2009 Quality and Productivity Research Conference

Neural network time series classification of changes in nuclear power plant processes

Karel Kupka
TriloByte Statistical Research,
Center for Quality and Reliability,
University Pardubice
Czech Republic

kupka@trilobyte.cz

Statistical models deployment in industry, technology and research

Quantification

Model building

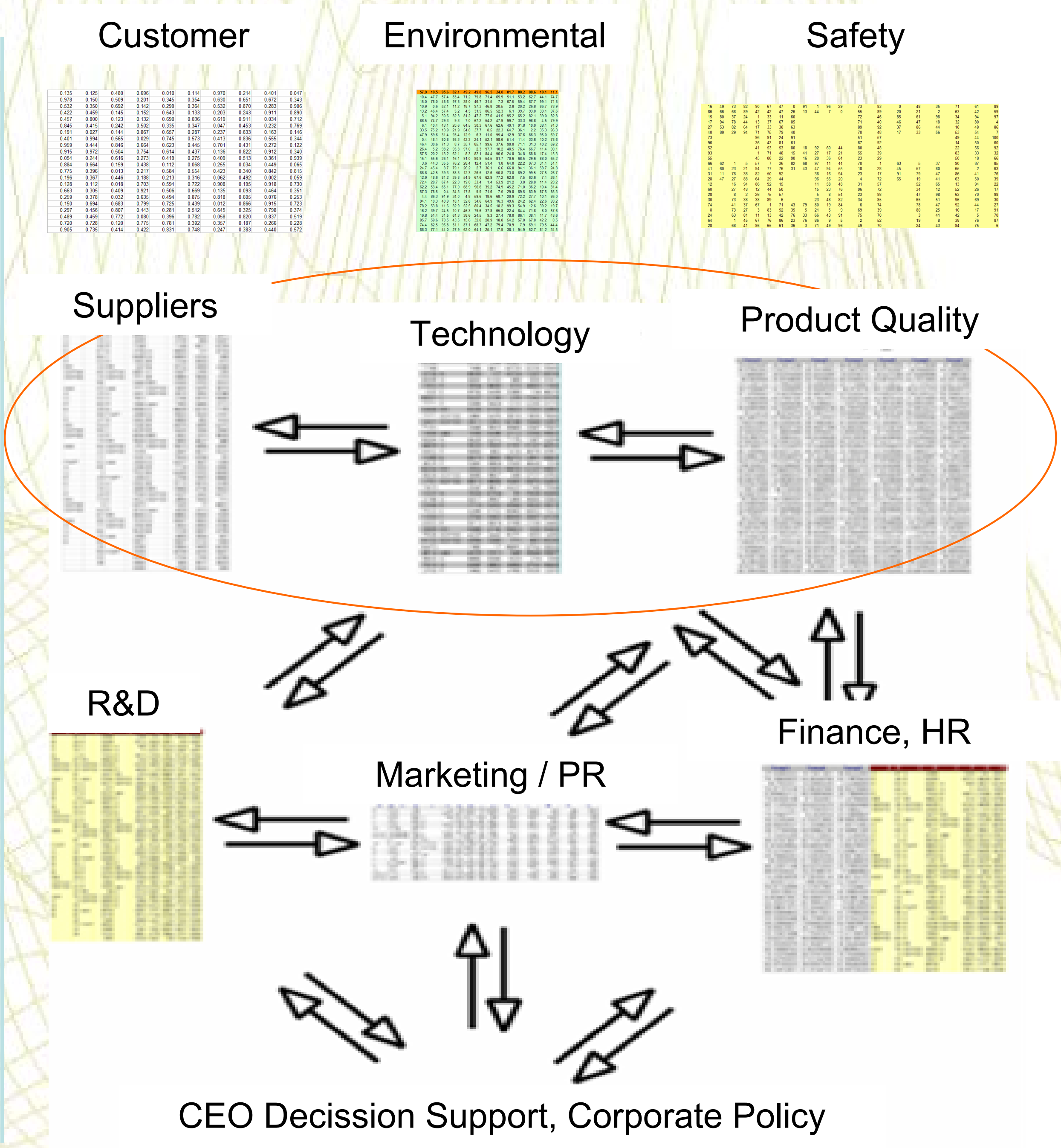
Evaluation

Interpretation

Decission Support

Profit

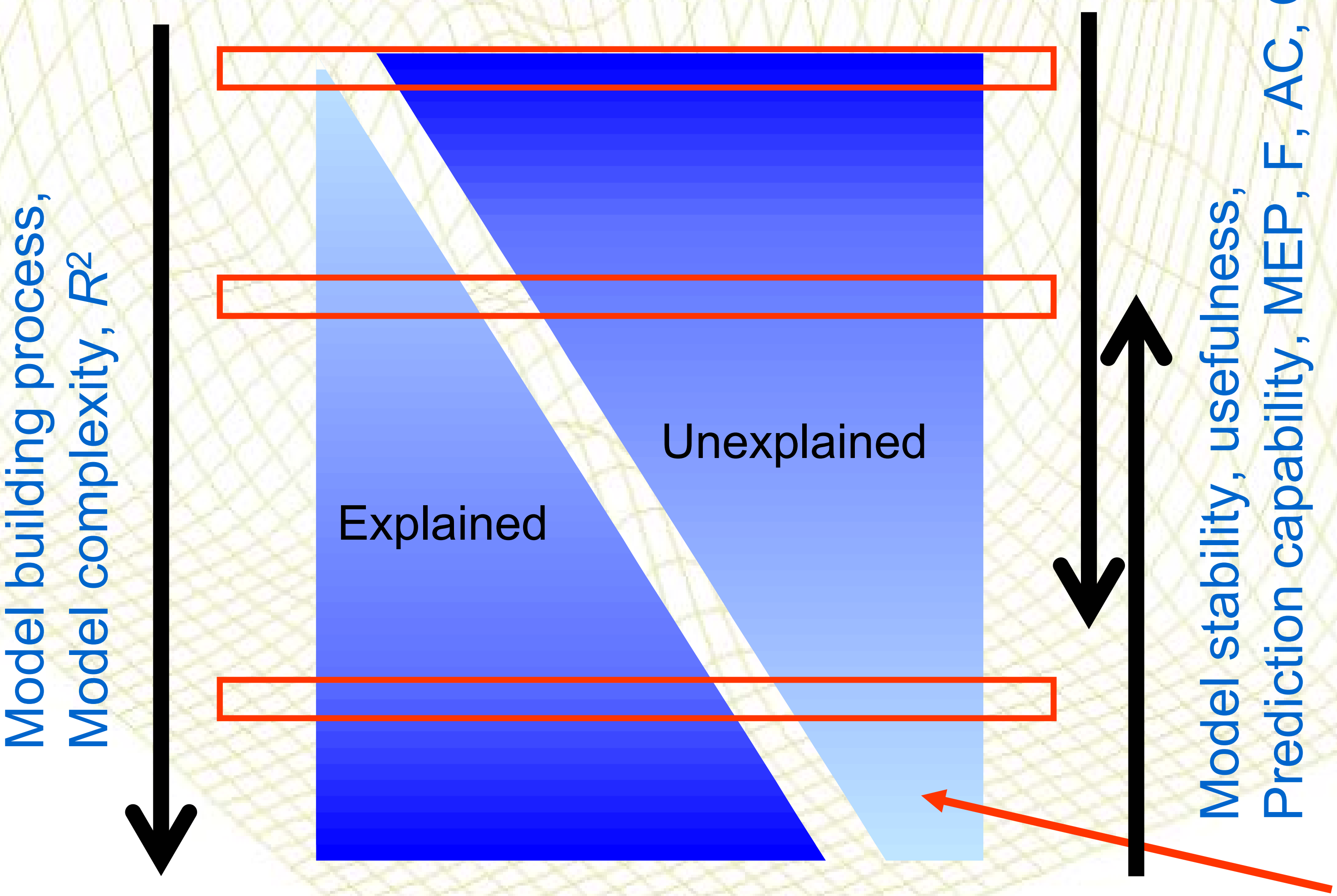
Data and Information Flow in a Company



Model building process

$$Q \propto \frac{1}{\sigma^2}$$

X = MODEL + NOISE



- Gradual variability explanation/ reduction
- Criterion: practically applicabe (empirical) prediction model

Unexplained (-able), inherent variability

Prediction model application examples

Steel works – Predict/classify mechanical&physical properties based on chemistry&technology parameters

Brewery – Find relationship between (bio-)chemistry and sensoric evaluation/classification of beer

Pharmaceutical/Cosmetics – Find quantitative structure – activity relationship (QSAR)

Financial – Classify loans, credit scoring, risk analysis

Climatology/Environment – Antarctica-based solar activity and ozone prediction, Industrial air pollution analysis

Regression, PCA, PLS, CART, ANN, ANNTS, SVM

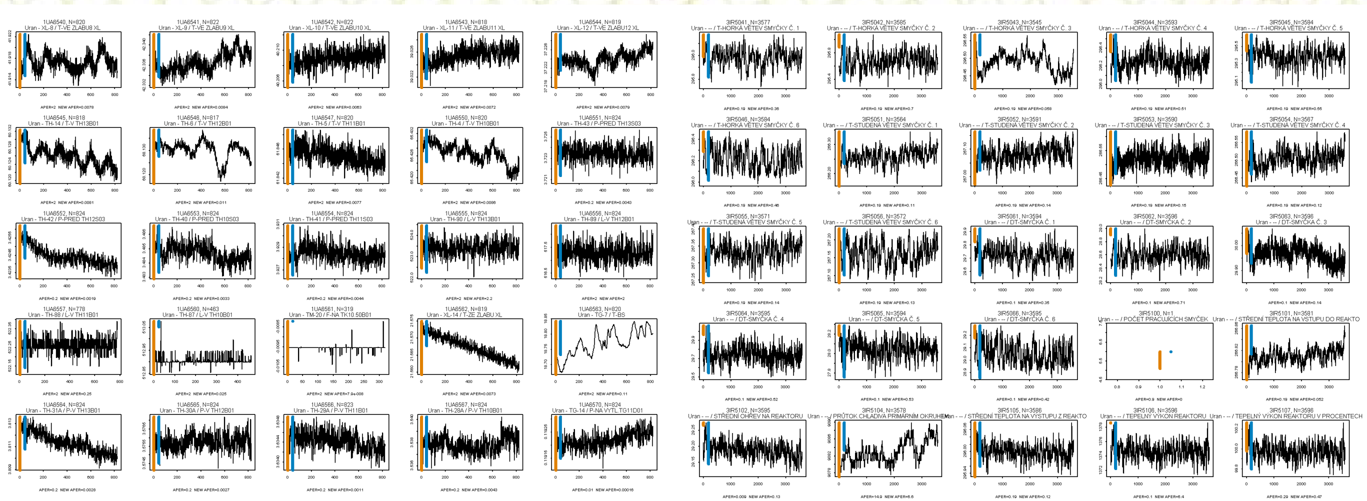
Nuclear power plant time series

The Plant

The Reactor

The Sensors

Temperatures, energy flow rates, meta-stable process modes



Driven by complicated nonlinear local physical processes

Some modes more important than others

Classify known metastable states *Y* of the process using process time series

The idea -

Sample space (*n*) → Variable space (*m*) → Model parameter space (*p*)

$n > m \geq p$

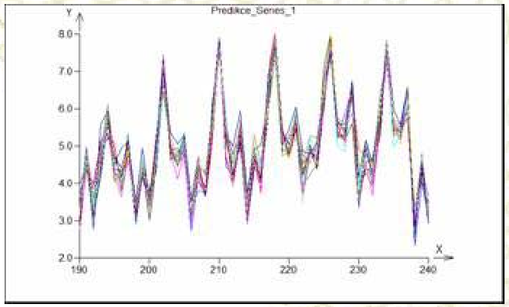
Use parameters as predictors in suitable classification model

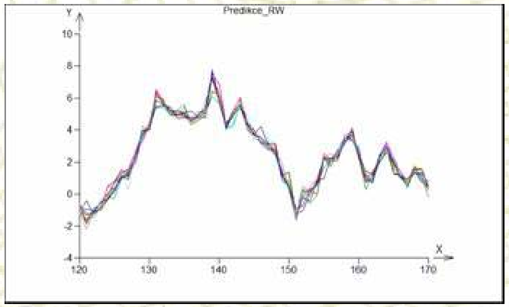
$$Y = \Theta\left(p, \alpha\right)$$

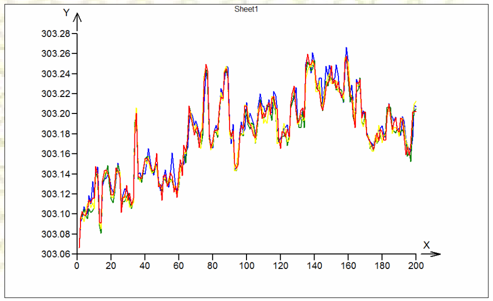
Time series model – (Differential) ANN Autoregresion

$\mathbf{x}_t = Z_{NNAR}\left(\mathbf{x}_{t-k}\right)$ $\quad$ $\mathbf{x}_t = Z_{NNAR}\left(\mathbf{x}_{t-k}, \Delta\mathbf{x}_{t-k}\right)$ $\quad$ $m \geq p$ is usualy not met

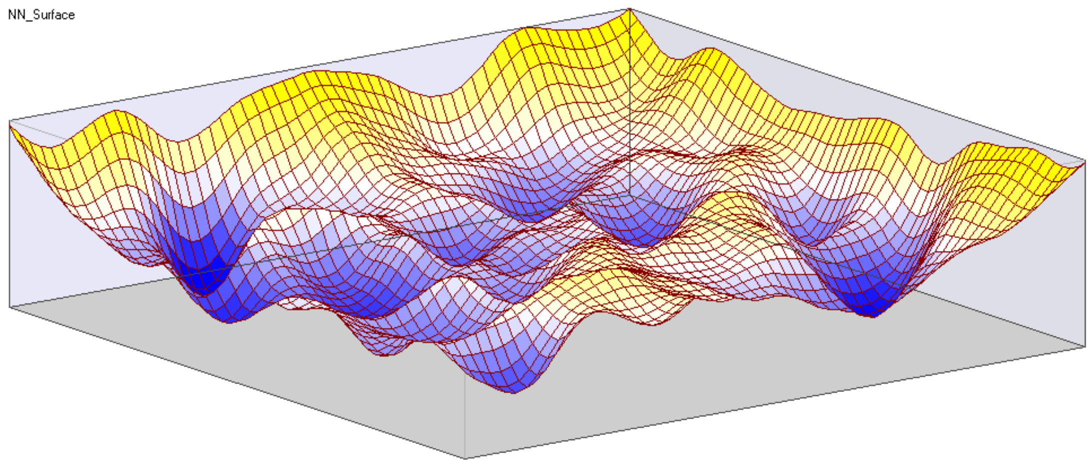
Time series examples -





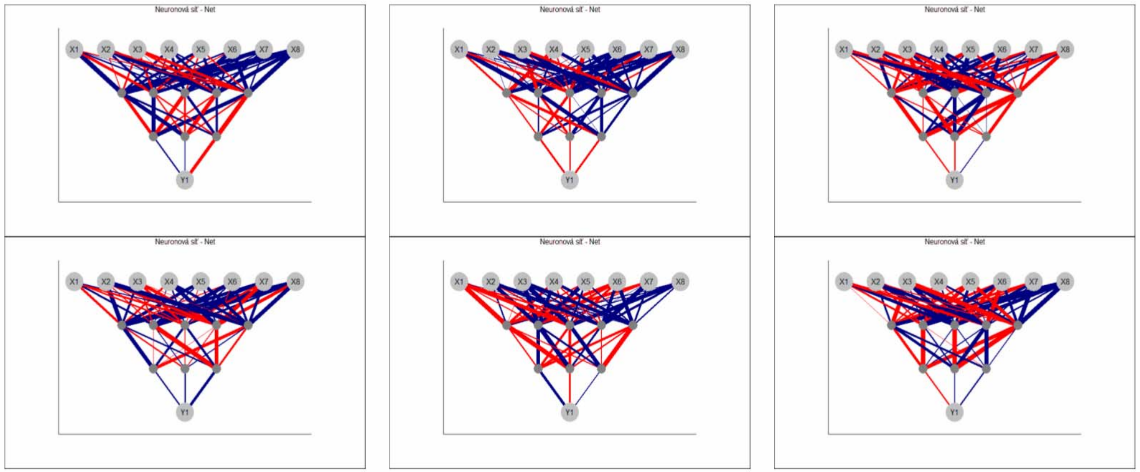


ANN solution stability problem



$$R(\theta) = \sum_j \left(y_j - f\left(\theta, \mathbf{x}_j\right)\right)^2$$

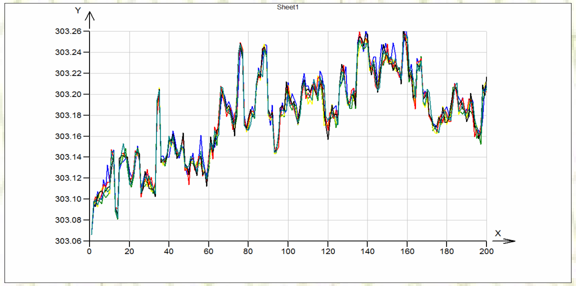
Adaptive Gauss-Newton optimization strategy

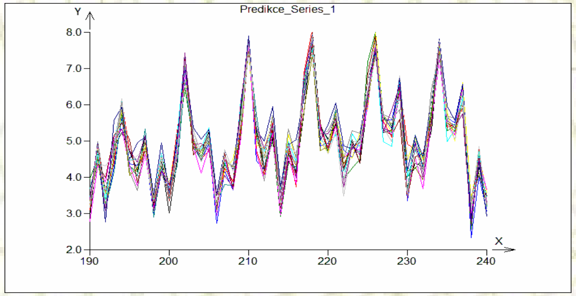


Non-convex error surface
Multiple local minima

Many different „good“ solutions to the same problem

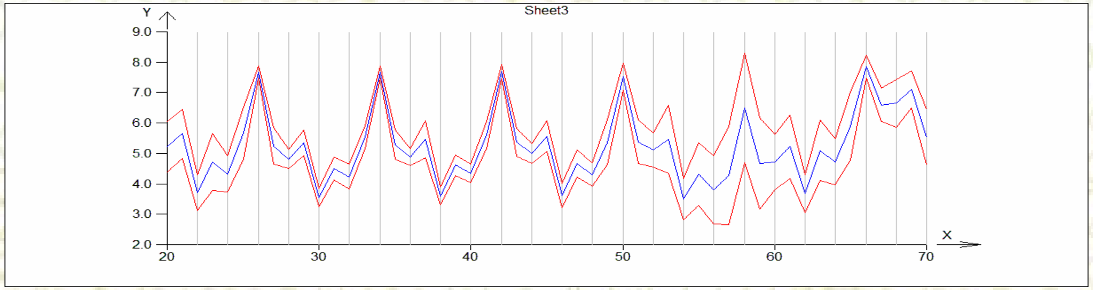
Error Distribution Estimation



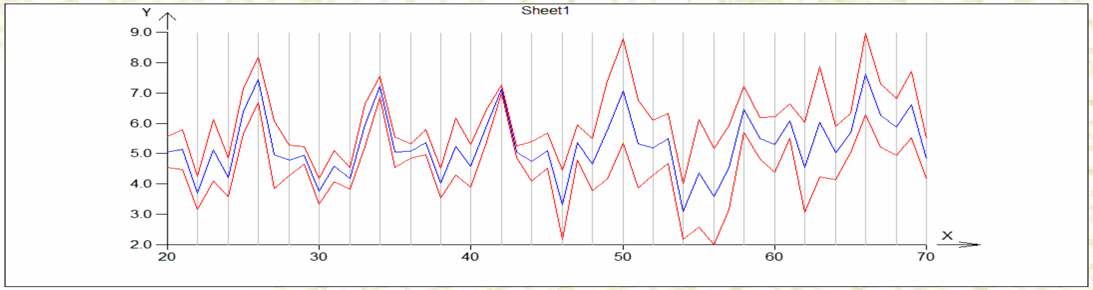


Take advantage of the ANN models drawback use unstability as „bootstrap“

Re-construct confidence intervals of prediction

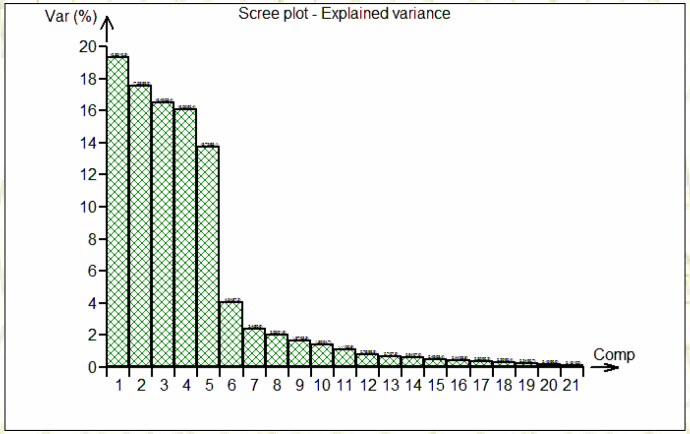


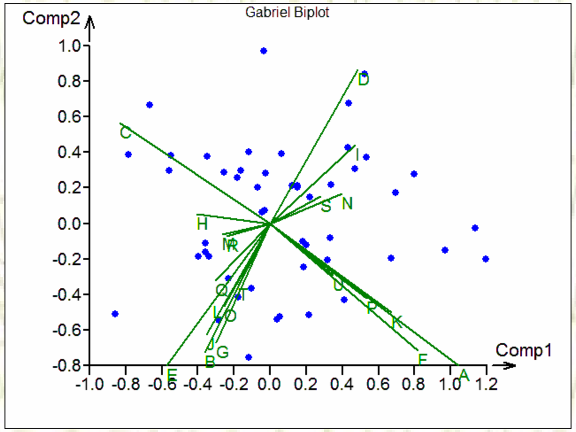
Robustification by choosing non-Eucleidian distance for optimizing the fit



$$R(\theta) = \sum_j \left(y_j - f\left(\theta, \mathbf{x}_j\right)\right)^r ; 1 < r < 2$$

Reducing parameter space dimension





PCA reveals inherent dimensionality of ANN weights $D(p) < p$

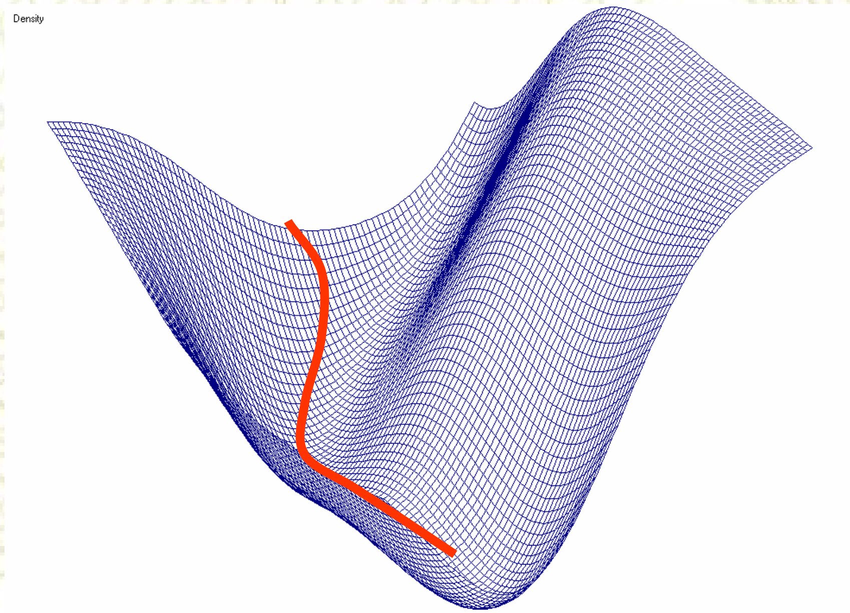
Significant principal components of model parameters appear to be a good predictor

$$\begin{pmatrix} \frac{\partial w_1}{\partial w_1} & \cdots & \frac{\partial w_1}{\partial w_p} \\ \vdots & \frac{\partial w_i}{\partial w_j} & \vdots \\ \frac{\partial w_p}{\partial w_1} & \cdots & \frac{\partial w_p}{\partial w_p} \end{pmatrix}$$

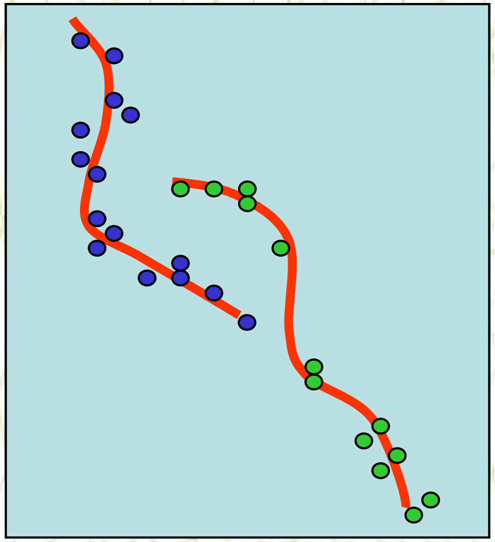
A nonlinear system defining a subspace in the parameter space with lower inherent dimensionality p' < p thus allowing stable prediction in p'.

subject to

$$\frac{\partial R(w)}{w_i} \le \delta; \delta \ge 0$$



Model classification



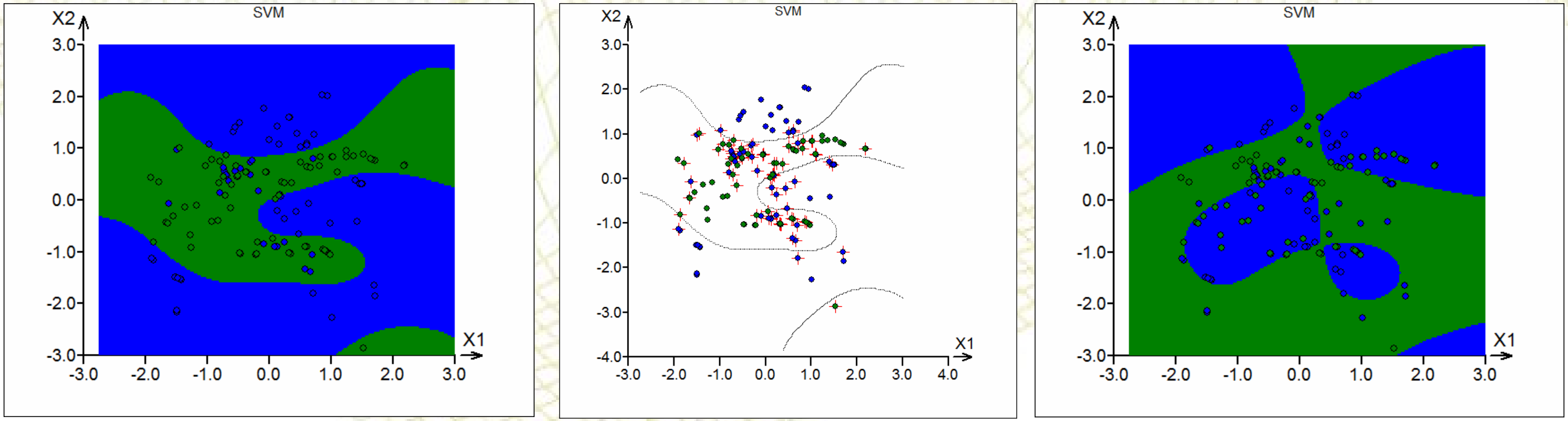
Linearly non-separable data on first p' components
Suitable for SVM separation

$$
\begin{aligned}
\min_{\mathbf{w},b,\boldsymbol{\xi},\rho} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} - \nu\rho + \frac{1}{l}\sum_{i=1}^{l}\xi_i \\
\text{subject to} \quad & y_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq \rho - \xi_i, \\
& \xi_i \geq 0, i = 1, \ldots, l, \rho \geq 0.
\end{aligned}
$$

svm_type	c_svc	
kernel_type	RBF	
Gamma	0.5	
cases	153	
variables	5	
nr_class	2	
total_sv	49	
label	1	2
nr_sv	26	23

Misclass Table	
Correct	151
Misclass	2
Misclass Rate	0.013

RBF Kernel SVM projection to pc_1, pc_2, m.r.=0.17



Conclusion

Applicability to static and dynamic multivariate problems

Significant dimension reduction

Applications in brewery, pharma, steel, medicine

(-) Computationally extensive

(-) No simple model

References

T. Hastie, R. Tibshirani, J. Friedman: The elements of statistical learning: Data mining, inference and prediction, Springer 2001

Chih-Chung Chang and Chih-Jen Lin: LIBSVM: a Library for Support Vector Machines

CHRISTOPHER J.C. BURGES: A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery 2, 121-167, 1998

K. Kupka, M. Meloun: The end-point estimation in instrumental titrations by nonlinear regression

E. Alpaydin, M. I. Jordan: Local linear perceptrons for classification. IEEE Transactions on Neural Networks, 7(3) May 1996

Stefan Ruping: SVM Kernels for Time Series Analysis, CS Dept, University of Dortmund, http://www-ai.cs.uni-dortmund.de/EVENTS/FGML2001/FGML2001-Paper-Rueping.pdf

Sayan Mukherjee, Edgar Osuna, Federico Girosi: Nonlinear Prediction of Chaotic Time Series Using Support Vector Machines, IEEE Workshop on Neural Networks for Signal Processing VII (1997)

W. A. C. Schmidt, J. P. Davis: Pattern Recognition Properties of Various Feature Spaces for Higher Order Neural Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 15 (8) 1993, 795 – 801

G. E. Hinton, R. R. Salakhutdinov: Reducing the Dimensionality of Data with Neural Networks, Science, Vol 313, July 2006, 504-507

Patterson D W, Chan K H, Tan C M.: Time Series Forecasting with neural nets: a comparative study. Proc. International Conference on Neural Network Applictions to Signal Processing. NNASP 1993 Singapore pp 269-274.

K. Muller, A. Smola, G. Ratsch, B. Scholkopf, J. Kohlmorgen, and V. Vapnik: Predicting time series with support vector machines. In Proceedings of the International Conference on Artificial Neural Networks, Springer Lecture Notes in Computer Science. Springer, 1997.

C. L. Giles, S. Lawrence, A. C. Tsoi: Noisy Time Series Prediction using Recurrent Neural Networks and Grammatical Inference, Machine Learning, 44, 161–183, 2001

Zhang Q, Benveniste A (1992) Wavelet Networks. IEEE Trans. Neural Netw. 3: 889-898.

V. Vapnik: Statistical Learning Theory. Wiley, Chichester, GB, 1998.