

## **INFORMATION TO USERS**

**The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.**

**The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.**

**In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.**

**Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.**

**Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.**



University Microfilms International  
A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA  
313/761-4700 800/521-0600



**Order Number 9109818**

**Chemometrics and infrared emission spectroscopy for remote  
analysis**

**Pell, Randall James, Ph.D.**

**University of Washington, 1990**

**Copyright ©1990 by Pell, Randall James. All rights reserved.**

**U·M·I**  
300 N. Zeeb Rd.  
Ann Arbor, MI 48106





**Chemometrics and Infrared Emission Spectroscopy  
for Remote Analysis**

by

RANDALL JAMES PELL

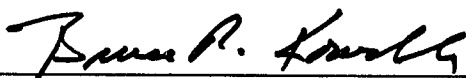
A dissertation submitted in partial fulfillment  
of the requirements for the degree of

Doctor of Philosophy

University of Washington

1990

Approved by

  
(Chairperson of Supervisory Committee)

Program Authorized  
to Offer Degree

chemistry

Date 11-2-90

© Copyright 1990

Randall James Pell

## Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Randy J. Bell

Date 11-2-90

University of Washington

Abstract

**Chemometrics and Infrared Emission Spectroscopy  
for Remote Analysis**

by Randall James Pell

Chairperson of Supervisory Committee:

Professor Bruce R. Kowalski  
Department of Chemistry

Chemometrics is a discipline of chemistry that uses mathematical and statistical techniques to extract useful information from chemical data. Infrared emission spectroscopy is one of a very few analytical techniques that allows noninvasive and remote measurements. This dissertation uses the chemometric methods of multivariate calibration and multiresponse nonlinear optimization to extract useful information from infrared emission data. Partial Least Squares and Locally Weighted Regression are used for prediction of chemical and physical properties of a free standing polymer system and a polymer coated aluminum system. Multiresponse nonlinear optimization is used to model the dynamic process of a polymer curing reaction. The choice of the proper objective function for multiresponse nonlinear optimization is also investigated.

## TABLE OF CONTENTS

List of Figures .....	vii
List of Tables.....	xi
Introduction .....	1
Chapter 1 Infrared Emission Spectroscopy .....	4
1.1 Introduction to Chapter 1 .....	4
1.2 Theory .....	4
1.2.1 Radiative Transfer Theory .....	8
1.3 Experimental Considerations .....	11
1.4 Survey of Infrared Emission Spectroscopy Applications .....	17
1.4.1 Gases .....	18
1.4.1.1 Laboratory Measurements.....	18
1.4.1.2 Remote Sensing and Atmospheric Applications.....	19
1.4.2 Condensed Phases .....	20
1.4.2.1 General Applications for Condensed Phases .....	20
1.4.2.2 Molten Salts.....	25
1.4.2.3 Glass.....	27
1.4.2.4 Surface Science .....	28
1.4.2.5 Reactions .....	30
1.4.3 Stimulated Infrared Emission.....	31
1.5 Conclusions .....	32
1.6 Notes to Chapter 1.....	33
Chapter 2 Multivariate Analysis .....	38
2.1 Introduction to Chapter 2 .....	38
2.2 Multivariate Calibration .....	40

2.2.1 Overview .....	40
2.2.2 Notation and Preprocessing Issues .....	41
2.2.3 Classical Least Squares (CLS) .....	42
2.2.4 Inverse Least Squares (ILS) .....	44
2.2.5 Biased Regression Methods .....	47
2.2.5.1 Principal Component Regression (PCR) .....	47
2.2.5.2 Partial Least Squares (PLS) .....	53
2.2.6 Rank Determination Methods .....	59
2.2.7 Locally Weighted Regression (LWR) .....	63
2.3 Multiresponse Nonlinear Optimization .....	67
2.3.1 Overview .....	67
2.3.2 Theory for Multiresponse Nonlinear Optimization .....	68
2.3.2.1 Multivariate Nonlinear Modeling Using Direct Observations .....	69
2.3.2.2 Multivariate Nonlinear Modeling Using Indirect Observations .....	70
2.3.3 Application of Multiresponse Nonlinear Optimization .....	76
2.3.3.1 Applications Involving Direct Observation Experiments .....	76
2.3.3.2 Applications Involving Indirect Observation Experiments .....	78
2.4 Conclusions .....	80
2.5 Notes to Chapter 2 .....	82
Chapter 3 Calibration Studies Using Infrared Emission for Remote Analysis .....	86
3.1 Introduction to Chapter 3 .....	86

3.2 Elvax® Polymer Calibration Study.....	88
3.2.1 Experimental and Data Analysis.....	88
3.2.2 Partial Least Squares Regression Results.....	90
3.2.2.1 Prediction of Vinyl Acetate Concentration .....	92
3.2.2.2 Prediction of Thickness.....	96
3.2.2.3 Prediction of Temperature.....	98
3.2.4 Locally Weighted Regression Results.....	101
3.3 Loading Weight Prediction of Thin Polymer Films on Metal Surfaces.....	105
3.3.1 Experimental and Data Analysis.....	105
3.3.2 Partial Least Squares Regression Results.....	111
3.3.2.1 Analysis of Sample Set #1 .....	111
3.3.2.2 Analysis of Sample Sets #2, #3, and #4 .....	123
3.3.3 Locally Weighted Regression Results.....	127
3.3.3.1 Analysis of Sample Set #1 .....	127
3.3.3.2 Analysis of Sample Sets #2, #3, and #4 .....	129
3.4 Conclusions .....	132
3.5 Notes to Chapter 3.....	134
Chapter 4 Polymer Cure Reaction Monitoring .....	135
4.1 Introduction to Chapter 4 .....	135
4.2 Experimental and Data Analysis.....	135
4.2.1 Rank Determination .....	137
4.2.2 Analysis of Closed Data.....	139
4.3 Results and Discussion.....	140
4.4 Conclusions .....	164

4.5 Notes to Chapter 4.....	166
Chapter 5 Multiresponse Fitting Criteria Tests.....	167
5.1 Introduction to Chapter 5 .....	167
5.2 Experimental and Data Analysis.....	168
5.3 Results and Discussion.....	175
5.3.1 Direct Observation Experiments .....	175
5.3.1.1 Experiment #1 .....	176
5.3.1.2 Experiment #2 .....	177
5.3.1.3 Experiment #3 .....	179
5.3.1.4 Experiment #4 .....	180
5.3.1.5 Experiment #5 .....	181
5.3.2 Indirect Observation Experiments.....	183
5.3.2.1 First-order Consecutive Reaction Results.....	183
5.3.2.1.1 Low overlap results .....	184
5.3.2.1.2 High overlap results .....	186
5.3.2.2 pH Titration Experiments.....	188
5.3.2.2.1 Low overlap results .....	189
5.3.2.2.2 High overlap results .....	191
5.4 Conclusions .....	193
5.5 Notes to Chapter 5.....	195
Bibliography.....	196
Appendix A. NMR Determination of Vinyl Acetate Concentration in Selected Elvax® samples.....	206
Appendix B Observations on the Effects of Closure, Linear Baselines and Mean Centering on Exploratory Multivariate Data .....	209



B.1 Introduction to Appendix B .....	209
B.2 Discussion.....	210
B.2.1 Distinction of Classes of Closure .....	212
B.2.2 Equivalence of Propagated Closure and a Baseline .....	214
B.2.3 Methodology for Rank Interpretation.....	217
B.3 Conclusions .....	228
B.4 Notes to Appendix B .....	229

## LIST OF FIGURES

Figure 1.1 Graphical representation of Planck's law with wavenumber axis .....	6
Figure 1.2 Physical model for radiative transfer theory discussion .....	8
Figure 2.1 Ideal PRESS plot .....	61
Figure 2.2 PRESS vs. number of latent variables added.....	62
Figure 2.3 Tricubic weight function used for locally weighted regression.....	66
Figure 2.4 Multiresponse nonlinear analysis.....	75
Figure 3.1 a) Single beam emission spectrum of 18 % vinyl acetate Elvax®, 0.102 mm at 120 °C b) Transmission measurement of the same material .....	90
Figure 3.2 Infrared emission spectra as a function of vinyl acetate concentration at a constant thickness of 0.102 mm and temperature of 120 °C.....	93
Figure 3.3 Infrared emission spectra as a function of thickness at a constant vinyl acetate concentration of 15 % and temperature of 120 °C .....	96
Figure 3.4 Infrared emission spectra as a function of temperature at a constant vinyl acetate concentration of 15 % and thickness of 0.102 mm.....	99
Figure 3.5 Graphical representation of Table 3.4.....	102
Figure 3.6 Graphical representation of Table 3.5.....	104
Figure 3.7 Relative standard deviation of 90 °C blackbody data. Inset is region of spectrum used for calibration .....	108
Figure 3.8 Fluctuation in the infrared emission spectra for a 5.78 MSI sample measured over a three day period .....	109

Figure 3.9 Fluctuation in the infrared emission spectra for 5.78 MSI sample, three spectra measured on a single day .....	110
Figure 3.10 Sample Set #1 emission spectra as a function of loading weight at a constant temperature of 80 °C and a zero degree angle of observation .....	112
Figure 3.11 Sample Set #1 emission spectra as a function of angle of observation at constant loading weight of 5.11 MSI and 80 °C.....	113
Figure 3.12 Sample Set #1 emission spectra as a function of temperature at a constant loading weight of 5.11 MSI and zero degrees from the norm observation angle .....	114
Figure 3.13 Scores from one factor model vs. the reference loading weights for Sample Set #1 .....	116
Figure 3.14 Spectra from low pigment concentration samples as a function of loading weight.....	124
Figure 3.15 Spectra from nominal pigment concentration samples as a function of loading weight .....	125
Figure 3.16 Spectra from high pigment concentration samples as a function of loading weight.....	125
Figure 4.1 Infrared absorption spectra from 4000 to 600 cm <sup>-1</sup> monitoring the reaction of a two component paint mixture at 100 °C .....	141
Figure 4.2 Infrared emittance spectra from 4000 to 600 cm <sup>-1</sup> monitoring the reaction of a two component paint mixture at 100 °C .....	141
Figure 4.3 Emittance versus absorbance computed from equation 4.2 for a single wavelength.....	144

Figure 4.4 a) Raw emittance versus absorbance for 100 °C curing data and b) linearized emittance data versus absorbance for 100 °C data.....	146
Figure 4.5 Time profiles for isocyanate emittance response at 100, 134 and 154 °C .....	147
Figure 4.6 Standard deviation spectrum for 100 °C absorbance data .....	148
Figure 4.7 Standard deviation spectrum for 100 °C emittance data .....	149
Figure 4.8 Time varying singular vectors one to four for 100 °C emittance data.....	153
Figure 4.9 Spectral singular vectors corresponding to time singular vectors displayed in Figure 4.8.....	153
Figure 4.10 Fit of the model to the first three singular vectors for the absorbance data .....	156
Figure 4.11 Fit of the absorption data in concentration and time space.....	157
Figure 4.12 Reconstructed "pure component" spectra for absorbance data in the 3684 to 3216 cm <sup>-1</sup> region, the alcohol and amine region.....	158
Figure 4.13 Reconstructed "pure component" spectra for absorbance data in the 2404 to 2164 cm <sup>-1</sup> region, the isocyanate region .....	159
Figure 4.14 Reconstructed "pure component" spectra for absorbance data in the 1840 to 1604 cm <sup>-1</sup> region, the urethane carbonyl region .....	159
Figure 4.15 Reconstructed "pure component" spectra for absorbance data in the 1604 to 1488 cm <sup>-1</sup> region, the Amide II region.....	160
Figure 4.16 Reconstructed "pure component" spectra for absorbance data in the 1284 to 1044 cm <sup>-1</sup> region, the Amide III and C-O-C region.....	160
Figure 4.17 Fit of model to raw emittance data at 100 °C .....	162

Figure 4.18 Change in third singular vector for emittance data as a function of R1 linearization parameter .....	163
Figure 5.1 Empirical response noise correlation matrix for infrared spectra, 2164 to 2404 $\text{cm}^{-1}$ , collected over time on a system with no systematic temporal changes.....	170
Figure 5.2 Pure concentration profiles for first-order consecutive reaction scheme.....	172
Figure 5.3 Concentration profiles for pure components obeying the Henderson-Hasselbalch equation.....	173
Figure 5.4 Pure spectral profiles used in the indirect observation experiment simulations. Peak maximum at 350, 450 and 550 nm, half width at half height is 50 nm, the low overlap experiment.....	174
Figure B.1 Graphical representation of three step rank determination procedure.....	218
Figure B.2 Possible matrix structures corresponding to Cases 1 and 2 .....	220
Figure B.3 Possible matrix structures corresponding to Case 3.....	221
Figure B.4 Possible matrix structures corresponding to Case 4.....	222
Figure B.5 Possible matrix structures corresponding to Case 5.....	223
Figure B.6 Possible matrix structures corresponding to Case 6.....	224

## LIST OF TABLES

Table 3.1	Standard error of prediction (SEP) results for vinyl acetate concentration (absolute %) in Elvax® samples using PLS.....	94
Table 3.2	Standard error of prediction (SEP) results for thickness (mm) in Elvax® samples using PLS.....	97
Table 3.3	Standard error of prediction (SEP) results for temperature (°C) in Elvax® samples using PLS.....	100
Table 3.4	Standard error of prediction results for vinyl acetate concentration, absolute % vinyl acetate, of Elvax® samples as a function of the model choice and fraction of the data ( $f$ ) used with locally weighted regression.....	102
Table 3.5	Standard error of prediction results for thickness (mm) of Elvax® samples as a function of the model choice and fraction of the data ( $f$ ) used with locally weighted regression.....	103
Table 3.6	Loading weight reference data for Sample Sets #1, #2, #3, and #4 in milligrams per square inch (MSI) .....	106
Table 3.7	Analysis of variance results for reference calibration loading weight data .....	111
Table 3.8	Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of temperature and the model complexity using PLS .....	115
Table 3.9	Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of temperature and the model complexity using PLS. (Outliers deleted) .....	117

Table 3.10	Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of angle of observation and the model complexity using PLS. (All samples used).....	118
Table 3.11	Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of angle of observation and the model complexity using PLS. (Outliers deleted).....	119
Table 3.12	Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of model complexity using PLS and data from all temperatures with and without outliers .....	120
Table 3.13	Standard error of prediction results for temperature ( $^{\circ}\text{C}$ ), for Sample Set #1 as a function of model complexity using PLS and data from all temperatures with and without outliers .....	121
Table 3.14	Standard error of estimate for simulated temperature variance experiments and associated standard deviation for 20 experiments .....	122
Table 3.15	Standard error of prediction results for loading weights, MSI, for Sample Sets #2, #3, and #4 as a function of model complexity using PLS .....	126
Table 3.16	Standard error of prediction results for loading weights, MSI, for Sample Set #1 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used. (Outliers not removed) .....	128
Table 3.17	Standard error of prediction results for loading weights, MSI, for Sample Set #1 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used. (Outliers removed) .....	129

Table 3.18 Standard error of prediction results for loading weights, MSI, for Sample Set #2 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used .....	130
Table 3.19 Standard error of prediction results for loading weights, MSI, for Sample Set #3 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used .....	130
Table 3.20. Standard error of prediction results for loading weights, MSI, for Sample Set #4 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used .....	131
Table 4.1 Band assignments for spectra displayed in Figures 4.1 and 4.2.....	142
Table 4.2 Experimental noise estimation and residual values used for rank determination of 100 °C absorbance data .....	150
Table 4.3 Experimental noise estimation and residual values used for rank determination of 100 °C emittance data.....	151
Table 4.4 Nonlinear model parameter estimates for absorption, emittance and linearized emittance data at 100 °C.....	155
Table 4.5 Nonlinear model parameter estimates for raw emittance data at three temperatures .....	164
Table 5.1 SEP for direct measure experiment, consecutive first-order reaction kinetics, $k_1=10$ , $k_2=5$ , diagonals of variance-covariance matrix are all ones.....	176
Table 5.2 SEP for direct measure experiment, consecutive first-order reaction kinetics experiment, $k_1=10$ , $k_2=5$ , diagonals of variance-covariance matrix are 3, 2, 1.....	178



Table 5.3 Reduced sample size. SEP for direct measure experiment, consecutive first-order reaction kinetics experiment, $k_1=10$ , $k_2=5$ , observed for first 35 % of time, diagonals of variance-covariance matrix are all ones .....	179
Table 5.4 Increased sample size. SEP for direct measure experiment, consecutive first-order reaction kinetics experiment, $k_1=10$ , $k_2=5$ , diagonals of variance-covariance matrix are all ones, 201 data points .....	181
Table 5.5 Poorer time resolution. SEP for direct measure experiment, consecutive first-order reaction kinetics experiment, $k_1=8$ , $k_2=6$ , diagonals of variance-covariance matrix are all ones .....	182
Table 5.6 SEP for indirect, consecutive first-order reaction kinetics experiment, low overlap, results for $k_1$ .....	185
Table 5.7 SEP for indirect, consecutive first-order reaction kinetics experiment, low overlap, results for $k_2$ .....	185
Table 5.8 SEP for indirect, consecutive first-order reaction kinetics experiment, high overlap, results for $k_1$ .....	187
Table 5.9 SEP for indirect, consecutive first-order reaction kinetics experiment, high overlap, results for $k_2$ .....	188
Table 5.10 SEP for indirect pK determination, low overlap, results for $k_1$ .....	190
Table 5.11 SEP for indirect pK determination, low overlap, results for $k_2$ .....	190
Table 5.12 SEP for indirect pK determination, low overlap, results for $k_3$ .....	191
Table 5.13 SEP for indirect pK determination, high overlap, results for $k_1$ .....	192
Table 5.14 SEP for indirect pK determination, high overlap, results for $k_2$ .....	192
Table 5.15 SEP for indirect pK determination, high overlap, results for $k_3$ .....	193

<b>Table A.1</b>	<b>Peak assignments for NMR analysis of Elvax® samples.....</b>	<b>207</b>
<b>Table A.2</b>	<b>Expected and measured vinyl acetate concentrations .....</b>	<b>208</b>

## Acknowledgements

I wish to acknowledge the following organizations for their financial support during my graduate career: the Center for Process Analytical Chemistry, Department of Chemistry at the University of Washington, American Chemical Society Analytical Division Full Year Fellowship sponsored by the Perkin-Elmer Corporation, the Dow Chemical Company Outstanding graduate student fellowship and the Symposium for Innovation in Measurement Science student paper award sponsored by the Instrument Society of America.

There are many people who have contributed to this work and I wish to sincerely thank them for their assistance. For his never ending enthusiasm and encouragement I would like to thank Professor Bruce Kowalski. Professor Jim Callis provided many helpful discussions that have made this a better work.

To the members and visitors to the Laboratory for Chemometrics I owe a great deal for their support and help during this work. Those persons include Ken Beebe, Pat Carey, Scott Ramos, Gene Sanchez, Bruce Wilson, Sandy Moore, Avi Lorber, Dave Haaland, Odd Borgen, Tormod Næs, Sharon Neal, Pentti Minkkinen, Jim Burger, Yongdong Wang and especially Mary Beth Seasholtz with whom I have had the pleasure of working and learning much about chemometrics. A special thanks to Brice Erickson for his help and encouragement in the early stages of this project. The CPAC computing staff, Dennis Gentry, Dave Veltkamp, Rob Etzel, Ken Case, and Pete Pulliam, have been very patience with my seemingly endless stream of mail messages to "help".

Several industrial sponsors have been very helpful, including Dr. Robert Hannah of the Perkin-Elmer Corporation who served on the advising committee and provided the use of his laboratories for the Elvax<sup>®</sup> study. Dr. Gil Sloan of the Du Pont Company provided the Elvax<sup>®</sup> samples and Drs. Jack Hartshorn and Don Carbaugh of the Du Pont Marshall Laboratories were very helpful in the polymer curing project. Ruth Roberts of ALCOA provided the polymer coated aluminum samples.

I thank the Chemistry Department Electronics and Machine Shop for their expert work in providing the needed instrumentation for the experiments conducted for this work.

I am grateful to my parents Glen and Darlean Pell for their support and encouragement throughout these studies. I am most grateful to my wife Karen, without whose encouragement, love and support this project would not have been completed. Her tireless reading and rereading of the dissertation has made it a better work.

*To my wife Karen  
and my Mother and Father*

## **Introduction**

The goal of process analytical chemistry is to supply quantitative and qualitative information about a chemical process (1). Five eras of process analytical chemistry have been identified: off-line, at-line, on-line, in-line and noninvasive. The first two eras, off-line and at-line, are distinguished by the requirement of manual removal of the sample from the process and transport to the measuring device. Off-line analysis is usually performed at a centralized facility with the advantages of sophisticated instrumentation and expert technical staff. The disadvantages are the delay time between sample submission and reporting of analysis results and the possible degradation or loss of sample integrity over time. In order to reduce this delay time, at-line analyzers are moved closer to the process. These analyzers must be simpler, easier to use, robust and rugged.

The at-line analyzer still requires manual sample removal and this is where on-line analysis is distinctly different. The on-line analyzer samples, conditions, and presents the sample to the instrument. An ubiquitous example of this is the process gas chromatograph in an oil refining operation. The disadvantage of the on-line analyzer is found in the problems associated with sampling and this has lead to in-line analysis.

In-line analyzers perform the analysis in situ with a chemically or physically sensitive probe inserted in the process stream. There are many types of transducers available for in-line sensors and microlithography and micromachining are believed to be the key to scaling these sensors down. The in-line approach has reduced the sampling problem, but the process stream and sensor are in direct contact which may be a problem for both the stream and the sensor.

This has led to the fifth era of process analytical chemistry, noninvasive analysis. This is the most desirable of the five eras because the probe is never in physical contact with the process stream. There are few techniques available for such an analysis and it is here that infrared emission spectroscopy can contribute to process analytical chemistry.

Of the various vibrational spectroscopic techniques available to the analytical chemist comparatively little use has been made of emission detection (2,3). In the first place, most analytical measurements are performed in the laboratory where there exists a variety of easily implemented alternatives to the emission technique, such as transmission and diffuse reflectance spectroscopy. These other, more conventional techniques, have better signal-to-noise ratios and do not have anomalous behavior arising from temperature gradients, which may be present in emission analysis due to sample heating (4). In the second place, the complex nonlinear behavior of the emission signal (5) and, until recently, the lack of available mathematical techniques, and computer hardware and software, have made the extraction of relevant information from these data difficult.

Recent interest in process analytical chemistry has provided a driving force for moving analytical measurements from the laboratory to the process line and has accelerated the search for techniques compatible with process situations (1). Advances in data reduction techniques from chemometrics now allow information to be derived from data previously considered to be of little value due to signal complexity (6, 7). In light of these developments, infrared emission deserves reconsideration as a process monitoring tool. Its simple experimental characteristics (the sample is the source) and noninvasive character are particularly desirable for certain process measurement situations. In combination with chemometric data

reduction strategies, infrared emission shows promise as a tool for process analytical chemistry (8).

Chapter one will review infrared emission theory and applications and make clear the need for multivariate analysis to handle the complex emission signals. Chapter two develops the theory of multivariate calibration and multiresponse nonlinear optimization techniques used to extract information from infrared emission signals. Chapter three will present two applications of infrared emission for the analysis of remote samples using multivariate calibration techniques. Chapter four presents the analysis of multiresponse data collected using infrared emission spectroscopy from a dynamic process. This study involves the application of multiresponse nonlinear optimization for the extraction of information from the dynamic process of a curing polymer on a metal surface. Chapter five will explore the details of the fitting criterion used in multiresponse nonlinear optimization.



# Chapter 1 Infrared Emission Spectroscopy

## 1.1 Introduction to Chapter 1

This chapter focuses on infrared emission spectroscopy. Section 1.2 is a review of the basic theory needed to understand infrared emission spectroscopy. Experimental considerations are discussed in Section 1.3. The final section surveys pertinent and interesting applications of infrared emission spectroscopy, especially as it relates to analytical chemistry.

## 1.2 Theory

The theory of infrared emission spectroscopy begins with the fact that all matter at a temperature above absolute zero is constantly absorbing and emitting radiation due to vibrations induced by thermal energies (9). Further consideration leads to the idealization of the blackbody as an object that absorbs all radiation incident upon it. For such a body to remain at thermal equilibrium it must also be a perfect emitter, thus

$$A = E \quad (1.1)$$

where  $A$  is the absorptance (the fraction of the incident radiation absorbed) and  $E$  is the emittance (the fraction of the radiation emitted compared to a blackbody) (10). This is Kirchoff's law and thus for a blackbody,  $A = E = 1$ .

The description of the frequency and temperature dependence of the radiation emitted by a blackbody is given by Planck's law (11). The mathematical formulation of Planck's law is

$$I_b(\lambda) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left\{\frac{hc}{\lambda kT}\right\} - 1} [\text{W m}^{-3} \text{ sr}^{-1}] \quad (1.2)$$

where

$h$  is Planck's constant,  $6.6262 \times 10^{-34}$  Js

$c$  is the velocity of light,  $2.9979 \times 10^8$  m/s

$\lambda$  is the wavelength in meters

$k$  is the Boltzman constant,  $1.308 \times 10^{-23}$  J/K

$T$  is absolute temperature in degrees Kelvin (K)

Planck's law expressed in wavenumbers is

$$I_B(\nu) = \frac{2hc^2\nu^3}{\exp\left\{\frac{h\nu}{kT}\right\} - 1} [\text{W m}^{-1} \text{ sr}^{-1}] \quad (1.3)$$

where  $\nu$  is wavenumbers.

The emission intensities from a blackbody carry no chemical information, only temperature information. Figure 1.1 displays the characteristic blackbody curves for several temperatures.

In reality no object is a perfect blackbody, but most objects reflect and/or transmit radiation. In considering these objects, conservation of energy principles dictate that

$$A + T + R = 1 \quad (1.4)$$

where  $A$  is the absorptance,  $T$  is the transmittance (fraction of incident radiation transmitted), and  $R$  is the reflectance (fraction of incident radiation reflected). If a local thermal equilibrium is assumed (12), then by using Kirchhoff's law, equation 1.4 becomes

$$E + T + R = 1 \quad (1.5)$$

From this simple equation some approximate behavior may be anticipated by investigating the extremes. For  $T = 0$ , an opaque sample,

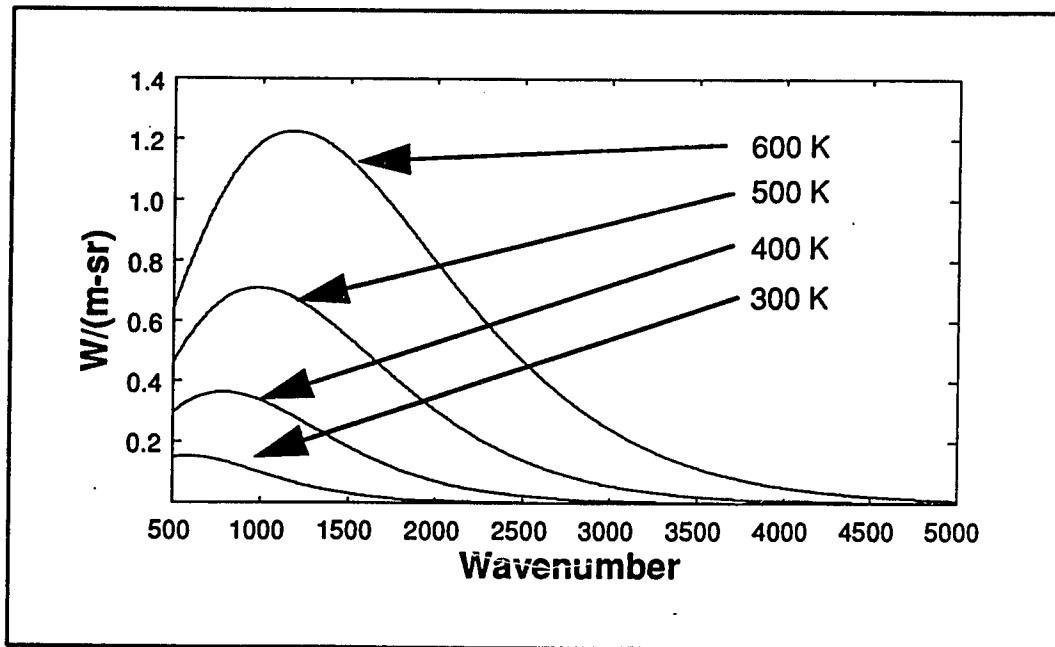


Figure 1.1 Graphical representation of Planck's law with wavenumber axis.

$$E = 1 - R \quad (1.6)$$

and there should be information available similar to that from a reflection experiment.

The other extreme is  $R = 0$ , a nonreflective sample,

$$E = 1 - T \quad (1.7)$$

and information complementary to that obtained from a standard transmission experiment should be available. Given  $T = e^{-\kappa L}$ , from Beer's law, where  $\kappa$  is the spectral absorption coefficient and  $L$  is the sample thickness, as  $\kappa L$  approaches zero,  $e^{-\kappa L}$  approaches  $1 - \kappa L$  and  $E \approx \kappa L$ . For substances with  $\kappa L < 0.05$ , it may be preferred to measure emissivity rather than the absorption in order to determine  $\kappa$ . This is due to the relative insensitivity of emission measurements to changes in the reflectivity as compared to transmission measurements (13).

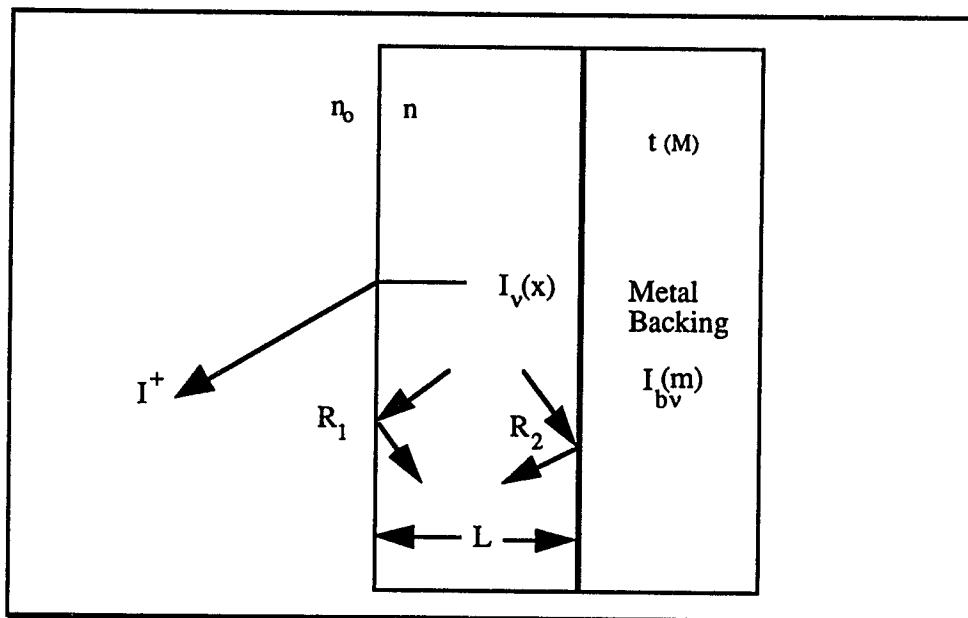
Most objects have transmission and reflection occurring simultaneously, and for solids there is the added complication of multiple internal reflections. A more complex theory is needed to account for these effects. McMahon (14) developed the theory for unidirectional emission normal to a surface from semitransparent materials and Gardon (15) extended this work to include hemispherical radiation. These papers rely on an accounting or ray tracing approach which has more recently been reiterated by Hvistendahl (16). A more flexible approach has been given by Lauer using the radiative transfer equation. This approach was developed by the astrophysicist and provides for inclusion of spatial and angular dependencies. The radiative transfer theory approach will be presented in an abbreviated form in Section 1.2.1.

### 1.2.1 Radiative Transfer Theory

The radiative transfer theory (17) begins with the radiative transfer equation

$$\frac{dI_v(x)}{dx} = -\kappa_v [n_v^2 I_{bv}(x) - I_v(x)] \quad (1.8)$$

where  $n$  is the refractive index of the medium,  $\kappa$  is the spectral absorption coefficient,  $I_{bv}(x)$  is the radiance given by the Planck equation, and  $I_v(x)$  is the monochromatic radiance at frequency  $v$ . This equation simply states that the change in intensity of radiation with distance is the difference between the radiation emitted (first term on the right of equation 1.8) and the radiation absorbed (second term on the right of equation 1.8). Figure 1.2 displays the physical model.



**Figure 1.2** Physical model for radiative transfer theory discussion.

This equation may be solved, given the proper boundary conditions, for a variety of experimental situations using the following assumptions: 1) Fresnel's equations apply, 2) the material investigated is isotropic, homogeneous and may emit and absorb, but not scatter thermal radiation, and 3) refraction of radiation within the sample due to variations of refractive index with temperature are negligible. The interested reader is referred to Lauer's discussion for the details (17).

Given the appropriate boundary conditions, the above equation may be solved for the radiation emitted in the positive direction,  $I^+$  (12),

$$I_V^+(0) = \left\{ (1 - R_1) \left[ \frac{n_0}{n} \right]^2 \beta \right\} X \left\{ (1 - R_2) n^2 I_{bv}(M) e^{-\kappa L} + \int_0^L n^2 I_{bv}(x) \kappa \left[ e^{-\kappa x} + R_2 e^{-\kappa(2L - x)} \right] dx \right\} \quad (1.9)$$

where  $\beta = (1 - R_1 R_2 \exp[-2\kappa L])^{-1}$ ,  $I_V^+(0)$  is the exiting radiance as shown in Figure 1.2,  $I_b(M)$  denotes the contribution by the metal backing at temperature  $t(M)$ , given by Planck's equation.  $R_1$  and  $R_2$  are the front surface and back surface reflectivities, and  $n$  is the index of refraction of the film which is related to  $n_0$ , the index of refraction of the boundary material (air), via Snell's law.

Consider the sampling situation of a free standing material. If it is known that  $n$  and  $\kappa$  are constant with respect to  $x$ , the position in the sample, and  $R_1 = R_2 = R$ , and the sample temperature is uniform, then  $I_{bv}(x)$  is constant. Since there is no external source (no metal backing),  $I_{bv}(M) = 0$  and equation 1.9 gives

$$I = \frac{(1 - T)(1 - R)}{(1 - RT)} I_B \quad (1.10)$$

Let  $E = I/I_B$ , and assume  $n_0 = 1$  then

$$E = \frac{(1 - T)(1 - R)}{(1 - RT)} \quad (1.11)$$

Equation 1.11 reduces to equation 1.6 as  $T$  approaches zero and to equation 1.7 as  $R$  approaches zero. The description of the transmissive and reflective behavior of transparent materials is also complicated by internal reflections which give rise to apparent transmissivity,  $T^*$ , and apparent reflectivity,  $R^*$ . These quantities may be related to the measured quantities (14),  $T$  and  $R$ , and are given here for completeness

$$R^* = R \left[ 1 + \frac{T^2 (1 - R)^2}{1 - T^2 R^2} \right] \quad (1.12)$$

$$T^* = T \frac{(1 - R)^2}{1 - R^2 T^2} \quad (1.13)$$

Another sampling situation that will be of interest, and easily investigated with the radiative transfer approach, is that of a thin film on a metal surface, with both the film and metal surface at the same temperature. This is the usual situation for obtaining emission spectra (18). As the layers of film increase in thickness, poorer spectral contrast (19) and anomalous spectral features have been observed (18). These anomalous features, characterized as dips in the spectra where peaks are expected, have been ascribed to temperature gradients by Griffiths (18), Lauer (17) and Bates and Boyd (20), while Hvistendahl and coworkers (16) maintain it is a reflective effect. It would appear that either explanation may be valid depending on the sample type. For a film at a uniform temperature,  $I_{bv}(x) = I_{bv}(M) = I_b$  and thus

$$I = (1 - R_1) \beta I_b [ (1 - R_2) e^{-\kappa L} - e^{-\kappa L} + 1 + R_2(e^{-\kappa L} - e^{-2 \kappa L}) ] \quad (1.14)$$

which gives

$$E = \frac{(1 - R_1)(1 - R_2T^2)}{1 - R_1R_2T^2} \quad (1.15)$$

As may be seen from equation 1.15, the emittance depends on the square of the transmissivity, which indicates the path length is essentially doubled by the reflective metal backing. Note that the reflectivity of the metal backing,  $R_2$ , is paired with the transmissivity. This indicates the importance of a highly reflective backing for maximum sensitivity to the film transmissivity.

### 1.3 Experimental Considerations

The experimental aspects of obtaining infrared emission spectra in the range from about 4000 to 450  $\text{cm}^{-1}$  (2.5 to 22.2 micron ( $\mu\text{m}$ )) are considered in this section. It is hoped that the reader will gain an appreciation for the advantages and disadvantages of using infrared emission spectroscopy and an understanding of some of the practical problems involved in measuring infrared emission spectra. This section will not discuss the experimental aspects of laser stimulated emission spectroscopy, but this subject will be briefly reviewed in Section 1.4.3.

One of the most fundamental considerations for obtaining an infrared emission signal is the temperature of the sample relative to the detector. In order to measure an emission signal the sample must be at a temperature different from the detector. The sample may be at a temperature lower or higher than the detector, as



demonstrated by Chase (21). Further, the sample must be at a temperature different from the surroundings, in order to differentiate it from the surroundings. This immediately points out one of the disadvantages of infrared emission, which is that the sample must be at a different temperature than the surroundings and thus must be heated or cooled. Process analysis conditions may readily meet this requirement, as many materials are produced at above ambient temperatures.

The surroundings that may contribute to the background signal include all the optical components within the spectrophotometer. In order to increase the signal-to-noise ratio in an emission experiment the sample temperature may be increased or the background temperature decreased. A decrease in the background temperature has been achieved by cooling the entire spectrophotometer to 77 K by Allara and coworkers (22). Cooling of the spectrophotometer to reduce the background was necessary in order to observe emission spectra of monolayer quantities of chemisorbed p-nitrobenzoic acid on a copper substrate. Reduction in the background temperature by cooling the spectrophotometer is particularly important for the application of Fourier transform infrared spectroscopy (FT-IR). This is because the multiplex advantage is realized only if the measurement is detector noise limited, which may not be the case depending on the region of the spectrum analyzed (23). Huppi discusses the details of cryogenic instrumentation for Fourier transform emission spectroscopy, especially as it relates to atmospheric studies (24).

The warmer the sample the greater the intensity of radiation emitted from the higher energy regions of the spectrum. This points out another disadvantage of infrared emission spectroscopy. In order to detect spectral features above  $2000\text{ cm}^{-1}$  (below  $5\text{ }\mu\text{m}$ ) sample temperatures must be relatively high, often approaching degradation temperatures for many organic compounds.

In order to compute the emittance spectrum from the single beam emission signal, the sample must be ratioed to the spectrum of a blackbody measured at the same temperature. Complications associated with this procedure have been investigated by Kember et al. (25). They found emission from a heated beam splitter introduced extraneous spectral features in the emittance spectrum. In order to overcome this problem a spectrum of the beam splitter was measured by placing a mirror at the sample position. The beam splitter emission spectrum was subtracted from both the sample emission signal and the blackbody emission signal before taking the ratio. This resulted in other spectral anomalies associated with the phase difference between the beam splitter emission and the sample emission. This phase difference problem was overcome by subtraction of interferograms before ratioing. Other complications were encountered when using a cooled detector because of the sensitivity of cooled detectors for background radiation reflected from the surroundings into the spectrophotometer. A four measurement technique was used to eliminate the reflected background radiation problem. The sample spectrum was measured at two temperatures without removing it from the sample holder and the blackbody was measured at the same two temperatures without removing it from the spectrophotometer. The interferograms of the samples at the two temperatures were subtracted from each other as were the interferograms for the blackbodies and the differences ratioed to compute the emittance spectrum. This is shown in equation 1.16 as

$$E = \frac{I_s(T_1) - I_s(T_2)}{I_b(T_1) - I_b(T_2)} \quad (1.16)$$

In equation 1.16,  $I_s(T_1)$  and  $I_s(T_2)$  are the sample interferograms at temperatures  $T_1$

and  $T_2$ , and  $I_b(T_1)$  and  $I_b(T_2)$  are the blackbody interferograms at temperatures  $T_1$  and  $T_2$ . This procedure had been described earlier by Weber (26).

Anomalous features in infrared emittance spectra have been attributed to temperature gradients (18) or selective reflection at the sample surface (16, 27). In order to eliminate the selective reflection problem, Hvistendahl et al. have proposed the ratioing of a single beam measurement to the single beam measurement of an optically thick sample (16). They support their contention with theoretical arguments, spectral simulations and experimental verification. Recently Rytter and Einarsrud have suggested that this procedure of ratioing to an opaque sample should be used to eliminate similar distortions in reflection spectroscopy (28).

Another method of background compensation involves the chopping of the input light so that the spectrophotometer alternately senses the signal of interest and a blackbody signal at the same temperature (17). A lock-in amplifier is tuned to the chopper frequency and phase locked in such a way that the detector sees a source and a reference signal at alternating half cycles. Thus only the difference between the signals is amplified.

A recent paper by Wadayama et al. has demonstrated the use of a double modulation technique in which the blackbody signal is mechanically chopped at 2.3 kHz, while the emission signal is modulated at 65 Hz using a wire-grid polarizer (29). The signals were combined at a beam splitter and wavelength selection was performed using a wavelength-variable wedge filter. Phase sensitive detection was used to demodulate the signals.

Chase has commented on another problem that may be encountered in the measurement of infrared emission spectra, that of multipassing of the radiation (21). If the sample is placed normal to the spectrophotometer, emitted radiation that is

returned to the sample from the beamsplitter of a Fourier transform instrument may be reflected from the sample back into the spectrophotometer. This effect can lead to higher order spectral features. Chase suggests either filtering the signal, mathematically compensating, or tilting the sample away from the norm, so that light returning from the beamsplitter is not reflected back into the spectrophotometer.

It has been calculated that optimal measurements of emission spectra of thin layers on metal surfaces are made at angles of 70-80° from the norm (30). In this instance thin means that the distance between the metal surface and the sample dipole is very small compared to the wavelength of the emitted radiation.

Sample shape and sample preparation are less critical for emission spectroscopy. Barr has demonstrated that rough surfaces have a minimal effect on emission spectra as compared to transmission spectra, which are severely effected by scattered light (13). Nagasawa and Ishitani have demonstrated that emission spectra of wire coatings may be measured and the requirement for a planar surface, as in reflection spectroscopy, is relaxed for emission measurements (31).

Because the sample must be at a temperature different from the surroundings, and because this is usually accomplished by heating the sample, the problem of temperature gradients in the sample may lead to spectral anomalies. These anomalies have been ascribed to selective reflection at the sample surface, but it would appear that for thick samples, temperature gradients must be considered. As the samples become thicker the spectra are more featureless and convey less information. Similar to an absorbance experiment, as the sample thickness increases less light is transmitted, and as the absorbance approaches one so must the emittance. Thus the sample thickness problem is shared by both absorbance and emission techniques, but laser heating of the surface may overcome this problem. This sampling technique is

discussed in Section 1.4.3.

One further experimental consideration is that of dynamic range. In an absorbance or reflection experiment, the resultant spectrum may be a small difference between two large numbers, while an emittance spectrum calculation may involve a small difference between small numbers. This may be particularly important in the analysis of trace quantities, for example monolayer surface species.

New hardware ideas recently introduced include the half aperture method of Ford and Spragg (32). This is a convenient method for measuring emission spectra using any FT-IR instrument. A mirror is inserted at a 45° angle at the Jacquinot stop so as to cover half the area of the Jacquinot stop image. The source is shut off and the emitted sample radiation directed into the interferometer via the 45° mirror. The modulated radiation that is reflected back from the interferometer is imaged at the other half of the Jacquinot stop and passes to the detector. Because the mirror covers half the Jacquinot stop, this experiment is less efficient than the normal method of measuring an emission spectrum, which is to replace the spectrophotometer source with the sample.

Harrick has introduced a new accessory for infrared emission measurements using an ellipsoidal mirror (33). By using an ellipsoidal mirror light is collected over a larger solid angle, increasing the signal and improving the signal-to-noise ratio. The sampling of small areas is facilitated by a microscope attachment mounted on the emission accessory. A blackbody mask with a selection of hole sizes allows one to control the sample area from which the emission spectrum is measured. The sample and mask may be tilted over a broad angle range in order to improve signal intensities for thin films on metal surfaces, as found by Greenler (30).

In review, infrared emission spectroscopy suffers from low signal levels that

may limit the useful frequency range, but at the same time this gives it a dynamic range advantage over more conventional techniques. Samples must be at a temperature different from the surroundings in order to differentiate the sample from the surroundings, but this may be less of a problem for process analytical chemistry because many materials are produced at above ambient temperatures. Emission spectroscopy is somewhat more flexible in terms of sample shape and sample preparation, but thicker samples show a loss of information, just as is experienced in absorption measurements. Due to the fact that emission samples are usually heated, spectral anomalies may be introduced, although with stimulated emission this feature can be turned to an advantage.

#### **1.4 Survey of Infrared Emission Spectroscopy Applications**

Reviews of the infrared emission spectroscopy literature through 1976 were presented by Bates (34) and Sheppard (27). Bates confines his review to FT-IR emission spectroscopy applications in chemistry and physics, excluding astronomy and astrophysics. He discusses thermally excited gases, solids and liquids, and infrared chemiluminescence from diatomic and polyatomic products of gas phase reactions. Sheppard's review is also confined to FT-IR emission applications, but does not offer as wide a coverage of applications outside the area of chemical analysis. Low, McManus and Abrams have provided a review of early non-FT emission work of a chemical analytical nature (35). Griffiths and deHaseth provide good background material in their book on FT-IR spectroscopy (23). The following discussion will summarize recent contributions, 1976 to the present. It will include only those papers relevant to the present research or of fundamental importance.

Some older material may be included if not referenced by Bates or Sheppard. The following is divided into the broad headings of gases and condensed phases with further sub-divisions as necessary.

#### **1.4.1 Gases**

The review of gas analysis using infrared emission spectroscopy is divided into laboratory measurements and remote sensing or atmospheric applications.

##### **1.4.1.1 Laboratory Measurements**

Bates has pointed out that laboratory measurements of the infrared emission spectra of gases are important not only for providing molecular structure information, but also because they can provide essential background information for interpreting the results of remote sensing measurements and the emission spectra of the atmosphere (34). Bates describes many applications of infrared emission spectroscopy for laboratory measurements of gaseous molecules. There has been a special interest in the spectra of molecules involved in high temperature flames or arc discharges in order to aid in interpreting the emission spectra of stellar atmospheres. Much work in this area has been performed by Mantz and coworkers (36,37).

Gross and Griffiths have reported on the estimation of temperatures of pure CO<sub>2</sub> samples using emission bands near 1000 cm<sup>-1</sup> (10 μm) and a method of nonlinear regression (38). This work is an extension of similar studies by Gross and Griffiths involving temperature estimation using absorption measurements (39). The estimation of temperature is important for calculation of the emissivity spectrum of

remote sources for pollutant studies.

The stimulation of infrared emission spectra of gases using a laser light source has been reported by Robinson and coworkers (40, 41). They have reported on quantitative analysis of simple organic molecules and have provided a catalog of laser-stimulated infrared emission spectra. More recently Belz et al. have used a continuous wave CO<sub>2</sub> laser to excite infrared emission spectra of nearly all molecular gases by adding small amounts of SF<sub>6</sub> (42).

Hudson and Busch have reported on the use of infrared emission for monitoring the emission from a flame, and thus provided a basis for detection of chromatographic profiles (43,44). Recently Fletcher and Leone have described the use of time resolved FT-IR emission to monitor the photodissociation of C<sub>2</sub>H<sub>2</sub> (45). Rotational state distribution of CCH was reported, as well as the rotational energy. Leone discusses further application of the technique for molecular photofragmentation studies in a subsequent publication (46).

#### **1.4.1.2 Remote Sensing and Atmospheric Applications**

Some of the first remote sensing measurements for environmental pollutants were made by Low and Clancy. They demonstrated the use of FT-IR for the monitoring of a power plant smoke stack (47). Bands due to CO<sub>2</sub> and SO<sub>2</sub> were clearly identifiable in the spectra. Methods for the determination of plume temperature and concentration have been investigated by Prengle and coworkers (48). Recently Small et al. have developed a methodology for the detection of atmospheric pollutants by direct analysis of infrared emission interferograms (49). With a digital filtering approach they were able to use a mobile spectrophotometer without the



problems associated with a changing background.

There has been much interest in the emission spectrum of the atmosphere. Mixing ratios and pollutant concentrations are of particular interest. Recent contributions are noted from Kendall et al., in which the emission spectrum of the lower stratosphere was measured (50). Another study by Park and Carli made use of nonlinear least squares fitting techniques to retrieve gas concentrations, instrument line shape, and instrument response functions simultaneously (51,52). Rinsland et al. have also used nonlinear least squares spectral curve fitting to derive stratospheric vertical temperature profiles from balloon-borne measurements of the  $962\text{ cm}^{-1}$  ( $10.4\text{ }\mu\text{m}$ ) band of  $\text{CO}_2$  (53).

#### **1.4.2 Condensed Phases**

Within the area of condensed phase samples there are several subtopics. A consideration of general applications is given first and more sample specific references are discussed in subsequent sections.

##### **1.4.2.1 General Applications for Condensed Phases**

The early work in this area, 1900-1910, was pioneered by Coblentz (54-57). He investigated the emission spectra of many materials in the  $700\text{-}1400\text{ }^{\circ}\text{C}$  temperature range and in the  $10,000\text{ - }1250\text{ cm}^{-1}$  ( $1\text{ - }8\text{ }\mu\text{m}$ ) wavenumber range.

In the late 50's and early 60's, Weber investigated the measurement of spectral emissivity of solids at low temperatures ( $30\text{ to }100\text{ }^{\circ}\text{C}$ ) (26,58). He introduced the method of calculating emissivity using four measurements that was later reintroduced

by Kember et al. (25).

Stierwalt et al. have demonstrated the utility of using emission measurements to determine absorption coefficients of near transparent materials (59). For near transparent materials, the reflections at the interfaces for transmittance measurements may greatly affect the computed absorption coefficients. There is much less of a problem in emission experiments with these materials. Stierwalt has also pointed out the utility of emission measurements for nearly opaque materials (60). He shows that by rearrangement of equation 1.11 to yield equation 1.16

$$\kappa d = \ln \left[ \frac{1 - R - ER}{1 - R - E} \right] \quad (1.16)$$

the calculation of the absorption coefficient,  $\kappa$ , is possible by the measurement of the emission at two thicknesses, even if the reflectivities are unknown. This gives rise to two simultaneous transcendental equations. Stierwalt and Potter also reported on the measurement of infrared emission spectra of Si, Ge and CdS in the 5000-400  $\text{cm}^{-1}$  (2 - 25  $\mu\text{m}$ ) range at temperatures between 50 and 200  $^{\circ}\text{C}$  (61).

Kapff measured emission spectra of liquids using a dispersive instrument showing the dependence of the spectra on thickness and temperature (62). The emission spectra compared well with absorption spectra for thin samples ( $2.54 \times 10^{-3}$  cm), while most of the bands were obscured by thicknesses of  $2.54 \times 10^{-2}$  cm.

Low has made numerous contributions to the field of infrared emission spectroscopy. He summarizes many of these applications in an applied spectroscopy review article (35). Other publications include the investigation of emission spectra of minerals (63,64), oleic acid on aluminum plates (65), and the remote sensing of stack gases (47).

Fabbri and Baraldi have measured emission spectra of vanadium oxides, copper sulfate and silver nitrate using a grating spectrophotometer in the temperature range of 100 to 400 °C (66). They have presented the results with respect to variations in support, temperature, sample fineness and layer thickness. Increased layer thickness and larger crystal size were found to give poorer spectral contrast. As the temperature of  $V_2O_2$  and  $V_2O_4$  was raised to 200 °C the emission spectrum was observed to undergo rapid and irreversible changes. The oxidation to  $V_2O_5$  was the apparent product, but it is suggested that monitoring of the process for nonstoichiometric oxide intermediates may lend valuable information with respect to the catalytic activity of these compounds.

Griffiths has commented on the basic considerations for infrared emission spectroscopy (18). He displays spectra of thin and thick films of silicon grease on heated aluminum foil and explains the distortion of the spectra in thicker samples as the result of temperature gradients.

Lauer and coworkers have made many contributions to the literature with respect to the use of infrared emission for the study of lubrication problems. In an initial study they used a slow scanning interferometer with the radiation from the sample chopped, the detector alternately seeing the sample and a blackbody reference (17). Phase sensitive detection is used to measure emission spectra from a lubricant trapped between a steel bearing and a diamond window. The importance of proper adjustment of the reference temperature is explained and it is demonstrated that the result of misadjustment gives rise to peak inversions. They state that they have learned to live with this problem by optically limiting the frequency and by obtaining duplicate spectra with slightly different reference temperatures. This is an interesting explanation because they claim in the theory section of their paper that peak inversion

can result from temperature gradients, and so it is unclear how they can tell the difference between the two effects. A subsequent publication presented dramatic spectral changes from a fluid trapped between a bearing and diamond window as the gap distance was decreased. They explained the observations as an indication of molecular alignment (67). Other publications have contributed further to the study of lubrication problems (68,69).

Frech and Bates have reported on the use of infrared emission in conjunction with polarized Raman and infrared reflection techniques in an extensive study of Na $\beta$ -Al<sub>2</sub>O<sub>3</sub> over a temperature range from 12 to 1073 K (70).

Chase has commented on the sensitivity and limitations of condensed phase infrared emission spectroscopy using a Fourier transform instrument (21). He discussed elimination of the instrumental background and illustrated the sensitivity of the technique by measuring microgram sample quantities. Spectra measured near and below room temperature using a room temperature detector are presented, demonstrating that only a temperature difference between the detector and the sample was required for the measurement of emission spectra.

A very interesting application of infrared emission has been published by Folberth and Heim (71). A Fourier transform spectrophotometer operated in the 190 to 420 cm<sup>-1</sup> (52.7 - 23.8  $\mu$ m) region was used to measure the emission spectra of human skin. The study involved the discrimination of patients with chronic diseases from a symptom free control group. The results were encouraging.

Solomon et al. have made use of the combination of emission and transmission in the infrared analysis of particulates (72). They used a Fourier transform instrument to demonstrate the feasibility of using this method for analysis of chemical composition and size of gas-suspended solids and liquids for on-line

applications. The general concepts of the emission/transmission (E/T) method were developed and validated with particles of varying sizes.

Recently Ohta et al. have simulated infrared emission spectra in order to evaluate space radiator performance (73). The method of simulation includes the measurement of the attenuated total reflectance (ATR) spectrum and then an iterative refinement procedure to compute the refractive index and absorption coefficient spectra that best reproduce the ATR spectrum. It is not clear that a unique result may be found from this procedure. Given the two basic spectral curves, the emission spectrum was then calculated based on exact optical theory (ie. ray tracing method). They have also recently published the results of a comparative study of the numerical integration methods that are best suited for the calculations involved in the Kramers-Konig relationship used for the emission spectra simulations (74). Also, recently published is a matrix formulation for calculation of emission spectra in stratified multilayered films (75).

Harrison and coworker have describe the measurement of thin polymer films using a double-beam grating spectrophotometer. They applied various corrections to the spectra to account for the blackbody reference being at a different temperature than the sample (76).

Nagasawa and Ishitani have reported on the measurement of infrared emission spectra of thin polymer layers on flat aluminum plates using an FT-IR instrument (31). Detection limits equivalent to Fourier transform infrared reflection absorption spectroscopy were found. An advantage of infrared emission was demonstrated by measurement on nonflat surfaces. A linear relationship was found between relative emission intensity and film thickness. This may be rationalized by recalling equation 1.15, describing emission from a thin film on a metal surface. If  $R_1 \sim 0$ , low front

surface reflectivity, then equation 1.15 may be approximated as  $E = 1 - R_2T^2$ . If the sample is thin, then  $T^2 = e^{-2\kappa d}$  may be approximated as  $T^2 = 1 - 2\kappa d$ , and thus a linear relationship between thickness and emission signal is found.

#### 1.4.2.2 Molten Salts

In the consideration of molten salts, the single largest advantage that emission spectroscopy has over other infrared techniques is that the optical apparatus need not come in contact with the corrosive high temperature melt. Molten salts provide an abundance of radiant energy and as a result were some of the first samples studied by emission spectroscopy using dispersive instrumentation.

Early work in this area was conducted by Wilmshurst (77). He measured emission spectra of silver chloride and sodium nitrate with a dispersive instrument in the range of  $3000 - 750 \text{ cm}^{-1}$  ( $3.3 - 13.3 \text{ }\mu\text{m}$ ), demonstrating the feasibility of the approach. He also observed splitting of peaks in the vicinity of strong absorption bands in the emission spectrum of thicker samples and attributed this to the changes in reflectivity due to the rapid changes in refractive index at frequencies associated with strong absorption bands (anomalous dispersion).

Kozlowski measured emission spectra of  $\text{NaNO}_3$ ,  $\text{KNO}_3$ , and  $\text{CsNO}_3$  near  $600^\circ\text{C}$  using a dispersive instrument (78). Measurements were also made on a  $\text{LiCl-KCl}$  eutectic at  $550^\circ\text{C}$ . It was noted that the spectra from the eutectic did not correspond well with previously reported absorption spectra and it was postulated that reaction with the Al blackplate had taken place giving rise to  $\text{AlCl}_4^-$ .

Bates and Boyd have reported the measurement of infrared emission spectra of alkali metal nitrates using a Fourier transform instrument (20). Spectra were

measured in the 50 to 3000  $\text{cm}^{-1}$  (200 - 3.3  $\mu\text{m}$ ) region. The distortion (splittings) of bands as the sample thickness was increased was noted and the explanation given was that the presence of a temperature gradient in the sample caused the outer cooler layers to reabsorb the emission from the inner warmer layers. Distortions of bands in the alkali-metal nitrates were interpreted as interionic interactions with some caution, because of the probability that the distortions were due to the postulated temperature gradient effect.

More recently Hvistendahl et al. have commented on the distortions observed in the spectra of thick molten salt samples (16). They maintain, as Wilmshurst had pointed out earlier, that the reflections at the surface in the vicinity of strong absorption peaks cannot be neglected and that this is, in fact, the source of the splittings observed by previous workers. They take issue with the previous explanation of temperature gradients, stating that excessive temperature gradients would probably have to be assumed in order to account for the observed distortions. Assuming a harmonic oscillator model, they simulate the distortions in the emission spectra assuming no temperature gradients. They then offer as a solution to the problem, ratioing to a thick sample rather than a blackbody in order to derive a more realistic representation of the internal transmittance of the sample. Experimental results computed with the following equation show that distortions in the vicinity of strong absorption bands are minimized.

$$E = \frac{I(\text{thin sample}) - I(\text{background})}{I(\text{thick/opaque}) - I(\text{background})} \quad (1.17)$$

### 1.4.2.3 Glass

There has been much interest by the glass industry in the use of infrared emission as concerns temperature measurement and heat transfer. Because the materials are partially transparent in the infrared, the modes of heat transfer must include "radiative conductance", as well as the more familiar convective and conductive modes.

Gardon has published on the emissivity of transparent materials and brought to light, from an engineer's point of view, the fact that the rate of emission is dependent on the refractive index of the medium surrounding the radiator (15). He points out that this was not new and that classical physicists had discussed this point some time ago (79). He has also published on the calculation of temperature distribution (80) in glass and has reviewed radiant heat transfer in glass (81).

Beattie et al. computed the emissivity spectrum of glass in the 10,000 to 1667  $\text{cm}^{-1}$  (1 to 6  $\mu\text{m}$ ) range given the absorption coefficients and a specified temperature distribution (82). They have assumed Wien's law in the derivation and integration of the differential equations. From the results of the calculations, they constructed a radiation pyrometer capable of measuring both the surface and inside temperatures of glass samples. Upon careful study of their paper, there appear to be errors in the equations for the linear temperature gradient case in their appendix (case b).

More recently Viskanta and coworkers have developed a method for recovering the temperature distribution in semi-transparent samples using infrared emission measurements (83). The method has been applied to dielectric coatings on conductor surfaces in which the gray case (absorption coefficient and indices of refraction assumed independent of wavelength) was assumed. Later the method was



further developed to enable the calculations of temperature gradients in glass samples, given the optical properties of the sample are known (84,85). The method involves iteratively solving the expressions for emerging intensity, and reconstructing the data that best fits the observed emission data.

Dvurechenskii et al. have reported on the measurement of spectral normal emittance of silica glass samples 2 to 10 mm thick in the  $5000$  to  $2000\text{ cm}^{-1}$  ( $2$  to  $5\text{ }\mu\text{m}$ ) region at temperatures from  $327$  to  $1427\text{ }^{\circ}\text{C}$  (86).

Borodai and Betina have presented results of the experimental investigation of the optical properties of quartz ceramic in the spectral region from  $25,000$  to  $2000\text{ cm}^{-1}$  ( $0.4$  to  $5.0\text{ }\mu\text{m}$ ) over a temperature range of  $800$  to  $1300\text{ }^{\circ}\text{C}$  (87).

#### 1.4.2.4 Surface Science

Some of the first attempts at the measurement of infrared emission spectra of surface absorbed molecules were those of Eischen and Pliskin (88). They attempted to measure the emission spectrum of thin films of oleic acid on aluminum, but were unable to observe any bands different from those of the bulk. Other attempts were made to measure carbon monoxide (CO) adsorbed on Pt, but these too failed. It is not clear why the experiments failed, but several speculations are given by Eischen and Pliskin. Some years later, Low and Inoue repeated the previous experiments of Eischens and Pliskin measuring the emission spectrum of oleic acid on aluminum and observed the expected shift in the carboxyl band at  $1718\text{ cm}^{-1}$  ( $5.82\text{ }\mu\text{m}$ ) to a longer wavelength of  $1563\text{ cm}^{-1}$  ( $6.4\text{ }\mu\text{m}$ ) indicative of the interaction with the metal surface (65).

Koga et al. have observed formic acid on aluminum by inducing a temperature

gradient in the sample (89). Using a diffuse reflectance attachment and filtering the radiation below  $2000\text{ cm}^{-1}$  (above  $5\text{ }\mu\text{m}$ ), the sample was heated by the radiation above  $2000\text{ cm}^{-1}$  (below  $5\text{ }\mu\text{m}$ ) and the emission from the sample below  $2000\text{ cm}^{-1}$  was measured. They observed two peaks at  $1585$  and  $1350\text{ cm}^{-1}$  ( $6.3$  and  $7.4\text{ }\mu\text{m}$ ) and noted the shift to lower frequency with deuteriated formic acid.

Several studies have dealt with metal oxides. Kember and Sheppard investigated copper surfaces (90) and Kember et al. worked with oxide films on aluminum (25). Zanzucchi and Yim have studied oxide coatings on iron and Invar (36% Ni-Fe alloy) in order to understand their heat dissipation characteristics (91).

Blanke et al. have described the instrumental conditions necessary for effective infrared emission spectroscopy. They made comparisons of the expected signal-to-noise ratio and dynamic range between emission and reflection for the observation of surface species (92).

Allara et al. cooled the entire spectrophotometer to liquid nitrogen temperatures in order to reduce the background contribution to the signal arising from the emission of the spectrophotometer itself (22). They reported spectra of monolayer quantities of chemisorbed p-nitrobenzoic acid on a thin, oxide coated, copper substrate at  $300\text{ K}$ .

Greenler has contributed to the study of absorbed molecules on metal surfaces by investigating the angular dependence of the emission from metal surfaces (30). He presented calculations that indicated the optimal viewing angle for infrared radiation was  $70$  to  $80^\circ$ .

More recently Conroy et al. have studied the emissivity of aluminum surfaces in order to better understand the difficulties involved in temperature determinations via remote sensing of infrared radiation (93).

#### 1.4.2.5 Reactions

Some investigations using infrared emission spectroscopy have focused on the monitoring of reactions. Primet et al. have investigated the interactions between propene and vanadium oxide (94). They concluded that infrared emission spectroscopy appears to be a very suitable method for studying the interactions of reactants with catalysts.

Van Woerkom and de Groot have reported on the design and construction of a microreactor assembly equipped with an IR transmitting window (95,96). They discussed the problems involved with the measurement of infrared emission spectra for the investigation of catalytic reactions. They concluded that the experimental difficulties were not prohibitive and presented emission spectra of powdered samples of Pt/gamma-Al<sub>2</sub>O<sub>3</sub> and methanol adsorbed on gamma-Al<sub>2</sub>O<sub>3</sub>. Van Woerkom has also described the use of infrared emission to investigate the curing behavior of organic coatings (95).

Gratton et al. have described the use of infrared emission for the study of the oxidation of metal surfaces (97). They showed good results using a standard dispersive spectrophotometer. They monitored the oxidation of molybdenum and observed increases in intensity and changes in bands shape during the reaction.

Jonson and coworkers have used infrared emission to monitor the spectral changes of a series of metal oxides, resulting from the oxidation of toluene (98). A comparison of a grating instrument and a Fourier transform instrument was made which showed the Fourier transform instrument provided superior performance.

Lauer et al. have used infrared emission spectroscopy to investigate the deposition from jet fuels on metal. Their results were consistent with the postulated

mechanism (99).

Wendt and coworkers have demonstrated the use of emission and absorption spectroscopy for characterization of a copper converting process (100). Because of the high temperatures used in the study (600 to 1200 °C), the wavelength region investigated was in the visible range from 400 to 700 nanometers (nm). A grating instrument was used with a linear diode array detector. Typical spectral acquisition times were from 10 to 100 milliseconds. One hundred scans were used to compute an average spectrum. Bands of PbS and PbO were observed at different times during the converting process and the ratio of the intensity of these bands was found to be in good agreement with thermodynamical calculations.

#### **1.4.3 Stimulated Infrared Emission**

Recently interest in obtaining emission spectra from opaque samples has lead to the investigation of stimulated infrared emission. Lin et al. used a continuous wave (CW) laser to excite surface emission from polymer materials at different thicknesses (101). They termed their approach "laser-induced thermal emission" (LITE) spectroscopy.

Jones and McClelland have demonstrated the use of CW and pulsed laser methods for stimulation of thermal emission (102,103). They have termed their approach "transient infrared emission spectroscopy" (TIRES). Laser saturation is overcome by rotation of the sample. They developed an approximate theory to explain the observations. They also made comparisons to photoacoustic absorption spectra measured on the same material.

## **1.5 Conclusions**

**This review of the literature has found few applications of infrared emission spectroscopy to quantitative analysis and no uses of multivariate data analysis. Because this technique is one of only a few that may be useful as a noninvasive analysis technique, the reinvestigation of it with the aid of multivariate data analysis is needed. The remainder of this dissertation will explore the application of chemometrics and infrared emission spectroscopy for remote noninvasive analysis.**

## 1.6 Notes to Chapter 1

1. Callis, J. B.; Illman, D. L.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 624A-637A.
2. McDonald, R. S. *Anal. Chem.* **1984**, *56*, 349R-372R.
3. McDonald, R. S. *Anal. Chem.* **1986**, *58*, 1906R-1925R.
4. Griffiths, P. R. *Appl. Spectrosc.* **1972**, *26*, 73-76.
5. Hirschfeld, T. *Appl. Opt.* **1978**, *17*, 1400-1412.
6. Ramos, L. S.; Beebe, K. R.; Carey, W. P.; Sanchez, E.; Erickson, B. C.; Wilson, B. E.; Wangen, L. E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 294R-315R.
7. Brown, S. D.; Barker, T. Q.; Larivee, R. J.; Monfre, S. L.; Wilk, H. R. *Anal. Chem.* **1988**, *60*, 252R-279R.
8. Pell, R. J.; Erickson, B. C.; Hannah, R. W.; Callis, J. B.; Kowalski, B. R. *Anal. Chem.* **1988**, *60*, 2824-2827.
9. Harrison, T. R. *Radiation Pyrometry and Its Underlying Principles of Radiant Heat Transfer*; Wiley: New York, 1960.
10. Hadni, A. *Essentials of Modern Physics Applied to the Study of the Infrared*; Pergamon Press: Oxford, 1967.
11. Wyatt, C. L. *Radiometric System Design*; MacMillan Publishing Company: New York, 1987.
12. Viskanta, R.; Anderson, E. E. *Advances in Heat Transfer* **1975**, *11*, 318.
13. Barr, J. K. *Infrared Physics* **1969**, *9*, 97-108.
14. McMahon, H. O. *J. Opt. Soc. Am.* **1950**, *40*, 376-380.
15. Gardon, R. J. *Am. Ceram. Soc.* **1956**, *39*, 278-287.
16. Hvistendahl, J.; Rytter, E.; Oye, H. A. *Appl. Spectrosc.* **1983**, *37*, 182-187.
17. Lauer, J. L.; King, V. W. *Infrared Physics* **1979**, *19*, 395-412.
18. Griffiths, P. R. *Appl. Spectrosc.* **1977**, *31*, 497-505.
19. VanKasteren, P. H. G.; Smeets, L. H. *SPIE - Fourier Transform Infrared Spectroscopy* **1981**, *289*, 83-86.
20. Bates, J. B.; Boyd, G. E. *Appl. Spectrosc.* **1973**, *27*, 204-208.

21. Chase, D. B. *Appl. Spectrosc.* **1981**, *35*, 77-81.
22. Allara, D. L.; Teicher, D.; Durana, J. F. *Chem. Phys. Lett.* **1981**, *84*, 20-24.
23. Griffiths, P. R.; deHaseth, J. A. *Fourier Transform Infrared Spectrometry*; John Wiley & Sons: New York, 1986.
24. Huppi, E. R. *Mikrochim. Acta* **1987**, *3*, 281-296.
25. Kember, D.; Chenery, D. H.; Sheppard, N.; Fell, J. *Spectrochim. Acta* **1977**, *35A*, 455-459.
26. Weber, D. *J. of Opt. Soc. Am.* **1959**, *49*, 815-820.
27. Sheppard, N. In *Analytical Applications of FT-IR to Molecular and Biological Systems: Proceedings of the NATO Advanced Study Institute held at Florence, Italy, August 31 to September 12, 1979*; Durig, J.R., Ed.; D. Reidel Pub. Co.: Hingham, MA, 1980; pp 125-140.
28. Rytter, E.; Einarsrud, M. *Mikrochim. Acta* **1988**, *2*, 307-309.
29. Wadayama, T.; Shiraishi, W.; Hatta, A.; Suetaka, W. *Appl. Surf. Sci.* **1990**, *44*, 43-47.
30. Greenler, R. G. *Surface Science* **1977**, *69*, 647-652.
31. Nagasawa, Y.; Ishitani, A. *Appl. Spectrosc.* **1984**, *38*, 168-173.
32. Ford, M. A.; Spragg, R. A. *Appl. Spectrosc.* **1986**, *40*, 715-716.
33. Handke, M.; Harrick, N. J. *Appl. Spectrosc.* **1986**, *40*, 401-405.
34. Bates, J. B. In *Fourier Transform Infrared Spectroscopy: Applications to Chemical Systems*; Vol. 1; Ferraro, J. R. and Basile, L. J., Eds.; Academic Press: New York; 1978; pp 99-142.
35. Low, M. J. D.; McManus, J. C.; Abrams, L. *Appl. Spectrosc. Rev.* **1971**, *5*, 171-210.
36. Mantz, A. W.; Maillard, J. P. *J. Mol. Spectrosc.* **1974**, *53*, 466-478.
37. Mantz, A. W.; Maillard, J. P.; Roh, W. B.; Rao, K. N. *J. Mol. Spectrosc.* **1975**, *57*, 155-159.
38. Gross, L. A.; Griffiths, P. R. *JQSRT* **1988**, *39*, 463-472.
39. Gross, L. A.; Griffiths, P. R. *JQSRT* **1988**, *39*, 131-138.
40. Robinson, J. W.; Woodward, C.; Barnes, H. M. *Anal. Chim. Acta* **1968**, *43*, 119-128.

41. Hailey, D. M.; Barnes, H. M.; Woodward, C.; Robinson, J. W. *Anal. Chim. Acta* **1971**, *56*, 161-174.
42. Belz, H. H.; Gutberlet, H.; Schallert, B.; Schrader, B. *Appl. Spectrosc.* **1987**, *41*, 1009-1019.
43. Hudson, K. M.; Busch, K. W. *Anal. Chem.* **1987**, *59*, 2603-2609.
44. Hudson, K. M.; Busch, K. W. *Anal. Chem.* **1988**, *60*, 2110-2115.
45. Fletcher, T. R.; Leone, S. R. *J. Chem. Phys.* **1989**, *90*, 871.
46. Leone, S. R. *Acc. Chem. Res.* **1989**, *22*, 139.
47. Low M. J. D.; Clancy F. K. *Environ. Sci. Tech.* **1967**, *1*, 73-74.
48. Prengle, H. W. J.; Morgan, C. A.; Fang, C.; Huang, L. *Current Research* **1973**, *7*, 417-423.
49. Small, G. W.; Kroutil, R. T.; Dittillo, J. T.; Loerop, W. R. *Anal. Chem.* **1988**, *60*, 264-269.
50. Kendall, D. J. W.; Clark, T. A. *Int. J. Infrared and Millimeter Waves* **1981**, *2*, 783.
51. Park, J. H.; Carli, B. *Appl. Opt.* **1986**, *25*, 3490-3501.
52. Carli, B.; Park, J. H. *J. Geo. Res.* **1988**, *93*, 3851-3865.
53. Rinsland, C. P.; Goldman, A.; Murcray, F. J.; Murcray, D. G.; Smith, M. A. H.; Seals, R. K. J.; Larsen, J. C.; Rinsland, P. L. *JQSRT* **1983**, *4*, 327-334.
54. Coblenz, W. W. *Tech. News Bull. Bur. Standards* **1908**, *5*, 159.
55. Coblenz, W. W. *Tech. News Bull. Bur. Standards* **1910**, *6*, 301.
56. Coblenz, W. W. *Tech. News Bull. Bur. Standards* **1910**, *7*, 243.
57. Coblenz, W. W. *Tech. News Bull. Bur. Standards* **1912**, *9*, 81.
58. Weber, D. *J. of Opt. Society of Am.* **1960**, *50*, 808-810.
59. Stierwalt, J. B.; Bernstein, J. B.; Kirk, D. D. *Appl. Opt.* **1963**, *2*, 1169-1173.
60. Stierwalt, D. L. *Appl. Opt.* **1966**, *5*, 1911-1915.
61. Stierwalt, D. L.; Potter, R. F. In *Report of the International Conference on the Physics of Semiconductors* **1962**, 513-520.



62. Kapff, S. F. *J. Chem. Phys.* **1948**, *16*, 446-453.
63. Coleman, I.; Low, M. J. D. *Spectrochim. Acta* **1966**, *22*, 1293-1298.
64. Low, M. J. D. *Nature* **1965**, *208*, 1089.
65. Low, M. J. D.; Inoue, H. *Anal. Chem.* **1964**, *36*, 2397-2399.
66. Fabbri, G.; Baraldi, P. *Appl. Spectrosc.* **1972**, *26*, 593-599.
67. King, V. W.; Lauer, J. L. *Journal of Lubrication Technology* **1981**, *103*, 65-73.
68. Lauer, J. L.; Keller, L. E. *SPIE - Fourier Transform Infrared Spectroscopy* **1981**, *289*, 87-93.
69. Lauer, J. L.; King, V. W. *Am. Soc. of Lubrication Engineers*; Preprint of paper presented at Anaheim Conference; **1980**.
70. Frech, R.; Bates, J. B. *Spectrochim. Acta* **1979**, *35A*, 685-694.
71. Folberth, W.; Heim, G. *SPIE - Fourier and Computerized Infrared Spectroscopy* **1985**, *553*, 129-130.
72. Solomon, P. R.; Carangelo, R. M.; Hamblen, D. G.; Best, P. E. *Appl. Spectrosc.* **1986**, *40*, 746-759.
73. Ohta, K.; Graf, R. T.; Ishida, H. *Appl. Spectrosc.* **1988**, *42*, 114-120.
74. Ohta, K.; Ishida, H. *Appl. Spectrosc.* **1988**, *42*, 952-957.
75. Ohta, K.; Ishida, H. *Appl. Opt.* **1990**, *29*, in press.
76. Harrison, N.; Bilen, C. S.; Morantz, D. J. *Polymer Communications* **1984**, *25*, 15-17.
77. Wilmshurst, J. K. *J. Chem. Phys.* **1963**, *39*, 2545-2548.
78. Kozlowski, T. R. *Appl. Opt.* **1968**, *7*, 795-800.
79. Drude, P. *Lehrbuch der Optik* Hirzel: Leipzig, **1912**; 494.
80. Gardon, R. *J. Am. Ceram. Soc.* **1958**, *41*, 200-209.
81. Gardon, R. *J. Am. Ceram. Soc.* **1961**, *44*, 305-312.
82. Beattie, J. R.; Cohen E. *British J. of Appl. Phys.* **1960**, *11*, 151-157.
83. Chupp, R. E.; Viskanta, R. *AIAA* **1970**, *8*, 551-557.
84. Hommert, P. J.; Viskanta, R.; Chupp, R. E. *J. Am. Ceram. Soc.* **1975**, *58*, 58-62.

85. Viskanta, R.; Hommert, P. J.; Groninger, G. L. *Appl. Opt.* **1975**, *14*, 428-437.
86. Dvurechenskii, A. V.; Petrov, V. A.; Reznik, V. Y. *High Temperatures-High Pressures* **1979**, *11*, 423-428.
87. Borodai, S. P.; Betina, T. A. *Zhurnal Prikladnoi Spektroskopii* **1988**, *48*, 114-119.
88. Eischen, R. P.; Pliskin, W. A. In *Advances in Catalysis and Related Subjects*; Vol 10; Frankenburg, Walter G., Ed.; Academic Press: New York; 1958, pp 2-54.
89. Koga, O.; Onishi, T.; Tamaru, K. *J.C.S. Chem. Comm.* **1974**, 464.
90. Kember, D.; Sheppard, N. *Appl. Spectrosc.* **1975**, *29*, 496-500.
91. Zanzucchi, P. J.; Yim, M. W. *Appl. Spectrosc.* **1986**, *40*, 1042-1046.
92. Blanke, J. F.; Vincent, S. E.; Overend, J. *Spectrochim. Acta* **1976**, *32a*, 163-173.
93. Conroy, C. M.; Guthrie, J. D.; Sharkins, A. J.; Sparr, B. J.; Crocombe, R. A.; Curbelo, R. *Appl. Spectrosc.* **1987**, *41*, 688-692.
94. Primet, M.; Fouilloux, P.; Imelik, B. *Surf. Sci.* **1979**, *85*, 457-470.
95. VanWoerkom, P. C. M.; deGroot, R. L. *Appl. Opt.* **1982**, *21*, 3114-3118.
96. VanWoerkom, P. C. M.; Blok, P.; VanVeenendaal, H. J.; deGroot, R. L. *Appl. Opt.* **1980**, *19*, 2546-2550.
97. Gratton, L. M.; Paglia, S.; Scattaglia, F.; Cavallini, M. *Appl. Spectrosc.* **1978**, *32*, 310.
98. Jonson, B.; Rebenstorf, B.; Larsson, R.; Primet, M. *Appl. Spectrosc.* **1986**, *40*, 798-803.
99. Lauer, J. L.; Vogel, P.; Seng, G. T. *Appl. Spectrosc.* **1985**, *39*, 997-1004.
100. Wendt, W.; Alden, M.; Persson, W. *Appl. Spectrosc.* **1988**, *42*, 128-133.
101. Lin, L. T.; Archibald, D. D.; Honigs, D. E. *Appl. Spectrosc.* **1988**, *42*, 477-483.
102. Jones, R. W.; McClelland, J. F. *Anal. Chem.* **1989**, *61*, 1810-1815.
103. Jones, R. W.; McClelland, J. F. *Anal. Chem.* **1989**, *61*, 650-656.

## **Chapter 2 Multivariate Analysis**

### **2.1 Introduction to Chapter 2**

Multivariate analysis is the discovery of regularities in the behavior of two or more variables and the testing of alternative models of association between two or more variables (1). The first pursuit is exploratory in nature, while the second is confirmatory. Multivariate analysis includes concepts and tools from applied mathematics, statistics, geometry, and calculus and involves a varied collection of many techniques. There are many references under the broad heading of multivariate analysis. A few have been particularly helpful and are highlighted here. Green has an excellent description of the mathematical tools necessary for multivariate analysis, including geometric interpretations (1). Jolliffe discusses principal component analysis in his book of the same name (2), while Mardia, Kent and Bibby (3) provide a more statistical treatment of multivariate analysis. Golub and Van Loan furnish the numerical analysis tools for the matrix computations involved in multivariate analysis (4).

Several factors have motivated the development of multivariate techniques for analytical chemistry. Multichannel detection and the coupling of techniques such as gas chromatography and mass spectrometry have made large quantities of multiresponse data reasonably easy and inexpensive to collect compared with acquiring new samples. Thus it is necessary to extract as much information as possible from the available data. Computing power, both in terms of hardware and software, is becoming more accessible, allowing multivariate analysis to be more

easily performed. For process analytical chemistry, multivariate analysis with non-selective robust sensors may be used to replace less robust, more selective sensors, making multivariate analysis indispensable. Not only is multivariate analysis able to increase productivity, it may provide information that can be found in no other way.

The application and development of multivariate techniques for use in chemistry are part of the mission of chemometrics. Since its inception in 1972, chemometrics has been growing steadily and has gained acceptance as a useful, if not required, area of study (5). There is a growing base of literature on the subject including Malinowski and Howery's book on factor analysis in chemistry (6), two general texts (7,8), and a fourth on the specific area of multivariate calibration (9). There are currently two journals devoted to chemometrics research (10,11). Chemometric literature reviews have been found in *Analytical Chemistry* every two years since 1980 (12-17) and *Applied Spectroscopy* provides examples of chemometrics applied to spectral data.

Multivariate analysis techniques for analytical chemistry include exploratory techniques such as cluster analysis and pattern recognition, and more quantitative techniques such as multivariate calibration and optimization. For this dissertation two of the quantitative techniques will be of interest, multivariate calibration and multiresponse nonlinear optimization. Multivariate calibration is an implicit or "soft" modeling technique with few assumptions made concerning the data structure. Multiresponse nonlinear optimization uses a "hard" theoretical model. It is concerned with estimation of the parameters associated with the theoretical model and resolution of the underlying basic waveforms representing the phenomena. The next two sections detail the theory needed to understand the techniques that are a part of

multivariate calibration and multiresponse nonlinear optimization used in this dissertation.

## 2.2 Multivariate Calibration

### 2.2.1 Overview

Multivariate calibration relates two or more indirectly measured variables (response measurements) from two or more objects (samples) with known properties via a mathematical model. The model is used to predict the same properties from the indirectly measured variables of unknown objects. There are many methods for multivariate calibration. Sanchez and Kowalski have recently provided a framework using tensorial notation, in which instruments are classified according to the data that they produce (18). For example, a simple pH meter produces a single number, a zero-order tensor. The measurement of an infrared spectrum produces a vector of numbers, a first-order tensor, and the measurement of a fluorescence excitation-emission spectrum provides a matrix of numbers, a second-order tensor. The calibration methods used for each type of data are different.

Univariate calibration is used for zero-order tensors, but this assumes that there are no interferences in the response (the sensor is completely selective). Interferences may invalidate the results of a calibration study using a zero-order instrument, unbeknownst to the analyst. For a first-order instrument an interference not accounted for in the calibration step may invalidate the prediction results, but proper outlier detection can alert the analyst to the problem (19). For bilinear second-order instruments the calibration method of generalized rank annihilation (GRAM)

can account for uncalibrated interferences with a single calibration sample (20).

For this work, only first-order methods of multivariate calibration will be of interest. Four methods will be discussed, classical least squares (CLS), inverse least squares (ILS), principal component regression (PCR), and partial least squares (PLS). Other references for first-order multivariate calibration are Beebe and Kowalski (21), Martens and Naes (9,22), Haaland and Thomas (23), Mandel (24), Sanchez and Kowalski (25) and Sundberg and Brown (26).

### 2.2.2 Notation and Preprocessing Issues

For this dissertation lower case bold face letters will represent column vectors and upper case bold face letters will represent matrices. The transpose is denoted by an upper case, superscript "T". Normal type face, lower case letters will represent scalars. Lower case, normal type face letters with a single subscript,  $x_i$ , will denote single entries in the  $i^{\text{th}}$  position of a vector,  $\mathbf{x}$ . Those with a double subscript,  $x_{ij}$ , will denote a single entry in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of a matrix  $\mathbf{X}$ . Lower case, bold face letters with a single subscript,  $\mathbf{x}_i$ , will denote the  $i^{\text{th}}$  column vector in the matrix  $\mathbf{X}$ .

All data matrices used for multivariate calibration are considered to be mean centered, that is the mean of each column is subtracted from every entry in the column. For more information on preprocessing see Geladi and Kowalski (27,28). Mean centering is not used for the multiresponse nonlinear study and this is discussed in more detail in Section 4.2.2.

The calibration problem will be discussed in terms of prediction of chemical concentrations from multivariate spectral measurements. Other constituent values

associated with chemical samples may be predicted as has been shown in many application studies. References may be found in the latest chemometrics review (17).

### 2.2.3 Classical Least Squares (CLS)

Classical least squares (CLS) may be defined with the following model in vector notation

$$\mathbf{R} = \mathbf{CK} + \mathbf{E}_R \quad (2.1)$$

where  $\mathbf{R}$  is a response matrix, (M-objects (samples) by N-variables (responses)),  $\mathbf{C}$  is a concentration matrix, (M-objects by L-components),  $\mathbf{K}$  is a sensitivity matrix, (L-components by N-variables), and  $\mathbf{E}_R$  is a matrix of residuals, (M-objects by N-variables), containing systematic and random errors not fit by the model. Component, as used here, is understood to mean chemical component and should not be confused with principal component discussed in

Section 2.2.5.1. The classical least squares method is also known as the K matrix method (29), as well as direct (30), total (31), and reverse calibration method (9). For spectroscopic applications CLS is one of a collection of full spectrum methods.

Responses are modeled as a function of concentrations, and as for most analytical applications this model represents the causal structure (9), that is the change in concentration (the cause) effects a change in response.

Given the classical model, the calibration step involves the measurement of N-variables on M-objects of which all L-component concentrations are known.

Measuring  $\mathbf{R}$  and knowing  $\mathbf{C}$ , the least squares solution for  $\mathbf{K}$  is found as

$$\hat{\mathbf{K}} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{R} \quad (2.2)$$

where  $\hat{\mathbf{K}}$  represents the matrix of pure-component spectra at unit concentration and unit path length. This approach minimizes the response or spectral errors. Haaland and Thomas have described this as a factor based method in which the rows of  $\mathbf{K}$  are the factor loadings (loading vectors, see Section 2.2.5.1) and the columns of  $\mathbf{C}$  are the score vectors of the response matrix,  $\mathbf{R}$  (see Section 2.2.5.1) (23).

For prediction of an L-component concentration vector,  $\hat{\mathbf{c}}$ , from an unknown response vector,  $\mathbf{r}^T$  (N-responses by 1), the following equation is solved

$$\mathbf{r}^T = \mathbf{c}^T \hat{\mathbf{K}} \quad (2.3)$$

giving

$$\hat{\mathbf{c}} = (\hat{\mathbf{K}} \hat{\mathbf{K}}^T)^{-1} \hat{\mathbf{K}} \mathbf{r} \quad (2.4)$$

Numerically the inverse exists only if the rank of  $(\hat{\mathbf{K}} \hat{\mathbf{K}}^T)$  is L, thus there must be more sensors than components, ( $N > L$ ), and the rows of  $\hat{\mathbf{K}}$  must be linearly independent, the pure spectra must be somewhat different, in order to solve equation 2.4.

The appealing features of the classical approach include improvement in precision due to the full spectrum nature of the method (32), i.e. signal averaging, the simultaneous fitting of spectral baselines (32,33), and the ability to examine pure component spectra and full spectral residuals. The drawback for this approach is that



all of the chemical components with response variance must be known and calibrated for during the calibration step. This is understood by noting that if the response matrix contains variance from other sources not accounted for in the concentration matrix, the calculation of  $\hat{\mathbf{K}}$  in equation 2.2 would be affected by those extraneous variations. In other words, the pure component spectra would be incorrectly estimated, and prediction of future samples using equation 2.2 would be incorrect. In order to overcome the drawback of having to know and calibrate for all of the components in a sample, the method of inverse least squares may be used.

#### 2.2.4 Inverse Least Squares (ILS)

Inverse least squares (ILS) is defined with the following model

$$\mathbf{C} = \mathbf{RB} + \mathbf{E}_c \quad (2.5)$$

where  $\mathbf{R}$  is a response matrix, (M-objects (samples) by N-variables (responses)),  $\mathbf{C}$  is a concentration matrix, (M-objects by L-components),  $\mathbf{B}$  is a matrix of coefficients that relate the L component concentrations to the N observed variables, (N by L), and  $\mathbf{E}_c$  is a matrix of concentration residuals, (M by L), containing systematic and random errors not fit by the model. This method has been variously known as the P matrix method (29), multiple linear regression (MLR) (30), as well as indirect (30), partial (31), and forward calibration method (9). The inverse least squares approach models the concentrations as a function of the responses, and thus does not follow the causal structure (9).

One very important aspect of the inverse least squares model is that the

columns of the **C** matrix may be treated individually, that is, each chemical component is treated individually. Assuming the columns of **E<sub>c</sub>** are independent, then

$$\mathbf{c} = \mathbf{R}\mathbf{b} + \mathbf{e}_c \quad (2.6)$$

where **c** is a (M by 1) vector of concentration values of the analyte of interest, **b** is the (N by 1) vector of calibration coefficients (the regression vector), and **e<sub>c</sub>** is the (M by 1) vector of concentration residuals not fit by the model.

For the calibration step in the inverse least squares method, equation 2.6 is solved for  **$\hat{\mathbf{b}}$**  as follows

$$\hat{\mathbf{b}} = (\mathbf{R}^T\mathbf{R})^{-1} \mathbf{R}^T \mathbf{c} \quad (2.7)$$

or

$$\hat{\mathbf{b}} = \mathbf{R}^\dagger \mathbf{c} \quad (2.8)$$

where **R<sup>†</sup>** is the pseudo-inverse of **R** (4). For prediction of a concentration value from a measured response vector, **r<sup>T</sup>**, the following equation is solved

$$\hat{\mathbf{c}} = \mathbf{r}^T \hat{\mathbf{b}} \quad (2.9)$$

or

$$\hat{\mathbf{c}} = \mathbf{r}^T \mathbf{R}^\dagger \mathbf{c} \quad (2.10)$$

For the inverse least squares approach only one matrix inversion is required

for prediction as contrasted with the classical least squares approach where two matrix inversions are required, equations 2.2 and 2.4. In practice the advantage of the inverse least squares approach is that quantitative analysis may be performed even if the concentration of only one component is known in the calibration mixtures. The concentration variation not explicitly included in the calibration model must be implicitly modeled during the calibration step. If some interferences are not implicitly modeled in the calibration step, outlier detection methods should alert the analyst to the fact that the unknown samples are different from the calibration samples (19).

The drawback for the inverse least squares method comes in the calibration step and the calculation of the inverse of  $(\mathbf{R}^T\mathbf{R})$  in equation 2.7. This inverse does not exist if  $N > M$ , more sensors than samples, as is the case for many analytical applications. Thus a sensor selection or dimension reduction must take place, especially when using spectral data. In fact, if linear additivity is obeyed, the rank of  $(\mathbf{R}^T\mathbf{R})$  is at most  $L$ , the number of independent chemical components, and thus using more variables than  $L$  leads to problems of multicollinearity. Variable selection for dimension reduction is not a trivial problem and will not be discussed here (34,35).

An alternative to variable selection is to use linear combinations of the variables, or what has been referred to as data compression (9). The use of linear combinations of the original variables leads to the biased regression or latent variable regression methods of principal component regression (PCR) and partial least squares (PLS).

### 2.2.5 Biased Regression Methods

Latent variable regression methods make use of the inverse least squares model. For this dissertation, only the single analyte calibration problem will be considered, as presented in equation 2.6. The latent variable methods most often used in analytical chemistry are principal component regression (PCR) and partial least squares (PLS). The difference between these methods is in how the pseudo-inverse,  $\mathbf{R}^\dagger$ , used in equation 2.8 is computed. As will be shown below, the basic difference between the way each method computes  $\mathbf{R}^\dagger$  is in how much influence the concentration vector has on the data compression. The concentration vector has no influence in the PCR method, while PLS allows some influence. These methods have been characterized as discrete members of a continuum of methods by Lorber and co-workers (31). This approach provides an interesting framework from which to study the first-order calibration methods and will be briefly reviewed later in this section.

#### 2.2.5.1 Principal Component Regression (PCR)

For principal component regression (PCR) the first step is principal component analysis (PCA). PCA was first described by Pearson (36), developed further by Hotelling (37) and was discussed in detail by Jolliffe (2). Wold et al. have provided a tutorial for PCA (38).

For a given data matrix,  $\mathbf{R}$  (M by N), a PCA decomposition results in a representation of the data matrix as a sum of vector outer products

$$\mathbf{R} = \sum_{i=1}^H \mathbf{t}_i \mathbf{p}_i^T \quad \max(H) \leq \min(M, N) \quad (2.11)$$

or in matrix notation

$$\mathbf{R} = \mathbf{TP}^T + \mathbf{E} \quad (2.12)$$

The maximum number of latent variables that may be computed is the minimum of the number of rows or columns in  $\mathbf{R}$ . The  $\mathbf{t}_i$  ( $M$  by 1) vectors are referred to as score vectors and the  $\mathbf{p}_i$  ( $N$  by 1) vectors are referred to as loading vectors. Other names that are associated with these vectors are latent variables or factors. There are many methods to decide on the number of outer products to retain (the pseudorank) and these are discussed below. The  $\mathbf{t}_i$  vectors and  $\mathbf{p}_i$  vectors are orthogonal to one another, i.e.  $\mathbf{t}_i^T \mathbf{t}_j = 0$  and  $\mathbf{p}_i^T \mathbf{p}_j = 0$  for  $i \neq j$ . Each score and loading vector pair describes successively less variance of the response matrix than the previous pair. The orthogonal nature of the score vectors imparts desirable numerical properties for the regression analysis described below. Two methods available to compute the PCA decomposition are the nonlinear iterative partial least squares algorithm, (NIPALS) (39) and singular value decomposition algorithm (SVD) (4).

The NIPALS algorithm used for calculation of the score and loading vectors is as follows :

Step 1. Set  $i = 1$ .

Step 2. Choose an initial guess for the first score vector,  $\mathbf{t}_i$ . Wold et al. suggest using the row from the response matrix with the highest variance (40).

Step 3. Compute a least squares estimate for the first loading vector,  $\mathbf{p}_i$ .

$$\hat{\mathbf{p}}_i^T = \frac{\mathbf{t}_i^T \mathbf{R}}{\mathbf{t}_i^T \mathbf{t}_i} \quad (2.13)$$

The elements of the loading vector,  $\mathbf{p}_i$ , may be considered as the slopes of the regression of  $\mathbf{t}_i$  on the corresponding columns of the  $\mathbf{R}$  matrix.

Step 4. Normalize the loading vectors to length 1.

$$\hat{\mathbf{p}}_{i,\text{norm}} = f \hat{\mathbf{p}}_i \quad \text{where } f = \frac{1}{\sqrt{(\hat{\mathbf{p}}_i^T \hat{\mathbf{p}}_i)}} \quad (2.14)$$

Step 5. Compute a new score vector using a least squares estimate.

$$\hat{\mathbf{t}}_{i,\text{new}}^T = \frac{\mathbf{R} \hat{\mathbf{p}}_i}{\hat{\mathbf{p}}_i^T \hat{\mathbf{p}}_i} \quad (2.15)$$

Step 6. Check for convergence. Compare the new score vector with the old score vector by computing a total sum of squares of the differences between the consecutively computed score vectors. If there is convergence, then continue with step 7 otherwise return to step 2. If convergence is not reached in a large number of iterations (> 50), then the data have no preferred direction of maximum variance.

Step 7. Compute a residual matrix by subtracting the outer product formed from the score and loading vectors

$$\mathbf{E} = \mathbf{R} - \mathbf{t}_i \mathbf{p}_i^T \quad (2.16)$$

and substitute **E** for **R** in step 2, set  $i = i + 1$ , and extract the next vector pair. This method has been described as a variant of the power method for eigenvalue analysis (31).

The singular value decomposition (SVD) represents the matrix as a product of three matrices

$$\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (2.17)$$

where the columns of **U**, (M by M), are the left singular vectors (eigenvectors of  $\mathbf{R}\mathbf{R}^T$ ), orthogonal and of unit length, representing the column space (sample or object domain if the spectra are in the rows of **R**). The columns of **V**, (N by N), are the right singular vectors (eigenvectors of  $\mathbf{R}^T\mathbf{R}$ ), orthogonal and of unit length, representing the row space (spectral domain if the spectra are in the rows of **R**) and **S**, (M by N), is a diagonal matrix of singular values (square root of eigenvalues of  $\mathbf{R}\mathbf{R}^T$  or  $\mathbf{R}^T\mathbf{R}$ ). The details of the implementation of the algorithm may be found in Golub and Van Loan (4). Comparing equation 2.17 with equation 2.12 shows **P** is equal to **V**, and **U** contains the same column vectors as **T**, but normalized to unit length. The length of each  $t_i$  is contained in the diagonal entries of the **S** matrix.

The SVD provides for more numerical stability than NIPALS, but the SVD requires computation of all of the principal components. The NIPALS algorithm allows less than the maximum number of principal components to be computed and may also be used on data with missing values (41).

After decomposition of the response matrix the number of latent variables or factors to retain must be chosen. The methods used to perform this selection are

discussed in Section 2.2.6.

After determination of the proper rank the principal component regression model is formulated as

$$\mathbf{c} = \mathbf{USV}^T\mathbf{b} + \mathbf{e}_c \quad (2.18)$$

where  $\mathbf{USV}^T$  has replaced  $\mathbf{R}$  with the understanding that an appropriate truncation of the data matrix has taken place. Note that the same model can be formulated with  $\mathbf{US}$  replaced by  $\mathbf{T}$ , and  $\mathbf{V}^T$  replaced by  $\mathbf{P}^T$ . The following definition is used in the development of the regression model by Mandel (24) in order to provide an analysis of variance of the regression coefficients,

$$\boldsymbol{\alpha} = \mathbf{SV}^T\mathbf{b} \quad (2.19)$$

Equation 2.18 may be rewritten as

$$\mathbf{c} = \mathbf{U}\boldsymbol{\alpha} + \mathbf{e}_c \quad (2.20)$$

which displays PCR as the inverse least squares problem with the normalized score vectors being used as artificial variables, linear combinations of the original variables (34). The  $\boldsymbol{\alpha}$  vector is the regression vector, (H by 1), relating the H variables (principal components) to the concentration vector. This vector,  $\boldsymbol{\alpha}$ , has been denoted as the concentration loadings by Martens and Naes (9). The least squares solution to equation 2.20 is



$$\hat{\alpha} = (U^T U)^{-1} U^T c \quad (2.21)$$

which, as a result of the properties of the U matrix, becomes

$$\hat{\alpha} = U^T c \quad (2.22)$$

Mandel has shown how this representation is useful for the discussion of the variance properties of  $\hat{\alpha}$  and  $\hat{b}$  (24).

The prediction of a concentration value,  $\hat{c}$ , from an unknown response vector,  $r^T$ , is as follows. First the unknown response vector must be centered or scaled, if the calibration data were centered or scaled, then the regression vector,  $b$ , is multiplied by the unknown response vector

$$\hat{c} = r^T \hat{b} \quad (2.23)$$

or from the solution of equation 2.19 for  $\hat{b}$

$$\hat{c} = r^T V S^{-1} \hat{\alpha} \quad (2.24)$$

The coordinates for the unknown response in the calibration space (the scores) are computed,  $u^T = r^T V S^{-1}$ , and then multiplied by the regression vector,  $\hat{\alpha}$ , to compute the concentration value. Note that the calculation of the spectral regression vector,  $\hat{b} = V S^{-1} \hat{\alpha}$ , involves the inverse of the S matrix and thus PCR gains stability by elimination of eigenvalues close to zero, i.e. data compression. Mandel discusses the "trading" of variance for bias by elimination of small eigenvalues and associated

eigenvectors in PCR (24).

The pseudo-inverse of the response matrix,  $\mathbf{R}^\dagger$ , as used in equation 2.10 is equal to  $\mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T$ , as shown by combining equations 2.22 and 2.24, and an alternative representation of the prediction equation for PCR consistent with equation 2.10 is

$$\hat{\mathbf{c}} = \mathbf{r}^T \mathbf{V} \mathbf{S}^{-1} \mathbf{U}^T \mathbf{c} \quad (2.25)$$

PCR uses no information from the concentration vector to influence the calculation of  $\mathbf{R}^\dagger$  and assumes latent variables associated with small variance or small singular values contain no useful information for prediction of  $\mathbf{c}$ , only noise. This assumption may not always be correct, because small variance components can have relevant information for prediction. These considerations have lead to the method of partial least squares (PLS) (39).

#### 2.2.5.2 Partial Least Squares (PLS)

The method of partial least squares (PLS) has been used for the prediction of chemical and physical properties from many types of multivariate observations, for many different sample types. References for such experiments may be found in the fundamental reviews (15,16,17). The PLS method has been thoroughly studied by Manne (42), Hoskuldsson (43), Lorber et al. (31), Helland (44), and summarized very well by Martens and Naes (9). The basic idea for PLS is that the decomposition of the data matrix should be guided to a solution that is best for prediction of the constituent value of interest. In order to do this, the PLS decomposition is conducted in such a way that the constituent value information is used to calculate the

decomposition.

For this discussion only PLS1 is considered, that is, only one constituent value at a time is used in the decomposition. The use of PLS2, PLS with two or more constituent values used in the decomposition simultaneously, may have some advantage if the precision of the constituent values differs widely. More precise constituent values may offer a stabilizing effect for the prediction of the less precise constituent values (9). The algorithm for PLS1 is discussed below, and is very similar to the PCA calculation shown in Section 2.2.5.1. The discussion draws largely from the work of Haaland and Thomas (23).

Step 1. Compute the weight loading vector for factor H,  $\hat{\mathbf{w}}_H$ , and normalize

$$\hat{\mathbf{w}}_H = \frac{\mathbf{R}^T \mathbf{c}}{\mathbf{c}^T \mathbf{c}} \quad (2.26)$$

$$\hat{\mathbf{w}}_H = \frac{\hat{\mathbf{w}}_H}{\sqrt{\hat{\mathbf{w}}_H^T \hat{\mathbf{w}}_H}} \quad (2.27)$$

The entries of the weight loading vector are the slopes of the simple regression of the constituent vector with the corresponding columns of the  $\mathbf{R}$  matrix, just as were found in the NIPALS calculation of the PCA decomposition.

Step 2. Formation of the score vector,  $\hat{\mathbf{t}}_H$

$$\hat{\mathbf{t}}_H = \mathbf{R} \hat{\mathbf{w}}_H \quad (2.28)$$

This is similar to the prediction of a constituent value in CLS, if  $\hat{\mathbf{w}}_H$  were considered a pure component spectrum for a single component system.

Step 3. Relate the score vector, to the concentrations

$$\hat{v}_H = \frac{\hat{\mathbf{t}}_H^T \mathbf{c}}{\hat{\mathbf{t}}_H^T \hat{\mathbf{t}}_H} \quad (2.29)$$

The scalar  $v$  is the regression coefficient relating  $\mathbf{t}$  and  $\mathbf{c}$ , sometimes referred to as the chemical loadings (9).

Step 4. Compute the loading vector,  $\mathbf{p}$

$$\hat{\mathbf{p}}_H = \frac{\mathbf{R}^T \hat{\mathbf{t}}_H}{\hat{\mathbf{t}}_H^T \hat{\mathbf{t}}_H} \quad (2.30)$$

Step 5. Calculate the residuals in  $\mathbf{R}$  and  $\mathbf{c}$

$$\mathbf{E}_R = \mathbf{R} - \hat{\mathbf{t}}_H^T \hat{\mathbf{p}}_H \quad (2.31)$$

$$\mathbf{e}_c = \mathbf{c} - v_H \hat{\mathbf{t}}_H^T \quad (2.32)$$

Step 6. The  $\mathbf{E}_R$  and  $\mathbf{e}_c$  vectors are used in step 1 as  $\mathbf{R}$  and  $\mathbf{C}$  and the next vector is extracted. Note that this algorithm is NOT iterative for each factor calculation, as was found for the PCA calculation.

Prediction using the PLS1 algorithm is somewhat more complex than for PCR

and this is summarized below. Given an unknown response vector,  $\mathbf{r}_{\text{un}}$

Step 1. Center and/or scale  $\mathbf{r}_{\text{un}}$  if the calibration data were centered and/or scaled.

Step 2. Compute the scalar score value for the unknown on the  $H^{\text{th}}$  factor

$$\hat{t}_H = \hat{\mathbf{w}}_H^T \mathbf{r}_{\text{un}} \quad (2.33)$$

Step 3. Compute the concentration contribution due to the  $H^{\text{th}}$  factor from the unknown score value and the chemical loading

$$\mathbf{c}_H = \mathbf{c}_{H-1} + \hat{\mathbf{v}}_H \hat{t}_H \quad (2.34)$$

Step 4. Compute the response residual

$$\mathbf{e}_H = \mathbf{e}_{H-1} + \hat{\mathbf{p}}_H \hat{t}_H \quad (2.35)$$

Step 5. Increment  $H$ , substitute  $\mathbf{e}_H$  for  $\mathbf{r}_{\text{un}}$ , and repeat with step 2 until the chosen rank is reached.

For this implementation of the PLS1 decomposition, the score vectors are orthogonal to one another as are the weight vectors, but the loading vectors are not. There is an alternative implementation of the PLS1 algorithm, given by Martens, in which neither the score vectors nor the loading vectors are orthogonal (19). A unifying and perhaps simpler approach to the latent variable regression models has

been given by Lorber et al. and is briefly reviewed here (31).

The following steps are performed for the Lorber implementation of PLS1:

Step 1. Preprocess the data matrix by centering and/or scaling.

Step 2. Compute two vectors  $\mathbf{t}_H$  and  $\mathbf{p}_H$

$$\mathbf{p}_{H+1} = \frac{\mathbf{R}_H^T \mathbf{c}}{\|\mathbf{R}_H^T \mathbf{c}\|} \quad (2.36)$$

$$\mathbf{t}_{H+1} = \frac{\mathbf{R}_H \mathbf{p}_{H+1}}{\|\mathbf{R}_H \mathbf{p}_{H+1}\|} = \frac{\mathbf{R}_H (\mathbf{R}_H^T \mathbf{c})}{\|\mathbf{R}_H (\mathbf{R}_H^T \mathbf{c})\|} \quad (2.37)$$

Step 3. Construct a residual matrix,  $\mathbf{R}_{H+1}$ , orthogonal to the two vectors

$$\mathbf{R}_{H+1} = (\mathbf{I}_M - \mathbf{t}_{H+1} \mathbf{t}_{H+1}^T) \mathbf{R}_H (\mathbf{I}_N - \mathbf{p}_{H+1} \mathbf{p}_{H+1}^T) \quad (2.38)$$

Step 4. If more latent variables are needed increment H and go to step 2.

At this point  $\mathbf{T}$  and  $\mathbf{P}$  are column orthonormal matrices representing the column and row space of  $\mathbf{R}$ , respectively. The PLS approximation to the data after appropriate truncation is

$$\hat{\mathbf{R}} = \mathbf{T} \mathbf{Q} \mathbf{P}^T \quad (2.39)$$

Because  $\mathbf{T}$  and  $\mathbf{P}$  are column orthonormal,  $\mathbf{Q}$  may be computed as

$$\mathbf{Q} = \mathbf{T}^T \hat{\mathbf{R}} \mathbf{P} \quad (2.40)$$

and is bidiagonal for PLS1 and diagonal for PCA.

The pseudo-inverse is required for prediction, as shown in equation 2.10, and this may be found by computing the SVD of  $Q$

$$Q = U_q S_q V_q^T \quad (2.41)$$

Then

$$R^\dagger = PQ^{-1}T^T = PV_q^T S_q^{-1} U_q T^T \quad (2.42)$$

An interesting observation is that as the score vector,  $t_{H+1}$ , is recycled through the algorithm, the PLS solution approaches a PCA solution. Mathematically

$$\sigma_1^2 t_1^{(n)} = (RR^T) t_1^{(n-1)} \quad (2.43)$$

so that  $t_1^{(n)}$ , the first score vector on its  $n^{\text{th}}$  iteration, approaches an eigenvector of  $RR^T$  and  $\sigma_1^2$  approaches the associated eigenvalue. What this means for intermediate vectors, i.e. vectors found before convergence, is that these vectors are determined by the power to which  $(RR^T)$  is raised. By computing the SVD of  $(RR^T)^n$  it may be shown that

$$(RR^T)^n = US^{2n}U^T \quad (2.44)$$

This suggests a modification to the PLS1 algorithm. At each factor extraction step,  $R_H$  is replaced by the SVD of  $R_H$  so

$$\mathbf{p}_{H+1} = \frac{\mathbf{V}\mathbf{S}\mathbf{U}^T\mathbf{c}}{\|\mathbf{V}\mathbf{S}\mathbf{U}^T\mathbf{c}\|} \quad (2.45)$$

$$\mathbf{t}_{H+1} = \frac{\mathbf{U}\mathbf{S}^2\mathbf{V}^T\mathbf{p}_{H+1}}{\|\mathbf{U}\mathbf{S}^2\mathbf{V}^T\mathbf{p}_{H+1}\|} = \frac{\mathbf{U}\mathbf{S}^2\mathbf{U}^T\mathbf{c}}{\|\mathbf{U}\mathbf{S}^2\mathbf{U}^T\mathbf{c}\|} \quad (2.46)$$

and the **S** matrix is raised to the power  $n$  within a given factor extraction

$$\mathbf{p}_{H+1} = \frac{\mathbf{V}\mathbf{S}^n\mathbf{U}^T\mathbf{c}}{\|\mathbf{V}\mathbf{S}^n\mathbf{U}^T\mathbf{c}\|} \quad (2.47)$$

$$\mathbf{t}_{H+1} = \frac{\mathbf{U}\mathbf{S}^{2n}\mathbf{V}^T\mathbf{p}_{H+1}}{\|\mathbf{U}\mathbf{S}^{2n}\mathbf{V}^T\mathbf{p}_{H+1}\|} = \frac{\mathbf{U}\mathbf{S}^{2n}\mathbf{U}^T\mathbf{c}}{\|\mathbf{U}\mathbf{S}^{2n}\mathbf{U}^T\mathbf{c}\|} \quad (2.48)$$

Now the **T** and **P** matrices may be computed for different values of  $n$  and predictions are performed by calculation of the pseudo-inverse as shown in equation 2.42.

The latent variable regression methods can be placed on a continuum, with PCR at  $n=\infty$ , PLS at  $n=1$ , and the "MLR" solution, if it exists (i.e.  $N>M$ ), is found at  $n=0$ .

### 2.2.6 Rank Determination Methods

After the matrix decomposition, the data are truncated by selecting only those latent variables that are "significant". There are many methods for selecting "significant" latent variables (6, 23, 41,45-49). These methods are based either on a fitting criterion, such as variance explained or standard error of estimate, or on a predictive criterion. The more relevant methods for rank determination for



quantitative analysis in analytical chemistry are those methods that use a prediction error as the criterion for latent variable selection since the goal of calibration is to predict future samples and not simply to fit the calibration data.

Using a predictive criterion presents the problem of how to assess the predictive behavior of the model. The best way to accomplish this is with independent, calibration and prediction data sets. Several models are constructed from the calibration data set using different numbers of latent variables. The independent prediction sample set is then treated as an unknown and predictions are made using the calibration models from the calibration data set. The predictive residual error sum of squares, PRESS, is the sum of the squared differences of the predicted concentrations,  $\hat{c}_i$ , from the known concentrations,  $c$ ,

$$\text{PRESS} = \sum_{i=1}^M (\hat{c}_i - c_i)^2 \quad (2.49)$$

PRESS is computed for each model tested. The PRESS values are plotted against the number of latent variables. Figure 2.1 displays an ideal PRESS plot shown as a dotted line.

As relevant latent variables are added to the model, the prediction error decreases until irrelevant information is drawn into the model, at which time the prediction error increases. A simple choice for the number of factors to select is at the minimum in the PRESS plot. Other methods to determine the number of factors, such as an F-test for the significance of added latent variables, may give a number different from the minimum (49). Another method is to compute the error in PRESS, as defined by Breiman et al. (50), using the following equation

$$\frac{\left[ \frac{\sum_{i=1}^M (\hat{c}_i - c_i)^4}{M} - \left[ \frac{\sum_{i=1}^M (\hat{c}_i - c_i)^2}{M} \right]^2 \right]^{1/2}}{M - \# \text{ of factors} - 1} * M \quad (2.50)$$

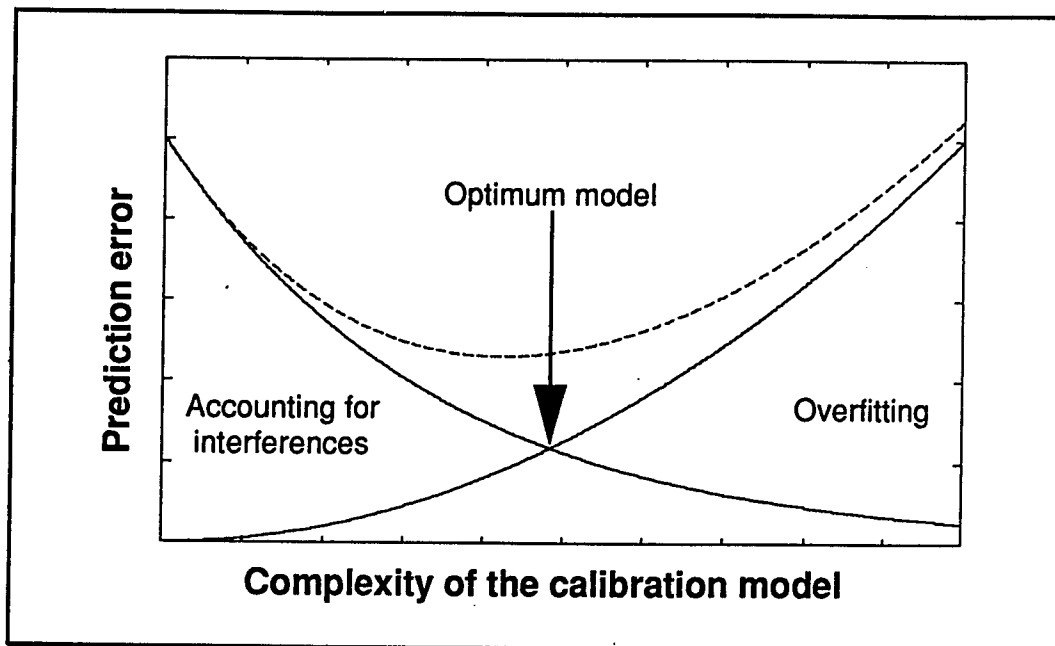
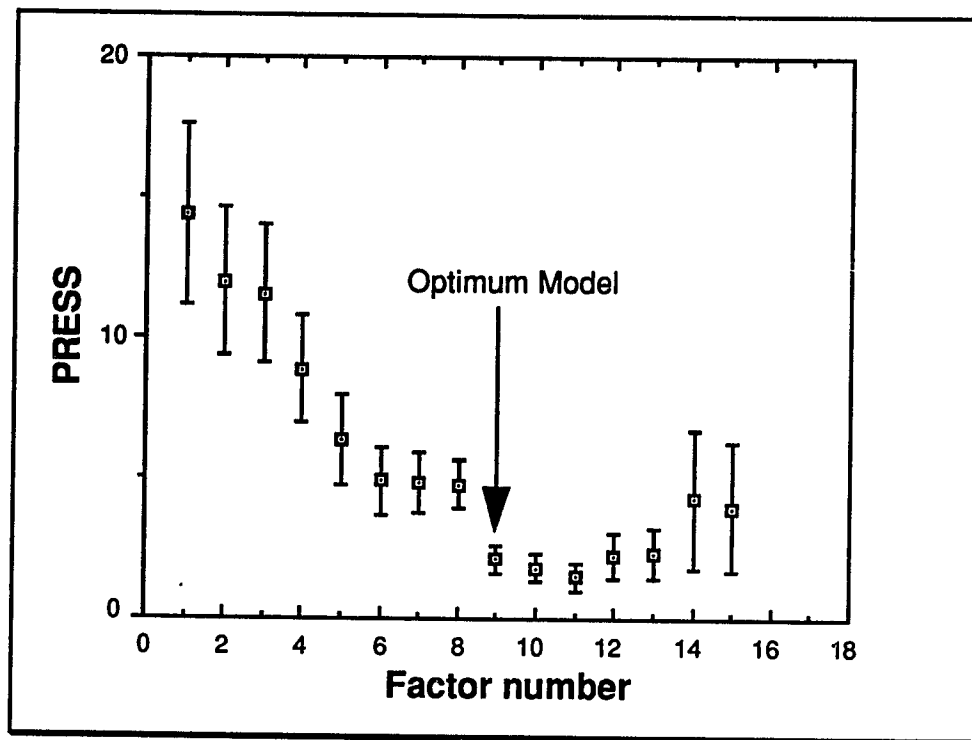


Figure 2.1 Ideal PRESS plot.

The model that gives the fewest number of factors which are within the error band overlap of the minimum in the PRESS plot is chosen as optimal. A rank of 9 is chosen for the example shown in Figure 2.2.

If a separate calibration and prediction data set are not available, then the method of cross-validation may be used (41, 45, 46). For cross-validation only a calibration data set is needed and from this data set each sample is removed, one at a time, and a separate calibration model constructed without that one sample.



**Figure 2.2** PRESS vs. number of latent variables added.

The excluded sample is predicted with the calibration model and this is repeated for all samples and for all the different models considered (all the different numbers of factors). The same PRESS plot is constructed and used to determine the pseudorank of the data matrix. A statistic related to PRESS is the standard error of prediction, SEP, and is computed as

$$SEP = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{c}_i - c_i)^2} \quad (2.51)$$

The SEP is a measure of how well the model is expected to perform for prediction of

unknowns from the same population of samples as the calibration set.

Another method for rank determination in PCA, where there are no concentration values to predict, is to explain the data matrix to the experimental noise level. This method requires an estimate of the experimental noise level which is compared with the residual computed after each latent variable is extracted. Chapter 4 uses this method to determine the rank of spectral data collected from a kinetics experiment and is discussed more fully there.

### **2.2.7 Locally Weighted Regression (LWR)**

Bilinear models such as PLS and PCR can handle nonlinear calibration data to some extent (9), but for large nonlinear effects an alternative method may be more appropriate. The method of locally weighted regression (LWR) is relatively new in the statistical literature and may offer certain advantages for multivariate calibration when the data are nonlinear (51). Naes et al. have recently shown that it offers improved prediction results for near infrared data, as well as reduced model complexity, and thus improved interpretability of the calibration model (52). Because of the complex models found using PLS and PCR for calibration with infrared emission data (see below), locally weighted regression was explored as an alternative. The theory for its implementation is as follows.

In the most general sense locally weighted regression is a procedure for relating an independent set of  $M$  measurements on  $N$  variables to a set of  $L$  dependent variables. The procedure forms local linear or nonlinear models, with the variables weighted in some fashion to estimate a regression surface. For the multivariate calibration problem the dependent variable may be considered as the concentration of

chemical components and the independent variable may be spectral measurements collected from each of the samples. The model used is the inverse least squares model

$$\mathbf{c} = \mathbf{R}\mathbf{b} + \mathbf{e}_c \quad (2.52)$$

where  $\mathbf{c}$  is (M by 1),  $\mathbf{R}$  is (M by N), and  $\mathbf{b}$  is (N by 1).

The implementation of the locally weighted regression procedure in this work will make use of the principal component scores as artificial variables to represent the response matrix giving

$$\mathbf{c} = \mathbf{U}\boldsymbol{\alpha} \quad (2.53)$$

where  $\mathbf{U}$  (M by L) is the orthonormal matrix from the singular value decomposition of the response matrix using L latent variables or L combinations of latent variables and  $\boldsymbol{\alpha}$  is the (L by 1) regression vector. Equation 2.53 is left multiplied by a diagonal weight matrix,  $\mathbf{W}$  (M by M), giving

$$\mathbf{W}\mathbf{c} = \mathbf{W}\mathbf{U}\boldsymbol{\alpha} \quad (2.54)$$

The regression vector,  $\hat{\boldsymbol{\alpha}}$ , is found with the usual least squares estimate as

$$\hat{\boldsymbol{\alpha}} = (\mathbf{U}^T \mathbf{W}^2 \mathbf{U})^{-1} \mathbf{U}^T \mathbf{W}^2 \mathbf{c} \quad (2.55)$$

Four important considerations in LWR are, 1) the initial number of principal

components to use, 2) the model, 3) the size of the local region, and 4) the weight function.

The number of principal components to be used is related to the model choice and may be investigated with cross-validation. Model choice is very flexible in this procedure in that a linear model making use of just the principal components may be used or a polynomial combination of the principal components may be used.

The choice of the local region to use requires a distance measure. For this work a simple Euclidian distance is used, computed with the principal component score vectors. The local region is considered in terms of  $f$ , the fraction of points (samples) used in the regression calculation. The fraction of points used must be considered relative to the number of independent variables,  $L$ . In this implementation  $L$  is the number of principal components plus the number of combinations of principal components if a polynomial model is considered. If the number of independent variables (and thus the number of parameters that must be estimated) is not allowed to become a large fraction of the number of data points, then the local regression analysis can be expected to behave reliably. Cleveland and Devlin refer to the increase in the number of independent variables for a fixed number of data points as the "curse of dimensionality" (51). Cross-validation is used to select  $f$  values in the calibration studies discussed below.

The issue of the weight function is also an important consideration for LWR. For this work the original function suggested by Cleveland and Devlin is used. This function, known as the tricubic function, calculates the weight for the  $j^{\text{th}}$  sample for a calibration about the  $i^{\text{th}}$  sample as

$$W(z) = (1-z^3)^3 \quad (2.56)$$

where

$$z = \frac{p(u_i, u_j)}{d(u_i)} \quad (2.57)$$

and  $p(u_i, u_j)$  is the distance of the  $j^{\text{th}}$  sample from the  $i^{\text{th}}$  sample and  $d(u_i)$  is the maximum of  $p(u_i, u_j)$  for all of the samples considered in the local region. Therefore samples close to  $u_i$  are heavily weighted and those far away have lower weights as shown in Figure 2.3.

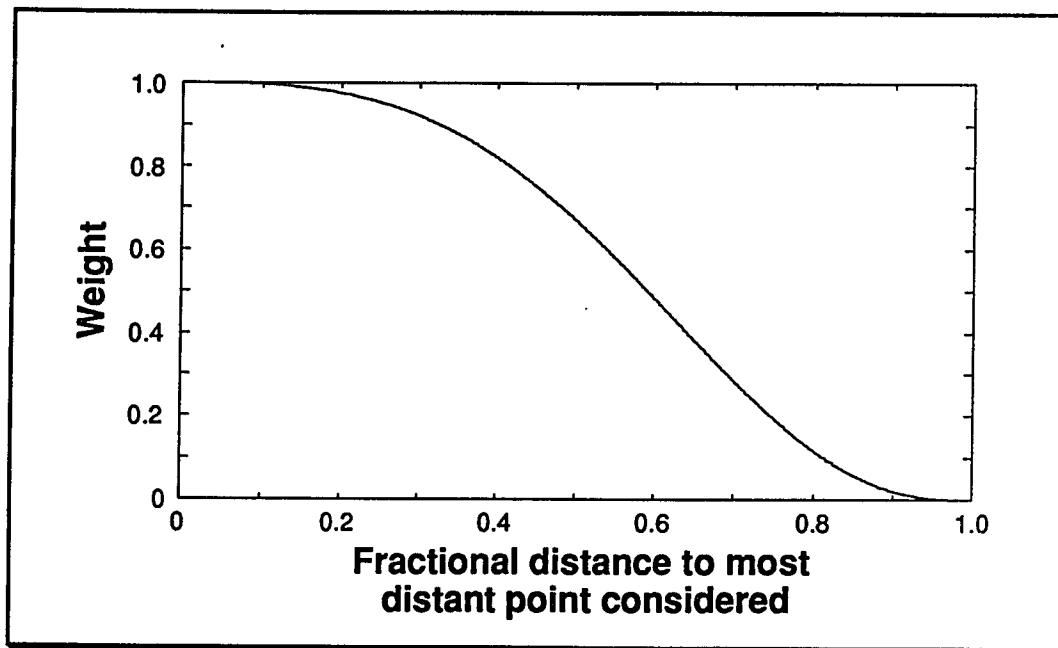


Figure 2.3 Tricubic weight function used for locally weighted regression.

In this implementation of LWR the weight function is directly connected to the distance measure, in that how far away a calibration sample is from an unknown dictates its weight. This may not be optimal for prediction. An alternative would be

to weight the samples based on their predictive ability. Cleveland and Devlin comment that asymptotic results for nonparametric regression show that the overall form of the weight function does not have an appreciable effect on the mean squared error. But they do state this does not imply the form of the weight function does not matter (51). It is beyond the scope of this work to explore the weight function issue, but it is suggested that a reiterative weighting scheme based on prediction ability may be an interesting area of research.

For prediction using LWR and principal components as artificial variables, a new calibration model is required for each unknown. The unknown response is projected onto the calibration space defined by the loading vectors (i.e. the scores for the unknown in the calibration space are computed) and the distance from the unknown to all the calibration samples is computed. The fraction of the calibration samples that are being considered is then weighted according to the weight function and the regression vector computed using equation 2.55. Prediction is performed using the computed regression vector and the scores from the unknown sample.

## **2.3 Multiresponse Nonlinear Optimization**

### **2.3.1 Overview**

Multiresponse nonlinear optimization is the use of multiresponse data to estimate model parameters via a nonlinear optimization technique. This differs from the multivariate calibration problem in that a "hard" parametric model is postulated to describe the observed data, and the parameters of the model are of interest. There is a further distinction for multiresponse nonlinear optimization in terms of the nature of



the data. The data may be from direct observations of the phenomenon of interest, for example the concentrations of the three chemical species in a first-order consecutive reaction, or the data may be from observations made in an indirect fashion, for example the measurement of infrared spectra of a reaction mixture following a first-order consecutive reaction scheme.

The details of the theory of multiresponse nonlinear optimization for direct observation and indirect observation experiments are discussed below. Applications of the methodology from the literature are discussed in 2.3.3. Chapter 5 details the proper choice of objective function for the nonlinear optimization.

### **2.3.2 Theory for Multiresponse Nonlinear Optimization**

For the development of the theory a distinction is drawn between the direct (full rank) and indirect (rank deficient) observation of the phenomenon of interest. An example of a direct observation experiment is the case of actual chemical concentrations measured as a function of time for a kinetics experiment. The response matrix,  $\mathbf{R}$ , ( $N$ -times by  $M$ -chemical components) is assumed to have  $N \geq M$  and be full rank, i.e.  $\text{rank} = M$ . Infrared spectra collected from the same reaction mixture as a function of time would be an example of an indirect observation experiment. The response matrix,  $\mathbf{R}$ , ( $N$ -times by  $L$ -wavenumbers) for the indirect observation experiments would most likely be rank deficient, i.e.  $\text{rank} < N$  and  $\text{rank} < L$ .

### 2.3.2.1 Multivariate Nonlinear Modeling Using Direct Observations

The development of the theory for multiresponse nonlinear modeling in the direct observation case is straightforward. Let  $\mathbf{R}$  be a response matrix, ( $N$ -objects by  $M$ -responses), with  $N > M$  and full rank, i.e.  $\text{rank} = M$ . For example,  $\mathbf{R}$  for a kinetics experiment would consist of the concentrations of the  $M$ -chemical components in the columns measured at  $N$ -times after mixing. A postulated parametric model is used to generate estimates of the observed data,  $\mathbf{F}$ . The differences between estimates from the model and the observed data are collected in a residual matrix,  $\mathbf{Z}$ . At each step in the parameter search a residual matrix is generated and it is the "size" of this matrix that is used to describe the response surface and guide the optimization algorithm.

The question of how to assess the "size" of the residual matrix and the choice of objective function are the same. The most commonly used objective function is the total sum of squares (TSS) given as the sum of the squared entries of the residual matrix. As pointed out by Hunter, the TSS objective function is valid only if the response errors have equal variance and are uncorrelated (53). A weighted total sum of squares objective function may be used, provided the weights are known or iteratively estimated, in order to account for the non-constant variance, but the correlation of errors may still be a problem (54).

Box and Draper have shown, using a Bayesian approach, that the determinant of  $\mathbf{Z}^T\mathbf{Z}$  is a more generally applicable objective function (55). This criterion, known as the determinant criterion, allows for non-constant response variance, as well as correlation in the response errors. When compared with the situation of a known variance-covariance matrix, it may be shown that the determinant criterion provides entries in the variance-covariance matrix that are proportional to the maximum-

likelihood estimates (55). A convenient way of considering the determinant criterion and total sum of squares is that the determinant of  $\mathbf{Z}^T\mathbf{Z}$  is the product of the eigenvalues of  $\mathbf{Z}^T\mathbf{Z}$ , while the total sum of squares is the sum of the eigenvalues of  $\mathbf{Z}^T\mathbf{Z}$ . The determinant of  $\mathbf{Z}^T\mathbf{Z}$  corresponds, geometrically, to the square of the volume of the M-dimensional parallelepiped spanned by the residual vectors. Minimizing the determinant corresponds to minimizing the volume enclosed by the residual vectors (56).

In order to successfully employ the determinant criterion there are dimensional restrictions on the residual matrix,  $\mathbf{Z}$ , and thus on the response matrix. For a residual matrix,  $\mathbf{Z}$  (N-objects by M-responses), the number of objects, N, must be greater than the number of responses, M, otherwise  $\mathbf{Z}^T\mathbf{Z}$ , which is (M by M) would have a rank of, at most N, and the determinant would be zero for all parameter values. This is understood by realizing that for a matrix,  $\mathbf{Z}^T\mathbf{Z}$ , of dimension (M by M) and of rank N, with  $M > N$ , there must be at least one zero eigenvalue and thus the product of the eigenvalues would be zero for all parameters. Another consideration is that the number of parameters, P, must be less than N, the number of objects, otherwise the criterion can be made zero by fitting any one response perfectly or even by fitting a linear combination of the responses perfectly.

### **2.3.2.2 Multivariate Nonlinear Modeling Using Indirect Observations**

For the indirect observation case the response matrix is usually underdetermined, in that there are more responses than objects, and is always rank deficient. In order to meet the dimensional constraints for the determinant criterion a data compression step is necessary. A data compression step will also offer a certain

amount of noise reduction.

Consider a measured matrix of spectral intensities,  $\mathbf{R}$  (N by L), as N spectra of L wavenumbers collected at N times during the course of a chemical reaction. Assuming the contribution of each component to the spectra is linear and independent of all others, the response matrix from the above experiment may be written in matrix notation as

$$\mathbf{R} = \mathbf{D}\mathbf{F}^T \quad (2.58)$$

Equation 2.58 states that the data matrix,  $\mathbf{R}$ , may be represented as the product of a pure concentration profiles matrix,  $\mathbf{D}$  (N by C), where C is the number of spectrally distinguishable chemical components in the reaction mixture, and a matrix,  $\mathbf{F}$  (L by C), representing the pure spectra of the C chemical components. A model may then be proposed to describe either the spectral or concentration profiles. For example, a specific kinetic model may be assumed in order to describe the pure concentration profiles contained in the  $\mathbf{D}$  matrix. That is, the concentrations in  $\mathbf{D}$  are not random, but follow the kinetic equations as a function of time. The optimal kinetic parameters may then be found for this model by minimizing the residuals,  $\mathbf{Z}$ , of the fit to the data given by

$$\mathbf{Z} = \mathbf{R} - \hat{\mathbf{D}}\hat{\mathbf{F}}^T = \mathbf{R} - \hat{\mathbf{D}}\hat{\mathbf{D}}^\dagger\mathbf{R} \quad (2.59)$$

where  $\hat{\mathbf{F}}^T = \hat{\mathbf{D}}^\dagger\mathbf{R}$ , and  $\hat{\mathbf{D}}^\dagger$  is the pseudo-inverse of  $\hat{\mathbf{D}}$ , the estimated kinetic model.

Shrager and Hendler describe this approach and point out the advantage of a data compression step involving the singular value decomposition as described below (57).

Note here that the determinant of  $\mathbf{Z}^T\mathbf{Z}$  would be zero for any parameter estimate because  $\mathbf{Z}^T\mathbf{Z}$  is  $M$  by  $M$  and of rank at most  $N$ , where  $N$  is less than  $M$ , thus the need for a data compression step in order to meet the dimensional constraints of the determinant criterion.

The response matrix,  $\mathbf{R}$  ( $N$  by  $L$ ), may be represented by a set of orthogonal basis vectors using the singular value decomposition as

$$\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (2.60)$$

where  $\mathbf{U}$  is ( $N$  by  $N$ ),  $\mathbf{S}$  is ( $N$  by  $L$ ) and  $\mathbf{V}$  is ( $L$  by  $L$ ) with the properties discussed in Section 2.2.5.1.

Data compression is achieved by selecting a reduced number of singular vectors to represent the data matrix. As discussed in Section 2.2.6, there are many methods to choose the appropriate number of singular vectors to retain (23,41,45-49). Common to most of the methods is that the singular vectors associated with small singular values are more likely due to noise than to chemical information and should be discarded. In Chapter 4 another method for rank determination using the known experimental error is described.

The compressed data matrix is given by

$$\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (2.61)$$

where it is understood that singular vector selection has taken place. Thus  $\mathbf{R}$  is still ( $N$  by  $L$ ), but  $\mathbf{U}$  is now ( $N$  by  $C$ ),  $\mathbf{S}$  is ( $C$  by  $C$ ) and  $\mathbf{V}$  is ( $L$  by  $C$ ), where  $C$  is the chosen rank of the data matrix. This data compression step offers a reduced noise

data matrix, assuming that small eigenvalues are associated more with noise than chemical information.

It is recognized that the columns of  $\mathbf{U}$  form an abstract linear combination of the time profiles and that this may be represented mathematically as

$$\mathbf{U} = \mathbf{D}\mathbf{H}^T \quad (2.62)$$

where the columns of  $\mathbf{D}$  contain, for example, the concentration profiles for the three species involved in a consecutive first-order reaction and  $\mathbf{H}^T$  is a rotation matrix. In order to generate an estimate of the concentration profiles,  $\hat{\mathbf{D}}$ , the kinetic parameters must be specified. Given an estimate for  $\mathbf{D}$ ,  $\hat{\mathbf{D}}$ , an estimate for  $\mathbf{H}$ ,  $\hat{\mathbf{H}}$ , may be calculated as

$$\hat{\mathbf{H}}^T = \hat{\mathbf{D}}^\dagger \mathbf{U} \quad (2.63)$$

where  $\hat{\mathbf{D}}^\dagger$  is the pseudo-inverse of  $\hat{\mathbf{D}}$ . The  $\mathbf{U}$  matrix may then be estimated as  $\hat{\mathbf{U}}$  by

$$\hat{\mathbf{U}} = \hat{\mathbf{D}}\hat{\mathbf{H}}^T \quad (2.64)$$

The residual matrix,  $\mathbf{Z}$ , for this process is given by

$$\mathbf{Z} = \mathbf{U} - \hat{\mathbf{U}} = \mathbf{U} - \hat{\mathbf{D}}\hat{\mathbf{D}}^\dagger \mathbf{U} \quad (2.65)$$

Lawton and Sylvestre pointed out the utility of performing a linear least squares calculation for the parameters of the  $\mathbf{H}$  matrix, equation 2.63, rather than

including them as nonlinear parameters, thus reducing the number of parameters to be estimated by the nonlinear search (58). The parameters used to generate  $\mathbf{D}$  are searched for by minimizing the residuals in a systematic fashion, by a modified SIMPLEX (59) or Gauss-Newton algorithm which has recently been implemented for linear systems of differential equations (56). It is at this point in the analysis that a choice of objective function must be made and this is discussed in Chapter 5 of this dissertation.

After  $\mathbf{D}$  has been estimated the corresponding space,  $\mathbf{F}$ , may be computed in two ways. From the original data,  $\mathbf{R}$ ,  $\mathbf{F}$  may be found as

$$\mathbf{F}^T = \hat{\mathbf{D}}^\dagger \mathbf{R} \quad (2.66)$$

or a more noise free estimate may be found as

$$\mathbf{F}^T = \hat{\mathbf{D}}^\dagger \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (2.67)$$

using the truncated data representation. Note also that  $\mathbf{H}$  from equation 2.63 may be thought of as a rotation of the singular vectors  $\mathbf{U}$ , into the proper space,  $\mathbf{D}$ . If  $\hat{\mathbf{H}}^T$  from 2.63 is substituted for  $\hat{\mathbf{D}}^\dagger \mathbf{U}$  in equation 2.67, then

$$\hat{\mathbf{F}}^T = \hat{\mathbf{H}}^T \mathbf{S} \mathbf{V}^T \quad (2.68)$$

and it may be seen that  $\hat{\mathbf{H}}^T$  is the rotation matrix for the unnormalized singular vectors  $\mathbf{S} \mathbf{V}^T$  into the space  $\hat{\mathbf{F}}^T$ .

The symmetry of the problem is illustrated graphically in Figure 2.4. Starting

with a data matrix  $\mathbf{R}$  the singular value decomposition is used to compress the data into the significant abstract singular vectors  $\mathbf{U}$  and  $\mathbf{V}$ . A parametric model may then be postulated for either the column space or the row space and the parameters of the model found by the nonlinear optimization technique. Once one of the spaces is found the other may be computed as shown in equations 2.66 or 2.67 for calculation of  $\hat{\mathbf{F}}$  from  $\hat{\mathbf{D}}$ . As will be seen in Chapter 4, the use of the two spaces to apply constraints to the nonlinear modeling can be very helpful in solving otherwise ambiguous kinetics problems.

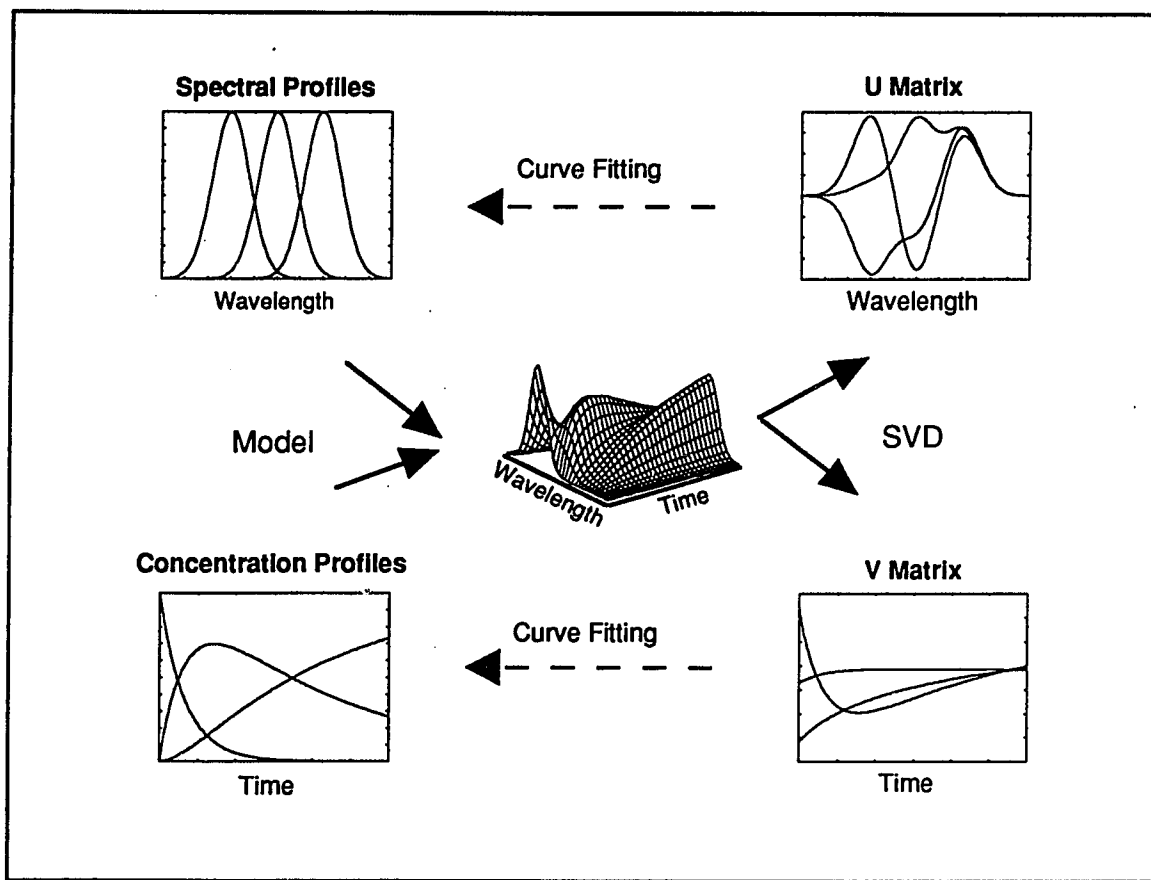


Figure 2.4 Multiresponse nonlinear analysis.



### **2.3.3 Application of Multiresponse Nonlinear Optimization**

The review of applications of multiresponse nonlinear optimization will be broken into those references concerning direct observation experiments and those focussed on indirect observation experiments.

#### **2.3.3.1 Applications Involving Direct Observation Experiments**

The primary paper dealing with the proper method of combining multiresponse data from direct observation experiments is that of Box and Draper (55). They showed that the proper choice of objective function for nonlinear optimization using multiresponse data is the determinant of  $\mathbf{Z}^T\mathbf{Z}$ , where  $\mathbf{Z}$  is the residual matrix. This criterion should provide for more precise parameter estimates than a simple total sum of squares (TSS) when the response errors are of a non-constant variance and/or there is a structure in the response variance-covariance matrix. Chapter 5 of this dissertation explores this topic further.

Hunter discusses and compares the criteria used for fitting multiresponse data and reiterates the understanding that the determinant criteria requires fewer assumptions about the error structure (53).

Mezaki and Butt evaluated the total sum of squares and the determinant criterion for estimation of kinetic parameters from multiresponse data in complex systems (60). They found that the determinant criterion allows precise estimation of constants which may have only a small influence in certain regions of the fitted data. Convergence was found to be much faster using the determinant criterion and is believed to be due to the nature of the response surface generated, rather than a

property of the minimization technique. They claimed to have shown better parameter precision using the determinant criterion, but it would appear on closer inspection of the paper that questions of accuracy are being addressed.

Eakman used a combination of total sum of squares and the determinant criterion to estimate rate constants from multiresponse isothermal time-concentration data (61). As the initial process for multiresponse nonlinear computations Eakman used the "approximate" TSS criterion, which may be readily differentiated, to provide expressions for the more efficient gradient search technique. The results from the TSS minimization were then used as a starting point for the determinant criterion minimization using a hill-climbing search method. Eakman found the determinant criterion to give better parameter estimates in most cases, but never poorer estimates than the TSS criterion.

Erjavec has criticized Eakman's paper pointing out that the response variances were not constant with time, as is required for both the TSS and determinant criterion, and suggested that the responses should have been weighted or a log transform of the responses used (62). Erjavec also comments, that if the point of the paper was a comparison of fitting criteria it used poor examples. In Eakman's response to Erjavec criticism, he agrees with Erjavec concerning the requirements of the error variance, but showed that there was little difference in the results using a log transform and stressed that his original paper was reporting on a computational procedure rather than a comparison of fitting criteria (63).

Practical problems using the determinant criterion have been presented by Box et al. (64). They demonstrated the care that must be taken in using the determinant criterion when certain exact linear dependencies are present in the data. They also presented a method for the detection of these dependencies. Although the

data set they used demonstrated the point of the paper, it has since been shown that the analysis was flawed (65). An incorrect model was used as indicated by the non-random behavior of the residuals which were not examined in the original Box et al. paper.

McLean et al. have furthered the investigation of Box et al. for the use of the determinant criterion with data that have exact dependencies (singularities) (66). They pointed out two other situations that the analyst must be aware of and suggested methods to deal with them.

Ziegel and Gorman presented an application of the determinant criterion for multiresponse parameter estimation using data from an oil shale pyrolysis experiment (67). They also investigated a problem in multiresponse nonlinear optimization that required the estimation of a large number of parameters.

Bates and Watts have contributed extensively to the nonlinear parameter estimation literature. Recently they have presented a generalized Gauss-Newton procedure for multiresponse parameter estimation (68) and detailed the algorithm in a second publication (69). They have also discussed the implementation of the Gauss-Newton algorithm for the special case of a linear system of differential equations (56). A separate discussion by McLean of Bates and Watts' work is very helpful for implementation purposes (70). Bates and Watts' recent book on nonlinear regression summarized much of this work and presents pseudo computer code for the implementation of the algorithms (65).

### **2.3.3.2 Applications Involving Indirect Observation Experiments**

Perhaps the first to demonstrate multiresponse nonlinear parameter estimation

for data from an indirect observation experiment were Sylvestre, Lawton, and Maggio (47). They used principal component analysis and nonlinear regression to estimate equilibrium constants from ultraviolet (UV) spectrophotometric measurements. They eliminated the rotational ambiguity of PCA, thus providing spectral curve resolution, as well as equilibrium constant parameter estimation. In this work they also described a method of rank determination that makes use of an estimate of the measurement noise as compared to the residual after each principal component is extracted.

Shrager and Hendler applied singular value decomposition to the analysis of spectra collected from pH and electrochemical titrations (57). They presented a method for rank determination involving the autocorrelation of the singular vectors. Shrager further investigated the use of SVD on synthetic data by simulating a pH titration experiment monitored with ultraviolet/visible (UV/VIS) spectroscopy (71). He explored the use of SVD as the pKs of the pure component titration profiles became less well separated and demonstrated the wealth of information that may be derived using a multivariate approach.

Frans and Harris also discussed the analysis of data consisting of UV/VIS spectra collected during a pH titration (72). Synthetic data were used to study their so called "reiterative least squares method" using different noise conditions and different spectral and pK resolution conditions. The power of the multivariate approach was demonstrated by resolving the spectra from a highly overlapped four component mixture of acid/base indicators. In a later paper, Frans and Harris investigated the use of SVD analysis on the same pH indicator data (72). They found the SVD least squares approach, which involved fitting of all the columns of  $U$  (see equation 2.62) simultaneously, used less computer time and provided superior parameter estimation

and spectral reconstruction than previous approaches.

Cochran and Horne presented a strategy for resolving rapid scanning wavelength experiments using PCA (73). Cochran et al. demonstrated an application of the strategy on the analysis of data from stopped-flow kinetics experiments on liver alcohol dehydrogenase catalyzed reduction of p-nitroso-N,N-dimethylaniline by 1,4-dihydronicotinamide adenine dinucleotide (74). They identified spectra and concentration-time profiles of the reactants and two intermediates demonstrating how information on transient species may be obtained.

Application of multiresponse nonlinear analysis to data from rapid scanning stopped-flow spectrophotometry was described by Halaka et al. (75). They used a weighted principal component analysis, presented by Cochran and Horne (76), to resolve the absorption surfaces generated during the reduction of cytochrome c oxidase by 5,10-dihydro-5-methylphenazine into individual spectral shapes and concentration-time profiles.

Shrager has recently summarized the analysis of rank deficient multiresponse data fitting and demonstrated the need for proper weighting of the singular vectors (77). Further comments on this problem are the subject of Chapter 5 of this dissertation.

## 2.4 Conclusions

Multivariate calibration and multiresponse nonlinear optimization have been discussed. Classical and inverse regression techniques were reviewed within the framework of multivariate calibration. The biased regression techniques of PLS and PCR are part of a continuum of regression techniques using the inverse regression

model. A relatively new technique known as locally weighted regression was reviewed. Multivariate calibration has been very successful in extracting information from data previously considered to be useless for analytical purposes. It is believed that by using these implicit modeling techniques in conjunction with infrared emission spectroscopy a powerful process analysis tool will be available.

Multiresponse nonlinear optimization was reviewed and a distinction was drawn between experiments in which the phenomenon were directly or indirectly observed. By combining multireponses in the estimation of nonlinear model parameters higher precision is expected. Introducing a hard model into the consideration of multivariate data provides a method to eliminate the rotational ambiguity of principal components analysis.

## 2.5 Notes to Chapter 2

1. Green, P. E. *Mathematical Tools for Applied Multivariate Analysis*; Academic Press, Inc.: San Diego, 1978.
2. Jolliffe, J. *Principal Component Analysis*; Springer: Berlin, 1986.
3. Mardia, K. V.; Kent, J. T.; Bibby, J. M. *Multivariate Analysis*; Academic Press: London, 1979.
4. Golub, G. H.; Vanloan, C. F. *Matrix Computations, Second Edition*; Johns Hopkins University Press: Baltimore, 1989.
5. Brereton, R. G. *Analyst* **1987**, *112*, 1635-1657.
6. Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; John Wiley & Sons: New York, 1980.
7. Sharaf, M. A.; Illman, D. L.; Kowalski, B. R. *Chemometrics*; John Wiley & Sons: New York, 1986.
8. Massart, D. L.; Vandeginste, B. G. M.; Deming, S. N.; Michotte, Y.; Kaufman, L. *Chemometrics: a textbook*; Elsevier: Amsterdam, 1988.
9. Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.
10. Massart, D. L., Editor-in-Chief; Hopke, P. K., Spiegelman, C. H., Wegscheider, W. E., Editors; Brereton, R. G., Dessy, R. E., Associate Editors *Chemometrics and Intelligent Laboratory Systems*; Elsevier: Amsterdam.
11. Kowalski, B. R., Editor-in-Chief; Brown, S. D., Vandeginste, B. G. M., Associate Editors *Journal of Chemometrics*; Wiley: Chichester.
12. Kowalski, B. R. *Anal. Chem.* **1980**, *52*, 112R-122R.
13. Frank, I. E.; Kowalski, B. R. *Anal. Chem.* **1982**, *52*, 232R-243R.
14. Delaney, M. R. *Anal. Chem.* **1984**, *56*, 261R-277R.
15. Ramos, L. S.; Beebe, K. R.; Carey, W. P.; Sanchez, E.; Erickson, B. C.; Wilson, B. E.; Wangen, L. E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 294R-315R.
16. Brown, S. D.; Barker, T. Q.; Larivee, R. J.; Monfre, S. L.; Wilk, H. R. *Anal. Chem.* **1988**, *60*, 252R-279R.
17. Brown, S. D. *Anal. Chem.* **1990**, *62*, 84R-101R.

18. Sanchez, E.; Kowalski, B. R. *J. Chemometrics* **1988**, *2*, 265-280.
19. Martens, H.; Naes, T. *Near-infrared Technology in the Agricultural and Food Industries*; Williams, P., Norris, K., Eds.; American Cereal Association, 1987; 57-87.
20. Sanchez, E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 496-499.
21. Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 1007A-1017A.
22. Martens, H.; Naes, T. *Trends Anal. Chem.* **1984**, *3*, 204-210.
23. Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193-1202.
24. Mandel, J. *American Statistician* **1982**, *36*, 15-24.
25. Sanchez, E.; Kowalski, B. R. *J. Chemometrics* **1987**, *2*, 247-263.
26. Sundberg, R.; Brown, P. J. *Technometrics* **1989**, *31*, 365-371.
27. Geladi, P.; MacDougall, D.; Martens, H. *Appl. Spectrosc.* **1985**, *39*, 491.
28. Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta.* **1986**, *185*, 1-17.
29. Brown, C. W.; Obremski, R. J. *Appl. Spect. Review* **1984**, *20*, 373-418.
30. Martens, H. A. "Multivariate Calibration: Quantitative Interpretation of Non-Selective Chemical Data, Parts I and II of the 1985 Ph.D. Dissertation"; The Norwegian Institute of Technology, University of Trondheim: Trondheim, Norway, NR-Report No.: 786 and 787, 1986, ISBN 82-539-0273 and 82-539-0274-3.
31. Lorber, A.; Wangen, L. E.; Kowalski, B. R. *J. Chemometrics* **1987**, *1*, 19-31.
32. Haaland, D. M.; Easterling, R. G. *Appl. Spectrosc.* **1980**, *34*, 539.
33. Haaland, D. M.; Easterling, R. G. *Appl. Spectrosc.* **1982**, *36*, 665.
34. Draper, N.; Smith, H. *Applied Regression Analysis, Second Edition*; John Wiley & Sons: New York, 1981, 294.
35. Tenge, B. J. "F-Test Significance in Near-infrared Calibration and Near-infrared and Mid-infrared Characterization and Calibration Studies of Matrix-immobilized and Matrix-dispersed Compounds Bearing Amide Groups." Ph.D. Dissertation, University of Washington: Seattle, WA (1989). Available from: University Microfilms, Ann Arbor, MI.
36. Pearson, K. *Philosophical Magazine* **1901**, *6*, 559-572.
37. Hotelling, H. *Journal of Educational Psychology* **1933**, *24*, 417-441.



38. Wold, S.; Esbensen, K.; Geladi, P. *Chemo. Intell. Lab. Sys.* **1987**, *2*, 37-52.
39. Wold, H. *Multivariate Analysis*; Krishnaiah, P. R. Ed.; Academic Press: New York, 1966; Vol. 6, 391-420.
40. Wold, S.; Albano, C.; Dunn, W. J.; Esbensen, K.; Hellberg, S.; Johansson, E.; Sjostrom, M. *Food Research and Data Analysis*; Martens, H. Russworm, Jr., H. Eds.; Applied Science Publishers: London, 1983; 147-199.
41. Wold, S. *Technometrics* **1978**, *20*, 397-405.
42. Manne, R. *Chemo. Intell. Lab. Sys.* **1987**, *2*, 187-197.
43. Hoskuldsson, A. *J. Chemometrics* **1988**, *2*, 211-228.
44. Helland, I. S. *Reports from the Departments of Mathematics and Statistics, Agricultural University of Norway*, **1986**, *21*, 44.
45. Stone, M. *J. Roy. Stat. Soc. Ser. B* **1974**, *36*, 111-148.
46. Eastment, H. T.; Krzanowski, W. J. *Technometrics* **1982**, *24*, 73-77.
47. Sylvestre, E. A.; Lawton, W. H.; Maggio, M. S. *Technometrics* **1974**, *16*, 353-368.
48. Tu, X. M.; Burdick, D. S.; Millican, D. W.; McGown, L. B. *Anal. Chem.* **1989**, *61*, 2219-2224.
49. Osten, D. W. *J. Chemometrics* **1988**, *2*, 39-48.
50. Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification and Regression Trees*; Wadsworth: Belmont, CA, 1984.
51. Cleveland, W. S.; Devlin, S. J. *J. American Statistical Association* **1988**, *83*, 596-610.
52. Naes, T.; Isaksson, T.; Kowalski, B. R. *Anal. Chem.* **1990**, *62*, 664-673.
53. Hunter, W. G. *Ind. Eng. Chem.* **1967**, *6*, 461-463.
54. Box, G. E. P.; Hunter, W. G.; Hunter, J. S. *Statistics For Experimenters*; John Wiley & Sons Inc.: New York, 1978.
55. Box, G. E. P.; Draper, N. R. *Biometrika* **1965**, *52*, 355-365.
56. Bates, D. M.; Watts, D. G. *Technometrics* **1985**, *27*, 329-339.
57. Shrager, R. I.; Hendler, R. W. *Anal. Chem.* **1982**, *54*, 1147-1152.
58. Lawton, W. H.; Sylvestre, E. A. *Technometrics* **1971**, *13*, 461-467.

59. Nelder, J. A.; Mead, R. *Computer Journal* **1965**, *7*, 308-313.
60. Mezaki, R.; Butt, J. B. *Ind. Eng. Chem.* **1968**, *7*, 120-125.
61. Eakman, J. M. *Ind. Eng. Chem. Fundamentals* **1969**, *8*, 53-58.
62. Erjavec, J. *Ind. Eng. Chem. Fundamentals* **1970**, *9*, 187.
63. Eakman, J. M. *Ind. Eng. Chem. Fundamentals* **1970**, *9*, 187-188.
64. Box, G. E. P.; Hunter, W. G.; MacGregor, J. F.; Erjavec, J. *Technometrics* **1973**, *15*, 33-51.
65. Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and Its Applications*; John Wiley & Sons: New York, 1988.
66. McLean, D. D.; Pritchard, D. J.; Bacon, D. W.; Downie, J. *Technometrics* **1979**, *21*, 291-298.
67. Ziegel, E. R.; Gorman, J. W. *Technometrics* **1980**, *22*, 139-151.
68. Bates, D. M.; Watts, D. G. *Siam. J. Sci. Stat. Comput.* **1987**, *8*, 49-55.
69. Bates, D. M.; Watts, D. G. *Commun. Statist.-Simula. Computa.* **1984**, *13*, 705-715.
70. McLean, D. D. *Technometrics* **1985**, *27*, 340-348.
71. Shrager, R. I. *Siam. J. Alg. Disc. Meth.* **1984**, *5*, 351-358.
72. Frans, S. D.; Harris, J. M. *Anal. Chem.* **1984**, *56*, 466-470.
73. Cochran, R. N.; Horne, F. H. *J. Chem. Phys.* **1980**, *84*, 2561-2567.
74. Cochran, R. N.; Horne, F. H.; Dye, J. L.; Ceraso, J.; Sueiter, C. H. *J. Chem. Phys.* **1980**, *84*, 230-245.
75. Halaka, F. G.; Babcock, B. T.; Dye, J. L. *J. Biophys.* **1985**, *48*, 209-219.
76. Cochran, R. N.; Horne, F. H. *Anal. Chem.* **1977**, *49*, 846-853.
77. Shrager, R. I. *Chemo. Intell. Lab. Sys.* **1986**, *1*, 59-70.

## **Chapter 3 Calibration Studies Using Infrared Emission for Remote Analysis**

### **3.1 Introduction to Chapter 3**

In a recent review of process analytical chemistry, noninvasive, remote methods of analysis were identified as highly desirable (1). Infrared emission spectroscopy (IES) is obviously such a technique, but its use has been limited due to the lack of suitable instrumentation of requisite sensitivity and ruggedness and the difficulty in deriving quantitative information from the observed spectra (2,3). The advent of process qualified Fourier transform infrared spectrophotometers has eliminated the instrumentation problem, but quantitation remains a challenge. Spectral interferences and anomalies due to blackbody emission, temperature gradients, self-reabsorption and internal reflections all contribute to the complications inherent in infrared emission spectroscopy (2,4).

Some of these complications may be overcome by careful experimental design, but these design considerations may not be possible in a process monitoring environment. An alternative is to mathematically extract information that may be available. Two approaches are possible, a "hard" modeling approach in which a mathematical model is derived from physical and chemical theory, or a "soft" modeling approach in which the data are analyzed with few assumptions about the data structure. Chapter one presented the "hard" model theory for infrared emission spectroscopy and highlighted some of the difficulties encountered when applying such an approach. Chapter two presented an alternative in terms of multivariate

calibration. This chapter explores the use of the multivariate calibration approach for the prediction of chemical and physical properties from the infrared emission spectra of polymers and polymers coated on aluminum.

The analysis of polymer materials for composition as well as other physical properties is an important task for many industries. Section 3.2 details the use of multivariate calibration for the prediction of vinyl acetate concentration, thickness, and temperature from the infrared emission spectra of moderately thick, free standing samples of Du Pont's Elvax® polymer product. PLS and locally weighted regression are used for the data analysis.

The application of organic coatings to metal surfaces is an important process in the aluminum industry. Many types of coating materials are applied including mineral oil, synthetic oil, bonded polymer coatings, and heavy wax coatings. Measurements of degree of coating cure, coating homogeneity and coating weight per unit surface area are important considerations.

Section 3.3 presents the results from the use of multivariate calibration and infrared emission spectroscopy for the determination of loading weight of polymer coating materials on metal surfaces. In the production facility the polymer is applied to the metal surface and cured by passage through recirculating hot air ovens. The coated aluminum sheet is coiled as it leaves the hot air ovens. The current technique for coating weight analysis involves the removal of a small sample from the end of a production coil and the measurement of the weight loss after stripping the polymer coating from the metal. This analysis method is invasive, destructive, and hardly compatible with a continuous coating operation producing coated materials at a rate of hundreds of feet per minute.

The ideal method would allow on-line measurement of coating weight, in a

noninvasive and nondestructive manner. Infrared emission would appear to be an ideal approach to obtain these on-line measurements for these process conditions. In order to evaluate IES for measuring coating weight on aluminum surfaces, laboratory experiments were conducted. These experiments were used to determine the loading weight prediction errors using infrared emission spectroscopy and multivariate calibration. The effect on prediction error of angle of observation and sample temperature were explored.

## **3.2 Elvax® Polymer Calibration Study**

### **3.2.1 Experimental and Data Analysis**

Samples used in this study consisted of a series of polyethylene and vinyl acetate co-polymers. Vinyl acetate monomer concentrations for the samples were 28, 18, 15, 9.5 and 9.0% by weight. These samples correspond to the Du Pont product numbers Elvax® 265, 450 (and 470), 550, 770 and 750. Vinyl acetate concentrations are mean values for the monomer, before polymerization, from the published nominal values with a range of  $\pm 1\%$  absolute. (Subsequent analysis of the polymers for vinyl acetate concentration using proton NMR was consistent with these data and showed a precision of about  $\pm 1\%$ , see Appendix A). The sample pellets were pressed into films of varying thickness using a Carver press with a heated plate attachment. Thicknesses were selected using calibrated metal shims. Samples were pressed for approximately 5 minutes at about 1000 psi, at 10 °C below the published softening point of the particular polymer.

Data were collected on a Perkin-Elmer Model 1800 FT-IR spectrophotometer

equipped with a liquid nitrogen cooled, wide band mercury-cadmium-telluride (MCT) detector in which the normal source was replaced by a heated cell containing the sample. Samples were held in the cell as polymer films. In contrast to most other studies, no reflecting backplate was employed. Single beam emission spectra were measured from 4000 to 450  $\text{cm}^{-1}$  (2.5 to 22.2 micrometers ( $\mu\text{m}$ )) at 2  $\text{cm}^{-1}$  intervals by the co-addition of 16 to 32 scans at 4  $\text{cm}^{-1}$  resolution. Data were collected on samples prepared from the five different concentrations of vinyl acetate (28, 18, 15, 9.5, and 9.0%) at four thicknesses (0.508, 0.305, 0.203, 0.102 mm) and three temperatures (120, 110, and 100  $^{\circ}\text{C}$ ). Temperature was controlled to  $\pm 1$   $^{\circ}\text{C}$  using a Perkin-Elmer temperature controller. Each concentration and thickness corresponded to a physically different pressing of the sample pellets. Temperature variations were induced in a systematic fashion from 100 to 120  $^{\circ}\text{C}$  on each pressed sample.

Data analysis was performed using partial least squares (PLS). The region from 2000 to 450  $\text{cm}^{-1}$  (5 to 22  $\mu\text{m}$ ) of the single beam emission spectra, consisting of 776 data points, was used in all cases and the data were mean centered before application of PLS. The 4000 to 2000  $\text{cm}^{-1}$  (2.5 to 5  $\mu\text{m}$ ) region was eliminated due to poor signal-to-noise ratio as discussed in Chapter 1. The selection of the optimal number of latent variables used in the PLS model was based on the estimated error of the predictive residual error sum of squares (PRESS, see equation 2.49) as calculated by Haaland et al. using equation 2.50 (5).

The standard error of prediction (SEP, see equation 2.51) was used to evaluate the predictive ability of the method. This gives a more conservative estimate of the predictive ability of the model (that is, the ability of the model to accurately predict an unknown sample) than if the predicted samples were included in the model.

### 3.2.2 Partial Least Squares Regression Results

Figure 3.1 compares a typical single beam emission spectrum from a thin sheet of Elvax® with the corresponding transmission spectrum. This particular sample contained 18% vinyl acetate, at 0.1 mm thickness measured at 120 °C. The transmission spectrum is the same material at approximately 50 to 100 µm thickness.

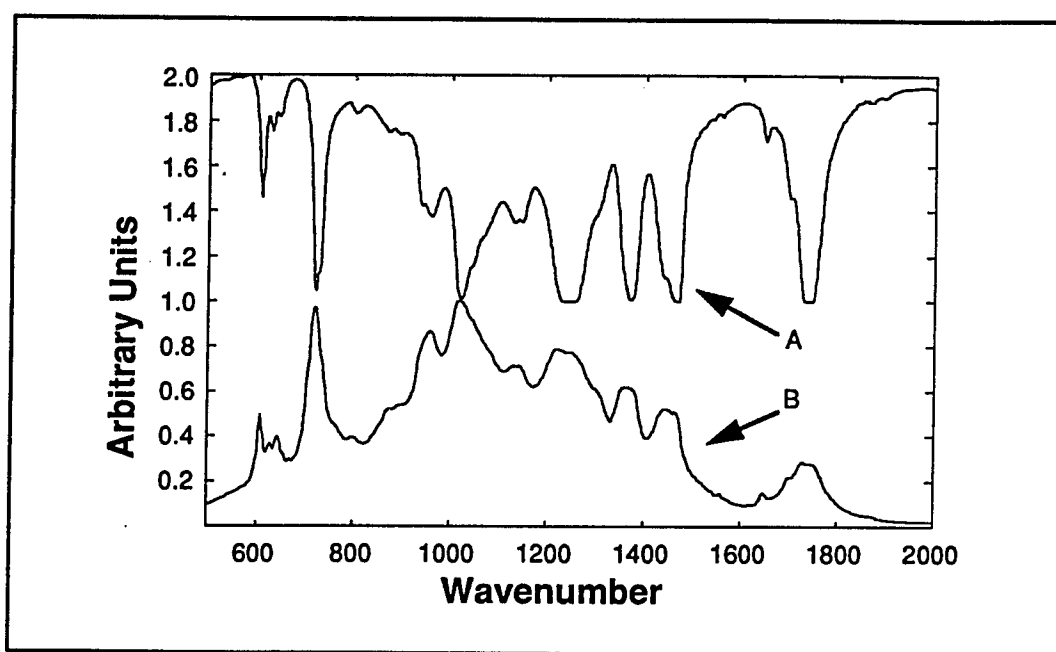


Figure 3.1 a) Single beam emission spectrum of 18 % vinyl acetate Elvax ®, 0.102 mm at 120 °C b) Transmission measurement of the same material.

No attempt was made to correct the emission spectra with respect to response of the instrument or temperature. This could have been done by ratioing to a blackbody at the sample temperature, but then the temperature information would be lost. The position of emission bands corresponds directly with bands

seen in the reference transmission spectrum in Figure 3.1b. Due to multiple radiation transfer events from thick films, the band shapes in the emission spectrum are distorted and intensities are dependent on temperature according to Planck's equation. Nevertheless, specific group frequencies and indicator (fingerprint) bands can be readily identified in the emission spectrum of this sample. Indicative of the acetate functional group are the C=O stretch at  $1740\text{ cm}^{-1}$  ( $5.7\text{ }\mu\text{m}$ ), and the  $1240\text{ cm}^{-1}$  ( $8.1\text{ }\mu\text{m}$ ) and  $1375\text{ cm}^{-1}$  ( $7.3\text{ }\mu\text{m}$ ) bands due to the protons of the  $\text{CH}_3\text{-(C=O)}$  group. The  $1450$  and  $720\text{ cm}^{-1}$  bands ( $6.9$  and  $13.9\text{ }\mu\text{m}$ ) are clearly identifiable, indicating deformation and rocking of  $\text{CH}_2$  groups arising from the (poly)ethylene contribution.

A single beam emission spectrum of each of the 20 sample films (five concentrations at four thicknesses) was measured at three temperatures ( $100$ ,  $110$  and  $120\text{ }^\circ\text{C}$ ). This gave a total of 60 possible spectra. The measurements were always made by increasing the sample temperature from  $100$  to  $120\text{ }^\circ\text{C}$  thus instrument drift over the time of the collection of the spectral data may be correlated with the temperature variance and may adversely affect the predictive ability of the method with respect to temperature. Seven of these spectra were eliminated before the data analysis phase. Six of the eliminated spectra were of the  $0.508\text{ mm}$  thickness samples at 28 and 18 % vinyl acetate and at all three temperatures. At this thickness these samples were not adequately held in the sample holder and were observed to have collapsed out of the focus point under their own weight. A seventh sample was eliminated (9% vinyl acetate at  $0.508\text{ mm}$  and  $110\text{ }^\circ\text{C}$ ) because of an obvious lack of temperature control adjustment. The remaining 53 single beam emission spectra were used as a training set to build a model using PLS analysis. A separate PLS model was built to predict



each of the variables of interest, namely vinyl acetate concentration, thickness, and temperature. Four additional samples were prepared as "unknowns" to further test the predictive ability of the method. The four samples contained 18% vinyl acetate (Du Pont product number 470 rather than product number 450 used in the calibration set). Three of the samples were pressed to 0.152 mm thickness and one to 0.127 mm thickness. Single beam emission spectra of the three 0.152 mm samples were measured at 105 °C and 115 °C. The emission spectra of the 0.127 mm samples were measured at 118, 113, 108 and 102 °C, yielding a total of 10 "unknown" spectra.

### **3.2.2.1 Prediction of Vinyl Acetate Concentration**

Prediction of vinyl acetate concentration of samples from a co-polymer of ethylene and vinyl acetate might be anticipated as an easy task if the spectroscopic responses used were linear with concentration. The infrared emission spectra observed in these experiments have a complex behavior with respect to vinyl acetate concentration even at a constant thickness and temperature as Figure 3.2 displays. Changes in refractive index due to changes in concentration may result in changes in selective reflections at the surface. Changing thermal properties of the samples may change the nature of possible temperature gradients in the samples. It is hoped that with the implicit modeling approach the detailed understanding and physical modeling of this complex system will not be necessary. The modeling and understanding of the physics that is the reality of the experiment are not unimportant, but it may not be practical to expect to use such knowledge in a process monitoring situation.

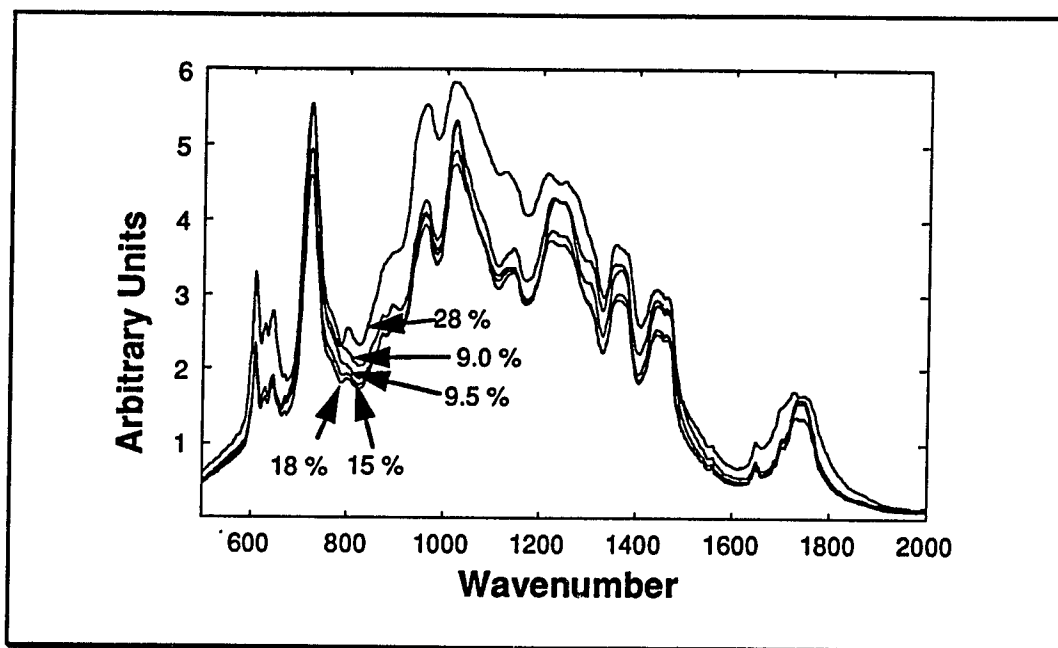


Figure 3.2 Infrared emission spectra as a function of vinyl acetate concentration at a constant thickness of 0.102 mm and temperature of 120 °C.

An additional complication for this calibration problem is that thickness and temperature variations are included in the data set. These variations have as much if not more of an effect on the observed spectra than does vinyl acetate concentration. Rather than attempting to remove the temperature information (which might be a valuable piece of data in a process situation), by ratioing to a blackbody at the same temperature, the raw uncorrected data were used and an implicit model built using PLS. In this situation an inverse least squares approach must be used because the "pure spectrum" of thickness and temperature are not known.

Two methods to assess the predictive ability of a calibration model are leave one (or many) out cross-validation or prediction of an independent data set. Because of the manner in which the data were collected (temperature variations induced on a

sample of given concentration and thickness without preparation of a completely new sample at each temperature), the leave one out cross-validation calculation of the SEP does not reflect all of the variations encountered due to sample positioning and pressing. This may result in an underestimation of the SEP. To provide an estimate of the predictive ability of the method with respect to all aspects of the analysis, specific cross sections of the data (i.e., all the 18 % samples) were excluded, and the remaining data used to build a calibration model and predict the left out cross section of data. A separate leave one out cross-validation on the remaining calibration data was also performed for comparison. Table 3.1 summarizes the results for these analyzes.

**Table 3.1** Standard error of prediction (SEP) results for vinyl acetate concentration (absolute %) in Elvax<sup>®</sup> samples using PLS.

Exp. #	Calibration Data	Validation Data	SEP with 0.5 mm samples Absolute %	SEP without 0.5 mm samples Absolute %
1	All 53 samples	Leave one out CV	1.1 (12)	1.2 (13)
2	Leave out 18 %	Leave one out CV	1.5 (10)	1.2 (9)
3	Leave out 18 %	Predict 18 %	1.5 (10)	2.9 (9)
4	Leave out 15 %	Leave one out CV	1.4 (10)	1.6 (7)
5	Leave out 15 %	Predict 15 %	2.3 (10)	1.0 (7)
6	Leave out 9.5 %	Leave one out CV	1.4 (8)	1.2 (10)
7	Leave out 9.5 %	Predict 9.5 %	2.0 (8)	1.9 (10)
8	All 53 samples	Ten "unknowns"	1.0 (12)	NA
9	All 53 samples	Best independent 4	0.6 (12)	NA
10	All 53 samples	Worst independent 4	1.3 (12)	NA

CV means cross-validation.

NA means not applicable.

The numbers in parentheses in Table 3.1 indicate the number of latent variables used for the optimal model choice using equation 2.50. Many of the 0.508

mm samples were initially eliminated from the data set and the remaining samples at this thickness appeared to be somewhat anomalous. Therefore separate predictions with and without the 0.508 mm (referred to as 0.5 mm in Table 3.1) samples were performed.

For the leave one out cross-validation using all 53 samples a 12 factor model resulted in an SEP of 1.1 % absolute. This was in reasonable agreement with the NMR analysis of these samples which had a precision of about 1.0 % absolute. It did not make a significant difference if the 0.508 mm data were included or not.

Other attempts to assess the predictive ability are shown in experiments 2 through 7 of Table 3.1 in which sections of the concentration data were removed (i.e., all the 18 % data at all thicknesses and temperatures, for example), and either a leave one out cross-validation on the remaining data or a separate prediction of the excluded data was used as the validation method. As may be seen there are some fluctuations in the SEPs and the number of factors used, depending on which section of the data was eliminated. Experiment 3 appears to be extreme, with an SEP of 2.9 % for the prediction of the left out 18 % data without the 0.5 mm samples. It is difficult to explain this result because most of the other experiments showed a decrease in the SEP upon removal of the 0.508 mm data.

Perhaps the best measure of predictive ability is the use of all 53 samples in the calibration set and the prediction of separate unknowns, experiments 8 through 10 in Table 3.1. Experiment 8 is in good agreement with the leave one out cross-validation of the calibration data with an SEP of 1.0 %. Even in the worst case (the four most poorly predicted, fully independent unknowns) the SEP only rises to 1.3 %. The above results show that by using PLS with a 12 factor model the prediction of vinyl acetate concentration may be performed to an SEP level of less than 1.3 %

absolute, even with changing sample thickness and temperature. These variations associated with thickness and temperature have been implicitly modeled in the calibration step.

### 3.2.2.2 Prediction of Thickness

The single beam spectra showed considerable variation with thickness. While a fair amount of structure is seen in thinner samples, the identifiable bands are broadened and the amount of visible structure decreases as thickness increases as seen in Figure 3.3.

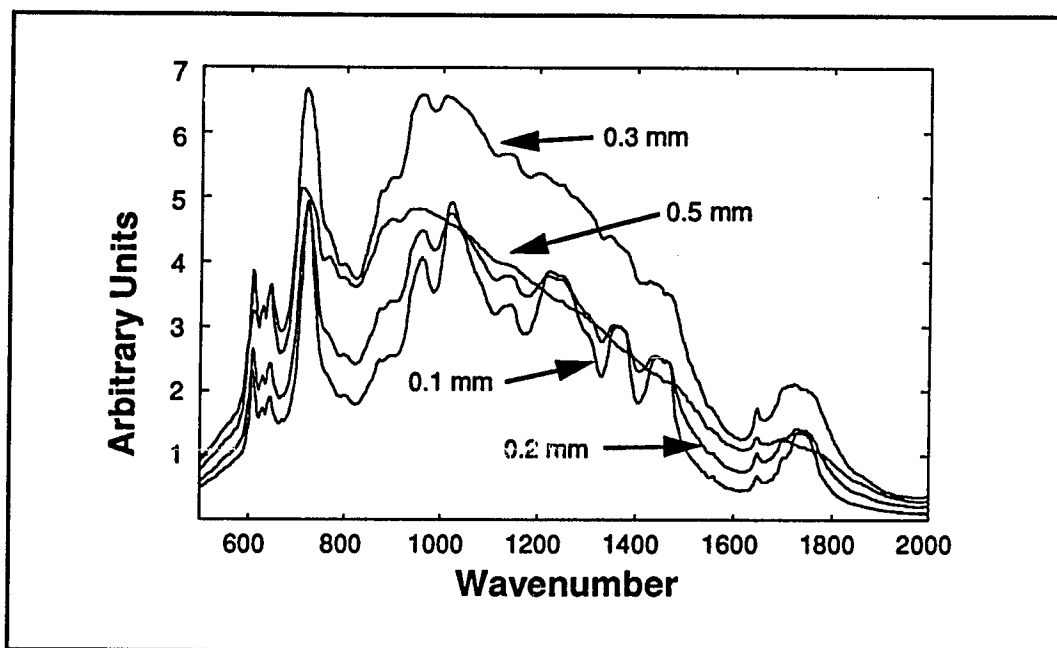


Figure 3.3 Infrared emission spectra as a function of thickness at a constant vinyl acetate concentration of 15 % and temperature of 120 °C.

The apparent anomalous behavior of the 0.508 mm spectrum is believed to be due to insufficient heating of the thicker samples. Although this may be adequate justification for removal of these samples from consideration, they have been retained, in order to provide a worst case scenario.

Given this variance with respect to thickness, a calibration model was developed to predict thickness with varying vinyl acetate concentration and temperature using the 53 spectra calibration set. Other combinations of the data were used to help assess the predictive ability and these results are summarized in Table 3.2. The numbers in parentheses indicate the number of factors used for an optimal calibration model.

**Table 3.2** Standard error of prediction (SEP) results for thickness (mm) in Elvax® samples using PLS.

Exp. #	Calibration Data	Validation Data	SEP (mm)
1	All 53 samples	Leave one out CV	0.018 (12)
2	Leave out 0.203 mm	Leave one out CV	0.015 (10)
3	Leave out 0.203 mm	Predict 0.203 mm	0.06 (10)
4	All 53 samples	Ten "unknowns"	0.06 (12)
5	All 53 samples	Best independent 4	0.05 (12)
6	All 53 samples	Worst independent 4	0.06 (12)

CV means cross-validation.

A 12 factor model was found to be optimal using leave one out cross-validation, giving an SEP of 0.018 mm. Removing all samples at 0.203 mm and performing a leave one out cross-validation analysis on the remaining data gave an SEP of 0.015 mm, although prediction of the 0.203 mm samples showed a 0.06 mm SEP. Prediction of "unknowns" using the entire calibration set showed SEP

values of 0.05 to 0.06 mm. This would indicate caution is required in the interpretation of SEP results from a leave one out cross-validation calculation if the data do not reflect all of the errors in the method. Reference thickness values were estimated to have a precision of about 0.03 mm and thus an SEP of 0.06 mm was not unreasonable considering that concentration and temperature were also varying in the calibration data set.

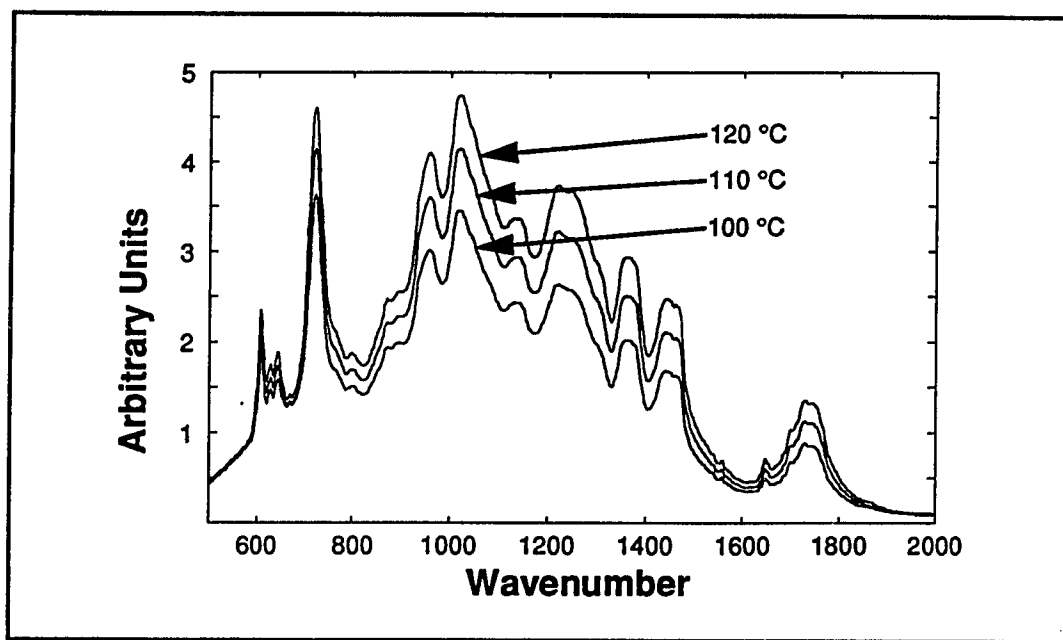
It is interesting to note that concentration and thickness variances have been separated in this analysis. If the system had followed a simple Beer's law model, or even the more complex equations for emission from a free standing polymer, the separation of the variance due to concentration from the variance due to thickness would be mathematically impossible. For this analysis the rather complex behavior of the emission signal has allowed the concentration and thickness variances to be separated.

### **3.2.2.3 Prediction of Temperature**

Because uncorrected single beam data were used, the spectra also contained temperature information and this is displayed in Figure 3.4 for a sample with a vinyl acetate concentration of 15 % and a thickness of 0.102 mm.

Spectra from the samples showed an increase in intensity with increasing temperature indicating temperature prediction may also be possible. Infrared emission is used extensively for temperature measurements, however some measurement or estimate of sample emissivity is normally required. For opaque (optically thick) samples this value might be assumed to be relatively constant over the range of samples analyzed, but with non-opaque samples the emissivity is

changing as concentration and thickness change, making temperature estimation more difficult. If sample composition and thickness were known, this could be estimated.



**Figure 3.4** Infrared emission spectra as a function of temperature at a constant vinyl acetate concentration of 15 % and thickness of 0.102 mm.

However, since the PLS models have been able to estimate composition and thickness fairly well, it seems reasonable that temperature may also be predicted. Again a calibration model was developed to predict temperature with varying composition and thickness using the same spectral data. The results of this analysis are shown in Table 3.3.



**Table 3.3** Standard error of prediction (SEP) results for temperature (°C) in Elvax® samples using PLS.

Exp. #	Calibration Data	Validation Data	SEP (mm)
1	All 53 samples	Leave one out CV	4 (10)
2	Leave out 0.203 mm	Leave one out CV	5.5 (13)
3	Leave out 0.203 mm	Predict 0.203 mm	3.8(13)
4	All 53 samples	Ten "unknowns"	9 (10)

CV means cross-validation.

The SEP from a leave one out cross-validation using the entire calibration set was  $\pm 4$  °C; while the total temperature range was only 20 °C. Prediction of the "unknowns" showed even poorer performance, giving an SEP of  $\pm 9$  °C. These results are not nearly as satisfying as the previous predictions of vinyl acetate concentration and thickness. Later investigation of the temperature controlling device showed large errors in the reproducibility of the temperature settings, and thus reasonable prediction errors for temperature would not be expected. Subsequent laboratory studies using liquid samples, in which temperature was controlled with greater precision, have shown that adequate prediction of temperature is possible even with thickness and concentration variations also present (6). In spite of not knowing the temperature reference values very well prediction of vinyl acetate concentration and thickness was accomplished. The temperature variance was implicitly modeled which may be the more probable course of action in a process measurement situation.

### 3.2.4 Locally Weighted Regression Results

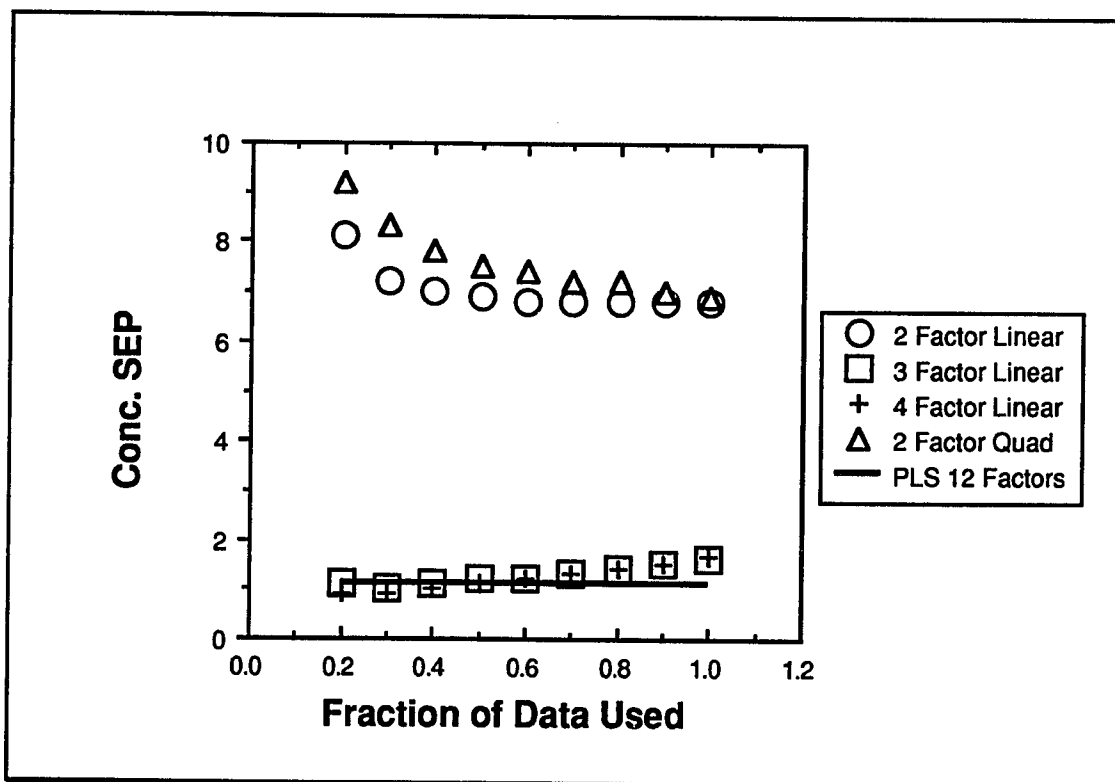
The very complex models required to predicted vinyl acetate concentration and thickness from infrared emission spectra are due to the linear nature of the calibration technique and the nonlinear nature of the data. By taking advantage of local linear regions, locally weighted regression (LWR) may be expected to perform as well if not better than PLS or PCR for data that show nonlinear behavior.

Locally weighted regression has been applied to the infrared emission data described in Section 3.2.1 using the principal components calculated from the entire calibration data set of 53 samples. Four different models were investigated (1) a two principal component linear model, (2) a two principal component quadratic model with cross terms, (3) a three principal component linear model, and (4) a four principal component linear model.

The fraction of data used,  $f$ , in each local regression is an important parameter, as described in Section 2.2.7. This fraction criterion was tested for each of the four models over the range of 0.20 to 1 at 0.10 intervals, using a leave one out cross-validation estimate of the SEP for comparisons. Table 3.4 displays the results for the prediction of vinyl acetate concentration and Figure 3.5 summarizes those results graphically.

**Table 3.4** Standard error of prediction results for vinyl acetate concentration, absolute % vinyl acetate, of Elvax<sup>®</sup> samples as a function of the model choice and fraction of the data ( $f$ ) used with locally weighted regression.

$f$	No. of points	SEP, 2 factor linear	SEP, 3 factor linear	SEP, 4 factor linear	SEP, 2 factor QUAD
0.20	11	8.1	1.1	0.9	9.2
0.30	16	7.2	1.0	0.9	8.3
0.40	21	7.0	1.1	1.0	7.8
0.50	27	6.9	1.2	1.1	7.5
0.60	32	6.8	1.2	1.2	7.4
0.70	37	6.8	1.3	1.3	7.2
0.80	42	6.8	1.4	1.4	7.2
0.90	48	6.8	1.5	1.5	7.0
1.00	53	6.8	1.6	1.7	6.9



**Figure 3.5** Graphical representation of Table 3.4.

As may be seen the two principal component linear and quadratic models performed much more poorly than the 12 factor PLS model. The three principal component and four principal component linear models perform as well as the PLS model up to a data fraction of about 0.6. This gives a sample to variable ratio of about 8:1 which suggest the method will behave reliably.

Table 3.5 and Figure 3.6 display the result for the thickness prediction. The two principal component models are again not found to perform as well as the PLS model, but the three and four principal component models do perform as well up to a data fraction of about 0.4 to 0.5.

**Table 3.5** Standard error of prediction results for thickness (mm) of Elvax<sup>®</sup> samples as a function of the model choice and fraction of the data ( $f$ ) used with locally weighted regression.

$f$	No. of points	SEP, 2 factor linear	SEP, 3 factor linear	SEP, 4 factor linear	SEP, 2 factor QUAD
0.20	11	0.057	0.019	0.018	0.074
0.30	16	0.055	0.020	0.018	0.062
0.40	21	0.055	0.022	0.018	0.059
0.50	27	0.056	0.025	0.020	0.058
0.60	32	0.057	0.026	0.021	0.059
0.70	37	0.057	0.028	0.023	0.059
0.80	42	0.059	0.031	0.025	0.058
0.90	48	0.060	0.033	0.028	0.057
1.00	53	0.061	0.038	0.033	0.056

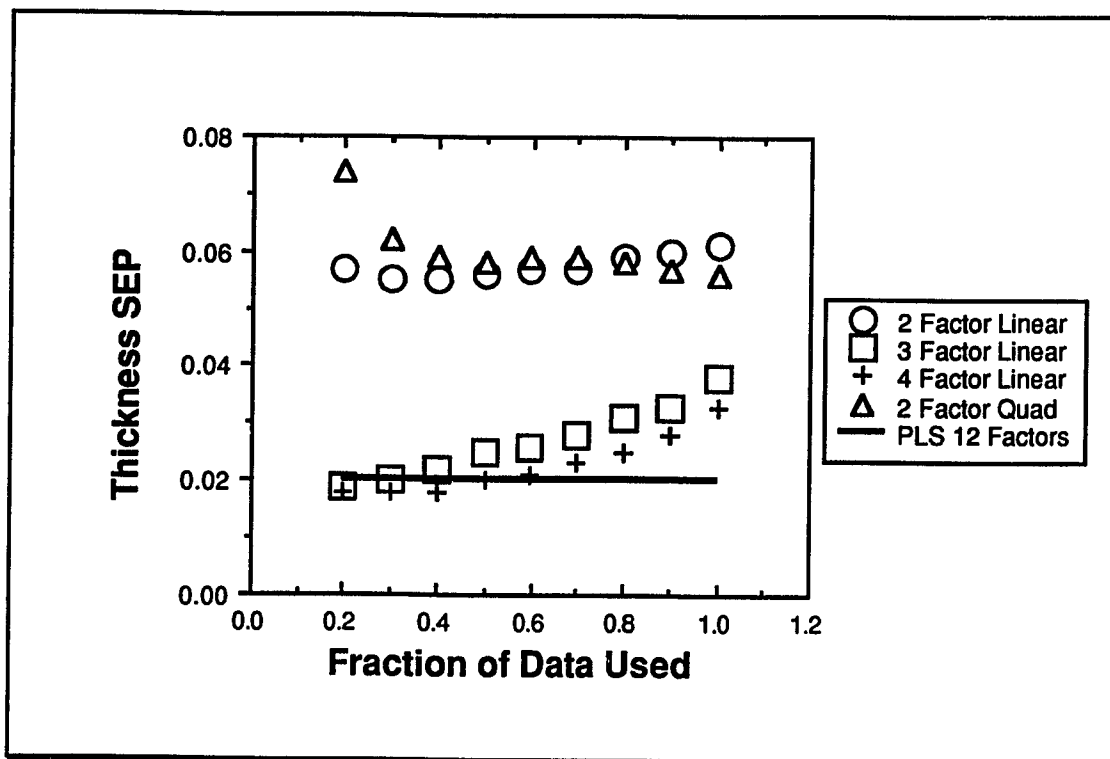


Figure 3.6 Graphical representation of Table 3.5.

This regression method provides an alternative to the standard PLS approach. By using fewer principal components from the entire data set, in a local fashion with proper weighting, a simpler model can be used to provide the same prediction error. A simpler model may be more easily interpreted, although the interpretation of multivariate regression models should be approached with caution, as has been recently shown by Seasholtz and Kowalski (7).

### **3.3 Loading Weight Prediction of Thin Polymer Films on Metal Surfaces**

Laboratory experiments were conducted to test the use of infrared emission spectroscopy and multivariate calibration for the prediction of film thickness of thin polymer coatings on metal surfaces. Thickness, angle of observation, and temperature were varied in these experiments.

#### **3.3.1 Experimental and Data Analysis**

Four sets of samples were provided by Ruth Roberts, Staff Scientist, of the Aluminum Company of America (ALCOA). The sample sets were designated as Set #1, #2, #3, and #4. Set #1 consisted of 11 samples of aluminum coated with a clear vinyl thermoset polymer ranging in loading weight from 0.75 to 9.15 milligrams per square inch (MSI). Table 3.6 displays the details of Sample Set #1.

**Table 3.6** Loading weight reference data for Sample Sets #1, #2, #3, and #4 in milligrams per square inch (MSI).

Sample #	Set #1, clear (MSI)	Set #2, low (MSI)	Set #3, nominal (MSI)	Set #4, high (MSI)
1	0.75	5.59	4.65	5.03
2	1.50	6.19	5.59	5.65
3	2.66	6.42	6.33 (2)	6.63
4	3.20	7.44	4.30 (3)	7.56
5	4.15	7.94	7.17	8.86
6	5.11	8.61	7.69	9.02
7	5.78	9.25	8.54	10.09
8	6.49	10.10	9.16	10.84
9	7.90	10.53 (2)	9.86	
10	8.30	10.89 (3)	10.56 (2)	
11	9.14	11.62	11.69	

Numbers in parentheses indicate multiple samples at that loading weight.

Sample Sets #2, #3, and #4 are aluminum coated with a different polymer than the one used in Sample Set #1. They are distinguished from each other by having low, nominal, and high pigment levels. The details are given in Table 3.6.

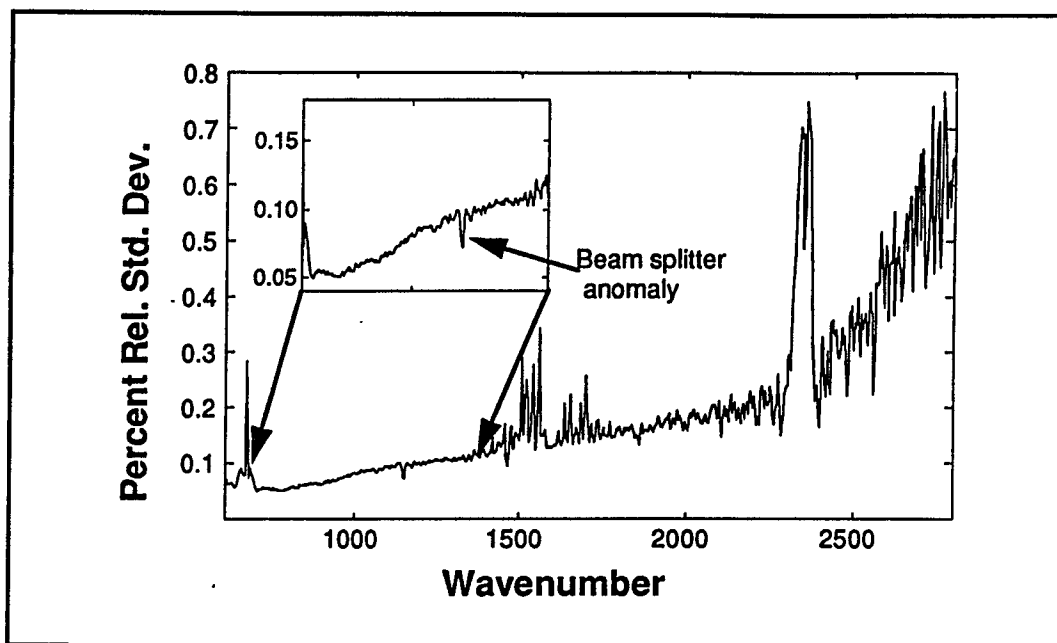
Emission spectra from 4000 to 450  $\text{cm}^{-1}$  (2.5 to 22.2  $\mu\text{m}$ ) at 4  $\text{cm}^{-1}$  resolution were measured using a Perkin-Elmer 1720 FT-IR spectrophotometer. Samples were mounted on a heated aluminum block and secured with spring steel clips. Emitted radiation was collected from a 16 mm spot approximately 30 inches from the spectrophotometer using a two mirror arrangement. Radiation was detected using a Judson wide band MCT liquid nitrogen cooled detector. In all cases each spectrum was the result of the co-addition of 10 spectra and was not corrected for instrument response or temperature. Acquisition time was approximately 12 seconds.

Data from 3000 to 450  $\text{cm}^{-1}$  (3.3 to 22.2  $\mu\text{m}$ ) were converted to the JCAMP (8) data format and transmitted to the Laboratory for Chemometrics Vax computer

system for further processing using the MATLAB (The MathWorks, Inc. Sherborn, MA) computing environment. Each spectrum was smoothed using the MATLAB DECIMATE function, reducing the number of data points by a factor of four from 2551 to 638 data points. This more reasonably reflects the true number of data points from a  $4\text{ cm}^{-1}$  resolution measurement.

As an initial characterization of the instrument, the spectrum of a blackbody (10 scans co-added) was measured repeatedly at 1.08 minute intervals over a twenty minute time period at temperatures of 70, 80 and 90 °C. These data may then be used to compute the standard deviation and relative standard deviation spectrum so the noise characteristics across the spectrum and at the different temperatures may be investigated. Figure 3.7 displays the relative standard deviation spectrum at 90 °C over the range from  $4000\text{ cm}^{-1}$  to  $604\text{ cm}^{-1}$  ( $2.5$  to  $16.5\text{ }\mu\text{m}$ ) and over the region used for calibration from  $1404\text{ cm}^{-1}$  to  $604\text{ cm}^{-1}$  ( $7.1$  to  $16.5\text{ }\mu\text{m}$ ) in the inset. Similar spectra were obtained for the lower temperature experiments, but with poorer signal-to-noise ratio.



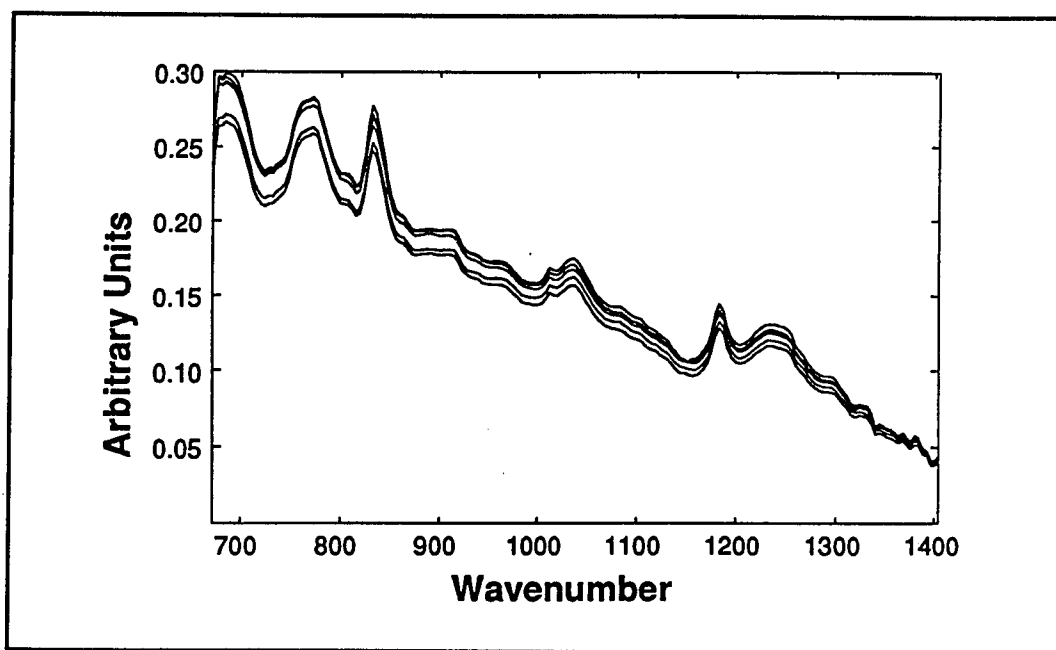


**Figure 3.7** Relative standard deviation of 90 °C blackbody data. Inset is region of spectrum used for calibration.

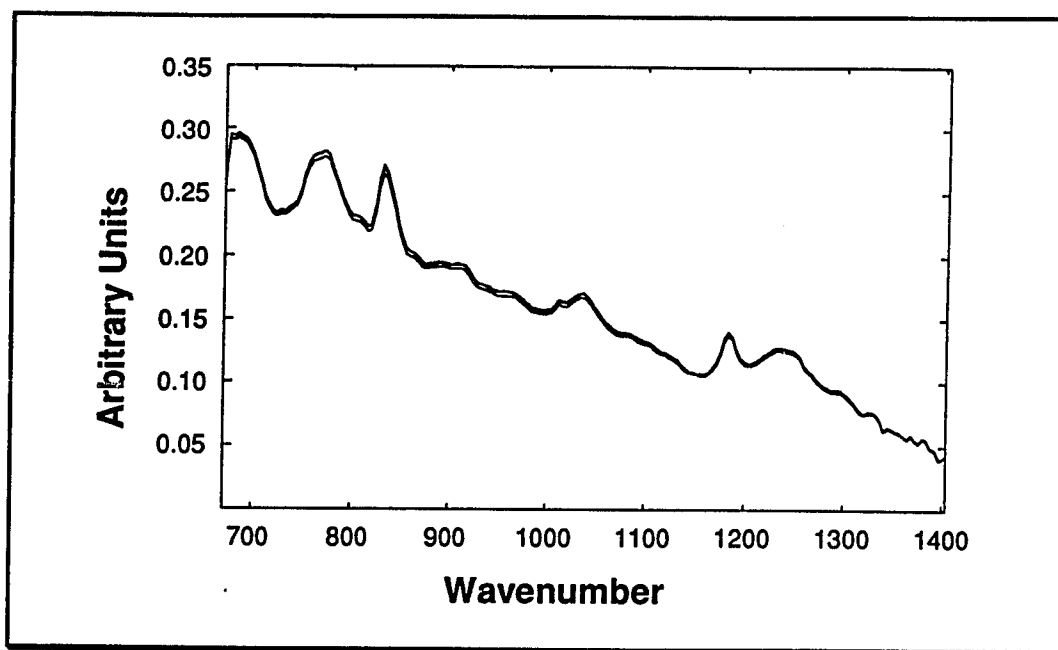
From the relative standard deviation spectrum it may be seen that the signal-to-noise ratio decreases at high wavenumbers. This is consistent with lower light levels at these frequencies due to a fall off in intensity as described by Planck's law. Noise due to water vapor may be seen in the 1800 to 1500  $\text{cm}^{-1}$  (5.5 to 6.6  $\mu\text{m}$ ) range and fluctuations due to  $\text{CO}_2$  are observed in the 2250  $\text{cm}^{-1}$  (4.4  $\mu\text{m}$ ) region. Within the spectral region used for calibration a spectral anomaly is noted. This is believed to be associated with the beam splitter coating and/or the products of the beam splitter reacting with moisture. Another anomalous band was found at approximately 1450  $\text{cm}^{-1}$  (6.9  $\mu\text{m}$ ), although it is not apparent in Figure 3.7. Noise in the calibration region ranges from 0.05 to 0.12 % for a signal-to-noise ratio of about 2000:1 to 833:1.

In the course of conducting these experiments over a three day period, three samples were measured at random times in order to assess the reproducibility of the

measurements. Six spectra were measured from samples containing 0.75, 5.78 and 9.14 MSI from Sample Set #1 and those from the 5.78 MSI sample are displayed in Figure 3.8. One measurement was made on day one, two measurements on day two and three measurements on day three. Measurements were from the middle of the sample at 80 °C and at an observation angle of zero degrees from the norm of the sample plane. There is considerable day to day variation, but less so on a given day (see Figure 3.9 for 3 spectra from 5.78 MSI sample collected during a single day). Any one calibration experiment was always completed in a single day thus minimizing the effect of these variations.



**Figure 3.8** Fluctuation in the infrared emission spectra for a 5.78 MSI sample measured over a three day period.



**Figure 3.9** Fluctuation in the infrared emission spectra for 5.78 MSI sample, three spectra measured on a single day.

The current method for measuring polymer loading weight is to weigh a sample of coated aluminum of a known area, strip the polymer layer off with methyl ethyl ketone, reweigh the bare aluminum, and compute the loading weight in milligrams per square inch (MSI). In order to gauge the precision of the reference method and the coating technique a study was conducted at the ALCOA research laboratories. The current procedure for the preparation of laboratory samples uses a draw bar. A draw bar is a steel rod with a wrapping of wire the thickness of which determines the loading weight of the coating. Three bars were evaluated in the ALCOA tests, bars #18, #36 and #38.

For each draw bar five panels of aluminum were coated yielding a total of 15 panels. From each panel, five samples were taken and strip weights determined, yielding a five by five matrix of strip weights for each of the three experiments. In

order to estimate the precision of the method, each matrix of data was subjected to an analysis of variance with the different panels considered as the different treatments. The following results for the within treatment variance estimates, related standard deviations, and grand means are shown in Table 3.7.

**Table 3.7** Analysis of variance results for reference calibration loading weight data.

Bar #	Variance (MSI) <sup>2</sup>	Standard Deviation (MSI)	Mean (MSI)
#18	0.006	0.08	5.57
#36	0.010	0.10	10.90
#38	0.016	0.13	11.64

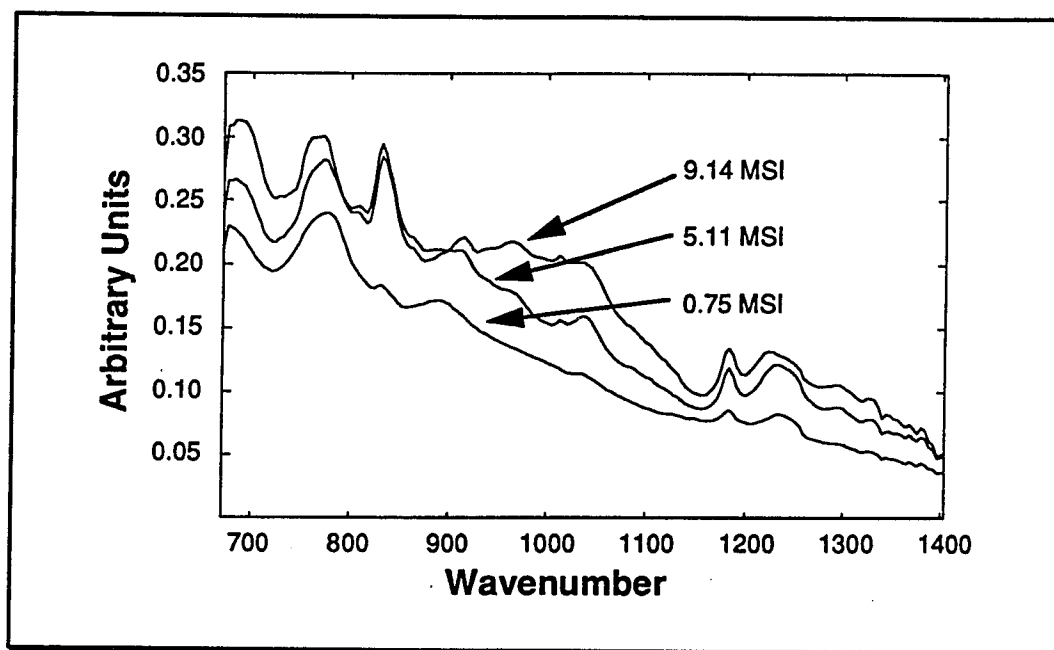
This provides an estimate of the experimental error in the reference method of about 0.1 MSI at one standard deviation. The on-line specification for measurement precision is 0.2 MSI (9). Further details for the infrared emission experiments are discussed below as pertains to each data set.

### 3.3.2 Partial Least Squares Regression Results

#### 3.3.2.1 Analysis of Sample Set #1

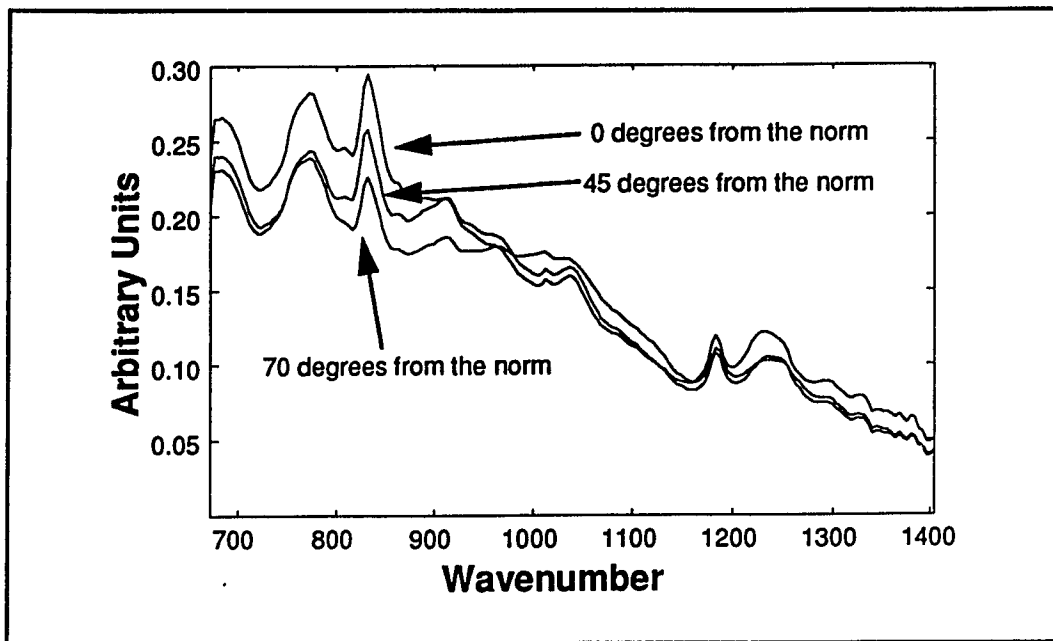
This sample set consisted of eleven samples coated with varying amounts of a clear vinyl thermoset polymer from 0.75 to 9.14 MSI, Table 3.6. Distinct spectral differences were noted with changes in polymer film loading weight. Figure 3.10 displays emission spectra (uncorrected for instrument response and temperature)

measured from three samples of different loading weights at 80 °C and at an angle of observation of zero degrees from the norm of the sample plane over the wavenumber range 2704 to 604  $\text{cm}^{-1}$  (3.7 to 16.6  $\mu\text{m}$ ). The intensity generally increases as the loading weight increases, except for some crossing of the 9.14 and 5.11 MSI sample spectra in the 800  $\text{cm}^{-1}$  (12.5  $\mu\text{m}$ ) region.



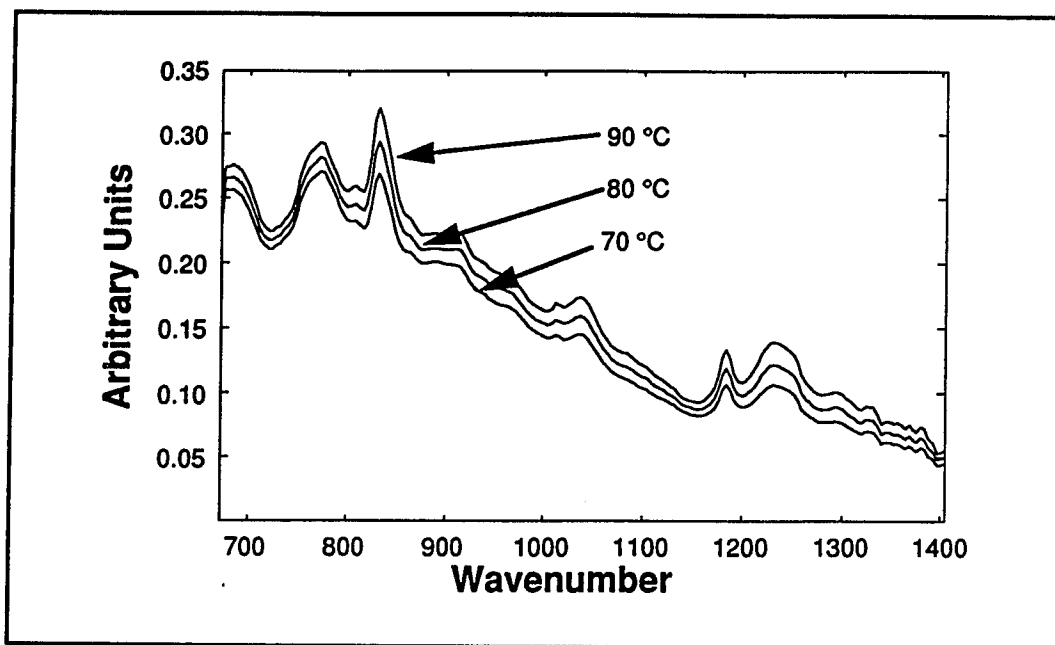
**Figure 3.10** Sample Set #1 emission spectra as a function of loading weight at a constant temperature of 80 °C and a zero degree angle of observation.

The spectral response is also a function of the angle of observation and the temperature as shown in Figures 3.11 and 3.12. Sample # 6 at 5.11 MSI is displayed at 0, 45 and 70 degrees from the norm in Figure 3.11.



**Figure 3.11** Sample Set #1 emission spectra as a function of angle of observation at constant loading weight of 5.11 MSI and 80 °C.

The same sample at zero degrees from the norm at temperatures of 70, 80 and 90 °C is displayed in Figure 3.12. The intensity as a function of angle of observation would appear to depend somewhat on the spectral location, but generally angles closer to the norm showed larger intensities. The estimated thickness of the films was 1 to 10  $\mu\text{m}$ . If the films had been thinner then larger intensities would have been expected at larger angles from the norm, as found by Greenler (10). One might have expected an intensity increase with increasing angles from the norm of the sample plane due to the increase in viewing area, but this was not observed. Spectral changes with temperature showed the expected response, increasing in intensity at all wavenumbers as the temperature was increased.



**Figure 3.12** Sample Set #1 emission spectra as a function of temperature at a constant loading weight of 5.11 MSI and zero degrees from the norm observation angle.

In order to investigate the effects that temperature and angle of observation have on a multivariate calibration approach for loading weight prediction, the following experiments were conducted. The temperatures chosen for these experiments were close to the expected process temperature of 80 °C.

Emission spectra were collected from three positions on each sample (low, medium and high) by repositioning the sample without removal from the sample holder, and at three temperatures, 70, 80 and 90 °C, with a constant angle of observation of zero degrees from the norm. This gives a total of 99 spectra for the eleven samples tested. The sample and temperature selections for each measurement were performed in a random fashion.

Multivariate calibration using PLS was applied to the mean centered data at each temperature setting. This partitions the data into three sets, each containing 33

samples and 638 responses. The region from 1404 to 672  $\text{cm}^{-1}$  (7.1 to 14.9  $\mu\text{m}$ ), containing 184 data points, was selected for calibration. This selection avoids the water vapor region of the spectrum and the low signal-to-noise region above 2000  $\text{cm}^{-1}$  (below 5  $\mu\text{m}$ ), as shown in Figure 3.7. Each temperature data set was first analyzed individually. For each of the temperature settings the standard error of prediction results (leave one out cross-validation) are presented in Table 3.8.

**Table 3.8** Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of temperature and the model complexity using PLS.

No. of Factors	SEP at 70 °C (MSI)	SEP at 80 °C (MSI)	SEP at 90 °C (MSI)
1	0.63	0.58	0.66
2	0.58	0.54	0.60
3	0.56	0.50	0.59
4	0.49	0.46	0.52
5	0.38	0.42	0.44
6	0.36	0.41	0.38
7	0.34	0.39	0.38
8	0.32	0.38	0.38
9	0.29	0.41	0.25 †
10	0.21 †	0.24 †	0.24
11	0.22	0.21	0.21
12	0.21	0.19	0.26
13	0.25	0.23	0.27
14	0.28	0.24	0.36
15	0.37	0.29	0.35

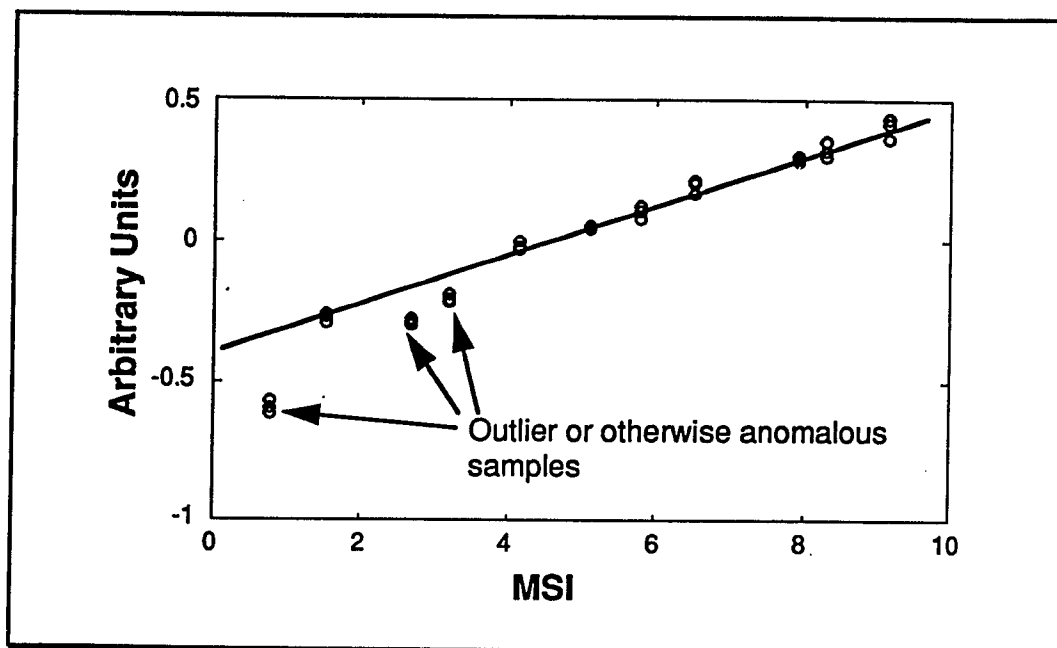
Eleven samples used for a total of 33 spectra.

† Indicate the optimal choice of the number of latent variables.

At each temperature a minimum of about 0.2 MSI in the SEP was found at about ten factors. This is not a very satisfying result considering only one experimental parameter is varying, the loading weight. Leverage plots indicated the 0.75 MSI sample might be an outlier, although deletion of the three samples at this



loading weight did not improve the prediction significantly. Several other combinations of samples were removed from the calibration without any significant effect on the prediction error. Inspection of the plot of the first score vector versus loading weight was very revealing, as shown in Figure 3.13.



**Figure 3.13** Scores from one factor model vs. the reference loading weights for Sample Set #1.

In a PLS analysis one expects maximum correlation between the PLS latent variables (the scores) and the constituent values. As may be seen in Figure 3.13, it appears as though the samples at 0.75, 2.66 and 3.2 MSI are outliers or otherwise anomalous samples. If these nine samples (three samples in triplicate) are removed the results of the analysis change dramatically as shown in Table 3.9.

**Table 3.9** Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of temperature and the model complexity using PLS. (Outliers deleted)

No. of Factors	SEP at 70 °C (MSI)	SEP at 80 °C (MSI)	SEP at 90 °C (MSI)
1	0.33 †	0.30 †	0.23 †
2	0.29	0.31	0.23
3	0.30	0.33	0.24
4	0.31	0.32	0.25
5	0.32	0.30	0.28
6	0.27	0.30	0.29
7	0.24	0.29	0.28
8	0.22	0.28	0.29
9	0.22	0.29	0.28
10	0.20	0.32	0.30
11	0.32	0.36	0.34
12	0.46	0.40	0.39
13	0.54	0.37	0.42
14	0.47	0.28	0.38
15	0.51	0.27	0.34

Eight samples used for a total of 24 spectra.

† Indicates optimal model choice.

Now in each case a one factor model is chosen and the SEP for the 90 °C data is close to that specified for an on-line instrument, 0.2 MSI. It is satisfying to see that the prediction error improves with an increase in temperature, as one might expect due to an increase in the signal-to-noise ratio. It should be noted that even with the deletion of these outlier samples nearly the entire original calibration range is still being covered.

The strip weights for several of the normal and outlier samples were reevaluated and found to agree with the reference values within the estimated 0.1 MSI precision of the reference method. At this point it is not clear why the three samples at 0.75, 2.66, and 3.20 MSI are anomalous.

A second experiment was performed to assess the effect of observation angle

on predictive ability. In this experiment a spectrum from a single position on each sample was measured at 0, 45 and 70 degrees from the norm. The sample and angle of measurement were selected in a random fashion. Table 3.10 displays the prediction results with the outliers included.

**Table 3.10** Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of angle of observation and the model complexity using PLS. All samples used

No. of Factors	SEP, 0 degrees from norm (MSI)	SEP, 45 degrees from norm (MSI)	SEP, 70 degrees from norm (MSI)
1	0.76 †	0.97 †	1.12 †
2	0.84	1.36	1.58
3	0.94	1.60	1.39
4	1.05	1.38	1.17
5	0.99	1.27	1.07
6	0.96	1.16	1.09
7	1.21	0.86	0.91
8	2.87	0.82	1.07
9	3.03	0.76	1.44

† Indicates the optimal model choice.

Table 3.11 displays the prediction results with the 0.75, 2.66, and 3.2 MSI outlier samples removed. The prediction error is significantly reduced. There also appears to be better prediction provided when using data from the zero degree angle of observation, which is consistent with the larger overall spectral intensities observed at the zero degree angle of observation, as shown previously in Figure 3.11

**Table 3.11** Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of angle of observation and the model complexity using PLS. (Outliers deleted)

No. of Factors	SEP, 0 degrees from norm (MSI)	SEP, 45 degrees from norm (MSI)	SEP, 70 degrees from norm (MSI)
1	0.43 †	0.62 †	0.72 †
2	0.50	0.66	1.00
3	0.58	0.72	0.98
4	0.65	0.80	0.85
5	0.72	0.92	0.83
6	0.63	0.96	0.93

† Indicates the optimal model choice.

All 99 samples were used in an attempt to predict loading weight. As seen in Table 3.12, the prediction of loading weight using all the samples gave an SEP of about 0.2 MSI, but the model is very complex, requiring 16 to 17 factors, and it does not appear to make much difference whether the outliers are excluded or not.

**Table 3.12** Standard error of prediction results for loading weights, MSI, for Sample Set #1 as a function of model complexity using PLS and data from all temperatures with and without outliers.

No. of Factors	SEP With outliers (MSI)	SEP Without outliers (MSI)
1	1.31	1.47
2	1.24	1.38
3	1.16	1.17
4	1.06	0.99
5	0.94	0.68
6	0.84	0.66
7	0.77	0.57
8	0.71	0.51
9	0.64	0.49
10	0.54	0.47
11	0.45	0.46
12	0.35	0.36
13	0.30	0.25
14	0.23	0.25
15	0.22	0.24
16	0.21	0.22
17	0.20 †	0.21
18	0.20	0.21 †
19	0.19	0.20
20	0.19	0.20

† Indicates optimal model choice.

A separate calibration was used to predict temperature from all 99 samples. As shown in Table 3.13 temperature prediction was optimal at 18 factors giving a SEP of 1.5 °C using all the data and 0.5 °C for the data with the outliers excluded.

**Table 3.13** Standard error of prediction results for temperature (°C), for Sample Set #1 as a function of model complexity using PLS and data from all temperatures with and without outliers.

No. of Factors	SEP With outliers (MSI)	SEP Without outliers (MSI)
1	7.53	6.60
2	7.29	6.09
3	5.86	4.56
4	6.31	4.15
5	5.97	2.74
6	5.28	2.64
7	5.13	2.11
8	4.66	1.67
9	4.41	1.41
10	3.74	1.27
11	3.48	1.13
12	2.70	0.77
13	2.30	0.73
14	2.09	0.72
15	1.90	0.66
16	1.79	0.59
17	1.65	0.56 †
18	1.54 †	0.55
19	1.44	0.55
20	1.43	0.53

† Indicates optimal model choice.

Many more latent variables are needed to model the loading weight variance when all the temperature information is included. This may in fact be somewhat unrealistic because temperature variations in a real process are probably much smaller. An experiment was conducted in order to assess what prediction error would be incurred given different levels of temperature fluctuation. For each concentration level, a temperature difference spectrum was constructed by subtracting the 70 °C spectrum from the 80 °C spectrum, thus creating a 10 °C difference spectrum. This spectrum was scaled and added in a random fashion to the 90 °C data to simulate one, two, three, four, and five degree temperature variances. This assumes the spectral

changes are linear with the temperature fluctuations, which is reasonable over small temperature changes.

Because of the computation time required to do a full cross-validation study the standard error of estimation (SEE) was used instead of SEP. SEE is computed just as SEP is, but the samples that are predicted are also used in the model building. Table 3.14 displays the SEE results for 20 repeats of the noise addition experiments using a one factor model, mean centered with outliers removed.

**Table 3.14** Standard error of estimate for simulated temperature variance experiments and associated standard deviation for 20 experiments.

Temperature variance	Mean SEE, 20 trials	Std. Dev. of SEE
0	0.20	0.03
1	0.25	0.06
2	0.39	0.06
3	0.51	0.07
4	0.70	0.12
5	0.81	0.10

A variance of 1 °C does not appreciably increase the SEE over the baseline SEE of 0.20 MSI. Further increases in the temperature fluctuations increased the SEE in a nearly linear fashion. It is not known what the real process temperature variations might be, but this experiment would indicate a greater than 1 °C variance level would result in unacceptably large loading weight estimation errors.

As was discussed at the beginning of this section, random fluctuations occurred in the emission spectrum of a single sample over the course of time. These deviations appeared to be simple instrumental offsets and elimination of these may

improve prediction. Attempts at unit area, unit length, first and second derivative transformation and fixing the baseline to an anchor point, showed no improvement in SEE. In fact in most instances the transformations significantly degraded the SEE.

In conclusion three of the samples appeared to be different from the rest of the samples in Sample Set #1. This was shown by the dramatic decrease in the complexity of the model when they were eliminated, although strip weight analysis of several samples indicated that the reference values were within the expected range. Prediction error was better for the higher temperatures, as would be expected, and approached that specified for on-line use (0.2 MSI). As for observation angle, it appeared that observations closer to the norm were better. Loading weight prediction was better at a single temperature setting. Temperature and loading weight prediction for all the data taken together over the three temperature ranges required a much more complex model for reasonable prediction errors.

### **3.3.2.2 Analysis of Sample Sets #2, #3, and #4**

Results from Sample Sets #2, #3, and #4 will be reviewed together. For each set the following experiment was performed. Three spectra from each sample were measured at three positions on the sample (low, medium and high) at 80 °C and zero degrees from the norm. The spectra were resampled and smoothed to reduce the number of responses from 2551 to 638. Figure 3.14 displays spectra from samples at 5.59, 9.25 and 11.62 MSI containing low pigment concentrations. Figure 3.15 displays spectra for samples at 5.59, 9.16 and 11.69 MSI and nominal pigment concentrations, while Figure 3.16 displays spectra for samples at 5.03, 9.02 and 10.84 MSI with high pigment concentrations, all measured at 80 °C and zero degrees from



the norm of the sample plane.

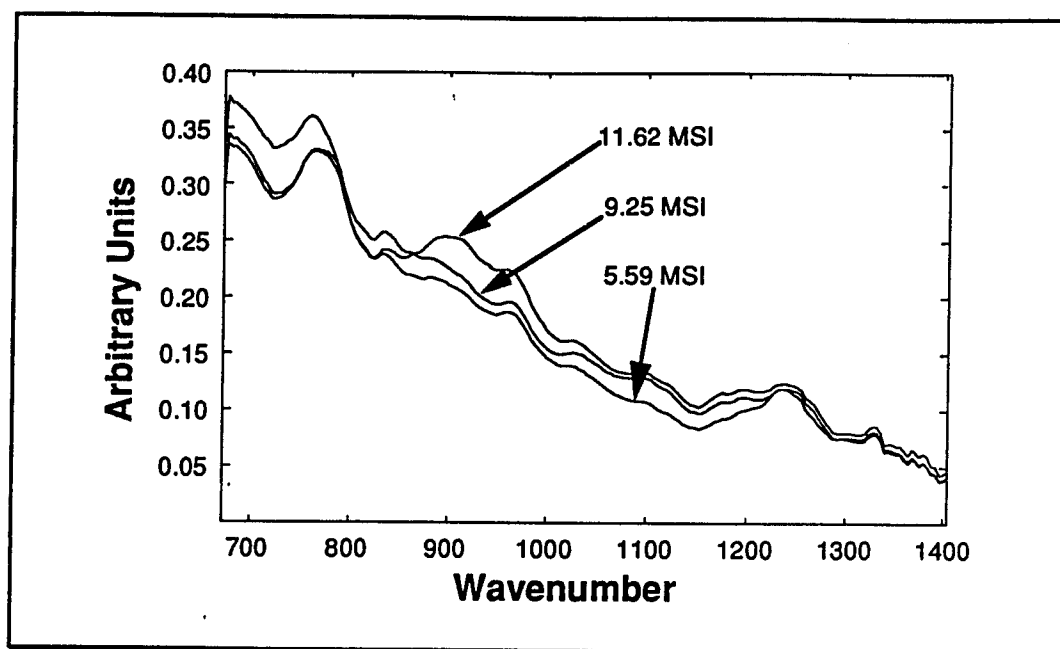


Figure 3.14 Spectra from low pigment concentration samples as a function of loading weight.

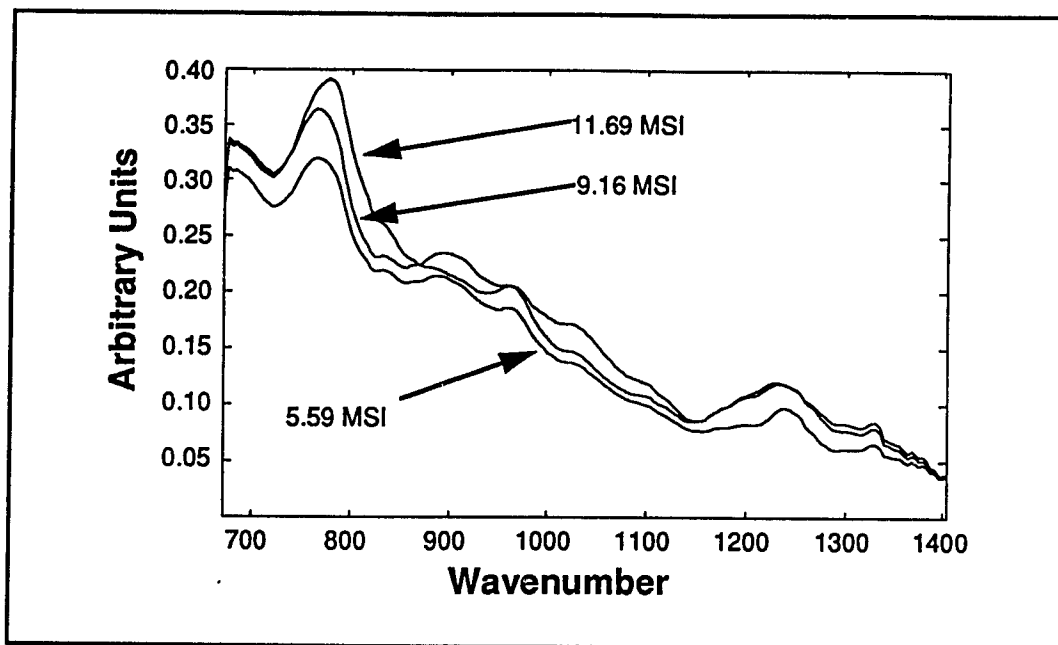


Figure 3.15 Spectra from nominal pigment concentration samples as a function of loading weight.

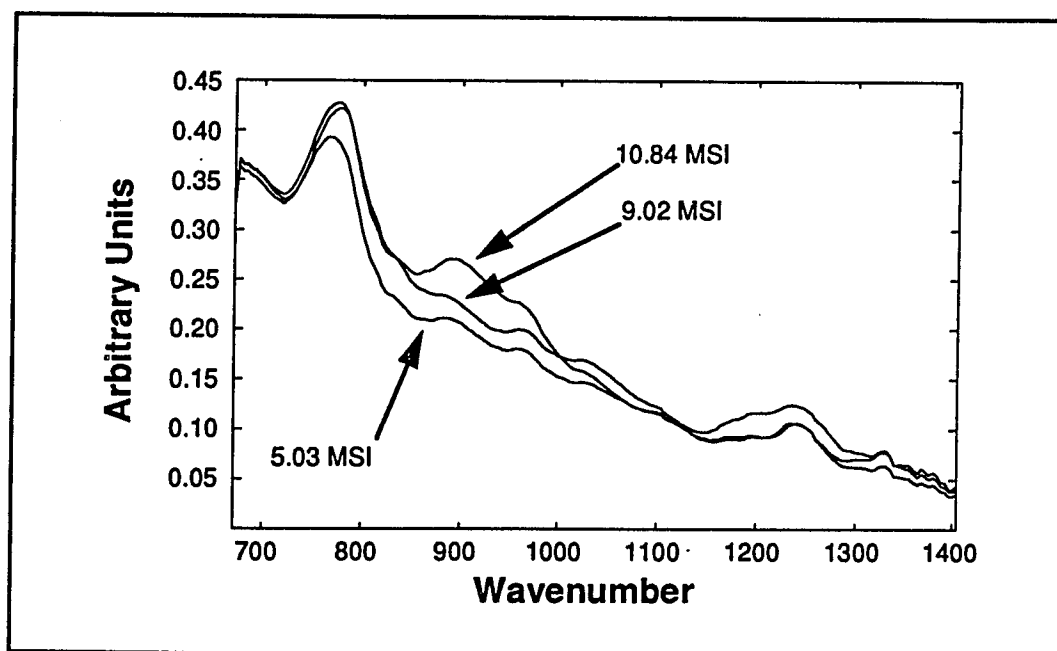


Figure 3.16 Spectra from high pigment concentration samples as a function of loading weight.

As may be seen there are still distinct spectral changes associated with loading weight changes, in spite of the presence of pigment. Table 3.15 displays the prediction results using PLS to calibrate for loading weight for each of the sample types. The data were mean centered before application of PLS.

**Table 3.15** Standard error of prediction results for loading weights, MSI, for Sample Sets #2, #3, and #4 as a function of model complexity using PLS.

No. of Factors	SEP for SET #2, LOW ~ (MSI)	SEP for SET #3 NOMINAL † (MSI)	SEP for SET #4, HIGH + (MSI)
1	0.66	0.80	0.55
2	0.49	0.64	0.44
3	0.44	0.51	0.41
4	0.41	0.45	0.38
5	0.39	0.40	0.30
6	0.33	0.39	0.15 *
7	0.24	0.36	0.13
8	0.18 *	0.20	0.13
9	0.15	0.17	0.12
10	0.15	0.13 *	0.12
11	0.14	0.10	
12	0.14	0.11	
13	0.14	0.10	
14	0.14	0.10	
15	0.12	0.11	

~ 15 samples, 45 spectra, range from 5.59 to 11.62 MSI.

† 15 samples, 45 spectra, range from 4.65 to 11.69 MSI.

+ 8 samples, 24 spectra, range from 5.03 to 10.84 MSI.

\* Indicates optimal model choice.

In this case the model is more complex than for Sample Set #1, probably due to the pigment and the associated scatter of the radiation. In each case the chosen number of factors for the model results in an SEP that approaches the reference

method error and is below the on-line specification error. No obvious outlier samples were identified.

In summary, the prediction of polymer coating weight from the infrared emission spectra using PLS is possible with a prediction error (for laboratory measurements) near that required for an on-line measurement. Clear samples are predicted with a simpler model than the pigmented samples. Prediction error variation for the clear samples with temperature and observation angle are consistent with the observed signal intensity changes with temperature and observation angle. That is, better prediction is observed for higher temperatures and observations angles closer to the norm where higher signal intensities are observed.

### **3.3.3 Locally Weighted Regression Results**

Locally weighted regression (LWR) analysis was applied to the infrared emission data collected from the four Sample Sets discussed in Section 3.3.1. In each case seven models were tested, one through six factor linear models and a two factor quadratic model. The models were compared using a leave one out cross-validated SEP over the data fraction range of 0.3 to 0.9.

#### **3.3.3.1 Analysis of Sample Set #1**

For Sample Set #1 the results for the LWR analysis of the 90 °C data at zero degrees from the norm are shown in Table 3.16. All the samples are used and the previously found "outliers" are not removed. As may be seen the locally weighted regression SEP approaches the PLS calibration SEP (0.25 at nine factors, Table 3.8)

using three factors and 0.3 of the points in the local regression.

**Table 3.16** Standard error of prediction results for loading weights, MSI, for Sample Set #1 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used. (Outliers not removed)

$f$	Pts.	1 factor SEP (MSI)	2 factor SEP (MSI)	3 factor SEP (MSI)	4 factor SEP (MSI)	5 factor SEP (MSI)	6 factor SEP (MSI)	Quadratic SEP (MSI)
0.3	10	0.52	0.53	0.28	0.33	0.43	0.41	3.68
0.4	13	0.54	0.52	0.35	0.34	0.30	0.44	2.67
0.5	17	0.53	0.55	0.37	0.35	0.30	0.47	1.68
0.6	20	0.53	0.59	0.42	0.40	0.34	0.42	1.40
0.7	23	0.56	0.60	0.46	0.43	0.41	0.44	1.52
0.8	26	0.58	0.60	0.48	0.44	0.44	0.42	1.42
0.9	30	0.58	0.61	0.49	0.45	0.45	0.43	0.63

A three factor model using 0.3 of the data is approaching the nine factor PLS results, indicating that the locally weighted regression surface is adapting to the regression surface containing the "outliers." Had the calibration space been more adequately sampled, the nonlinearities of the regression surface introduced by the "outliers" could have been better approximated. Table 3.17 displays the locally weighted regression results for Sample Set #1 with the outliers removed.

**Table 3.17** Standard error of prediction results for loading weights, MSI, for Sample Set #1 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used. (Outliers removed)

$f$	Pts.	1 factor SEP (MSI)	2 factor SEP (MSI)	3 factor SEP (MSI)	4 factor SEP (MSI)	5 factor SEP (MSI)	6 factor SEP (MSI)	Quadratic SEP (MSI)
0.3	10	0.24	0.37	0.28	0.77	0.56	0.53	1.49
0.4	13	<b>0.22</b>	0.31	0.24	0.67	0.39	0.36	1.04
0.5	17	0.22	0.29	0.23	0.25	0.33	0.29	0.45
0.6	20	0.23	0.29	0.23	0.23	0.29	0.28	0.31
0.7	23	0.24	0.28	0.22	0.23	0.28	0.29	0.29
0.8	26	0.24	0.26	0.22	0.23	0.28	0.31	0.30
0.9	30	0.25	0.26	0.24	0.23	0.27	0.30	0.27

The prediction error using LWR for the data with the outliers removed is not improved significantly over the PLS results. This is because the nonlinearities introduced by the outlier samples have been removed. A minimum in the prediction error is found at a data fraction of 0.4 with one factor, but it does not appear to be significantly different from the prediction error at a data fraction of 0.9 and one factor.

### 3.3.3.2 Analysis of Sample Sets #2, #3, and #4

For Sample Set #2 the locally weighted regression results for five factors and a data fraction of 0.3, shown in Table 3.18, are comparable to the results for the eight factor PLS model shown in Table 3.15.

**Table 3.18** Standard error of prediction results for loading weights, MSI, for Sample Set #2 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used.

$f$	Pts.	1 factor SEP (MSI)	2 factor SEP (MSI)	3 factor SEP (MSI)	4 factor SEP (MSI)	5 factor SEP (MSI)	6 factor SEP (MSI)	Quadratic SEP (MSI)
0.3	10	0.46	0.36	0.28	0.28	<b>0.18</b>	0.36	0.36
0.4	13	0.51	0.42	0.35	0.30	0.19	0.24	0.32
0.5	17	0.53	0.44	0.37	0.30	0.19	0.20	0.29
0.6	20	0.54	0.45	0.38	0.31	0.21	0.17	0.31
0.7	23	0.55	0.47	0.41	0.33	0.24	0.18	0.35
0.8	26	0.56	0.48	0.43	0.36	0.28	0.21	0.38
0.9	30	0.56	0.50	0.43	0.38	0.31	0.27	0.40

The LWR results for Sample Set #3 are shown in Table 3.19 and may be compared to the PLS results in Table 3.15. The locally weighted regression SEP at five factors and a data fraction of 0.3 is comparable to the ten factor PLS model.

**Table 3.19** Standard error of prediction results for loading weights, MSI, for Sample Set #3 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used.

$f$	Pts.	1 factor SEP (MSI)	2 factor SEP (MSI)	3 factor SEP (MSI)	4 factor SEP (MSI)	5 factor SEP (MSI)	6 factor SEP (MSI)	Quadratic SEP (MSI)
0.3	14	0.52	0.56	0.39	0.32	<b>0.12</b>	0.12	0.63
0.4	18	0.56	0.56	0.41	0.31	0.19	0.13	0.59
0.5	23	0.61	0.58	0.40	0.31	0.25	0.19	0.60
0.6	27	0.65	0.59	0.41	0.33	0.27	0.23	0.60
0.7	32	0.70	0.60	0.41	0.35	0.31	0.27	0.59
0.8	36	0.73	0.60	0.46	0.37	0.31	0.29	0.56
0.9	41	0.73	0.61	0.52	0.39	0.32	0.32	0.58

The LWR results in Table 3.20 for Sample Set #4 show little improvement over the PLS results. Blank entries in the table denote a poorly posed regression problem.

**Table 3.20.** Standard error of prediction results for loading weights, MSI, for Sample Set #4 using locally weighted regression as a function of model choice and fraction of the data ( $f$ ) used.

$f$	Pts.	1 factor SEP (MSI)	2 factor SEP (MSI)	3 factor SEP (MSI)	4 factor SEP (MSI)	5 factor SEP (MSI)	6 factor SEP (MSI)	Quadratic SEP (MSI)
0.3	14	0.54	0.33	0.31	0.24	0.23		0.58
0.4	18	0.52	0.40	0.32	0.25	0.20		0.54
0.5	23	0.53	0.43	0.35	0.26	0.22	0.19	0.37
0.6	27	0.55	0.43	0.35	0.28	0.24	0.17	0.40
0.7	32	0.56	0.43	0.36	0.28	0.26	0.16	0.38
0.8	36	0.57	0.43	0.36	0.30	0.30	0.15	0.37
0.9	41	0.58	0.44	0.37	0.34	0.34	0.15	0.38

For the low and nominally pigmented samples there was a reduction in model complexity using LWR with similar prediction results as those obtained using PLS. The high pigment samples showed similar LWR model complexity and prediction error as compared to the PLS model. It is possible that the regression surface for Sample Set #4 is much more complex than the other sample set regression surfaces and more samples would be needed to accurately represent a complex surface, as pointed out by Cleveland (11). Note that Sample Set #4 contains the fewest samples of all the sample sets.

In conclusion, the locally weighted regression approach provided for similar prediction error for Sample Set #1 even with the outliers included when compared with the PLS model with the outliers included. Naes et al. have found similar



behavior with the LWR approach accurately predicting samples that had been previously designated as outliers in a PCR analysis of near infrared data (12).

Whether or not LWR is a substitute for outlier deletion will probably depend on the particular data set and moreover on the experimental design. If an outlier is to be considered as a sample that introduces a significant nonlinearity into the data set, then rigorous sampling of the regression surface is imperative if that part of the regression surface is to be modeled properly. These considerations are beyond the scope of this research and as was suggested by Naes et al. the LWR approach should be tested more thoroughly to determine to which data sets it may be appropriately applied (12).

### 3.4 Conclusions

The ethylene and vinyl acetate co-polymer system was chosen due to the interest in process analysis applications. Concentration prediction errors were comparable to the precision from the NMR analysis method. Reference thickness values were estimated to have a precision of 0.03 mm and thus the SEP of 0.06 mm does not seem unreasonable, given that concentration and temperatures were also varying in the calibration set. Temperature prediction was not as good and it is believed this was due to the poor temperature reproducibility of the temperature controller. Nevertheless the challenge of prediction of composition of samples which are not simple mixtures of the components of interest, but the product of a polymerization reaction, has been met.

The number of factors used for the models may seem excessive. But it must be remembered that the uncorrected single beam spectral data were used, in which nonlinear variations are observed, forcing the model to use more linear factors than

one might expect from such a chemically simple system. The major problems arise from the lack of linearity of intensity due to multiple radiation transfer events, which are in turn dependent on sample temperature and geometry (thickness). As with any calibration method, care must be taken to see that the unknown sample is adequately represented in the calibration set.

The locally weighted regression approach provided a simpler calibration model with the same prediction error as found from a PLS analysis. This provides for the possibility of more easily interpreted models, but interpretation of multivariate calibration models should be treated with some caution (7).

Loading weight of a bonded polymer on a metal surface was predicted using IES and PLS to within the precision needed for on-line measurements, after suspected outliers were removed. Better results were found for the polymer sample set containing no pigment. Higher temperatures and observation angles closer to zero degrees from the norm of the sample plane gave smaller prediction errors. Locally weighted regression showed some reduction in model complexity for the same prediction errors as found using PLS.

### 3.5 Notes to Chapter 3

1. Callis, J. B.; Illman, D. L.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 624A-637A.
2. Griffiths, P. R. *Appl. Spectrosc.* **1972**, *26*, 73-76.
3. Hirschfeld, T. *Appl. Opt.* **1978**, *17*, 1400-1412.
4. Hvistendahl, J.; Rytter, E.; Oye, H. A. *Appl. Spectrosc.* **1983**, *37*, 182-187.
5. Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193-1202.
6. Erickson, B. C. "Multicomponent Flow Injection Analysis and Quantitative Infrared Emission Spectroscopy: Chemometric Applications" Ph. D. Dissertation, University of Washington: Seattle, WA (1988). Available from: University Microfilms, Ann Arbor, MI.
7. Seasholtz, M. B.; Kowalski, B. R. *Appl. Spectrosc.* **1990**, *44*, 1337-1347.
8. McDonald, R. S.; Wilks, J. P. A. *Appl. Spectrosc.* **1988**, *42*, 151-162.
9. Roberts, R. ALCOA, Private Communication, 1988.
10. Greenler, R. G. *Surface Science* **1977**, *69*, 647-652.
11. Cleveland, W. S.; Devlin, S. J. *J. American Statistical Association* **1988**, *83*, 596-610.
12. Naes, T.; Isaksson, T.; Kowalski, B. R. *Anal. Chem.* **1990**, *62*, 664-673.

## **Chapter 4 Polymer Cure Reaction Monitoring**

### **4.1 Introduction to Chapter 4**

The cure process of a polymeric coating on a metal surface is an obvious application for infrared emission spectroscopy. A thin organic film applied to a highly reflective surface provides good spectral contrast (1) and low background emission. Van Woerkom has reported monitoring the cure reaction of hexamethoxymethylmelamine with an esterdiol using infrared emission spectroscopy (2). He monitored the methoxy group using the  $915\text{ cm}^{-1}$  band and applied a univariate approach to the data analysis.

There has also been considerable interest in polyurethane cure monitoring (3). This chapter details an investigation into the use of infrared emission spectroscopy for noninvasive remote monitoring of the cure reaction of a commercial urethane paint product. Results from absorption spectroscopy data analysis are compared to emission spectroscopy results and the relationship of the two techniques is discussed. Multiresponse nonlinear curve fitting is used to model the absorption, raw emittance and a linearized emittance data set. Model parameters for the raw emittance data collected at different temperatures are compared.

### **4.2 Experimental and Data Analysis**

The materials used for these experiments consisted of a two component Du Pont urethane paint product, Centari<sup>®</sup>, product numbers 780s and 782s. In order to insure complete mixing, the manufactures' suggested proportions of each component were

mixed for 10 minutes using a magnetic stirrer. Initial experiments had indicated that insufficient mixing of the materials resulted in very poor reproducibility. The mixture was then spin coated onto a one inch aluminum disk for the emission experiments, or onto a KBr window for the absorption experiments. The aluminum disk was placed on a heated stage and the KBr disk was placed in a heated sample holder. The data collection was started immediately.

A Perkin-Elmer 1720 FT-IR spectrophotometer, modified by Perkin-Elmer for emission measurements, was used for all data collection. The heating unit for the emission experiments was an aluminum block, 9 cm by 9 cm, heated with two 200 watt cartridge heaters; 7.6 cm long and 0.6 cm in diameter, and controlled by an Omega controller. A temperature probe in the center of the heated block continually monitored the block temperature; this was fed back to the Omega controller for temperature control to  $\pm 1$  °C. Radiation from the emitting sample was collected at a distance of 30 inches from the spectrophotometer from a spot size of 16 mm using an optics extension. The heated stage for the absorption measurements was a standard Perkin-Elmer heating unit.

The collected radiation was modulated with an interferometer and the signal detected using a Judson narrow band mercury-cadmium-telluride (MCT) liquid nitrogen cooled detector for the emission measurements. A triglycine sulfate (TGS) detector was used for the absorption measurements. All emission spectra were the result of the coaddition of 10 single beam uncorrected spectra collected at 4  $\text{cm}^{-1}$  resolution requiring approximately 12 seconds to acquire. The ratio of the spectra with that from a blackbody spectrum at the same temperature was then computed. The blackbody consisted of a heated block of aluminum with a cone bored into it 80 mm deep with a 20 mm opening coated with soot. The emittance calculated for such a configuration is 0.999 (4).

Absorption spectra were collected in the scan mode providing data that were the ratio with a blank spectrum, in this case the blank was air. No spectral averaging was performed for the absorption data because one scan had a sufficient signal-to-noise ratio. Also because of the longer scan time using the TGS detector, significant changes in the reaction system would take place over the time for multiple scans, especially at the beginning of the reaction. Data were collected as a function of time using a Model 7700 Perkin-Elmer Professional computer running an "OBEY" program that records the time of the data collection as part of the header file saved to disk. The data were converted to the JCAMP file format (5), transmitted to a VAX computer system and analyzed with programs developed with the MATLAB (The MathWorks, Inc. Sherborn, MA) computing environment.

#### 4.2.1 Rank Determination

Determination of the pseudorank of a data matrix is one of the most fundamental questions in multivariate analysis. This information can give insight into the number of varying and distinguishable chemical components. Also, compression of data from numerical rank to pseudorank also offers a reduction in noise and results in a better posed problem. There are many methods to assess the pseudorank of a data matrix (6, 7, 8). The method of Lawton, Sylvestre and Maggio is particularly appealing because it is less arbitrary than others. This method involves the calculation of the following test statistic

$$\sigma^2 = \frac{\sum_{i=1}^M \sum_{j=1}^N (Y_{ij} - \hat{Y}_{ij})^2}{(M - K)(N - K)} \quad (4.1)$$

where  $Y_{ij}$  is the  $i,j$  entry in the data matrix and  $\hat{Y}_{ij}$  is the estimated  $i,j$  matrix entry using  $K$  singular vectors. The values of  $N$  and  $M$  are the number of rows and columns in the data matrix. The test statistic,  $\sigma^2$ , is the total sum of squares of the residuals divided by the number of degrees of freedom, adjusted for the number of factors used. This test statistic is compared with the estimated noise in the experiment. The number of singular vectors to include in the data compression step is that number which explains the data to the estimated noise level.

Because of the integer nature of matrix rank the test statistic from equation 4.1 will not agree exactly with the experimental error estimates when the correct model is found. It must be decided if it is best to remain above the noise level, choosing a less complex model, rank  $K$ , or select a more complex model, rank  $K+1$ , and thus explain the data to below the estimated noise level. In this chapter it has been decided that choosing a model that may include some noise is better than leaving useful information out, and thus a more complex model will be chosen.

Experimental error estimates were made as follows. A matrix of data, spectra as a function of time, was collected on a well cured polymer under the same experimental conditions as the analytical data, except that no reaction was taking place. The experimental noise level was then estimated by performing an ANOVA (analysis of variance) on the static data set over the same wavelength and time range as the analytical data. The different wavelengths are designated as the "treatments" and the within treatment variance computed and used as an experimental error estimate (9). The test statistic using equation 4.1 is computed for several different models (different number of factors or pseudoranks) and compared with the within treatment standard deviation estimate from the ANOVA calculation. Simulations were conducted on rank three data matrices which showed the method performed well up to about 5 % added noise. Added

noise in this case is a matrix of normally distributed random numbers with a standard deviation of a given percentage of the largest intensity in the data matrix. The noise estimates will have some variability and thus this procedure should be treated as a guide for rank selection and used in conjunction with visual inspection of the singular vector plots.

#### 4.2.2 Analysis of Closed Data

This section will review the special problems associated with closed data sets and the use of closed models in multiresponse nonlinear optimization. A familiar form of closure is the constant sum relationship found, for instance in mineral concentrations of geologic samples, in which the constituent concentrations sum to 100 %. In terms of the analysis of multiresponse data from indirect experimental measurements the concepts and effects of closure are somewhat more subtle.

The details of closure and its effect on the pseudorank of a data matrix are given in Appendix B. The following discussion concerns the proper modeling of multiresponse data that are suspected of having a closure relationship. Given a data matrix,  $\mathbf{R}$ , that is the product of two matrices,  $\mathbf{D}$  and  $\mathbf{F}$ , and further that the columns of  $\mathbf{D}$  are related by  $\mathbf{j}$ , a vector of linking coefficients (a closure relationship), such that  $\mathbf{Dj} = \mathbf{c}$ , where  $\mathbf{c}$  is a constant vector, then the matrix  $\mathbf{R}$  will have characteristics of the closure relationship in  $\mathbf{D}$  propagated through to it from the relationship  $\mathbf{R} = \mathbf{DF}^T$ .

In practice this situation might arise if a two step consecutive first-order reaction is monitored using a spectral measurement. The columns of  $\mathbf{D}$ , which are the concentration profiles, are constrained by a constant sum equal to the initial concentration of species A, and thus the matrix  $\mathbf{F}$  would represent the pure spectra. As shown in



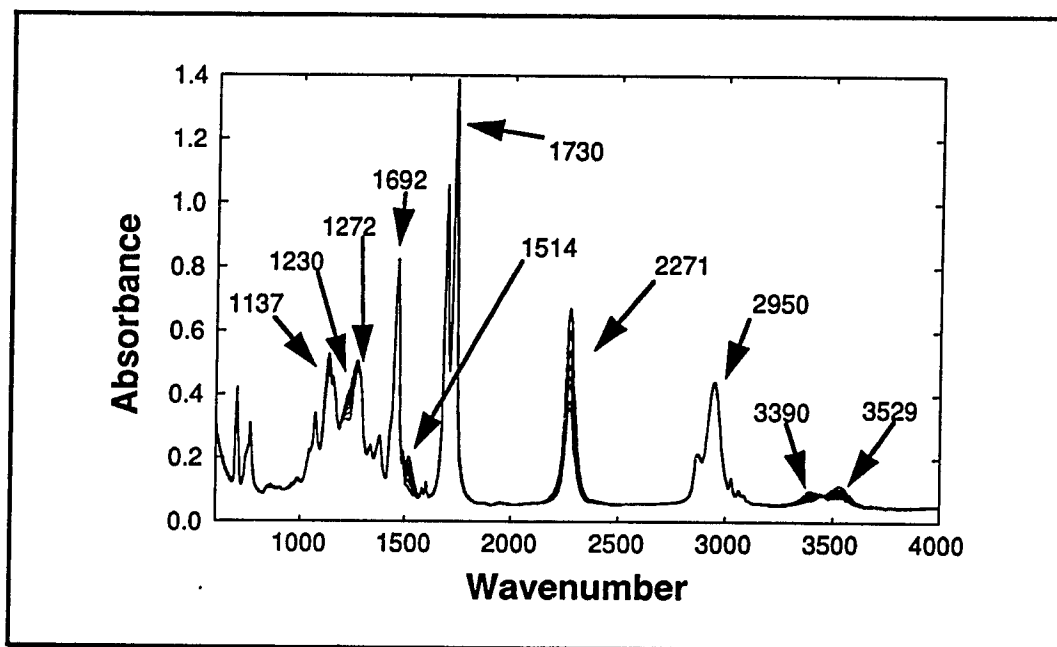
Appendix B this type of closure relationship is mathematically equivalent to the addition of a constant background spectrum to each row of  $\mathbf{R}$ .

The problem arises in the multiresponse analysis when choosing whether or not to model a baseline by the addition of a column of ones to the  $\mathbf{D}$  matrix. Addition of a column of ones will make the columns of the  $\mathbf{D}$  matrix linearly dependent if there already exists a relationship of the form  $\mathbf{D}\mathbf{j} = \mathbf{c}$ . The pseudoinverse may still be computed, but the rotation matrix,  $\hat{\mathbf{H}}$ , from equation 2.63 will have one column of zeros (in the noise free case). Because of this the reconstructed spectra found using equation 2.68 will be linear combinations of the three pure spectra and the background. The pure spectra, convolved with the background, may be recovered from these linear combinations by subtracting the reconstructed background from each of the reconstructed pure spectra.

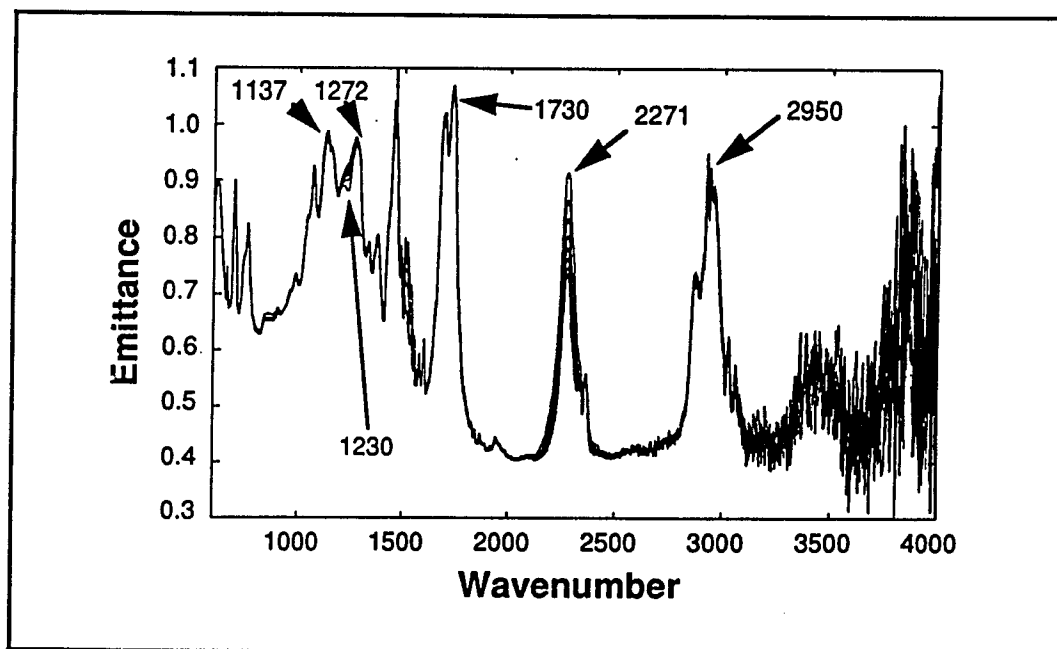
What this means in practice is that a simple background offset and the equivalent closure relationship may not be resolved from each other. For the analysis discussed in Section 4.3 no baseline modeling is attempted because of this closure constraint on the model.

### 4.3 Results and Discussion

Figures 4.1 and 4.2 display the infrared absorption and emittance spectra measured as a function of time while monitoring the reaction of the two component commercial paint product at 100 °C. The spectral bands noted in these figures are indicative of the formulations used for urethane coating preparation. The band assignments are summarized in Table 4.1.



**Figure 4.1** Infrared absorption spectra from 4000 to 600  $\text{cm}^{-1}$  monitoring the reaction of a two component paint mixture at 100 °C.



**Figure 4.2** Infrared emittance spectra from 4000 to 600  $\text{cm}^{-1}$  monitoring the reaction of a two component paint mixture at 100 °C.

**Table 4.1** Band assignments for spectra displayed in Figures 4.1 and 4.2.

Wavenumber	Behavior With Time	Assignment
1137	increasing	C - O - C <sup>a</sup>
1230	increasing	N - C = O Amide III <sup>a</sup>
1465	static	CH scissor <sup>b</sup>
1514	increasing	Amide II <sup>c</sup>
1692	static	carbonyl
1730	increasing	urethane carbonyl Amide I <sup>c</sup>
2271	decreasing	N=C=O isocyanate <sup>c</sup>
2800-2900	static	CH stretching <sup>b</sup>
3390	increasing	NH stretching <sup>b</sup>
3529	decreasing	OH stretching <sup>b</sup>

a) Ref. (10).

b) Ref. (11).

c) Ref. (12).

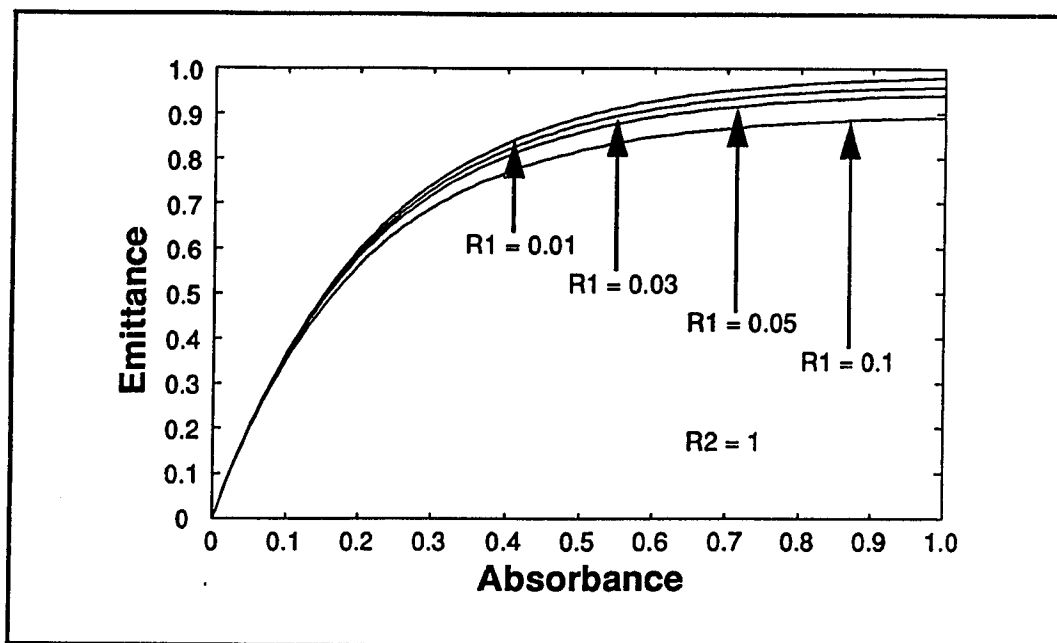
The formation of the urethane linkage is noted by an increase in the 1730 and 3390  $\text{cm}^{-1}$  bands (5.78 and 2.95  $\mu\text{m}$ ) corresponding to the amide I and amine functionality, as well as the decrease of the isocyanate and alcohol bands at 2271 and 3529  $\text{cm}^{-1}$  (4.4 and 2.83  $\mu\text{m}$ ), respectively.

Comparisons of the absorption spectra in Figure 4.1 with the emittance spectra in Figure 4.2 show a marked difference in the signal-to-noise ratio, especially above 2500  $\text{cm}^{-1}$  (below 4  $\mu\text{m}$ ). This is characteristic of the Planck blackbody function peaking in intensity at approximately 1000  $\text{cm}^{-1}$  (10  $\mu\text{m}$ ) at 100 °C and rapidly decreasing at higher spectral energies. In spite of this discrepancy in the signal-to-noise ratio there are still discernible changes in the emittance signal at the 2271  $\text{cm}^{-1}$  (4.4  $\mu\text{m}$ ), 1514  $\text{cm}^{-1}$  (6.6  $\mu\text{m}$ ), 1230  $\text{cm}^{-1}$  (8.1  $\mu\text{m}$ ) and 1137  $\text{cm}^{-1}$  (8.8  $\mu\text{m}$ ) as observed in the absorbance spectra.

The spectra from the emittance experiment appear to be much less well defined than from the absorption experiment. Less overall change in signal intensity for an equivalent time interval is also noted for the emittance data when compared to the absorbance data. This discrepancy between the absorption and emittance data sets is due to the nonlinear relationship between emittance intensity and concentration. This may be readily understood by referring to equation 1.15 and recalling  $A = -\log(T)$ , where  $T = e^{-\kappa cl}$ , and thus  $\kappa cl = A/\log(e)$  where  $\kappa$  is the extinction coefficient,  $c$  is concentration and  $l$  is path length. Note that  $A$  is absorbance and not absorptance as used in Chapter 1. Substituting into equation 1.15, the following relationship between absorbance and emittance may be derived

$$\epsilon = \frac{(1 - R_1)(1 - R_2 e^{-2A/\log(e)})}{1 - R_1 R_2 e^{-2A/\log(e)}} \quad (4.2)$$

Figure 4.3 displays the emittance calculated using equation 4.2 versus absorbance for a single wavelength with  $R_2=1$  and several values of  $R_1$ . Recall  $R_1$  is the front surface reflectivity and  $R_2$  is the reflectivity of the backplate. From Figure 4.3 it may be seen that the relationship between absorbance and emittance is nonlinear for experiments in which absorbance measurements exceed approximately 0.2 absorbance units. As the front surface reflectivity becomes larger the onset of the nonlinear behavior is found at smaller absorbance values. Since absorbance is linearly related to concentration (if Beer's law is followed) the above analysis indicates that the emittance response will be nonlinearly related to concentration at an optical density that would normally show linear behavior for absorbance measurements.



**Figure 4.3** Emittance versus absorbance computed from equation 4.2 for a single wavelength.

From Figure 4.1 it may be seen that for the isocyanate functionality the absorbance at  $2271\text{ cm}^{-1}$  ( $4.4\text{ }\mu\text{m}$ ) has a maximum of about 0.6 absorbance units. Because the absorbance and the emittance experiments were conducted under the same experimental conditions (same measured weight of material applied to the same area) this suggests that the emittance measurements for these experiments will be nonlinearly related to concentration, especially at the beginning of the reaction when the absorbance is high. If equation 4.2 is solved for absorbance in terms of emittance,

$$A = -\frac{1}{2} \log \left\{ \frac{\epsilon - 1 + R_1}{\epsilon R_1 R_2 - R_2 + R_1 R_2} \right\} \quad (4.3)$$

the linearization of the emittance data may be carried out. Equation 4.3 indicates that for the emittance data to be properly linearized  $R_1$  and  $R_2$  must be known. It may be assumed that the back surface reflectivity is relatively high and constant with wavelength, characteristic of a metal, but the front surface reflectivity will be a function of wavelength, showing a maximum near an absorbance maximum. In a curing process the concentration of the chemical components will also be changing with time and thus the front surface reflectivity may be a function of time as well as wavelength. An exact correction for this behavior would appear to be impossible without the corresponding absorbance data and thus an approximate linearization is suggested. If the front surface reflectivity is ignored, equation 4.3 simplifies to

$$A = -\frac{1}{2} \log \left\{ \frac{1 - \epsilon}{R_2} \right\} \quad (4.4)$$

and a linear relationship between absorbance and emittance is found if the logarithm of one minus the emittance is computed. Figure 4.4a displays the emittance versus the absorbance for the maximum of the isocyanate peak at 100 °C. As may be seen there is a distinct curvature to the plot as predicted by equation 4.2. Figure 4.4b displays the emittance data linearized as  $-\frac{1}{2} \log \left\{ \frac{1 - \epsilon}{R_2} \right\}$  versus the absorbance data. The transformed emittance data show a much more linear relationship with the absorbance data. The computed slope of the plot in Figure 4.4b is 0.6 indicating that  $R_1$  is not negligible (expected slope is one if  $R_1=0$ ). If the data are linearized with increasing levels of  $R_1$ , the slope of the linearized emittance versus absorbance plot approaches one at about  $R_1=0.07$ .

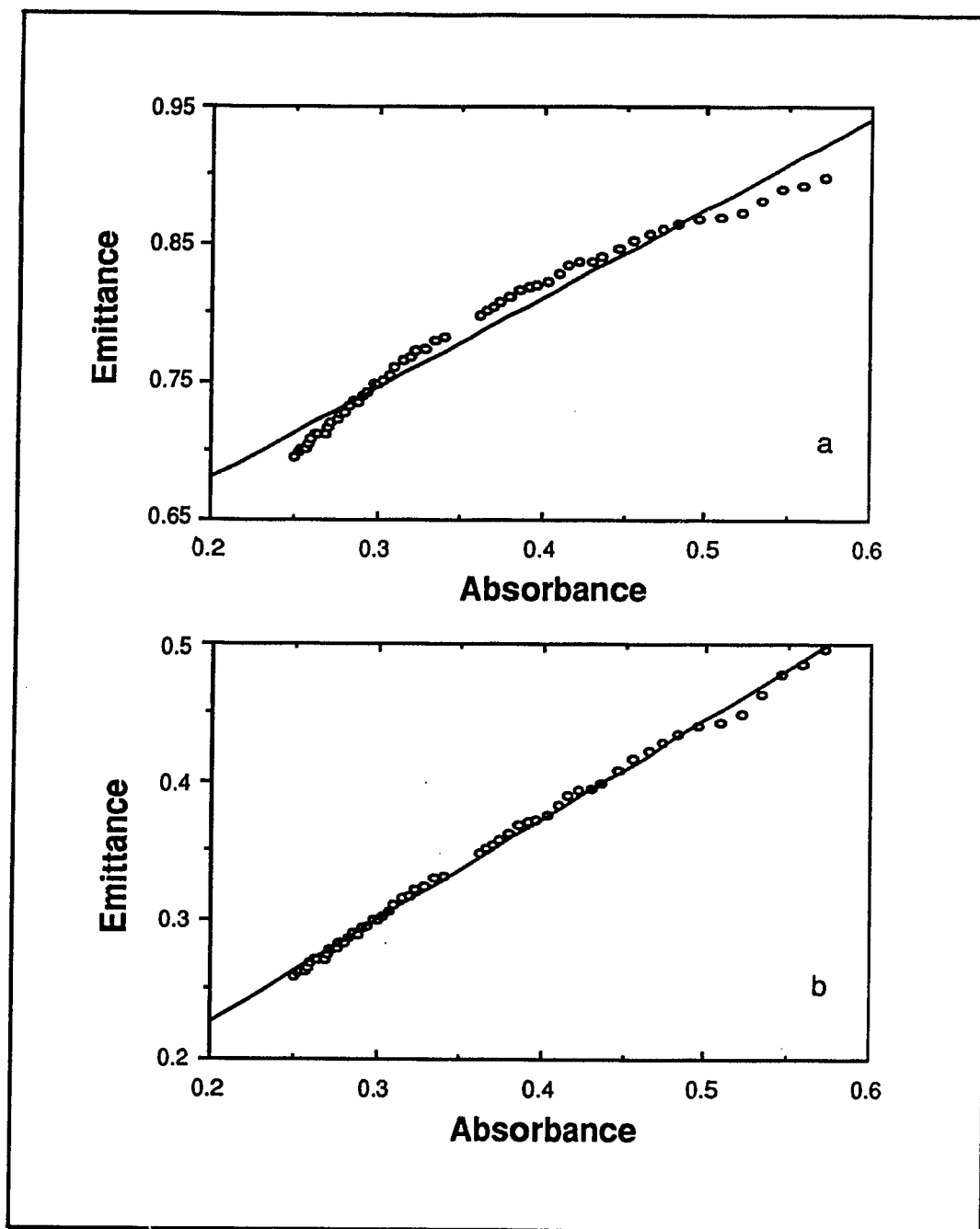
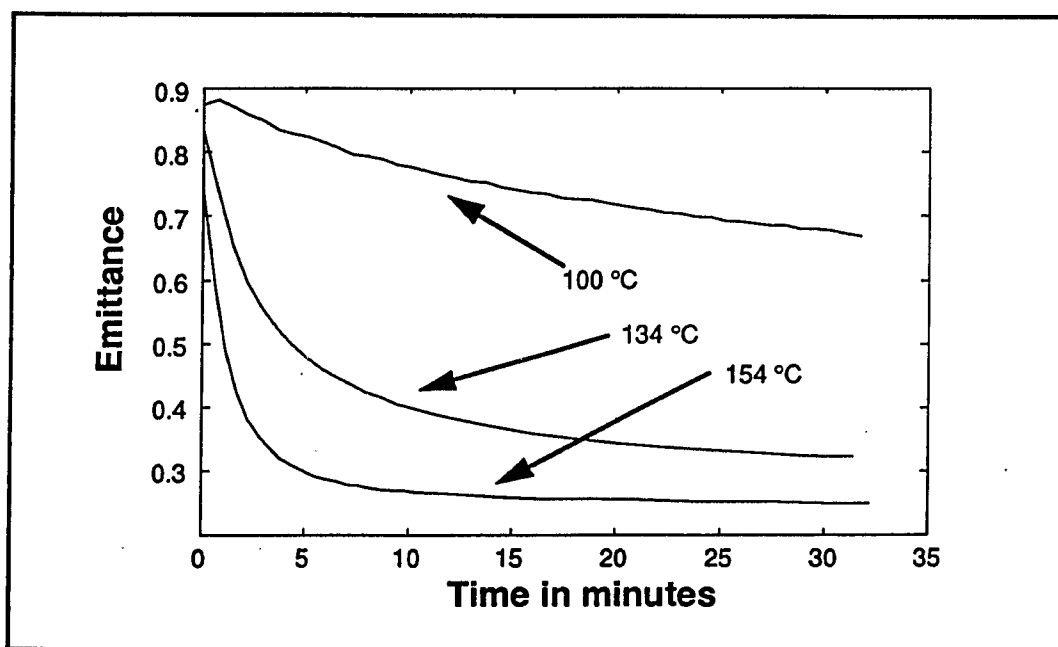


Figure 4.4 a) Raw emittance versus absorbance for 100 °C curing data and  
b) linearized emittance data versus absorbance for 100 °C data.

Other emittance data were collected at higher temperatures to observe the variation in the emittance intensities as a function of time under differing process conditions. Figure 4.5 displays the time profiles of the emittance response at  $2271\text{ cm}^{-1}$  ( $4.4\text{ }\mu\text{m}$ ) corresponding to the isocyanate functionality at process temperatures of 100, 134 and  $154\text{ }^{\circ}\text{C}$ . As may be seen an increase in temperature shows an increased rate of consumption of isocyanate demonstrating the ability of the emission measurement to distinguish changes in the process over time with changes in temperature.



**Figure 4.5** Time profiles for isocyanate emittance response at 100, 134 and  $154\text{ }^{\circ}\text{C}$ .

In order to identify other sources of variation in the spectra the standard deviation is computed at each wavenumber over time. Figure 4.6 displays the standard deviation spectrum for absorbance data measured at  $100\text{ }^{\circ}\text{C}$ . There are five regions of significant variation that will be considered from these data as indicated on the plot by the



superimposed circles. Those regions are 1) 3684 to 3204  $\text{cm}^{-1}$ , 2) 2404 to 2164  $\text{cm}^{-1}$ , 3) 1844 to 1604  $\text{cm}^{-1}$ , 4) 1604 to 1488  $\text{cm}^{-1}$ , and 5) 1284 to 1044  $\text{cm}^{-1}$ .

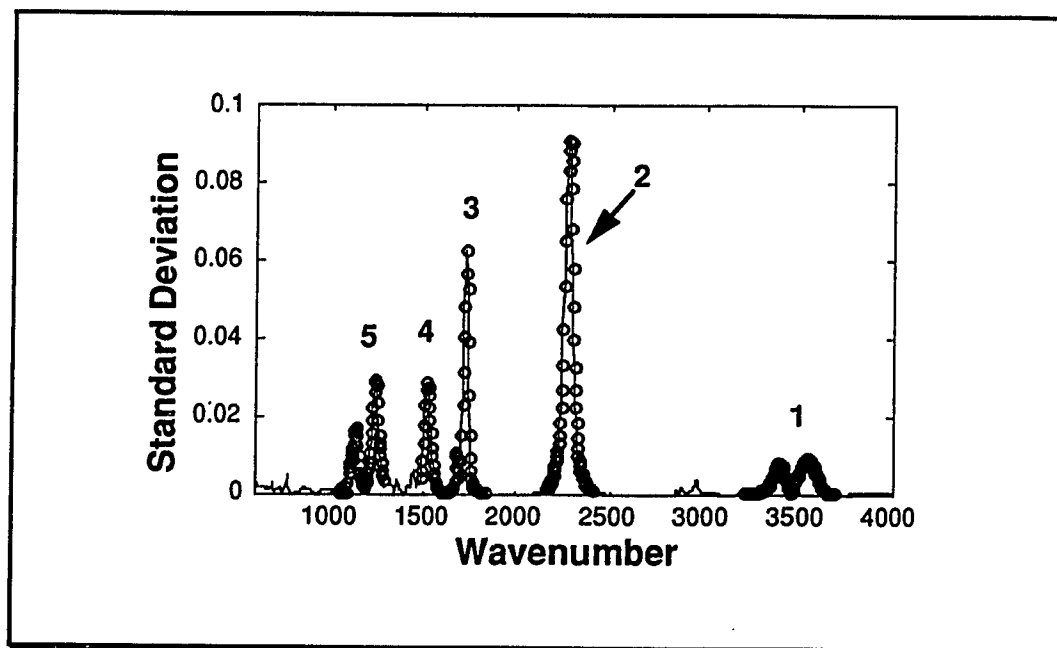


Figure 4.6 Standard deviation spectrum for 100 °C absorbance data.

Figure 4.7 displays the same plot for the emittance data at 100 °C with the corresponding areas of significant variance from the absorbance data denoted by the superimposed circles. As may be seen Region 1 from the emission experiment is of little use and will not be considered.

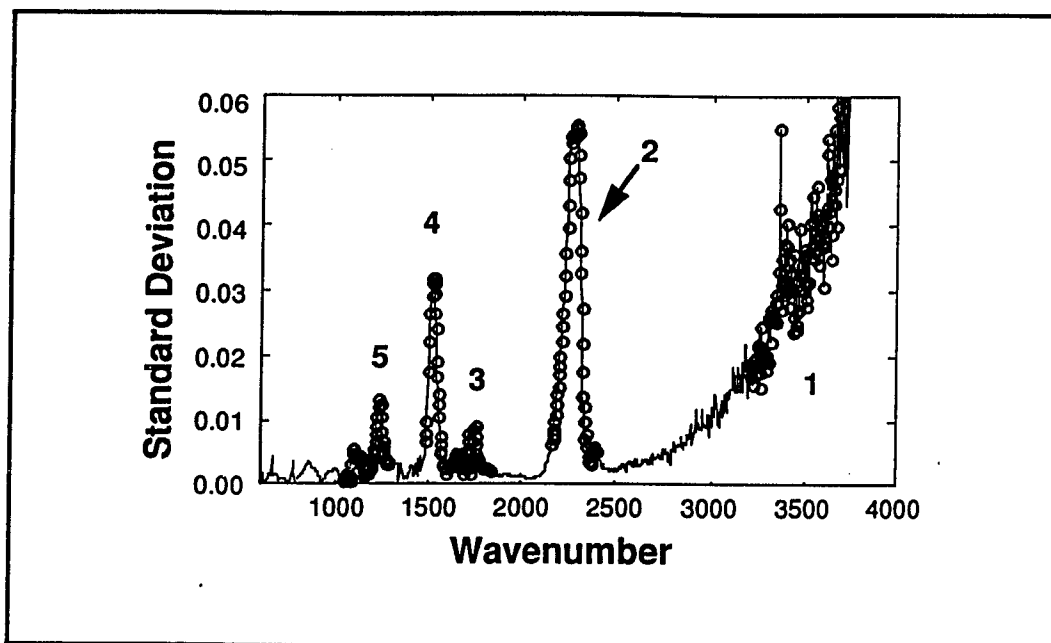


Figure 4.7 Standard deviation spectrum for 100 °C emittance data.

The isocyanate band has been used as a measure of reactivity for urethane formation reactions by many other workers (11, 13, 14, 15). Most of these studies have made use of only one or two spectral intensities or an integrated area of the isocyanate band and assumed a simple second-order reaction scheme to fit the data. Instead of pursuing a univariate analysis it was decided to explore the use of multivariate analysis methods, making use of all the sources of variance.

When using a multivariate measurement more information can be derived from the data than for a univariate measurement. One of the most informative pieces of information is the pseudorank of the data matrix. For absorbance measurements in the range of zero to one absorbance units, Beer's law should be satisfied and thus the pseudorank can be related to the number of chemical components in the mixture, provided there are no interactions. The relationship of the pseudorank of the emittance

data matrix and the number of chemical components is not as apparent, because of the nonlinearities, and will require some care in interpretation.

Rank analysis was initially performed on the isocyanate region of the absorbance data, Region 2, as this region was expected to give a simple interpretation and provide a basis of comparison for the emittance data analysis. The results for the rank determination step are summarized in Table 4.2, where a rank of three is indicated. Further rank analysis was performed on the data by successively augmenting other regions of the spectrum to Region 2. The results for the absorbance regions of interest, that is, Regions 2, 2345 and 12345 are given in Table 4.2. Table 4.3 gives the rank analysis results for the emittance data of Regions 2 and 2345.

**Table 4.2** Experimental noise estimation and residual values used for rank determination of 100 °C absorbance data.

Region	Noise estimate	Residual (Eq. 4.1 )	Rank
2	2.8 e-7	5.9 e-5	1
		1.3 e-6	2
		7.0 e-8	3
		4.9 e-8	4
		2.7 e-8	5
12345	1.8 e-6	4.0 e-4	1
		3.4 e-6	2
		6.3 e-7	3
		1.5 e-7	4
		1.0 e-7	5
2345	2.4 e-6	5.4 e-4	1
		7.3 e-7	2
		1.5 e-7	3
		1.0 e-7	4
		8.9 e-8	5

**Table 4.3** Experimental noise estimation and residual values used for rank determination of 100 °C emittance data.

Region	Noise estimate	Residual (Eq. 4.1)	Rank
2	6 e-6	2.6 e-4	1
		7.2 e-6	2
		4.2 e-6	3
		3.5 e-6	4
		3.2 e-6	5
2345	3 e-6	4.1 e-4	1
		6.8 e-6	2
		2.3 e-6	3
		1.5 e-6	4
		1.2 e-6	5

The analysis for Region 2345 for the absorbance data suggests a rank of 2, but a visual examination of the singular vector plots indicates that the third singular vector is clearly non-random indicating a rank of three. This rank determination method is not foolproof and must be used in conjunction with a visual assessment of the singular vectors. In each case the singular vectors clearly show non-random behavior through the third singular vector.

The need for three singular vectors to explain the absorbance data is consistent with the known chemistry of these materials. The system proceeds through three distinct physical states during the course of the reaction (16). The first physical state is a liquid, the second state is a gel and the final state is a crosslinked polymer. Further, it is known that approximately one third of the isocyanate is consumed at the gel point. If these physical states are distinct, then it would be expected that the functional groups on the

molecule would be presented with distinct environments which would in turn affect the spectral characteristics. It follows from the principal component analysis and the known behavior of the system that the pseudorank is revealing the three physical states. Feger has presented the description of an ideal isocyanate and alcohol system in terms of pregel and postgel periods, but applies only a univariate data analysis (13).

The same rank analysis was performed for the 100 °C emittance data and a rank of three was determined. A graphical representation of the singular vectors from the emittance data is shown in Figure 4.8. Note that singular vector four is offset for visual clarity. This clearly shows the non-random behavior of the third singular vector indicating a rank of three. Figure 4.9 displays the corresponding spectral singular vectors for these data. These vectors are offset for visual clarity and show non-random behavior through vector three with some slight non-random behavior in vector four which may be associated with carbon dioxide in the beam path. Considering the nonlinear concentration/response relationship of the emittance data the rank was expected to be somewhat higher than the corresponding absorbance data. It is not, and this suggests that not as many features are being detected in the emittance experiment as compared with the absorbance experiment and it may be that the nonlinearity is inflating the rank.

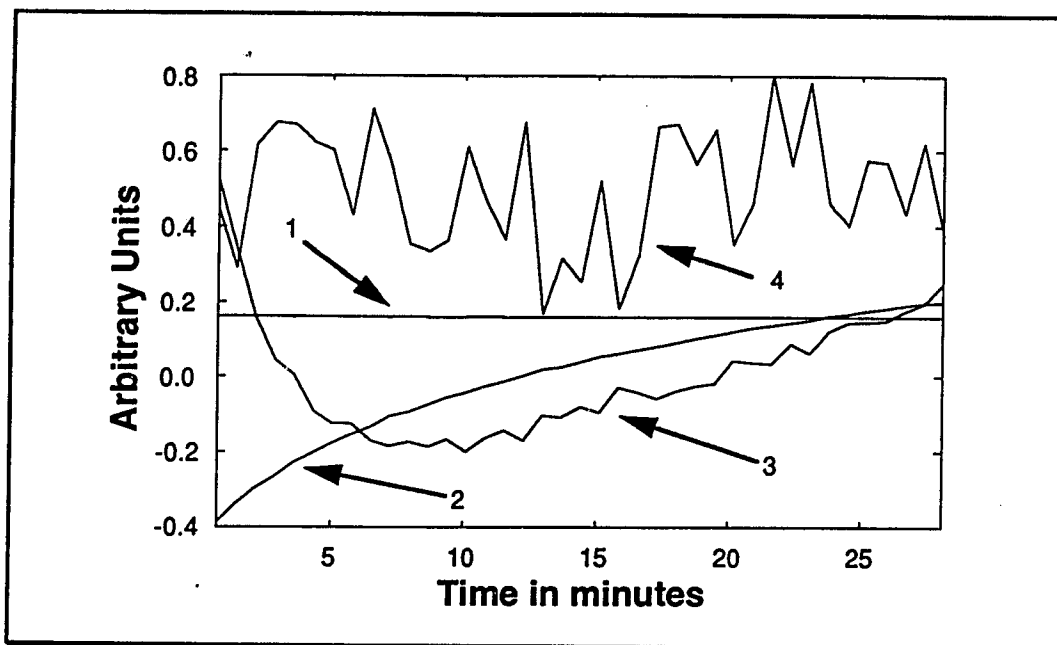


Figure 4.8 Time varying singular vectors one to four for 100 °C emittance data.

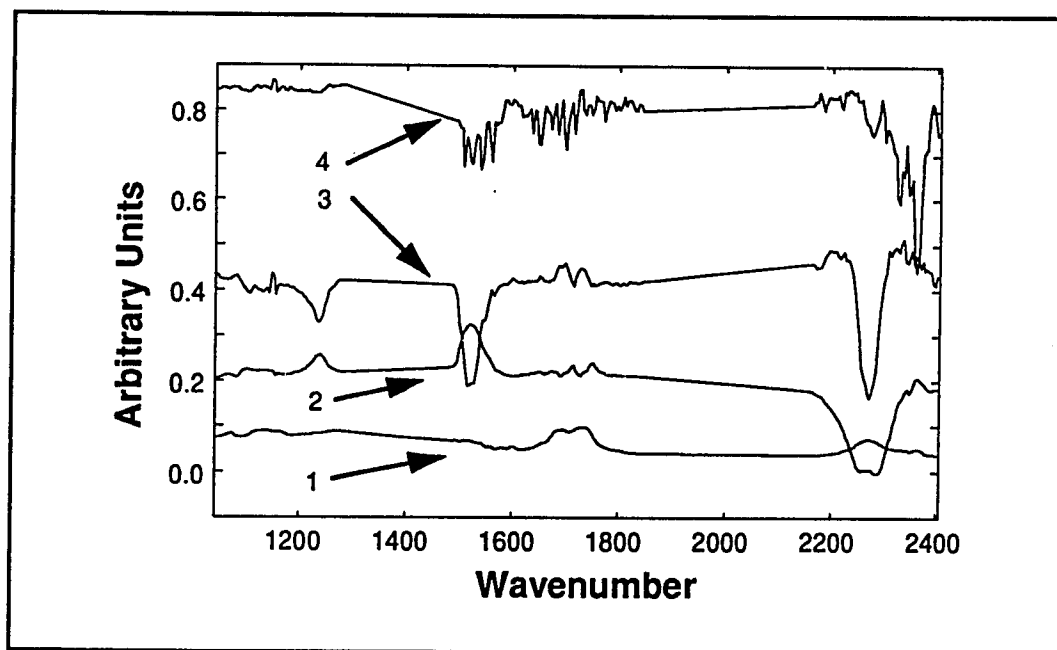
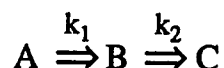


Figure 4.9 Spectral singular vectors corresponding to time singular vectors displayed in Figure 4.8.

Given an estimate of the pseudorank of the data matrix and prior information concerning the reaction course (i.e., the change of state) a model describing the time course of the reaction may be postulated. A simple scheme consistent with this information is



in which the different states A, B and C are associated with the different physical states by the model parameters  $k_1$  and  $k_2$ .

Given a model for the time behavior, multivariate nonlinear optimization was used to estimate the model parameters. A model was built for the absorption data, including regions one through five, giving 334 spectral responses and a time range of 1.8 to 22 minutes yielding 35 time points. The hydrocarbon band at  $2950\text{ cm}^{-1}$ , which should remain unchanged during the course of the reaction, was observed to decrease rapidly in intensity at time point one and leveled off to a constant by time point number five. Thus the data analysis used information beginning with time point five. The nonlinear model parameters and associated parameter error estimates are summarized in Table 4.4.

**Table 4.4** Nonlinear model parameter estimates for absorption, emittance and linearized emittance data at 100 °C.

Data set	R <sub>1</sub>	R <sub>2</sub>	k <sub>1</sub>	k <sub>2</sub>
Absorption	--	--	0.80 +/- 0.04	0.08 +/- 0.01
Emittance	--	--	0.24 +/- 0.03	0.03 +/- 0.01
Emittance (linearized *)	0	1	1.4 +/- 0.5	0.10 +/- 0.02
Emittance (linearized)	0.07	1	1.2 +/- 0.4	0.08 +/- 0.01

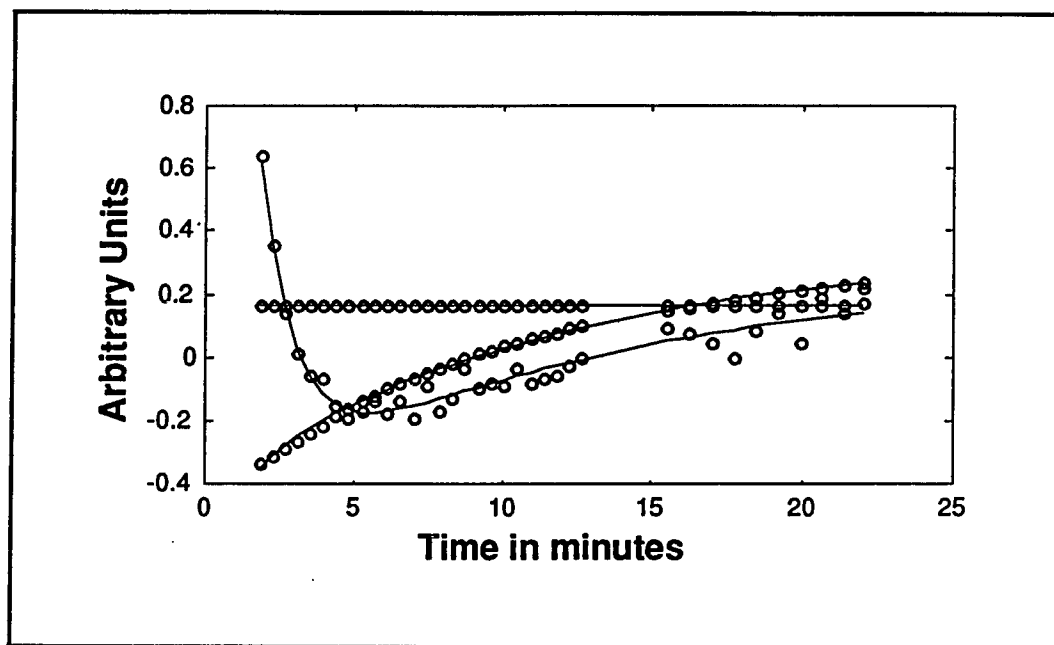
\* k<sub>1</sub> and k<sub>2</sub> are reported after removal of two outliers.

These model parameters and error estimates are computed as the mean and standard deviation from Monte Carlo calculations. For each Monte Carlo calculation the nonlinear optimization was performed after the addition of a matrix of normally distributed random numbers of which the standard deviation was scaled to the experimental error estimate. For each parameter and error estimate 50 Monte Carlo calculations were performed. The estimated parameters are consistent with a rapid change from a liquid state to a gel state and a slower reaction to the final crosslinked state.

Also note that a delay time parameter is used in the data fitting in order to adjust the isocyanate pure component band ratio for A and B to 3:2. This is consistent with the consumption of one third of the isocyanate at the gel point. The delay parameter does not affect the fit of the data to this model, only the ratio of the reconstructed spectra. The delay parameter was determined in a separate optimization, with k<sub>1</sub> and k<sub>2</sub> fixed, that gave a ratio of A to B of 3:2. Then the delay parameter was fixed for the Monte Carlo error estimate calculations for k<sub>1</sub> and k<sub>2</sub>. The invariance of the fit of the data to the model with changes in the delay parameter was an empirical observation using synthesized data.



From the multivariate analysis several plots are available to help assess the quality and reasonableness of the fit. Figure 4.10 displays the fit of the rotated model to the first three absorbance data singular vectors. The fit is generally quite good with only slight systematic deviations noted in the second singular vector in the early part of the reaction. Figure 4.11 shows the fit in the more familiar concentration and time domain.



**Figure 4.10** Fit of the model to the first three singular vectors for the absorbance data.

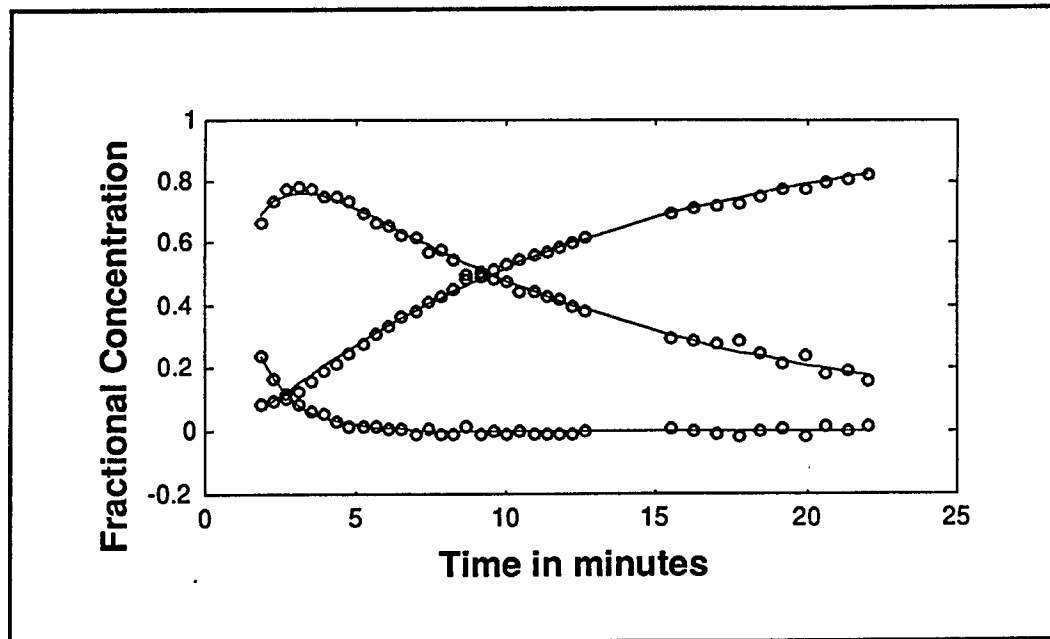


Figure 4.11 Fit of the absorption data in concentration and time space.

Figures 4.12 through 4.16 display the reconstructed pure component spectra for each of the five regions. The error bands indicate the extremes of the reconstruction from all possible combinations of parameters and associated standard deviations. For example there are three ways to associate the standard deviation with the parameter estimates, either add it, subtract it or add nothing. Thus for a two parameter problem there are nine ways to incorporate the error information. The nine spectra for each 'pure component' reflect the nine combinations of parameters and associated standard deviations. Figure 4.12 displays the alcohol and amine region and as expected the 'A' state contains mostly alcohol while the 'B' state shows less alcohol and more of the amine functionality. The 'C' state is dominated by mostly the amine functional group. Figure 4.13 displays the reconstructed spectra of the isocyanate region where the ratio of 'A' to 'B' has been adjusted to 3:2 by a delay time parameter of 0.9 minutes. The 'C' spectrum shows even

less intensity which is of course consistent with the consumption of isocyanate as the reaction proceeds from the gelled to the crosslinked state. Each of the other regions show a consistent behavior with respect to what is known about the chemistry of the system except for the urethane carbonyl region in Figure 4.14. Here there appears to be somewhat less of a clear distinction between the 'A' and 'B' states with 'B' of a similar intensity as 'A', but slightly broadened. This is most likely a result of the shifting of the urethane carbonyl band with the changes of state (17).

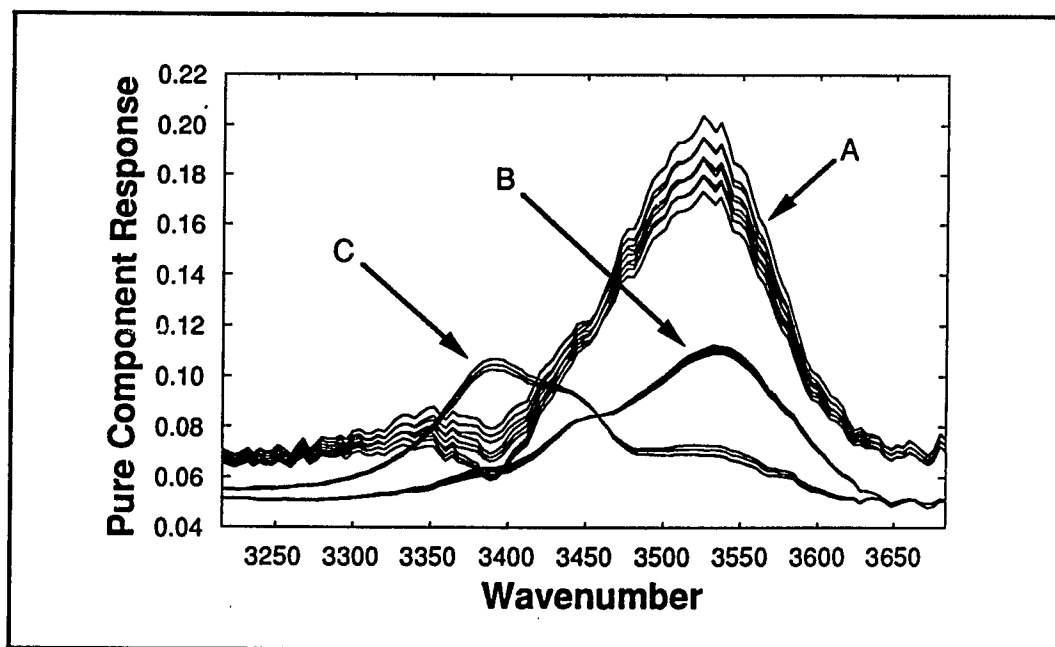


Figure 4.12 Reconstructed "pure component" spectra for absorbance data in the 3684 to 3216  $\text{cm}^{-1}$  region, the alcohol and amine region.

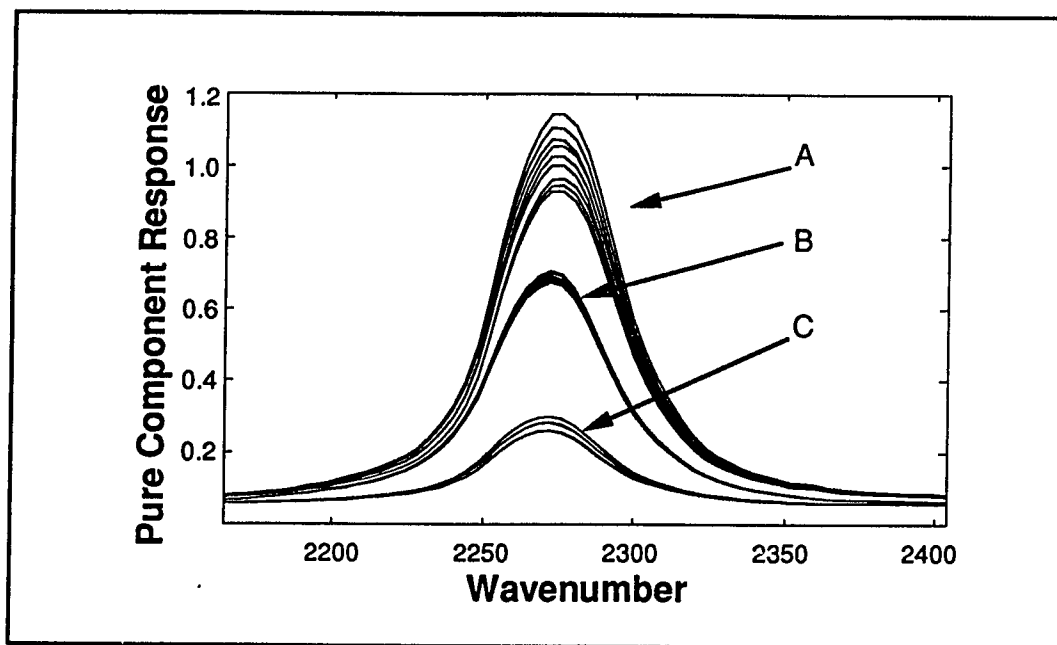


Figure 4.13 Reconstructed "pure component" spectra for absorbance data in the 2404 to 2164  $\text{cm}^{-1}$  region, the isocyanate region.

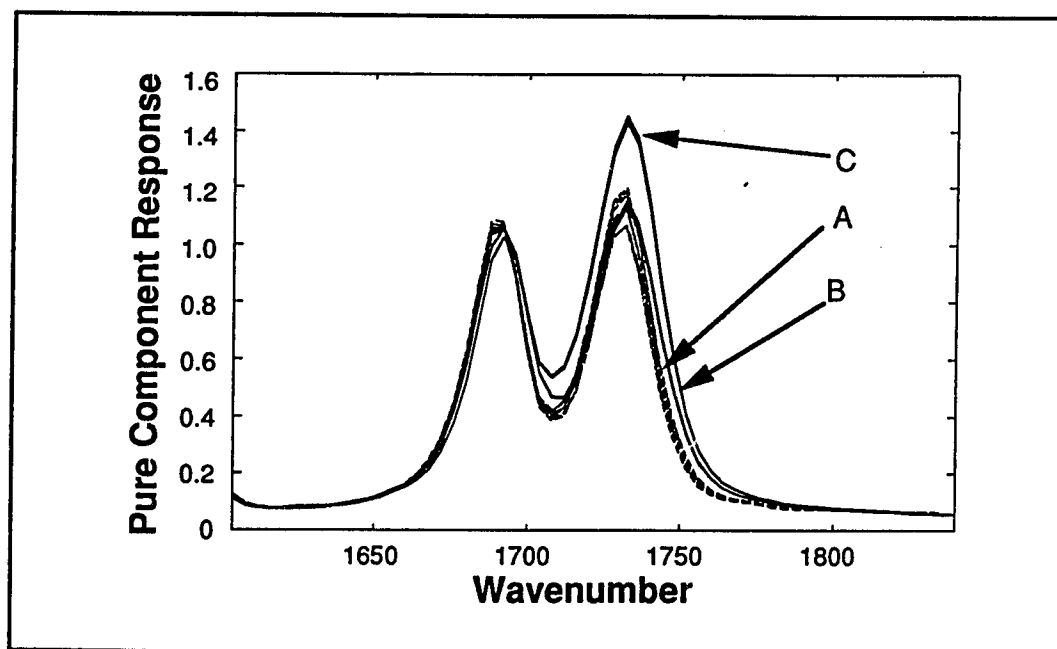


Figure 4.14 Reconstructed "pure component" spectra for absorbance data in the 1840 to 1604  $\text{cm}^{-1}$  region, the urethane carbonyl region.

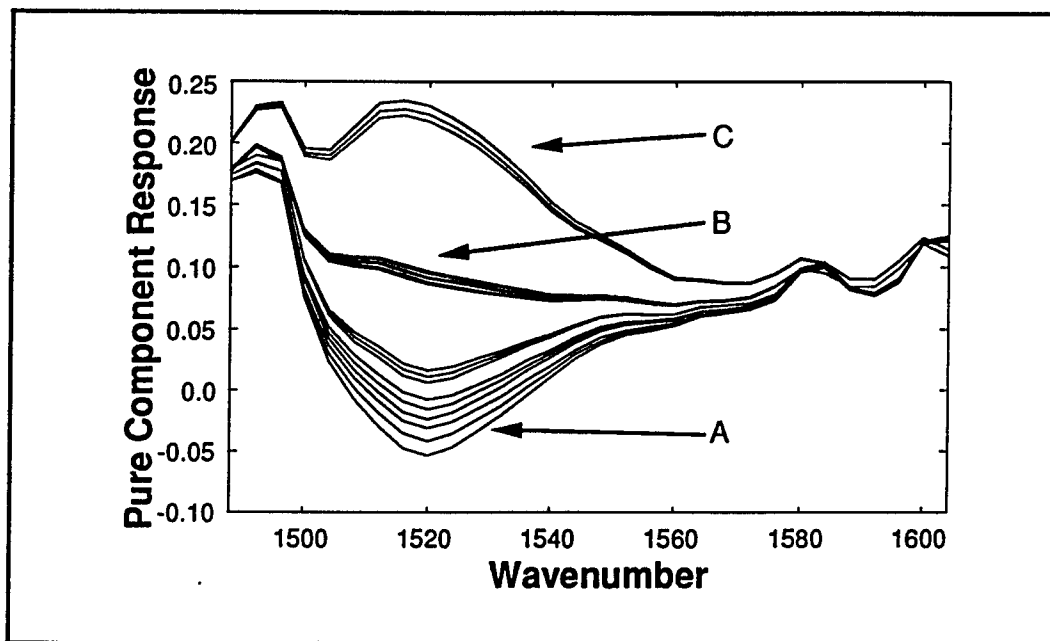


Figure 4.15 Reconstructed "pure component" spectra for absorbance data in the 1604 to 1488  $\text{cm}^{-1}$  region, the Amide II region.

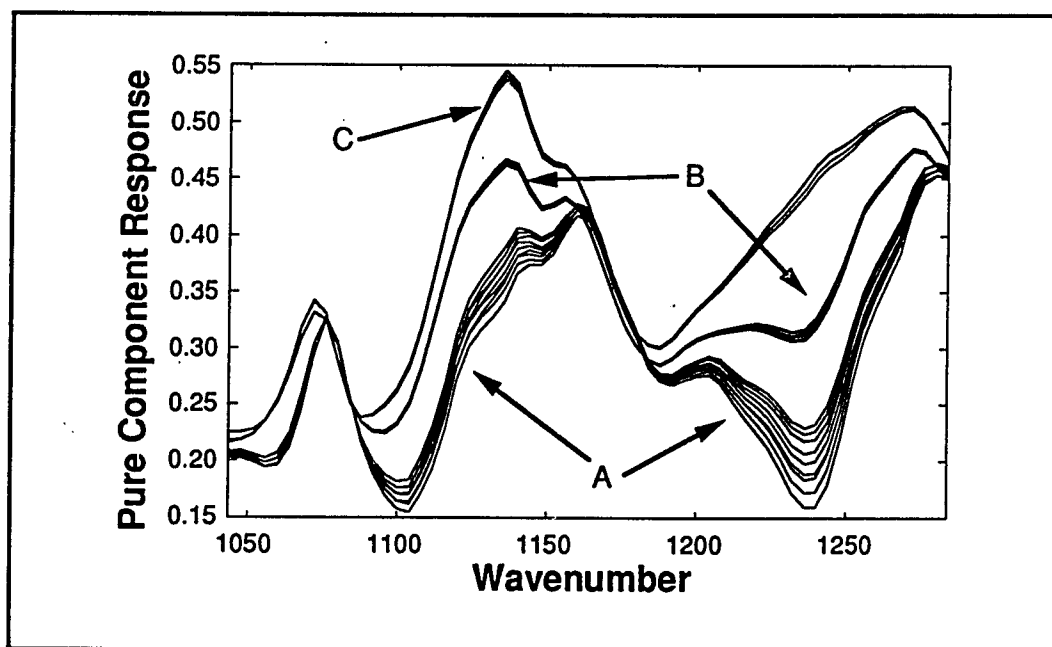


Figure 4.16 Reconstructed "pure component" spectra for absorbance data in the 1284 to 1044  $\text{cm}^{-1}$  region, the Amide III and C-O-C region.

The reconstructed spectra may also be used as a guide for the nonlinear optimization, in that physically meaningful results require that the reconstructed spectra have all positive intensities. This is particularly important for the above postulated model because this model has two equally valid minima on a residual error surface. These minima correspond to the same model parameters, which have simply been exchanged with each other. This feature of the model has been observed by others and is known as the slow-fast ambiguity (18). For the above example, if the model parameters are exchanged the spectrum for component B in Region 1 and 2 will show negative features indicating physically meaningless results. Band inversions will not always occur upon insertion of the parameters in the incorrect order, but there may be enough of a difference in the reconstructions from what is expected that an analyst would be alerted to the problem.

In order to investigate the emittance data, only Region 2 was considered. Region one is clearly dominated by noise as shown in Figure 4.7 and in anticipation of the nonlinear emittance response relationship it was felt that the approximate linearization would be most plausible in a restricted spectral region. The same multiresponse nonlinear analysis was performed on the raw emittance data. The Monte Carlo model parameters and associated error estimates are displayed in Table 4.4. The fit of the model to the raw data is displayed in Figure 4.17 and appears to be very good. There is poor correspondence between the estimated model parameters for the 100 °C absorption and the corresponding emittance data, as might be expected considering the nonlinear emittance and concentration relationship. A simple linearization of the data was attempted using equation 4.4. The model parameters using the linearized data given in Table 4.4 are shown to correspond more closely to the absorbance data, but only because of larger errors associated with them. This decrease in the precision of the parameter

estimates is the result of either noise introduction when performing the linearization step or a loss of rank via linearization. As the  $R_1$  parameter is changed from 0 to 0.08, the third singular vector from the singular value decomposition of the linearized emittance data is characterized by less random behavior as shown in Figure 4.18. Recall that univariate manipulation of the emittance data indicated that an  $R_1$  value of 0.07 gave a slope of nearly one for the absorbance versus linearized emittance. The parameter estimates with  $R_1$  equal to 0.07 are also displayed and show better correspondence with the absorbance data with some reduction in the error estimates.

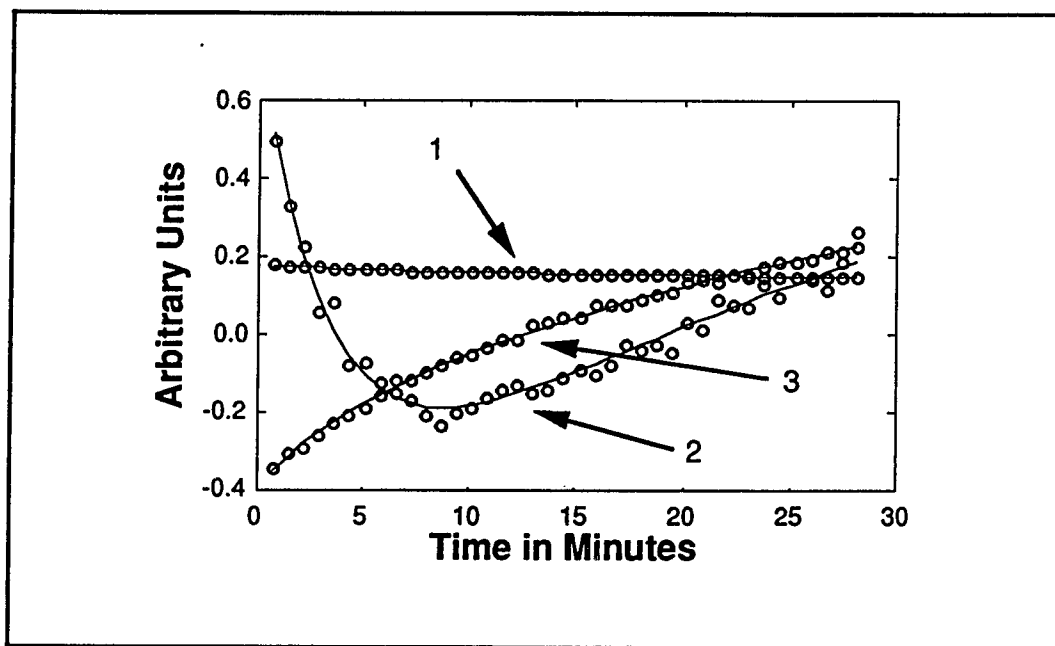
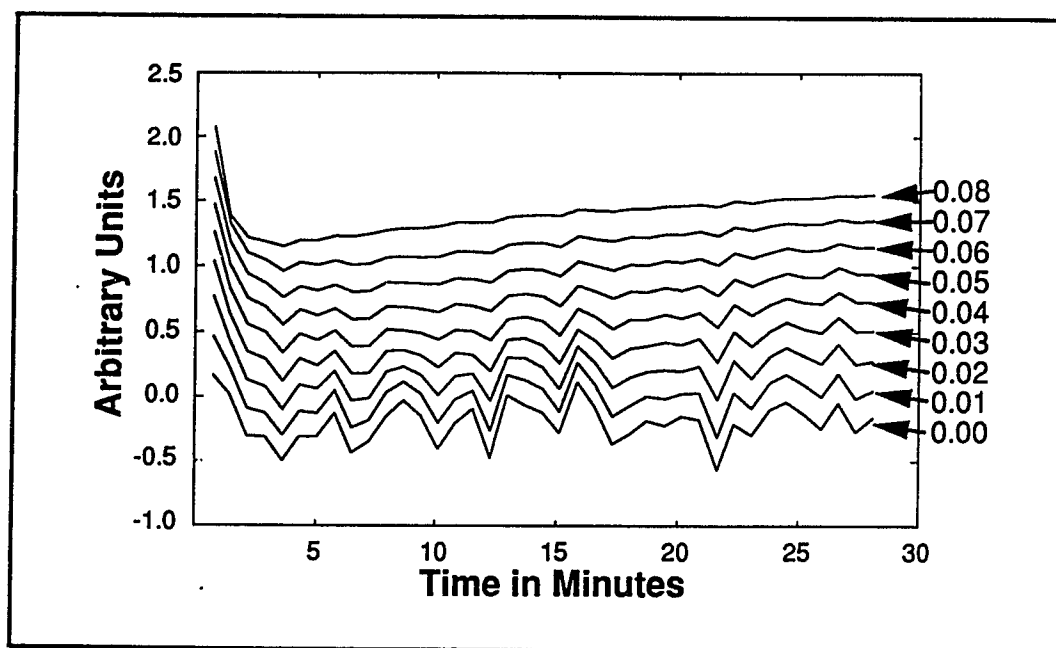


Figure 4.17 Fit of model to raw emittance data at 100 °C.



**Figure 4.18** Change in third singular vector for emittance data as a function of  $R_1$  linearization parameter.

Even though the emittance data in the raw form do not give readily interpretable model parameters, as might be found from an absorbance experiment, and linearization introduces large uncertainties into the parameter estimates, the model parameters from the raw emittance data may still be useful as a method of comparing systems and summarizing data. Table 4.5 displays the model parameters computed from the raw emittance data for replicate experiments at 100 °C and 134 °C and a single experiment at 154 °C.



**Table 4.5.** Nonlinear model parameter estimates for raw emittance data at three temperatures.

DATA	$k_1$ (min <sup>-1</sup> )	$k_2$ (min <sup>-1</sup> )
Emittance (raw 100 °C)	0.24 +/- 0.03	0.03 +/- 0.01
Emittance (raw 100 °C)	0.19 +/- 0.03	0.05 +/- 0.03
Emittance (raw 100 °C)	0.22 +/- 0.02	0.04 +/- 0.01
Emittance (raw 134 °C)	0.76 +/- 0.06	0.09 +/- 0.01
Emittance (raw 134 °C)	0.66 +/- 0.06	0.08 +/- 0.01
Emittance (raw 134 °C)	0.68 +/- 0.04	0.08 +/- 0.01
Emittance (raw 154 °C)	2.2 +/- 0.8	0.4 +/- 0.1

As may be seen the model parameters from the replicate analyses are within the estimated error of each analysis, indicating the reproducible nature of the experiments and supporting the validity of the error estimates. The  $k_1$  model parameter appears to be the most affected by temperature, while the  $k_2$  parameter remains fairly insensitive to changes in temperature until 154 °C. These parameters may give useful insight into the reacting system or at the least give a convenient method of summarizing the observations.

#### 4.4 Conclusions

Infrared absorption and emittance data have been collected from a reacting polymer system. It is found that a univariate inspection of the isocyanate band shows distinct differences in the time profiles for different temperature experiments. A comparison of the emittance data with the absorption data indicates that the emittance data require a preprocessing linearization step before results comparable to the absorbance data may be found. The standard deviation spectrum of the data over time indicates five regions of variance. Rank analysis of the absorption data indicates three

singular vectors are needed to describe the observed spectral/temporal variance. The three singular vectors are interpreted as revealing the three physical states of the reacting polymer system.

Multivariate nonlinear curve fitting is applied to the 100 °C absorption and emittance data yielding quite different nonlinear model parameter estimates due to the nonlinear concentration/response relationship for the emittance data. A simple linearization applied to the emittance data shows improved comparability with the absorption data, but at the expense of poorer model parameter precision. Modeling of the raw emittance data for several tests at the same temperature indicates that the procedure is reproducible and shows changes in model parameters for experiments at different temperatures. Although the computed parameters from the raw emittance data may not be as readily interpretable as those from an absorption experiment they may still be useful as a basis for comparison of experiments.

## 4.5 Notes to Chapter 4

1. Van Kasteren, P. H. G.; Smeets, L. H. *SPIE - Fourier Transform Infrared Spectroscopy* **1981**, 289, 83-86.
2. Van Woerkom, P. C. M.; de Groot, R. L. *Appl. Opt.* **1982**, 21, 3114-3118.
3. Provder, T. J. *Coatings Technology* **1989**, 61, 33.
4. Stewart, J. E. *Infrared Spectroscopy; Experimental Methods and Techniques*; M. Dekker: New York, 1970.
5. McDonald, R. S.; Wilks, J. P. A. *Appl. Spectrosc.* **1988**, 42, 151-162.
6. Sylvestre, E. A.; Lawton, W. H.; Maggio, M. S. *Technometrics* **1974**, 16, 353-368.
7. Shrager, R. I. *Chemo. Intell. Lab. Sys.* **1986**, 1, 59-70.
8. Tu, X. M.; Burdick, D. S.; Millican, D. W.; McGown, L. B. *Anal. Chem.* **1989**, 61, 2219-2224.
9. Box, G. E. P.; Hunter, W. G.; Hunter, J. S. *Statistics For Experimenters*; John Wiley & Sons Inc.: New York, 1978.
10. Alpert, N. L.; Keiser, W. E.; Szymanski, H. A. *IR Theory and Practice of Infrared Spectroscopy*; Plenum Publishing Corporation: New York, 1964.
11. Bauer, D. R.; Dickie, R. A. *Ind. Eng. Chem. Prod. Res. Dev.* **1986**, 25, 289-296.
12. Chadwick, D. H.; Cleveland, T. H. In *Encyclopedia of Chemical Technology* Third Edition Vol. 13; Grayson, M. Ed., John Wiley & Sons: New York; p 810.
13. Feger, C.; Molis, S. E.; Hsu, S. L.; MacKnight, W. J. *Macromolecules* **1984**, 17, 1830-1834.
14. Vander Ven, L. G. J.; Van Dijk, J. H. *ACS Polym. Mat. Sci. and Eng.* **1988**, 58, 267.
15. Yang, Y. S.; Lee, L. J. *J. Appl. Polym. Sci.* **1988**, 36, 1325.
16. Hartshorn, J. H. , Du Pont Marshall Labs., *Personal communication* **1989**.
17. Bellamy, L. J. *The Infrared Spectra of Complex Molecules: 2nd ed.*; Chapman and Hall Ltd.: New York, 1980.
18. Alcock, N. W.; Benton, D. J.; Moore, P. *Trans. Farad. Soc.* **1970**, 66, 2210-2213.

## **Chapter 5 Multiresponse Fitting Criteria Tests**

### **5.1 Introduction to Chapter 5**

Multivariate measurements are common place in many research laboratories and there is a realization that more information is made available using multivariate mathematical and statistical techniques to analyze these data (1). A number of methods are available to extract useful information from multivariate data and some of these techniques were discussed in Chapter 2. This chapter is concerned with the proper choice of objective function when performing multiresponse nonlinear optimization analysis. As explained in Chapter 2 the multiresponse nonlinear optimization problem may be analyzed by discussing two data types: direct observation (full rank) data and indirect observation (rank deficient) data.

Statistical considerations for properly combining multiresponse data for nonlinear modeling when there is a direct observation of the phenomena of interest and the data are full rank have been previously presented by Box and Draper (2). They employed a Bayesian approach and found a generally applicable objective function to be the determinant of the dispersion of the residual matrix rather than the more commonly used total sum of squares of the residual matrix.

Recently the analysis of this type of problem, in which nonlinear parameters are found via an optimization algorithm for rank deficient data, has been explored by Shrager and Hendler (3) and Frans and Harris (4,5). Shrager has also recently summarized the analysis of this type of problem and demonstrated the need for proper statistical analysis of the data (6). He finds improved parameter precision using a

weighted total sum of squares (TSSW) objective function compared with a simple total sum of squares (TSS) objective function.

Use of the determinant criterion (DET) for multiresponse nonlinear analysis has focused on full rank, overdetermined data matrices (7). Multiresponse nonlinear analysis of rank deficient, underdetermined data matrices has employed a total sum of squares (4,5,8) or a weighted total sum of squares objective function (6). This apparent disjoint treatment of multiresponse nonlinear parameter estimation has motivated the current investigation. It is the intention of this chapter to examine both the full rank and rank deficient multiresponse parameter estimation problems. A specific goal is to compare and contrast different objective functions and recommend the optimal criterion.

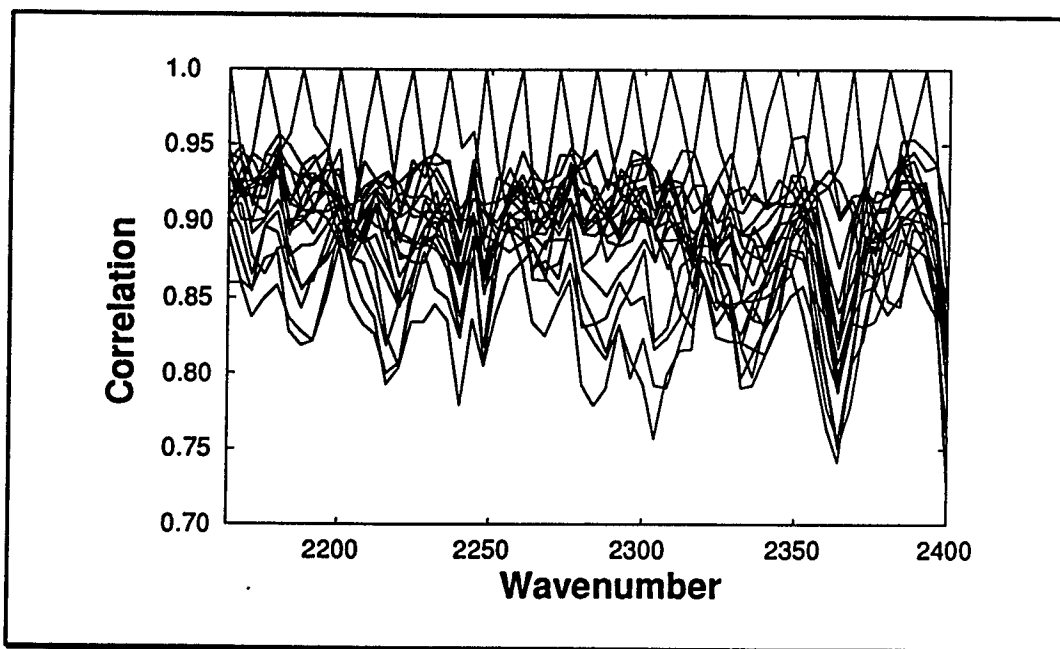
## 5.2 Experimental and Data Analysis

A general description of the experimental conditions common to all experiments will be given below, with the specifics for each unique experiment to be discussed in Section 5.3. All experiments were conducted using synthetic data. The generation of added noise for each experiment was accomplished using the RAND function from MATLAB (The MathWorks, Inc. Sherborn, MA) which generates normally distributed numbers with zero mean and variance of one. Signal-to-noise was computed as a ratio of the largest entry in the data matrix to the standard deviation of the added noise.

The noise structure was controlled in the following manner. For a given response matrix,  $\mathbf{R}$  ( $N$ -objects by  $L$ -responses), a target variance-covariance matrix,  $\mathbf{B}$  ( $L$  by  $L$ ), was identified (structures used are described below). The Cholesky

decomposition was performed on the target variance-covariance matrix giving a matrix  $C$  ( $L$  by  $L$ ) such that  $C^T C = B$ . A noise matrix,  $E$  ( $N$  by  $L$ ), is then created using the RAND function. The variance associated with  $E$  was adjusted by scalar multiplication of all entries in  $E$ . A structured error matrix,  $E_s$ , was then created by right multiplication of  $E$  by  $C$ . The resulting matrix,  $E_s$ , will then have a variance-covariance structure similar to the target matrix. The extent to which the variance-covariance structure of the resulting matrix,  $E_s$ , reproduces the target matrix variance-covariance structure will depend on the number of rows in  $E$ , where more rows will result in a better approximation.

There are many choices for variance-covariance structure, but some are more interesting than others. In order to have a guide for this selection an empirical noise matrix was measured. Infrared spectra were collected over time measuring the transmission of a polyurethane film in which no systematic temporal changes were taking place. A calculation of the response correlation structure for the spectral region from  $2164$  to  $2404\text{ cm}^{-1}$  at  $4\text{ cm}^{-1}$  intervals is shown in Figure 5.1 (every third response correlation is shown).



**Figure 5.1** Empirical response noise correlation matrix for infrared spectra, 2164 to 2404  $\text{cm}^{-1}$ , collected over time on a system with no systematic temporal changes.

As may be seen, there are high and relatively constant off diagonal terms in the correlation matrix. Given this empirical result it was decided to test variance-covariance structures with constant and equal off diagonal terms. The off diagonal terms used were 0.0, 0.1, 0.2, 0.4, 0.5, 0.6, or 0.8 and the diagonal elements were equal to one in most cases. The designations for these structures in the tables are 0.0, 0.1, 0.2, 0.4, 0.5, 0.6, 0.8. One other interesting noise structure is the "diagonal ridge" structure (9). In this structure the biggest correlations are next to the principal diagonal, tapering off toward the left and right and upward and downward. This structure is denoted by 0.9-0.0, indicating that the off diagonals closest to the principal diagonal are 0.9 and taper off in a linear fashion to 0.0 furthest away from the main diagonal.

For the direct observation experiments the independent responses were the concentrations of the species (A, B and C) involved in a two step consecutive first-order reaction scheme following the model

$$A = A_0 e^{-k_1 t} \quad (5.1)$$

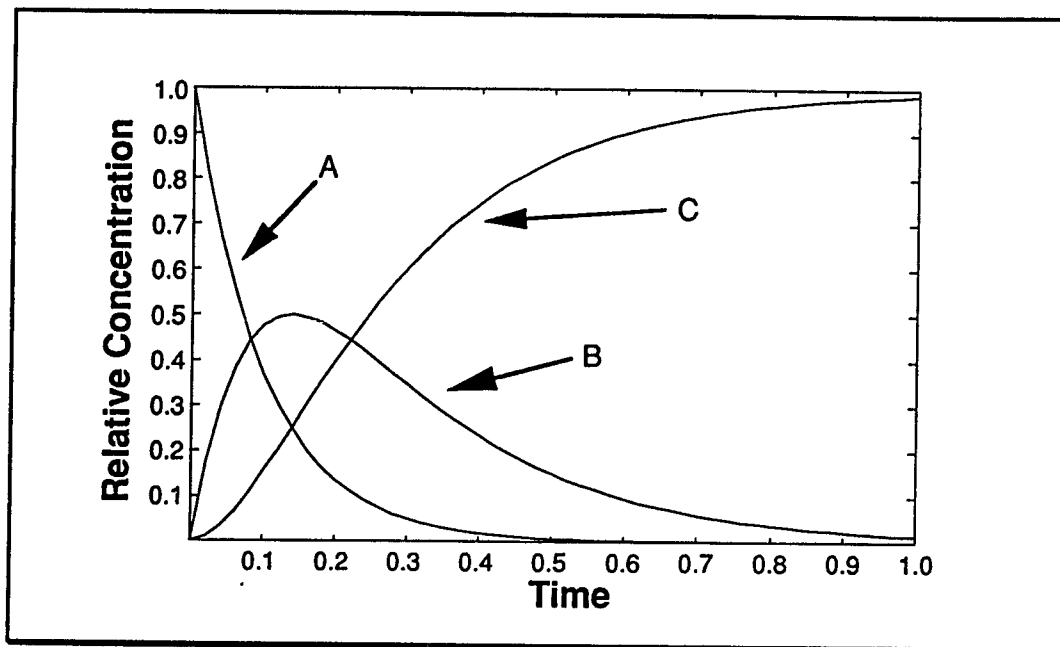
$$B = \frac{k_1 A_0}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}) \quad (5.2)$$

$$C = A_0 - B - A \quad (5.3)$$

where  $k_1$  and  $k_2$  are the kinetic parameters,  $t$  is time and  $A_0$  is the initial concentration of A.

For the indirect experiment simulations, there are two data matrices required to construct the observed data as shown in equation 2.58. One of the matrices corresponds to the parameterized model. Two parametric models were considered for these experiments. The first was a consecutive first-order reaction scheme, as was considered above for the direct observation experiments, with a time axis from 0 to 1 and an interval such that 40 time points were generated, as shown in Figure 5.2.



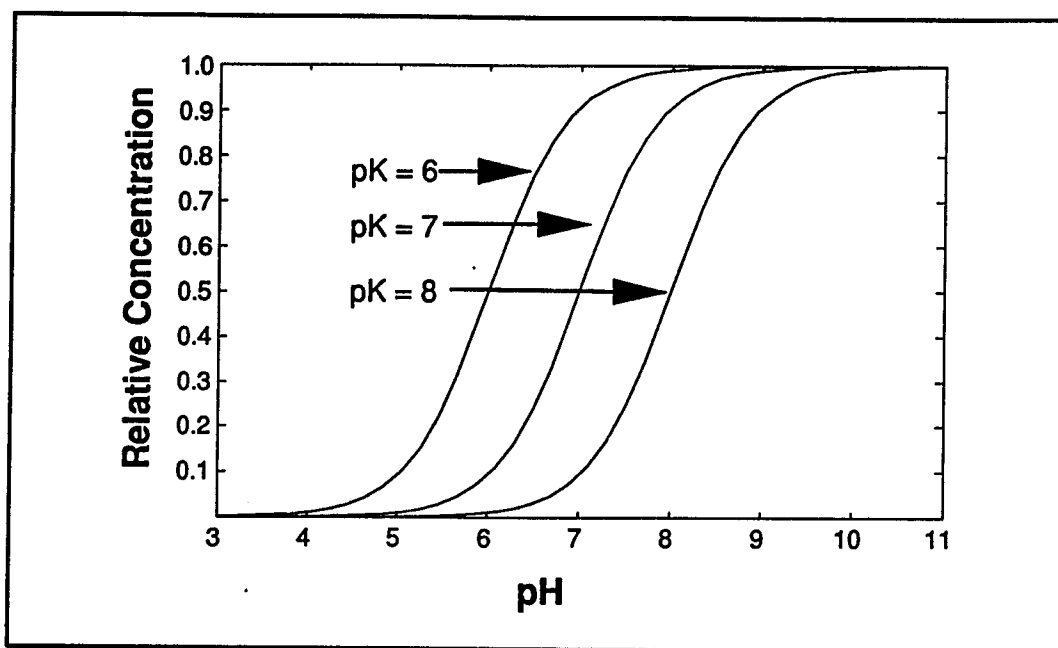


**Figure 5.2** Pure concentration profiles for first-order consecutive reaction scheme.

The second parameterized model used the Henderson-Hasselbalch equation for pH titrations

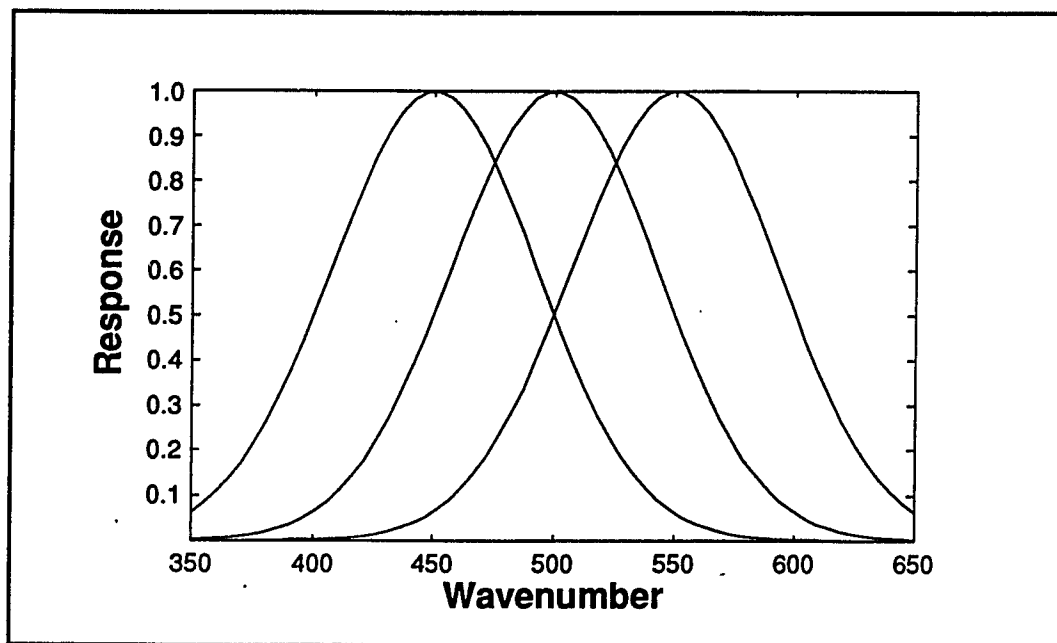
$$f(\text{pH}, \text{pK}) = \frac{1}{1 + 10^{(\text{pK} - \text{pH})}} \quad (5.4)$$

where the pH range considered was from 3 to 11, with an interval such that 40 data points were generated. Three chemical features were generated using the Henderson-Hasselbalch equation with pKs of 6, 7 and 8. This is shown graphically in Figure 5.3.



**Figure 5.3** Concentration profiles for pure components obeying the Henderson-Hasselbalch equation.

Gaussian peak models, simulating pure spectra, were used with the consecutive reaction model and the pH titration model to generate the indirect observation data using equation 2.58. Three gaussian peaks, with mean values of 450, 500 and 550 nm over a spectral axis of 350 to 650 nm at an interval giving 60 data points, were used as shown in Figure 5.4.



**Figure 5.4** Pure spectral profiles used in the indirect observation experiment simulations. Peak maximum at 350, 450 and 550 nm, half width at half height is 50 nm, the low overlap experiment.

The spectral features for the pH titrations are characterized by reaching a maximum intensity at the high pH region showing no absorbance in the acidic form. For each type of indirect experiment the spectral features are the same, but the parameterized models are a consecutive first-order reaction and a pH titration.

A modified SIMPLEX, implemented in MATLAB, was used as the optimization routine with the known true values of the parameters used as starting values. For each of the different objective functions several noise levels and variance-covariance noise structures were tested. At each of the different noise levels and variance-covariance structures 50 repeats of the optimization were performed and the standard error of prediction, SEP, defined as

$$SEP = \sqrt{\frac{1}{m} \sum_{i=1}^m (p_i - \hat{p}_i)^2} \quad (5.5)$$

was computed. In equation 5.5  $p_i$  is the known parameter value used to create the data and  $\hat{p}_i$  is the value returned by the optimization routine. Even though SEP is referred to as a standard error of prediction and is computed as an accuracy measure, it reflects the precision of the parameter estimates because in all cases the correct model is known and used.

The SIMPLEX search was terminated when the maximum of the sum of absolute differences between parameter values at the lowest residual node and all other nodes was less than a preset tolerance. A tolerance level of 0.0001 was used for all experiments except for the indirect observation experiment when the first-order consecutive reaction model was used with a tolerance of 0.001. All simulations were performed in the MATLAB computing environment on either a VAX running VMS, a SUN running UNIX or a VAX running ULTRIX.

## 5.3 Results and Discussion

### 5.3.1 Direct Observation Experiments

Five different experiments were performed for the direct observation case. For each experiment a comparison of TSS and DET objective functions was made at four signal-to-noise ratios and five variance-covariance noise structures. The simulation parameters unique to each experiment are detailed below.

### 5.3.1.1 Experiment #1

For Experiment #1 the time axis was from 0 to 1 at an interval of 0.02, giving 51 time data points and thus a response matrix of 51 rows by 3 columns. The true kinetic parameters were  $k_1 = 10$  and  $k_2 = 5$ . The four signal-to-noise ratios were 50:1, 25:1, 16.7:1 and 12.5:1. The five variance-covariance structures have ones on the diagonal and either 0.0, 0.2, 0.4, 0.6 or 0.8 on the off diagonals.

The results for these experiments, 1000 in all, are shown in Table 5.1.

**Table 5.1** SEP for direct measure experiment, consecutive first-order reaction kinetics,  $k_1=10$ ,  $k_2=5$ , diagonals of variance-covariance matrix are all ones.

Results for  $k_1$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.13	0.13	0.10	0.13	0.10	0.13	0.08	0.09	0.07	0.07
25:1	0.26	0.30	0.25	0.30	0.22	0.26	0.15	0.16	0.13	0.12
16.7:1	0.41	0.40	0.35	0.38	0.31	0.28	0.27	0.27	0.18	0.21
12.5:1	0.55	0.68	0.38	0.50	0.48	0.39	0.41	0.34	0.23	0.23

Results for  $k_2$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.05	0.05	0.04	0.05	0.05	0.04	0.04	0.03	0.02	0.02
25:1	0.09	0.12	0.10	0.10	0.09	0.07	0.05	0.07	0.04	0.05
16.7:1	0.16	0.17	0.16	0.14	0.12	0.11	0.10	0.11	0.06	0.07
12.5:1	0.24	0.23	0.15	0.18	0.17	0.17	0.14	0.15	0.11	0.10

The SEPs using the TSS and DET criteria for  $k_1$  and  $k_2$  are shown as a function of signal-to-noise ratio down the rows and as a function of noise structure across the columns. The one immediately obvious trend is that the SEP increases as

the signal-to-noise ratio decreases for a given noise structure, which is not unexpected. At a given noise level the SEP for both criteria tends to decrease as the variance-covariance off diagonal entries increase. Overall there appears to be little difference between the determinant criterion and the total sum of squares criterion for these experimental parameters.

#### **5.3.1.2 Experiment #2**

For this experiment the time axis was from 0 to 1 at an interval of 0.02, giving 51 time data points as in Experiment #1 and the kinetic parameters were  $k_1 = 10$  and  $k_2 = 5$ . The variance-covariance structure used the same set of off diagonal entries as in Experiment #1. The diagonal entries were changed from all ones to three, two, and one in order to test for differences in objective function performance under conditions of non-constant variance. The simulation results are summarized in Table 5.2.

**Table 5.2** SEP for direct measure experiment, consecutive first-order reaction kinetics experiment,  $k_1=10$ ,  $k_2=5$ , diagonals of variance-covariance matrix are 3, 2, 1.

Results for  $k_1$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.20	0.20	0.20	0.19	0.16	0.20	0.18	0.21	0.17	0.19
25:1	0.44	0.48	0.43	0.47	0.40	0.47	0.39	0.44	0.33	0.33
16.7:1	0.68	0.64	0.66	0.62	0.57	0.60	0.55	0.50	0.53	0.55
12.5:1	0.99	1.07	0.97	1.05	0.65	0.77	0.86	0.71	0.83	0.71

Results for  $k_2$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.07	0.07	0.07	0.07	0.05	0.07	0.06	0.06	0.06	0.05
25:1	0.13	0.16	0.13	0.15	0.13	0.12	0.13	0.11	0.09	0.10
16.7:1	0.22	0.23	0.21	0.22	0.20	0.19	0.17	0.16	0.15	0.17
12.5:1	0.33	0.31	0.32	0.30	0.20	0.24	0.25	0.23	0.23	0.22

As observed for Experiment #1 there is a general increase in the SEP values for increasing noise levels at a constant noise structure. There is also a trend toward smaller SEP values for a given noise level as the off diagonal entries in the variance-covariance matrix are increased, although it is not as pronounced as in Experiment #1. Overall the SEPs are larger for Experiment #2 than for Experiment #1, as would be expected because of the larger diagonal entries in the variance-covariance matrix which lead to a larger total variance. There appears to be no advantage of one criterion over the other for these experimental parameters.

### 5.3.1.3 Experiment #3

In order to investigate the effect of observing fewer data points on the performance of the two objective functions only the first 35 % of the data were used in Experiment #3. The time axis was from 0 to 0.35 at an interval of 0.02, giving 18 time data points. The kinetic parameters were  $k_1 = 10$  and  $k_2 = 5$ . The variance-covariance structure was all ones on the diagonal and 0.0, 0.2, 0.4, 0.6 or 0.8 on the off diagonals. Table 5.3 displays the results for these simulations.

**Table 5.3** Reduced sample size. SEP for direct measure experiment, consecutive first-order reaction kinetics experiment,  $k_1=10$ ,  $k_2=5$ , observed for first 35 % of time, diagonals of variance-covariance matrix are all ones.

Results for  $k_1$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.13	0.13	0.09	0.16	0.10	0.12	0.09	0.09	0.06	0.06
25:1	0.23	0.29	0.24	0.25	0.21	0.26	0.17	0.18	0.12	0.14
16.7:1	0.38	0.55	0.39	0.42	0.35	0.36	0.24	0.29	0.21	0.20
12.5:1	0.60	0.56	0.46	0.66	0.44	0.51	0.38	0.38	0.25	0.25

Results for  $k_2$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.07	0.07	0.06	0.07	0.06	0.06	0.05	0.05	0.04	0.04
25:1	0.14	0.14	0.15	0.14	0.13	0.14	0.10	0.08	0.08	0.06
16.7:1	0.16	0.22	0.25	0.24	0.20	0.21	0.16	0.14	0.09	0.09
12.5:1	0.30	0.30	0.27	0.33	0.25	0.27	0.22	0.24	0.15	0.09

Similar observations are found for these experiments as were found for Experiments #1 and #2. The SEP increases with increasing noise level and decreases with increasing size of the off diagonal variance-covariance entries. Experiment #1



showed similar results for  $k_1$  as those found in Experiment #3, while for  $k_2$  higher error estimates in Experiment #3 are noted. This is to be expected because of fewer data points defining the time at which species C would be observed. Again there appears to be no clear distinction between the two criteria.

#### **5.3.1.4 Experiment #4**

Because the accuracy with which the noise variance-covariance structure is generated depends on sample size (in this case on the number of time data points), a fourth experiment was conducted using 201 time data points in order to have better control over the noise structure. For this experiment the time axis was from 0 to 1 at an interval of 0.005, giving 201 time data points. The kinetic parameters were  $k_1 = 10$  and  $k_2 = 5$ . The variance-covariance structure was all ones on the diagonal and 0.0, 0.2, 0.4, 0.6 or 0.8 on the off diagonals. The simulation results are summarized in Table 5.4.

**Table 5.4** Increased sample size. SEP for direct measure experiment, consecutive first-order reaction kinetics experiment,  $k_1=10$ ,  $k_2=5$ , diagonals of variance-covariance matrix are all ones, 201 data points.

Results for  $k_1$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.06	0.07	0.06	0.05	0.06	0.05	0.05	0.05	0.03	0.03
25:1	0.13	0.14	0.10	0.14	0.10	0.10	0.08	0.08	0.06	0.06
16.7:1	0.22	0.20	0.20	0.17	0.18	0.17	0.12	0.14	0.09	0.09
12.5:1	0.25	0.27	0.28	0.26	0.23	0.26	0.18	0.19	0.13	0.13

Results for  $k_2$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.01
25:1	0.03	0.04	0.04	0.04	0.03	0.04	0.03	0.03	0.03	0.02
16.7:1	0.09	0.07	0.07	0.07	0.05	0.06	0.06	0.05	0.04	0.03
12.5:1	0.11	0.10	0.12	0.09	0.09	0.08	0.07	0.07	0.04	0.04

The same SEP trends are present as were found in Experiment #1. The SEP values are lower in this experiment when compared with Experiment #1, as would be expected, because of the increased number of time points observed. There still remains an unclear distinction between the two criteria.

### 5.3.1.5 Experiment #5

A fifth experiment was conducted in order to test if a difference in the objective function performance could be detected with poorer time resolution. The kinetic parameters were set to  $k_1 = 8$  and  $k_2 = 6$ . The time vector was from 0 to 1 at an interval of 0.02, giving 51 time data points. The variance-covariance structure was all ones on the diagonal and 0.0, 0.2, 0.4, 0.6 or 0.8 on the off diagonals.

Table 5.5 displays the simulation results for these 1000 experiments. Similar observations, as noted in the other experiments, may be made about the results from this experiment. The SEP increases for a decrease in signal-to-noise ratio and decreases as the noise structure is increased.

**Table 5.5** Poorer time resolution. SEP for direct measure experiment, consecutive first-order reaction kinetics experiment,  $k_1=8$ ,  $k_2=6$ , diagonals of variance-covariance matrix are all ones.

Results for  $k_1$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.08	0.09	0.06	0.09	0.07	0.08	0.06	0.07	0.05	0.05
25:1	0.18	0.22	0.18	0.23	0.16	0.18	0.10	0.12	0.10	0.09
16.7:1	0.28	0.29	0.25	0.28	0.23	0.20	0.19	0.19	0.13	0.16
12.5:1	0.39	0.50	0.29	0.34	0.37	0.27	0.29	0.23	0.16	0.16

Results for  $k_2$ .

	0.0		0.2		0.4		0.6		0.8	
S/N	TSS	DET	TSS	DET	TSS	DET	TSS	DET	TSS	DET
50:1	0.07	0.06	0.06	0.07	0.06	0.06	0.05	0.04	0.03	0.03
25:1	0.13	0.16	0.14	0.14	0.12	0.11	0.07	0.09	0.06	0.06
16.7:1	0.22	0.23	0.22	0.20	0.16	0.16	0.13	0.14	0.09	0.10
12.5:1	0.33	0.32	0.21	0.25	0.24	0.23	0.18	0.20	0.14	0.13

For the indirect observation experiments it has been shown that the SEP for the parameter estimates generally increases with decreasing signal-to-noise ratio and decreases for increasing variance-covariance structure. Better parameter estimates are found if more time data points are used because of the signal averaging. When the diagonal terms in the variance-covariance noise structure are larger, the noise variance is larger and the parameter estimates are poorer. No clear distinction is found between the two criteria for this model and these experiments.

### 5.3.2 Indirect Observation Experiments

For the indirect observation study two different experiments were simulated. One in which spectral measurements were made as a first-order consecutive reaction occurred and a second in which spectral measurements were made during a pH titration. For each experiment the same three spectral features were used as described in the Section 5.2 and shown in Figure 5.4. In each case four signal-to-noise ratios and six variance-covariance structures were tested. Three singular vectors were used to represent the compressed data. A comparison was made between the total sum of squares (TSS), total sum of squares with weights (TSSW) and the determinant criterion (DET). The details unique to each experiment are presented below.

#### 5.3.2.1 First-order Consecutive Reaction Results

For the first-order consecutive reaction study two experiments were performed, one at low spectral overlap and one at high spectral overlap. For the low spectral overlap experiments each of the three spectral features has a half width at half height of 50 nm. For the high overlap experiments each has a half width at half height of 73.5 nm. Each response matrix is 40 time points by 60 spectral points. The kinetic parameters were  $k_1 = 10$  and  $k_2 = 5$  in both cases. The SIMPLEX stopping tolerance was set to 0.001.

#### **5.3.2.1.1 Low overlap results**

The results for the low overlap experiments are displayed in Tables 5.6 and 5.7. Table 5.6 displays the results for  $k_1$  and Table 5.7 displays the results for  $k_2$ . A decrease in the signal-to-noise ratio leads to an increase in the parameter SEP, as expected. The SEP for the TSS criterion is highest in nearly all cases. There is a general increase in SEP with an increase in the variance-covariance structure from 0.0 to 0.8 for the TSS criterion at a constant noise level. There is a general trend of decreasing SEP as the variance-covariance structure increases for a given noise level for the DET criterion. This is especially evident at the lower signal-to-noise ratios. The SEP for the TSSW criterion is fairly constant with noise structure at low noise intensity, but at the highest noise level TSSW shows an increase in SEP with increasing noise structure.

**Table 5.6** SEP for indirect, consecutive first-order reaction kinetics experiment, low overlap, results for  $k_1$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.29	0.16	0.18	0.47	0.16	0.16	0.57	0.14	0.14
50:1	0.63	0.30	0.33	0.96	0.30	0.30	1.19	0.28	0.26
25:1	1.29	0.58	0.79	1.64	0.57	0.70	2.01	0.60	0.61
12.5:1	2.25	1.44	1.72	4.35	1.54	1.59	4.73	4.62	1.24

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.63	0.13	0.13	0.88	0.09	0.09	0.72	0.32	0.30
50:1	1.31	0.27	0.24	1.46	0.26	0.15	1.29	0.75	0.64
25:1	2.14	0.65	0.55	3.01	0.93	0.31	2.53	1.41	1.33
12.5:1	5.44	5.88a	1.17	22.73	15.38b	0.77	4.87	159.06c	2.71

- a. Two results over 20. New SEP after eliminating two outliers is 2.8.  
 b. Two results over 20. New SEP after eliminating two outliers is 2.9.  
 c. Two results over 20. New SEP after eliminating two outliers is 4.0.

**Table 5.7** SEP for indirect, consecutive first-order reaction kinetics experiment, low overlap, results for  $k_2$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.14	0.04	0.06	0.22	0.04	0.05	0.28	0.04	0.04
50:1	0.34	0.09	0.12	0.53	0.09	0.11	0.69	0.07	0.09
25:1	0.77	0.18	0.26	0.81	0.19	0.22	0.95	0.17	0.19
12.5:1	1.25	0.43	0.78	1.42	0.44	0.73	1.43	0.45	0.46

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.31	0.04	0.04	0.60	0.02	0.03	0.33	0.13	0.13
50:1	0.79	0.07	0.09	0.87	0.05	0.05	0.79	0.30	0.29
25:1	1.00	0.16	0.17	1.33	0.13	0.10	1.20	0.79	0.72
12.5:1	1.32	0.43	0.41	1.46	0.77	0.23	1.41	1.13	1.24

The overall result is that the TSSW criterion may be somewhat better at the lowest variance-covariance structure, but at high noise and high variance-covariance structure the determinant criterion is superior. In fact, in some cases ( $S/N = 12.5:1$  and noise structure at 0.6 and 0.8), the TSSW diverged to grossly incorrect estimates. Even if these outliers are eliminated from the SEP calculation, the DET criterion was still a better performer.

For the diagonal ridge noise structure, 0.9-0.0, the TSSW for  $k_1$  performs very poorly at the lowest signal-to-noise ratio, converging to grossly incorrect estimates in two cases and giving an SEP of 159.06. The determinant criterion performs better even when these outlier results were eliminated from the TSSW SEP calculation. For the  $k_2$  parameter very similar performance for the DET and TSSW objective functions was found for the diagonal ridge structure.

#### **5.3.2.1.2 High overlap results**

For the high spectral overlap experiments the results are displayed in Tables 5.8 and 5.9. For these experiments the overall SEP increased as compared to the low overlap experiments due to poorer spectral resolution. The same trends were found as were noted for the low overlap case except that they became evident with less structure in the noise. For example, the TSSW criterion diverges to grossly incorrect parameter estimates for  $k_1$  at a noise structure level of 0.2 for the high overlap case, while the difference does not become evident in the low overlap case until a noise structure level of 0.4 or 0.5. Note also that the determinant criterion was never found to diverge to grossly incorrect parameter estimates.

**Table 5.8** SEP for indirect, consecutive first-order reaction kinetics experiment, high overlap, results for  $k_1$ .

	0.0			0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.45	0.24	0.25	0.83	0.24	0.22	1.12	0.23	0.20
50:1	1.09	0.52	0.48	1.66	0.53	0.44	1.96	0.50	0.39
25:1	1.65	1.04	1.15	2.23	0.94	0.98	2.71	1.13	0.83
12.5:1	3.34	3.29	2.58	4.73	6.52a	2.27	6.90	10.73b	1.93

- a. Three results over 20. New SEP after eliminating 3 outliers is 3.0.  
 b. Three results over 20. New SEP after eliminating 3 outliers is 4.4.

	0.5			0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	1.17	0.23	0.18	1.31	0.25	0.12	1.34	0.54	0.49
50:1	1.82	0.49	0.35	2.01	0.49	0.23	1.92	1.27	1.03
25:1	2.94	1.23	0.74	6.61	1.77	0.47	3.20	2.13	1.95
12.5:1	8.18	13.77a	1.78	304.6	262.62b	1.03	6.59	173.81c	3.99

- a. Six results over 20. New SEP after eliminating six outliers is 3.8.  
 b. Seven results over 20. New SEP after eliminating seven outliers is 3.1.  
 c. Eleven results over 20. New SEP after eliminating eleven outliers is 4.5.



**Table 5.9** SEP for indirect, consecutive first-order reaction kinetics experiment, high overlap, results for  $k_2$ .

S/N	0.0			0.2			0.4		
	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.24	0.05	0.06	0.43	0.05	0.05	0.64	0.04	0.04
50:1	0.63	0.11	0.12	1.05	0.11	0.10	1.18	0.09	0.09
25:1	0.87	0.19	0.27	1.08	0.22	0.22	1.32	0.25	0.19
12.5:1	1.51	0.62	0.78	1.51	0.52	0.72	1.39	0.46	0.57

S/N	0.5			0.8			0.9-0.0		
	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
100:1	0.68	0.04	0.04	0.74	0.02	0.02	0.75	0.17	0.17
50:1	1.00	0.08	0.08	1.11	0.06	0.05	1.10	0.44	0.43
25:1	1.30	0.22	0.17	1.46	0.32	0.10	1.41	1.14	0.93
12.5:1	1.40	0.70	0.52	1.60	0.46	0.21	1.63	1.52	1.43

The performance of the two criteria is more comparable for the  $k_2$  parameter at the low noise structure experiments. At the highest noise structure level, 0.8, and two lowest signal-to-noise ratios, 25:1 and 12.5:1, the determinant criterion is better by a factor of three and two, respectively.

### 5.3.2.2 pH Titration Experiments

For the indirect observation of a changing equilibrium, two experiments were performed, one at low spectral overlap and a second at high spectral overlap, as in the consecutive first-order reaction simulations discussed in Section 5.3.2.1. The true  $pK_a$  values were  $k_1 = 6$ ,  $k_2 = 7$ , and  $k_3 = 8$ . Six different variance-covariance structures were tested at four noise levels.

### 5.3.2.2.1 Low overlap results

Standard error prediction results for  $k_1$ ,  $k_2$  and  $k_3$  for the low overlap experiments, are displayed in Tables 5.10, 5.11 and 5.12, respectively. The familiar trend of increasing SEP with increasing noise may be seen in these tables. The TSS criterion is the poorest performer under all conditions tested. There appears to be a general trend of increasing SEP for the TSS and TSSW with increasing noise structure, for a constant signal-to-noise ratio. The DET criterion SEP generally decreases as the noise structure increases for a constant signal-to-noise ratio. These trends give the determinant criterion a distinct advantage with as little noise structure as 0.2. The  $k_1$  SEP using the determinant criterion is as much as ten times better than the TSSW SEP at the 0.8 noise structure level with a 25:1 signal-to-noise ratio. For the diagonal ridge structure the determinant criterion is more advantageous for determining  $k_2$  at high noise levels.

**Table 5.10** SEP for indirect pK determination, low overlap, results for  $k_1$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.04	0.01	0.01	0.06	0.01	0.01	0.09	0.02	0.01
100:1	0.08	0.01	0.02	0.09	0.03	0.02	0.19	0.03	0.02
50:1	0.27	0.03	0.03	0.35	0.05	0.03	0.34	0.06	0.03
25:1	0.33	0.06	0.08	0.44	0.13	0.07	0.44	0.17	0.05

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.10	0.02	0.01	0.13	0.02	0.01	0.14	0.02	0.02
100:1	0.23	0.04	0.01	0.19	0.05	0.01	0.23	0.05	0.04
50:1	0.37	0.07	0.02	0.33	0.17	0.02	0.47	0.10	0.08
25:1	0.41	0.19	0.05	0.62	0.29	0.03	0.50	0.31	0.24

**Table 5.11** SEP for indirect pK determination, low overlap, results for  $k_2$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.06	0.01	0.02	0.09	0.02	0.02	0.15	0.03	0.01
100:1	0.12	0.02	0.03	0.12	0.04	0.02	0.27	0.05	0.02
50:1	0.35	0.04	0.06	0.47	0.08	0.06	0.59	0.10	0.05
25:1	1.16	0.10	0.12	0.92	0.21	0.13	0.95	0.46	0.11

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.11	0.03	0.01	0.16	0.03	0.01	0.20	0.04	0.02
100:1	0.33	0.07	0.02	0.36	0.08	0.01	0.37	0.08	0.05
50:1	0.54	0.12	0.05	0.84	0.21	0.03	0.54	0.16	0.11
25:1	0.70	0.61	0.08	0.86	0.50	0.07	1.18	0.73	0.34

**Table 5.12** SEP for indirect pK determination, low overlap, results for  $k_3$ .

S/N	0.0			0.2			0.4		
	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.04	0.01	0.01	0.06	0.01	0.01	0.12	0.02	0.01
100:1	0.09	0.01	0.01	0.14	0.02	0.02	0.29	0.03	0.02
50:1	0.40	0.02	0.03	0.66	0.05	0.04	0.68	0.08	0.03
25:1	0.70	0.07	0.06	0.72	0.23	0.07	0.96	0.34	0.07

S/N	0.5			0.8			0.9-0.0		
	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.09	0.02	0.01	0.11	0.02	0.01	0.13	0.03	0.02
100:1	0.36	0.05	0.01	0.46	0.05	0.01	0.50	0.05	0.06
50:1	0.74	0.10	0.03	0.69	0.12	0.02	0.67	0.13	0.09
25:1	0.74	0.35	0.05	0.65	0.62	0.04	0.62	0.45	0.37

**5.3.2.2.2 High overlap results**

The results for the high overlap experiments are displayed in Tables 5.13, 5.14, and 5.15. They show the same general trends as observed in the low overlap case. Overall the SEPs are larger than the lower overlap experiments, as expected.

Although there were no gross divergence problems experienced by the TSSW criterion in these experiments, the determinant criterion is found to have superior performance especially at the low signal-to-noise ratios and the high variance-covariance noise structure.

**Table 5.13** SEP for indirect pK determination, high overlap, results for  $k_1$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.08	0.01	0.01	0.17	0.01	0.01	0.19	0.02	0.01
100:1	0.16	0.01	0.02	0.20	0.03	0.02	0.20	0.05	0.02
50:1	0.49	0.03	0.03	0.32	0.06	0.03	0.39	0.07	0.03
25:1	0.43	0.07	0.06	0.53	0.13	0.06	0.43	0.17	0.05

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.14	0.02	0.01	0.18	0.02	0.01	0.18	0.02	0.03
100:1	0.24	0.05	0.01	0.31	0.05	0.01	0.20	0.06	0.04
50:1	0.39	0.07	0.02	0.33	0.09	0.02	0.81	0.10	0.07
25:1	0.42	0.17	0.05	0.59	0.24	0.03	0.63	0.26	0.35

**Table 5.14** SEP for indirect pK determination, high overlap, results for  $k_2$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.10	0.01	0.02	0.19	0.02	0.02	0.23	0.03	0.02
100:1	0.27	0.02	0.04	0.28	0.06	0.04	0.30	0.07	0.03
50:1	0.56	0.05	0.06	0.58	0.09	0.08	0.65	0.11	0.06
25:1	0.84	0.12	0.12	0.98	0.27	0.14	1.00	0.36	0.12

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.25	0.03	0.01	0.27	0.04	0.01	0.26	0.05	0.03
100:1	0.36	0.08	0.03	0.50	0.08	0.02	0.47	0.09	0.07
50:1	0.72	0.11	0.05	0.70	0.18	0.04	0.58	0.19	0.17
25:1	0.77	0.39	0.09	0.84	0.53	0.07	1.01	0.56	0.35

**Table 5.15** SEP for indirect pK determination, high overlap, results for  $k_3$ .

0.0				0.2			0.4		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.07	0.01	0.01	0.10	0.01	0.01	0.13	0.02	0.01
100:1	0.17	0.01	0.02	0.39	0.03	0.02	0.39	0.05	0.02
50:1	0.75	0.03	0.03	0.72	0.06	0.04	0.84	0.08	0.04
25:1	0.80	0.07	0.07	1.44	0.18	0.08	0.79	0.31	0.07

0.5				0.8			0.9-0.0		
S/N	TSS	TSSW	DET	TSS	TSSW	DET	TSS	TSSW	DET
200:1	0.20	0.03	0.01	0.18	0.03	0.01	0.18	0.03	0.02
100:1	0.49	0.05	0.01	0.69	0.05	0.01	0.56	0.05	0.05
50:1	0.85	0.09	0.03	0.77	0.17	0.02	0.73	0.16	0.23
25:1	0.97	0.52	0.06	0.68	0.46	0.04	1.14	0.38	0.45

## 5.4 Conclusions

The choice of the proper objective function for multiresponse nonlinear parameter estimation has been considered. A distinction was drawn between the full rank or direct observation experiment, and the rank deficient or indirect observation experiment. The model for a first-order consecutive reaction was used in the full rank experiments and two objective functions were tested, the TSS and the DET criteria. There was no clear distinction between the two objective functions based on parameter SEP comparisons. While these simulations do not support the use of one criterion over the other for the full rank case, they do suggest that the determinant criterion be used because, in theory, the determinant criterion should perform better and has been shown not to perform significantly worse than the TSS in these studies.

For the rank deficient experiments, three objective functions were considered, (1) TSS, (2) TSSW and (3) DET. Two experimental situations were considered, a consecutive first-order reaction and a pH titration, using spectral measurements as the indirect observation technique in both cases. The TSS is never recommended as an objective function. The TSSW has some advantage at low noise intensity and structure, but the determinant criterion is a better performer in the high noise intensity, high noise structure situations. In light of the high noise structure observed experimentally it is suggested that the DET criterion be used for highest parameter precision.

## 5.5 Notes to Chapter 5

1. Brown, S. D. *Anal. Chem.* **1990**, *62*, 84R-101R.
2. Box, G. E. P.; Draper, N. R. *Biometrika* **1965**, *52*, 355-365.
3. Shrager, R. I.; Hendler, R. W. *Anal. Chem.* **1982**, *54*, 1147-1152.
4. Frans, S. D.; Harris, J. M. *Anal. Chem.* **1984**, *56*, 466-470.
5. Frans, S. D.; Harris, J. M. *Anal. Chem.* **1985**, *57*, 1718-1721.
6. Shrager, R. I. *Chemo. Intell. Lab. Sys.* **1986**, *1*, 59-70.
7. Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and Its Applications*; John Wiley & Sons: New York, 1988.
8. Halaka, F. G.; Babcock, B. T.; Dye, J. L. *J. Biophys.* **1985**, *48*, 209-219.
9. Takeuchi, K.; Yanai, H.; Mukherjee, B. N. *Foundations of Multivariate Analysis A Unified Approach by Means of Projection Onto Linear Subspaces*; Wiley Eastern Limited: New Delhi, 1982, 425.



## Bibliography

- Aitchison, J. *J. R. Statist. Soc. B* **1982**, *44*, 139-177.
- Aitchison, J. *The Statistical Analysis of Compositional Data*; Chapman and Hall: London, 1986.
- Alcock, N. W.; Benton, D. J.; Moore, P. *Trans. Farad. Soc.* **1970**, *66*, 2210-2213.
- Allara, D. L.; Teicher, D.; Durana, J. F. *Chem. Phys. Lett.* **1981**, *84*, 20-24.
- Alpert, N. L.; Keiser, W. E.; Szymanski, H. A. *Theory and Practice of Infrared Spectroscopy*; Plenum Publishing Corporation: New York, 1964.
- Barr, J. K. *Infrared Physics* **1969**, *9*, 97-108.
- Bates, D. M.; Watts, D. G. *Siam. J. Sci. Stat. Comput.* **1987**, *8*, 49-55.
- Bates, D. M.; Watts, D. G. *Commun. Statist.-Simula. Computa.* **1984**, *13*, 705-715.
- Bates, D. M.; Watts, D. G. *Technometrics* **1985**, *27*, 329-339.
- Bates, D. M.; Watts, D. G. *Nonlinear Regression Analysis and Its Applications*; John Wiley & Sons: New York, 1988.
- Bates, J. B. In *Fourier Transform Infrared Spectroscopy: Applications to Chemical Systems*; Vol. 1; Ferraro, J. R. and Basile, L. J., Eds.; Academic Press: New York; 1978; pp 99-142.
- Bates, J. B.; Boyd, G. E. *Appl. Spectrosc.* **1973**, *27*, 204-208.
- Bauer, D. R.; Dickie, R. A. *Ind. Eng. Chem. Prod. Res. Dev.* **1986**, *25*, 289-296.
- Beattie, J. R.; Cohen E. *British J. of Appl. Phys.* **1960**, *11*, 151-157.
- Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 1007A-1017A.
- Bellamy, L. J. *The Infrared Spectra of Complex Molecules: 2nd ed.*; Chapman and Hall Ltd.: New York, 1980.
- Belz, H. H.; Gutberlet, H.; Schallert, B.; Schrader, B. *Appl. Spectrosc.* **1987**, *41*, 1009-1019.
- Blanke, J. F.; Vincent, S. E.; Overend, J. *Spectrochim. Acta* **1976**, *32a*, 163-173.
- Borgen, O. S.; Kowalski, B. R. *Anal. Chim. Acta.* **1985**, *174*, 1-26.
- Borodai, S. P.; Betina, T. A. *Zhurnal Prikladnoi Spektroskopii* **1988**, *48*, 114-119.

- Box, G. E. P.; Draper, N. R. *Biometrika* **1965**, *52*, 355-365.
- Box, G. E. P.; Hunter, W. G.; MacGregor, J. F.; Erjavec, J. *Technometrics* **1973**, *15*, 33-51.
- Box, G. E. P.; Hunter, W. G.; Hunter, J. S. *Statistics For Experimenters*; John Wiley & Sons Inc.: New York, 1978.
- Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification and Regression Trees*; Wadsworth: Belmont, CA, 1984.
- Brereton, R. G. *Analyst* **1987**, *112*, 1635-1657.
- Brown, C. W.; Obremski, R. J. *Appl. Spect. Review* **1984**, *20*, 373-418.
- Brown, S. D.; Barker, T. Q.; Larivee, R. J.; Monfre, S. L.; Wilk, H. R. *Anal. Chem.* **1988**, *60*, 252R-279R.
- Brown, S. D. *Anal. Chem.* **1990**, *62*, 84R-101R.
- Callis, J. B.; Illman, D. L.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 624A-637A.
- Carli, B.; Park, J. H. *J. Geo. Res.* **1988**, *93*, 3851-3865.
- Chadwick, D. H.; Cleveland, T. H. In *Encyclopedia of Chemical Technology* Third Edition Vol. 13; Grayson, M. Ed., John Wiley & Sons: New York; p 810.
- Chase, D. B. *Appl. Spectrosc.* **1981**, *35*, 77-81.
- Chupp, R. E.; Viskanta, R. *AIAA* **1970**, *8*, 551-557.
- Cleveland, W. S.; Devlin, S. J. *J. American Statistical Association* **1988**, *83*, 596-610.
- Coblentz, W. W. *Tech. News Bull. Bur. Standards* **1908**, *5*, 159.
- Coblentz, W. W. *Tech. News Bull. Bur. Standards* **1910**, *6*, 301.
- Coblentz, W. W. *Tech. News Bull. Bur. Standards* **1910**, *7*, 243.
- Coblentz, W. W. *Tech. News Bull. Bur. Standards* **1912**, *9*, 81.
- Cochran, R. N.; Horne, F. H. *Anal. Chem.* **1977**, *49*, 846-853.
- Cochran, R. N.; Horne, F. H.; Dye, J. L.; Ceraso, J.; Sueiter, C. H. *J. Chem. Phys.* **1980**, *84*, 230-245.
- Cochran, R. N.; Horne, F. H. *J. Chem. Phys.* **1980**, *84*, 2561-2567.
- Coleman, I.; Low, M. J. D. *Spectrochim. Acta* **1966**, *22*, 1293-1298.

- Conroy, C. M.; Guthrie, J. D.; Sharkins, A. J.; Sparr, B. J.; Crocombe, R. A.; Curbelo, R. *Appl. Spectrosc.* **1987**, *41*, 688-692.
- Delaney, M. R. *Anal. Chem.* **1984**, *56*, 261R-277R.
- Draper, N.; Smith, H. *Applied Regression Analysis, Second Edition*; John Wiley & Sons: New York, 1981.
- Drude, P. *Lehrbuch der Optik Hirzel*: Leipzig, **1912**; 494.
- Dvurechenskii, A. V.; Petrov, V. A.; Reznik, V. Y. *High Temperatures-High Pressures* **1979**, *11*, 423-428.
- Eakman, J. M. *Ind. Eng. Chem. Fundamentals* **1969**, *8*, 53-58.
- Eakman, J. M. *Ind. Eng. Chem. Fundamentals* **1970**, *9*, 187-188.
- Eastment, H. T.; Krzanowski, W. J. *Technometrics* **1982**, *24*, 73-77.
- Eischen, R. P.; Pliskin, W. A. In *Advances in Catalysis and Related Subjects*; Vol 10; Frankenburg, Walter G., Ed.; Academic Press: New York; 1958, pp 2-54.
- Erickson, B. C. "Multicomponent Flow Injection Analysis and Quantitative Infrared Emission Spectroscopy: Chemometric Applications" Ph. D. Dissertation, University of Washington: Seattle, WA (1988). Available from: University Microfilms, Ann Arbor, MI.
- Erjavec, J. *Ind. Eng. Chem. Fundamentals* **1970**, *9*, 187.
- Fabbri, G.; Baraldi, P. *Appl. Spectrosc.* **1972**, *26*, 593-599.
- Feger, C.; Molis, S. E.; Hsu, S. L.; MacKnight, W. J. *Macromolecules* **1984**, *17*, 1830-1834.
- Fletcher, T. R.; Leone, S. R. *J. Chem. Phys.* **1989**, *90*, 871.
- Folberth, W.; Heim, G. *SPIE - Fourier and Computerized Infrared Spectroscopy* **1985**, *553*, 129-130.
- Ford, M. A.; Spragg, R. A. *Appl. Spectrosc.* **1986**, *40*, 715-716.
- Frank, I. E.; Kowalski, B. R. *Anal. Chem.* **1982**, *52*, 232R-243R.
- Frans, S. D.; Harris, J. M. *Anal. Chem.* **1984**, *56*, 466-470.
- Frans, S. D.; Harris, J. M. *Anal. Chem.* **1985**, *57*, 1718-1721.
- Frech, R.; Bates, J. B. *Spectrochim. Acta* **1979**, *35A*, 685-694.
- Gardon, R. J. *Am. Ceram. Soc.* **1958**, *41*, 200-209.

- Gardon, R. *J. Am. Ceram. Soc.* **1956**, *39*, 278-287.
- Gardon, R. *J. Am. Ceram. Soc.* **1961**, *44*, 305-312.
- Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta.* **1986**, *185*, 1-17.
- Geladi, P.; MacDougall, D.; Martens, H. *Appl. Spectrosc.* **1985**, *39*, 491.
- Golub, G. H.; Vanloan, C. F. *Matrix Computations Second Edition*; Johns Hopkins University Press: Baltimore, 1989.
- Gratton, L. M.; Paglia, S.; Scattaglia, F.; Cavallini, M. *Appl. Spectrosc.* **1978**, *32*, 310.
- Green, P. E. *Mathematical Tools for Applied Multivariate Analysis*; Academic Press, Inc.: San Diego, 1978.
- Greenler, R. G. *Surface Science* **1977**, *69*, 647-652.
- Griffiths, P. R. *Appl. Spectrosc.* **1977**, *31*, 497-505.
- Griffiths, P. R.; deHaseth, J. A. *Fourier Transform Infrared Spectrometry*; John Wiley & Sons: New York, 1986.
- Griffiths, P. R. *Appl. Spectrosc.* **1972**, *26*, 73-76.
- Gross, L. A.; Griffiths, P. R. *JQSRT* **1988**, *39*, 463-472.
- Gross, L. A.; Griffiths, P. R. *JQSRT* **1988**, *39*, 131-138.
- Haaland, D. M.; Easterling, R. G. *Appl. Spectrosc.* **1980**, *34*, 539.
- Haaland, D. M.; Easterling, R. G. *Appl. Spectrosc.* **1982**, *36*, 665.
- Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193-1202.
- Hadni, A. *Essentials of Modern Physics Applied to the Study of the Infrared*; Pergamon Press: Oxford, 1967.
- Hailey, D. M.; Barnes, H. M.; Woodward, C.; Robinson, J. W. *Anal. Chim. Acta* **1971**, *56*, 161-174.
- Halaka, F. G.; Babcock, B. T.; Dye, J. L. *J. Biophys.* **1985**, *48*, 209-219.
- Handke, M.; Harrick, N. J. *Appl. Spectrosc.* **1986**, *40*, 401-405.
- Harrison, N.; Bilen, C. S.; Morantz, D. J. *Polymer Communications* **1984**, *25*, 15-17.
- Harrison, T. R. *Radiation Pyrometry and Its Underlying Principles of Radiant Heat Transfer*; Wiley: New York, 1960.
- Hartshorn, J. H. , Du Pont Marshall Labs., *Personal communication* **1989**.

- Helland, I. S. *Reports from the Departments of Mathematics and Statistics, Agricultural University of Norway*, 1986, 21, 44.
- Hirschfeld, T. *Appl. Opt.* 1978, 17, 1400-1412.
- Hommert, P. J.; Viskanta, R.; Chupp, R. E. *J. Am. Ceram. Soc.* 1975, 58, 58-62.
- Hoskuldsson, A. *J. Chemometrics* 1988, 2, 211-228.
- Hotelling, H. *Journal of Educational Psychology* 1933, 24, 417-441.
- Hudson, K. M.; Busch, K. W. *Anal. Chem.* 1988, 60, 2110-2115.
- Hudson, K. M.; Busch, K. W. *Anal. Chem.* 1987, 59, 2603-2609.
- Hunter, W. G. *Ind. Eng. Chem.* 1967, 6, 461-463.
- Huppi, E. R. *Mikrochim. Acta* 1987, 3, 281-296.
- Hvistendahl, J.; Rytter, E.; Oye, H. A. *Appl. Spectrosc.* 1983, 37, 182-187.
- Johansson, E.; Wold, S.; Sjodin, K. *Anal. Chem.* 1984, 56, 1685-1688.
- Jolliffe, J. *Principal Component Analysis*; Springer: Berlin, 1986.
- Jones, R. W.; McClelland, J. F. *Anal. Chem.* 1989, 61, 1810-1815.
- Jones, R. W.; McClelland, J. F. *Anal. Chem.* 1989, 61, 650-656.
- Jonson, B.; Rebenstorf, B.; Larsson, R.; Primet, M. *Appl. Spectrosc.* 1986, 40, 798-803.
- Kapff, S. F. *J. Chem. Phys.* 1948, 16, 446-453.
- Kember, D.; Chenery, D. H.; Sheppard, N.; Fell, J. *Spectrochim. Acta* 1977, 35A, 455-459.
- Kember, D.; Sheppard, N. *Appl. Spectrosc.* 1975, 29, 496-500.
- Kendall, D. J. W.; Clark, T. A. *Int. J. Infrared and Millimeter Waves* 1981, 2, 783.
- King, V. W.; Lauer, J. L. *Journal of Lubrication Technology* 1981, 103, 65-73.
- Koga, O.; Onishi, T.; Tamaru, K. *J.C.S. Chem. Comm.* 1974, 464.
- Kowalski, B. R. *Anal. Chem.* 1980, 52, 112R-122R.
- Kowalski, B. R., Editor-in-Chief; Brown, S. D., Vandeginste, B. G. M., Associate Editors *Journal of Chemometrics*; Wiley: Chichester.

- Kozlowski, T. R. *Appl. Opt.* **1968**, *7*, 795-800.
- Lauer, J. L.; Keller, L. E. *SPIE - Fourier Transform Infrared Spectroscopy* **1981**, 289, 87-93.
- Lauer, J. L.; King, V. W. *Infrared Physics* **1979**, *19*, 395-412.
- Lauer, J. L.; King, V. W. *Am. Soc. of Lubrication Engineers*; Preprint of paper presented at Anaheim Conference; **1980**.
- Lauer, J. L.; Vogel, P.; Seng, G. T. *Appl. Spectrosc.* **1985**, *39*, 997-1004.
- Lawton, W. H.; Sylvestre, E. A. *Technometrics* **1971**, *13*, 461-467.
- Leone, S. R. *Acc. Chem. Res.* **1989**, *22*, 139.
- Lin, L. T.; Archibald, D. D.; Honigs, D. E. *Appl. Spectrosc.* **1988**, *42*, 477-483.
- Lorber, A.; Wangen, L. E.; Kowalski, B. R. *J. Chemometrics* **1987**, *1*, 19-31.
- Low, M. J. D. *Nature* **1965**, *208*, 1089.
- Low, M. J. D.; Inoue, H. *Anal. Chem.* **1964**, *36*, 2397-2399.
- Low, M. J. D.; McManus, J. C.; Abrams, L. *Appl. Spectrosc. Rev.* **1971**, *5*, 171-210.
- Low M. J. D.; Clancy F. K. *Environ. Sci. Tech.* **1967**, *1*, 73-74.
- Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; John Wiley & Sons: New York, 1980.
- Mandel, J. *American Statistician* **1982**, *36*, 15-24.
- Manne, R. *Chemo. Intell. Lab. Sys.* **1987**, *2*, 187-197.
- Mantz, A. W.; Maillard, J. P. *J. Mol. Spectrosc.* **1974**, *53*, 466-478.
- Mantz, A. W.; Maillard, J. P.; Roh, W. B.; Rao, K. N. *J. Mol. Spectrosc.* **1975**, *57*, 155-159.
- Mardia, K. V.; Kent, J. T.; Bibby, J. M. *Multivariate Analysis*; Academic Press: London, 1979.
- Martens, H. A. "Multivariate Calibration: Quantitative Interpretation of Non-Selective Chemical Data, Parts I and II of the 1985 Ph.D. Dissertation"; The Norwegian Institute of Technology, University of Trondheim: Trondheim, Norway, NR-Report No.: 786 and 787, 1986, ISBN 82-539-0273 and 82-539-0274-3.
- Martens, H.; Naes, T. *Near-infrared Technology in the Agricultural and Food Industries*; Williams, P., Norris, K., Eds.; American Cereal Association, 1987; 57-87.

Martens, H.; Naes, T. *Trends Anal. Chem.* **1984**, *3*, 204-210.

Martens, H.; Naes Tormod *Multivariate Calibration*; John Wiley & Sons: New York, 1989.

Massart, D. L., Editor-in-Chief; Hopke, P. K., Spiegelman, C. H., Wegscheider, W. E., Editors; Brereton, R. G., Dessy, R. E., Associate Editors *Chemometrics and Intelligent Laboratory Systems*; Elsevier: Amsterdam.

Massart, D. L.; Vandeginste, B. G. M.; Deming, S. N.; Michotte, Y.; Kaufman, L. *Chemometrics: a textbook*; Elsevier: Amsterdam, 1988.

McDonald, R. S. *Anal. Chem.* **1984**, *56*, 349R-372R.

McDonald, R. S. *Anal. Chem.* **1986**, *58*, 1906R-1925R.

McDonald, R. S.; Wilks, J. P. A. *Appl. Spectrosc.* **1988**, *42*, 151-162.

McLean, D. D.; Pritchard, D. J.; Bacon, D. W.; Downie, J. *Technometrics* **1979**, *21*, 291-298.

McLean, D. D. *Technometrics* **1985**, *27*, 340-348.

McMahon, H. O. *J. Opt. Soc. Am.* **1950**, *40*, 376-380.

Mezaki, R.; Butt, J. B. *Ind. Eng. Chem.* **1968**, *7*, 120-125.

Naes, T.; Isaksson, T.; Kowalski, B. R. *Anal. Chem.* **1990**, *62*, 664-673.

Nagasawa, Y.; Ishitani, A. *Appl. Spectrosc.* **1984**, *38*, 168-173.

Nelder, J. A.; Mead, R. *Computer Journal* **1965**, *7*, 308-313.

Ohta, K.; Graf, R. T.; Ishida, H. *Appl. Spectrosc.* **1988**, *42*, 114-120.

Ohta, K.; Ishida, H. *Appl. Spectrosc.* **1988**, *42*, 952-957.

Ohta, K.; Ishida, H. *Appl. Opt.* **1990**, *29*, in press.

Osten, D. W. *J. Chemometrics* **1988**, *2*, 39-48.

Park, J. H.; Carli, B. *Appl. Opt.* **1986**, *25*, 3490-3501.

Pearson, K. *Philosophical Magazine* **1901**, *6*, 559-572.

Pell, R. J.; Erickson, B. C.; Hannah, R. W.; Callis, J. B.; Kowalski, B. R. *Anal. Chem.* **1988**, *60*, 2824-2827.

Prengle, H. W. J.; Morgan, C. A.; Fang, C.; Huang, L. *Current Research* **1973**, *7*, 417-423.

- Primet, M.; Fouilloux, P.; Imelik, B. *Surf. Sci.* **1979**, *85*, 457-470.
- Provder, T. J. *Coatings Technology* **1989**, *61*, 33.
- Ramos, L. S.; Beebe, K. R.; Carey, W. P.; Sanchez, E.; Erickson, B. C.; Wilson, B. E.; Wangen, L. E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 294R-315R.
- Reyment, R. A. *Chemo. Intell. Lab. Sys.* **1988**, *3*, 254.
- Rinsland, C. P.; Goldman, A.; Murcray, F. J.; Murcray, D. G.; Smith, M. A. H.; Seals, R. K. J.; Larsen, J. C.; Rinsland, P. L. *JQSRT* **1983**, *4*, 327-334.
- Roberts, R. ALCOA, Private Communication, 1988.
- Robinson, J. W.; Woodward, C.; Barnes, H. M. *Anal. Chim. Acta* **1968**, *43*, 119-128.
- Rytter, E.; Einarsrud, M. *Mikrochim. Acta.* **1988**, *2*, 307-309.
- Sanchez, E.; Kowalski, B. R. *J. Chemometrics* **1987**, *2*, 247-263.
- Sanchez, E.; Kowalski, B. R. *J. Chemometrics* **1988**, *2*, 265-280.
- Sanchez, E.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 496-499.
- Seasholtz, M. B.; Kowalski, B. R. *Appl. Spectrosc.* **1990**, *44*, 1337-1347.
- Sharaf, M. A.; Illman, D. L.; Kowalski, B. R. *Chemometrics*; John Wiley & Sons: New York, 1986.
- Sheppard, N. In *Analytical Applications of FT-IR to Molecular and Biological Systems: Proceedings of the NATO Advanced Study Institute held at Florence, Italy, August 31 to September 12, 1979*; Durig, J.R., Ed.; D. Reidel Pub. Co.: Hingham, MA, 1980; pp 125-140.
- Shrager, R. I. *Chemo. Intell. Lab. Sys.* **1986**, *1*, 59-70.
- Shrager, R. I. *Siam J. Alg. Disc. Meth.* **1984**, *5*, 351-358.
- Shrager, R. I.; Hendler, R. W. *Anal. Chem.* **1982**, *54*, 1147-1152.
- Small, G. W.; Kroutil, R. T.; Ditillo, J. T.; Loerop, W. R. *Anal. Chem.* **1988**, *60*, 264-269.
- Solomon, P. R.; Carangelo, R. M.; Hamblen, D. G.; Best, P. E. *Appl. Spectrosc.* **1986**, *40*, 746-759.
- Stewart, J. E. *Infrared Spectroscopy; Experimental Methods and Techniques*; M. Dekker: New York, 1970.
- Stierwalt, D. L. *Appl. Opt.* **1966**, *5*, 1911-1915.



- Stierwalt, D. L.; Potter, R. F. *In Report of the International Conference on the Physics of Semiconductors 1962*, 513-520.
- Stierwalt, J. B.; Bernstein, J. B.; Kirk, D. D. *Appl. Opt.* **1963**, *2*, 1169-1173.
- Stone, M. J. *Roy. Stat. Soc. Ser. B* **1974**, *36*, 111-148.
- Strang, G. *Linear Algebra and Its Application*; Academic Press, Inc.: 1988.
- Sundberg, R.; Brown, P. J. *Technometrics* **1989**, *31*, 365-371.
- Sylvestre, E. A.; Lawton, W. H.; Maggio, M. S. *Technometrics* **1974**, *16*, 353-368.
- Takeuchi, K.; Yanai, H.; Mukherjee, B. N. *Foundations of Multivariate Analysis A Unified Approach by Means of Projection Onto Linear Subspaces*; Wiley Eastern Limited: New Delhi, 1982, 425.
- Tenge, B. J. "F-Test Significance in Near-infrared Calibration and Near-infrared and Mid-infrared Characterization and Calibration Studies of Matrix-immobilized and Matrix-dispersed Compounds Bearing Amide Groups." Ph.D. Dissertation, University of Washington: Seattle, WA (1989). Available from: University Microfilms, Ann Arbor, MI.
- Tu, X. M.; Burdick, D. S.; Millican, D. W.; McGown, L. B. *Anal. Chem.* **1989**, *61*, 2219-2224.
- Van der Ven, L. G. J.; Van Dijk, J. H. *ACS Polym. Mat. Sci. and Eng.* **1988**, *58*, 267.
- Van Kasteren, P. H. G.; Smeets, L. H. *SPIE - Fourier Transform Infrared Spectroscopy* **1981**, *289*, 83-86.
- Van Woerkom, P. C. M.; Blok, P.; Van Veenendaal, H. J.; deGroot, R. L. *Appl. Opt.* **1980**, *19*, 2546-2550.
- VanWoerkom, P. C. M.; deGroot, R. L. *Appl. Opt.* **1982**, *21*, 3114-3118.
- Viskanta, R.; Anderson, E. E. *Advances in Heat Transfer* **1975**, *11*, 318.
- Viskanta, R.; Hommert, P. J.; Groninger, G. L. *Appl. Opt.* **1975**, *14*, 428-437.
- Wadayama, T.; Shiraishi, W.; Hatta, A.; Suetaka, W. *Appl. Surf. Sci.* **1990**, *44*, 43-47.
- Weber, D. J. *J. of Opt. Society of Am.* **1960**, *50*, 808-810.
- Weber, D. J. *J. of Opt. Soc. Am.* **1959**, *49*, 815-820.
- Wendt, W.; Alden, M.; Persson, W. *Appl. Spectrosc.* **1988**, *42*, 128-133.
- Wilmschurst, J. K. *J. Chem. Phys.* **1963**, *39*, 2545-2548.

Wold, H. *Multivariate Analysis*; Krishnaiah, P. R. Ed.; Academic Press: New York, 1966; Vol. 6, 391-420.

Wold, S. *Technometrics* **1978**, *20*, 397-405.

Wold, S.; Albano, C.; Dunn, W. J.; Esbensen, K.; Hellberg, S.; Johansson, E.; Sjoström, M. *Food Research and Data Analysis*; Martens, H. Russworm, Jr., H. Eds.; Applied Science Publishers: London, 1983; 147-199.

Wold, S.; Esbensen, K.; Geladi, P. *Chemo. Intell. Lab. Sys.* **1987**, *2*, 37-52.

Wyatt, C. L. *Radiometric System Design*; MacMillan Publishing Company: New York, 1987.

Yang, Y. S.; Lee, L. J. *J. Appl. Polym. Sci.* **1988**, *36*, 1325.

Zanzucchi, P. J.; Yim, M. W. *Appl. Spectrosc.* **1986**, *40*, 1042-1046.

Ziegel, E. R.; Gorman, J. W. *Technometrics* **1980**, *22*, 139-151.

## Appendix A NMR Determination of Vinyl Acetate Concentration in Selected Elvax® samples

Proton nuclear magnetic resonance spectroscopy (NMR) was used to confirm the vinyl acetate concentration in the Elvax® samples used in the infrared emission study discussed in Chapter 3. Samples analyzed were 28, 18, 15, 9.5 and 9.0% vinyl acetate concentration corresponding to the Du Pont product numbers 265, 450, 550, 770, and 750, respectively. Each sample was analyzed in triplicate by dissolving approximately 0.18 grams in enough o-dichlorobenzene to produce a solution of approximately 2.5 wt. %. All samples were soluble at 80 °C and the NMR spectra (200 MHz) were collected at 80 °C.

The spectrum showed four peaks at 0.5, 0.9, 1.1 and 1.8 ppm. The 1.8 ppm resonance is associated with the methyl protons on the acetate group, while the other three resonances are associated with the ethylenic protons, although only half of the 1.1 ppm peak area is associated with the ethylene portion of the molecule. The peak assignments are displayed in Table A.1

Given these resonances one third of the P4 area is proportional to the number of moles of vinyl acetate and the sum of one half P3 and P2 and P1 divided by four is proportional to the number of moles of ethylene. The following formulas state this mathematically

$$\text{Moles of vinyl acetate} \propto \frac{1}{3} * P2 \quad (\text{A.1})$$

$$\text{Moles of ethylene} \propto \frac{\frac{P3}{2} + P2 + P1}{4} \quad (\text{A.2})$$

The molar ratio of vinyl acetate to ethylene may be found by dividing equation A.1 by A.2. The weight percent is then computed as

$$\text{wt \% vinyl acetate} = \frac{86 * \text{mole ratio}}{86 * \text{mole ratio} + 28} * 100\% \quad (\text{A.3})$$

where 86 is the molecular weight of vinyl acetate and 28 is the molecular weight of ethylene.

**Table A.1** Peak assignments for NMR analysis of Elvax® samples.

Peak Label	Resonance Frequency (ppm)	Assignment
Peak 1 (P1)	0.5	Methyl associated with ethylene
Peak 2 (P2)	0.9	Ethylenic protons associated with polyethylene backbone
Peak 3 (P3)	1.1	Ethylenic
Peak 4 (P4)	1.8	Methyl protons of vinyl acetate group

Table A.2 summarizes the predicted and expected weight percent. An ANOVA was performed on the data where the different samples were the treatments. The within treatment variance estimate was found to be 0.83 wt%<sup>2</sup> giving an estimated standard deviation of 0.91 wt%.

**Table A.2** Expected and measured vinyl acetate concentrations.

<b>Du Pont Product Number</b>	<b>wt % Vinyl Acetate Expected</b>	<b>wt % Vinyl Acetate Found</b>
265	28	27.99
265	28	29.38
265	28	28.86
450	18	18.56
450	18	17.08
450	18	20.14
550	15	16.94
550	15	15.90
550	15	16.15
770	9.5	11.73
770	9.5	10.40
770	9.5	10.00
750	9	9.78
750	9	10.15
750	9	10.70

# **Appendix B Observations on the Effects of Closure, Linear Baselines and Mean Centering on Exploratory Multivariate Data Analysis**

## **B.1 Introduction to Appendix B**

Closed data sets are conventionally characterized by the unit-sum constraint. They occur in many areas of the natural sciences, for example, mineral compositions summing to 100 % in geochemical samples. According to Reyment (1) too little attention has been paid to the proper statistical analysis of these data, save a recent contribution by Aitchison (2,3). Reyment states that the incorrect application of statistical techniques to closed data sets will at best lead to misleading results and at worst to just plain wrong results. He characterizes the situation as grave. Misleading results are especially prevalent in multivariate statistical analysis of closed data sets. Aitchison has given the first unified treatment of the analysis of compositional data and offers alternative approaches which are distortion free. He demonstrates the utility of the log-ratio transformation applied to closed data for increased interpretability of principal components.

Modern analytical measurements may give rise to data that are closed in the conventional sense and further may give rise to a more subtle closure effect when systems are observed by making indirect multivariate measurements. If, for example, spectral measurements are recorded for a series of multicomponent samples, then the resulting data matrix may be represented as an outer product of two matrices, call them submatrices (assuming linearly additive data), either of which may have closure

in the conventional sense, and thus the product matrix (the observed data matrix ) will exhibit properties resulting from a "propagation" of the conventional closure.

Johansson et al. considers the consequence of normalization of the data (a class of closure as discussed below) with respect to information loss and the effects on the interpretability of principal component plots (4). Martens and Naes have recently commented on closure and its relationship to multivariate calibrations (5). The determination of pseudorank and its relationship to closure have been discussed by Cochran and Horne for the specific example of a kinetics experiment (6). They analyzed the dispersion and covariance matrix and gave a set of rules for relating the pseudorank as found from each analysis to the number of chemical species in a series of samples.

The purpose of this appendix is to discuss the closure problem within the context of analytical chemistry, specifically with respect to indirect multivariate measurements. The main thrust of this work is two fold 1) discuss the effect closure can have on data matrices that result from indirect multivariate observations, i.e., spectral data, specifically with respect to the interpretation of the pseudorank and 2) offer a simple systematic approach for the characterization of the submatrices making up the indirect observation matrix. The characterization procedure will clarify what statements may be made about the submatrices with respect to closure and/or baseline properties.

## B.2 Discussion

In this appendix scalars will be denoted by plain lowercase letters, vectors by bold lowercase letters, and matrices by bold uppercase letters. Vectors will be

column vectors, and row vectors will be represented as a column vector transposed (using a superscript T). The inner product of two vectors with an equal number of elements,  $\mathbf{x}^T\mathbf{y}$ , is a scalar, and the outer product,  $\mathbf{xy}^T$ , is a matrix.

The concept of rank will be involved in the discussion and so as to orient the reader the following definitions are given. Rank is the maximum number of outer products that may be computed from the decomposition of a data matrix and is less than or equal to the smaller of the number of rows or columns of the data matrix. A single outer product has rank one. Pseudorank is the number of "significant" outer products that may be computed and is therefore always less than or equal to the rank. A matrix is said to be full rank if the pseudorank = rank = min(rows, columns). In this appendix a matrix is considered rank deficient if the pseudorank < min (rows, columns).

It may be noted that most data matrices dealt with in analytical chemistry are rank deficient, in that the pseudorank for a 50 by 500 matrix may be only three. The pseudorank is associated with the number of chemical species for data that are linearly additive (e.g. obeys Beer's law in spectroscopy). For nonlinear data the pseudorank may or may not be very informative. There are many methods for pseudorank determination which will not be explored here (7-14). The following will distinguish two types of closure, demonstrate the equivalence of a baseline with one of the closure types, and propose a method to characterize data matrices with respect to closure and baseline properties.



### B.2.1 Distinction of Classes of Closure

The first class of closure that has been observed will be denoted as imposed closure. It is characterized by the imposition of a constant norm on each of the rows or columns of a data matrix or the imposition of an arithmetic relationship on the rows or columns that is the result of a known or suspected theoretical behavior, for instance a stoichiometric relationship. This type of closure is usually the result of a preprocessing step and is the conventional closure type. Three common processing procedures are 1) dividing each element in a row or column by the sum of the absolute values of all elements, 2) dividing each element by the square root of the sum of the squared element entries, and 3) dividing each element by the largest element. All of these operations may be thought of as imposing a particular unit vector norm.

In general, vector norms, known as the p-norms, are given for a vector  $x$  by

$$\|x\|_p = \left[ \sum_i (|x_i|)^{1/p} \right] \quad p = 1, 2 \dots \neq \infty ;$$

$$\|x\|_\infty = \max (|x_i|) \quad (B.1)$$

It may be seen that the three common preprocessing steps given above are simply the 1, 2 and infinity p norms. The imposition of a unit 1-norm, or division of each row or column entry by the sum of absolute values of all row or column entries, may be thought of as a unit area norm and is most commonly found in self-modeling curve resolution (15, 16). The 2-norm, division of each entry by the square root of

the sum of squared entries is the unit length norm and is commonly found as part of the autoscaling process. Autoscaling is used if the response variables have different units of measure. The infinity norm, division of each entry by the largest entry is common practice in the reporting of mass spectral data. The detection of the 1 and 2 norms is a trivial observation, the sums of the rows or columns add to a constant for the 1-norm or the diagonal entries for  $XX^T$  or  $X^TX$  are unity for a matrix with a constant 2-norm imposed on the rows or columns, respectively. Imposition of the infinity norm results in a loss of more information than imposition of the other two norms and may be observed by noting that the maximum entry in each row or column is one.

The imposition of an arithmetic relationship such as that resulting from stoichiometry may be detected by an eigenvector analysis. Box et al. have shown how exact linear relationships may be found by a comparison of the eigenvalues with the expected value of the eigenvalues (17).

The second class of closure will be denoted as propagated closure. This class of closure is found in matrices that can be written as the product of two other matrices, either or both of which have an equation that links the entries in the rows or columns. A linking equation for the columns of a matrix may be represented as

$$\sum_{j=1}^N a_{ij}\beta_j = 1 \quad (\text{for all } i) \quad (\text{B.2})$$

or in matrix notation

$$A\beta = \mathbf{1} \quad (\text{B.3})$$

where  $\mathbf{1}$  is a vector of ones and  $\beta$  is a vector of linking coefficients, the entries of which are not necessarily all equal. Note that a specific example of this representation of closure is the unit 1-norm provided the data matrix ( $A$  in equation B.3) has all positive elements and the  $\beta$  is a vector of ones. The 2-norm does not fall into this type of relationship as it is not possible to express the relationship between the rows or columns of a matrix that has a 2-norm relationship using equation B.3 because the linking coefficients are different for each row or column. Relationships involving stoichiometric consideration can be represented by equation B.3.

In practice this type of closure can be found, for example, in multivariate calibration where the response matrix  $X$  has propagated closure because it can be written as the product of a concentration matrix and a pure component response (spectral) matrix, in which the experiment has been designed such that the concentrations of the species in each sample sum to the same constant. A completely analogous situation occurs for the multivariate observation of kinetic experiments in which the reacting species always sum to a constant initial concentration of reactants. A matrix with propagated closure does not display any of the obvious properties of imposed closure (i.e. constant row or column sums etc.), but does display other more subtle effects.

### **B.2.2 Equivalence of Propagated Closure and a Baseline**

The detection of propagated closure is not as obvious as imposed closure and is further complicated by its equivalence to a simple baseline offset. A simple baseline offset is defined as the addition of the same vector to each of the rows or columns of a data matrix and may be represented as an additional outer product or

component of the form  $\mathbf{1}\mathbf{b}_r^T$ , for adding the vector  $\mathbf{b}_r$  to all the rows, and  $\mathbf{b}_c\mathbf{1}^T$  for adding the vector  $\mathbf{b}_c$  to all the columns. A more complex baseline offset would involve the addition of different amounts of a constant vector to each of the rows or columns. This type of offset may still be represented as an outer product, but with the vector of ones replaced by a nonconstant vector. Either type of offset may be treated as an additional component, but only the simple offset can be shown to be mathematically equivalent to propagated closure.

The equivalence of simple baseline offsets and propagated closure may be shown as follows. A matrix  $\mathbf{X}$ , with  $M$  rows and  $N$  columns representing two chemical components and containing a simple baseline offset added to each row may be written as

$$\mathbf{X} = \mathbf{C}\mathbf{R}^T + \mathbf{B}_r = \mathbf{c}_1\mathbf{r}_1^T + \mathbf{c}_2\mathbf{r}_2^T + \mathbf{1}\mathbf{b}_r^T \quad (\text{B.4})$$

This equation, as well as all subsequent equations, may represent either a noise free data matrix, where exactly three outer products can represent the data, or an approximation to a data matrix with measurement noise, where three "significant" outer products can represent the data. Compare equation B.4 to a matrix possessing a linking equation across the rows of the  $\mathbf{C}$  matrix containing three components which may be written as

$$\mathbf{X} = \mathbf{C}\mathbf{R}^T = \mathbf{c}_1\mathbf{r}_1^T + \mathbf{c}_2\mathbf{r}_2^T + \mathbf{c}_3\mathbf{r}_3^T \quad (\text{B.5})$$

Because the linking equation is given by  $\beta_1\mathbf{c}_1 + \beta_2\mathbf{c}_2 + \beta_3\mathbf{c}_3 = \mathbf{1}$ ,  $\mathbf{c}_1 = (\mathbf{1} - \beta_2\mathbf{c}_2 - \beta_3\mathbf{c}_3)/\beta_1$ , equation B.5 becomes

$$\mathbf{X} = \left( \frac{\mathbf{1} - \beta_2 \mathbf{c}_2 - \beta_3 \mathbf{c}_3}{\beta_1} \right) \mathbf{r}_1^T + \mathbf{c}_2 \mathbf{r}_2^T + \mathbf{c}_3 \mathbf{r}_3^T \quad (\text{B.6})$$

$$= \mathbf{c}_2 \left( \mathbf{r}_2^T - \frac{\beta_2}{\beta_1} \mathbf{r}_1^T \right) + \mathbf{c}_3 \left( \mathbf{r}_3^T - \frac{\beta_3}{\beta_1} \mathbf{r}_1^T \right) + \frac{1}{\beta_1} \mathbf{1} \mathbf{r}_1^T \quad (\text{B.7})$$

$$= \mathbf{c}_2 \tilde{\mathbf{r}}_2^T + \mathbf{c}_3 \tilde{\mathbf{r}}_3^T + \mathbf{1} \tilde{\mathbf{b}}_r^T \quad (\text{B.8})$$

which shows that mathematically closure linking the columns of  $\mathbf{C}$  (which span the column space of  $\mathbf{X}$ ) and a baseline added to every row of the matrix  $\mathbf{X}$  are equivalent (18). Similarly, a linking equation in the columns of  $\mathbf{R}$  (which span the row space of  $\mathbf{X}$ ) is equivalent to a baseline offset added to each column.

The above representation of closure as an outer product with a vector of ones suggests a method for detection. This may be performed by the projection of a column or row of ones onto the appropriate space. For a rank deficient matrix the projection may be computed by using a suitable representation of the column or row space, for example, the columns of  $\mathbf{U}$  or  $\mathbf{V}$ , respectively from the singular value decomposition of  $\mathbf{X} = \mathbf{USV}^T$  that has been appropriately truncated. If the operation of  $\mathbf{UU}^T$  or  $\mathbf{VV}^T$  on a vector of ones returns a vector of ones, then the column or row space, respectively has either closure or a baseline.

If the data were corrected for the baseline offset, then the above test could be used to indicate the presence of closure. This may be very helpful in a kinetics experiment in which the observation should be closed and a positive indication of closure for baseline corrected data would demonstrate that all the species in the observed reaction are giving rise to a signal. Clarification of what may be interpreted

from a change in pseudorank upon mean centering is found by the application of the following methodology.

### **B.2.3 Methodology for Rank Interpretation**

The pseudorank of a data matrix may or may not be related to the number of chemical components present in the samples depending on the structure of the matrix. A method for exploratory data analysis using mean centering is presented to aid in determining the properties of the submatrices making up a data matrix and to show how the pseudorank is related to the number of chemical components. It is appropriate at this point to precisely define a few terms. Column centering a matrix is defined as the subtraction of the mean of each column from every entry in that column. The result is a zero sum for each column. Similarly, row centering a matrix is defined as the subtraction of the mean of each row from every entry in that row. The result is a zero sum for each row.

Mean centering is used in the following three step procedure summarized in graphical form in Figure B.1. First, determine the pseudorank of the raw data matrix. This will determine if the matrix is full rank or rank deficient. Second, determine the pseudorank of the column centered matrix, and third, determine the pseudorank of the row centered matrix. The procedure can be followed in Figure B.1 for a matrix (M by N) with  $M \leq N$ . The results for a matrix (M by N) with  $M \geq N$  are obvious. An important distinction is drawn between full rank and rank deficient matrices, since each have different geometric properties with respect to centering.



The behavior of the full rank matrix with respect to centering is governed by a type of degrees of freedom restraint since the pseudorank equals the rank. This is in contrast to the rank deficient case where the pseudorank is less than the rank. As may be seen in Figure B.1 one of the possibilities for the full rank case is missing. This missing option corresponds to not having a rank reduction upon centering along the rank determining dimension (the lesser of the rows or columns). The reason the rank must always be reduced is a demonstration of the degree of freedom constraint mentioned above in that after centering,  $\sum_{i=1}^m \tilde{x}_{i,j} = 0$  for  $j = 1, \dots, N$ , which is a linear dependence, and will force rank reduction. Also, case two shown in Figure B.1 will only arise with nonsquare matrices for the same reason.

In Figure B.1 the row and column centering results produce six submatrix classifications. The closure and baseline properties associated with each position are shown pictorially in Figures B.2 through B.6.



$$\mathbf{B}_c = \left[ \begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \right] \left[ \begin{array}{ccc} 1 & \dots & 1 \end{array} \right]$$

$$\mathbf{B}_r = \left[ \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right] \left[ \begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \right]$$

### Case 1

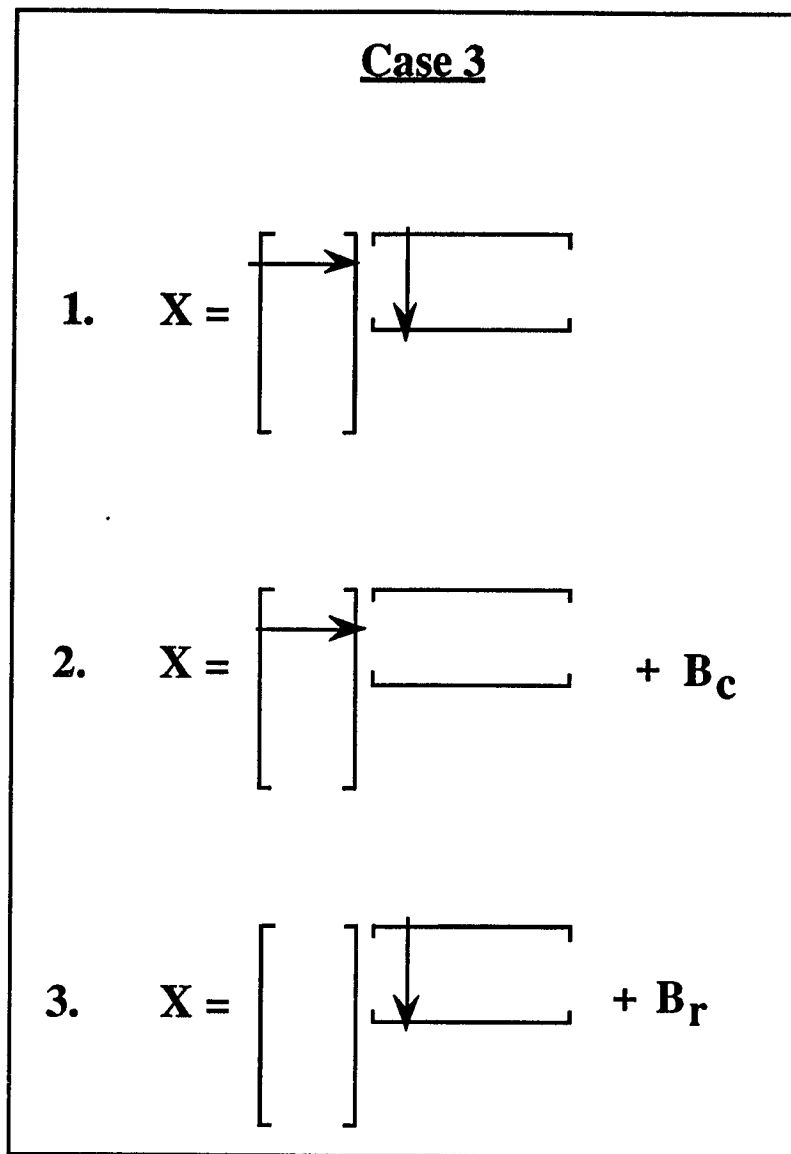
$$1. \quad \mathbf{X} = \left[ \begin{array}{c} \text{ } \\ \text{ } \end{array} \right] \left[ \begin{array}{c} \text{ } \\ \text{ } \end{array} \right] (+ \mathbf{B}_r) (+ \mathbf{B}_c)$$

$$2. \quad \mathbf{X} = \left[ \begin{array}{c} \text{ } \\ \text{ } \end{array} \right] \left[ \begin{array}{c} \text{ } \\ \text{ } \end{array} \right] (+ \mathbf{B}_c)$$

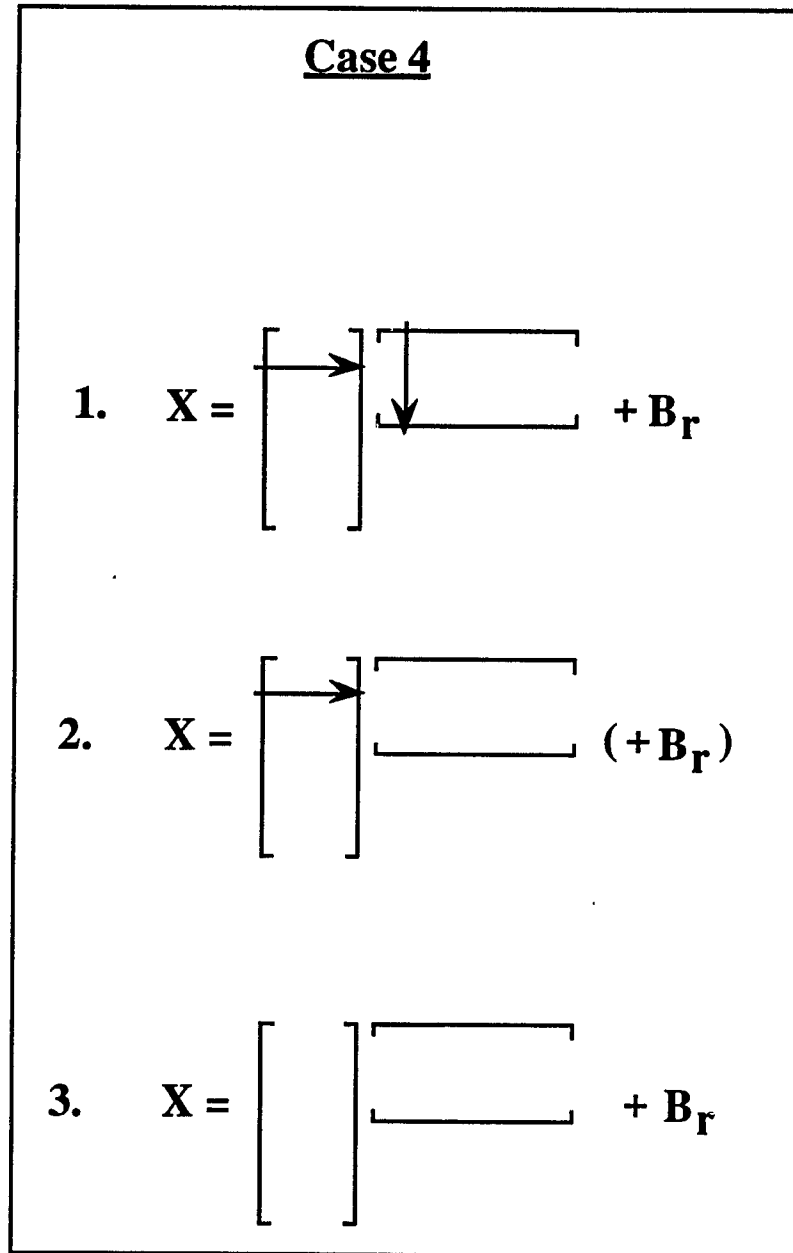
### Case 2

**All other full rank matrices**

**Figure B.2** Possible matrix structures corresponding to Cases 1 and 2.



**Figure B.3** Possible matrix structures corresponding to Case 3.



**Figure B.4** Possible matrix structures corresponding to Case 4.

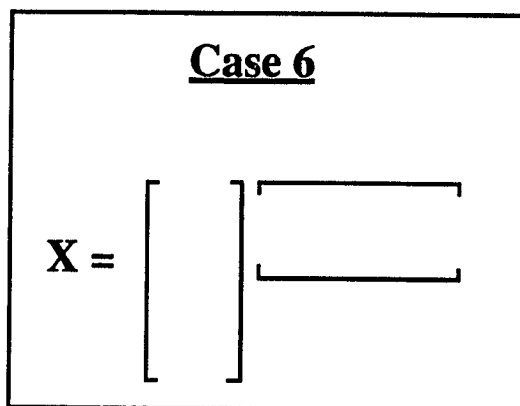
**Case 5**

$$1. \quad \mathbf{X} = \begin{bmatrix} \rightarrow & \downarrow \\ \downarrow & \rightarrow \end{bmatrix} + \mathbf{B}_c$$

$$2. \quad \mathbf{X} = \begin{bmatrix} & \downarrow \\ & \rightarrow \end{bmatrix} (+ \mathbf{B}_c)$$

$$3. \quad \mathbf{X} = \begin{bmatrix} & \\ & \end{bmatrix} + \mathbf{B}_c$$

**Figure B.5** Possible matrix structures corresponding to Case 5.



**Figure B.6** Possible matrix structures corresponding to Case 6.

The background matrices  $\mathbf{B}_r$  and  $\mathbf{B}_c$  are given as a vector added equally to all of the rows and columns respectively, and are defined pictorially in Figure B.2. The arrows through the matrices indicate where the closure relationship is located. For example, in Figure B.2, case 1 number 1, the horizontal arrow means that there is closure linking the vectors spanning the column space (the columns of  $\mathbf{C}$ ). Similarly, the vertical arrow means that there is closure linking the vectors spanning the row space (the rows of  $\mathbf{R}^T$ ). The parenthesis around the addition of the background indicates that the matrix will be in this category whether the background is present or not. Note that closure is not indicated along any of the non-rank-determining dimensions. Interpretation of the pseudorank however, does depend on the knowledge of the presence of the background. In all cases the smaller calculated (pseudo)rank represents the number of *independently* varying chemical components.

The statements that may be made regarding the closure and baseline characteristic of the submatrices after application of the three step procedure to the raw data matrix may be derived mathematically. These mathematical manipulations aid in the understanding of how the pseudorank can be interpreted for a given

situation. For example, consider the matrix  $\mathbf{X}$  in equation B.5, and assume the  $\beta_i = 1$  for all  $i$ . The rank of the raw matrix is three, since it may be written as the sum of three outer products (see equation B.8). A matrix with the mean of the columns in every row may be written as

$$\bar{\mathbf{X}}_c = \bar{c}_1 \mathbf{1} \mathbf{r}_1^T + \bar{c}_2 \mathbf{1} \mathbf{r}_2^T + \bar{c}_3 \mathbf{1} \mathbf{r}_3^T \quad (\text{B.9})$$

where  $\bar{c}_i$  are the means of the  $c_i$ . The column centered matrix  $\tilde{\mathbf{X}}_c$  is

$$\tilde{\mathbf{X}}_c = \mathbf{c}_1 \mathbf{r}_1^T + \mathbf{c}_2 \mathbf{r}_2^T + \mathbf{c}_3 \mathbf{r}_3^T - \left( \bar{c}_1 \mathbf{1} \mathbf{r}_1^T + \bar{c}_2 \mathbf{1} \mathbf{r}_2^T + \bar{c}_3 \mathbf{1} \mathbf{r}_3^T \right) \quad (\text{B.10})$$

Using the given closure information,  $\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 = \mathbf{1}$  and substituting for  $\mathbf{c}_1$ ,

$$\tilde{\mathbf{X}}_c = \left( \mathbf{1} - \mathbf{c}_2 - \mathbf{c}_3 - \bar{c} \mathbf{1} \right) \mathbf{r}_1^T + \left( \mathbf{c}_2 - \bar{c}_2 \mathbf{1} \right) \mathbf{r}_2^T + \left( \mathbf{c}_3 - \bar{c}_3 \mathbf{1} \right) \mathbf{r}_3^T \quad (\text{B.11})$$

$$\tilde{\mathbf{X}}_c = \mathbf{c}_2 \left( -\mathbf{r}_1^T + \mathbf{r}_2^T \right) + \mathbf{c}_3 \left( -\mathbf{r}_1^T + \mathbf{r}_3^T \right) + \mathbf{1} \left( \mathbf{r}_1^T - \bar{c}_1 \mathbf{r}_1^T - \bar{c}_2 \mathbf{r}_2^T - \bar{c}_3 \mathbf{r}_3^T \right) \quad (\text{B.12})$$

It can be shown that if  $\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 = \mathbf{1}$ ,  $\bar{c}_1 + \bar{c}_2 + \bar{c}_3 = 1$ . Substituting for  $\bar{c}_1$ ,

$$\begin{aligned} \tilde{\mathbf{X}}_c = & \mathbf{c}_2 \left( -\mathbf{r}_1^T + \mathbf{r}_2^T \right) + \\ & \mathbf{c}_3 \left( -\mathbf{r}_1^T + \mathbf{r}_3^T \right) + \mathbf{1} \left( \mathbf{r}_1^T - \left( 1 - \bar{c}_2 - \bar{c}_3 \right) \mathbf{r}_1^T - \bar{c}_2 \mathbf{r}_2^T - \bar{c}_3 \mathbf{r}_3^T \right) \end{aligned} \quad (\text{B.13})$$

$$\tilde{\mathbf{X}}_c = \left( \mathbf{c}_2 - \mathbf{1} \bar{c}_2 \right) \left( -\mathbf{r}_1^T + \mathbf{r}_2^T \right) + \left( \mathbf{c}_3 - \mathbf{1} \bar{c}_3 \right) \left( -\mathbf{r}_1^T + \mathbf{r}_3^T \right) \quad (\text{B.14})$$

which is rank 2. A matrix with the mean of the rows in every column may be written as

$$\bar{\mathbf{X}}_r = \bar{r}_1 \mathbf{c}_1 \mathbf{1}^T + \bar{r}_2 \mathbf{c}_2 \mathbf{1}^T + \bar{r}_3 \mathbf{c}_3 \mathbf{1}^T \quad (\text{B.15})$$

Row centering gives

$$\tilde{\mathbf{X}}_r = \mathbf{c}_1 \mathbf{r}_1^T + \mathbf{c}_2 \mathbf{r}_2^T + \mathbf{c}_3 \mathbf{r}_3^T - \left( \bar{r}_1 \mathbf{c}_1 \mathbf{1}^T + \bar{r}_2 \mathbf{c}_2 \mathbf{1}^T + \bar{r}_3 \mathbf{c}_3 \mathbf{1}^T \right) \quad (\text{B.16})$$

$$\tilde{\mathbf{X}}_r = \mathbf{c}_1 \left( \mathbf{r}_1^T - \bar{r}_1 \mathbf{1}^T \right) + \mathbf{c}_2 \left( \mathbf{r}_2^T - \bar{r}_2 \mathbf{1}^T \right) + \mathbf{c}_3 \left( \mathbf{r}_3^T - \bar{r}_3 \mathbf{1}^T \right) \quad (\text{B.17})$$

Using the closure information,

$$\tilde{\mathbf{X}}_r = (\mathbf{1} - \mathbf{c}_2 - \mathbf{c}_3) \left( \mathbf{r}_1^T - \bar{r}_1 \mathbf{1}^T \right) + \mathbf{c}_2 \left( \mathbf{r}_2^T - \bar{r}_2 \mathbf{1}^T \right) + \mathbf{c}_3 \left( \mathbf{r}_3^T - \bar{r}_3 \mathbf{1}^T \right) \quad (\text{B.18})$$

$$\begin{aligned} \tilde{\mathbf{X}}_r = & \mathbf{c}_2 \left( \mathbf{r}_2^T - \mathbf{r}_1^T + \left( \bar{r}_1 - \bar{r}_2 \right) \mathbf{1}^T \right) + \\ & \mathbf{c}_3 \left( \mathbf{r}_3^T - \mathbf{r}_1^T + \left( \bar{r}_1 - \bar{r}_3 \right) \mathbf{1}^T \right) + \mathbf{1} \left( \mathbf{r}_1^T - \bar{r}_1 \mathbf{1}^T \right) \end{aligned} \quad (\text{B.19})$$

which is rank 3. The matrix  $\mathbf{X}$  therefore is a member of case 4, specifically matrix 2 shown in Figure B.4. In this case, the pseudorank calculated for the raw matrix 3 is the number of chemical species giving rise to a signal. In the case that there is no

closure and a baseline is added to every row (case 4, matrix 3), the pseudorank of the raw matrix would be one more than the number of chemical species.

The utility of these results is best demonstrated by an example. Consider monitoring a chemical reaction at  $J$  wavelengths over a period of time collecting  $I$  spectra and saving the results in a matrix dimensioned  $I$  rows by  $J$  columns ( $I \leq J$ ). Suppose it is known that the background spectra in this particular type of spectroscopy are constant over all wavelengths, i.e. are simple offsets from zero, and so the first derivative of all the spectra are taken to remove the baseline. If the background spectra were linear the second derivative would be necessary to completely remove the baseline. Another method for removing baselines is to subtract an observed baseline. This is often used if the baseline is complex, and therefore the first or second derivatives cannot remove it. However, one must caution against using a noisy measurement of the baseline, as well as subtracting an incorrect amount of the baseline from the sample.

In this example, after completing the three step procedure described above, it was found that this matrix is an example of case 4. The matrices belonging to case 4 are shown in Figure B.4. Because all the background was removed by the derivative, the first and third matrix structures can be eliminated from consideration. Therefore, the matrix must have imposed closure linking the vectors spanning the column space, and the pseudorank calculated for the raw matrix represents the number of chemical components giving rise to a signal (recall equations (B.5 through B.8)). The number of chemical components is NOT equal to the column centered pseudorank because the rank reduction was due to the closure property.

From this analysis, the experimenter knows that all components involved in the reaction have been observed, which is information necessary for further modeling.



If the observed data fit into case 6 instead, this would indicate that there is no closure, and not all of the species in the reaction were observed. If the background had not been removed, the detection of closure would not be possible. If there was a nonzero background in all rows of the matrix (all spectra) the matrix would have belonged to case 4 whether there was closure or not (assuming there could not be closure linking the pure spectra which are the basis of the row space).

### B.3 Conclusions

In chemometrics the interpretation of the pseudorank can provide valuable information regarding the observed chemistry. If linear additivity of the data holds the pseudorank can represent the number of independently varying chemical components. However, without considering the possibility of a background or closure, incorrect conclusions may be drawn. In this discussion two classes of closure were presented. Further, the second class of closure, propagated closure, was shown to be equivalent to a simple background. A three step procedure was presented as a tool in exploratory data analysis to aid in the interpretation of the pseudorank.

## B.4 Notes to Appendix B

1. Reyment, R. A. *Chemo. Intell. Lab. Sys.* **1988**, *3*, 254.
2. Aitchison, J. *J. R. Statist. Soc. B* **1982**, *44*, 139.
3. Aitchison, J. *The Statistical Analysis of Compositional Data*; Chapman and Hall: London, 1986.
4. Johansson, E.; Wold, S.; Sjodin, K. *Anal. Chem.* **1984**, *56*, 1685.
5. Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.
6. Cochran, R. N.; Horne, F. H. *Anal. Chem.* **1977**, *49*, 846.
7. Stone, M. *J. Roy. Stat. Soc. Ser. B* **1974**, *36*, 111.
8. Wold, S. *Technometrics* **1978**, *20*, 397.
9. Eastment, H. T.; Krzanowski, W. J. *Technometrics* **1982**, *24*, 73.
10. Sylvestre, E. A.; Lawton, W. H.; Maggio, M. S. *Technometrics* **1974**, *16*, 353-368.
11. Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193-1202.
12. Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; John Wiley & Sons: New York, 1980.
13. Tu, X. M.; Burdick, D. S.; Millican, D. W.; McGown, L. B. *Anal. Chem.* **1989**, *61*, 2219-2224.
14. Osten, D. W. *J. Chemometrics* **1988**, *2*, 39-48.
15. Lawton, W. H.; Sylvestre, E. A. *Technometrics* **1971**, *13*, 461-467.
16. Borgen, O. S.; Kowalski, B. R. *Anal. Chim. Acta.* **1985**, *174*, 1.
17. Box, G. E. P.; Hunter, W. G.; MacGregor, J. F.; Erjavec, J. *Technometrics* **1973**, *15*, 33-51.
18. Strang, G. *Linear Algebra and Its Application*; Academic Press, Inc.: Orlando, 1988.

## **Vita**

Randall James Pell was born in Yakima, Washington on April 5, 1956, the second son to Glen and Darlean Pell. He graduated from West Valley High School in Yakima in 1974 and spent two years at Yakima Valley Junior College before transferring to Washington State University (W.S.U.) in 1976. He graduated from W.S.U. in 1979 with a B.S. in Chemistry and obtained an M.S. degree in Chemistry from W.S.U in 1982 studying under the direction of Dr. John Hunt.

In April of 1982 Randy accepted the position of Chemist with Conoco Inc. in the Gas Chromatography Group of the Research Services Division, located in Ponca City, Oklahoma. He was promoted to Research Chemist in 1986 and at that time joined the Mass Spectrometry Group. His responsibilities included high resolution capillary gas chromatography and capillary gas chromatography/mass spectrometry for the analysis of petroleum related materials.

In August of 1986 Randy married Karen Acs, a Research Chemist in the Liquid Chromatography Group at Conoco Inc.. In July of 1986 Randy and Karen resigned from Conoco and moved to Seattle, Washington, where Randy joined Dr. Bruce Kowalski's Chemometrics research group. In August of 1988 Randy was awarded Best Student paper in the Symposium for Innovation in Measurement Science (SIMS) in Geneva, New York. In September of 1988 Randy was awarded the Outstanding Graduate Student Award sponsored by the Dow Chemical Company. Randy was awarded an ACS Analytical Division Full Year Fellowship (1989-1990) sponsored by the Perkin Elmer Corporation. Randy completed his studies with Dr. Kowalski in September of 1990 and he and Karen accepted positions as Senior Research Chemists with the Dow Chemical Company, Michigan Division, in Midland, Michigan.