

Basic Chemometrics with PLS_Toolbox 5.0

©Copyright 1996-2008
Eigenvector Research, Inc.
No part of this material may be
photocopied or reproduced in any form
without prior written consent from
Eigenvector Research, Inc.

Barry M. Wise
Rasmus Bro



Contact Information

Eigenvector Research, Inc.
3905 West Eaglerock Drive
Wenatchee, WA 98801 USA
web: www.eigenvector.com

Barry M. Wise
e-mail: bmw@eigenvector.com
phone: 509-662-9213

Rasmus Bro
e-mail: rb@life.ku.dk
web: www.models.kvl.dk/users/rasmus/



Table of Contents

- Introduction/Motivation
- Linear Algebra for Chemometricians
- MATLAB for Chemometricians
- Principal Components Analysis (PCA)
- Clustering
- Regression and Partial Least Squares (PLS)
- Classification
- Variable Selection
- Common Mistakes in Chemometrics
- PLS_Toolbox Tips and Tricks



Definition of Chemometrics

Chemometrics is the chemical discipline that uses mathematical and statistical methods to

- 1) relate *measurements* made on a *chemical* system to the *state* of the system
- 2) design or select optimal *measurement* procedures and experiments.

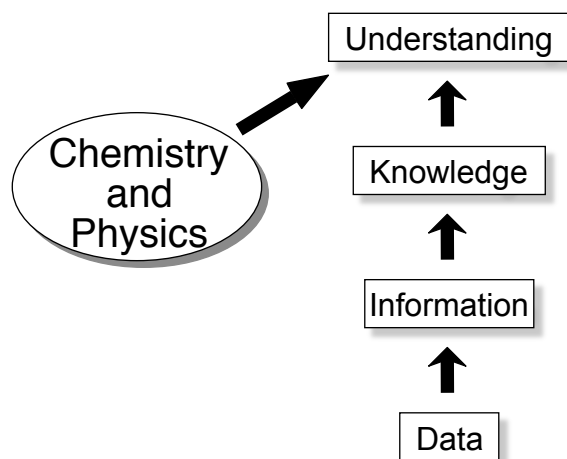


Multivariate Analysis

Multivariate Statistical Analysis is concerned with data that consists of *multiple measurements* on a number of individuals, objects, or data samples. The measurement and analysis of *dependence between variables* is fundamental to multivariate analysis.



Information Hierarchy

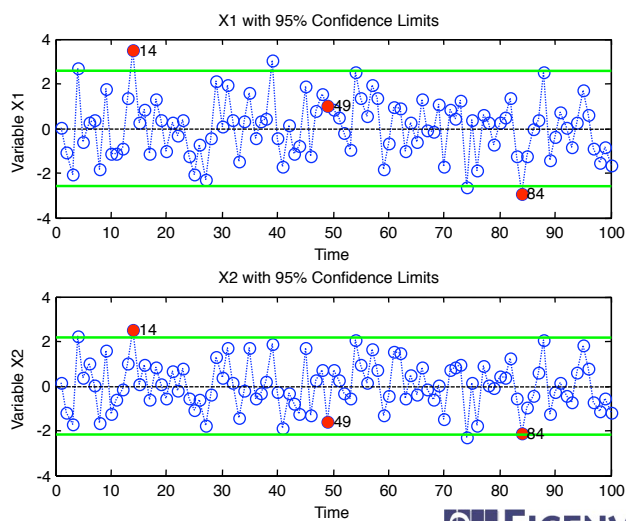


Evolution of Chemometrics

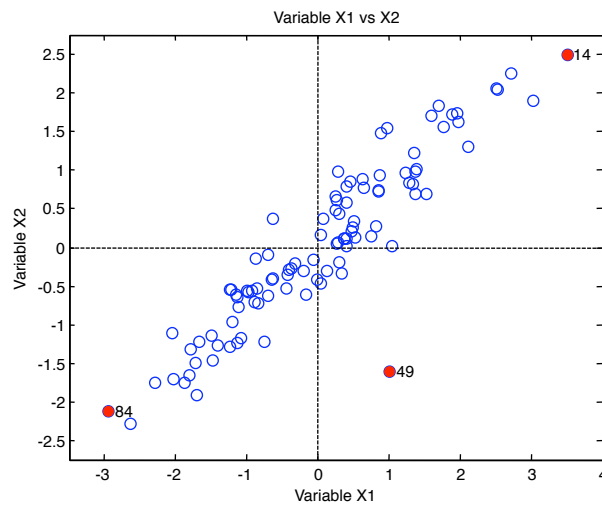
- Methods borrowed from Economics, Psychology and Sociology
 - Lots of data with many variables
 - Data has underlying structure
 - Measurements often very noisy
- Chemometrics has evolved from:
 - Laboratory analytical instruments
 - Process analytical instruments
 - Chemical processes



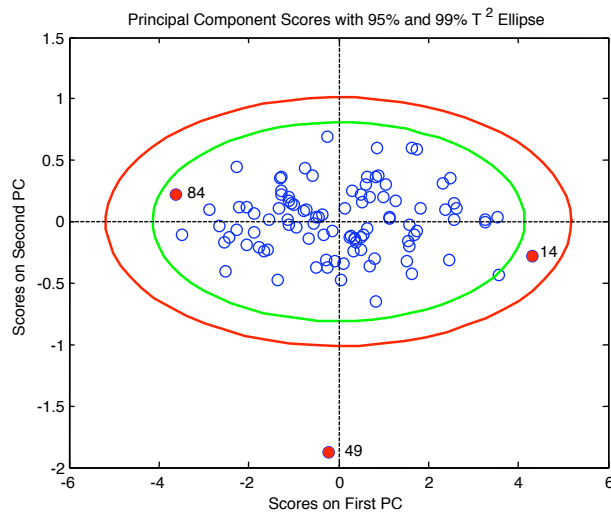
Motivation: Which Point is Out of Control?



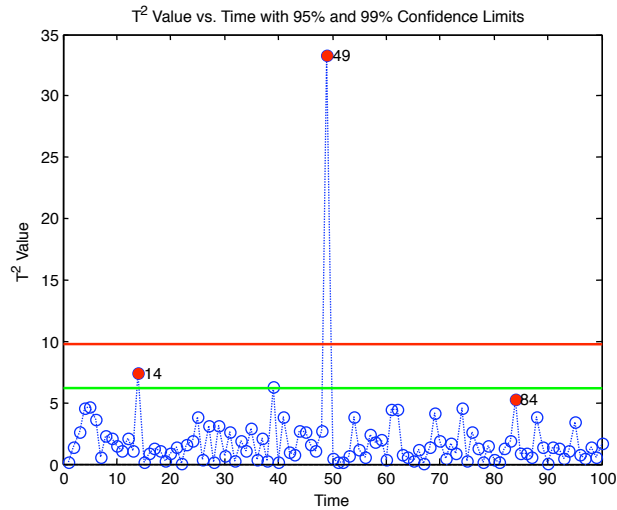
Plot X2 versus X1



Principal Component Scores

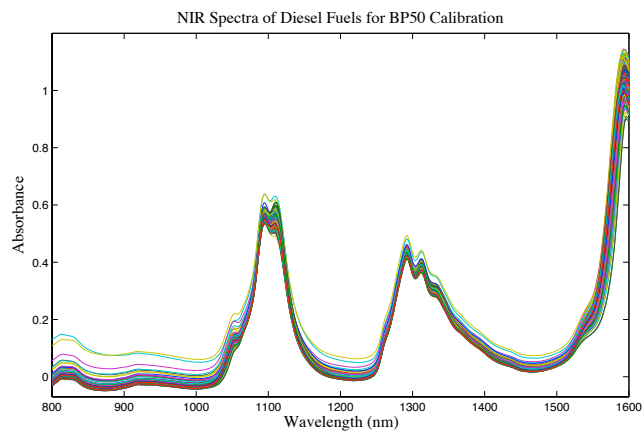


Monitor Single T^2 Chart



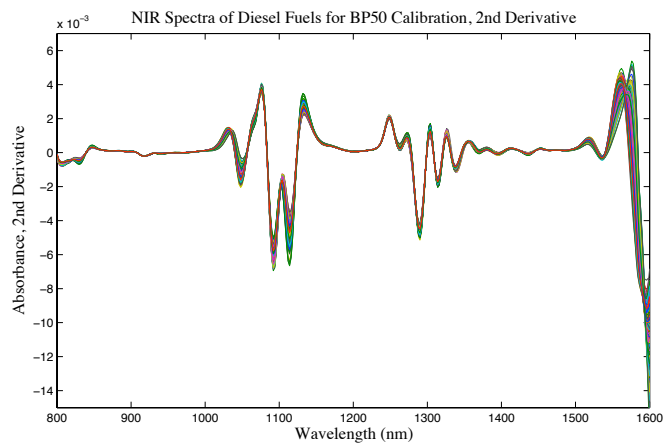
 **EIGENVECTOR**
RESEARCH INCORPORATED

Motivation: Can NIR be used to Estimate BP50 in Diesel Fuel?

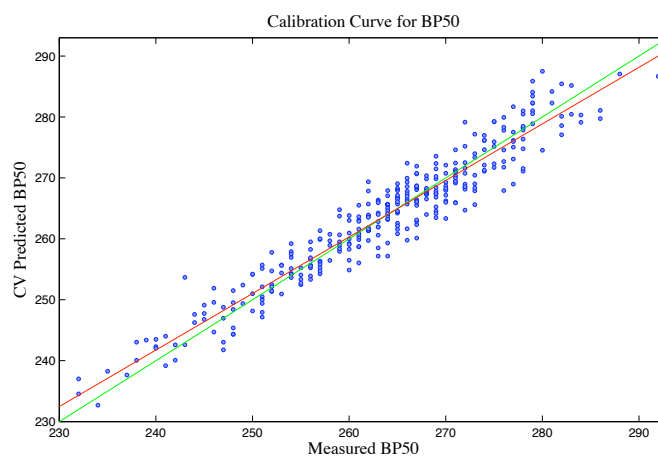


 **EIGENVECTOR**
RESEARCH INCORPORATED

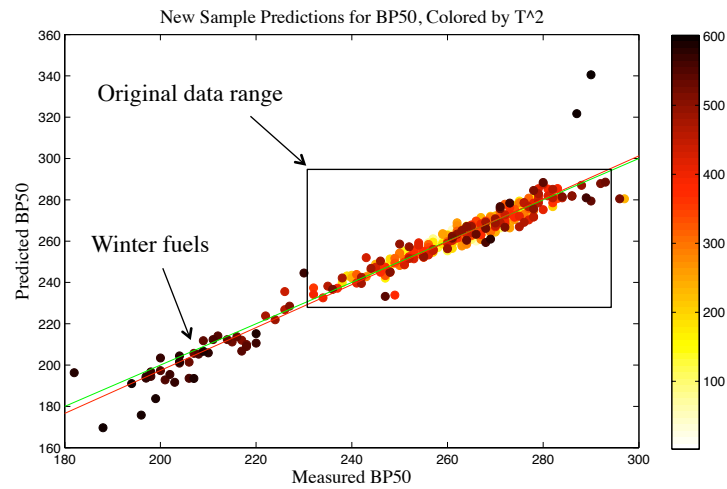
Use Smoothed Second Derivative Spectra



Calibration Curve

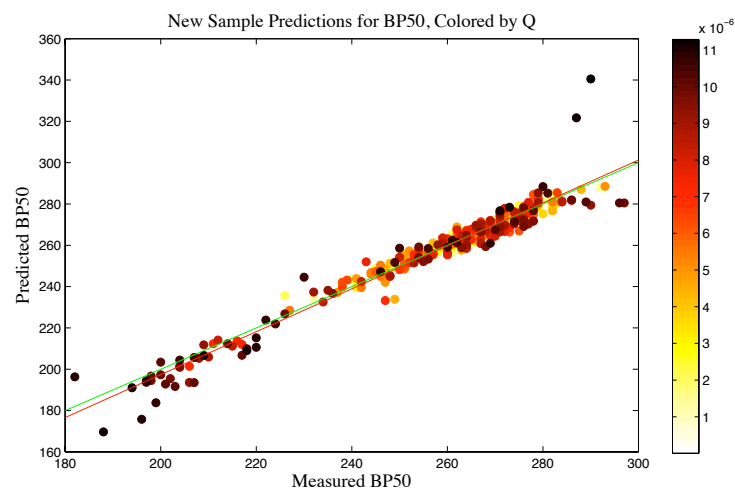


Will Model Work on New Data?



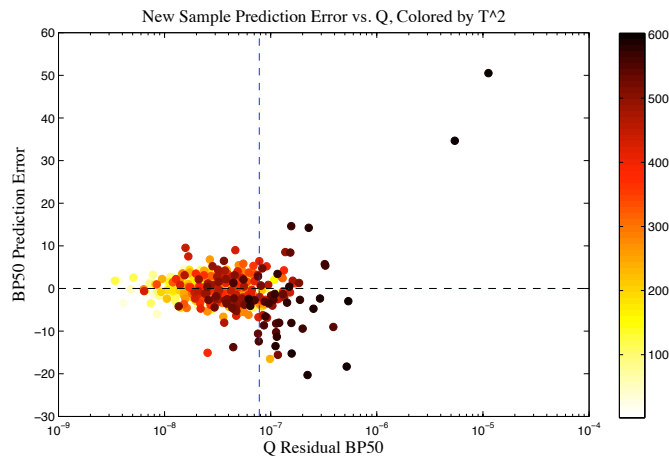
EIGENVECTOR
RESEARCH INCORPORATED

Colored by Residual Q



EIGENVECTOR
RESEARCH INCORPORATED

Higher Q Samples Have Larger Prediction Error



Why Multivariate and Factor Based Methods?

- Noise filtering
- Selectivity enhancement
- Interpretation
- It's a multivariate world!



Resources

- **Books**

- *Chemometrics*, M.A. Sharaf, D.L. Illman and B.R. Kowalski, Wiley-Interscience (1986) ISBN 0-471-83106-9
- *Multivariate Analysis*, K.V. Mardia, J.I. Kent and J.M. Bibby, Academic Press, (1979) ISBN 0-12-471252-2
- *Multivariate Calibration*, H. Martens and T. Næs, John Wiley & Sons Ltd. (1989) ISBN 0-471-90979-3
- *Chemometrics: a textbook*, D.L. Massart et al., Elsevier (1988) ISBN 0-444-42660-4
- *Chemometrics: A Practical Guide*, K.R. Beebe, R.J. Pell, M.B. Seasholtz, Wiley (1998) ISBN 0-471-12451-6
- *Multivariate Data Analysis In Practice*, Kim H. Esbensen, CAMO ASA (2000), ISBN 82-993330-2-4
- *A user-friendly guide to Multivariate Calibration and Classification*, T. Næs, T. Isaksson, T. Fearn, T. Davies, NIR Publications(2002), ISBN 0-9528666-2-5
- *Multivariate Image Analysis*, Paul Geladi and Hans Grahn, Wiley (1996), ISBN 0-471-93001-6
- *Multivariate Analysis of Quality: An Introduction*, H. Martens and M. Martens, Wiley (2001), ISBN 0-471-97428-5

- **Journals**

- Journal of Chemometrics
- Chemometrics and Intelligent Laboratory Systems
- Analytical Chemistry
- Analytica Chimica Acta
- Applied Spectroscopy
- Critical Reviews in Analytical Chemistry
- Journal of Process Control
- Computers in Chemical Engineering
- Technometrics
-

