Computational Intelligence in the Chemical Industry

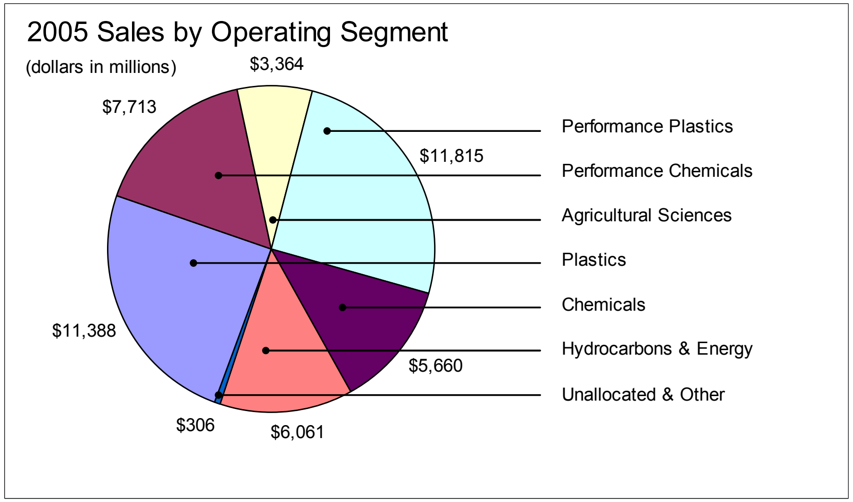
Elsa Jordaan

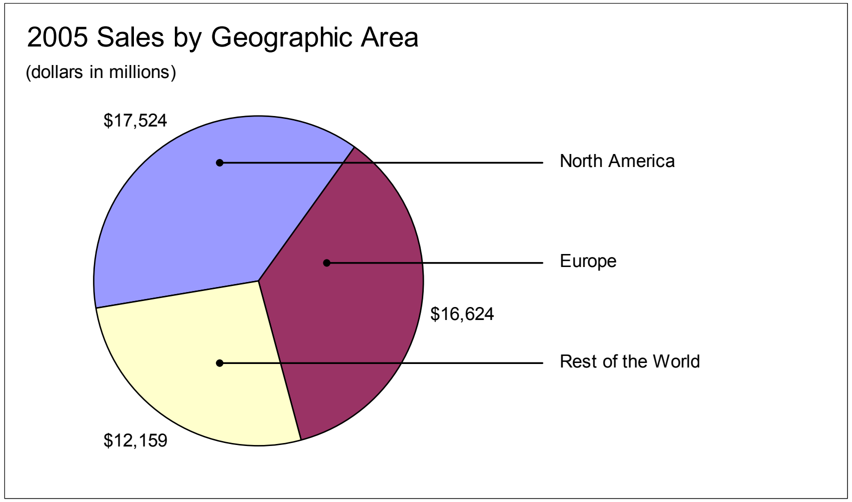
The Dow Chemical Company
Core R&D, Engineering & Process Sciences
Modelling

About Dow…

- Leading science and technology company
- 43 000 employees (5 600 in Global R&D)
- 2005 Record Sales of $46.3 billion
- Customers in more than 180 countries
- Wide range of markets: food, transportation, health and medicine, personal and home care, and building and construction.

- Global R&D:
  - Business R&D
    - dedicated to an operating segment
  - Core R&D
    - supports all operating segments and functions





* 2005 Corporate Report

Before 1990

- Fundamental modelling
  - *a priori* knowledge of processes
  - years of research
- Statistical modelling and data analysis (chemometrics)
  - availability of clean data
  - low-dimensional
  - linearizable

After 1990

- Highly nonlinear processes
- High-dimensional data
- Noisy data
- Fast development of models
- New set of tools
  - Neural Networks
  - Genetic Algorithms
  - Genetic Programming
  - Support Vector Machines
  - Particle Swarm Optimization

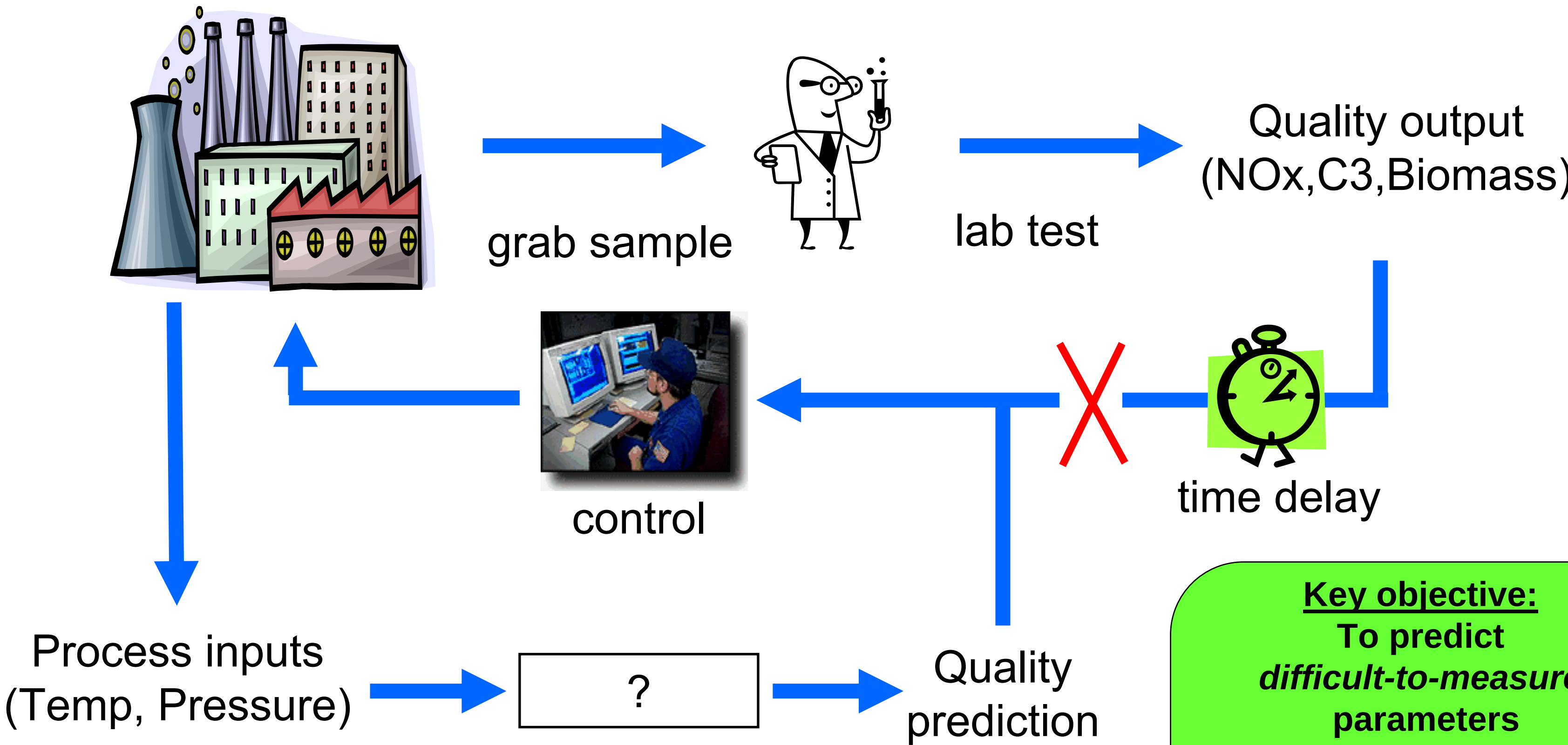
Computational Intelligence at Dow

Application Domains	Examples
Material Design	• Color Matching • Appearance Engineering • Polymer Design • Synthetic Leather
Materials Research	• Diverse Chemical Library Selection • Fundamental Model Building • Reaction Kinetics Modeling • Combi-Chem Catalyst Exploration • Combi-Chem Data Analysis
Production Design	• Nonlinear DOE • Bioreactor Optimization
Production Monitoring & Analysis	• Critical Parameter Monitoring • Calibration Variable Selection • Intelligent Alarm Processing • Emulator for Online Optimization • Emissions Monitoring
Business Modeling	• Diffusion of Innovation • Hydrocarbon Trading & Energy Systems Optimization • Scheduling Heuristics • Plant Capacity Drivers

Specific Examples

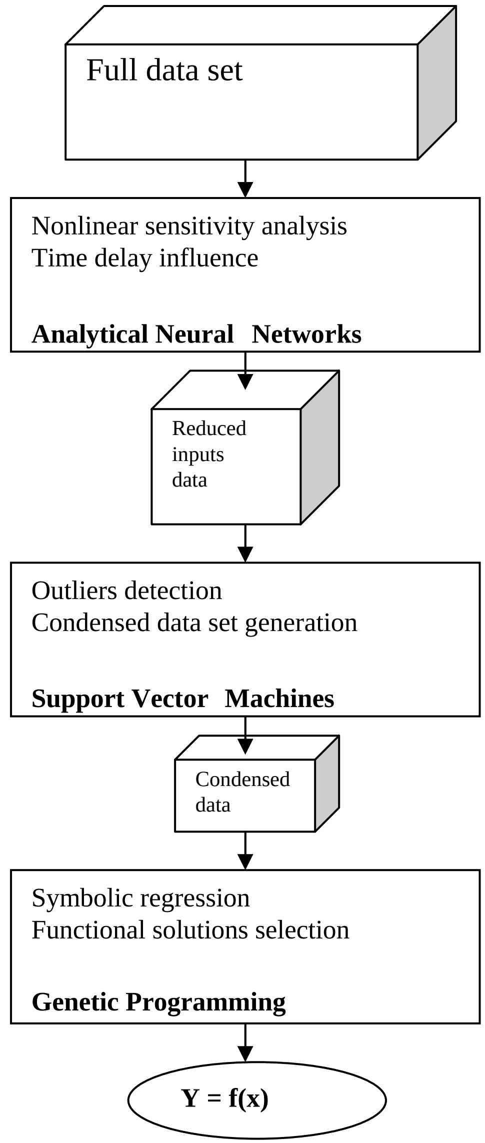
- Inferential Sensors (NN, SVM, GP)
  - NOx-Emissions predictions
  - Biomass concentration prediction in batch fermentation processes
  - Distillation impurity estimation
- Fault and Drift Detection of inline GC
- Knowledge discovery
  - New rheological insights
- Optimization
  - Multi-objective of plastic properties

Inferential Sensor

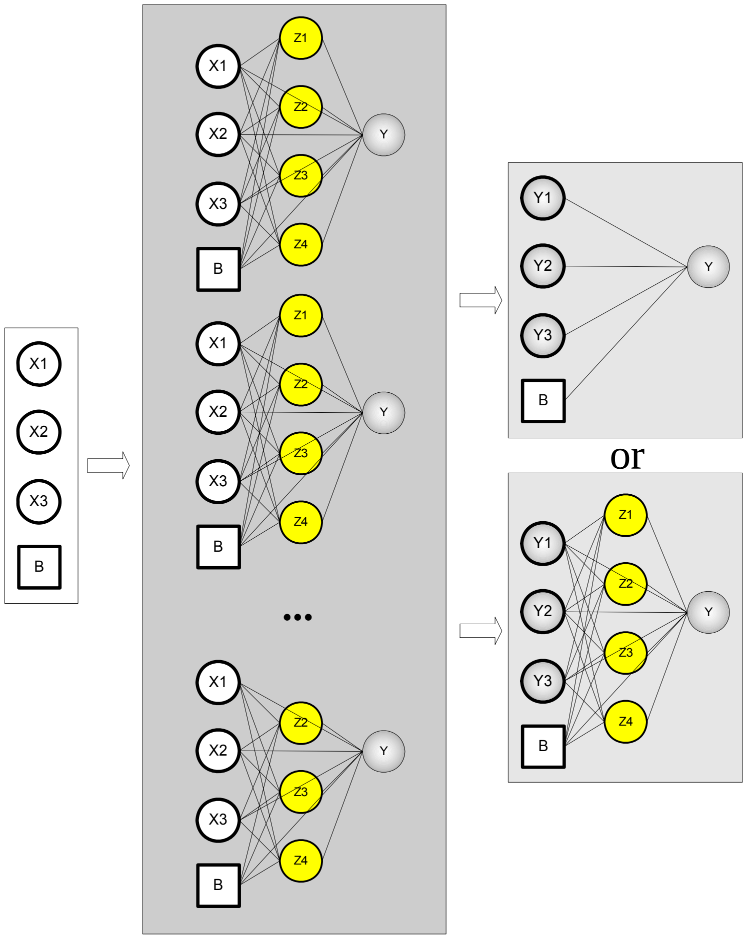


<u>Key objective:</u>
To predict
difficult-to-measure
parameters
(NOx, Hg, Melt Index)
from
***easy-to-measure* data**
(temperature, pressure, flow, etc.)

Inferential Sensor Development



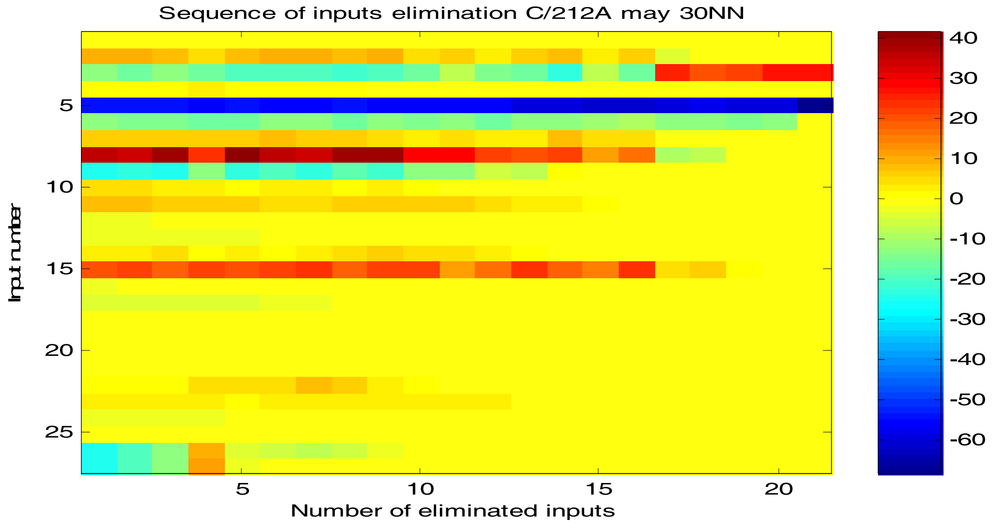
Stacked Analytical Neural Networks



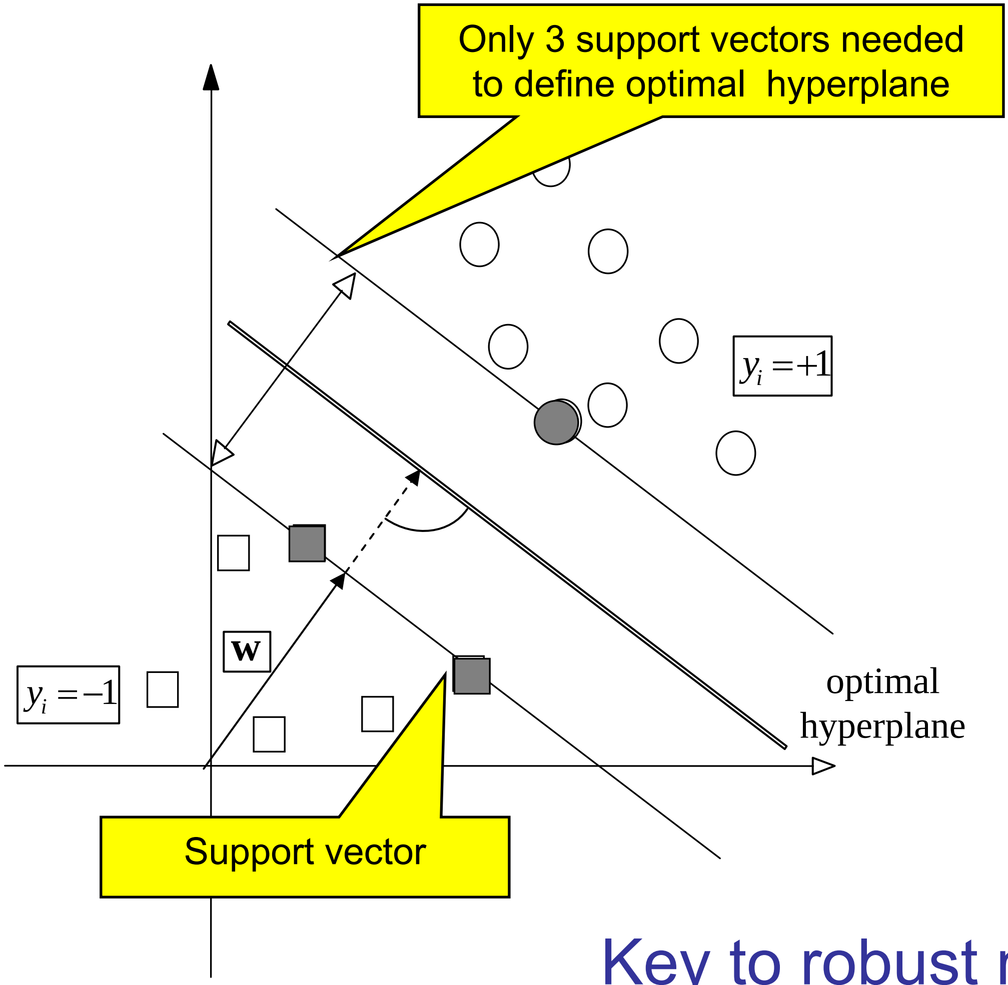
- Fast development
- Diverse subnet consensus indicator of model output quality
- Allows explicit calculations of input/output sensitivity
- Can handle time-delayed inputs by convolution functions
- Gives more reliable estimates based on multiple models statistics

Nonlinear Sensitivity Analysis

- Iteratively eliminate input variables
- Sensitivity of a variable is the impact to the quality of the total model



Support Vector Machines

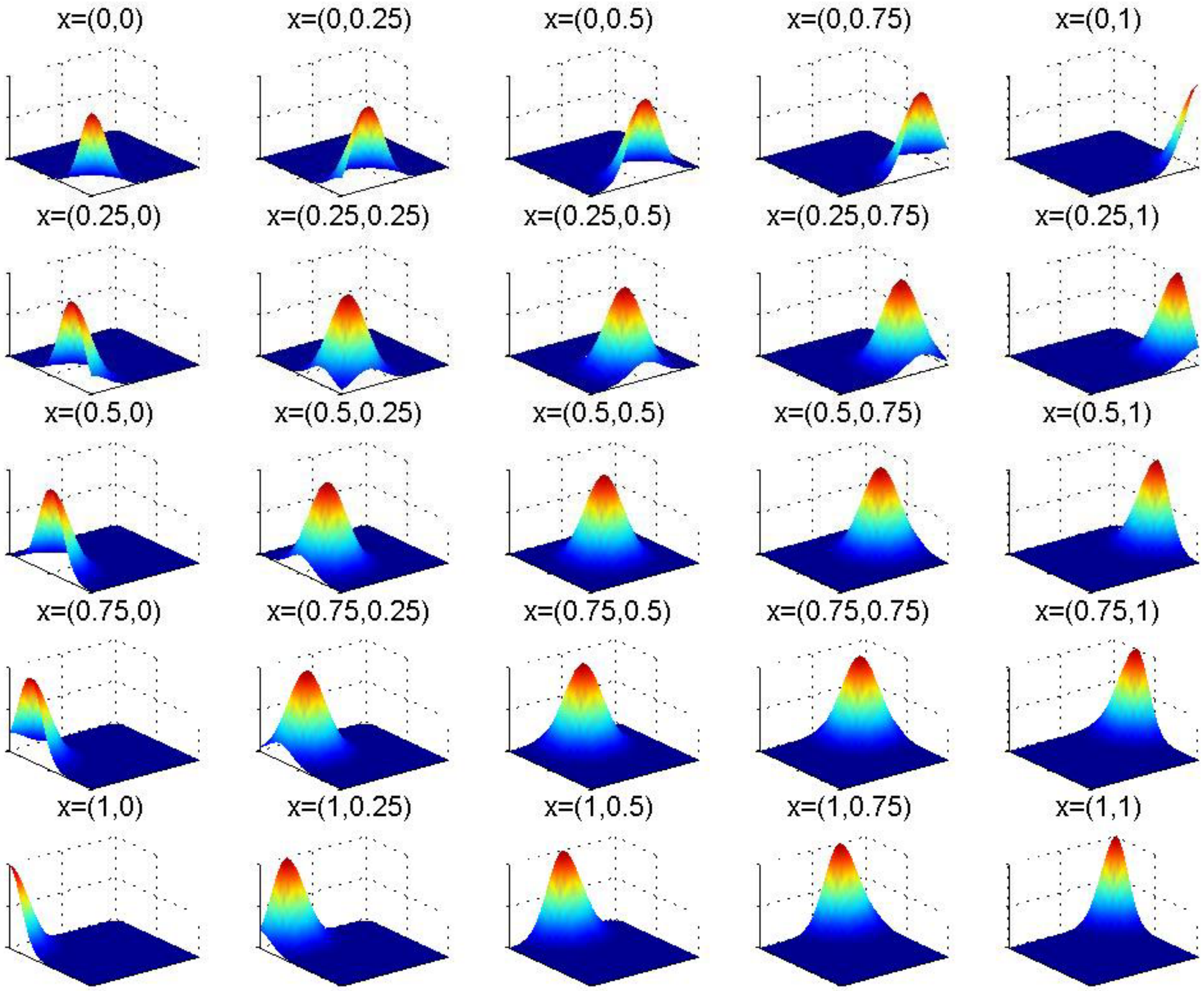


- Solid theoretical basis => Statistical Learning Theory
- Model building is based on global optimum
- Explicit control over model complexity

- *ad hoc* Kernel selection
- Complex theory
- No commercial software
- Computationally intensive

Key to robust modeling

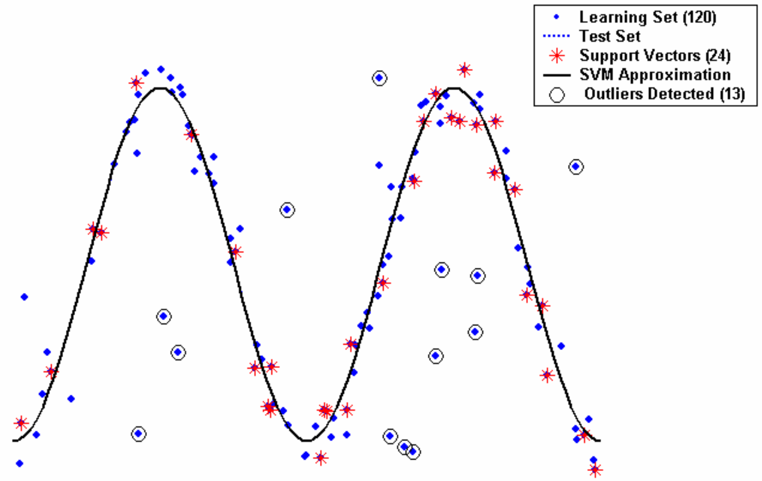
Kernel functions in SVM



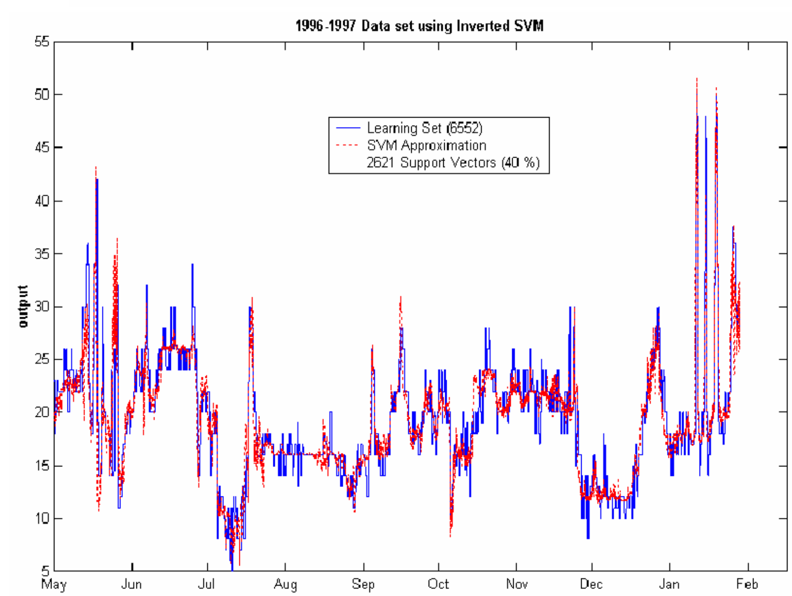
RBF Kernel with σ=0.2

Outlier detection using SVM

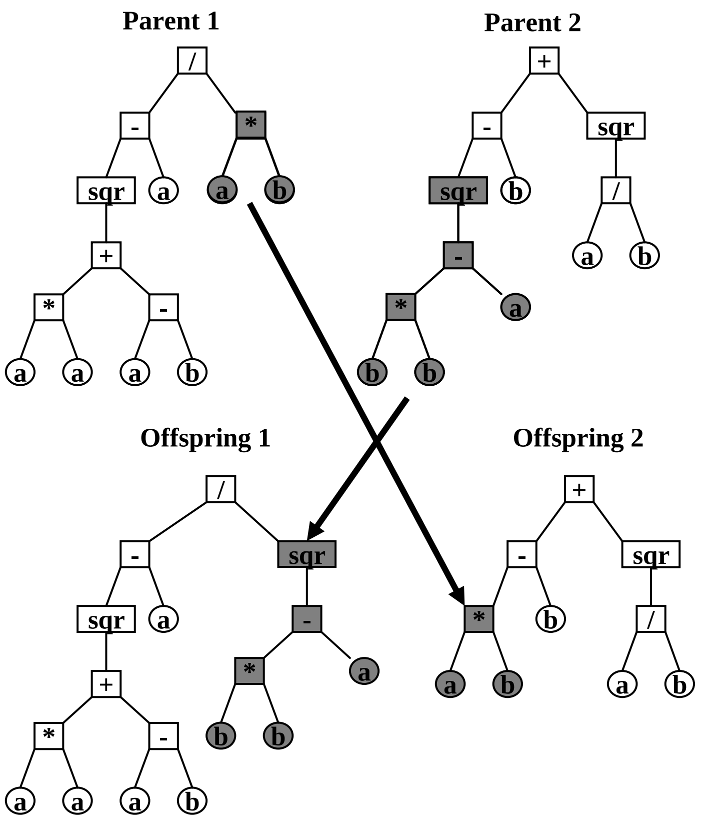
- Support vectors are 'unusual' data points
- Set of sv's contain also the outliers



- Inversed reasoning can be used for data compression



Genetic Programming



Example of GP-generated function

y = 7.05 - 0.024*((x2 + x3/x4))

How does it work?

- Based on artificial evolution of millions of potential nonlinear functions
- Start with population of basic operators and functions:

  + - / * sqrt(x) exp(x) sin(x) x^a
- Let new functions (offspring) “evolve” through mutation, cross-over, etc.
- Next generation => survival of the fittest

What is the result?

- Many (thousands) possible solutions with different levels of complexity
- The final result is an explicit nonlinear function

NOx-Emissions via NN

- Environmental laws limits amount yearly emitted by a production site.
- US/EU requires an accurate estimate of the emissions.
- The hardware solution CEMS available at a cost of $300 000 per unit.
- May use predictive model
  - accuracy
  - yearly audited
- Commercial system based on NN
- Dow-internally developed model (based on statistical model)

Training & Performance

	Commercial	Inferential sensor
Model	Neural network with 14 inputs	GP-model with 5 inputs
Maintenance	Complete model re-training done by vendor.	Parameter fit in statistical software
On-line implementation	Requires specialized run-time licenses	Direct implementation in control
RATA 2003	1.8 %	1.3 %
RATA 2004	5.8 %	3.6 %

- **Better accuracy**
- **Low maintenance cost**
- **No software licenses**
- **Estimated savings ~ $100M/sensor**

[Biomass] via GP-ensemble

- Bacterial fermentation processes in batch reactors
- Need to know when to stop growth-phase and go into production phase
- Biomass-concentration indicates when enough organisms are present for production
- Concentrations are measured by weighing the dried biomass taken from samples every 2-4 hours
- Online hardware equipment (e.g. optical density) is very expensive ($100 000 per unit)
- Typical difficulties for modeling
  - High batch-to-batch variation
  - New cultures requires retraining

Experimental Data

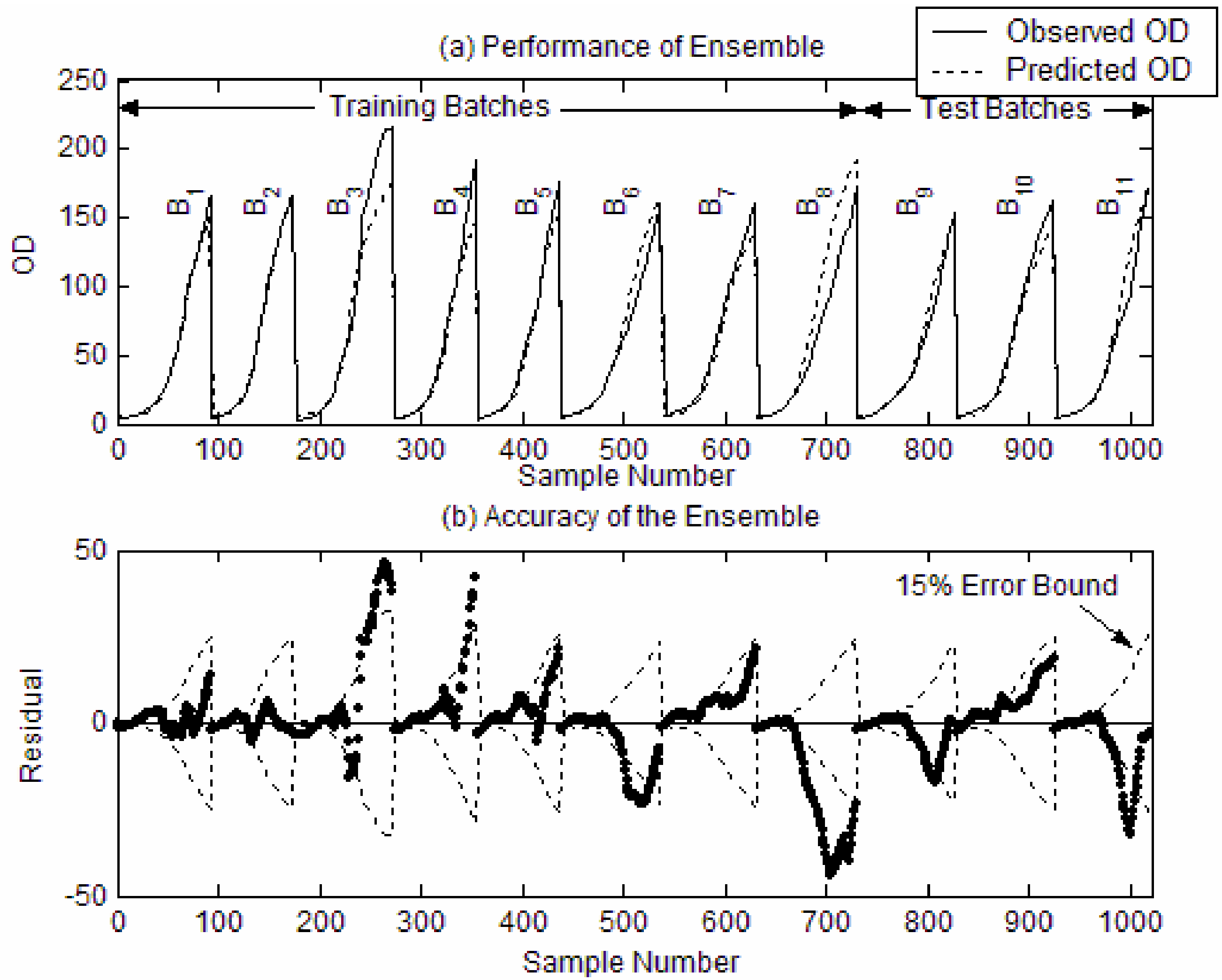
- Input variables: 18 process variables
- Output variable: measured OD (proportional to biomass concentration)
- 8 batches for training
- 3 batches for validation

Performance of soft sensor

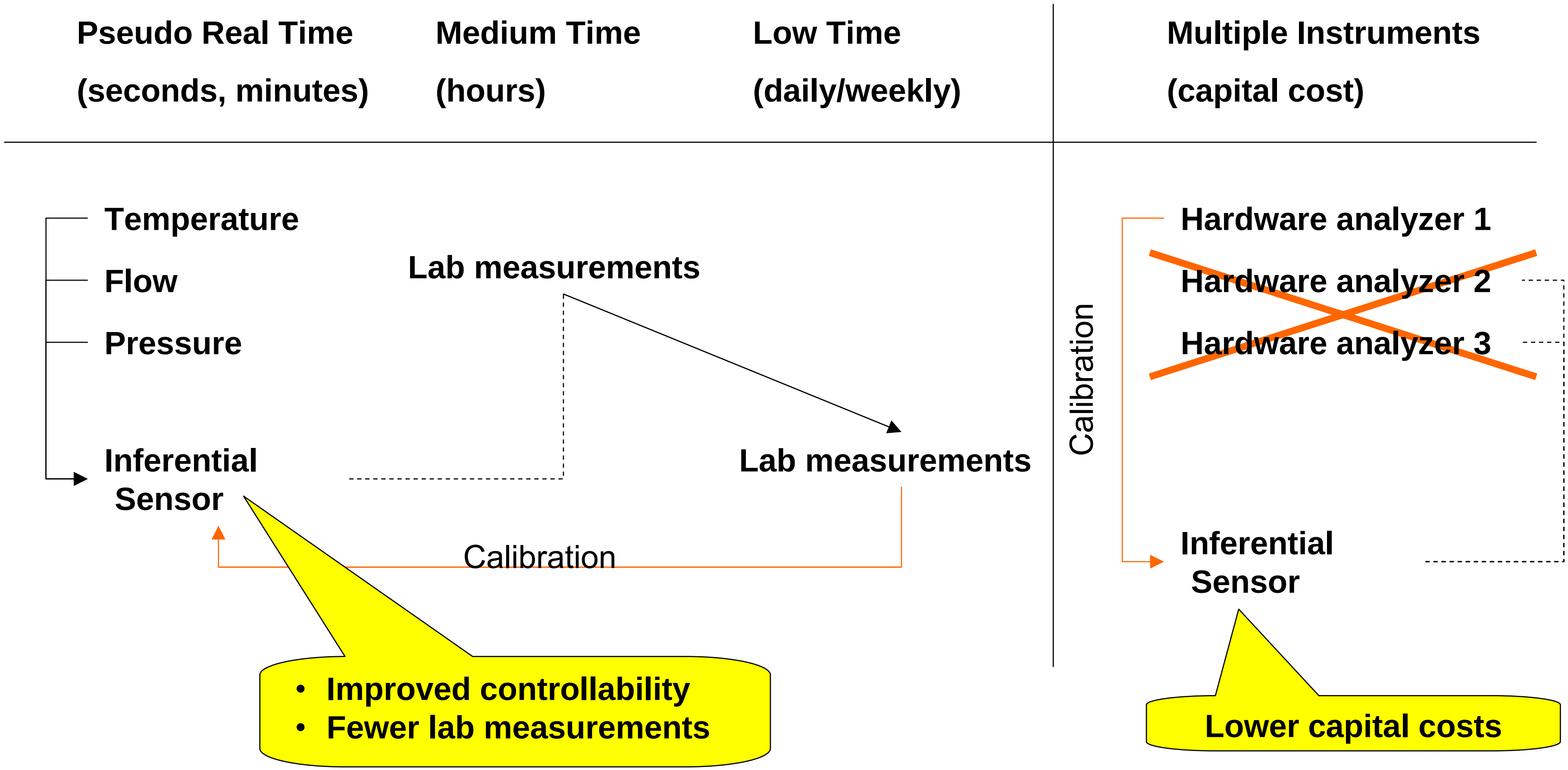
Robust w.r.t batch-to-batch variation

Accuracy

Detection of inconsistencies

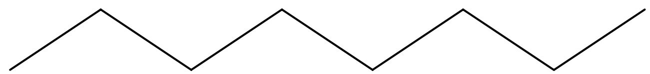


Impact: $ and quality



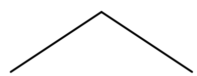
C3-estimation using GP

- Raw material (long-chain polymers) are broken in crackers under high T and high P.



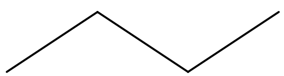
C_8 (octane)

- Shorter chain polymers needed for other processes



C_3 (propane)

C_2 (ethane)



C_4 (buthane)

- Need estimation of fraction of C3
- Measured by Gas-chromatograph
- Problems: GC-measurement frequently unavailable (drifting, faulty measurements)

GP model development

x1
X2
X3
X4
X5
X6
X7
X8
X9
X10
X11
X12
X13
X14
X15
X16
X17
X18
X19
X20
X21
X22
x23

- Model 1 (GP selection based on x1 to x23) :

$$-0.004339-0.002665\left(\frac{\mathbf{x}_8^{\ 2}}{\mathbf{x}_8-\mathbf{x}_{11}}\right)$$

$R^2 = 0.941$
RMSE = 0.0404

- Model 2 (GP selection based on x1 to x23 and x8-x11)

$$-0.02352+1.2697\cdot 10^{-9}\left(\frac{\mathbf{x}_8^{\ 4}\cdot\mathbf{x}_{21}^{\ 3}}{\mathbf{x}_4^{\ 2}}\right)$$

$R^2 = 0.941$
RMSE = 0.0409

Strongly supported by the HC engineers

- Model 3 (GP selection based on x1 to x7, x9 to x23):

$$-0.09153+1.6187\cdot 10^{-10}\left(\frac{\mathbf{x}_5^{\ 3}\cdot\mathbf{x}_9^{\ 4}\cdot\mathbf{x}_{21}^{\ 2}}{\mathbf{x}_4^{\ 4}}\right)$$

$R^2 = 0.910$
RMSE = 0.0497

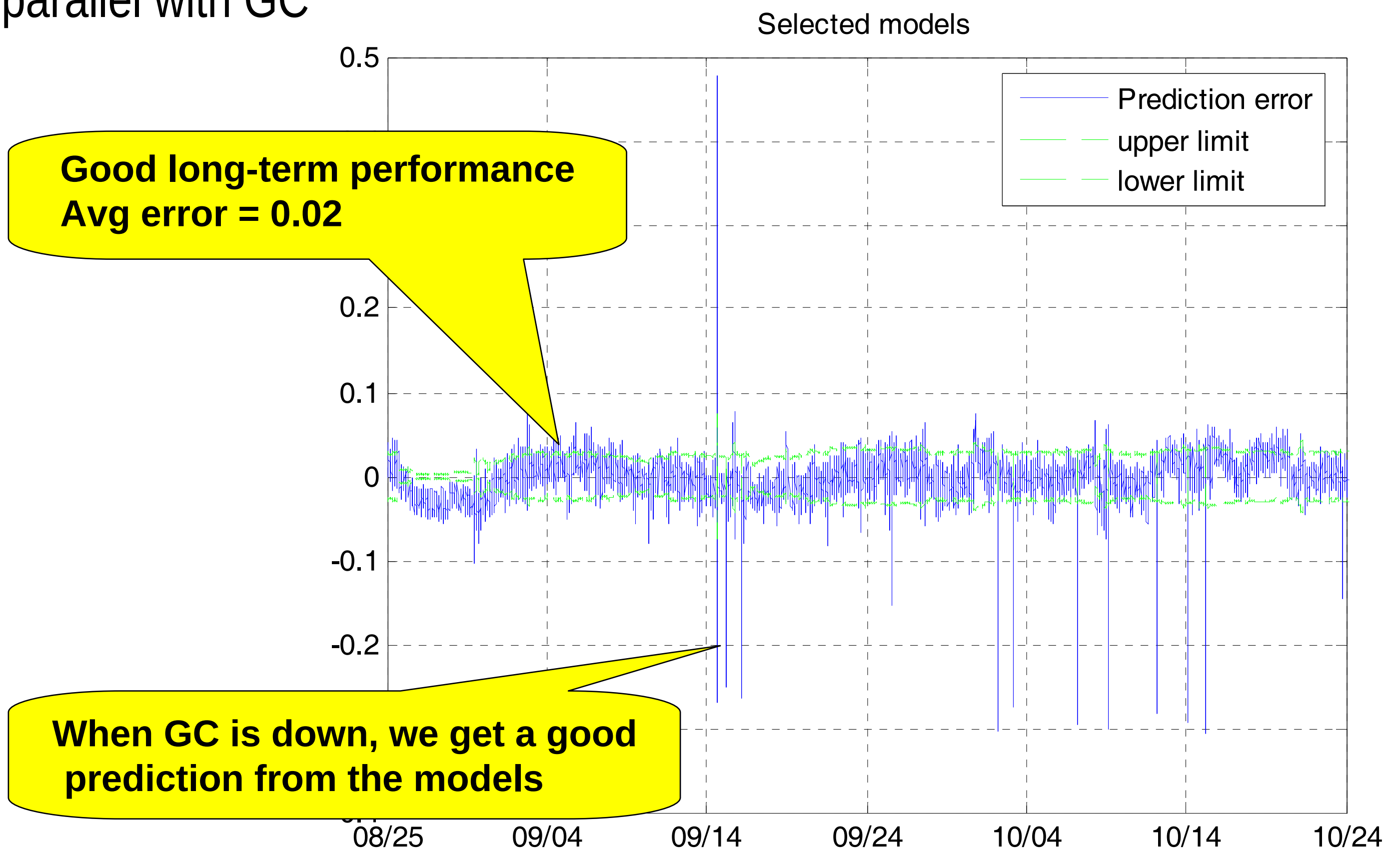
- Model 4 (GP selection based on x1 to x23, x8-x11, and log y)

$$\exp\left(2.022-88.7054\left(\frac{\sqrt{\mathbf{x}_{11}-\mathbf{x}_8}}{\mathbf{x}_8+\mathbf{x}_{12}}\right)\right)$$

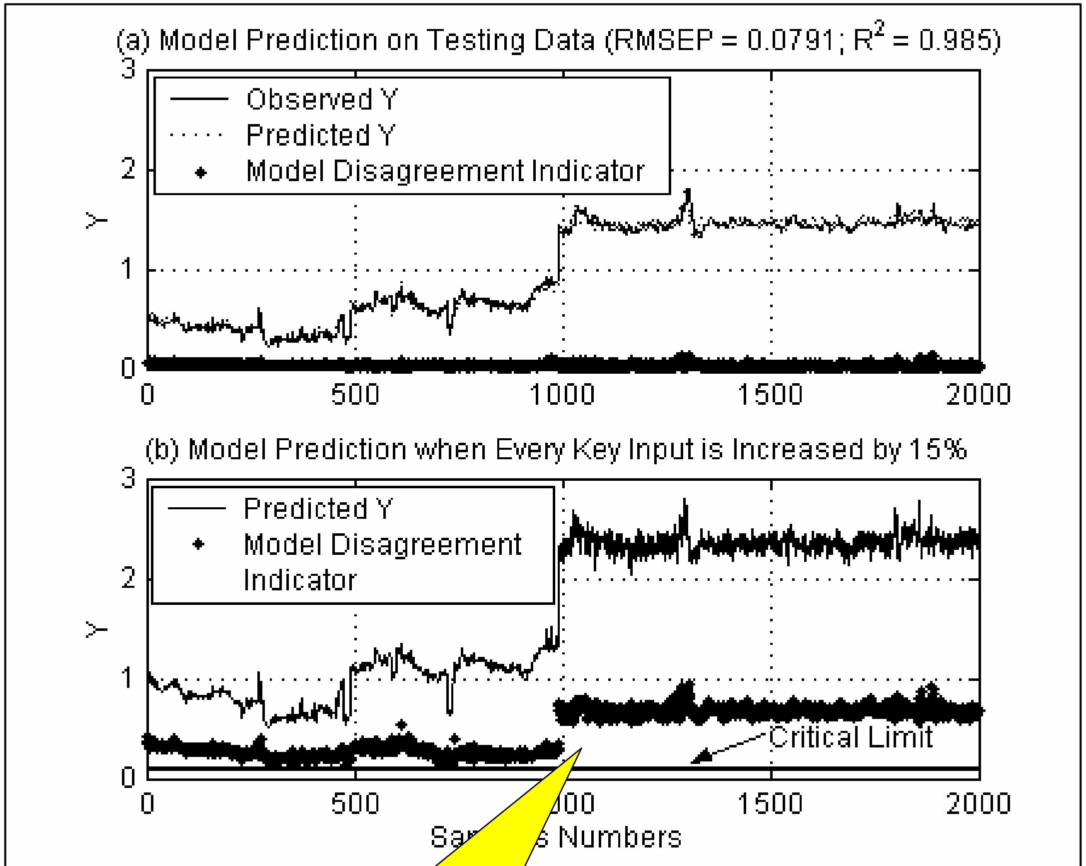
$R^2 = 0.942$
RMSE = 0.0406

Online Performance

- In operation since May 2004
- In parallel with GC

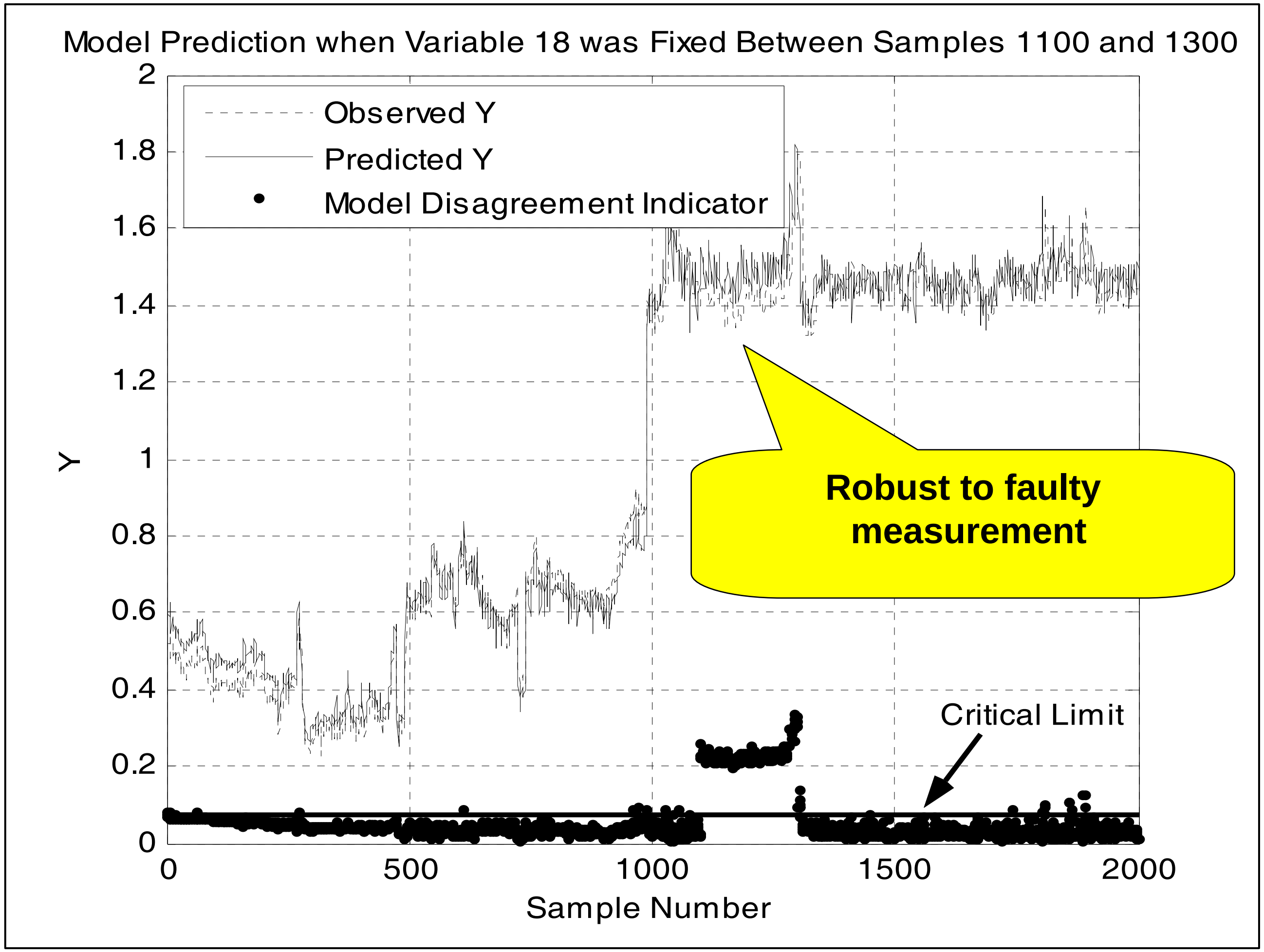


Fault & Drift Detection (1)

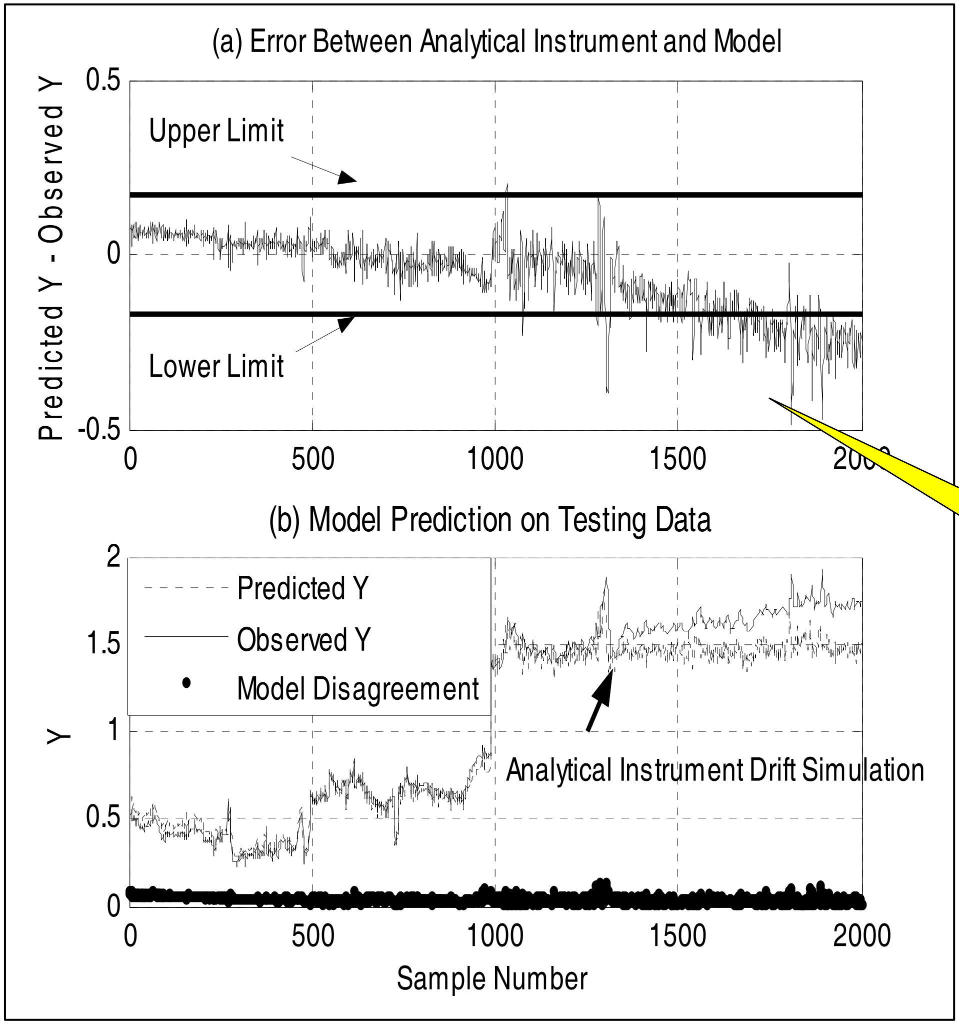


Detect abnormal process conditions

Fault & Drift Detection (2)



Fault & Drift Detection (3)



Detect drift of analytical instrument

Knowledge discovery via GP

Understanding the complex connection between molecular structure and rheological properties is key to new product development. (MWD ↔ rheology)

Empirical Models

- Measure MWD indirectly
- Moments: Mw, Mn and Mw/Mn
- rheology = f(moments)
- Only valid for limited distribution parameter ranges

Fundamental Models

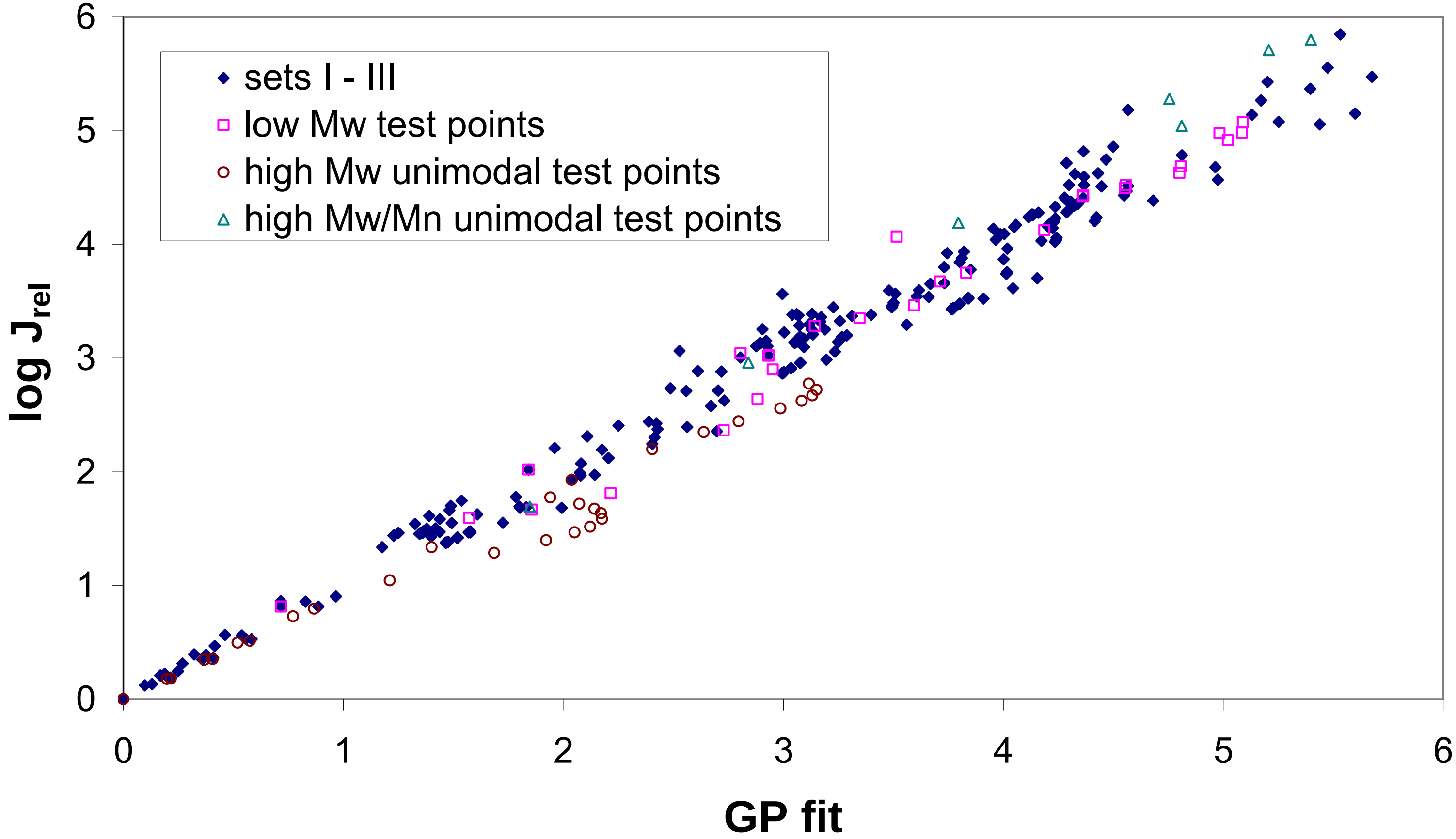
- Theoretically derived
- MWD theoretically determined
- rheology = f(MWD)
- Valid for large ranges
- However, MWD not known in practice

Desire: rheology = f(moments), but also valid for large ranges

Virtual Experimental Design

- Use fundamental models to generate rheology data (output)
  – Set I: Monodisperse systems
  – Set II: Polydisperse systems - unimodal
  – Set III: Polydisperse systems – bimodal
- Derive moments data from theoretical distributions (inputs)
- Use GP to develop large range model

Performance New Model



Particle Swarm Optimization

- An efficient technique to find the global optimum for model inversion and non-linear parameter estimation

At each time step *t*

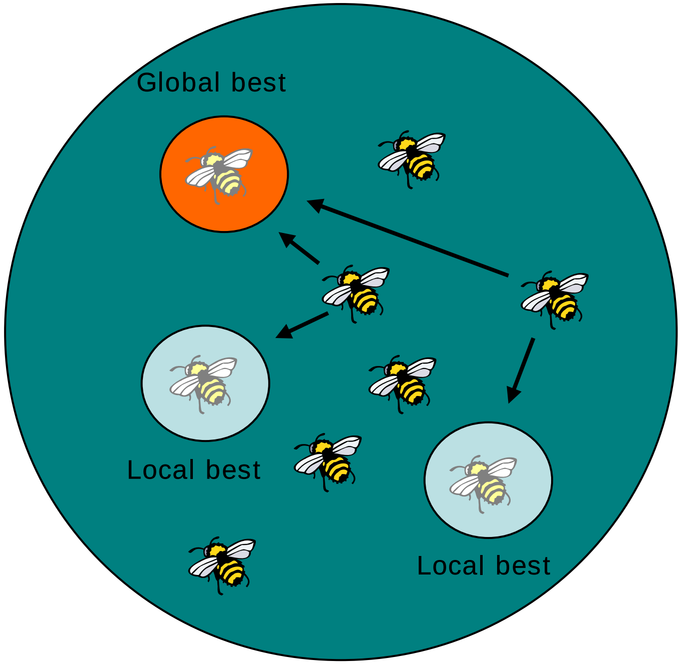
For each particle *i*

Update the position change (velocity)

$$V_i(t+1) = \chi \cdot (V_i(t) + c_1 \cdot rand(0,1) \cdot (P_i(t) - X_i(t)) + c_2 \cdot rand(0,1) \cdot (P_g(t) - X_i(t))$$

Then move

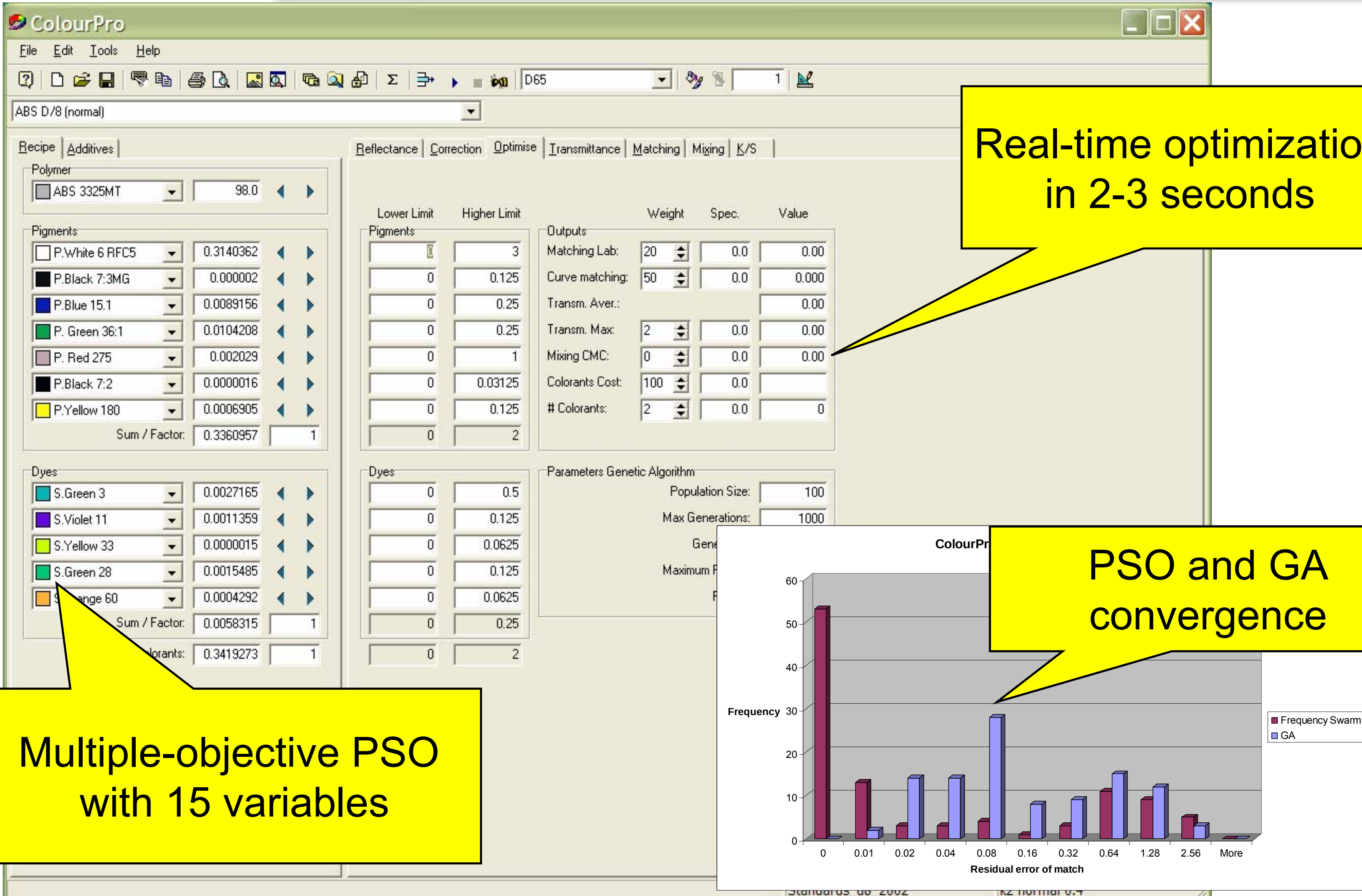
$$X_i(t+1) = X_i(t) + V_i(t+1)$$



Note: - stochastic component

- parameters c_1, c_2, χ default values (2.05, 2.05, 0.73)

Optimized color spectrum of plastics



Real-time optimization in 2-3 seconds

PSO and GA convergence

Multiple-objective PSO with 15 variables

Requirements for success

- **Accuracy** of prediction
- **Robustness** in the production environment where the quality of the input data is not guaranteed
- **Extrapolation** (generalization) outside the known operating conditions
- **Self-assessment** capabilities to warn operators of low confidence of the model
- **Fault and drift detection** of hardware sensors (analytical instruments)
- **Interpretability** of the model by the process engineers
- **Ease of implementation** in the current control systems
- **Cheap and fast** to development and maintain

Key messages to take home

- (Chemical) Industry needs a broad toolset (no “silver bullet” technique)
  - different techniques are complementary
  - hybrid solutions are the most effective
- Accuracy of a model is not the only requirement for successful applications
  - interpretability
  - ease of implementation
  - low cost of maintenance

External Publications of Applications

Application	Initial data size	Reduced data size	Model structure	Reference
Inferential sensors				
Interface level prediction	(25 inputs x 6500 data pts)	(2 inputs x 2000 data pts)	3 models, 2 inputs	Kordon and Smits, 2001
Interface level prediction	(28 inputs x 2850 data pts)	(5 inputs x 2850 data pts)	One model, 3 inputs	Kalos *et al*, 2003
Emissions prediction	(8 inputs x 251 data pts)	(4 inputs x 34 data pts)	Two models, 4 inputs	Kordon *et al*, 2003b
Biomass prediction	(10 inputs x 705 data pts)	(10 inputs x 705 data pts)	9 models ens, 2-3 inputs	Jordaan *et al* , 2004
Propylene prediction	(23 inputs x 6900 data pts)	(7 inputs x 6900 data pts)	4 models ens, 2-3 inputs	Jordaan *et al* , 2004
Emulators				
Chemical reactor	(10 inputs x 320 data pts)	(10 inputs x 320 data pts)	5 models, 8 inputs	Kordon *et al*, 2003a
Accelerated modeling				
Structure-property	(5 inputs x 32 data pts)	(5 inputs x 32 data pts)	One model,4 inputs	Kordon *et al*, 2002
Structure-property	(9 inputs x 24 data pts)	(9 inputs x 24 data pts)	7 models , 3 -5 inputs	Kordon and Lue, 2004
Structure-property	(4 inputs x 289 data pts)		one model, 3 inputs	Jordaan et al. 2006
Linearized transforms				
Chemical reactor model	(4 inputs x 19 data pts)	(4 inputs x 19 data pts)	3 transforms	Castillo *et al*, 2002