1996

FLOW INJECTION AND CHEMOMETRIC TECHNIQUES FOR THE ON-LINE MONITORING OF INDUSTRIAL LIQUID EFFLUENTS

ANDREW, KEVIN NEIL

FLOW INJECTION AND CHEMOMETRIC TECHNIQUES FOR THE ON-LINE MONITORING OF INDUSTRIAL LIQUID EFFLUENTS

by

KEVIN NEIL ANDREW

A thesis submitted to the University of Plymouth

in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

Department of Environmental Sciences

Faculty of Science

In collaboration with

Brixham Environmental Laboratory (Zeneca Limited)

ICI Chemicals & Polymers (Runcorn)

ICI Engineering (Winnington)

February 1996

ABSTRACT

FLOW INJECTION AND CHEMOMETRIC TECHNIQUES FOR THE ON-LINE MONITORING OF INDUSTRIAL LIQUID EFFLUENTS

KEVIN NEIL ANDREW

The legal requirement to monitor discharges of harmful substances in industrial waste waters is presented in Chapter One, which also discusses the merits of using automated on-line analytical instruments for this purpose. Flow injection analysis with solid-state UV-visible detection is proposed as a potential on-line effluent monitoring technique, and the principles and advantages of this methodology are summarised.

Chapter Two describes the development of a portable, automated FI monitor for on-line determination of ammonia in liquid effluents. The development process culminates with deployments of the system at two chemical production sites, and validated results are presented for on-line analyses of real effluents.

The principles of multivariate calibration of spectrophotometric data are summarised in Chapter Three, and five commonly applied techniques (DMA, SMLR, PCR, PLS1 and PLS2) are described and compared. These multivariate calibration techniques are then applied in Chapter Four for the quantification of metal ions in model effluent systems, using diode-array spectral data sets. The relative predictive performances of the techniques are compared for both simple and more complex multicomponent systems.

Flow injection and multivariate calibration techniques are combined in Chapter Five, in which the development of a method for the determination of BTEX compounds in effluents is described. UV absorbance spectra are obtained for synthetic aqueous mixtures using an FI-diode array system, and SMLR, PCR, PLS1 and PLS2 are employed to quantify individual and total BTEX compounds. An FI solvent extraction method is also described for the analysis of a real effluent matrix.

The thesis concludes with an examination of a recursive digital filtering technique which has potential applications for on-line effluent monitoring. Chapter Six describes the principles of the Kalman filter, and presents results for both multivariate calibration and baseline drift correction of multicomponent spectral data sets, performed using different forms of the Kalman filter algorithm.

ACKNOWLEDGEMENTS

Many people have advised and supported me during the course of this work, and none more so than Paul Worsfold, whose guidance for the last three and a half years and regular prodding during the last few months are much appreciated.

I am very grateful for the financial support provided by ICI Chemicals & Polymers, ICI Engineering and Brixham Environmental Laboratory (Zeneca Limited), and for the considerable assistance offered by a number of people at the three sites. I would particularly like to acknowledge Dave Barnes, Mike Comber, Malcolm McIvor, Mike Crane, Brian Harland and Phil Norman for this.

My gratitude is also extended to Paul MacLaurin (Zeneca FCMO), whose advice has been given freely at various stages of the work, and whose expense account has been diminished on my behalf at several fine Indian restaurants in Plymouth and Huddersfield.

I would like to thank Sarah Rutan for giving me the opportunity to visit Virginia Commonwealth University for two months, and for her valuable advice and guidance both during and since my time in Richmond.

My work has been made all the more enjoyable by the various people I have been lucky enough to share a lab with at Plymouth, including Nick (Monica's dad), Bloxy (Monica's new dad), Simon L., Wavey (and WPL), Anthony, Trevvy-Trev (my brother in writing-up torment), Tony, Heft, Bert, Dazz, Matt, Andy-Mare, Acid-Man Tom, Simon C. and Cristina.

I also want to thank Carol for her invaluable support and friendship during the three years, and to say "merci beaucoup - tu es vachement chouette" to Albine.

Lastly I want to thank my parents for their unending support, and for continually asking me why I still hadn't finished for the last four months!

AUTHOR'S DECLARATION

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

This study was financed with the aid of a Polytechnics and Colleges Funding Council (PCFC) Research Initiative Research Studentship, and with CASE equivalent support contributed by three collaborating organisations: Brixham Environmental Laboratory (Zeneca Limited), ICI Chemicals and Polymers (The Heath, Runcorn) and ICI Engineering (Winnington, Northwich).

A programme of advanced study was undertaken, which included a two-month study visit to Virginia Commonwealth University (Richmond, VA, USA), and short courses on Unscrambler multivariate analysis software and digital signal processing methods. Relevant scientific seminars and conferences were regularly attended at which work was often presented; external institutions were visited for consultation purposes, and five papers were prepared for publication.

Publications:

1 Andrew, K.N, and Worsfold, P. J., *Analyst*, 1994, **119**, 1541-1546.

2 Andrew, K. N., Blundell, N. J., Price, D., and Worsfold, P. J., *Anal. Chem.*, 1994, **66**, 916A-922A.

3 MacLaurin, P., Andrew, K. N., and Worsfold, P. J., in *Process Analytical Chemistry*, eds. McLennan, F., and Kowalski, B. R., Chapman & Hall, London, 1995, ch. 5, pp. 159-182.

4 Andrew, K. N., Worsfold, P. J., and Comber, M., *Anal. Chim. Acta*, 1995, **314**, 33-43.

5 Andrew, K. N., and Worsfold, P. J., *Anal. Proc.*, 1995, **32**, 507-510.

Conferences Attended and Presentations:

1 ICI Chemometrics Club meeting, ICI Agrochemicals, Jealotts Hill, October 1992.

2 RSC Molecular Spectroscopy Group meeting, Perkin Elmer, Seer Green, November 1992.

3 RSC Chemometrics Group meeting, University of Wolverhampton, February 1993.

4 RSC Analytical Division meeting, ICI Chemicals and Polymers, Wilton, May 1993.

5 R and D Topics Meeting, University of Bradford, July 1993, *poster presentation.*

6 SCI Environment and Water Group/Institute of Measurement and Control/RSC Analytical Division meeting, SCI, London, March 1994.

7 R and D Topics Meeting, University of Hertfordshire, July 1994, *poster presentation.*

8 Chemistry Dept. research colloquium, Virginia Commonwealth University, Sept. 1994, *oral presentation.*

9 208th American Chemical Society National Meeting and Exposition, Washington, DC, USA, August 1994.

10 R and D Topics Meeting, University of Hull, July 1995, *oral presentation.*

11 SAC 95, University of Hull, July 1995, *poster presentation.*

External Contacts:

Dr Michael Comber, Brixham Environmental Laboratory;

Dr David Barnes, ICI Chemicals & Polymers;

Mr Malcolm McIvor, ICI Engineering.

Signed.....

Date.....29/3/96.....

For Mum and Dad

CONTENTS

Abstract i

Acknowledgements iii

Author's declaration iv

List of Tables xi

List of Figures xiv

Chapter 1 INTRODUCTION

1.1 THE ANALYSIS OF INDUSTRIAL LIQUID EFFLUENTS 1

1.1.1 Liquid effluent monitoring and environmental control 1

1.1.2 Legislation regarding liquid effluent discharges 1

1.1.3 On-line monitoring 8

1.2 FLOW INJECTION ANALYSIS 10

1.2.1 Fundamental principles 10

1.2.2 FI instrumentation and methodologies 12

1.2.3. Advantages of FI for on-line analysis 21

1.2.4 Industrial applications for on-line FI monitoring 22

1.3 SOLID-STATE UV/VISIBLE SPECTROPHOTOMETRY 27

1.3.1 Light emitting diode photometry 27

1.3.2 Photodiode array spectrophotometry 30

1.3.3 Charge transfer devices 33

1.4 RESEARCH OBJECTIVES 35

Chapter 2 FLOW INJECTION DETERMINATION OF AMMONIA IN INDUSTRIAL LIQUID EFFLUENTS

2.1 INTRODUCTION 36

2.2 THE PRINCIPLES OF GAS DIFFUSION IN FI METHODS 38

2.3 EXPERIMENTAL 42

2.3.1 Reagents 42

2.3.2 Instrumentation 43

2.3.3 Portable FI monitor 44

2.3.4 Procedures 48

2.3.5 The indophenol blue reference method 49
2.4 RESULTS AND DISCUSSION 49
2.4.1 Method optimisation 49
2.4.2 Alternative acid-base indicators 54
2.4.3 Buffering capacity of sodium hydroxide reagent on acidic samples ... 58
2.4.4 Interferences 59
2.4.5 Effect of membrane variability on response 61
2.4.6 Portable monitor development 62
2.4.7 Stability trial 68
2.4.8 On-line industrial site trials 71
2.5 CONCLUSIONS 76

Chapter 3 MULTIVARIATE CALIBRATION

3.1 INTRODUCTION 77
3.2 UNIVARIATE VERSUS MULTIVARIATE CALIBRATION 78
3.2.1 Univariate calibration 78
3.2.2 Multivariate calibration 80
3.3 MULTIVARIATE CALIBRATION TECHNIQUES 81
3.3.1 Direct multicomponent analysis 81
3.3.2 Multiple linear regression 84
3.3.3 Principal components regression 86
3.3.4 Partial least-squares regression 90
3.3.5 Summary of multivariate calibration techniques 94

Chapter 4 COMPARISON OF MULTIVARIATE CALIBRATION TECHNIQUES FOR THE QUANTIFICATION OF METAL IONS IN MODEL EFFLUENT STREAMS

4.1 INTRODUCTION 95
4.2 EXPERIMENTAL DESIGN 96
4.3 EXPERIMENTAL 98
4.3.1 Reagents 98
4.3.2 Instrumentation 98
4.3.3 Software 99
4.3.4 Design of sample sets 99
4.3.5 Procedures 105

4.3.6 Unscrambler model validation 106
4.4 RESULTS AND DISCUSSION 108
4.4.1 Spectral characteristics 108
4.4.2 Three-component system 109
4.4.3 Four-component system 112
4.4.4 Five-component system 114
4.4.5 Copper-iron interaction 116
4.4.6 Summary of overall precisions for multicomponent absorbance data sets 116
4.4.7 First-derivative data 117
4.5 CONCLUSIONS 119

Chapter 5 FI WITH MULTIVARIATE CALIBRATION FOR THE QUANTIFICATION OF BTEX COMPOUNDS IN MODEL SYSTEMS AND EFFLUENTS

5.1 INTRODUCTION 120
5.2 EXPERIMENTAL 121
5.2.1 Reagents 121
5.2.2 Instrumentation 121
5.2.3 Software 122
5.2.4 Procedures 125
5.3 RESULTS AND DISCUSSION 129
5.3.1 Spectral characteristics 129
5.3.2 Static determination of toluene, ethylbenzene and *o*-xylene in hexane solution 132
5.3.3 Static determination of toluene, ethylbenzene and *o*-xylene in aqueous solution 132
5.3.4 Simplex optimisation of a single-channel FI manifold 134
5.3.5 FI determination of toluene, ethylbenzene and *o*-xylene in aqueous solution 138
5.3.6 On-line FI-solvent extraction procedure 141
5.4 CONCLUSIONS 145

Chapter 6 THE APPLICATION OF KALMAN FILTERING METHODS TO MULTIVARIATE CALIBRATION AND DRIFT CORRECTION

6.1 INTRODUCTION 146

6.2 THEORY OF THE KALMAN FILTER 147

6.3 EXPERIMENTAL 151

6.3.1 Software 151

6.3.2 Procedures 151

6.4 RESULTS AND DISCUSSION 155

6.4.1 Multivariate calibration of multicomponent transition metal mixtures 155

6.4.2 Detection of baseline drift in multicomponent calibration spectra 160

6.5 CONCLUSIONS 168

Chapter 7 CONCLUSIONS AND FUTURE WORK

7.1 FINAL CONCLUSIONS 170

7.2 SUGGESTIONS FOR FUTURE WORK 172

REFERENCES 174

APPENDICES

1 BASIC Program for Automated Control of the Portable FI Ammonia Monitor 184

2 Matlab Program for the Standard Kalman Filter 192

3 Matlab Program for the Modified Kalman Filter 195

4 Matlab Program for the Determination of Baseline Drift Using the Kalman Filter 199

LIST OF TABLES

1.1 List I and List II families of substances as defined by European Community Directive 76/464/EEC. 3

1.2 Emission standards for mercury from the chloralkali industry, as specified by daughter Directive 82/176/EEC 4

1.3 Prescribed substances for discharge to water, as defined by the Environmental Protection Act 1990. 6

1.4 Various options for the three principal FI components. 13

1.5 Process FI applications classified by area and analyte 23

1.6 Characteristics of some commercially available LEDs 29

2.1 Summary of published flow injection methods for the analysis of ammonia in liquid samples 37

2.2 Power consumption of the portable monitor components 47

2.3 FI manifold conditions used by Clinch *et al.* for the determination of ammonia in river water 50

2.4 Results for the simplex optimisation of the FI manifold using 10 mg l^{-1} NH_3-N standard 51

2.5 Effect of injection volume on linear response range 52

2.6 Effect of relative flow rates on linear response range 52

2.7 Optimised FI manifold conditions and linear regression parameters for linear response in the range 0-100 mg l^{-1} NH_3-N 53

2.8 Working ranges and pK values for three sulfonephthalein indicators 54

2.9 Two-variable optimisation results for phenol red and bromocresol purple indicators 56

2.10 Comparison of linear ranges obtained using different sulfonephthalein indicators 57

2.11 Some physicochemical properties of methyl and ethyl amines in comparison to those of ammonia 59

2.12 Summary of analysis of variance for the membrane variability assessment 61

2.13 Effect of reducing carrier and indicator flow rates on the linear response of the portable monitor 63

2.14 Monitor reagent consumption for a typical (6-injection) analytical cycle 65
2.15 Linear regression parameters for determination of portable monitor detection limits 66
2.16 Portable monitor precision 66
2.17 Predictive performance of the monitor for ammonia standard solutions 67
2.18 Predictive performance of the monitor for spiked industrial effluent samples, compared with the indophenol blue method 67
2.19 Effluent pH and a comparison of monitor and indophenol blue results for Site 2 74

3.1 Definition of notation used for multivariate calibration algorithms 82
3.2 The principal characteristics of DMA, MLR, PCR and PLS 94

4.1 Factorial design of the calibration and test sets for the two-component system 100
4.2 Factorial design of the calibration and test sets for the three-component system 100
4.3 Fractional factorial design of the calibration and test sets for the four-component system 102
4.4 Fractional factorial design of the calibration and test sets for the five-component system 103

5.1 Concentration data (μmol l^{-1}) for the three-component system in hexane 126
5.2 Concentration data (mg l^{-1}) for the three-component system in aqueous solution for static determinations 127
5.3 Concentration data (mg l^{-1}) for the three-component system in aqueous solution for FI determinations 128
5.4 Results for the simplex optimisation of the FI manifold using a 1 mg l^{-1} toluene standard in aqueous solution 137
5.5 Linear regression results for absorbances at 210 nm of extracted *o*-xylene standard solutions and *o*-xylene-spiked effluent solutions 144

6.1 Precisions for the three-component system 156
6.2 Precisions for the three-component system with $BaSO_4$ interference 156

6.3 Precisions for the four-component system 157
6.4 Precisions for the four-component system with $BaSO_4$ interference 157
6.5 Precisions for the two-component system 158
6.6 Precisions for the five-component system 159
6.7 Precisions for the five-component system with $BaSO_4$ interference 159
6.8 One-component system with linear baseline drift component 161
6.9 One-component system with random baseline drift component 162
6.10 Three-component system with linear baseline drift component 164
6.11 Three-component system with random baseline drift component 165
6.12 Three-component training set (including replicates) with random baseline drift component 167
6.13 RRMSEC values for a three-component calibration model with random drift and RRMSEP values for predictions of new sample concentrations 168

LIST OF FIGURES

1.1 Schematic diagram of a simple single-channel flow injection manifold, showing the transient nature of the signal output. 11

1.2 Effect of injection volume on sample dispersion for a 0.05 g l^{-1} bromothymol blue solution at pH 11.0 (tube length = 50 cm). 12

1.3 Effect of tube length on sample dispersion for a 0.05 g l^{-1} bromothymol blue solution at pH 11.0 (injected sample volume = 70 µl). 12

1.4 Diagram of an eight-roller peristaltic pump unit 14

1.5 Schematic diagram of a six-port rotary injection valve 15

1.6 Schematic diagram of an automated FI monitor 17

1.7 Schematic diagram of a four-channel manifold 18

1.8 Four-channel manifold with reagent injection 18

1.9 Schematic diagram of a two-channel/three-channel parallel manifolds system for simultaneous determinations 19

1.10 Schematic diagram of a multi-channel manifold for simultaneous determination, using post-injection sample splitting 20

1.11 Schematic diagram of an automated process FI monitoring system 22

1.12 Pie chart showing distribution of published process FI methods by area of application 26

1.13 Normalised emission spectra of some commercially available light emitting diodes 28

1.14 Schematic diagram of a photodiode array spectrophotometer 31

2.1 Diagram of a typical gas diffusion cell for FI analysis 40

2.2 Effect of pH on dissolved ammonia speciation at 25°C 40

2.3 Structural formulae of the acidic (HI^-) and basic (I^{2-}) forms of bromothymol blue indicator 41

2.4 FI manifold for the determination of total ammonia in liquid effluents 43

2.5 Diagram of the gas diffusion cell 44

2.6 Basic layout of the portable FI monitor 46

2.7 Linear response range achieved using optimised manifold parameter settings 53

2.8 Absorbance spectra of (a) bromothymol blue, (b) bromocresol purple and (c) phenol red at various pH levels, showing acidic, mid-transition and basic forms of the indicator 55
2.9 Effect of decreasing sample pH on response 58
2.10 Comparison of blank-subtracted absorbances for pure and mixed solns. of ammonia, methylamines and ethylamines 60
2.11 Mean absorbance values obtained for a 2 mg l^{-1} NH_3-N solution using various membrane strips 62
2.12 Plot of baseline noise for the portable monitor during an 8-hour period of continuous operation 64
2.13 Linear calibration for determination of portable monitor detection limits 65
2.14 Results for continuous analyses of a 20 mg l^{-1} NH_3-N solution during a seven-day stability trial of the portable monitor 68
2.15 Schematic diagram of the sample delivery system for on-line effluent monitoring 72
2.16 Typical absorbance profile of the liquid industrial effluent sample used to evaluate the predictive performance of the monitor 72
2.17 Results of the on-line monitor trial at Site 1 73
2.18 Results of the on-line monitor trial at Site 2, compared with those obtained manually by the indophenol blue method 73
2.19 Correlation between the monitor and the indophenol blue results for Site 2 75

3.1 (a) Matrix **A** plotted in column space; (b) first and second principal components for **A** following mean-centring and variance scaling of the columns 87

4.1 Optimal model dimensionality (PCR, PLS1 and PLS2) defined as (a) the first local minimum in the value of PRESS, and (b) according to an *F*-statistic comparison of significant differences 107
4.2 Absorbance spectral profiles for the five metal ion solutions 108
4.3 Precisions for the three-component system 110
4.4 (a) Absorbance spectra of the three-component calibration samples; (b) spectra for the same solutions after the addition of varying amounts of $BaCl_2$, indicating the effect of $BaSO_4$ interference 111
4.5 Precisions for the three-component system with $BaSO_4$ interference 112

4.6 Precisions for the four-component system 113
4.7 Precisions for the four-component system with $BaSO_4$ interference 113
4.8 Precisions for the five-component system 114
4.9 Precisions for the five-component system with $BaSO_4$ interference 115
4.10 Precisions for the two-component (copper and iron) system 116
4.11 Comparison of the overall precisions for each multicomponent system..... 117
4.12 Comparison of overall precisions for the five-component system with barium sulfate interference using first-derivative and absorbance data 118

5.1 Schematic diagram of the single-channel FI manifold used for the determination of BTEX compounds in aqueous model systems..... 123
5.2 Schematic diagram of the modified FI manifold used for solvent extraction of aqueous *o*-xylene solutions and spiked effluent samples..... 123
5.3 Diagram of the solvent extraction cell..... 124
5.4 FI response profile (absorbance against wavelength and time) for a solution containing toluene, ethylbenzene and *o*-xylene each at 20 mg l^{-1}..... 129
5.5 UV absorbance spectral profiles for toluene, ethylbenzene and *o*-xylene in hexane (20 μmol l^{-1} solutions)..... 130
5.6 UV absorbance spectral profiles for toluene, ethylbenzene and *o*-xylene in aqueous (10 mg l^{-1}) solutions..... 131
5.7 Predictive precisions for static determinations of the three-component model system in hexane..... 133
5.8 Predictive precisions for static determinations of the aqueous three-component model system 135
5.9 Simplex optimisation for a two-parameter system..... 136
5.10 Results for a two-parameter optimisation of the FI manifold using 1 mg l^{-1} toluene standard in aqueous solution 138
5.11 Predictive precisions for FI determinations of the aqueous three-component model system 139
5.12 Predictive precisions for total TEX in the three-component aqueous model system (FI determination)..... 140
5.13 UV spectrum of a typical industrial effluent sample..... 141
5.14 UV spectra for aqueous *o*-xylene standards following on-line FI extraction 143

5.15 UV spectra for effluent samples spiked with *o*-xylene following on-line FI extraction 143
5.16 UV spectra for three-component aqueous TEX solutions following on-line FI extraction 144

6.1 Graphic representation of the Kalman filter algorithm 149
6.2 RRMSEC for drift-corrected one-component calibration models (linear drift)... 161
6.3 RRMSEC for drift-corrected one-component calibration models (random drift) 162
6.4 RRMSEC for drift-corrected three-component calibration models (linear drift)..... 166
6.5 RRMSEC for drift-corrected three-component calibration models (random drift) 166

Chapter 1

Introduction

1.1 THE ANALYSIS OF INDUSTRIAL LIQUID EFFLUENTS

1.1.1 Liquid effluent monitoring and environmental control

Water is one of the most important commodities for many manufacturing industries, and its varied uses include incorporation in products as a raw material, cooling of manufacturing processes and washing of production vessels. Such applications inevitably result in the production of waste water, the composition of which will be dependant on the nature of the industry concerned. Industrial waste water effluent is normally discharged to natural water bodies (*e.g.* rivers and estuaries) or sewerage systems [1].

The last 20-30 years have witnessed a growing concern over the quality of the natural environment, and a consequent desire to improve or preserve the quality of aquatic, atmospheric and terrestrial systems. Increasingly stringent legislation has been (and continues to be) introduced for this purpose. Industrial waste waters are subject to such legislative control, and an industrial site wishing to discharge a liquid effluent into the environment must adhere to certain conditions relating to effluent composition and rate of discharge.

Regular sampling and analysis of liquid effluents are necessary to ensure that the terms of the legislation are being met. As a growing number of parameters of effluent composition become subject to legislative control, it is necessary to develop analytical methods that are capable of accurately quantifying these parameters in potentially harsh and complex sample matrices.

1.1.2 Legislation regarding liquid effluent discharges

In the United Kingdom, the earliest significant article of legislation referring to industrial waste water discharges was the Rivers Pollution Prevention Act 1876, which specified that it was a criminal offence to allow "any poisonous, noxious or polluting liquid" resulting

from an industrial processes to enter a natural freshwater body [2 -3]. The circumstances under which this could be enforced in an area of manufacturing industry were considerably restricted however. This Act remained in force until it was superseded by the Rivers (Prevention of Pollution) Act 1951, which gave River Boards the power to grant effluent discharge consents, specifying conditions to be met by sites wishing to emit a liquid effluent into an inland water body (this applied to new rather than existing discharges). The Clean Rivers (Estuaries and Tidal Waters) Act 1960 extended these powers to cover some estuaries and tidal bodies. The concept of discharge consents was further refined by the Control of Pollution Act 1974 (commonly referred to as COPA), which empowered the ten regional Water Authorities formed in England and Wales the previous year (and the River Purification Boards in Scotland) to issue consents for effluent discharges (excluding sewage) to surface, ground and coastal waters. Consents for sewage effluents were granted by the Secretary of State, as sewage treatment and disposal was the responsibility of the water authorities.

Since the introduction of COPA (which was not fully implemented until the latter half of the 1980s), it has been necessary for new environmental legislation in the UK to reflect the demands of European Community Directives. Community Member States are required to implement legislation within a given time period to enforce the stipulations of each Directive. The most significant Directive applying to industrial effluents is 76/464/EEC, commonly referred to as the Dangerous Substances in Water Directive [2, 4]. This was first notified on 5 May 1976, and is applicable to:

> "...the discharge by man, directly or indirectly, of substances or energy into the aquatic environment, the results of which are such as to cause hazards to human health, harm to living resources and to aquatic ecosystems, damage to amenities or interference with other legitimate uses of water."

This Directive classifies families of potentially polluting compounds as List I and List II substances (see Table 1.1). List I contains substances considered most dangerous owing to their toxicity, persistence and bioaccumulation properties, while List II comprises

substances considered less harmful, but suspected nonetheless of having a deleterious effect on the aquatic environment.

Table 1.1 List I and List II families of substances as defined by European Community Directive 76/464/EEC [5].

List I substances	List II substances
Organohalogen compounds (and substances which may form organohalogens in the aquatic environment)	The following metals/metalloids and their compounds: Zn, Cu, Ni, Cr, Pb, Se, As, Sb, Mo, Ti, Sn, Ba, Be, B, U, V, Co, Tl, Te, Ag
Organophosphorus compounds	Biocides (and derivatives) not given in List I
Organotin compounds	Substances with a deleterious effect on taste/smell of products for human consumption derived from natural waters
Compounds with proven carcinogenic properties in or *via* the aquatic environment	Toxic or persistent organic compounds of silicon (and substances which may produce the aforementioned compounds in water), apart from those which are biologically harmless
Mercury and its compounds	Phosphorus and inorganic phosphorus compounds
Cadmium and its compounds	Cyanides and fluorides
Persistent mineral oils and petroleum hydrocarbons	Non-persistent mineral oils and petroleum hydrocarbons
Persistent synthetic substances which float, remain in suspension or sink, and therefore interfere with any water usage.	Substances with an adverse effect on aquatic oxygen balance (particularly ammonia and nitrites)

The Dangerous Substances in Water Directive requires Member States to take appropriate action to eliminate emissions of List I substances, and to introduce programmes to reduce emissions of List II substances. However, all List I substances are treated as List II substances until such time as secondary (or 'daughter') Directives are introduced which

specify emission standards for particular List I substances. Discharges of any listed substance must be authorised by a competent licensing authority. Emission standards for discharges of List II substances are determined with reference to quality objectives set out by other Directives according to the type of receiving water body (*e.g.* bathing water; surface water intended for the abstraction of drinking water). In the case of List I substances, daughter Directives allow each Member State the choice of either imposing limit values which emission standards must not exceed regardless of the type of receiving body, or to set emission standards relating to quality objectives specified by the daughter Directive. An example of this is given in Table 1.2.

Table 1.2 Emission standards for mercury from the chloralkali industry, as specified by daughter Directive 82/176/EEC [2].

Limit Values[a]	
Concentration in all Hg-containing discharges:	50 $\mu g\ l^{-1}$
Quantity per tonne of installed Cl capacity	
'Recycle brine' process (production unit discharges):	0.5 g
'Recycle brine' process (total Hg in all mercury-containing discharges):	1.0 g
'Lost brine' process (total Hg in all mercury-containing discharges):	5.0 g
Quality Objectives[b]	
Fish flesh:	0.3 mg kg^{-1} (wet flesh)
Inland surface waters:	1.0 $\mu g\ l^{-1}$
Estuary waters:	0.5 $\mu g\ l^{-1}$
Sea and coastal waters:	0.3 $\mu g\ l^{-1}$

[a] monthly average limit values to be met following 1 July 1986

[b] arithmetic mean values obtained for 12 months to be met following 1 July 1983

In the UK, many of the requirements of Directive 76/464/EEC were fulfilled by the existing COPA 1974 legislation, with regional Water Authorities and Scottish River Purification Boards responsible for authorising discharges of the listed dangerous substances using the consent procedure. The UK chose to use environmental quality objectives relating to intended water usage as the appropriate means of determining discharge limits for effluents.

The Water Act 1989 created the National Rivers Authority (NRA), a new regulatory authority for England and Wales, with the regional Water Authorities becoming privatised Water Services. The Water Services were responsible for water supply and sewage collection/disposal, while the NRA adopted all regulatory functions, including the authorisation of discharge consents. Further implementations of the requirements of the Dangerous Substances in Water Directive and its subsequent daughter Directives have been brought about with the introductions of the Surface Waters (Dangerous Substances) (Classification) Regulations 1989 and 1992, the Environmental Protection Act 1990, and the Water Resources Act 1991 [3-4].

The Environmental Protection Act 1990 controls discharges of "prescribed" (*i.e.* dangerous) substances from pre-defined "prescribed" industrial processes to atmospheric, aquatic and terrestrial systems, with Her Majesty's Inspectorate of Pollution (HMIP) responsible for enforcing the Act. Prescribed substances for release into water are those listed in Table 1.3, while prescribed industrial processes include fuel and power (*e.g.* gasification and combustion processes), waste disposal (*e.g.* incineration), minerals (*e.g.* cement and asbestos), chemical (*e.g.* petrochemicals and pharmaceuticals), metal production and certain other (*e.g.* paper and uranium) manufacturing industries. In cases of discharges to controlled waters (see definition below), HMIP is required to consult with the NRA to establish appropriate discharge limits.

Table 1.3 Prescribed substances for discharge to water, as defined by the Environmental Protection Act 1990 [3].

Mercury and its compounds	Aldrin, Dieldrin and Endrin	Tributyltin compounds
Cadmium and its compounds	Polychlorinated biphenyls	Triphenyltin compounds
Hexachlorocyclohexane isomers	Dichlorvos	Trifluralin
DDT isomers	1,2-Dichloroethane	Fenitrothion
Pentachlorophenol and its compounds	Trichlorobenzene isomers	Azinphos-methyl
Hexachlorobenzene	Atrazine	Malathion
Hexachlorobutadiene	Simazine	Endosulfan

The Water Resources Act 1991 is the principal legislation currently applied to the control of water pollution. Section 85 of the Act states that it is an offence to cause or knowingly permit any poisonous, noxious or polluting matter or solid waste to enter any controlled water. Controlled waters are defined as territorial waters extending seawards for three miles, coastal waters extending landwards as far as the highest tide limit or the fresh water limit of a river or watercourse, inland fresh waters (lakes, ponds and rivers/watercourses above the fresh water limit) and ground waters. Liability under section 85 is excluded under section 88 if the emission into a controlled water is the subject of a discharge consent, a prescribed process authorisation under the Environmental Protection Act 1990, or any other statutory power of discharge. Discharge consents, normally issued by the NRA, are defined in section 91 as:

> "...a consent for any discharge or description of discharges given for the purposes of section 88(1) either on application or.....without application."

Effluent discharges in breach of or not covered by a consent renders the responsible party liable to pollution offence proceedings under section 85 [4]. Depending on the individual case, a consent may include conditions relating to the location of the discharge, the design

of the outlet, the type, composition, temperature, volume and rate of the discharge, requirements for sampling and/or *in situ* monitoring, and the keeping of records and information relating to the discharge [3].

In the case of industrial effluent discharges to sewerage systems rather than controlled waters, it is the sewerage undertaker (*i.e.* the regional Water Service) which is authorised to issue discharge consents, under the terms of the Water Industry Act 1991. An industrial site applying for such a consent must state the composition, maximum daily volume and highest proposed rate of discharge for the effluent. The consent, if granted, will state conditions reflecting the details supplied in the application, and may include additional terms relating to permitted times of day for discharges, payment to the Water Service for reception and disposal of the effluent, requirements for sampling and monitoring, and the keeping of records. If the proposed discharge is to include prescribed substances as defined by the Environmental Protection Act 1990, the sewerage undertaker must consult HMIP prior to granting the consent [4].

It is proposed that the current functions of the NRA and HMIP, together with those of regional Waste Regulation Authorities, will eventually be undertaken by the Environmental Protection Agency, a new unitary authority which it is estimated will be established by 1996. The Environmental Protection Agency will therefore be responsible for authorising industrial discharges under the terms of the Environmental Protection Act 1990 and the Water Resources Act 1991.

A further modification to the current system of discharge regulation has been proposed in a recent report by the Royal Commission on Environmental Pollution [1]. The report recommends the introduction of an incentive charging scheme to accompany discharge consents, in order to reinforce existing legislation, encourage new emission control technology and provide a further incentive to dischargers to reduce their harmful emissions below the regulatory limits. This system, if adopted, would apply to all point

source discharges which are subject to consent, with charges determined according to the volume and composition of the effluent. Discharges in excess of the consent levels would automatically incur higher charges, while lower charge rates would be offered as an incentive to reduce discharges to levels significantly below those specified by the consent.

1.1.3 On-line monitoring

In the fields of both environmental and industrial process monitoring, the analysis of liquid samples has traditionally entailed manual collection of samples from the point of interest. These samples are transported to a central laboratory facility, where they are logged and stored until such time as an analysis can be performed by the appropriate technician. The advantage of this approach is that it permits the use of sophisticated analytical instrumentation, operated by highly trained individuals with an expert knowledge of both the instruments and the samples, and an ability to interpret any unusual analytical results which may be encountered. However, a number of disadvantages are also associated with manual sampling and analysis [6 -7]. The delay between the times of sample collection and laboratory analysis, which may be hours or even days, can result in losses of volatile components or degradation of unstable sample determinands prior to analysis. This delay in obtaining analytical results also precludes the option for interactive control of dynamic systems such as industrial process and effluent streams, which may result in unnecessary wastage of materials and energy in the case of process streams, and breaches of discharge consent conditions in the case of effluents. A further problem is that manual sampling and analysis is both time-consuming and costly, particularly if samples need to be collected or analysed outside normal working hours. The cost factor will therefore restrict the number of samples collected, and this will limit the information available for a rapidly changing system.

To overcome these problems, automated on-line analysers are now increasingly being used to monitor process and effluent streams [8 -9]. These enable effective control of the system of interest by providing regular information on system composition, thus allowing prompt remedial action to be taken if undesirable conditions are detected. On-line analysis can therefore improve process efficiency and reduce costs, minimise waste production and ensure that effluent discharges are maintained within legal limits.

An on-line analyser can be defined as an instrumental system installed at some point alongside a process or effluent stream, which automatically draws samples from the stream and performs the required measurement [10]. Analytical results can then be communicated to a central process control computer, which makes adjustments to the manufacturing process or effluent treatment system as appropriate.

A number of laboratory analytical techniques have been adopted for on-line monitoring, including chromatography (gas, liquid and supercritical fluid methods), optical spectroscopy (near infrared, Fourier transform infrared, ultraviolet/visible and Raman) nuclear magnetic resonance spectroscopy, X-ray fluorescence and mass spectroscopy [9-10]. In all cases, on-line analysers must be of sufficiently robust and rugged design to withstand potentially harsh sample matrices, corrosive or dusty atmospheric conditions and fluctuations in temperature and humidity, and still be capable of providing precise and reliable analytical data. Other important design criteria are self-calibration procedures, minimal maintenance requirements and an appropriate selection of sampling point (and hence analyser location) [8, 11].

In addition to the techniques mentioned above, flow injection analysis, in combination with a wide range of detection methods, is increasingly being applied to on-line or *in situ* monitoring of process and environmental parameters. The principles of this technique and its suitability for on-line monitoring are discussed in the following section.

1.2 FLOW INJECTION ANALYSIS

1.2.1 Fundamental principles

Since its development in the mid-1970s, flow injection (FI) analysis has become a routine laboratory technique for sample presentation and on-line sample treatment [12 , 13 , 14]. FI was first described by Růžička and Hansen in 1975 [15], and is essentially a technique involving the insertion of a volume of liquid sample (typically 10-200 µl) into an unsegmented, continually flowing liquid carrier stream. Following injection, the sample zone undergoes physical dispersion as it is pumped along a narrow-bore tube (typically 0.5-0.8 mm) to a flow-through detector for measurement of a specific physico-chemical parameter. If the carrier stream also contains a reagent, then a zone of dispersed reaction product is formed. The detected response is in the form of a transient peak, the height of which is usually directly related to analyte concentration. A schematic representation of a simple, single-channel FI manifold is given in Figure 1.1, which illustrates the basic components of such a system. These typically comprise a propulsion unit (*e.g.* a peristaltic pump), a six-port rotary injection valve and a flow-through detector (*e.g.* a spectrophotometer). Poly(tetrafluoroethylene) (PTFE) tubing is commonly used, with tightly-wound coils often included to aid mixing

The sample dispersion process is highly reproducible, and can be controlled by adjusting operating variables such as flow rate, manifold geometry, tubing length and diameter. The degree of dispersion is quantified in terms of the dispersion coefficient (D):

$$D = C^{o}/C^{max}$$

where C^{o} is the concentration of an analyte in the sample prior to dispersion, and C^{max} is the maximum concentration in the dispersed sample zone at the time of detection. In terms of spectrophotometric absorbance measurements, C^{o} therefore corresponds with the absorbance of a pure sample stream, while C^{max} corresponds with the absorbance peak

maximum of an injected sample. Dispersion is generally defined as limited for D = 1-3, medium for D = 3-10 and large for D > 10 [14]. The degree of dispersion is manipulated to optimise analytical performance (*e.g.* to reduce detection limits), with limited dispersion generally used to increase sensitivity, while large dispersion is employed when on-line sample dilution is required prior to measurement. Figure 1.2 demonstrates that larger injection volumes produce more limited dispersion, while Figure 1.3 shows dispersion increasing with the length of tubing along which the sample zone travels prior to detection. Higher dispersion coefficients are produced by more rapid flow rates, which increase axial dispersion owing to the frictional forces generated between the flowing stream and the tubing. The incorporation of coiled or knitted lengths of tubing increases the degree of radial mixing between the sample and the carrier/reagent, but minimises axial dispersion.

Figure 1.1 Schematic diagram of a simple single-channel flow injection manifold, showing the transient nature of the signal output. ——— represents liquid flow; ····· represents data flow.

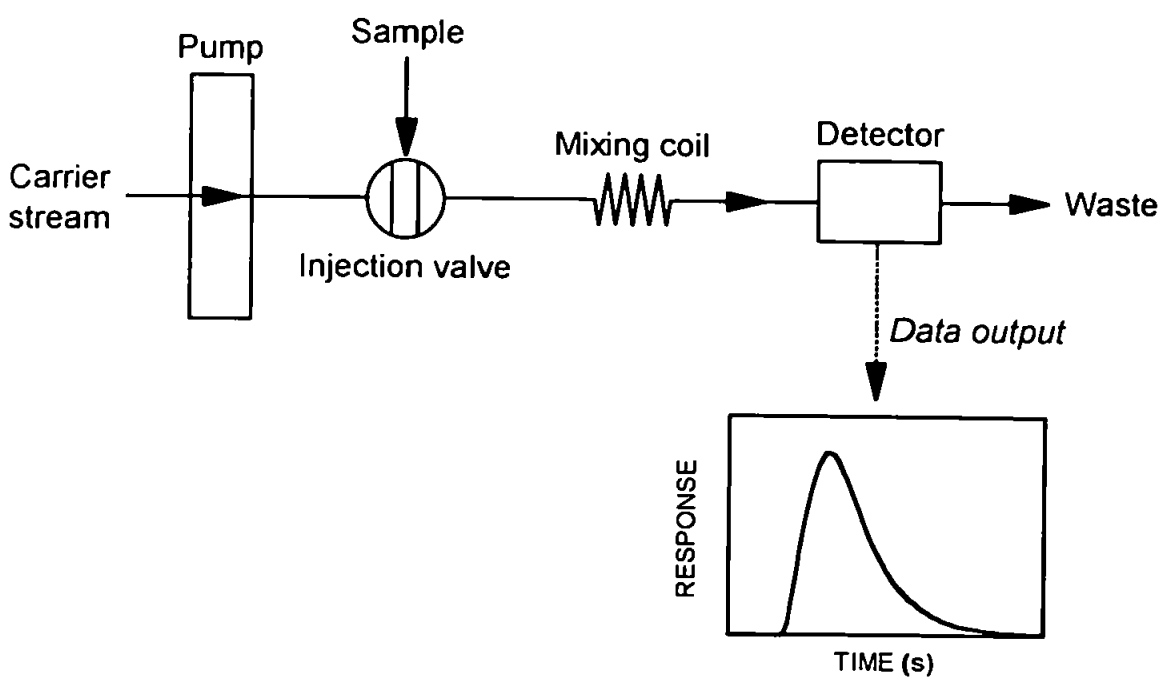


Figure 1.2 Effect of injection volume on sample dispersion for a 0.05 g l^{-1} bromothymol blue solution at pH 11.0 (tube length = 50 cm).

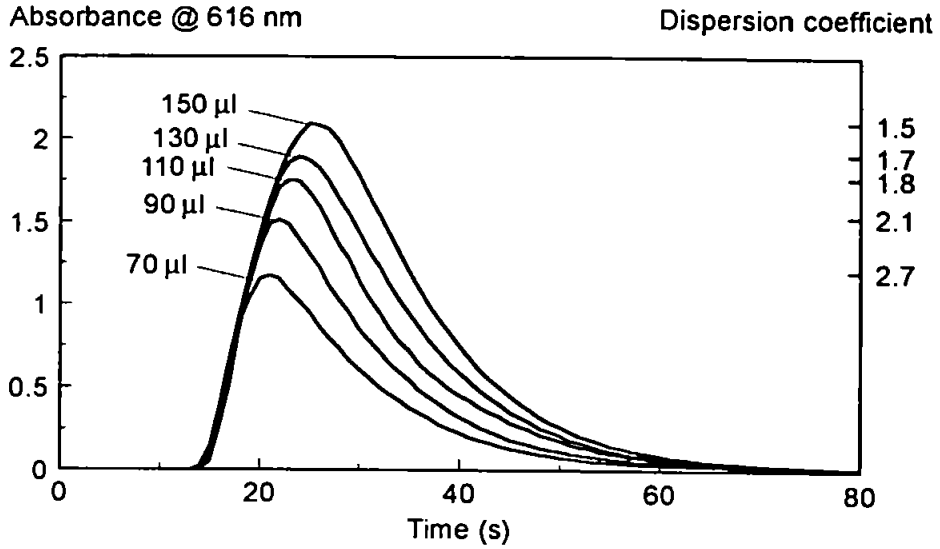
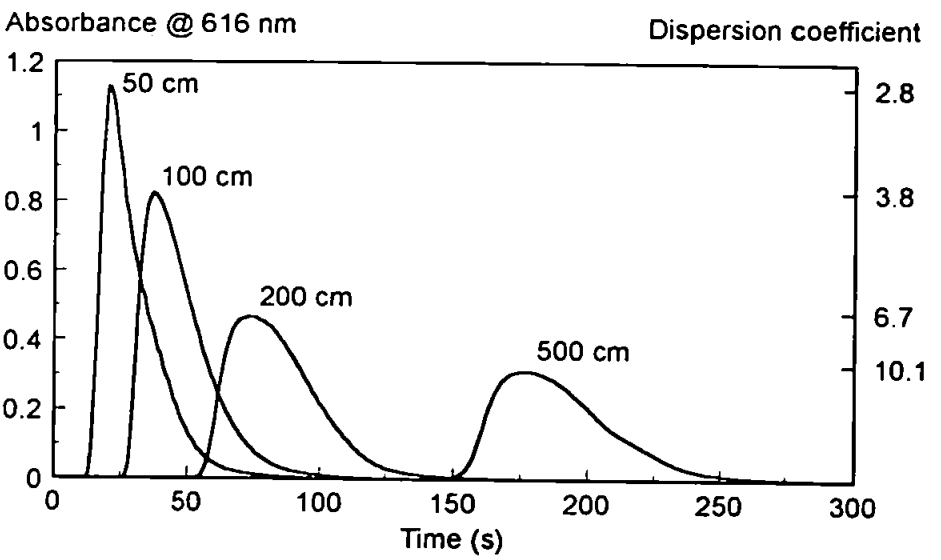


Figure 1.3 Effect of tube length on sample dispersion for a 0.05 g l^{-1} bromothymol blue solution at pH 11.0 (injected sample volume = 70 µl).



1.2.2 FI instrumentation and methodologies

As mentioned above, three essential components of an FI manifold are the propulsion, injection and detection systems. A number of options are available for each of these basic elements, and a summary is given in Table 1.4.

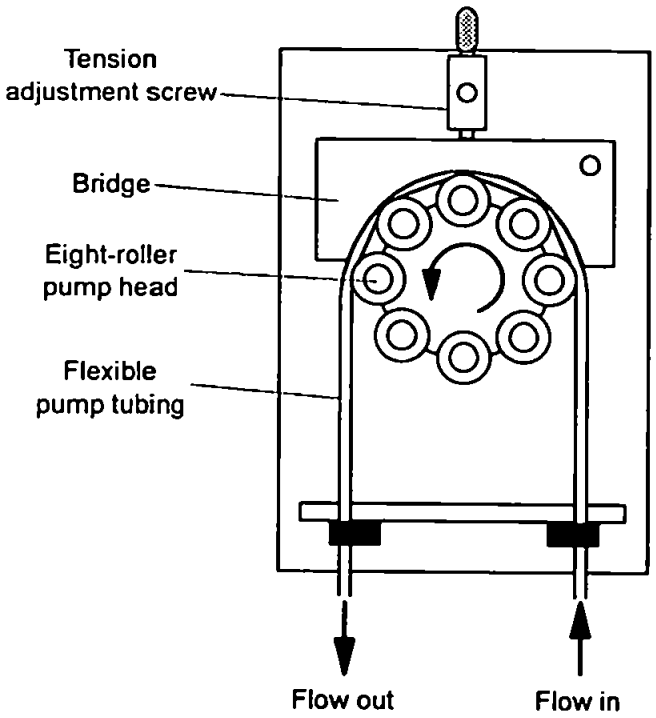
Table 1.4 Various options for the three principal FI components.

Component	Options	Description
Propulsion system	Peristaltic pump	Set of rollers on a revolving drum, which squeeze flexible tubing to produce a constant, pulsing flow.
	Gas pressurised vessel	Pressurised inert gas vessel connected via a flow regulator to each reagent/carrier reservoir, producing pulseless flow.
	Reciprocating pump	Reciprocating piston pumping fluid through a small chamber, with valves alternately opening and closing to control flow through the chamber. Produces pulsing flow.
	Piston pump	Computer-controlled, cam-driven piston , which produces bi-directional, variable speed, precise and pulseless flow.
Injection system	Rotary valve	Six-port unit incorporating a sample loop, which can be switched between filling and emptying positions. Electric or pneumatic operation.
	Hydrodynamic injection	Involves the selective stopping and starting of a sample pump and a reagent pump, with sample entering the reagent stream while the latter is stopped, then transported into the manifold when it is restarted.
	Multiposition selector valve	Multi-port unit allowing sequential selection of a number of flow streams (*e.g.* sample, standard and reagent streams). Electric operation.
Detection system	Optical	*e.g.* UV-visible spectrophotometry; solid-state photometry; diode array spectrophotometry; IR spectrophotometry; fluorimetry; chemiluminescence, atomic spectrometry.
	Electrochemical	*e.g.* Potentiometry (ion-selective and pH electrodes); conductimetry; amperometry; coulometry; voltammetry.

Since reproducible dispersion is a fundamental aspect of FI, the **propulsion system** must be capable of providing consistent flow patterns throughout the manifold. The most frequently used flow pattern is continuous forward linear flow, although alternative methods include stopped-flow (in which flow is halted when the sample zone reaches the detector), intermittent flow and flow reversal, all of which may be used with non-linear flow patterns. The choice of propulsion system must therefore reflect the desired flow type. Peristaltic pumps are most frequently used, since these are relatively inexpensive, reliable and robust, and can be applied to continuous flow, stopped/intermittent flow and flow reversal methods.

Propulsion is achieved by the peristaltic action of a rotating multi-roller drum (typically 8-10 rollers in close proximity) compressing a length (or several parallel lengths) of flexible tubing against a bridge (see Figure 1.4). A constant but slightly pulsing flow is produced, although the pulsing effect can be minimised by using pumps with no less than eight rollers. Flow rates are typically in the range 0.1-5.0 ml min^{-1}, and are proportional to both the rate of drum revolution and the internal diameter of the pump tubing, which is typically fabricated from poly(vinyl chloride) (PVC). Modified PVC, silicone rubber and thermally set fluorine rubber are also used as pump tubing materials in cases where the stream is an organic solvent or a concentrated mineral acid. A large peristaltic unit can be used for pumping as many as 16 separate flow channels, each at different rates.

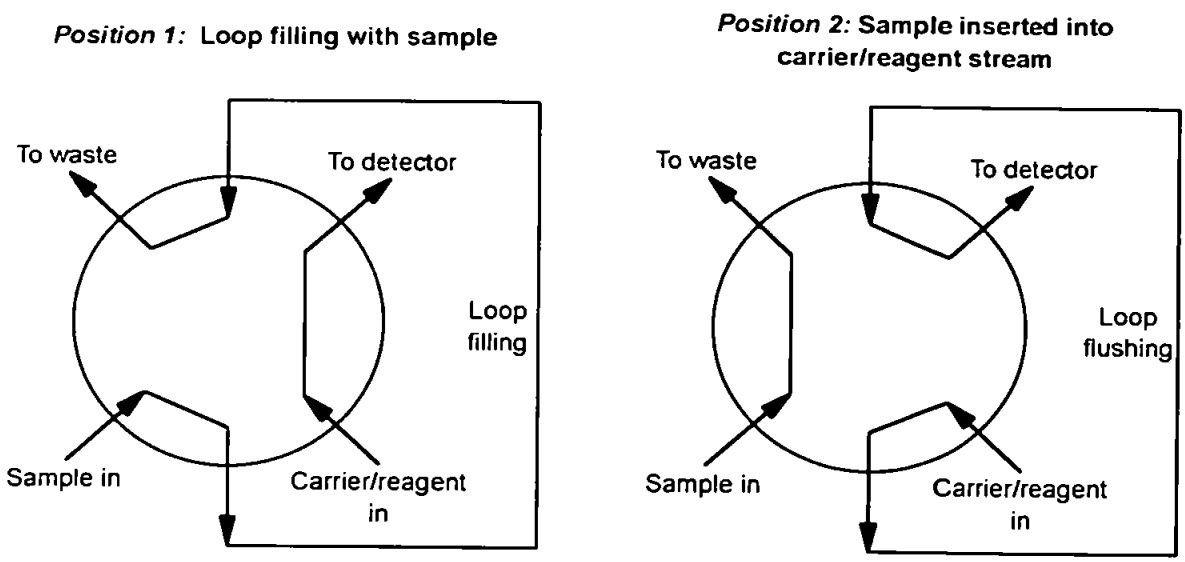
Figure 1.4 Diagram of an eight-roller peristaltic pump unit.



Analytical reproducibility in FI is also dependent upon the precision and accuracy of the **injection system**, which must be capable of inserting a pre-determined volume of sample into the carrier or reagent stream. Low pressure rotary injection valves are

commonly used for this purpose, since these are of low cost and can provide a high degree of precision and very low maintenance requirement. Rotary injection devices are usually constructed from PTFE, and have six ports for flow input or output. Two ports are used for the connection of an external sample loop, the length and internal diameter of which determines the volume to be injected. The other four ports are used for carrier and sample stream inputs, and outputs to the remainder of the flow injection manifold and to waste (see Figure 1.5). The device has two operational positions, one for charging the sample loop with fresh sample and the other for flushing the sample from the loop into the carrier or reagent stream. Switching between the two positions is performed by a rapid rotary action, which can be activated either pneumatically or electrically.

Figure 1.5 Schematic diagram of a six-port rotary injection valve.



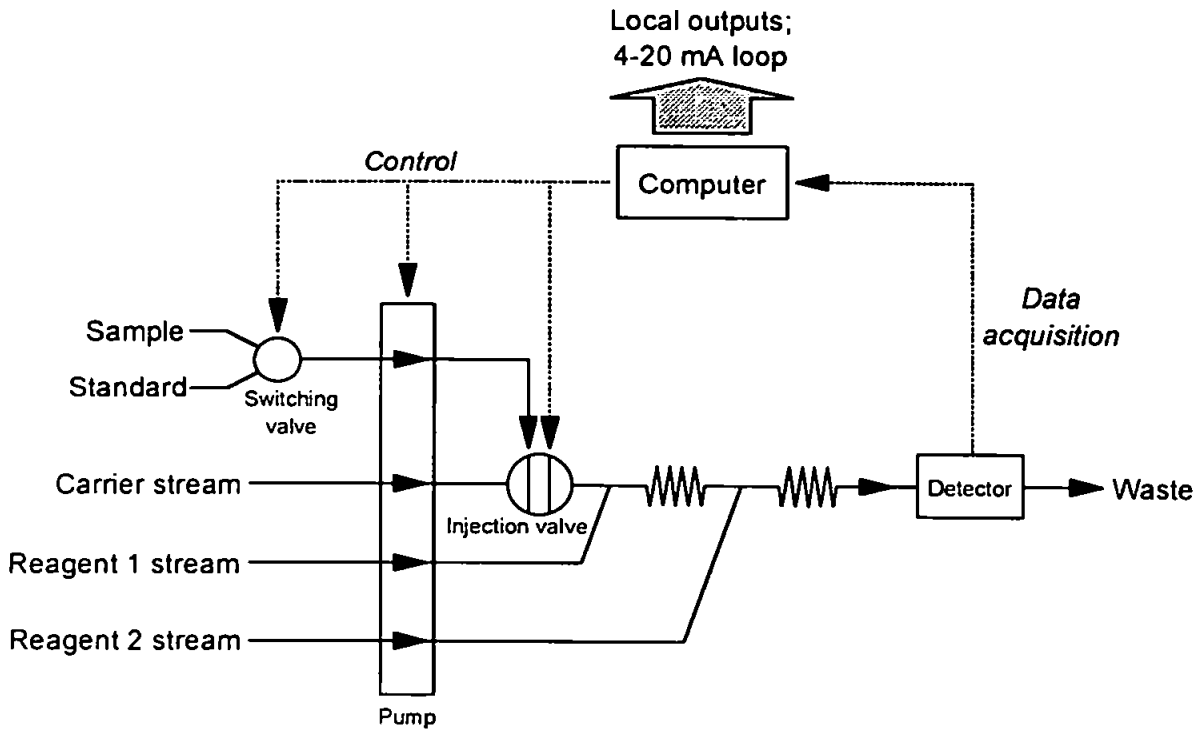
As shown in Table 1.4 previously, FI can be applied to a very wide range of both optical and electrochemical **detection systems**, which reflects the versatility of the technique. In addition to the normal analytical requirements of accuracy, precision and sensitivity, an essential aspect of an FI detector is its compatibility with flowing liquid media. Spectrophotometric methods have been widely applied in FI, using instruments

adapted with appropriate flow cells. UV/visible spectrophotometry is most commonly used, and this method of detection has become more popular with the development of solid-state photometers, incorporating light emitting diode (LED) light sources and photodiode detectors [16]. These systems are very low cost, robust and compact, and ideally suited for many *in situ* environmental or process monitoring applications. A detailed discussion of solid-state UV/visible detection systems for FI is given in Section 1.3. FI methods which combine on-line derivitisation procedures with fluorescence or chemiluminescence detection can provide a high degree of selectivity and sensitivity for certain analytes, and FI is also applicable to on-line analyte preconcentration and sample delivery for atomic spectroscopy. In the case of electrochemical methods, flow-through potentiometric electrodes for pH or selective ion determinations have been most frequently used in FI, with conductimetry, coulometry, voltammetry and amperometry less commonly applied.

FI methods readily lend themselves to automation, which is an essential requirement of a remote or on-line analytical system. The instrumental components of an FI manifold can be controlled by a simple computer board, which is also responsible for the tasks of data acquisition and processing. Control software can be stored on electrically programmable read-only memory (EPROM) chips, with analytical data communicated to an external computer or chart recorder *via* 4-20 mA loops or RS-232 serial connections. Automated self-calibration can be performed with the inclusion of switching valves, which allow regular injections of calibration standards in place of the sample. Figure 1.6 provides a schematic illustration of a typical automated FI system.

A wide range of manifold configurations are possible in FI, allowing the technique to automate almost any wet chemical reaction procedure. The most simple FI configuration is the single-channel manifold, as shown previously in Figure 1.1. This is used when an inherent property of the sample is being measured, or if only a single reagent stream is required. A two-channel manifold may be used for merging a reagent stream with the

Figure 1.6 Schematic diagram of an automated FI monitor. ——— represents liquid flow; ····· represents communication links.



carrier stream following sample injection, thereby providing a constant concentration of reagent for reaction throughout the dispersed sample zone, or for two-step reaction procedures. More sophisticated chemistries can be accommodated using multi-channel configurations, such as the four-channel manifold shown in Figure 1.7, in which three reagent streams merge in sequence with the carrier stream. Reagent consumption is generally low in FI systems, but can be reduced still further by use of reagent injection manifolds, as shown in Figure. 1.8. Reverse FI, as this configuration is often called, is suitable for applications in which the sample is in abundant supply (as is often the case for environmental and industrial analyses), and is particularly useful when expensive reagents are necessary. Reverse FI also minimises the quantity of reagent(s) discharged to waste, which is advantageous if environmentally-sensitive reagents are used. The manifold shown

Figure 1.7 Schematic diagram of a four-channel manifold.

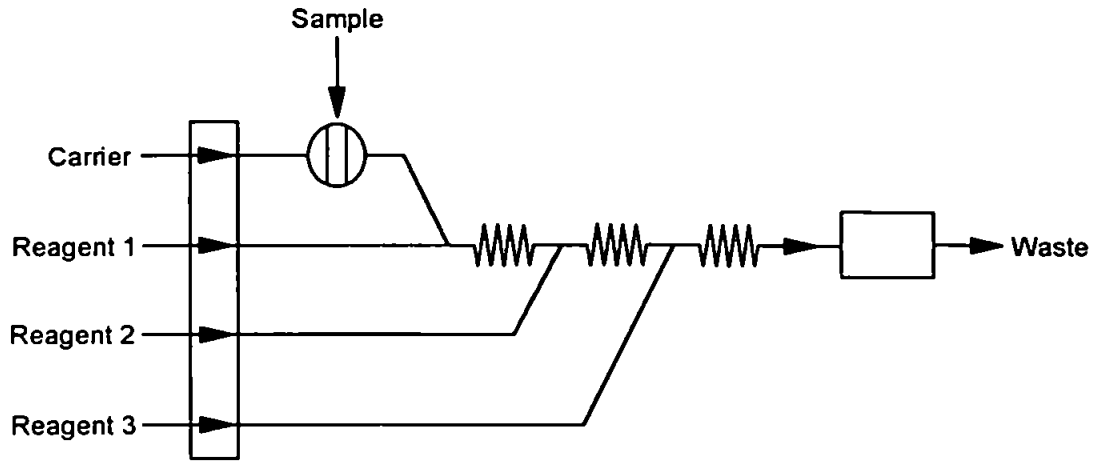
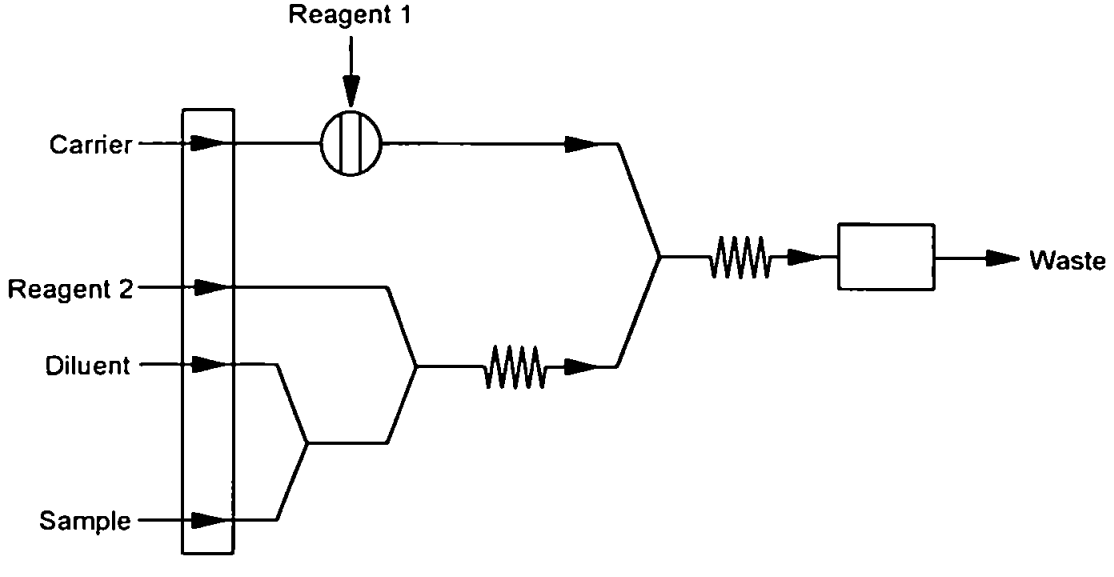


Figure 1.8 Four-channel manifold with reagent injection.



in Figure 1.8 is one used for the determination of sulfite in a high ionic strength process stream [17], and includes a diluent stream for dilution of the sample prior to measurement.

FI is not restricted to the analysis of single analytes in a given sample. Manifolds can be configured to perform simultaneous determinations of two or more analytes, and this can be achieved in several ways. One option is the combination of two reaction chemistries and manifolds into a single system, as shown in Figure 1.9. This method has been employed

for the determination of iron(II) and iron(III) in process liquors [18]. Here, the sample is injected into two parallel manifolds, undergoing a different reaction in each. Another approach is shown in Figure 1.10, and involves splitting the carrier stream after injection to undergo different treatments, then recombining for detection of two reaction products in sequence (*e.g.* for the determination of iron(II) and total iron [19]). The incorporation of multichannel detection systems, such as diode-array spectrophotometers or electrochemical sensor arrays, provides a further option for simultaneous determination of multiple analytes [20]. The multivariate data obtained by these instruments can be calibrated with respect to several analytes in a sample using chemometric routines such as principal components regression (PCR) and partial least squares regression (see Chapter 3 for a full discussion of multivariate calibration techniques).

Figure 1.9 Schematic diagram of a two-channel/three-channel parallel manifolds system for simultaneous determinations.

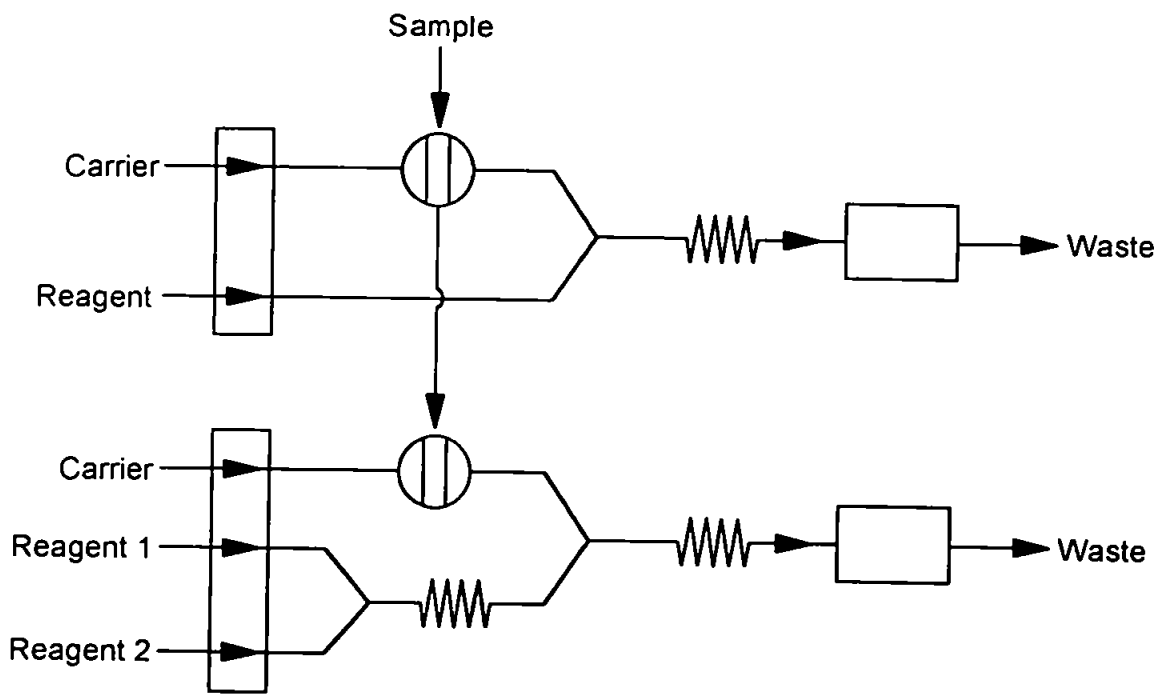
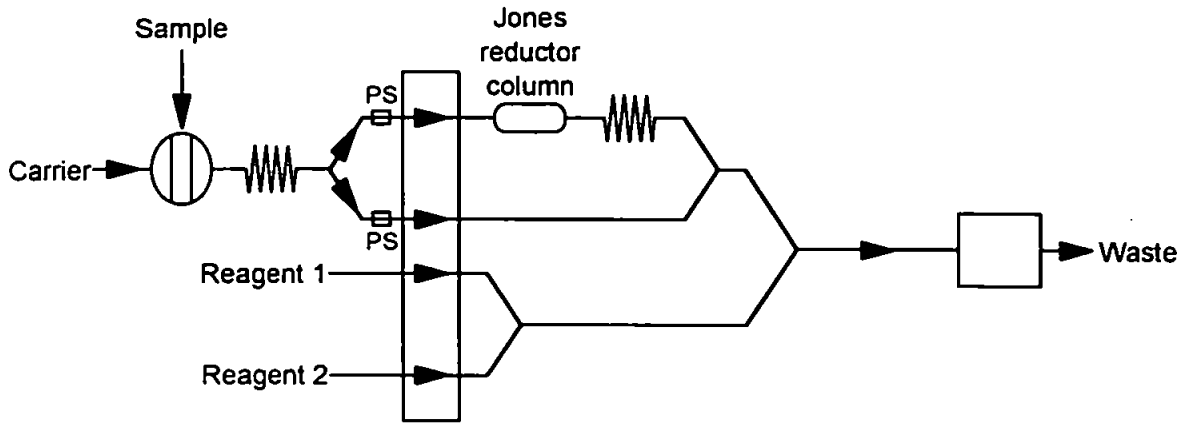


Figure 1.10 Schematic diagram of a multi-channel manifold for simultaneous determination, using post-injection sample splitting. PS represents pulse suppressor.



Other components which can easily be incorporated into FI systems include gas dialysis units, for the diffusion of a gaseous analyte from a carrier (donor) stream through a microporous membrane into a reagent (acceptor) stream (see Chapter 2), and solid phase reaction columns, in which the injected sample reacts with (or selected components are retained by) a column packed with solid material (*e.g.* the Jones reductor column shown in Figure 1.10).

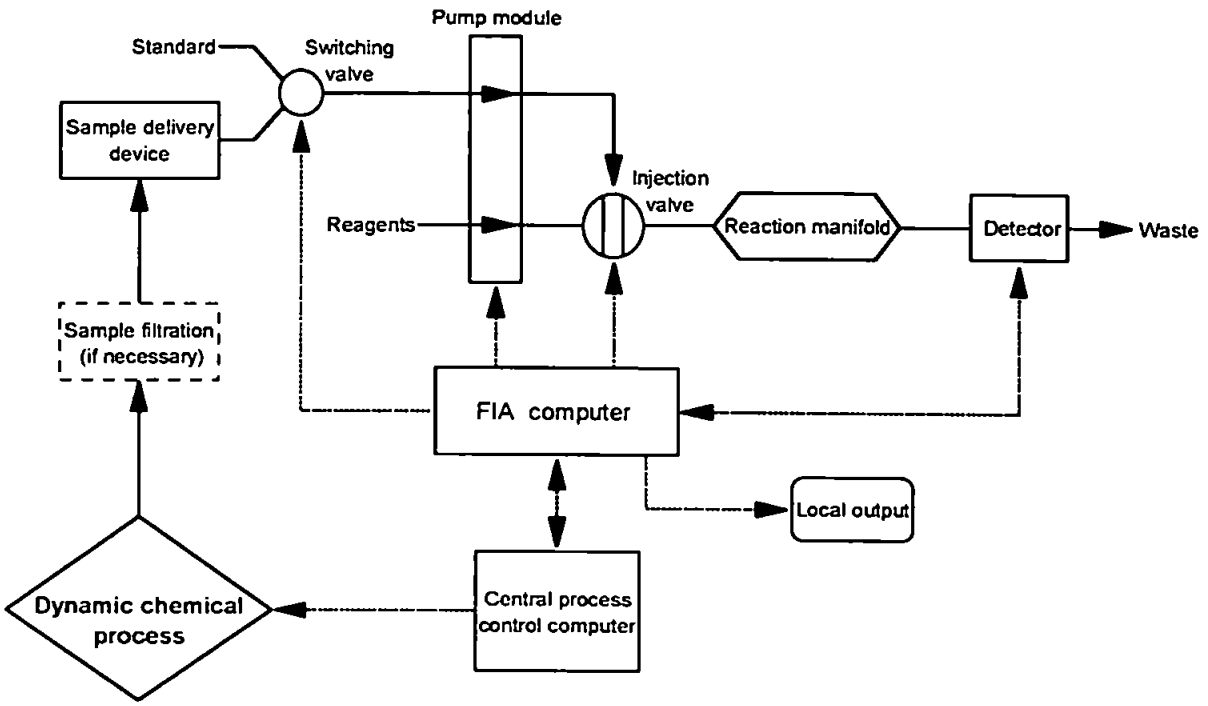
Advances have recently been made in the miniaturisation of FI components and manifolds. Examples are flow channels etched onto chrome-plated glass plates, using electrokinetic and electro-osmotic flow to mobilise reagent and sample streams [21], and micro-machined silicon structures incorporating piezoelectric membrane pumps and flow manifolds [22]. These systems have the advantages of extremely low reagent consumption (total flow channel volumes of approximately 5 µl) and compact size, and it is feasible that they will eventually be an important part of *in situ* environmental and process monitoring strategies.

1.2.3. Advantages of FI for on-line analysis

The requirements of an analytical system for on-line monitoring of process and effluent streams include an ability to interface with liquid-phase samples; rapid analysis and high sampling frequency to provide near-continuous information about the sample stream; robust construction to withstand harsh chemical matrices; a simple design that can be easily maintained; the ability to perform automated, unattended analyses and undertake regular self-calibration, and minimal capital and operating costs. FI can meet all these requirements, with sample response times of typically 10-120 s and sample throughputs in the range 30-120 h^{-1}. Reagent consumption is low (typically 30-180 ml h^{-1} for each stream during continuous operation), particularly in the case of reverse FI, and this helps to minimise operating costs and waste production. The simplicity of FI construction and its automation have been demonstrated in the previous section. The characteristic features of FI are obviously well suited to on-line analysis, and a number of publications have consequently discussed its potential for on-line, *in situ* monitoring of process streams [20, 23 , 24 , 25 , 26 , 27 , 28] and natural waters [29 -30].

A typical on-line FI process monitor is represented in Figure 1.11, which demonstrates how each component is controlled automatically by a simple single-board computer, which in turn communicates with a central process control computer. As mentioned in Section 1.1.2, this arrangement allows appropriate adjustments to be made to process variables (or to the effluent treatment system) in response to feedback from the on-line monitor. The monitor can be designed to incorporate sample pre-treatment procedures such as filtration, dilution and preconcentration, and to include a self-calibration protocol. The interface between the process/effluent stream and the monitor is an important aspect of on-line analysis, and will typically take the form of a series of coarse and/or membrane filters, draining into a constant head vessel from which samples are drawn by the monitor.

Figure 1.11 Schematic diagram of an automated process FI monitoring system. ——— represents liquid flow; ····· represents communication links.



1.2.4 Industrial applications for on-line FI monitoring

To date, the number of reported applications for on-line FI monitoring of industrial process and effluent streams remains relatively small considering the suitability of the technique. The reason for this may partly be due to industrial confidentiality, and also to the fact that the potential of FI for on-line monitoring is not yet fully realised. However, the number of publications appears to be steadily growing as the technique receives more widespread acceptance, and these are listed in Table 1.5. The primary applications are in the areas of biotechnology, industrial chemical processes and water quality monitoring. The distribution of publications with respect to the area of process application is illustrated by Figure 1.12. These applications demonstrate both the versatility of FI for monitoring a wide range of diverse analytes and its ability to withstand harsh sample matrices such as dye production liquors and fermentation broths. An FI instrument has been demonstrated to be capable of

long periods of continuous, unattended operation by a system which has been installed for several years at a remote site, performing *in situ* analyses of nitrate in river water [57].

The application of FI to on-line process monitoring will increase as FI technology develops (*e.g.* miniaturisation of FI components), commercial process systems become available, and quantitative chemometrics become more routinely applied to the interpretation of on-line data.

Table 1.5 Process FI applications classified by area and analyte.

Area	Analyte	Comments	Ref.
Chemical production	Sulfuric acid, ammonia and caustic solutions		31
	Sulfide in di-isopropanolamine solutions		24
	HCl in concentrated hydrochloric acid		32
	Azo dyes		33
	Sulfite in KCl brine	On-line process monitoring	17, 34
	Salicylic/acetylsalicylic acids in pharmaceutical preparations	Continuous monitoring of tablet dissolution tests	35
	Morphine		36
	Hydrogen cyanide in process gas streams	On-line monitoring of industrial process gas streams	37
	Ammonium sulfite	On-line process monitoring	38
Metal production	Iron(II) and iron(III)in mineral process liquors		18
	Soluble aluminium in steels		39
	Thiocyanate in metallurgical process solutions		40
	Trace gold in cyanide process solutions		41
Paper production	Calcium in paper machine back water		42
Fish farming	Ammonia	On-line monitoring of tanks containing fish farming plant sea water	43
	Ammonia and nitrite	On-line monitoring of sea and tap water tanks containing suspended fish feed	44

Table 1.5 (continued)

Area	Analyte	Comments	Ref.
Hydroponic cultivation	Nitrate	On-line monitoring of outflow water from a hydroponic water cress bed	45
Wastewater monitoring	Sulfates and phosphates		33
	Chloride and ammoniacal-N		46
	Phosphate, ammonia and nitrogen	On-line monitoring of a pilot-scale wastewater treatment process	47
	Total phosphorus		48
	Glucose	On-line monitoring of a laboratory-scale waste whey treatment process	49
Treated water monitoring	Fluoride	On-line monitoring of a simulated fluoridation process	50
	Aluminium	On-line monitoring of potable water	51
	Aluminium and iron	On-line monitoring of potable water	52
Power-plant/ cooling water monitoring	Ammonia, hydrazine, copper, iron, silicon and pH		53
	Phosphate and chlorine		54
Freshwater monitoring	Phosphate		55
	Nitrate	On-line monitoring of river water	27, 56 -57
	Nitrate	On-line monitoring of tap water	58
	Ammonia	On-line monitoring of river water	59
Biotechnology	Protein	On-line monitoring of micro-organism cultivation and disruption processes.	60
	Formate dehydrogenase and L-leucine dehydrogenase	On-line monitoring of micro-organism disintegration and diafiltration processes.	61
	L-Phenylalanine	On-line monitoring of micro-organism cultivation processes.	62
	Glucose, lactic acid and protein	On-line monitoring of lactic acid fermentation.	63
	Oxidase	On-line monitoring of enzyme purification LC eluent.	64
	Extracellular proteins	On-line monitoring of cellulase fermentation processes.	65
	Glucose	On-line monitoring of micro-organism cultivation processes.	66

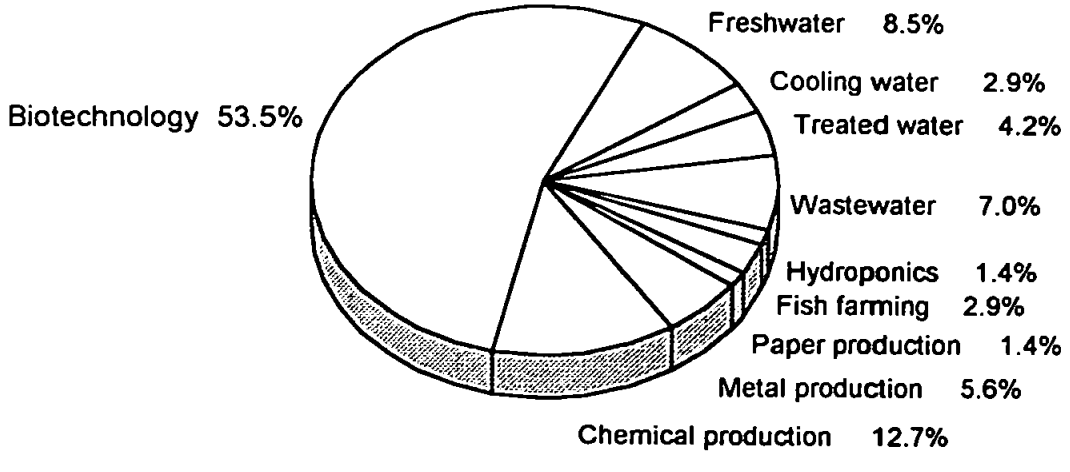
Table 1.5 (continued)

Area	Analyte	Comments	Ref.
Biotechnology (continued)	Ethanol	On-line monitoring of bioethanol production.	67
	Glucose, dimethylformamidase and protein	On-line monitoring of micro-organism cultivation processes + enzyme purification LC eluent.	68
	Alanine dehydrogenase, formate dehydrogenase and phenylalanine dehydrogenase	On-line monitoring of enzyme purification LC eluent.	69
	Cellulase		70
	Ammonium, glucose and proteins	On-line monitoring of fermentation processes.	71
	Ammonium and glucose	On-line monitoring of penicillin fermentation processes.	72
	Glucose	On-line monitoring of micro-organism cultivation processes.	73
	Acetate and phosphate	On-line monitoring of fermentation processes.	74
	Proteins	On-line monitoring of cell-culture and micro-organism fermentation processes.	75
	β-Galactosidase	On-line monitoring of micro-organism cultivation processes.	76
	Immunoglobulin		77
	Glucose and ethanol	On-line monitoring of yeast fermentation processes.	78
	Total acidity, reducing sugars, ethanol and pH	On-line monitoring of fermentation processes.	79
	Penicillin, ethanol, glucose, maltose and sucrose	On-line monitoring of micro-organism cultivation processes.	80
	Ammonium, glucose, maltose, amino acids, lactose, lactate and glutamine	On-line monitoring of alkaline protease and penicillin fermentation processes	81
	Antithrombin III, immunoglobulin and pullulanase	On-line monitoring of simulated and real (cell-culture and micro-organism) cultivation processes.	82
	Pullulan and glucose		83
	Serum albumin, immunoglobulin and peroxidase		84
	Amylase, xylanase, polygalacturonase and protease		85

Table 1.5 (continued)

Area	Analyte	Comments	Ref.
Biotechnology (continued)	Acetic acid	On-line monitoring of vinegar production.	86
	Glucose and lactate	On-line monitoring of cell-culture fermentation processes.	87
	Urea and glucose	On-line monitoring of micro-organism cultivation processes.	88
	Penicillin V	On-line monitoring of penicillin fermentation processes.	89
	Glucose	On-line monitoring of microbial gluconic acid production.	90
	Pullulanase and immunoglobulin	On-line monitoring of micro-organism and hybridoma cultivation processes.	91
	α-Amylase		92
	Ethanol	On-line monitoring of yeast fermentation processes.	93
	pH, urea, penicillin V and immunoglobulin		94
	Glucose, disaccharides and β-galactosidase	On-line monitoring of recombinant protein production	95
	Formate dehydrogenase and malate dehydrogenase	On-line monitoring of yeast fermentation processes.	96

Figure 1.12 Pie chart showing distribution of published process FI methods by area of application.



1.3 SOLID-STATE UV/VISIBLE SPECTROPHOTOMETRY

UV/visible spectrophotometry continues to be the most commonly applied detection method in FI analysis, owing to the diverse range of analytes which can be determined this way. Solid-state instruments, using either simple and low-cost LED/photodiode components or more sophisticated one- and two-dimensional photodiode arrays, offer the additional advantages of robustness and mechanical simplicity, and are therefore an attractive option for on-line industrial monitoring. Array detectors also enable the rapid acquisition of multiwavelength data, which can be used for simultaneous multicomponent analyses.

This section describes the theoretical principles of solid-state UV/visible detectors, and gives examples of their application to on-line determinations.

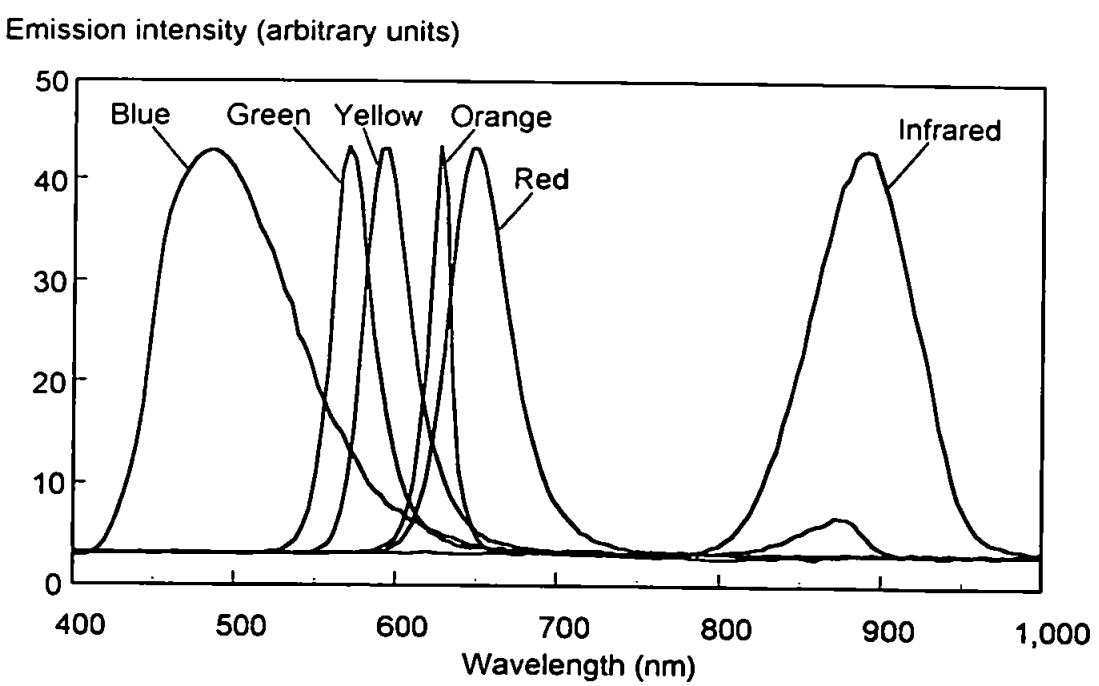
1.3.1 Light emitting diode photometry

The possibility of using light emitting diodes (LEDs) and photodiodes as the optical components of visible absorption photometers was first discussed by Flaschka *et al* in 1973 [97]. LED and photodiode components offer the advantages of minimal cost, high power efficiency and small size, and can therefore be used to construct compact, robust and portable photometric detectors. The application of LED photometers as FI detectors has been described by several publications [16, 98 , 99 , 100]

As can be seen in Figure 1.13, LEDs which span much of the visible through to the near-infrared region are now commercially available. LEDs are generally constructed from gallium arsenide (GaAs), gallium phosphide (GaP) or gallium arsenide phosphide (GaAsP) [99]. These crystalline material are semiconductors, which exhibit electrical conductivity less than that of a metal but greater than that of an electrical insulator [101]. A *p-n* junction is fabricated within the diode, which permits conduction in one direction only. This

is the junction between a negatively charged semiconductor in which an excess of unbonded electrons are present (the *n* region) and a positively charged semiconductor which possesses an excess of 'holes' (the *p* region). A hole is an area of positive charge, produced by the liberation of an electron from the crystal lattice. When the *p* and *n* regions are connected respectively to the positive and negative terminals of a d.c. source ('forward bias'), the excess electrons of the *n* region and the excess holes of the *p* region travel toward the junction, where they combine and neutralise each other. New electrons travel from the negative terminal of the source, and continue conduction towards the *p-n* junction, while electrons are drawn from the *p* region by the positive terminal, thereby creating new holes which migrate in the opposite direction. If the source terminals are connected in the reverse direction ('reverse bias'), then electrons and holes in the *n* and *p* regions are drawn away from the *p-n* junction, creating a non-conductive depletion layer. Conduction is thus permitted in one direction only.

Figure 1.13 Normalised emission spectra of some commercially available light emitting diodes (obtained from RS Components, Corby, Northants, UK).



In the case of LEDs, a major proportion of the hole-electron recombination energy is released as light, the wavelength of which is dependent upon the band gap (*i.e.* the gap between the valence and conduction electron energy bands) of the semiconductor. Table 1.6 lists semiconductor materials, peak emission wavelengths and output intensities for some common LEDs.

Table 1.6 Characteristics of some commercially available LEDs.

LED type	Construction materials	Peak emission wavelength (nm)[a]	Typical output intensity (mcd)[a]
Blue	SiN	470	13[b]
Green	GaP	563	200[c]
Yellow	GaAsP layer on GaP substrate	585	160[c]
Orange	GaAsP layer on GaP substrate	620	1500[c]
Red	GaAsP layer on GaP substrate	650	160[c]
Near-infrared	GaAlAs	880	[d]

[a] At 25 °C

[b] At 50 mA forward input current

[c] At 20 mA forward input current

[d] Radiant power = 16 mW sr^{-1} at 100 mA forward input current

Both photodiodes and phototransistors can be used in LED photometers to detect the radiant light of the LED [16, 99]. Phototransistors can provide greater sensitivity at lower visible wavelengths, although they exhibit a slower response than photodiodes. Photodiodes also offer the advantage of a much wider range of linear response to transmitted light. Photodiodes are generally constructed using silicon, which is partially 'doped' with a Group V element (*e.g.* As) to produce excess free electrons in the *n* region, and with a Group III element (*e.g.* Ga) to produce excess holes in the *p* region. When the photodiode is connected to a d.c. source in reverse bias, any UV or visible photons impinging on the photodiode will possess sufficient energy to liberate additional electrons

(and therefore create additional holes) within the depletion layer at the *p-n* junction. This results in an increase in conductivity which is directly proportional to the radiant light reaching the photodiode [101].

The absorbance path length for an LED photometer in FI is generally the diameter of the manifold tubing, with the LED and photodiode positioned directly opposite each other, either side of the flow stream within a Perspex® or aluminium block. Single beam designs are frequently used, although the inclusion of a second LED/photodiode pair in the flow cell produces a double beam instrument, which allows drift compensation by measuring absorbance in both the sample stream and a reference stream. A further option is the incorporation of LEDs of more than one emission wavelength, in order to produce a multi-wavelength photometer [102 -103]. In this case, light is transmitted from the LEDs to the flow cell *via* fibre optic cables.

A number of on-line monitoring applications involving FI in combination with LED photometers have been reported. On-line FI determination of nitrate in the outflow water of a hydroponic water cress bed was achieved using an automated, portable photometer fitted with a green LED [45]. A similar system was used for on-line analyses of aluminium (yellow LED) [51] and iron (red LED) [52] in drinking water. Protein was determined in a lactic acid fermentation broth using an on-line FI system with a green LED photometer. Glucose and lactic acid were determined in the same broth by an FI chemiluminescence method, in which a photodiode was used as the detector [63].

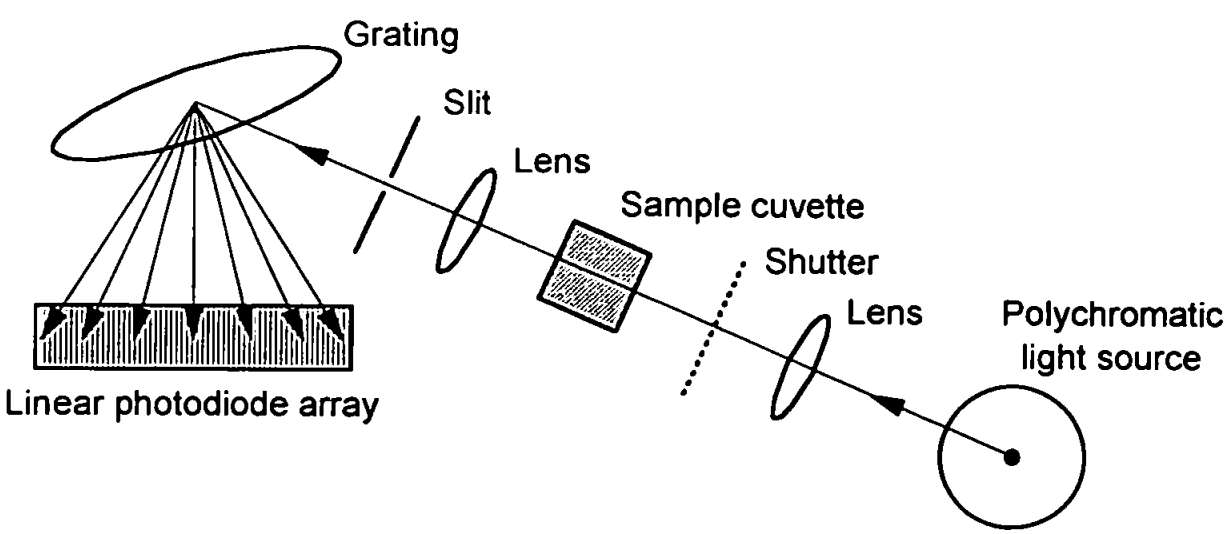
1.3.2 Photodiode array spectrophotometry

Photodiode array (PDA) spectrophotometers have been available commercially since the late 1970s, and have been widely used as UV/visible detectors in liquid chromatography, and increasingly for simultaneous multicomponent determinations in FI [104 -105]. A PDA typically comprises up to 1024 photodiode elements fabricated on a single silicon chip

in a one-dimensional linear series. A typical photodiode element is 15-50 μm wide, and the dimensions of a PDA chip are 2.5 × 10-60 mm.

In a PDA spectrophotometer [101] (see Figure 1.14), polychromatic light from a tungsten or deuterium lamp is focused upon a sample cuvette, and the fraction of the incident light which is not absorbed by the sample passes into a polychromator with a fixed, holographic grating. The grating produces spectral dispersion of the light in such a way that light of a different wavelength impinges on each of the photodiodes. Each diode is connected to a dedicated capacitor, which is charged to -5 V by the momentary closing of a solid-state switch (controlled sequentially by a shift-register). If photons impinge on a given diode, this causes the capacitor to partially discharge, and the lost charge is restored during the following switching cycle. The degree of current necessary to recharge the capacitor is directly proportional to the intensity of light reaching the surface of the diode. This whole process occurs in a matter of milliseconds, and PDA spectrophotometers can therefore record an entire absorbance spectrum in as little as 0.1 s.

Figure 1.14 Schematic diagram of a photodiode array spectrophotometer



PDA spectrophotometers offer a number of advantages over conventional UV/visible spectrophotometers [106]. Full absorbance spectra can be acquired in a minimum of 0.1 s with no significant loss in sensitivity. This very fast data acquisition time is ideally suited to the analysis of dynamic systems, such as those of LC and FI determinations, process analysis and kinetic measurements. Simultaneous measurement of multiple wavelengths enables the spectrophotometer to perform internal referencing, which compensates for fluctuations in lamp output, and also wavelength averaging across adjacent photodiodes, which reduces signal noise. Reduction of noise has the additional advantage of extending the measurement range of the instrument at low absorption levels. Wavelength resettability is greater for PDA spectrophotometers than for conventional instruments, since no moving parts are required to change or scan wavelengths. This eliminates the potential problems of mechanical error and wavelength drift over time. The speed of data acquisition also allows a number of spectra to be measured in the space of one second or less, and from these, a calculation of standard deviation for each data point can be provided, thus providing an indication of data quality at each wavelength. Perhaps the most significant advantage from the perspective of on-line industrial monitoring is the mechanical simplicity of PDA instruments, which provides a high degree of robustness and reliability.

Miniature PDA systems are now commercially available, *e.g.* the S1000 series of 1024-element PDA spectrophotometers (Ocean Optics, Dunedin, FA, USA), which have dimensions no greater than 15 × 14 × 6 cm and are fully compatible with fibre optic cables. These instruments represent a very promising option for on-line multiwavelength monitoring, since they are compact, robust, and relatively low cost (<£4000 for a complete system).

An example of an on-line FI-PDA determination is that reported for the single-component analyses of ammonium, glucose and proteins in a fed-batch fermentation process [71]. The rate of fermentation was monitored with respect to the three analytes at either

one or two wavelengths. An FI-PDA method was also used for the simultaneous determination of phosphate and chlorine in simulated industrial cooling water samples [54]. Visible absorbance spectra were acquired and analysed using multivariate calibration techniques to quantify the individual components. Although not performed on-line, this method demonstrated the feasibility of applying an FI-PDA-multivariate data analysis combination to a process monitoring situation.

1.3.3 Charge transfer devices

Although PDA spectrophotometers offer many advantages over conventional UV/visible instruments, they are limited in terms of sensitivity to very low intensity light inputs and in terms of wavelength resolution (typically 1-2 nm). Owing to the electronic design of the PDA, it can only be cooled to temperatures achievable using thermoelectric cooling (-50 °C), at which a relatively high dark current (500 counts/diode/second) is still produced [107]. It is this noise which limits sensitivity at very low light intensities. PDA resolution is dependent on the number and size of photodiodes per spectral range, which in turn is related to the performance of the polychromator and the required total spectral coverage [106].

PDA sensitivity has been partly improved by the development of intensified PDAs, in which a multichannel plate intensifier is positioned in front of the PDA chip to perform photomultiplication. However, this approach is limited in terms of spectral range and resolution, and it is only with the development of charge transfer devices (CTDs) that very high sensitivity and resolution have become available in multichannel solid-state spectroscopy. CTDs are two-dimensional arrays of photosensitive metal oxide semiconductor capacitors arranged within a single, solid-state integrated circuit, and are capable of collecting and quantifying photo-generated electrical charge [108]. Since CTDs are two-dimensional arrays, they can obtain information on variations in light intensity with

wavelength and with slit height. In addition, multiple sources may be determined simultaneously using different regions of the array (*e.g.* inputs from several vertically aligned fibre optic cables).

Two forms of CTD exist, namely the charge injection device (CID) and the charge coupled device (CCD). In the case of the CID, photons striking the surface of each detector element (or pixel) generate a proportional degree of electrical charge, which is shifted between two electrodes within the pixel (*i.e.* intra-cell transfer). Voltage fluctuations resulting from this transfer are detected and are proportional to the intensity of light striking the pixel. In a CCD detector, photo-generated charge accumulating within each pixel is transferred to a serial register, and then to a charge-sensing output amplifier (*i.e.* inter-cell transfer) [107-108]. CIDs have been used for wide dynamic range imaging applications (*e.g.* in atomic spectroscopy), whereas CCDs are preferred for low intensity spectroscopic (*e.g.* Raman and fluorescence) and spatial (*e.g.* astronomical) imaging owing to their superior signal-to-noise ratio.

The sensitivity of a CCD detector is enhanced by cooling to -130 °C with liquid nitrogen, at which temperature the dark current is <1 electron/pixel/hour. A pixel is typically about 22 × 22 μm in size, while CCD arrays are typically arranged as 578 × 385, 512 × 512 or 1152 × 298 pixels. CCD detectors are sensitive to visible and near-IR radiation in the range 400-1100 nm, and if the CCD is coated with a UV-sensitive fluorescent dye, this range is extended to 200 nm. CCD spectrophotometers are also capable of a very high degree of spectral resolution throughout the wavelength range, *e.g.* 0.1 nm at 546 nm when using a 1200 groove mm^{-1} grating [109].

CCD spectrophotometers share the advantages of mechanical simplicity, reliability and fast, full spectrum data acquisition that PDA instruments provide. In addition, they are capable of much higher sensitivity and resolution, and can simultaneously measure multiple spectra. In technical terms, CCD detectors are therefore also ideally suited to on-line

monitoring, although their high capital costs may restrict the number of practical applications. No examples of on-line FI-CCD spectrophotometric determinations have been reported as yet, although CCD detectors have been applied to on-line monitoring of C_8 aromatics separation processes using Raman spectroscopy [110], and to on-line image analysis for determining granule size distribution in pharmaceutical granulation processes [111].

1.4 RESEARCH OBJECTIVES

The general aim of this research was to investigate the potential of applying FI in combination with UV/visible detection and multivariate calibration techniques to the on-line monitoring of single and multiple analytes in industrial waste waters.

The specific objectives were as follows:

1. To develop an automated FI method for the determination of a single analyte in effluents, and to test this system on-line in a real process environment.
2. To investigate the relative performances of different multivariate calibration techniques for quantifying individual components in multicomponent mixtures, analysed by PDA spectrophotometry.
3. To develop a method combining FI with PDA detection and multivariate calibration for the simultaneous multicomponent analysis of effluents.
4. To investigate the potential of Kalman filtering techniques for multivariate calibration and drift correction of multicomponent data.

Chapter 2

Flow Injection Determination of Ammonia in Industrial Liquid Effluents

2.1 INTRODUCTION

Ammonia is often found at mg l^{-1} or higher levels in wastewaters discharged from a variety of industrial and agricultural activities. It can be formed by the biodegradation of organic nitrogen compounds (*e.g.* in sewage or agricultural wastes), and is also commonly used as a raw material by the chemical industry (*e.g.* in the production of fertilisers and biocides). However, owing to its adverse effect on oxygen balance in the aquatic environment, ammonia is included in List II of the Dangerous Substances in Water Directive (76/464/EEC), as indicated in Table 1.1 of the previous chapter. Discharges of wastewaters containing ammonia are therefore subject to the stipulations of discharge consent agreements within the UK, and regular sampling and analysis are necessary in order to ensure compliance with the terms of the consent. As discussed in the previous chapter, these requirements are most satisfactorily achieved through the use of automated, on-line effluent analysers.

The potential merits of flow injection (FI) as a method of on-line analysis were discussed in Section 1.2.3. FI is also well suited to the determination of ammonia, with a number of methods reported for the analysis of ammonia in a diverse range of liquid samples (as summarised in Table 2.1). Spectrophotometric methods are frequently applied, including those utilising the Berthelot reaction, in which ammonia reacts with phenol and hypochlorite to produce indophenol blue, and gas diffusion methods in which gaseous ammonia diffuses from the sample stream across a microporous membrane into typically an acid-base indicator stream. The latter method is particularly well suited to complex sample matrices such as effluents since the membrane provides a physical barrier, excluding potential interferences (*e.g.* suspended solids and non-volatile ionic species) from the measured stream.

This chapter describes the development of an automated gas diffusion-FI monitor for the on-line determination of ammonia in industrial liquid effluent streams.

Table 2.1 Summary of published flow injection methods for the analysis of ammonia in liquid samples.

Analytical method	Comments	Sample type	Reference
Visible spectrophotometry	Gas diffusion (bromothymol blue indicator)	Canal water	112
	Gas diffusion (bromothymol blue indicator)	River water	59
	Gas diffusion (bromothymol blue indicator)	Seawater, haemolymph	113
	Gas diffusion (phenol red indicator)	Whole blood, plasma	114
	Gas diffusion (phenol red indicator)	Seawater	115
	Gas diffusion (Tecator NH_3 mixed indicator)	Industrial effluent water	46
	Gas diffusion (bromocresol purple indicator)	Aq. standard solutions	116
	Berthelot reaction	Aq. standard solutions	116
	Berthelot reaction	Aq. standard solutions	117
	Berthelot reaction	Fish tank water	44
	Berthelot reaction	Plant Kjeldahl distillates, soil extracts	118
	Modified Berthelot reaction	Fish farm sea water	43
	Gas diffusion (Nessler's reagent)	Aq. standard solutions	119
	Nessler's reagent	Drinking and river water	120
	Nessler's reagent	Irrigation waters	121
	Nessler's reagent	Natural waters	122
	Enzymatic determination (GIDH and NADPH)[a]	Aq. extracts of cheese and ham	123

[a] Glutamate dehydrogenase and nicotinamide adenine dinucleotide phosphate.

Table 2.1 (continued)

Analytical method	Comments	Sample type	Reference
Fluorimetry	OPA[b] + mercaptoethanol	Aq. standards	124
	OPA + mercaptoethanol	Synthetic aq. ammonia/ hydrazine mixtures	125
	OPA + mercaptoethanol	Natural waters	126
	OPA + sulfite	Tap, rain and lake water	127
	Gas diffusion; NH_3-ISE[c]	Aq. standard solutions	128
	Gas diffusion; pH-ISFET[d]	River water	129
Conductimetry	Gas diffusion	Aq. standard solutions	131
	Gas diffusion	Kjeldahl digests (leaves, fertilisers, animal feeds)	130
	Gas diffusion	River and lake waters, soil extracts	131
	Gas diffusion	Kjeldahl digests (vegetable tissue)	132
Chemiluminescence	Gas diffusion (luminol)	River water	133
	Hypobromite reaction	Rain and fog water	134

[b] *o*-Phthalaldehyde;
[c] Ion-selective electrode;
[d] Ion-selective field-effect transistor.

2.2 THE PRINCIPLES OF GAS DIFFUSION IN FI METHODS

Gas diffusion represents an important technique for separation and preconcentration of volatile analytes in flow injection analysis. It can be used to remove potential matrix interferences and to enhance both selectivity and sensitivity of analyses [135]. Gas diffusion involves the reproducible transport of gaseous analytes from a sample (or 'donor') stream through a hydrophobic, microporous membrane into a detector (or 'acceptor') stream. This diffusive transport produces a change in the physico-chemical nature of the acceptor stream, which is proportional to the concentration of the gaseous analyte. The

hydrophobic nature of the membrane permits only the exchange of gaseous molecules, and prevents the transfer of liquids, dissolved ionic species and particulates. Volatile analytes determined by gas diffusion-flow injection (GD-FI) methods have included ammonia (see Table 2.1), methylamines [136], carbon dioxide [137] and sulfite [138].

A typical gas diffusion cell is fabricated from Perspex® or poly(tetrafluoroethylene), and is comprised of two blocks with identical (but mirror image) flow channels on the surface, which form the donor and acceptor halves of the cell. The two blocks are fastened together so that the flow channels face each other, with the microporous membrane placed between the channels to form a barrier between the donor and acceptor streams (see Figure 2.1). The direction of flow is usually the same (*i.e.* concurrent) for both the acceptor and donor streams, although countercurrent flow (as shown in Figure 2.1) can be employed if a reduction in sensitivity is required. The efficiency of diffusion is also influenced by the residence time of the sample in the diffusion cell (which is dependent on the rate of flow and the length of the diffusion channel), the surface area of the sample at the membrane in relation to the volume of the diffusion channel, and the porosity of the membrane [119]. The rate of diffusion is optimal when the pressure on each side of the membrane is equal (*i.e.* equal flow rates and volumes in both the donor and acceptor streams) [139].

The pH of the sample stream is a fundamental factor in gas diffusion separations, since pH conditions determine whether the solution equilibrium favours the gaseous or ionic species of a particular analyte. Maximum diffusion efficiency is achieved when the sample pH is such that the dissolved analyte is fully converted to its gaseous form. In the case of ammonia in aqueous solution at 25°C, the ionic NH_4^+ species accounts for virtually all dissolved ammonia at pH 7.0, but is almost completely converted to the gaseous NH_3 species at pH 11.0 (see Figure 2.2) [140].

Figure 2.1 Diagram of typical gas diffusion cell for FI analysis.

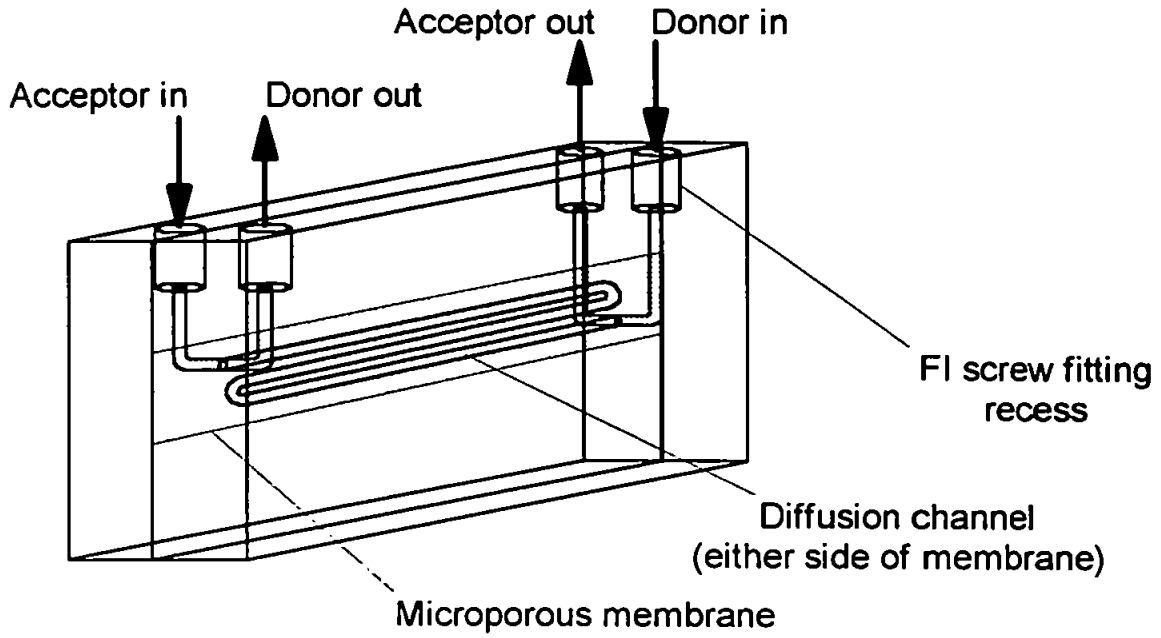
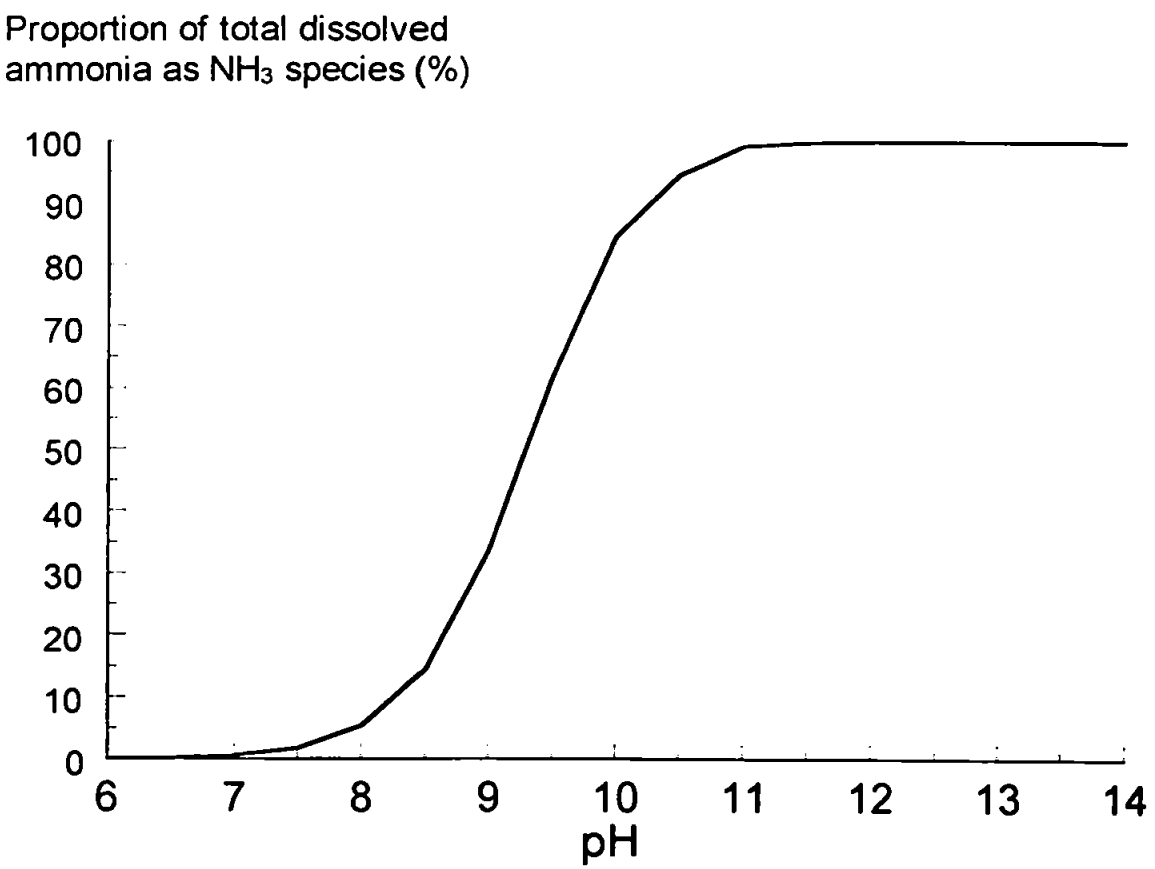


Figure 2.2 Effect of pH on dissolved ammonia speciation at 25°C.

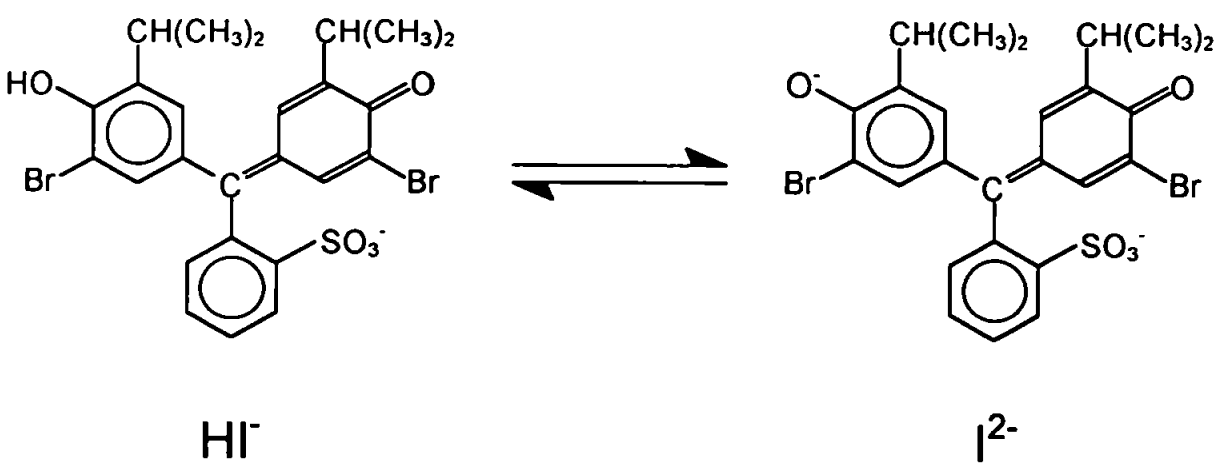


Acid-base indicator solutions are frequently used as the acceptor streams in GD-FI analyses. In the case of ammonia determinations, the acidic form of the indicator (HI^-) is used as the initial acceptor reagent. As ammonia diffuses across the membrane, a proportion of the indicator is converted to its basic form (I^{2-}), as shown in equation 2.1:

$$NH_3 + HI^- \rightarrow NH_4^+ + I^{2-} \qquad (2.1)$$

The formation of I^{2-} in the acceptor stream is then measured photometrically, with the height of the I^{2-} absorbance peak proportional to the concentration of ammonia originally present in the sample stream. The relationship between ammonia concentration and the absorbance change produced in the indicator stream is linear over a certain range, depending on the concentration and the initial pH of the indicator. For example, the change in absorbance of a 10^{-4} mol l^{-1} bromothymol blue solution (initially adjusted to pH 6.5) is linear for total ammonia concentrations of up to 3×10^{-5} mol l^{-1} in the indicator solution [112]. Figure 2.3 illustrates the structural formulae for the acidic and basic forms of bromothymol blue (3′,3″-dibromothymolsulfonephthalein) indicator.

Figure 2.3 Structural formulae of the acidic (HI^-) and basic (I^{2-}) forms of bromothymol blue indicator.



HI^- I^{2-}

2.3 EXPERIMENTAL

2.3.1 Reagents

All solutions were prepared using Milli-Q water (Millipore, Milford, MA, USA) and all reagents were of AnalaR grade (Merck, Darmstadt, Germany) unless otherwise indicated. A stock ammonia solution (1000 mg l^{-1} NH_3-N) was prepared by dissolving 3.819 g of ammonium chloride (previously dried at 105 °C) in 1 l of water. Ammonia calibration standards were prepared by serial dilution of the stock solution. A 1 mol l^{-1} sodium hydroxide stock solution was prepared by dissolving 40 g of sodium hydroxide pellets in 1 l of water, with serial dilution used to produce 0.1 and 0.01 mol l^{-1} working solutions. A stock bromothymol blue solution was prepared by dissolving solid bromothymol blue (0.4 g; Merck indicator grade) in 64 ml of 0.01 mol l^{-1} sodium hydroxide and diluting to 1 l with water. Stock solutions (0.4 g l^{-1}) of bromocresol purple and phenol red (Merck indicator grade) were prepared similarly, but using 74 ml and 113 ml respectively of 0.01 mol l^{-1} sodium hydroxide. Working indicator solutions in the range 0.05-0.35 g l^{-1} were prepared by serial dilution of the stock solutions, and adjusted to varying pH levels using 0.1 mol l^{-1} sodium hydroxide or hydrochloric acid solutions.

Stock solutions (1000 mg l^{-1} amine-N) were prepared for methyl and ethyl primary, secondary and tertiary amines by diluting or dissolving 40 % w/v methylamine (6.15 ml; Aldrich, Gillingham, Dorset, UK), dimethylammonium chloride (5.821 g; Merck GPR grade), 45 % w/v trimethylamine (10.05 ml; Aldrich), 70 % w/v ethylamine (5.75 ml; Sigma, Poole, Dorset, UK), diethylamine (7.45 ml; Merck GPR grade) and triethylamine (9.90 ml; Sigma) respectively in 1 l of water. Working solutions (2 mg l^{-1} amine-N) were produced by serial dilution of the stock solutions.

A 0.05 mol l^{-1} EDTA solution was prepared by dissolving 18.6 g of ethylenediaminetetra-acetic acid disodium salt in 1 l of water. Sodium phenate solution was

prepared by dissolving 62.5 g of phenol in 18.5 ml of acetone and making up to 100 ml with 96 % v/v ethanol. This solution was then mixed with 100 ml of 270 g l^{-1} sodium hydroxide solution and diluted to 500 ml with water. Stock sodium hypochlorite solution was prepared by diluting 25 ml of sodium hypochlorite solution (GPR grade, Merck) to 1 l with water. This solution was standardised for available chlorine by titration against 0.05 mol l^{-1} iodine (AnalaR Volumetric Solution, Merck), using a mixture of 50 ml of stock sodium hypochlorite solution, 50 ml of 0.05 mol l^{-1} sodium arsenite (AnalaR Volumetric Solution, Merck), 5 g of sodium hydrogen carbonate, and iodine indicator (BDH grade, Merck). A blank determination was also conducted using 50 ml of water in place of the stock solution. A working sodium hypochlorite solution was prepared freshly as required by diluting x ml (where $x \leq 38$) of stock solution to 250 ml with water. The volume of x was determined according to equation 2.2:

$$x = \frac{2250}{z \times 2.85} \tag{2.2}$$

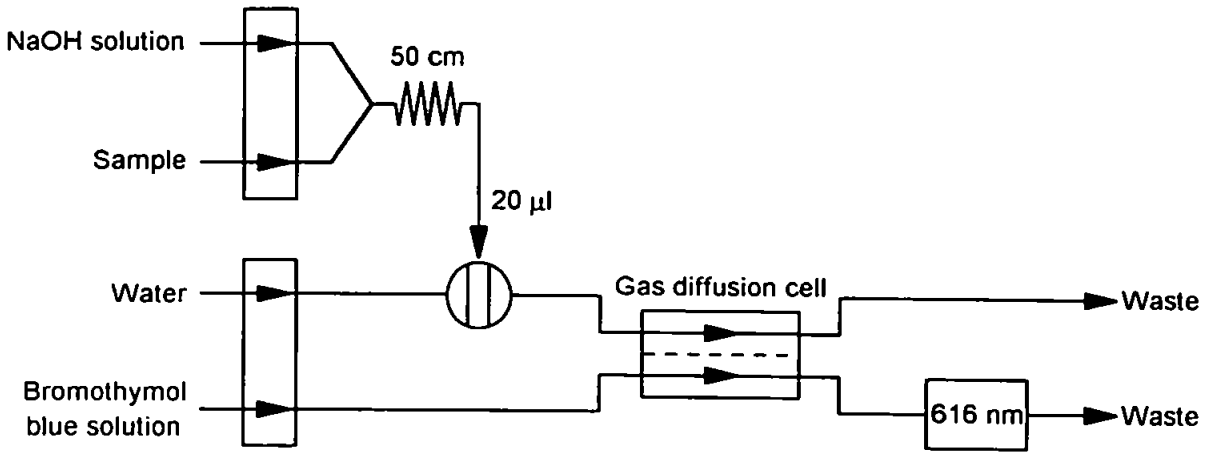
where z = the difference between the stock solution and blank titre values for the stock solution standardisation.

2.3.2 Instrumentation

The flow injection manifold used for initial method development is shown in Figure 2.4. Poly(tetrafluoroethylene) (PTFE) tubing of 0.8 mm i.d. (Anachem, Luton, Beds., UK) was used in the construction of the manifold. A peristaltic pump (Ismatec Mini-S 820, Ismatec, Carshalton, Surrey, UK) with poly(vinyl chloride) (PVC) pump tubing (Elkay, Basingstoke, Hants., UK) was used to propel the sample and sodium hydroxide streams through a mixing coil to a pneumatic six-port rotary valve unit (PS Analytical, Sevenoaks, Kent, UK). A second peristaltic pump was used to propel the water carrier (*i.e.* donor) and the indicator (*i.e.* acceptor) streams through an in-house Perspex® gas diffusion cell (see Figure 2.5) with

a diffusion path of 240 mm × 1.5 mm (volume = 72 µl on either side of the membrane). General purpose PTFE tape (width = 22 mm, thickness = 8-9 µm; RS Components, Corby, Northants, UK) was used for the gas diffusion membrane.

Figure 2.4 FI manifold for the determination of total ammonia in liquid effluents.



The indicator stream was passed to a Hewlett-Packard (Avondale, PA, USA) 8451A photodiode array spectrophotometer fitted with an 18 µl flow cell of 1 cm pathlength (Hellma, Westcliff-on Sea, Surrey, UK). Carrier and indicator stream flow rate optimisations were performed using two Minipuls 2 (Gilson, Villiers-le-Bel, France) variable-speed peristaltic pumps.

2.3.3 Portable FI monitor

The manual FI manifold was incorporated into a portable, automated monitor, designed and built by Blundell *et al.* for the determination of nutrients in natural waters [58, 141]. The basic layout of the monitor is shown schematically in Figure 2.6, and comprised an FI manifold and a microcomputer system housed within two polycarbonate boxes. Reagents

Figure 2.5 Diagram of the gas diffusion cell: (a) side view; (b) plan view; (c) end view.

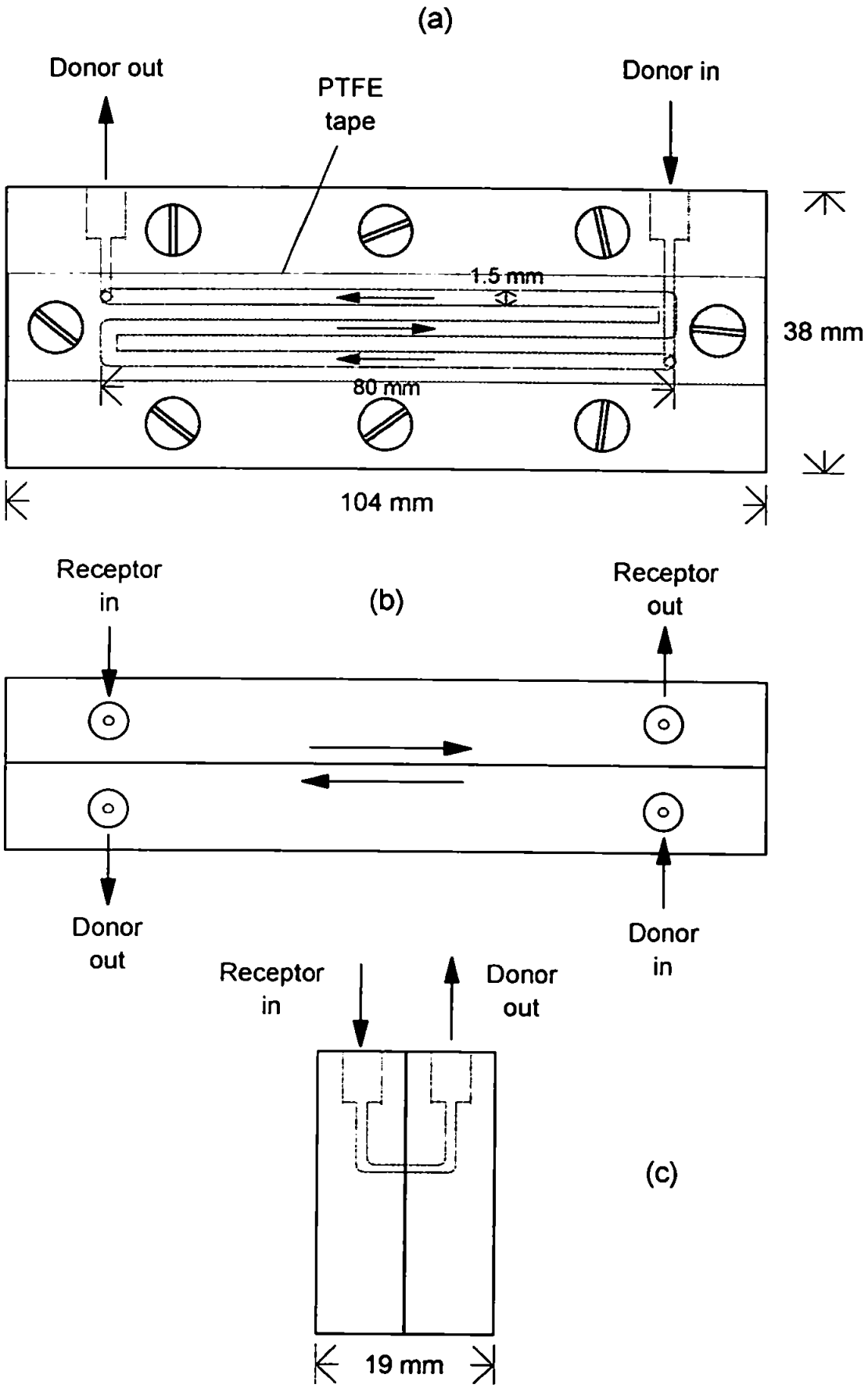
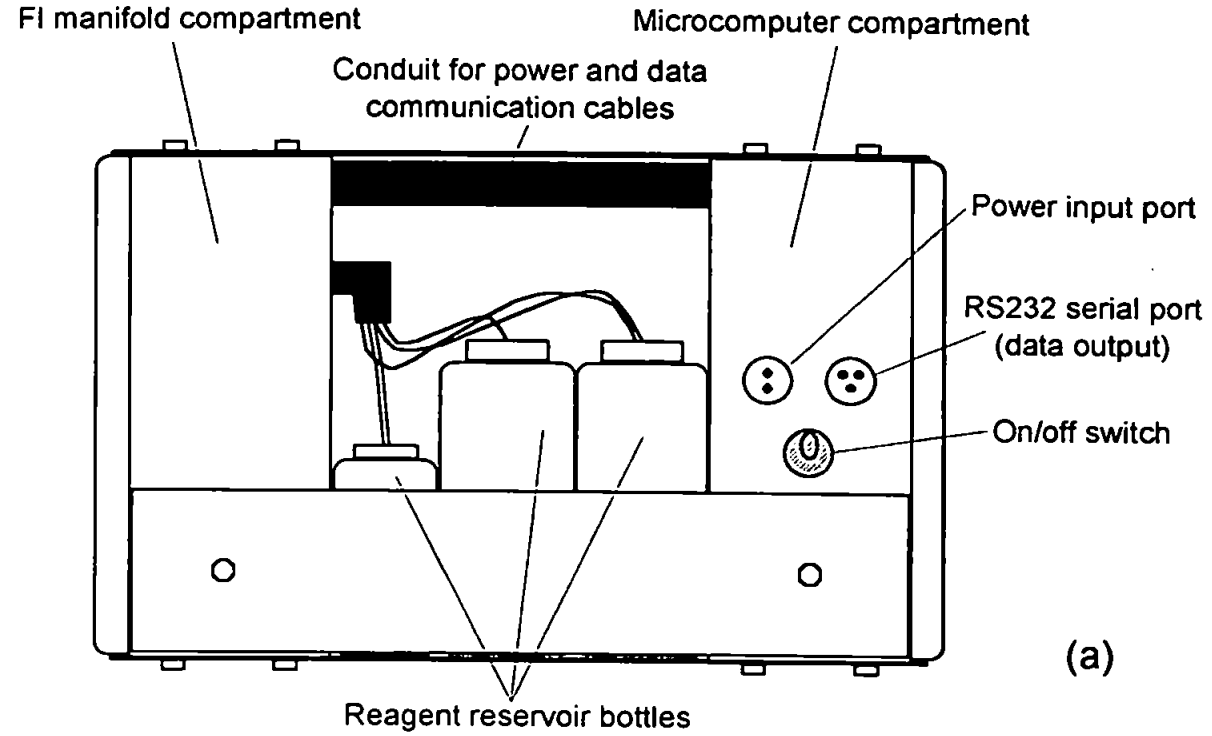
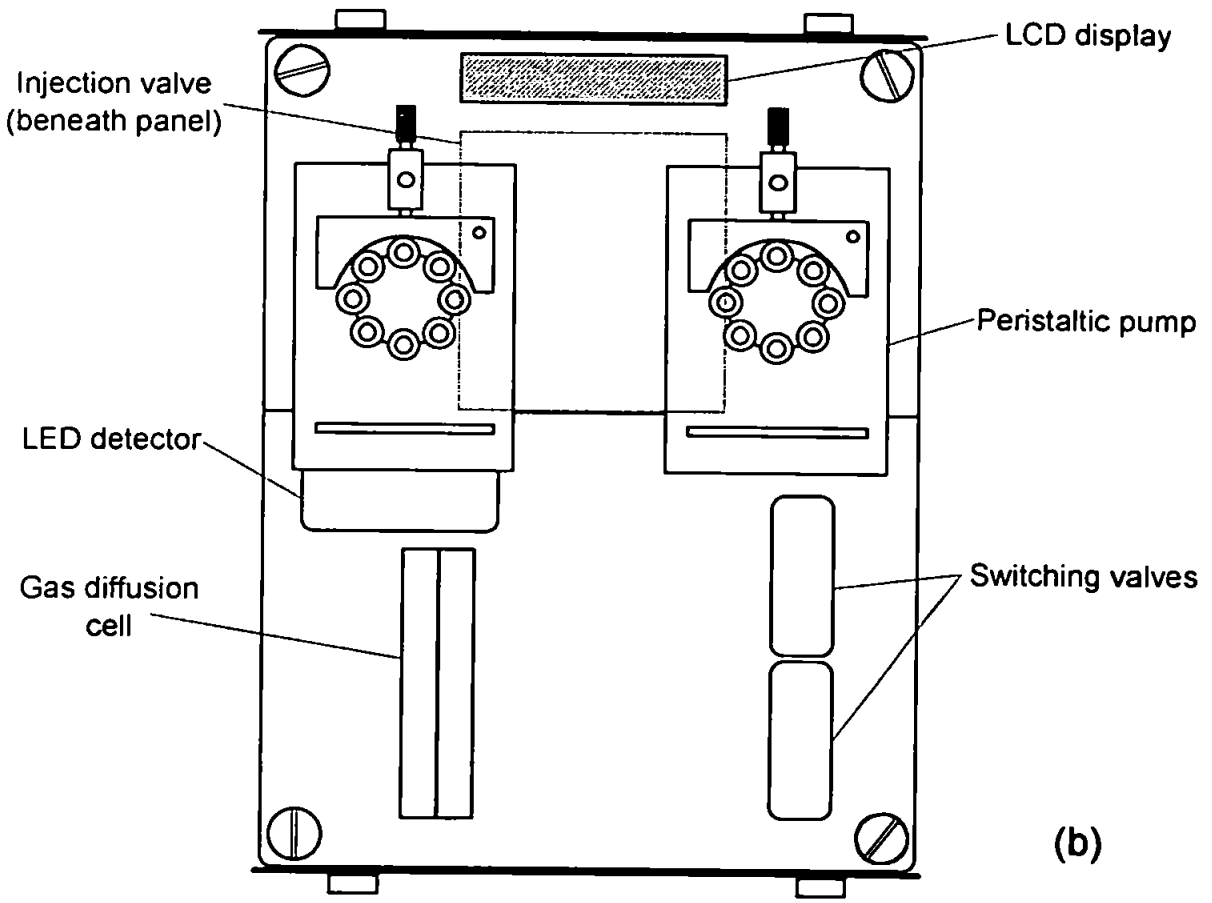


Figure 2.6 Basic layout of the portable FI monitor: (a) side view; (b) end view of the FI manifold compartment (tubing omitted for clarity).





were stored in a central area between the manifold and computer compartments. The weight of the monitor with full reagent bottles was 20 kg.

The microcomputer system was designed in-house, and based around the 8-bit Intel 80C32 microcontroller. The computer performed acquisition, processing and storage of data, and enabled automated control of the FI manifold components. The system was designed with 32 Kbytes of battery-backed random access memory (RAM) and 24 Kbytes of storage within electronically programmable read-only memory (EPROM) space. Control software was written in BASIC (see Appendix 1) and permanently stored in EPROM, executing automatically on power-up. Monitor operating variables were stored in battery-backed RAM. Processed analytical data were logged in RAM, then downloaded *via* a three-line RS232 connection to a PC when required.

The FI manifold was comprised of two peristaltic pumps (Ismatec Mini-S-E/8/12VDC/60:1), a solenoid operated injection valve (Burkard Scientific, Uxbridge, Middlesex, UK) and a pair of two-way solenoid operated switching valves (Biochem, East Hanover, NJ, USA) for selection of sample or calibration standards. A solid state photometric detector was used, which incorporated a single-channel aluminium flow cell block (in-house design and construction), a red light emitting diode source (λ_{max} = 635 nm; stock no. 590-480, RS Components) and a photodiode light detecting component (1.75 mm; stock no. 194-290, RS Components). A liquid crystal display (LCD) screen was positioned above the peristaltic pumps, and used to display simple messages describing monitor status during operation.

Power was supplied from an external 12 V lead-acid vehicle battery (89 Ah), which had an operational lifetime of over three weeks. The FI manifold components accounted for the majority of power consumption, as shown in Table 2.2.

Table 2.2 Power consumption of the portable monitor components [58].

Component	Power consumption (mA)
Peristaltic pumps (5 V)	200 each
Injection valve (12 V)	400
Switching valves (12 V)	300 each
Microcomputer (12 V)	40 active 20 idle

2.3.4 Procedures

A manual FI system was optimised with respect to the manifold variables in order to produce the widest possible range of linear response. The optimisation process included an investigation of the effect on linear range of using different sulfonephthalein indicators for the acceptor stream. The buffering capacity of the sodium hydroxide reagent on acid samples was assessed, and the effect of alkyl amine interference on the method was investigated. The degree of variability between different samples of PTFE membrane was also examined.

Having determined the optimal manifold variables, the method was adapted to the portable, automated monitor. This system was subjected to a stability trial to determine whether any significant changes in response occurred during an eight-day period of unattended, continuous operation. The effect of ambient temperature fluctuations on monitor performance was investigated, and the degree of variability in the baseline signal was assessed. Validation of the monitor was performed by comparing results for analyses of aqueous ammonia standard solutions and spiked liquid effluent samples with those obtained using a standard laboratory spectrophotometric method (the indophenol blue method), which is described below.

In order to assess the performance of the portable ammonia monitor under real industrial process conditions, it was deployed at two chemical production sites to perform

automated, on-line analyses of the effluent streams. The results at one of the sites were validated by comparison with those obtained by laboratory analyses using the indophenol blue spectrophotometric method.

2.3.5 The indophenol blue reference method

The version of the indophenol blue method used for validation of the monitor performance was the standard method employed for determination of ammonia in effluents at one of the sites where on-line monitor trials were performed. A series of 50 ml volumetric flasks were used, one for a reagent blank, one for each sample and six for calibration standards. To each sample flask was added 1.0 ml of well mixed sample, and to the calibration flasks, 0.5, 1.0, 2.0, 4.0, 6.0 and 8.0 ml of 10 mg l^{-1} NH_3-N standard solution (giving calibration standards in the range 5 - 80 mg l^{-1} NH_3-N when made up to volume). To all the flasks were then added 1.0 ml of 0.05 mol l^{-1} EDTA, approximately 30 ml of Milli-Q water and 8.0 ml of sodium phenate solution. Having mixed the contents of each flask, 6.0 ml of freshly prepared sodium hypochlorite solution was added, the flasks were made up to volume with water, then allowed to stand for at least 30 minutes for full colour development to occur. The solutions were measured for absorbance at 620 nm in a 1 cm glass cell, using the reagent blank as a reference.

2.4 RESULTS AND DISCUSSION

2.4.1 Method optimisation

The GD-FI method used for this work was adapted from that used by Clinch *et al.* for the *in situ* determination of ammonia in river water [59], in which bromothymol blue solution is used for the indicator stream. One of the most important requirements for an on-line

effluent monitor is to be able to accurately determine concentrations of the selected analyte over a range which spans the discharge consent level, so that effluents containing the analyte at concentrations in excess of the stipulated maximum can be detected and appropriate action taken (*i.e.* treatment prior to discharge). In this work, 80 mg l^{-1} NH_3-N was the target consent level, and the FI-gas diffusion method was therefore required to produce an appropriate linear response.

Table 2.3 FI manifold conditions used by Clinch *et al.* for the determination of ammonia in river water [59].

FI parameter	Setting
Indicator[a] flow rate	0.7 ml min^{-1}
Carrier flow rate	0.7 ml min^{-1}
Type of flow in GD cell	Concurrent
Injection volume	180 µl
Indicator[a] pH	6.5
Indicator[a] concentration	0.5 g l^{-1}
NaOH concentration	0.4 g l^{-1}

[a] Bromothymol blue

The FI manifold conditions used by Clinch *et al.* (see Table 2.3) were quoted as having a linear calibration in the range 0-5 mg l^{-1} NH_3-N , since the method was optimised for determinations of ammonia at the concentration levels typically found in freshwaters (≤ 200 µg l^{-1}) [59]. A calibration was performed in this work using the same FI parameters, but with a diode-array spectrophotometer used instead of the solid-state photometric detector described by Clinch and co-workers, and 0.05 g l^{-1} bromothymol blue solution (adjusted to pH 6.5) used as the indicator stream. The lower indicator concentration was used since it was found that a very noisy signal was produced when using a 0.5 g l^{-1} solution, owing to a very high degree of light absorbance. This produced a linear response

up to 20 mg l^{-1} NH_3-N (R^2 = 0.9977, gradient = 0.047 and *y*-intercept = 0.055 absorbance units) for absorbances measured at 634 nm. This wavelength corresponded with the maximum emission of a red LED, as used in the solid-state photometer employed by Clinch *et al.* However, response was not linear up to the desired minimum of 80 mg l^{-1} NH_3-N, and the manifold was therefore optimised to extend the linear range.

A full simplex optimisation (see Chapter 5 for a description of the principles of this technique) was performed for six manifold variables, namely the carrier and indicator stream flow rates, the volume of injected sample, the concentration of sodium hydroxide solution, and the concentration and pH of the bromothymol blue solution. The simplex procedure was conducted using 10 mg l^{-1} NH_3-N standard solution, and the results are given in Table 2.4. In all cases, countercurrent flow was used for the carrier and indicator streams when passing through the gas diffusion cell, thereby minimising the period during which ammonia could diffuse across the membrane. Absorbance was measured at 616 nm, which is the wavelength of maximum absorbance for the basic form of bromothymol blue.

Table 2.4 Results for the simplex optimisation of the FI manifold using 10 mg l^{-1} NH_3-N standard.

Variable	Precision	Range		Optimum value[a]
		Upper value	Lower value	
Flow rate (ml min^{-1}):				
Indicator[b]	0.1	1.8	0.8	1.5
Carrier	0.1	1.8	0.8	0.9
Injection volume (µl)	10	150	10	150
Indicator[b] pH	0.2	5.6	7.6	5.6
Indicator[b] conc. (g l^{-1})	0.05	0.50	0.05	0.3
NaOH conc. (mol l^{-1})	0.005	0.050	0.005	0.010

[a] Optimisation procedure ended after 33 runs
[b] Bromothymol blue

Simplex optimisation was able to determine the FI manifold conditions which would produce an optimal response at the 10 mg l^{-1} NH_3-N level. However, these conditions produced a response which was only linear over the concentration range 0-40 mg l^{-1} NH_3-N (R^2 = 0.998, gradient = 0.082 and *y*-intercept = 0.025 absorbance units). For this reason, further univariate optimisations of the manifold parameters were performed to extend the linear range. One of the simplest ways of achieving a longer range of linear response was to reduce the volume of injected sample. Table 2.5 indicates that the linear range was doubled by decreasing the sample volume from 150 to 30 µl. It was decided also to reduce the concentration of the bromothymol blue solution from 0.3 to 0.1 g l^{-1}, since this reduced the operating costs of the method (an important consideration for an on-line method) and was found to have no adverse effect on the linear range. Increasing the flow rates of the countercurrent indicator and carrier streams to 1.6 ml min^{-1} each was found to have no significant effect on linear range (see Table 2.6), which indicates that the diffusion process was very rapid. However, the increased flow rates offered the advantage of reducing the time of analysis by approximately 15 s. It was discovered that linearity up to 100 mg l^{-1} NH_3-N (see Figure 2.7 and Table 2.7) could be achieved by a further small reduction in the injection volume to 20 µl and by adjusting the indicator to pH 5.4.

Table 2.5 Effect of injection volume on linear response range[a].

Injection volume (µl)	Linear range (mg l^{-1} NH_3-N)	R^2	Gradient	*y*-Intercept (abs. units)
150	0-40	0.998	0.082	0.025
70	0-60	0.999	0.059	0.051
30	0-80	0.998	0.036	0.074

[a] All other FI parameters set at optimal levels determined by simplex.

Table 2.6 Effect of relative flow rates on linear response range[a].

Flow rate (ml min^{-1})	Linear range (mg l^{-1} NH_3-N)	R^2	Gradient	*y*-Intercept (abs. units)
Indicator stream = 1.5 Carrier stream = 0.9	0-80	1.000	0.035	-0.003
Indicator stream = 1.6 Carrier stream = 1.6	0-80	0.997	0.030	-0.031

[a] Injection volume = 30 µl; bromothymol blue = 0.1 g l^{-1} at pH 5.6; sodium hydroxide = 0.01 mol l^{-1}.

Figure 2.7 Linear response range achieved using optimised manifold parameter settings (error bars represent ± 3σ for each mean absorbance value, where n = 3).

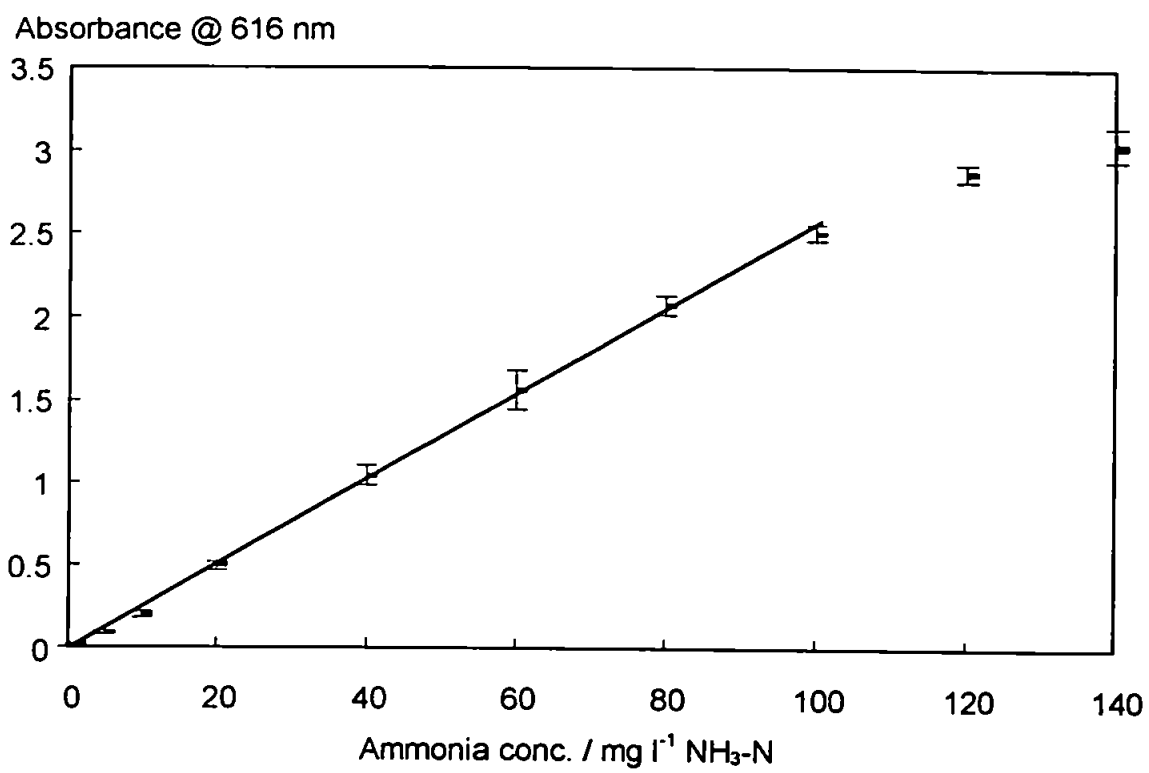


Table 2.7 Optimised FI manifold conditions and linear regression parameters for linear response in the range 0-100 mg l^{-1} NH_3-N.

FI parameter	Level/value
Indicator[a] flow rate	1.6 ml min^{-1}
Carrier flow rate	1.6 ml min^{-1}
Type of flow in GD cell	Countercurrent
Injection volume	20 µl
Indicator[a] pH	5.4
Indicator[a] concentration	0.1 g l^{-1}
NaOH concentration	0.01 mol l^{-1}
R^2	0.999
Gradient	0.026
y-Intercept	-0.013

[a] Bromothymol blue.

2.4.2 Alternative acid-base indicators

The use of other indicators for the determination of ammonia by FI-gas diffusion methods has also been reported, particularly phenol red (phenolsulfonephthalein) and bromocresol purple (5',5''-dibromo-*o*-cresolsulfonephthalein) [114-116], which have working ranges overlapping with that of bromothymol blue. The working range of an indicator is the pH range over which it changes from the acidic to the basic form, and undergoes a contrasting change in colour. The transition range is approximately equal to pK ± 1, where pK is the apparent indicator constant shown in equation 2.3:

$$pH = pK + \log \frac{[I^{2-}]}{[HI^{-}]} \qquad (2.3)$$

Actual working ranges (as shown in Table 2.8 for the three indicators used in this work) do not necessarily correspond with pK ± 1 however, since the human eye is not sensitive to some colour changes. The quoted limits therefore relate to *visual* transition ranges. The

absorbance spectra of the acidic and basic forms of the three indicators are shown in Figure 2.8 (overleaf), together with the spectra at pH levels in the middle of the working ranges.

Table 2.8 Working ranges and pK values for three sulfonephthalein indicators [142].

Indicator	Working pH range	Colour change	pK (at 20°C)
Bromocresol purple	5.2-6.8	Yellow-red	6.40
Bromothymol blue	6.2-7.6	Yellow-blue	7.30
Phenol red	6.4-8.0	Yellow-red	8.00

Phenol red and bromocresol purple solutions of varying concentrations and pH levels were used in place of bromothymol blue for the indicator stream, in order to determine whether any improvement in the linear response range could be achieved. Two-variable optimisations were first conducted to determine which combination of concentration and pH level produced the optimum response for a 10 mg l^{-1} NH_3-N solution. In each case, the other manifold parameters were maintained at the settings determined by simplex optimisation using bromothymol blue (as indicated in Table 2.4 previously). The results of this optimisation procedure is summarised for both indicators in Table 2.9.

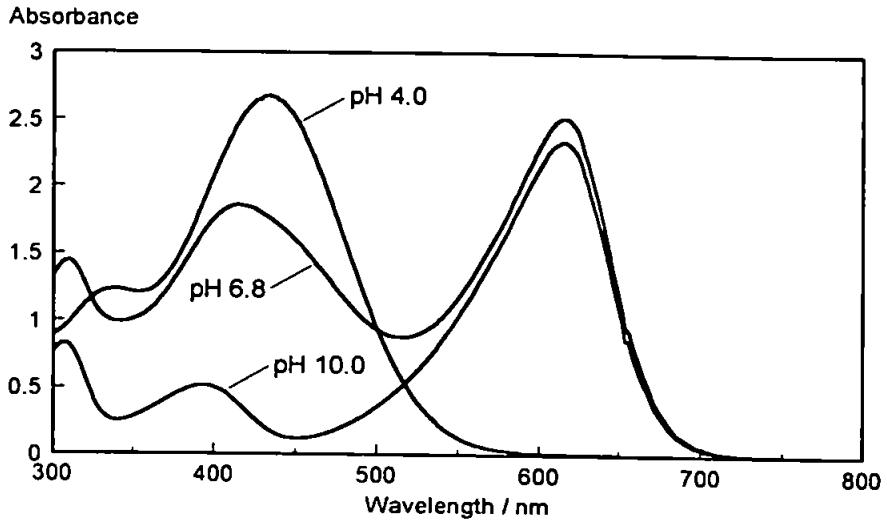
Table 2.9 Two-variable optimisation results for phenol red and bromocresol purple indicators[a].

Indicator	Optimisation range	Absorbance wavelength (nm)[b]	Optimal levels
Phenol red	Conc. = 0.05-0.20 g l^{-1} pH = 6.4-7.2	552	Conc. = 0.1 g l^{-1} pH = 6.6
Bromocresol purple	Conc. = 0.05-0.25 g l^{-1} pH = 4.6-6.0	586	Conc. = 0.2 g l^{-1} pH = 4.8

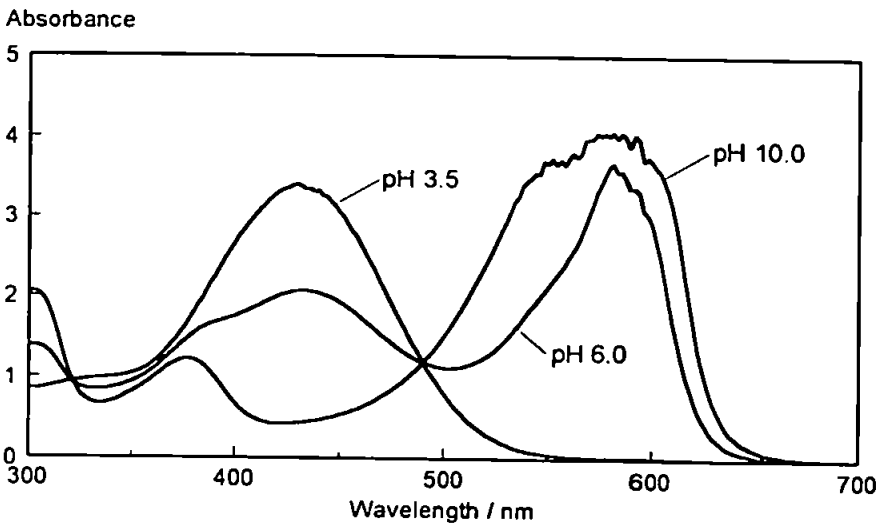
[a] Injection volume = 150 µl; indicator and carrier flow rates = 1.5 and 0.9 ml min^{-1} respectively; sodium hydroxide = 0.01 mol l^{-1};
[b] $\lambda_{maximum}$ values for the basic form of each indicator.

Figure 2.8 Absorbance spectra of (a) bromothymol blue, (b) bromocresol purple and (c) phenol red at various pH levels, showing acidic, mid-transition and basic forms of the indicator.

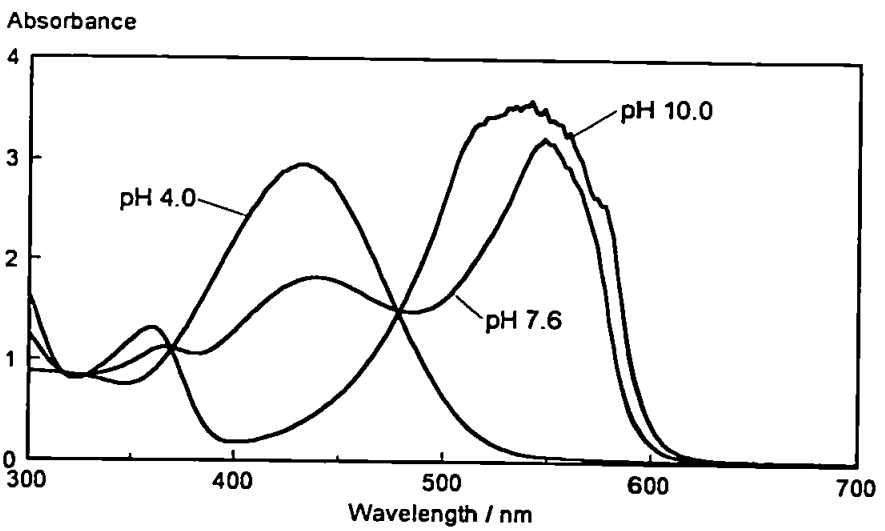
(a)



(b)



(c)



Calibrations were then performed using the optimal indicator levels, and for comparison, using the combination of levels which produced a low response for each indicator during the optimisation procedure. In this case, the injection volume was reduced to 30 µl, but flow rates and sodium hydroxide concentration were maintained as previously. Absorbances were measured at the $\lambda_{maximum}$ values for the basic form of each indicator (*i.e.* phenol red = 552 nm; bromocresol purple = 586 nm; bromothymol blue = 616 nm). Table 2.10 indicates that no improvement in linear range was obtained using either phenol red or bromocresol purple (either at the optimal or sub-optimal settings) when compared to that obtained using bromothymol blue. It was therefore concluded that no advantage was offered for this work through the use of phenol red or bromocresol purple.

Table 2.10 Comparison of linear ranges obtained using different sulfonephthalein indicators[a].

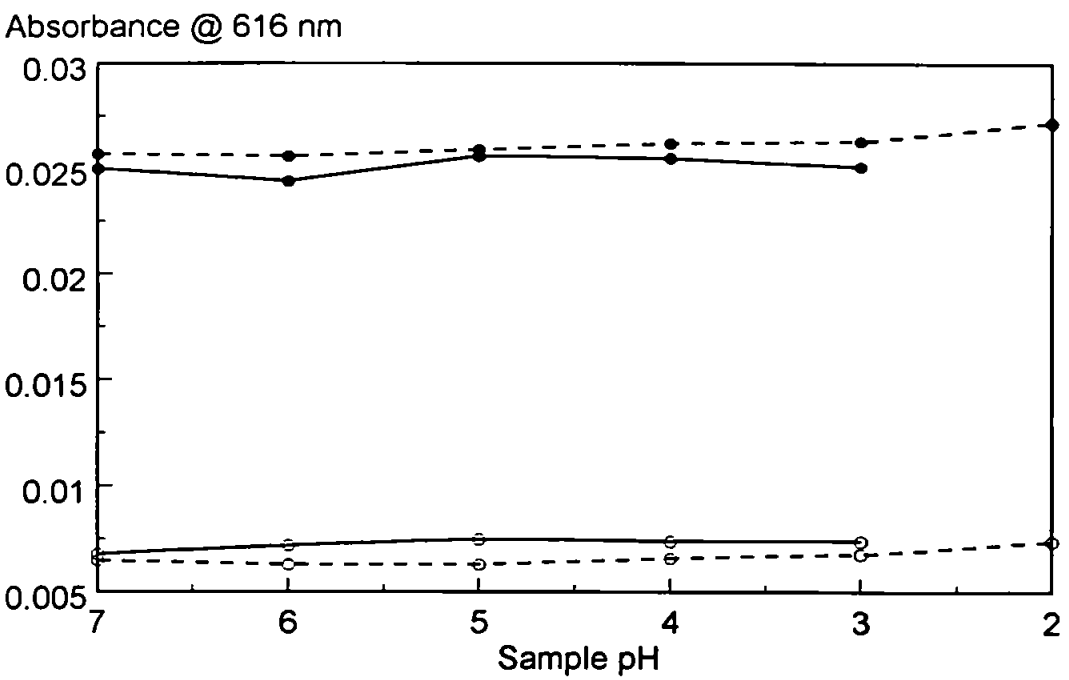
Indicator	Linear range (mg l^{-1} NH_3-N)	R^2	Gradient	*y*-Intercept (abs. units)
Phenol red:				
0.1 g l^{-1}; pH 6.6[b]	0-20	0.999	0.059	0.065
0.05 g l^{-1}; pH 6.4[c]	0-40	0.998	0.050	0.057
Bromocresol purple:				
0.2 g l^{-1}; pH 4.8[b]	0-40	0.997	0.067	0.035
0.05 g l^{-1}; pH 4.8[c]	0-60	0.997	0.053	-0.006
Bromothymol blue[d]:				
0.1 g l^{-1}; pH 5.6	0-80	1.000	0.035	-0.003

[a] Injection volume = 30 µl; indicator and carrier flow rates = 1.5 and 0.9 ml min^{-1} respectively; sodium hydroxide = 0.01 mol l^{-1};
[b] Optimal levels;
[c] Sub-optimal levels;
[d] Linear range determination shown in Table 2.6.

2.4.3 Buffering capacity of sodium hydroxide reagent on acidic samples

The pH levels of industrial liquid effluents can vary considerably with time, particularly at sites with multiple production streams and/or reactors. Since the gas diffusion method used for this work was dependent upon raising sample pH in order to convert all NH_4^+ to gaseous NH_3 it was important to establish the tolerance of the method to samples of low pH. Standard solutions (1.0 mg l^{-1} NH_3-N) and water blanks were adjusted to varying pH levels between 1 and 7, then analysed. Figure 2.9 gives the results for NH_3-N analyses using both 0.01 and 0.10 mol l^{-1} sodium hydroxide. No significant change in response was observed for solutions ≥ pH 3.0 when using 0.01 mol l^{-1} sodium hydroxide, and ≥ pH 2.0 in the case of 0.10 mol l^{-1} sodium hydroxide. Below these pH levels, there was an observable negative response due to the presence of residual acidic gases (*e.g.* CO_2) in the sample stream diffusing into the indicator stream.

Figure 2.9 Effect of decreasing sample pH on response. —○— represents blank soln./0.01 ml l^{-1} NaOH; —●— represents 1 mg l^{-1} NH_3-N/0.01 ml l^{-1} NaOH; – –○– – represents blank soln./0.1 ml l^{-1} NaOH; – –●– – represents 1 mg l^{-1} NH_3-N/0.1 ml l^{-1} NaOH.



2.4.4 Interferences

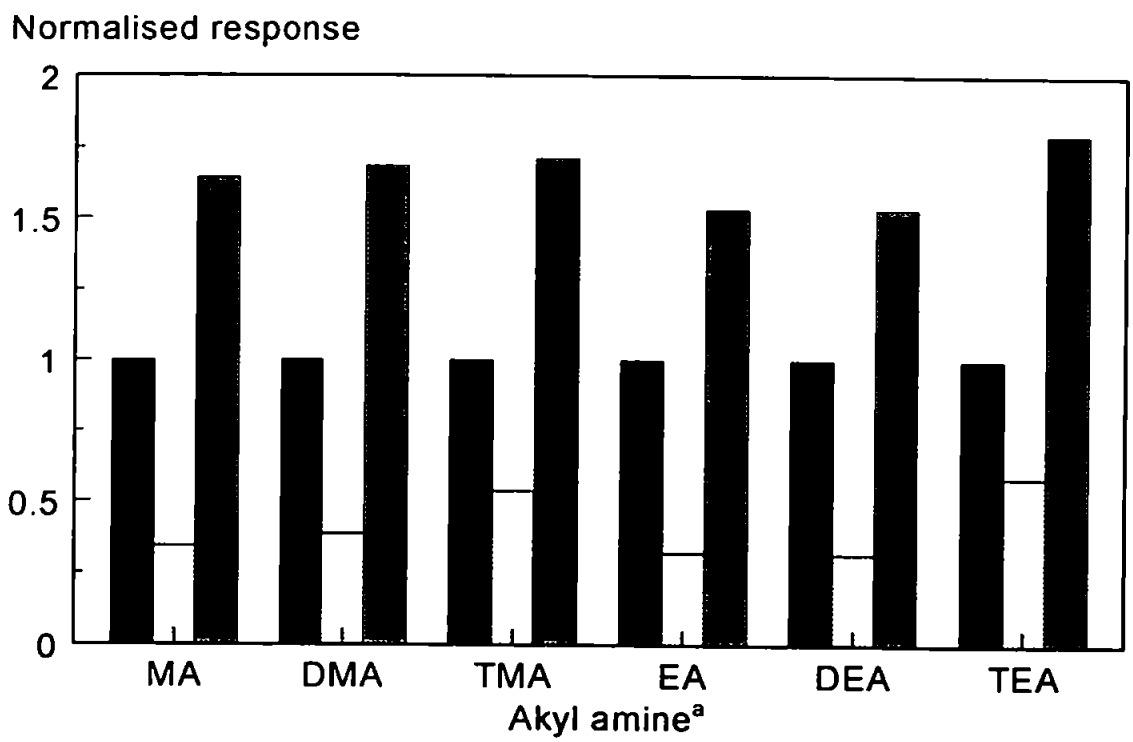
Methods aimed at the determination of low concentrations of ammonia (*e.g.* seawater ammonia levels rarely exceed 1.0 μmol l^{-1}, or 14 μg l^{-1} NH_3-N [115]) must take care to exclude atmospheric CO_2 from the indicator solutions, otherwise the buffering capacity of the indicator will be increased and analytical sensitivity will be reduced [112]. However, the method reported here is intended to monitor ammonia concentrations at the mg l^{-1} level, and for this reason CO_2 does not significantly interfere. Determinations of ammonia by gas diffusion methods can also be affected by alkyl amines [143], particularly methyl and ethyl amines which exhibit similar physicochemical properties to ammonia (see Table 2.11). To assess the effect of alkyl amine interference, two-component mixtures of ammonia with each of the methyl and ethyl primary, secondary and tertiary amines (all at 2 mg l^{-1} with respect to N, and hence of equal molarity) were analysed by the FI-gas diffusion method. The absorbances obtained for ammonia/alkyl amine mixtures were compared with those for pure, single-component solutions.

Table 2.11 Some physicochemical properties of methyl and ethyl amines in comparison to those of ammonia [143, 144].

Compound	Boiling point °C	pK_a (25°C)	pK_b (25°C)
Ammonia	-33.4	9.24	4.76
Methylamine	-7.5	10.68	3.32
Dimethylamine	7.5	10.77	3.23
Trimethylamine	3.0	9.80	4.20
Ethylamine	17	10.63	3.37
Diethylamine	55	10.93	3.07
Triethylamine	89	10.75	3.25

Figure 2.10 gives blank-subtracted absorbance values for each solution, normalised to the absorbance for a 2 mg l^{-1} NH_3-N solution (0.062 @ 616 nm). The six alkyl amine species produced significant responses both individually and in the presence of ammonia. The relative magnitudes of these responses can be related to differences in both the basicity of the amines and their diffusivity across the membrane. The presence of alkyl amines in the sample would therefore interfere with the ammonia response. This is not surprising when the physicochemical similarities between the species are considered, and such interference is also reported for potentiometric and other photometric methods (*e.g.* the indophenol blue reaction) of ammonia analysis [145]. Therefore if amines are present in the effluent stream then the method will give an integrated amine-N response.

Figure 2.10 Comparison of blank-subtracted absorbances for pure and mixed solns. of ammonia, methylamines and ethylamines. ■ represents 2 mg l^{-1} NH_3-N; □ represents 2 mg l^{-1} amine-N; ▬ represents 2 mg l^{-1} NH_3-N/2 mg l^{-1} amine-N.



[a] MA = methylamine; DMA = dimethylamine; TMA = trimethylamine; EA = ethylamine; DEA = diethylamine; TEA = triethylamine

2.4.5 Effect of membrane variability on response

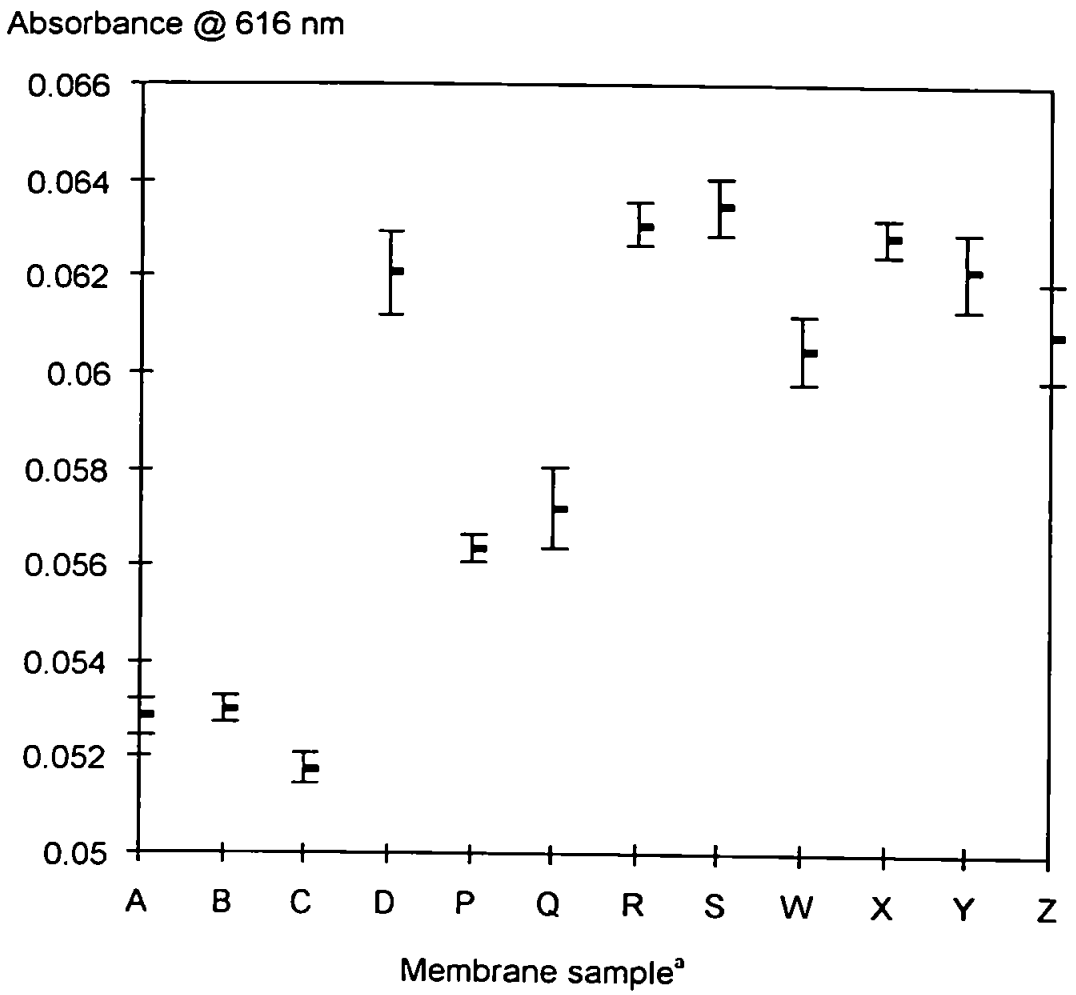
As described earlier, general purpose PTFE tape was used for the gas diffusion membrane in this work. A certain degree of variability was known to exist in the porosity of this material, and an investigation was therefore made to determine the effect of different membrane strips on instrumental response. Three rolls of PTFE tape were used (two obtained from RS Components and one of unknown origin), and four separate strips were taken from each roll. Ten replicate analyses of a 2 mg l^{-1} NH_3-N solution were performed for each membrane strip using the optimised conditions listed previously in Table 2.7. An analysis of variance test was then used to determine any significant differences between the mean absorbances obtained with each strip.

Table 2.12 summarises the results of the analysis of variance test, while Figure 2.11 indicates the spread of mean absorbance values for each membrane strip. The very high F-statistic indicated that very significant differences existed between some of the mean values. A subsequent least significant difference test revealed that significant differences existed both between and within the three different rolls. These results indicated that the inherent variability of the materials used for the membrane would adversely affect the reproducibility of the response signal (particularly at low absorbance levels) unless the system was recalibrated when the membrane was changed. A regular self-calibration protocol would circumvent this problem in an on-line monitor however.

Table 2.12 Summary of analysis of variance for the membrane variability assessment.

Source of variance	Degrees of freedom	Sum of squares	Mean square	F-statistic
Between samples	11	0.0022	2.0×10^{-4}	160.8
Within samples	108	1.3×10^{-4}	1.2×10^{-6}	
Total	119	0.0023		

Figure 2.11 Mean absorbance values obtained for a 2 mg l^{-1} NH_3-N solution using various membrane strips (error bars represent the 95 % confidence interval of the means).



[a] A-D = RS Components roll 1; P-S = unknown supplier; W-Z = RS Components roll 2.

2.4.6 Portable monitor development

Having developed the method using a laboratory FI system, it was then adapted to a portable, automated FI monitor (illustrated schematically in Figure 2.5 previously). Two ammonia standard solutions (2 and 80 mg l^{-1} NH_3-N) were included in the system for automated two-point calibration. The indicator and the carrier streams were each pumped at a flow rate of 1.2 ml min^{-1} using 1.14 mm i.d. pump tubing. A flow rate of 1.5 ml min^{-1}

could be achieved using 1.30 mm i.d. pump tubing, which more closely matched the optimum value determined previously. However the slower flow rate reduced indicator and carrier consumption by 20 % without significantly affecting the linear response. In both cases a linear range of 0-100 mg l^{-1} NH_3-N was achieved, as shown in Table 2.13. The standard/sample and sodium hydroxide streams were pumped at a flow rate of 1.3 ml min^{-1}.

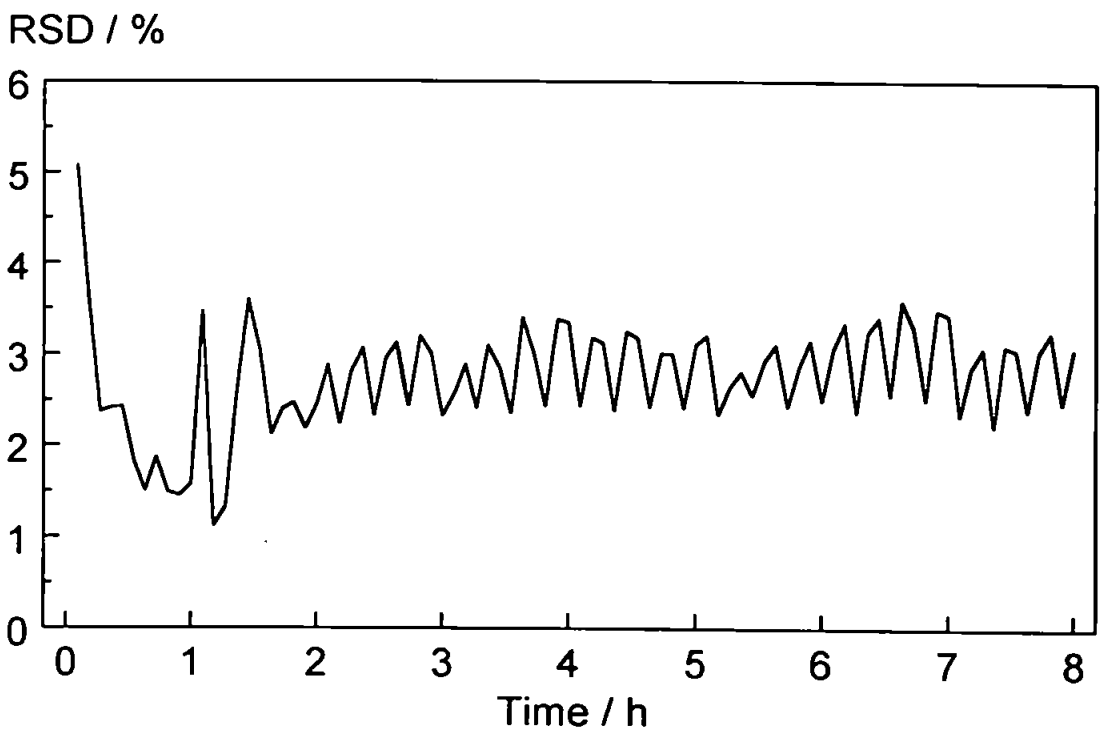
Table 2.13 Effect of reducing carrier and indicator flow rates on the linear response of the portable monitor.

Flow rates (ml min^{-1})	Linear range (mg l^{-1} NH_3-N)	R^2	Gradient	*y*-Intercept (counts)
1.2/1.2	0-100	0.996	42.5	18.1
1.5/1.5	0-100	0.996	37.9	18.3

During normal automated operation, two duplicate injections of the sample and the two calibration standards were made for each analytical cycle. The sample loop was filled for 90 s for the first of each duplicate set of injections, in order to flush the previous sample/standard through the system, and for 20 s in the case of each second duplicate injection. Following injection, the detector output signal was sampled for 60 s, during which time 150 data points were recorded (each point was an average of five A/D integrations). The acquired data was digitally filtered using a moving median algorithm in order to remove any sharp peaks produced by air bubbles in the indicator stream. The first 10 data points were averaged to produce the baseline response, which was subtracted from the maximum filtered value to give the peak height in arbitrary counts. Baseline drift between injections was automatically corrected by adjusting the current supplied to the LED to bring the baseline response within a software defined window (600-1600 counts).

The degree of baseline noise was determined by allowing the monitor to operate continuously during an 8 hour period, but using Milli-Q water in place of the sample, calibration standards and sodium hydroxide reagent. Relative standard deviations were then calculated for each set of 150 data points (*i.e.* the baseline signals produced following each injection). Figure 2.12 demonstrates that baseline noise was typically lower than ± 5 %.

Figure 2.12 Plot of baseline noise for the portable monitor during an 8 hour period of continuous operation.



The total run time for each analytical cycle (*i.e.* six injections) was 28 mins., approximately half of which was required for on-board data processing and analysis, *i.e.* data filtering, calculation of peak height, calculation of replicate mean and RSD values, and calculation of sample analyte concentration. Reagent consumption for one analytical cycle is shown in Table 2.14.

Table 2.14 Monitor reagent consumption for a typical (6-injection) analytical cycle.

Reagent	Consumption (ml)
Calibration standards	2.4 each
Sodium hydroxide	6.9
Bromothymol blue	10.4
Water (carrier)	10.8

A calibration in the range 0-2 mg l^{-1} NH_3-N (see Figure 2.13) was performed to determine the limit of detection, which was defined as $y_b + 3S_{x/y}$ [146], where y_b is the blank response and $S_{x/y}$ is the standard error of the y estimate. The results are summarised in Table 2.15, and the detection limit was determined to be 0.6 mg l^{-1} NH_3-N for this system.

Figure 2.13 Linear calibration for determination of portable monitor detection limits (error bars represent ±3σ for each mean absorbance value, where n = 10).

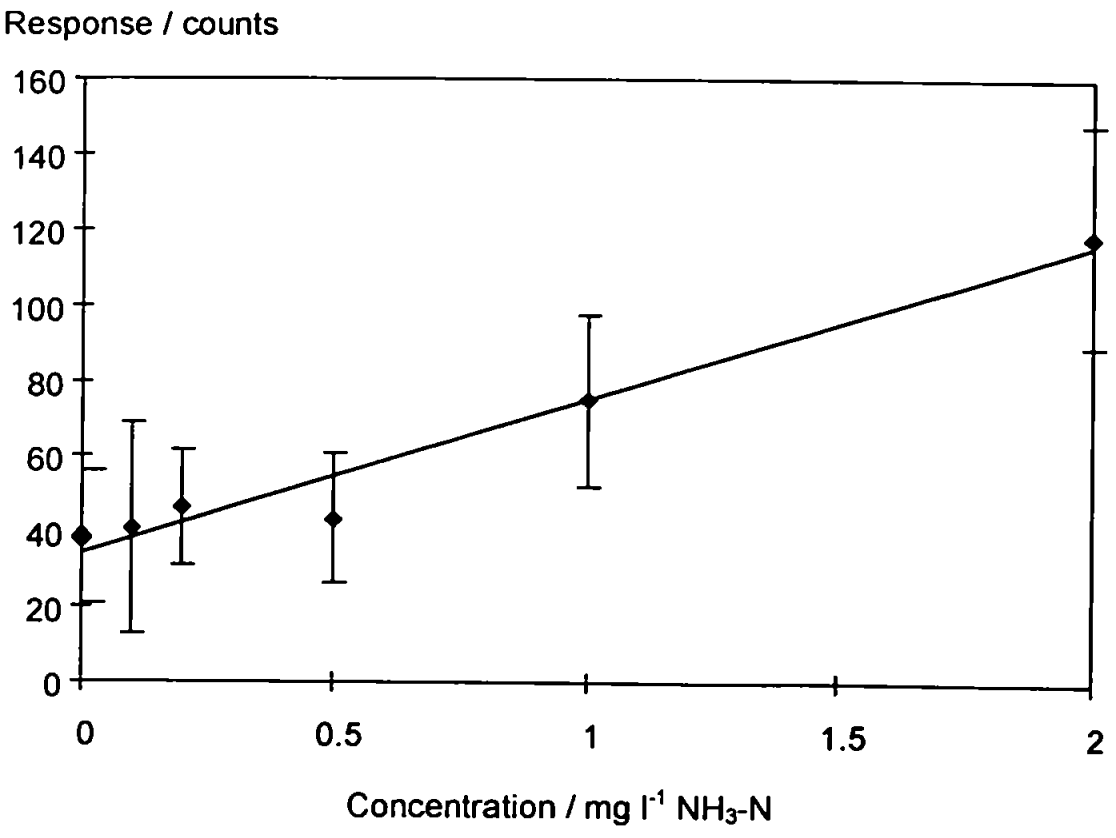


Table 2.15 Linear regression parameters for determination of portable monitor detection limits.

Linear range (mg l^{-1} NH_3-N)	R^2	Gradient	*y*-Intercept (counts)	y_b (counts)	$S_{x/y}$	Limit of detection (mg l^{-1} NH_3-N)
0-2	0.965	40.6	34.7	38.3	6.55	0.6

The precision of the monitor was determined by 10 replicate analyses of both the low and high calibration standards (2 and 80 mg l^{-1} NH_3-N), the results of which (Table 2.16) demonstrated a high degree of precision.

Table 2.16 Portable monitor precision.

Concentration (mg l^{-1} NH_3-N)	Response (counts)		RSD (%)[a]
	Range[a]	Mean[a]	
2	68-75	72.2	3.4
80	2997-3053	3031	0.7

[a] $n = 10$

The predictive performance of the monitor was evaluated using both ammonia standard solutions and samples of a filtered (0.45 μm cellulose nitrate membrane) industrial effluent spiked with varying concentrations of ammonia. Table 2.17 lists results for aqueous ammonia solutions, for which the monitor bias did not exceed ± 10 % for any solutions within the linear response range. Table 2.18 compares the results obtained for analyses of spiked effluent samples using the portable monitor and the manual indophenol blue spectrophotometric method. In this case the monitor bias was < ± 12 %.

Table 2.17 Predictive performance of the monitor for ammonia standard solutions.

Actual conc. (mg l^{-1} NH_3-N)	Predicted conc. (mg l^{-1} NH_3-N)	Bias[a] (%)
Blank	0.6	∝
1	1.1	+10
2	2.1	+5.0
5	5.3	+6.0
10	9.5	-5.0
20	20.9	+4.5
40	41.0	+2.5
60	62.3	+3.8
80	80.3	+0.4
100	92.3	-7.7
120	103.4	-13.8

[a] Bias = {(predicted conc. - actual conc.)/actual conc.} x 100

Table 2.18 Predictive performance of the monitor for spiked industrial effluent samples, compared with the indophenol blue method.

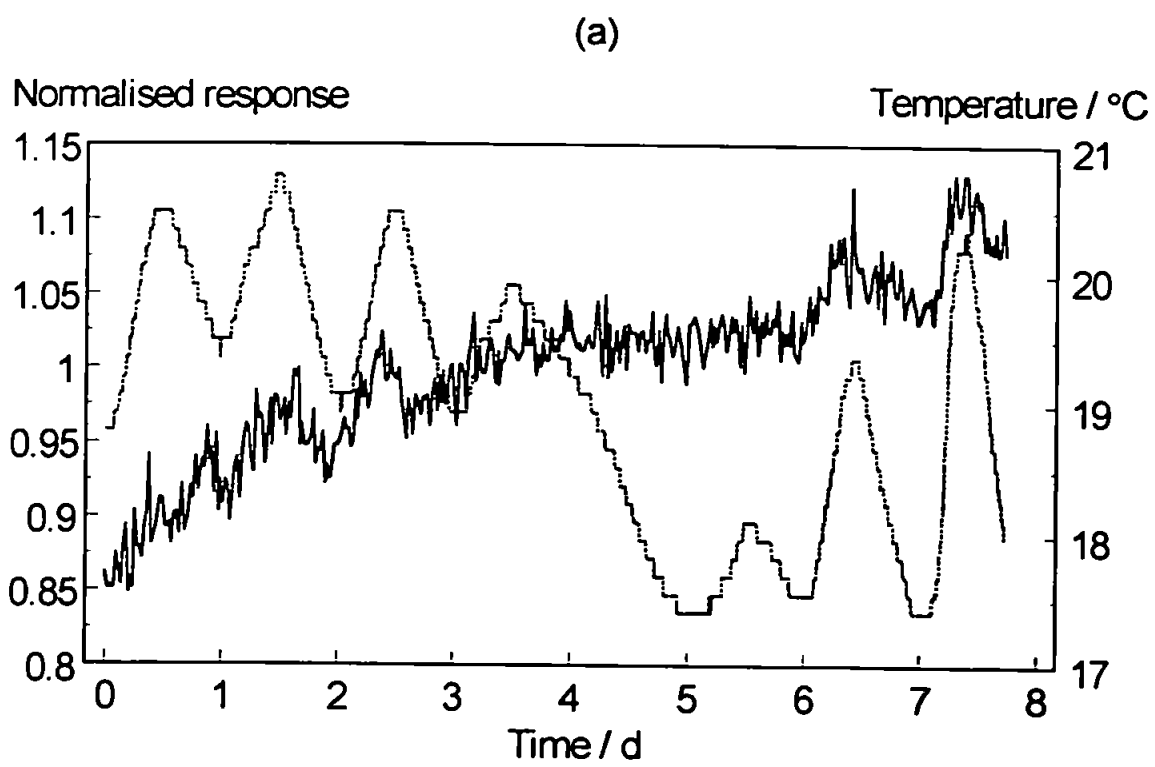
Sample number	Conc. by indophenol blue method (mg l^{-1} NH_3-N)	Conc. by portable monitor (mg l^{-1} NH_3-N)	Bias[a] (%)
1	9.2	9.0	-2.2
2	14.2	15.6	+9.9
3	19.8	21.7	+9.6
4	29.5	32.9	+11.5
5	45.1	49.0	+8.6
6	81.9	78.8	-3.8

[a] Bias = {(monitor conc. - indophenol conc.)/indophenol conc.} x 100

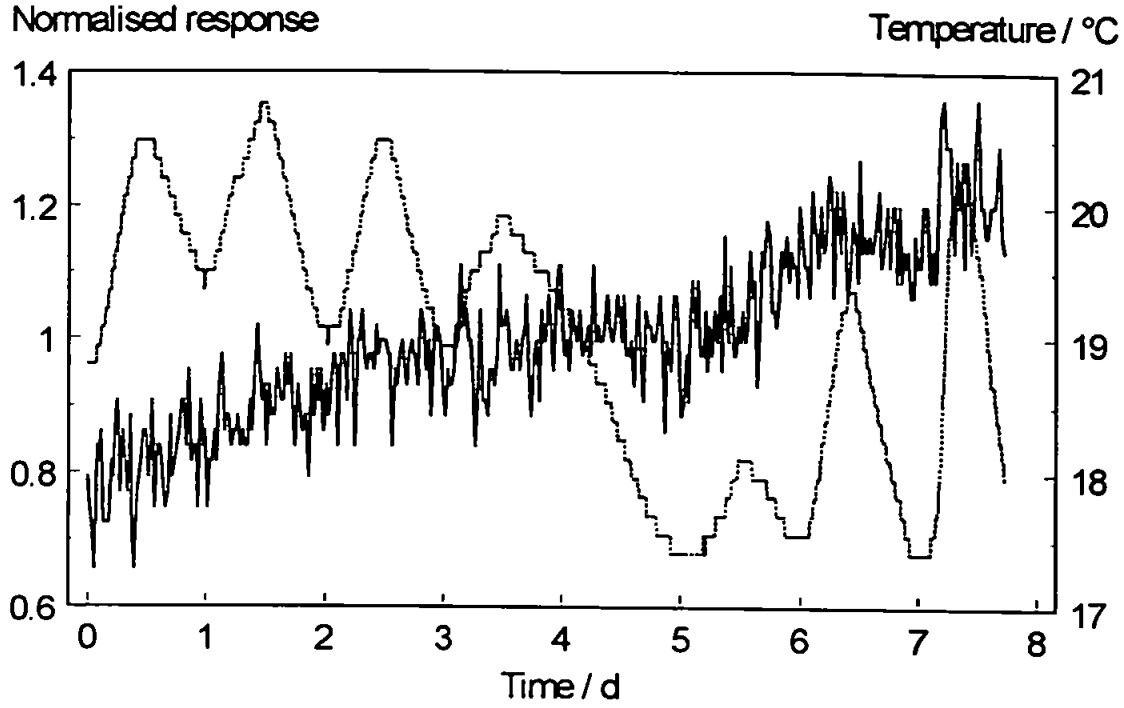
2.4.7 Stability trial

A stability test was undertaken to determine whether any significant changes in monitor response occurred during a week of unattended, continuous operation. A 20 mg l^{-1} NH_3-N solution was used as the sample. Approximately 380 analytical cycles were performed over the period of the trial. The temperature of water in the carrier stream reservoir bottle was recorded every 10 minutes, using a Squirrel SQ32-2U/2V data logger (Grant Instruments, Cambridge, Cambs., UK) with a temperature probe, in order to observe whether monitor response was significantly affected by ambient temperature changes. The results of the trial are displayed in Figures 2.14(a) to 2.14(d), which plot the changes in normalised response, calculated sample concentration and temperature with time.

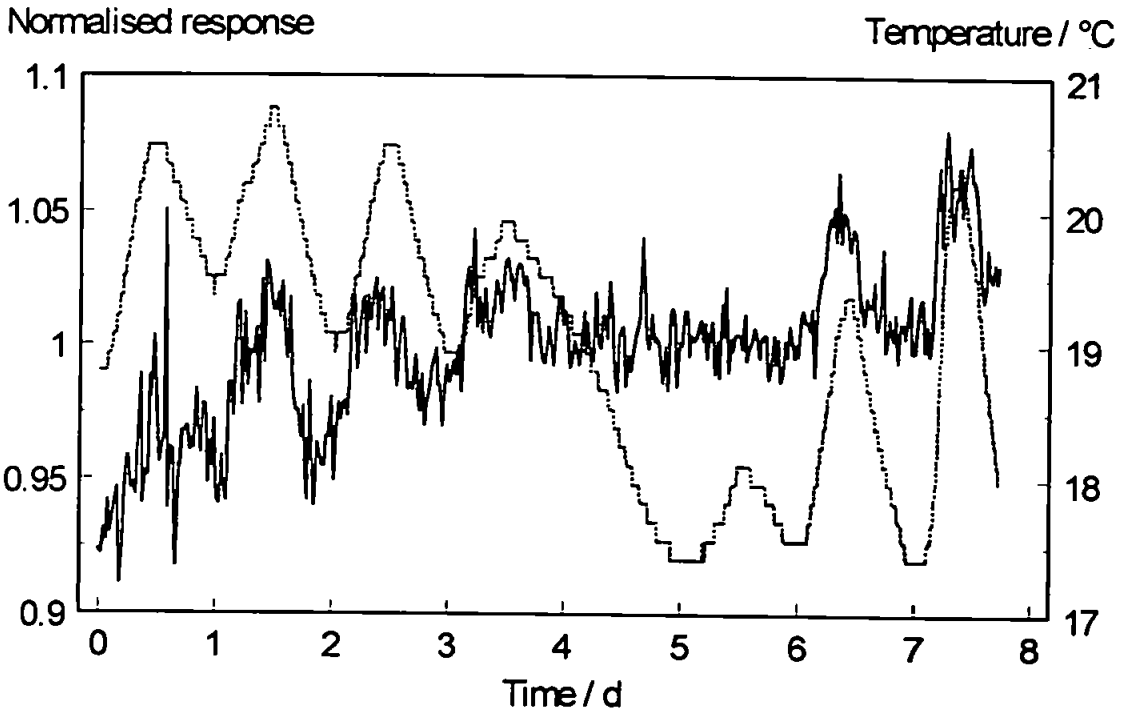
Figure 2.14 Results for continuous analyses of a 20 mg l^{-1} NH_3-N solution during a seven-day stability trial of the portable monitor: (a) normalised 20 mg l^{-1} sample response; (b) normalised 2 mg l^{-1} standard response; (c) normalised 80 mg l^{-1} standard response; (d) calculated sample concentration (dotted line represents temperature; solid line represents monitor response/concentration).



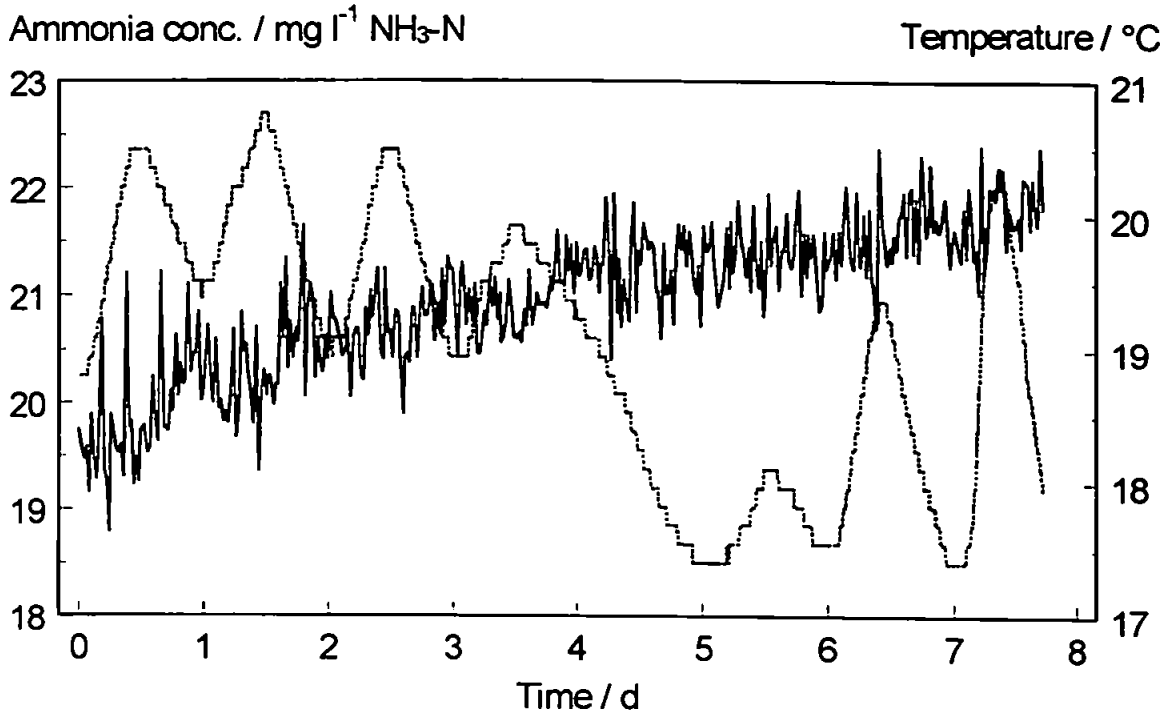
(b)


Normalised response
Temperature / °C
1.4
1.2
1
0.8
0.6
21
20
19
18
17
0
1
2
3
4
5
6
7
8
Time / d

(c)


Normalised response
Temperature / °C
1.1
1.05
1
0.95
0.9
21
20
19
18
17
0
1
2
3
4
5
6
7
8
Time / d

(d)



Fluctuations in ambient temperature exerted a very small influence on the monitor response signals, as shown in Figures 2.14(a) to 2.14(c). However Figure 2.14(d) indicates that temperature changes had no noticeable influence on the calculated ammonia concentration, thus demonstrating that the self-calibration procedure had compensated for this effect. A slight upward drift was observed for the three response signals during the period of the trial, which reflected small changes in pump tubing elasticity and reagent quality. Again this effect is less noticeable in the calculated concentration values, the mean value of which was 21.0 mg l^{-1} NH_3-N, with a relative standard deviation of 3.4 %. These results also show that over the 7-8 day period of the trial the two-point self-calibration procedure enabled the monitor to meet the required level of precision and accuracy (typically 5-10 %) for on-line process monitoring [17, 147]. Replacement of the PTFE membrane was not required during the stability trial as its physical characteristics remained constant.

2.4.8 On-line industrial site trials

The monitor was deployed for on-line analyses of effluent streams at two chemical production facilities to test its operation under real industrial process conditions. At Site 1 the monitor performed automated, unattended analyses for four days. The results were compared with those obtained by an on-line Skalar segmented flow analyser, used at the site as the standard method for ammonia analyses. At Site 2, automated analyses were performed over a two-day period, during which time samples were also collected manually for laboratory analysis by the indophenol blue method. The results obtained by the latter method were used to validate those produced by the monitor. At both sites the monitor was installed in an analyser house alongside existing on-line effluent analysers. At Site 1 the samples were drawn from treated effluent leaving the plant, and at Site 2 the analyser house was located at a point between the main pump house for the works effluent and the on-site effluent management plant. In both cases the sample stream was filtered into a constant head vessel, from which samples were abstracted into the portable monitor, as shown in Figure 2.15. Self-calibration using 2 and 80 mg l^{-1} NH_3-N standard solutions was performed during every analytical cycle.

A typical absorbance spectrum for the effluents under investigation is shown in Figure 2.16, which demonstrates that the highest absorbance was found in the UV region. The results for the on-line trial at Site 1 (Figure 2.17) show that ammonia levels during the period of operation were variable, but did not exceed 70 mg l^{-1} NH_3-N. An accurate comparison with the results obtained by the on-line segmented flow analyser could not be made however, since the analyser was calibrated in the range 0-2000 mg l^{-1} NH_3-N, and the results were plotted on a chart recorder with a resolution of only ± 50 mg l^{-1}. It was therefore impossible to perform a full validation of the portable monitor using this method, but the results appeared to be in general agreement, with none above the 50-100 mg l^{-1} level.

Figure 2.15 Schematic diagram of the sample delivery system for on-line effluent monitoring. ——▶ represents liquid flow; - - - ▶ represents power/data transmission; SV = switching valve.

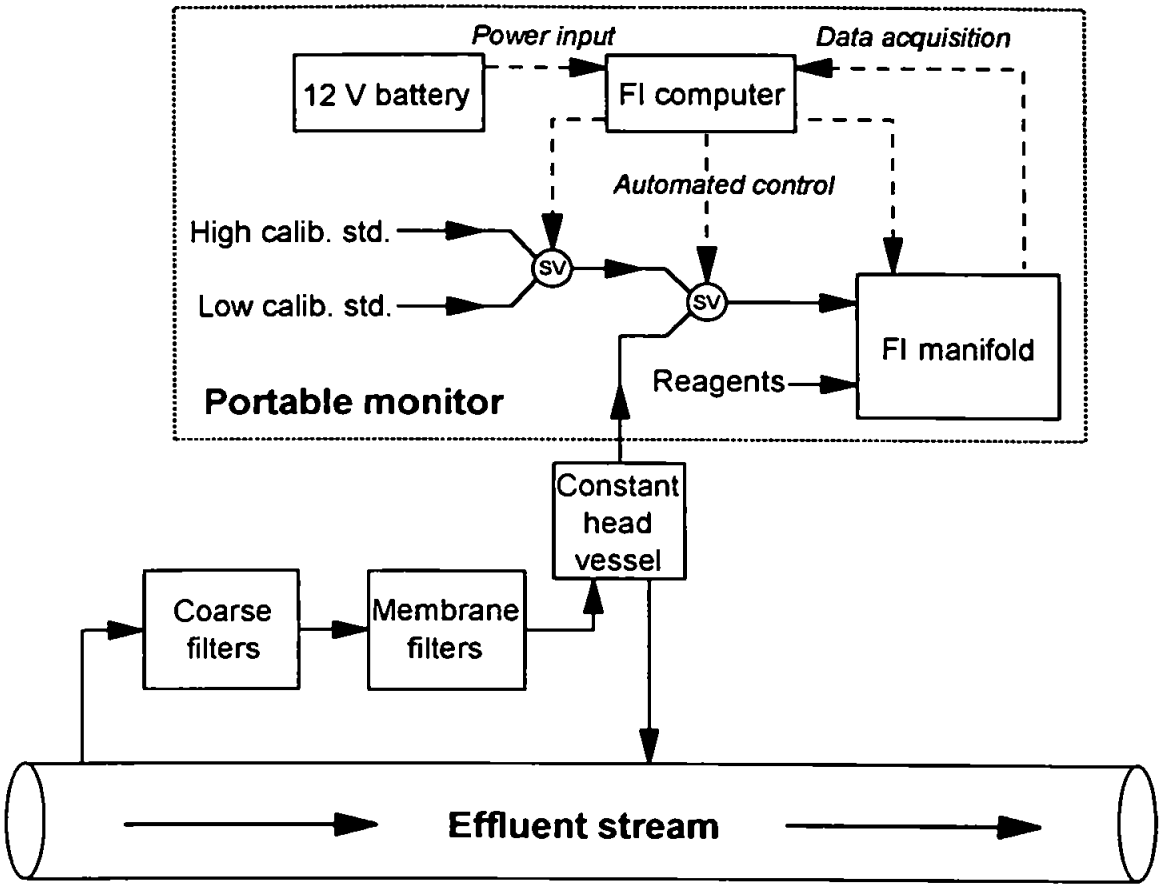


Figure 2.16 Typical absorbance profile of the liquid industrial effluent sample used to evaluate the predictive performance of the monitor.

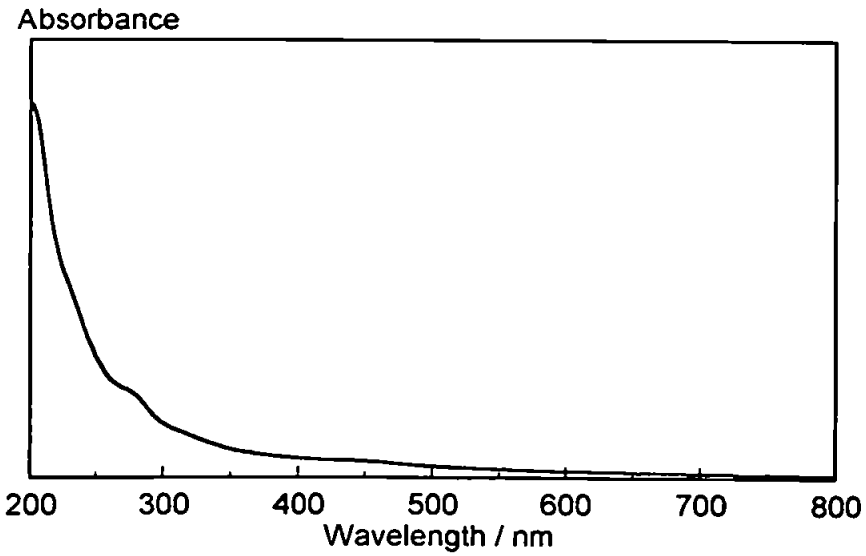


Figure 2.17 Results of the on-line monitor trial at Site 1.

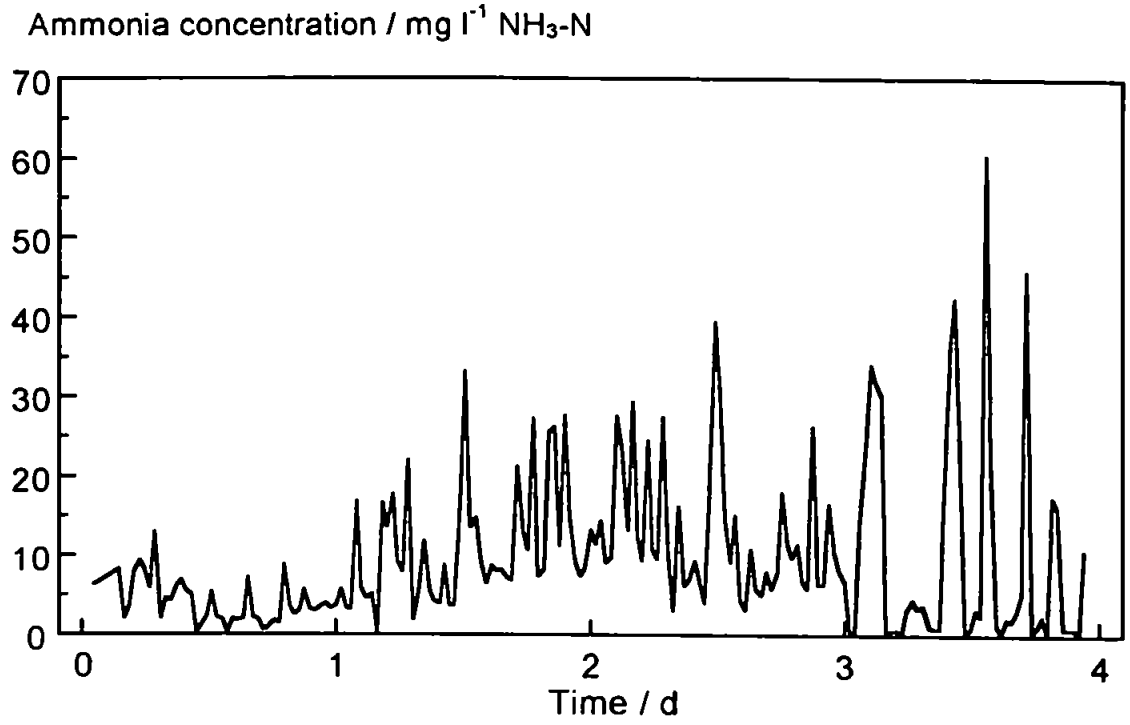


Figure 2.18 Results of the on-line monitor trial at Site 2, compared with those obtained manually by the indophenol blue method (one result above the linear range of the monitor omitted). Solid line represents the monitor results; ● represents indophenol blue results.

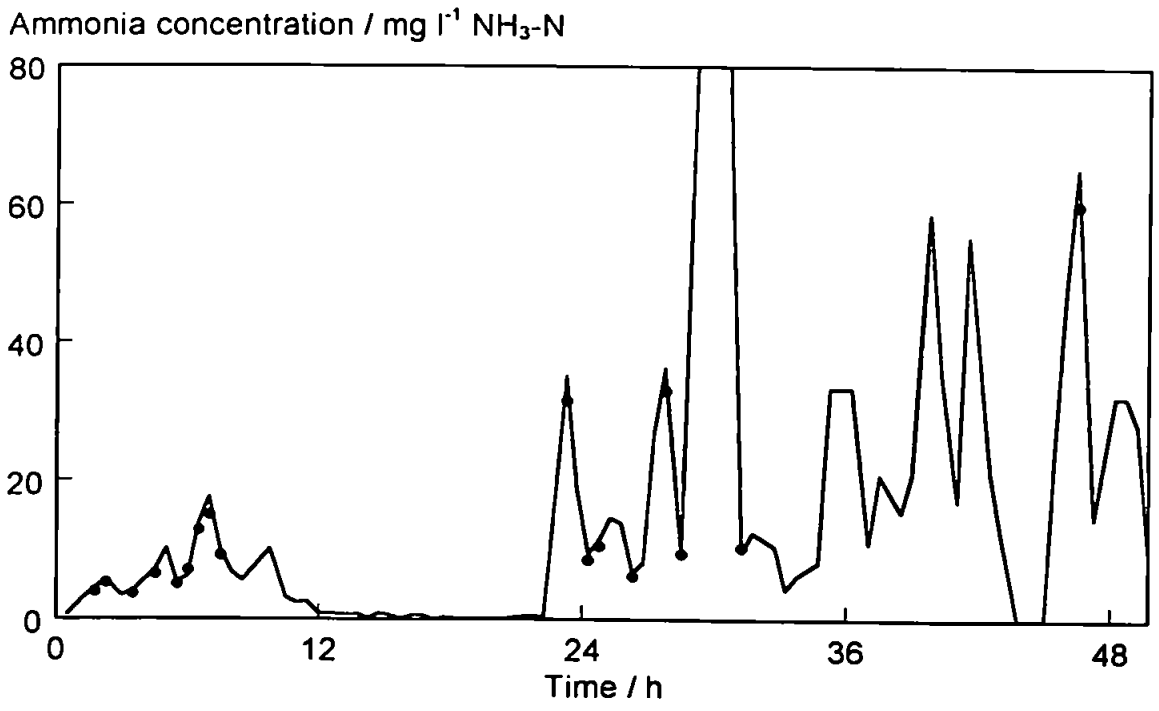


Table 2.19 Effluent pH and a comparison of monitor and indophenol blue results for Site 2.

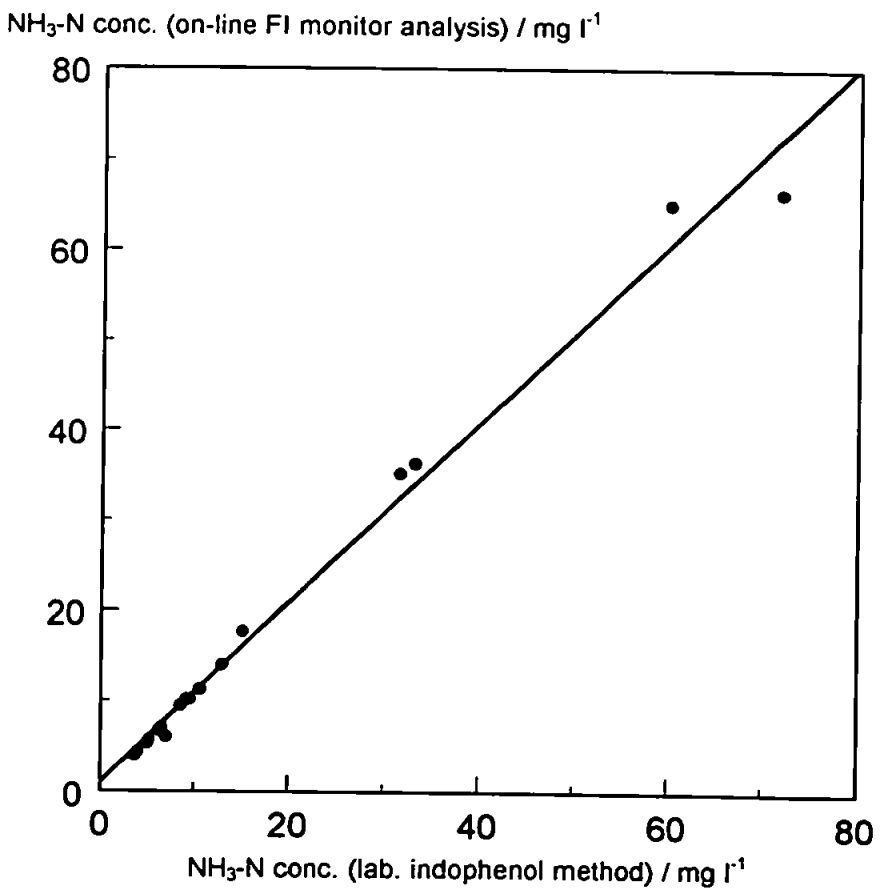
Time of sampling	Effluent pH	Ammonia conc. (mg l^{-1} NH_3)		Bias[a]
		Indophenol blue	On-line monitor	
Day 1:				
11.20	8.2	3.9	4.4	+12.8
11.50	6.4	5.2	5.8	+11.5
13.05	8.1	3.6	4.0	+11.1
14.26	10.7	6.5	7.1	+9.2
15.12	10.9	5.0	5.4	+8.0
15.41	11.8	7.0	6.1	-12.9
16.08	11.6	12.9	14.0	+8.5
16.38	12.2	15.1	17.7	+17.2
17.03	11.4	9.1	10.2	+12.1
Day 2:				
08.43	9.3	31.6	35.2	+11.4
09.46	6.9	8.5	9.5	+11.8
10.15	6.9	10.5	11.3	+7.6
11.46	9.6	6.2	6.8	+9.7
13.24	9.3	33.1	36.3	+9.7
13.58	8.3	9.5	10.2	+7.4
14.28	10.7	71.9	66.5	-7.5
15.31	11.7	238	161	-32.5[b]
16.45	12.0	10.3	9.6	-6.8
Day 3:				
08.08	12.2	60.0	65.2	+8.7

[a] Bias = {(monitor conc. - indophenol conc.)/indophenol conc.} × 100
[b] Concentration above linear response range of monitor

The results for the on-line trial at Site 2 are shown in Figure 2.18 (previous page), together with those obtained by the indophenol blue method in manual laboratory analyses. The period between 12 and 24 h when no ammonia was detected was caused by an interruption in the supply of effluent to the constant head vessel. The result obtained after 30 h of the trial period was above the linear range of the monitor, and was found to be 238

mg l^{-1} NH_3-N by laboratory analysis. Details of the monitor performance as compared with the laboratory analyses are given in Table 2.19 (previous page), together with the effluent pH values. Most results had a positive bias ≤ 17 %, which reflects the fact that the manual samples were stored for up to 12 h prior to analysis, by which time a proportion of the gaseous ammonia originally present in the samples would have been lost (analysis of a 20 mg l^{-1} NH_3-N standard adjusted to pH 9, using the indophenol method, revealed a 3.5 % reduction in the concentration determined initially when analysed after standing in a closed bottle for 24 h). Despite this fact, a good correlation (R^2 = 0.988, gradient = 0.993 and y-intercept = 1.06 mg l^{-1} NH_3-N) was observed between the results for the two methods, as shown in Figure 2.19. Furthermore the deployment was 100 % successful in terms of providing meaningful results over 50 h of continuous, unattended operation.

Figure 2.19 Correlation between the monitor and the indophenol blue results for Site 2 (one result above the linear range of the monitor omitted).



2.5 CONCLUSIONS

A portable, automated FI monitor was developed for the analysis of ammonia in industrial liquid effluent streams. The method employed a gas diffusion cell to enable the transfer of gaseous ammonia through a PTFE membrane into a bromothymol blue indicator stream, with the resulting colour change measured photometrically. The monitor had a linear response for ammonia in the range 1-100 mg l^{-1} NH_3-N, and was tolerant of acidic samples $\geq$ pH 3. If necessary, ammonia concentrations in excess of 100 mg l^{-1} NH_3-N could be accommodated by incorporating an on-line sample dilution step. Similarly, higher concentrations of sodium hydroxide reagent would permit analysis of samples below pH 3. Alkyl amines represented the only significant interferences for this method. The stability of the monitor was demonstrated over a 7-8 day period of continuous operation in the laboratory, and it was shown to be capable of reliable, on-line operation within real industrial process environments. On-line results showed a good correlation with a standard indophenol blue laboratory reference method.

An important advantage of this approach to effluent monitoring is the fact that results are obtained on a near-continuous basis, and therefore provide time-integrated loadings for ammonia discharges. The portability of this monitor would also lend itself to site assessment applications (*i.e.* short-term deployment at a variety of remote sites) in addition to its reported use as an on-line system based within an effluent analyser house.

Chapter 3

Multivariate Calibration

3.1 INTRODUCTION

The development of computer-controlled laboratory instrumentation has provided the analytical chemist with ever larger and more complex data sets. As a result of this, increasingly sophisticated mathematical and statistical methods have been required to derive useful information from the data. This trend first became apparent in the early 1970s with the introduction of a number of new mathematical techniques, such as pattern recognition and multivariate statistics [148]. In 1972 the term 'chemometrics' was proposed by the Swedish physical organic chemist, Svante Wold, as a generic name for the discipline of chemistry in which mathematical and statistical techniques are used for the purposes of optimising experimental design procedures and maximising the information obtainable from analytical data [149 -150]. Two years later he formed the Chemometrics Society in association with the American analytical chemist, Bruce R. Kowalski, in order to provide an international forum for chemists applying formal logic to chemical analyses. Prominent members of the society included D. Luc Massart, Stan N. Deming and Sergio Clementi. [151 -152]. Since this time, chemometrics has expanded into a very prominent area of chemical research, and a growing number of textbooks [149, 153 , 154 , 155 , 156 , 157] and two specialist journals (*Journal of Chemometrics*, Elsevier; *Chemometrics and Intelligent Laboratory Systems*, Wiley) are now available.

One of the most important applications of chemometrics in the field of analytical chemistry is multivariate calibration [158 -159], which can be applied to the quantification of single or multiple analytes when more than one data point is acquired for each sample (*i.e.* multivariate data). This is particularly appropriate in the case of multiwavelength spectroscopic techniques. This chapter discusses the advantages and applications of multivariate calibration, and explains the operation of the more frequently used linear calibration techniques.

3.2 UNIVARIATE VERSUS MULTIVARIATE CALIBRATION

Calibration is the process of determining a mathematical function to relate measured instrumental response (*e.g.* absorbance) to a known parameter (*e.g.* concentration) of a sample analyte, and using this function to predict the same parameter in unknown samples. Typically, response measurements are obtained for a series of samples in which accurate analyte concentration values have been determined independently. The measured response and concentration data are then used to construct a model which relates one to the other, and the model can be used to predict analyte concentrations in new samples on the basis of their measured instrumental response [146, 160].

The simplest form of calibration in analytical chemistry is univariate calibration, in which a single instrumental measurement is used to determine the level of a single analyte. However, with the development of instrumentation capable of rapidly obtaining multiple response data (*e.g.* full spectrum absorbance measurements), it has become desirable to adopt calibration techniques which can fully utilise the available multivariate data to quantify both single and multiple sample analytes. The relative merits of univariate and multivariate calibration are discussed below.

3.2.1 Univariate calibration

Univariate calibration in analytical chemistry involves the measurement of a single variable (*e.g.* absorbance at a particular wavelength) to predict a level (typically concentration) of a single analyte. Calibration is actually a two-stage procedure, involving firstly the construction of a calibration model and secondly the prediction of analyte levels in new samples. In the calibration stage, univariate instrumental measurements are acquired for a series of samples spanning a range of analyte levels. The levels are accurately determined by an independent assay technique, and a least-squares regression procedure is used to

produce a calibration model which relates instrumental response to analyte level. In the prediction stage, instrumental measurements obtained for new samples are incorporated into the calibration model in order to determine levels for the analyte of interest.

One of the most commonly applied univariate calibration procedures is the 'classical' model, which assumes a linear relationship between instrumental measurement and analyte level [160]:

$$y_i = b \cdot x_i + e_i \tag{3.1}$$

where y_i and x_i are the instrumental response and analyte level respectively for sample i, b is the calibration coefficient determined by the least-squares regression of instrumental response on analyte level for the calibration sample set, and e_i is the measurement error associated with y_i.

The principal advantages of univariate calibration techniques are their simplicity of application and ability to produce accurate calibration models using a relatively small number of standards. However, in order to obtain accurate predictions with the univariate approach, instrumental measurements must be highly selective with respect to the analyte of interest, with no interferences affecting instrumental response [161]. If these requirements cannot be met, then predictions of new sample levels are likely to be unreliable. To remove potential interferences, samples may require purification (*e.g.* solvent extraction) or stabilisation (*e.g.* pH buffering) prior to analysis. If interferences cannot be effectively eliminated, then the instrumental measurement must be highly selective for the desired analyte. This can only be achieved if the sample matrix is of low complexity. A further limitation to univariate calibration is the fact that unknown interferences and unreliable prediction values cannot be detected on the basis of single-point measurements.

As discussed in Chapter 1, environmental and industrial process systems are increasingly being analysed *in situ*, and by the nature of these analytical techniques and the complex systems under investigation, it is often impossible to obtain highly selective

measurements or to separate the analyte of interest from all potential interferences. Univariate calibration is generally inappropriate under these circumstances, and for this reason multivariate calibration methods are increasingly being applied to environmental and process analyses.

3.2.2 Multivariate calibration

Multivariate and univariate calibration are similar insofar as they both involve the construction of a calibration model relating instrumental response to analyte level for a set of known standards, and use this model to predict analyte levels in new samples. As the name implies however, multivariate calibration incorporates multiple instrumental measurements of each sample (*e.g.* the spectral data obtained by multiwavelength spectrometers) into the calibration model.

Multivariate calibration has two significant advantages over the univariate approach. Firstly, multivariate instrumental response can be related to the levels of more than one analyte in a sample, thereby enabling simultaneous determination of multiple sample components. Secondly, it follows that instrumental response does not have to be selective for only one analyte, and complete separation of the analyte(s) of interest is therefore unnecessary. Depending on the multivariate calibration technique employed, the effect of both known and unknown matrix interferences can be modelled to a greater or lesser degree, thus providing accurate predictions of multiple analytes in complex samples without the need for elaborate sample preparation. In addition, errors produced in the instrumental response of new samples by interferences not present in the calibration standards can be detected, and the sample rejected as an 'outlier' [162].

The least complex, most widely available and therefore most frequently applied multivariate calibration techniques are those which assume a linear relationship between response signal and analyte level. These include classical least squares (otherwise known as

direct multicomponent analysis), multiple linear regression, principal components regression and partial least-squares regression [163 -164], and are discussed in detail in the following section. In recent years, a number of non-linear multivariate calibration techniques have also been developed, including locally weighted regression, projection pursuit regression and artificial neural networks [165]. These offer the ability to model complex, non-linear relationships between analyte levels and instrumental response, and their application in analytical chemistry will become more widespread as the availability of appropriate commercial software increases.

3.3 MULTIVARIATE CALIBRATION TECHNIQUES

The theory behind four commonly applied linear multivariate calibration methods is discussed in this section, including a description of the algorithms used in the modelling process. In all cases, the following format will be used for algorithmic expressions:

MATRICES written in bold upper case;

vectors written in bold lower case (all vectors are *column* vectors, and all transposed vectors are *row* vectors);

scalars written in italics (upper and lower case).

The algebraic notation used throughout this section is defined in Table 3.1.

3.3.1 Direct multicomponent analysis

Direct multicomponent analysis (DMA) is a relatively simple form of linear multivariate calibration based upon Beer's law. In analytical spectroscopy the Beer's law model states that absorbance at each wavelength is proportional to the sum of the concentrations of each component present multiplied by their molar absorptivities at that wavelength [164].

Table 3.1 Definition of notation used for multivariate calibration algorithms.

$\mathbf{C}$	Component concentrations matrix (*i.e.* analyte concentrations for all samples in the calibration set)
$\mathbf{A}$	Instrumental response matrix (*i.e.* spectra for all samples in calibration set)
$\mathbf{c}_i$	Component concentrations vector for sample *i*
$\mathbf{a}_i$	Instrumental response vector (*i.e.* spectrum) for sample *i*
$\mathbf{T}$	Principal components scores matrix
$\mathbf{B}$	Principal components loadings matrix
I	Total number of samples
J	Total number of analytes
M	Total number of wavelengths
H	Total number of principal components/factors
ε	Molar absorptivity (sensitivity coefficient)
$\mathbf{K}$	Molar absorptivities matrix
β	Regression coefficient (relating concentration to instrumental response)
$\mathbf{P}$	Regression coefficients matrix
$\mathbf{Q}$	Principal components regression coefficients matrix
$\mathbf{w}_h$	Partial least squares loading weights vector for factor *h*
v_h	Partial least squares loading coefficient (*i.e.* the regression coefficient relating scores to concentrations for factor *h*)
$\mathbf{E}_A$	Spectral residuals matrix
$\mathbf{E}_C$	Concentration residuals matrix
$\hat{}$	Estimated parameter
^{T}	Transpose of a matrix or vector

Therefore, for a given sample measurement at wavelength k in a cuvette of unit path length:

$$a_k = \varepsilon_{k1}c_{k1} + \varepsilon_{k2}c_{k2} + \ldots\ldots + \varepsilon_{kj}c_{kj} \tag{3.2}$$

and for a set of calibration standards:

$$\mathbf{A} = \mathbf{CK} + \mathbf{E_A} \tag{3.3}$$

where $\mathbf{A}$ is the $I \times M$ matrix of absorbance spectra, $\mathbf{C}$ is the $I \times J$ matrix of component concentrations, $\mathbf{K}$ is the $J \times M$ matrix of molar absorptivities (*i.e.* pure-component spectra at unit concentration and unit path length) and $\mathbf{E_A}$ is the $I \times M$ matrix of spectral residuals. If $\mathbf{K}$ is unknown, it is estimated during calibration in the following way:

$$\hat{\mathbf{K}} = (\mathbf{C}^T\mathbf{C})^{-1} \bullet \mathbf{C}^T\mathbf{A} \tag{3.4}$$

No estimation of $\mathbf{K}$ is required if the calibration model is built using pure component spectra. The least-squares estimate of component concentration during prediction is given by:

$$\hat{\mathbf{c}}_i^T = (\mathbf{K}\mathbf{K}^T)^{-1} \bullet \mathbf{K}\mathbf{a}_i^T \tag{3.5}$$

DMA is often referred to as the classical least squares (CLS) or K-matrix method. It has the advantage of being a full-spectrum calibration technique (*i.e.* instrumental response measured at all wavelengths can be included in the calibration model), and can offer greater precision than models restricted to a smaller number of response data owing to its signal averaging capabilities [166]. In addition, DMA can offer statistical estimates of pure component spectra which cannot be determined by other means. However, DMA is a *direct* calibration technique in the sense that the model must include pure component spectra or concentration data for all sample components exerting an influence on instrumental response within the required spectral range. This is a significant limitation of the technique, since it is seldom possible to provide the model with information for all the species within a complex sample matrix, and unmodelled spectral interferences will produce large residual errors in predictions of new sample concentrations. These errors can be

minimised by selecting regions of the desired spectral range in which unknown components do not significantly interfere with the response of the analytes of interest, although in many cases the entire spectral range can be subject to interference effects [156, 164].

3.3.2 Multiple linear regression

Multiple linear regression (MLR) is another linear calibration technique related to Beer's law, although in this case the inverse relationship is assumed, *i.e.* concentration is a linear function of instrumental response [158]. Therefore, for analyte *j* of sample *i*, concentration is equal to the sum of the products of absorbance and regression coefficients measured at *M* wavelengths:

$$c_{ij} = \sum_{m=1}^{M} a_{im} \beta_{jm} \tag{3.6}$$

Since this relationship is the inverse of Beer's law, MLR is also known as the inverse least squares (ILS) or P-matrix method [164, 166]. The MLR model can be expressed in matrix terms for all components in a set of calibration standards as shown in equation 3.7:

$$\mathbf{C} = \mathbf{AP} + \mathbf{E_C} \tag{3.7}$$

In this case, $\mathbf{P}$ is the $M \times J$ matrix of unknown regression coefficients which relate analyte concentrations to instrumental response, and $\mathbf{E_C}$ is the $I \times J$ matrix of concentration residuals ($\mathbf{A}$ and $\mathbf{C}$ are the same as for equation 3.3). The least squares estimate of the regression coefficients matrix is determined during calibration as:

$$\hat{\mathbf{P}} = (\mathbf{A}^T\mathbf{A})^{-1} \bullet \mathbf{A}^T\mathbf{C} \tag{3.8}$$

with the term describing the sum of squares of the deviations between predicted concentration (using $\hat{\mathbf{P}}$) and true concentration (*i.e.* equation 3.9) being minimised:

$$\sum_{i=1}^{I}\sum_{j=1}^{J}(c_{ij} - \hat{c}_{ij})^2 = \sum_{i=1}^{I}\sum_{j=1}^{J} e_{ij}^2 \tag{3.9}$$

(where e_{ij} is an element of the matrix $\mathbf{E_C}$). Analyte concentration in a new sample *i* is simply predicted as:

$$\hat{\mathbf{c}}_i^T = \mathbf{a}_i^T \hat{\mathbf{P}} \tag{3.10}$$

MLR is termed an *indirect* calibration technique since it does not require pure component spectra to build the calibration model. In addition, the model does not require concentration data for every analyte present in multicomponent calibration samples in order to perform accurate predictions of a given analyte in a new sample. In other words, prior knowledge of interferences is not required, although these interferences must be present in the calibration samples and therefore implicitly modelled [164]. As in the case of DMA, MLR is a relatively simple multivariate calibration routine, and is often the method of choice when the system under investigation is 'well behaved' with few or no overlapping signals. It is important that the instrumental response is also low in noise, since MLR will attempt to use all the data present in the $\mathbf{A}$ matrix to model concentration, including any irrelevant information. The inclusion of signal noise in the calibration model (*i.e.* overfitting of the data) may result in erroneous predictions for new samples [158].

Collinearity in the response data can pose a problem for MLR, particularly in the calibration of multiwavelength spectrophotometric data. A data set is collinear if at least one variable is an exact or approximate linear combination of the others (*i.e.* a linear or near-linear relationship exists between the data points), and this is frequently encountered within absorbance spectra. A generalised inversion of the response matrix $\mathbf{A}$ is performed in the calculation of the estimated regression coefficients matrix $\hat{\mathbf{P}}$ (as shown in equation 3.8), and inversion of a collinear matrix will produce instability in the coefficients of $\hat{\mathbf{P}}$. The resulting calibration model will have a poor predictive ability [164].

The problem of collinearity can be overcome by selecting a suitable subset of the response data to include in the model. This can be determined statistically by a number of techniques, an example of which is stepwise multiple linear regression (SMLR) [159].

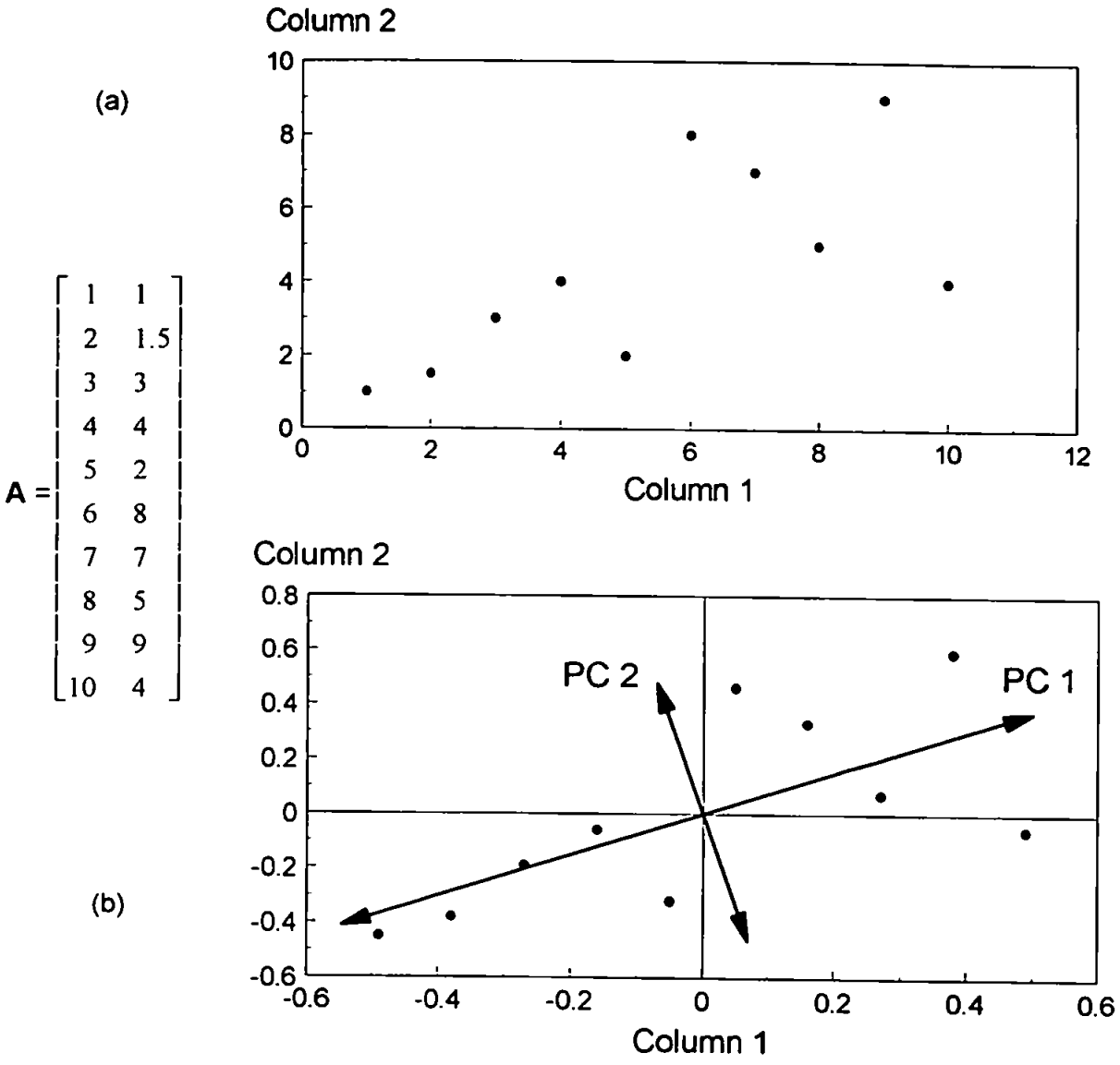
SMLR can be performed as either forward selection, which begins with one wavelength variable and progressively incorporates more wavelengths into the model until a certain criterion is met, or backwards elimination, which starts with the full spectrum and deletes one wavelength from the model at each step until the predefined criterion is achieved. The stopping criterion is typically an *F*-statistic, which tests the significance of the regression coefficients for each wavelength variable. In forward selection, the wavelength with the most significant coefficient at each step is added to the model, and this continues until no added wavelength is significant. In backwards elimination, the variable with the lowest *F*-statistic at each step is removed until the point is reached when all the remaining variables are significant. In this way, SMLR can circumvent the problem of collinear data, although the signal-averaging capabilities of full-spectrum techniques such as DMA are lost.

3.3.3 Principal components regression

Principal components regression is a method of calibration derived from factor analysis, a technique first developed in the fields of psychology and sociology to describe patterns in large data sets in terms of a much smaller number of underlying factors (*i.e.* to reduce the dimensionality of the data set) [167]. Factors are linear combinations of the original variables which describe correlations within the data set. The method of factor analysis most frequently used in chemistry is principal components analysis (PCA) [168 -169].

As with MLR, PCA assumes that concentration is a function of instrumental response, although in this case the problem of collinearity is overcome by decomposing the response matrix **A** into its most dominant factors, or 'principal components' (PCs) as they are also termed. The first PC is that which best describes the variability within the matrix, while the second and subsequent PCs successively describe the remaining variance, with the proviso that each PC is orthogonal (*i.e.* perpendicular) to the previous one. This is illustrated in Figure 3.1, which in part (a) represents the matrix **A** as ten points when plotted

Figure 3.1 (a) The matrix A plotted in column space; (b) first and second principal components (PC 1 and PC 2) for A following mean-centring and variance scaling of the columns.



in two-dimensional column space, and in part (b) indicates the first and second principal components. Prior to performing PCA, the columns of A have been mean-centred (*i.e.* the column mean subtracted from each column element) and scaled (*i.e.* each column element divided by the variance of the column to ensure comparable levels of noise between the columns). Figure 3.1 illustrates PCA in terms of only two dimensions, although it is

important to realise that this technique is equally applicable to much larger matrices, for which many PCs are required to describe the variability.

The process of PCA involves the approximation of the response matrix in terms of two smaller matrices:

$$\mathbf{A} = \mathbf{TB} + \mathbf{E_A} \tag{3.11}$$

where $\mathbf{T}$ is an $I \times H$ matrix of PC scores, $\mathbf{B}$ is an $H \times M$ matrix of PC loadings and $\mathbf{E_A}$ is an $I \times M$ matrix of spectral residuals not fitted by the PCA model. Scores are the values of the original variables when projected onto each principal component, while loadings represent the coefficients for the regression of $\mathbf{A}$ on $\mathbf{T}$ (the rows of $\mathbf{B}$ are in fact the principal components). The columns of $\mathbf{T}$ and the rows of $\mathbf{B}$ are both orthogonal (thus eliminating the problem of collinearity), with each score vector $\mathbf{t}_h$ describing the concentration patterns for the samples in $\mathbf{A}$, and each loading vector $\mathbf{b}_h^T$ describing pure-component spectra for the analytes contributing to instrumental response.

A number of algorithms can be used to perform the decomposition of the response matrix into PCs. Non-iterative partial least squares (NIPALS) is one of the more frequently applied methods, owing to its simplicity and speed of computation [159]. For each successive PC from $h = 1$ to H, $\hat{\mathbf{t}}_h$ and $\hat{\mathbf{b}}_h^T$ are calculated from $\mathbf{A}_{h-1}$, as follows:

(a) The initial score vector $\hat{\mathbf{t}}_h$ is selected as the column of $\mathbf{A}_{h-1}$ with the largest remaining variance;

(b) A new loading vector is estimated for this PC by projecting $\mathbf{A}_{h-1}$ onto $\hat{\mathbf{t}}_h$:

$$\hat{\mathbf{b}}_h^T = (\hat{\mathbf{t}}_h^T \hat{\mathbf{t}}_h)^{-1} \hat{\mathbf{t}}_h \mathbf{A}_{h-1} \tag{3.12}$$

(c) The length of $\hat{\mathbf{b}}_h^T$ is scaled to 1.0:

$$\hat{\mathbf{b}}_h^T = (\hat{\mathbf{b}}_h^T \hat{\mathbf{b}}_h)^{-0.5} \hat{\mathbf{b}}_h^T \tag{3.13}$$

(d) A new score vector is estimated by projecting $\mathbf{A}_{h-1}$ onto $\hat{\mathbf{b}}_h^T$:

$$\hat{\mathbf{t}}_h = (\hat{\mathbf{b}}_h^T \hat{\mathbf{b}}_h)^{-1} \hat{\mathbf{b}}_h^T \mathbf{A}_{h-1} \tag{3.14}$$

(e) If the difference between the newly estimated $\hat{\mathbf{t}}$ and the previously estimated $\hat{\mathbf{t}}$ is less than a pre-defined criterion, then the method has achieved convergence with respect to this PC. If not, then repeat process from step (b);

(f) Subtract the effect of this PC:

$$\mathbf{A}_h = \mathbf{A}_{h-1} - \hat{\mathbf{t}}_h \hat{\mathbf{b}}_h^T \tag{3.15}$$

(g) Repeat the process from step (a) for the next principal component.

PCA is very useful for qualitative analysis of a data set. For example, plotting scores for PC 1 against scores for PC 2 will reveal clusters of samples which have a similar analyte composition, and can therefore be used to identify similarities which are not apparent from a visual inspection of the data, and to detect outlying samples. The number of PCs explaining the total variance in the data can give an indication of the number of analytes contributing to response (although the presence of physical and chemical interferences will require additional PCs). Plots of the loading vectors can reveal which variables are contributing most significantly to each PC, and in the case of multicomponent spectral data will indicate the pure spectra of the individual analytes.

If quantitative information is required, then principal components regression (PCR) is used. PCR is conceptually similar to MLR, but the calibration model is constructed using the matrix of PC scores in place of the original response matrix:

$$\mathbf{C} = \mathbf{TQ} + \mathbf{E_C} \tag{3.16}$$

where $\mathbf{Q}$ is the $H \times J$ matrix of regression coefficients relating scores to concentrations. The regression coefficients are estimated by a least squares procedure, in which the residuals in $\mathbf{E_C}$ are minimised:

$$\hat{\mathbf{Q}} = \hat{\mathbf{B}}\,(\hat{\mathbf{T}}^{T}\hat{\mathbf{T}})^{-1}\mathbf{T}^{T}\mathbf{C} \tag{3.17}$$

Prediction of concentration in a new sample involves an initial calculation of its scores vector (equation 3.18), which is then multiplied by the regression coefficients matrix to provide an estimate of concentrations (equation 3.19):

$$\hat{\mathbf{t}}^{T} = \mathbf{a}\mathbf{B}^{T} \tag{3.18}$$

$$\hat{\mathbf{c}}^{T} = \hat{\mathbf{t}}^{T}\mathbf{Q} \tag{3.19}$$

It should be noted that the number of PCs used in prediction (*i.e.* the dimension H in $\mathbf{t}$, $\mathbf{B}$ and $\mathbf{Q}$) is selected as that which provides an optimal description of the variance in $\mathbf{A}$ without overfitting for noise. The criteria used to select the optimal number of PCs are described in Chapter 4.

PCR possesses the same advantages as MLR (*i.e.* prior knowledge of interferences and pure component spectra is not required), but is a more robust calibration technique since full-spectrum information is used to build the model. For this reason, PCR is often used in preference to MLR in the calibration of collinear spectral response data. A potential drawback to the PCR approach is the fact that the PCs which best describe the variance in the response matrix may not also be the best description of the variance in the analyte concentrations matrix (*e.g.* instrumental noise may be responsible for the largest component of the measurement variance). If this is the case, then the resulting calibration model will produce erroneous predictions of new samples [164].

3.3.4 Partial least squares regression

Partial least squares (PLS) regression is a conceptually similar technique to PCR, and is based upon a decomposition of original, full-spectrum data into dominant factors. Unlike PCR however, PLS involves simultaneous analyses of both the response and the concentration matrices to calculate scores and loading vectors for each. In this way, it is

able to determine which factors in the response matrix are most relevant to variance in the concentration matrix, thereby reducing the influence of irrelevant factors upon the calibration model [164, 170 -171].

The technique of PLS was first introduced in 1977, following development work carried out largely by Herman Wold during the 1960s and 1970s. Since that time, variations of the technique have been applied in areas of research as diverse as economics, psychology, politics and a number of the natural sciences [172 -173]. In analytical chemistry it is increasingly being applied to the calibration of multicomponent spectroscopic data, particularly UV-visible [174 -175], NIR [176 -177] and FT-IR [178].

The form of PLS employed in chemometrics is actually a modification of the NIPALS algorithm used in PCA, as described above. The modifications enable PLS to calculate loading vectors which contain the maximum amount of predictive information in the earlier vectors. This is achieved by using information from the concentration matrix when performing the decomposition of the response matrix, so that the loading vectors are concentration-dependent [164].

The algorithm used for PLS calibration with respect to a single-analyte concentration vector $\mathbf{c}_j$ can be summarised as follows. $\mathbf{A}$ and $\mathbf{c}_j$ are initially mean-centred and $\mathbf{A}$ is scaled. Then, for the calculation of each PLS factor in the range $h = 1$ to H (where H is greater than the number of expected components contributing to $\mathbf{A}$):

(a) The loading weight vector $\hat{\mathbf{w}}_h$ is calculated by regressing $\mathbf{A}_{h-1}$ onto $\mathbf{c}_j$:

$$\mathbf{A}_{h-1} = \mathbf{c}_j \hat{\mathbf{w}}_h^{\mathsf{T}} + \mathbf{E}_{\mathbf{A}} \tag{3.20}$$

using a least squares estimation of $\hat{\mathbf{w}}_h$ to minimise the residuals in $\mathbf{E}_{\mathbf{A}}$:

$$\hat{\mathbf{w}}_h = \mathbf{A}_{h-1}^{\mathsf{T}} \mathbf{c}_j (\mathbf{c}_j^{\mathsf{T}} \mathbf{c}_j)^{-1} \tag{3.21}$$

(b) Following normalisation of $\hat{\mathbf{w}}_h$, a new scores vector $\hat{\mathbf{t}}_h$ is estimated by regression of $\mathbf{A}_{h-1}$ onto $\hat{\mathbf{w}}_h$:

$$\mathbf{A}_{h-1} = \hat{\mathbf{t}}_h \hat{\mathbf{w}}_h^T + \mathbf{E}_A \tag{3.22}$$

using a least squares estimation of $\hat{\mathbf{t}}_h$:

$$\hat{\mathbf{t}}_h = \mathbf{A}_{h-1}\hat{\mathbf{w}}_h(\hat{\mathbf{w}}_h^T\hat{\mathbf{w}}_h)^{-1} = \mathbf{A}_{h-1}\hat{\mathbf{w}}_h \tag{3.23}$$

(c) The chemical loading $\hat{v}_h$ (*i.e.* the regression coefficient relating the new scores vector to analyte concentrations) is estimated by regression of $\mathbf{c}_j$ onto $\hat{\mathbf{t}}_h$:

$$\mathbf{c}_j = \hat{v}_h \hat{\mathbf{t}}_h + \mathbf{e}_c \tag{3.24}$$

using a least squares estimate of $\hat{v}_h$:

$$\hat{v}_h = \hat{\mathbf{t}}_h^T \mathbf{c}_j (\hat{\mathbf{t}}_h^T \hat{\mathbf{t}}_h)^{-1} \tag{3.25}$$

(d) The loadings vector $\hat{\mathbf{b}}_h$ for the response matrix is estimated by regression of $\mathbf{A}_{h-1}$ onto $\hat{\mathbf{t}}_h$:

$$\mathbf{A}_{h-1} = \hat{\mathbf{t}}_h \hat{\mathbf{b}}_h^T + \mathbf{E}_A \tag{3.26}$$

using a least squares estimate of $\hat{\mathbf{b}}_h$:

$$\hat{\mathbf{b}}_h = \mathbf{A}_{h-1}\hat{\mathbf{t}}_h(\hat{\mathbf{t}}_h^T\hat{\mathbf{t}}_h)^{-1} \tag{3.27}$$

(e) New response and concentration residuals are calculated by subtracting the estimated effect of this factor:

$$\mathbf{E}_A = \mathbf{A}_{h-1} - \hat{\mathbf{t}}_h \hat{\mathbf{b}}_h^T \tag{3.28}$$

$$\mathbf{e}_c = \mathbf{c}_j - \hat{v}_h \hat{\mathbf{t}}_h \tag{3.29}$$

(f) The new values for $\mathbf{E}_A$ and $\mathbf{e}_c$ are substituted for $\mathbf{A}_{h-1}$ and $\mathbf{c}_j$ prior to the calculation of the next factor.

Once the optimal number of factors for retention in the model has been determined (see Chapter 4), then prediction of new samples can be performed in the following way.

The new sample response vector $\mathbf{a}_i$ is initially mean-centred and scaled. Then, for the calculation of each PLS factor in the range $h = 1$ to H (where H is the optimal number of factors determined during calibration):

(i) The score $\hat{t}_{i,h}$ is estimated for the new sample using the same procedure as step (b) of calibration:

$$\hat{t}_{i,h} = \mathbf{a}_i^T \hat{\mathbf{w}}_h \tag{3.30}$$

(ii) A new residual vector $\mathbf{e}_h$ is calculated:

$$\mathbf{e}_{i,h} = \mathbf{e}_{i,h-1} - \hat{\mathbf{b}}_h t_{i,h} \tag{3.31}$$

then h is incremented by 1, $\mathbf{a}_i$ is reassigned as $\mathbf{e}_{i,h}$ and the algorithm is repeated from step (i) until $h = H$.

(iii) Analyte concentration for sample i is then predicted as:

$$\hat{c}_i = \bar{c} + \sum_{h=1}^{H} \hat{t}_{i,h} \hat{v}_h \tag{3.32}$$

where $\bar{c}$ is the mean analyte concentration in the calibration samples.

Two forms of the PLS algorithm are commonly applied in chemometrics [179]. The algorithm shown above is PLS1, which performs calibration and prediction with respect to one analyte only (*i.e.* a separate calibration model is required for each analyte in the sample set). An alternative method is PLS2, in which modifications to the algorithm used for PLS1 permit two or more analytes to be modelled simultaneously. In practice, PLS2 can represent the most convenient and rapid method of calibration and prediction in cases where the sample matrix is complex and the calibration set is large. However, PLS1 tends to provide more accurate predictions for multicomponent samples, since PLS2 is restricted to a single optimal number of factors to represent all the components, and in many cases the optimal number is found to be different for each individual component.

3.3.5 Summary of multivariate calibration techniques

The four linear multivariate calibration techniques described in this chapter have been shown to have both shared and unique characteristics. These are summarised in Table 3.2.

Table 3.2 The principal characteristics of DMA, MLR, PCR and PLS.

Technique	Description
Direct multicomponent analysis	Direct calibration (requires prior knowledge of pure spectra or concentration data for all components present); Beer's law model (response a linear function of concentration); Full-spectrum modelling; Simultaneous calibration of all components.
Multiple linear regression	Indirect calibration (no prior knowledge of pure component spectra required); Inverse Beer's law model (concentration a linear function of response); Collinearity prevents full-spectrum modelling, therefore a subset of response variables must be selected (*e.g.* by stepwise regression); Simultaneous calibration of all components.
Principal components regression	Indirect calibration and inverse Beer's law relationship; Collinearity problem overcome by decomposition of response matrix into dominant factors (principal components). Principal component scores are then regressed by an MLR procedure; Full-spectrum modelling; Simultaneous calibration of all components.
Partial least squares regression	Indirect calibration and inverse Beer's law relationship; Similar data decomposition and regression procedures to PCR, but involves simultaneous determination of dominant factors in both response and concentration matrices, to determine which response factors are most relevant to concentration variance; Full-spectrum modelling; Single-component (PLS1) or simultaneous multicomponent (PLS2) calibration.

Chapter 4

Comparison of Multivariate Calibration Techniques for the Quantification of Metal Ions in Model Effluent Streams

4.1 INTRODUCTION

As discussed in the previous chapter, multivariate calibration techniques are gaining importance as a means of deriving greater amounts of information from complex analytical data. This is particularly relevant to industrial process and effluent monitoring, where the combination of multiwavelength spectrophotometers and multivariate calibration can offer a method of simultaneously quantifying a number of analytes in complex sample matrices [20, 180]. Since multivariate calibration routines can implicitly model the effect of potential interferents, enhanced analytical sensitivity and selectivity can be achieved without the need for more time-consuming and expensive physico-chemical extraction procedures. Multivariate calibration enables the resolution of multicomponent spectral data in terms of the individual components, and in this way can improve the selectivity of instrumental techniques such as UV-visible spectrophotometry [181].

In Chapter 2, an on-line method for determining a single analyte (*i.e.* ammonia) in effluent streams was described. The aim of this chapter is to expand on work previously reported by MacLaurin *et al.* [182], in order to investigate the application of diode-array spectrophotometry in combination with multivariate calibration to simultaneously quantify multiple analytes in model effluent systems, and to compare the predictive abilities of the five linear multivariate calibration methods described in the previous chapter (*i.e.* DMA, SMLR, PCR, PLS1 and PLS2). Spectral data were analysed for seven multicomponent sample systems, which contained mixtures of up to five transition metal salts, and incorporated physical and chemical interferences to simulate the inter-analyte interactions and suspended solids often encountered in real effluent matrices.

4.2 EXPERIMENTAL DESIGN

In general, a multivariate calibration model is constructed from instrumental response data collected for a set of multicomponent samples of known concentrations with respect to the analytes of interest. With experimental time and cost in mind, the calibration sample set should comprise the smallest number of samples necessary to encompass the important variabilities expected in any new sample for prediction [162]. This can be achieved by constructing the calibration set according to systematic experimental design principles. In this work, sample sets for both calibration and prediction were designed according to factorial or fractional factorial design principles [183 , 184 , 185 , 186].

Factorial designs are used for systems in which two or more analytes (or *factors*) are to be investigated, and involve the construction of a sample set in which all possible combinations of the analytes at two or more concentration levels are present. For example, if three components are to be studied at two concentration levels, then a total of $2^3 = 8$ samples will be required to provide the full range of component and concentration combinations. Factorial designs are generally the most simple and efficient form of experimental design for multi-analyte systems, and encompass the effects on instrumental response of the individual analytes (*i.e.* the *main effects*) and any inter-analyte interactions. Their main limitation is the large number of samples required if more than two concentration levels are examined (*e.g.* four components at three levels would require $3^4 = 81$ samples to span all possible combinations). Two-level (2^k) factorial designs are often used when a large number of components are to be investigated, but assume that response is approximately linear over the range of the chosen concentration levels. A minimum of three levels (*i.e.* 3^k designs) is required if potential non-linear effects are to be examined in the sample set [186].

In cases where a larger number of components and/or levels are involved, fractional factorial designs allow a reduced number of calibration samples to span the main effects and low-order (*i.e.* two- or three-component) interactions. Fractional factorial designs assume that high-order interactions have a negligible effect on response, and therefore do not require every possible combination of analyte levels in the sample set (*i.e.* only a fraction of the complete factorial experiment is used). The fractional design is generated by *confounding* the highest-order interactions with the main effects (*i.e.* the effects on response of the highest-order interactions become indistinguishable from that of the main effects). An estimate of a given main effect is therefore actually a linear combination of the main effect and a high-order interaction (*i.e.* the main effect and the interaction are said to be *aliases* of each other), but as stated before, the contribution of the high-order interaction is assumed to be negligible. For example, if the five analytes A, B, C, D and E are to be examined at two levels, then a fraction of the full $2^5 = 32$ design can be generated by confounding the effect of A with the BCDE four-factor interaction. The design is represented as 2^{5-1}, and in this case is termed a *half-fraction* factorial design since only 16 samples are required. The relationship A = BCDE is termed the *generator* of the design, and if these are multiplied, the result I is called the defining relation:

$$I = ABCDE \tag{4.1}$$

The aliases within the design are obtained by multiplying the defining relation with each of the effects and interactions, according to standard algebraic rules. An additional condition is applied whereby for an *n*-level design, the terms A^n, B^n, etc., are cancelled in the product [187]. For example, the aliases of B and AB are determined for a two-level design as follows:

$$B = B \times ABCDE = AB^2CDE = ACDE \tag{4.2}$$

$$AB = AB \times ABCDE = A^2B^2CDE = CDE \tag{4.3}$$

Two generators can be used to reduce the number of experimental runs still further. A 2^{6-2} *quarter fraction* design can be produced for example in the case of a six-component system by using the generators A = BCDE and F = BCD. The defining relations in this case are I = ABCDE, I = BCDF and I = AEF (the latter being the product of the first two defining relations). However, if more than one generator is used to produce a fractional design, there is a risk of main effects being aliased with potentially significant two-component interactions. If this is the case, the problem can be circumvented by the use of alternative generators. An examination of the aliases produced in each case will determine whether a particular fractional design is suitable for the system under investigation [185, 186].

4.3 EXPERIMENTAL

4.3.1 Reagents

All solutions were prepared using Milli-Q water (Millipore) and all reagents were of AnalaR grade (Merck) unless otherwise indicated. Stock solutions (0.1 mol l^{-1}) of chromium(III) potassium sulfate dodecahydrate, iron(II) sulfate heptahydrate, cobalt(II) sulfate heptahydrate, nickel(II) sulfate heptahydrate and copper(II) sulfate pentahydrate were prepared in 1% v/v sulfuric acid. Calibration and test set solutions were prepared by serial dilution of these stock solutions. Barium chloride dihydrate (Fisons AR) was added, where indicated below.

4.3.2 Instrumentation

Absorbance and first-derivative spectra were obtained using a Hewlett-Packard 8451A photodiode array spectrophotometer fitted with a 1 cm pathlength silica cuvette. Raw data were initially stored using an HP 9121 disk drive unit, then transferred in ASCII format via

an HP 82939A serial interface to a personal computer. Multivariate analysis of the data was performed on a 486DX personal computer with 8 Mbyte of RAM.

4.3.3 Software

Two programs written in HP BASIC were used, firstly to measure and record absorbance and first-derivative spectra, and secondly to download this data to the personal computer. Kermit version 3.01 serial interface software was used to receive and store this data in ASCII format on the personal computer. DMA was performed using "maximum likelihood" weighted least squares software supplied on board the diode-array spectrophotometer. PCR, PLS1 and PLS2 were conducted using Unscrambler-II version 4.00 multivariate analysis software (Camo A/S). This package incorporates matrix handling routines, enabling manipulation of the ASCII data files. SMLR was carried out in two stages within Minitab version 8.2 statistical software (Minitab Inc.), using an initial stepwise regression procedure followed by an MLR calibration of the selected wavelength data.

4.3.4 Design of Sample Sets

Calibration and test sets were designed for two-, three-, four- and five-component systems. Solutions containing transition metal salt concentrations in the range 0.005 - 0.025 mol l^{-1} were produced by serial dilution of the stock solutions.

For the two-component system (Cu^{2+} and Fe^{2+}), a five-level factorial design (5^2) was used to produce a full set of 25 samples. A three-level set (3^2) was derived from this to produce a calibration set of nine samples, with the remaining 16 samples used for an independent test set, as shown in Table 4.1.

The three-component system (Co^{2+}, Cu^{2+} and Ni^{2+}) employed a three-level factorial design (3^3) to produce a full set of 27 solutions. A two-level design (2^3) was used to derive

a calibration set of eight samples from the full set, with the remaining 19 samples used as a test set, as shown in Table 4.2.

Table 4.1 Factorial design of the calibration and test sets for the two-component system. ● represents calibration set sample; ○ represents test set sample.

Iron concentration (mol l^{-1})	0.020	●	○	●	○	●
	0.015	○	○	○	○	○
	0.010	●	○	●	○	●
	0.005	○	○	○	○	○
	0	●	○	●	○	●
		0	0.005	0.010	0.015	0.020

Copper concentration (mol l^{-1})

Table 4.2 Factorial design of the calibration and test sets for the three-component system. ● represents calibration set sample; ○ represents test set sample.

Cobalt concentration (mol l^{-1})						Nickel concentration (mol l^{-1})
		●	○	●	0.025	
	0.025	○	○	○	0.015	
		●	○	●	0.005	
		○	○	○	0.025	
	0.015	○	○	○	0.015	
		○	○	○	0.005	
		●	○	●	0.025	
	0.005	○	○	○	0.015	
		●	○	●	0.005	
		0.005	0.015	0.025		

Copper concentration (mol l^{-1})

In the case of the four-component system (Cr^{3+}, Ni^{2+}, Co^{2+} and Cu^{2+}), two-level and three-level fractional factorial designs (2^{4-1} and 3^{4-1}) were used to produce calibration and test sets of 8 and 27 samples respectively. The 2^{4-1} half-fraction and the 3^{4-1} third-fraction designs (see Table 4.3) were produced by confounding the copper main effect with the chromium-nickel-cobalt three-component interaction, ensuring that no main effects were confounded with each other or with two-component interactions. The following fractional design generators were used:

Calibration set: $D = ABC$ (4.4)

Test set: $D = W(ABC)$ (4.5)

where A, B, C and D represent Cr, Ni, Co and Cu respectively, and W(ABC) is an element of the ABC interaction accounting for an independent pair of degrees of freedom, as described by Yates [188].

For the five-component system (Cr^{3+}, Ni^{2+}, Co^{2+}, Cu^{2+} and Fe^{2+}), calibration and test sets, each comprising 27 samples, were formed from two different three-level, ninth-fraction factorial designs (3^{5-2}), as shown in Table 4.4. The designs were generated by confounding the cobalt and the nickel main effects with iron-copper-chromium three-component interactions. The only significant two-component interaction was assumed to be that between copper and iron, since Fe^{2+} is partially oxidised to Fe^{3+} in the presence of Cu^{2+} ions. With this in mind, the chosen fractional designs were those in which no main effects were aliased with each other or with copper-iron two-component interactions. The fractional design generators used in this case were as follows:

Calibration set: $D = X(ABC)$ $E = Y(ABC)$ (4.6)

Test set: $D = W(ABC)$ $E = X(ABC)$ (4.7)

where A, B, C, D and E represent Fe, Cu, Cr, Co and Ni respectively, and W(ABC), X(ABC) and Y(ABC) are three elements of the ABC interaction, accounting for three independent pairs of degrees of freedom [183, 188].

Table 4.3 Fractional factorial design of the calibration set (samples 1-8) and test set (samples 9-35) for the four-component system (concentration in mol l^{-1}).

Sample	Chromium	Nickel	Cobalt	Copper
1	0.005	0.005	0.005	0.005
2	0.025	0.005	0.005	0.025
3	0.005	0.025	0.005	0.025
4	0.025	0.025	0.005	0.005
5	0.005	0.005	0.025	0.025
6	0.025	0.005	0.025	0.005
7	0.005	0.025	0.025	0.005
8	0.025	0.025	0.025	0.025
9	0.005	0.005	0.005	0.005
10	0.015	0.005	0.015	0.005
11	0.025	0.005	0.025	0.005
12	0.005	0.015	0.025	0.005
13	0.015	0.015	0.005	0.005
14	0.025	0.015	0.015	0.005
15	0.005	0.025	0.015	0.005
16	0.015	0.025	0.025	0.005
17	0.025	0.025	0.005	0.005
18	0.005	0.005	0.025	0.015
19	0.015	0.005	0.005	0.015
20	0.025	0.005	0.015	0.015
21	0.005	0.015	0.015	0.015
22	0.015	0.015	0.025	0.015
23	0.025	0.015	0.005	0.015
24	0.005	0.025	0.005	0.015
25	0.015	0.025	0.015	0.015
26	0.025	0.025	0.025	0.015
27	0.005	0.005	0.015	0.025
28	0.015	0.005	0.025	0.025
29	0.025	0.005	0.005	0.025
30	0.005	0.015	0.005	0.025
31	0.015	0.015	0.015	0.025
32	0.025	0.015	0.025	0.025
33	0.005	0.025	0.025	0.025
34	0.015	0.025	0.005	0.025
35	0.025	0.025	0.015	0.025

Table 4.4 Fractional factorial design of the calibration set (samples 1-27) and test set (samples 28-54) for the five-component system (conc. in mol l^{-1}).

Sample	Iron	Copper	Chromium	Cobalt	Nickel
1	0.005	0.005	0.005	0.005	0.005
2	0.005	0.005	0.015	0.015	0.025
3	0.005	0.005	0.025	0.025	0.015
4	0.015	0.005	0.025	0.005	0.025
5	0.015	0.005	0.005	0.015	0.015
6	0.015	0.005	0.015	0.025	0.005
7	0.025	0.005	0.015	0.005	0.015
8	0.025	0.005	0.025	0.015	0.005
9	0.025	0.005	0.005	0.025	0.025
10	0.005	0.015	0.015	0.005	0.005
11	0.005	0.015	0.025	0.015	0.025
12	0.005	0.015	0.005	0.025	0.015
13	0.015	0.015	0.005	0.005	0.025
14	0.015	0.015	0.015	0.015	0.015
15	0.015	0.015	0.025	0.025	0.005
16	0.025	0.015	0.025	0.005	0.015
17	0.025	0.015	0.005	0.015	0.005
18	0.025	0.015	0.015	0.025	0.025
19	0.005	0.025	0.025	0.005	0.005
20	0.005	0.025	0.005	0.015	0.025
21	0.005	0.025	0.015	0.025	0.015
22	0.015	0.025	0.015	0.005	0.025
23	0.015	0.025	0.025	0.015	0.015
24	0.015	0.025	0.005	0.025	0.005
25	0.025	0.025	0.005	0.005	0.015
26	0.025	0.025	0.015	0.015	0.005
27	0.025	0.025	0.025	0.025	0.025

Table 4.4 (continued).

Sample	Iron	Copper	Chromium	Cobalt	Nickel
28	0.005	0.005	0.005	0.005	0.005
29	0.015	0.005	0.015	0.005	0.025
30	0.025	0.005	0.025	0.005	0.015
31	0.005	0.015	0.025	0.005	0.015
32	0.015	0.015	0.005	0.005	0.005
33	0.025	0.015	0.015	0.005	0.025
34	0.005	0.025	0.015	0.005	0.025
35	0.015	0.025	0.025	0.005	0.015
36	0.025	0.025	0.005	0.005	0.005
37	0.005	0.005	0.025	0.015	0.025
38	0.015	0.005	0.005	0.015	0.015
39	0.025	0.005	0.015	0.015	0.005
40	0.005	0.015	0.015	0.015	0.005
41	0.015	0.015	0.025	0.015	0.025
42	0.025	0.015	0.005	0.015	0.015
43	0.005	0.025	0.005	0.015	0.015
44	0.015	0.025	0.015	0.015	0.005
45	0.025	0.025	0.025	0.015	0.025
46	0.005	0.005	0.015	0.025	0.015
47	0.015	0.005	0.025	0.025	0.005
48	0.025	0.005	0.005	0.025	0.025
49	0.005	0.015	0.005	0.025	0.025
50	0.015	0.015	0.015	0.025	0.015
51	0.025	0.015	0.025	0.025	0.005
52	0.005	0.025	0.025	0.025	0.005
53	0.015	0.025	0.005	0.025	0.025
54	0.025	0.025	0.015	0.025	0.015

4.3.5 Procedures

The order in which the samples of each multicomponent system were measured was randomised to reduce the risk of obtaining biased results. Absorbance spectra were measured over the wavelength range 302 - 800 nm, with measurements taken at 2 nm intervals to produce 250 data points per spectrum. Each solution was measured in triplicate against a 1% v/v sulfuric acid blank, using an integration time of 25 s. The triplicate sets were averaged to produce mean spectra, which were stored for use in calibration or prediction.

Having obtained and stored the spectra, small, varying quantities of barium chloride were added in a non-quantitative fashion to all calibration and test samples in the three- and five-component systems. These solutions were then measured as before, with the new spectra stored for subsequent data analysis. The addition of barium chloride resulted in precipitation of barium sulfate, which caused varying degrees of absorbance and scattering of incident light, thereby simulating the effect of suspended solids in the sample matrix.

Pure spectra of each metal salt were obtained for 0.015 mol l^{-1} solutions, and used as calibration standards in DMA. The predictive ability of this technique was determined for each multicomponent system by calibrating with the pure spectra of the metal ions present, then using the model to predict the concentrations of metal salts in every sample in the full set (*i.e.* calibration and test samples). The spectrum of barium sulfate precipitate in 1% v/v sulfuric acid was also used as a DMA calibration standard for samples into which barium chloride had been added. The predictive abilities of SMLR, PCR, PLS1 and PLS2 were determined by using the sample data of each calibration set to construct models, which were used to predict the metal concentrations of the respective test set solutions. Mean-centring was applied to all variables used for PCR, PLS1 and PLS2 calibrations.

The precision of each multivariate calibration technique is expressed here in terms of the relative root-mean-square error of prediction (RRMSEP), as shown in equation 4.8:

$$\text{RRMSEP (\%)} = \frac{100}{\bar{y}_j}\sqrt{\sum_{i=1}^{I}\frac{(y_{ij} - \hat{y}_{ij})^2}{I}} \tag{4.8}$$

where $\bar{y}_j$ = the true concentration mean of component j in the test set, I = the number of samples in the set, y_{ij} = the true concentration of component j in sample i, and $\hat{y}_{ij}$ = the predicted concentration of component j in sample i. An averaged, overall precision for each technique was obtained using equation 4.9:

$$\text{RRMSEP (\%)} = \frac{100}{\bar{y}}\sqrt{\sum_{j=1}^{J}\sum_{i=1}^{I}\frac{(y_{ij} - \hat{y}_{ij})^2}{I \times J}} \tag{4.9}$$

where $\bar{y}$ = the true concentration mean of all components in the test set, and J = the number of components in the test set. In each case lower RRMSEP values indicate better precision.

4.3.6 Unscrambler model validation

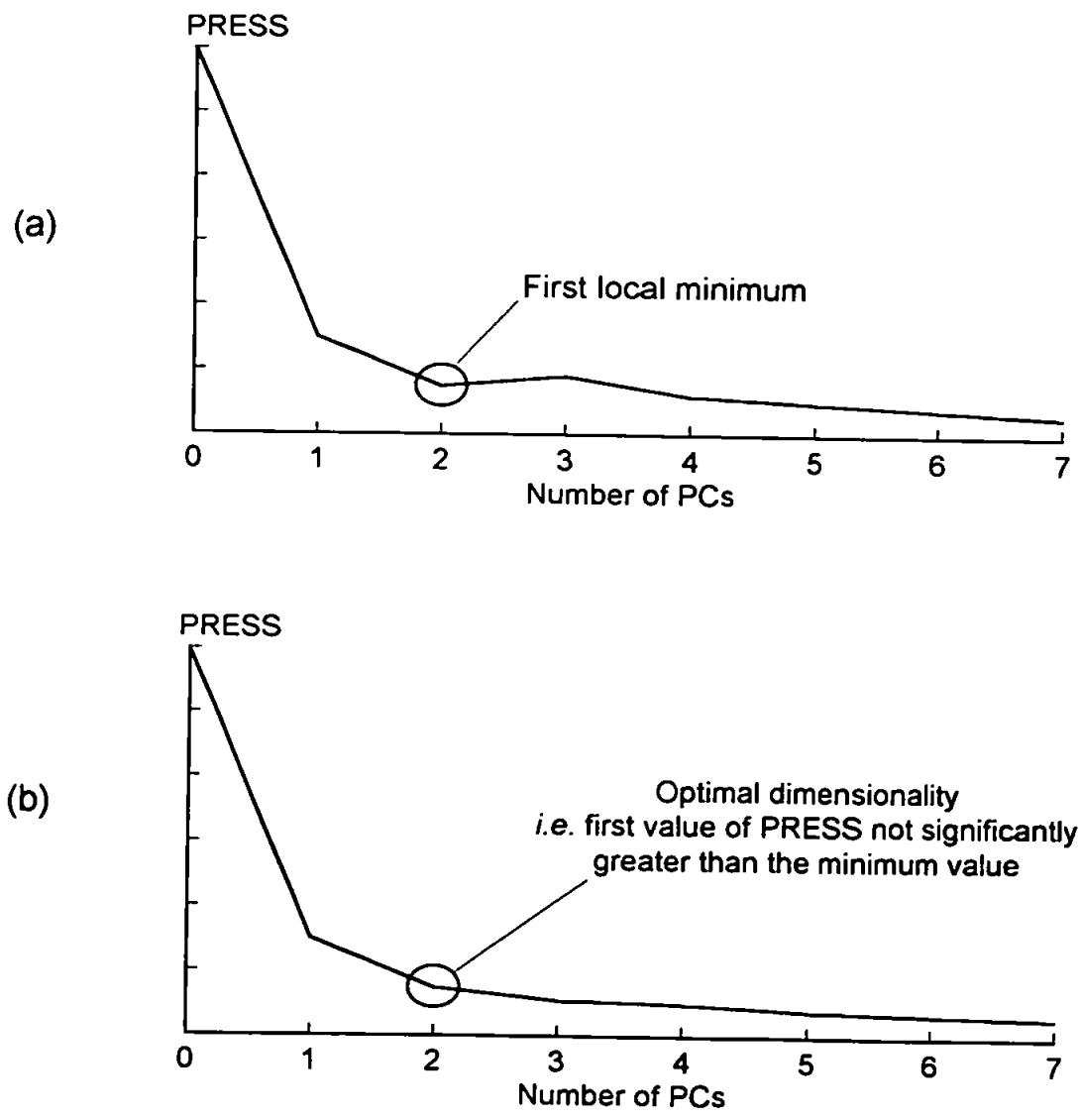
As discussed in the previous chapter, the optimal number of principal components (PCs) used for prediction with PCR, PLS1 and PLS2 models must be carefully selected in order to fully describe the variance in the response data matrix without overfitting for noise. For this purpose, the Unscrambler software compares the predictive ability of a given model at various PC dimensionalities during the modelling process. This process is termed full internal cross-validation, and successively predicts each sample in the calibration set (I samples) using a subset model constructed from the remaining I - 1 samples. This process is performed at each PC dimension, and the predictive ability is calculated in terms of the prediction error sum of squares (PRESS) for each PC, as shown in equation 4.10:

$$\text{PRESS} = \sum_{i=1}^{I}(y_i - \hat{y}_i)^2 \tag{4.10}$$

Unscrambler automatically selects the optimal model dimensionality for prediction of new samples as that which produces the first local minimum value for PRESS, as shown in

Figure 4.1(a). This is generally an acceptable compromise between under- and overfitting, although in some cases a local minimum may not be achieved for the total number of PCs included in the model. To avoid overfitting of data, the optimal dimensionality in this work was defined either as that corresponding with the first local minimum in the value of PRESS, or as the fewest number of PCs yielding a value of PRESS not significantly greater, by using an *F*-statistic comparison ($\alpha = 0.05$), than the minimum PRESS [164]. The latter criterion is represented in Figure 4.1 (b).

Figure 4.1 Optimal model dimensionality (PCR, PLS1 and PLS2) defined as (a) the first local minimum in the value of PRESS, and (b) according to an *F*-statistic comparison of significant differences.

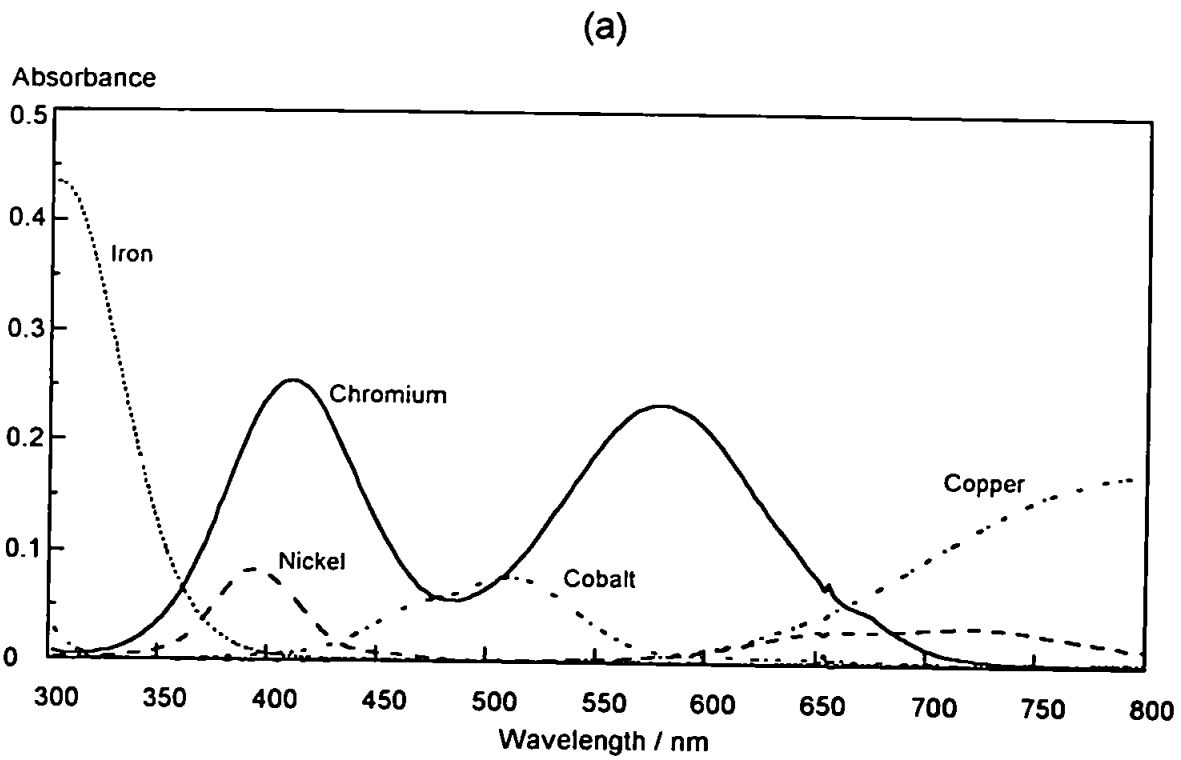


4.4 RESULTS AND DISCUSSION

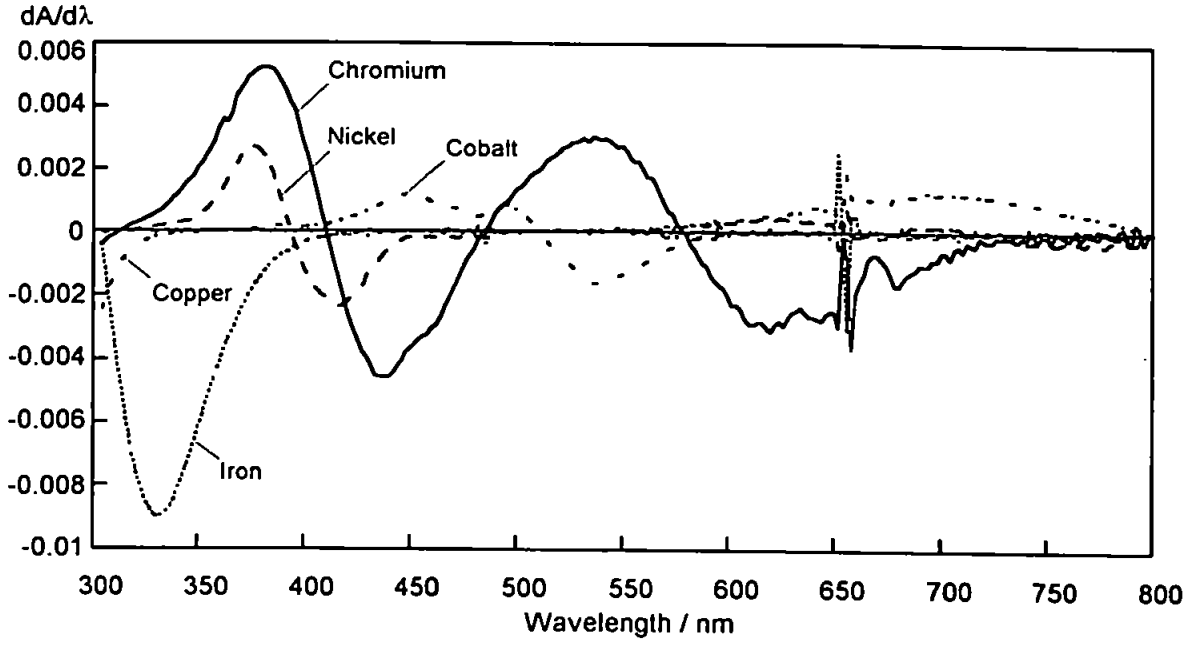
4.4.1 Spectral Characteristics

Absorbance and first-derivative spectra for pure solutions of the five metal salt solutions are shown in Figures 4.2(a) and 4.2(b) respectively. Between 302 and 800 nm it can be seen that very little overlapping occurs for the spectra of copper and iron in the two-component system, and cobalt, copper and nickel in the three-component system. This is in contrast to the four- and five-component systems, in which a high degree of overlapping exists between the chromium, cobalt and nickel absorbance spectra, and the chromium and nickel first-derivative spectra.

Figure 4.2 Absorbance spectral profiles for the five metal ion solutions: (a) absorbance data; (b) first-derivative data.



(b)

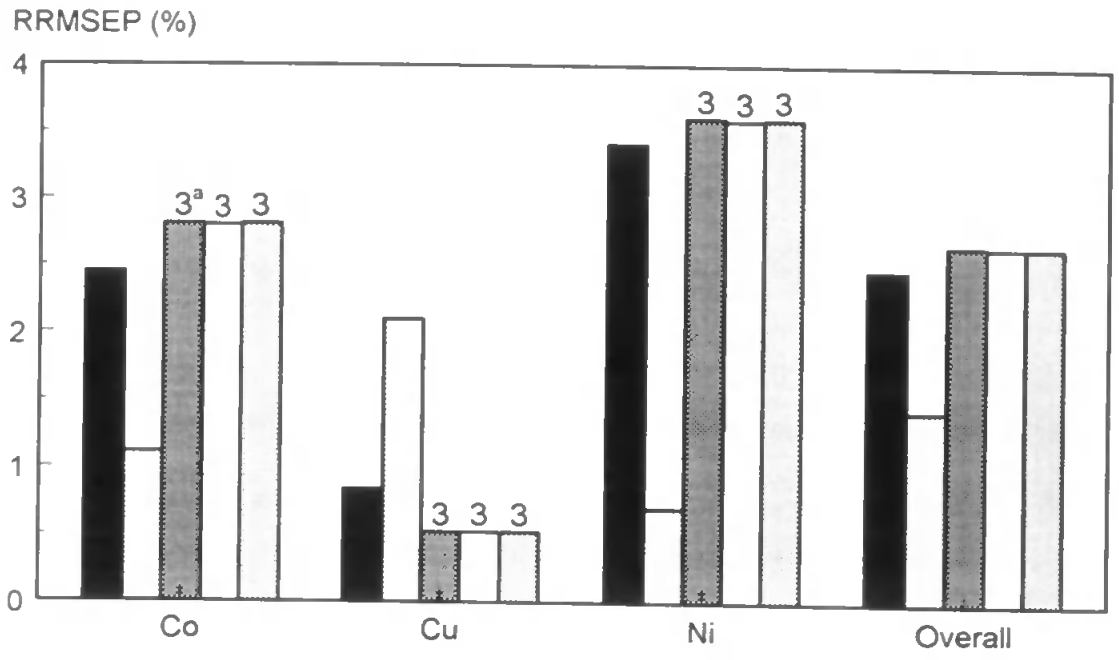


4.4.2 Three-component System

Relative root-mean-square errors of prediction for the three-component system are represented as a bar chart in Figure 4.3. This was a "well behaved" system, involving no interferences or notable chemical interactions between the components, and as a consequence it can be seen that no significant differences were found in the overall precisions of DMA, PCR, PLS1 and PLS2. A significant improvement in the overall precision was evident however in the case of SMLR, which used four wavelengths to calibrate for copper and cobalt, and three for nickel. The wavelengths selected in the stepwise regression procedure corresponded with the maxima and minima of the most intense spectral peaks (*e.g.* 362, 386 and 458 nm for nickel). Although SMLR lacks the full-spectrum capabilities of DMA, PCR and the two PLS algorithms, this may be an advantage in a simple system where the spectra of the components are well-defined and have no significant overlap, since only the most salient information is modelled. In addition, there was no difference in the concentration ranges used for the calibration and test set

solutions, and this may have minimised the potential advantage of incorporating full spectral data in the models. As mentioned in Chapter 2, the required level of precision for on-line monitoring is typically in the range 5-10% relative standard deviation [17, 147], and all five methods of calibration produced overall precisions of < 3% (and precisions of < 4% for individual metal ions) in this simple, well behaved system.

Figure 4.3 Precisions for the three-component system. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; □, PLS2.



[a] Number of principal components used for predictions.

The effect of physical interference (simulating suspended solids in a real process system) on the spectral response of the three-component system is illustrated by Figures 4.4 (a) and 4.4 (b), which respectively show the calibration set spectra before and after the addition of varying quantities of barium chloride. The results given in Figure 4.5 show that barium sulfate interference resulted in a deterioration in the overall precisions of all techniques. SMLR again displayed a significantly better precision than the full-spectrum techniques for the reasons given above, and was the only technique capable of an overall

precision of < 5%. However the overall precisions of DMA, PCR, PLS1 and PLS2 were all < 10%, and are therefore within the range required for on-line monitoring.

Figure 4.4 (a) Absorbance spectra of the three-component calibration samples (labelled A to H); (b) spectra for the same solutions after the addition of varying amounts of $BaCl_2$, indicating the effect of $BaSO_4$ interference.

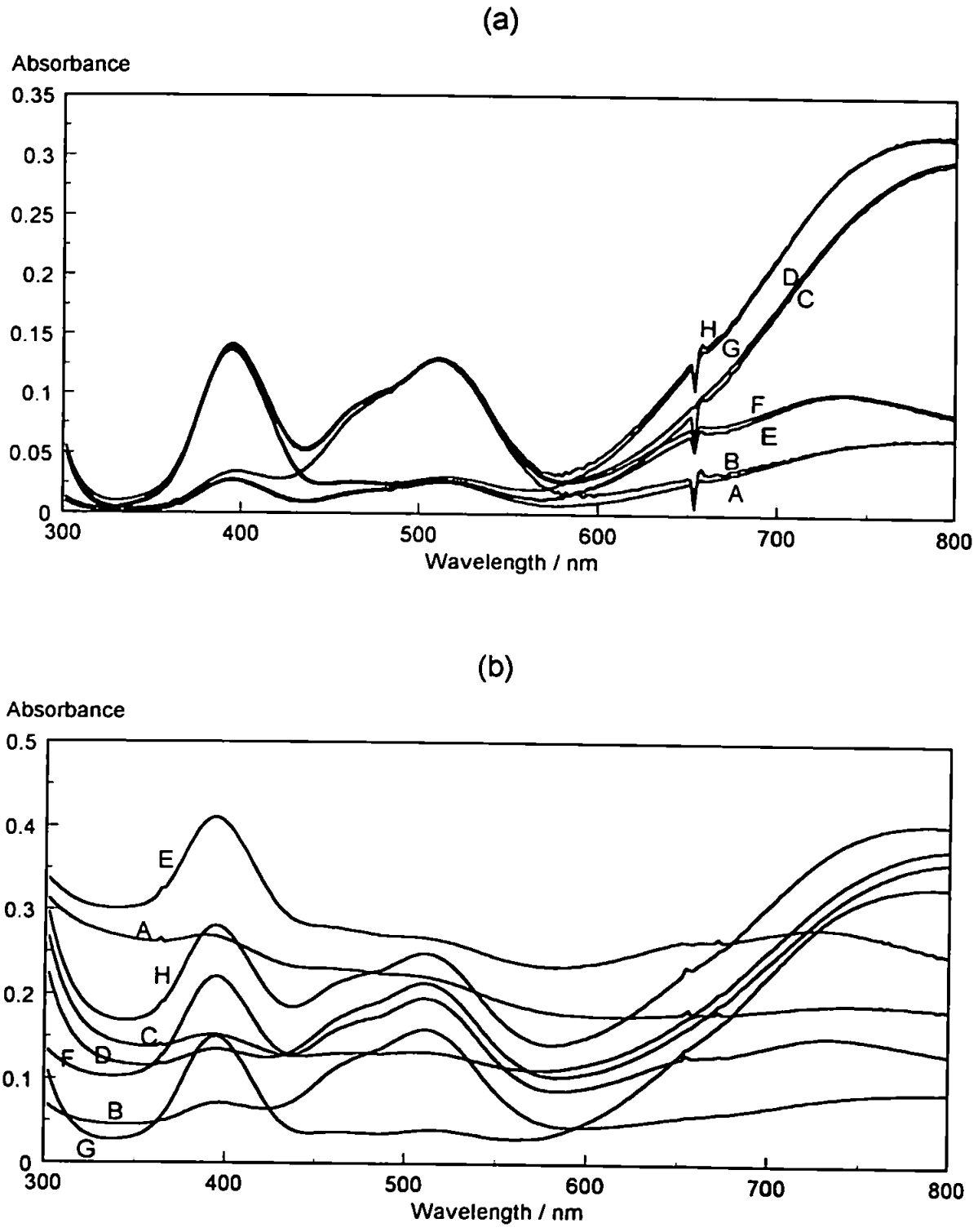
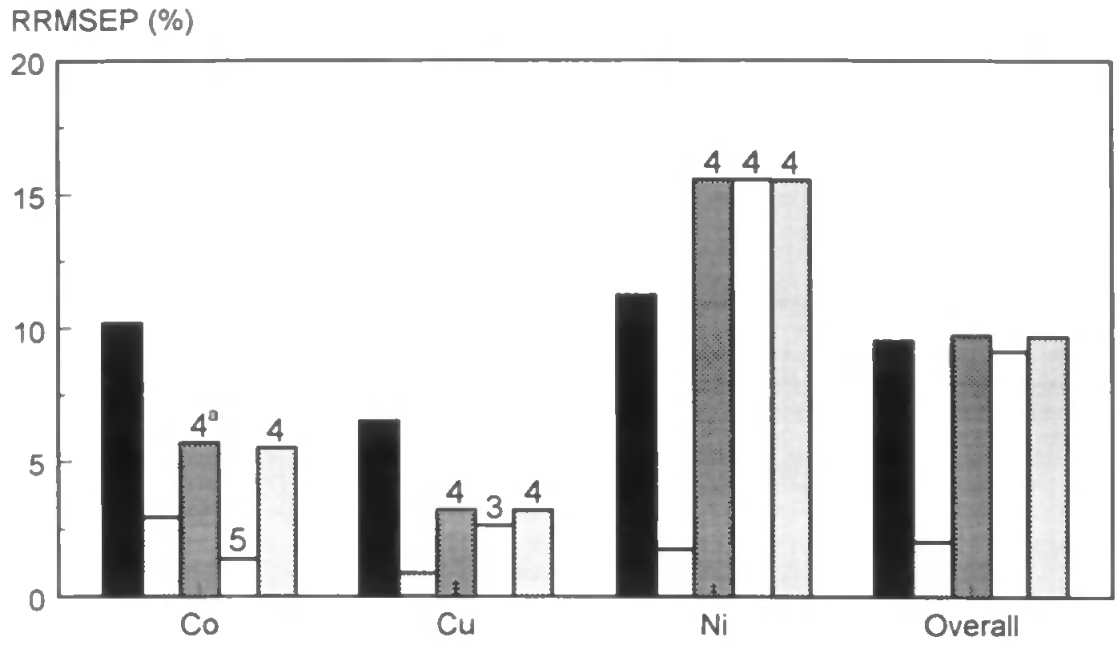


Figure 4.5 Precisions for the three-component system with barium sulfate interference. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; □, PLS2.



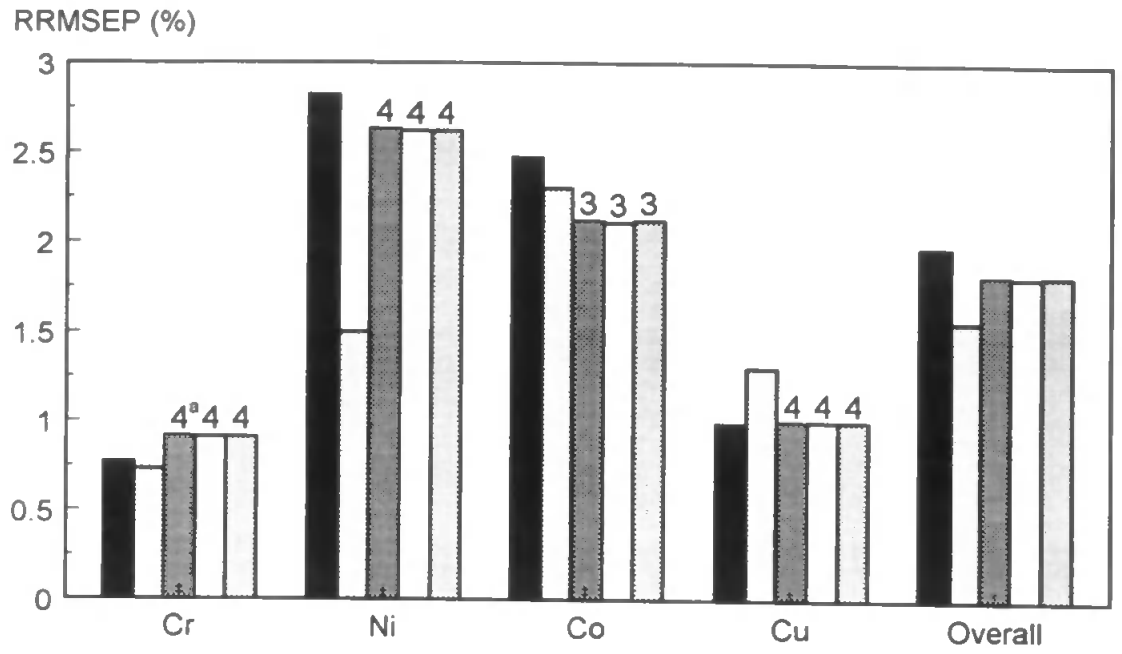
[a] Number of principal components used for predictions.

4.4.3 Four-component system

Figure 4.6 summarises the prediction errors for the well behaved four-component system. As in the case of the three-component system, very low RRMSEP values (< 3%) were recorded for the five calibration techniques with respect to all metal ions. No significant difference can be seen in the overall precisions when using absorbance data.

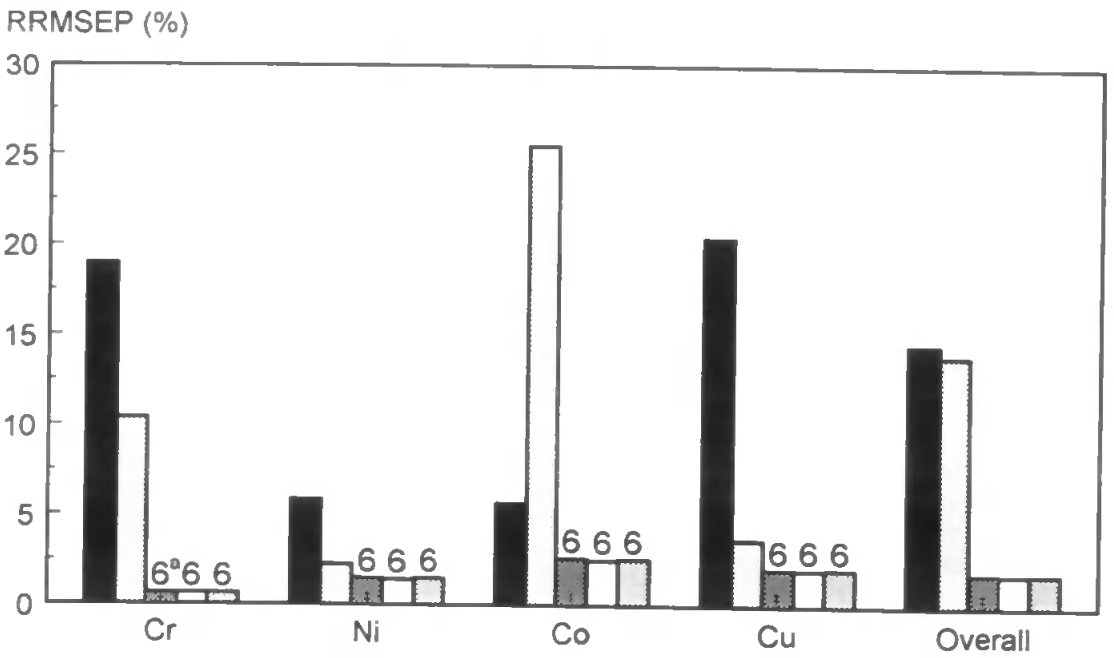
The results for the four-component system with barium sulfate interference are given in Figure 4.7. While no significant differences were observed between the low predictive errors of PCR, PLS1 and PLS2 (< 3% with respect to the four metal ions), their overall precisions were considerably superior to those of DMA and SMLR when interference effects were present in the system. RRMSEP values were particularly high for chromium and copper in the case of DMA, and chromium and cobalt with respect to SMLR.

Figure 4.6 Precisions for the four-component system. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; □, PLS2.



[a] Number of principal components used for predictions.

Figure 4.7 Precisions for the four-component system with barium sulfate interference. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; □, PLS2.

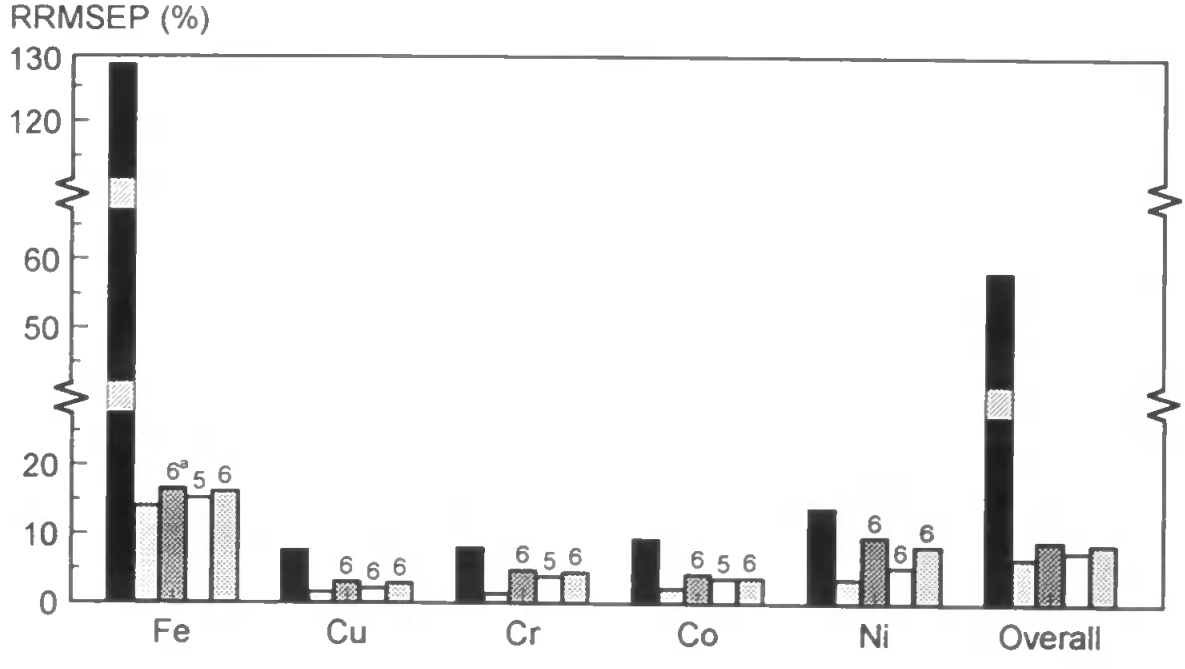


[a] Number of principal components used for predictions.

4.4.4 Five-component system

Figure 4.8 presents the results for the five-component system, which incorporated a strong chemical interaction between copper and iron. Partial oxidation of Fe^{2+} ions in the presence of Cu^{2+} (simulating inter-analyte interactions in a real effluent system) had a considerable effect on the visible spectrum of iron. It can be seen that the precisions for each technique with respect to iron were considerably poorer than those for the other metal ions. In this case DMA was unable to take account of the interaction, with a precision of > 120% for iron (overall precision > 50%), and is therefore unsuitable for on-line quantification of such a system. In contrast, the precisions of SMLR, PCR, PLS1 and PLS2 for iron were all in the range 14-17%, with overall precisions all < 10%.

Figure 4.8 Precisions for the five-component system. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; ▨, PLS2.

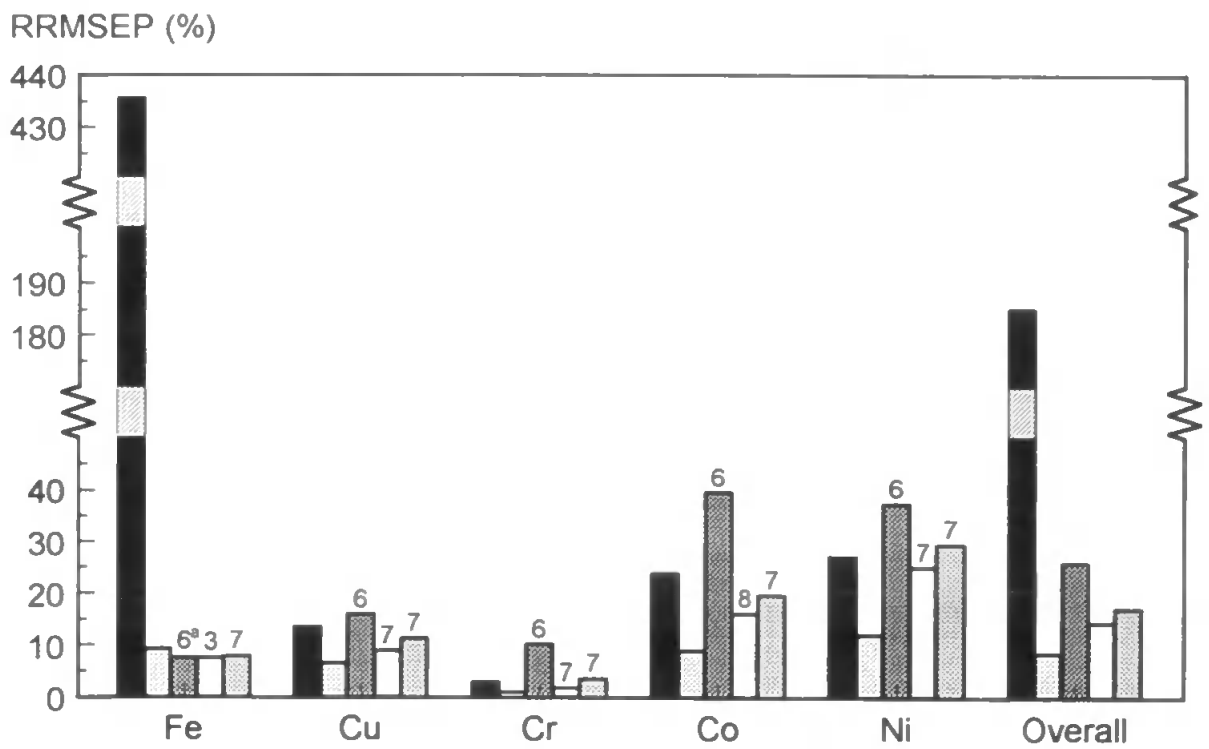


[a] Number of principal components used for predictions.

The most rigorous test of calibration and prediction procedures was presented by the five-component system with barium sulphate interference and the results are shown in

Figure 4.9. Again SMLR, PCR and the PLS algorithms produced much better precisions than DMA. The best overall precision was produced by SMLR (< 10%), which demonstrates that this technique can potentially accommodate the effects of suspended solids and a chemical interaction within the same system. However it should be remembered that the calibration and test sets were of a very similar nature. In some cases real effluent streams would be inherently more difficult to model, in which instances one would expect the signal averaging and data decomposition capabilities of PCR, PLS1 and PLS2 to produce more robust calibrations than SMLR. In addition, the high prediction errors produced by SMLR for the four-component system with barium sulfate interference demonstrate that PCR and the two PLS algorithms are more reliable in terms of predictive precision.

Figure 4.9 Precisions for the five-component system with barium sulfate interference. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; ▨, PLS2.

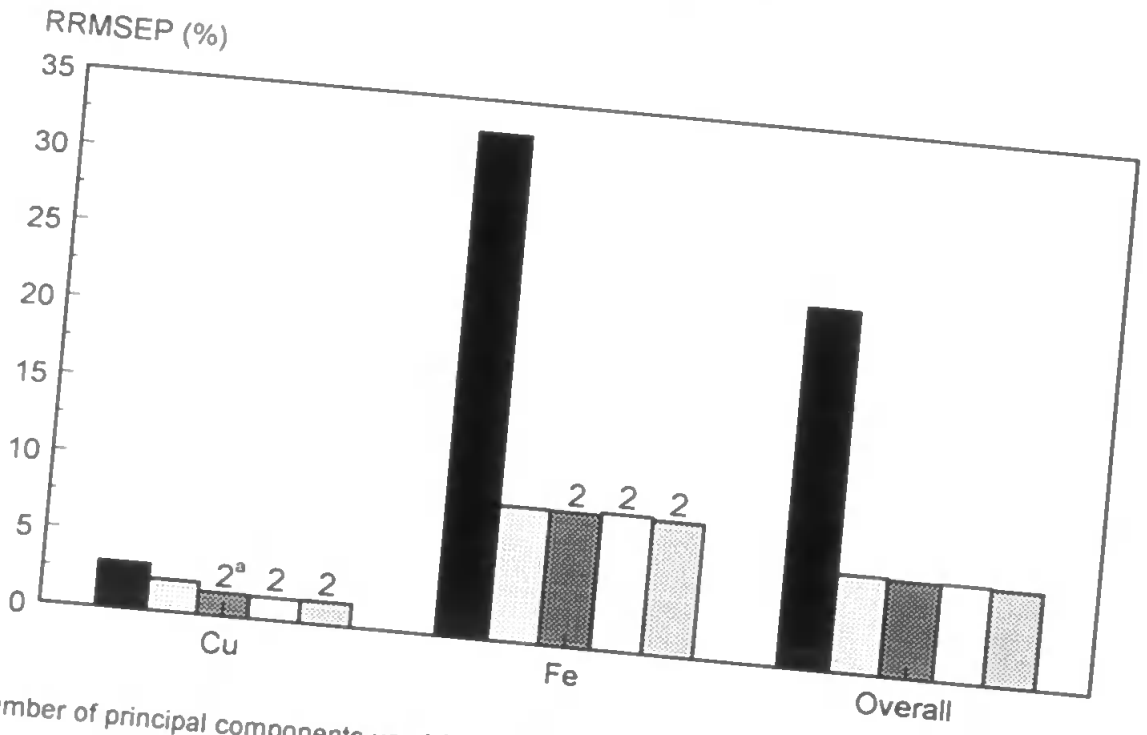


[a] Number of principal components used for predictions.

4.4.5 Copper-iron Interaction

The interaction between copper and iron was examined in the absence of the other metal ions, using a two-component system. The results given in Figure 4.10 show that predictions for copper were in all cases more precise than those for the partially oxidised iron(II). As with the five-component system, SMLR, PCR, PLS1 and PLS2 produced significantly better precisions (all < 10%) for iron than DMA (> 30%). The overall precisions of all calibration techniques with the exception of DMA were in the range 6-7%, and therefore appropriate to the requirements of on-line monitoring.

Figure 4.10 Precisions for the two-component (copper and iron) system. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; ▨, PLS2.



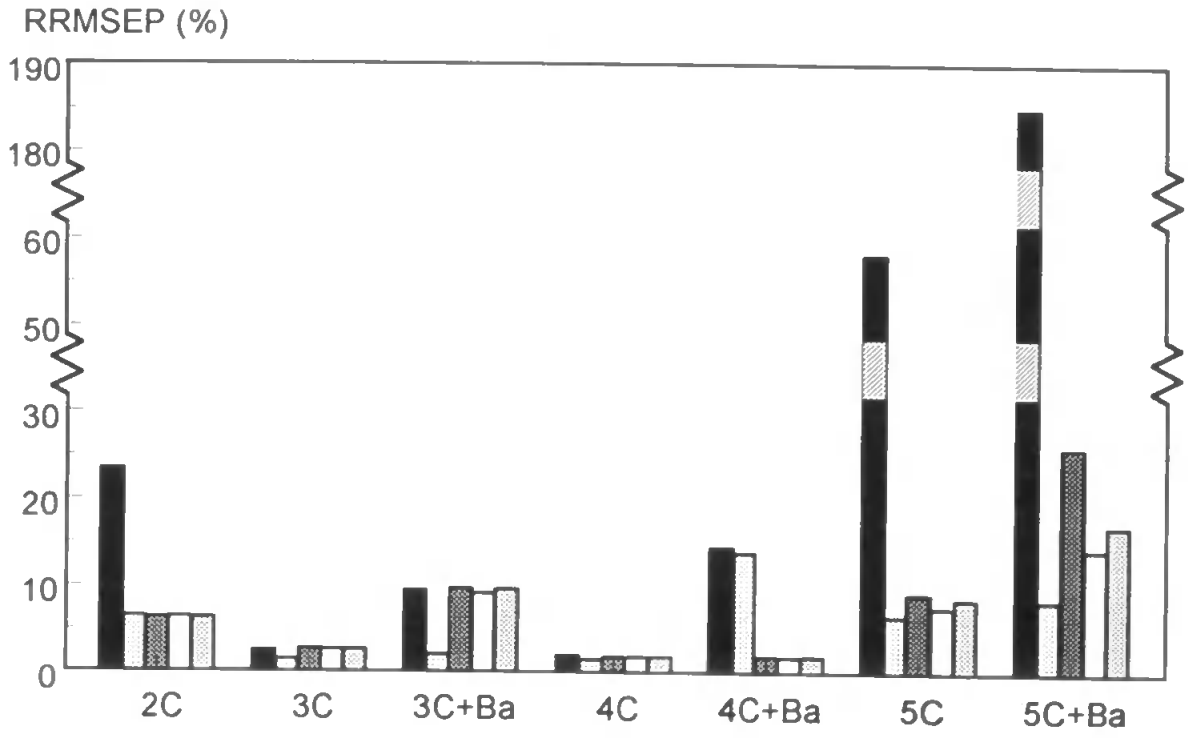
[a] Number of principal components used for predictions.

4.4.6 Summary of overall precisions for multicomponent absorbance data sets

The overall precisions obtained for the five multivariate calibration techniques when using absorbance data are summarised in Figure 4.11 for all the multicomponent systems. This

clearly demonstrates that DMA was unable to match the predictive precision of the other four techniques when chemical and/or physical interferences were present. The results also show that PLS1 offered marginally better precisions than PLS2 or PCR for the more complex systems, which reflects the fact that individual PLS1 models were constructed for each component, and therefore were not restricted to using the same optimal dimensionality for every component of a given system.

Figure 4.11 Comparison of the overall precisions for each multicomponent system. ■, DMA; □, SMLR; ▩, PCR; □, PLS1; ▨, PLS2.

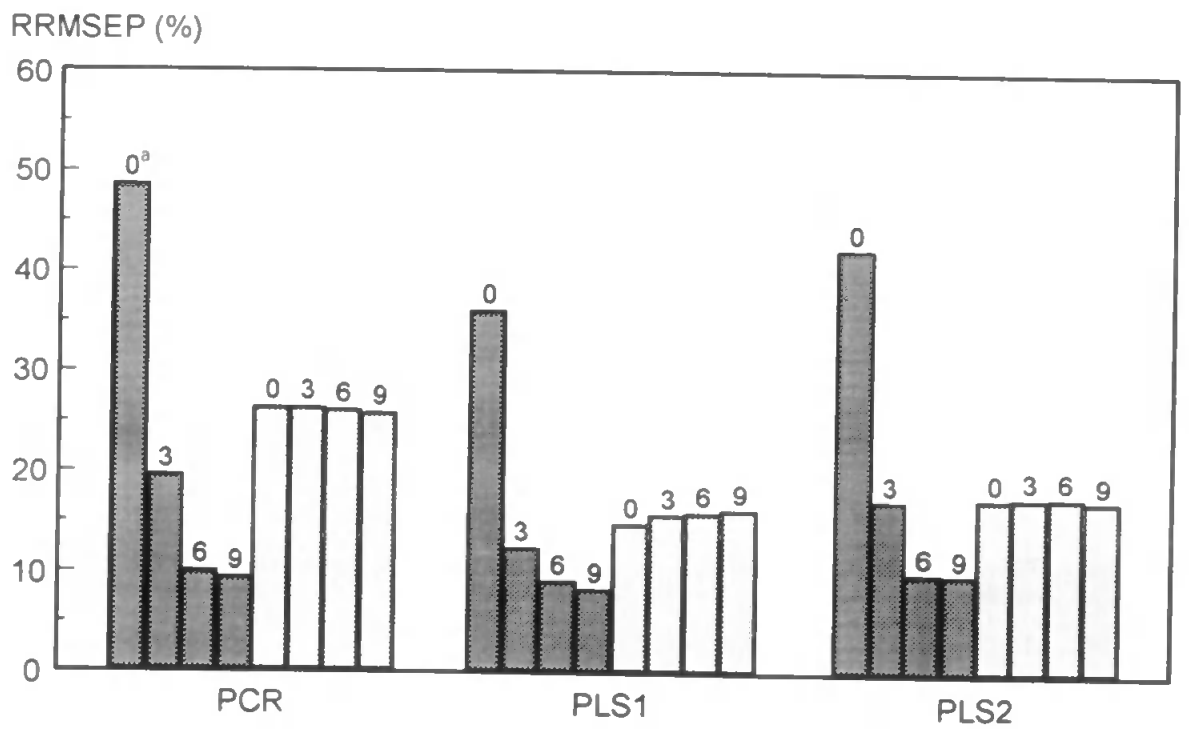


4.4.7 First-derivative data

It has been reported that improved multivariate calibration precisions can be obtained by applying derivatisation techniques [189]. Therefore the PCR, PLS1 and PLS2 calibrations were repeated for the five-component system with barium sulfate interference, by using first-derivative spectra with varying degrees of wavelength averaging (0, 3, 6 and 9 data

points). Figure 4.12 gives the overall precisions for these calibrations, together with those obtained using absorbance data with the same wavelength averaging steps. For the absorbance data, wavelength averaging had no significant effect on the calibration precisions. When using first-derivative data with no wavelength averaging, a significant deterioration in precision resulted, since greater levels of noise were introduced into the models. However, precisions improved progressively to 10% for each calibration method as increasing degrees of wavelength averaging were applied to the data. These results demonstrate that first-derivative data can produce considerably better precisions than absorbance data, provided that an optimum level of averaging is used. If more wavelengths were averaged however, a point would be reached where the advantages of reduced noise would be outweighed by the disadvantages of lost spectral resolution.

Figure 4.12 Comparison of overall precisions for the five-component system with barium sulfate interference using first-derivative and absorbance data. ■, First-derivative data; □, absorbance data.



[a] Number of wavelengths averaged.

4.5 CONCLUSIONS

The five multivariate calibration techniques examined all offered a high degree of precision when making predictions for the well behaved three- and four-component systems. However, when physical or chemical interferences were incorporated, SMLR, PCR and the two PLS routines provided significantly more robust calibrations than those of DMA. No significant differences were observed between the overall precisions of PCR, PLS1 and PLS2, other than in the case of the most challenging five-component system, in which both chemical and physical interferences were present. SMLR often provided the best precisions, in both well behaved and more complex systems.

It was demonstrated that the precisions of multivariate calibrations could be significantly improved by using first-derivative data, provided that an optimum level of wavelength averaging was applied.

Chapter 5

Flow Injection with Multivariate Calibration for the Quantification of BTEX Compounds in Model Systems and Effluents

5.1 INTRODUCTION

In Chapter 2 an FI monitor was successfully applied to the automated, on-line determination of a single analyte in effluent streams, while Chapter 4 demonstrated the potential of using multivariate calibration to quantify complex, multi-analyte systems. The aim of this chapter is to develop a system combining FI with diode array detection and multivariate calibration to simultaneously determine a group of effluent analytes.

Multivariate calibration is increasingly being used in conjunction with FI techniques, and reported methods include those in which DMA [190], MLR [191 , 192], PCR [189, 193 , 194] and PLS [195 , 196 , 197] have been applied to the quantification of a wide range of multicomponent systems. It has been suggested [10, 20, 198] that the combination of multivariate calibration and multichannel detection, interfaced with FI and/or fibre-optic technologies, will play an increasingly important role in the on-line monitoring of industrial process and effluent streams. With miniature diode array systems now commercially available, this combination offers the advantages of robustness, rapid sample throughput, low maintenance requirement and low capital/operating costs.

This chapter describes the development of an FI-diode array-multivariate calibration method for the simultaneous determination of BTEX compounds. BTEX is a generic name given to benzene, toluene, ethylbenzene and the isomers of xylene, which are an important group of aromatic hydrocarbons currently under discussion for listing within the framework of the EC Dangerous Substances in Water directive (76/464/EEC) [2].

Capillary column gas chromatography, following a purge-and-trap [199] or solvent extraction [200] procedure, is often used for the laboratory analysis of BTEX compounds in wastewaters. However, this technique is not an attractive option for continuous on-line monitoring owing to its relative complexity and low sampling frequency (>40 min per sample).

The method described here applies SMLR, PCR, PLS1 and PLS2 to the calibration of UV spectral data sets for synthetic mixtures of toluene, ethylbenzene and *o*-xylene in hexane and aqueous solutions. DMA is not used in this work since it was often shown in the previous chapter to have an inferior predictive performance. Analysis of a real effluent matrix spiked with the analytes of interest is also performed, with a solvent extraction procedure incorporated into the FI method in order to minimise the effect of sample matrix interferences.

5.2 EXPERIMENTAL

5.2.1 Reagents

Stock solutions (1000 μmol l^{-1}) were initially prepared in hexane (Rathburn, Walkerburn, Scotland) for toluene (AnalaR grade; Merck, Darmstadt, Germany), ethylbenzene (Sigma, Poole, Dorset, UK) and *o*-xylene (HPLC grade; Sigma-Aldrich, Gillingham, Dorset, UK). Calibration and test set solutions in the range 10-60 μmol l^{-1} were prepared by serial dilution of the stock solutions with hexane. Aqueous stock solutions (100 mg l^{-1}) were also prepared with Milli-Q water (Millipore, Milford, MA, USA) for toluene, ethylbenzene and *o*-xylene, with calibration and test set solutions in the range 1-20 mg l^{-1} prepared by serial dilution. Solvent extraction of the aqueous solutions was performed using hexane.

5.2.2 Instrumentation

A Hewlett-Packard (Avondale, PA, USA) 8451A photodiode array fitted with either a 1 cm pathlength silica cell (for static measurements) or a silica flow cell (18 μl, 1 cm pathlength; Hellma, Westcliff-on-Sea, Surrey, UK) was used to obtain absorbance and first and second derivative UV spectra in the range 200-300 nm for all samples. Raw data were initially

stored on an HP 9121 disk drive unit, then transferred in ASCII format *via* an HP 82939A serial interface to a personal computer.

A single-channel flow injection manifold (see Figure 5.1) was used for the analysis of aqueous solutions, and was constructed using poly(tetrafluoroethylene) (PTFE) tubing of 0.75 mm i.d. (Fisons, Loughborough, Leicestershire, UK). Two peristaltic pumps (Mini-S 820; Ismatec, Carshalton, Surrey, UK) with poly(vinyl chloride) (PVC) pump tubing (1.30 mm i.d.; Elkay, Basingstoke, Hampshire, UK) were used to transport a water carrier stream and the sample stream at 1.5 ml min^{-1} each to a pneumatic six-port rotary injection valve unit (PS Analytical, Sevenoaks, Kent, UK) and on to the flow cell. The sample injection volume was 280 µl.

A modified, two-channel FI manifold (Figure 5.2) was used for solvent extraction of aqueous samples. An on-line solvent extraction cell was constructed in-house using a PTFE block in two halves, as shown in Figure 5.3. When joined, a PTFE microporous membrane (0.085 mm thickness, 0.02 µm pore size; Goodfellow, Cambridge, Cambridgeshire, UK) partitioned a central flow channel (2 × 3 × 70 mm in each half). Two peristaltic pumps (Ismatec) were used to pump a water carrier stream at 2.3 ml min^{-1} (1.52 mm i.d. PVC pump tubing; Elkay) and a hexane stream at 0.36 ml min^{-1} (0.635 mm i.d. Viton® pump tubing; Ismatec), *via* a PTFE T-piece (in-house construction), to the extraction cell. A Minipuls 2 peristaltic pump (Gilson, Villiers-le-Bel, France) with 0.635 mm i.d. Viton® pump tubing was used to draw the hexane fraction from the extraction cell at 0.26 ml min^{-1}. In this case a sample injection volume of 200 µl was used.

5.2.3 Software

A program written in HP BASIC was used to automate the FI manifold components and measure/record the UV absorbance/derivative spectra. SMLR was performed using Minitab v. 8.2 statistical software (Minitab, State College, PA, USA), while PCR, PLS1 and

PLS2 were conducted using Unscrambler v. 5.03 multivariate analysis software (Camo A/S, Trondheim, Norway).

Figure 5.1 Schematic diagram of the single channel FI manifold used for the determination of BTEX compounds in aqueous model systems.

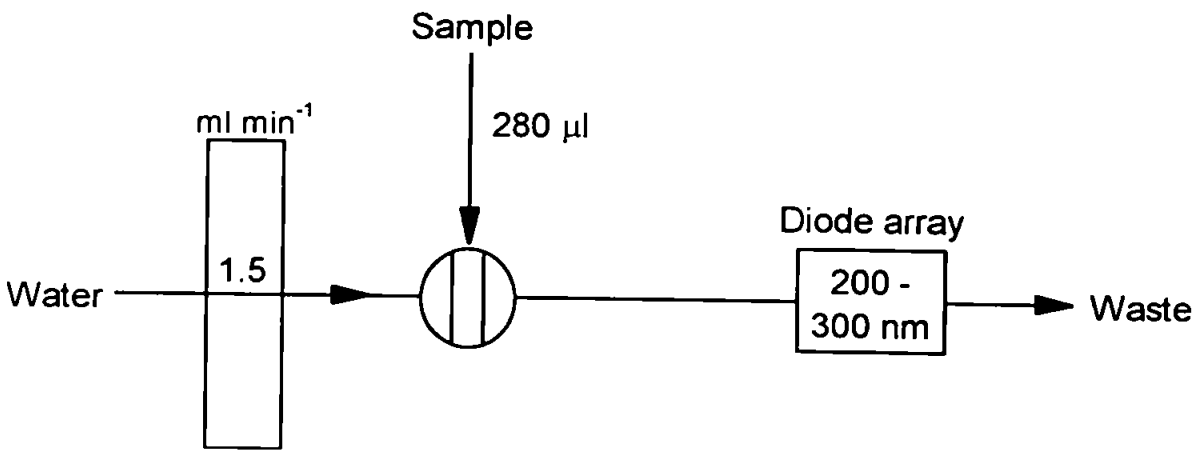


Figure 5.2 Schematic diagram of the modified FI manifold used for solvent extraction of aqueous *o*-xylene solutions and spiked effluent samples.

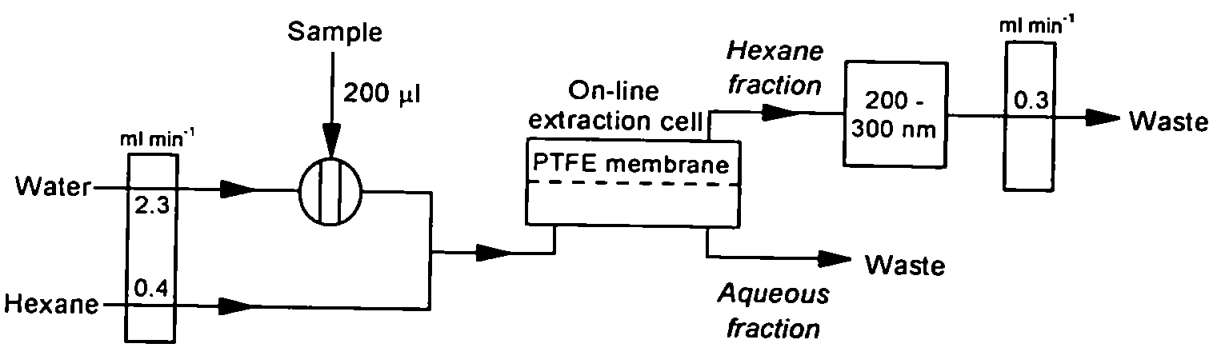
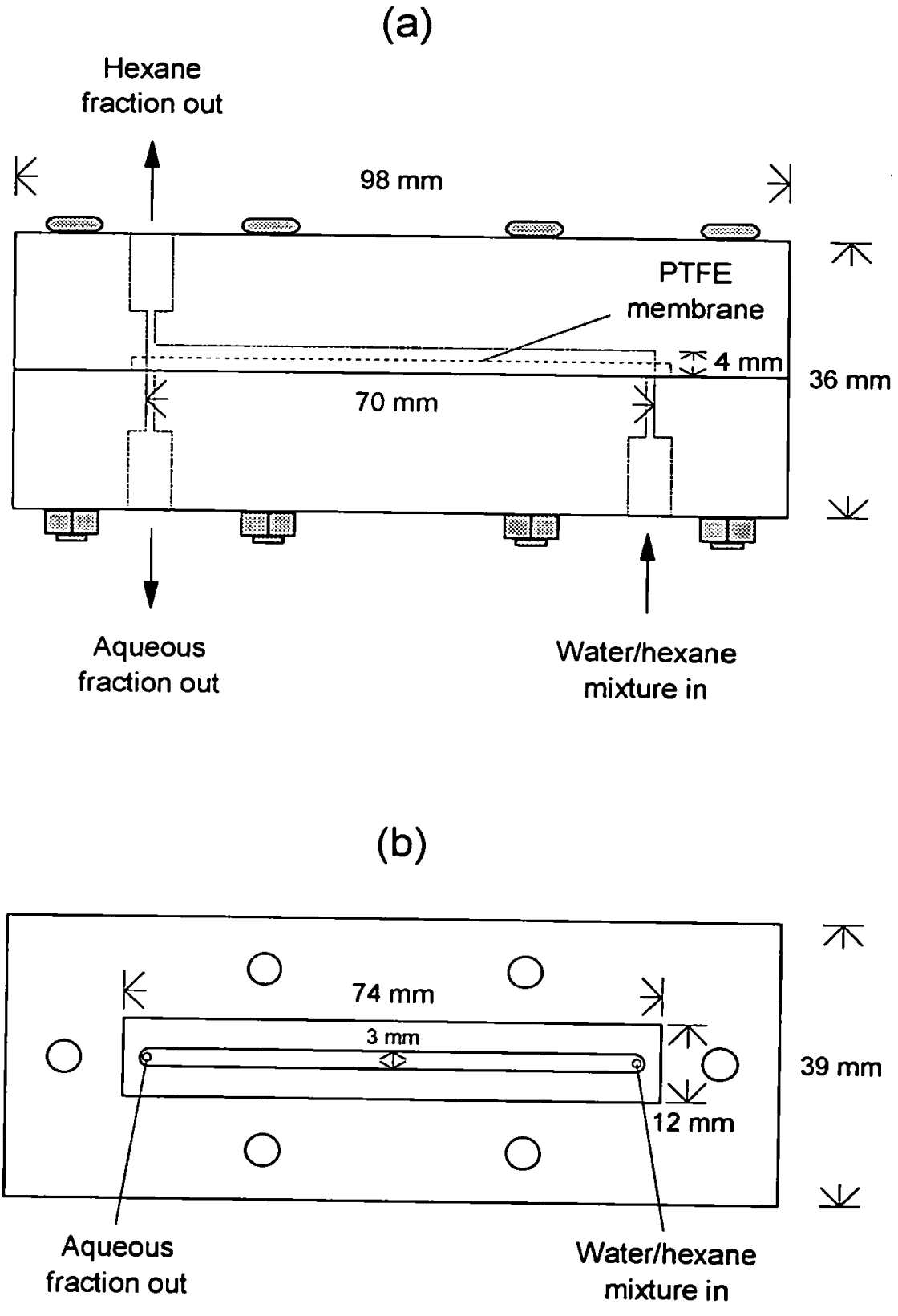


Figure 5.3 Diagram of the solvent extraction cell: (a) side view; (b) plan view of lower half (inner face).



5.2.4 Procedures

Three-component model systems (toluene, ethylbenzene and *o*-xylene) were prepared in both hexane and aqueous solution. In both cases, a 3^3 factorial design was used to construct a 27-sample calibration set and random concentrations were used to produce a 20-sample independent test set (see Table 5.1 and Table 5.2). The samples were measured in randomised order to reduce the risk of bias. UV absorbance spectra were recorded over the wavelength range 200-300 nm, with measurements taken at 2 nm intervals to produce 50 data points per spectrum. Each solution was measured in triplicate against a solvent (*i.e.* hexane or Milli-Q water) reference, using an integration time of 5 s. The triplicate sets were averaged to produce mean spectra, which were stored for use in calibration or prediction. First- and second-derivative spectra were calculated with three-point wavelength smoothing (according to the Savitsky-Golay algorithm) using software supplied on-board the diode array.

The spectral data for the calibration sets were used to construct SMLR, PCR, PLS1 and PLS2 calibration models, as described in Chapter 4. These were used to predict analyte concentrations in the respective test set solutions. As in Chapter 4, mean-centring was applied to all variables used for PCR, PLS1 and PLS2 calibrations, and the precision of each calibration technique was again expressed in terms of the relative root-mean-square error of prediction (equation 4.8). The criteria used to define optimal dimensionality for PCR, PLS1 and PLS2 models was that described in Section 4.3.6.

In addition to the static measurements described above, three-component aqueous calibration and test sets (Table 5.3) were analysed using an automated, single-channel FI manifold to deliver the samples to the diode array For all samples, spectra were measured every 1 s following injection and at 2 nm intervals over the range 200-300 nm against a water reference (see Figure 5.4). The recorded spectrum in each case was taken as the difference between the average of three spectra immediately following injection (the

baseline) and the average of three spectra around the peak maximum. The solutions were again measured in triplicate, with the averaged spectra used in calibration.

Table 5.1 Concentration data (μmol l^{-1}) for the three-component system in hexane (3^3 factorial design calibration set and random design test set).

Sample number	Calibration set			Test set		
	Toluene	Ethylbenz.	o-Xylene	Toluene	Ethylbenz.	o-Xylene
1	0	0	0	30	30	0
2	20	0	0	10	20	10
3	50	0	0	50	20	10
4	0	0	20	20	50	0
5	20	0	20	20	40	20
6	50	0	20	20	40	50
7	0	0	50	10	40	10
8	20	0	50	50	0	40
9	50	0	50	0	40	0
10	0	20	0	0	20	30
11	20	20	0	20	10	10
12	50	20	0	10	50	0
13	0	20	20	0	20	10
14	20	20	20	0	10	20
15	50	20	20	40	30	0
16	0	20	50	10	10	40
17	20	20	50	10	0	0
18	50	20	50	40	10	30
19	0	50	0	0	20	30
20	20	50	0	20	50	20
21	50	50	0	-	-	-
22	0	50	20	-	-	-
23	20	50	20	-	-	-
24	50	50	20	-	-	-
25	0	50	50	-	-	-
26	20	50	50	-	-	-
27	50	50	50	-	-	-

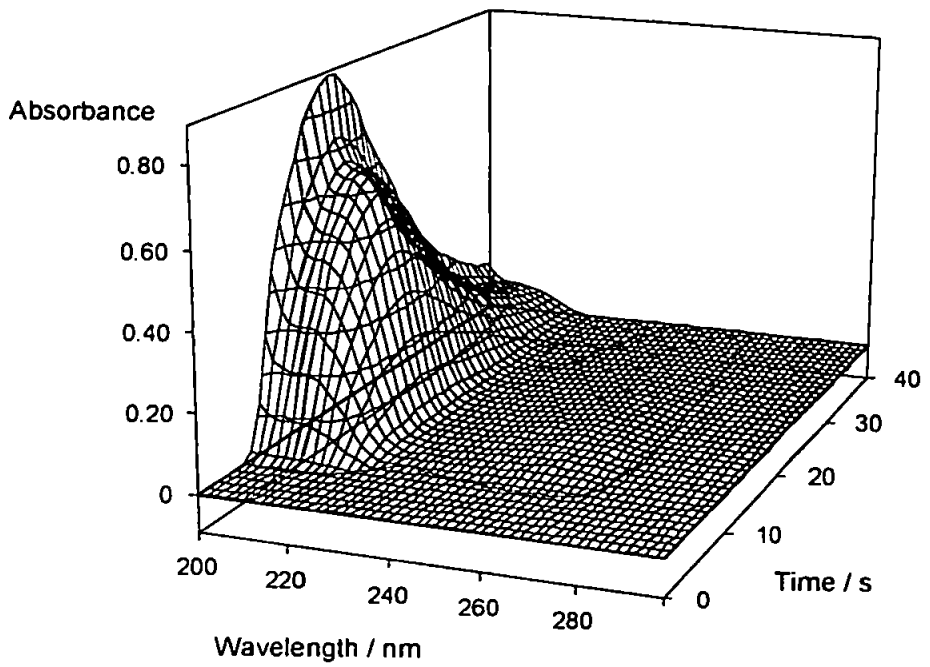
Table 5.2 Concentration data (mg l^{-1}) for the three-component system in aqueous solution for static determinations (3^3 factorial design calibration set and random design test set).

Sample number	Calibration set			Test set		
	Toluene	Ethylbenz.	o-Xylene	Toluene	Ethylbenz.	o-Xylene
1	0	0	0	2	10	4
2	5	0	0	10	0	1
3	10	0	0	5	0	7
4	0	0	5	6	8	5
5	5	0	5	10	2	0
6	10	0	5	10	8	10
7	0	0	10	1	5	7
8	5	0	10	10	6	7
9	10	0	10	4	7	2
10	0	5	0	8	2	4
11	5	5	0	7	3	2
12	10	5	0	8	5	5
13	0	5	5	4	2	5
14	5	5	5	6	0	9
15	10	5	5	1	8	3
16	0	5	10	1	5	9
17	5	5	10	4	8	3
18	10	5	10	7	1	0
19	0	10	0	7	10	7
20	5	10	0	7	8	5
21	10	10	0	-	-	-
22	0	10	5	-	-	-
23	5	10	5	-	-	-
24	10	10	5	-	-	-
25	0	10	10	-	-	-
26	5	10	10	-	-	-
27	10	10	10	-	-	-

Table 5.3 Concentration data (mg l^{-1}) for the three-component system in aqueous solution for FI determinations (3^3 factorial design calibration set and random design test set).

Sample number	Calibration set			Test set		
	Toluene	Ethylbenz.	o-Xylene	Toluene	Ethylbenz.	o-Xylene
1	0	0	0	4	2	9
2	10	0	0	20	14	11
3	20	0	0	17	12	13
4	0	0	10	3	18	11
5	10	0	10	3	19	13
6	20	0	10	7	6	15
7	0	0	20	14	7	11
8	10	0	20	12	9	8
9	20	0	20	4	18	4
10	0	10	0	17	17	16
11	10	10	0	2	5	20
12	20	10	0	7	0	17
13	0	10	10	11	1	11
14	10	10	10	20	1	15
15	20	10	10	0	2	4
16	0	10	20	6	19	2
17	10	10	20	2	5	8
18	20	10	20	5	13	2
19	0	20	0	7	20	1
20	10	20	0	8	12	17
21	20	20	0	-	-	-
22	0	20	10	-	-	-
23	10	20	10	-	-	-
24	20	20	10	-	-	-
25	0	20	20	-	-	-
26	10	20	20	-	-	-
27	20	20	20	-	-	-

Figure 5.4 FI response profile (absorbance against wavelength and time) for a solution containing toluene, ethylbenzene and *o*-xylene each at 20 mg l^{-1}.



The potential of incorporating an on-line solvent extraction procedure to the FI-diode array method was also investigated. A modified FI manifold was used to analyse both *o*-xylene calibration standards and solutions incorporating a real effluent matrix spiked with either one or three components.

5.3 RESULTS AND DISCUSSION

5.3.1 Spectral characteristics

Absorbance, first-derivative and second-derivative spectra for pure 20 μmol l^{-1} solutions of toluene, ethylbenzene and *o*-xylene in hexane are shown in Figures 5.5(a) to 5.5(c). Pure 10 mg l^{-1} aqueous solutions of the same three compounds are shown in Figure 5.6(a) to 5.6(c). The spectral profiles were very similar in both solvents, and a high degree of spectral

Figure 5.5 UV absorbance spectral profiles for toluene (————), ethylbenzene (·····) and *o*-xylene (– – –) in hexane (20 μmol l^{-1} solutions): (a) absorbance data; (b) first-derivative data; (c) second derivative data.

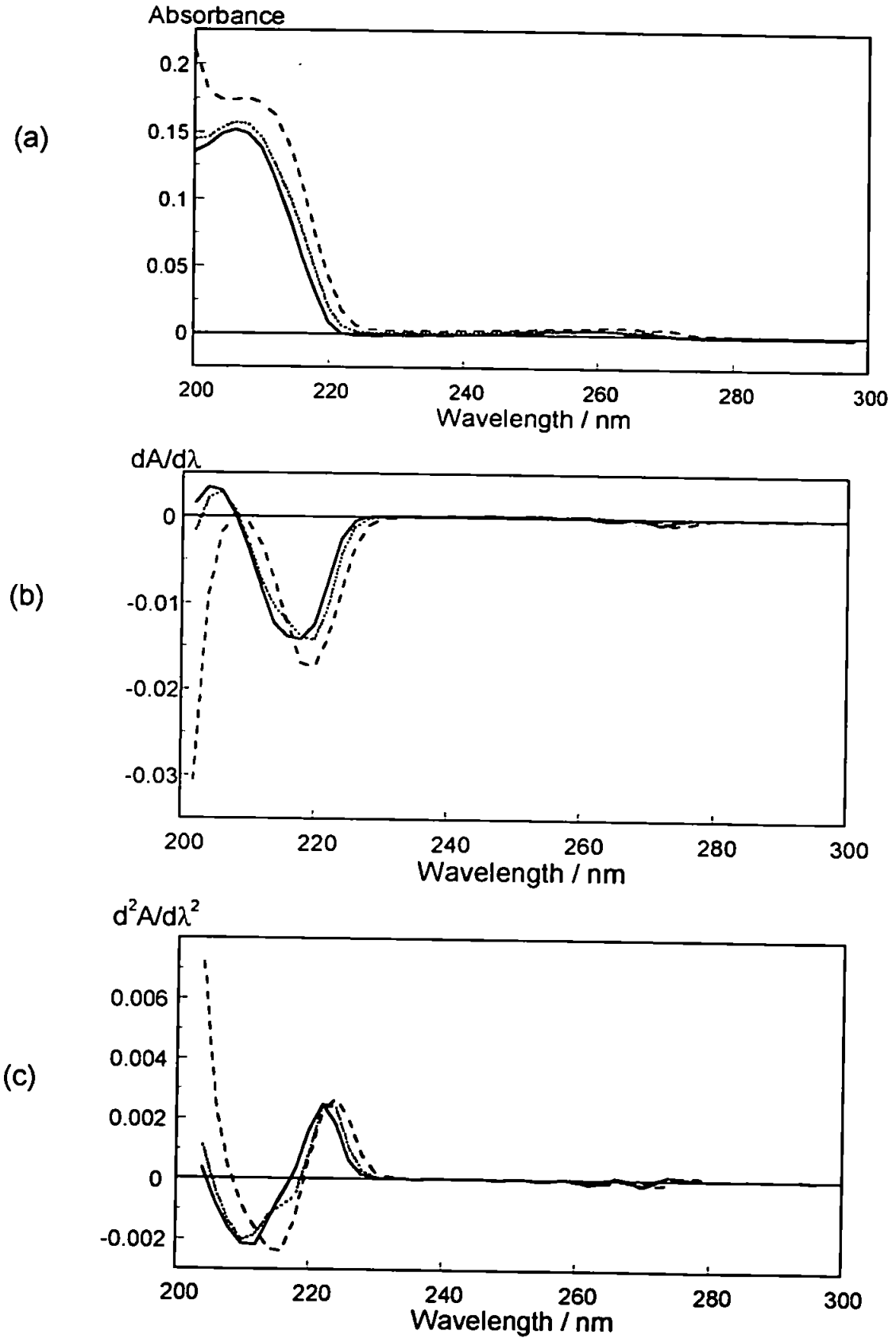
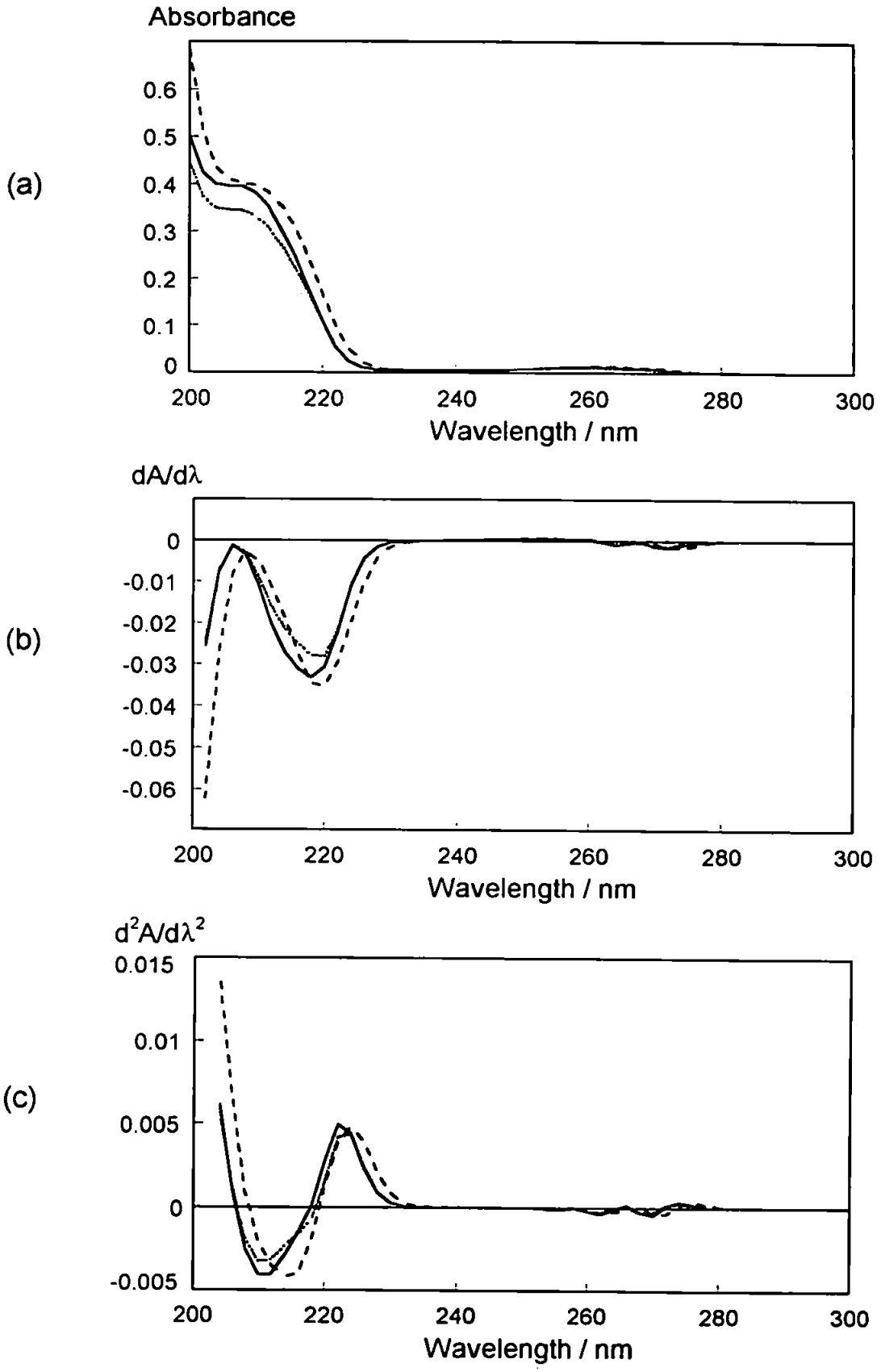


Figure 5.6 UV absorbance spectral profiles for toluene (———), ethylbenzene (·····) and *o*-xylene (– – –) in aqueous (10 mg l^{-1}) solutions: (a) absorbance data; (b) first-derivative data; (c) second derivative data.



overlap was evident in each case, particularly in the case of toluene and ethylbenzene. Derivatisation was shown to partially resolve the spectra.

5.3.2 Static determination of toluene, ethylbenzene and *o*-xylene in hexane solution

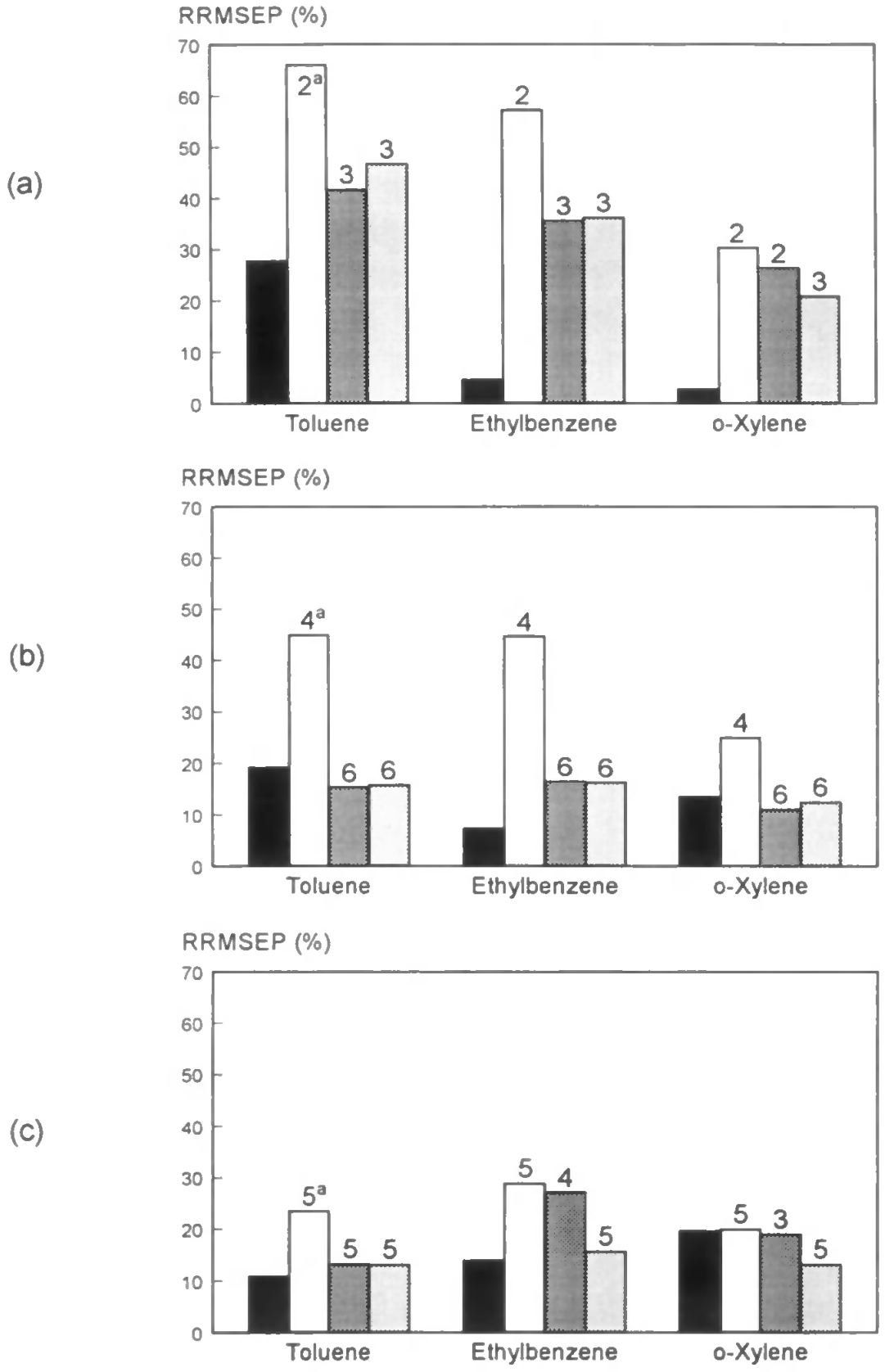
An absorbance versus concentration plot for pure solutions of *o*-xylene in hexane was linear over the range 0-80 μmol l^{-1} (0-8.5 mg l^{-1}) at 206 nm (R^2 = 0.999, gradient = 0.006 and *y*-intercept = 0.003 absorbance units). The concentration range used for the three components in the calibration and test set solutions was therefore within this linear range.

Relative root-mean-square errors of prediction for the three-component system in hexane solution are shown for absorbance, first-derivative and second-derivative data in Figures 5.7(a) to 5.7(c). The predictive precisions were generally poor for all the calibration techniques, which reflected the stern challenge presented by the high degree of spectral similarity between the three compounds. SMLR was the only method capable of producing prediction errors < 10%, while PCR gave the least precise results (errors > 20% in all cases). Predictive performances were improved by derivatisation of the data, particularly in the case of PCR and the two PLS techniques. The optimal dimensionalities of these calibration models increased from that used for absorbance data in order to incorporate the effects of derivatisation. Smoothing was restricted to three-point wavelength averaging for both first- and second-derivative data, since this found to be sufficient to reduce noise without resulting in a loss of spectral information.

5.3.3 Static determination of toluene, ethylbenzene and *o*-xylene in aqueous solution

The aqueous solubilities (at 20°C) of toluene, ethylbenzene and *o*-xylene are approximately 507, 170 and 170 mg l^{-1} respectively [201 , 202], and concentrations of BTEX compounds in industrial wastewaters are typically in the range 0 - 10 mg l^{-1}. Absorbance versus

Figure 5.7 Predictive precisions for static determinations of the three-component model system in hexane: (a) absorbance data; (b) first-derivative data; (c) second derivative data. ■, SMLR; □, PCR; ▩, PLS1; ▨, PLS2.



[a] Number of principal components used for predictions.

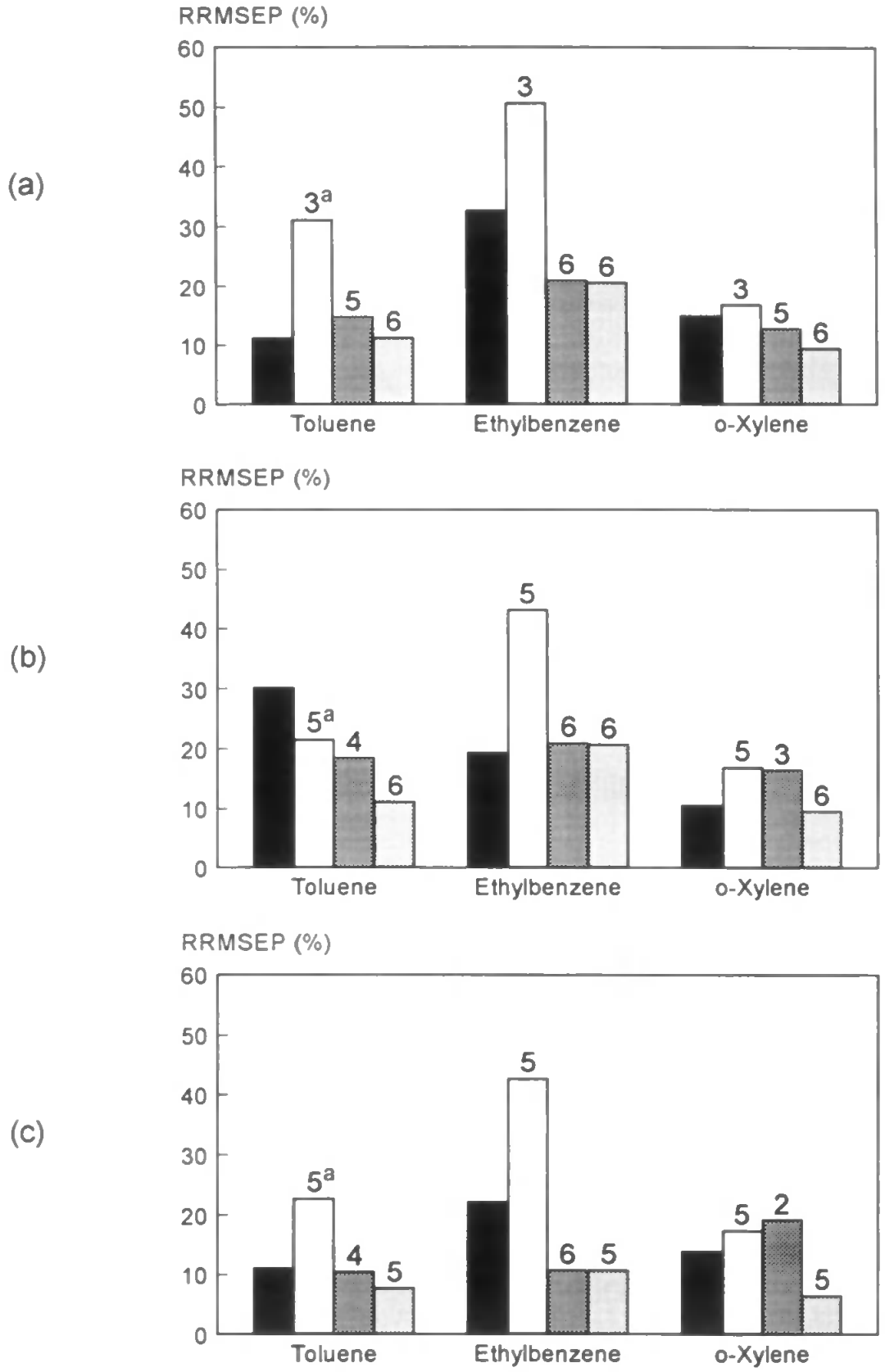
concentration plots for pure aqueous solutions of the three components were linear over the range 0-20 mg l^{-1} at 206 nm (*e.g.* for *o*-xylene, $R^2 = 0.999$, gradient = 0.057 and *y*-intercept = 0.005 absorbance units). The concentration range used for the calibration and test set solutions was therefore within this linear range, and spanned the range of interest for effluents.

RRMSEP results for the aqueous three-component system are shown in Figures 5.8(a) to 5.8(c). Once again the predictive precisions were generally poor for all the calibration techniques, although in this case only PLS2 was capable of producing errors < 10%. SMLR offered no advantages over PLS1 or PLS2, although PCR again displayed the poorest predictive performance. Derivatisation resulted in some improvements in predictive precision, although these were less significant than in the case of the hexane model system. The optimal dimensionalities for PLS1 and PLS2 models when using absorbance data were higher than for the hexane system, which reflects an extra degree of variability in the aqueous system resulting from small evaporative losses of the analytes from the solutions.

5.3.4 Simplex optimisation of a single-channel FI manifold

Simplex optimisation is a multivariate technique often used to configure the operating variables of an analytical system in order to maximise the response signal [203]. The term *simplex* refers to a geometrical figure which has $n + 1$ vertices when a response is being optimised with respect to n parameters [146]. For a simple two-parameter system, this will be a triangle, as shown in Figure 5.9. The points labelled as 1, 2 and 3 in this diagram represent the initial simplex, while the contours are lines of iso-response forming a *response surface* for the two parameters X and Y. The central contour represents the *summit* of the response surface (*i.e.* the highest response level).

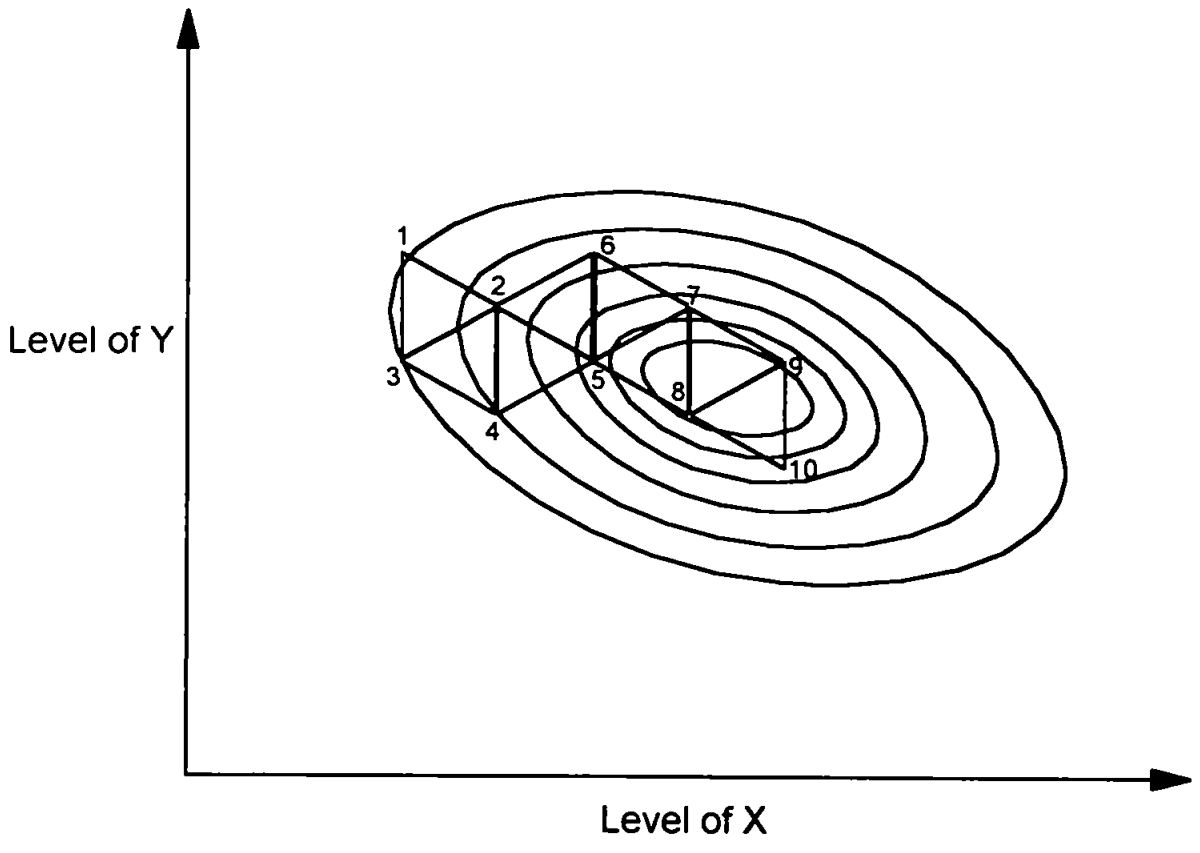
Figure 5.8 Predictive precisions for static determinations of the aqueous three-component model system: (a) absorbance data; (b) first-derivative data; (c) second derivative data. ■, SMLR; □, PCR; ▩, PLS1; ▢, PLS2.



[a] Number of principal components used for predictions.

In the first optimisation experiment, response is measured at each combination of the parameter levels given at points 1, 2 and 3. Since the lowest response for the initial simplex is that located at point 1, the next simplex is chosen as a mirror image of the initial simplex across the line facing the point of lowest response (*i.e.* the line connecting 2 and 3). The new simplex is therefore formed by the points 2, 3 and 4. This procedure is repeated until no further improvement in response can be made. The optimum conditions for the simplex shown in Figure 5.8 will be those defined by point 8, since points 9 and 10 both give lower responses.

Figure 5.9 Simplex optimisation for a two-parameter system [146].



Simplex optimisation was used in this work to optimise three FI manifold parameters, namely injection volume, carrier flow rate and path length (*i.e.* the distance

from the injector to the flow cell). The optimisation procedure was performed for a 1 mg l^{-1} toluene standard in aqueous solution, measuring absorbance at 206 nm. The results of the optimisation procedure are summarised in Table 5.4. It can be seen that maximum response was obtained using the upper levels of injection volume and flow rate and the shortest path length, which is in accordance with the principles of sample dispersion, as discussed in Chapter 1.

Table 5.4 Results for the simplex optimisation of the FI manifold using a 1 mg l^{-1} toluene standard in aqueous solution.

Variable	Precision	Range		Optimum value[a]
		Upper value	Lower value	
Injection volume (µl)	40	320	160	320
Flow rate (ml min^{-1})	0.7	3.4	0.7	3.4
Path length (cm)	50	200	50	50

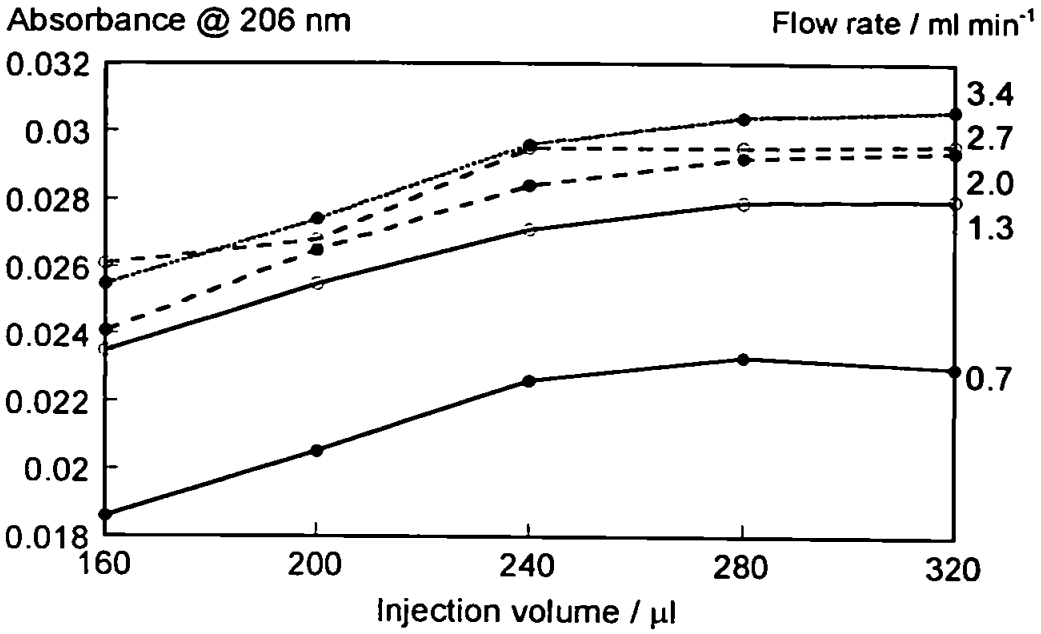
[a] Optimisation procedure ended after 12 runs

A two-variable optimisation for injection volume and flow rate (with the path length fixed at 50 cm) revealed that no significant increase in absorbance was achieved above 280 µl and 1.3 ml min^{-1} respectively, as shown in Figure 5.10.

FI manifold conditions of 50 cm path length, 280 µl injection volume and 1.5 ml min^{-1} flow rate (achieved using 1.30 mm i.d. pump tubing and an Ismatec fixed-speed peristaltic pump) were used to determine the linear response ranges for each analyte. Absorbance versus concentration plots at 206 nm were linear over the range 0-50 mg l^{-1} for the three components (*e.g.* for *o*-xylene, R^2 = 0.999, gradient = 0.011 and *y*-intercept = 0.014 absorbance units). Calibrations in the range 0-1 mg l^{-1} were performed to determine the limits of detection (as defined in Chapter 2), which were found to range from 0.13 mg l^{-1}

for toluene (R^2 = 0.987, gradient = 0.015 and *y*-intercept = 0.007 absorbance units) to 0.19 mg l^{-1} for *o*-xylene (R^2 = 0.987, gradient = 0.010 and *y*-intercept = 0.008 absorbance units).

Figure 5.10 Results for a two-parameter optimisation of the FI manifold using 1 mg l^{-1} toluene standard in aqueous solution.

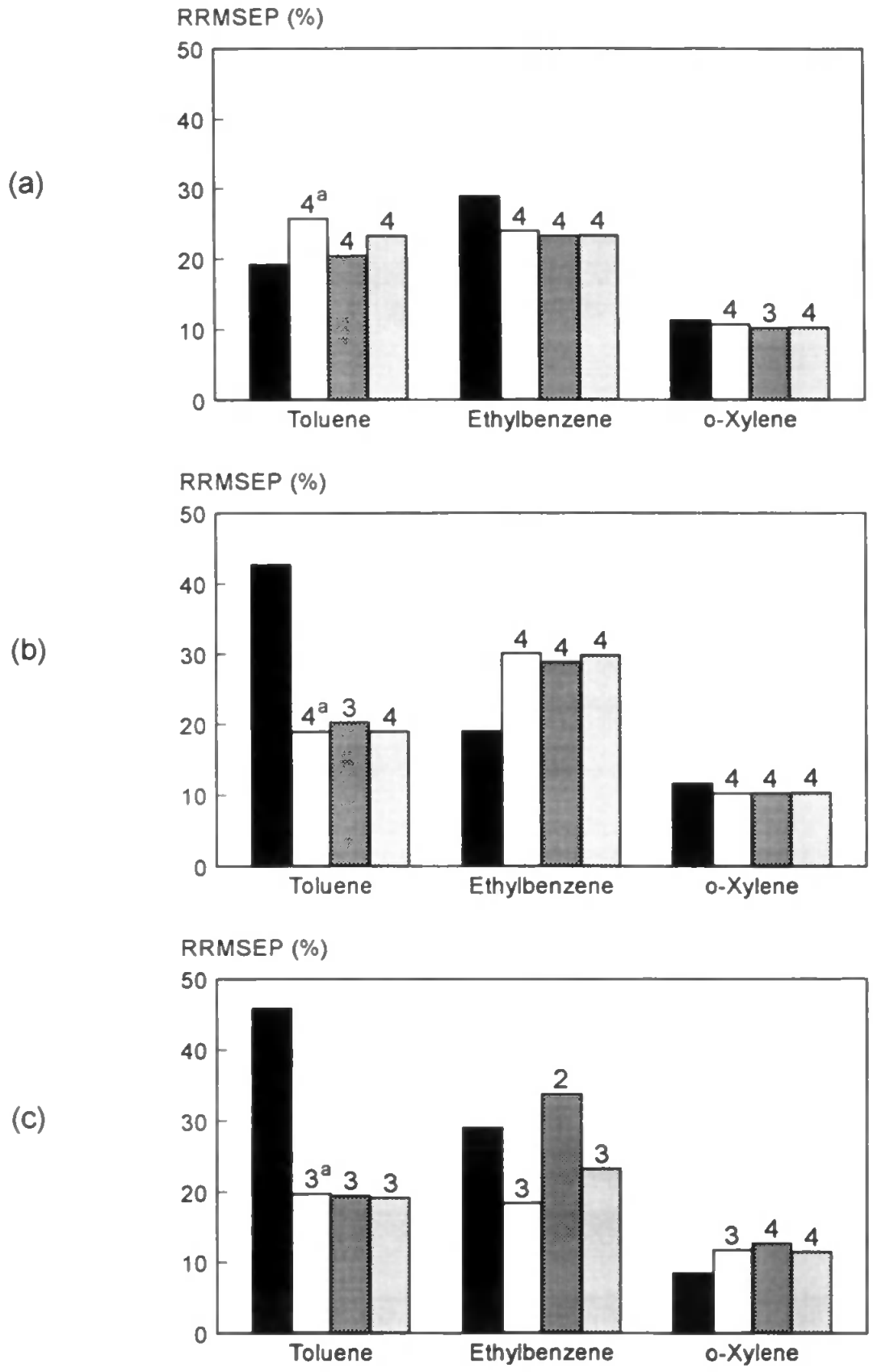


5.3.5 FI determination of toluene, ethylbenzene and *o*-xylene in aqueous solution

The optimised FI manifold was used to perform automated analyses of the aqueous three-component calibration and test set solutions shown previously in Table 5.3. The prediction errors produced with the four multivariate calibration techniques are summarised in Figures 5.11(a) to 5.11(c). The trends observed for the previous static determinations were again observed in this case, with prediction errors in the ranges 8.5-12.6% for *o*-xylene, 19.0-33.7% for ethylbenzene and 18.9-45.8% for toluene.

Only the precisions for *o*-xylene approached the requirements of on-line monitoring when calibrations were performed with respect to the individual components. For this reason, calibrations and predictions were also performed in terms of total TEX compounds

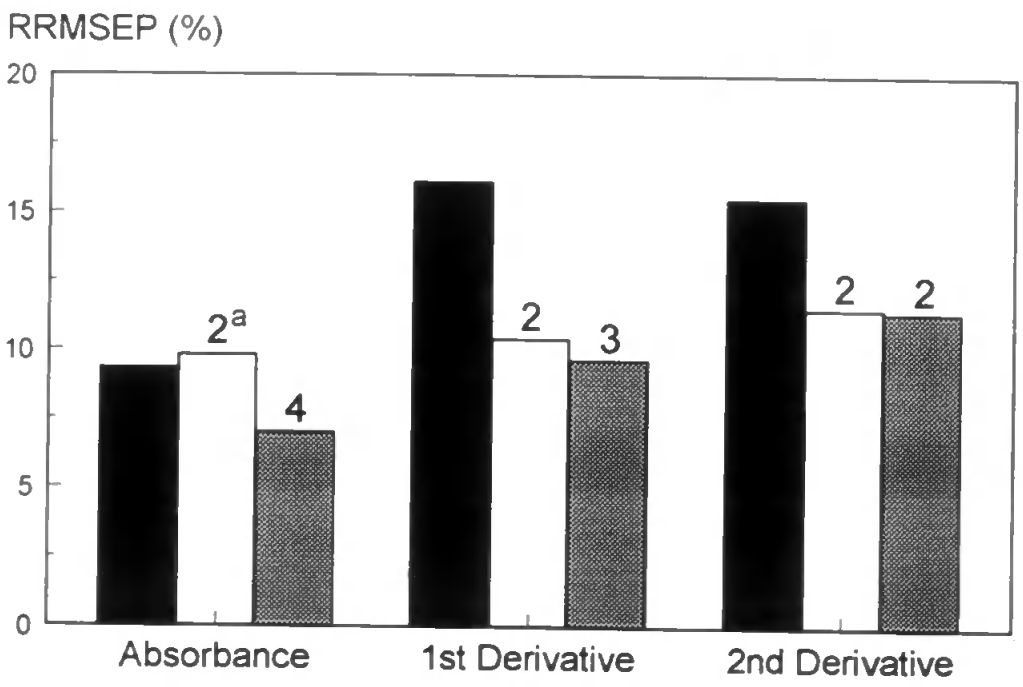
Figure 5.11 Predictive precisions for FI determinations of the aqueous three-component model system: (a) absorbance data; (b) first-derivative data; (c) second derivative data. ■, SMLR; □, PCR; ▩, PLS1; ▨, PLS2.



[a] Number of principal components used for predictions.

(*i.e.* the three components were treated as a single component by summing the concentrations of toluene, ethylbenzene and *o*-xylene in each solution). The predictive errors produced for total TEX are given in Figure 5.12. In this instance PLS2 was not used since the technique is only applicable when more than one component is being modelled. Prediction errors were lower for PCR and PLS1 than those obtained for the individual components, with the best precisions provided by absorbance data (7.0-9.8%) and PLS1 calibrations (7.0-11.4%). These results suggest that in many cases the quantification of total BTEX compounds in effluents would be the most suitable application for an on-line FI-diode array monitor. In this case, predictive performances were least precise when using second-derivative data, which indicates that some spectral information was lost as a result of the derivatisation.

Figure 5.12 Predictive precisions for total TEX in the three-component aqueous model system (FI determination). ■, SMLR; □, PCR; ▩, PLS1.



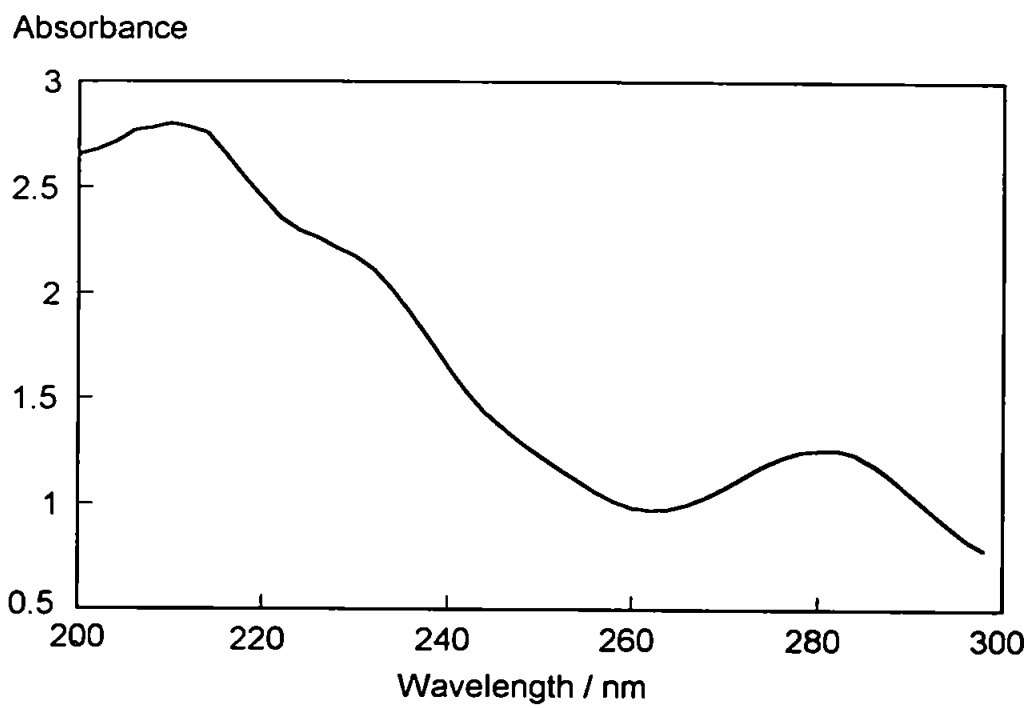
[a] Number of principal components used for predictions.

Data were also obtained for a replicate test set, using identical concentration levels to the first. Predictions were made using the same calibration models as before, and the resulting precisions for total TEX compounds were found to be within 3% RRMSEP of the errors obtained for the first test set (average difference = 1.1% RRMSEP), thus demonstrating that an acceptable degree of between-batch reproducibility was achievable for these determinations.

5.3.6 On-line FI-Solvent Extraction Procedure

Figure 5.13 shows the UV absorbance spectrum for a typical industrial wastewater sample. A very high absorbance is evident in the UV region, which indicates that for real effluent monitoring it will in many cases be necessary to extract BTEX compounds from the effluent matrix prior to measurement to reduce matrix interferences.

Figure 5.13 UV spectrum of a typical industrial effluent sample.



For this reason, the FI manifold was modified to incorporate a hexane stream and an on-line extraction cell (as shown in Figures 5.2 and 5.3 previously). Samples were injected into a water carrier stream, which mixed with the hexane stream before entering the extraction cell. The hexane fraction of the mixed stream was then drawn across the microporous PTFE membrane by the action of a third peristaltic pump, and transported to the diode array detector. The third pump was positioned after the detector, in order to minimise the degree of pulsing in the flow cell.

Two sets of single-component calibration solutions were analysed using this method. Aqueous *o*-xylene solutions in the range 0-20 mg l^{-1} were measured initially, then samples of a real effluent spiked with *o*-xylene to produce the same range of concentrations (20 ml of effluent diluted to 25 ml in each case). The absorbance spectra (measured against a hexane reference) for the aqueous standards and the spiked effluent solutions are given in Figures 5.14 and 5.15 respectively. The results shown in Table 5.5 demonstrate that this approach quantitatively extracted *o*-xylene from an aqueous matrix into a hexane matrix over the range 0-20 mg l^{-1}.

Automated FI extractions were also performed for effluent solutions (17.5 ml of effluent diluted to 25 ml) spiked with three-component mixtures (2^3 factorial design, using 0 and 10 mg l^{-1} concentration levels). The absorbance spectra of these solutions are given in Figure 5.16, which indicates that the method was also able to quantitatively extract three analytes from an effluent matrix, and would therefore be suitable for on-line determinations of total TEX compounds in effluent streams.

Figure 5.14 UV spectra for aqueous *o*-xylene standards following on-line FI extraction.

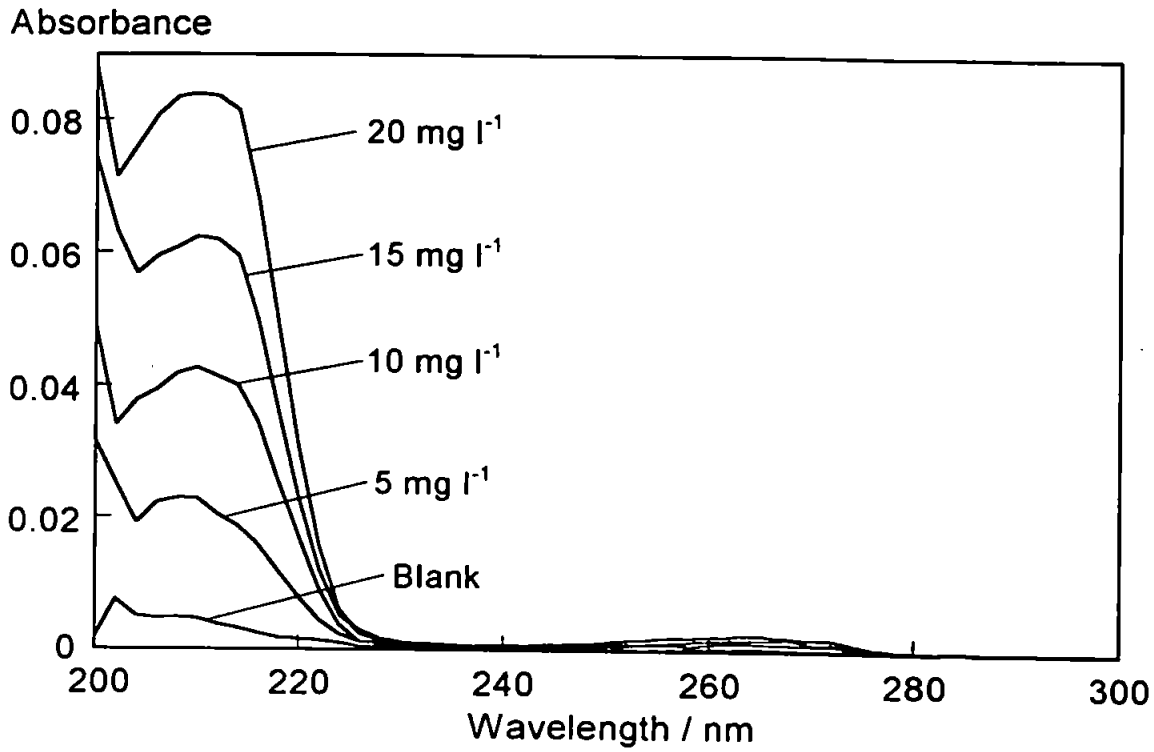


Figure 5.15 UV spectra for effluent samples spiked with *o*-xylene following on-line FI extraction.

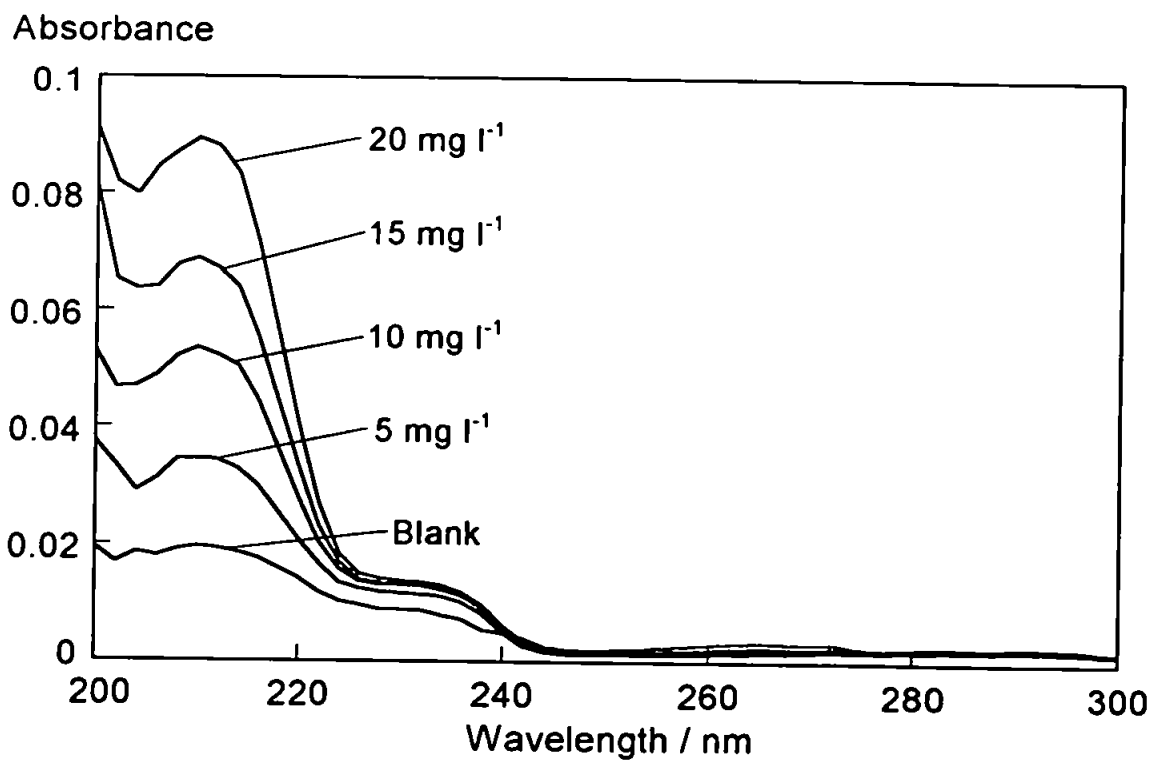
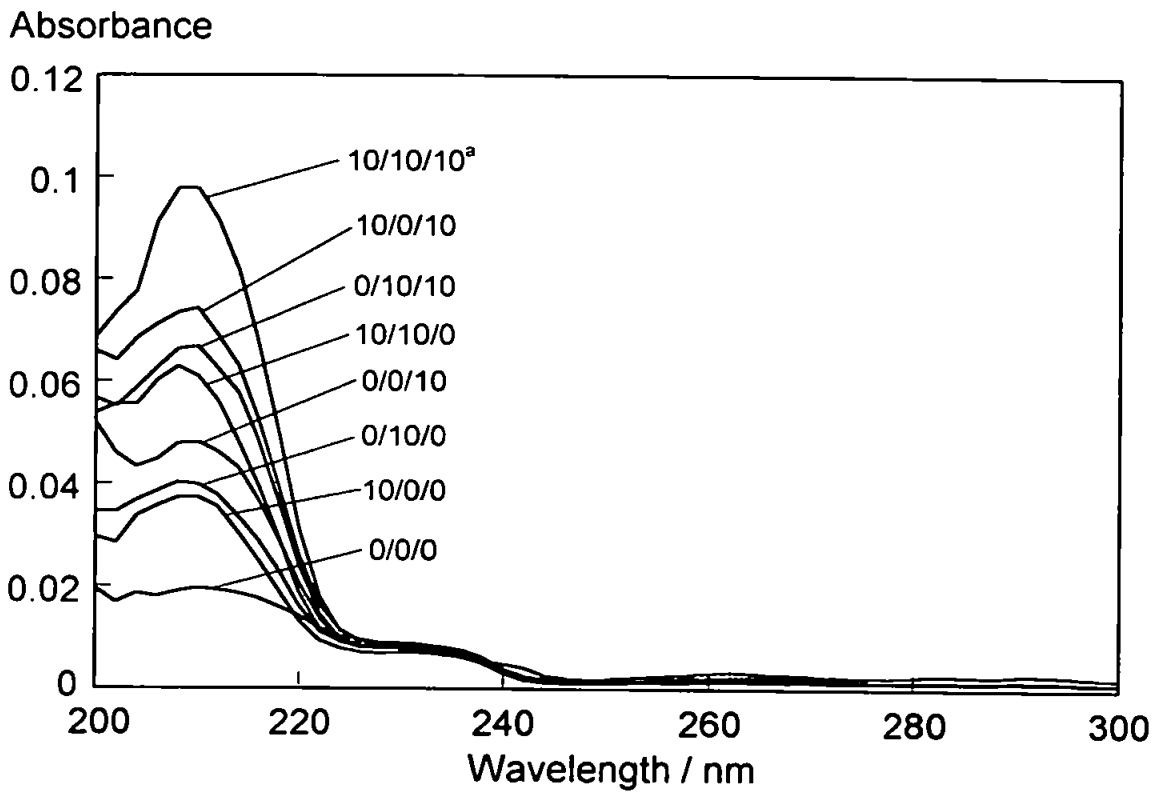


Table 5.5 Linear regression results for absorbances at 210 nm of extracted *o*-xylene standard solutions and *o*-xylene-spiked effluent solutions.

Sample type	R^2	Slope (conc./abs. units)	*y*-Intercept (absorbance units)
Standard solutions	0.999	0.004	0.004
Spiked effluent solutions	0.997	0.003	0.018

Figure 5.16 UV spectra for three-component aqueous TEX solutions following on-line FI extraction.



[a] Concentration (mg l^{-1}) of toluene/ethylbenzene/*o*-xylene.

5.4 CONCLUSIONS

Predictions of individual analyte concentrations in synthetic mixtures (for solutions in both hexane and water) produced the lowest errors for *o*-xylene in the case of all multivariate calibration techniques. The use of first- and second-derivative data produced better precisions than for those when using absorbance data in some (but not all) cases, while the two PLS calibration techniques tended to offer the most robust calibration models. SMLR sometimes produced lower prediction errors than PLS1 and PLS2, but tended to be less consistent than the latter techniques.

The lowest prediction errors were produced when calibrating in terms of total TEX compounds. Quantification of total TEX compounds gave acceptable precisions for the requirements of on-line effluent monitoring when using absorbance data.

Monitoring of real effluent samples may require solvent extraction in order to reduce matrix interferences, and for this purpose an FI manifold incorporating a solvent extraction cell was successfully applied to the analysis of both aqueous *o*-xylene standards and solutions of a real effluent spiked with one or three analytes.

Chapter 6

The Application of Kalman Filtering Methods to Multivariate Calibration and Drift Correction

6.1 INTRODUCTION

The accuracy of an on-line analytical system is often dependent on the robustness of its calibration model, and it has been shown in Chapter 4 that this can vary according to the type of calibration technique being employed and the complexity of the sample system under investigation. Another factor to consider is the stability of instrumental response over time, which in the case of an *in situ* monitor can be affected by fluctuations in ambient temperature (as shown in Chapter 2) and reagent quality. For this reason, on-line analytical systems should be capable of regular, automated recalibration routines to compensate for drifting response signals. In addition, chemometric methods can be employed to determine and correct for drift in instrumental response.

The Kalman filter is a recursive, digital filtering algorithm which can be used for a variety of applications in analytical chemistry, including multivariate calibration and the determination of instrumental response drift. The latter application provides both a method of correcting for drift in a series of calibration spectra, and a means of determining when the precision of a calibration parameter (*e.g.* baseline or sensitivity) falls below a desired level, which can be used to trigger instrument recalibration.

This chapter describes an investigation of the Kalman filter as both a multivariate calibration technique (in comparison to results obtained in Chapter 4), and as a method of determining drift in multicomponent spectral data. The latter application has been discussed in the literature with respect to univariate FI measurements of single-component samples [204 , 205 , 206 , 207], but here the approach is extended to multivariate data and multicomponent samples.

6.2 THEORY OF THE KALMAN FILTER

The Kalman filter was developed by R. E. Kalman in 1960 [208] as a digital filter for processing complex data in electrical engineering applications. In recent years the technique has increasingly been applied to the solution of a number of problems in analytical chemistry, including multicomponent curve resolution, removal of variable background responses, calibration with drift correction and estimation of kinetic parameters [209]. The Kalman filter is a recursive technique in which one data point is processed at a time, with the previous best estimate of the parameter of interest being used to calculate an updated estimate as each new data point is obtained. This approach allows given parameters of an analytical system (*e.g.* concentration) to be estimated in real time from noisy measurements, and is thus potentially well-suited to the requirements of on-line analyses [210].

A number of variations of the Kalman filter algorithm have been developed for different applications. The simplest form of the algorithm is the original filter (hereafter referred to as the standard Kalman filter), and is used in the estimation of linear parameters [209, 211 , 212]. This algorithm is based on two linear models, which respectively describe the dynamics of the chemical system under investigation (equation 6.1) and the measurement process itself (equation 6.2):

$$\mathbf{x}(k) = \mathbf{F}(k, k\text{-}1)\bullet\mathbf{x}(k\text{-}1) + \mathbf{w}(k) \qquad (6.1)$$

$$z(k) = \mathbf{h}^{T}(k)\bullet\mathbf{x}(k) + v(k) \qquad (6.2)$$

In the system model, $\mathbf{x}(k)$ is an $L \times 1$ vector representing the best estimates of the system parameters of interest (*e.g.* the slope and intercept for a univariate linear calibration or analyte concentrations for a multicomponent data set) after k measurements have been obtained (L is the total number of parameters being estimated). $\mathbf{F}(k, k\text{-}1)$ is the $L \times L$ system transition matrix, which describes how $\mathbf{x}(k)$ changes from time t_{k-1} to time t_k, while the $L \times 1$ vector $\mathbf{w}(k)$ describes the noise contribution to the system parameters. In the case

of the measurement model, $z(k)$ is the instrumental measurement performed at time t_k, while $\mathbf{h}^T(k)$ is the $1 \times L$ measurement function vector, describing the relationship between the measurement and the system parameters at time t_k. $v(k)$ represents measurement noise.

The standard Kalman filter algorithm consists of the following five equations:

(i) System parameter extrapolation:

$$\mathbf{x}(k \mid k-1) = \mathbf{F}(k,\ k-1) \bullet \mathbf{x}(k-1 \mid k-1) \tag{6.3}$$

where $\mathbf{x}(k \mid k-1)$ is the best estimate for $\mathbf{x}(k)$ based on all measurements up to and including $z(k-1)$, and $\mathbf{x}(k-1 \mid k-1)$ is the previous estimate of $\mathbf{x}$.

(ii) Covariance extrapolation:

$$\mathbf{P}(k \mid k-1) = \mathbf{F}(k,\ k-1) \bullet \mathbf{P}(k-1 \mid k-1) \bullet \mathbf{F}^T(k,\ k-1) + \mathbf{Q}(k) \tag{6.4}$$

where $\mathbf{P}(k \mid k-1)$ is the best estimate for the $L \times L$ system parameter covariance matrix and $\mathbf{Q}(k)$ is the $L \times L$ system noise covariance matrix.

(iii) Kalman gain:

$$\mathbf{k}(k) = \frac{\mathbf{P}(k \mid k-1) \bullet \mathbf{h}(k)}{\mathbf{h}^T(k) \bullet \mathbf{P}(k \mid k-1) \bullet \mathbf{h}(k) + R(k)} \tag{6.5}$$

where $\mathbf{k}(k)$ is the $L \times 1$ vector describing the Kalman gain (a weighting factor for the next processed measurement) and $R(k)$ is the measurement noise variance.

(iv) System parameter update:

$$\mathbf{x}(k \mid k) = \mathbf{x}(k \mid k-1) + \mathbf{k}(k) \bullet \left[z(k) - \mathbf{h}^T(k) \bullet \mathbf{x}(k \mid k-1) \right] \tag{6.6}$$

where $\mathbf{x}(k \mid k)$ is the updated estimate of $\mathbf{x}(k)$ based on all measurements up to and including $z(k)$.

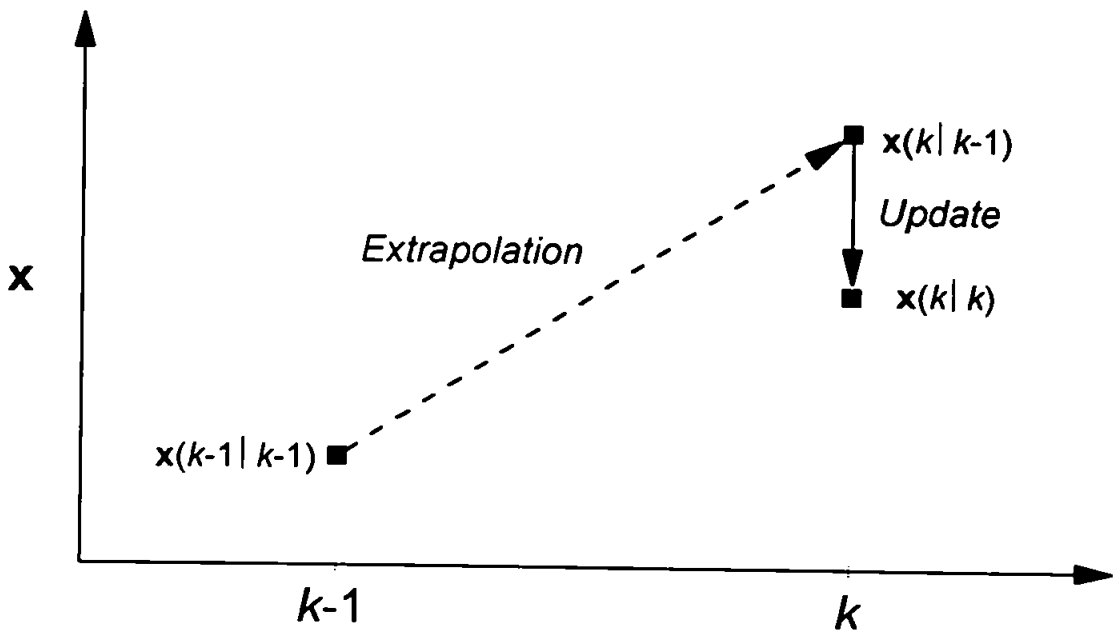
(v) Covariance update:

$$\mathbf{P}(k \mid k) = \left[\mathbf{I} - \mathbf{k}(k) \bullet \mathbf{h}^T(k) \right] \bullet \mathbf{P}(k \mid k-1) \tag{6.7}$$

where $\mathbf{P}(k \mid k)$ is the updated estimate of system parameter covariance.

It can therefore be seen that the algorithm is comprised of two extrapolation equations (6.3 and 6.4) which predict values for the system parameter variables and covariances at data point k based on measurements obtained up to point k-1, an equation for the calculation of Kalman gain (6.5), and a final pair of equations (6.6 and 6.7) which calculate updated values for the parameter variables and covariances based on all measurements up to and including point k. This process can be represented graphically as shown in Figure 6.1.

Figure 6.1 Graphic representation of the Kalman filter algorithm [210].



Owing to the recursive nature of the Kalman filter, initial guesses for the system parameter variables $\mathbf{x}(0 \mid 0)$ and covariances $\mathbf{P}(0 \mid 0)$ are required in equations 6.3 and 6.4 to begin the algorithm. Zeros are often used for $\mathbf{x}(0 \mid 0)$, while an $L \times L$ identity matrix with diagonal values of 10^4 (or a similarly large number) is used for $\mathbf{P}(0 \mid 0)$. The eventual

results are not dependent on the initial guesses, providing $\mathbf{P}(0 \mid 0)$ is sufficiently large to prevent biased parameter estimates. Initial estimates are also required for $\mathbf{F}(1, 0)$ (typically an $L \times L$ identity matrix), $\mathbf{Q}(1)$ (generally assumed to be zero) and $R(1)$ (typically $< 10^{-6}$ for spectroscopic data). The contents of $\mathbf{h}^T(1)$ will be dependent on the application (*e.g.* pure component spectra in the case of multicomponent spectral resolution, and component concentrations in the case of calibration drift correction).

This standard Kalman filter is approximately equivalent to classical least-squares regression, but has the advantages of greater model flexibility and real-time data processing. Accurate results are produced in cases where the model is fully characterised, *i.e.* accurate information is available for $\mathbf{F}(k, k\text{-}1)$, $\mathbf{Q}(k)$, $\mathbf{h}^T(k)$ and $R(k)$. However, if the model is incomplete (*e.g.* an unknown component is contributing to the response signal), then data points inconsistent with the model will adversely affect the calculation of the updated parameter estimates. For this reason another form of the algorithm, referred to as the adaptive Kalman filter, has been developed.

The adaptive Kalman filter [213] recursively calculates the measurement variance $R(k)$ during data processing. Data points which are inconsistent with the model information are then attributed to random noise by artificially increasing their $R(k)$ values. This recursive estimation is performed according to:

$$R(k) = \frac{1}{W}\left(\sum_{w=1}^{W} v(k-w) \bullet v(k-w)\right) - \mathbf{h}^T(k) \bullet \mathbf{P}(k \mid k-1) \bullet \mathbf{h}(k) \qquad (6.8)$$

where W is the total number of points used for a smoothing window and $v(k)$ is the innovations sequence shown below:

$$(k) = z(k) - \mathbf{h}^T(k) \bullet \mathbf{x}(k \mid k-1) \qquad (6.9)$$

The innovations sequence is the difference between the actual and the predicted measured response (*i.e.* the on-line residuals), and is an indicator of model errors. Since the Kalman gain factor is inversely proportional to measurement noise, data points with a large value of

$R(k)$ will receive a low weighting in the update calculation. The limitation of the adaptive Kalman filter is that it requires model information to be accurate for at least some of the processed data points. In addition, the adaptive filter is sensitive to the initial guesses for parameter variables and covariances in cases where model errors affect the first few processed data points. This problem can be circumvented by the use of a simplex optimisation procedure [214], which generates initial guess values for which the error of each system parameter is minimised, thereby permitting the estimation of system parameters using the maximum amount of data consistent with the model.

6.3 EXPERIMENTAL

The data used for this work were those originally obtained for the work detailed in Chapter 4 (*i.e.* visible absorbance spectra for multicomponent mixtures of transition metal salts). Details of reagents, instrumentation and experimental design are therefore not repeated here.

6.3.1 Software

All Kalman filter data analysis was performed using programs written within the Matlab environment (Matlab for Windows version 4.0; Mathworks Inc., Natwick, MA, USA) on Pentium® and 486 personal computers. The programs are listed in Appendices 2-4.

6.3.2 Procedures

Multivariate calibration

The standard Kalman filter (see Appendix 2) was used to resolve sets of multicomponent spectra for two-, three-, four- and five-component mixtures of transition metal salts, in

order to predict the individual analyte concentrations. Separate calibration models were built for each sample, using pure component spectra obtained for 0.015 mol dm^{-3} solutions of the metal salts as the rows of the $L \times M$ measurement function matrix **H**, and each unresolved multicomponent spectrum as the $M \times 1$ measurement vector **z** (where L is the number of analytes and M is the number of wavelengths). An $L \times 1$ column of zeros and an $L \times L$ identity matrix (with values of 10^4 for the diagonal elements) were used as the initial guesses of analyte concentrations (**x**) and covariance (**P**) respectively. A value of 10^{-6} and an $L \times L$ identity matrix were used as the respective initial guesses of measurement variance R and the system transition matrix **F**. The system noise covariance **Q** was assumed to be zero. As described in Chapter 4, the multicomponent systems used for prediction again included those in which inter-analyte interactions and barium sulfate precipitate were present, creating chemical and physical interferences in the absorbance spectra. The precisions of the models were calculated in terms of the relative root-mean-square error of prediction (RRMSEP), and were compared with values obtained in Chapter 4 using DMA and PLS1 multivariate calibration techniques.

In addition, the standard Kalman filter was modified to simultaneously model data for a set of multicomponent calibration standards (as shown in Appendix 3). In this way calibration constants were derived for each analyte in the calibration standards, which were then used to simultaneously predict concentrations in sets of new samples. **H** was formed by an $L \times I$ matrix of analyte concentrations for each sample in the calibration set, while the $I \times M$ matrix **Z** contained their respective absorbance spectra (where I is the number of samples in the calibration set). Initial guesses of **x**, **P**, **F**, R and **Q** were the same as used above. The modified algorithm performed two processing cycles: an inner cycle which recursively processed the calibration samples at a given wavelength point, and an outer cycle which incrementally stepped through the wavelengths. The final updated values of **x** (in this case the regression coefficients relating concentrations to absorbance at each

wavelength) and **P** obtained by each of the inner cycles were stored and recalled when the outer cycle advanced to the next wavelength. When all M wavelengths had been processed, the values of **x** determined during each inner cycle were used to form the $L \times M$ calibration constants matrix β. These calibration constants were equivalent to the pure spectra of each component. Analyte concentrations in new samples were estimated by multiplying the new absorbance spectra matrix by the pseudoinverse of β, as shown in equation 6.10:

$$\hat{\mathbf{C}}_{new} = \mathbf{A}_{new}(\beta^T\beta)^{-1}\beta^T \qquad (6.10)$$

where $\hat{\mathbf{C}}_{new}$ is the $I \times L$ matrix of estimated concentrations and $\mathbf{A}_{new}$ is the $I \times M$ matrix of new absorbance spectra.

Detection of baseline drift:

The Kalman filter program described above for the simultaneous modelling of multiple calibration samples was further modified in order to determine and correct for baseline drift within the calibration set (see Appendix 4). A synthetic baseline drift component was added to spectra for both single- and three-component calibration samples (*i.e.* a given value added to absorbance at all wavelengths for a given sample), in order to represent time-based instrumental drift. Both linear and random baseline drift components were investigated. The **Z** matrix and the initial guess of R were the same as those used for the modified Kalman filter for multivariate calibration. In this case, the **H** matrix was again comprised of the component concentrations in the calibration set, but included a row of ones to represent the offset of each spectrum (*i.e.* the baseline), and a row of zeros to represent the drift component affecting the baseline, *i.e.* an $(L+2) \times I$ matrix, as shown in equation 6.11 (in the form used for the three-component system). An $(L+2) \times 1$ vector of zeros and an $(L+2) \times (L+2)$ identity matrix were used as the initial guesses of **x** and **P** respectively (the additional elements in each case again representing the baseline and its drift component). The initial

guesses of **Q** and **F** were formed as shown in equations 6.12 and 6.13 (as used for the three-component system):

$$\mathbf{H} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,I} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,I} \\ c_{3,1} & c_{3,2} & \cdots & c_{3,I} \\ 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{6.11}$$

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.12}$$

$$\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q \end{bmatrix} \tag{6.13}$$

where c is analyte concentration and q is a scalar in the range 0-1. The arrangement of **Q** is based on the assumption that system noise affects only the drift parameter. This version of the Kalman filter performed two processing cycles similar to those of the modified filter used in multivariate calibration. The final updated values of **x** determined at each wavelength were again used to form the calibration constants matrix β, which in this case also included rows describing the contributions of the baseline and its drift component to the calibration spectra. As described previously in equation 6.10, analyte concentrations in new samples were estimated by multiplying the matrix of new absorbance spectra with the pseudoinverse of β, which automatically compensated for the baseline drift component present in the calibration spectra. Similarly, estimates of calibration error were obtained by multiplying the original **Z** matrix by the pseudoinverse of β, and comparing the resulting predictions of calibration sample concentrations with the actual concentrations. Quantitative estimates for the drift contribution to each spectrum were obtained by storing

the (L+2)th element of each updated **x** vector determined recursively during the final inner cycle of processing (*i.e.* values for all I samples obtained for the final Mth wavelength).

6.4 RESULTS AND DISCUSSION

6.4.1 Multivariate calibration of multicomponent transition metal mixtures

Predictive precisions are quoted here for both the standard and the modified Kalman filter algorithms, and for those obtained in Chapter 4 using DMA and PLS1.

Table 6.1 lists RRMSEP values for the three-component system, in which no chemical or physical interferences were present. In Chapter 4 it was shown that no significant differences were evident in the overall precessions of the five multivariate calibration techniques used for this system. This pattern was repeated here for the two Kalman filter methods, which gave very similar prediction errors to both DMA and PLS1 for the three components. In the case of the three-component system in which barium sulfate was present (Table 6.2), a deterioration was apparent in the overall precisions of both Kalman filtering techniques, as was the case for DMA and PLS1. The standard Kalman filter gave results very similar to DMA, which reflects the conceptual similarity between the two methods of calibration, in that both models assume the Beer's law relationship between instrumental response and component concentrations, and require prior knowledge of pure component spectra. For this system, the standard Kalman filter model included a pure spectrum for barium sulfate precipitate in 1% v/v sulfuric acid solution in the **H** matrix. Predictive errors obtained by the modified Kalman filter were considerably higher with respect to cobalt and nickel when only the concentrations of the three metal salts were included in the **H** matrix. While it was not possible to accurately quantify the barium sulfate component in **H** (original additions of barium chloride to the

samples were non-quantitative, as described in Chapter 4), it was discovered that predictive precisions could be significantly improved by adding a row of arbitrary values to **H** to represent the concentration of barium sulfate. The results given in Table 6.2 were obtained using a row of barium sulfate 'concentrations' equal to the mean concentration of the metal salts in the calibration set (0.015 mol dm^{-3}), although no significant differences were observed when other values were used. This enabled the modified filter to implicitly model the effect of barium sulfate interference, and the precisions obtained were a significant improvement on those of the standard filter.

Table 6.1 Precisions for the three-component system.

Calibration method	RRMSEP (%)			
	Co	Cu	Ni	Overall
Standard KF	2.5	0.93	3.4	2.5
Modified KF	2.8	0.52	3.6	2.7
DMA	2.5	0.84	3.4	2.5
PLS1	2.8	0.52	3.6	2.6

Table 6.2 Precisions for the three-component system with $BaSO_4$ interference.

Calibration method	RRMSEP (%)			
	Co	Cu	Ni	Overall
Standard KF	10.4	5.2	10.2	8.9
Modified KF	22.0[a]	2.8	34.9	23.9
	5.6[b]	3.4	7.5	5.7
DMA	10.2	6.5	11.2	9.5
PLS1	1.4	2.6	15.5	9.1

[a] $BaSO_4$ 'concentrations' not included in **H**.

[b] Including a row of 0.015 values in **H** to represent the 'concentrations' of $BaSO_4$.

Table 6.3 gives results for a four-component system, in which again no significant interferences were present. A similar trend was observed as found in the case of the three-component system, with no significant difference between the two Kalman filters, DMA and PLS1, and very good precisions obtained for the four metal salts. When barium sulfate precipitate was present (Table 6.4), both versions of the Kalman filter produced errors considerably higher than those of PLS1 but generally similar to those of DMA. In this case, the overall precision of the modified Kalman filter was not significantly different to that of the standard algorithm, since the higher degree of spectral overlap present in the four-component system adversely affected the ability of the modified filter to model barium sulfate interference.

Table 6.3 Precisions for the four-component system.

Calibration method	RRMSEP (%)				
	Cr	Ni	Co	Cu	Overall
Standard KF	0.77	3.0	2.6	0.96	2.1
Modified KF	0.92	2.6	2.1	0.97	1.8
DMA	0.77	2.8	2.5	0.99	2.0
PLS1	0.91	2.6	2.1	1.0	1.8

Table 6.4 Precisions for the four-component system with $BaSO_4$ interference.

Calibration method	RRMSEP (%)				
	Cr	Ni	Co	Cu	Overall
Standard KF	11.5	13.2	30.6	23.5	21.2
Modified KF	11.6[a]	63.9	59.1	22.5	46.3
	10.4[b]	5.6	24.9	24.5	18.4
DMA	19.0	5.9	5.7	20.4	14.5
PLS1	0.65	1.4	2.6	2.1	1.8

[a] $BaSO_4$ 'concentrations' not included in **H**.

[b] Including a row of 0.015 values in **H** to represent the 'concentrations' of $BaSO_4$.

Table 6.5 summarises prediction errors for the four techniques when calibrating and predicting samples in which a strong chemical interaction was present between the two components (partial oxidation of Fe^{2+} in the presence of Cu^{2+}, as described in Chapter 4). Predictions for copper were in all cases more precise than for iron as a result of the partial oxidation effect on the iron spectrum. The standard Kalman filter again produced results almost identical to DMA, as expected, but the modified filter offered a very significant improvement in precision, equivalent to that of PLS1 (all prediction errors <10%).

Table 6.5 Precisions for the two-component system.

Calibration method	RRMSEP (%)		
	Cu	Fe	Overall
Standard KF	3.1	33.1	23.5
Modified KF	1.5	9.9	7.1
DMA	3.0	32.9	23.4
PLS1	1.5	8.9	6.4

Predictive precisions for a five-component system (which also incorporated the iron-copper interaction) are listed in Table 6.6. As before, the similarity between the results for the standard Kalman filter and DMA was very evident, and the highest prediction errors were produced for iron. The modified filter was again better able to model the chemical interference effect than the standard filter, and produced significantly better precisions for all components. In this case it was unable to match the overall precision of PLS1 however. Table 6.7 gives prediction errors for the five-component system with barium sulfate precipitate. This was the most difficult system to model, with a high degree of spectral overlap and the presence of both physical and chemical interferences. The standard filter was unable to accurately predict the components of this system, and could not match the precision of DMA with respect to all metals other than the partially oxidised iron. A very

Table 6.6 Precisions for the five-component system.

Calibration method	RRMSEP (%)					
	Fe	Cu	Cr	Co	Ni	Overall
Standard KF	123	8.6	7.0	9.6	13.5	55.8
Modified KF	44.2	4.2	4.6	3.0	10.1	20.5
DMA	129	7.6	7.9	9.3	13.8	58.2
PLS1	15.2	2.1	3.9	3.6	5.2	7.6

Table 6.7 Precisions for the five-component system with $BaSO_4$ interference.

Calibration method	RRMSEP (%)					
	Fe	Cu	Cr	Co	Ni	Overall
Standard KF	108	28.2	113	113	105	99.1
Modified KF	14.8[a]	44.5	38.9	61.6	94.5	57.4
	14.8[b]	43.5	36.9	83.6	53.6	51.6
DMA	435	13.5	2.9	23.8	27.1	185
PLS1	7.6	8.9	1.8	16.1	25.1	14.4

[a] $BaSO_4$ 'concentrations' not included in **H**.

[b] Including a row of 0.015 values in **H** to represent the 'concentrations' of $BaSO_4$.

badly characterised Kalman filter model was produced as a result of the combined chemical and physical interferences, which appeared to have a more adverse effect on the recursive calibration process than on the non-recursive DMA technique. The modified filter was able to offer a considerably better precision for iron and a significant improvement in the precisions of chromium, cobalt and nickel. However, its prediction errors were still significantly higher than those of DMA and PLS1 for all metals other than iron, which again indicated the limitation of the recursive process when attempting to model a very badly characterised system. The prediction error for iron was lower than that of the other metals for the modified filter since it produced the largest absorbance peaks in the set of calibration

spectra, which were least affected by the high background absorbance of barium sulfate, and therefore had the strongest influence on the recursive modelling process.

6.4.2 Detection of baseline drift in multicomponent calibration spectra

The results for a set of single-component copper sulfate pentahydrate solutions with linear baseline drift are summarised in Table 6.8. The table lists the actual synthetic drift component added to each spectrum, and also the estimated drift component determined by the Kalman filter routine, using values of q between 1 and 0. Estimates of linear drift were more accurate for smaller values of q, since in this case the added drift component was completely linear (*i.e.* no random element is present in the drift, a situation best described by $q = 0$). The estimated drift values for the first and second samples were less accurate than for the subsequent samples, which reflected the recursive nature of the filtering technique. Baseline drift was calculated relative to the previous sample, therefore the estimate for the first sample had no meaning. The filter was able to produce accurate estimates after only three samples had been processed however. The effect of drift correction on the accuracy of the calibration models is shown in Figure 6.2, which shows that relative root-mean-square errors of calibration (RRMSEC: the difference between actual concentrations in the calibration set and those estimated by the drift-corrected Kalman model) decreased from 0.67% for $q = 1$ to 1.79×10^{-5}% for $q = 0$. This compared very favourably with the RRMSEC value of 24.8% obtained when no drift correction procedure was employed (*i.e.* using the modified Kalman filter described in Section 6.3.1, with $q = 0$).

Table 6.9 summarises results for the same single-analyte system with a random baseline drift component. In this case, calibration models with values of $q \geq 0.1$ appeared to offer more accurate estimations of the incremental drift component, although the RRMSEC values shown in Figure 6.3 indicate that the lowest calibration errors ($< 2.1\times10^{-5}$%) were produced for models with $q \leq 1 \times 10^{-5}$. All the drift-corrected models exhibited very low

Table 6.8 One-component system with linear baseline drift component.

Cu conc. (mol dm^{-3})	Added drift[a]	Baseline drift determined by Kalman filter				
		q=1	q=0.1	q=1×10^{-5}	q=1×10^{-10}	q=0
0.015	0	0.09	0.09	0.09	0.09	0.09
0.045	0.01	0.37	0.37	0.37	0.37	0.37
0.04	0.02	0.01	0.01	0.01	0.01	0.01
0.025	0.03	0.00	0.01	0.01	0.01	0.01
0.03	0.04	0.01	0.01	0.01	0.01	0.01
0.02	0.05	0.00	0.01	0.01	0.01	0.01
0.05	0.06	0.02	0.01	0.01	0.01	0.01
0.01	0.07	0.00	0.01	0.01	0.01	0.01
0.035	0.08	0.01	0.01	0.01	0.01	0.01
0.005	0.09	0.01	0.01	0.01	0.01	0.01
0	0.1	0.01	0.01	0.01	0.01	0.01

[a] Cumulative drift component

Figure 6.2 RRMSEC for drift-corrected one-component calibration models (linear drift).

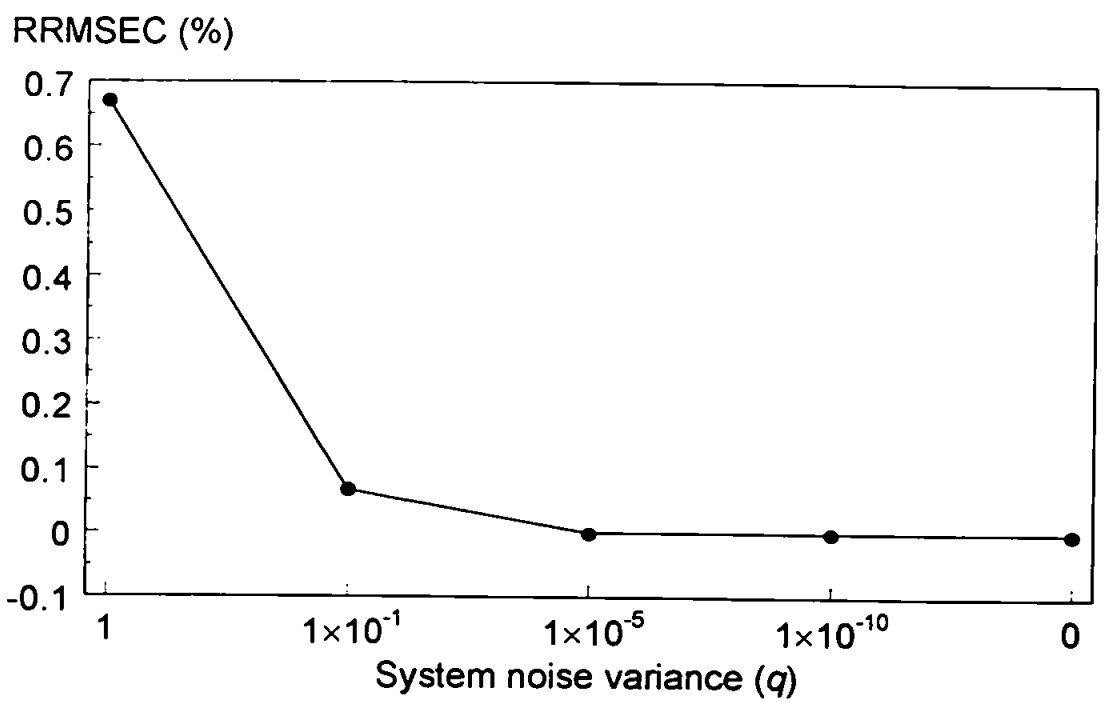
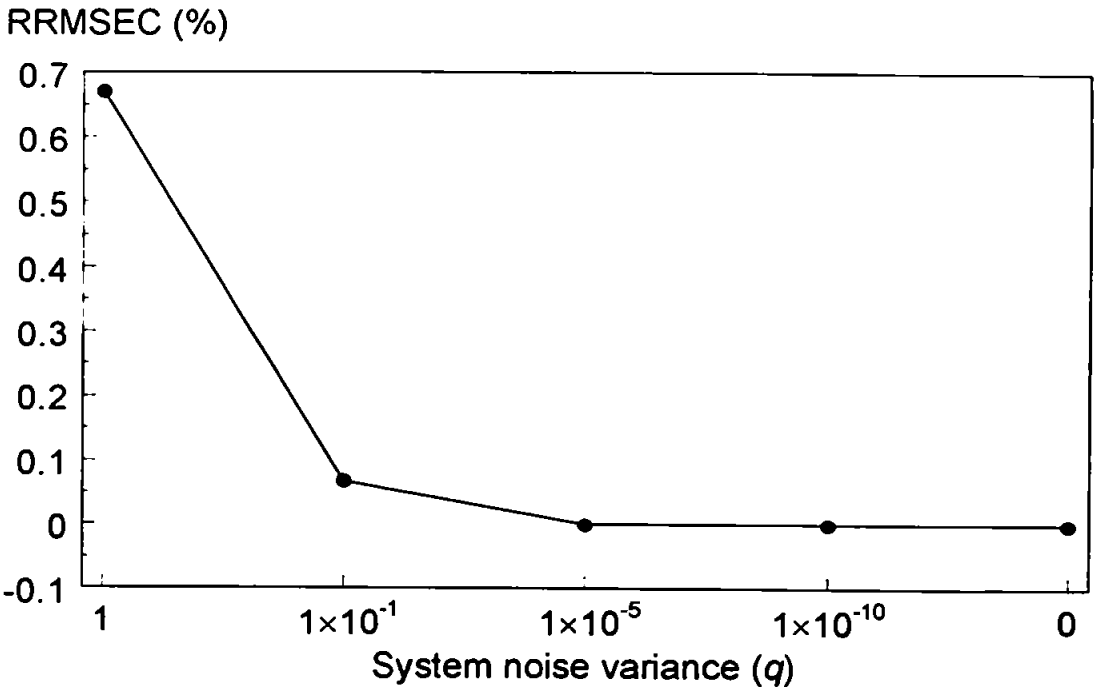


Table 6.9 One-component system with random baseline drift component.

Cu conc.	Added drift		Baseline drift determined by Kalman filter				
(mol dm^{-3})	Cum.[a]	Incr.[b]	q=1	q=0.1	q=1×10^{-5}	q=1×10^{-10}	q=0
0.015	0	0	0.09	0.09	0.09	0.09	0.09
0.045	0.07	0.07	0.43	0.43	0.43	0.43	0.43
0.04	0.1	0.03	0.03	0.04	0.04	0.04	0.04
0.025	0.15	0.05	0.05	0.06	0.06	0.05	0.05
0.03	0.21	0.06	0.06	0.06	0.06	0.05	0.05
0.02	0.28	0.07	0.07	0.08	0.07	0.05	0.05
0.05	0.32	0.04	0.05	0.04	0.04	0.05	0.05
0.01	0.42	0.1	0.08	0.08	0.08	0.06	0.06
0.035	0.43	0.01	0.03	0.03	0.04	0.06	0.06
0.005	0.5	0.07	0.04	0.04	0.04	0.06	0.06
0	0.55	0.05	0.05	0.05	0.05	0.05	0.05

[a] Cumulative drift component
[b] Incremental drift component

Figure 6.3 RRMSEC for drift-corrected one-component calibration models (random drift).



calibration errors however, and were again significantly more precise than calibration models with no drift correction (RRMSEC = 99.0% when $q = 0$).

Results for baseline drift determination in multicomponent calibration samples are given in Table 6.10, which summarises actual and calculated linear drift in a three-component set. The trend was found to be very similar to that observed in the single-analyte set, with values of $q \leq 1\times10^{-5}$ again producing the most accurate estimations of incremental drift, and (as shown in Figure 6.4) the lowest calibration errors. RRMSEC values for $q = 0$ were in the range 0.44-1.0%, which was again a very significant improvement on calibration errors obtained when no drift correction was employed (4.8-53.3%).

When random drift was added to the three-component system (see Table 6.11), calibration models with $q \geq 1\times10^{-5}$ appeared to offer slightly better estimations of the incremental drift, as was discovered for the single-component system. Once again however, calibration errors decreased as q approached zero, as shown in Figure 6.5. RRMSEC values for $q = 1$ were in the range 1.7-46.1%, and for $q = 0$ were in the range 0.45-1.1%, indicating that non-zero values of q were having an increasingly detrimental effect on the precision of the calibration models. Drift correction was again able to produce considerably more precise calibration models, with RRMSEC values in the range 19.7-92.9% when no drift correction was employed.

As a final test of the improved calibration precision offered by the drift-correction procedure, a calibration model was built using a set of 24 three-component spectra (comprised of 3 replicates of 8 samples) to which random baseline drift had been added. The drift-corrected calibration model was then used to predict analyte concentrations in a set of 19 new samples (the three-component test set described in Chapter 4) to which no baseline drift had been added. Table 6.12 lists the actual and calculated drift component in the calibration set, while Table 6.13 summarises the calibration and prediction errors

Table 6.10 Three-component system with linear baseline drift component.

Conc. (mol dm^{-3})			Added	Baseline drift determined by Kalman filter				
Co	Cu	Ni	drift[a]	q=1	q=10^{-1}	q=10^{-5}	q=10^{-10}	q=0
0.005	0.005	0.005	0	0.03	0.03	0.03	0.03	0.03
0.025	0.005	0.005	0.01	0.01	0.01	0.01	0.01	0.01
0.015	0.005	0.005	0.02	0.01	0.01	0.01	0.01	0.01
0.005	0.025	0.005	0.03	0.06	0.02	0.01	0.01	0.01
0.025	0.025	0.005	0.04	0.02	0.01	0.01	0.01	0.01
0.015	0.025	0.005	0.05	0.01	0.01	0.01	0.01	0.01
0.005	0.015	0.005	0.06	-0.01	0.01	0.01	0.01	0.01
0.025	0.015	0.005	0.07	0.01	0.01	0.01	0.01	0.01
0.015	0.015	0.005	0.08	0.01	0.01	0.01	0.01	0.01
0.005	0.005	0.025	0.09	0.01	0.01	0.01	0.01	0.01
0.025	0.005	0.025	0.1	0.01	0.01	0.01	0.01	0.01
0.015	0.005	0.025	0.11	0.01	0.01	0.01	0.01	0.01
0.005	0.025	0.025	0.12	0.03	0.01	0.01	0.01	0.01
0.025	0.025	0.025	0.13	0.01	0.01	0.01	0.01	0.01
0.015	0.025	0.025	0.14	0.01	0.01	0.01	0.01	0.01
0.005	0.015	0.025	0.15	0.00	0.01	0.01	0.01	0.01
0.025	0.015	0.025	0.16	0.01	0.01	0.01	0.01	0.01
0.015	0.015	0.025	0.17	0.01	0.01	0.01	0.01	0.01
0.005	0.005	0.015	0.18	0.00	0.01	0.01	0.01	0.01
0.025	0.005	0.015	0.19	0.01	0.01	0.01	0.01	0.01
0.015	0.005	0.015	0.2	0.01	0.01	0.01	0.01	0.01
0.005	0.025	0.015	0.21	0.02	0.01	0.01	0.01	0.01
0.025	0.025	0.015	0.22	0.01	0.01	0.01	0.01	0.01
0.015	0.025	0.015	0.23	0.01	0.01	0.01	0.01	0.01
0.005	0.015	0.015	0.24	0.00	0.01	0.01	0.01	0.01
0.025	0.015	0.015	0.25	0.01	0.01	0.01	0.01	0.01
0.015	0.015	0.015	0.26	0.01	0.01	0.01	0.01	0.01

[a] Cumulative drift component

Table 6.11 Three-component system with random baseline drift component.

Conc. (mol dm^{-3})			Added drift		Drift determined by Kalman filter				
Co	Cu	Ni	Cum.[a]	Incr.[b]	q=1	q=10^{-1}	q=10^{-5}	q=10^{-10}	q=0
0.005	0.005	0.005	0.00	0	0.03	0.03	0.03	0.03	0.03
0.025	0.005	0.005	0.02	0.02	0.02	0.02	0.02	0.02	0.02
0.015	0.005	0.005	0.01	-0.01	0.00	0.00	0.00	0.00	0.00
0.005	0.025	0.005	0.02	0.01	0.05	0.01	0.00	0.00	0.00
0.025	0.025	0.005	0.05	0.03	0.02	0.01	0.01	0.00	0.00
0.015	0.025	0.005	0.10	0.05	0.05	0.05	0.05	0.02	0.02
0.005	0.015	0.005	0.12	0.02	0.01	0.02	0.03	0.02	0.02
0.025	0.015	0.005	0.09	-0.02	-0.02	-0.02	-0.01	0.02	0.02
0.015	0.015	0.005	0.12	0.03	0.02	0.02	0.01	0.02	0.02
0.005	0.005	0.025	0.09	-0.03	0.01	0.02	0.01	0.02	0.02
0.025	0.005	0.025	0.04	-0.05	-0.03	-0.04	-0.03	0.02	0.02
0.015	0.005	0.025	0.04	0.00	-0.01	-0.01	-0.01	0.02	0.02
0.005	0.025	0.025	-0.01	-0.04	-0.02	-0.03	-0.03	0.02	0.02
0.025	0.025	0.025	-0.04	-0.03	-0.02	-0.02	-0.02	0.02	0.02
0.015	0.025	0.025	0.00	0.04	0.03	0.03	0.02	0.02	0.02
0.005	0.015	0.025	0.03	0.03	0.00	0.01	0.01	0.01	0.01
0.025	0.015	0.025	0.05	0.02	0.04	0.04	0.03	0.01	0.01
0.015	0.015	0.025	0.05	-0.01	-0.01	-0.01	0.00	0.01	0.01
0.005	0.005	0.015	0.00	-0.05	-0.08	-0.07	-0.06	0.00	0.00
0.025	0.005	0.015	0.01	0.01	0.01	0.01	0.00	0.00	0.00
0.015	0.005	0.015	0.01	0.00	0.00	0.00	0.00	0.00	0.00
0.005	0.025	0.015	-0.03	-0.04	-0.03	-0.04	-0.03	0.00	0.00
0.025	0.025	0.015	-0.03	0.00	0.01	0.01	0.00	0.00	0.00
0.015	0.025	0.015	-0.04	-0.01	-0.01	-0.01	-0.01	0.00	0.00
0.005	0.015	0.015	-0.07	-0.03	-0.04	-0.04	-0.03	0.00	0.00
0.025	0.015	0.015	-0.08	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
0.015	0.015	0.015	-0.07	0.01	0.01	0.01	0.01	-0.01	0.00

[a] Cumulative drift component
[b] Incremental drift component

Figure 6.4 RRMSEC for drift-corrected three-component calibration models (linear drift). —●— represents Co; —●— represents Cu; - -⊖- - represents Ni.

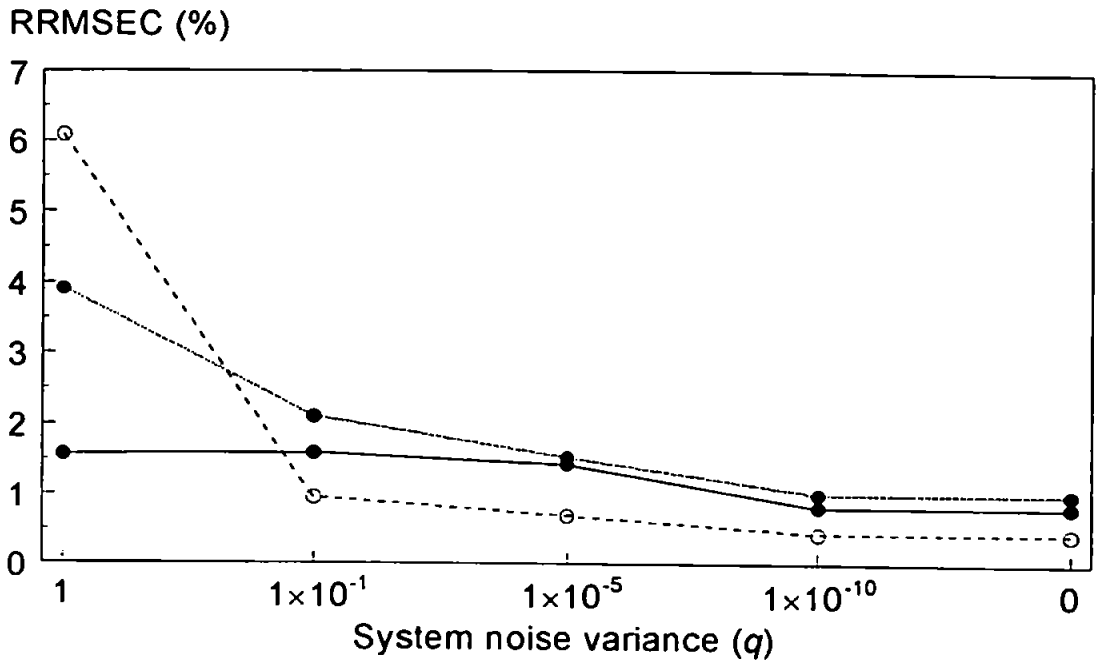


Figure 6.5 RRMSEC for drift-corrected three-component calibration models (random drift; values for q = 1 omitted for clarity). —●— represents Co; —●— represents Cu; - -⊖- - represents Ni.

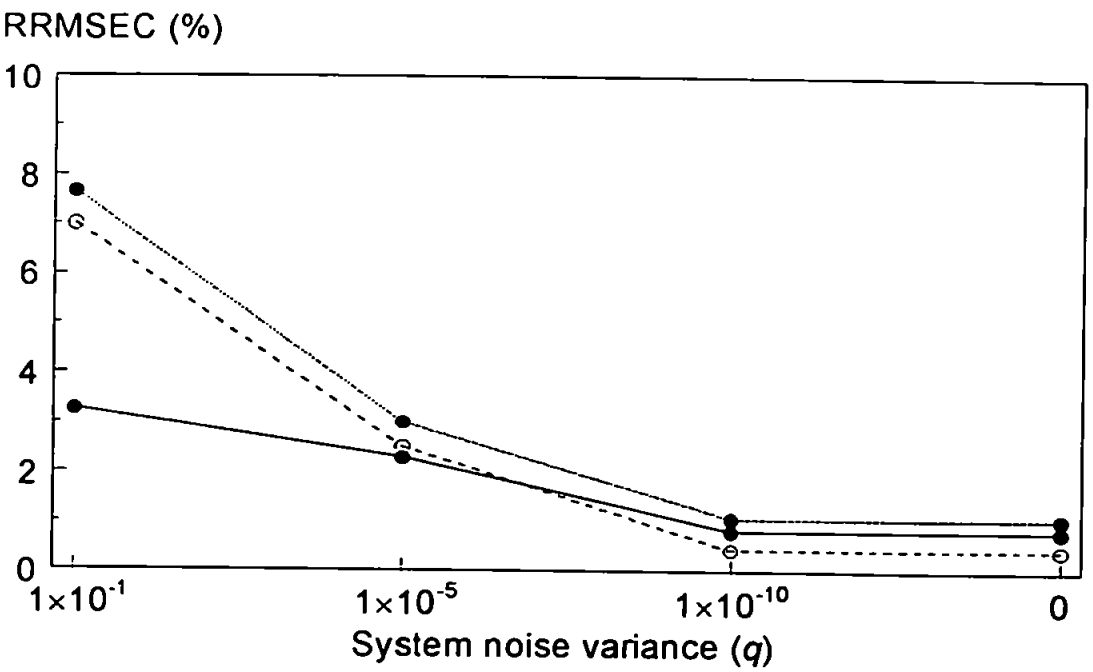


Table 6.12 Three-component training set (including replicates) with random baseline drift component.

Conc. (mol dm^{-3})			Added	Baseline drift
Co	Cu	Ni	drift[a]	(Kalman filter)[b]
0.005	0.005	0.005	0.00	0.03
0.005	0.005	0.005	0.02	0.02
0.005	0.005	0.005	0.01	0.01
0.025	0.005	0.005	0.02	0.01
0.025	0.005	0.005	0.05	0.01
0.025	0.005	0.005	0.10	0.02
0.005	0.025	0.005	0.12	0.02
0.005	0.025	0.005	0.09	0.02
0.005	0.025	0.005	0.12	0.02
0.025	0.025	0.005	0.09	0.02
0.025	0.025	0.005	0.04	0.01
0.025	0.025	0.005	0.04	0.01
0.005	0.005	0.025	-0.01	0.01
0.005	0.005	0.025	-0.04	0.00
0.005	0.005	0.025	0.00	0.00
0.025	0.005	0.025	0.03	0.00
0.025	0.005	0.025	0.05	0.00
0.025	0.005	0.025	0.05	0.01
0.005	0.025	0.025	0.00	0.01
0.005	0.025	0.025	0.01	0.01
0.005	0.025	0.025	0.01	0.01
0.025	0.025	0.025	-0.03	0.01
0.025	0.025	0.025	-0.03	0.01
0.025	0.025	0.025	-0.04	0.00

[a] Cumulative drift component.
[b] $q = 0$.

obtained using both a drift-corrected model and a model without drift correction. Very low calibration errors (all < 1%) were again obtained using the drift-corrected Kalman model, while those produced using the modified Kalman filter with no drift correction were all very significantly higher (> 19%). The trend was repeated for predictions of new sample

concentrations, with RRMSEP values all < 1% for the drift-corrected model and all > 12% for the model with no drift correction. These results demonstrate that the drift-correction filter was capable of producing precise multicomponent calibration models in which the effect of baseline drift was successfully compensated for, and that very significant improvements in predictive performance were obtained when compared with the modified Kalman filter used in Section 6.3.1.

Table 6.13 RRMSEC values for a three-component calibration model with random drift and RRMSEP values for predictions of new sample concentrations.

Analyte	RRMSEC (%)[a]		RRMSEP (%)	
	With drift correction	Without drift correction	With drift correction	Without drift correction
Cobalt	0.80	67.8	0.26	43.3
Copper	0.85	21.0	0.74	13.9
Nickel	0.62	19.7	0.49	12.3

[a] $q = 0$.

6.5 CONCLUSIONS

When applied to multivariate calibration of multicomponent absorbance spectra, the standard Kalman filter tended to produce predictive precisions very similar to those of DMA, owing to the conceptual similarity of the two methods of calibration. Improved prediction errors were often obtained by using the modified Kalman filter algorithm, particularly when chemical interferences were present in the absorbance spectra. The most precise predictions of analyte concentrations were obtained for well-characterised systems, in which no unmodelled interferences were present. However, the precisions of both

recursive techniques were adversely affected by the presence of non-quantified amounts of barium sulfate precipitate, which created interference across the full spectral range.

The version of the Kalman filter used to determine and correct for baseline drift in calibration sample sets was able to produce very precise calibration models for both single- and three-component systems. The calibration errors obtained were much lower than those obtained using models with no drift correction, and were all less than 1% when a value of zero was used for the system noise variance. A drift-corrected calibration model was also shown to produce a highly significant improvement in predictive precision for new samples when compared with a Kalman filter model which was not corrected for baseline drift.

Chapter 7

Conclusions and Future Work

7.1 FINAL CONCLUSIONS

The following general conclusions can be drawn from the work discussed in the preceding chapters:

1 Flow injection was shown to be a suitable technique for the on-line monitoring of a single analyte in industrial liquid effluent streams. An automated, portable FI monitor was successfully deployed for the on-line determination of wastewater ammonia within real industrial process environments, using a gas diffusion method. The monitor was capable of linear response in the range 1-100 mg l^{-1} NH_3-H with a precision of ± 3.4%, and was tolerant of acidic samples ≥ pH 3. A good correlation was obtained with a standard indophenol blue laboratory reference method.

2 Multivariate calibration techniques enabled the quantification of multicomponent diode array spectrophotometric data obtained for synthetic model systems. These systems represented effluent matrices and contained mixtures of up to five transition metal ions. The five calibration techniques examined (DMA, SMLR, PCR, PLS1 and PLS2) were all capable of a high degree of precision (errors < 5%) when applied to the quantification of simple systems in which no interferences were present. However, SMLR, PCR and the two PLS techniques were shown to be significantly more robust calibration techniques than DMA when physical and/or chemical interferences were incorporated. The use of derivatised spectral data was shown to improve the precisions of multivariate calibration.

3 The combination of flow injection with diode array detection and multivariate calibration was shown to be a potential technique for the simultaneous determination of groups of analytes in effluent streams. An on-line, process version of this system would

offer the advantages of relative simplicity, robustness and low capital and operating costs. SMLR, PCR, PLS1 and PLS2 were able to quantify total concentrations of toluene, ethylbenzene and *o*-xylene in aqueous mixtures, analysed using an automated, single-channel FI manifold with diode array detection. The precisions obtained for total TEX compounds were acceptable for the requirements of on-line effluent monitoring. An FI manifold incorporating a solvent extraction cell was successfully applied to the analysis of both aqueous *o*-xylene standards and solutions of a real effluent spiked with one or three analytes. This method reduced potential matrix interferences, and would therefore be suitable for on-line determinations of total TEX compounds in effluent streams.

4 The Kalman filter was shown to be a technique with potential applications for on-line effluent monitoring. A modified version of the Kalman filter algorithm was used for the multivariate calibration of multicomponent diode array spectral data sets, and was able to provide good precisions (errors < 10%) for a range of multicomponent systems, including those in which chemical interferences were present. Another version of the Kalman filter was able to determine and correct for baseline drift in single- and three-component calibration sets, producing a highly significant improvement in predictive precision when compared with uncorrected calibration models. This ability to correct for response drift over time offers significant benefits for on-line analyses.

7.2 SUGGESTIONS FOR FUTURE WORK

The work described in the preceding chapters could be developed in a number of ways. Possible areas for further investigations are summarised below:

Short term projects

1 Extended on-line operation of the portable ammonia monitor in a process environment, and an investigation of alternative membrane materials and on-line sample dilution (to extend the linear range).

2 Further investigation of interference effects for the FI-diode array method used to quantify BTEX compounds. This would include determining the effect of compounds which may be co-extracted from the effluent matrix into the hexane fraction.

3 The use of a high resolution (0.1 nm) CCD spectrophotometer to obtain derivative spectral data for the multicomponent BTEX systems.

4 Multivariate calibration of multicomponent BTEX data sets obtained using the FI-solvent extraction method.

5 Application of the Kalman filter to drift correction of single- and multiple-analyte calibration data obtained using an automated FI system.

Long term projects

1 Development of a portable, automated FI monitor incorporating a multichannel detection system (*e.g.* a miniaturised diode array spectrophotometer) for on-line monitoring of multiple analytes in effluents.

2 Investigation of the use of non-linear multivariate calibration techniques (*e.g.* artificial neural networks) for the resolution of multi-analyte spectrophotometric data obtained for effluent systems.

3 Incorporation of a Kalman filter algorithm into the operating software of an on-line effluent monitor, to perform either multivariate calibration in real time, or to determine when monitor recalibration is necessary.

4 Investigation of the potential applications of miniaturised FI systems for on-line effluent monitoring.

References

1 Royal Commission on Environmental Pollution, *Sixteenth Report: Freshwater Quality*, HMSO, London, 1992.

2 Haigh, N, *EEC Environmental Policy and Britain*, Longman, Harlow, 2nd rvsd. edn., 1990.

3 *Discharge Consents and Compliance. The NRA 's Approach to Control of Discharges to Water*, National Rivers Authority, Bristol, 1994.

4 Leeson, J. D., *Environmental Law*, Pitman Publishing, London, 1995.

5 Gardiner, J., and Mance, G., *UK Water Quality Standards Arising from EEC Directives. Technical Report TR204*, WRC, Medmenham, 1984.

6 Hunt, D. T. E., and Wilson, A. L., *The Chemical Analysis of Water. General Principles and Techniques*, Royal Society of Chemistry, London, 2nd edn., 1986.

7 Clarke, J. R. P., *Anal. Chim. Acta*, 1986, **190**, 1.

8 Blaser, W. W., Ruhl, H. D., and Bredeweg, R. A., *Amer. Lab.*, 1989, Jan., 69.

9 Reeve, A., *Chem. Ind.* (London), 1993, March, 35.

10 Callis, J. B., Illman, D. L., and Kowalski, B. R., *Anal. Chem.*, 1987, **59**, 624A.

11 Riebe, M. T., and Eustace, D. J., *Anal. Chem.*, 1990, **62**, 65A.

12 Růžička, J., *Fresenius Z. Anal. Chem.*, 1986, **324**, 745.

13 Valcárcel, M., and Luque de Castro, M. D., *Flow Injection Analysis: Principles and Applications*, Ellis Horwood, Chichester, 1987.

14 Růžička. J., and Hansen, E. H., *Flow Injection Analysis*, Wiley-Interscience, New York, 2nd edn., 1988.

15 Růžička, J, and Hansen, E. H., *Anal. Chim. Acta*, 1975, **78**, 145.

16 Trojanowicz, M., Worsfold, P. J., and Clinch, J. R., *Trends Anal. Chem.*, 1988, **7**, 301.

17 MacLaurin, P., Worsfold, P. J., Townshend, A., Barnett, N. W., and Crane, M., *Analyst*, 1991, **116**, 701.

18 Lynch, T. P., Kernoghan, N. J., and Wilson, J. N., *Analyst*, 1984, **109**, 843.

19 Faizullah, A. T., and Townshend, A., *Anal. Chim. Acta*, 1985, **167**, 225.

20 MacLaurin, P., and Worsfold, P. J., *Microchem. J.*, 1992, **45**, 178.

21 Daykin, R. N. C., and Haswell, S. J., *Anal. Chim. Acta*, 1995, **313**, 155.

22 van der Schoot, B. H., Jeanneret, S., van den Berg, A., and de Rooij, N. F., *Anal. Method. Instrumen.*, 1993, **1**, 38.

23 Mowery Jr., R. A., *Instr. Technol.*, 1984, **31**, 51.

24 van der Linden, W. E., *Anal. Chim. Acta*, 1986, **179**, 91.

25 Růžička, J, *Anal. Chim. Acta*, 1986, **190**, 155.

26 Gisin, M. and Thommen, C., *Anal. Chim. Acta*, 1986, **190**, 165.

27 Benson, R. L., Worsfold, P. J., and Sweeting, F. W., *Anal. Proc.*, 1989, **26**, 385.

28 MacLaurin, P., Andrew, K. N., and Worsfold, P. J., in *Process Analytical Chemistry*, eds. McLennan, F., and Kowalski, B. R., Chapman & Hall, London, 1995, ch. 5.

29 Trojanowicz, M., Benson, R. L., and Worsfold, P. J., *Trends Anal. Chem.*, 1991, **10**, 11.

30 Andrew, K. N., Blundell, N. J., Price, D., and Worsfold, P. J., *Anal. Chem.*, 1994, **66**, 916A.

31 Schick, K. G., *ISA Transactions Adv. Instr.*, 1984, **39**, 279.

32 van Staden, J. F., *Fres. Z. Anal. Chem.*, 1987, **328**, 68.

33 Gisin, M., and Thommen, C., *Trends Anal. Chem.*, 1989, **8**, 62

34 MacLaurin, P., Parker, K. S., Townshend, A., Worsfold, P. J., Barnett, N. W., and Crane, M., *Anal. Chim. Acta*, 1990, **238**, 171.

35 López Fernández, J. M., Luque de Castro, M. D., and Valcárcel, M., *J. Automatic Chem.*, 1990, **12**, 263.

36 Barnett, N. W., Rolfe, D. G., Bowser, T. A., and Paton, T. W., *Anal. Chim. Acta*, 1993, **282**, 551.

37 Olson, D. C., Bysouth, S. R., Dasgupta, P. K., and Kuban, V., *Process Qual. Cont.*, 1994, **5**, 259.

38 Cosano, J. S., Izquierdo, A., Luque de Castro, M. D., Valcárcel, M., Aguilar, C., and Penelas, G., *Anal. Chim. Acta*, 1995, **308**, 187.

39 Bergamin, H., Krug, F. J., Zagatto, E. A. G., Arruda, E. C., and Coutinho, C. A., *Anal. Chim. Acta*, 1986, **190**, 177.

40 Jones, E. A., and Hemmings, M.-J., *S. Afr. J. Chem.*, 1989, **42**, 6.

41 Taylor, M. J. C., Barnes, D. E., Marshall, G. D., *Anal. Chim. Acta*, 1992, **265**, 71.

42 Nyman, J., and Ivaska, A., *Talanta*, 1993, **40**, 95.

43 Muraki, H., Higuchi, K., Sasaki, M., Korenaga, T., and Tôei, K., *Anal. Chim. Acta*, 1992, **261**, 345.

44 Ariza, A. C., Linares, P., Luque de Castro, M. D., Valcárcel, M., *J. Automatic Chem.*, 1992, **14**, 181.

45 Clinch, J. R., Worsfold, P. J., Casey, H., and Smith, S. M., *Anal. Proc.*, 1988, **25**, 71.

46 van Staden, J. F., *Anal. Chim. Acta*, 1992, **261**, 453.

47 Pedersen, K. M., Kümmel, M., and Søeberg, H., *Anal. Chim. Acta*, 1990, **238**, 191.

48 Benson, R. L., McKelvie, I. D., Hart, B. T., and Hamilton, I. C., *Anal. Chim. Acta*, 1994, **291**, 233.

49 Pilloton, R., Mignogna, G., Fortunato, A., *Anal. Letters*, 1994, **27**, 833.

50 Chen, D., Luque de Castro, M. D., and Valcárcel, M., *Anal. Chim. Acta*, 1990, **230**, 137.

51 Benson, R. L., Worsfold, P. J., and Sweeting, F. W., *Anal. Chim. Acta*, 1990, **238**, 177.

52 Benson, R. L., and Worsfold, P. J., *Sci. Tot. Env.*, 1993, **135**, 17.

53 Balconi, M. L., Sigon, F., Borgarello, M., Ferraroli, R., and Realini, F., *Anal. Chim. Acta*, 1990, **234**, 167.

54 MacLaurin, P., Worsfold, P. J., Norman, P., and Crane, M., *Analyst*, 1993, **118**, 617.

55 Worsfold, P. J., Clinch, J. R., and Casey, H., *Anal. Chim. Acta*, 1987, **197**, 43.

56 Clinch, J. R., Worsfold, P. J., and Casey, H., *Anal. Chim. Acta*, 1987, **200**, 523.

57 Casey, H., Clarke, R. T., Smith, S. M., Clinch, J. R., and Worsfold, P. J., *Anal. Chim. Acta*, 1989, **227**, 379.

58 Blundell, N. J., Hopkins, A., Worsfold, P. J., and Casey, H., *J. Auto. Chem.*, 1993, **15**, 159.

59 Clinch, J. R., Worsfold, P. J., and Sweeting, F. W., *Anal. Chim. Acta*, 1988, **214**, 401.

60 Recktenwald, A., Kroner, K.-H., Kula, M.-R., *Enzyme Microb. Technol.*, 1985, **7**, 146.

61 Recktenwald, A., Kroner, K.-H., and Kula, M.-R., *Enzyme Microb. Technol.*, 1985, **7**, 607.

62 Nalbach, U., Schiemenz, H., Stamm, W. W., Hummel, W., and Kula, M.-R., *Anal. Chim. Acta*, 1988, **213**, 55.

63 Nikolajsen, K., Nielsen, J., and Villadsen, J., *Anal. Chim. Acta*, 1988, **214**, 137.

64 Künnecke, W., Kalisz, H. M., and Schmid, R., D., *Anal. Letters*, 1989, **22**, 1471.

65 Stamm, W. W., Pommerening, G., Wandrey, C., and Kula, M.-R., *Enzyme Microb. Technol.*, 1989, **11**, 96.

66 Brand, U., Reinhardt, B., Rüther, F., Scheper, T., and Schügerl, K., *Anal. Chim. Acta*, 1990, **238**, 201.

67 Künnecke, W., and Schmid, R. D., *J. Biotechnol.*, 1990, **14**, 127.

68 Lüdi, H., Garn, M. B., Bataillard, P., and Widmer, H. M., *J. Biotechnol.*, 1990, **14**, 71.

69 Stamm, W. W., and Kula, M.-R., *J. Biotechnol.*, 1990, **14**, 99.

70 Worsfold, P. J., Whiteside, I. R. C., Pfeiffer, H. F., and Waldhoff, H., *J. Biotechnol.*, 1990, **14**, 81.

71 Chung, S., Wen, X., Vilholm, K., de Bang, M., Christian, G., andRůžička, J., *Anal. Chim. Acta*, 1991, **249**, 77.

72 Christensen, L. H., Nielsen, J., and Villadsen, J., *Anal. Chim. Acta*, 1991, **249**, 123.

73 Filippini, C., Sonnleitner, B., Fiechter, A., Bradley, J., and Schmid, R., *J. Biotechnol.*, 1991, **18**, 153.

74 Forman, L. W., Thomas, B. D., Jacobson, F. S., *Anal. Chim. Acta*, 1991, **249**, 101.

75 Freitag, R., Fenge, C., Scheper, T., Schügerl, K., Spreinat, A., Antranikian, G., and Fraune, E., *Anal. Chim. Acta*, 1991, **249**, 113.

76 Kracke-Helm, H.-A., Brandes, L., Hitzmann, B., Rinas, U., and Schügerl, K., *J. Biotechnol.*, 1991, **20**, 95.

77 Nilsson, M., Håkanson, H., and Mattiasson, B., *Anal. Chim. Acta*, 1991, **249**, 163.

78 Ogbomo, I., Kittsteiner-Eberle, R., Englbrecht, U., Prinzing, U., Danzer, J., and Schmidt, H.-L., *Anal. Chim. Acta*, 1991, **249**, 137.

79 Peris-Tortajada, M., Maquieira, A., López, E., and Ors, R., *Analusis*, 1991, **19**, 266.

80 Scheper, T., Brandes, W., Grau, C., Hundeck, H. G., Reinhardt, B., Rüther, F., Plötz, F., Schelp, C., Schügerl, K., Schneider, K. H., Giffhorn, F., Rehr, B., and Sahm, H., *Anal. Chim. Acta*, 1991, **249**, 25.

81 Schügerl, K., Brandes, L., Dullau, T., Holzhauer-Rieger, K., Hotop, S., Hübner, U., Wu, X., Zhou, W., *Anal. Chim. Acta*, 1991, **249**, 87.

82 Degelau, A., Freitag, R., Linz, F., Middendorf, C., Scheper, T., Bley, T., Müller, S., Stoll, P., and Reardon, K. F., *J. Biotechnol.*, 1992, **25**, 115.

83 Englbrecht, U., and Schmidt, H.-L., *J. Chem. Tech. Biotechnol.*, 1992, **53**, 397.

84 Nilsson, M., Håkanson, H., and Mattiasson, B., *J. Chromatography*, 1992, **597**, 383.

85 Pfeiffer, H. F., Waldhoff, H., Worsfold, P. J., and Whiteside, I. R. C., *Chromatographia*, 1992, **33**, 49.

86 Becker, T., Kittsteiner-Eberle, R., Luck, T., and Schmidt, H.-L., *J. Biotechnol.*, 1993, **31**, 267.

87 Becker, T., Schuhmann, W., Betken, R., Schmidt, H.-L., Leible, M., and Albrecht, A., *J. Chem. Tech. Biotechnol.*, 1993, **58**, 183.

88 Busch, M., Höbel, W., and Polster, J., *J. Biotechnol.*, 1993, **31**, 327.

89 Carlsen, M., Johansen, C., Min, R. W., Nielsen, J., Meier, H., and Lantreibecq, F., *Anal. Chim. Acta*, 1993, **279**, 51.

90 Gründig, B., Strehlitz, B., Kotte, H., and Ethner, K., *J. Biotechnol.*, 1993, **31**, 277.

91 Middendorf, C., Schulze, B., Freitag, R., Scheper, T., Howaldt, M., and Hoffman, H., *J. Biotechnol.*, 1993, **31**, 395.

92 Nilsson, M., Mattiasson, G., and Mattiasson, B., *J. Biotechnol.*, 1993, **31**, 381.

93 Ogbomo, I., Steffl, A., Schuhmann, W., Prinzing, U., Schmidt, H.-L., *J. Biotechnol.*, 1993, **31**, 317.

94 Scheper, T., Brandes, W., Maschke, H., Plötz, F., and Müller, C., *J. Biotechnol.*, 1993, **31**, 345.

95 Schügerl, K., Brandes, L., Wu, X., Bode, J., Ree, J. I., Brandt, J., and Hitzmann, B., *Anal. Chim. Acta*, 1993, **279**, 3.

96 Steube, K., and Spohn, U., *Analyst*, 1994, **287**, 235.

97 Flaschka, H., McKeithan, C., and Barnes, R. M., *Anal. Lett.*, 1973, **6**, 585.

98 Trojanowicz, M., Pobozy, E., and Szpunar, J., *Chemia Anal.*, 1990, **35**, 661.

99 Dasgupta, P. K., Bellamy, H. S., Liu, H., Lopez, J. L., Loree, E. L., Morris, K., Petersen, K., and Mir, K. A., *Talanta*, 1993, **40**, 53.

100 Hauser, P. C., and Chiang, D. W. L., *Talanta*, 1993, **40**, 1193.

101 Skoog, D. A., West, D. M., and Holler, F. J., *Fundamentals of Analytical Chemistry*, Saunders College Publishing, New York, 5th edn., 1988.

102 Liu, H., and Dasgupta, K., *Anal. Chim. Acta*, 1994, **289**, 347.

103 Hauser, P. C., Rupasinghe, T. W. T., and Cates, N. E., *Talanta*, 1995, **42**, 605.

104 Lázaro, F., Ríos, A., Luque de Castro, M. D., and Valcárcel, M., *Anal. Chim. Acta*, 1986, **179**, 279.

105 Blanco, M., Gene, J., Iturriaga, H., Maspoch, S., and Riba, J., *Anal. Chim. Acta*, 1987, **34**, 987.

106 Owen, A. J., *The Diode-array Advantage in UV/Visible Spectroscopy*, Hewlett-Packard, Waldbronn, 1988.

107 Falkin, D. and Vosloo, M., *Spec. Europe*, 1993, **5**, 16.

108 Earle, C. W., Baker, M. E., Denton, M. B., and Pomeroy, R. S., *Trends Anal. Chem.*, 1993, **12**, 395.

109 *270M Rapid Scanning Imaging Spectrograph User Manual*, Spex Industries, Edison, NJ, 1992.

110 Marteau, P., Zanier-Szydlowski, N., Aoufi, A., Hotier, G. and Cansell, F., *Vib. Spec.*, 1995, **9**, 101.

111 Watano, S., and Miyanami, K., *Powder Technol.*, 1995, **83**, 55.

112 van Son, M., Schothorst, R. C., and den Boef, G., *Anal. Chim. Acta*, 1983, **153**, 271.

113 Hunter, D. A., and Uglow, R. F., *Ophelia*, 1993, **37**, 31.

114 Svensson,, G., and Anfält, T., *Clin. Chim. Acta*, 1982, **119**, 7.

115 Willason, S. W., and Johnson, K. S., *Marine Biol.*, 1986, **91**, 285.

116 Schulze, G., Liu, C. Y., Brodowski, M., Elsholz, O., Frenzel, W., and Möller, J., *Anal. Chim. Acta*, 1988, **214**, 121.

117 Georgiou, C. A., and Koupparis, M. A., *J. Auto. Chem.*, 1991, **13**, 199.

118 de la Torre, M., Luque de Castro, M. D., and Valcarcel, M., *Talanta*, 1992, **39**, 869.

119 van der Linden, W. E., *Anal. Chim. Acta*, 1983, **151**, 359.

120 Cañete, F., Ríos, A., Luque de Castro, M. D., and Valcárcel, M., *Analyst*, 1988, **113**, 739.

121 Cosano, J. S., Calle, J. L., Pinillos, J. L., Linares, P., and Luque de Castro, M. D., *Anal. Chim. Acta*, 1989, **221**, 173.

122 Santos Filha, M. M., dos Reis, B. F., Bergamin F°., H., and Baccan, N., *Anal. Chim. Acta*, 1992, **261**, 339.

123 Canale-Gutierrez, L., Maquieira, A., and Puchades, R., *Analyst*, 1990, **115**, 1243.

124 Goyal, S. S., Rains, D. W., and Huffaker, R. C., *Anal. Chem.*, 1988, **60**, 175.

125 Rios, A., Luque de Castro, M. D., and Valcárcel, M., *Anal. Chim. Acta*, 1986, **187**, 139.

126 Izquierdo, A., Linares, P., Luque de Castro, M. D., and Valcárcel, M., *Fres. J. Anal. Chem.*, 1990, **336**, 490.

127 Genfa, Z., and Dasgupta, P. K., *Anal. Chem.*, 1989, **61**, 408.

128 Frenzel, W., and Liu, C.-Y., *Fres. J. Anal. Chem.*, 1992, **342**, 276.

129 Alegret, S., Alonso, J., Bartroli, J., del Valle, M., Jaffrezic-Renault, N., and Duvault-Herrera, Y., *Anal. Chim. Acta*, 1990, **231**, 53.

130 Pasquini, C., and Cardoso de Faria, L., *Anal. Chim. Acta*, 1987, **193**, 19.

131 Cardoso de Faria, L., and Pasquini, C., *Anal. Chim. Acta*, 1991, **245**, 183.

132 Rohwedder, J. J. R., and Pasquini, C., *Analyst*, 1991, **116**, 841.

133 Kraus, P. R., and Crouch, S. R., *Anal. Lett.*, 1987, **20**, 183.

134 Hu, X., Takenaka, N., Takasuna, S., kitano, M., Bandow, H., Maeda, Y., and Hattori, M., *Anal. Chem.*, 1993, **65**, 3489.

135 Frenzel, W., *Lab. Rob. Autom.*, 1993, **5**, 245.

136 Gibb, S. W., *PhD Thesis*, University of East Anglia, 1994.

137 Baadenhuijsen, H., and Seuren-Jacobs, H. E. H., *Clin. Chem.*, 1979, **25**, 443.

138 García Prieto, A. M., Pérez Pavón, J. L., and Moreno Cordero, B., *Analyst*, 1994, **119**, 2447.

139 Canham, J. S., Gordon, G., and Pacey, G. E., *Anal. Chim. Acta*, 1988, **209**, 157.

140 Smith, R. M., and Martell, A. E., *Critical Stability Constants: Volume 2*, Plenum Press, New York, 1975.

141 Blundell, N. J., Worsfold, P. J., Casey, H. and Smith, S., *Environ. Inter.*, 1995, **21**, 205.

142 Krishnankutty, K., in *Encyclopedia of Analytical Sciences*, ed. Townshend, A., Academic Press, London, 1995. pp2109-2139.

143 Aoki, T., Uemura, S., and Munemori, M., *Anal. Chem.*, 1983, **55**, 1620.

144 Morrison, R. T., and Boyd, R. N., *Organic Chemistry*, Boston, Allyn and Bacon, 5th edn., 1987.

145 The Standing Committee of Analysts, *Ammonia in Waters 1981*, Methods for the Examination of waters and associated Materials, HMSO, London, 1981.

146 Miller, J. C., and Miller, J. N., *Statistics for Analytical Chemistry*, Ellis Horwood, Chichester, 3rd edn., 1993.

147 *Water Industry Specifications Information and Guidance Notes*, WAA Process Systems Committee, WRC, Swindon, 1989.

148 Kowalski, B. R., *Chem. & Ind.*, 1978, Nov., 882.

149 Massart, D. L., Vandeginste, B. G. M., Deming, S. N., Michotte, Y., and Kaufman, L., *Chemometrics: a Textbook*, Elsevier, Amsterdam, 1988.

150 Brereton, R. G., *Analyst*, 1987, **112**, 1635.

151 Geladi, P., and Esbensen, K., *J. Chemometr.*, 1990, **4**, 337.

152 Esbensen, K., and Geladi, P., *J. Chemometr.*, 1990, **4**, 389.

153 Kowalski, B. R., *Chemometrics: Theory and Applications*, ACS, Washington, D.C., 1977.

154 Sharaf, M. A., Illman, D. L., and Kowalski, B. R., *Chemometrics*, Wiley, New York, 1986.

155 Brereton, R. G., *Chemometrics: Applications of Mathematics and Statistics to Laboratory Systems*, Ellis Horwood, Chichester, 1990.

156 Haswell, S. J., *Practical Guide to Chemometrics*, Marcel Dekker, New York, 1992.

157 Adams, M. J., *Chemometrics in Analytical Spectroscopy*, Royal Society of Chemistry, Cambridge, 1995.

158 Beebe, K. R., and Kowalski, B. R., *Anal. Chem.*, 1987, **59**, 1007A.

159 Martens, H., and Næs, T., *Multivariate Calibration*, Wiley, Chichester, 1989.

160 Thomas, E. V., *Anal. Chem.*, 1994, **66**, 795A.

161 Martens, H., and Næs, T., *Trends Anal. Chem.*, 1984, **3**, 204.

162 Sanchez, E., and Kowalski, B. R., *J. Chemometr.*, 1988, **2**, 247.

163 Næs, T., and Martens, H., *Trends Anal. Chem.*, 1984, **3**, 266.

164 Haaland, D. M., and Thomas, E. V., *Anal. Chem.*, 1988, **60**, 1193

165 Sekulic, S., Seasholtz, M. B., Wang, Z., Kowalski, B. R., Lee, S. E., and Holt, B. R., *Anal. Chem.*, 1993, **65**, 835A.

166 Thomas, E. V., and Haaland, D. M., *Anal. Chem.*, 1990, **62**, 1091.

167 Howery, D. G., *Am. Lab.*, 1976, Feb., 14.

168 Wold, S. Esbensen, K. and Geladi, P. *Chemometr. Intell. Lab. Syst.*, 1987, **2**, 37.

169 Meglen. R. R., *Marine. Chem.*, 1992, **39**, 217.

170 Geladi, P., and Kowalski, B. R., *Anal. Chim. Acta*, 1986, **185**, 1.

171 Höskuldsson, A., *J. Chemometr.*, 1988, **2**, 211.

172 Clementi, S., Cruciani, G, and Curti, G., *Anal. Chim. Acta*, 1986, **191**, 149.

173 Geladi, P., *J. Chemometr.*, 1988, **2**, 231.

174 Blanco, M., Coello, J. Iturriaga, H., Maspoch, S., Redón, M., and Riba, J., *Anal. Chim. Acta*, 1992, **259**, 219.

175 Lehrmann, R., Bürck, J., and Ache, H.-J., *Process Qual. Cont.*, 1993, **4**, 139.

176 Lew, R., and Balke, S. T., *Applied Spec.*, 1993, **47**, 1747.

177 Small, G. W., Arnold, M. A., and Marquardt, L. A., *Anal. Chem.*, 1993, **65**, 3279.

178 Haaland, D. M., and Thomas, E.V., *Anal. Chem.*, 1988, **60**, 1202.

179 Manne, R., *Chemometr. Intell. Lab. Syst.*, 1987, **2**, 187.

180 Wise, B. M., and Kowalski, B. R., in *Process Analytical Chemistry*, eds. McLennan, F., and Kowalski, B. R., Chapman & Hall, London, 1995, ch. 8.

181 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and Redón, M., *App. Spec.*, 1994, **48**, 37.

182 MacLaurin, P., Worsfold, P. J., Crane, M., and Norman, P., *Anal. Proc.*, 1992, **29**, 65.

183 *The Design and Analysis of Industrial Experiments*, ed. Davies, O. L., Oliver and Boyd, London, 1956, pp247-494.

184 John, P. W. M., *Statistical Design and Analysis of Experiments*, MacMillan, New York, 1971, pp123-192.

185 Montgomery, D. C., *Design and Analysis of Experiments*, Wiley, New York, 2nd edn., 1984, pp189-356.

186 Morgan, E., *Chemometrics: Experimental Design*, Wiley, Chichester, 1991, pp81-188.

187 Finney, D. J., *Annals of Eugenics*, 1945, **12**, 291.

188 Yates, F., *The Design and Analysis of Factorial Experiments*, Imperial Bureau of Soil Science, Harpenden, 1937.

189 Durán-Merás, I., Muñoz de la Peña, A., Espinosa-Mansilla, A., and Salinas, F., *Analyst*, 1993, **118**, 807.

190 Blanco, M., Gene, J., Iturriaga, H., Maspoch, S., and Riba, J., *Talanta*, 1987, **34**, 987.

191 Kuban, V., and Dolezel, P., *Collect. Czech. Chem. Commun.*, 1988, **53**, 543.

192 Blanco, M., Coello, J., Gene, J., Iturriaga, H., and Maspoch, S., *Fres. J. Anal. Chem.*, 1990, **338**, 831.

193 Whitman, D. A., Seasholtz, M. B., Christian, G. D., Růžička, J., and Kowalski, B. R., *Anal. Chem.*, 1991, **63**, 775.

194 Saurina, J., and Hernández-Cassou, S., *Anal. Chim. Acta*, 1993, **281**, 593.

195 Lukkari, I., and Lindberg, W., *Anal. Chim. Acta*, 1988, **211**, 1.

196 Blanco, M., Coello, J., Iturriaga, H., Maspoch, S., and Riba, J., *Anal. Chem.*, 1994, **66**, 2905.

197 Hernández, O., Jiménez, A. I., Jiménez, F., and Arias, J. J., *Anal. Chim. Acta*, 1995, **310**, 53.

198 Brown, S. D., *Bioprocess Technol.*, 1990, **6**, 293.

199 *Standard Methods for the Examination of Water and Wastewater*, Am. Pub. Health Assoc., Am. Water Works Assoc., Water Poll. Cont. Fed., Washington DC, 17th edn., 1989.

200 The Standing Committee of Analysts, *Determination of Very Low Concentrations of Hydrocarbons and Halogenated Hydrocarbons in Water 1984-5*, Methods for the Examination of Waters and Associated Materials, HMSO, London, 1987.

201 Rossi, S. S., and Thomas, W. H., *Environ. Sci. Technol.*, 1981, **15**, 715.

202 Cline, P. V., Delfino, J. J., and Rao, P. S. C., *Environ. Sci. Technol.*, 1991, **25**, 914.

203 Deming, S. N., and Parker, Jr., L. R., *CRC Critical Reviews in Analytical Chemistry*, 1978, 7, 187.

204 Thijssen, P. C., Wolfrum, S. M., Kateman, G., and Smit, H. C., *Anal. Chim. Acta*, 1984, **156**, 87.

205 Thijssen, P. C., *Anal. Chim. Acta*, 1984, **162**, 253.

206 Thijssen, P. C., Kateman, G., and Smit, H. C., *Anal. Chim. Acta*, 1985, **173**, 265.

207 Thijssen, P. C., Prop, L. T. M., Kateman, G., and Smit, H. C., *Anal. Chim. Acta*, 1985, **174**, 27.

208 Kalman, R. E., *Trans. ASME, Ser. D.: J. Basic Eng.*,1960, **82**, 34.

209 Rutan, S. C., *J. Chemometr.*, 1987, **1**, 7.

210 Rutan, S. C., *Chemometr. Intell. Lab. Syst.*, 1989, **6**, 191.

211 Brown, S. D., *Anal. Chim. Acta*, 1986, **181**, 1.

212 Rutan, S. C., *J. Chemometr.*, 1990, **4**, 103.

213 Rutan, S. C., *Anal. Chem.*, 1991, **63**, A1103.

214 Rutan, S. C., and Brown, S. D., *Anal. Chim. Acta*, 1985, **167**, 39.

Appendices

APPENDIX 1

BASIC Program for Automated Control of the Portable FI Ammonia Monitor

```
1000 XBY(94)=52
1010  ONEX1 31000
1020  ONERR 27000
1030  GOSUB 30000
1040  PRINT  TAB (11),"---- Ammonia Monitor ----"
1050 XBY(57347)=128 : XBY(57345)=0 : XBY(57344)=0 : XBY(32766)=0
1060 XBY(61450)=32 : XBY(61451)=70
1070 XBY(59392)=56 : XBY(59392)=14 : XBY(59392)=6 : XBY(59394)=128
1080  DIM ADC(150),SAMPLE(10),INJECT(10),RSD(10),MV(10),SIGNAL(10)
1090  DIM ERCODE(10),MAX(10)
1100  DIM IN(20),OUT(20),LCD(250),TIM(20),ARRAY(9)
1110 LOMEM=MTOP : HIMEM=32768
1120  GOSUB 2000 :  GOSUB 3000
1130  IF XBY(32766)>=5 THEN  GOTO 40000
1140  REM ---------- ANALYSIS CYCLE ----------
1150 SAM=0
1160  DO
1170 SAM=SAM+1
1180  PRINT  :  PRINT "INJECTION CYCLE",SAM
1190 XBY(61454)=0
1200 ALLINJ=0 : MEAN=0
1210 INJ=0
1220  DO
1230 INJ=INJ+1
1240 FIRST=219 : LAST=222 : BASE=192 :  GOSUB 13000
1250  PRINT "SAMPLE ",SAM,"INJECTION",INJ
1260 FIRST=34 : LAST=50
1270 LCD(50)=48+INJ : LCD(38)=48+SAM : BASE=128 :  GOSUB 13000
1280  GOSUB 4000 :  GOSUB 9000 :  GOSUB 12000
1290 NUMBER=MV(INJ) :  GOSUB 20000
1300  UNTIL INJ>=INJECT
1310  PRINT "INJECTION CYCLE ",SAM," COMPLETE"
1320  GOSUB 18000 :  GOSUB 16000 :  GOSUB 19000
1330  GOSUB 6140 :  GOSUB 21000
1340 NUMBER=OUT(8) :  GOSUB 20000
1350  IF SAMPLE=3 THEN  GOTO 1380
1360  PRINT "Press return to continue"
1370 X=GET :  IF X<>13 THEN  GOTO 1370
1380  UNTIL SAM>=SAMPLE
1390 XBY(57344)=0 : XBY(57345)=8 : XBY(57346)=255
1400  GOSUB 19000 :  GOSUB 6000
1410  GOSUB 22000 :  GOSUB 23000
1420  PRINT "ANALYSIS CYCLE COMPLETE " :  PRINT CHR(38)
1440  GOSUB 19000
1450 MINS=OUT(3) : HOUR=OUT(1)
1460 MINS=MINS+SFREQ
1470  IF MINS=0 THEN MINS=60
1480  IF MINS>59 THEN  GOSUB 50000
1490 FIRST=219 : LAST=222 : BASE=192 :  GOSUB 13000
1500  PRINT "NEXT ANALYSIS = ",HOUR," : ",MINS
1510 NUMBER=HOUR :  GOSUB 29000 : LCD(16)=X+48 : LCD(17)=Z+48 : LCD(18)=58
1520 NUMBER=MINS :  GOSUB 29000 : LCD(19)=X+48 : LCD(20)=Z+48
1530 FIRST=1 : LAST=20 : BASE=128 :  GOSUB 13000
1540  PRINT "SYSTEM PAUSED#"
```

```
1550  REM ------- SYSTEM IDLE -------
1560  CLOCK 1
1570 TIME=0
1580  ONTIME 20,1620
1590  IDLE
1600  GOTO 1550
1610  REM ------- DISPLAY LAST SAMPLE  -------
1620 NUMBER=ABS(INT(OUT(9)))
1630  GOSUB 28000
1640 LCD(194)=W+48 : LCD(195)=X+48 : LCD(196)=Y+48 : LCD(197)=Z+48
1650 FIRST=183 : LAST=202 : BASE=192
1660  GOSUB 13000
1670 DELAY=15 :  GOSUB 14000
1680  GOSUB 19000
1690  GOSUB 17000 :  REM ------- DISPLAY TIME -------
1700  IF HOUR<>OUT(1) THEN  RETI
1710  IF MINS<>OUT(3) THEN  RETI
1720  RROM 1
2000  REM ---------- READ  VARIABLES ----------
2010 LOAD=61457
2020  FOR LOOP=1 TO 11
2030 DRESS=LOAD+LOOP : IN(LOOP)=XBY(DRESS)
2040  NEXT LOOP
2050 DRESS=DRESS+6 :  LD@ DRESS :  POP IN(12)
2060 DRESS=DRESS+6 :  LD@ DRESS :  POP IN(13)
2070 DRESS=DRESS+6 :  LD@ DRESS :  POP IN(14)
2080 DRESS=DRESS+6 :  LD@ DRESS :  POP IN(15)
2090 STADD=DRESS+6 :  LD@ STADD :  POP IN(16)
2100 SAMPLE=IN(1) : INJECT=IN(2) : FILLTME=IN(3) : DELTME=IN(4)
2110 FLUSH=IN(5) : PRECIS=IN(6) : MAXFAIL=IN(7)
2120 SFREQ=IN(8) : ALBINE=IN(9) : REFLAG=IN(10)
2130 CNC1=IN(12) : CNC2=IN(13)
2140 HIGHLIM=IN(14) : LOWLIM=IN(15) : SAVE=IN(16)
2150  RETURN
3000  REM ---------- INIT SCREEN ----------
3010  FOR LOOP=128 TO 148 : XBY(59392)=LOOP : XBY(59394)=32
3020 XBY(59392)=LOOP+62 : XBY(59394)=32 :  NEXT LOOP
3030  FOR LOOP=1 TO 4 :  READ VOID :  NEXT
3040 LOOP=0
3050  DO
3060 LOOP=LOOP+1
3070  READ LCD(LOOP)
3080  UNTIL LCD(LOOP)=255
3090  FOR LOOP=1 TO 20 :  READ TIM(LOOP) :  NEXT LOOP :  RESTORE
3100  DATA 61444,61442,61447,61448
3110  DATA 78,69,88,84,32,65,78,65,76,89,83,73,83,58
3120  DATA 32,32,32,32,32,32
3130  DATA 83,89,83,84,69,77,32,80,65,85,83,69,68
3140  DATA 83,65,77,32,32,32,32,32,32,32,32,32,73,78,74,32,32
3150  DATA 70,76,85,83,72,73,78,71,32,83,89,83,84,69,77
3160  DATA 83,65,77,80,76,73,78,71,32,68,65,84,65
3170  DATA 80,82,79,67,69,83,83,73,78,71,32,68,65,84,65
3180  DATA 70,73,78,68,73,78,71,32,80,69,65,75
3190  DATA 73,78,74,69,67,84
3200  DATA 73,78,71,32,32,83,65,77,32
3210  DATA 82,69,73,78,74,69,67,84,73,78,71,32,83
3220  DATA 65,77,80,76,69
3230  DATA 77,69,77,66,82,65,78,69,32
3240  DATA 70,65,73,76,85,82,69,32
3250  DATA 70,73,76,76,73,78,71,32,76,79,79,80
3260  DATA 83,89,83,84,69,77,32,83,72,85,84,68,79,87,78
3270  DATA 76,65,83,84,32,83,65,77,32,61,32,32,32,32,32,32
3280  DATA 109,103,47,108
3290  DATA 76,65,83,84,32,73,78,74,32,61,32,32,32,32,32
3300  DATA 32,32,32,32,32
```

```
3310  DATA 65,68,74,32,66,65,83,69,76,73,78,69,32
3320  DATA 255
3330  DATA 36,25,29,21,-16,13,-16,-16,-16,-16,-16,-16,-16,-16
3340  DATA -16,-16,-16,-16,-16
3350  DATA 77,69,77,66,82,65,78,69,32,68,69,70,69,67,84,73,86,69
3360  RETURN
4000  REM ---------- INJECTION CYCLE ----------
4010 XBY(57345)=8
4020  IF (SAMPLE=3).AND.(SAM=2) THEN G=1 : J=0 :  GOSUB 15000
4030  IF (SAMPLE=3).AND.(SAM=3) THEN G=1 : J=1 :  GOSUB 15000
4040 A=1 :  GOSUB 15000
4050 DELAY=4 :  GOSUB 14000
4060 E=1 :  GOSUB 15000
4070  IF INJ=1 THEN DELAY=FLUSH :  GOSUB 14000
4080 DELAY=10 :  GOSUB 14000
4090  GOSUB 5000
4100 DELAY=4 :  GOSUB 14000
4110 XBY(57345)=0
4120 C=1 :  GOSUB 15000
4130 DELAY=4 :  GOSUB 14000
4140 E=0 :  GOSUB 15000
4150 FIRST=156 : LAST=167 : BASE=192 :  GOSUB 13000
4160  IF INJ=1 THEN DELAY=FILLTME :  GOSUB 14000
4165  IF INJ>1 THEN DELAY=FILLTME/5 :  GOSUB 14000
4170 XBY(57345)=8
4180 DELAY=4 :  GOSUB 14000
4190 E=1 :  GOSUB 15000
4200 DELAY=4 :  GOSUB 14000
4210 C=0 :  GOSUB 15000
4220 FIRST=106 : LAST=120 : BASE=192 :  GOSUB 13000
4230 DELAY=DELTME :  GOSUB 14000
4240  GOSUB 7000
4250 FIRST=51 : LAST=65 : BASE=192 :  GOSUB 13000
4260 DELAY=FLUSH :  GOSUB 14000
4270 A=0 : E=0 : G=0 : J=0 :  GOSUB 15000
4280  RETURN
5000  REM ---------- ADJUST BASELINE ----------
5010 XBY(57346)=175 : TEMP=0 :  IF XBY(57346)=255 THEN  GOTO 27000
5020 ADCLOW=600 : ADCHI=1600 : WART=0
5030  FOR LOOP=1 TO ALBINE :  GOSUB 8000 : TEMP=TEMP+CHAP :  NEXT
5040 TREV=INT(TEMP/ALBINE)
5050 FIRST=223 : LAST=235 : BASE=192 :  GOSUB 13000
5060  IF XBY(61457)=1 THEN NUMBER=TREV :  GOSUB 10020
5070  PRINT "Bkg =",TREV,"  DAC =",XBY(57346),"No Adjusts = ",WART
5080  DO
5090 INC=0 : TEMP=0
5100  IF TREV>ADCHI THEN INC=-1 : XBY(57346)=XBY(57346)+INC : WART=WART+1
5110  IF TREV<ADCLOW THEN INC=1 : XBY(57346)=XBY(57346)+INC : WART=WART+1
5120 DELAY=1 :  GOSUB 14000
5130  IF INC=0 THEN  GOTO 5200
5140  FOR LOOP=1 TO ALBINE :  GOSUB 8000 : TEMP=TEMP+CHAP :  NEXT
5150 TREV=INT(TEMP/ALBINE)
5160  IF (XBY(57346)<=0).OR.(XBY(57346)>=255) THEN  GOTO 27000
5170  IF (WART>100) THEN  GOTO 27000
5180  PRINT "Bkg =",TREV,"  DAC =",XBY(57346),"No Adjusts = ",WART
5190  IF XBY(61457)=1 THEN NUMBER=TREV :  GOSUB 10020
5200  UNTIL (TREV<=ADCHI).AND.(TREV>=ADCLOW)
5210  RETURN
6000  REM ---------- FORMAT OUTPUT ----------
6010  PRINT "FORMATTING DATA"
6020  REM ------ CALIBRATED OUTPUT -------
6030  IF SAMPLE<>3 THEN  GOTO 6140
6040 CEPT=SIGNAL(3)-CNC2*((SIGNAL(3)-SIGNAL(2))/(CNC2-CNC1))
6050 GRAD=(SIGNAL(3)-SIGNAL(2))/(CNC2-CNC1)
6060 OUT(9)=(SIGNAL(1)-CEPT)/GRAD
```

```
6070 OUT(7)=ERCODE(1)+ERCODE(2)+ERCODE(3) : OUT(7)=INT(OUT(7))
6080 TRAFF=(RSD(1)*RSD(1))+(RSD(2)*RSD(2))+(RSD(3)*RSD(3))
6090 OUT(8)=SQR(TRAFF) : OUT(8)=INT(OUT(8))
6100 OUT(10)=INT(SIGNAL(1))
6110 OUT(11)=INT(SIGNAL(2))
6120 OUT(12)=INT(SIGNAL(3))
6130  GOTO 6180
6140  REM ------- GENERAL OUTPUT -------
6150  REM *** OUT(1,3,4,6) = HOURS, MINUTES, DATE, MONTH
6160 OUT(7)=INT(ERCODE(SAM)) : OUT(8)=INT(SIGNAL(SAM))
6170 OUT(9)=INT(RSD(SAM))
6180  RETURN
7000  REM ---------- SAMPLE SIGNAL ----------
7010  PRINT "SAMPLING SIGNAL ",CHR(33) : FIRST=66 : LAST=78 : BASE=192
7020  GOSUB 13000
7030 LOOP=0
7040  DO
7050 LOOP=LOOP+1
7060 TEMP=0
7070  FOR DELAY=1 TO ALBINE
7080  GOSUB 8000
7090 TEMP=TEMP+CHAP
7100  NEXT DELAY
7110 ADC(LOOP)=INT(TEMP/ALBINE)
7120  PRINT ADC(LOOP)
7130  IF XBY(61457)=1 THEN  GOSUB 10000
7140  UNTIL LOOP>=149
7150  PRINT CHR(64)
7160 FIRST=219 : LAST=222 : BASE=192 :  GOSUB 13000
7170 XBY(57345)=8
7180  RETURN
8000  REM ---------- SAMPLE ADC ----------
8010 XBY(57345)=24 : XBY(57345)=8 : XBY(57345)=24
8020  FOR DICK=1 TO 20 :  NEXT
8030 LOW=XBY(60416)
8040 HIGH=XBY(60417)
8050  IF HIGH>31 THEN SUB=8191
8060  IF HIGH<32 THEN SUB=-8191
8070 NICK=LOW+(256*HIGH)
8080 CHAP=16382-(NICK-SUB)
8090  RETURN
9000  REM ---------- PROCESS DATA ----------
9010 MAX(INJ)=-1
9020 N=1
9030  DO
9040 N=N+1
9050  IF ADC(N)>MAX(INJ) THEN MAX(INJ)=ADC(N)
9060  UNTIL N>=149
9070  PRINT "MAX BEFORE FILTERING =",MAX(INJ)
9080  PRINT "FILTERING DATA" : FIRST=79 : LAST=93 : BASE=192 :  GOSUB 13000
9090  REM  ------ MEDIAN FILTER DATA -------
9100 LOOP=5
9110  DO
9120 LOOP=LOOP+1
9130 COUNT=1
9140  FOR I=LOOP-4 TO LOOP+4
9150 ARRAY(COUNT)=ADC(I)
9160 COUNT=COUNT+1
9170  NEXT I
9180  GOSUB 11000 :  REM ----- SORT ARRAY -----
9190 ADC(LOOP)=MED
9200  UNTIL LOOP>=146
9210 MAX(INJ)=-1
9220 N=5
9230  DO
```

```
9240  N=N+1
9250   IF ADC(N)>MAX(INJ) THEN MAX(INJ)=ADC(N)
9260   UNTIL N>=146
9270   PRINT "MAX AFTER FILTERING =",MAX(INJ)
9280   RETURN
10000  REM ---------- SHOW ADC ON DISPLAY ----------
10010 NUMBER=ADC(LOOP)
10020  GOSUB 28000
10030 LCD(239)=V+48 : LCD(240)=W+48 : LCD(241)=X+48
10040 LCD(242)=Y+48 : LCD(243)=Z+48
10050 FIRST=239 : LAST=243 : BASE=211 : CURSE=59392
10060  FOR LOP=FIRST TO LAST
10070 POS=LOP-LAST
10080 LCD(LOP)=ABS(INT(LCD(LOP)))
10090 XBY(CURSE)=BASE+POS
10100  IF LCD(LOP)<256 THEN XBY(59394)=LCD(LOP)
10110  NEXT LOP
10120 XBY(59392)=148
10130  RETURN
11000  REM ---------- SORT ARRAY ----------
11010  FOR LOOP1=2 TO 9
11020 TEMP=ARRAY(LOOP1)
11030 BOT=1 : TUP=LOOP1-1
11040  DO
11050 MIDDLE=INT((BOT+TUP)/2)
11060  IF TEMP<ARRAY(MIDDLE) THEN TUP=MIDDLE-1 ELSE BOT=MIDDLE+1
11070  WHILE BOT<=TUP
11080  FOR J=LOOP1-1 TO BOT STEP -1
11090 ARRAY(J+1)=ARRAY(J)
11100  NEXT J
11110 ARRAY(BOT)=TEMP
11120  NEXT LOOP1
11130 MED=ARRAY(5)
11140  RETURN
12000  REM ---------- PEAK FINDING ROUTINE ----------
12010  PRINT "FINDING PEAK" : FIRST=94 : LAST=105 : BASE=192 :  GOSUB 13000
12020 I=0
12030  FOR LOOP=3 TO 12
12040 I=I+ADC(LOOP)
12050  NEXT LOOP
12060 BGD=INT(I/10) : MAX(INJ)=-1 : ELEM=1
12070 LOOP=18
12080  DO
12090 LOOP=LOOP+1
12100  IF ADC(LOOP)>MAX(INJ) THEN MAX(INJ)=ADC(LOOP) : ELEM=LOOP
12110  UNTIL LOOP>=144
12120 MAX(INJ)=0
12130  IF ELEM>144 THEN MAX(INJ)=ADC(ELEM) :  GOTO 12180
12140  FOR LOOP=(ELEM-4) TO (ELEM+4)
12150 MAX(INJ)=MAX(INJ)+ADC(LOOP)
12160  NEXT LOOP
12170 MAX(INJ)=INT(MAX(INJ)/9)
12180 MV(INJ)=MAX(INJ)-BGD :  PRINT
12190  PRINT "BACKGROUND = ",BGD," mV" :  PRINT "MAX = ",MAX(INJ)," mV"
12200  PRINT "INJECTION  = ",MV(INJ)," mV" :  PRINT
12210 ALLINJ=ALLINJ+MV(INJ)
12220  RETURN
13000  REM ---------- DRIVE LCD DISPLAY ----------
13010  FOR LOOP=BASE TO BASE+20 : XBY(59392)=LOOP : XBY(59394)=32 :  NEXT
13020 CURSE=59392
13030  FOR LOOP=FIRST TO LAST
13040 POS=LOOP-FIRST
13050 XBY(CURSE)=BASE+POS
13060 LCD(LOOP)=INT(LCD(LOOP))
13070  IF (LCD(LOOP)<1).OR.(LCD(LOOP)>255) THEN LCD(LOOP)=255
```

```
13080  IF LCD(LOOP)<255 THEN XBY(59394)=LCD(LOOP)
13090  NEXT LOOP
13100 XBY(59392)=148
13110  RETURN
14000  REM ---------- DELAY LOOP ----------
14010  CLOCK 1 : TIME=0
14020  DO  :  UNTIL TIME>=DELAY
14030  RETURN
15000  REM ---------- CALC OUTPUT CODE ----------
15010 OP1=(1*A)+(2*B)+(4*C)+(8*D)+(16*E)+(32*G)+(64*H)+(128*J)
15020 XBY(57344)=OP1
15030  RETURN
16000  REM ---------- CALC ERROR CODES ----------
16010  PRINT "CALCULATING ERROR CODE"
16020 NUMFAIL=XBY(61454) : FAIL=0
16030 ERCODE(SAM)=0 : FLAG1=0 : FLAG2=0 : FLAG3=0 : FLAG4=0
16040 FLAG5=0 : FLAG6=0 : FLAG7=0 : FLAG8=0
16050  IF SAM=1.AND.MV(SAM)<MINEPT THEN FLAG1=1
16060  IF SAM=2.AND.MV(SAM)<MINEPT THEN FLAG2=1
16070  IF SAM=3.AND.MV(SAM)<MINEPT THEN FLAG3=1
16080  IF SAM=1.AND.MV(SAM)>HIGHLIM THEN FLAG4=1
16090  IF SAM=1.AND.MV(SAM)<LOWLIM THEN FLAG5=1
16100  IF SAM=1.AND.ABS(RSD(SAM))>(100-PRECIS) THEN FLAG6=1
16110  IF SAM=2.AND.ABS(RSD(SAM))>(100-PRECIS) THEN FLAG7=1
16120  IF SAM=3.AND.ABS(RSD(SAM))>(100-PRECIS) THEN FLAG8=1
16130 ERCODE(SAM)=(1*FLAG1)+(2*FLAG2)+(4*FLAG3)+(8*FLAG4)+(16*FLAG5)
16140 ERCODE(SAM)=ERCODE(SAM)+(32*FLAG6)+(64*FLAG7)+(128*FLAG8)
16150  IF ERCODE(SAM)>REFLAG THEN FAIL=1
16160  IF FAIL=1 THEN NUMFAIL=NUMFAIL+1
16170  IF NUMFAIL=MAXFAIL THEN  GOTO 27000
16180  IF FAIL=0 THEN NUMFAIL=0
16190 XBY(61454)=NUMFAIL
16200  IF FAIL=1 THEN  GOSUB 19000 :  GOSUB 6140 :  GOSUB 21000
16210  IF FAIL=1 THEN  PRINT "REINJECTING SAMPLE",SAM
16220  IF FAIL=1 THEN FIRST=121 : LAST=138 : BASE=192 :  GOSUB 13000
16230  IF FAIL=1 THEN  GOTO 1200
16240  RETURN
17000  REM ---------- DISPLAY TIME ----------
17010  PRINT "TIME = ",OUT(1)," : ",OUT(3),
17020  PRINT "   DATE = ",OUT(4)," / ",OUT(6), CR ,
17030 CURSE=59392 : BASE=191
17040  GOSUB 24000
17050  FOR LOOP=1 TO 20 : XBY(CURSE)=BASE+LOOP : XBY(59394)=48+TIM(LOOP)
17060  IF TIM(LOOP)<207 THEN XBY(59394)=48+TIM(LOOP)
17070  NEXT LOOP
17080  RETURN
18000  REM ---------- CALC STATS ----------
18010  PRINT "CALCULATING STATS"
18020 FIRST=139 : LAST=155 : BASE=192 :  GOSUB 13000
18030 MEAN=ALLINJ/INJECT
18040 SIGNAL(SAM)=MEAN
18050  IF INJECT=1.OR.MEAN=0 THEN  GOTO 18130
18060 MANU=0
18070  FOR Y=1 TO INJECT
18080 THIS=(MV(Y)-MEAN)*(MV(Y)-MEAN)
18090 MANU=MANU+THIS
18100  NEXT Y
18110 SD=SQR(MANU/(INJECT-1))
18120 RSD(SAM)=ABS((SD/MEAN)*100)
18130  RETURN
19000  REM ---------- GET TIME AND DATE ----------
19010  FOR LOOP=1 TO 6
19020  IF LOOP=2.OR.LOOP=5 THEN  GOTO 19070
19030  READ DRESS
19040  IF XBY(61451)>128 THEN  GOTO 19040
```

```
19050 OUT(LOOP)=XBY(DRESS)
19060  IF OUT(LOOP)>60 THEN  GOTO 19040
19070  NEXT LOOP
19080 OUT(2)=ASC( : ) : OUT(5)=ASC( : )
19090  RESTORE
19100  RETURN
20000  REM ---------- DISPLAY INJECTIONS ----------
20010  GOSUB 28000
20020 LCD(214)=W+48 : LCD(215)=X+48 : LCD(216)=Y+48 : LCD(217)=Z+48
20030 FIRST=203 : LAST=222 : BASE=192
20040  GOSUB 13000
20050  RETURN
21000  REM ---------- OUTPUT DATA ----------
21010  PRINT "************************************************"
21020  PRINT "SAMPLE",SAM :  PRINT
21030  PRINT "TIME =",OUT(1),CHR(OUT(2)),OUT(3),
21040  PRINT "   DATE = ",OUT(4),CHR(OUT(5)),OUT(6)
21050  FOR LOOP=1 TO INJECT
21060  PRINT "INJECTION ",LOOP," = ",MV(LOOP),"mV"
21070  NEXT LOOP
21080  PRINT "MEAN = ",OUT(8)," mV" :  PRINT USING(#####.#)
21090  PRINT USING(0) :  PRINT "RSD = ",OUT(9),"%"
21100  PRINT "ERROR CODE :",OUT(7)
21110  PRINT "************************************************"
21120  RETURN
22000  REM ---------- CALIBRATED OUTPUT ----------
22010  PRINT "TIME =",OUT(1),CHR(OUT(2)),OUT(3),
22020  PRINT "DATE = ",OUT(4),CHR(OUT(5)),OUT(6)
22030  PRINT "SAMPLE   = ",OUT(10),"mV"
22040  PRINT "STANDARD1 = ",OUT(11),"mV"
22050  PRINT "STANDARD2 = ",OUT(12),"mV"
22060  PRINT USING(####.##)
22070  PRINT "CONCENTRATION = ",OUT(9)," mg/l      RSD = ",OUT(8),"%"
22080  PRINT USING(0)
22090  PRINT "ERROR CODE :",OUT(7)
22100  PRINT "************************************************"
22110  RETURN
23000  REM ---------- STORE DATA ----------
23010  PRINT "STORING DATA"
23020  LD@ STADD :  POP SAVE
23030 PROP=1
23040  FOR LOOP=1 TO 6
23050 DRESS=SAVE+LOOP
23060  IF LOOP=2.OR.LOOP=4 THEN PROP=PROP+1
23070  IF ABS(OUT(PROP))>255 THEN OUT(PROP)=255
23080 XBY(DRESS)=INT(ABS(OUT(PROP))) : PROP=PROP+1
23090  NEXT LOOP
23100 SAVE=DRESS
23110  FOR LOOP=1 TO 4
23120  PUSH ABS(OUT(LOOP+8)) : DRESS=SAVE+(6*LOOP) :  ST@ DRESS
23130  NEXT LOOP
23140  IF DRESS>HIMEM-30 THEN  GOTO 40000
23150  PUSH DRESS :  ST@ STADD
23160  RETURN
24000  REM ---------- CHANGE TIME FORMAT FOR LCD ----------
24010 NUMBER=OUT(1) :  GOSUB 29000 : TIM(8)=X : TIM(9)=Z : TIM(10)=10
24020 NUMBER=OUT(3) :  GOSUB 29000 : TIM(11)=X : TIM(12)=Z
24030 NUMBER=OUT(4) :  GOSUB 29000 : TIM(16)=X : TIM(17)=Z : TIM(18)=-1
24040 NUMBER=OUT(6) :  GOSUB 29000 : TIM(19)=X : TIM(20)=Z
24050  RETURN
25000  REM ---------- FORMAT NUMBERS ON DISPLAY ----------
25010  FOR LOOP=FIRST TO LAST
25020 POS=LOOP-FIRST
25030  IF ABS(LCD(LOOP)<208) THEN XBY(CURSE)=BASE+POS : XBY(59394)=
       48+LCD(LOOP
```

```
25040  NEXT  : XBY(59392)=148
25050  RETURN
27000  REM ---------- INCREMENT SYSTEM FAILURES ----------
27010 XBY(32766)=XBY(32766)+1
27020  PRINT "SYSTEM FAILURE",XBY(32766)
27030  RROM 1
28000  REM ---------- FORMAT SIGNAL DISPLAY ----------
28010 V=0 : W=0 : X=0 : Y=0 : Z=0 : FAN=NUMBER
28020  IF NUMBER<10000 THEN V=-16 :  GOTO 28050
28030 V=INT(NUMBER/10000)
28040 NUMBER=NUMBER-INT(V*10000)
28050  IF NUMBER<1000 THEN W=0 :  GOTO 28080
28060 W=INT(NUMBER/1000)
28070 NUMBER=NUMBER-INT(W*1000)
28080  IF NUMBER<100 THEN X=0 :  GOTO 28110
28090 X=INT(NUMBER/100)
28100 NUMBER=NUMBER-INT(X*100)
28110  IF NUMBER<10 THEN Y=0 :  GOTO 28140
28120 Y=INT(NUMBER/10)
28130 Z=NUMBER-INT(Y*10) :  GOTO 28150
28140 Z=NUMBER
28150  IF FAN<1000 THEN W=-16
28160  IF FAN<100 THEN X=-16
28170  IF FAN<10 THEN Y=-16
28180  RETURN
29000  REM ---------- FORMAT TIME ARRAY ----------
29010  IF NUMBER<10 THEN X=NUMBER : Z=-16 :  GOTO 29030
29020 X=INT(NUMBER/10) : Y=10*X : Z=NUMBER-Y
29030  RETURN
30000  REM ---------- USER ESCAPE FROM RUNTRAP MODE ----------
30010 LOOP=1
30020  DO
30030 X=GET :  IF X=0 THEN LOOP=LOOP+1
30040  IF X=27 THEN XBY(94)=0 : DBY(38)=DBY(38).AND.0FEH
30050  IF X=27 THEN  PRINT "RUN TRAP OFF" :  PRINT
30060  UNTIL (X=27).OR.(LOOP>=200)
30070  RETURN
31000  REM ---------- EXTERNAL INTERUPT ----------
31010  REM DOESN'T DO ANYTHING YET
31020  RETI
40000  REM ---------- SHUTDOWN SYSTEM ----------
40010 XBY(57344)=0 : XBY(57345)=8 : XBY(57346)=255 : XBY(94)=0
40020  PRINT "SYSTEM SHUTDOWN"
40030 BASE=128 : FIRST=168 : LAST=182 :  GOSUB 13000
40040  GOSUB 19000 :  GOSUB 17000
40050  END
45000  REM ---------- MEMBRANE CHECK ----------
45010 TEFLN=OUT(12)/OUT(11)
45020  IF TEFLN<30 THEN 45040
45030  RETURN
45040 XBY(57344)=0 : XBY(57345)=8 : XBY(57346)=255 : XBY(94)=0
45050  PRINT "****MEMBRANE DEFECTIVE****"
45060  PRINT "    SYSTEM SHUTDOWN"
45070 BASE=128 : FIRST=139 : LAST=155 :  GOSUB 13000
45080 BASE=192 : FIRST=168 : LAST=182 :  GOSUB 13000
45090  END
50000  REM ---------- TIME CORRECTION ----------
50010 MINS=MINS-60
50020 HOUR=HOUR+1
50040  IF HOUR>23 THEN HOUR=HOUR-24
50050  RETURN
```

APPENDIX 2

Matlab Program for the Standard Kalman Filter

```
function [newx,fit]=regkal2(HH,zz) ;

% This program uses the standard Kalman filter.
%[newx,fit]=regkal2(HH,zz,).
% n: number of known component in the system.
% HH: The model component spectra (row vectors).
% Create HH as "HH=[x;y;z;...] " where x,y,z,... are model components.
% p: The initial guess of the variance.
% R: The measurement variance.
% zz: The unresolved spectra (column vector).



        if nargin ~=2;
          error(' Kalman filter - wrong number of parameters');
        end

%  Get input from user

n=input('Enter the number of known components in the system.');

p=input('Enter the initial guess of variance. (Default = 10000)');
        if isempty(p);
                p=10000;
        end

R=input('Enter the Measurement variance. (Default = 0.000001)');
        if isempty(R);
                R=1e-6;
        end

        end
clg
var=input('Enter the column you wish to analyse. ');
zz=zz(:,var);
echo on
plot(zz','r')
hold
plot(HH')
title('Display of the measurement(red) and the model spectra')
ylabel('Absorbance (arbitrary unit)')
xlabel('Wavelength (nm)')
hold off

[range,yvalue]=ginput(2)
[lower]=input('Enter the lower value, use the value display under range')
[upper]=input('Enter the higher value, use the value display under range')
zzsub=zz(lower:upper,:);
```

```
HHsub=HH(:,lower:upper);
zz=zzsub;
HH=HHsub;
plot(zzsub')
title('display of the selected measurement spectra')
ylabel('Absorbance (arbitrary unit)')
xlabel('Wavelength (nm)')

% Determine the direction of filter.

forwrev=input('Enter the direction of the fit [forward=1/reverse=-1] (Default is forward)');
        if isempty(forwrev);
                forwrev = 1;
        end

[row,column]=size(zz);

  if forwrev==1;
     initial=1;
     num=row;

     else
     num=row;
  end ;

%Extract initial guess for concentration, oldx.
oldx=zeros(n,1);

%Extract initial guess for variance, oldp.
identity = eye(n);
oldp =p*identity;

% Determine System Transition Matrix

F = input('Enter system transition matrix, F. (Default is identity.)');
        if isempty(F)
                F = identity;
        end

        for i=1:num,
          back=num-i+1;
            if forwrev==-1 ;
              k=back;
               else
              k=i;
            end

                H = HH(:,k);

% Determine innovation sequence,
          v(k) = (zz(k) - H'*oldx);


% State estimate extrapolation.

              oldx = F*oldx;
```

```
% Covariance estimate extrapolation.

                    oldp = F*oldp*F';

% Weighting factor, Kalman gain.
                Kk=oldp*H*(1/(H'*oldp*H +R));

%Parameter update,
                newx(:,i) = oldx + Kk*v(k);


%Covariance matrix
                C=(identity - (Kk*H')) ;
                newp= C* oldp*C' + Kk*R*Kk';
                oldp=newp;
                oldx=newx(:,i);
        end

fit =HH'*newx(:,i);
      if forwrev==-1
            newx=fliplr(newx);
      end

subplot(221),plot(newx');
subplot(221), title('New x values')
subplot(223),plot(zz')
subplot(223), title('Fit(red) result and original zz')
hold on;
subplot(223),plot(fit','r')
hold off;
subplot(224),plot(HH')
subplot(224), title('HH, the model spectra')
pause
subplot(111)

        if forwrev == 1
                newx(:,upper-lower)
        else
                newx(:,1)
        end
pause


end;
```

APPENDIX 3

Matlab Program for the Modified Kalman Filter

```
function [calibconst,RRMSEC,newconc,RRMSEP]=regkal3(HH,zz,newabs,realconc) ;

% This program uses the standard Kalman filter modified to use sets of multicomponent
%       calibration samples.
% Optional step included for prediction of new sample concs.
%[calibconst,RRMSEC,newconc,RRMSEP]=regkal3(HH,zz,newabs,realconc).
% calibconst: the calibrated model spectra for the system components.
% n: number of known components in the system.
% RRMSEC: relative root-mean-squared error of calibration;
%       (error between predicted and actual concs. for calibration set samples)
% RRMSEC gives errors for individual components and overall error
%       (n+1 vector, i.e. [C1, C2....Cn, overall])
% newconc: predicted concentrations for test set samples;
% RRMSEP: relative root-mean-squared error of prediction for test set samples
% HH: the calibration set component concentrations;
%       rows = no. components;  columns = no. samples.
% p: the initial guess of the variance.
% R: the measurement variance.
% zz: the calibration set absorbance spectra;
%       rows = no. samples;  columns = no. wavelength points.
% newabs: the test set absorbance spectra;
%       rows = no. samples;  columns = no. wavelength points.
% realconc: actual concentrations for test set samples.;
%       rows = no. components;  columns = no. samples.



        if nargin ~=4;
          error(' Kalman filter - wrong number of parameters');
        end

%  Get input from user

n=input('Enter the number of known components in the system.');

p=input('Enter the initial guess of variance. (Default = 10000)');
        if isempty(p);
                p=10000;
        end

R=input('Enter the measurement variance. (Default = 0.000001)');
        if isempty(R);
                R=0.000001;
        end

        end

clf

figure(1)
```

```
subplot(211),plot(zz')
subplot(211),title('zz: calibration set spectra')
subplot(211),ylabel('Absorbance')
subplot(211),xlabel('Wavelength')

subplot(212),bar(HH')
subplot(212),title('HH: calibration set concentrations')
subplot(212),ylabel('mol/l')
subplot(212),xlabel('Sample number')

pause

% Determine the direction of filter.

forwrev=input('Enter the direction of the fit [forward=1/reverse=-1] (Default is forward)');
        if isempty(forwrev);
                forwrev = 1;
        end

[row,column]=size(zz);

  if forwrev==1;
     initial=1;
     num=row;

     else
     num=row;
  end ;

%Extract initial guess for concentration, oldx.
oldx=zeros(n,1);

%Extract initial guess for variance, oldp.
identity = eye(n);
oldp =p*identity;

% Determine System Transition Matrix

F = input('Enter system transition matrix, F. (Default is identity.)');
        if isempty(F)
                F = identity;
        end


% Begin outer loop (by wavelength)
        for ii = 1:column
                z=zz(:,ii);
                oldx=zeros(n,1);
                oldp=p*identity;

% Begin inner loop (by sample)
                for i=1:num,
                  back=num-i+1;
                if forwrev==-1 ;
                     k=back;
                 else
                  k=i;
```

```
                end

                    H = HH(:,k);

% Determine innovation sequence,
            v(k) = (z(k) - H'*oldx);


% State estimate extrapolation.

                    oldx = F*oldx;

% Covariance estimate extrapolation.

                        oldp = F*oldp*F';

% Weighting factor, Kalman gain.
                    Kk=oldp*H*(1/(H'*oldp*H +R));

% Parameter update,
             newx(:,i) = oldx + Kk*v(k);


% Covariance matrix
           C=(identity - (Kk*H')) ;
               newp= C* oldp*C' + Kk*R*Kk';
                  oldp=newp;
             oldx=newx(:,i);

             end     % inner loop

    calibconst(:,ii)=newx(:,num);

    end             % outer loop

  if forwrev==-1
       newx=fliplr(newx);
  end

    if forwrev == 1
          newx(:,num);
    else
          newx(:,1);
    end

% Calculate RRMSEC

pred=zz/calibconst;

    for c=1:n
          RRMSEC(1,c)=(100/mean(HH(c,:)))*sqrt(sum(((pred(:,c)-
HH(c,:)').^2)/row));
    end

RRMSEC(1,n+1)=(100/mean(mean(HH)))*sqrt((sum(sum((pred(:,:)-HH(:,:)').^2))/(row*n)));

clf
```

```
figure(1)

axes('position',[.2 .05 .6 .4])
plot(calibconst')
title('Calibration constants')

axes('position',[.15 .6 .3 .3])
plot(newx');
title('New x values')

axes('position',[.55 .6 .3 .3])
bar(RRMSEC)
title('RRMSEC (%)')

pause

% Optional prediction routine for new samples

[newpred]=input('Do you want to predict new samples? (yes=1, no=0; default=no)');
        if isempty(newpred);
                end
        if newpred==0;
                end
        if newpred==1;
                newconc=newabs/calibconst;
                [row,col]=size(realconc);
                        for c=1:n

        RRMSEP(1,c)=(100/mean(realconc(c,:)))*sqrt(sum(((newconc(:,c)-
realconc(c,:)').^2)/col));
                        end
        RRMSEP(1,n+1)=(100/mean(mean(realconc)))*sqrt((sum(sum((newconc(:,:)-
realconc(:,:)').^2))/(col*n)));


                figure(2)
                subplot(221),bar(RRMSEP)
                subplot(221),title('RRMSEP (%)')
                subplot(224),plot(newconc',realconc,'x')
                subplot(224),title('Correlation')
                subplot(224),xlabel('Predicted concentration')
                subplot(224),ylabel('Actual concentration')
                subplot(111)
                pause

                end

end;
```

APPENDIX 4

Matlab Program for Determination of Baseline Drift Using the Kalman Filter

```
function [calibconst,RRMSEC,drift]=rkfd3(HH,zz) ;

% This program uses the KF to determine drift in multicomponent calibration spectra.
% [calibconst,RRMSEC,drift]=rkfd3(HH,zz).
% calibconst: the calibrated model spectra for the system components.
% n: number of known components in the system (analytes + spectral offset)
% RRMSEC: relative root-mean-squared error of calibration;
%       (error between predicted and actual concs. for calibration set samples)
% HH: the calibration set component concentrations;
%       rows = no. components;  columns = no. samples.
% p: the initial guess of the system variance.
% R: the measurement variance.
% zz: the calibration set absorbance spectra;
%       rows = no. samples;  columns = no. wavelength points.


        if nargin ~=2;
          error(' Kalman filter - wrong number of parameters');
        end

% Get input from user

 n=input('Enter no. of known components in system (analytes + spectral offset)');

 p=input('Enter the initial guess of variance. (Default = 10000)');
        if isempty(p);
                p=10000;
        end

 R=input('Enter the Measurement variance. (Default = 1e-6)');
        if isempty(R);
                R=1e-6;
        end

 q=input('Enter the system noise. (Default = 0.0000001)');
        if isempty(q);
                q=0.0000001;
        end

 clf
 figure(1)
 subplot(211),plot(zz')
 subplot(211),title('zz: calibration set spectra')
 subplot(211),ylabel('Absorbance')
 subplot(211),xlabel('Wavelength')
 subplot(212),bar(HH')
 subplot(212),title('HH: calibration set concentrations')
 subplot(212),ylabel('mol/l')
 subplot(212),xlabel('Sample number')
 pause
```

```
% Determine the direction of filter.

 forwrev=input('Enter the direction of the fit [forward=1/reverse=-1] (Default is forward)');
        if isempty(forwrev);
                forwrev = 1;
        end

 [row,column]=size(zz);

        if forwrev==1;
        initial=1;
                num=row;

                else
                num=row;
        end ;

%Extract initial guess for concentration, oldx.
 oldx=zeros(n+1,1);

%Extract initial guess for variance, oldp.
 identity = eye(n+1);
 oldp =p*identity;

% System noise covariance matrix
 Q=zeros(n+1,n+1);
 Q(n+1,n+1)=q;

% System transition matrix
 F=eye(n+1,n+1);
 F(n,n+1)=1;

% Begin outer loop (for all wavelengths)
        for ii = 1:column
                z=zz(:,ii);

% Begin inner loop (for all calib stds)
          for i=1:num,
                back=num-i+1;
                if forwrev==-1 ;
                     k=back;
                 else
                  k=i;
                   end

                H = [HH(:,k);0];


% State estimate extrapolation.
                oldx = F*oldx;

% Covariance estimate extrapolation.
                oldp = F*oldp*F' + Q;

% Determine inovation sequence,
           v(k) = (z(k)-(H'*oldx));
```

```
% Weighting factor, Kalman gain.
                Kk=oldp*H*(1/(H'*oldp*H +R));

% Parameter update.
                newx(:,i) = oldx + Kk*v(k);


% Covariance matrix.
                C=(identity - (Kk*H')) ;
                newp= C* oldp*C' + Kk*R*Kk';
                oldp=newp;
                oldx=newx(:,i);
                drift(i,1)=newx(n+1,i);

          end   % inner loop

          oldp=[(p*eye(n+1)];
          oldx=zeros(n+1,1);

          calibconst(:,ii)=newx(1:n+1,num);

        end     % outer loop

        if forwrev==-1
                newx=fliplr(newx);
        end

        if forwrev == 1
                newx(:,num);
        else
                newx(:,1);
        end

% Calculate RRMSEC

 pred=zz/calibconst;

        for c=1:(n-1)
            RRMSEC(1,c)=(100/mean(HH(c,:)))*sqrt(sum(((pred(:,c)-HH(c,:)').^2)/row));
        end

 clf

 figure(1)
 axes('position',[.2 .05 .6 .4])
 plot(calibconst')
 title('Calibration constants')
 axes('position',[.15 .6 .3 .3])
 plot(newx');
 title('New x values')
 axes('position',[.55 .6 .3 .3])
 bar(RRMSEC)
 title('RRMSEC (%)')
 pause


end;
```