

Hybrid Intelligent Systems for Data-Driven Monitoring and Optimization

Arthur Kordon

Applied Math&Physics Group
Physical Sciences
The Dow Chemical Company



IMA Workshop Data-driven Control and Optimization
December 4, 2002



Applied Math & Physics Group

- Flor Castillo – Ph.D. Industrial Eng. – Applied Statistics & Data Analysis
- Elsa Jordaan - Ph.D. Applied Mathematics - Statistical Learning Theory
- Mike Durisin – M.S. Chemical Eng. – Scientific Computing
- Bob Haley – M.S. Physics – Physics Modeling, Control Engineering, Mathematical Analysis
- Alex Kalos – Ph.D. Chemistry – Intelligent Systems, Simulation, Modeling
- Arthur Kordon – Ph.D. Electrical Eng – Industrial Data Analysis, Intelligent Systems, Adaptive Control
- Mark Kotanchek – Ph.D. Aerospace Eng. – Signal Processing, Systems Engineering
- Kip Mercure – Ph.D. Chemical Eng. – Control Systems, Dynamical Systems
- Dan Obermiller – M.S. Statistics – Data Mining, Non-parametric Statistics
- Guido Smits – Ph.D. Chemistry – Modeling, Computational Intelligence
- Jeff Sweeney – M.S. Applied Statistics – Experimental Design, Multivariate Statistics
- David West – B.S. Chemistry – Reaction Engineering, Dynamical Systems,...



Outline

- Why Data-driven Models?
- Hybrid Intelligent Systems Classics
- Hybrid Intelligent Systems State of the Art
- Hybrid Intelligent Systems at Dow Chemical
- Hybrid Intelligent Systems for Process Monitoring and Optimization

Academic vs. Industrial Data Analysis



Transfer data into knowledge



Transfer data into value



Economic advantage of data-driven models

Expensive hardware analyzers

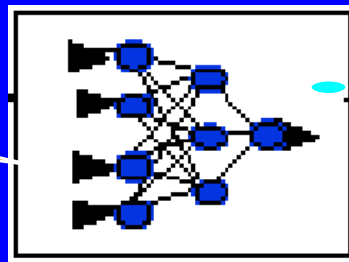


More expensive fundamental models

$$\rho C_p \frac{\partial T}{\partial t} + \nabla \cdot (-k \nabla T + \rho C_p T \mathbf{u}) = Q$$



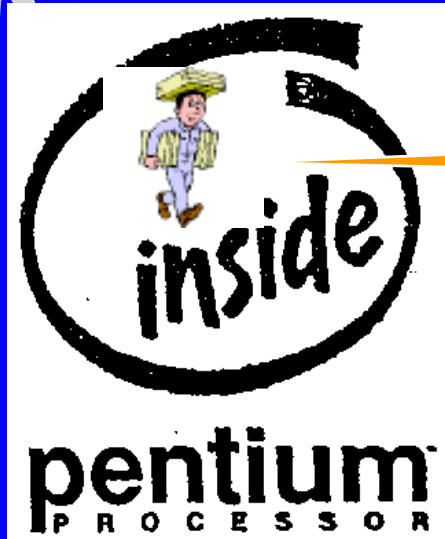
Robust empirical models are often at the economic optimum



Key issue:
Models credibility

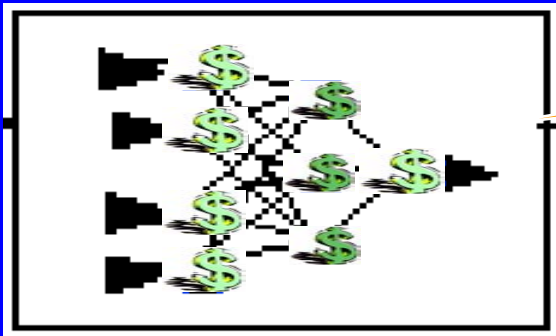


Intelligent Systems in Industry: Lessons From the Past



The Expert Systems campaign (late 80s)
“We’ll put engineers in the box”

- the politics of knowledge acquisition
- static rule-based models not linked to numerical world



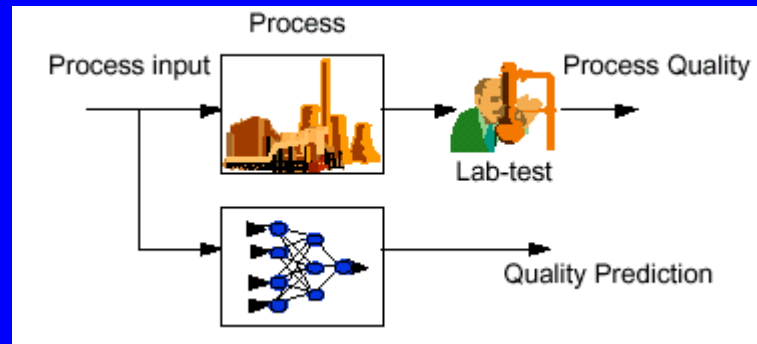
The Neural Networks campaign (early 90s)
“We’ll turn data into gold”

- black-box models with inefficient structure
- maintenance nightmare

What is a soft sensor?

A boring definition:

An empirical model to infer process state and product quality variables that are difficult to measure on-line (composition, melt index, molecular distribution, etc.) from readily available process measurements (temperature, pressure, flow, etc.)



- It is data-driven
- It requires hardware sensors for model development and validation
- It assumes a nonlinear functional relationship
- It is a mature technology
- It is NOT magic
- It is NOT a cause-effect relationship



Lessons from ‘classical’ neural net soft sensors

OR

from “irrational exuberance” to realistic expectations

- Lesson #1

Not a silver bullet

- Lesson #2

Data quality is of crucial importance

- Lesson #3

Soft sensor development requires process knowledge and *ad hoc* tuning

- Lesson #4

Changing operating conditions require frequent re-training of the model

- Lesson #5

Long-term maintenance becomes the main issue

Based on ~ 25 applications in The Dow Chemical Company

Requirements for successful industrial data-driven modeling

Objective function:

Minimizing modeling cost and maximizing data analysis efficiency under broad range of operating conditions

- Robust, fast, and cost effective development process
- Low sensitivity to process changes
- Performance self-assessment capability
- Low cost of ownership and maintenance

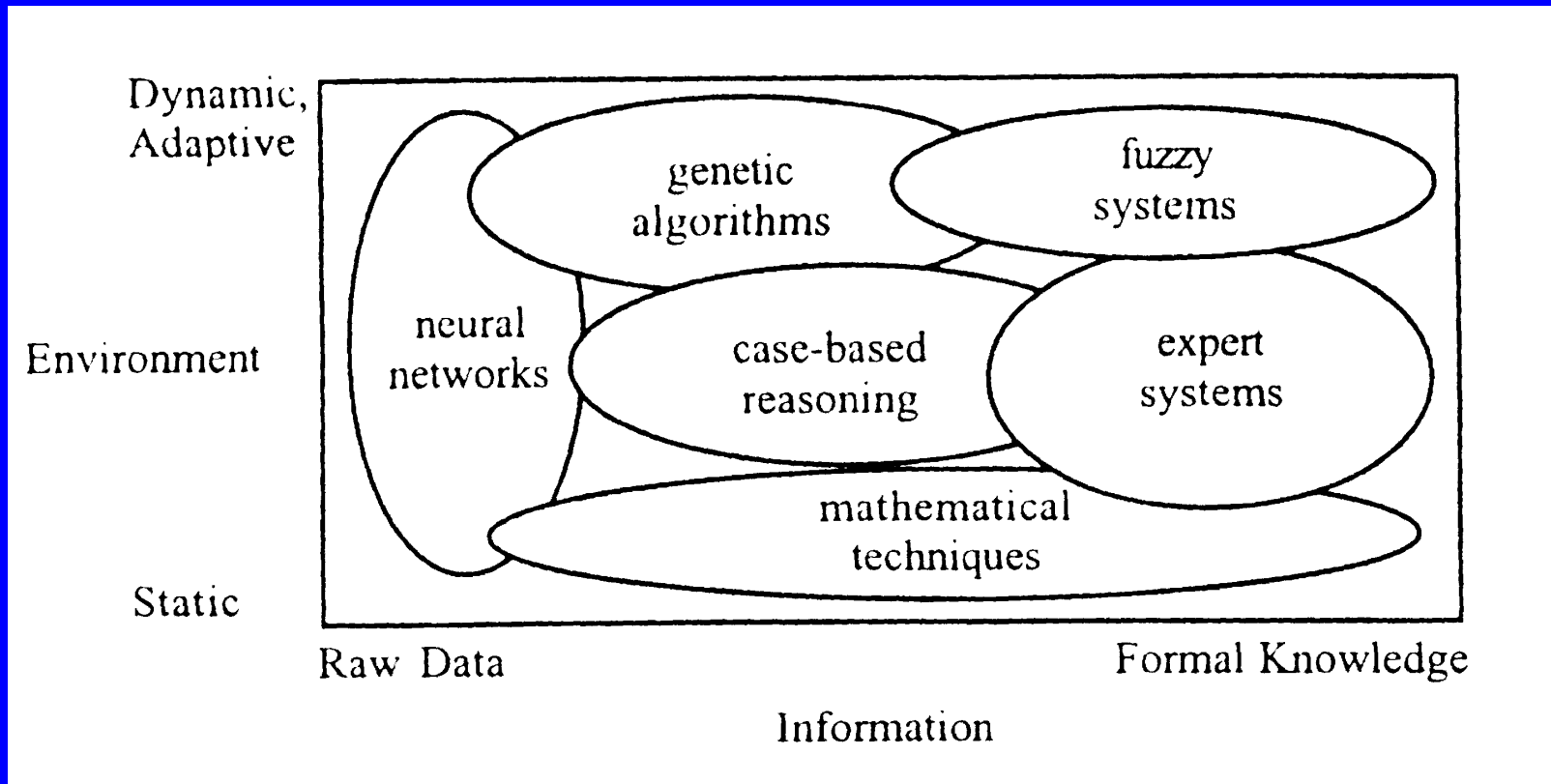
How?
One technology is not enough!



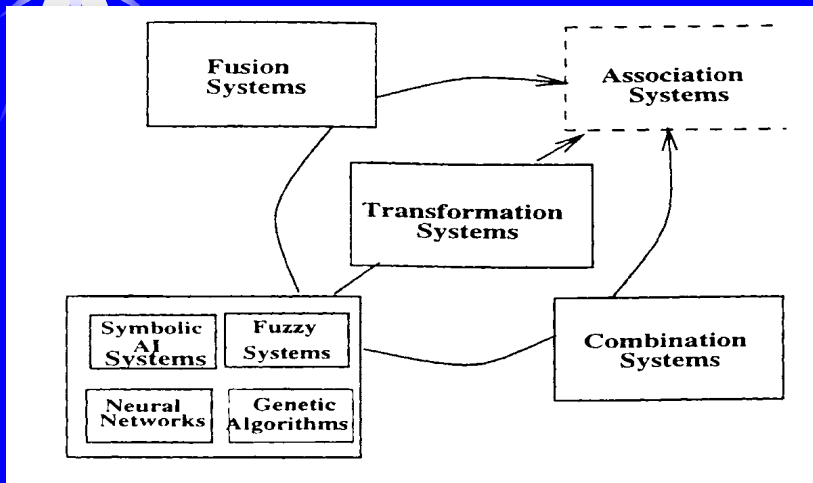
Hybrid Intelligent Systems Classics

Or

The Genetic, The Back-propagated, and The Fuzzy



Hybrid modeling classics



Neuro-Fuzzy System

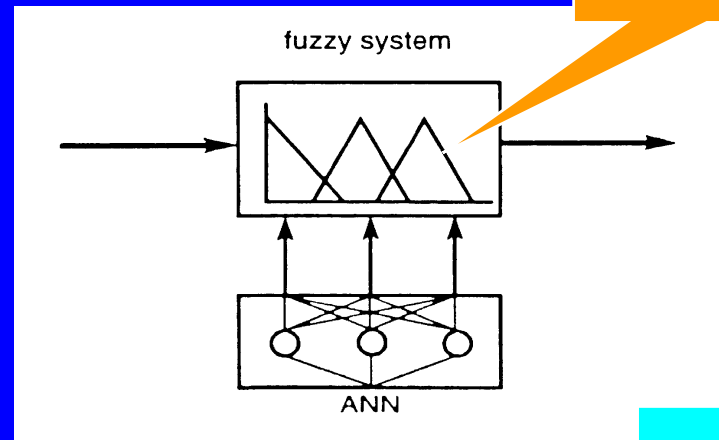
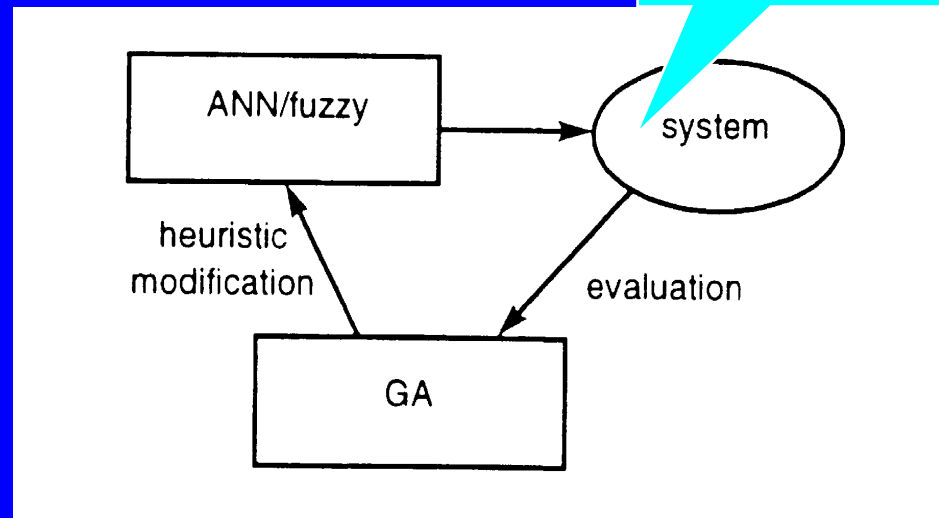
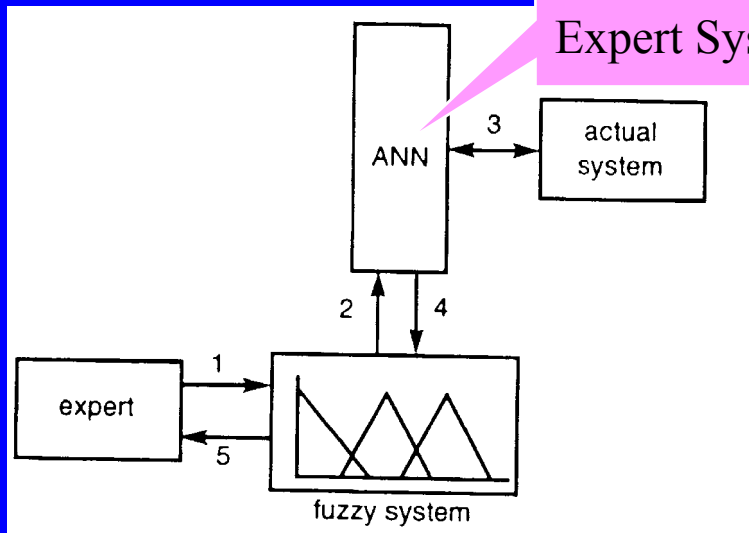


Fig. 6. ANN for tuning membership function

Neuro-Fuzzy Genetic System

Neuro-Fuzzy Expert System



Cat-Dog Hybridization



PERGAMON

Expert Systems with Applications 21 (2001) 63

Expert Systems
with Applications

www.elsevier.com/locate/eswa

Editorial

If you are a dog lover, build expert systems; if you are a cat lover, build neural networks

Most dog lovers are people persons. They like companionship, expressiveness, and empathy. Cat lovers are typically independent types of persons who don't yearn for attention. Cats typically watch life pass by, but dogs get involved with it!

You may wonder what in the world does this have to do with developing intelligent systems like expert systems and neural networks. Expert systems development is very humanistic whereby frequent human interaction is critical to the success of the project. Knowledge acquisition is usually a labor-intensive activity in interviewing the domain experts and eliciting their knowledge. The knowledge engineering life cycle process is an iterative, rapid prototyping approach

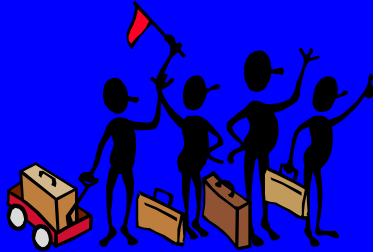
getting at least one paper per day worldwide for possible submission to the Expert Systems With Applications international journal. Many expert systems, and neural networks for that matter, have been integrated within larger systems and have matured so that many commercial applications use these technologies. If we had an "AI Inside" stamp, like the "Intel Inside" stamp, embossed on various information systems, there would be greater awareness and recognition of AI being used. Even the Microsoft wizards use case-based reasoning technology, which flows from the expert system and machine learning fields. Expert systems are being used in almost any industry. Neural networks are being utilized in many applications requiring noisy data

Hybrid Intelligent Systems State of the Art

or

The New Kids on the Block

Stacked Analytic Neural Nets

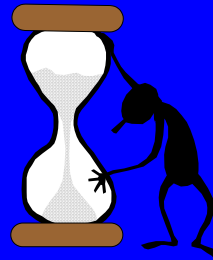


- Sensitivity analysis
- Dealing with time delays
- Very fast development (< 1hr)
- Model confidence indicator

Extract relevant inputs

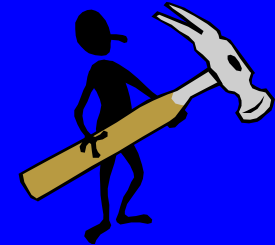
Support Vector Machines

Extract relevant data points



- Outlier detection
- Condensed data set
- Best generalization capability
- Unifying framework

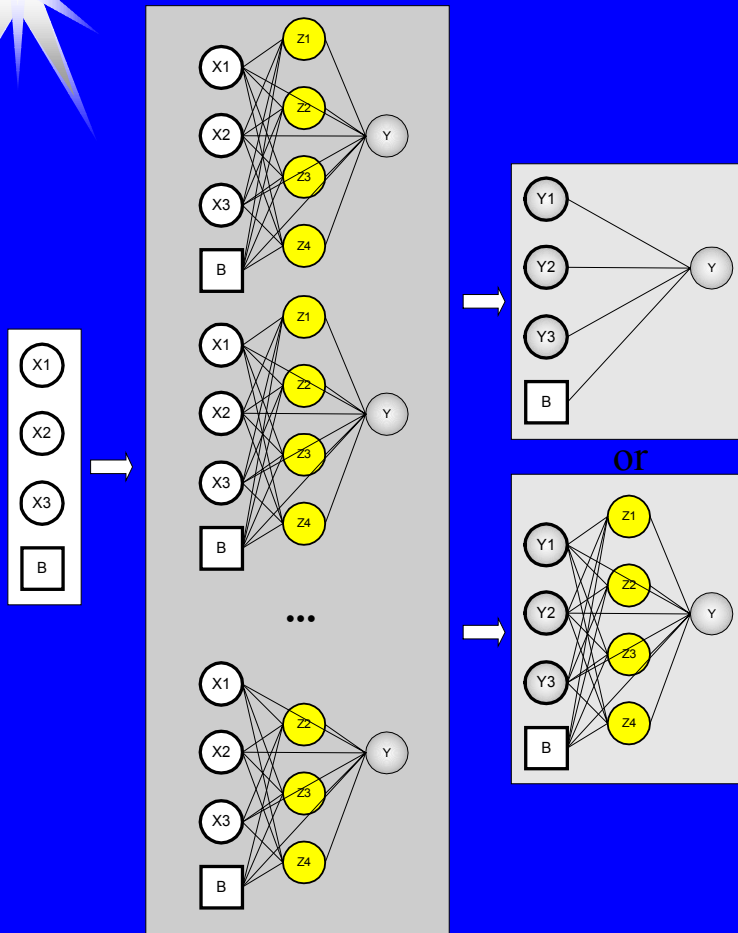
Genetic Programming



- Symbolic regression
- Complexity control
- Easy implementation
- Low maintenance

Delivers final analytical function

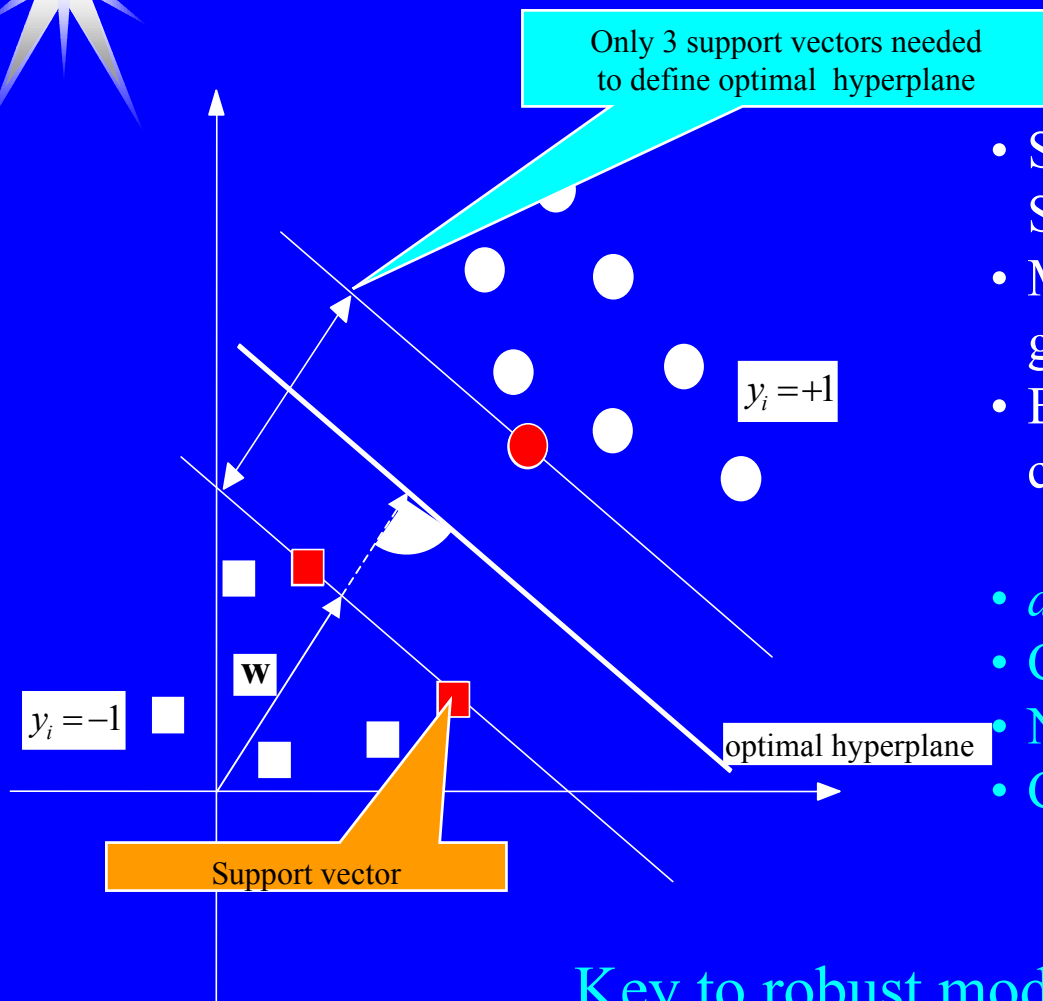
Stacked Analytic Neural Nets



- Fast, not based on back propagation
- Has well defined single global optimum
- Allows explicit calculations of input/output sensitivity
- Can handle time-delayed inputs by convolution functions
- Gives more reliable estimates based on multiple models statistics

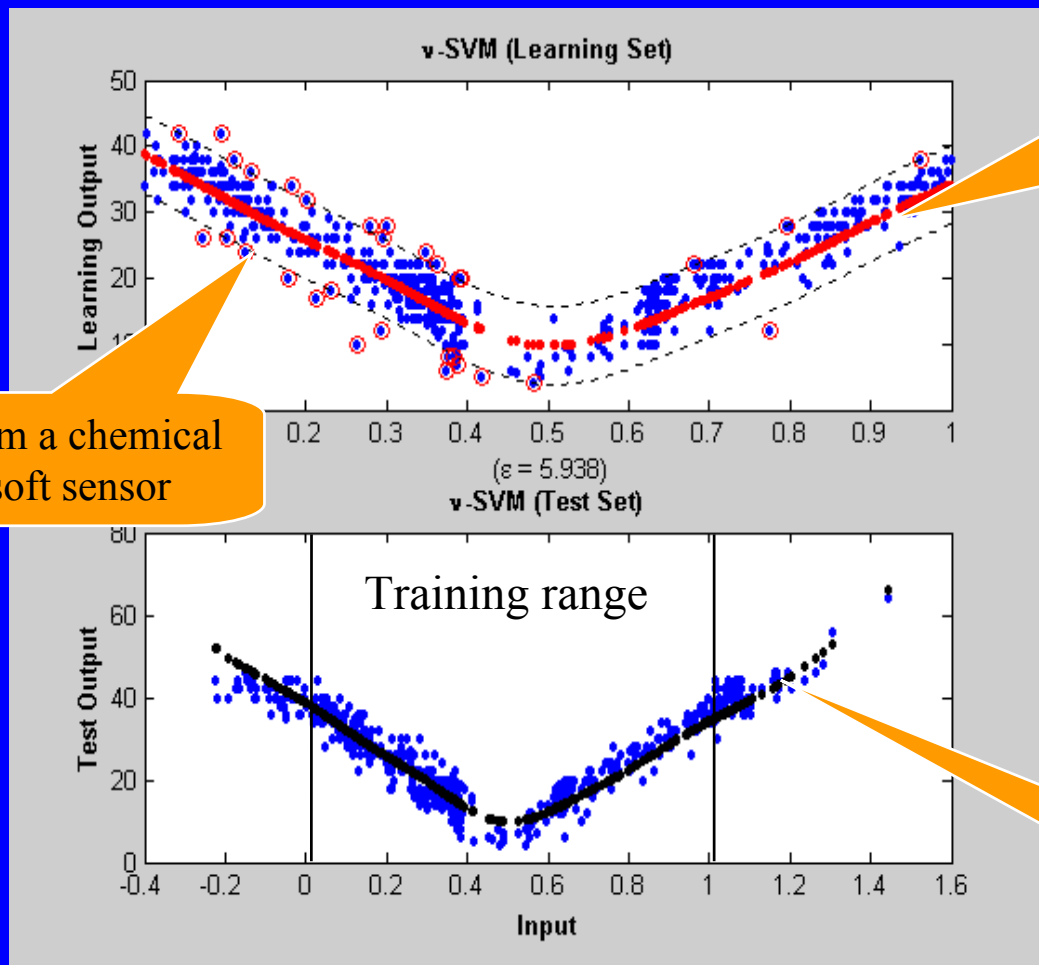
Internally developed in Dow Chemical
by Guido Smits

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

Generalization capabilities of SVM based on mixed kernels



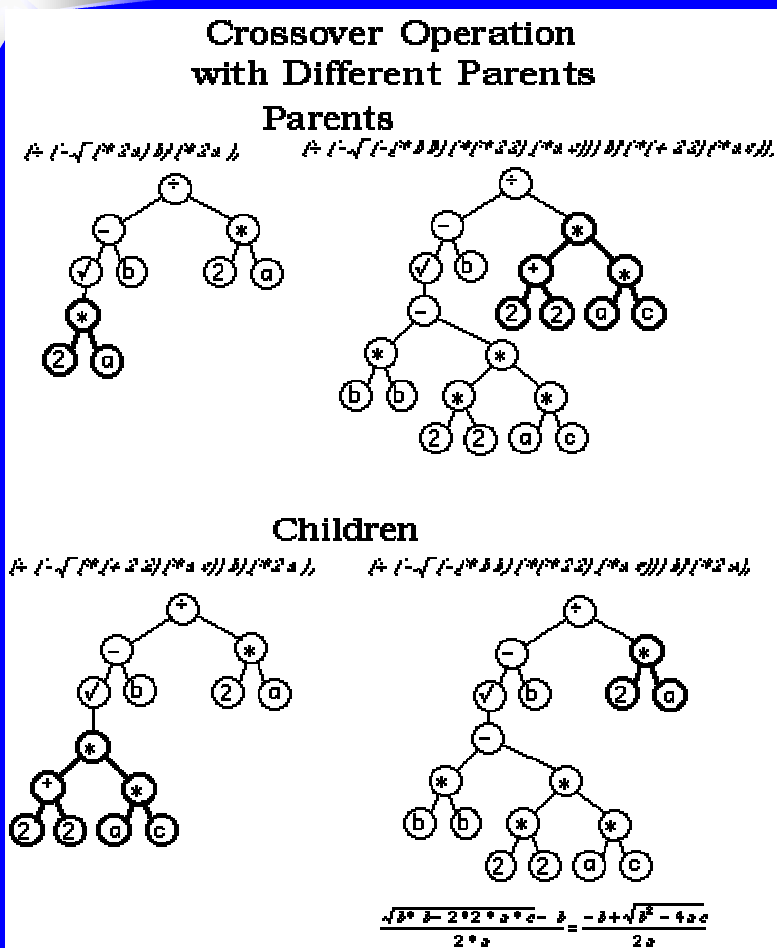
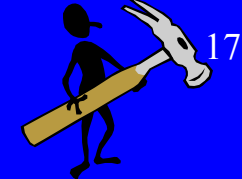
Support Vector Machine model based on mixed polynomial and radial bases function kernels

Data set from a chemical reactor soft sensor

Almost perfect generalization 50% outside the training range!

Reliable model scale-up is feasible

Genetic Programming



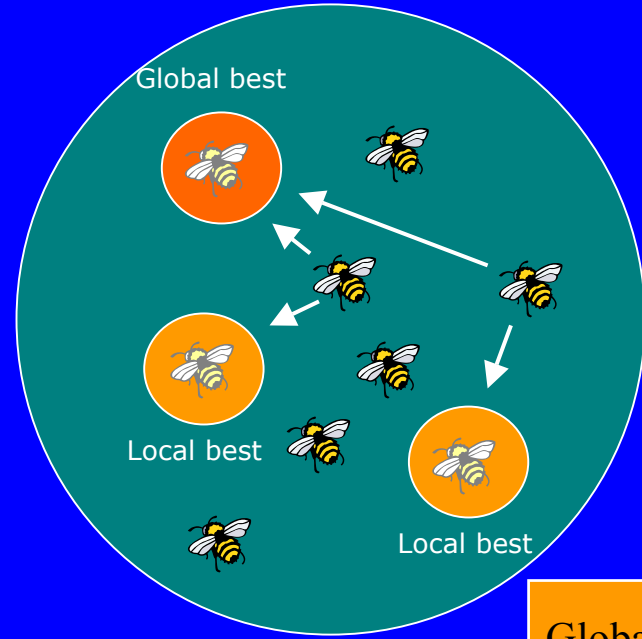
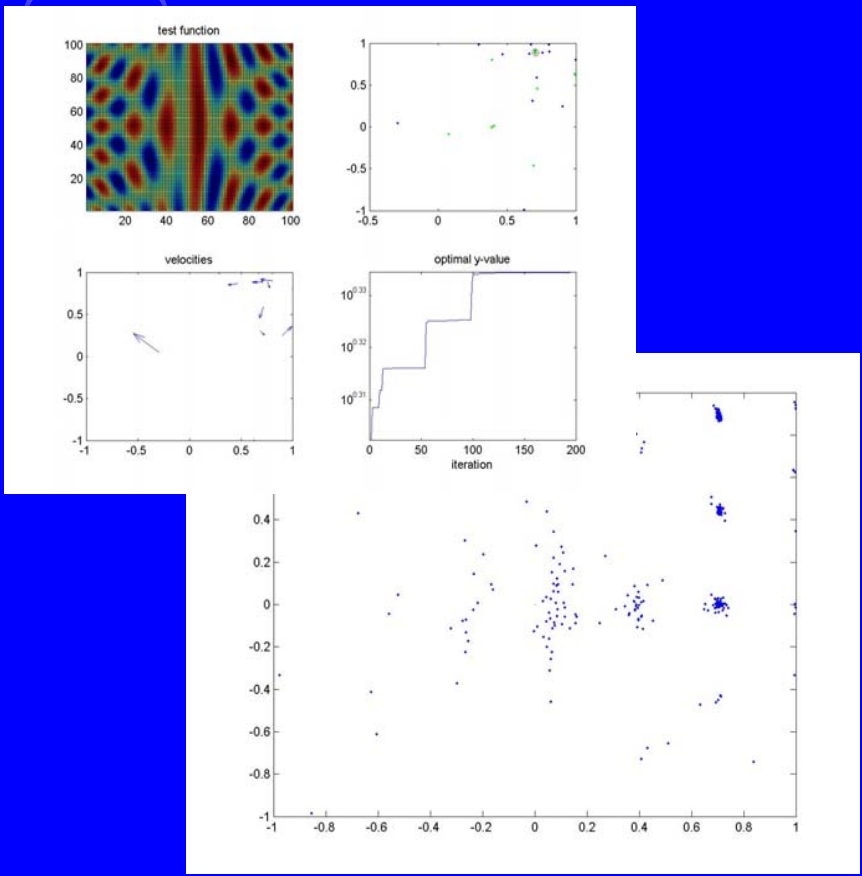
Example of a GP-generated function

$$y = 7.05 - 0.024 * ((x^2 + x^3/x^4))$$

- Based on artificial evolution of millions of potential nonlinear functions => survival of the fittest
- Many possible solutions with different levels of complexity
- The final result is an explicit nonlinear function
- Better generalization capabilities than neural nets
- Low implementation requirements
- Time delays
- Sensitivity analysis of large data sets
- Slow (several days of computational time)

Particle Swarm Optimization

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



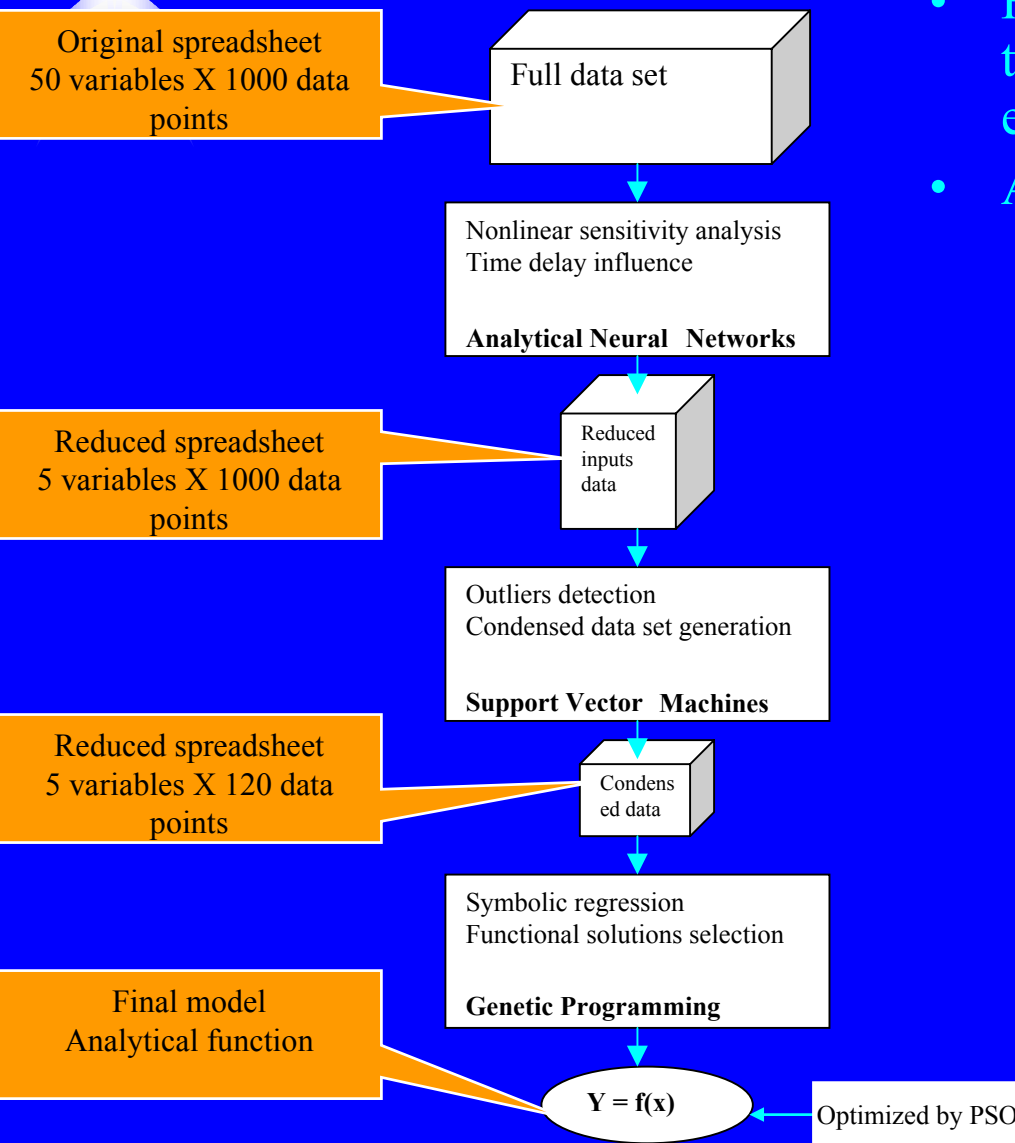
Global best term

$$v_i \rightarrow w_i(v_i + c_i \phi_i(h_i - x_i)) + c_g \phi_g(\xi_g - x_i)$$

Local best term

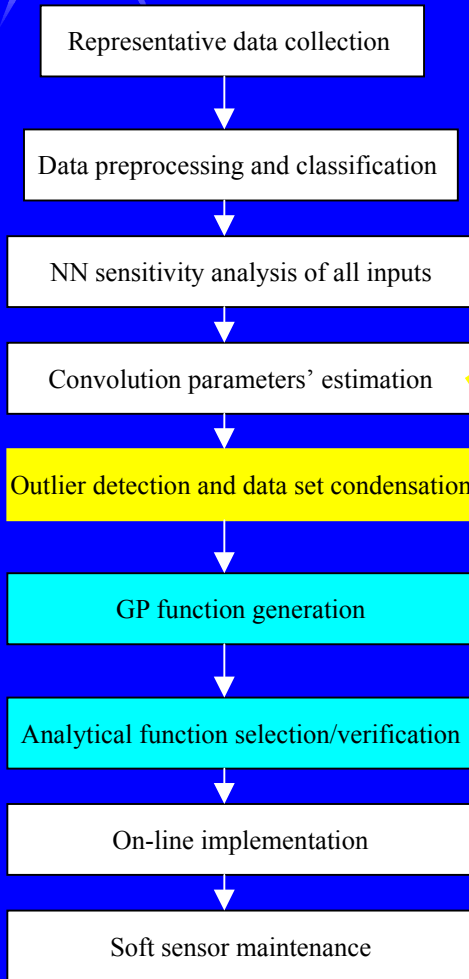
$$x_i \rightarrow x_i + v_i$$

Integrated Methodology for Empirical Models Development

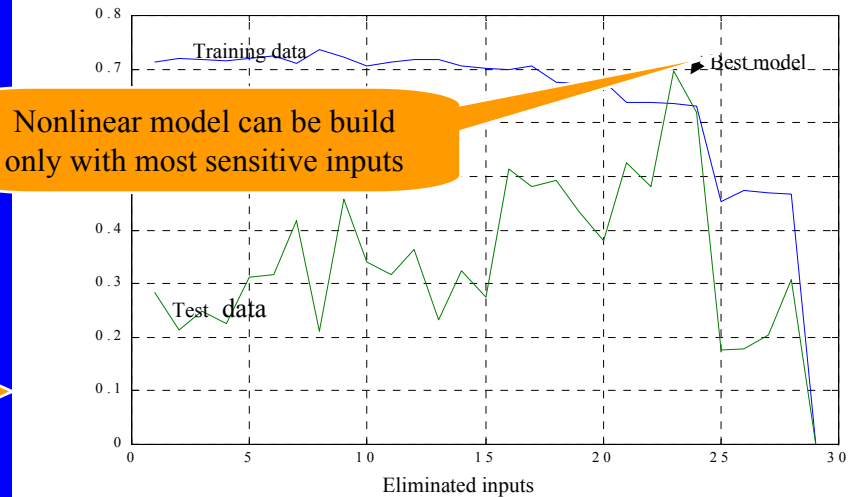


- Hybrid approach integrating multiple technologies exploits the strengths of each
- Advantages:
 - Fast development (5-10 days)
 - Robust performance (compact models)
 - Direct implementation in any Distributed Control System (no need for specialized software)
 - Very low capital cost (only if hardware for data collection is unavailable)
 - Low average cost of ownership (reduced development and maintenance cost - \$7,500/year)
 - Process engineers like it (preferable to black-box models)

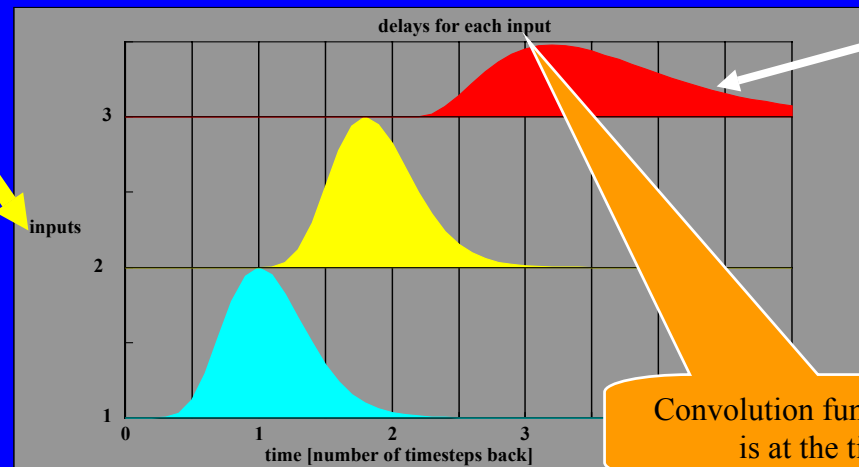
Steps Based on Analytical Neural Nets



Nonlinear model can be build only with most sensitive inputs



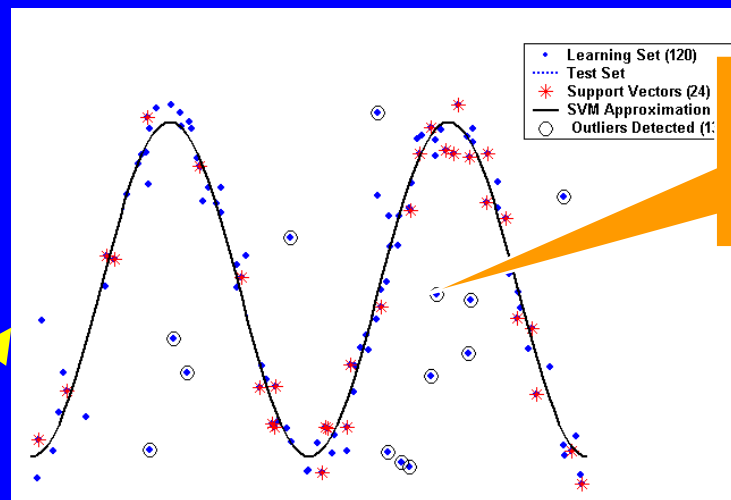
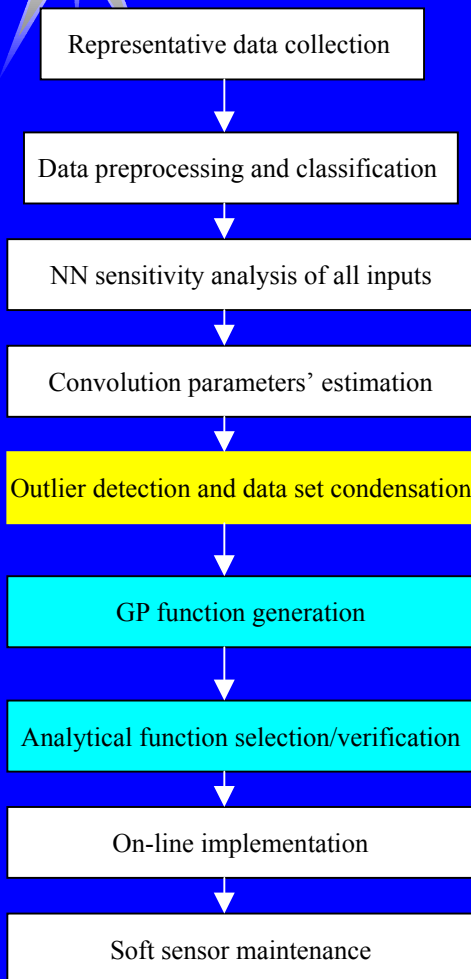
$$f_k(t) = \beta \left(\frac{t}{k} \right)^n e^{-n \left(\frac{t}{k} \right)}$$



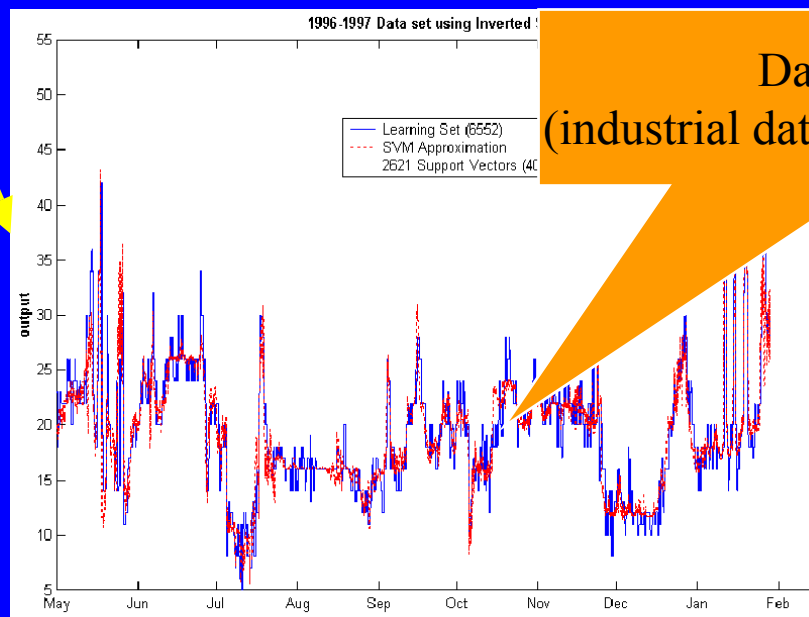
Convolution function maximum is at the time delay

Objective: to supply GP with clean, informative, and parsimonious data set

Steps Based on Support Vector Machines

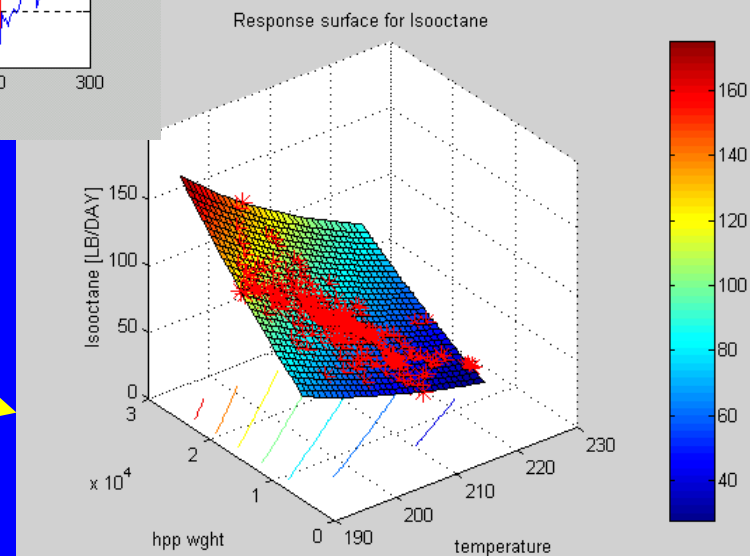
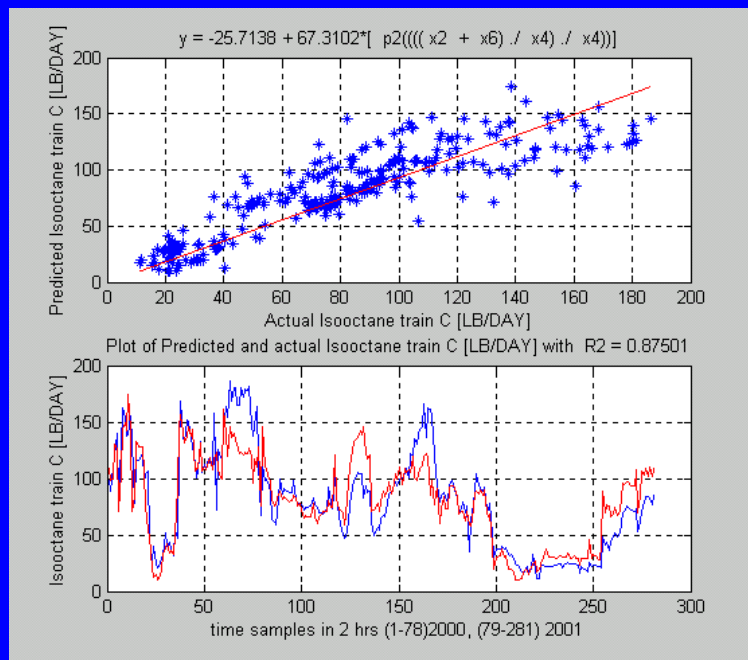
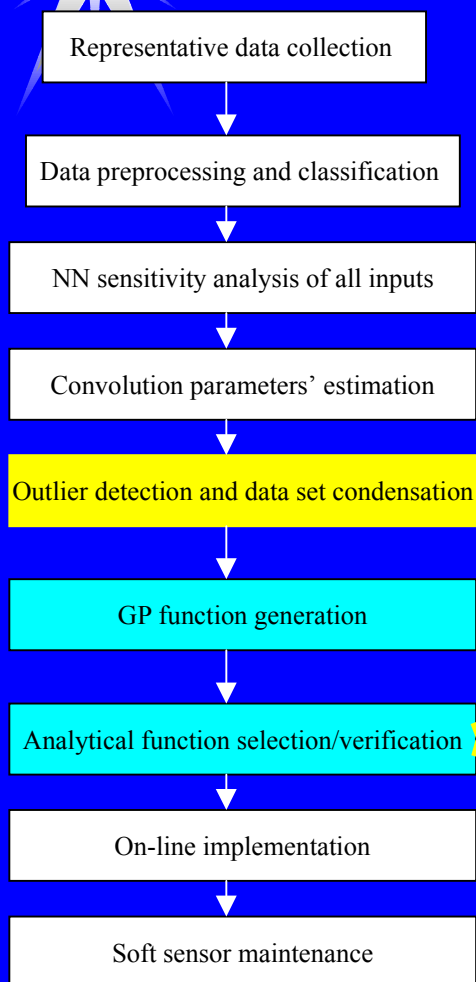


Reliable outlier
detection



Data compression
(industrial data from a chemical reactor)

Steps Based on Genetic Programming



Software tools

Excel, JMP, SIMCA

DAP, Cave, IP21

Representative data collection

Data preprocessing and classification

MATLAB, Excel

NN sensitivity analysis of all inputs

Convolution parameters' estimation

MATLAB Toolbox

Outlier detection and data set condensation

GP function generation

MATLAB & MATHEMATICA Toolboxes

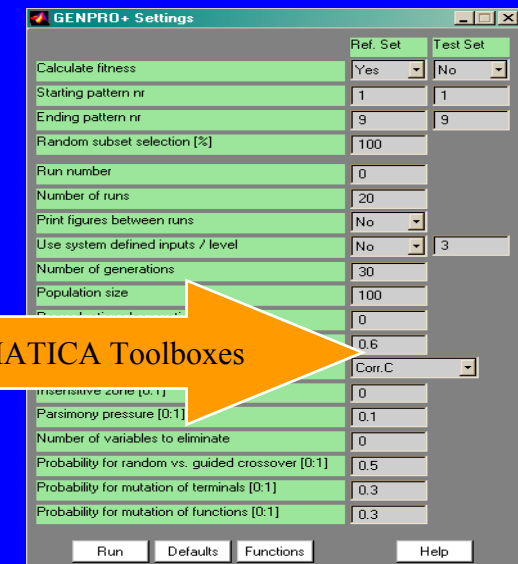
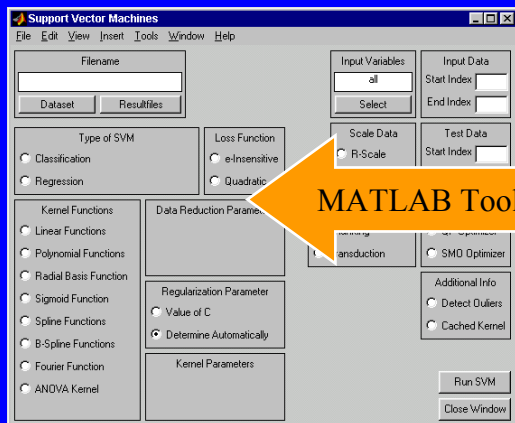
Analytical function selection/verification

G2, MOD, IP21

On-line implementation

Soft sensor maintenance

DAP, Cave, IP21



Robust soft sensors projects work process



Project owner
 Defines requirements and final users
 Delivers data/data access
 Organizes access to experts
 Organizes run-time model implementation



Soft sensor developer
 Builds the soft sensor
 Delivers functional models to the owner
 Analyzes performance for the first year



Final user
 Use soft sensor as any other
 sensors in GPI/IP21

The Dow Chemical Company

The biggest chemical company in USA
(based on \$27.5 billions sales in 2001)

141 in Forbes 500 companies ranking (based on
\$30.1 billions market capitalization on 06/30/2002)



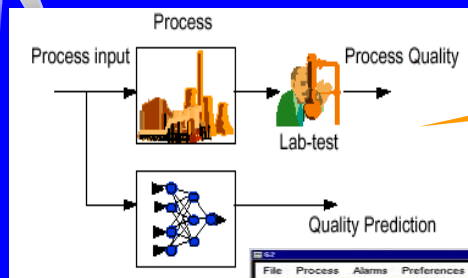
Living.
Improved daily.

> \$1.1 billion R&D

>6,000 researchers

Over 3000 products in 259 plants across 170
countries

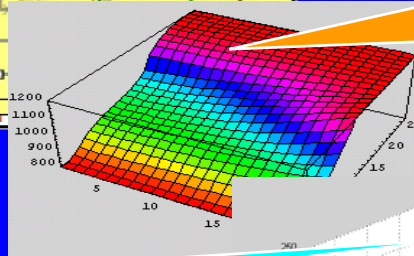
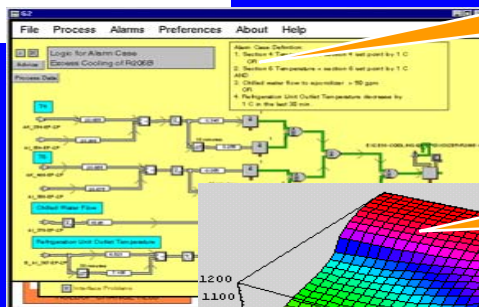
Hybrid Intelligent Systems in Dow



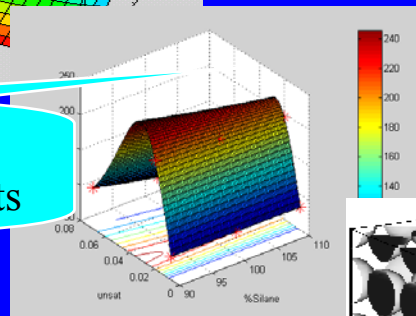
Robust Soft Sensors
Mass-scale on-line empirical models

Automated Operating Discipline
Consistent intelligent on-line supervision

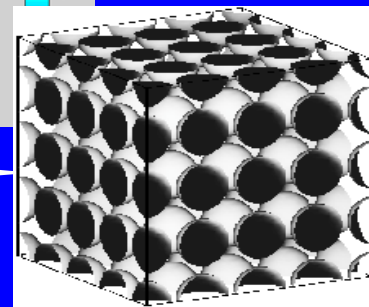
Empirical Emulators of Fundamental Models
Effective on-line process optimization



Nonlinear DOE based on GP
Minimizing expensive process experiments



Fundamental model building based on GP
Accelerated new product development



Robust soft sensors in action: a soft sensor for holdup prediction in a chemical reactor

Representative data collection

Data preprocessing and classification

NN sensitivity analysis of all inputs

Convolution parameters' estimation

Final GP-based sensitivity analysis

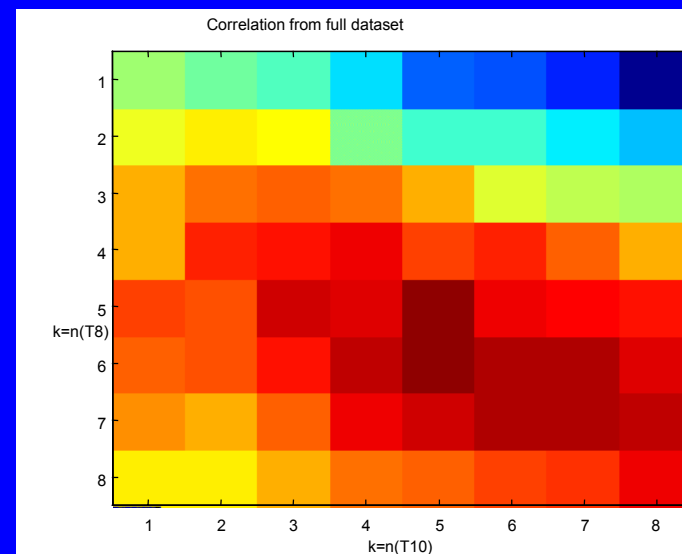
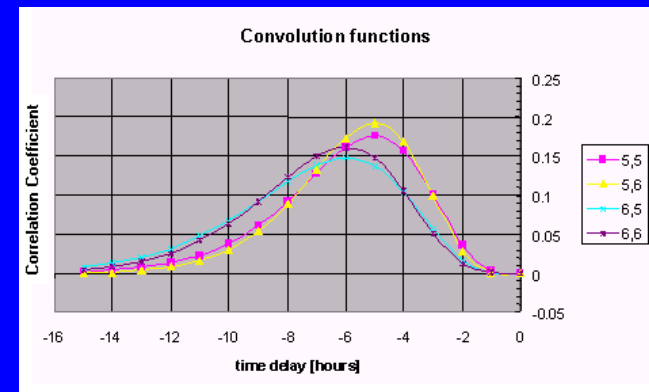
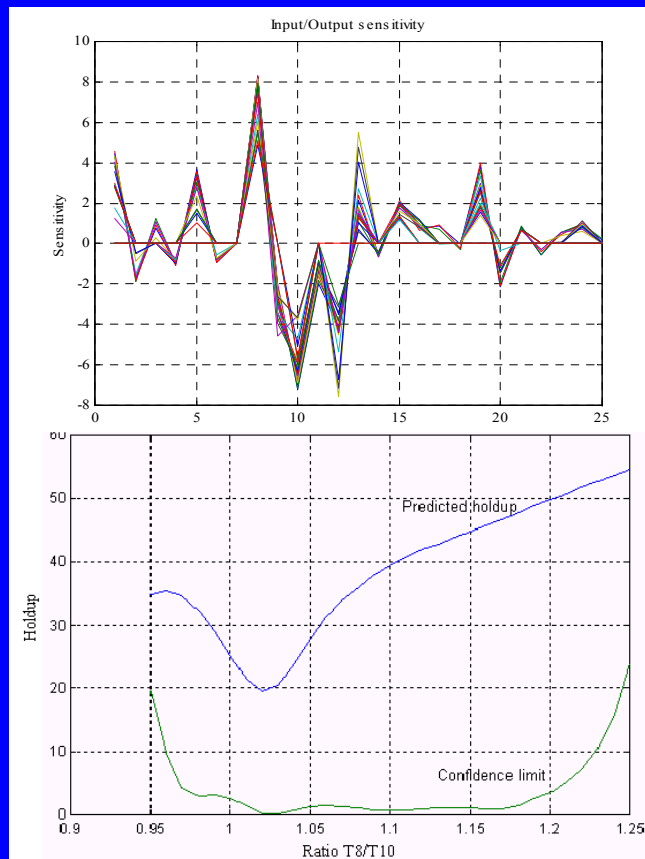
GP function generation

Analytical function selection/verification

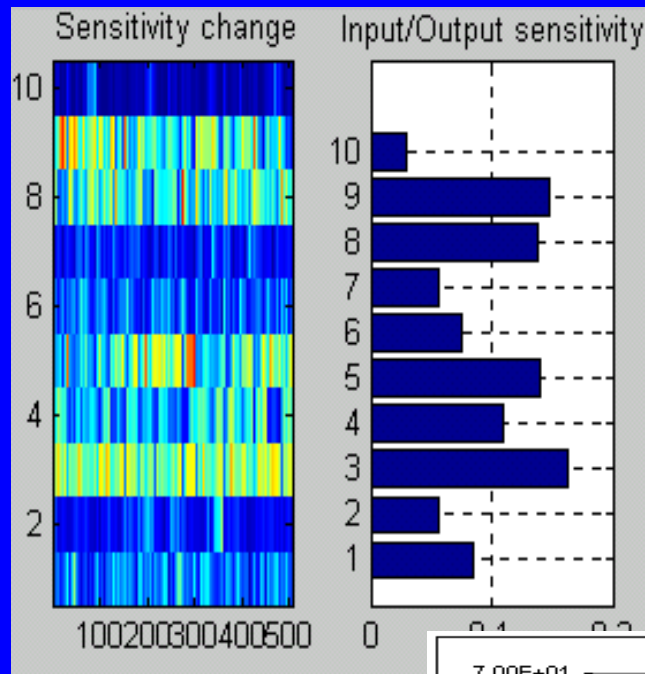
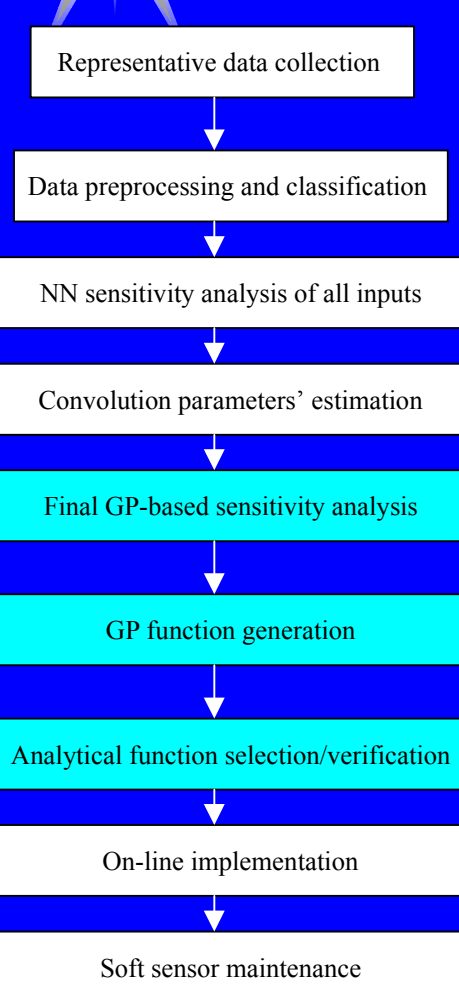
On-line implementation

Soft sensor maintenance

25 inputs (temperatures, flows, concentrations)
1 output (% holdup)
~6500 data points

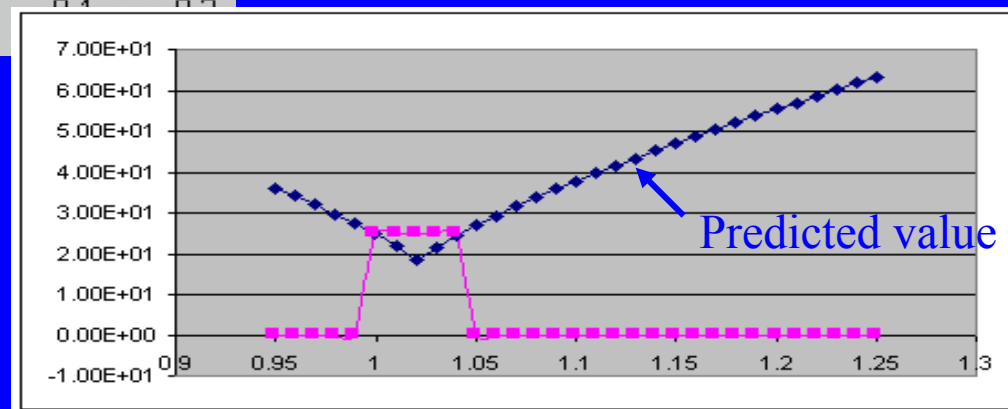


Robust soft sensors in action: a soft sensor for holdup prediction in a chemical reactor



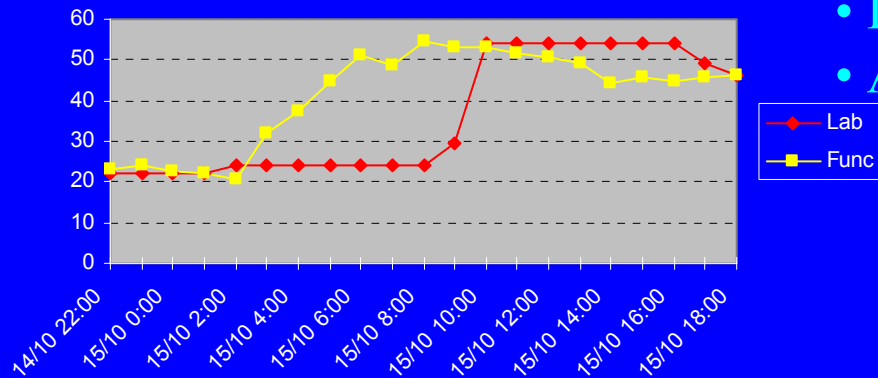
Selected function:

$$y = a + b \cdot \left(e \cdot \left| \frac{x^3}{x^5} - d \right| \right)^c$$



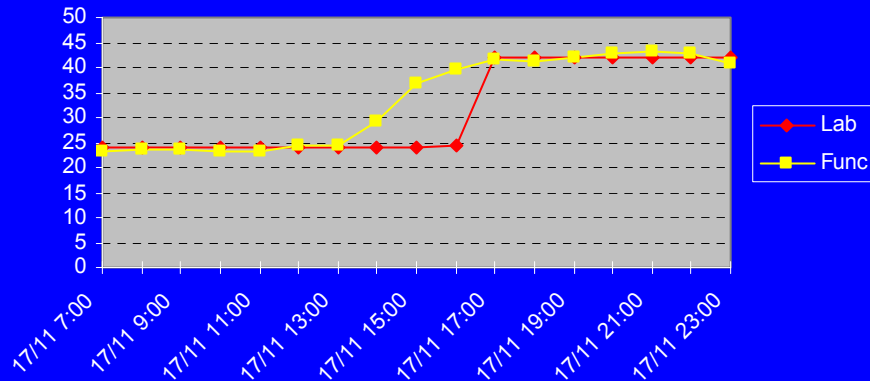
Robustness of GP-generated soft sensor when implemented in other chemical reactors

R206B holdup Oct 15, 1997

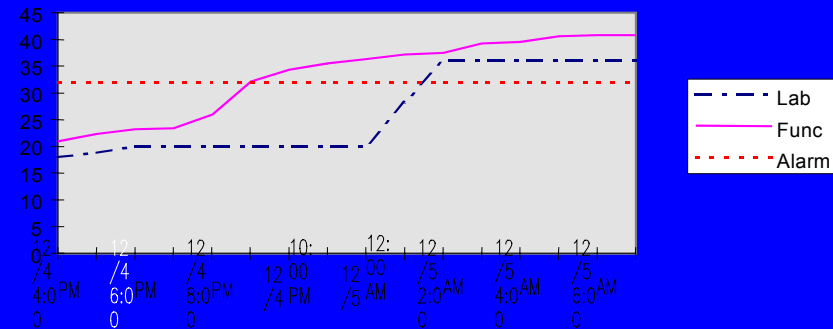


- Implemented in all three similar reactors
- Acceptable performance
 - std Soft sensor => 2.9 - 4.1%
 - std Lab samples => 2%

R206A holdup Nov 17, 1997

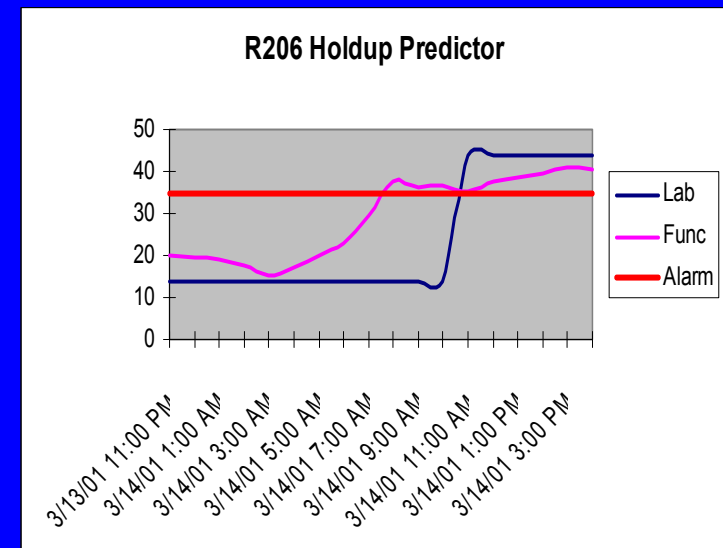
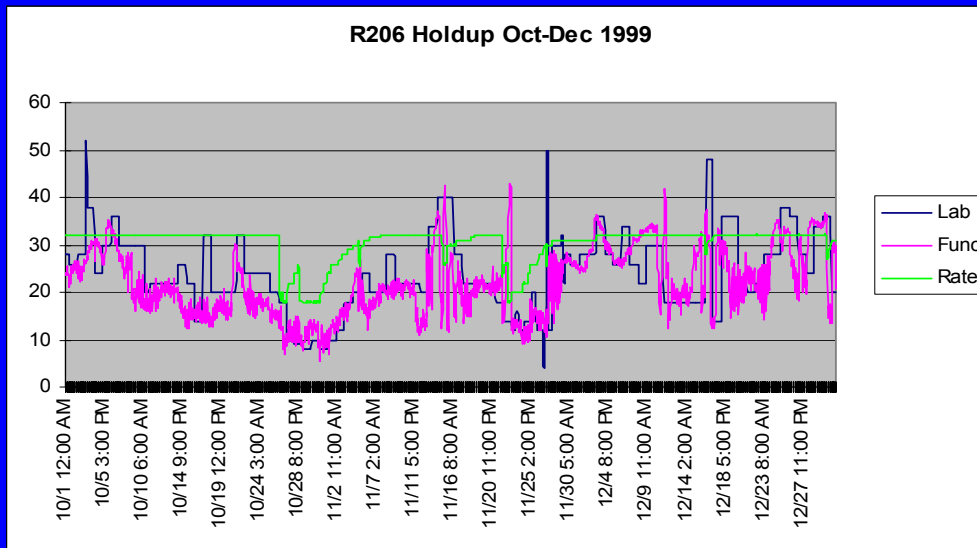


R206 holdup Dec 4-5 1997



Long-term performance

- **In operation since Oct 1997**
- Captured all high holdup cases
- Demonstrated robust performance even after internals changes
- Earned high credibility in plant operation (lab sampling rate was reduced after July'99 => \$190M/year)
- Requires very low maintenance (**no re-training since 1998!!!**)



Another robust soft sensor application: a soft sensor for interface level estimation

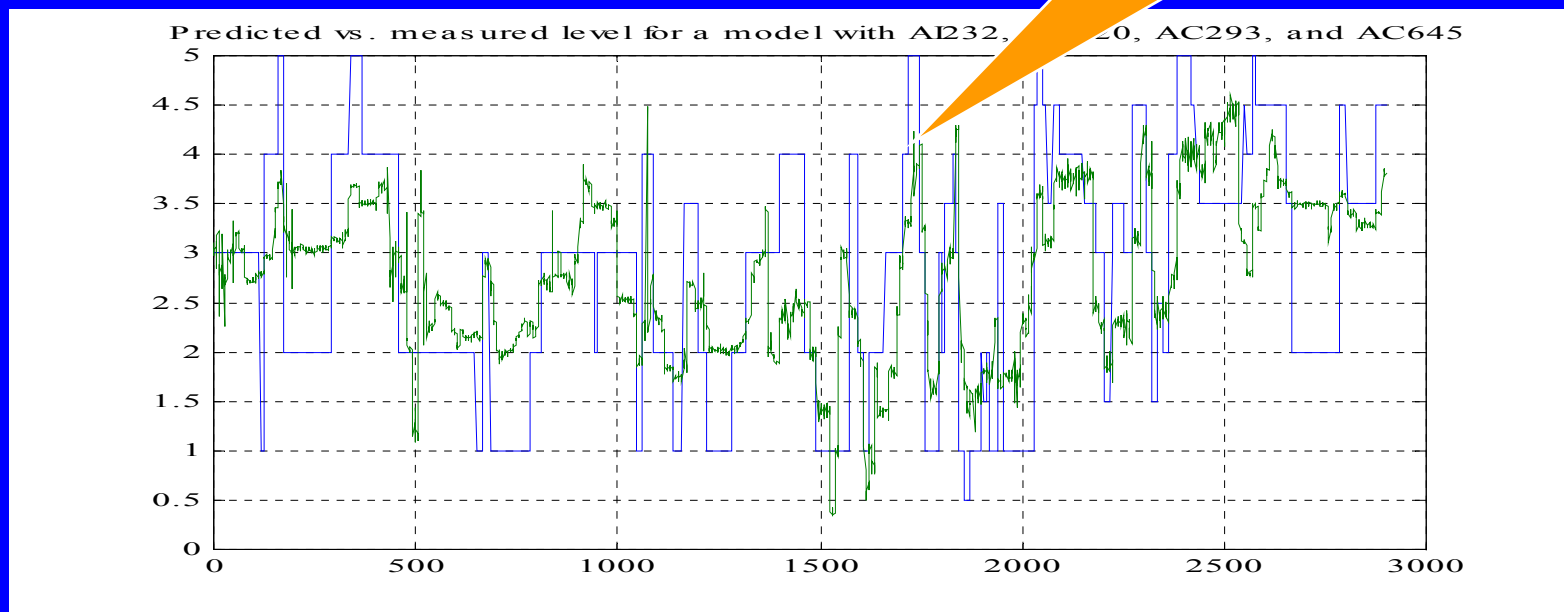
x23 - flow2 x1 - flow1

x25 - ratio setpoint

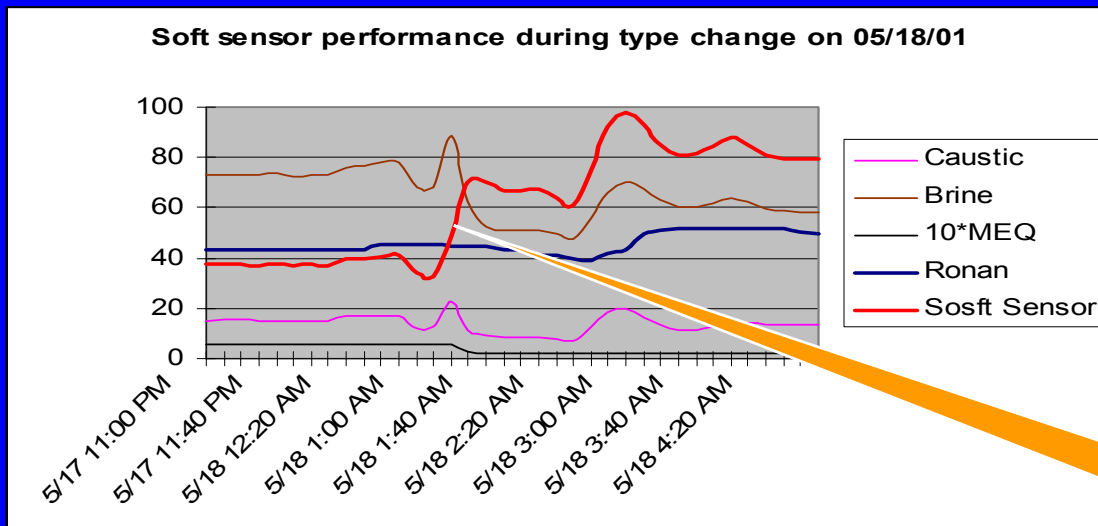
x28 - flow3

$$y = 7.37 - 0.00678 * ((x1 * x23) / (\ln(x28) * x25))$$

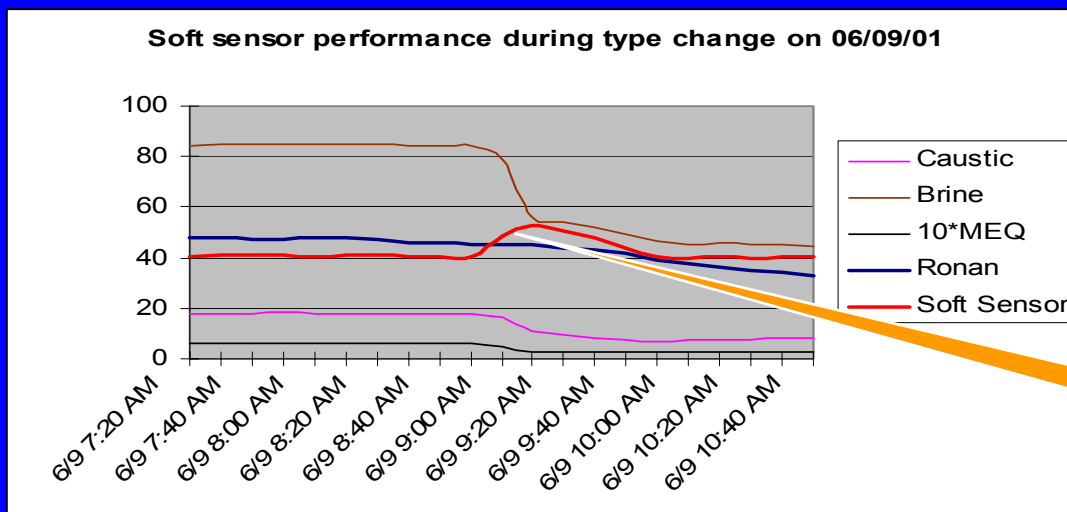
Blue - operators estimates
Green - soft sensor predictions



Robust soft sensor performance during product type transitions



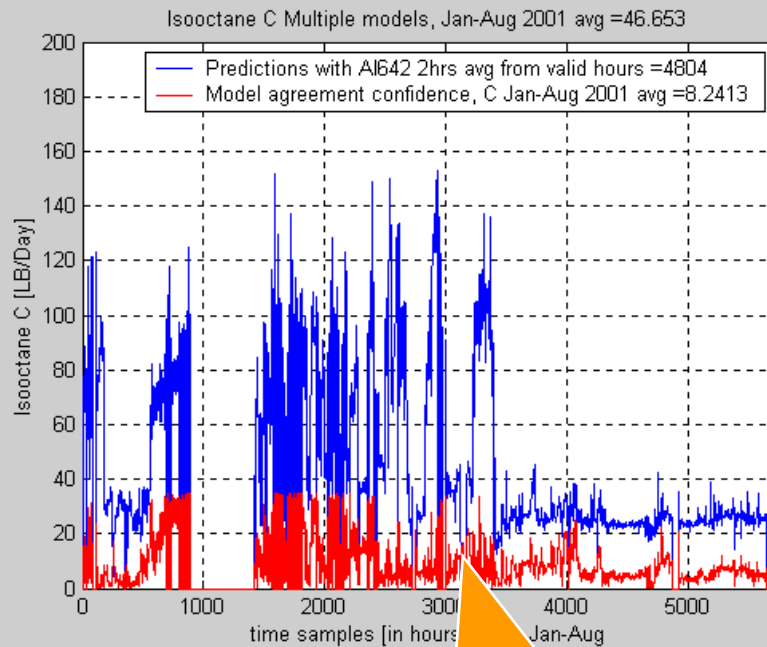
Non-smooth transition



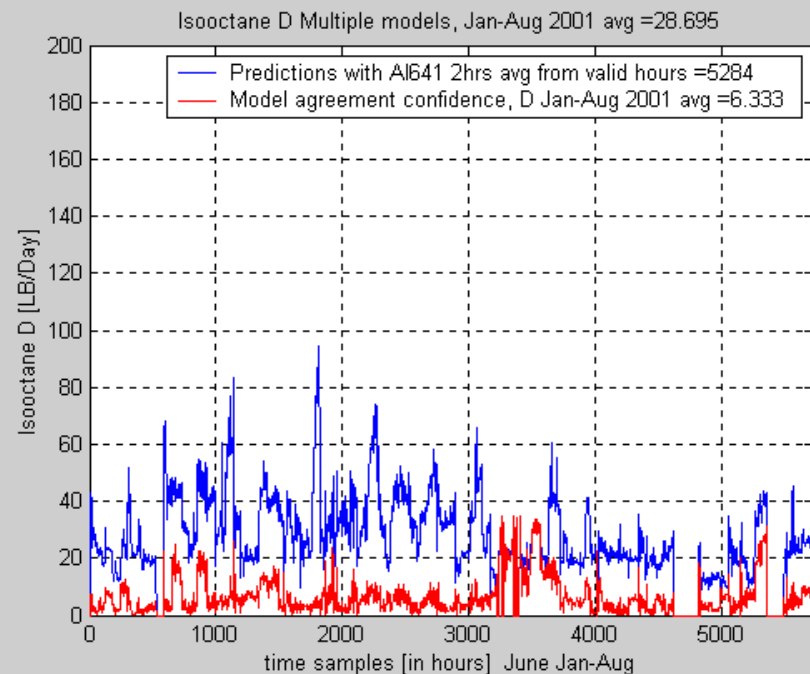
Smooth transition

Robust soft sensor for emission estimation based on multiple linear and nonlinear models

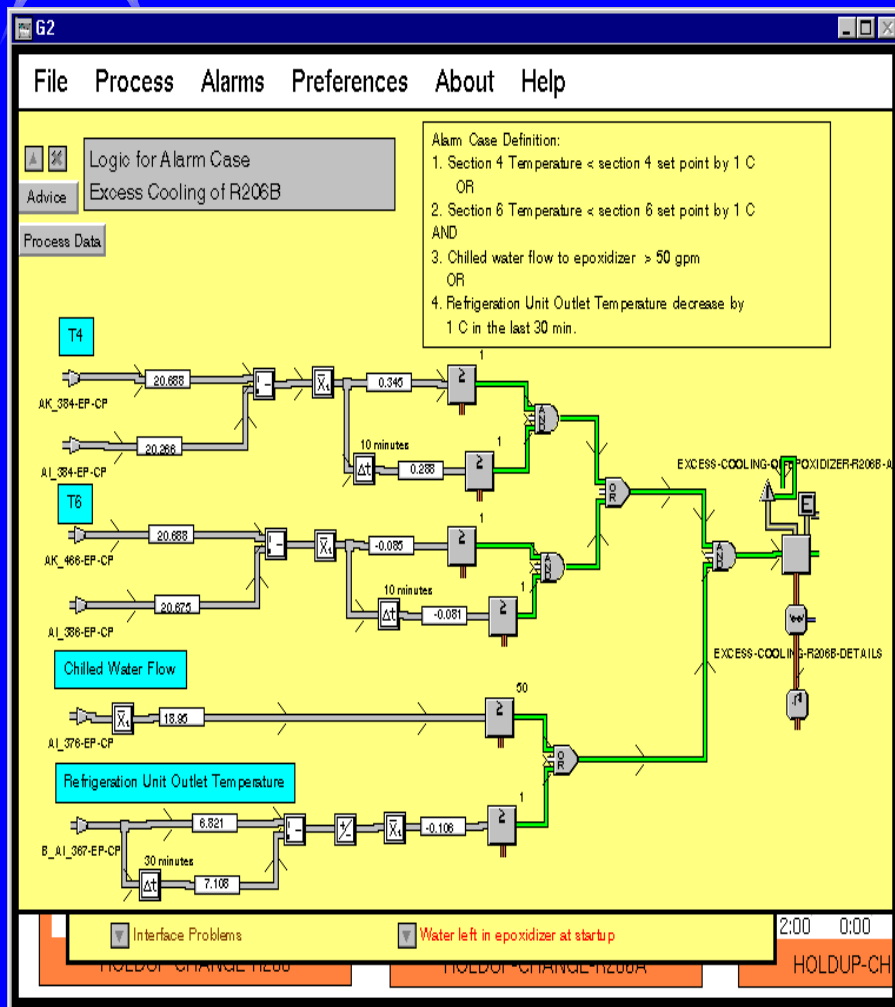
Implemented on-line on 08/29/01



Novelty:
model agreement
confidence indicator



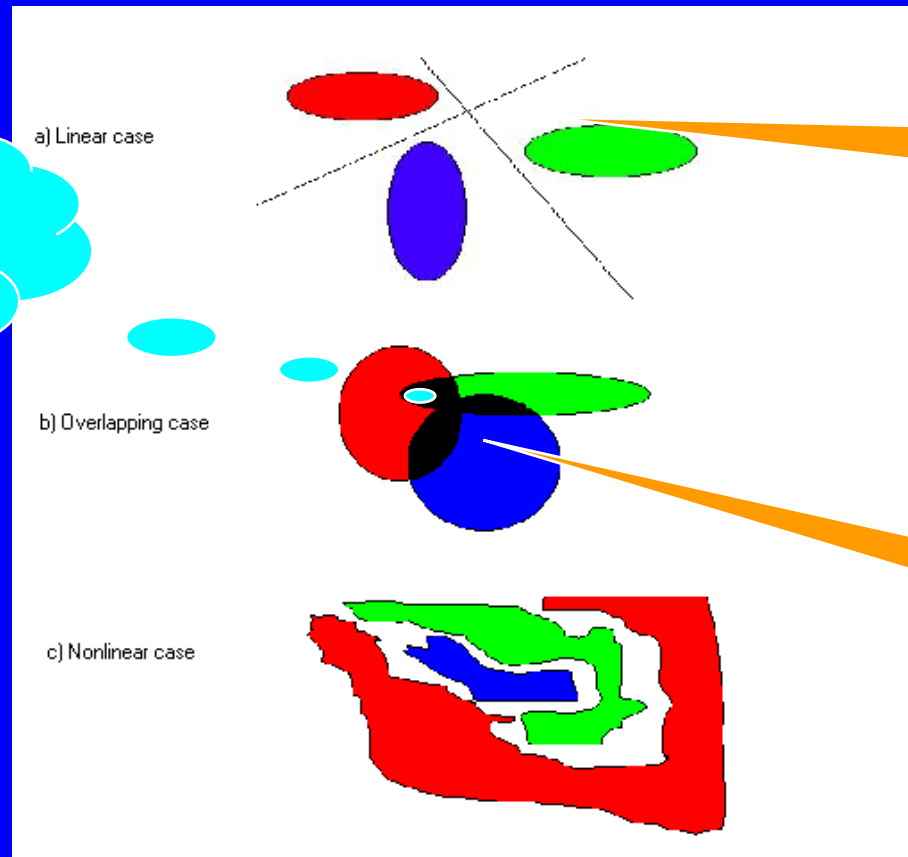
Hybrid Intelligent Systems in Dow: Automating Operating Discipline



- Heuristic rules defined verbally by process engineers/operators
- holdup predictor designed by stacked analytic NN and GP
- all decision blocks have fuzzy thresholds defined by membership functions
- simple empirical models and mass balances
- fundamental model predictions are used in the heuristic rules
- reduced major shutdowns
- reduced lab sampling

Fault diagnosis based on genetic algorithms and support vector machines (project with Leo Chiang)

Criterion:
misclassification rate
for overlapped
data

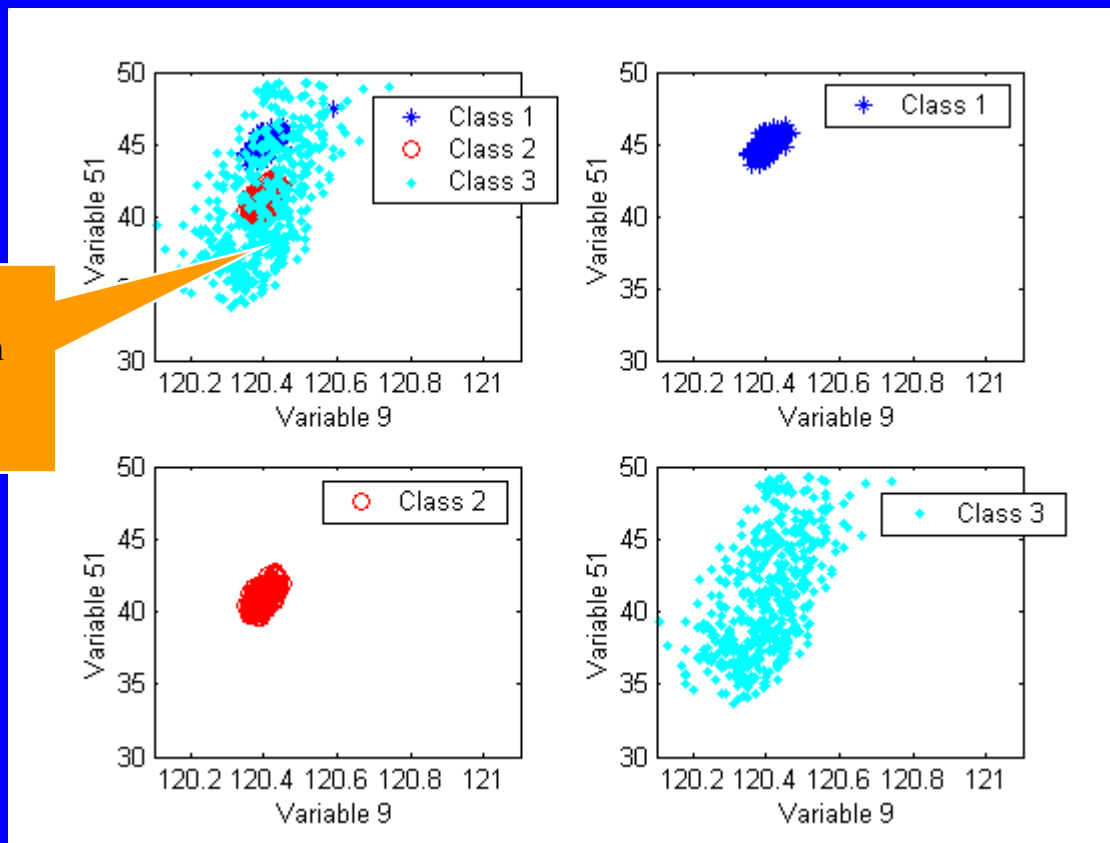


Fisher Discriminant Analysis,
Principle Component Analysis,
K-nearest neighbor, etc

Neural Networks,
Support Vector Machines

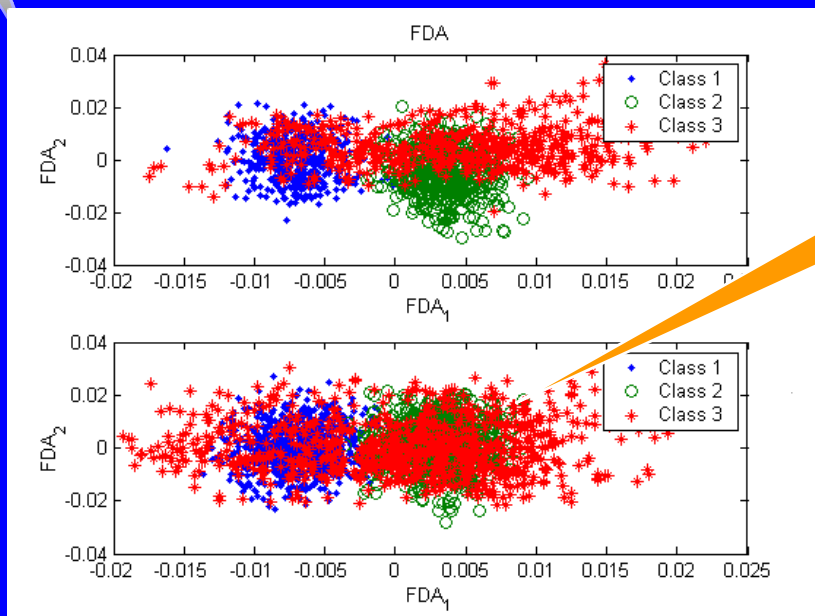
Case study for fault diagnosis based on Tennessee Eastman Process

- Three classes of faulty data, generated by Tennessee Eastman Process (Faults 4, 9, and 11)
- 800 training and 480 test data points
- 52 process variables but only two variables (9 and 51) are related to the faults



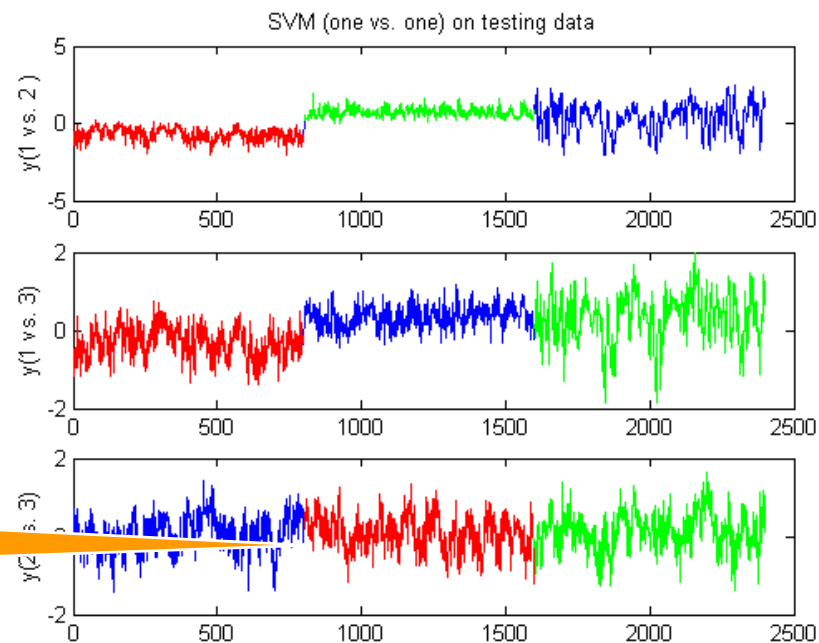
Class 3 overlaps with
Class 1 and Class 2

Fault diagnosis without variable selection



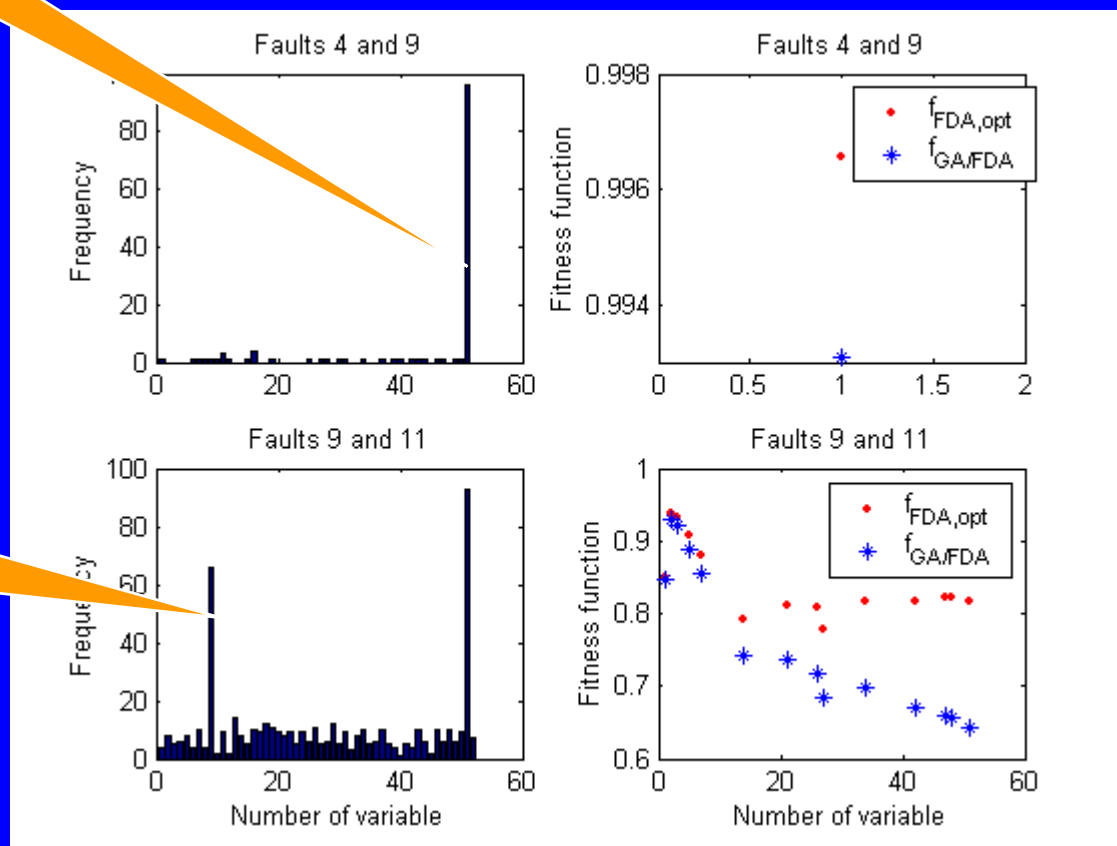
FDA
38% misclassification
on test data

SVM
44% misclassification
on test data



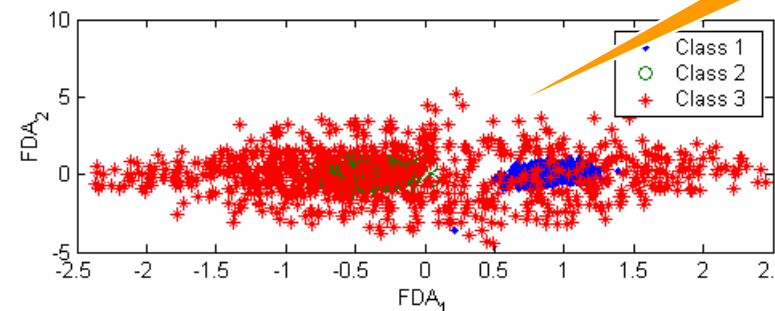
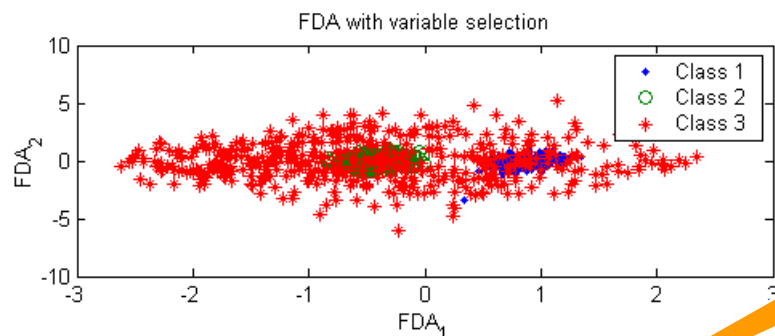
Variable selection based on GA/FDA

Variable 51 is important
for Faults 4 and 9



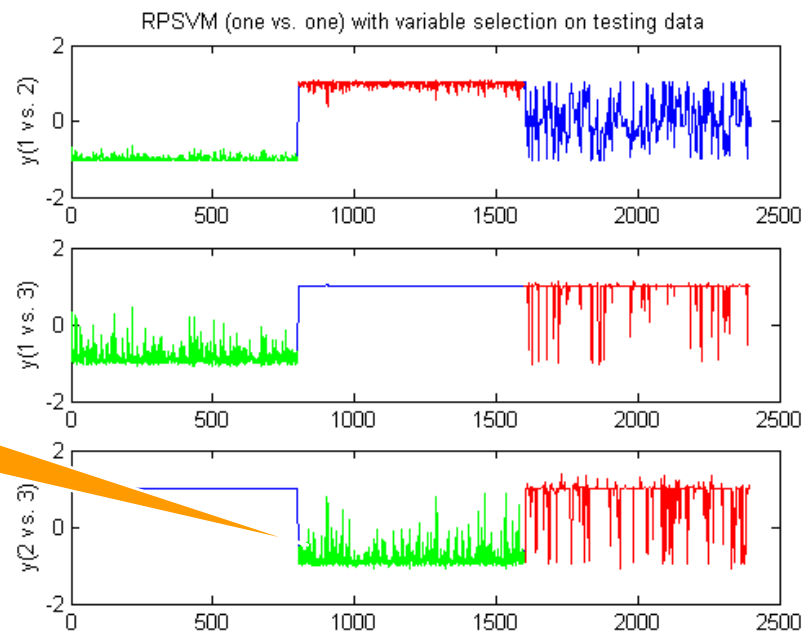
Variables 9 and 51 are
important
for Faults 9 and 11

Fault diagnosis with variable selection based on GA

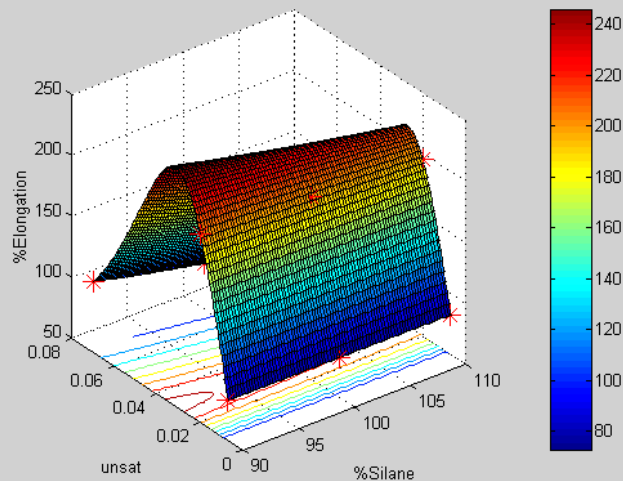


FDA
18% misclassification
on test data

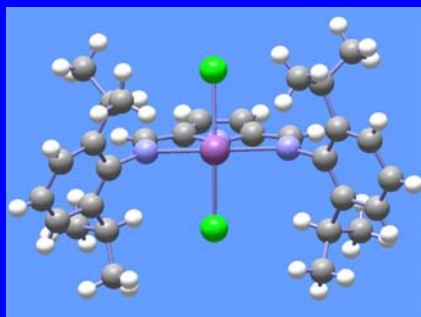
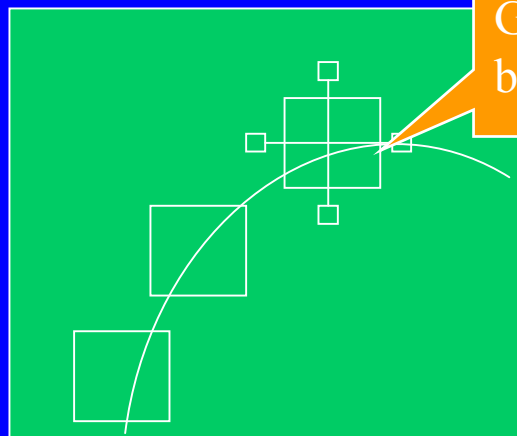
SVM
6% misclassification
on test data



Hybrid Intelligent Systems in Dow: Optimal Design of Experiments



$$S_k = \frac{3.13868 \times 10^{-17} e^{\sqrt{2x_1}} \ln[(x_3)^2] x_2}{x_4} + 1.00545$$



Intelligent CombiChem



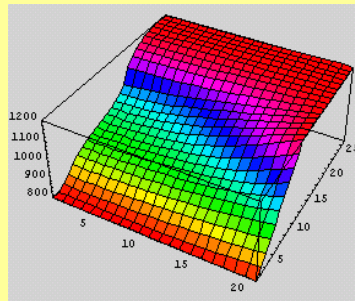
Hybrid Intelligent Systems in Dow: Empirical Emulators

Fundamental models

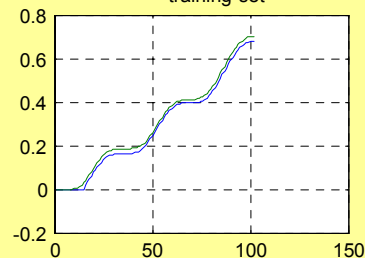
$$\frac{dT}{dt} = \alpha \frac{\partial^2 T}{\partial z^2} - \frac{\Delta H}{C_p \rho} \frac{dc}{dt}$$

$$\frac{dc}{dt} = 3k(1 - c[t])^{\frac{2}{3}} e^{-\frac{E}{RT}}$$

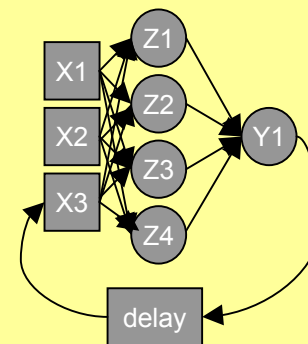
Generate training data



recursive model vs. actual training-set

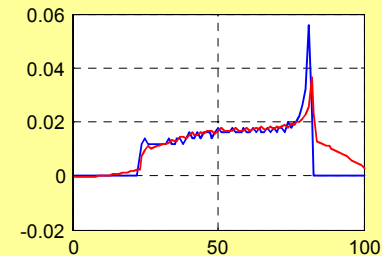
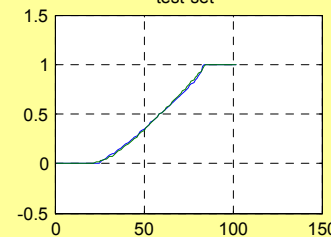


For Neural Nets or Genetic Programming

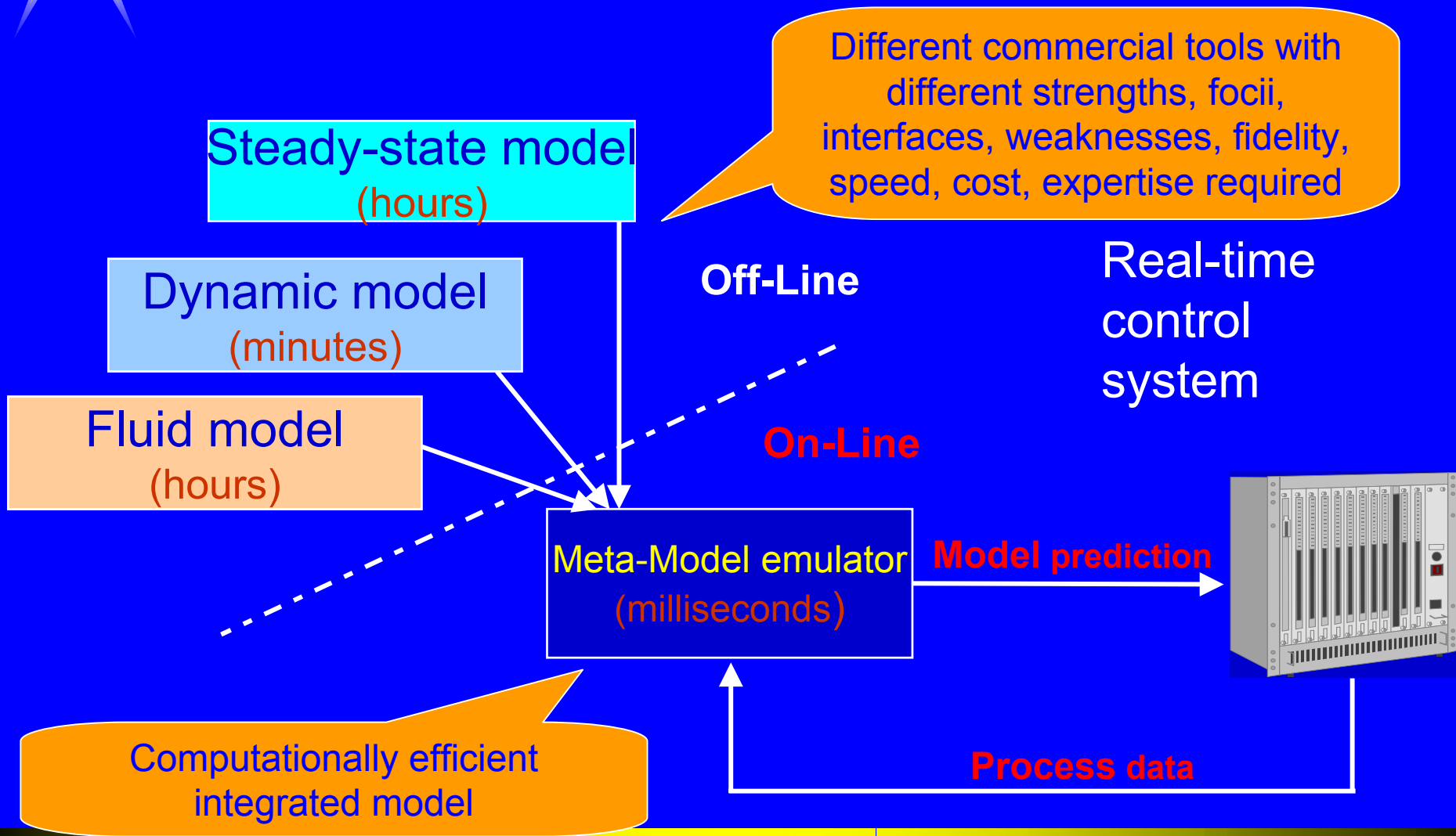


To produce rapid emulation of process behavior

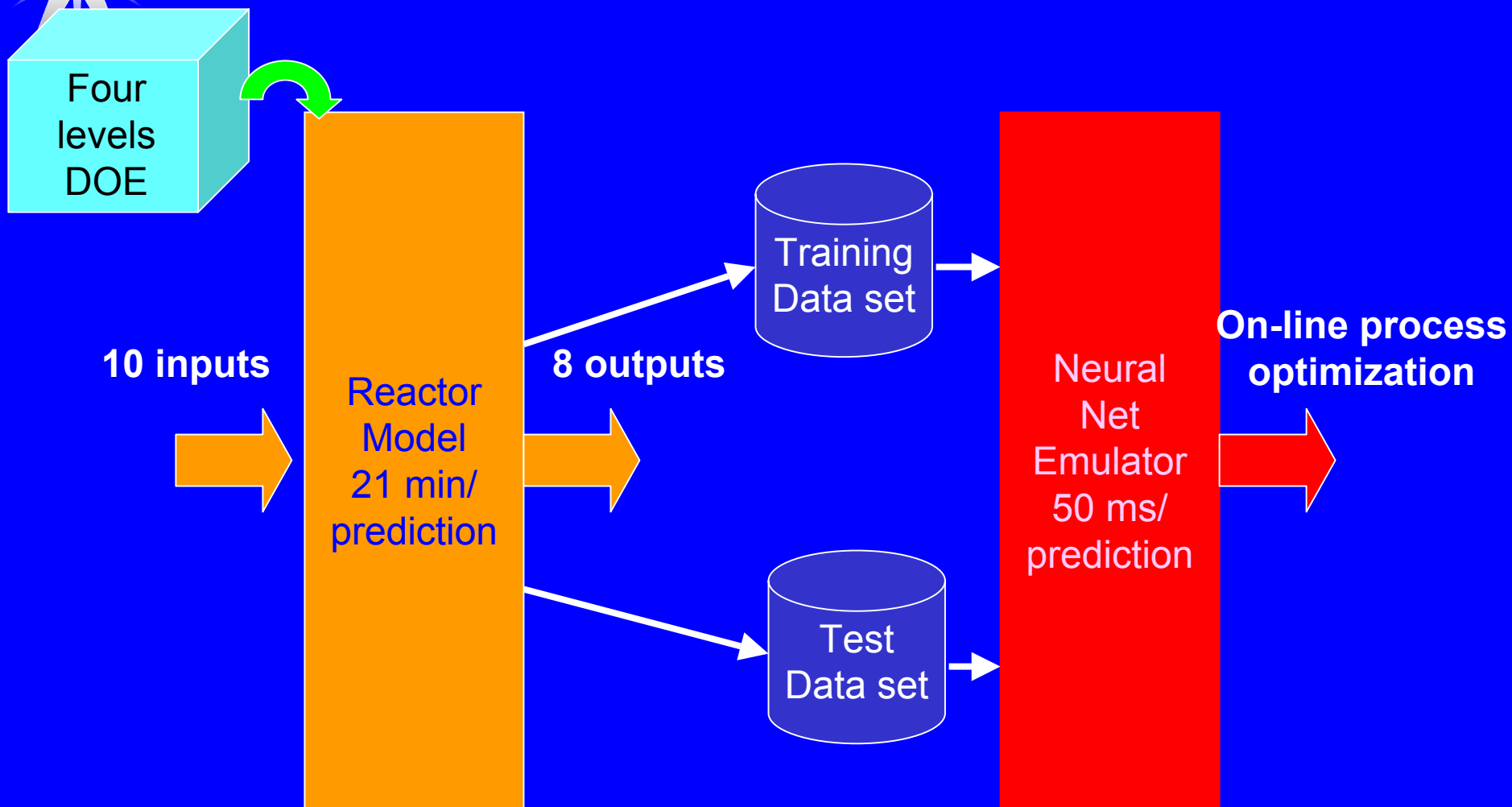
test-set



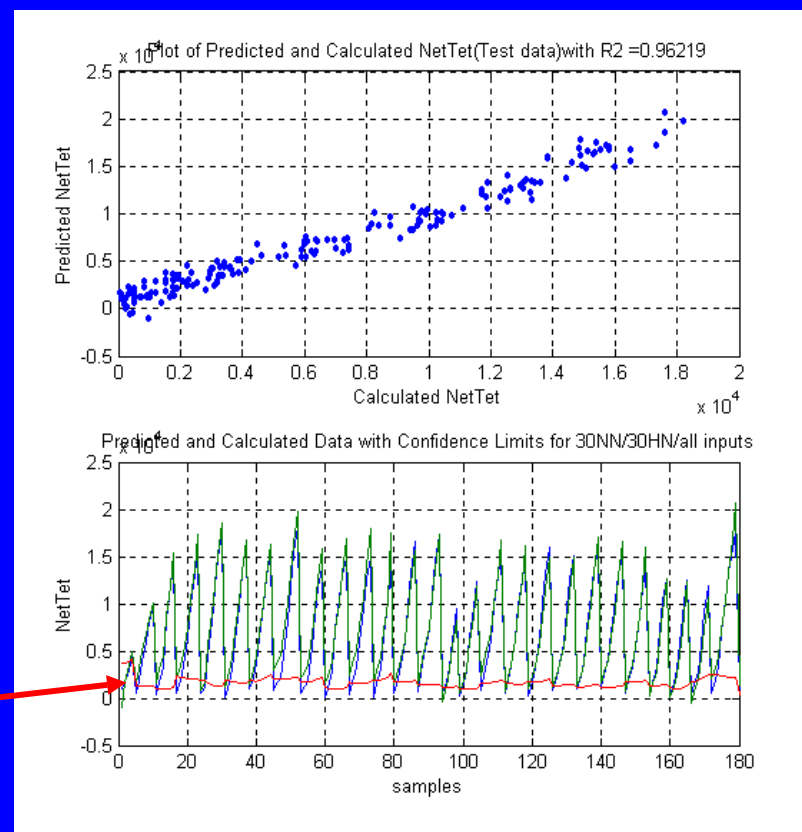
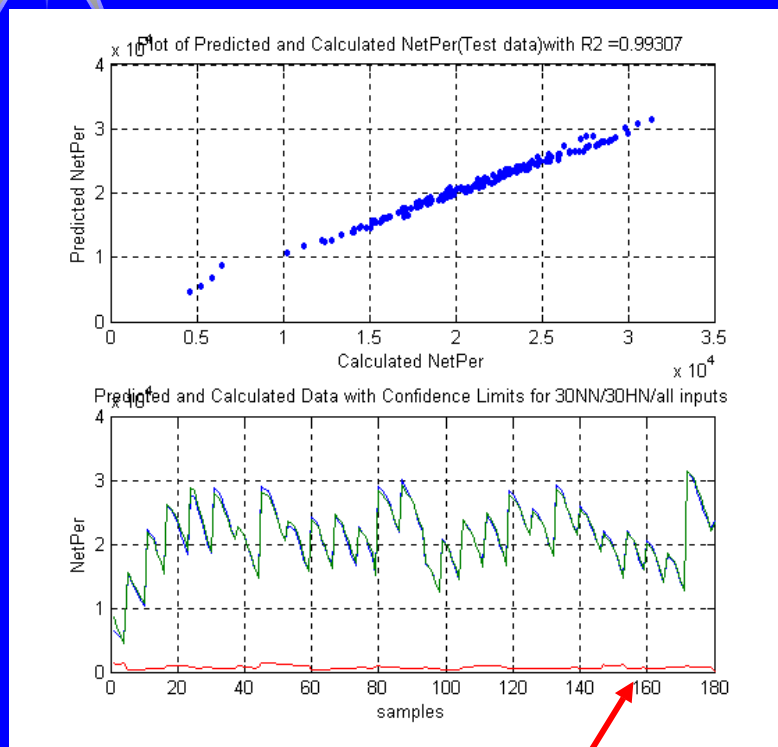
Emulator As a Model Integrator



Emulator of a chemical reactor



Emulator of a chemical reactor based on Stacked Analytic Neural Nets

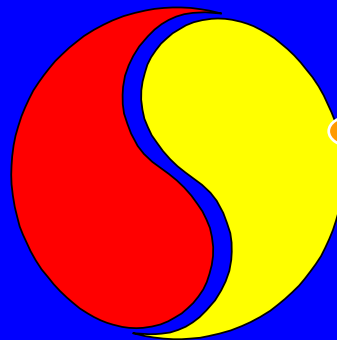


Model agreement indicator



Industrial Data Analysis: Recommended Areas for Future Research

- SVM for regression
- Intelligent GP
- Swarm Intelligence
- Adaptive models
- Intelligent integrators



Integrate and
Conquer!