
Chemometrics: A Practical Guide

**KENNETH R. BEEBE
RANDY J. PELL
MARY BETH SEASHOLTZ**
The Dow Chemical Company



A Wiley-Interscience Publication

JOHN WILEY & SONS, INC.

New York · Chichester · Weinheim · Brisbane · Singapore · Toronto

Contents

Preface	xi
1 The Six Habits of an Effective Chemometrician	1
1.1 Using this Book / 2	
1.2 The Six Habits of an Effective Chemometrician / 3	
1.3 Definitions / 5	
1.4 Summary of the Six Habits of an Effective Chemometrician / 8	
Reference / 8	
2 Defining the Problem	9
2.1 Decision Tree / 9	
2.2 Communication, Communication, Communication / 11	
2.2.1 Understanding the System / 11	
2.2.2 Understanding the Problem / 12	
2.2.3 Understanding the Chemometrics View / 16	
2.3 Experimental Planning / 16	
2.3.1 Benefits of Experimental Planning and Design / 17	
2.3.2 Steps for Designing the Experiment / 18	
2.4 Summary of Defining the Problem / 23	
References / 23	
Appendix A. Special Considerations for Inverse Chemometric Modeling / 23	
3 Preprocessing	26
3.1 Preprocessing the Samples / 26	
3.1.1 Normalization / 27	
3.1.2 Sample Weighting / 29	
3.1.3 Smoothing / 30	
3.1.3.1 Mean Smoother / 30	
3.1.3.2 Running Mean Smoother / 30	

3.1.3.3	<i>Running Median Smoother</i>	/ 31
3.1.3.4	<i>Running Polynomial Smoother</i>	/ 32
3.1.3.5	<i>Fourier Filter Smoother</i>	/ 35
3.1.4	Baseline Corrections	/ 36
3.1.4.1	<i>Explicit Modeling Approach</i>	/ 36
3.1.4.2	<i>Derivatives</i>	/ 41
3.1.4.3	<i>Multiplicative Scatter Correction</i>	/ 46
3.2	Preprocessing the Variables	/ 47
3.2.1	Mean Centering	/ 48
3.2.2	Variable Weighting	/ 50
3.2.2.1	<i>A Priori Information Weighting</i>	/ 50
3.2.2.2	<i>Variable Selection</i>	/ 50
3.2.2.3	<i>Variance Scaling</i>	/ 51
3.2.2.4	<i>Autoscaling</i>	/ 52
3.3	Summary of Preprocessing	/ 53
	References	/ 55

4 Pattern Recognition

56

4.1	Decision Tree	/ 62
4.2	Unsupervised Learning	/ 64
4.2.1	Hierarchical Cluster Analysis (HCA)	/ 65
4.2.1.1	<i>HCA Example 1</i>	/ 71
4.2.1.2	<i>HCA Example 2</i>	/ 76
4.2.2	Principal Components Analysis (PCA)	/ 81
4.2.2.1	<i>PCA Example 1</i>	/ 87
4.2.2.2	<i>PCA Example 2</i>	/ 99
4.2.3	Summary of Unsupervised Pattern Recognition	/ 111
4.3	Supervised Learning	/ 112
4.3.1	K-Nearest Neighbor (KNN)	/ 113
4.3.1.1	<i>KNN Example 1</i>	/ 116
4.3.1.2	<i>KNN Example 2</i>	/ 125
4.3.2	SIMCA (Soft Independent Modeling of Class Analogies)	/ 134
4.3.2.1	<i>SIMCA Example 1</i>	/ 138
4.3.2.2	<i>SIMCA Example 2</i>	/ 164
4.3.3	Summary of Supervised Pattern Recognition Methods	/ 180
4.4	Summary of Pattern Recognition	/ 181
	References	/ 182

5	Multivariate Calibration and Prediction	183
5.1	Decision Tree / 186	
5.2	Classical Methods / 188	
5.2.1	Direct CLS / 192	
5.2.1.1	<i>DCLS Example 1 / 192</i>	
5.2.1.2	<i>DCLS Example 2 / 210</i>	
5.2.2	Indirect CLS / 218	
5.2.2.1	<i>ICLS Example 1 / 218</i>	
5.2.2.2	<i>ICLS Example 2 / 227</i>	
5.2.3	Summary of Classical Methods / 243	
5.3	Inverse Methods / 245	
5.3.1	Multiple Linear Regression (MLR) / 250	
5.3.1.1	<i>MLR Example 1 / 251</i>	
5.3.1.2	<i>MLR Example 2 / 269</i>	
5.3.2	Partial Least Squares/Principal Components Regression / 278	
5.3.2.1	<i>PLS Example 1 / 284</i>	
5.3.2.2	<i>PLS Example 2 / 315</i>	
5.3.2.3	<i>PLS Example 3 / 327</i>	
5.3.3	Summary of Inverse Methods / 335	
5.4	Summary of Multivariate Calibration and Prediction / 337	
	References / 338	
	Index / 341	