

(c) 2004-2008, ProSensus, Inc.

OUTLINE

- Presentation:
  - Will be conceptual in nature
  - Will cover many areas of Process Systems Engineering
  - Will be illustrated with numerous industrial examples
  - But will not cover any topic in much detail

- Objective:
  - Provide a feel for Latent Variable (LV) models, why they are used, and their great potential in many important problems

(c) 2004-2008, ProSensus, Inc.

Process Systems Engineering?

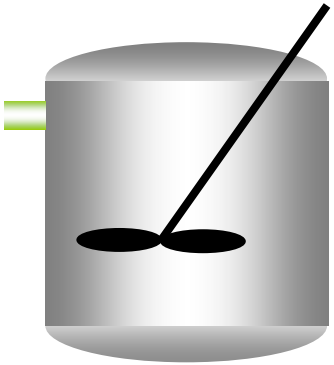
- Process modeling, simulation, design, optimization, control.
- But it also involves data analysis
  - learning from industrial data
    - An area of PSE that is poorly taught in many engineering programs
- This presentation is focused on this latter topic
  - The nature of industrial data
  - Latent Variable models
  - How to extract information from these messy data bases for:
    - Passive applications: Gaining process understanding, process monitoring, soft sensors
    - Active applications: Optimization, Control, Product development
  - Will illustrate concepts with industrial applications

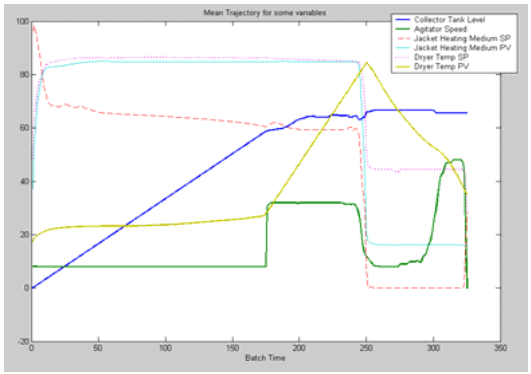
(c) 2004-2008, ProSensus, Inc.

A. Types of Processes and Data Structures

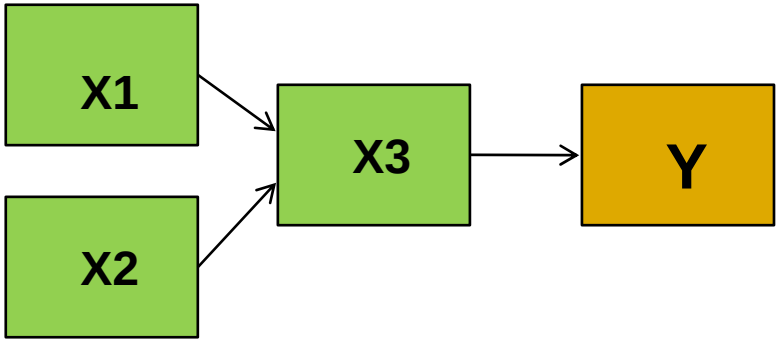
- Continuous Processes

- Batch Processes

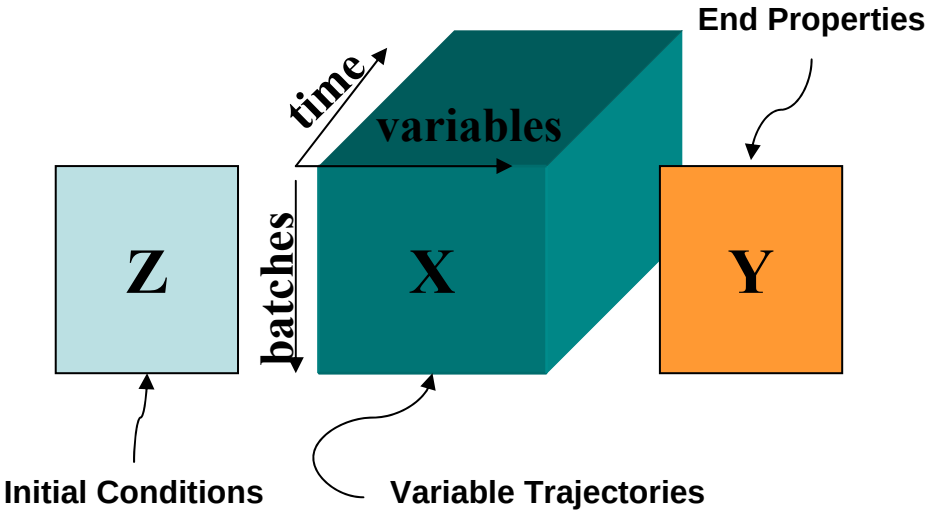




- Data structures



- Data structures



(c) 2004-2008, ProSensus, Inc.

Nature of process data

- High dimensional
  - Many variables measured at many times
- Non-causal in nature
  - No cause and effect information among individual variables
- Non-full rank
  - Process really varies in much lower dimensional space
- Missing data
  - 10 – 30 % is common (with some columns/rows missing 90%)
- Low signal to noise ratio
  - Little information in any one variable

- Latent variable models are ideal for these problems

(c) 2004-2008, ProSensus, Inc.

B. Concept of latent variables

Measurements are available on K physical variables: matrix=**X**

K columns

	1	2	3	4	5	6	7	8	9	1
1	Primary ID	Prim In T	Sec In T	Prim Out T	Feed Flow	Chamb P	Diff P Bag/h	System P	Exhaust P	Sec
2	2006-04-05 16:35:00.00	119.049	116.541	41.1646	76.5042	320.199	126.565	66.401	-61.6004	41.
3	2006-04-05 16:35:05.00	119.046	116.532	41.1979	76.4959	325.755	126.636	95.8617	-43.3963	41.
4	2006-04-05 16:35:10.00	119.044	116.523	41.1626	76.4875	321.37	126.708	82.759	-52.5372	41.
5	2006-04-05 16:35:15.00	119.041	116.514	41.1274	76.4792	327.09	126.78	80.6494	-51.5954	41.
6	2006-04-05 16:35:20.00	119.039	116.505	41.101	76.4709	326.797	126.851	94.5307	-43.7692	41.
7	2006-04-05 16:35:25.00	119.036	116.497	41.0367	76.4625	318.052	126.923	85.1925	-50.9631	41.
8	2006-04-05 16:35:30.00	119.034	116.488	41.281	76.4542	323.099	126.995	72.5004	-56.6797	41.

= **X**

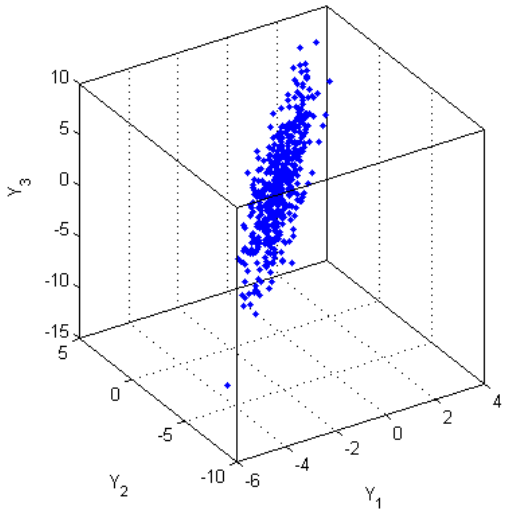
But, the process is actually driven by small set of "A" (A « K) *independent* latent variables, **T.**

- Raw material variations
- Equipment variations
- Environmental (temp, humidity, etc.) variations

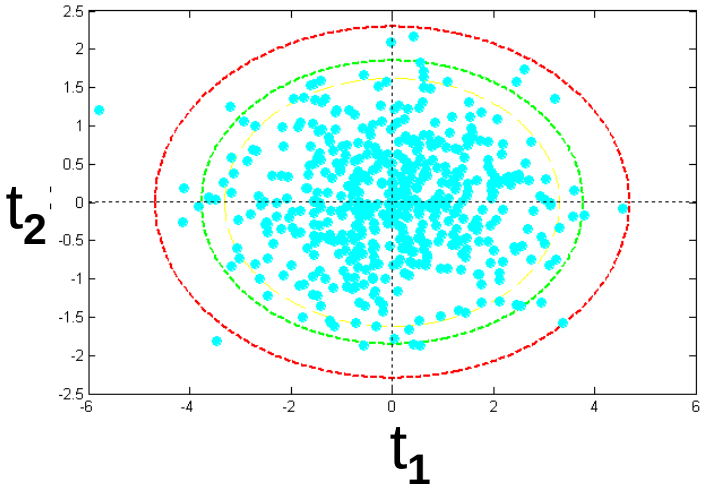
(c) 2004-2008, ProSensus, Inc.

Projection of data onto a low dimensional latent variable space (T)

Measured variables



Latent variable space



X

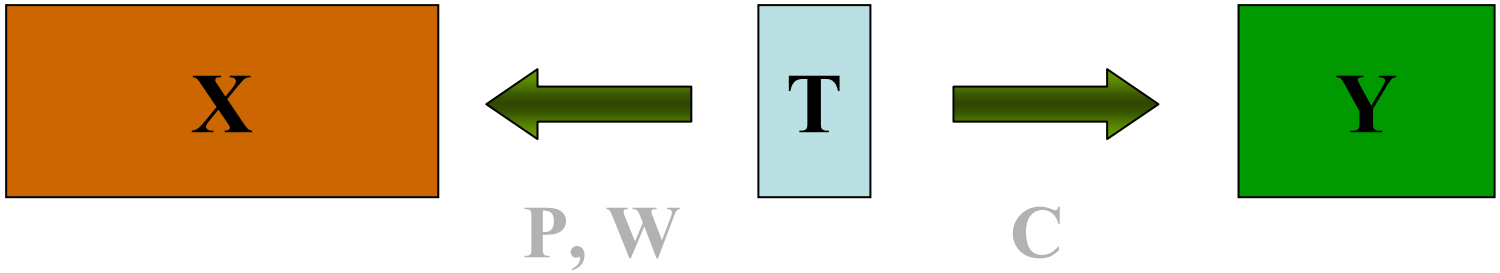
P

T

Television analogy

(c) 2004-2008, ProSensus, Inc.

Latent variable regression models



$X = TP^T + E$

$Y = TC^T + F$

$T = XW^*$

Symmetric in X and Y

- Both X and Y are functions of the latent variables, T
- No hypothesized relationship between X and Y
- Choice of X and Y is arbitrary (up to user)
- A model exists for the X space as well as for Y (a key point)

(c) 2004-2008, ProSensus, Inc.

Estimation of LV Model Parameters

- Parameters: W*, C, P

- Principal Component Analysis
  - Single matrix X: Maximizes the variance explained
- PLS (Projection to Latent Structures / Partial Least Squares)
  - Maximizes covariance of (X, Y)
- Reduced Rank Regression
  - Maximizes Var(Y) explained by X
- Canonical Variate Analysis (CVA)
  - Maximizes correlation (X, Y)

- Appear to be subtle differences, but method used is often critical to the application

(c) 2004-2008, ProSensus, Inc.

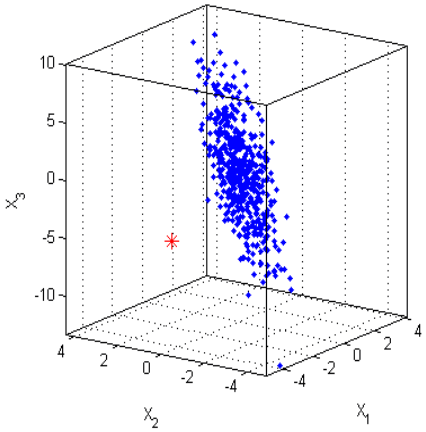
Subspace Identification
Latent variable methods you may be familiar with.

- Subspace identification methods are latent variable methods
  - N4SID – Equivalent to Reduced Rank Regression (RRR) (maximizes the variance in Y explained thru correlation with X
  - CVA – Canonical Correlation Analysis (maximizes the correlation between X and Y)
  - State variables are the latent variables.

(c) 2004-2008, ProSensus, Inc.

Important Concepts in Latent Variable Models

- Handle reduced rank nature of the data
  - Work in new low dimensional orthogonal LV space (t1, t2,…)
- Model for X space as well as Y space (PLS)
  - $X = TP^T + E$; $Y = TC^T + F$
    - Unique among regression methods in this respect
  - X space model will be the key to all applications in this talk
    - Essential for uniqueness and for interpretation
    - Essential for checking validity of new data
    - Essential to handle missing data



- Provide causal models in LV space
  - Optimization & control can be done in this space
    - only space where this is justified

(c) 2004-2008, ProSensus, Inc.

Use of LV Models

- Multivariate latent variable (LV) methods have been widely used in passive chemometric environments
  - A passive environment is one in which the model is only used to interpret data arising from a constant environment
    - Calibration
    - Inferential models (soft sensors)
    - Monitoring of processes
- Used much less frequently in an active environment
  - An active environment is one in which the model will be used to actively adjust the process environment
    - Optimization
    - Control
    - Product Development

(c) 2004-2008, ProSensus, Inc.

Causality in Latent Variable models

- In the passive application of LV models no causality is required
  - Model use only requires that future data follow the same structure
  - No causality is implied or needed among the variables for use of the model
    - Calibration; soft sensors; process monitoring
- For active use such as in optimization and control one needs causal models
  - For empirical models to be causal in certain x-variables – we need to have independent variation (DOE's) in those x's.
  - But much process modeling uses "happenstance data" that arise in the natural operation of the process
    - These models do not yield causal effects of individual x's on the y's
  - But LV models do provide causal models in the low dimensional LV space
    - ie. if we move in LV space (t1, t2, …) we can predict the causal effects of these moves on X and Y thru the X and Y space models
    - Will use this fact together with the model of the X-space to perform optimization and control in the LV spaces

(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- Analysis of historical data
- Process monitoring
- Inferential models / Soft sensors

Passive applications

- Optimization of process operation
- Control
- Scale-up and transfer between plants
- Rapid development of new products

Active applications

(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- **Analysis of historical data**
- Process monitoring
- Inferential models / Soft sensors

Passive applications

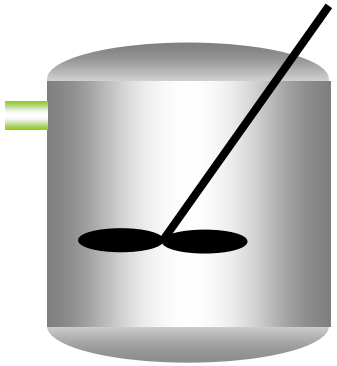
- Optimization of process operation
- Control
- Scale-up and transfer between plants
- Rapid development of new products

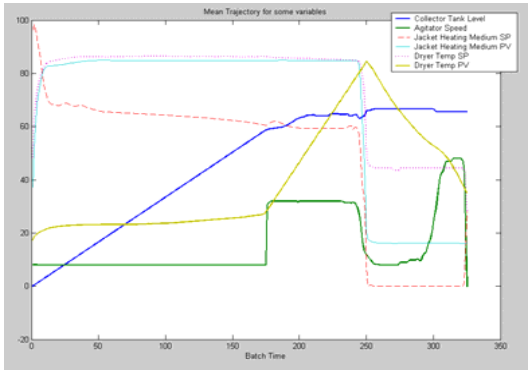
Active applications

(c) 2004-2008, ProSensus, Inc.

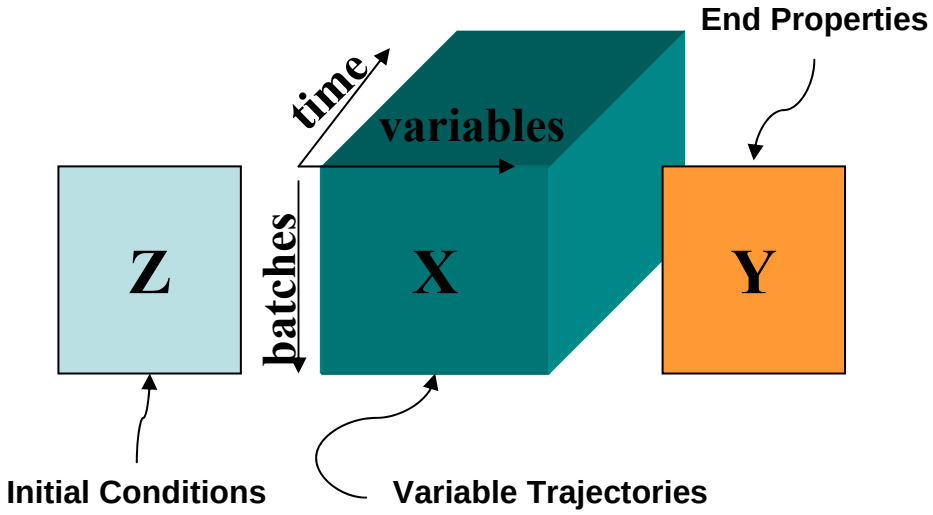
Analysis of Historical Batch Data

- Batch Processes





- Data structure



Herbicide Manufacture

Z - Chemistry of materials
- Discrete process events

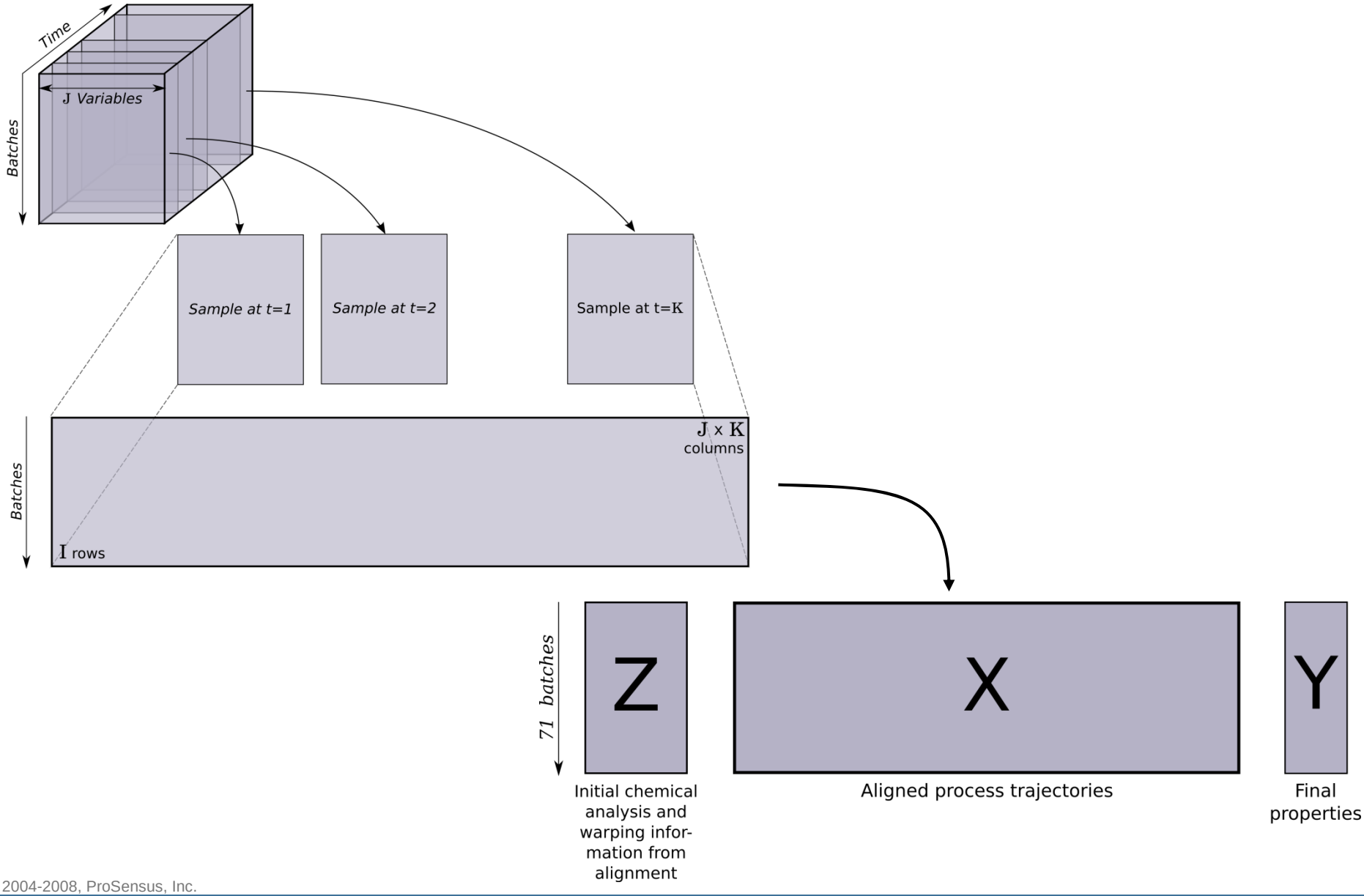
X - Process variable trajectories

Y - Final quality

- 71 batches
~ 400,000 data points

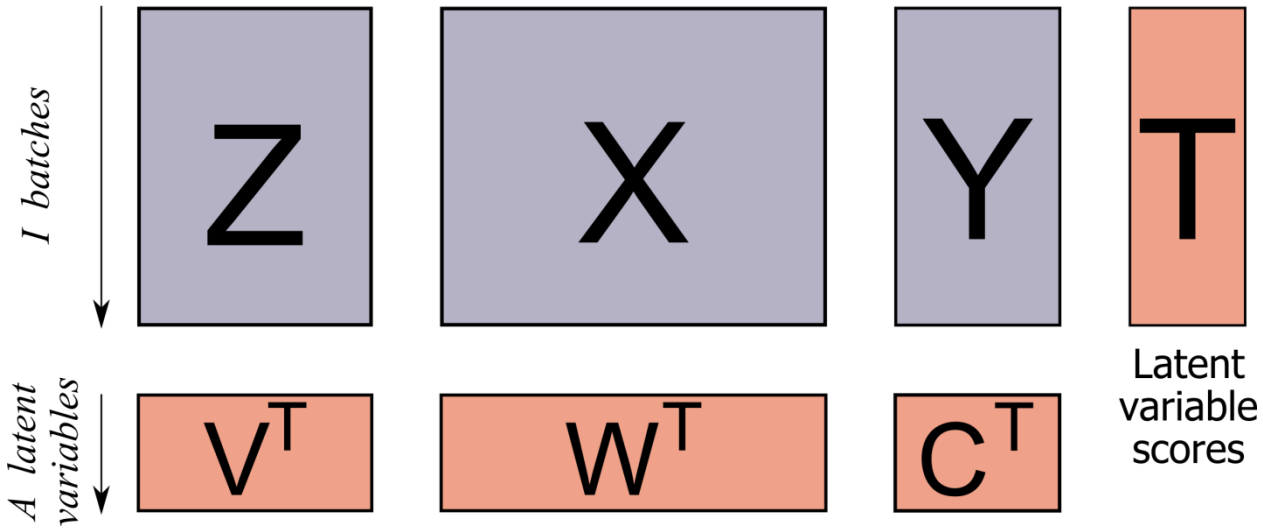
(c) 2004-2008, ProSensus, Inc.

Unfolding and blocking the data



(c) 2004-2008, ProSensus, Inc.

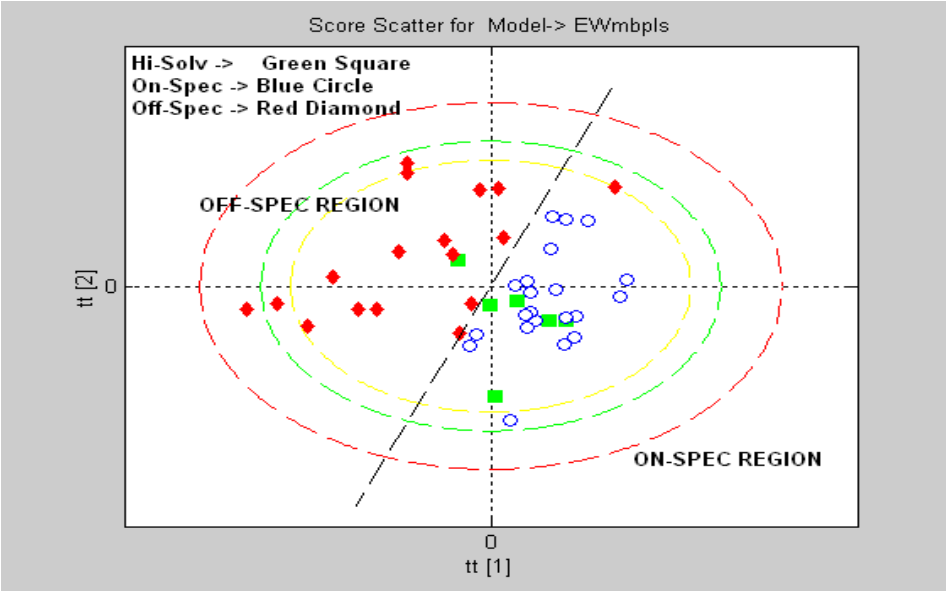
Multi-way PLS for batch data



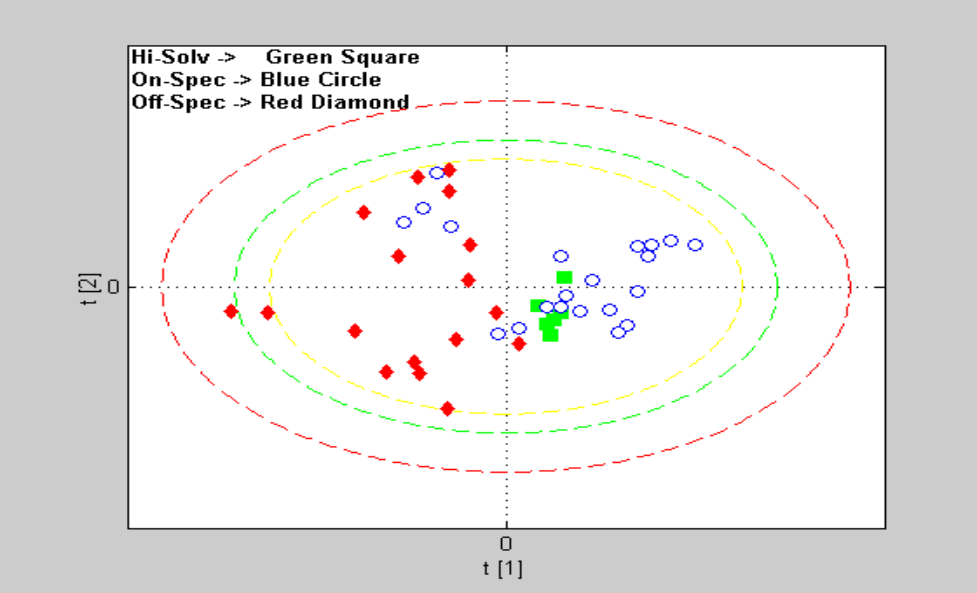
- Mean centering removes the average trajectories
- Models the time varying covariance structure among all the process variables over the entire time history of the batch
- Every batch summarized by a few LV scores (t_1, t_2, t_3)
- Relates the IC's (Z) and time varying trajectory information (X) to the final product quality (Y)

(c) 2004-2008, ProSensus, Inc.

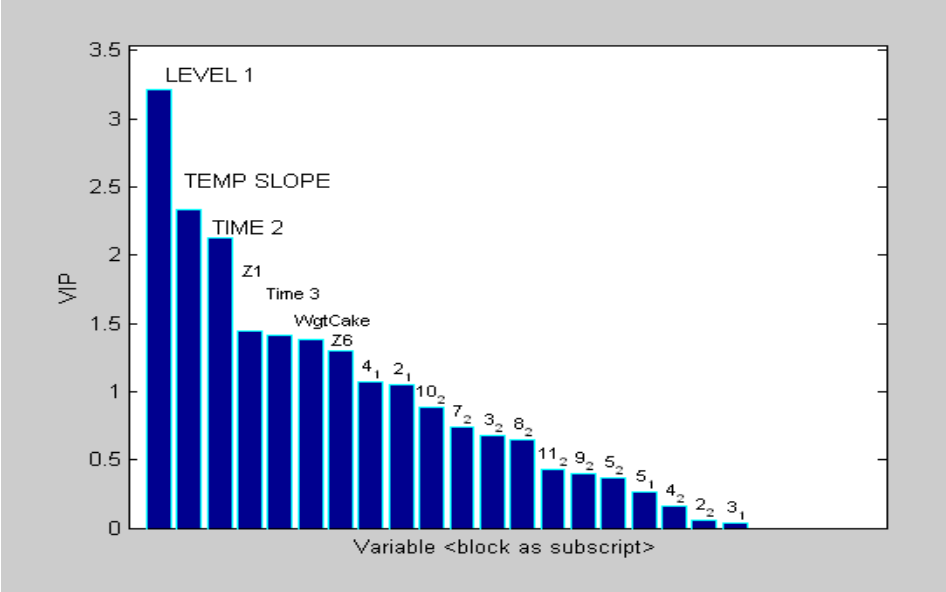
LV score plot for Z



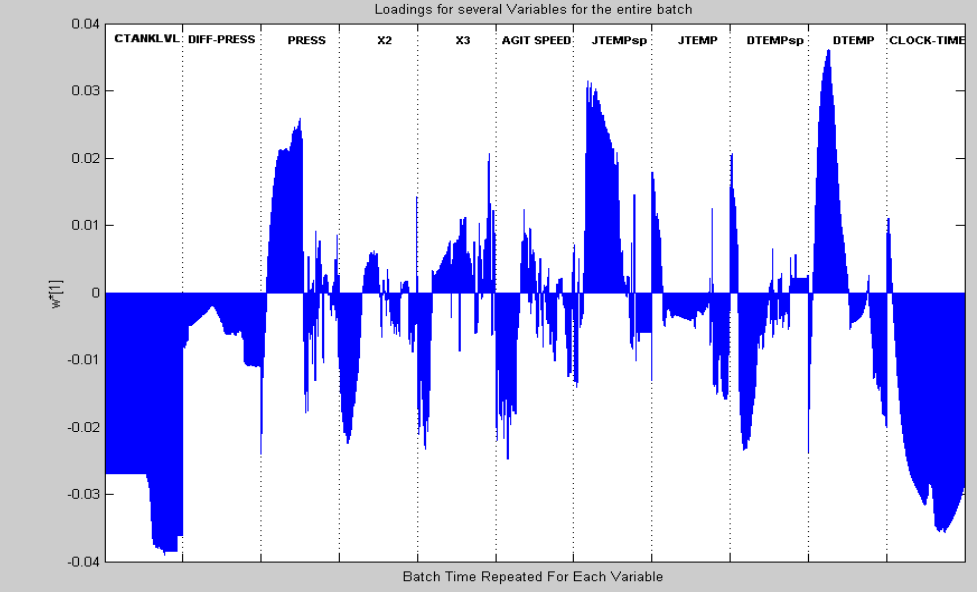
LV score plot for X



VIP's for Z model



Loading vector w^*_1 for X model



C. Industrial applications

- Analysis of historical data
- **Process monitoring**
- Inferential models / Soft sensors

Passive applications

- Optimization of process operation
- Control
- Scale-up and transfer between plants
- Rapid development of new products

Active applications

(c) 2004-2008, ProSensus, Inc.

On-line Monitoring of New Batches

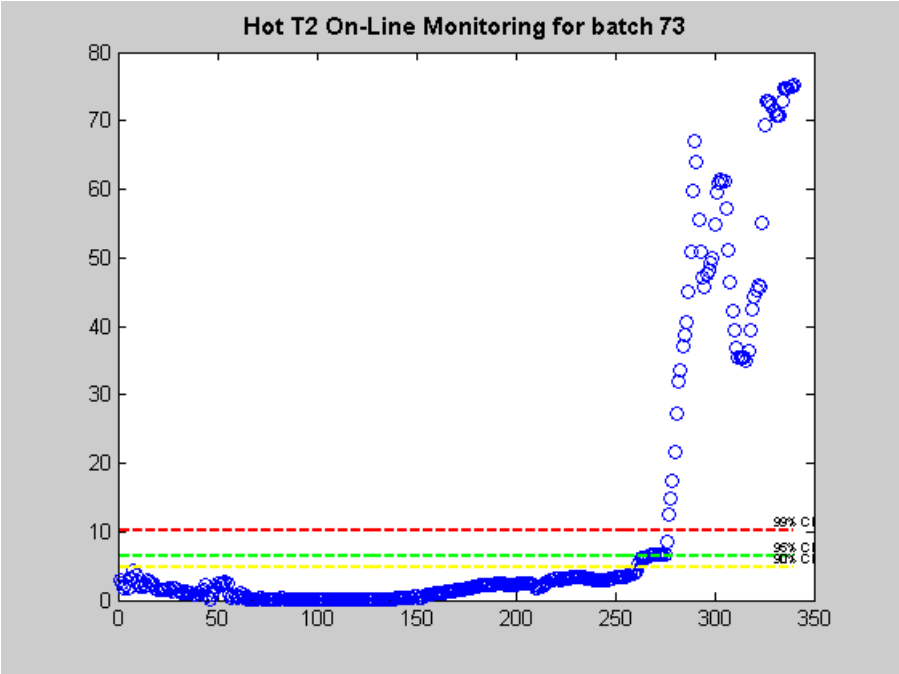
- Multivariate Statistical Process Control
  - Build LV model on all acceptable operational data
  - Statistical tests to see if new batches remain within that model space
    - Hotelling's T^2 shows movement within the LV plane
    - SPE shows movement off the plane

(c) 2004-2008, ProSensus, Inc.

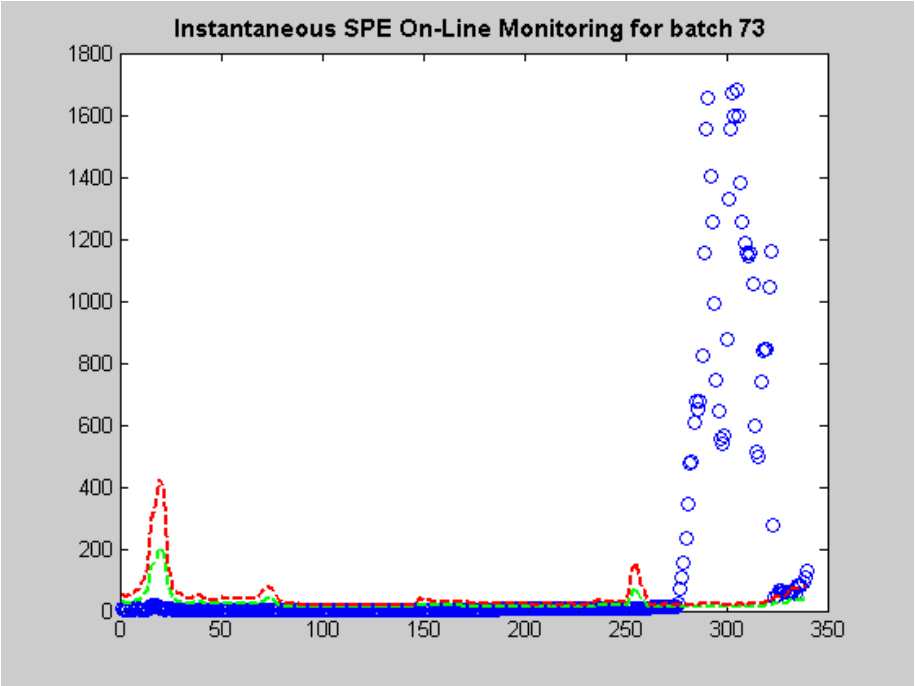
Process monitoring: Herbicide process

Monitoring of new batch number 73

T^2 plot

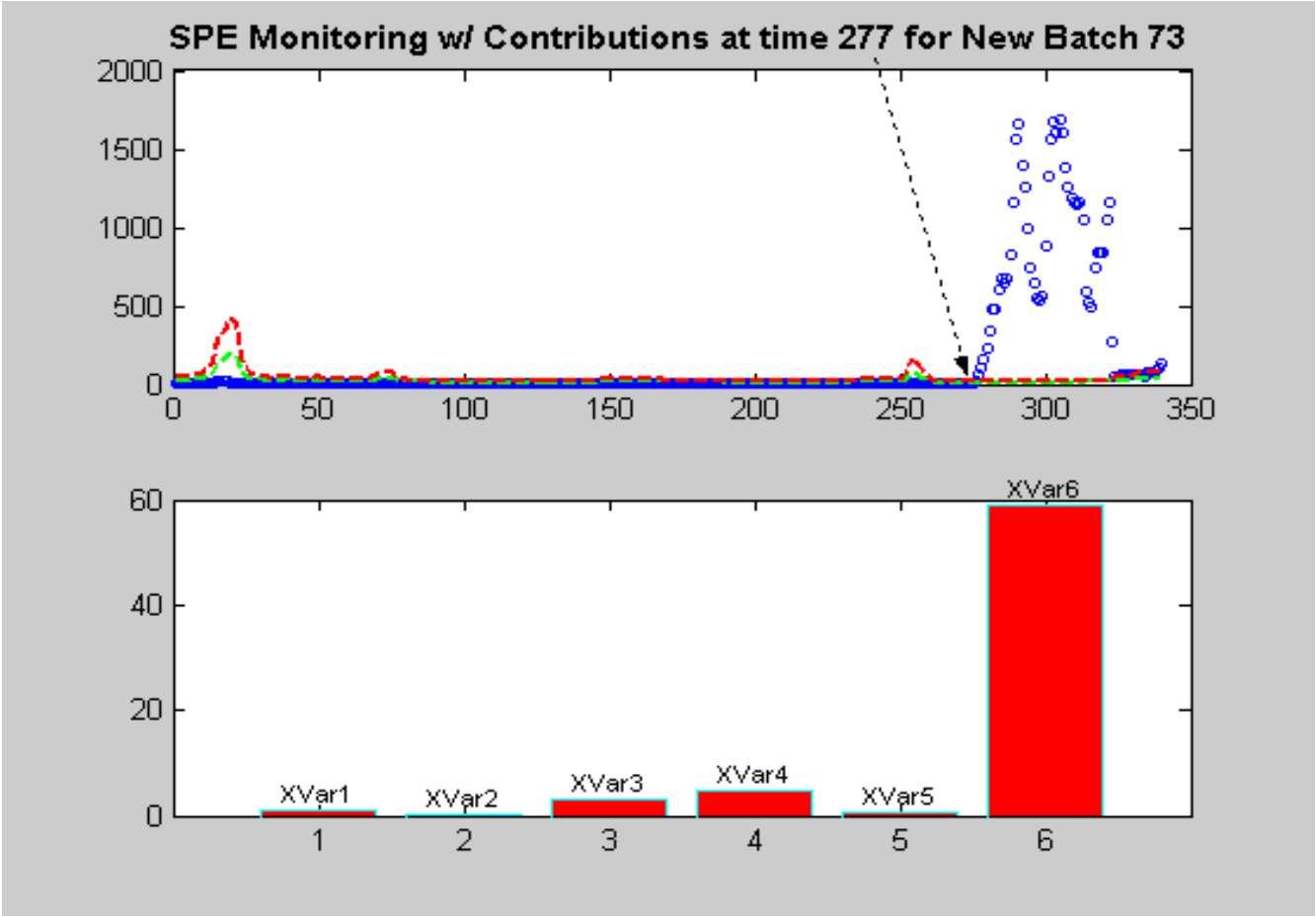


SPE plot



(c) 2004-2008, ProSensus, Inc.

Contribution plots to diagnose the problem



Problem: Variable x_6 diverged above its nominal trajectory at time 277

(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- Analysis of historical data
- Process monitoring
- **Inferential models / Soft sensors**

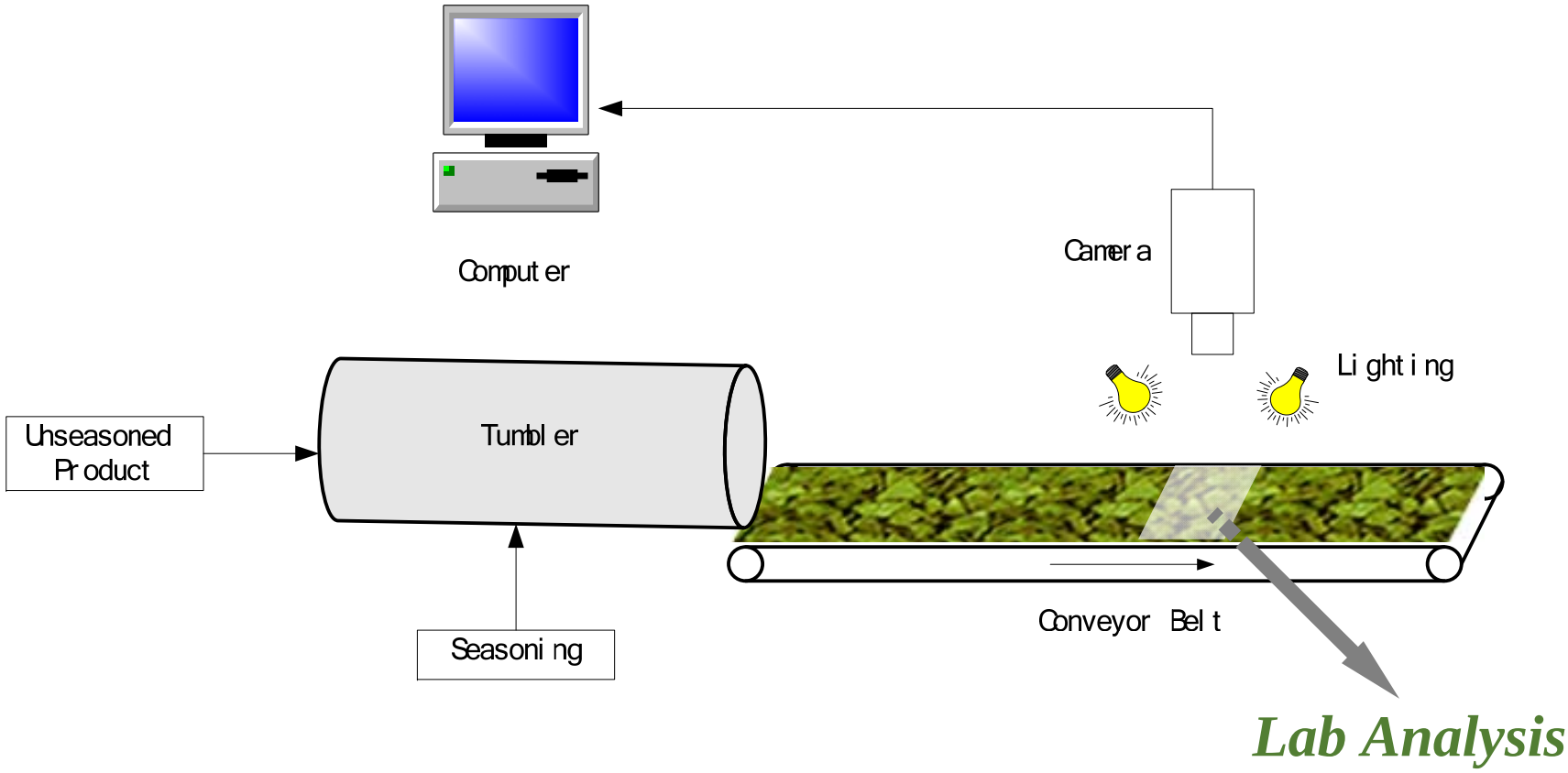
Passive applications

- Optimization of process operation
- Control
- Scale-up and transfer between plants
- Rapid development of new products

Active applications

(c) 2004-2008, ProSensus, Inc.

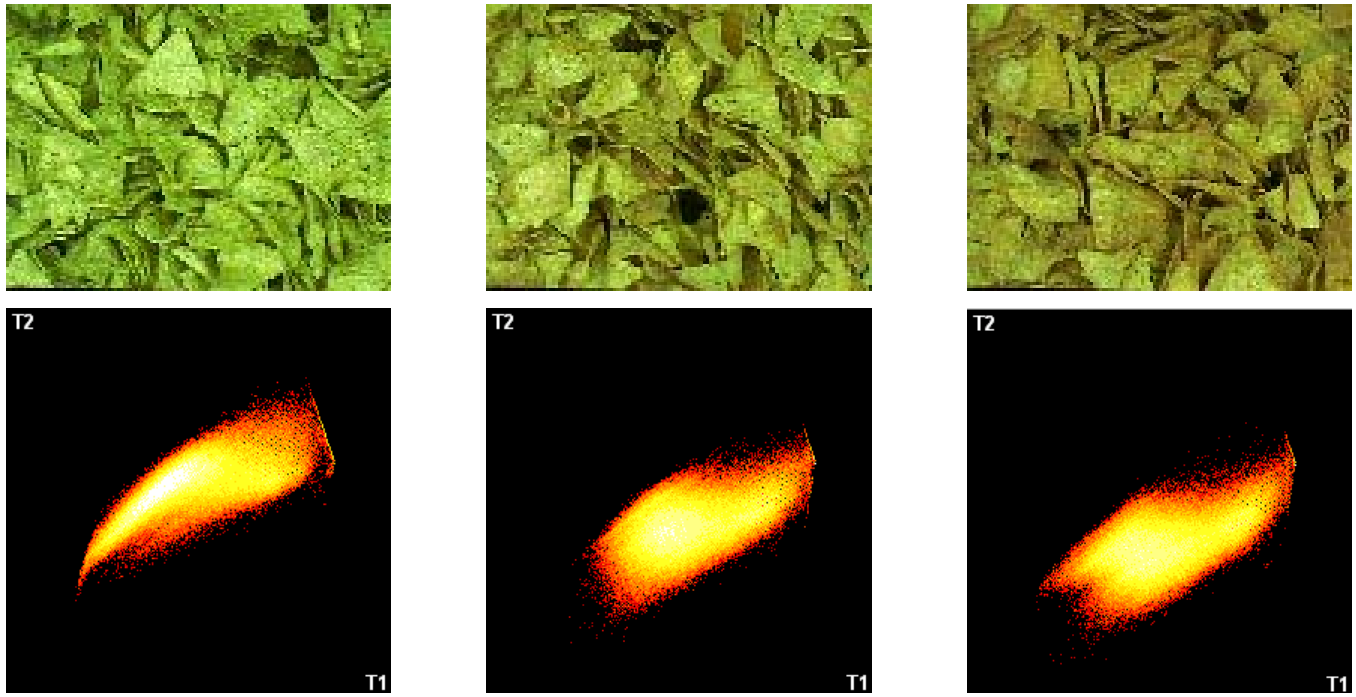
Image-based Soft Sensor for Monitoring and Feedback Control of Snack Food Quality



(c) 2007, ProSensus, Inc.

(c) 2004-2008, ProSensus, Inc.

PCA Score Plot Histograms



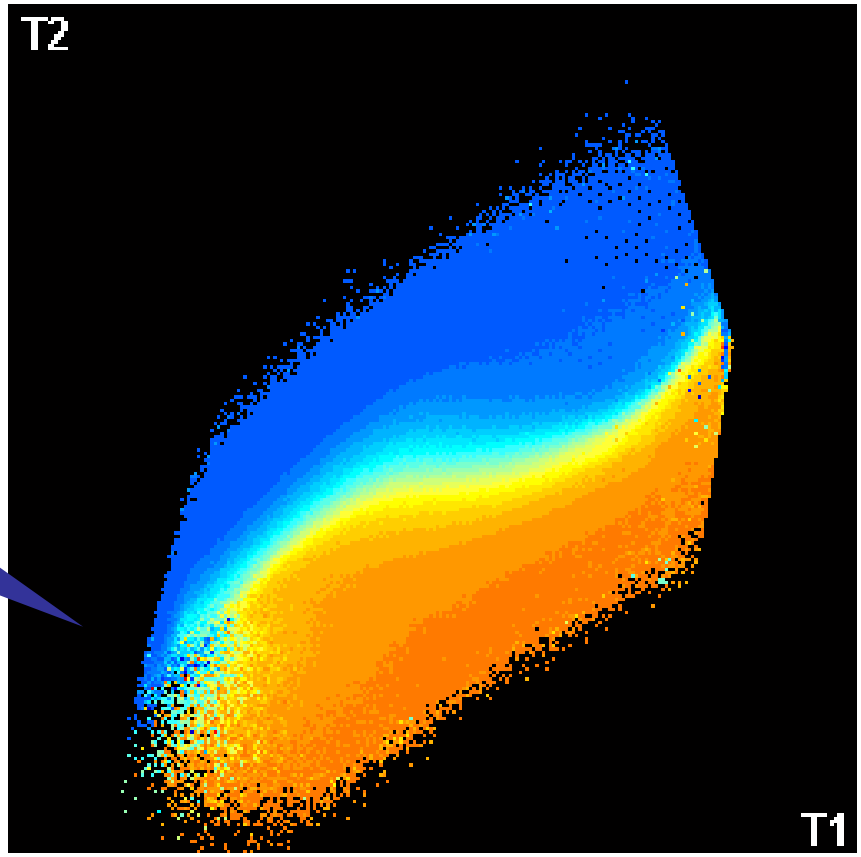
Non-seasoned **Low-seasoned** **High-seasoned**

(c) 2007, ProSensus, Inc.

(c) 2004-2008, ProSensus, Inc.

Segment Score Space into Multi-mask Region Based on Covariance with Quality

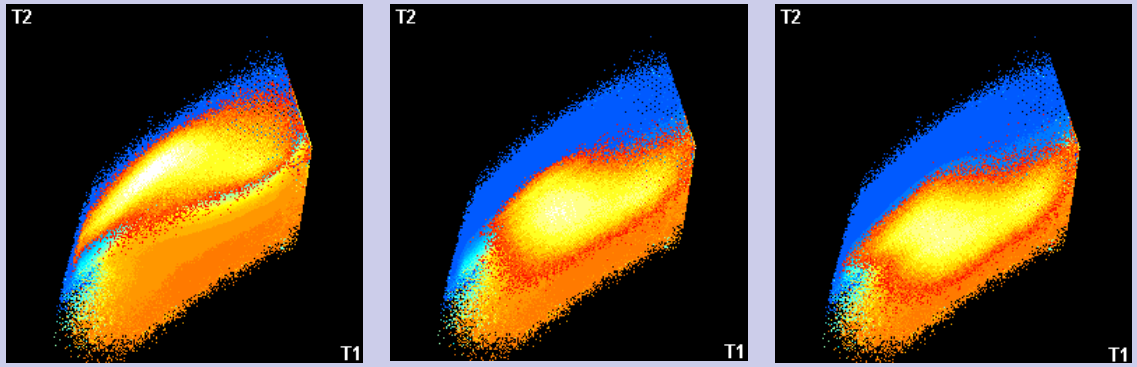
Score space divided up into 32 regions corresponding to various coating levels



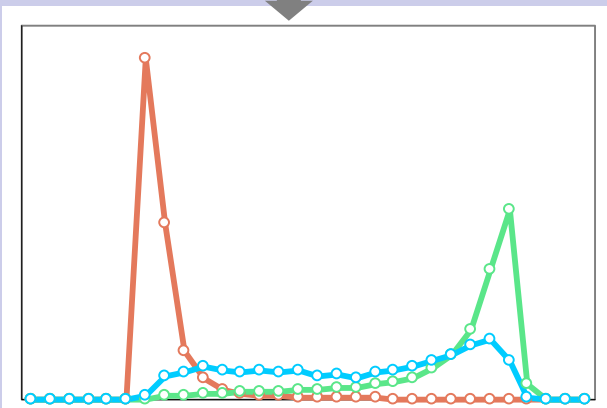
(c) 2007, ProSensus, Inc.

(c) 2004-2008, ProSensus, Inc.

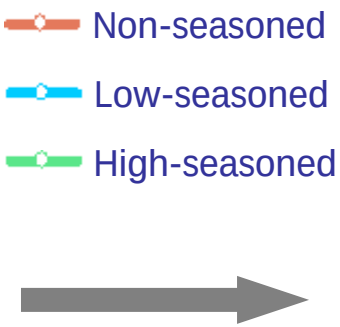
Distribution of Pixels

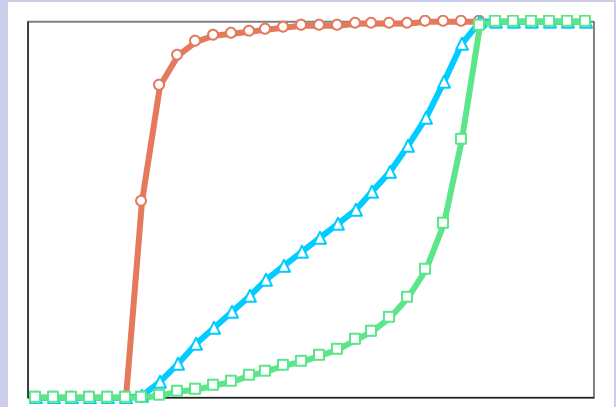


Superposition of the score plots from 3 sample images on top of the mask



Distribution of pixels

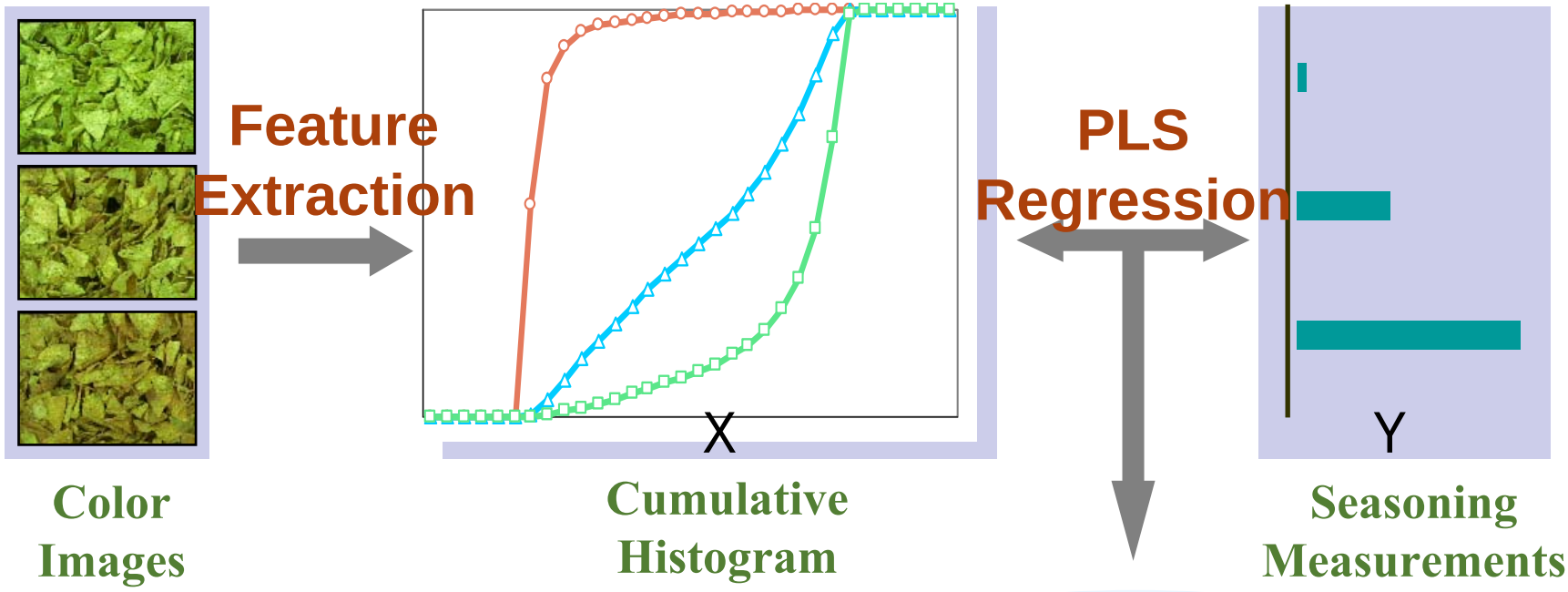




Cumulative distribution of pixels

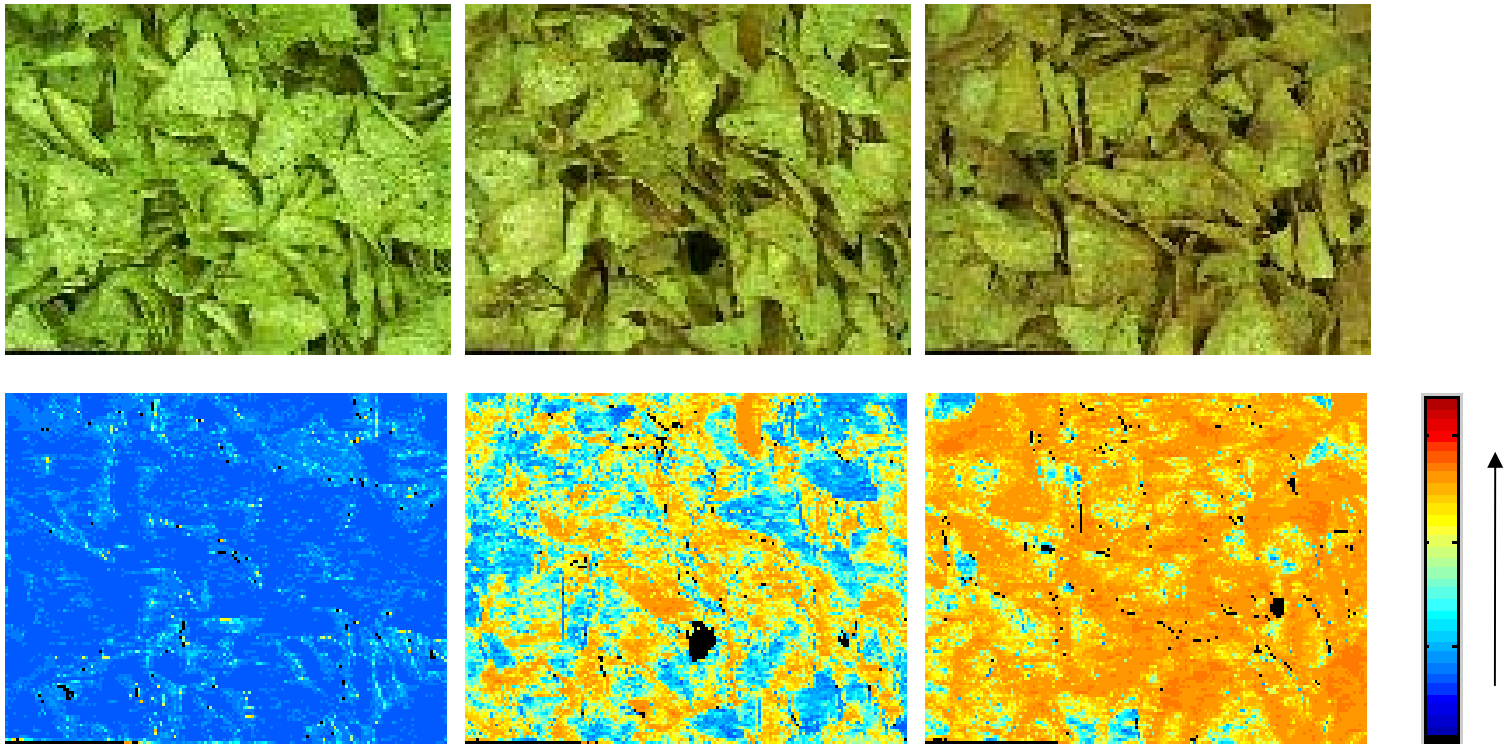
(c) 2004-2008, ProSensus, Inc.

Model to Predict Seasoning Concentration



(c) 2004-2008, ProSensus, Inc.

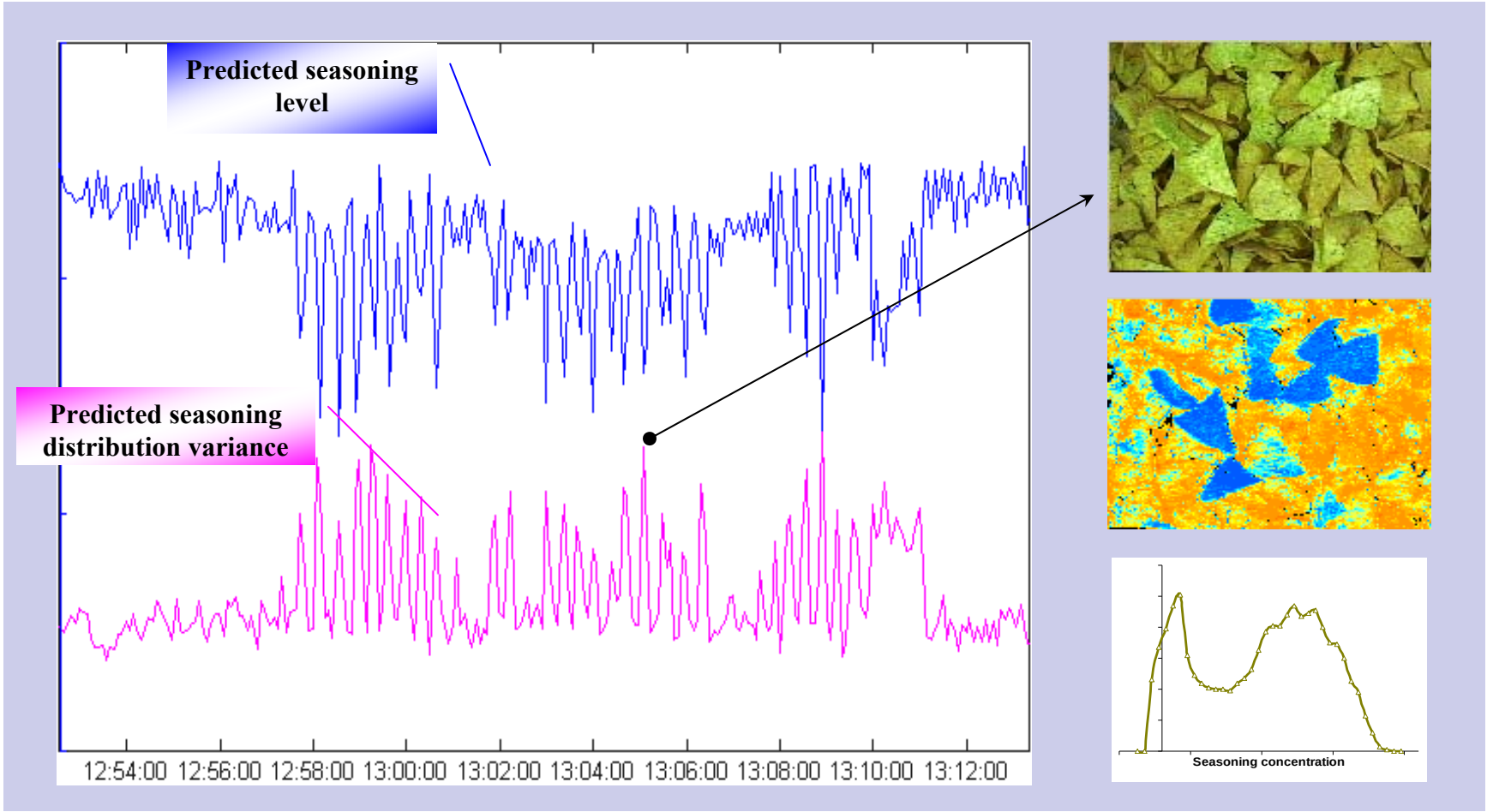
Visualize Images: Pixel by pixel Prediction of Seasoning Concentration



Non-seasoned Low-seasoned High-seasoned

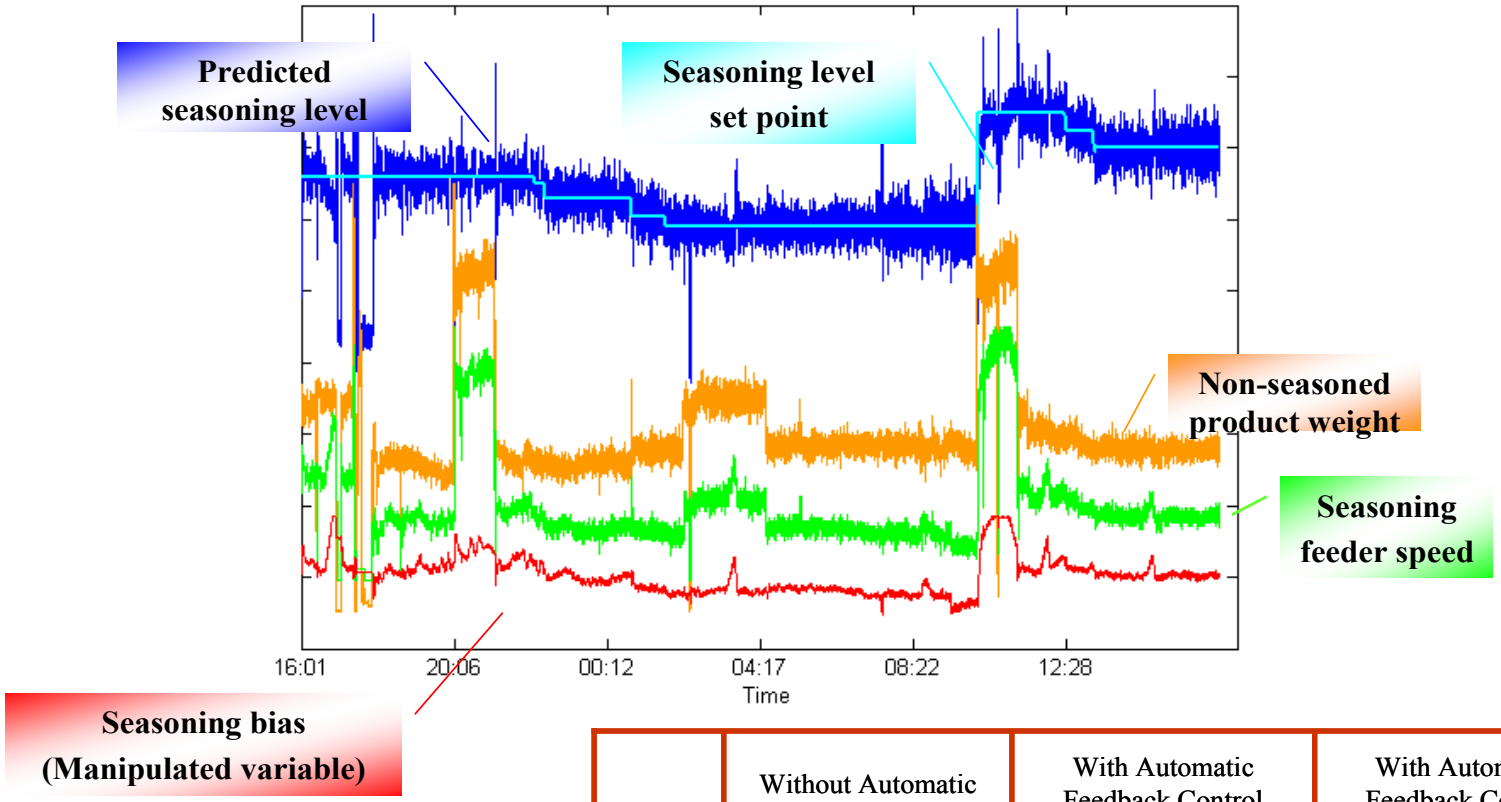
(c) 2004-2008, ProSensus, Inc.

Online Results: Mixed Product Experiment



(c) 2004-2008, ProSensus, Inc.

Closed-loop Control of Seasoning Level



	Without Automatic Feedback Control	With Automatic Feedback Control (constant set point)	With Automatic Feedback Control (set point tracking)
MAE	0.8523	0.4481	0.4769

(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- Analysis of historical data
- Process monitoring
- Inferential models / Soft sensors

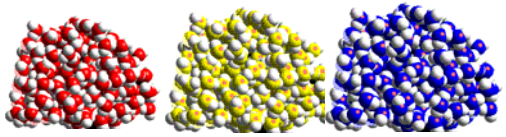
Passive applications

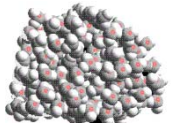
- **Optimization of process operation**
- Control
- Scale-up and transfer between plants
- Rapid development of new products

Active applications

(c) 2004-2008, ProSensus, Inc.

Optimizing operating policies for new products





X

Temperatures
Pressures
Concentrations
Recipes
Flows
Trajectories

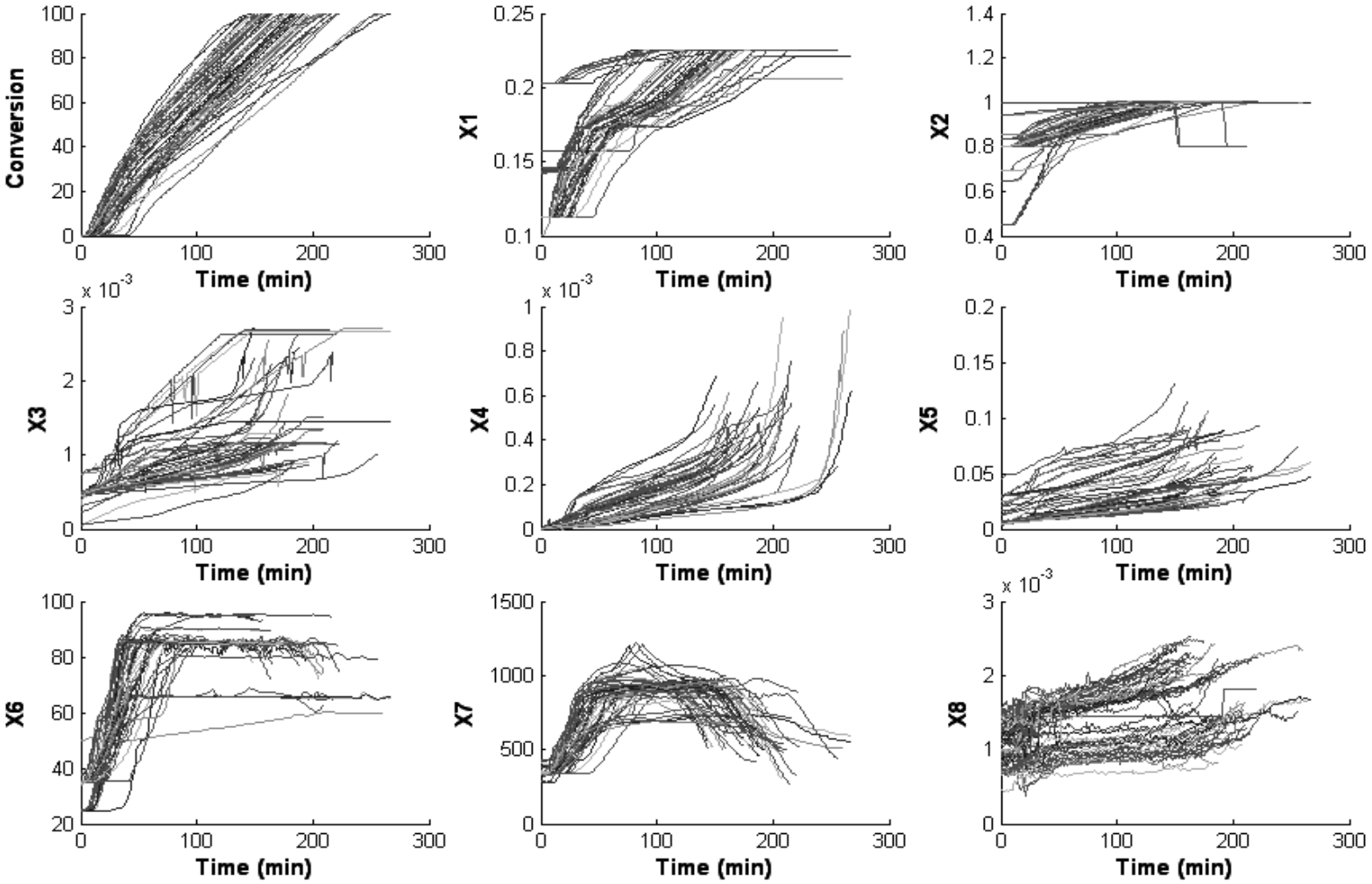
Y

Density
Tensile strength
Mw, Mn
Transparency
Biological activity
Toxicity
Hydrophobicity

(c) 2004-2008, ProSensus, Inc.

Batch polymerization: Process trajectory data (X)

Batch polymerization – Air Products & Chemicals



(c) 2004-2008, ProSensus, Inc.

Batch polymerization data

- 13 variables in Y

  Desire a new product with the following final quality attributes (Y's):

  Maintain in normal ranges: Y_1 Y_2 Y_3 Y_4 Y_5 Y_6 Y_8

  Constraints:

  $Y_7 = Y_{7des}$

  $Y_9 = Y_{9des}$

  $Y_{10} < Y_{10const}$

  $Y_{11} < Y_{11const}$

  $Y_{12} < Y_{12const}$

  $Y_{13} < Y_{13const}$

  … and with the minimal possible batch time (*)

- Solution
  - Build batch PLS latent variable model on existing data
  - Perform an optimization in LV space to find optimal LV's
  - Use LV model of X-space to find the corresponding recipes and process trajectories

(c) 2004-2008, ProSensus, Inc.

Unconstrained Solution

- Design via PLS model inversion (no constraints)

PLS Model:

$$\hat{\mathbf{Y}} = \mathbf{T}\mathbf{Q}^{\mathbf{T}}$$

$$\hat{\mathbf{y}}_{\mathbf{des}} = \mathbf{Q}\,\tau_{\mathbf{new}}$$

Step 1

$$\tau_{new} = inv\left(Q^T Q\right) Q^T y_{des}$$

$$\hat{\mathbf{X}} = \mathbf{T}\mathbf{P}^{\mathbf{T}}$$

$$\hat{\mathbf{x}}_{\mathbf{new}} = \mathbf{P}\,\tau_{\mathbf{new}}$$

Step 2

- If dim(Y) < dim(X) then there is a null space
  - A whole line or plane of equivalent solutions yielding the same y_{des}

(c) 2004-2008, ProSensus, Inc.

Solution with constraints: Formulate inversion as an optimization

- **Step 1**: Solve for $\hat{\boldsymbol{\tau}}_{new}$ with constraints on T^2 and on y's

$$\min_{\hat{\boldsymbol{\tau}}_{xnew}} \left\{ (\mathbf{y}_{des} - \mathbf{Q}\,\hat{\boldsymbol{\tau}}_{xnew})^{T} \mathbf{G}_1 (\mathbf{y}_{des} - \mathbf{Q}\,\hat{\boldsymbol{\tau}}_{xnew}) + \rho \left(\sum_{a=1}^{A} \frac{\hat{\tau}^2_{xnew,a}}{s_a^2} \right) \right\}$$

$$s.t$$

$$\mathbf{B}\,\mathbf{Q}\,\hat{\boldsymbol{\tau}}_{\mathbf{xnew}} < \mathbf{b}$$

- **Step 2**: Solve for x_{new} that yields $\hat{\boldsymbol{\tau}}_{new}$ subject to certain constraints on SPE and x's.

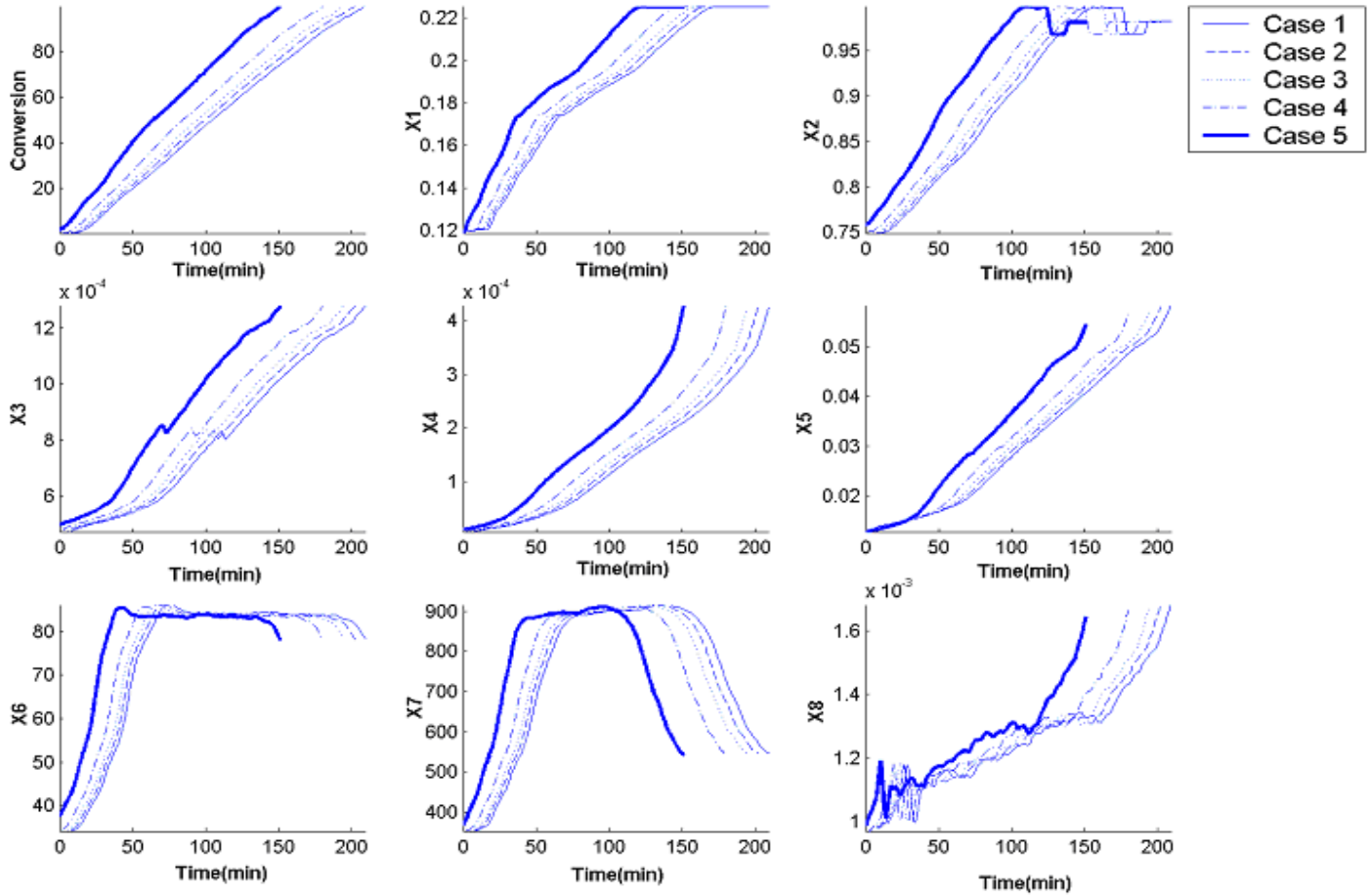
$$\min_{\mathbf{x}_{new}} \left\{ \left(W^{*}\mathbf{x}_{new} - \boldsymbol{\tau}_{new} \right)^{T} \mathbf{G}_2 \left(W^{*}\mathbf{x}_{new} - \boldsymbol{\tau}_{new} \right) + \left(\mathbf{x}_{new} - PW^{*}\mathbf{x}_{new} \right)^{T} \boldsymbol{\Lambda} \left(\mathbf{x}_{new} - PW^{*}\mathbf{x}_{new} \right) + \boldsymbol{\eta}\mathbf{x}_{new} \right\}$$

(c) 2004-2008, ProSensus, Inc.

Different solutions: change the penalty (η) on time usage

All solutions satisfy the requirements on y_{des}

Case 1 to 5: weight on time-usage is gradually increased



Garcia-Munoz, S., J.F. MacGregor, D. Neogi, B.E. Latshaw and S. Mehta, "Optimization of batch operating policies. Part II: Incorporating process constraints and industrial applications", Ind. & Eng. Chem. Res., 2008

(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- Analysis of historical data
- Process monitoring
- Inferential models / Soft sensors

Passive applications

- Optimization of process operation
- **Control**
- Scale-up and transfer between plants
- Rapid development of new products

Active applications

(c) 2004-2008, ProSensus, Inc.

Control of batch product quality

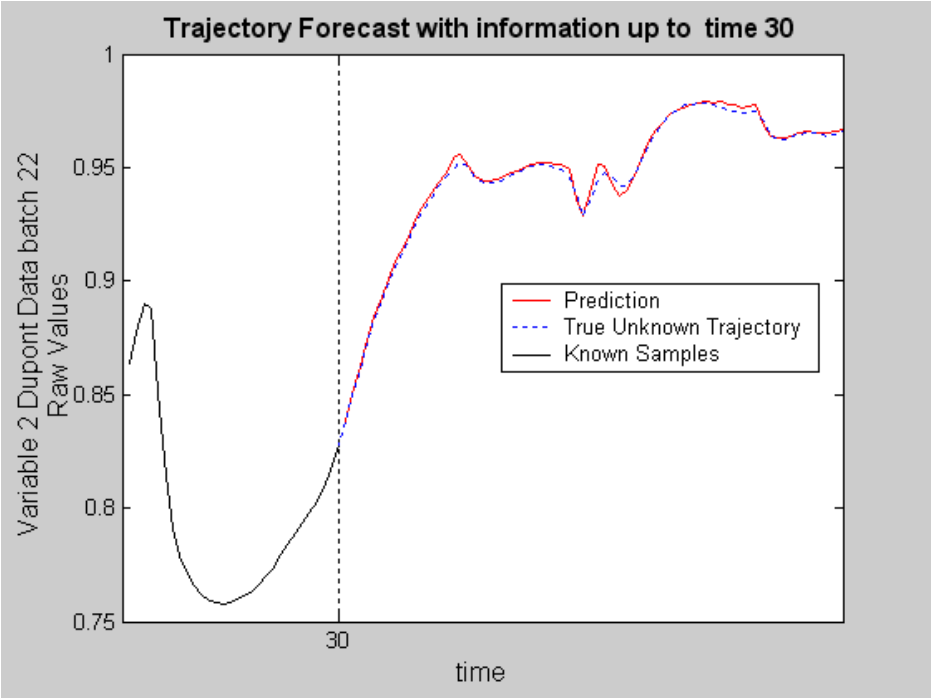
- Objective is to control final product quality
  - e.g. control of final particle size distribution (PSD)
- Using all data up to some decision time, predict final quality with latent variable model
  - All prediction done in low dimensional latent variable space (y's then calculated from t's)
- If predicted quality is outside a desired window, then make a mid-course correction to the batch
  - Analogy to NASA mid-course rocket trajectory adjustment in moon missions
- Data requirement: Historical batches + few with DOE on corrective variables

(c) 2004-2008, ProSensus, Inc.

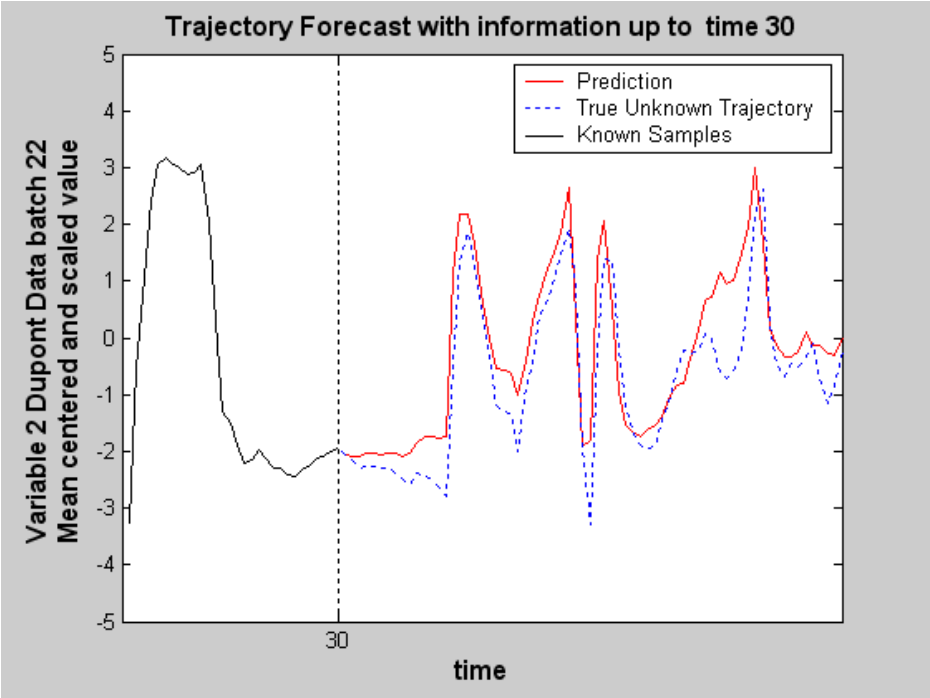
LV models provide accurate adaptive trajectory predictions

- **Use various missing data imputation methods**
  - Equivalent of Kalman Filter

Prediction of a variable trajectory using information up to time 30 (DuPont)



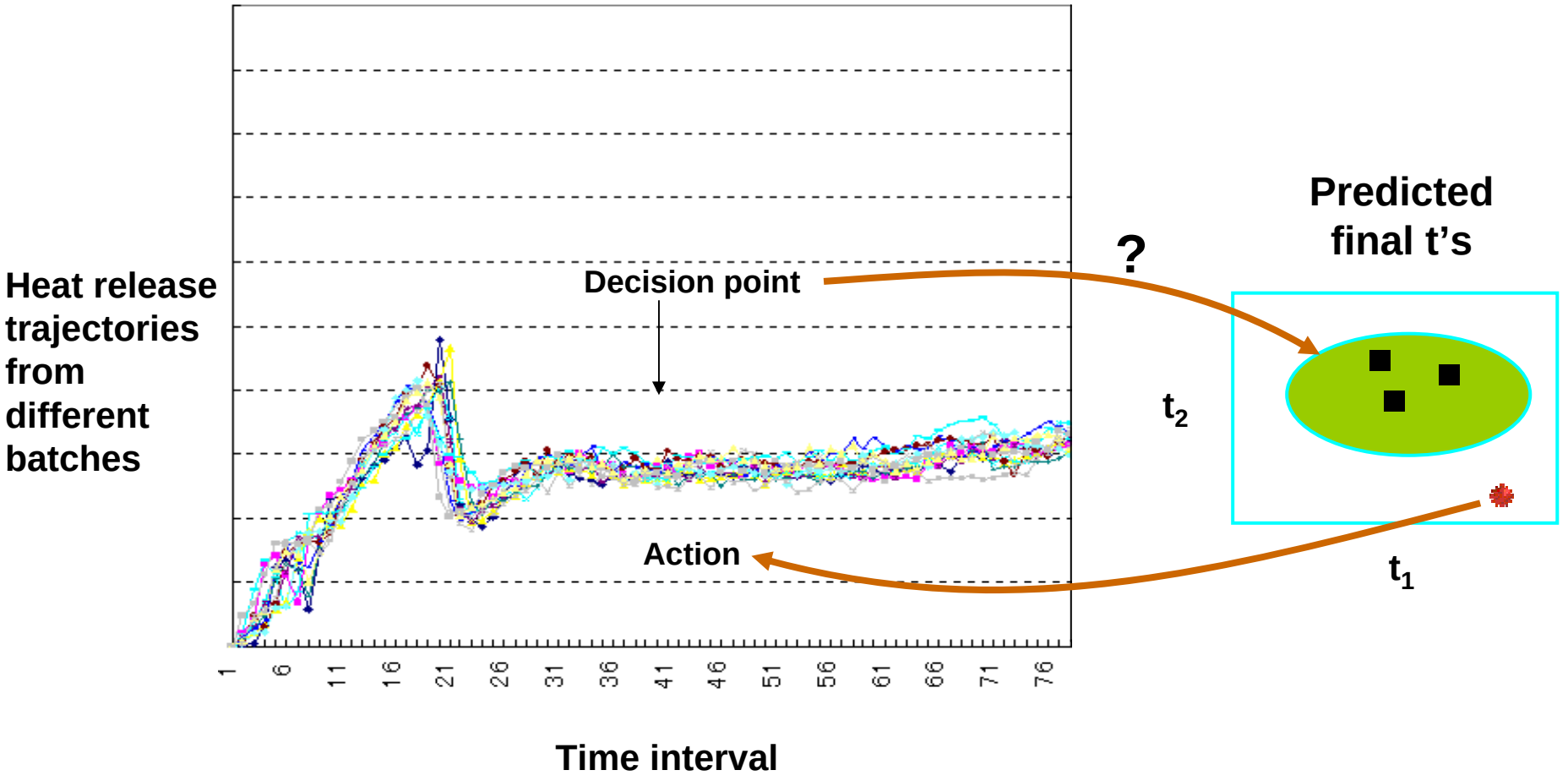
Deviations from the mean trajectory – Prediction vs actual



(c) 2004-2008, ProSensus, Inc.

Control of PSD via mid-course correction

- At decision point – predict t's (Y's) – if outside target region – take action

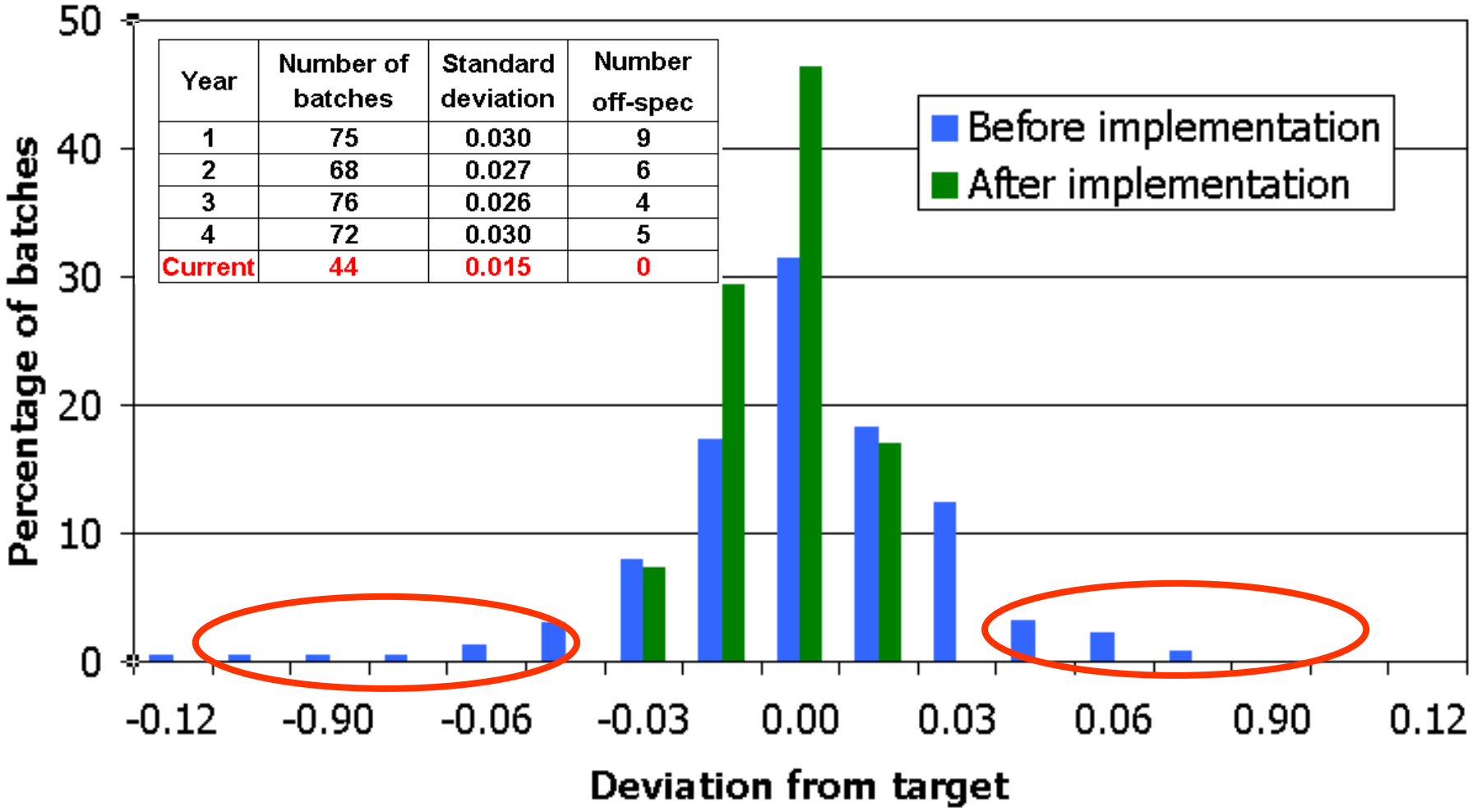


(c) 2004-2008, ProSensus, Inc.

Industrial results (Mitsubishi Chemicals)

Mid-course control : before and after implementation

Year	Number of batches	Standard deviation	Number off-spec
1	75	0.030	9
2	68	0.027	6
3	76	0.026	4
4	72	0.030	5
Current	44	0.015	0



(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- Analysis of historical data
- Process monitoring
- Inferential models / Soft sensors

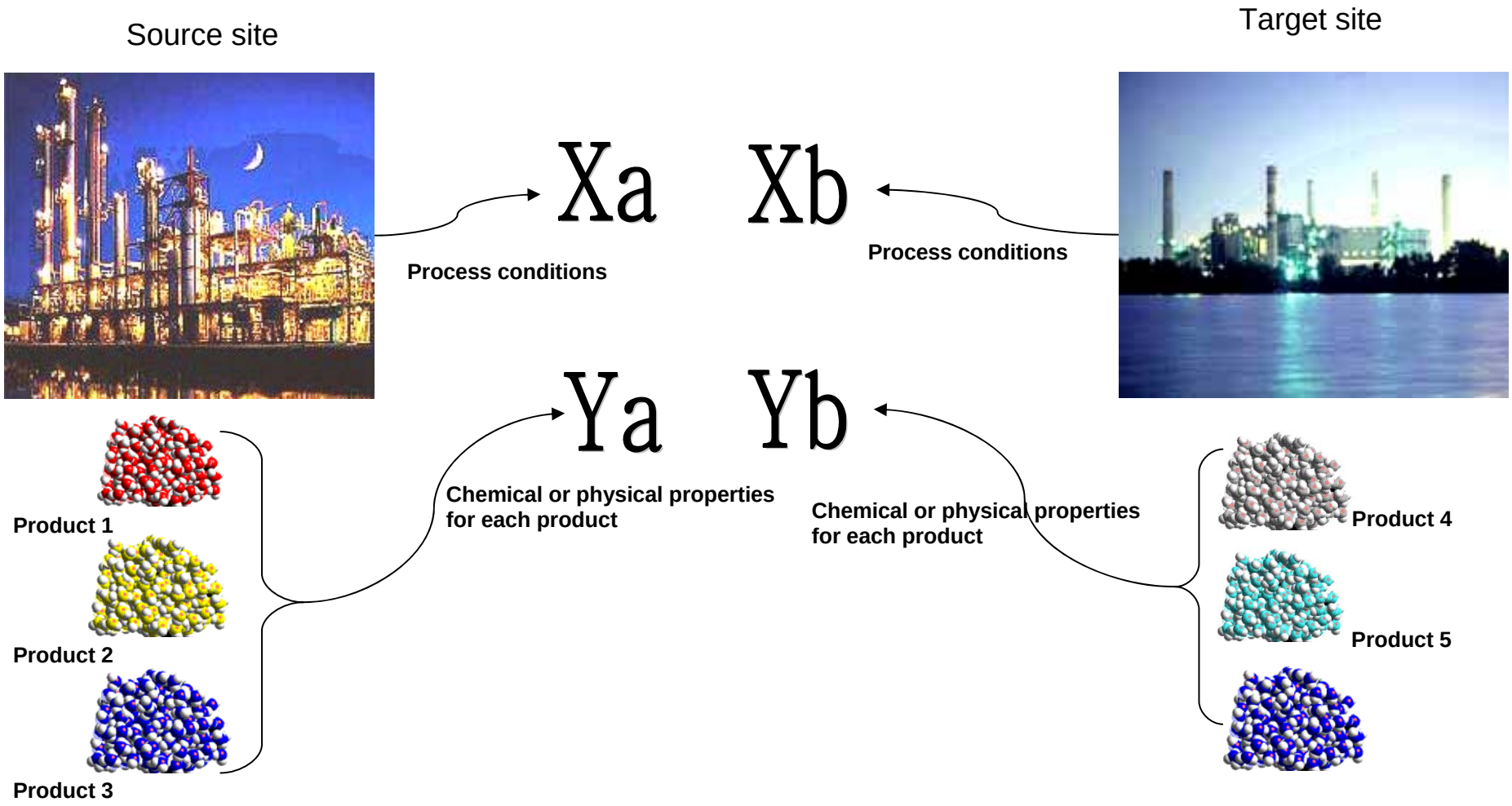
Passive applications

- Optimization of process operation
- Control
- **Scale-up and transfer between plants**
- Rapid development of new products

Active applications

(c) 2004-2008, ProSensus, Inc.

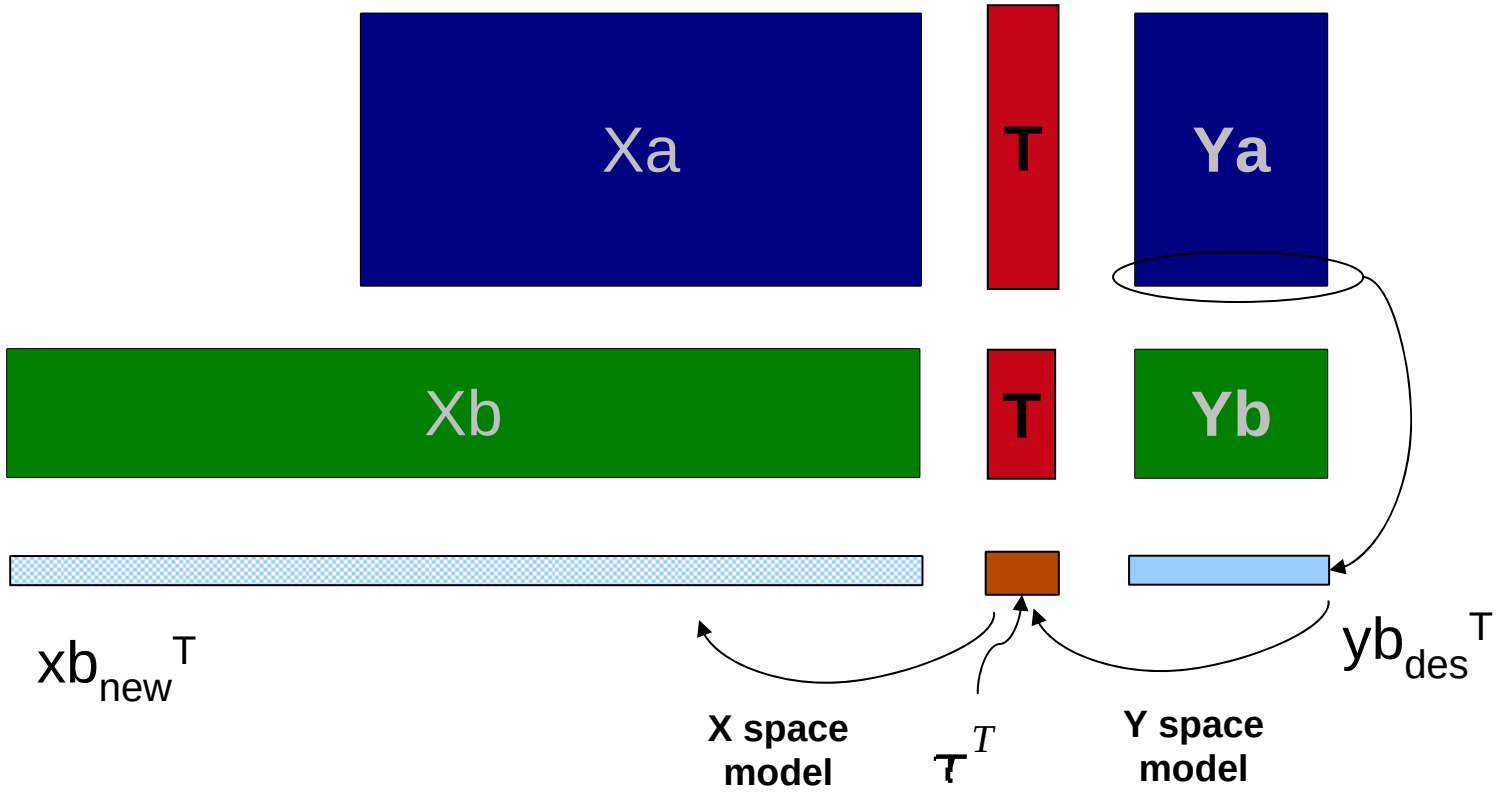
Product transfer between plants and scale-up



(c) 2004-2008, ProSensus, Inc.

Product transfer and scale-up

Historical data from the 2 plants. Build JYPLS model



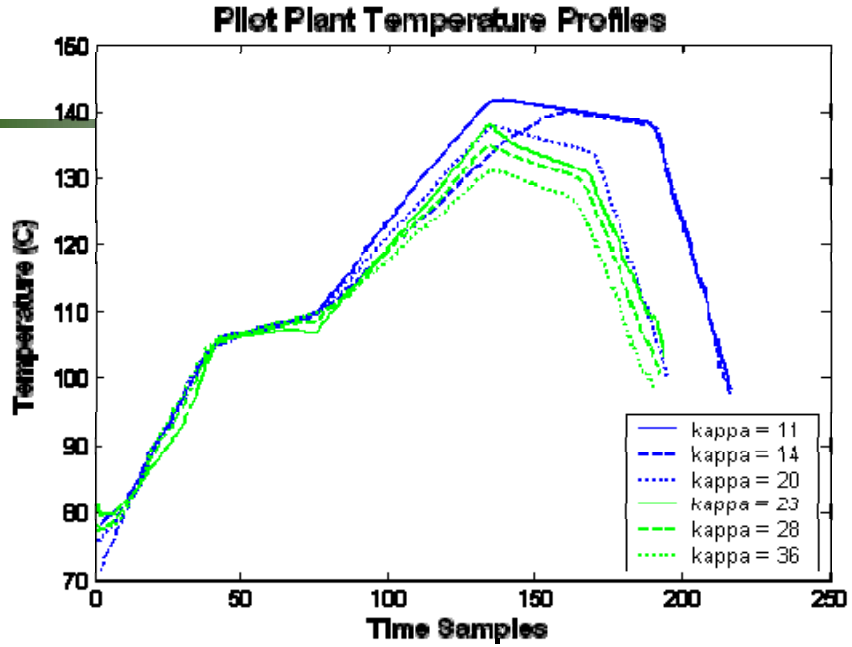
Garcia-Munoz, S., T.Kourti and J.F. MacGregor, "Product Transfer Between Sites using Joint-Y PLS", Chemometrics & Intell. Lab. Systems, 79, 101-114, 2005.

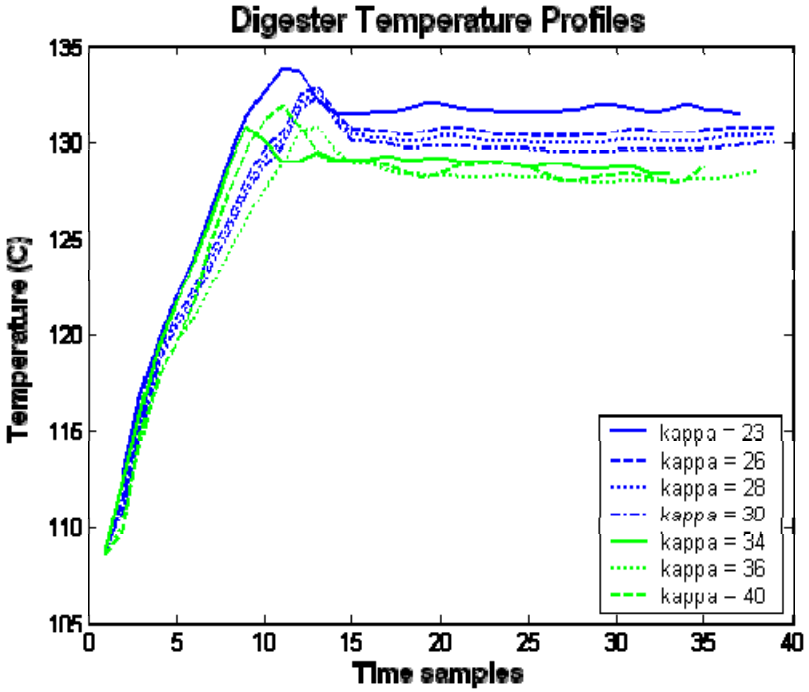
(c) 2004-2008, ProSensus, Inc.

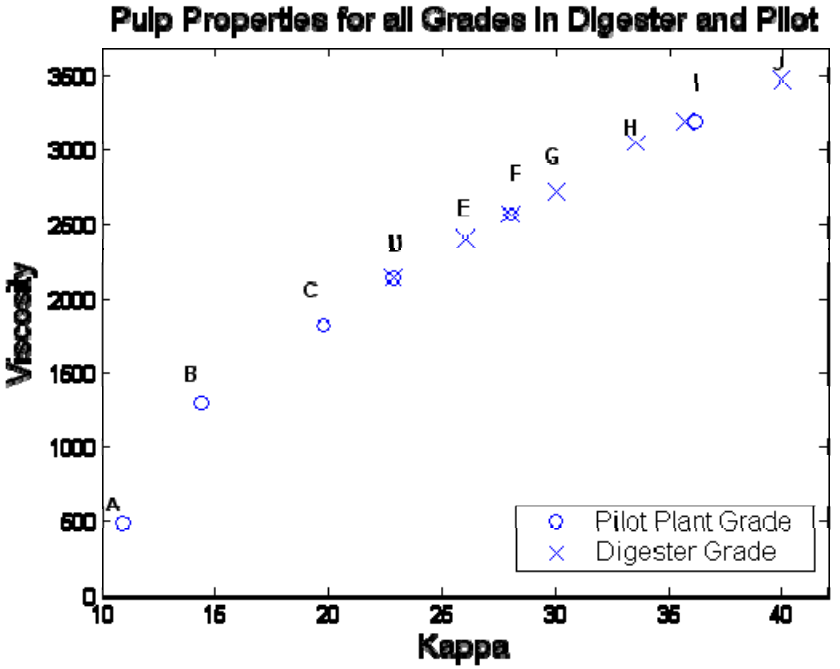
Industrial Scale-up Example

Tembec - Cdn. pulp & paper company:

Pilot plant and full scale digesters





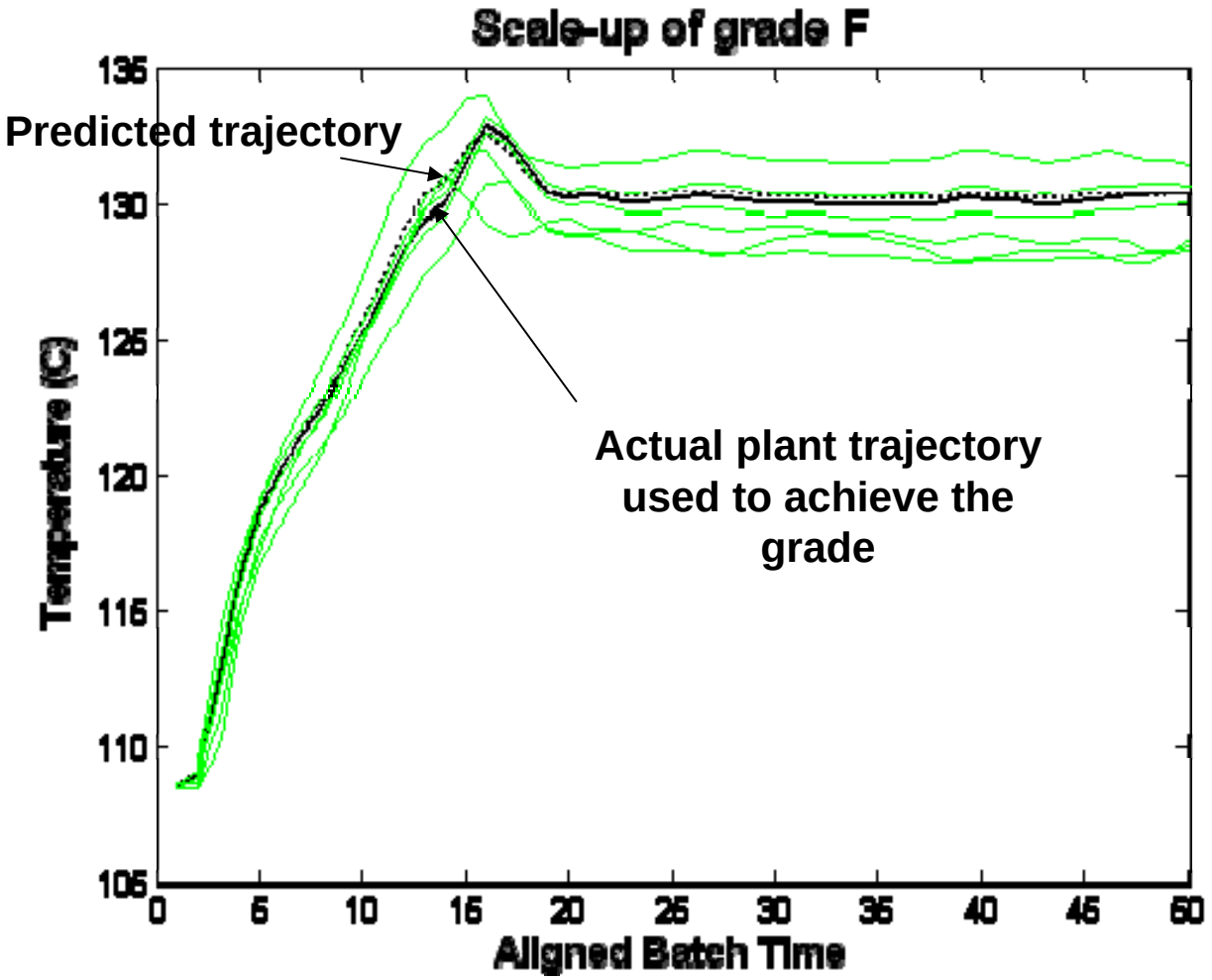


(c) 2004-2008, ProSensus, Inc.

Scale up for grade F – pulp digester

Build models on all pilot plant data and all plant data (ex F)

Design operating profiles to achieve grade F in plant.



(c) 2004-2008, ProSensus, Inc.

C. Industrial applications

- Analysis of historical data
- Process monitoring
- Inferential models / Soft sensors

Passive applications

- Optimization of process operation
- Control
- Scale-up and transfer between plants
- **Rapid development of new products**

Active applications

(c) 2004-2008, ProSensus, Inc.

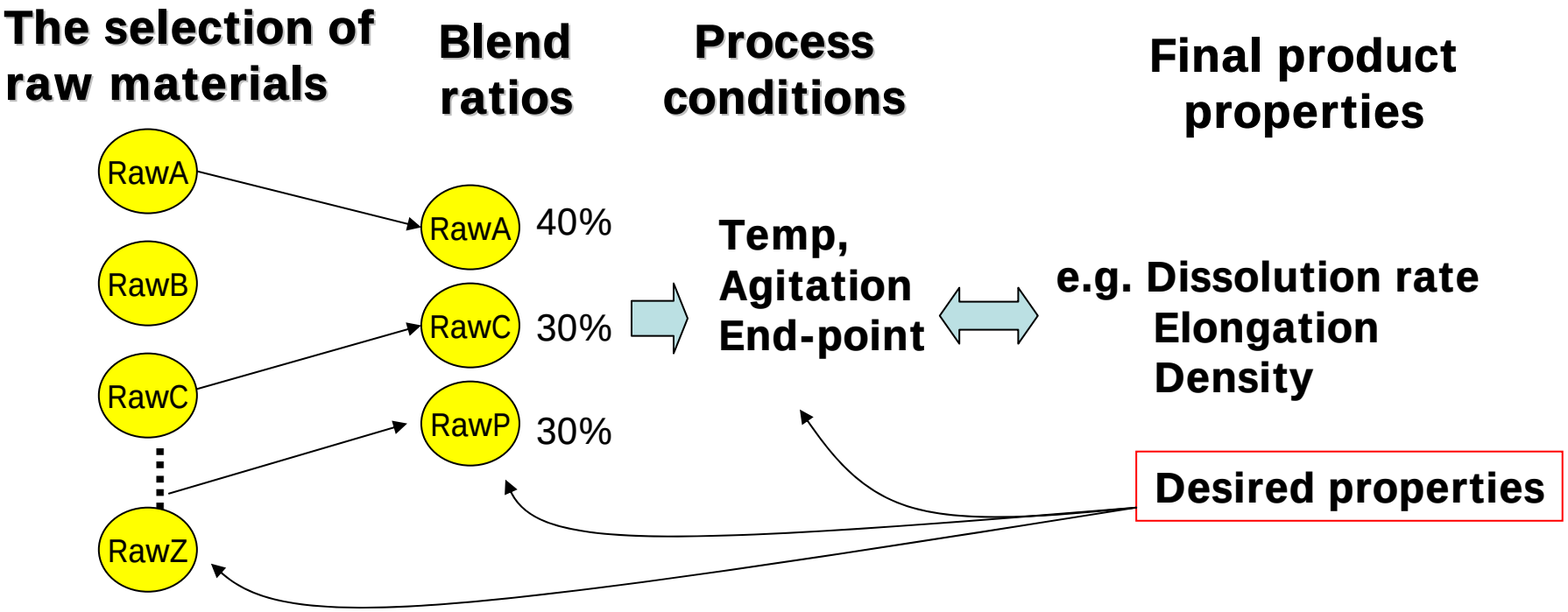
Data Mining for Product Engineering

- Companies accumulate lot of data on their products and processes
- Can we use it to rapidly develop new products?
- Three general degrees of freedom for developing new products:
  - Raw material selection
  - Ratios in which to use raw materials (formulation)
  - Process conditions for manufacturing

    - Relative importance of these three depends on the industry and the product
    - Huge synergisms among these

(c) 2004-2008, ProSensus, Inc.

What is the problem ?

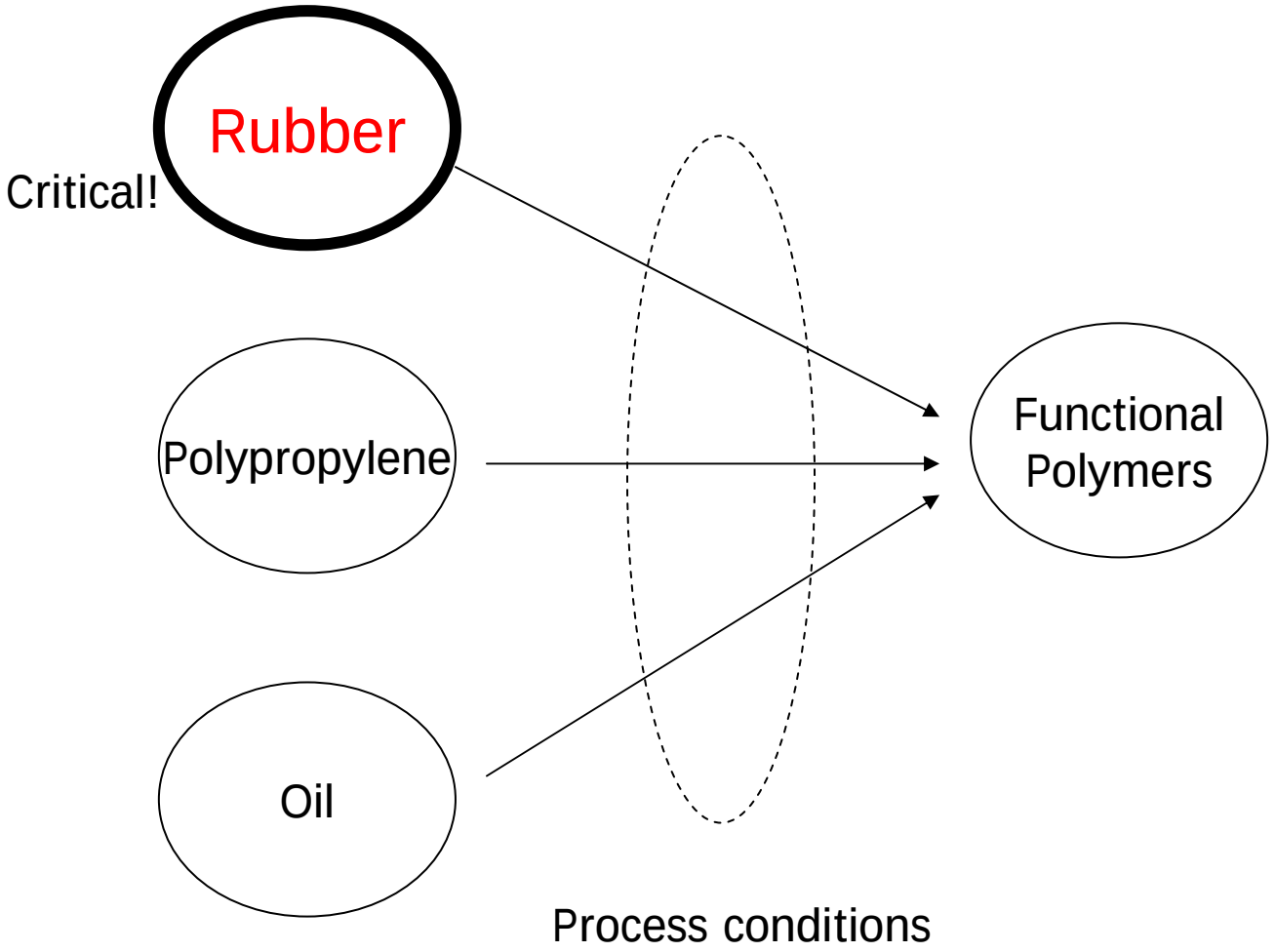


Traditional approaches tend to treat each step separately➔ inefficient as they miss synergism among these degrees of freedom

(c) 2004-2008, ProSensus, Inc.

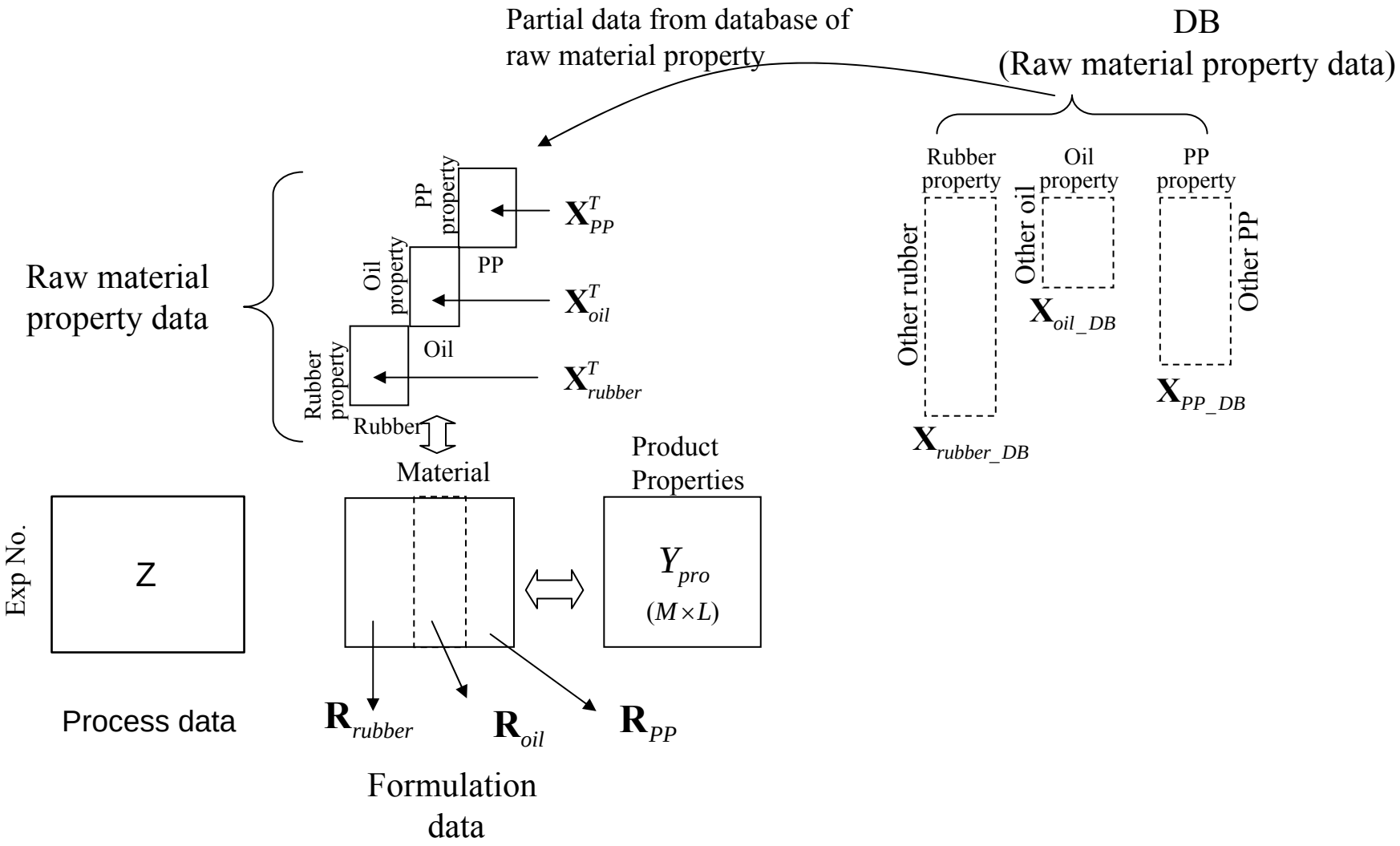
Example: Functional Polymer Development

Mitsubishi Chemicals



(c) 2004-2008, ProSensus, Inc.

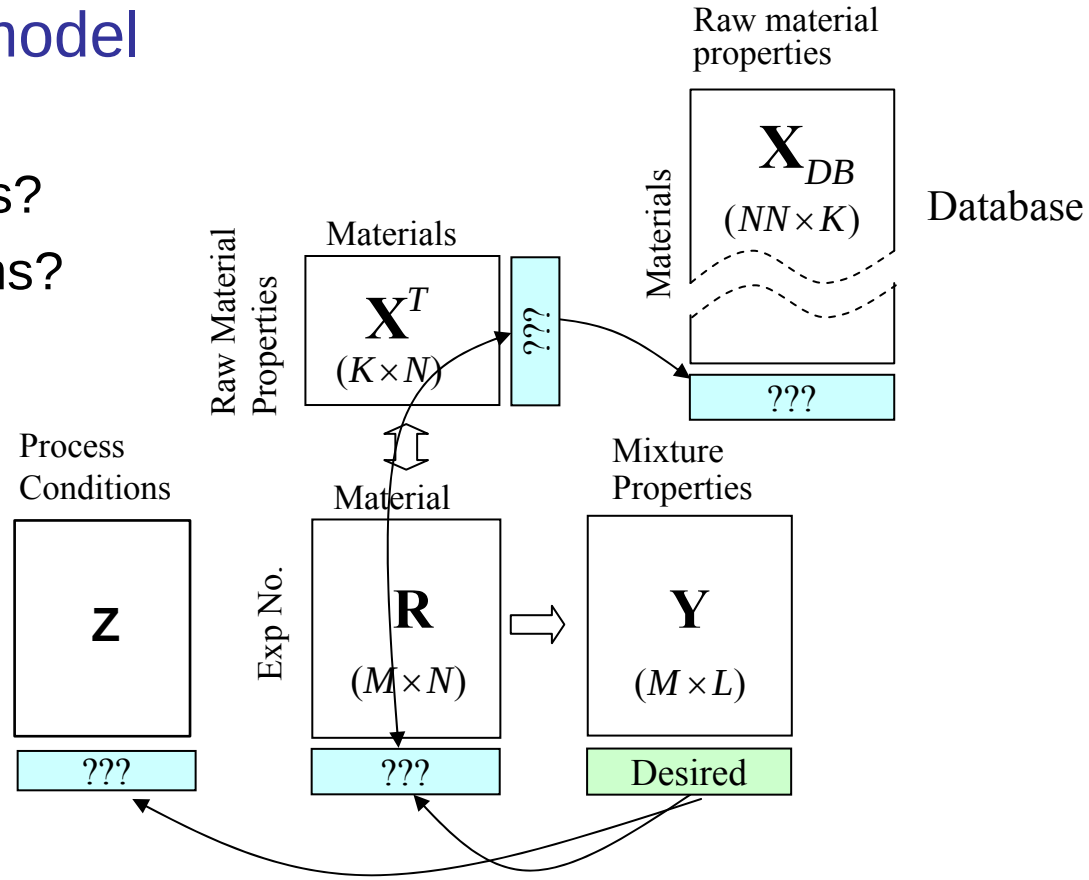
Data structure



(c) 2004-2008, ProSensus, Inc.

Methodology

- Build a multi-block PLS model that relates all the databases together and predicts the final quality attributes
- Perform an optimization in the latent variable space of the multi-block PLS model
  - Which materials?
  - Formulation ratios?
  - Process conditions?
  - Minimum cost



(c) 2004-2008, ProSensus, Inc.

Formulation of the Optimization

Estimation error | Total material cost | The number of materials

$$\underset{r_{new}}{Min} \quad (y_{des} - x_{mix\,new} \cdot B_{PLS})^T \cdot W_1 \cdot (y_{des} - x_{mix\,new} \cdot B_{PLS}) + w_2 \cdot \sum_{j=1}^{NN} r_{new,j} \cdot c_j + w_3 \cdot \sum_{j=1}^{NN} \delta_j$$

s.t.

Ideal mixing rule

$$x_{mix\,new} = r_{new} \cdot \mathbf{X}_{DB}$$

PLS model constraint

$$SPE_{new} = \sum_{k=1}^{k=K} (x_{mix\,new} - \hat{x}_{mix\,new})^2 \cong 0$$

$$T_{new}^{\ 2} = \sum_{a=1}^{A} \frac{\tau_{new,a}^2}{s_a} \le const$$

Mixture constraint

$$\sum_{j=1}^{NN} r_{new,j} = 1, \quad 0 \le r_{new,j} \le 1$$

Binary variable constraint

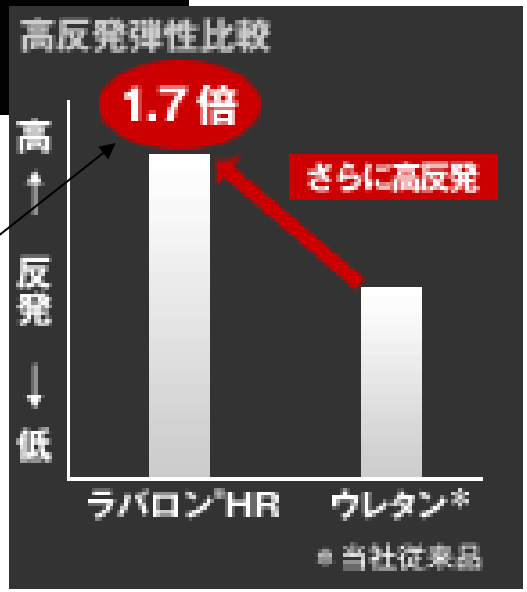
$$\delta_j = \begin{cases} 1 & r_{new,j} > 0 \\ 0 & r_{new,j} = 0 \end{cases}$$

Optimized variables:
The mixture ratios of all the raw materials available on the database X_{DB} (and process variables Z)

Nonlinear, Constrained, Mixed Integer Optimization Problem

(c) 2004-2008, ProSensus, Inc.

One Example: Golf ball development



Approach to golf ball core design increased the resilience 1.7 times compared to previous products

(c) 2004-2008, ProSensus, Inc.

Summary

- Latent variable models:
  - Important concepts:
    - Low dimensional Latent Variable spaces
    - Models for both X and Y
    - Causality only in LV space
- Passive applications
  - Analysis, soft sensors, monitoring
- Active applications
  - Optimization, control, product development
- All work done in LV space (t1,t2, ..)
- Models for X and Y used to get back to these variables.

(c) 2004-2008, ProSensus, Inc.

Conclusions

- If you use historical process data, then Latent Variable approaches are essential
  - Allow one to uniquely interpret and use the data in both passive and active applications

(c) 2004-2008, ProSensus, Inc.


Thank You
?

Some References on topics in the presentation

- Latent variable methods (general)
  - Eriksson L., Johansson, E., Kettaneh-Wold, N. and Wold, S., 1999. “Introduction to Multi- and Megavariate Data Analysis using Projection Methods (PCA & PLS), Umetrics AB, Umea, Sweden
  - Kourti, T. (2002). Process Analysis and Abnormal Situation Detection: From Theory to Practice. IEEE Control Systems, 22(5), 10-25.
- Software
  - SIMCA_P (Umetrics); Unscrambler (Camo); Matlab toolbox (Eigenvector Technologies), ProMV (ProSensus)
- Analysis of historical data
  - Garcia-Munoz, S., T. Kourti and J.F. MacGregor, A.G.. Mateos and G. Murphy, “Trouble-shooting of an industrial batch process using multivariate methods”, Ind. & Eng. Chem. Res., 42, 3592-3601, 2003
- Monitoring
  - T. Kourti and J.F. MacGregor, 1995. "Process Analysis, Monitoring and Diagnosis Using Multivariate Projection Methods", J. Chemometrics and Intell. Lab. Systems, 28, 3-21.
- Control
  - Flores-Cerillo, J. and J. F. MacGregor, “Within-batch and batch-to-batch inferential adaptive control of semi-batch reactors: A Partial Least Squares approach”, Ind. & Eng. Chem. Res., 42, 3334-3345, 2003.
- Image-based soft sensors
  - Yu, H., J.F. MacGregor, G. Haarsma, and W. Bourg, “Digital imaging for on-line monitoring and control of industrial snack food processes”, Ind. & Eng. Chem. Res., 42, 3036-3044, 2003
  - Yu, H. and J.F. MacGregor, “Multivariate image analysis and regression for prediction of coating content and distribution in the production of snack foods”, Chem. & Intell. Lab. Syst., 67, 125-144, 2003

(c) 2004-2008, ProSensus, Inc.

References, continued

- Optimization
  - Jaeckle, J.M., and MacGregor, J.F. (1998). Product Design Through Multivariate Statistical Analysis of Process Data. AIChE Journal, 44, 1105-1118.
  - Jaeckle, J.M., and MacGregor, J.F. (2000). Industrial Applications of Product Design through the Inversion of Latent Variable Models. Chemometrics and Intelligent Laboratory Systems, 50, 199-210.
  - Yacoub, F. and J.F. MacGregor, "Product optimization and control in the latent variable space of nonlinear PLS models", Chemometrics & Intell. Lab. Syst., 70, 63-74, 2004.
  - Garcia-Munoz, S., J.F. MacGregor, D. Neogi, B.E. Latshaw and S. Mehta, "Optimization of batch operating policies. Part II: Incorporating process constraints and industrial applications", Ind. & Eng. Chem. Res., Published on-line, May, 2008
- Product development
  - Muteki, K., J.F. MacGregor and T. Ueda, "On the Rapid development of New Polymer Blends: The optimal selection of materials and blend ratios", Ind. & Eng. Chem. Res., 45, 4653-4660, 2006.
  - Muteki, K. and J.F. MacGregor, "Multi-block PLS Modeling for L-shaped Data Structures, with Applications to Mixture Modeling", Chemometrics & Intell. Lab. Systems, 85, 186-194, 2006
- Design of Experiments
  - Muteki, K., J.F. MacGregor, and T. Ueda, "Mixture designs and models for the simultaneous selection of ingredients and their ratios", Chemometrics & Intell. Lab. Systems, 86, 17-25, 2007.
  - Muteki, K. and J.F. MacGregor, "Sequential design of mixture experiments for the development of new products", Chemometrics & Intell. Lab Sys., 2007.

(c) 2004-2008, ProSensus, Inc.