Forensic comparison of unevaporated and evaporated automotive gasoline samples from Australia and New Zealand

by

Paul Mark Lyne Sandercock

A thesis
submitted for the Degree of
Doctor of Philosophy (Science)

University of Technology, Sydney
December, 2002

Certificate of authorship and originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as acknowledged within the text.

I certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. I certify that all information sources and literature used are indicated in the thesis.

P. Mark L. Sandercock

November, 2002

Acknowledgements

Before I could undertake a major project like this one I needed financial support and time away from my duties at the Royal Canadian Mounted Police (RCMP) Forensic Laboratory, Winnipeg. Thanks are due my Master's supervisor, Professor Josef Takats, University of Alberta, Edmonton, who supported my bid for a scholarship from the Natural Sciences and Research Engineering Council (NSERC), Canada. I am indebted to NSERC for granting me a two year post-graduate scholarship. I wish to thank my supervisor at the RCMP Forensic Laboratory Services, Winnipeg, Mr. Ron Hrynchuk, who wholeheartedly supported my desire to take an unpaid leave from the RCMP and pursue research in this area.

A big thank you is extended to Mr. Gary Bunio, Imperial Oil (ESSO) Canada, Calgary whose genuine desire to assist me led to my first research grant. Mr. Bunio put me in contact with Dr. Roger W. Cohen at Exxon Research and Engineering Company, Annandale, New Jersey, USA. Through Dr. Cohen, ExxonMobil generously gave me a research grant to support this project.

Associate professor Claude Roux, University of Technology, Sydney, was instrumental in guiding my application for an International Postgraduate Research Scholarship (IPRS) through the university bureaucracy. I was awarded this full fee paying scholarship so generously provided by the Department of Education, Training and Youth Affairs (DETYA), Commonwealth of Australia.

Although money is important in seeing my project to completion, my acknowledgements are not limited to the amount of money awarded. The people my family and I have had contact with have demonstrated that God can turn seemingly impossible situations into times of blessing. The people who come to mind when we first arrived in Australia are Mal and Heather York, and Emma and Ric Van Wachem. Then there is our family in Christ at St. Andrew's Abbottsford: Nicola and Eric Lewis, Rev. Jacinth Myles, Rev. Jackie Stoneman, Barbara and Tony Robinson, Greg and Carmel Jarmain, Dr. Coral Chamberlain, and Bill (the dairy man) Steele, to name only a few, all of whom gave generously of themselves in time, possessions and often both. I would like to thank the entire congregation for their prayers at specific times of need.

Thanks are due to Mr. Kevan Walsh from the New Zealand forensic science service, ESR. Kevan kindly offered to drive around Auckland and collect gasoline samples for this project. I would also like to thank Dr. Peter Petocz, Department of Mathematical Sciences, UTS for contributing to my understanding of PCA and the use of Minitab, and Dr. Philip Doble for his encouragement during this project.

I would like to acknowledge the work of Dr. Eric Du Pasquier, colleague, friend and research supervisor. Eric brought a calm, sober influence to this project. I am indebted to the hard work he did supporting me with research grants, helping me keep the project focused and reviewing this thesis. Thanks, Eric!

I could never have even gotten up the courage to pursue my dream of doing full time research (and completing a PhD) without the encouragement of my best friend, my wife Sandi. To say "thank you" does not capture my gratitude towards you, Sandi, for always supporting me, even in a move half-way around the world. No acknowledgement would be complete without mentioning my children, Evan (7) and Claire (4). No one else can lay the seriousness of life aside and remind an adult that there is more to life than work. Yes, Claire, we are here to swing. Thank you to you both.

I dedicate this research to the One who is the Author of all inspiration. To God be the glory.

"In his heart a man plans his course, but the Lord determines his steps."
Proverbs 16:9 (NIV)

Preface

July, 1994. It was early afternoon on a fine summer day in a small city in Western Canada. The sidewalk cafes were doing a brisk business, and residents were coming and going from the neighbouring apartment buildings. Within this bustling block was a laundromat and drycleaning shop, closed for a lunch break. A natural gas leak was reported by passers-by and the gas company was called in. The leak was traced to the back of the laundromat/drycleaner where it was discovered that the gas mains had been tampered with, filling the shop with natural gas. The area was evacuated, the gas supply shut off and the gas allowed to dissipate. Meanwhile, the Fire Commissioner and the Royal Canadian Mounted Police (RCMP) were called in to investigate the tampering. The investigation led to the discovery of a failed arson attempt at the laundromat. Thirty liquid filled bottles, some stoppered and others with cloth wicks, were found in the crawl space beneath the business. The investigators believed that had the fire not gone out soon after it was lit, the entire city block would have been destroyed by the inevitable gas explosion. The bottles were found to contain automotive gasoline, aviation gasoline, tetrachloroethylene, or combinations of these three liquids. It was likely that the mixtures of tetrachloroethylene and gasoline had caused the fire to go out shortly after it was lit. At the suspect's residence two jerry cans, one with automotive gasoline and one with aviation gasoline, were found. Clearly, it was important to make a comparison between the two types of gasoline found at the suspect's residence and the gasoline present in the bottles. The lack of a scientific method (validated by the RCMP Forensic Laboratory Service) for the comparison of liquid gasoline samples prevented a meaningful comparison being made between the known liquids and those recovered from the scene. Thus began my interest in comparing refined petroleum products, and in particular, comparing automotive gasoline.

Table of contents

Chapter 3. Data analysis **62**

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

List of tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Abbreviations

ADG	Australian Dangerous Goods
AFIRS	Australian Fire Incident Reporting System
ASTM	American Society for Testing and Materials
BP	British Petroleum
bp	boiling point
cf.	Latin: *confer*, “compare”
ci	confidence interval
CSIRO	Commonwealth Scientific and Industrial Research Organisation
e.g.	Latin: *exempli gratia*, “for example”
EIC	extracted ion chromatogram
et al.	Latin: *et alia*, “and others”
FTIR	Fourier transform infrared
g	gram
GC	gas chromatography
GC-MS	gas chromatography-mass spectrometry
HPLC	high performance liquid chromatography
i.d.	inside diameter
i.e.	Latin: *id est*, “that is”
ISO	International Organisation for Standardisation
km	kilometre
L	litre
LDA	linear discriminant analysis
m	metre
mg	milligram
mL	millilitre
mm	millimetre
MS	mass spectrometry
ms	millisecond
m/z^+	mass to charge ratio
NFIRS	National Fire Incident Reporting System

NIST	National Institute of Standards and Technology
o.d.	outside diameter
PAH	polycyclic aromatic hydrocarbon
PC	principal component
PCA	principal component analysis
ppm	part per million
RON	research octane number
SIM	selected ion monitoring
SIMCA	soft independent model classification analogy
v/v	volume per volume
w/v	weight per volume
w/w	weight per weight
UK	United Kingdom
µL	microlitre
USA	United States of America (also abbreviated as US)

Abstract

The comparison of two or more samples of gasoline (petrol) to establish a common origin is a difficult problem in the forensic investigation of arsons and suspicious fires. The high-boiling fraction of the gasoline was targeted with a view to apply the techniques described herein to evaporated gasoline samples in the future. A novel micro solid phase extraction technique using activated alumina was developed to isolate the polar compounds and the polycyclic aromatic hydrocarbons from a 200 μL sample of gasoline. This technique was applied to 35 randomly collected samples of unevaporated gasoline, covering three different grades (regular unleaded, premium unleaded and lead replacement), collected in Sydney, Australia. The samples were analysed using full-scan GC-MS; potential target compounds identified were the C_0- to C_2-naphthalenes. The samples were then re-analysed directly, without prior treatment, using GC-MS in selected ion monitoring (SIM) mode for target compounds that exhibited variation between gasoline samples. Multivariate statistical analysis (principal component and linear discriminant analysis) was applied to the chromatographic data. The first two principal components described approximately 90% of the variation in the data and showed that the majority of the 35 samples could be differentiated using the method developed. A comparison of unevaporated samples collected in Auckland, New Zealand to those collected in Sydney was also made. Most of the samples could be differentiated based on their country of origin.

The variation of unevaporated regular unleaded and premium unleaded gasoline over time at three different service stations was studied. Ninety-six samples of gasoline were collected over a 16 week period and analysed for their C_0- to C_2-naphthalene content using the GC-MS (SIM) method that was developed. In most cases it was found that the C_0- to C_2-naphthalene profile in gasoline changed from week to week, and from station to station.

Samples of 25%, 50%, 75% and 90% evaporated gasoline (w/w) were generated from the 35 randomly collected samples of unevaporated gasoline. The C_0- to C_2-naphthalene content of all unevaporated and evaporated gasoline samples was determined using the GC-MS (SIM) method. Analysis of the data by principal components followed by linear discriminant analysis showed that the 35 samples formed 18 unique groups, irrespective of the level of evaporation. The application of the method to forensic casework is discussed.

Chapter 1. Introduction

1.1 Introduction

Arson is the deliberate or malicious act of setting fire to property, including buildings and vehicles. Although arson is itself a criminal act, it may also be used either to cover up a crime such as burglary, fraud or homicide, or to perpetrate a crime such as assault (intent to injure) or homicide.[1,2] Arson is classified as a property crime. Despite this classification, it is regarded by the courts as a serious offence; the most serious of all property crimes. In Australia arson is ranked 17th out of 123 offence types, appearing with the majority of the violent crimes and other so-called "crimes against persons" (Table 1.1).

Table 1.1 Order of seriousness of offence types.

Offence Type	Seriousness
Murder	1
Attempted murder	2
Manslaughter (excludes driving)	3
Drug trafficking	4
Manufacturing or growing drugs	5
Importing or exporting drugs	6
Possession of drugs	7
Sexual assault - rape	8
Assault causing grievous bodily harm	9
Unlawful carnal knowledge	10
Incest	11
Robbery with major assault	12
Robbery, armed	13
Extortion and blackmail	14
Robbery with minor assault	15
Robbery, other	16
Arson	17
Manslaughter (driving)	18
Conspiracy to murder	19
Dangerous driving causing death	20

Source: Crime and Justice Statistics Queensland 1999-2000. Office of the Government Statistician, Queensland Government, 2001.

One difficulty in the investigation of arson is that this crime, by its very nature, destroys much of the evidence. A lack of physical evidence combined with the criminal court's requirement that the prosecution demonstrate that the defendant had exclusive opportunity to set the fire while disproving other explanations (*e.g.* accidental cause), makes arson one of the most difficult crimes to prosecute successfully.[3]

The conclusion of a police investigation by an arrest or other means, such as a warning or caution, is used as a measure of police activity and is known as the "clearance rate". Recent clearance rates for arson have been published by some Australian states and the Australian Capital Territory (ACT) and are summarised in Table 1.2.

Table 1.2 Arson clearance rates for selected Australian states and the ACT.

State	Clearance Rate (%)	Reporting Year	Clearance Rate (%)	Reporting Year
Australian Capital Territory[1]	8.2	1999/2000	4.9	2000/2001
New South Wales[2]	6.0	2000	5.7	2001
Queensland[3]	16.3	1999/2000	--	--
Victoria[4]	12.2	1999/2000	11.1	2000/2001
Western Australia[5]	--	--	26.0	2001/2002

1. ACT Policing Annual Report 1999-2000. Australian Capital Territory Police Service, 2001.
ACT Policing Annual Report 2000-2001. Australian Capital Territory Police Service, 2002.
2. Allen, J.; Chilvers, M.; Doak, P.; Goh, D.; Painting, T.; Ramsay, M. NSW Recorded Crime Statistics 2001. NSW Bureau of Crime Statistics and Research, Statistical Services Unit. Attorney General's Department, Sydney, Australia, 2002.
3. Crime and Justice Statistics Queensland 1999-2000. Office of the Government Statistician, Queensland Government, 2001.
4. Victoria Police 2000/20001 Provisional Crime Statistics, Summary of Offences Recorded and Cleared. <http://police.vic.gov.au> Accessed June 26, 2002.
5. Western Australia Police Service. Facts and Figures. <http://www.police.wa.gov.au/AboutUs/AboutUs.asp?FactsandFigures#CrimeStatistics> Accessed October 4, 2002.

It is important to note that the definition of a "cleared offence" differs between Australian states (and territories) making it difficult to compare data between the regions.[4] Clearance rates for arson in other countries are also typically low: 6.8% in Canada in 1999[5] and 2000;[6] an average of 26.0% in New Zealand[7] from 1998 to 2001; 8.8% in the United Kingdom[8] in 1999/2000; and, 16.5% in the USA[9] in 2000. It is also difficult to obtain accurate arson clearance rate statistics because the definition of arson may differ between police forces and/or between countries. In the United Kingdom, for example, arson causing injury is recorded as a violent crime and not as arson.[10] Furthermore crime rates may not be reported uniformly between jurisdictions. Some police forces do not report an arson clearance rate and instead combine arson with other property crimes to produce a combined clearance rate.

1.2 Fire and arson

1.2.1 Fire and arson in Australia

The Australian Fire Incident Reporting System (AFIRS) is used to collect data from fires that occur around the country. This data collection and reporting system, described by the Australian Standard AS 2577.1-1992, is based on the National Fire Incident Reporting System (NFIRS) developed in the United States. It has been estimated that data collected by the AFIRS during the early 1990s covered approximately 85% of Australia.[11,12] During the mid-1990s the Division of Building, Construction and Engineering at the Commonwealth Scientific and Industrial Research Organisation (CSIRO) undertook a study of the fire data collected by the AFIRS between July 1, 1992 and June 30, 1994.[11,12] Although dated, this data represents the most recent and comprehensive examination of Australian fire data ever published. To date, attempts by the author to access the AFIRS for more current data have been unsuccessful.

Despite the fact that published data from the AFIRS represents fires that occurred approximately a decade ago, this data is still useful for describing the problem of arson in Australia. During the two year period from July 1, 1992 to June 30, 1994, fires in Australia caused 2,651 injuries and 272 deaths to civilians, and 971 injuries and two deaths to fire brigade personnel.[11,12] The AFIRS data was used to estimate that during this same two year period, all fires in Australia have resulted in $887,000,000 of property damage, $164,000,000 of which was directly attributed to arson (Table 1.3). The estimated property loss in each column in Table 1.3 is based on the number of fires multiplied by the mean cost of those fires.

Table 1.3 Estimated property loss ($'000) resulting from incendiary (arson) and suspicious fires, as well as all fires reported in Australia between July 1, 1992 to June 30, 1994.

1992-1993

	$0-10	$10-50	$50-250	$250-1000	$1000-5000	>$5000	Totals
incendiary	3,105	6,270	25,050	14,375	36,000	0	84,800
suspicious	3,875	6,480	15,750	12,500	18,000	5,000	61,605
total, all fires	46,835	39,960	108,000	66,875	99,000	20,000	380,670

1993-1994

	$0-10	$10-50	$50-250	$250-1000	$1000-5000	>$5000	Totals
incendiary	1,815	1,320	8,700	25,625	27,000	15,000	79,460
suspicious	8,235	1,530	5,100	23,125	9,000	15,000	61,990
total, all fires	64,620	14,370	59,550	130,000	153,000	85,000	506,540

Dollar losses estimated using the data presented in:

King, M. Australian fire incident statistics 1992-1993. Division of Building, Construction and Engineering, CSIRO, 1995.

King, M. Australian national fire incident statistics 1993-1994. Division of Building, Construction and Engineering, CSIRO, 1997.

This data estimates direct property loss due to fires, and so it must be regarded as a conservative estimate of the actual cost of arson to Australian society. In a report prepared by the Home Office in the United Kingdom,[13] the five major types of cost related to fire were identified as follows:

- cost of physical damage to property;
- cost of fire brigade operations (does not include cost of police investigations);
- cost of injuries and deaths;
- cost of fire prevention programs and research; and,
- cost of business interruptions.

Using data from fires that occurred in the UK in 1993, this report estimated that physical damage to property accounted for approximately 19% of the total of the five major costs related to a fire. When this information is combined with the data in Table 1.3, the cost of arson to Australia can be calculated to be more than 400 million dollars each year ($446,000,000 for 1992-1993 and $418,000,000 for 1993-1994).

The most common type of fire in Australia is the "bush fire" where grass, brush and/or trees are ignited (Table 1.4). The second most common type of fire in Australia is the structure fire (Table 1.4).

Table 1.4 Number of incendiary (arson) and suspicious fires, as well as all fires reported in Australia, by type, between July 1, 1992 to June 30, 1994.

1992-1993

Number of fires

	Structure fire	Outside structure fire	Vehicle fire	Brush, grass fire	Refuse fire	Explosion, no fire	Spill, ensuing fire	Other fire	Total
incendiary	774	499	706	3941	1498	2	4	1	7425
suspicious	3434	1247	2999	11048	3727	34	25	77	22591
total, all fires	12359	2423	7350	26054	8442	80	142	10072	66922

Number of fires as percent of total (all fires)

	Structure fire	Outside structure fire	Vehicle fire	Brush, grass fire	Refuse fire	Explosion, no fire	Spill, ensuing fire	Other fire	Total
incendiary	1.2%	0.7%	1.1%	5.9%	2.2%	0.0%	0.0%	0.0%	11.1%
suspicious	5.1%	1.9%	4.5%	16.5%	5.6%	0.1%	0.0%	0.1%	33.8%
total, all fires	18.5%	3.6%	11.0%	38.9%	12.6%	0.1%	0.2%	15.1%	100.0%

1993-1994

Number of fires

	Structure fire	Outside structure fire	Vehicle fire	Brush, grass fire	Refuse fire	Explosion, no fire	Spill, ensuing fire	Other fire	Total
incendiary	568	557	534	3846	1764	5	2	11	7287
suspicious	1861	1028	2080	5981	2207	9	7	25	13198
total, all fires	19181	4503	6589	24770	8957	103	115	5649	69867

Number of fires as percent of total (all fires)

	Structure fire	Outside structure fire	Vehicle fire	Brush, grass fire	Refuse fire	Explosion, no fire	Spill, ensuing fire	Other fire	Total
incendiary	0.8%	0.8%	0.8%	5.5%	2.5%	0.0%	0.0%	0.0%	10.4%
suspicious	2.7%	1.5%	3.0%	8.6%	3.2%	0.0%	0.0%	0.0%	18.9%
total, all fires	27.5%	6.4%	9.4%	35.5%	12.8%	0.1%	0.2%	8.1%	100.0%

Sources:

King, M. Australian fire incident statistics 1992-1993. Division of Building, Construction and Engineering, CSIRO, 1995.

King, M. Australian national fire incident statistics 1993-1994. Division of Building, Construction and Engineering, CSIRO, 1997.

Using AFIRS data, the CSIRO reported that in 1992-1993 12,273 structure fires involved 7,664 residential properties (62.4%),[11] and in 1993-1994 19181 structure fires involved 11,980 residential properties (62.5%).[12] Yet, from 1992 to 1994 less than 4% of residential fires were determined to be arson: 3.6% in 1992-1993[11] (275 of 7,664); and, 1.8% in 1993-

1994[12] (220 of 11,980). The majority of structure fires and arson involving residential property occurred in one or two family dwellings (Table 1.5). In contrast, fires to educational institutions (*e.g.* public schools) accounted for less than 4% of all structure fires: 2.9% in 1992-1993[11] (353 of 12,273), and 3.6% in 1993-1994[12] (681 of 19,181). However, of these school fires, 186 (52.7%) in 1992-1993,[11] and 197 (28.9%) in 1993-1994,[12] were determined to be either arson or suspicious in nature. In a report on arson at public schools in New South Wales it was stated that 14% of all school arsons involved the use of accelerants such as gasoline and kerosene.[14] Unfortunately the analysis of the AFIRS data by the CSIRO did not include data identifying the type of accelerant used (if any) in those fires that were determined to be either arson or suspicious.

Table 1.5 Number of incendiary (arson) and suspicious residential structure fires, as well as all structure fires reported in Australia between July 1, 1992 and June 30, 1994.

1992-1993

	1 or 2 family dwelling	Apart-ment, flat	Rooming, boarding house	Hotel, motel	Dorm	Self-contained unit	Other	Not classified	Total
Incendiary	364	52	19	6	0	0	1	0	442
Suspicious	351	95	4	14	2	5	5	7	483
total, all fires	6716	1176	88	192	30	40	24	35	8301

1993-1994

	1 or 2 family dwelling	Apart-ment, flat	Rooming, boarding house	Hotel, motel	Dorm	Self-contained unit	Other	Not classified	total
Incendiary	171	31	5	6	1	1	1	4	220
Suspicious	475	165	15	31	2	9	6	9	712
total, all fires	6645	1547	117	292	40	49	31	3259	11980

Sources:
King, M. Australian fire incident statistics 1992-1993. Division of Building, Construction and Engineering, CSIRO, 1995.
King, M. Australian national fire incident statistics 1993-1994. Division of Building, Construction and Engineering, CSIRO, 1997.

1.2.2 Fire and arson in other countries

The availability and consistency of fire statistics varies widely from country to country and year to year. Uniform statistical data on fire and arson was obtained from Canada, the United Kingdom (UK) and the United States of America (USA). Monetary loss is quoted as originally reported with no adjustments made for inflation or currency fluctuations.

The total number of fires reported in Canada (Table 1.6) is similar to the number reported in Australia (Table 1.4). A notable difference in the fire statistics between Canada and Australia is that the number of residential dwelling fires is significantly higher in Canada (*cf.* Tables 1.5 and 1.6). Note that the estimated value of property loss due to arson in Canada rose significantly after 1996. This change occurred after the responsibility of reporting Canadian fire data was transferred from the Council of Canadian Fire Marshals and Fire Commissioners to a federal government agency (Ministry for Human Resources and Development Canada).

Table 1.6 Fires and arson in Canada.

	1996[1]	*1997*[2]	*1998*[3]	*1999*[4]
total number of fires	60,138	no data	57,602	55,169
injuries, all fires	3,152	no data	2,697	2,287
deaths, all fires	374	no data	337	388
dwellings, all fires	24,367	no data	23,546	22,150
number of arsons reported	11999	no data	10624	10170
loss, arson (CA$ estimated)	85,000,000	no data	258,230,339	221,990,022

Sources:

1. Council of Canadian Fire Marshals and Fire Commissioners. Fire Losses in Canada Annual Report 1996. Ottawa, Canada 1997.
2. Unable to obtain published data.
3. Human Resources and Development Canada. Fire Losses in Canada Annual Report 1998. Ottawa, Canada, 2000.
4. Human Resources and Development Canada. Fire Losses in Canada Annual Report 1999. Ottawa, Canada, 2001.

The United Kingdom's Home Office estimated that arson cost UK society £1.3 billion a year, and resulted in 22,000 injuries and 1,100 deaths between 1986 and 1996.[15] Recently released statistics from the United Kingdom on fires that occurred between 1990 and 2000 revealed that the occurrence of arson in dwellings is of particular concern. In 2000 one fifth (14,200 out of 70,900) of all dwelling fires were recorded as arson or suspected arson. These malicious dwelling fires were estimated to have caused 51 deaths and 3,400 injuries.[16] A summary of the available data on fires and arson in the United Kingdom over several years is given in Table 1.7.

Table 1.7 Fires and arson in the United Kingdom.

	1996	*1997*	*1998*	*1999*	*2000*
total number of fires[1]	532,000	469,000	410,000	469,000	476,000
injuries, all fires[1]	18,221	18,583	18,198	18,198	17,640
deaths, all fires[1]	709	723	656	623	595
structures, all fires[1]	118,800	116,900	112,900	115,900	112,600
number of arsons reported[1]	88,400	82,600	86,200	103,000	111,100
loss, arson (UK£ estimated)[2]	707,000,000	739,000,000	864,000,000	no data	no data

Sources:

1. Fire Statistics United Kingdom, 2000. Department for Transport, Local Government and the Regions, Wetherby, United Kingdom. February 20, 2002.
2. Arson Prevention Bureau. Annual Report 1999. United Kingdom, 2000.

In the USA, arson and suspicious structure fires typically cost approximately one billion dollars each year in direct property losses (Table 1.8). The U.S. Federal Bureau of

Table 1.8 Fires and arson in the United States of America.

	1998[1]	*1999*[2]	*2000*[3]
total number of fires	1,756,000	1,823,000	1,708,000
injuries, all fires	23,100	21,875	23,350
deaths, all fires	4,035	3,570	4,045
structures, all fires	684,443	523,000	505,500
monetary loss (US$), all fires	8,600,000,000	10,024,000,000	11,207,000,000
number of arsons, structures only†	97,875	72,000	75,000
loss, arson (structures only, US$ estimated)†	970,920,000	1,281,000,000	1,340,000,000

1. United States Fire Administration. Fire in the United States 1989-1998. 12th Edition, August 2001. This data includes suspicious structure fires as well as those fires that were determined to be arson.
2. Karter, M.J. Fire Loss in the United States During 1999. Fire Analysis and Research Division, National Fire Protection Association, USA. September 2000.
3. Karter, M.J. Fire Loss in the United States During 2000. Fire Analysis and Research Division, National Fire Protection Association, USA. September 2001.

† Number of arsons and estimated dollar loss for arson includes suspicious fires.

Investigation reported that, in 2000, structure fires accounted for 43.8% of all recorded arsons, with residential property comprising 60.3% of all arsons on structures.[17]

A general summary of the fire and arson situation in Australia, Canada, the United Kingdom and the United States of America over a two year period is presented in Table 1.9. It is important to note the great disparity between each nation's estimates of property loss. This highlights how each nation estimates loss differently and suggests an undervaluing of the cost of arson in Australia during the early 1990s. Note that the estimated dollar loss for arson in the USA is only for structure fires and includes suspicious fires.

Table 1.9 Comparison of fire and arson data from Australia, Canada, the United Kingdom and the United States of America.

Country	Australia[2]	Canada[3]	UK[4]	USA[5]
year range[1]	1992-1994	1998-1999	1997-1998	1998-1999
Mean population[6]	17,687,977	30,792,972	59,096,924	271,846,088
	Rates per 100,000 population per year			
total number of fires	387	183	744	658
Injuries	8.9	8.1	31	8.3
Deaths	2.1	1.2	1.2	1.4
structure fires	89	no data	194	222
number of incidents of arson reported	42	34	143	31
property loss due to arson†	$464,430	$780,092	£1,356,088	$413,930

1. Australian data - June 30, 1992 to July 31, 1994; Canadian and USA data – January 1, 1998 to December 31, 1999; UK data – January 1, 1997 to December 31, 1998.
2. See references in Table 1.5
3. See references in Table 1.6
4. See references in Table 1.7
5. See references in Table 1.8; number of incidents of arson and property loss due to arson are for structure fires only and include suspicious fires.
6. Population data obtained from U.S. Census Bureau, International Data Base (IDB) On-line Access. <http://www.census.gov/ipc/www/idbass.html>, accessed September 30, 2002.

† Estimates are in each country's respective currency.

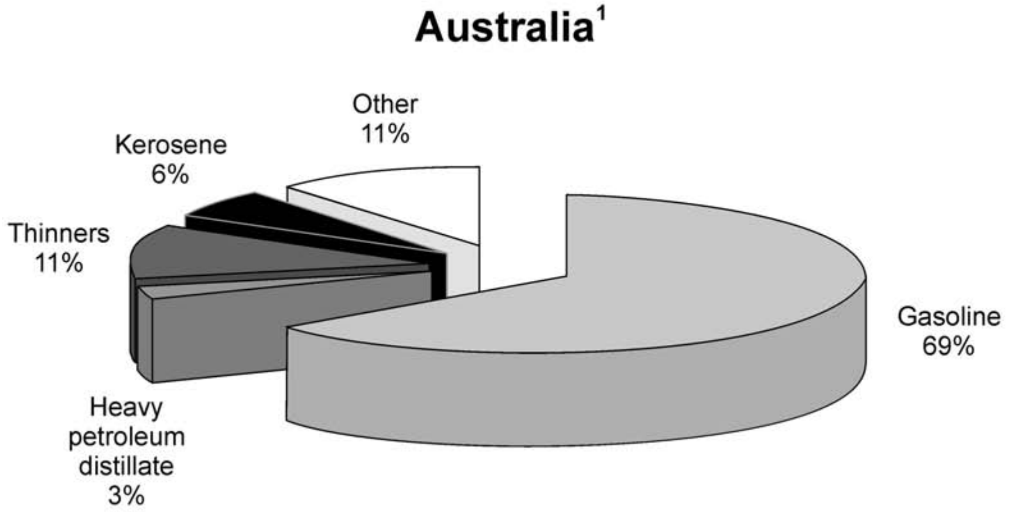
1.3 Forensic comparison of fuels and other petroleum products

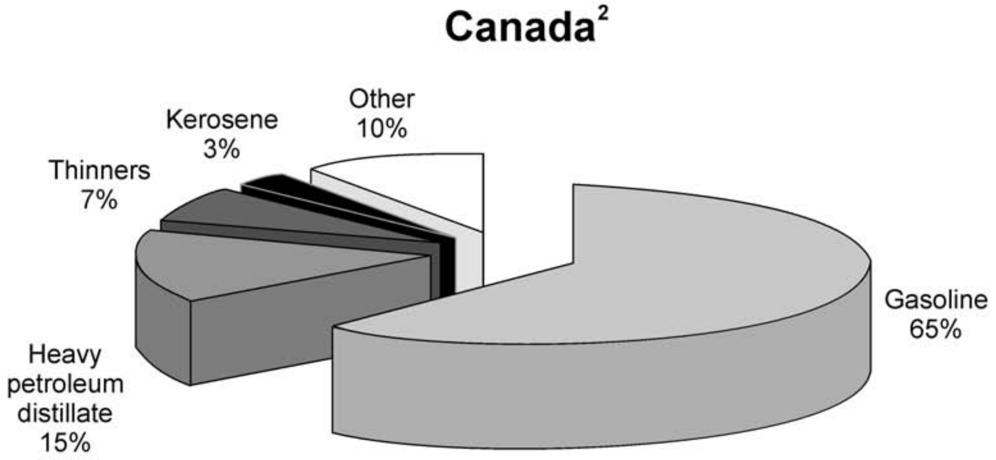
1.3.1 Review of the scientific literature

The comparison of petroleum products in order to determine whether two samples have a common origin has a long history in forensic science. One of the earliest publications in the forensic literature dealt with the comparison of motor oils using column chromatography and infrared spectroscopy.[18] Around the same time Cain compared samples of home heating kerosene for suspected adulteration.[19] Petroleum products such as gasoline, paint thinners, and kerosene are often used in arson to accelerate the fire with the intent of causing the greatest amount of damage in the least amount of time.[20] Petroleum products are, therefore, likely to be encountered in the investigation of suspicious fires and arson.

In the modern forensic laboratory the focus of fire investigation is on the analysis of fire debris for the presence of petroleum based products. The standard methods employed for fire debris analysis are well documented.[21,22,23,24,25,26,27,28,29] and the way in which two laboratories deal with this type of analysis has been recently described in the literature.[30,31] Although liquid samples are encountered in routine casework,[32,33] the analysis of liquid petroleum products is usually limited to their identification.[23] Two studies have reported that by far the most commonly found petroleum product in fire debris was gasoline (see Figure 1.1).[30,33] These findings are supported by this author's casework where 25 of 38 fire debris cases (65.8%) that tested positive for a petroleum product contained gasoline.[34] Mann[32] reported that for the fire debris cases that tested positive for gasoline, 20% also had a liquid sample available for comparison. In this author's casework seven of the 38 cases (18.4%) also had a liquid sample available for comparison.[34] It is, therefore, of particular interest to the forensic scientist to be able to compare two samples of gasoline. A forensic comparison between two samples of liquid gasoline is usually limited to finding differences between the samples. When a difference is found the samples are reported to have originated from different sources. When liquid-liquid comparisons of petroleum products are attempted, and no differences are found, the similarity between the two samples is

Figure 1.1 The occurrence of selected petroleum products found in fire debris.





1. Jackowski, J.P. The incidence of ignitable liquid residues in fire debris as determined by a sensitive and comprehensive analytical scheme. Journal of Forensic Sciences, 1997; 42(5): 828-832.

2. Sutherland, D.A. Fire debris analysis statistics and the use of the latest analytical tools. The/Le Journal: Canadian Association of Fire Investigators, 1999: 11-13.

reported as inconclusive. Some forensic science laboratories have made it a policy not to perform liquid-liquid petroleum comparisons in order to avoid this dilemma.

In the past it has been demonstrated that significant differences may be found between samples of leaded gasoline, [35,36,37] allowing strong conclusions to be drawn regarding the common origin of the samples found to have the same chemical composition. Chemical differences in unleaded gasoline samples are well known and have been reported in several publications.[38,39,40,41,42] Unfortunately, the methods described in these publications, while routinely used in the petroleum industry, are not directly transferable to a forensic laboratory. During the 1980s and 1990s, as leaded gasoline was phased out in many countries, new fingerprinting techniques used to compare gasoline samples emerged.[37,43,44,45,46,47] The observation that gasoline samples could be discriminated based on their chromatographic profile was the focus of neither Hennig's[43] nor Thatcher's[44] work and so was not pursued. The methods used by Hirz[37] and Mann[45] had an element of subjectivity because they involved a visual comparison between chromatograms. Hirz[37] only compared the chromatograms of unevaporated gasoline samples. The method described by Mann[45] was found to be suitable for gasoline samples that were evaporated to no more than half their original volume (*i.e.* no more than 50% evaporated by volume). The use of three-dimensional fluorescence spectroscopy to compare gasoline samples has been reported by Alexander *et al.*[46] and Sheff and Siegel.[47] The fluorescence equipment described is not commonly available in forensic science laboratories.[46,47] Furthermore, this spectroscopic method may not work well if the samples being compared are contaminated with soot containing large polycyclic aromatic hydrocarbons because these compounds will change the fluorescence characteristics of the sample.

1.3.2 Application of chemometrics and other methods for the discrimination of petroleum products

One of the early applications of chemometric methods to petrochemical data was for the discrimination of heavy fuel oils, crude oil residues and asphalts.[48] The petroleum industry has long been interested in calculating the octane number of a gasoline sample. In 1971

Walsh and Mortimer[49] divided the chromatogram of a gasoline sample into 31 time segments, assigned an empirical weight to each segment, then summed the values of each segment to determine the research octane number (RON) of the sample. A chemometric method was later introduced by Crawford and Hellmuth[50] to calculate the RON for a gasoline sample. These authors reduced a chromatogram containing approximately 30,000 data points to a data set containing 666 time segments, then used principal component regression to calculate the octane number. Finally, Ichikawa *et al.*[51] demonstrated that chemometrics could be used to differentiate between premium and regular unleaded gasoline samples.

In a forensic context, the discrimination of gasoline using chemometric methods was first explored in 1981 by Bertsch *et al.*.[52] This work was subsequently followed by the extensive research of Mayfield and Henley,[53] and later Lavine *et al.*,[54,55,56,57] into the classification of jet fuels. These authors were able to demonstrate that the judicious use of chemometrics could often allow petroleum products within the same class to be distinguished. Potter[58] applied a chemometric technique to discriminate between a small number of gasoline samples. Tan *et al.*[59] demonstrated the utility of chemometric methods to discriminate between petroleum products based on their class. In addition to using chemometrics, Andrews and Lieberman[60] also applied artificial neural networks to fluorescence spectra to discriminate petroleum products by class.

Rogers *et al.*[61] have shown that high-resolution mass spectrometry might be used to discriminate between and within classes of petroleum products. Smallwood *et al.*[62] have used small differences in carbon isotope ratios found to exist between different samples of gasoline as a way of comparing both evaporated and unevaporated samples. Neither high-resolution mass spectrometers, nor gas chromatography-carbon isotope ratio mass spectrometers are at the disposal of most forensic scientists. Forensic science laboratories around the world typically use gas chromatography-mass spectrometry (GC-MS) for fire debris analysis. The primary aims of this study are two-fold: to develop methods that will discriminate between samples of gasoline using GC-MS equipment already commonly used for fire debris analysis in forensic laboratories; and, to validate these methods objectively using chemometrics.

1.4 Conclusions

Arson is, by its very nature, a destructive crime. The cost of this crime to Australian society has been estimated to be in the order of hundreds of millions of dollars every year in lost property, in interruptions to business, and in fire fighting and fire prevention costs. Arson is a world wide problem. The arsonist often chooses to accelerate the fire by using a liquid petroleum product. The most commonly used accelerant is gasoline. The ability to uniquely characterise gasoline would allow a comparison between a gasoline sample obtained from a suspect to be made with a gasoline sample recovered from the fire scene. Where gasoline is used as an accelerant, such a comparison would prove to be very helpful in the implication or exoneration of a suspect. In the development of a comparison method, the repeatability of the analysis, the effects of evaporation on a sample, the type of equipment available to the forensic scientist, and the speed and simplicity of the analytical method(s) are some of the factors that must be taken into consideration.

1.5 References

1. Homicide trends in the U.S. Bureau of Justice Statistics, U.S. Department of Justice. <http://www.ojp.usdoj.gov.> Accessed September 20, 2002.

2. "Forum of arson investigators" seminar 2001. Arson Intelligence Newsletter. Arson Prevention Bureau, 2002: 50; 2.

3. Arson prosecution: issues and strategies. Federal Emergency Management Agency, United States Fire Administration. Report FA-78; June 1988.

4. Mukherjee, S.; Carcash, C.; Higgins, K. A statistical profile of crime in Australia. Australian Institute of Criminology Research and Public Policy Series, Australian Institute of Criminology, ACT; 1997.

5. Canadian Crime Statistics, 1999. Canadian Centre for Justice Statistics, Statistics Canada. Ottawa, Canada; 2000.

6. Canadian Crime Statistics, 2000. Canadian Centre for Justice Statistics, Statistics Canada. Ottawa, Canada; 2001.

7. Report of the New Zealand Police for the Year Ended 30 June 2001. New Zealand Police, Wellington, New Zealand.

8. Arson Prevention Bureau. Arson offences on the rise. Arson Intelligence Newsletter, 2000: 35; 1.

9. Crime in the United States 2000. Federal Bureau of Investigation, Criminal Justice Information Services Division. Clarksburg, WV, USA; 2001.

10. Measuring arson: recorded crime figures and police counting rules. Crime Reduction Toolkits – Arson. <www.crimereduction.gov.uk/toolkits.an020202.htm> Home Office Crime Reduction College. Accessed September 20, 2002.

11. King, M. Australian fire incident statistics 1992-1993. Division of Building, Construction and Engineering, CSIRO. Technical report TR95/1, Australia; 1995.

12. King, M. Australian national fire incident statistics 1993-1994. Division of Building, Construction and Engineering, CSIRO. Technical report TR97/2, Australia; 1997.

13. Roy, D. The cost of fires: a review of the information available. Home Office, London, UK; 1997.

14. Burns, L. School arson in New South Wales. NSW Bureau of Crime Statistics and Research, Sydney, Australia; 1991.

15. Arson Prevention Bureau. Safer communities: towards effective arson control. The Report of the Arson Scoping Study. Home Office, UK; 1998.

16. Fire Statistics United Kingdom, 2000. Department for Transport, Local Government and the Regions, Wetherby, United Kingdom. February 20, 2002.

17. Federal Bureau of Investigation. Crime in the United States 2000 Uniform Crime Reports. U.S. Department of Justice, Washington, D.C. 2002.

18. Yip, H.L. The identification of lubricating oils on clothing by column chromatography, infrared spectroscopy, and refractometry. Journal of Forensic Sciences, 1973; 18: 263-270.

19. Cain, P.M. Comparison of kerosenes using capillary column gas chromatography. Journal of the Forensic Science Society, 1975; 15: 301-308.

20. DeHann, J.D. Arson as a crime. In: Kirk's Fire Investigation, 4th Edition. Brady, Prentice Hall, New Jersey; 1997.

21. American Society for Testing and Materials. E1385-00 Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Steam Distillation. Philadelphia, USA.

22. American Society for Testing and Materials. E1386-00 Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Solvent Extraction. Philadelphia, USA.

23. American Society for Testing and Materials. E1387-01 Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography. Philadelphia, USA.

24. American Society for Testing and Materials. E1388-00 Standard Practice for Sampling of Headspace Vapors from Fire Debris Samples. Philadelphia, USA.

25. American Society for Testing and Materials. E1389-95 Standard Practice for Cleanup of Fire Debris Sample Extracts by Acid Stripping. Philadelphia, USA.

26. American Society for Testing and Materials. E1412-00 Standard Practice for Separation of Ignitable Liquid Residues from Fire Debris Samples by Passive Headspace Concentration with Activated Charcoal. Philadelphia, USA.

27. American Society for Testing and Materials. E1413-00 Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Dynamic Headspace Concentration. Philadelphia, USA.

28. American Society for Testing and Materials. E1618-01 Standard Guide for Identification of Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry. Philadelphia, USA.

29. American Society for Testing and Materials. E2154-01 Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Passive Headspace Concentration with Solid Phase Microextraction (SPME). Philadelphia, USA.

30. Jackowski, J.P. The incidence of ignitable liquid residues in fire debris as determined by a sensitive and comprehensive analytical scheme. Journal of Forensic Sciences, 1997; 42(5): 828-832.

31. Touron, P.; Malaquin, P.; Gardebas, D.; Nicolai, J.P. Semi-automatic analysis of fire debris. Forensic Science International, 2000; 110(1): 7-18.

32. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography II: limitations of automotive gasoline comparisons in casework. Journal of Forensic Sciences, 1987; 32(3): 616-628.

33. Sutherland, D.A. Fire debris analysis statistics and the use of the latest analytical tools. The/Le Journal: Canadian Association of Fire Investigators, 1999: 11-13.

34. Sandercock, P.M.L. Royal Canadian Mounted Police Forensic Laboratory, Winnipeg. Unpublished results from 1993-1997 casework.

35. Frank, H.A. Lead alkyl components as discriminating factors in the comparison of gasolines. Journal of the Forensic Science Society, 1980; 20(4): 285-292.

36. Chan, L. The determination of tetraalkyl lead compounds in petrol using combined gas chromatography atomic absorption spectrometry. Forensic Science International, 1981; 18: 57-62.

37. Hirz, R. Gasoline brand identification and individualization of gasoline lots. Journal of the Forensic Science Society, 1989; 29(2): 91-101.

38. Johnson, G.V. Routine analysis of oxygenates and benzene in retail motor fuel: a survey by the state of Alabama. Journal of Chromatographic Science, 1987; 25: 65-70.

39. Canadian Petroleum Products Institute. Composition of Canadian summer and winter gasolines 1993. CPPI Report No.94-5, Ottawa, ON, Canada; June 1994.

40. Kaplan, I.R.; Galperin, Y.; Alimi, H.; Lee, R.P.; Lu, S.T. Patterns of chemical changes during environmental alteration of hydrocarbon fuels. Ground Water Monitoring & Remediation, 1996; 16(4): 113-124.

41. Kaplan, I.R.; Galperin, Y.; Lu, S.T.; Lee, R.P. Forensic environmental geochemistry - differentiation of fuel-types, their sources and release time. Organic Geochemistry, 1997; 27(5-6): 289-317.

42. Patel, V. Petroleum fuel quality in New Zealand 1992-1999. Ministry of Economic Development, Wellington, New Zealand. July; 2000.

43. Hennig, H.J. Möglichkeiten der differenzierung von mineralölen, gasölen und vergaserkraftstoffen mit hilfe der gaschromatographie. Archive Fur Kriminologie, 1982; 170: 12-20.

44. Thatcher, P.J. The identification of petroleum residues in arsons. Thesis (Ph.D.) University of Melbourne; 1982.

45. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography I: Comparison methodology. Journal of Forensic Sciences, 1987; 32(3): 606-615.

46. Alexander, J.; Mashak, G.; Kapitan, N.; Siegel, J.A. Fluorescence of petroleum products II. Three-dimensional fluorescence plots of gasolines. Journal of Forensic Sciences, 1987; 32(1): 72-86.

47. Sheff, L.M. and Siegel, J.A. Fluorescence of petroleum products V: Three-dimensional fluorescence spectroscopy and capillary gas chromatography of neat and evaporated gasoline samples. Journal of Forensic Sciences, 1994; 39(5): 1201-1214.

48. Kawahara, F.K. and Yang, Y.Y. Systems chemical analysis of petroleum pollutants. Analytical Chemistry, 1976; 48(4): 651-655.

49. Walsh, R.P. and Mortimer, J.V. New way to test product quality. Hydrocarbon Processing, 1971; 50: 153-158.

50. Crawford, N.R. and Hellmuth, W.W. Refinery octane blend modelling using principal components regression of gas chromatography data. Fuel, 1990; 69(4): 443-447.

51. Ichikawa, M.; Nonaka, N.; Takada, I.; Ishimori, S. Mass spectrometric analysis for distinction between regular and premium motor gasolines. Analytical Sciences, 1993; 9: 261-266.

52. Bertsch, W.; Mayfield, H.; Thomason, M.M. Application of pattern recognition to high resolution GC and GC/MS. Part I. Basic studies. Proceedings of the International Symposium on Capillary Chromatography, 1981: 313-334.

53. Mayfield, H.T. and Henley, M.V. Classification of jet fuels using high resolution gas chromatography and pattern recognition, ASTM Special Technical Publications, 1991; 1102: 578-597.

54. Lavine, B.K.; Stine, A.; Mayfield, H.T. Gas chromatography-pattern recognition techniques in pollution monitoring. Analytica Chimica Acta, 1993; 277: 357-367.
55. Lavine, B.K.; Mayfield, H.T.; Kromann, P.R.; Faruque, A. Source identification of underground fuel spills by pattern recognition analysis of high-speed gas chromatograms. Analytical Chemistry, 1995; 67: 3846-3852.

56. Lavine, B.K.; Moores, A.J.; Mayfield, H.T.; Faruque, A. Fuel spill identification by gas chromatography genetic algorithms/pattern recognition techniques. Analytical Letters, 1998; 31(15): 2805-2822.

57. Lavine, B.K.; Ritter, J.; Moores, A.J.; Wilson, M.; Faruque, A.; Mayfield, H.T. Source identification of underground fuel spills by solid-phase microextraction/high-resolution gas chromatography/genetic algorithms. Analytical Chemistry, 2000; 72: 423-431.

58. Potter, T.L. Fingerprinting petroleum products: unleaded gasoline. In: Petroleum Contaminated Soils; Volume 3. Lewis Publishers, Chelsea, MI; 1990.

59. Tan, B.J.; Hardy, J.K.; Snavely, R.E. Accelerant classification by gas chromatography/mass spectrometry and multivariate pattern recognition. Analytica Chimica Acta, 2000; 422(1): 37-46.

60. Andrews, J.M. and Lieberman, S.H. Neural network approach to qualitative identification of fuels and oils from laser induced fluorescence spectra. Analytica Chimica Acta, 1994; 285: 237-246.

61. Rodgers, R.P.; Blumer, E.N.; Freitas, M.A.; Marshall, A.G. Compositional analysis for identification of arson accelerants by electron ionization Fourier transform ion cyclotron resonance high-resolution mass spectrometry. Journal of Forensic Sciences, 2001; 46(2): 268-279.

62. Smallwood, B.J.; Philp, R.P.; Allen, J.D. Stable carbon isotope composition of gasolines determined by isotope ratio monitoring gas chromatography mass spectrometry. Organic Geochemistry, 2002; 33(2): 149-159.

Chapter 2. Gasoline distribution network in Australia

2.1 Introduction

The forensic examination of physical evidence may be divided into two general parts:

- the identification and comparison of two or more man-made materials (*e.g.* samples of gasoline from an arson investigation); and,
- the interpretation of the results.

The results obtained during the first part of the examination must be interpreted in the context of how the materials are distributed to and, eventually, consumed by the public. When a gasoline comparison is made during an arson investigation, the forensic scientist must have knowledge of the fuel distribution system in order to provide a context in which to interpret the results of the comparison.

Recognising the need to understand the refining process and the supply chain to service stations, Hirz[1] gave a brief description of the distribution of gasoline from refineries to service stations in parts of Europe. Mann[2] also recognised that the distribution mechanism may affect the results of a gasoline comparison and so intentionally sought out gasoline samples delivered from a single distribution terminal. There is little published information available to the forensic scientist interested in understanding gasoline distribution from the refinery to the service station. This chapter discusses oil refining and the distribution of gasoline in the context of Australia and the state of New South Wales. Many of the aspects of modern gasoline distribution in Australia are mirrored in countries and regions around the world.

2.2 Materials and methods

By agreement with Caltex Australia, delivery schedules were obtained for all deliveries from the Caltex Banksmeadow distribution terminal during the 30 working day period from August 31 to October 1, 2001. This time period was chosen at the discretion of Caltex Australia and was selected to represent, as much as possible, a typical delivery schedule free of delivery anomalies due to public and/or school holidays. Delivery schedules were received from Caltex Australia on December 12, 2001.

Pertinent information from the schedules was scanned into a text file and parsed into a spreadsheet (Microsoft Excel®). The information was then checked manually against the originals and corrections were made. At Caltex Australia's request, the delivery locations were not identified and all data was kept in confidence.

2.3 Gasoline distribution network

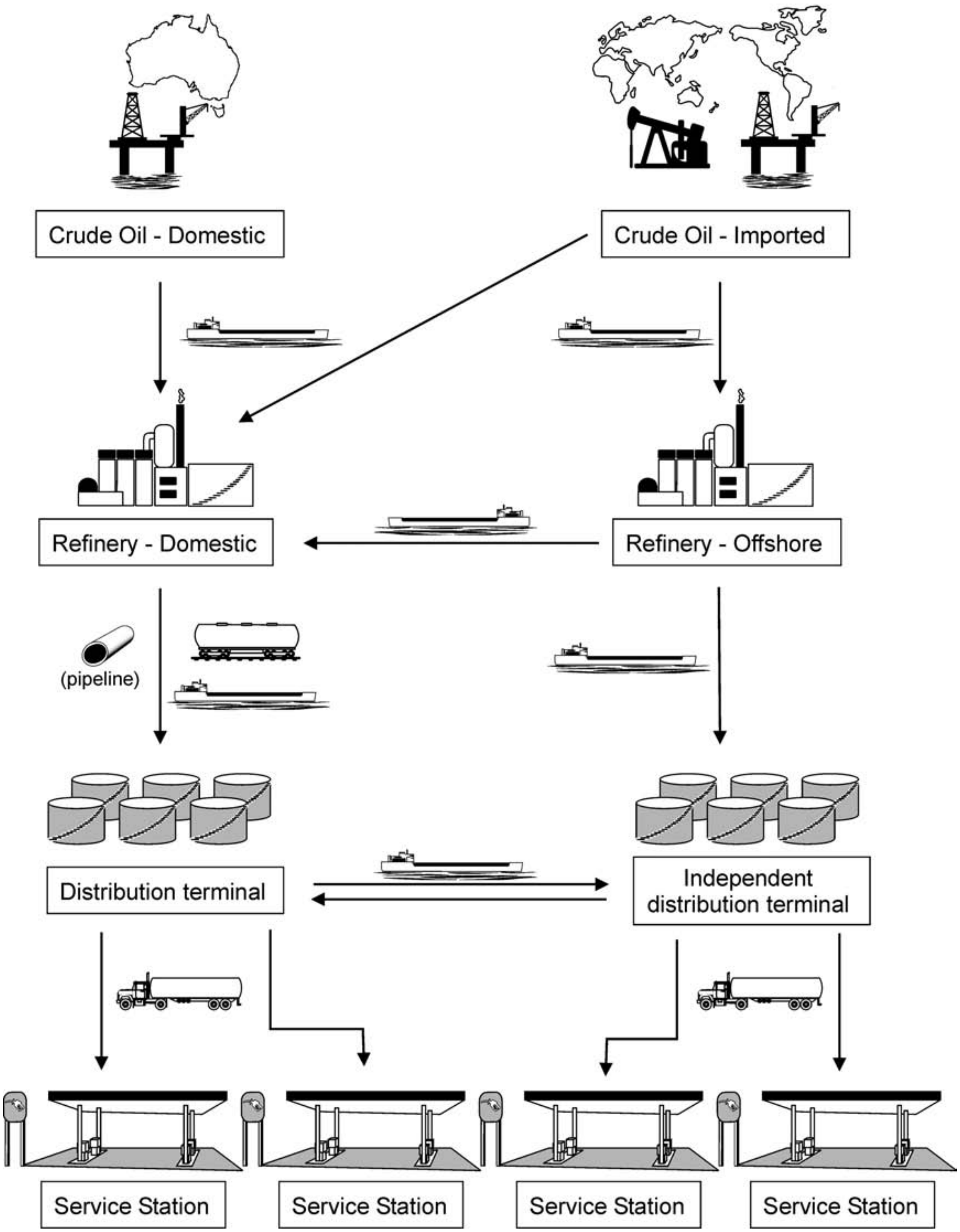
2.3.1 Overview

The composition of a gasoline sold to the public at a service station will be determined by such factors as the source of the crude oil, the refining processes used (including blending), the mixing of product at the bulk distribution terminal and, finally, the mixing of the product in the underground storage tank at the service station. Each of these steps may potentially alter the chemical composition of the final product purchased by the consumer and may contribute to the uniqueness of a gasoline sold at one service station compared to gasoline sold at another service station down the street. The steps taken to get from crude oil to gasoline at a service station forms the distribution network (Figure 2.1). Knowledge of this network will help the forensic scientist understand how differences between gasoline samples arise.

2.3.2 Refineries

Australian refineries receive crude oil for processing from either domestic sources (typically oil wells in offshore basins) or by importing crude oil from other countries. Historically Australia has always imported some crude oil and other refinery feedstock for processing in domestic refineries (Table 2.1). For example, in 2001 approximately 30% of the crude oil processed by the Caltex refinery at Kurnell was from Australian oil fields with

Figure 2.1 Flow diagram of distribution network for gasoline in Australia.



the balance being imported from primarily Indonesia and Vietnam with some from Saudi Arabia.[3]

Table 2.1 Top ten countries from which Australia imported crude oil and other refinery feedstock, 1995 to 1998.

1995[1]	*1996*[2]	*1997*[3]	*1998*[4]
Saudi Arabia	Indonesia	Indonesia	Indonesia
Papua New Guinea	Papua New Guinea	Saudi Arabia	Vietnam
United Arab Emirates	United Arab Emirates	United Arab Emirates	Saudi Arabia
Indonesia	Saudi Arabia	Papua New Guinea	Papua New Guinea
Vietnam	Vietnam	Vietnam	United Arab Emirates
Qatar	New Zealand	New Zealand	New Zealand
New Zealand	Singapore	Qatar	Malaysia
Singapore	Qatar	Malaysia	Qatar
Malaysia	Malaysia	Nigeria	China
Kuwait	Nigeria	Congo	Iraq

1. Oil and Australia statistical review. Petroleum Gazette. 1996; 3: 16.
2. Oil and Australia statistical review. Petroleum Gazette. 1997; 3: 16.
3. Australian petroleum in facts and figures. Statistical review. Petroleum Gazette, 1998; 3: 42.
4. Australian petroleum in facts and figures. Statistical review. Petroleum Gazette, 1999; 3: 46.

Note that *Petroleum Gazette* ceased publication after 1999.

Australia currently has eight refineries producing gasoline. These refineries are owned and operated by the four major oil companies: BP, Caltex, Mobil, and Shell. The locations and output of these refineries are listed in Table 2.2 and the locations are shown in Figure 2.2. All Australian refineries are relatively old which means that, compared to newer refineries, they are also relatively small. All Australian refineries have upgraded their refining processes since they were first built and now refine crude oil using the same processes as more recently constructed refineries elsewhere in the world. Australia also imports refined products such as gasoline from offshore refineries. These imports may be received by an Australian refinery for re-processing to meet company standards, or are shipped directly to an independent bulk distribution terminal for delivery to independent retailers such as Woolworths.[4]

Figure 2.2 Location of Australian refineries and seaboard distribution terminals. See Table 2.2 and Table 2.4 for further details.

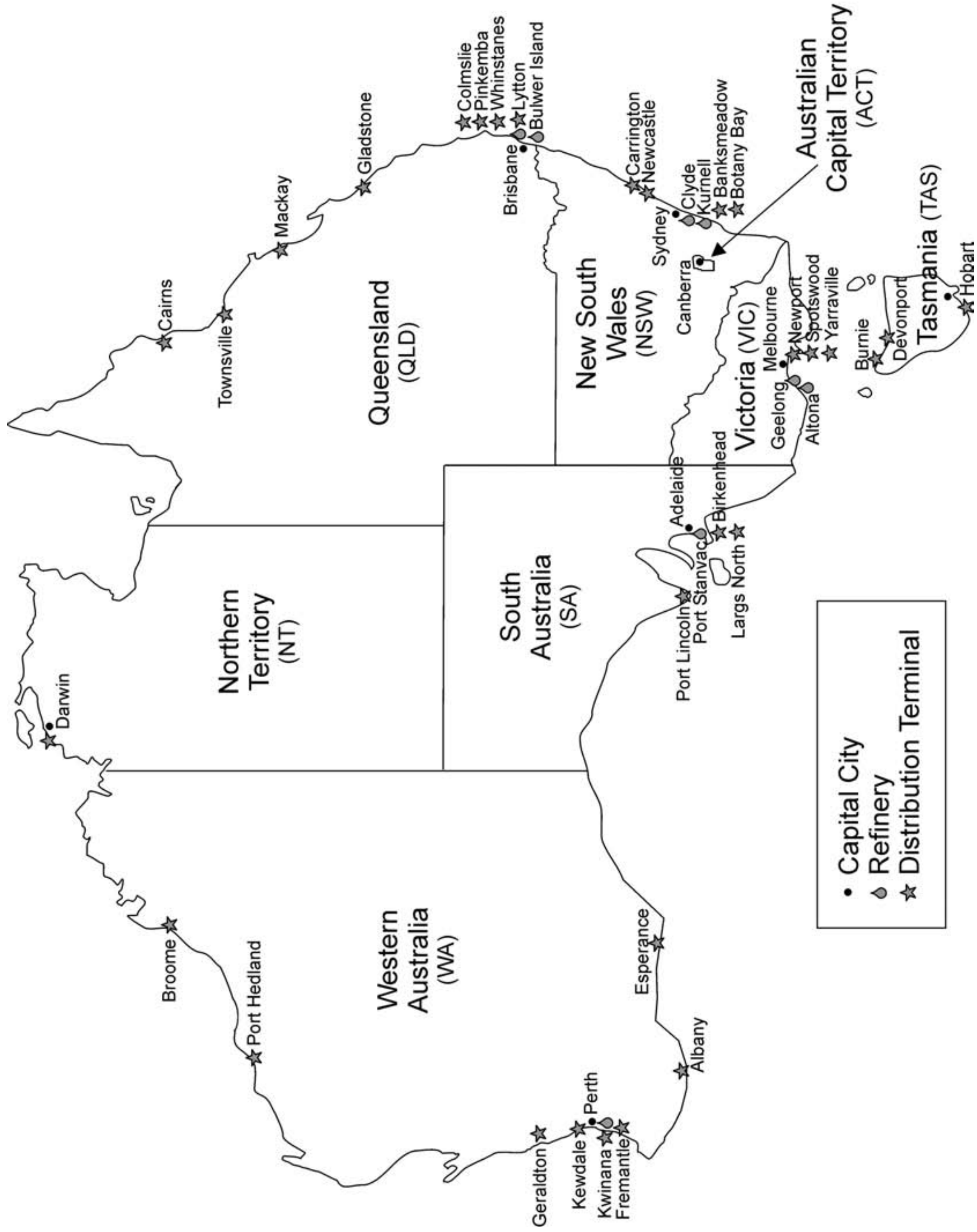


Table 2.2 Oil refineries in Australia (data current as of December 31, 1998).

State	Details	Major oil companies			
		BP	Caltex	Mobil	Shell
NSW	Location		Kurnell		Clyde
	Maximum capacity†		18,444,000		13,674,000
	Year plant first opened		1956		1928
QLD	Location	Bulwer Island	Lytton		
	Maximum capacity†	11,686,500	15,900,000		
	Year plant first opened	1965	1965		
SA	Location			Port Stanvac	
	Maximum capacity†			12,402,000	
	Year plant first opened			1963	
VIC	Location			Altona	Geelong
	Maximum capacity†			21,465,000	18,921,000
	Year plant first opened			1949	1954
WA	Location	Kwinana			
	Maximum capacity†	22,021,500			
	Year plant first opened	1955			
	Total maximum capacity†	33,708,000	34,344,000	32,595,000	33,867,000

† Maximum capacity is typically measured in units of b/sd (barrels per stream day of crude oil). Data converted to litres per day using the conversion of 159 litres per barrel. Source: Measures and conversions. Petroleum Communication Foundation. <http://www.pcf.ab.ca/> Accessed August 22, 2002.

Note that refineries typically operate, on an annual basis, at 85% to 88% of their maximum capacity. Source: Australian Institute of Petroleum. Australian Refinery Capacity. <http://www.aip.com/info> Accessed January 17, 2001.

The processing of crude oil into gasoline at a modern refinery is a complex, multi-step process. A crude oil is first separated into different boiling range fractions by distillation. The distillation fractions then undergo “conversion” where they are further refined along different processing streams that will either crack, reform, alkylate, polymerise, or isomerise the fraction to achieve the desired composition (see Table 2.3).[5,6] The next step, purification, is employed to remove impurities (*e.g.* metals) or undesirable compounds (*e.g.* sulphur). A refinery will typically have several processing streams for each refining process (*i.e.* more than one cracking or alkylation unit operating simultaneously). The final product, gasoline, is created by blending different processing streams together. The number of streams blended, and the amount of each product from each stream, is dependent on the requirements of the finished gasoline and will typically involve the blending of six or more streams.[5] The Marsden Point refinery in New Zealand (opened in 1964) has reported that it blends gasoline from three to seven processing streams.[7]

Table 2.3 Refinery conversion processes used to make gasoline.

Name	*Conversion process*[1]
Cracking	Breaking down of larger hydrocarbons into smaller ones.
Reforming	Conversion of straight-chain hydrocarbons into aromatic compounds to increase octane rating.
Alkylation	Combining of small, gaseous C_3- to C_5-hydrocarbons to form higher octane branched hydrocarbons (*e.g.* trimethylpentanes).
Polymerisation	Polymerisation of short chain unsaturated hydrocarbons (*e.g.* polyproplyene) into longer chain unsaturated hydrocarbons (*e.g.* C_6-, C_9-, C_{12}-olefins). Products of polymerisation must be hydrolysed to yield saturated hydrocarbons prior to blending into gasoline.
Isomerisation	Conversion of straight-chain hydrocarbons (typically n-C_5 and n-C_6) into branched-chain isomers to increase octane rating.

1. Motor Gasolines Technical Review. Chevron Products Company, Chevron USA Inc. 1996.

2.3.3 Distribution terminals

Bulk distribution terminals are storage facilities where refined products such as gasoline and diesel fuel are held pending distribution to market. The main distribution terminals in Australia are owned and operated by one of the four major oil companies (BP, Caltex, Mobil or Shell), by two of the four major oil companies in a joint venture, or by an independent company. The distribution terminals owned and operated by the four major oil companies receive shipments of fuel and other refined products by pipeline, ship or rail. Delivery to smaller, inland distribution terminals is typically made by road or rail. The independent distribution terminals are located in ports and receive shipments of fuel exclusively by ship.[8] A listing of the bulk distribution terminals along the Australian seaboard is given in Table 2.4. It has been estimated that there are approximately 600 inland wholesalers operating in Australia who distribute motor fuels.[9] For the sake of brevity, the many smaller, inland distribution terminals have not been listed.

Table 2.4 Gasoline bulk distribution terminals along the Australian seaboard.

State	*Bulk terminal location by oil company*†				
	BP[1]	*Caltex*[2]	*Mobil*[3]	*Shell*[4]	*Independents*[‡,5]
NSW	Carrington	Banksmeadow Newcastle Silverwater (Mobil)†	Botany Bay Silverwater	Parramatta	Botany Bay
NT	Darwin	Darwin	Darwin	Darwin	
QLD	Gladstone Mackay Townsville Whinstanes	Cairns Gladstone (Mobil)† Lytton Mackay Townsville (Shell)†	Colmslie	Pinkenba	Brisbane
SA	Largs North	Birkenhead Port Lincoln	Birkenhead	Birkenhead	
TAS	Burnie Hobart	Hobart (Shell)†		Devonport (Caltex)†	
VIC		Newport	Spotswood Yarraville	Newport	Hastings
WA	Broome E. Fremantle Esperance Geraldton Kewdale N. Fremantle Port Hedland	Albany Fremantle (Shell)† Port Hedland		N. Fremantle	Kwinana

† Where disclosed, the major oil companies that have an agreement to share facilities and/or supply fuel are indicated in parentheses.

‡ The owner/operator of each independent seaboard distribution terminal is: Botany Bay and Hastings by Trafigura/Roayl Vopak; Brisbane by Neumanns; and, Kwinana by Gull Petroleum.

1. BP Terminals, BP Australia. <http://www.bp.com.au> Accessed July 11, 2002.
2. Caltex Terminals, Caltex Australia. <http://www.caltex.com.au> Accessed September 14, 2001.
3. Mobil Terminals, Mobil Australia. <http://www.mobil.com.au> Accessed September 14, 2001.
4. Shell Terminals, Shell Australia. <http://www.shell.com.au> Accessed September 14, 2001.
5. Personal communication. Mr. Len Daly, Director, Royal Vopak Distribution Terminal, Botany

In those states and territories where one of the four major oil refining companies does not operate an oil refinery (Table 2.2), a product exchange arrangement is entered into for fuel to be delivered to that company's distribution terminal from another company's refinery. Once at the distribution terminal, proprietary additives may be injected into the fuel while it is being loaded into the tanker truck prior to delivery to a service station.[10]

2.3.4 Service stations

There are over 8000 service stations across Australia (Table 2.5). The number of service stations in each state or territory is related to the size of the population being served. The number of rural and urban service stations in Australia, by state or territory is listed in Table 2.5. Service stations receive delivery of fuel from the distribution terminal by truck. The most common trailer used for delivering fuel has a barrel (tank) with a capacity of approximately 40,000 litres and is known as a "maxi barrel" (see Figure 2.3) The barrel on a maxi barrel trailer is divided into five equal compartments, thus each compartment holds approximately 8,000 litres of fuel.

For larger volume deliveries to more remote areas, two smaller trailers are coupled together and are known as "B-doubles".[11] B-double trailers have a combined capacity of 50,000 to 55,000 litres.[10] Smaller capacity trucks have a barrel fixed directly to the chassis. These rigid frame trucks, without a fifth wheel, have a capacity ranging from 10,000 to 27,000 litres and are occasionally used to haul fuel to smaller independent retailers.[10]

Gasoline is stored in underground storage tanks located under the service station forecourt (the filling area). Underground storage tank capacity will vary depending on such factors as the number of pumps, the product throughput of the service station, and the age of the service station's storage tanks. The Steel Tank Institute in the United States has reported that in 1999, the average capacity of underground gasoline storage tanks in the U.S. was approximately 37,800 litres (10,000 US gallons); nearly double the mean capacity of the tanks used in the mid to late 1980s, which was approximately 20,000 litres (5,500 US gallons).[12]

Bay, NSW. September 13, 2001.

Figure 2.3 Diagram of maxi barrel trailer for transporting liquid fuels. Barrel exterior is shown (top) as well as cut-away view of the five compartments inside (bottom). Drawings are based on information contained in US Patents 3,131,949 and 3,187,766, and personal communication with Mr. John Morgan, Banksmeadow distribution terminal, Caltex Australia Limited.

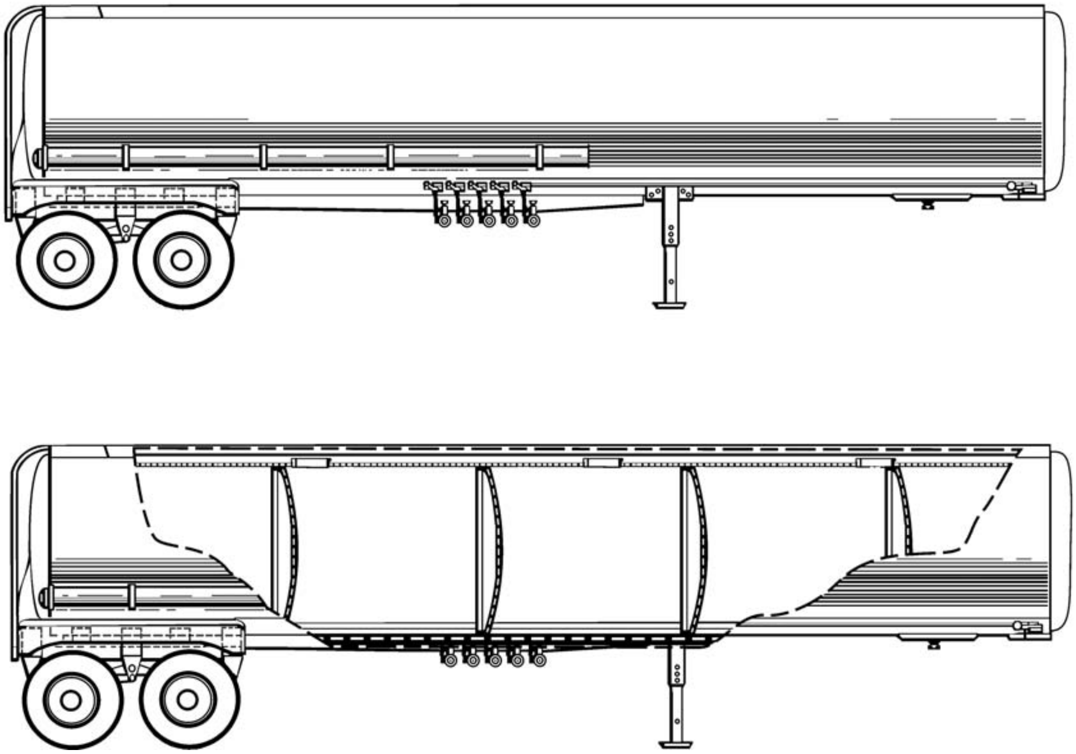


Table 2.5 Number of rural and urban service stations in Australia by state or territory (current as of January 1, 2001).

State or territory	*Number of service stations*[1]	
NSW	1632	
VIC	1038	
QLD	925	
SA	495	
WA	465	
TAS	299	
NT	84	
Total number of rural stations	4938	4938
NSW	1051	
VIC	928	
QLD	596	
WA	383	
SA	312	
ACT†	75	
TAS	62	
NT	25	
Total number of urban stations	3432	3432
Total number of service stations in Australia		8370

† All service stations in the ACT were considered urban.

1. Service station survey. Australian Institute of Petroleum. <http://www.aip.com.au> Accessed August, 20 2002.

The amount of gasoline consumed in Australia in recent years is shown in Table 2.6. The two types of higher octane gasoline, premium unleaded and lead replacement, have been grouped together. It is anticipated by the Australian petroleum refineries that lead replacement gasoline will be phased out of the market over the next five years.[13] In 1990 nearly 70% of all gasoline consumed in Australia contained lead.[14] The market share of leaded gasoline had dropped to 28% in 1998,[15] and was completely phased out of the Australian market by January 1, 2002.[16] Leaded gasoline was banned in Western Australia on January 1, 2000 and in Queensland on March 1, 2001.[16] Approximately 70% of all

Table 2.6 Gasoline consumption in Australia by state/territory, 1995-1998. Volumes are in megalitres.

Regular unleaded gasoline

	NSW	*VIC*	*QLD*	*WA*	*SA*	*TAS*	*NT*	Total
1995[1]	3542.5	2435.1	1999.7	1028.9	712.2	218.3	101.5	10038.2
1996[2]	3791.1	2644.8	2189.5	1120.8	755.3	237.4	108.5	10847.4
1997[3]	3982.7	2849.0	2341.0	1203.8	813.1	253.1	122.4	11565.1
1998[4]	4326.7	3069.1	2420.6	1290.5	882.4	268.3	127.4	12385.0

Premium unleaded and lead replacement gasoline

	NSW	*VIC*	*QLD*	*WA*	*SA*	*TAS*	*NT*	Total
1995[1]	116.0	45.9	52.6	32.0	11.2	2.8	0.0	260.5
1996[2]	153.5	59.8	66.5	40.9	13.8	3.8	0.1	338.4
1997[3]	190.9	76.1	82.7	48.8	16.5	5.4	0.4	420.8
1998[4]	234.4	99.8	95.5	61.1	18.7	8.6	1.2	519.3

Leaded gasoline[†]

	NSW	*VIC*	*QLD*	*WA*	*SA*	*TAS*	*NT*	Total
1995[1]	2219.6	1989.5	1497.9	772.0	652.2	254.6	85.9	7471.7
1996[2]	2001.1	1799.4	1387.3	700.7	595.5	239.8	57.7	6781.5
1997[3]	1693.1	1577.4	1221.5	616.1	537.8	217.4	52.0	5915.3
1998[4]	1456.0	1358.7	1050.7	536.2	489.8	193.8	43.9	5129.1

† Leaded gasoline is no longer marketed in Australia. See: National phase out of leaded petrol. Environment Australia. <http://www.ea.gov.au/atmosphere/airquality/lead/qu.htlm> Accessed October 30, 2002.

1. Oil and Australia statistical review. Petroleum Gazette. 1996; 3: 10-11.
2. Oil and Australia statistical review. Petroleum Gazette. 1997; 3: 10-11.
3. Australian petroleum in facts and figures. Statistical review. Petroleum Gazette, 1998; 3: 36.
4. Australian petroleum in facts and figures. Statistical review. Petroleum Gazette, 1999; 3: 40.

Note that *Petroleum Gazette* ceased publication after 1999.

gasoline consumed in Australia is sold to the public through retail outlets (service stations). The volume and market share of each type of gasoline, together with the total volume sold through service stations, is given in Table 2.7. Gasoline that was not sold through retailers to the general public for use in motor vehicles was distributed to and consumed by other users such as the fishing industry and marinas (for pleasure vessels), as well as to organisations such as government road maintenance departments and the military.

Table 2.7 Market share of gasoline consumed in Australia by type, including total volume sold through service stations (retailers). Volumes are in megalitres.

All gasoline consumed in Australia by type

	1995[1]		1996[2]		1997[3]		1998[4]	
	Volume	*%*	*Volume*	*%*	*Volume*	*%*	*Volume*	*%*
regular unleaded	10038.2	56.5%	10847.4	60.4%	11565.1	64.6%	12385	68.7%
premium/lead replacement	260.5	1.5%	338.4	1.9%	420.8	2.4%	519.3	2.9%
leaded	7471.7	42.0%	6781.5	37.7%	5915.3	33.0%	5129.1	28.4%
Total	17770.4		17967.3		17901.2		18033.4	

All gasoline consumed in Australia by consumer

	1995[1]		1996[2]		1997[3]		1998[4]	
	Volume	*%*	*Volume*	*%*	*Volume*	*%*	*Volume*	*%*
total sales to retailers	12197.6	68.6%	12208.5	67.9%	12261.2	68.5%	12109.8	67.2%
total sales to non-retailers	5572.8	31.4%	5758.8	32.1%	5640.0	31.5%	5923.6	32.8%
Total	17770.4		17967.3		17901.2		18033.4	

1. Oil and Australia statistical review. Petroleum Gazette. 1996; 3: 10-11.
2. Oil and Australia statistical review. Petroleum Gazette. 1997; 3: 10-11.
3. Australian petroleum in facts and figures. Statistical review. Petroleum Gazette, 1998; 3: 36.
4. Australian petroleum in facts and figures. Statistical review. Petroleum Gazette, 1999; 3: 40.

Note that *Petroleum Gazette* ceased publication after 1999.

2.4 Caltex Banksmeadow distribution terminal

It was an aim of this study to obtain information about gasoline deliveries to service stations in order to help forensic scientists understand the number, frequency, and volume of fuel deliveries to retail and commercial outlets and the concomitant changes in gasoline composition. This information could be used to help direct police during an investigation and could assist the court of law in interpreting evidence. Initially gasoline distribution information was sought from retailers at the time a sample was collected from the pump. This method of data collection proved to be inaccurate or did not work because some service station employees or managers were unwilling to discuss this information, or did not have the information readily available. Caltex Australia was approached for assistance and they agreed to provide the schedules for all deliveries made from the Caltex Banksmeadow distribution terminal during the 30 working day period from August 31 to

October 1, 2001. This distribution terminal is located on Botany Bay and is opposite its supplier, the Caltex Kurnell refinery.

2.4.1 Number of deliveries and delivery locations

Fuel deliveries from the Caltex Banksmeadow distribution terminal were made six days per week from Monday to Saturday. During the 30 working day period from August 31 to October 1, 2001, a total of 2,199 deliveries of both gasoline and diesel fuel were made from this terminal: 2,029 (92.3%) to retail sites (*i.e.* service stations), and 170 (7.7%) to commercial enterprises (*e.g.* freight haulers). Although the majority of retail gasoline and diesel fuel deliveries (1861 or 91.7%) were made in the Sydney metropolitan area (urban NSW) some deliveries were made elsewhere in the state of New South Wales (rural NSW) (Table 2.8). The occasional delivery was also made into the Australian Capital Territory and into the state of Victoria (Table 2.8). Most commercial fuel deliveries were within the Sydney metropolitan area (Table 2.9). All commercial fuel deliveries were confined to the state of New South Wales. The locations of the various districts within the state of New South Wales are shown on the map in Figure 2.4.

2.4.2 Delivery frequency

The number of deliveries made by the truck fleets operating from the Caltex Banksmeadow distribution terminal was examined for any trends that may assist in the collection of samples or the interpretation of results during an investigation. For the purpose of this study, a delivery is defined as a truck movement to a delivery site and back to the distribution terminal. At the distribution terminal a truck fills its barrel with the fuel(s) required for a particular delivery. The truck is driven to a single site, discharges the fuel(s) and returns to the terminal. It is standard practise for a truck to deliver to one site only and then return to the terminal. A delivery of one or multiple fuel types, by one truck, to a single site was counted as one delivery. An example of a multiple fuel delivery would be a maxi barrel trailer with three compartments of regular unleaded gasoline, one of premium unleaded gasoline, and one of diesel fuel. One site (*i.e.* a retail service station) may receive

Table 2.8 Retail fuel deliveries (gasoline and diesel) made from the Caltex Banksmeadow distribution terminal into local council areas from August 31 to October 1, 2001.

Australian Capital Territory
39 deliveries

	Local council area	*State*
1	Canberra Central	ACT
2	Majura	ACT
3	Tuggeranong	ACT

New South Wales – rural
127 deliveries

	Local council area	*District‡*
4	Blue Mountains	SS
5	Gosford	SS
6	Goulburn	SE
7	Hawkesbury	SS
8	Queanbeyan	SE
9	Wagga Wagga	MB
10	Wollondilly	SS
11	Wyong	SS

New South Wales – urban
1861 deliveries

	Local council area	*District‡*
12	Ashfield	IS
13	Auburn	OS
14	Bankstown	OS
15	Baulkham Hills	OS
16	Blacktown	OS
17	Botany Bay	IS
18	Burwood	IS
19	Camden	OS
20	Campbelltown	OS
21	Canada Bay	IS
22	Canterbury	IS
23	Fairfield	OS
24	Holroyd	OS
25	Hornsby	OS
26	Hurstville	IS
27	Kogarah	IS
28	Ku-ring-gai	OS
29	Lane Cove	IS
30	Leichhardt	IS
31	Liverpool	OS
32	Manly	IS
33	Marrickville	IS
34	Mosman	IS
35	North Sydney	IS
36	Parrmatta	OS
37	Penrith	OS
38	Pittwater	OS
39	Randwick	IS
40	Rockdale	IS
41	Ryde	OS
42	South Sydney	IS
43	Strathfield	IS
44	Sutherland	OS
45	Sydney	IS
46	Warringah	OS
47	Waverley	IS
48	Willoughby	IS
49	Woolahra	IS

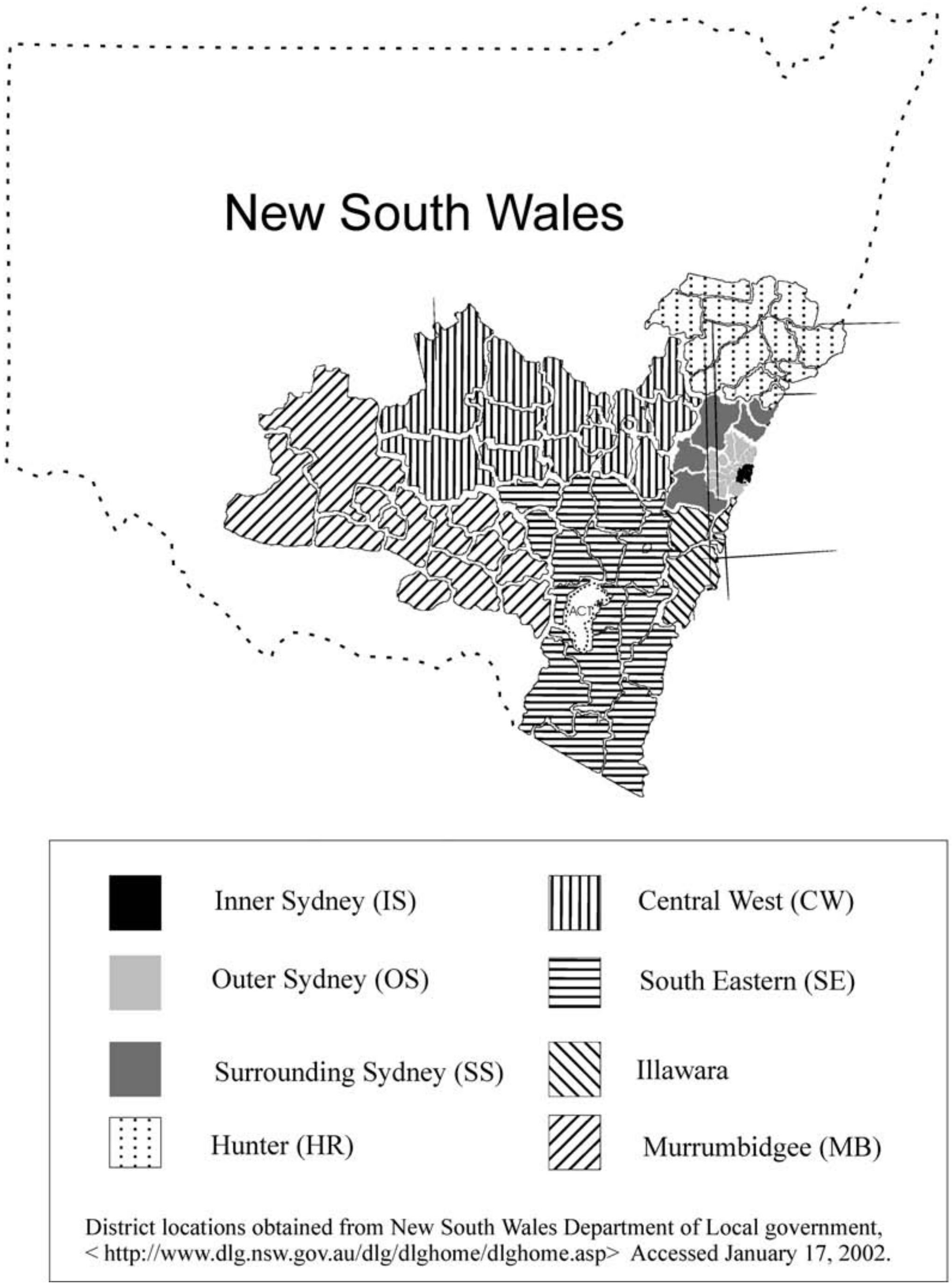
Victoria
2 deliveries

	Local council area	*State*
50	Monash	VIC

‡District abbreviations: IS = inner Sydney; MB = Murrumbidgee; OS = outer Sydney; SE = South East; SS = surrounding Sydney.

Local council area boundaries obtained from NSW Department of Local Government, Local Council Search. <http://www.dlg.nsw.gov.au/dlg/dlghome/dlg_home.asp> Accessed May 10, 2002.

Figure 2.4 Map showing districts of NSW into which retail and commercial deliveries of gasoline and diesel fuel were made from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. Note that no deliveries were made into the Illawara district during this time period.



District locations obtained from New South Wales Department of Local government, < http://www.dlg.nsw.gov.au/dlg/dlghome/dlghome.asp> Accessed January 17, 2002.

Table 2.9 Commercial fuel deliveries (gasoline and diesel) made from the Caltex Banksmeadow distribution terminal into local council areas from August 31 to October 1, 2001.

New South Wales – rural
19 delivieries

	Local Council	*District‡*
1	Cessnock	HR
2	Gosford	SS
3	Parkes	CW
4	Wyong	SS

New South Wales – urban
151 deliveries

	Local Council	*District‡*
5	Auburn	OS
6	Bankstown	OS
7	Blacktown	OS
8	Campbelltown	OS
9	Fairfield	OS
10	Holroyd	OS
11	Leichhardt	IS
12	Liverpool	OS
13	Marrickville	IS
14	Parramatta	OS
15	Penrith	OS
16	Strathfield	IS
17	Sutherland	OS
18	Warringah	OS

‡District abbreviations: CW = Central West; HR = Hunter; IS = inner Sydney; OS = outer Sydney; SS = surrounding Sydney.
Local council area boundaries obtained from NSW Department of Local Government, Local Council Search. <http://www.dlg.nsw.gov.au/dlg/dlghome/dlg_home.asp> Accessed May 10, 2002.

more than one delivery on a given day. The number of deliveries made each day of the week was counted and the results presented in Figure 2.5. The average number of deliveries made each day of the week is summarised in Table 2.10.

Figure 2.5 Number of deliveries of gasoline and diesel fuel made each day of the week from the Caltex Banksmeadow distribution terminal from Friday, August 31 to Monday, October 1, 2001.

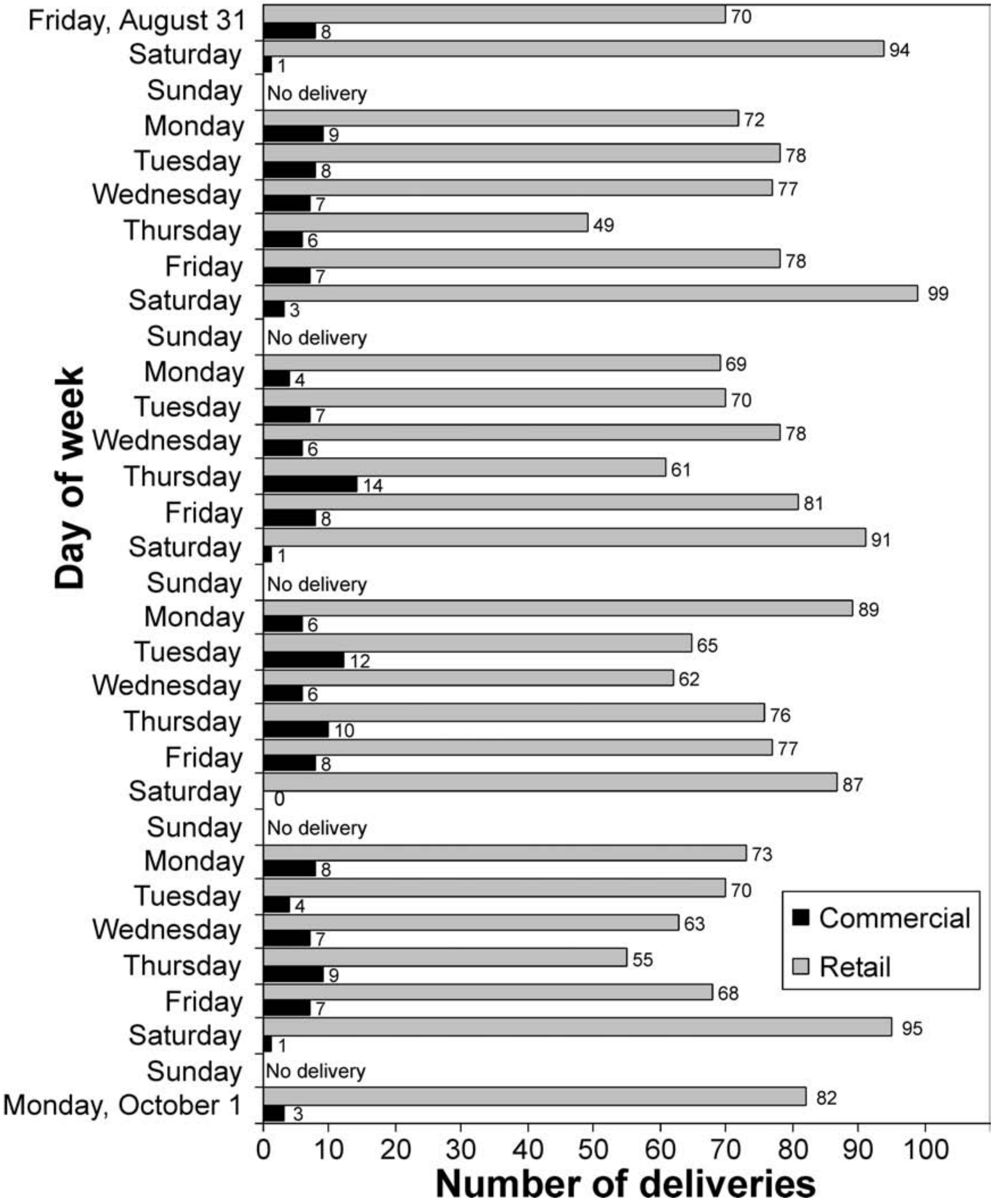


Table 2.10 Average number of deliveries of gasoline and diesel fuel made each day of the week from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. The 90% confidence interval about the mean number of deliveries is given.

Retail deliveries (*i.e.* service stations)			Commercial deliveries (*e.g.* freight haulers)		
day of week	*n*	*mean*	*day of week*	*n*	*mean*
Sunday	5	no deliveries	Sunday	5	no deliveries
Monday	5	77.0 ± 7.9	Monday	5	6.0 ± 2.4
Tuesday	4	70.8 ± 6.3	Tuesday	4	7.8 ± 3.9
Wednesday	4	70.0 ± 10.2	Wednesday	4	6.5 ± 0.7
Thursday	4	60.3 ± 13.6	Thursday	4	9.8 ± 3.9
Friday	5	74.8 ± 5.3	Friday	5	7.6 ± 0.5
Saturday	5	93.2 ± 4.3	Saturday	5	1.2 ± 1.0

Analysis of the data revealed that more retail deliveries occurred on Saturdays than on any other day of the week. On average, fewer deliveries were made to retail service stations on Thursdays than on any other day of the week. Fewer deliveries were made to commercial sites on Saturdays and more deliveries were made on Thrusdays than on any other day of the week. In general terms these results show that, on average, the chemical composition of gasoline at a service station is more likely to change on a Saturday, least likely to change on a Thursday and, for deliveries made from the distribution terminal studied here, could not change between Saturday evening and the following Monday morning as no fuel deliveries were made. The days of the week on which deliveries are made to a specific retail site will have a unique pattern over the course of time, a fact that will be illustrated later.

The number of trucks operating out of a distribution terminal is limited; therefore, each truck will have to make multiple trips from the terminal each day to fill all fuel orders. The number of trips a truck will make during a shift is mostly dependent on travel time. The closer the destination to the distribution terminal the faster the turn around time of the truck. The work day at the Banksmeadow distribution terminal is divided into two shifts, one in the morning and one in the afternoon. The number of deliveries a truck made during any given shift was counted over the entire delivery period and the results are given in Figure 2.6. During a typical shift a truck made four deliveries to four different sites.

Figure 2.6 Number of deliveries made per truck per shift from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001 (gasoline and diesel fuel deliveries).

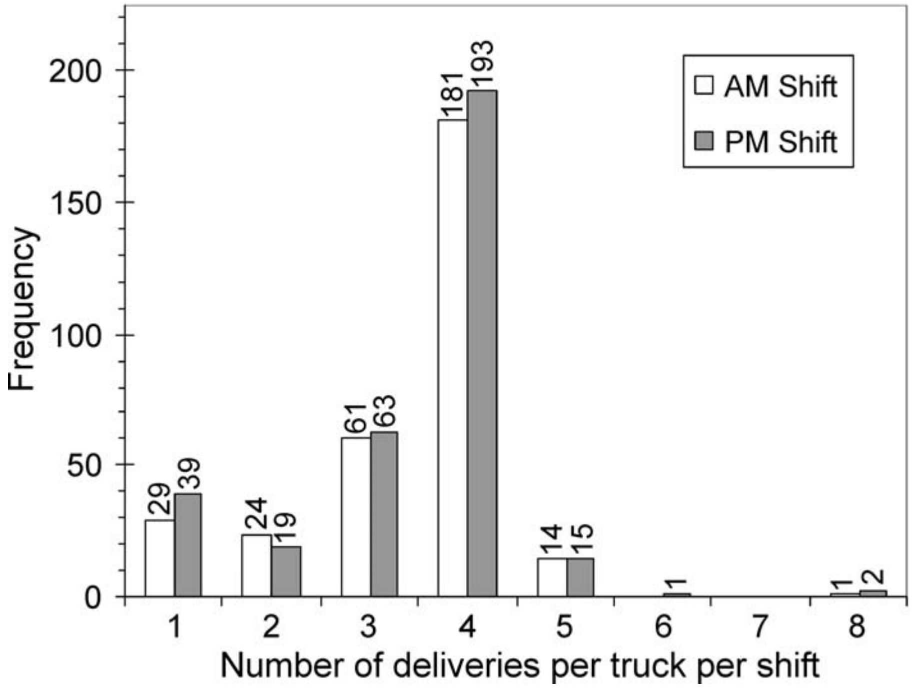


Figure 2.6 only takes into account the two truck fleets (one owned and operated by Caltex, the other owned and operated by an independent carrier) operating out of the Banksmeadow terminal and so 2,197 of the 2,199 deliveries are shown. The remaining two deliveries were small volume deliveries made by a third, independent owner/operator on two separate days and so were not included in Figure 2.6.

The turn-over rate of a given type of fuel at a retail service station is of particular interest to the forensic scientist engaged in an investigation involving fuel. The data supplied by Caltex Australia gave the dates and shift of each delivery, and the corresponding coded location of each retail site, therefore, a "day" was arbitrarily chosen to be a calendar day. The number of days between deliveries to a specific site is simply the numerical difference between calendar days. Thus, in those situations where two deliveries made on the same day to the same site (*e.g.* one in the morning and one in the afternoon), the number of days between deliveries was counted as zero. Deliveries made to a site on consecutive days to a site were counted as one day between deliveries and so on. The number of days between deliveries of a given type of fuel for each retail site serviced by the Banksmeadow distribution terminal was calculated and the frequency of the number of days between deliveries for all sites was determined and is shown in Figure 2.7.

From Figure 2.7 it can be seen that most retail sites received a delivery of regular unleaded gasoline every one to four days, with the majority having received regular gasoline every two or three days. It is more difficult to generalise the delivery frequency for premium unleaded and lead replacement gasoline, and diesel fuel. Premium unleaded gasoline was typically delivered every two to seven days, lead replacement gasoline was usually delivered every three to seven days, and diesel fuel was delivered every one to five days. It is worth noting that there were some retail sites that received a fuel delivery only once every seven or more days.

Commercial sites, which almost exclusively received delivery of diesel fuel, tended to take delivery of fuel between zero (two deliveries on the same day to the same site) and three days (Figure 2.8).

Figure 2.7 Frequency of the number of days between deliveries of fuel, by fuel type, for all retail sites supplied by the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. See text for explanation of how the number of days between deliveries was deteremined for each site.

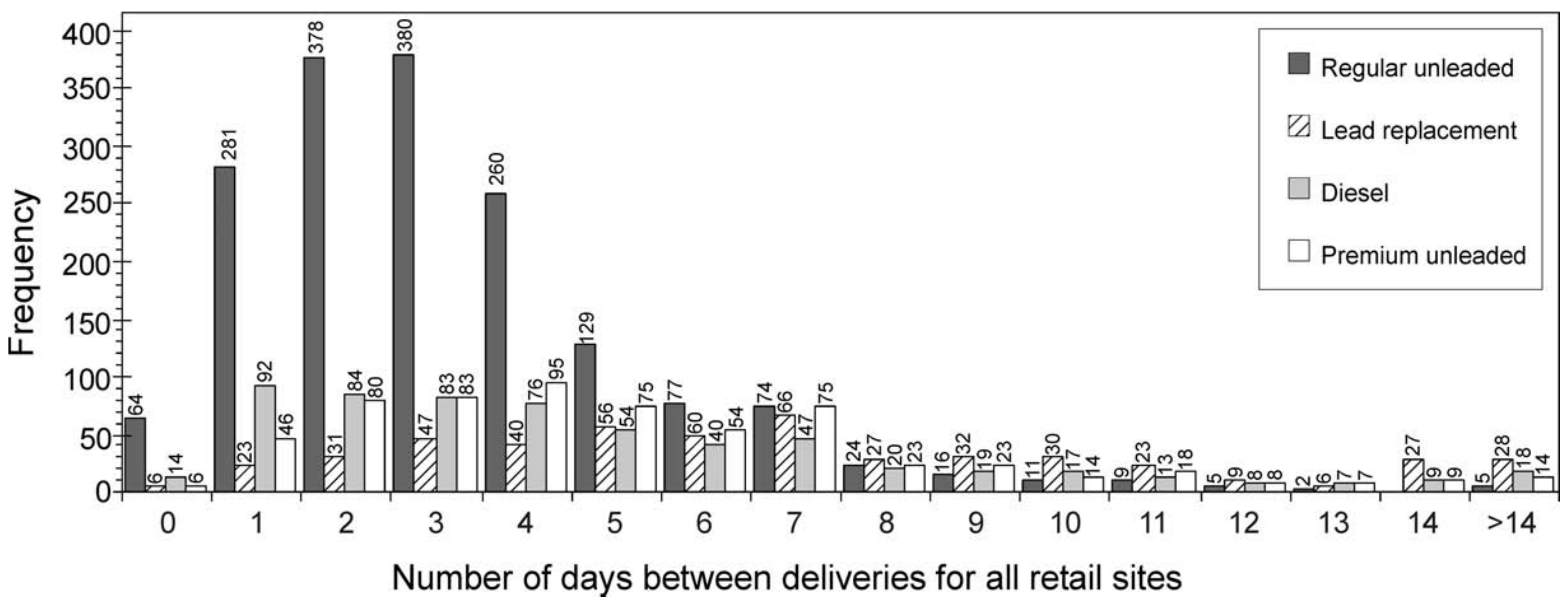
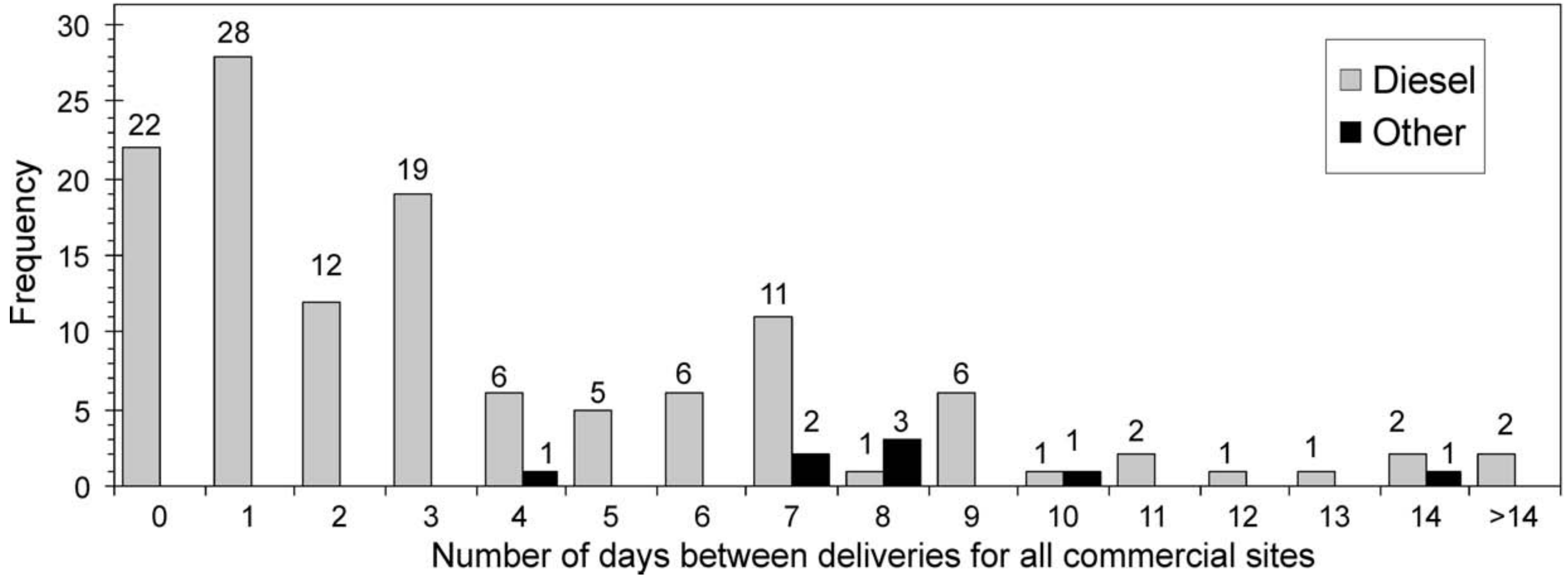


Figure 2.8 Frequency of the number of days between deliveries of fuel, by fuel type, for all commercial sites supplied by the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. See text for explanation of how the number of days between deliveries was determined.



2.4.3 Delivery volumes

Of the 2,199 fuel deliveries made from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001, a total of 70.3 megalitres of gasoline were moved in 1,964 deliveries, while 14.3 megalitres of diesel fuel were moved in 765 deliveries. The proportion of each fuel type delivered to either retail or commercial sites is shown in Figure 2.9.

The vast majority of fuel delivered to retail sites was regular unleaded gasoline (72.7%), followed by diesel fuel (10.9%), then premium unleaded gasoline (8.7%) and lead replacement gasoline (7.8%). Commercial sites almost exclusively took delivery of diesel fuel (94.0%), with a small amount of regular unleaded gasoline (<1%). Each type of fuel is delivered each day of the week (except Sunday) as shown in Figure 2.10. The data for regular unleaded gasoline in Figure 2.10 clearly mirrors the delivery frequency data presented in Figure 2.5 with the largest volumes of regular gasoline being delivered on Saturdays and the smallest volumes being delivered on Thursdays. The volume of a given type of fuel delivered by truck to a retail service station was also examined and the data presented in Figure 2.11.

Taking into account that each of the five compartments in a maxi barrel trailer, the work horse of the truck fleets, holds approximately 8,000 litres it can be seen in Figure 2.11 that the delivered volume of a given type of fuel is a multiple of approximately 8,000 litres. Using regular unleaded gasoline as an example it can be seen in Figure 2.11 that the majority of trucks transporting this fuel to retail stations had three or four of the barrel's compartments filled, that is, approximately 24,000 or 32,000 litres of fuel, respectively. The remaining compartments in the barrel may or may not have contained another type of fuel. It can also be seen that only one of the barrel's compartments was filled with one of the other three types of fuel (lead replacement, premium unleaded and diesel fuel), that is, approximately 8,000 litres of these fuels were delivered at any one time. This information suggests that many deliveries consisted of three of four compartments of regular unleaded

Figure 2.9 Proportion of each fuel type delivered from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. Volumes are in kilolitres.

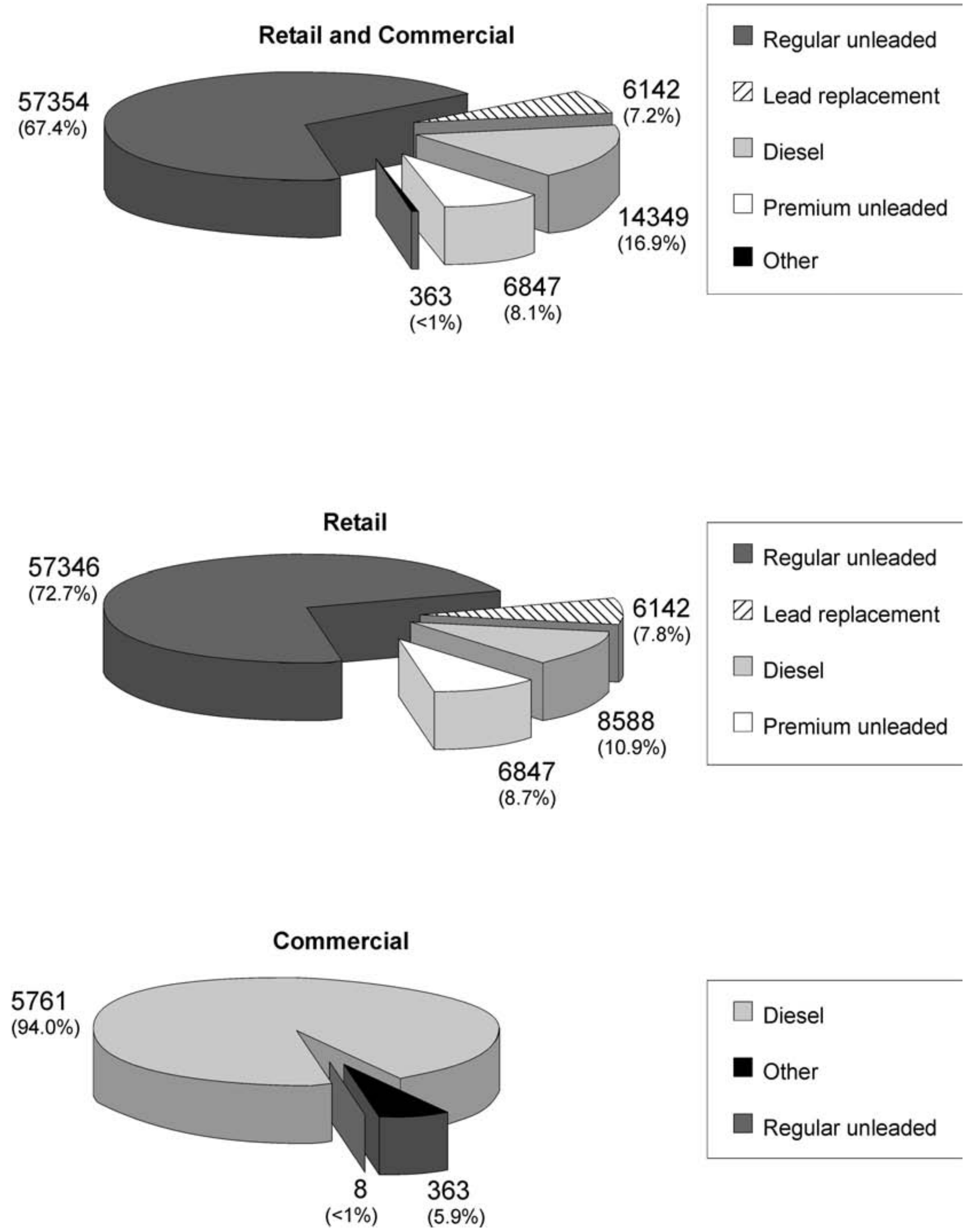


Figure 2.10 Total volumes of gasoline and diesel fuel (retail and commercial) delivered each day from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. All volumes are in litres.

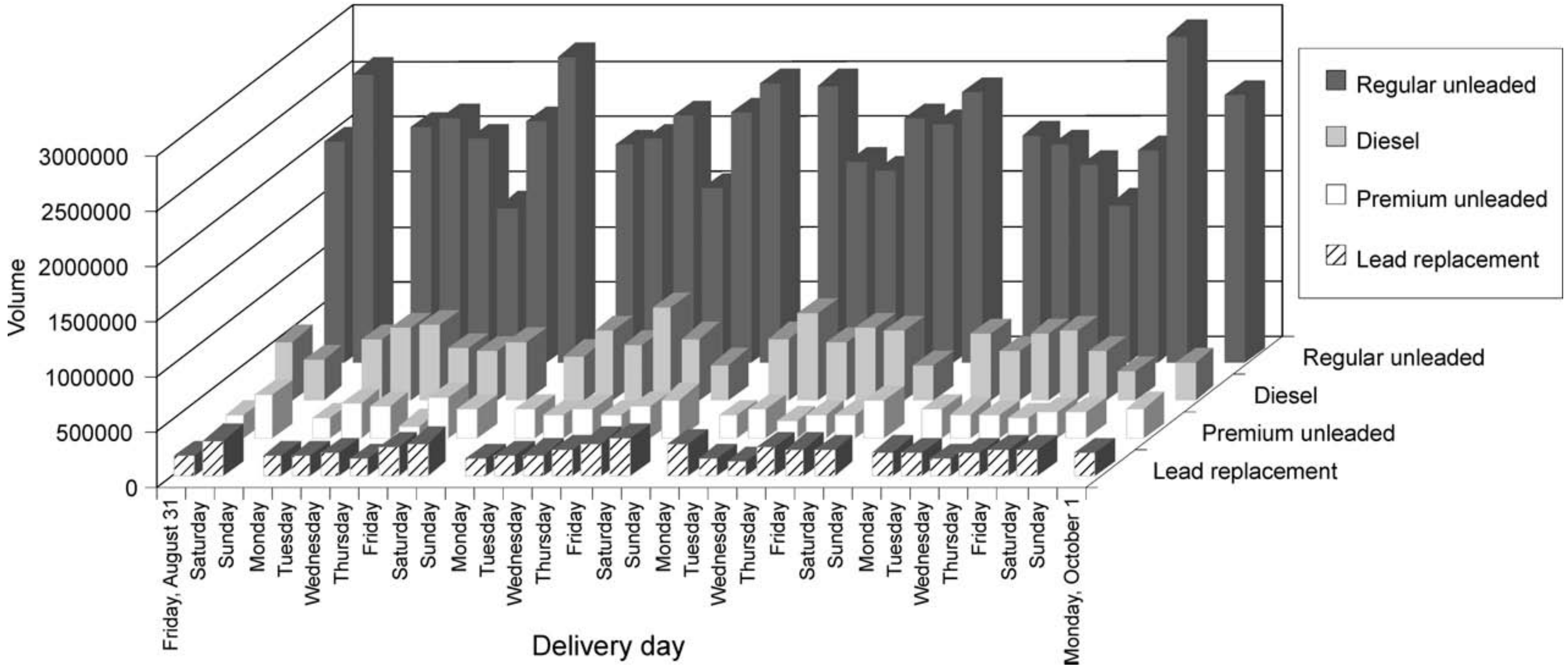
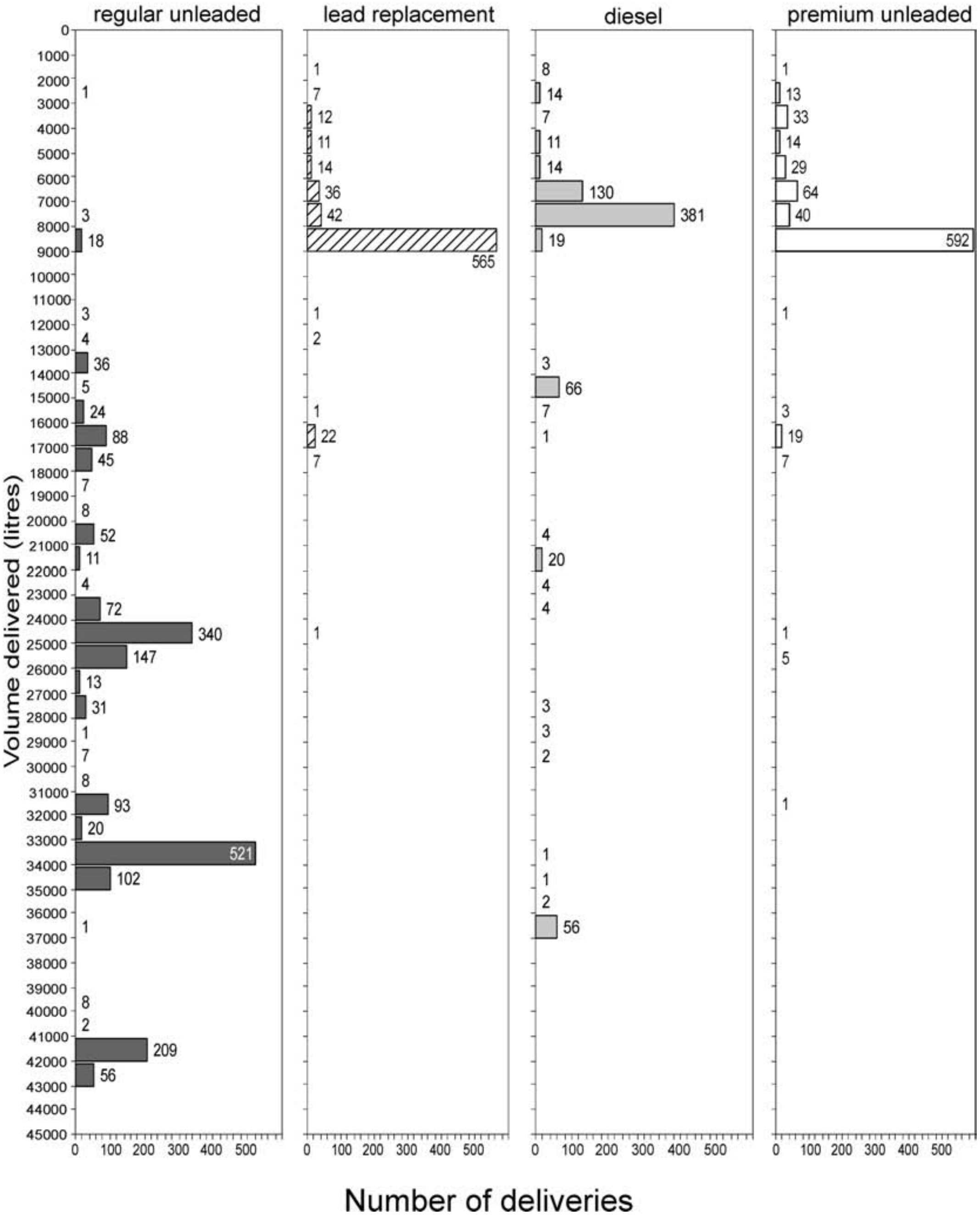


Figure 2.11 Number of deliveries, by volume, of different fuels to retail service stations from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001.



gasoline plus one of the other types of fuel. The number of retail deliveries of single and multiple fuel types is shown in Figure 2.12.

Most deliveries to retail service stations were either regular unleaded gasoline (386 deliveries), or two fuel types: regular unleaded and premium unleaded gasoline (376 deliveries); regular unleaded and lead replacement gasoline (331 deliveries); or, regular unleaded gasoline and diesel fuel (276 deliveries).

2.4.4 Example of fuel distribution to five service stations in one local council area

From the preceding data it is clear that there is a large variation in fuel deliveries, and therefore fuel turnover rates, at individual retail service stations. In order to more thoroughly investigate the fuel turnover rate at a retail service station, the day to day delivery data of five different service stations in one council area serviced by the Banksmeadow distribution terminal was examined.

2.4.4.1 The service stations

Five service stations, referred to as site numbers 1 to 5, were chosen because they were all within the same local council area and could be identified as receiving their deliveries from the Banksmeadow terminal from August 31 to October 1, 2001. Each service station was located on a major road and all had a convenience store on site. These five sites range in size from a relatively small station having four pumps and two types of fuel for sale (site number 4), to a large station having 17 pumps and offering all four fuel types for sale (site number 2). Schematic diagrams of each of the five service station sites, the types of fuel dispensed and the number of underground fuel storage tanks are presented in Figures 2.13 to 2.17.

Figure 2.12 Number of retail deliveries of single and multiple fuel types from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001.

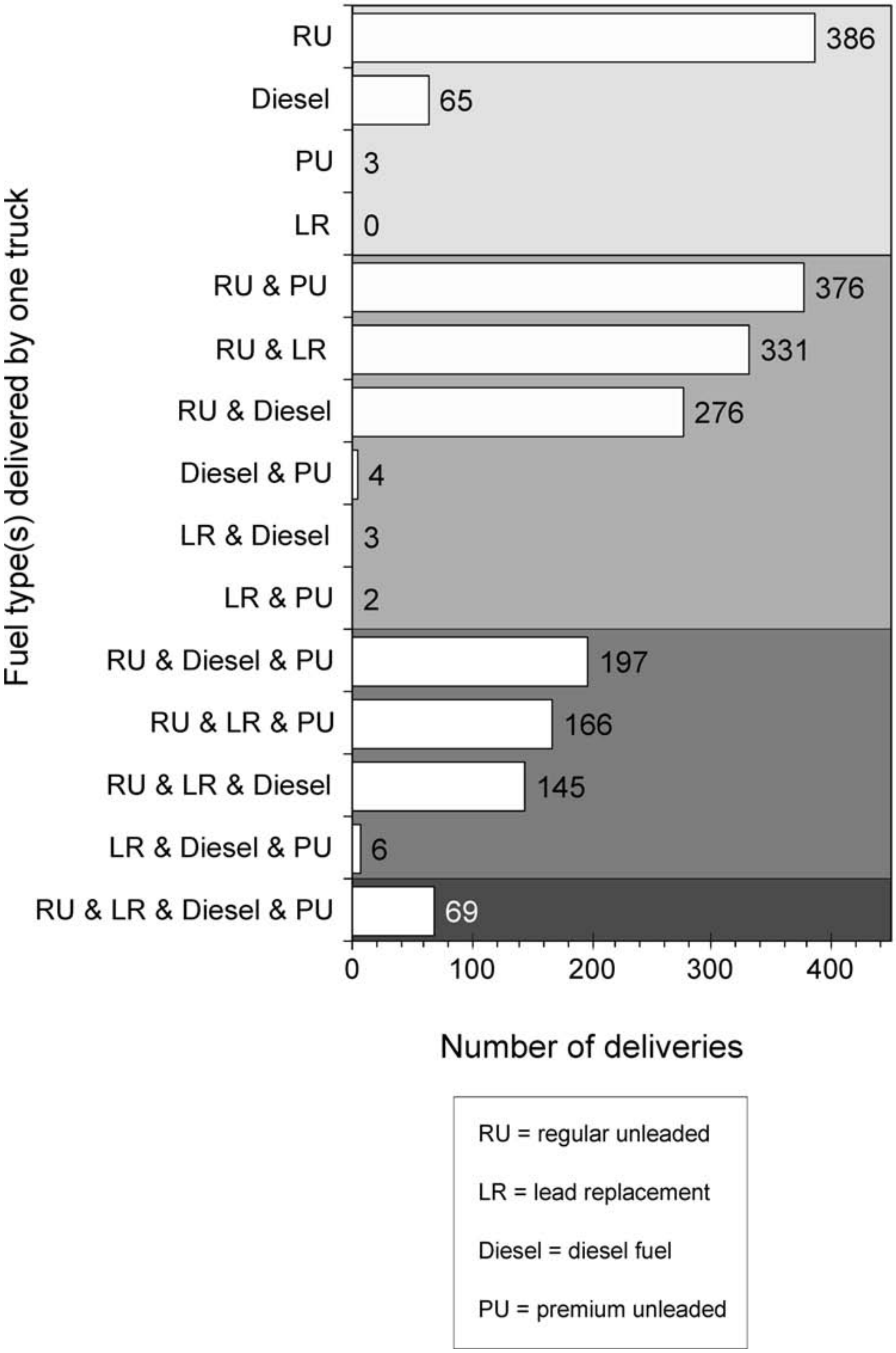
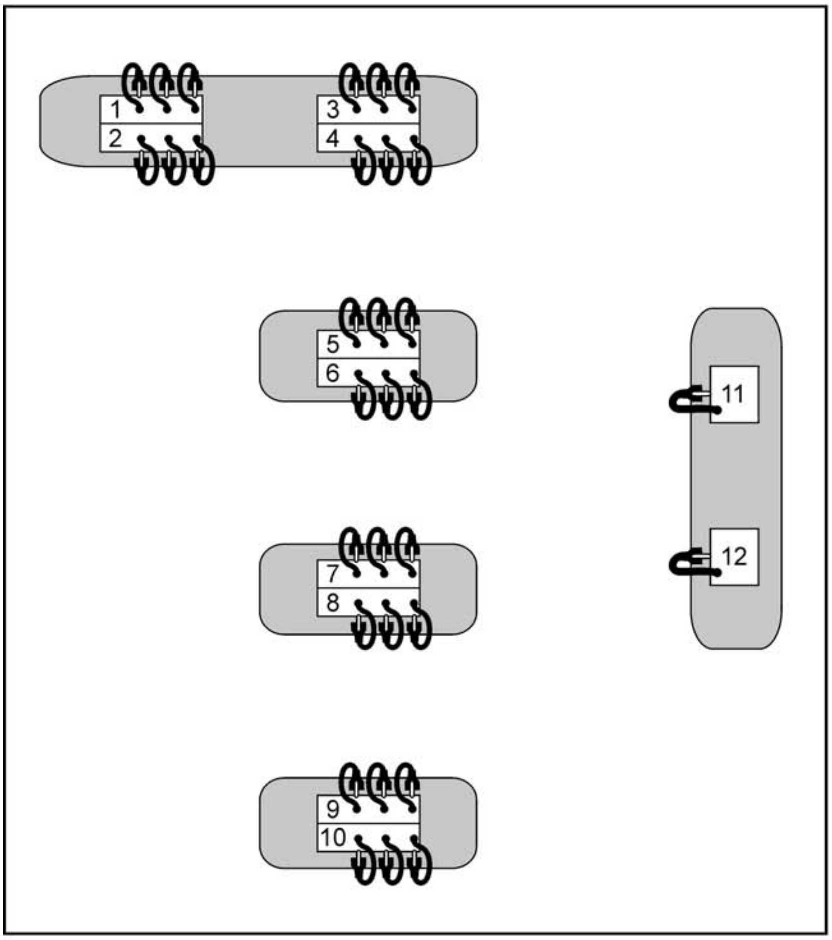


Figure 2.13 Service station forecourt details, site number 1.



Fuel dispensed by each pump

Pump	RU	PU	LR	Diesel
1	⛽	⛽	⛽	
2	⛽	⛽	⛽	
3	⛽	⛽	⛽	
4	⛽	⛽	⛽	
5	⛽	⛽	⛽	
6	⛽	⛽	⛽	
7	⛽	⛽	⛽	
8	⛽	⛽	⛽	
9	⛽	⛽	⛽	
10	⛽	⛽	⛽	
11				⛽
12				⛽
Total	10	10	10	2

Number of underground storage tanks

	Number	
RU	▭ ▭ ▭ ▭	4
PU	▭	1
LR	▭ ▭	2
Diesel	▭	1

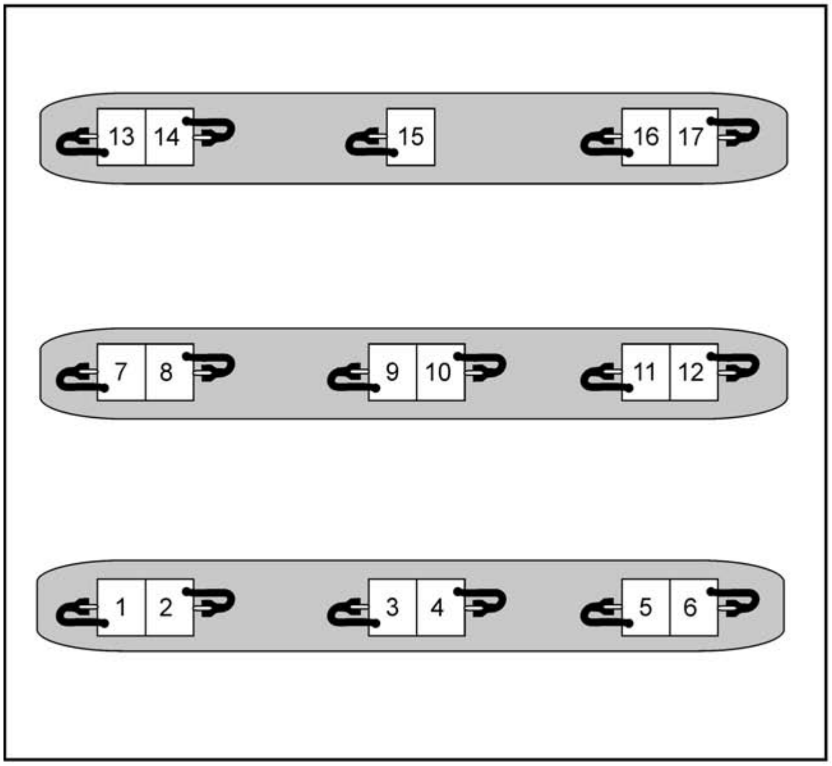
RU = regular unleaded gasoline

PU = premium unleaded gasoline

LR = lead replacement gasoline

Diesel = diesel fuel

Figure 2.14 Service station forecourt details, site number 2.



Fuel dispensed by each pump

Pump	RU	PU	LR	Diesel
1	⛽			
2		⛽		
3	⛽			
4				⛽
5			⛽	
6	⛽			
7		⛽		
8		⛽		
9			⛽	
10	⛽			
11	⛽			
12			⛽	
13			⛽	
14	⛽			
15	⛽			
16		⛽		
17			⛽	
Total	7	4	5	1

Number of underground storage tanks

	Number	
RU	⬭ ⬭ ⬭	3
PU	⬭	1
LR	⬭ ⬭	2
Diesel	⬭	1

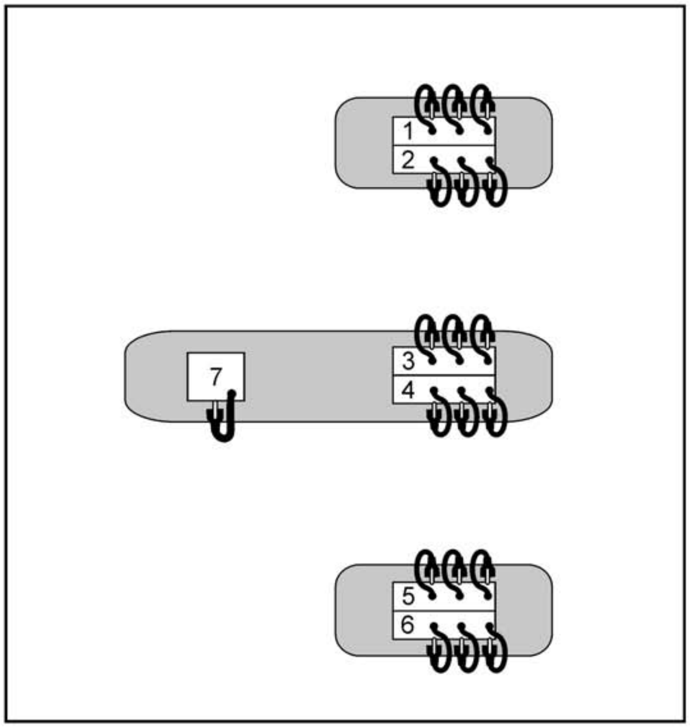
RU = regular unleaded gasoline

PU = premium unleaded gasoline

LR = lead replacement gasoline

Diesel = diesel fuel

Figure 2.15 Service station forecourt details, site number 3.



Fuel dispensed by each pump

Pump	RU	PU	LR	Diesel
1	⛽	⛽	⛽	
2	⛽	⛽	⛽	
3	⛽	⛽	⛽	
4	⛽	⛽	⛽	
5	⛽	⛽	⛽	
6	⛽	⛽	⛽	
7				⛽
Total	6	6	6	1

Number of underground storage tanks

	Number	
RU	▭ ▭	2
PU	▭	1
LR	▭ ▭	2
Diesel	▭	1

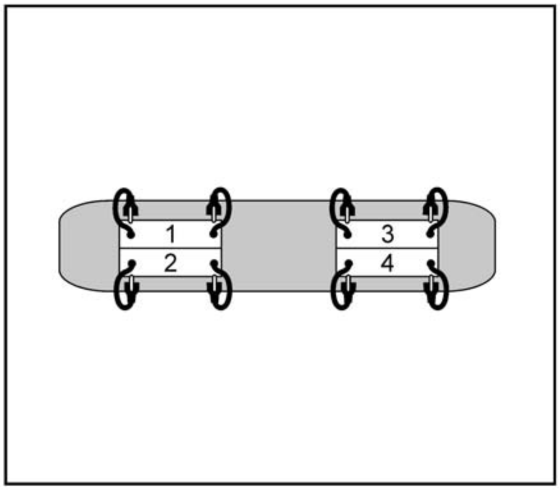
RU = regular unleaded gasoline

PU = premium unleaded gasoline

LR = lead replacement gasoline

Diesel = diesel fuel

Figure 2.16 Service station forecourt details, site number 4.



Fuel dispensed by each pump

Pump	RU	PU	LR	Diesel
1	⛽		⛽	
2	⛽		⛽	
3	⛽		⛽	
4	⛽		⛽	
Total	4	0	4	0

Number of underground storage tanks

	Number	
RU	▭ ▭ ▭	3
PU		0
LR	▭	1
Diesel		0

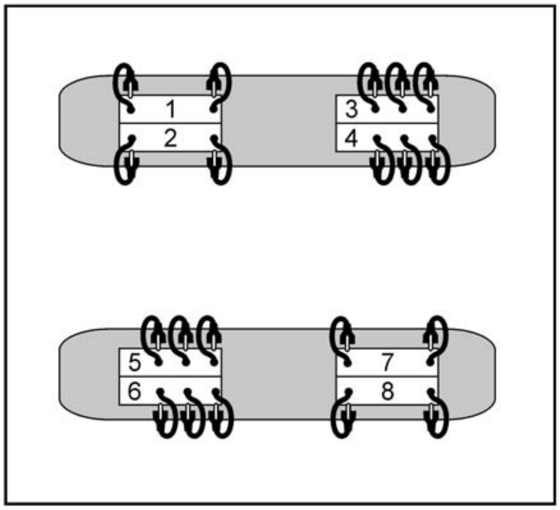
RU = regular unleaded gasoline

PU = premium unleaded gasoline

LR = lead replacement gasoline

Diesel = diesel fuel

Figure 2.17 Service station forecourt details, site number 5.



Fuel dispensed by each pump

Pump	RU	PU	LR	Diesel
1	⛽		⛽	
2	⛽		⛽	
3	⛽	⛽	⛽	
4	⛽	⛽	⛽	
5	⛽	⛽	⛽	
6	⛽	⛽	⛽	
7	⛽		⛽	
8	⛽		⛽	
Total	8	4	8	0

Number of underground storage tanks

	Number	
RU	▭ ▭ ▭	3
PU	▭	1
LR	▭ ▭	2
Diesel		0

RU = regular unleaded gasoline

PU = premium unleaded gasoline

LR = lead replacement gasoline

Diesel = diesel fuel

2.4.4.2 Delivery frequency and volumes

The volume of each fuel type delivered each day to each site is provided in Figures 2.18 to 2.22. Only one retail site (site number 2) received two deliveries in one day (Figure 2.19). Sites 1 and 4 have relatively consistent delivery frequencies. For example, site number 1 often received a delivery of regular unleaded gasoline every three to four days (Figure 2.18). Data from site number 1 also shows that a typical delivery consists of a full maxi barrel trailer of approximately 40,000 litres of one or more types of fuel. The volume of each type of fuel is a multiple of approximately 8,000 litres. Examples of one, two, three or four fuel types being delivered at one time can be found at this service station. Site number 5 (Figure 2.22) showed a relatively large turnover rate of regular unleaded gasoline in the first three weeks of the time period studied. This data (Figures 2.18 to 2.22) demonstrates the high degree of variability in fuel turnover rates between service stations.

2.5 Conclusions

The data presented here gives a glimpse of the complexities of the retail and commercial fuel distribution system and provides the forensic scientist with a valuable framework with which to interpret the results of a gasoline comparison. For example, it has been shown here that the turnover rate of the most common fuel, regular unleaded gasoline, at a retail service station is usually every one to four days. Retail stations most commonly receive regular unleaded gasoline on a Saturday, while Thursdays appear to be the least likely day for delivery of regular unleaded gasoline.

The use of crude oil from geographically distinct oil basins and the blending of different processing streams at the refinery will continually impart unique chemical differences to the gasoline. Gasoline produced at a refinery on different days is mixed in the storage tanks at bulk distribution terminals and in the underground storage tanks at service stations. These factors contribute to the variation in the chemical composition of the gasoline that is ultimately purchased by the consumer. The turnover rate of gasoline at the retail site will also impact on the rate of change of the gasoline composition.

Figure 2.18 Deliveries to service station site number 1 from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001.

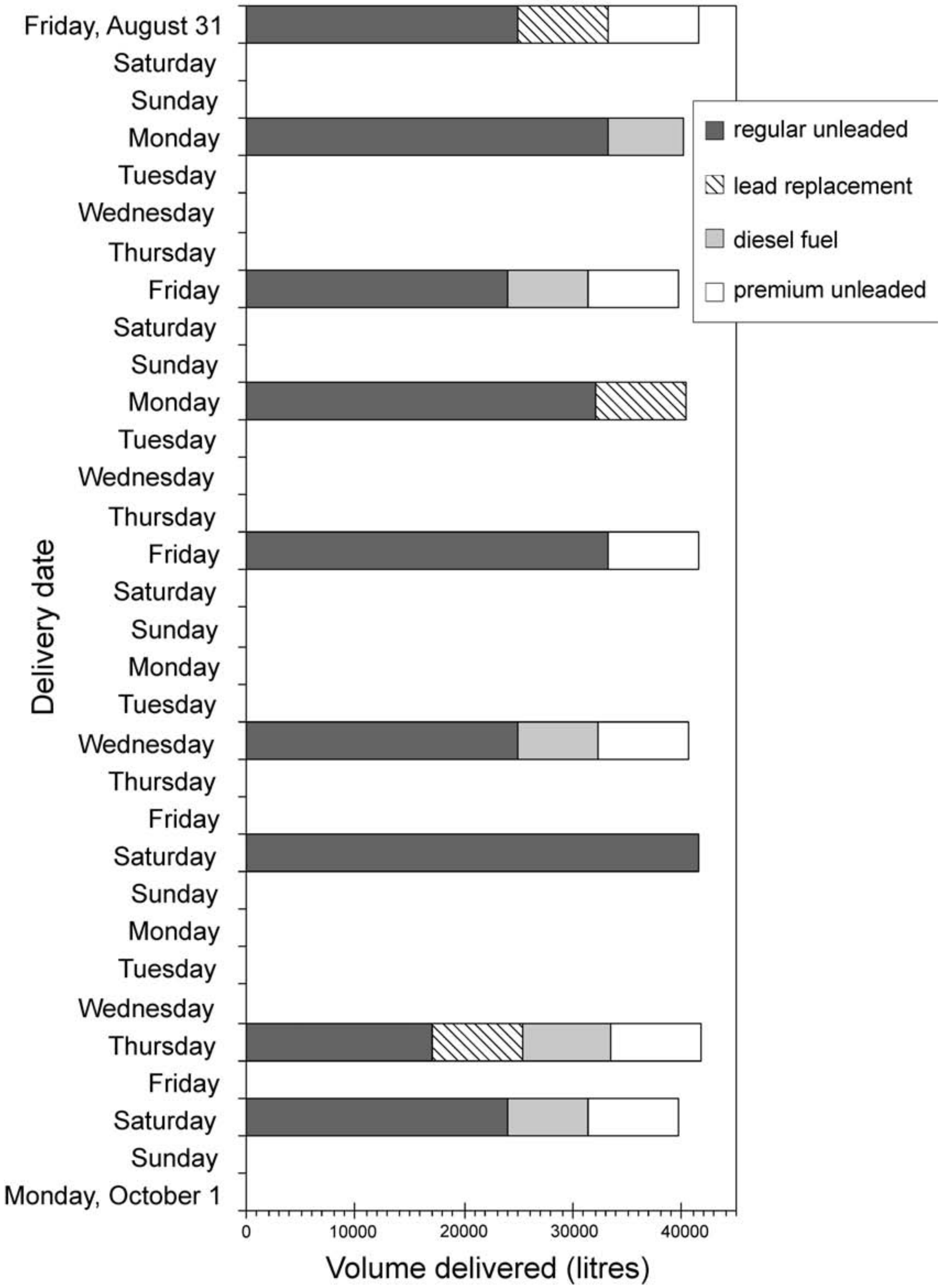


Figure 2.19 Deliveries to service station site number 2 from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. Note that although this station sells diesel fuel, there were no diesel fuel deliveries to this site during this time period.

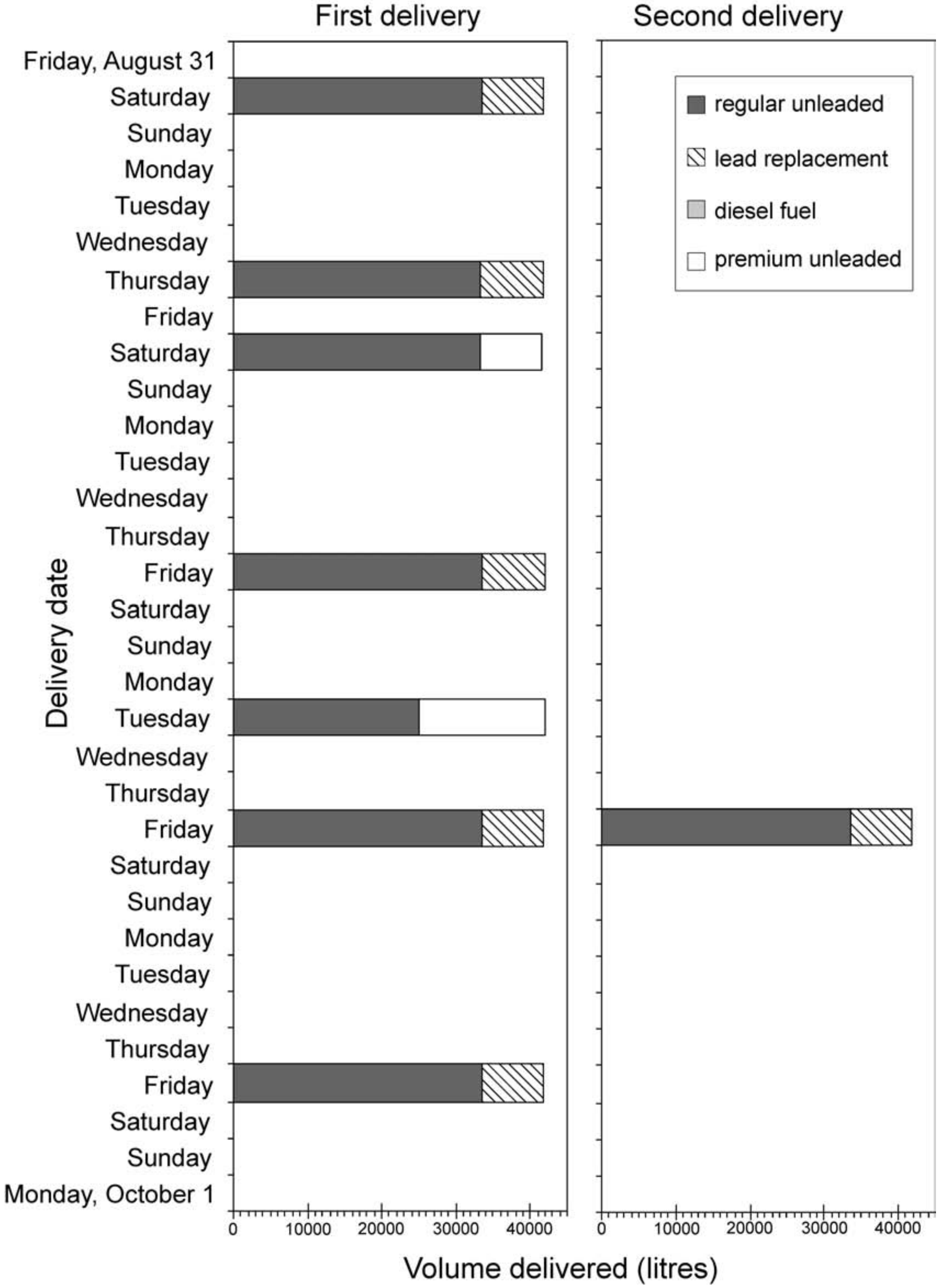


Figure 2.20 Deliveries to service station site number 3 from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. Note that the only delivery of diesel fuel to this site was made on Saturday, September 29.

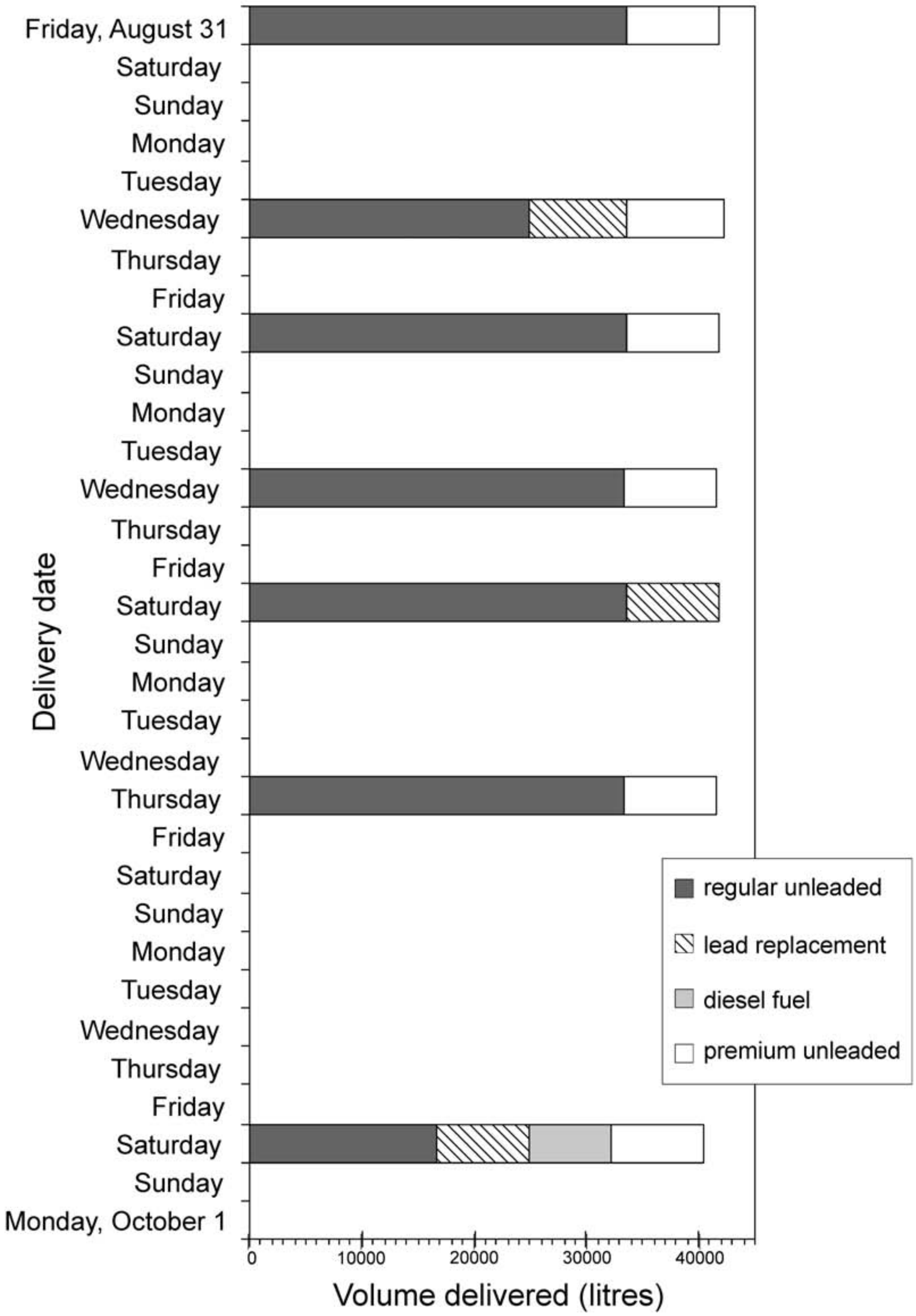


Figure 2.21 Deliveries to service station site number 4 from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. Note that this site does not sell any premium unleaded gasoline or diesel fuel.

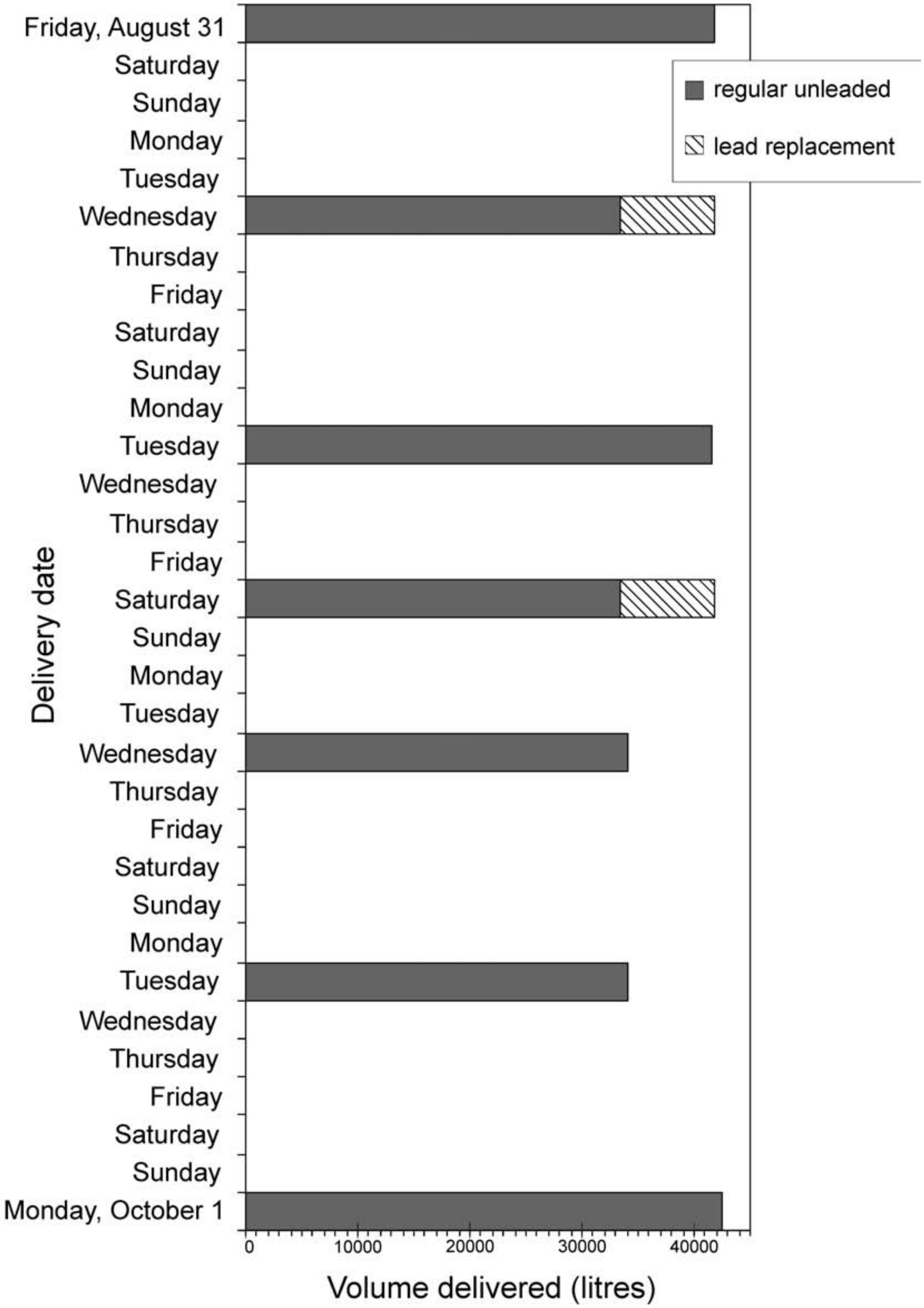
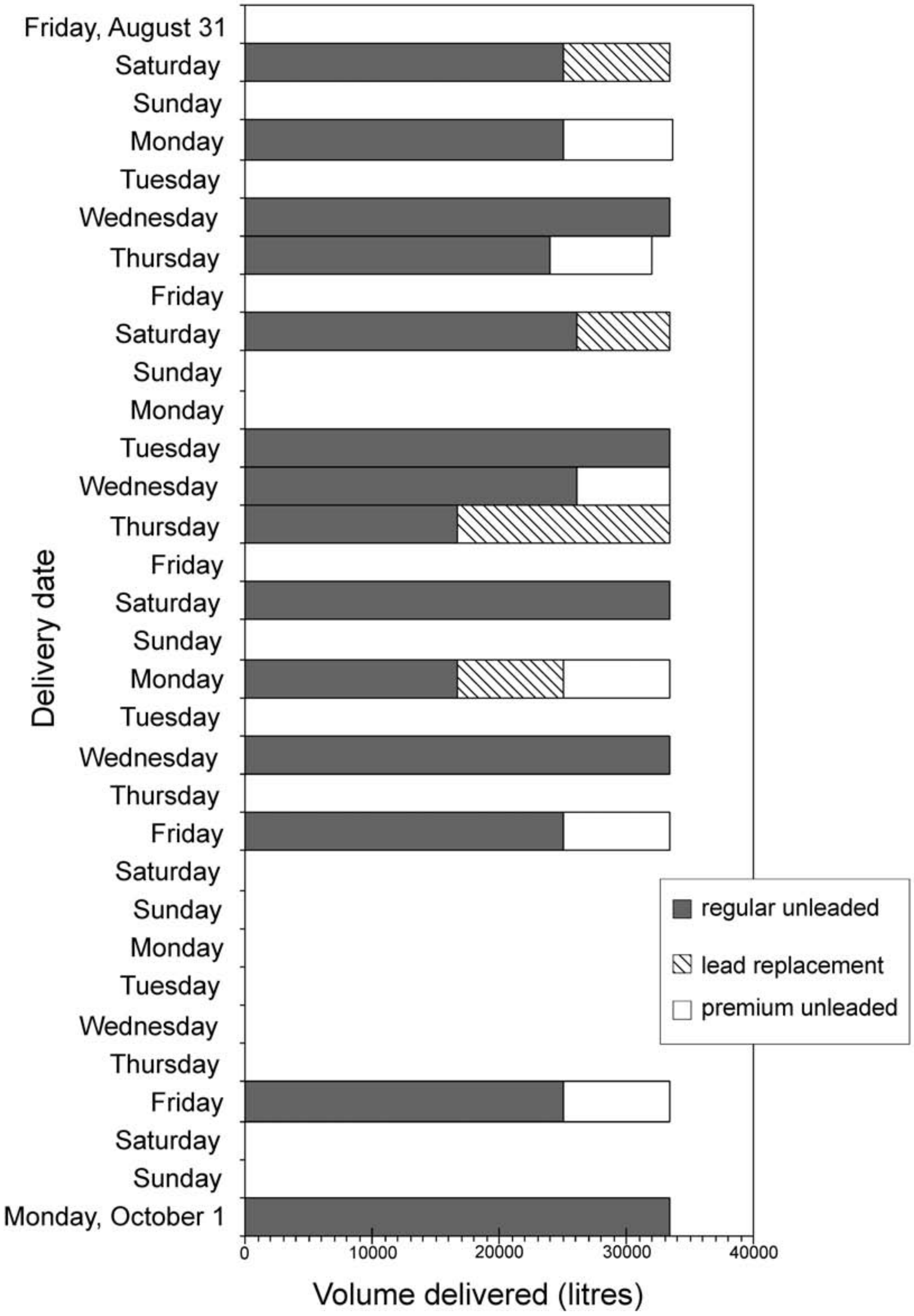


Figure 2.22 Deliveries to service station site number 5 from the Caltex Banksmeadow distribution terminal from August 31 to October 1, 2001. Note that this site does not sell any diesel fuel.



2.6 References

1. Hirz, R. Gasoline brand identification and individualization of gasoline lots. Journal of the Forensic Science Society, 1989; 29(2): 91-101.

2. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography I: comparison methodology. Journal of Forensic Sciences, 1987; 32(3): 606-615.

3. Mr. Paul Seage, Chief Chemist, Caltex Kurnell Refinery, Caltex Australia Limited. Personal Communication.

4. Wilson, N. Woolies wants change to fuel rules. The Australian, The Nation, page 2. May 8, 2001.

5. Motor Gasolines Technical Review. Chevron Products Company, Chevron USA Inc., San Ramon, CA, USA; 1996.

6. A comprehensive exploration of the refining process of crude oil and its many by-products. New Zealand Refining Company. <www.nzrc.co.nz/education/documents/> Accessed August 26, 2002.

7. Patel, P. Pumping allocation at a New Zealand oil refinery. The 34th Annual Conference of the Operational Research Society of New Zealand (ORSNZ). University of Waikato, Hamilton, New Zealand. December 10-11, 1999.

8. Mr. Len Daly, Commercial Manager, Royal Vopak (Australia) Limited. Personal Communication.

9. Petroleum products. Industry Commission Report No. 40, Commonwealth of Australia, Canberra, ACT; July 5, 1994.

10. Mr. John Morgan, Director of Operations, Banksmeadow Distribution Terminal, Caltex Australia Limited. Personal Communication.

11. Road Traffic Act (1961). Government of South Australia. <http://www.transport.gov.sa.au/permits> Accessed 22 August, 2002.

12. Geyer, W. Where has our petroleum storage capacity gone? NPN: National Petroleum News. December, 1999.

13. Mark Deacon, Senior Operations Analyst, Caltex Australia Limited. Personal Communication.

14. Oil and Australia statistical review 1991. Petroleum Gazette, 1991; 2: 11.

15. Australian Energy News. Petrol demand continues to grow. 1998; 10.

16. National phase out of leaded petrol. Environment Australia. <http://www.ea.gov.au/atmosphere/airquality/lead/qa.htlm> Accessed October 30, 2002.

Chapter 3. Data analysis

3.1 Overview of chemometrics and principal component analysis (PCA)

Chemometrics is the finding of mathematical relationships between the chemical composition of a substance (measured data) and a property of interest that may be inferred from the measured data.[1] Chemometrics has been successfully applied to several diverse problems:[2]

- multivariate calibration;
- structure-(re)activity modelling;
- pattern recognition, classification, and discriminant analysis;
- multivariate process modelling and monitoring.

Chemometrics works well with a data set that has many collinear variables. Collinear means to lie along the same line. In terms of variables such as peaks in a gas chromatogram, the peaks are collinear if the information they contain is largely redundant, that is, if they are highly correlated.[3] One chemometric tool that is widely applied to chemical data is principal component analysis (PCA). Jolliffe describes principal component analysis as "probably the oldest and best known of the techniques of multivariate analysis."[4]

An example of a use of PCA to examine chromatographic data was recently published by Tan *et al.*[5] This study attempted to classify 51 petroleum products according to their ASTM Class.[6] The authors extracted seven ions of interest from each sample using selected ion monitoring (SIM) then created 19 variables by summing the abundance of these ions in each of 19 equal time segments over the entire chromatogram. A plot of the first two principal component scores, accounting for 57.1% of the total variance in the data, demonstrated that different petroleum classes could be distinguished from one another. Tan *et al.*[5] also investigated a supervised learning technique, soft independent model classification analogy (SIMCA), to create a PCA model for each class. Once the model had been developed it was used to predict the class of an unknown sample.

Another use of PCA in the classification of petroleum products was reported by Andrews and Lieberman.[7] These authors examined the fluorescence spectra of seven different petroleum products. In this case the first three principal components accounted for 97.3% of the variance in the data. A plot of the first three scores made it possible to visually distinguish between classes of petroleum products (*e.g.* jet fuel could be distinguished from gasoline). However, it was not always possible to distinguish different samples within the same class (*e.g.* some regular unleaded and premium unleaded gasoline samples overlapped in the same visual space).

3.2 Data pre-treatment

3.2.1 Heteroscedastic and homoscedastic error

Prior to the use of a chemometric method such as principal component analysis, the raw data must often be pre-treated (*e.g.* normalised) in some way. One common method of data pre-treatment is to normalise the data to a constant sum. Normalising the data from one object (sample) is performed by dividing each variable by the total sum of all variables for the object (constant sum normalisation). This type of normalisation is said to "close" the data set and is commonly used to remove variations between samples that are due to small differences in sample dilution or differences in sample volume injected into the instrument. However, Rietjens[8] has identified two independent errors that are introduced when constant sum normalisation is applied to a data set. The first of these is error propagation caused by heteroscedastic noise.[8] For raw data the noise in gas chromatography and mass spectrometry is heteroscedastic, that is, the error is proportional to the area of the peak. Constant sum normalisation in the presence of heteroscedastic noise may propagate a significant error in smaller peaks if both large and small peaks are present. Homoscedastic noise is uniform, that is, it is constant over all variables (peaks) irrespective of their sizes (peak areas). In an object (chromatogram) where all of the variables (peaks) are approximately the same size, the errors, too, will all be approximately the same size (homoscedastic) and so a constant sum normalisation will not introduce spurious errors.

The second type of error introduced by normalisation is closure of the data set.[8] These two types of errors, in turn, may cause subsequent analysis by principal components to return poor or incorrect results due to the introduction of false correlations from constant sum normalisation. It is important, therefore, to eliminate heteroscedastic noise and replace it with homoscedastic noise.

The detection of a heteroscedastic noise pattern in a data set of eight replicates of 1,2,3,4-tetrachlorobenzene analysed by GC-MS has been described by Kvalheim *et al.*[9] A plot of the standard deviation of the selected ion fragments versus the mean ion intensity clearly demonstrated the presence of heteroscedastic noise. After constant sum normalisation, a plot of the standard deviation against the mean signal intensity will typically show a reduction in heteroscedasticity for the smaller peaks but this normalisation method failed to remove heteroscedasticity from the larger peaks.[9] Kvalheim *et al.*[9] demonstrated that log row centring of the data prior to normalisation may be used in some situations to eliminate heteroscedastic noise. It is very important to determine the type of noise present in the data set prior to attempting to remove the noise. For example, Toft and Kvalheim[10] have shown that the noise in transmittance FTIR data is homoscedastic and that log row centring will actually introduce heteroscedastic noise.

3.2.2 Normalisation of the data set

Normalisation to a constant sum and normalisation to the largest peak are two simple normalisation methods.[11] Normalisation to a constant sum is simply making the sum of all the variables for a given object equal to a constant (*e.g.* 100 or 10000). Normalisation to the largest peak makes the variable with the largest value equal to a constant (*e.g.* 100%), and it is against this peak that all other variables are compared. However, in gas chromatography and mass spectrometry data there is often at least one very large variable (peak) which, when a data set is closed, will introduce a dependence between the variables such that when the large variable increases, all the other variables must concomitantly decrease.[12] This situation is particularly true when normalising to the largest peak, a practise commonly used for mass spectrometry data. In order to avoid the effects that a dominant peak will have on other, smaller peaks, Johansson *et al.*[12] proposed the use of a method they called "selective closure." In this method the criterion for variable selection is to choose those variables whose means and standard deviations are approximately the same size while allowing the greatest number of variables to be selected. The selected variables are then normalised with a normalisation constant such that their sum equals 100. The remaining variables are then scaled with the same normalisation constant. The authors concede that this method is only a "partial solution" noting that a drawback to this method is its element of subjectivity where one must decide which variables have approximately the same mean and standard deviation, while attempting to keep the number of variables retained as large as possible.[12] A variation on the idea of selective closure of the data prior to principal component analysis has been applied by Wilcken and Schulten[13] to pyrolysis-gas chromatography data obtained from paint samples. Wilcken and Schulten[13] selected the 30 most intense gas chromatographic peaks for PCA and pre-treated the data by normalising these peaks relative to another peak. In this case they did not select the largest peak to normalise against but an "averaged sized" peak.[13] As with the selective closure method for constant sum normalisation described by Johansson *et al.*[12], Wilcken and Schulten's method[13] also contains an element of subjectivity that makes it difficult to reproduce from one data set to another.

Other normalisation methods have been described in the literature. Wold *et al.*[14] described a simple data transformation method they referred to as "autoscaling" which they used to convert a heteroscedastic noise pattern into a homoscedastic one. Autoscaling of a data set is performed as follows: for each variable, k, in a given sample (*e.g.* chromatogram) a scaling weight, w_k, is calculated such that $w_k = 1 / s_k$, where s_k is the standard deviation of variable k. This normalisation procedure was called autoscaling because the scaling weight will scale all the variables in a sample such that they have unit variance; that is, the variables became homoscedastic. Rietjens[8] applied a logarithmic and a weight factor normalisation method to simulated data and compared the error propagation in the results to the errors obtained using either constant sum normalisation or a normalisation to the largest peak. The weight factor method, which uses a flexible weighting for each variable, was developed based on the work of Kvalheim *et al.*[9] Rietjens[8] recommended that either the weight factor or the logarithmic method be applied to data showing heteroscedastic noise in order to reduce the chances of introducing spurious correlations prior to principal component analysis. Rietjens[8] also demonstrated that the logarithmic transformation works as well as the weight factor normalisation, even for a relatively small number of variables (*e.g.* ten variables or less).

3.2.3 Closure of the data set

As discussed above, closure of a data set by normalisation (*e.g.* normalising to a constant sum) can introduce a false dependence between the variables, particularly when the number of variables is small and there is one large variable dominating the data set.[12] It has been known for a long time that normalisation to a constant sum may bias the data by introducing negative correlations between variables, particularly when the number of variables is small (*i.e.* up to six).[15] Negative correlations result from strong differences between correlations in open and closed data sets. Skala[16] investigated the effects of constant sum normalisation on randomly generated data sets having three to ten variables. Skala's data represented a special case in which all variables had equal means and variances. Skala[16] found that distortion of correlation coefficients due to normalisation rapidly decreases in data sets when the number of variables is increased, and that no significant distortion occurs in data sets having seven or more variables. For real data sets, where the variables have unequal means and variances, the situation is much more complex and could not be easily simulated. Thus for real data sets it is important to select a robust normalisation procedure that propagates the fewest errors.[8]

Deming *et al.*[11] have investigated the effects on PCA results when constant sum normalisation and normalisation to the largest peak are used on a data set. They called normalisation to the largest peak "the maximum value transformation."[11] In their example they constructed a data set to visually resemble a flat spoked wheel, tilted such that it occupied three dimensions. They used this data set so that the effects of closure and PCA could be easily observed. They found that constant sum normalisation of the data set resulted in a reasonable, flat two-dimensional projection of the data, that is, most of the original structure of the data set was retained. They noted that some distortion of the data set did occur during the projection because the constant sum plane was not parallel to the original plane of the wheel. When normalising to the largest peak, Deming *et al.*[11] found that the dimensionality of the data set changed such that the data set occupied two planes instead of one. Visually, the flat wheel had become "draped" over the intersection of these two planes. Principal component analysis of this normalised data set rendered a result that was very distorted compared with the original data set, clearly showing the drawbacks of this normalisation method prior to multivariate data analysis.

3.3 Scope of principal component analysis

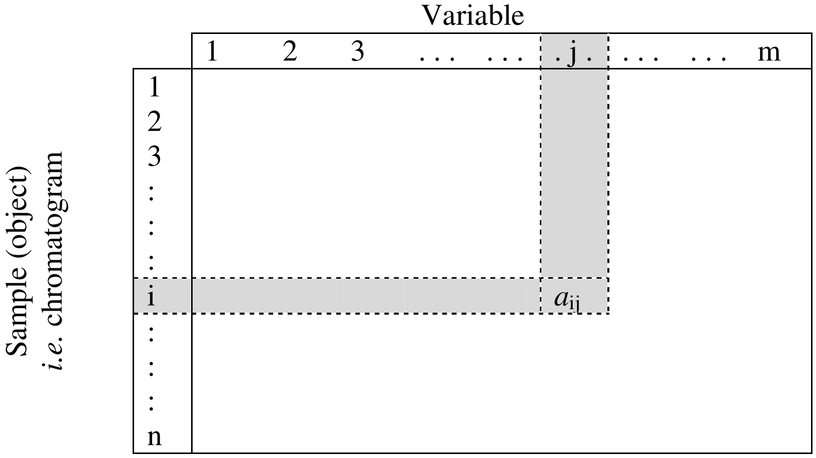
The scope of principal component analysis (PCA) is to obtain an overview of the dominant "patterns" in a data set.[14] By identifying these dominant patterns PCA will tend to reduce the dimensionality of a data set; that is, it will reduce a large number of variables, m, to a smaller number of variables, r, in the form of principal components while retaining as much of the variation in the original m variables as possible. The direct observation (measurement) of a sample, such as the near-infrared spectrum of a gasoline, can allow one to infer the value of another property, such as, the gasoline's octane number.[17] The use of directly measurable properties to observe, or measure, another property of interest is called indirect observation, or indirect measurement. A property such as octane number can be measured directly,[18] but it may be more convenient, less expensive and/or quicker to measure it indirectly. Properties that cannot be measured directly are sometimes called latent (hidden) properties.[19] By definition, then, a latent property can only be determined by indirect observations made on the original data set.[19] Thus one may use chemometrics to find a latent property, or latent variable, in a set of chemical data. Once found, latent variables may be used to extract information, such as relationships (regression/calibration) or patterns (classification), from the chemical data.

Principal component analysis examines the variances in the original data set.[4] This is an important concept. Wold *et al*.[14] have pointed out that in the past, data sets were typically analysed one variable at a time, that is, by examining each variable's mean and standard deviation. For a data set containing m variables this would yield 2m parameters. Principal component analysis also uses the covariance between the variables to yield a total of {m(m-1)/2} parameters.[14] For example, for 25 variables (m = 25) there will be 300 parameters available for analysis, but only 50 parameters from the means and standard deviations alone. Therefore more information is obtained by including the covariance (or correlation) data in the analysis. Thus PCA by itself can be a valuable exploratory technique that may be used to reveal important relationships in a data set.[3] This information, in turn, can be used to refine a hypothesis or an analytical approach.

3.3.1 Description of principal component analysis

The following is a brief description of principal component analysis (PCA). Suppose one has a data set of n samples and m variables. This data set can be represented by an n x m matrix, noted **A** (Figure 3.1). Here each row is an object, or sample, and is represented by the letter i; each column is a variable, or peak, and is represented by the letter j.

Figure 3.1 An n x m data matrix, noted **A**.



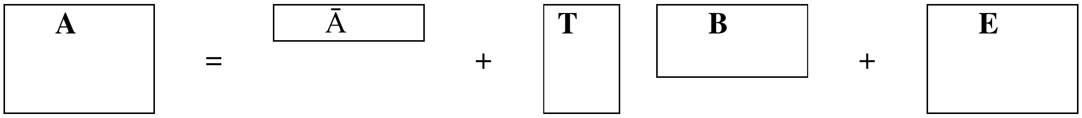
The first step in PCA is to look for a linear function that explains the maximum variation in the data set. The second step is to look for a linear function, uncorrelated (orthogonal) with the first function, which represents the second largest variation in the data. The process continues until the j^{th} linear function has been calculated. The maximum number of linear functions is equal to the original number of variables, m, however, it is generally expected that most of the variation in the data set will be accounted for by j principal components (PCs), where j is much less than m.

The principal component (PC) model is given mathematically in Equation [1].[20]

$$\mathbf{A} = \bar{a} + \mathbf{TB} + \mathbf{E} \qquad \textbf{[1]}$$

Equation [1] may be represented diagrammatically[21] as shown in Figure 3.2.

Figure 3.2 The principal component model.



Here **A** is the n x m data matrix of n objects and m variables, ā is the 1 x m matrix (vector) of the m variable means, **T** is the n x j matrix of PC scores, **B** is the i x m loading matrix and **E** is the n x m matrix of the residuals.[21] The residuals are the part of **A** not explained by the PC model, that is, the deviations between the original values and the projections.

Each PC consists of a set of coefficients, y, that form a linear combination of all the variables, m, in the original data set. For example, the first principal component, PC1, from the n x m matrix can be written as shown in Equation [2]:

$$PC1 = y_1 a_{i1} + y_2 a_{i2} + \ldots + y_k a_{ij} + \ldots + y_m a_{im} \qquad \textbf{[2]}$$

where $y_1 \ldots y_m$ are the coefficients (loadings) for the principal components, which are found in matrix **B**.

The scores are the new variables obtained from the principal component analysis and are represented by matrix **T** in Figure 3.2. Each score is calculated from its corresponding principal component.[21] As with the principal components, the first score has the greatest variation, the second score the next greatest variation, and so on.

Principal component analysis can be performed using either a correlation matrix or a covariance matrix.[4] As discussed by Jolliffe,[4] a drawback of using a covariance matrix is the resulting sensitivity of the principal components to the units of measurement and to the relative variance in each variable. Using a covariance matrix will tend to give more weight to the larger, more variable measurements and less weight to the smaller, less variable measurements. A correlation matrix, on the other hand, will tend to treat all variables equally.

After principal component analysis, the first principal component (PC) describes the greatest variation in the data set, the second PC describes the second greatest variation orthogonal to the first, and so on. Therefore the first few PCs are expected to describe virtually all of the variation in the data set and it is unnecessary to retain all of the PCs. Jolliffe[4] has described a number of subjective methods to help one decide how many PCs to retain; one simple method is known as "Kaiser's Rule."[22] Kaiser's Rule retains only those PCs that have variances equal to or greater than one (that is, for a correlation matrix, an eigenvalue equal to or greater than one) because any PC with a variance less than one is considered to have less information than one of the original variables and so should not be retained. When PCA is performed on a small number of variables, Kaiser's Rule can actually lead to the retention of too few PCs.[23] Another method is to construct a "scree plot" where the PC is plotted against the eigenvalue.[23] The scree plot is examined for the point at which the change in eigenvalue remains approximately the same from one PC to the next. This usually occurs when the eignenvalues for a PC are close to one (*cf.* Kaiser's Rule), however, it does not arbitrarily prevent one from retaining a PC whose eigenvalue is slightly less than one. The scree plot method is reproducible and is considered the most helpful when using PCA to explore relationships in chemical data.[4,23]

Although PCA is used to reduce the dimensionality of a data set by finding a smaller number of latent variables, the PC loadings (coefficients) may be used to identify those original variables which contribute to most of the variation (*i.e.* information) held in the data set;[4] the other variables may be considered redundant or irrelevant. Wold and Sjostrom[2] caution against using chemometrics as a means to identify which original variables to discard, particularly when the number of variables available is large. Wold and Sjostsrom[2] have recognised that traditional methods of viewing the data, such as plots of the PC coefficients or the score plots, may be difficult to interpret when the number of variables exceeds 100, and can become impossible to interpret with a thousand or more variables.

3.4 Materials and methods

Evaporated samples of a gasoline were prepared and analysed. Simulated data was created from these samples and the data sets used to demonstrate the effects of normalisation on principal component analysis.

3.4.1 Sample preparation

A sample of regular unleaded gasoline (sample RU03) in an amber Boston round glass bottle fitted with a Teflon-lined, polypropylene cap was weighed on an analytical balance (initial weight of sample was 94.4 grams). An aliquot (50 µL) of the gasoline was removed and transferred to a vial containing 5 mL of dichloromethane (HPLC grade, BDH). The bottle was capped and re-weighed on the analytical balance. The cap was removed and the bottle placed on a triple beam balance in a fume cupboard and the weight monitored until approximately 4 grams (*i.e.* approximately 4% by weight) had been lost. The bottle was then capped and its weight on an analytical balance recorded. A 50 µL aliquot of gasoline was removed and transferred to a vial containing 5 mL of dichloromethane. The bottle was then capped and its weight determined on the analytical balance. The process of evaporating approximately 4 grams of sample, weighing, removing an aliquot and weighing again was repeated. The amount of evaporation was determined from the weights, taking into account the weight lost when each aliquot was removed. Each 5 mL dichloromethane/gasoline solution was divided between five GC vials and analysed.

3.4.2 Gas chromatography-mass spectrometry

All samples were analysed using an Agilent 6890 gas chromatograph fitted with a 5973 mass selective detector and autosampler. A 1 µL sample was injected onto a DB-5ms fused silica capillary column (30 m x 0.32 mm i.d. x 0.25 µm film thickness; Agilent). Helium was used as the carrier gas at a constant flow rate of 1.2 mL/min. The inlet temperature was held at 280 °C and the sample was split with a ratio of 1:15. The oven temperature was programmed from 40 °C (initial hold 3 minutes), ramped at 5 °C /min to 60 °C, ramped at 10 °C /min to 190 °C, then ramped at 20 °C /min to 290 °C (final hold 5 minutes). Mass spectral data was collected from 30 to 350 daltons at a rate of 4.45 scans/second. A total of ten samples were prepared from each aliquot. Selected aromatic ions (m/z^+ 78, 91, 92, 105, 106, 119, 120, 133 and 134) were extracted from each chromatogram using a macro (Appendix 3.1) and peak areas integrated. The integration results were written to an ASCII text file in a format suitable for export to a spreadsheet.

3.4.3 Data analysis

Integrated gas chromatographic data was loaded into a spreadsheet (Microsoft Excel®) and normalised. All simulated samples were constructed and normalised in the same spreadsheet. Data analysis by principal components was made using Minitab® (version 13.1 for Windows) loaded on a personal computer. Error ellipses were calculated using the method described by Jackson.[24] Error ellipse calculations were performed using a Minitab macro modified for this work to allow user input of the F-statistic.[25] The macro is provided in Appendix 3.1.

3.5 The effects of normalisation on a simulated data set

The importance of correctly pre-treating a data set prior to multivariate analysis has been the subject of a number of articles.[8,9,11,12,14,15,16,26] The methods available for data normalisation (transformation) span a wide range of computational complexity. Both Deming *et al.*[11] and Rietjens[8] used simulated data to explore the effects of different transformations on a data set. The effects of different normalisation procedures on the PCA results for gas chromatography data may be best explored using simulated data.

3.5.1 Simulated data

Two data sets were constructed based on the simulated data set of ten chromatography peaks used by Rietjens[8] and are shown in Table 3.1. Sample 1A is the original data set published by Rietjens.[8] Sample 2A is a slight modification of Sample 1A and is intended to represent a sample that is distinct from Sample 1A while maintaining a similar pattern overall. Note that the simulated data sets contain "peaks" in a wide range of sizes (areas).

Table 3.1 Basis sets for two different samples, 1A and 2A, used for simulations.

	Peak area	
Peak number	Sample 1A	Sample 2A
1	30000	28000
2	1000	1000
3	60000	65000
4	14000	13800
5	5000	5000
6	10000	10000
7	20000	21000
8	500	400
9	100	100
10	400000	401000

The values listed in Table 3.1 were used as a basis for the construction of different chromatographic runs of the same sample. A relative error, E_r, of ±2% was calculated for each peak using the following equation:

$$E_r = (S_{ign} \times R_{Er}) \qquad \textbf{[3]}$$

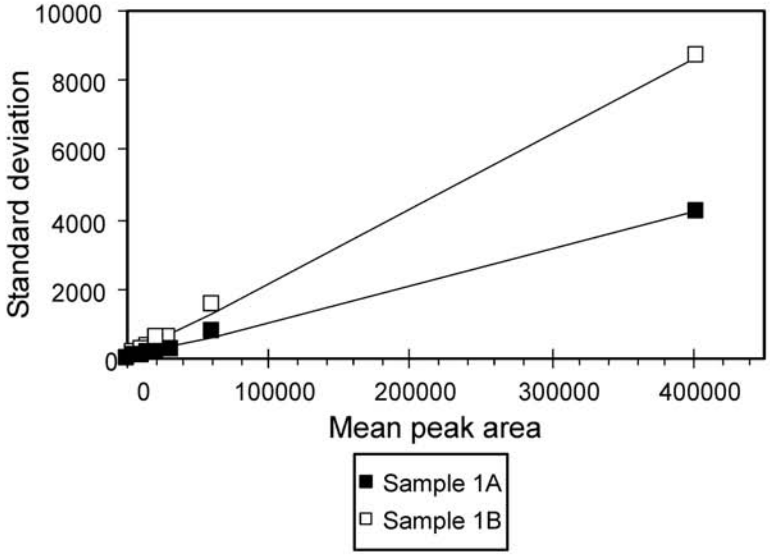
where R_{Er} is the random relative error and S_{ign} has a value of either +1 or −1. In this case R_{Er} is a random number having any value between zero and two. S_{ign} was initially calculated as a random number having a value of either 0 or 1, where the 0 values were first converted to −1 prior to multiplication with R_{Er}. In this way S_{ign} was used to generate positive and negative values of E_r.

These random numbers, E_r, were then applied to the data in Table 3.1 to construct new chromatograms using the following equation:

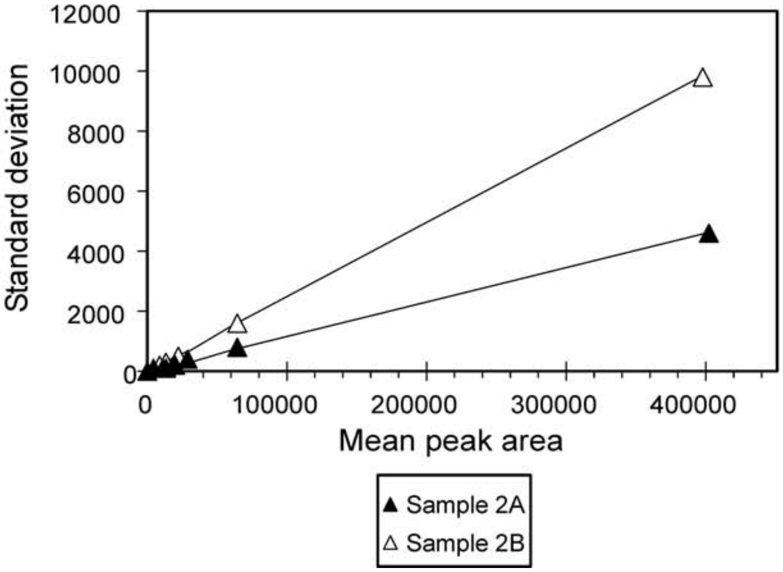
$$Pk_{new} = Pk_{\mathrm{j}} - \left\{ Pk_{\mathrm{j}} \times \frac{E_r}{100} \right\} \qquad \textbf{[4]}$$

where Pk_{j} is the peak area for the j^{th} peak in Table 3.1, and the relative error, E_r, is given by Equation [3]. A total of ten chromatograms were constructed for Sample 1A and ten for Sample 2A. Two new samples, Sample 1B and 2B, each having a relative error of ±4%, were constructed by applying Equations [3] and [4] to the values in the basis set for Samples 1A and 2A respectively (Table 3.1). Ten chromatographic runs were simulated for both Samples 1B and 2B. The "peaks" in all four samples (1A, 1B, 2A and 2B) have a heteroscedastic noise pattern which can be illustrated by plotting the mean against the standard deviation for each peak (Figure 3.3).

Figure 3.3 (a) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 1A and 1B without prior treatment.



(b) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 2A and 2B without prior treatment.



3.5.2 Normalisation methods

Several normalisation methods that are commonly applied to chromatographic data will be described and applied to the simulated data. In the discussion that follows, the subscript i refers to the samples (rows in the data matrix) and the subscript j refers to the variables (columns in the data matrix). The first normalisation method considered is normalisation to a maximum value. This method is commonly applied to mass spectrometry data in which all peak intensities are presented relative to the peak with the greatest intensity. For chromatography data, this method measures the peak area for each peak, a_{ij}, in chromatogram i relative to the peak with the greatest area in that chromatogram, a'_{ij}. This method will be referred to as MS normalisation, or simply MS, and is described by Equation [5].

$$\mathrm{MS} = \frac{a_{ij}}{a'_{ij}} \qquad \textbf{[5]}$$

where, a'_{ij} is the largest peak in chromatogram i, that is, peak number 10 in Table 3.1.

The second method under consideration is normalisation to a constant sum and will be referred to as CS normalisation, or CS (Equation [6]). This method simply divides the peak area of each peak in chromatogram i by the sum of all peaks in that chromatogram:

$$\mathrm{CS} = \frac{a_{ij}}{\sum_{j=1}^{m} a_{ij}} \qquad \textbf{[6]}$$

where j is peak 1 through m (m = 10) in chromatogram i.

Another, more complicated normalisation procedure is described by Wold *et al.*[14] In this method a scaling weight is calculated for each variable. The scaling weight, w_j, for a given variable, j, is the standard deviation of all peak areas for all samples being examined. Each element in the data matrix is divided by its corresponding scaling weight (Equation [7]). Wold *et al.*[14] termed this method "autoscaling" because each variable in the data matrix is scaled to unit variance.

$$\text{Autoscale} = \frac{a_{ij}}{w_j} \qquad \textbf{[7]}$$

where, for a given variable, j, i = 1 to n. The scaling weights are flexible in the sense that each variable, j, has its own unique scaling weight.

The last two normalisation methods were described by Rietjens.[8] These two methods require more complex calculations to be performed on the data. One method will be termed "weight factor" normalisation and is described by Equations [8] and [9].

$$w_f = \frac{50000}{\sum_{i=1}^{n}\left[\frac{a_{ij}}{n}\right]} \qquad \textbf{[8]}$$

In Equation [8], 50000 is an arbitrary value, and $\sum_{i=1}^{n}\left[\frac{a_{ij}}{n}\right]$ is the mean value of variable j over all chromatograms in the data matrix where n is the total number of chromatograms in the data matrix. The arbitrary value of 50000 was chosen by Rietjens[8] because it is the rounded mean value of peaks 1 through 10 in Sample 1A, Table 3.1. The Weight Factor is then calculated as shown in Equation [9]:

$$\text{Weight Factor} = \frac{w_f \times a_{ij}}{\sum_{j=1}^{m}\left[w_f \times a_{ij}\right]} \qquad \textbf{[9]}$$

where w_f is the weighting factor from Equation [8] and $\sum_{j=1}^{m}\left[w_f \times a_{ij}\right]$ is the sum of all weighted peaks in chromatogram i.

The last method is a natural logarithmic normalisation and will be referred to as LN. This method is described by the following equation:

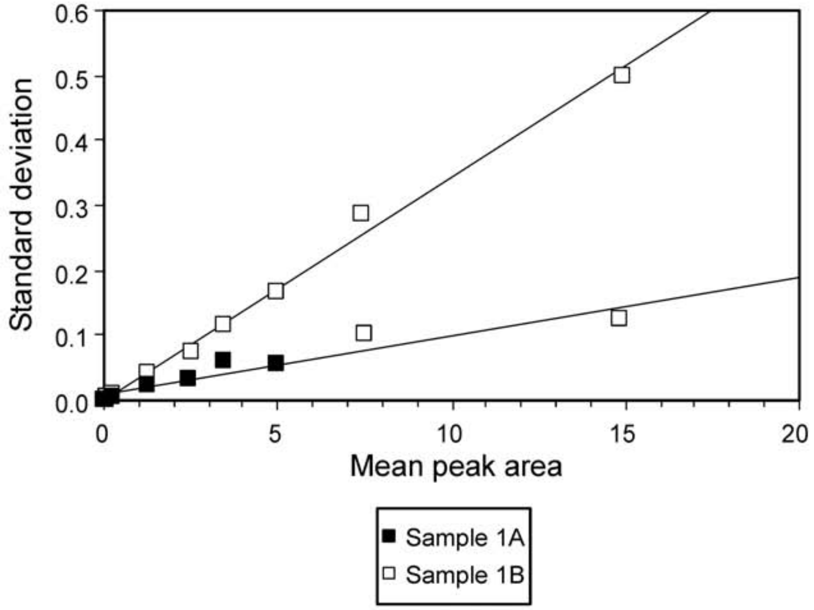
$$\mathrm{LN} = \ln(a_{\mathrm{ij}}) - \left\{ \sum_{\mathrm{j=1}}^{\mathrm{m}} \ln(a_{\mathrm{ij}}) \Big/ \mathrm{m} \right\} \qquad \textbf{[10]}$$

where $\left\{ \sum_{\mathrm{j=1}}^{\mathrm{m}} \ln(a_{\mathrm{ij}}) \Big/ \mathrm{m} \right\}$ is the mean natural logarithmic value of all variables j in chromatogram i.

3.5.3 Comparison of normalisation methods

As a first step, each normalisation method must reduce, as much as possible, the heteroscedastic noise (see Figure 3.3) to homoscedastic noise. The mean versus standard deviation of the peak areas for each of the ten chromatograms in the four different samples (1A, 1B, 2A, and 2B) were plotted after each of the normalisation methods were performed on the data (Figures 3.4 to 3.8). Although each normalisation method significantly reduced the magnitude of the standard deviation range in the data set, not all transformations performed equally. The size of the standard deviation range for the MS, CS and Autoscale transformations is one order of magnitude greater than the range for the LN transformation and two orders of magnitude greater than the range for the Weight Factor transformation. Clearly the Weight Factor and LN transformations were more successful in eliminating heteroscedastic noise than the other procedures.

Figure 3.4 (a) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 1A and 1B after MS transformation.



(b) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 2A and 2B after MS transformation.

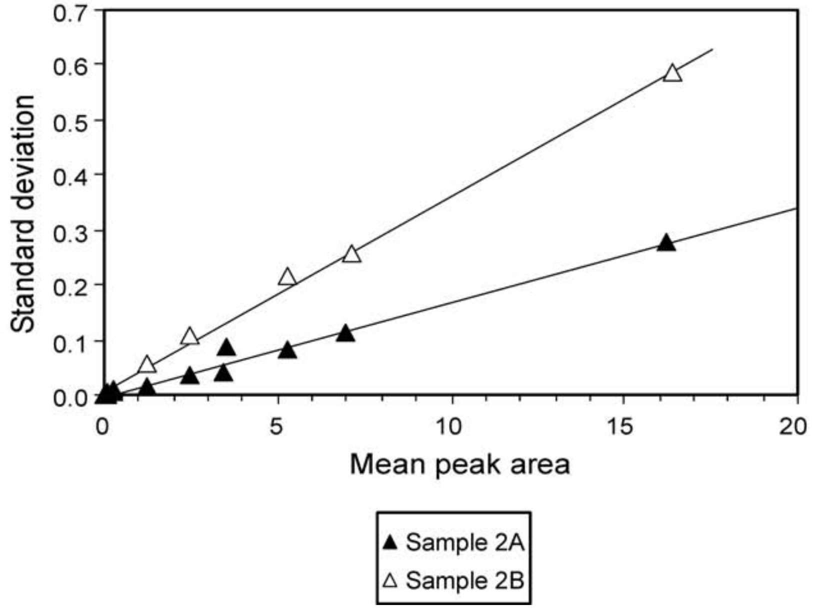
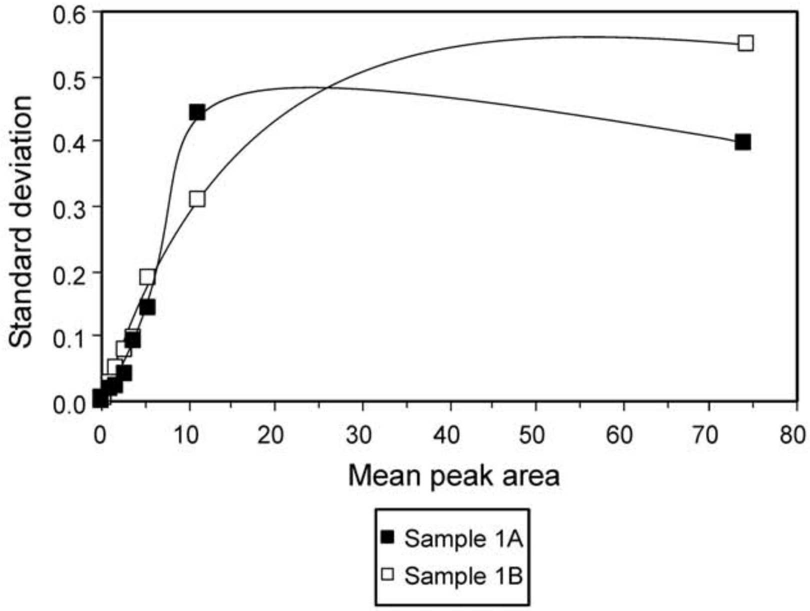


Figure 3.5 (a) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 1A and 1B after CS transformation.



(b) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 2A and 2B after CS transformation.

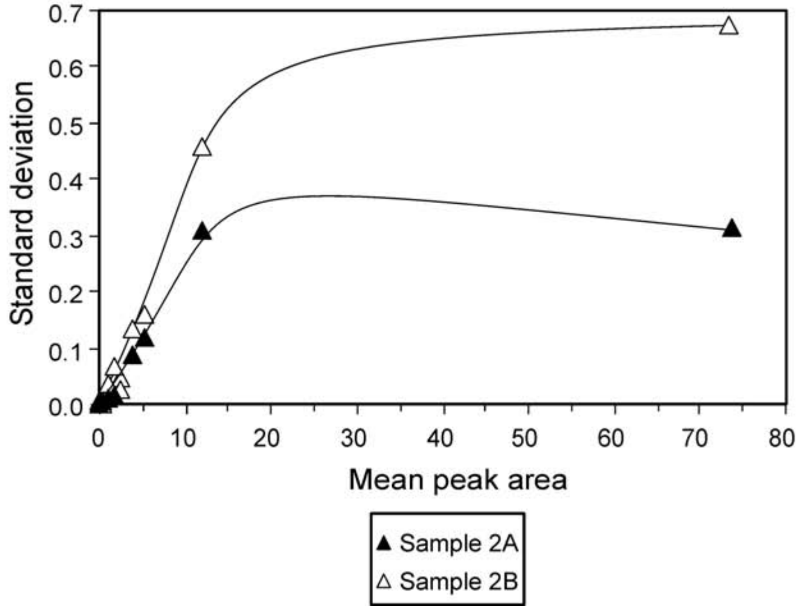
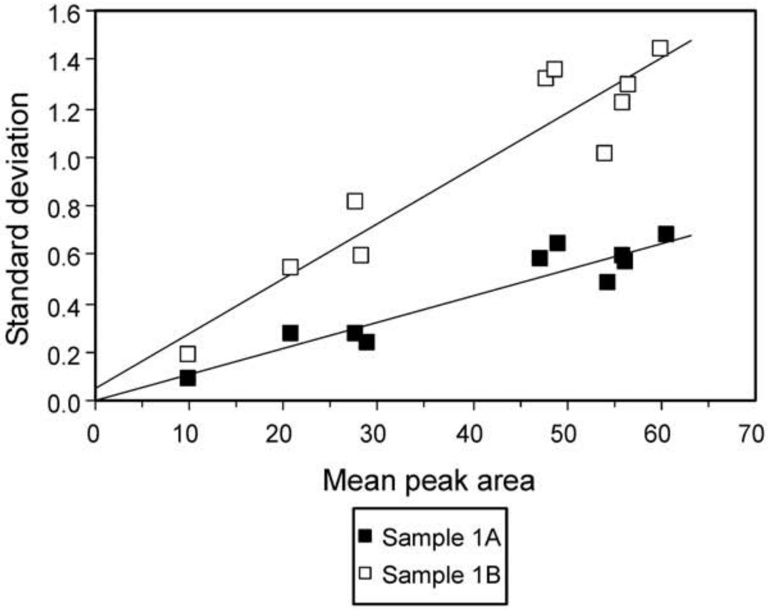


Figure 3.6 (a) Plot of mean peak area against standard deviation for ten simulated chromatograms from Sample 1A and 1B after Autoscale transformation.



(b) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 2A and 2B after Autoscale transformation.

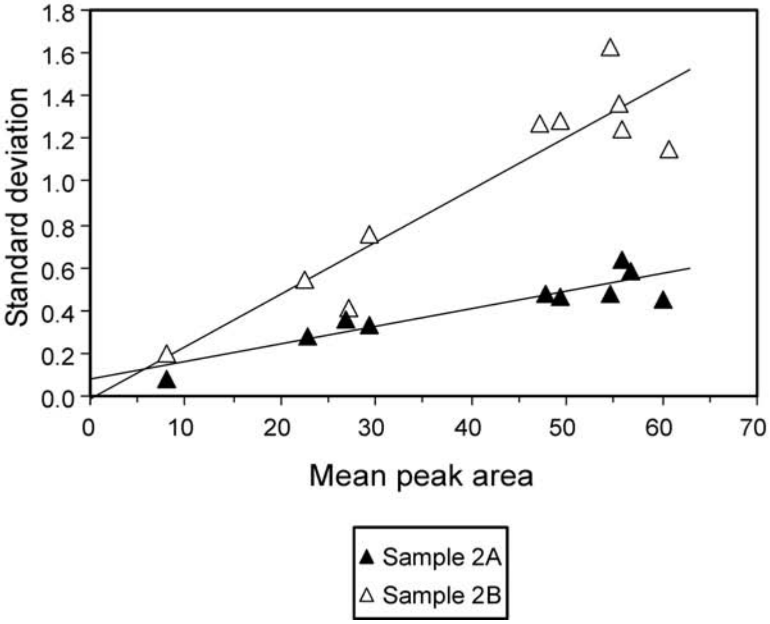
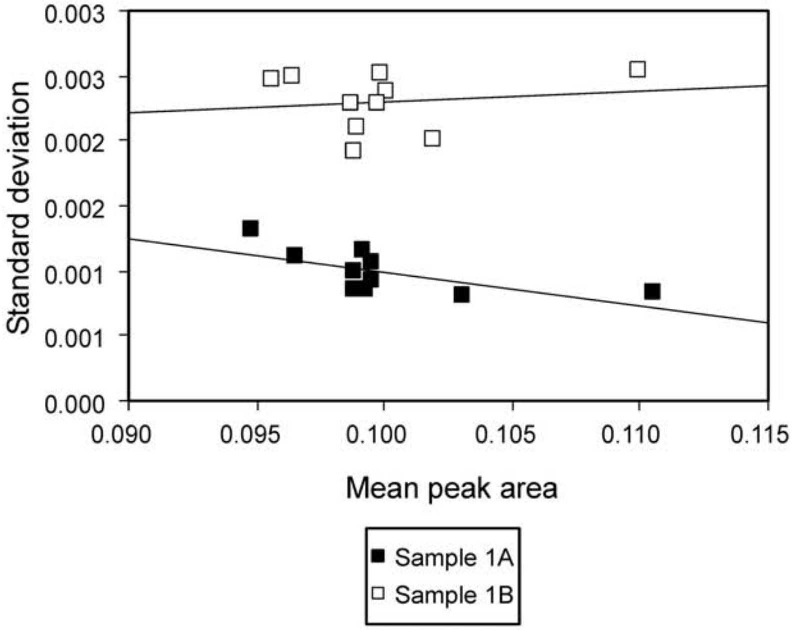


Figure 3.7 (a) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 1A and 1B after Weight Factor transformation.



(b) Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 2A and 2B after Weight Factor transformation.

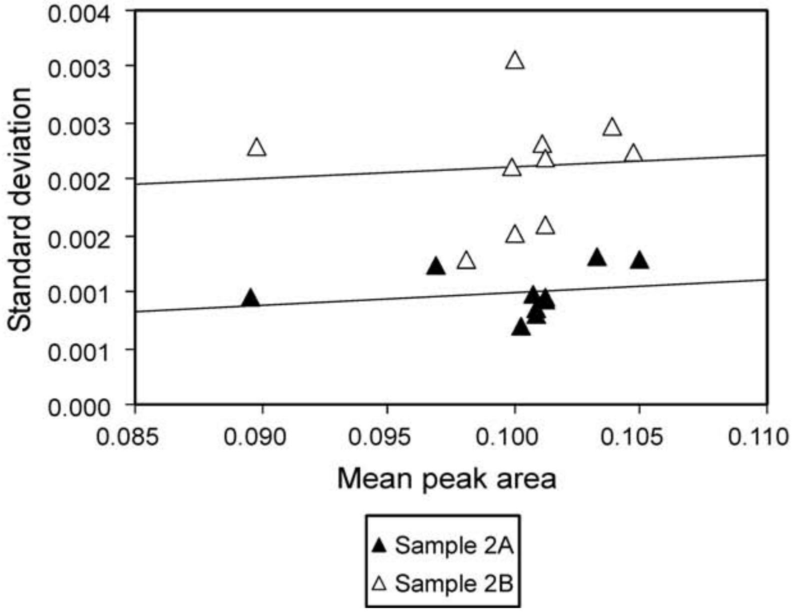
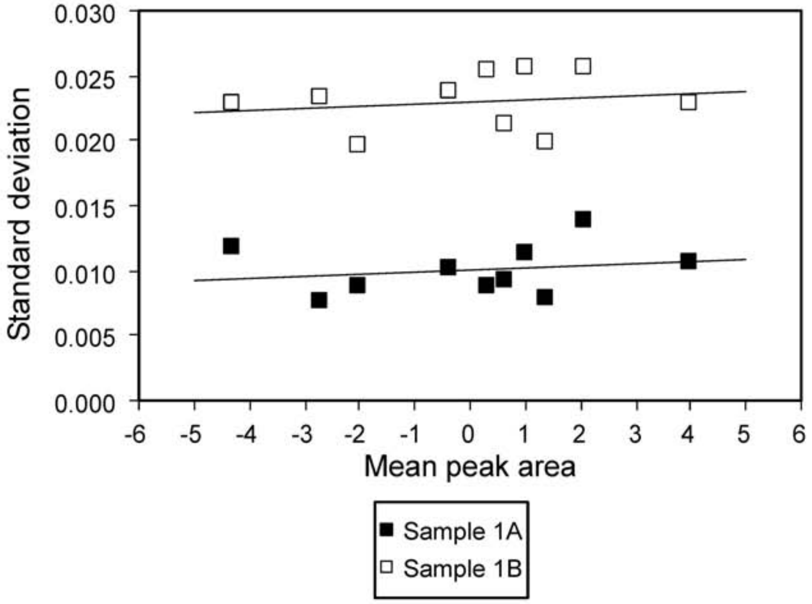
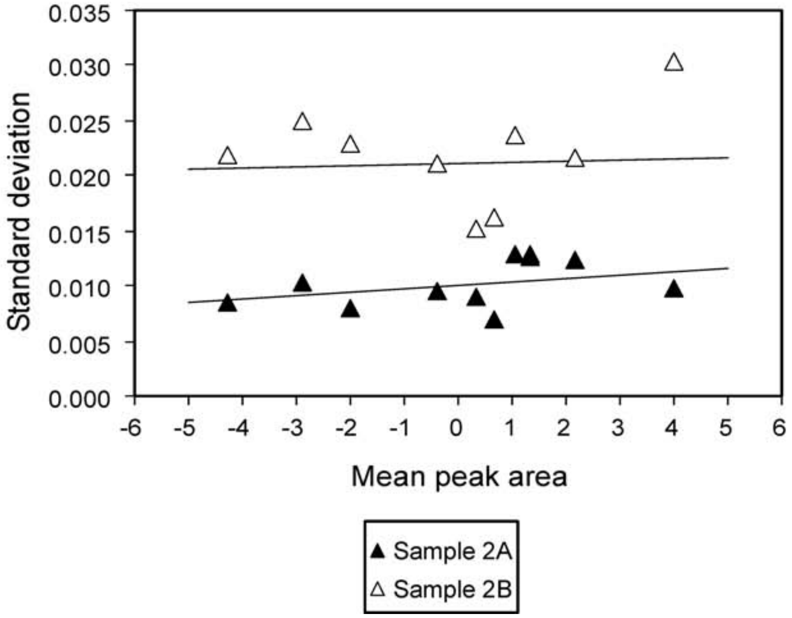


Figure 3.8 Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 1A and 1B after LN transformation.



Plot of mean peak area against standard deviation for ten simulated chromatograms from Samples 2A and 2B after LN transformation.



In order to more fully compare the normalisation methods, the different effects that can be applied to chromatographic data must be considered. Three different effects have been identified for study as follows:

- relative error;
- dilution;
- evaporation.

A good normalisation method must be able to account for each of these three effects without compromising the underlying correlation structure of the data. Any spurious changes to the correlation structure of the data caused by a normalisation method are undesirable and may have a deleterious effect on the PCA results.

3.5.3.1 Changes to the relative error

3.5.3.1.1 Changing the relative error between the same samples

The effects of normalisation on chromatograms having different relative errors is explored first. Samples 1A and 2A have a relative error of ±2%. Two new samples, Sample 1B and 2B, each having a relative error of ±4%, were constructed by applying Equations [3] and [4] to the values in the basis set for Samples 1A and 2A respectively (Table 3.1). Ten chromatographic runs were simulated for both Samples 1B and 2B. The resulting data matrix consisted of 40 chromatograms ($n = 40$ rows) each with 10 variables ($m = 10$ columns). That is, each of the four samples, Samples 1A, 2A, 1B and 2B, contributed ten chromatograms to the data matrix. The samples in the data matrix were normalised using each of the five methods described above. Principal component analysis was performed on the raw data as well as on the results from each of the five normalisation methods. The amount of variance in the data that the first five PCs account for is tabulated in Table 3.2, and the PC score plots are shown in Figure 3.10.

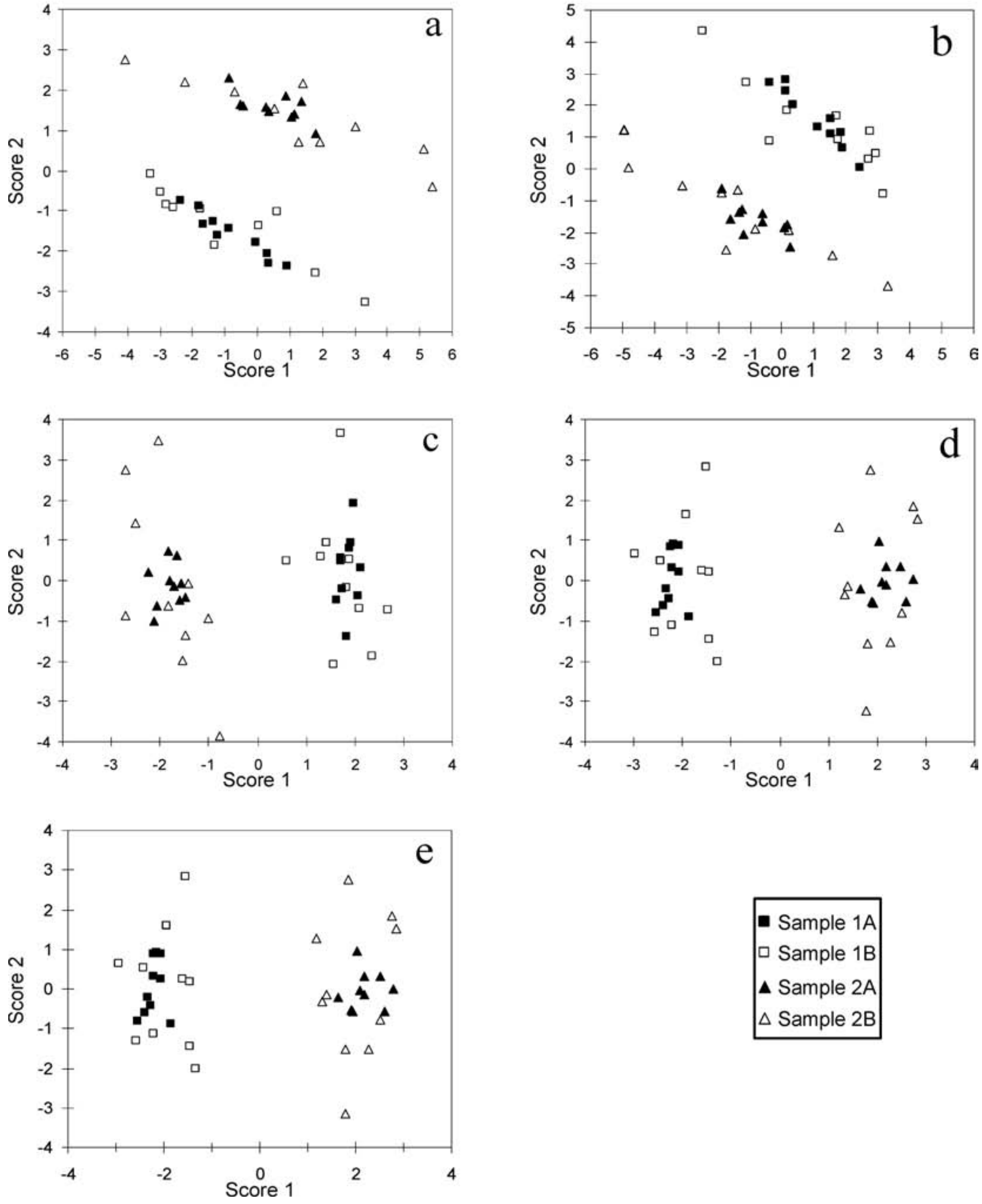
Table 3.2. PCA results from each normalisation method for simulated data where peaks in Samples 1A and 2A have a relative error of ±2% and peaks in Samples 1B and 2B have a relative error of ±4%.

PC	Eigenvalue for each PC					
	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	3.5436	4.4709	4.0698	3.5436	4.6343	4.6446
2	2.0325	2.2779	3.4257	2.0325	1.4912	1.4841
3	1.1181	0.4523	0.6254	1.1181	1.0121	1.0181
4	0.9113	0.3924	0.5344	0.9113	0.9241	0.9200
5	0.8565	0.3579	0.4942	0.8565	0.7444	0.7500

PC	Cumulative proportion of variance accounted for by each PC (%)					
	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	35.4	49.7	40.7	35.4	46.3	46.4
2	55.8	80.0	75.0	55.8	61.3	61.3
3	66.9	85.0	81.2	66.9	71.4	71.5
4	76.1	89.4	86.6	76.1	80.6	80.7
5	84.6	93.3	91.5	84.6	88.1	88.2

From the plots of the first two scores from the PC analysis it can be seen that Sample 1 and Sample 2 can be differentiated. For a given sample pair (*e.g.* Samples 1A and 1B) the expectation is that the data will not be separated by PCA because the only difference between the two samples is the amount of scatter in the data. This is what is observed for each sample pair in the score plots in Figure 3.9 indicating that all of the normalisation methods performed correctly. Note that PCA of the data that had been normalised using the Autoscale method gave identical results to PCA of the raw data (Table 3.2). Also, the PCA results for the Weight Factor and LN normalisation methods are virtually identical. Rietjens[8] found that after applying both the Weight Factor and the LN normalisation methods to real pyrolysis gas chromatography data, the PCA results were virtually identical as well.

Figure 3.9 Change in relative error: Samples 1A and 2A have ±2% relative error, Samples 1B and 2B have ±4% relative error. Plots of first two PC scores using the following data transformations: MS (a); CS (b); Autoscale (c); Weight Factor (d); and LN (e).



3.5.3.1.2 Increasing the relative error of the largest peak

The effect of normalisation on the PCA results when the largest peak has a relative error that is greater than the error in the other peaks was also examined. Chromatograms were constructed from the basis set (Table 3.1) such that the first nine peaks were each given a random relative error of ±2%. The last and largest peak was given a random relative error between ±2% and ±4%. Ten chromatographic runs were simulated for each sample resulting in a data matrix consisting of 40 chromatograms (40 rows) of 10 variables each (10 columns). The relative standard deviations calculated from the raw data for Sample 1 (n = 20) and Sample 2 (n = 20) are presented in Table 3.3.

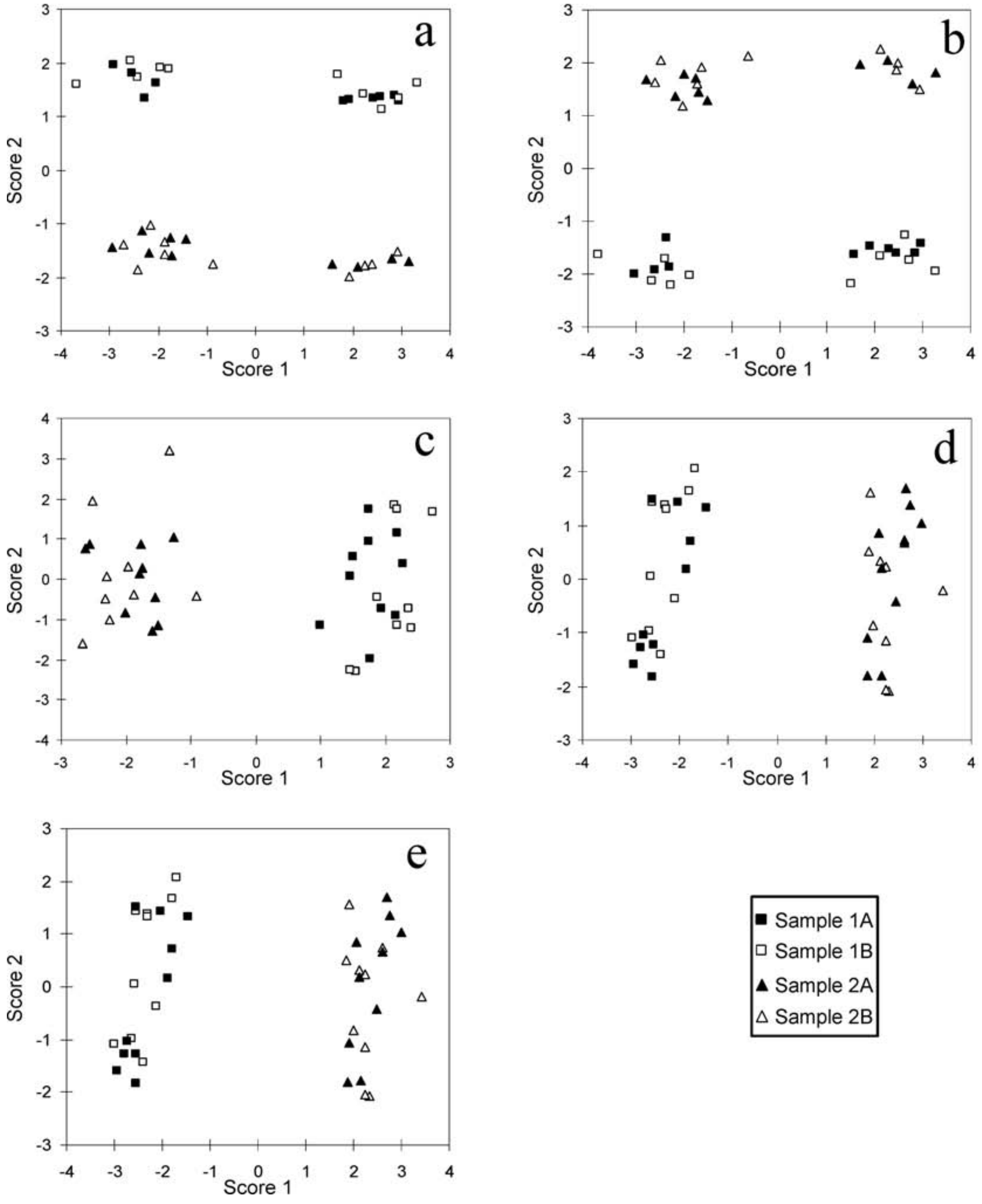
Table 3.3 Relative standard deviations for Sample 1 and Sample 2.

	Relative standard deviation (%)			
Peak	Sample 1A	Sample 2A	Sample 1B	Sample 2B
1	0.97	1.08	1.16	1.08
2	1.18	1.37	1.26	1.02
3	1.35	1.33	0.93	1.19
4	0.70	0.82	0.95	1.16
5	1.26	1.11	1.08	1.31
6	0.78	1.22	0.80	0.75
7	1.36	1.25	1.07	1.19
8	1.09	1.39	1.15	1.11
9	1.18	1.33	1.22	1.08
10	3.35	2.97	3.08	2.93

The samples in the data matrix were normalised using each of the five methods described in detail above. Principal component analysis was performed on the raw data as well as on the results from each of the five normalisation methods. The amount of variance in the data that the first five PCs accounted for is tabulated in Table 3.4 and the five PC score plots are shown in Figure 3.11.

In the original data matrix the two samples that make up one pair (*e.g.* Sample 1A and 1B) are the same except for minor differences in the scatter of data from the random number generation. Therefore one would expect the PCA results to show two groups, one for Samples 1A and 1B and another for Samples 2A and 2B. The plots of the first two PC scores (Figure 3.10) show that Sample 1 and Sample 2 can be differentiated by Autoscale,

Figure 3.10 Change in relative error: all peaks in all samples have ±2% relative error except peak 10 which has a relative error between ±2% and ±4%. Plots of first two PC scores using the following data transformations: MS (a); CS (b); Autoscale (c); Weight Factor (d); and LN (e).



Weight Factor and LN normalisation methods. Application of the MS and CS normalisation methods gave erroneous results by causing the data for the first PC score to split depending on whether the relative error for peak 10 was between +2% and +4% or between −2% and −4%.

Table 3.4 PCA results from each normalisation method for simulated data where peaks 1 through 9 have a relative error of ±2% and peak 10 has a relative error between ±2% and ±4% for all samples.

	Eigenvalue for each PC					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	4.3182	5.8483	5.8393	4.3182	5.7351	5.751
2	1.5983	2.5499	3.1987	1.5983	1.6082	1.606
3	1.2140	0.1777	0.2810	1.2140	0.9409	0.9353
4	0.8631	0.1374	0.2217	0.8631	0.5445	0.541
5	0.8308	0.0868	0.1485	0.8308	0.4656	0.4635

	Cumulative proportion of variance accounted for by each PC (%)					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	43.2	65.0	58.4	43.2	57.4	57.5
2	59.2	93.3	90.4	59.2	73.4	73.6
3	71.3	95.3	93.2	71.3	82.8	82.9
4	79.9	96.8	95.4	79.9	88.3	88.3
5	88.2	97.8	96.9	88.2	92.9	93.0

3.5.3.2 Differences in dilution between samples

When preparing real samples for analysis, slight differences in dilution between samples is an unavoidable source of experimental error. The effects of normalisation on the PCA results for chromatograms that have different dilutions were examined. Two new samples, Samples 1B and 2B, were made by “diluting” the corresponding basis set values in Table 3.1. A dilution factor of 10% was applied to the basis set in Table 3.1 to create a new basis set of four samples, as shown in Table 3.5. Ten chromatograms were constructed for each sample in the new basis set (Table 3.5) by applying a random relative error of ±2% to each sample.

Table 3.5 Simulated data sets, undiluted and 10% diluted, for two different samples.

	Peak area				
	Undiluted		*10% diluted*		
Peak number	Sample 1A	Sample 2A	Sample 1B	Sample 2B	Dilution factor
1	30000	28000	27000	25200	0.10
2	1000	1000	900	900	0.10
3	60000	65000	54000	58500	0.10
4	14000	13800	12600	12420	0.10
5	5000	5000	4500	4500	0.10
6	10000	10000	9000	9000	0.10
7	20000	21000	18000	18900	0.10
8	500	400	450	360	0.10
9	100	100	90	90	0.10
10	400000	401000	360000	360900	0.10

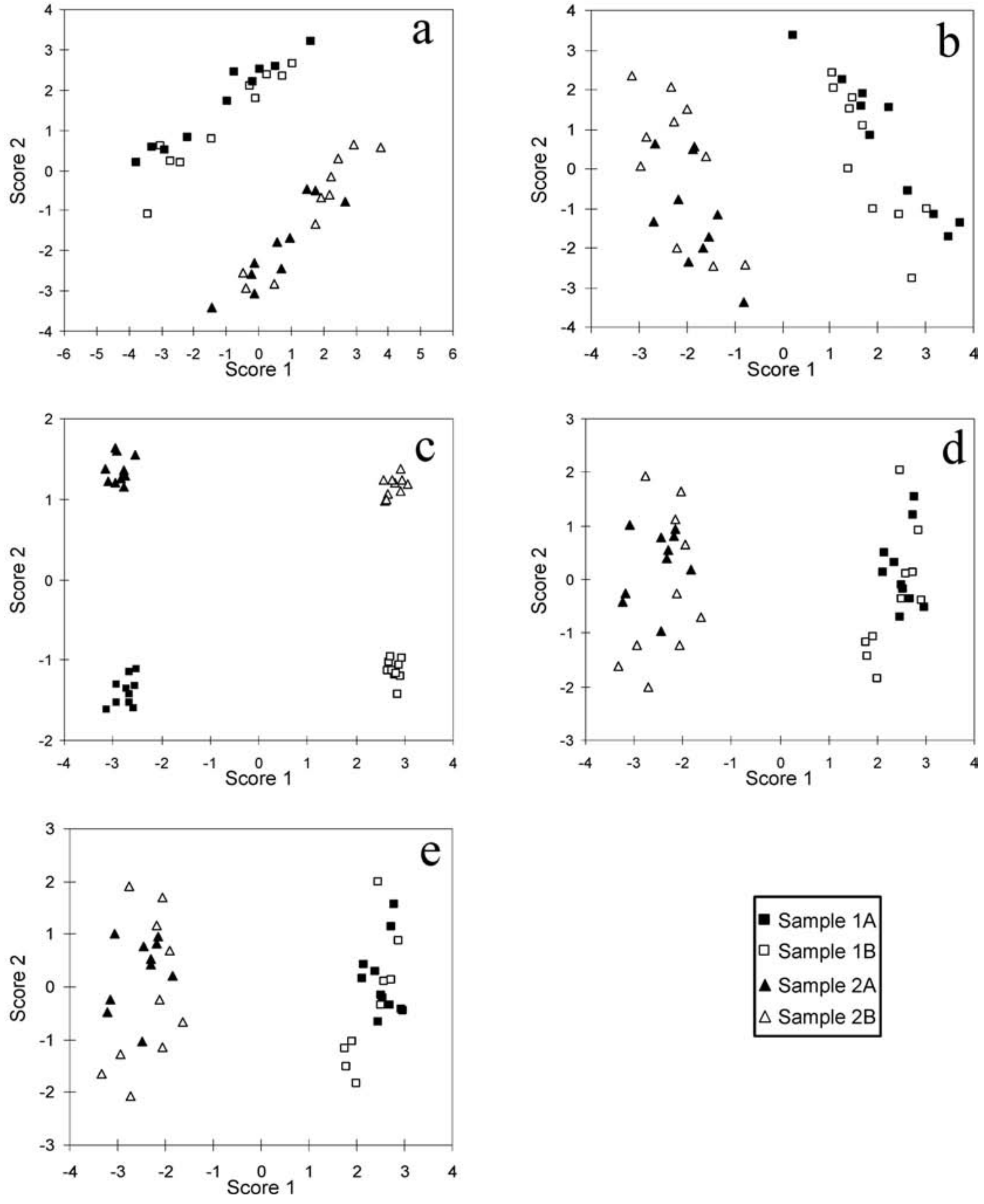
The data matrix consisted of 40 chromatograms (40 rows) of 10 variables each (10 columns), where each of the four samples, Samples 1A, 2A, 1B and 2B, contributed ten chromatograms. Principal component analysis was performed on the raw data as well as on the results from each of the five normalisation methods and the results are given in Table 3.6. The PC score plots for the undiluted and 10% diluted samples are shown in Figure 3.11.

Table 3.6 PCA results from each normalisation method for simulated undiluted Samples 1A and 2A and 10% diluted Samples 1B and 2B.

	Eigenvalue for each PC					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	8.0609	3.6650	4.7412	8.0609	6.3083	6.3252
2	1.6707	3.5600	3.0411	1.6707	1.0744	1.0653
3	0.0676	0.5892	0.7374	0.0676	0.9131	0.9118
4	0.0516	0.4293	0.5140	0.0516	0.5683	0.5671
5	0.0383	0.2616	0.3244	0.0383	0.4773	0.4750

	Cumulative proportion of variance accounted for by each PC (%)					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	80.6	40.7	47.4	80.6	63.1	63.3
2	97.3	80.3	77.8	97.3	73.8	73.9
3	98.0	86.8	85.2	98.0	83.0	83.0
4	98.5	91.6	90.3	98.5	88.6	88.7
5	98.9	94.5	93.6	98.9	93.4	93.4

Figure 3.11 Differences in dilution: Samples 1A, 1B, 2A and 2B have ±2% relative error. Samples 1B and 2B have been diluted by 10% relative to samples 1A and 2A, respectively. Plots of first two PC scores using the following data transformations: MS (a); CS (b); Autoscale (c); Weight Factor (d); and LN (e).



The two samples in a given sample pair (*e.g.* Sample 1A and 1B) are actually the same, except that one has been diluted relative to the other, therefore, only two groups, one containing Samples 1A and 1B, and the other containing Samples 2A and 2B, are expected after normalisation and PCA. The PC score plot for the raw data is the same as the PC score plot for the Autoscale normalisation method (Figure 3.11c) and shows the importance of taking dilution into account. When differences in dilution are not taken into account an erroneous separation of the samples in the PCA results will occur.

Another basis set of four samples was made by "diluting" the corresponding basis set values in Table 3.1 by a dilution factor of 50% (Table 3.7). Ten chromatograms were constructed for each sample in the second basis set (Table 3.7) by applying a random relative error of ±2% to each sample.

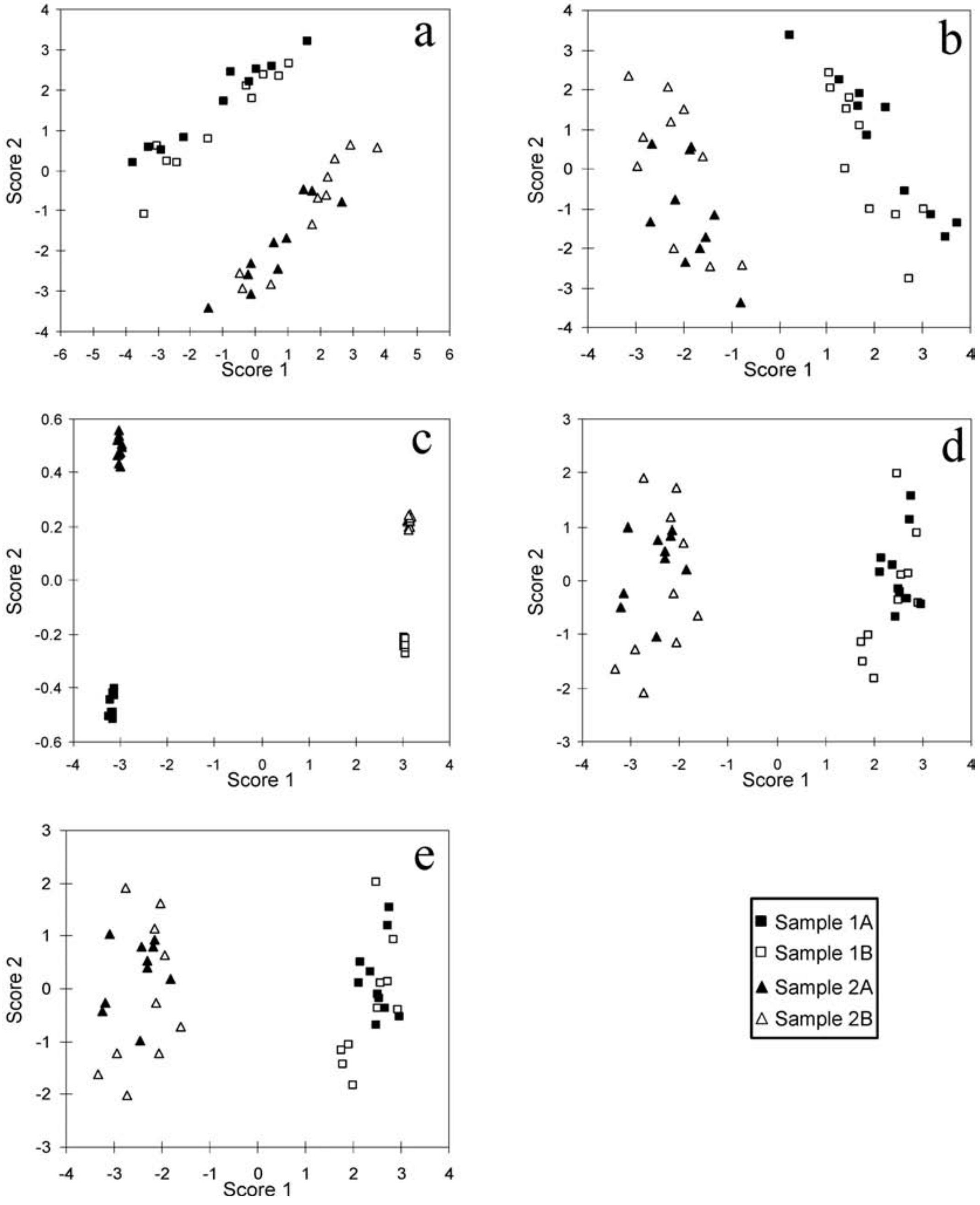
Table 3.7 Simulated data sets, undiluted and 50% diluted, for two different samples.

	Peak area				
	Undiluted		*50% diluted*		
Peak number	Sample 1A	Sample 2A	Sample 1B	Sample 2B	Dilution factor
1	30000	28000	15000	14000	0.50
2	1000	1000	500	500	0.50
3	60000	65000	30000	32500	0.50
4	14000	13800	7000	6900	0.50
5	5000	5000	2500	2500	0.50
6	10000	10000	5000	5000	0.50
7	20000	21000	10000	10500	0.50
8	500	400	250	200	0.50
9	100	100	50	50	0.50
10	400000	401000	200000	200500	0.50

Principal component analysis was performed on the raw data, as well as on the results of each of the five normalisation methods. The PCA results are given in Table 3.8 and the PC score plots for the undiluted and 50% diluted samples are shown in Figure 3.12.

Comparison of Figures 3.11 and 3.12 shows that the degree of dilution had no affect on the PCA results for the MS, CS, Weight Factor and LN normalisation methods. Increasing the dilution factor from 10% to 50% caused the four samples to separate further after PCA was

Figure 3.12 Differences in dilution: Samples 1A, 2A,1B and 2B have ±2% relative error. Samples 1B and 2B have been diluted by 50% relative to samples 1A and 2A, respectively. Plots of first two PC scores using the following data transformations: MS (a); CS (b); Autoscale (c); Weight Factor (d); and LN (e).



performed on the results obtained from the Autoscale transformation.

Table 3.8 PCA results from each normalisation method for simulated undiluted Samples 1A and 2A and 50% diluted Samples 1B and 2B.

PC	Eigenvalue for each PC					
	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	9.8455	3.6650	4.7412	9.8455	6.3078	6.3252
2	0.1459	3.5600	3.0411	0.1459	1.0746	1.0653
3	0.0022	0.5892	0.7374	0.0022	0.9133	0.9118
4	0.0016	0.4293	0.5140	0.0016	0.5684	0.5671
5	0.0012	0.2616	0.3244	0.0012	0.4773	0.4750

PC	Cumulative proportion of variance accounted for by each PC (%)					
	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	98.5	40.7	47.4	98.5	63.1	63.3
2	99.9	80.3	77.8	99.9	73.8	73.9
3	99.9	86.8	85.2	99.9	83.0	83.0
4	100.0	91.6	90.3	100.0	88.6	88.7
5	100.0	94.5	93.6	100.0	93.4	93.4

3.5.3.3 Effects of evaporation

The effect of normalisation on the PCA results for chromatograms whose first few peaks have undergone a simulated “evaporation” was explored. Two new samples in the basis set, Samples 1B and 2B, were made by “evaporating” the corresponding basis set values in Table 3.1. A simplified evaporation was simulated for each sample in Table 3.1 by applying an evaporation factor of 10% to the first peak and then gradually reducing the factor over subsequent peaks to create a new basis set of four samples. No evaporation factor was applied to the last four peaks. The new basis set is shown in Table 3.9. Ten chromatograms were constructed for each sample in the new basis set (Table 3.9) by applying a random relative error of ±2% to each sample.

Table 3.9 Simulated data sets, unevaporated and evaporated, for two different samples.

	Peak area				
	Unevaporated		*Evaporated*		
Peak number	Sample 1A	Sample 2A	Sample 1B	Sample 2B	Evaporation factor
1	30000	28000	27000	25200	0.10
2	1000	1000	920	920	0.08
3	60000	65000	57000	61750	0.05
4	14000	13800	13580	13386	0.03
5	5000	5000	4900	4900	0.02
6	10000	10000	9900	9900	0.01
7	20000	21000	20000	21000	0.00
8	500	400	500	400	0.00
9	100	100	100	100	0.00
10	400000	401000	400000	401000	0.00

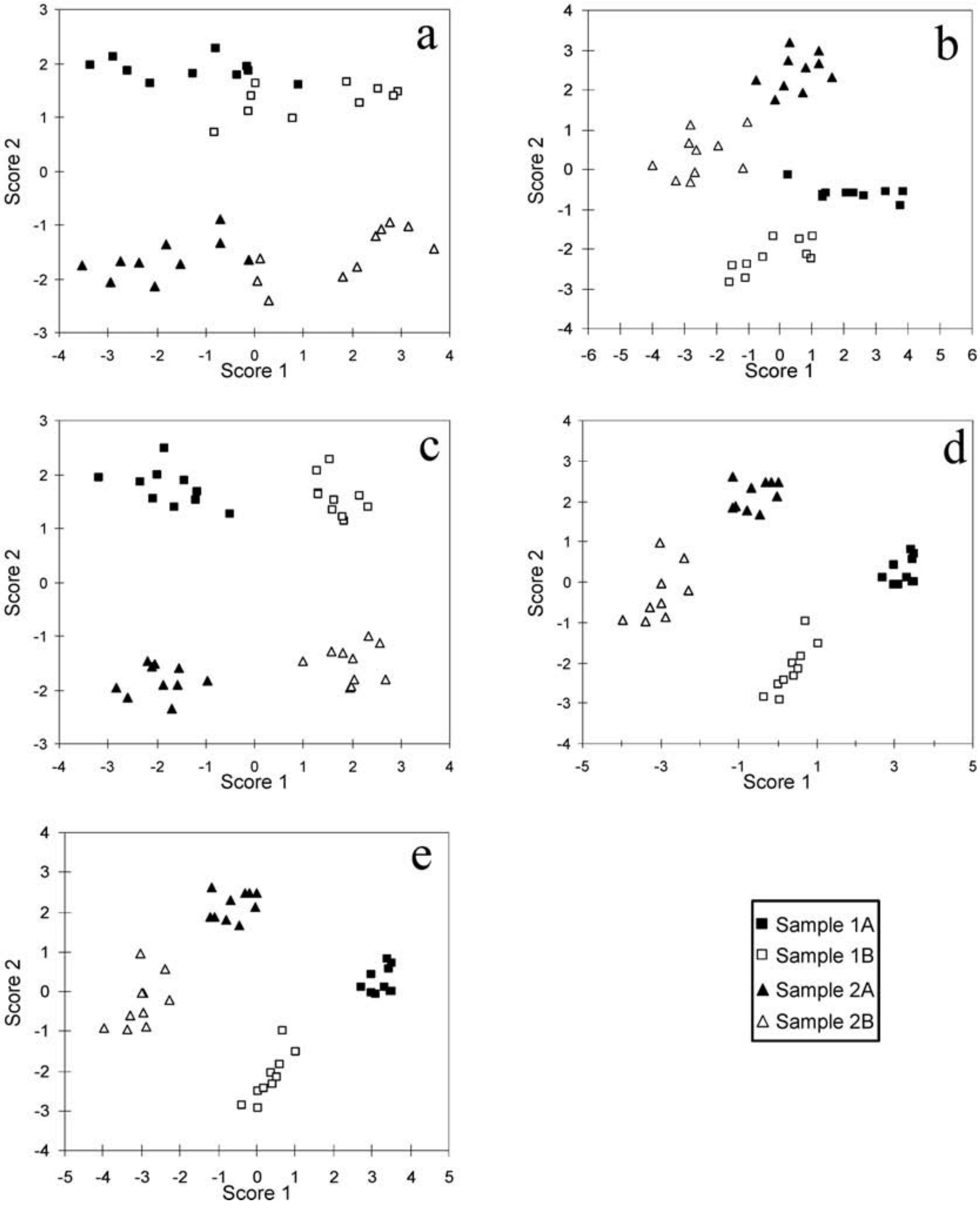
The data matrix consisted of 40 chromatograms (40 rows) of 10 variables each (10 columns), where each of the four samples, Samples 1A, 2A, 1B and 2B, contributed ten chromatograms. Principal component analysis was performed on the raw data as well as on the results of each of the five normalisation methods and the results are given in Table 3.10. The PC score plots for the unevaporated and evaporated samples are shown in Figure 3.13.

Table 3.10 PCA results from each normalisation method after simulated evaporation of Samples 1A and 2A to produce Samples 1B and 2B respectively.

	Eigenvalue for each PC					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	3.7652	4.2028	3.8621	3.7652	5.3083	5.3166
2	2.9683	2.7577	3.0834	2.9683	2.6712	2.6694
3	1.2797	1.2263	1.9320	1.2797	0.7465	0.7418
4	0.9759	0.2843	0.4341	0.9759	0.5704	0.5658
5	0.3872	0.1868	0.2397	0.3872	0.3152	0.3183

	Cumulative proportion of variance accounted for by each PC (%)					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	37.7	46.7	38.6	37.7	53.1	53.2
2	67.3	77.3	69.5	67.3	79.8	79.9
3	80.1	91.0	88.8	80.1	87.3	87.3
4	89.9	94.1	93.1	89.9	93.0	92.9
5	93.8	96.2	95.5	93.8	96.1	96.1

Figure 3.13 Effects of evaporation: Samples 1A, 2A, 1B and 2B have ±2% relative error. Samples 1B and 2B have undergone a simulated evaporation relative to samples 1A and 2A, respectively. Plots of first two PC scores using the following data transformations: MS (a); CS (b); Autoscale (c); Weight Factor (d); and LN (e).



After evaporation, all four samples (two unevaporated and two evaporated) are expected to be different from each other. The expected result after PCA is four groups, one for each sample, which is observed in the plots of the first two scores in Figure 3.13. The PC score plots demonstrate that the MS and CS transformations were not as successful as the other normalisation methods in eliciting the true differences between sample pairs.

3.6 Application to real data

The five normalisation procedures were applied to real gas chromatography data obtained from a sample of gasoline. A sample of regular unleaded gasoline was taken to different levels of evaporation and analysed by GC-MS. Selected ions characteristic of some of the lower molecular weight aromatic compounds were extracted from the chromatograms and the integrated areas for eleven of these compounds were obtained. Six levels of evaporation were examined: 0%, 7.2%, 9.0%, 13.2%, 17.5% and 21.7% by weight. The aromatic compounds were chosen because they span a large boiling point range and a large concentration range in gasoline. The peaks selected for analysis are listed in Table 3.11 and are shown in Figure 3.14.

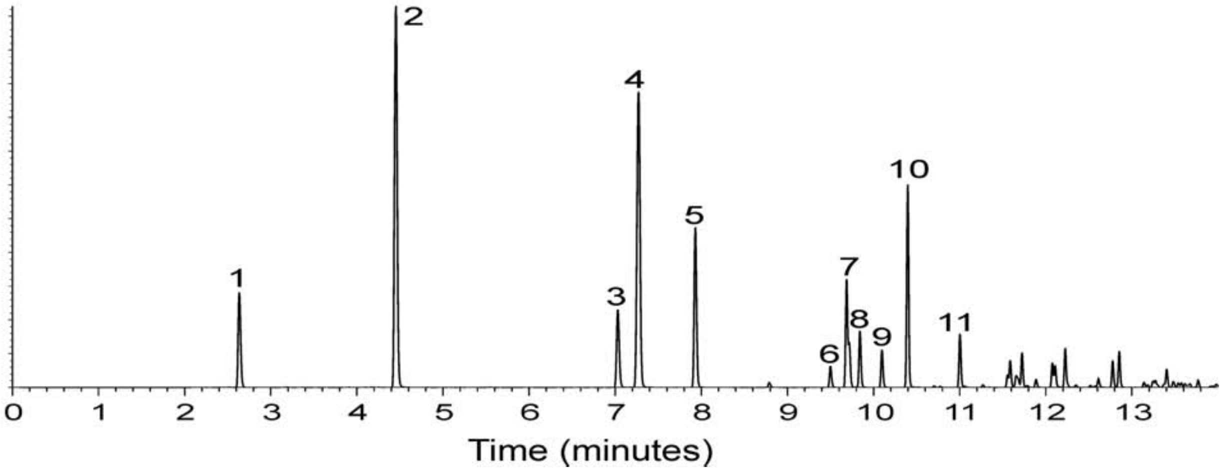
Table 3.11 Selected peaks from aromatic extracted ion chromatograms.

Peak Number	Retention Time	Assignment[1]	Boiling Point[2]
1	2.64	Benzene	80.1
2	4.46	Toluene	110.6
3	7.03	ethyl benzene	136.0
4	7.27	p+m-xylene	138.3 + 139.1
5	7.93	o-xylene	144.4
6	9.50	n-propylbenzene	159.2
7	9.69+	3-ethyl-1-methylbenzene +	161.3
	9.72	4-ethyl-1-methylbenzene	162.0
8	9.84	1,3,5-trimethylbenzene	164.7
9	10.10	2-ethyl-1-methylbenzene	165.2
10	10.40	1,2,4-trimethyl benzene	169.3
11	11.01	1,2,3-trimehtylbenzene	176.1

1. Johansen, N.G.; Ettre, L.S.; Miller, R.L. Quantitative analysis of hydrocarbons by structural group type in gasolines and distillates. Journal of Chromatography, 1983; 256(3): 393-417.

2. Weast, R.C. and Graselli, J.G. CRC Handbook of Data on Organic Compounds, 2nd Edition. CRC Press, Boca Raton, USA; 1989.

Figure 3.14 Extracted ion chromatogram (EIC) of selected aromatic ions from a regular unleaded gasoline sample. Peak assignments are given in Table 3.11.



Five aliquots of gasoline were analysed at each level of evaporation. For each level of evaporation, a mean extracted ion chromatogram was calculated from the results of the five corresponding aliquots (Figure 3.15). In order to clearly demonstrate the effect of each normalisation method, a large number of chromatograms are required. It is not practical to make and analyse a large number of aliquots at each level of evaporation so extracted ion chromatograms were simulated for each level of evaporation by using the mean extracted ion chromatograms as the basis set. The simulation was based on a random error of ±1.05%. This error was calculated by finding the percent deviation from the mean for each variable at each evaporation level ($n = 330$, 5 aliquots per evaporation level x 6 evaporation levels x 11 variables). The standard deviation, s, of the percent deviations from the means is 0.00537, thus 95% ($\alpha = 0.05$) of the percent deviations from the mean fall within $1.96s$, or ±1.05%. An 11 x 330 data matrix consisting of the eleven variables (peaks) by 55 chromatograms for each of the six levels of evaporation was constructed. The 55 chromatograms for each evaporation levels were a combination of the original five chromatograms plus 50 simulated chromatograms. Plots of the mean versus standard deviation of each variable (peak) demonstrate that the simulated data is heteroscedastic (Figure 3.16). All five normalisation procedures were applied to the resulting data matrix. The results of the PCA are given in Table 3.12 and in Figures 3.17 and 3.18.

It is clear from Figure 3.15 that there are only subtle differences between the chromatograms at each of the six evaporation levels. Normalisation by MS has already been shown to be unsuitable[11] whilst constant sum normalisation has also been shown to propagate errors in the data set.[8] The PCA score plots for these two methods are shown in Figure 3.17. The PCA results for the raw data and the Autoscale transformation are the same (Table 3.12). The Autoscale transformation was affected by the dilution rather than by any true chemical differences between the different levels of evaporation (Figure 3.18). It can be seen from examination of the vertical axis (amplitude) in Figure 3.15 that the chromatographic amplitude increased slightly with increased evaporation. The change in amplitude is an artefact of the dilution made for each level of evaporation. Normalisation using the Weight Factor and the LN transformations yielded nearly identical results (Table 3.12), therefore, only the PC scores for the LN transformation were plotted (Figure 3.18b).

Figure 3.15 Mean extracted ion chromatograms (EICs) of an unevaporated and evaporated gasoline sample.

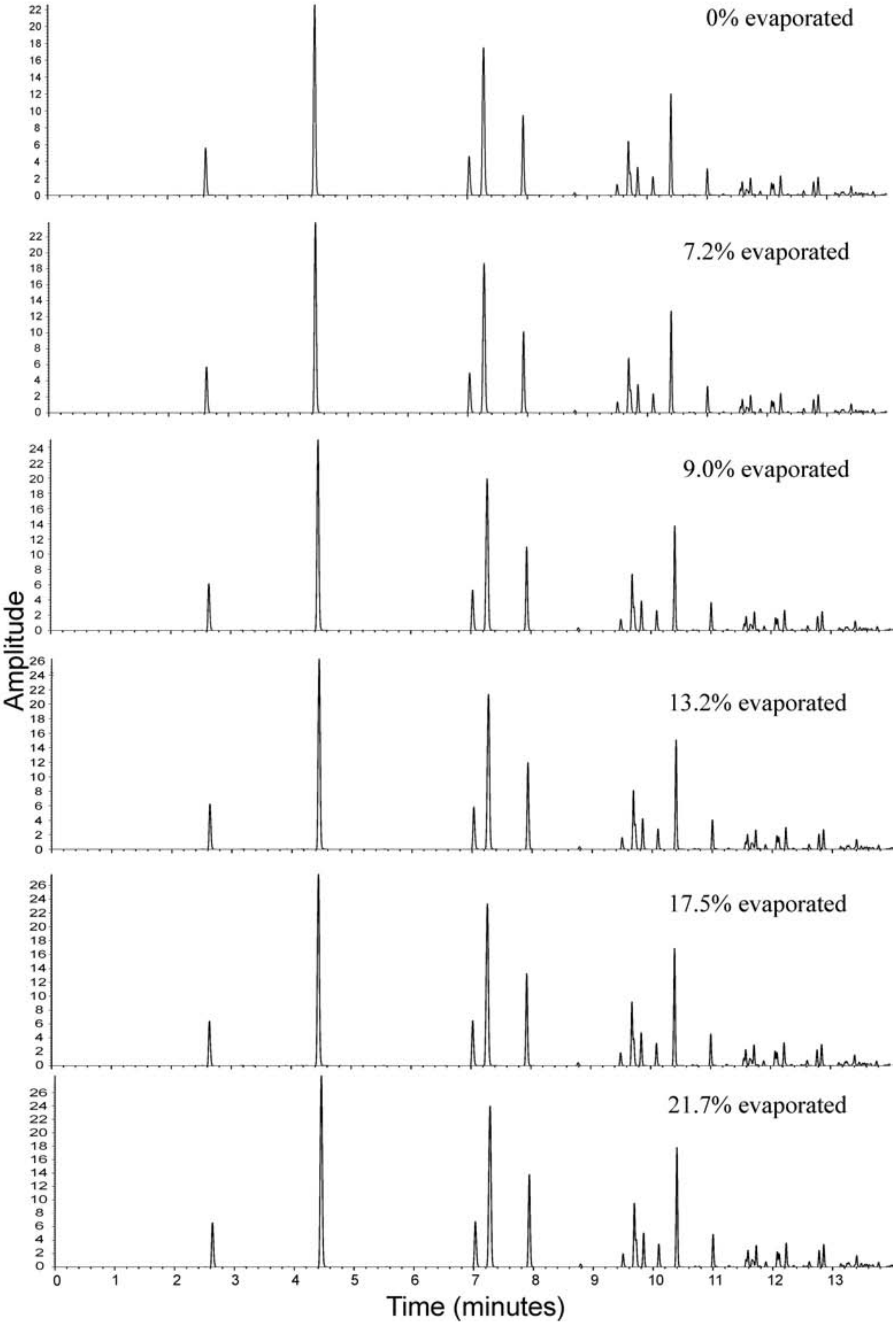


Figure 3.16 Plot of mean peak area against standard deviation for simulated extracted ion chromatograms obtained from the six evaporation levels of a regular unleaded gasoline sample.

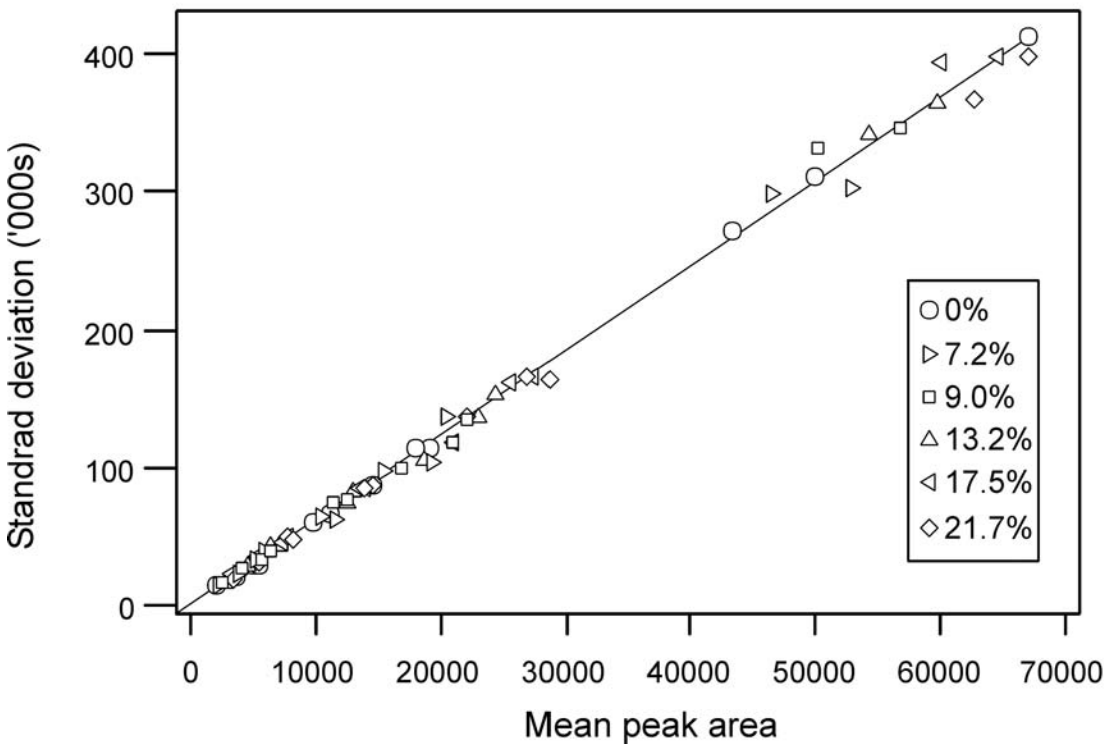


Table 3.12 PCA results from each normalisation method for real and simulated data based on a regular unleaded gasoline sample at the following evaporation levels: 0%, 7.2%, 9.0%, 13.2%, 17.5%, and 21.7% by weight.

	Eigenvalue for each PC					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	10.944	9.6884	9.6308	10.944	8.5942	8.6002
2	0.037	0.1418	0.7748	0.037	0.7509	0.7429
3	0.005	0.0500	0.1684	0.005	0.5167	0.5195
4	0.003	0.0288	0.1312	0.003	0.4171	0.4154
5	0.002	0.0213	0.0750	0.002	0.2427	0.2435

	Cumulative proportion of variance accounted for by each PC (%)					
PC	Raw data	MS	CS	Autoscale	Weight Factor	LN
1	99.5	96.9	87.6	99.5	78.1	78.2
2	99.8	98.3	94.6	99.8	85.0	84.9
3	99.9	98.8	96.1	99.9	89.7	89.7
4	99.9	99.1	97.3	99.9	93.4	93.4
5	99.9	99.3	98.0	99.9	95.7	95.7

The PC score plots for the LN transformed data (Figure 3.18) show not only the chemical similarity between the different evaporation levels, but also demonstrate the sensitivity of PCA in detecting small chemical differences due to evaporation. It can be seen from the 95% error ellipse envelopes that 0% and 9.0% levels can be differentiated, whilst the 7.2% evaporation level overlaps both. The 13.2% evaporation level is different from all the others. The 17.5% and 21.7% evaporation levels are very similar chemically. There is a progression from right to left across the plot as the amount of evaporation in each sample increases. This indicates a progressive compositional change in the monoaromatic profile from one evaporation level to the next.

Figure 3.17 Effects of evaporation on real and simulated data. Plots of first two PC scores after the following transformations: MS (a); and CS (b). Real data is denoted by solid markers whilst simulated data is denoted by the corresponding unfilled markers. 95% error ellipses are shown.

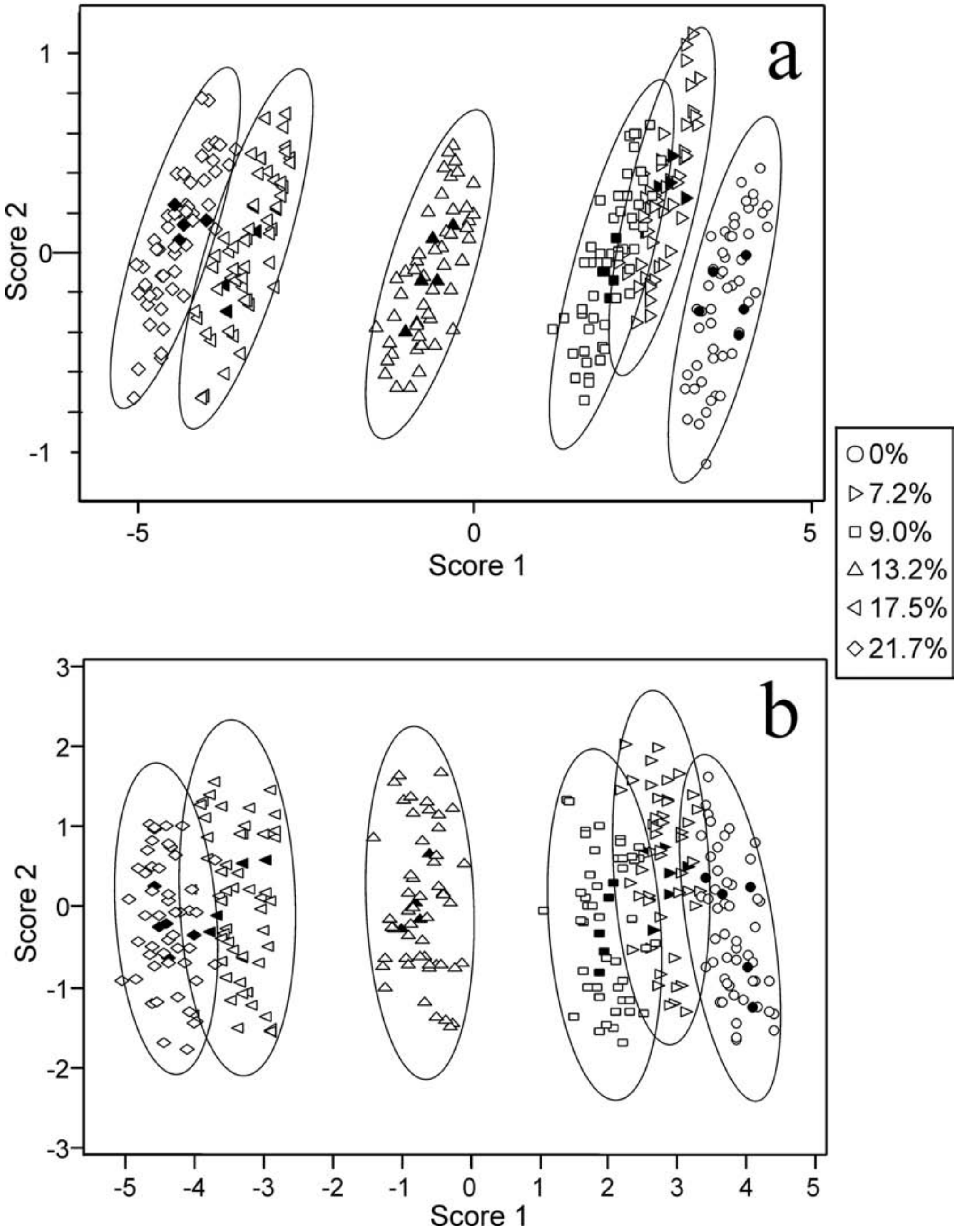
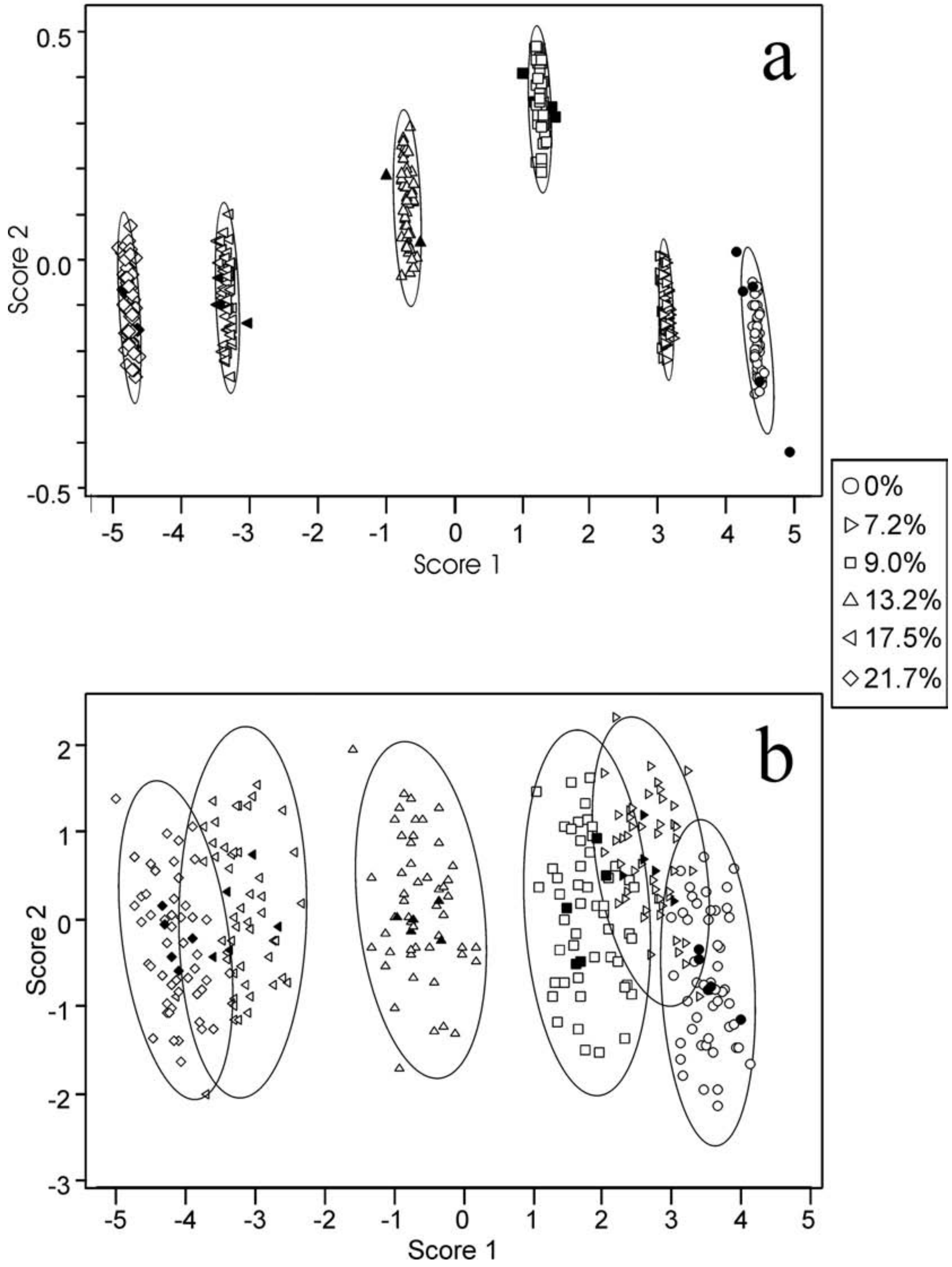


Figure 3.18 Effects of evaporation on real and simulated data. Plot of first two PC scores after the following transformations: Autoscale (a); and LN (b). Real data is denoted by solid markers whilst simulated data is denoted by the corresponding unfilled markers. 95% error ellipses are shown.



3.7 Conclusions

It is important to apply a normalisation method that will not propagate error in the data set,[8] will not introduce spurious correlations, and will not distort the data[11] prior to principal component analysis (PCA). The MS transformation was found by Rietjens,[8] by Deming *et al.*,[11] and the simulation performed in this chapter, to generate errors in the data set and affect the PCA results. The CS transformation has been shown to introduce negative correlations between variables.[15,16] Neither MS nor CS transformations can be considered robust enough to use to normalise any data set containing a heteroscedastic noise pattern. Application of the Autoscale normalisation method to simulated and to real data demonstrated that this method is susceptible to differences in dilution between samples. From the results reported in this chapter using simulated data, and the work of Rietjens,[8] it was found that the most robust normalisation methods are the Weight Factor method and LN method. These two methods give nearly identical PCA results. The fact that the analyst must choose the size of the weighting factor in the Weight Factor method introduces an element of subjectivity. The LN normalisation method, on the other hand, is as robust at the Weight Factor method, does not require any additional subjective input from the analyst, and is more simple computationally. Selection of higher boiling compounds may remove the effect that evaporation has on the PCA results, and suggests an approach that may allow different gasoline samples to be distinguished.

3.8 References

1. Lavine, B.K. Chemometrics. Analytical Chemistry, 1998; 70: 209R-228R.

2. Wold, S. and Sjostrom, M. Chemometrics, present and future success. Chemometrics & Intelligent Laboratory Systems, 1998; 44(1-2): 3-14.

3. Lavine, B.R. Clustering and classification of analytical data. In: Encyclopedia of Analytical Chemistry. Myers, R.A., Editor; v. 11, Wiley and Sons: Chichester, UK; 2000.

4. Jolliffe, I.T. Principal Component Analysis; Springer Series in Statistics, Springer Verlag, New York, 1986.

5. Tan, B.J.; Hardy, J.K.; Snavely, R.E. Accelerant classification by gas chromatography/mass spectrometry and multivariate pattern recognition. Analytica Chimica Acta, 2000; 422(1): 37-46.

6. American Society for Testing and Materials. E1618-97 Standard Guide for Identification of Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry. Philadelphia, USA.

7. Andrews, J.M. and Lieberman, S.H. Neural network approach to qualitative identification of fuels and oils from laser induced fluorescence spectra. Analytica Chimica Acta, 1994; 285: 237-246.

8. Rietjens, M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytical Chemistry, 1995; 316: 205-215.

9. Kvalheim, O.M.; Brakstad, F.; Laing, Y-Z. Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. Analytical Chemistry, 1994; 66: 43-51.

10. Toft, J. and Kvalheim, O.M. Eigenstructure tracking analysis for revealing noise pattern and local rank in instrumental profiles: application to transmittance and absorbance IR spectroscopy. Chemometrics and Intelligent Laboratory Systems, 1993; 2: 79-91.

11. Deming, S.N.; Palasota, J.A.; Nocerino, J.M. The geometry of multivariate object preprocessing. Journal of Chemometrics, 1993; 7: 393-425.

12. Johansson, E.; Wold, S.; Sjodin, K. Minimizing effects of closure on analytical data. Analytical Chemistry, 1984; 56: 1685-1688.

13. Wilcken, H. and Schulten, H.R. Differentiation of resin-modified paints by pyrolysis-gas chromatography mass spectrometry and principal component analysis. Fresenius' Journal of Analytical Chemistry, 1996; 355(2): 157-163.

14. Wold, S.; Albano, C.; Dunn, W.J. III; Edlund, U.; Esbensen, K.; Geladi, P.; Hellberg, S.; Johansson, E.; Lindberg, W.; Sjostrom, M. Multivariate data analysis in chemistry. In: Chemometrics. Mathematics and Statistics in Chemistry. Kowalski, B.R., Editor. NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences, Volume 138. D. Reidel Publishing Company, Dordrecht, Holland; 1984.

15. Chayes, F. On correlation between variables of constant sum. Journal of Geophysical Research, 1960; 65(12): 4185-4193.

16. Skala, W. A mathematical model to investigate distortions of correlation coefficients in closed arrays. Journal of the International Association for Mathematical Geology, 1977; 9(5): 519-528.

17. Kelly, J.L.; Barlow, C.H.; Jungui, T.M.; Callis, J.B. Prediction of gasoline octane numbers from near-infrared spectral features in the range 660-1215 nm. Analytical Chemistry, 1989; 61: 313-320.

18. American Society for Testing and Materials. D 2699-01 Standard Test Method for Research Octane Number of Spark-Ignition Engine Fuel. Philadelphia, USA.

19. Brown, S.D. Chemical systems under indirect observation: latent properties and chemometrics. Applied Spectroscopy, 1995; 49(12):14A-31A.

20. Wold, S.; Sjostrom, M.; Carlson, R.; Lundstedt, T.; Hellberg, S.; Skagerberg, B.; Wikstrom, C.; Ohman, J. Multivariate design. Analytica Chimica Acta, 1986; 191: 17-32.

21. Wold, S.; Albano, C.; Dunn, W.J.; Esbensen, K.; Hellberg, S.; Johansson, E.; Sjostrom, M. Pattern recognition; finding and using regularities in multivariate data. In: Food Research and Data Analysis. Martens, H. and Russwurm, H. Jr., Editors. Elsevier Science; 1983.

22. Kaiser, H.F. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 1960; 20: 141-151.

23. Cattell, R.B. The scree test for the number of factors. Multivariate Behaviour Research, 1966; 1: 245-276.

24. Jackson, J.E. Principal components and factor analysis: part I. Journal of Quality Technology, 1980; 12(4): 201-213.

25. Lindsay, W.S. Bivariate control chart: uses Hotelling's T in combination with principal components analysis. January 1, 1984. <http://www.minitab.com/support/macros/index.asp?cat=QC_DOE#17> Accessed May 8, 2002.

26. Reyment, R.A. Multivariate analysis in geoscience: fads, fallacies, and the future. Chemometrics and Intelligent Laboratory Systems, 1987; 2: 79-91.

Chapter 4. Sampling protocol

4.1 Introduction

When a complex mixture of hydrocarbons evaporates, the low boiling more volatile compounds are lost first. It is expected that when a gasoline pump at a retail station remains unused for a period of time there will be some loss of the low boiling compounds from the gasoline held in the dispensing nozzle and hose leading from the pump. Previous studies that involved the collection of gasoline samples from retail stations,[1,2,3,4,5,6,7] or refineries,[8,9] did not state whether or not the dispensing pump or tap was purged prior to sample collection. Kanai *et al.*[10] described their sampling protocol in detail. They purged the retail pump with five U.S. gallons (18.9 litres) of fuel prior to sample collection. Samples were also collected directly from the top, middle and bottom of the underground storage tanks at retail stations. They found no differences between samples obtained from the top, middle and bottom of the storage tanks and concluded, therefore, that samples taken directly from a purged pump were representative of the gasoline held in the storage tanks. No justification was given for the purge volume used in their study.[10] The need to purge a retail station pump prior to sample collection, and the volume of that purge, if necessary, has not been addressed in the scientific literature.

American Society for Testing and Materials (ASTM) method D5842[11] and Interntional Organisation for Standardisation (ISO) method 3170[12] provide guidelines on how to obtain a representative sample from a gasoline pump at a retail station. The ASTM method provides a more detailed procedure than the ISO method, therefore, the ASTM method will be discussed. The scope of the ASTM method is sample collection and handling as it relates to volatility measurements (*e.g.* vapour pressure) of light fuels such as gasoline. The method recommends that fuel be collected in an appropriate glass (preferably amber glass), plastic or metal container. In order to collect a sample from a retail station pump the nozzle must first be purged by dispensing a quantity of fuel from the pump, the pump is then reset and a sample is collected via a reusable, metal (non-ferrous) nozzle extension. The purge volume is not specified in the method. The recommended use of a reusable nozzle extension opens up the possibility of contamination between samples, an issue that would not likely be a concern when collecting a sample for volatility measurements but is imperative to control for when collecting samples for chemical comparison.

It is important to ascertain whether or not a significant loss of volatile hydrocarbons occurs from an idle retail station pump because such a loss may lead to significant differences between consecutive samples from the same pump. If significant evaporation is found to occur between consecutive samples from a pump, then a sampling protocol will have to take this into account. For example, it may be found that all samples must be taken immediately after a specified amount of gasoline has been passed through the pump. However, if no significant evaporation is found to occur between consecutive samples then it will not be necessary to purge the pump prior to sample collection.

4.2 Materials and methods

4.2.1 Sample collection

In order to determine whether differences exist between consecutive samples, three samples were collected from a retail station's pump dispensing regular unleaded gasoline and three samples were collected from a pump dispensing premium unleaded gasoline. The retail station used in this study was closed during the night from 19:30 to 07:30 hours the next day, therefore, both pumps had been idle for a minimum of 12 hours. For each sample, approximately 100 mL of gasoline was collected in a 120 mL amber-coloured Boston round glass bottle (Wheaton, USA. Dimensions – bottle: 48 mm x 112 mm (diameter x height); finish: 22 mm o.d. (includes threads), 15 mm i.d., 8 mm (height)) sealed with a Teflon-lined polyethylene screw cap (Redi-Cap, USA). The gasoline sample were delivered from the nozzle into the glass bottle via a disposable plastic funnel made by cutting the bases from 45 mL Nylex (high density polyethylene) plastic bottles designed for pharmaceuticals (The Bottle People, Sydney, Australia). The plastic bottles used were oval, (45 mm x 24 mm (base) x 66 mm (height)) with finish dimensions of 14 mm o.d. (including threads) and 11 mm i.d.. Glass bottles and funnels were selected at random and checked for the presence of contaminants by washing each one to two millilitres dichloromethane (HPLC grade, BDH). The wash solvent was collected and checked by gas chromatography-mass spectrometry for possible contaminants. No contaminants that could potentially interfere with this study were found.

The number of samples and the amount of fuel purged between each sample is given in Table 4.1. The first sample (Sample 1) was collected immediately after the pump was unlocked by the attendant, before any customers had purchased fuel. The pump was then purged (Purge 1) by dispensing gasoline into a five litre fuel container until a total of 1 litre had been dispensed from the pump. The second sample (Sample 2) was then collected. The pump was purged a second time (Purge 2) until a total of 4.8 litres had been dispensed. A third sample was then taken (Sample 3). Three samples were collected from a pump dispensing regular unleaded gasoline and three samples were collected from a pump dispensing premium unleaded gasoline for a total of six samples. The purged gasoline was not analysed.

Table 4.1 Retail service station pump sampling regime.

	Gasoline dispensed (L)	Cumulative total gasoline dispensed (L)	
Pump unlocked	0.0	0.0	
Sample 1	0.1	0.1	sample for analysis
Purge 1	0.9	1.0	
Sample 2	0.1	1.1	sample for analysis
Purge 2	3.7	4.8	
Sample 3	0.1	4.9	sample for analysis

Approximately five litres was chosen as the final volume to dispense from the pump because it is approximately five times the volume held by a typical pump hose. The volume held by a typical pump hose is based on a standard hose inside diameter of 1.90 cm (0.75 inches) and a hose length of 3.65 m (12 feet) (radius, r = 0.95 cm, length is 365 cm, Volume = length x πr^2 = 365 x 3.14 x 0.95^2 = 1034 cm^3). Flexible gasoline pump hoses do come in a range of sizes, the most common having volumes ranging from 694 cm^3 (1.90 cm i.d. x 245 cm (0.75 inch i.d. x 8 feet)) up to 2993 cm^3 (2.5 cm i.d. x 610 cm (1 inch i.d. x 20 feet)).[13]

4.2.2 Instrument conditions

A 10 µL aliquot of each sample was diluted in methylene chloride (1:100, v/v) and 1.0 µL was injected on an Agilent 6890 gas chromatograph fitted with a 5973 mass selective detector. The components were separated on a HP-5ms fused silica capillary column (30 m x 0.32 mm i.d. x 0.25 µm film thickness; Agilent). Helium was used as the carrier gas at a constant flow rate of 1.2 mL/min. The inlet temperature was held at 280 °C and the sample was split with a ratio of 1:15. The oven temperature program was 35 °C (initial hold 3 minutes), ramped at 2 °C/min to 43 °C, then ramped at 50 °C/min to 290 °C (final hold 15 minutes). Since the greatest changes in the concentration of monoaromatics is expected to occur for the most volatile compounds, that is, benzene and toluene, the temperature program was designed to give the best chromatography of these two compounds. The remaining monoaromatics were eluted using a faster temperature ramp rate, however, resolution of these compounds from one another was maintained. Mass spectral data was collected from 30 to 350 daltons at a rate of 4.4 scans/second. A total of five aliquots were prepared and analysed for each sample. Selected aromatic ions (m/z^+ 78, 91, 92, 105, 106, 119, 120, 133 and 134) were extracted from each chromatogram using a macro (Appendix 3.1) and peak areas integrated. The resulting integration results were written to an ASCII text file in a format suitable for export to a spreadsheet.

4.2.3 Data analysis

Integrated gas chromatographic data was loaded into a spreadsheet (Microsoft Excel®) and normalised using the LN transformation.[14] Data analysis by principal components was performed using Minitab® (version 13.1 for Windows) loaded on a personal computer. Error ellipses were calculated using the method described by Jackson.[15] Error ellipse calculations were performed using a Minitab macro modified for this work to allow user input of the F-statistic.[16] The macro is provided in Appendix 3.2.

4.3 Results and discussion

When a complex mixture of hydrocarbons such as gasoline evaporates, the more volatile compounds will be lost first. On a non-polar (*e.g.* DB-1), or a slightly polar (*e.g.* DB-5 or HP-5ms), fused silica capillary column, the hydrocarbon species typically elute from the lowest boiling point to the highest. The evaporation of the more volatile compounds will be reflected as a greater loss of the earlier eluting compounds relative to the later eluting, less volatile compounds in the chromatogram.

In order to compare consecutive samples from a retail service station pump, it was decided to monitor the loss of the monoaromatic species present in the fuel and to compare the loss using principal component analysis (PCA). The monoaromatics selected for this purpose span a wide boiling range from benzene to 1,2,3-trimethylbenzene (Table 4.2).

Table 4.2 Selected aromatic peaks from extracted ion chromatograms.

Peak Number	Retention Time	Assignment[1]	Boiling Point[2]
1	2.64	benzene	80.1
2	4.46	toluene	110.6
3	7.03	ethyl benzene	136.0
4	7.27	p+m-xylene	138.3 + 139.1
5	7.93	o-xylene	144.4
6	9.50	n-propylbenzene	159.2
7	9.69+	3-ethyl-1-methylbenzene +	161.3
	9.72	4-ethyl-1-methylbenzene	162.0
8	9.84	1,3,5-trimethylbenzene	164.7
9	10.10	2-ethyl-1-methylbenzene	165.2
10	10.40	1,2,4-trimethyl benzene	169.3
11	11.01	1,2,3-trimethylbenzene	176.1

1. Johansen, N.G.; Ettre, L.S.; Miller, R.L. Quantitative analysis of hydrocarbons by structural group type in gasolines and distillates. Journal of Chromatography, 1983; 256(3): 393-417.
2. Weast, R.C. and Graselli, J.G. CRC Handbook of Data on Organic Compounds, 2nd Edition. CRC Press, Boca Raton, USA; 1989.

The results of the PC analysis of the regular and premium gasoline samples are given in Table 4.3 and Figure 4.1.

Table 4.3 PCA results for the regular unleaded and premium unleaded gasoline samples taken from idle pumps. Variance refers to the cumulative variance in the data set accounted for by each principal component.

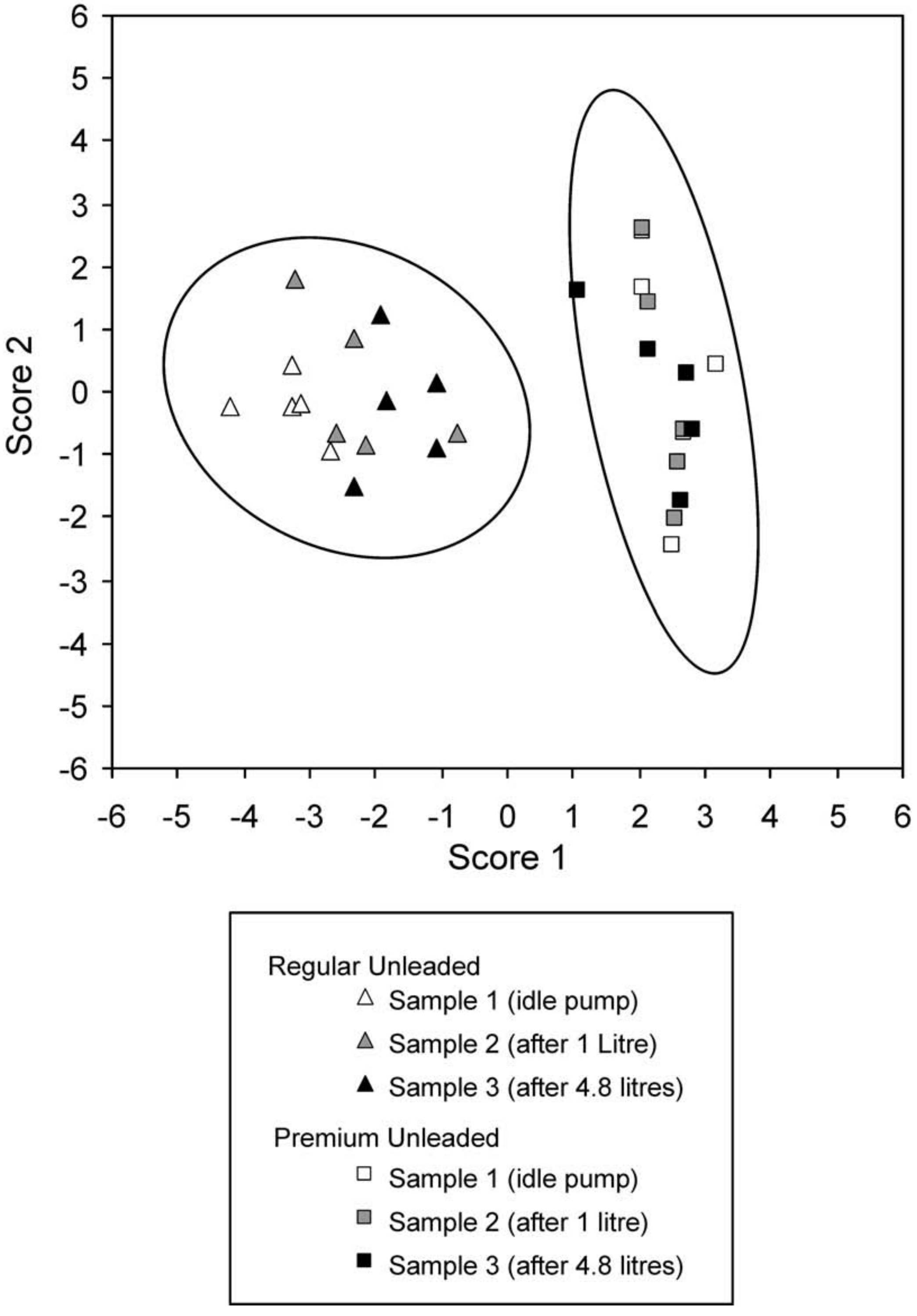
PCA Results		
PC	Eigenvalue	Variance (%)
1	6.5125	59.2
2	1.6573	74.3
3	1.3270	86.3
4	0.6429	92.2
5	0.3297	95.2

In Figure 4.1 it can be seen that no difference was detected in the data set between the three samples of premium unleaded gasoline. For the samples collected from the regular unleaded gasoline pump there appears to be a slight separation of the samples from the idle pump (Sample 1) compared with the samples taken after 1 and 4.8 litres of fuel were dispensed from this pump (Samples 2 and 3). However, three of the data points, two from Sample 2 and one from Sample 3, are so close to the cluster of Sample 1 data points (idle pump) that there is no significant difference between the fuel samples dispensed before and after purging the regular unleaded gasoline pump.

4.4 Sampling protocol

Full details of the sample collection protocol at a retail service station is given in Appendix 4.1. The salient details of the protocol are as follows: use of a clean, new amber Boston round bottle and a clean, new disposable plastic funnel for each sample collected; recording of all sample information on the data collection sheet (Appendix 4.1); transporting of all samples in accordance with the Australian code for the transport of dangerous goods; and, storage of all samples in approved flammable liquid storage cabinets.

Figure 4.1 PCA score plot of selected aromatics in regular and premium unleaded gasoline samples taken from a service station pump that had been idle for a minimum of 12 hours, after 1 litre of gasoline was pumped, and after 4.8 litres of gasoline were pumped. 95% error ellipses shown.



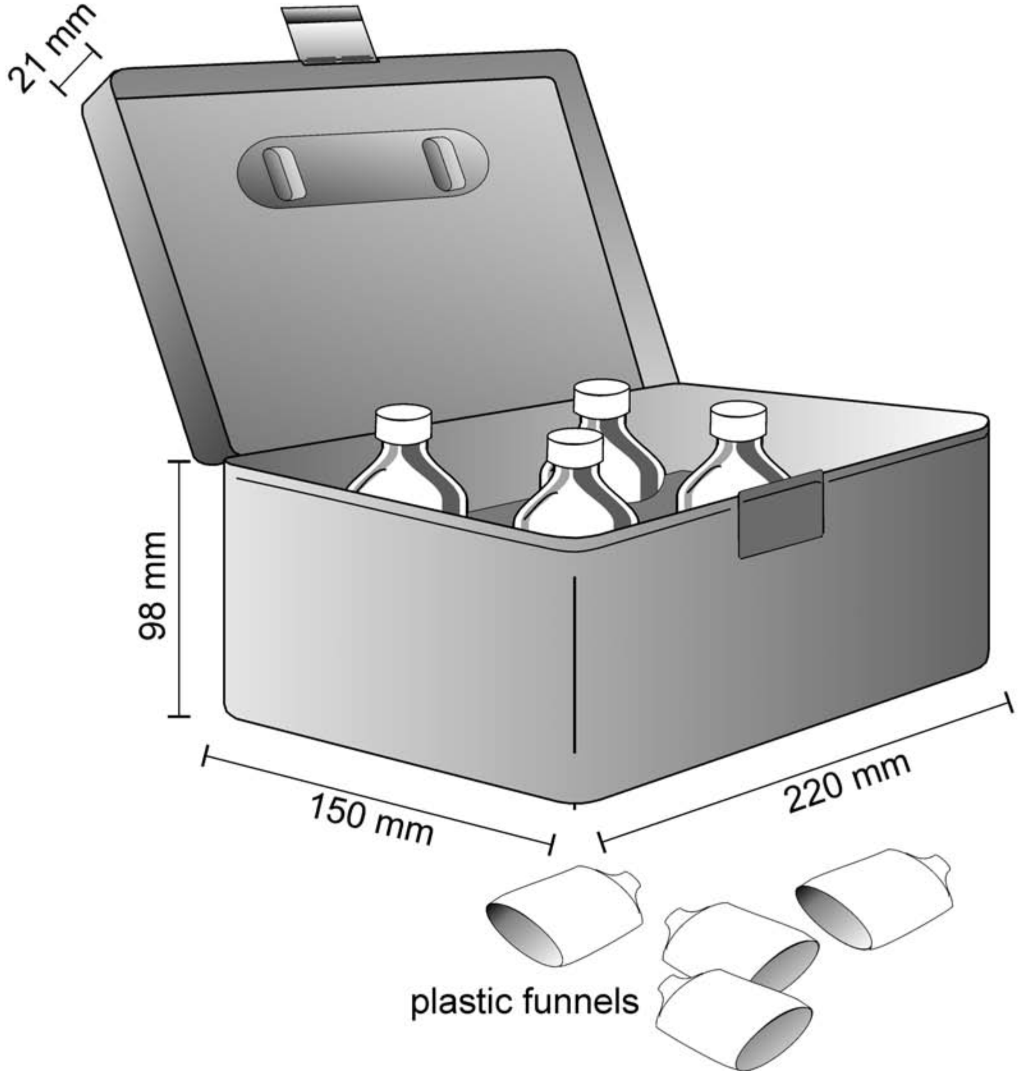
4.4.1 Safety standards

All gasoline samples were collected in accordance with internationally accepted guidelines and standards, transported in compliance with the Australian dangerous goods code, and stored appropriately. The glass containers (amber Boston round bottles) and closures (teflon-lined polypropylene caps) conform to two internationally accepted standards, ASTM D 5842-95, *Standard Practise for Sampling and Handling of Fuels for Volatility Measurement*[11] and ISO 3170:1988, *Petroleum Liquids Manual Sampling*.[12] Samples were transported in accordance with the Australian code for the transport of dangerous goods (ADG code) Section 1.2.1, Division 1.2 (*Consumer Commodity Loads*), and Section 3.8.3, Division 3.8 (*Special Provision for Packaging of Particular Classes of Dangerous Goods*).[17]

The ADG code requires that the packaging for Class 3 (flammable liquid) dangerous goods must consist of an inner package that is transported in an approved outer combination package. The inner package was an amber Boston round bottle described above. The outer packaging was fashioned from a small polyethylene lunch box (Sistema, New Zealand) fitted with a high density polyurethane foam liner cut to accept four sample bottles. The plastic box had a clasp to secure the lid and a carrying handle on the top of the lid. The packaging is illustrated in Figure 4.2. All samples were stored at the University of Technology, Sydney in an approved flammable liquid storage cabinet located in a locked room with limited access rights.

The nozzles used to deliver gasoline from service station pumps had a diameter that was larger than the diameter of the opening of the sample collection bottles used, therefore, it was necessary to use a funnel to avoid spilling gasoline while collecting samples. The funnel had to fulfil two requirements: it had to be small enough to fit into the neck of the sample collection bottle; and, in order to avoid cross-contamination between samples, the funnel had to be relatively inexpensive so that it could be discarded after the sample was collected. Funnels were made by cutting the bases from 45 mL Nylex (high density polyethylene) plastic bottles designed for pharmaceuticals (The Bottle People, Sydney,

Figure 4.2 Illustration of polyethylene plastic box with high density polyurethane foam liner used in the collection of gasoline samples from service stations. Inner packaging (glass bottles) and outer packaging (plastic box with liner) conforms with the Australian code for the Transport of dangerous goods. Polyethylene bottles, with their bases cut off, were used as disposable funnels and are shown in the foreground.



Australia). The plastic bottles used were oval, (45 mm x 24 mm (base) x 66 mm (height)) with finish dimensions of 14 mm o.d. (including threads) and 11 mm i.d.

4.4.2 Record keeping

Appendix 4.1 contains the two page form completed for each sample collected. The information described in Table 4.4 was recorded for each sample. Note that fuel distributor information was obtained on a strictly voluntary basis and experience showed that many station attendants either did not know the delivery information or were reluctant to provide it without approval from the station manager. When information was provided verbally it was difficult to verify its accuracy, except in rare instances when stations mangers were present and willingly provided delivery dockets for inspection. As a result of these

Table 4.4 Information recorded for each gasoline sample collected.

Reference number (a unique number assigned to each sample and written on the sample bottle and cap)

Collection details
- collection date
- collection time
- name of person who collected sample
- volume of fuel pumped immediately prior to sample collection (if known)

Petrol station information
- name of station (supplier/brand recorded for independent retailers)
- street address
- suburb
- post code

Petrol sample (unleaded only)
- type (regular, premium, lead replacement)

Fuel distributor information
- date of last delivery
- last volume delivered
- frequency of deliveries

difficulties, Caltex Australia was approached for delivery schedule information from their Banksmeadow distribution terminal. Caltex Australia agreed to provide this information which is discussed in detail in Chapter 2.

4.5 Conclusions

The change in the monoaromatic composition of gasoline from a service station pump which has been idle for a minimum of twelve hours, has been examined. Two different pumps were studied, one dispensing regular unleaded gasoline and the other dispensing premium unleaded gasoline. It was found that for both types of fuel the monoaromatic content of the first 0.1 litre of fuel collected from the idle pump (Sample 1) was not significantly different from the next two samples (Samples 2 and 3) that were taken after purging the pump with 1.0 litre and 4.8 litres of fuel, respectively. This result demonstrated that, under the conditions of the study, significant evaporation did not occur from the nozzle and hose of an idle pump. Therefore, it is not necessary to purge a service station pump prior to collecting a gasoline sample.

4.6 References

1. Chan, L. The determination of tetraalkyl lead compounds in petrol using combined gas chromatography atomic absorption spectrometry. Forensic Science International, 1981; 18: 57-62.

2. Bertsch, W.; Mayfield, H.; Thomason, M.M. Application of pattern recognition to high resolution GC and GC/MS. Part I. Basic studies. Proceedings of the International Symposium on Capillary Chromatography, 1981: 313-334.

3. Frank, H.A. Lead alkyl components as discriminating factors in the comparison of gasolines. Journal of Forensic Sciences, 1983; 28: 330-334.

4. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography I: Comparison methodology. Journal of Forensic Sciences, 1987; 32(3): 606-615.

5. Potter, T.L. Fingerprinting petroleum products: unleaded gasoline. Petroleum Contaminated Soils: Volume 3. Lewis Publishers, Chelsea, MI; 1990.

6. Ichikawa, M.; Nonaka, N.; Takada, I. and Ishimori, S. Mass spectrometric analysis for distinction between regular and premium motor gasolines. Analytical Science, 1993: 9; 261-266.

7. Sheff, L.M. and Siegel, J.A. Fluorescence of petroleum products V: three-dimensional fluorescence and capillary gas chromatography of neat and evaporated gasoline samples. Journal of Forensic Sciences, 1994: 39(5); 1201-1214.

8. Thatcher, P.J. The identification of petroleum residues in arsons. Thesis (Ph.D.) University of Melbourne, 1982.

9. Hirz, R. Gasoline brand identification and individualization of gasoline lots. Journal of the Forensic Science Society, 1989; 29(2): 91-101.

10. Kanai, H.; Yazawa, L.; Maka, J. Baseline data on the quality of gasolines in Hawaii and their implication in preparing alcohol blends. Hawaii State Department of Agriculture, Honolulu, HI; 1989.

11. American Society for Testing and Materials. D5842-95(2000), Standard Practice for Sampling and Handling of Fuels for Volatility Measurement. Philadelphia, USA.

12. International Organisation for Standardisation. ISO 3170:1988, Petroleum Liquids Manual Sampling. Switzerland; 1988.

13. DurEquip – Gas Pump Equipment Specialists. 2000/2001 Reference Guide and Parts Catalogue. Minneapolis, MN, USA; 2000.

14. Rietjens, M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytical Chemistry, 1995; 316: 205-215.

15. Jackson, J.E. Principal components and factor analysis: part I. Journal of Quality Technology, 1980; 12(4): 201-213.

16. Lindsay, W.S. Bivariate control chart: uses Hotelling's T in combination with principal components analysis. January 1, 1984. Accessed May 8, 2002. http://www.minitab.com/support/macros/index.asp?cat=QC_DOE#17>

17. Australian dangerous goods code: Australian code for the transport of dangerous goods by road and rail. Commonwealth of Australia, Federal Office of Road Safety. Canberra, ACT; 1998.

Chapter 5. Trace compounds in gasoline

5.1 Introduction

The forensic characterisation of trace compounds in gasoline has been the subject of several studies.[1,2,3,4,5] During the 1980s, several authors reported the forensic comparison of gasoline based on trace levels of alkyl lead additives.[1,2,3] The gasoline samples examined by Frank[1] had alkyl lead concentrations between 0.3 and 0.4 mg/L (0.3 – 0.4 ppm). At this time the average lead content of gasoline in the USA was between 0.5 and 0.8 g/US gallon (130 – 210 ppm),[6] while the lead concentration allowed in New South Wales, Australia was between 0.4 and 0.65 g/L (400 – 650 ppm).[2] Frank,[1] Chan,[2] and Hirz[3] each found that the different alkyl lead ratios present in gasoline allowed them to identify, at worst, the source refinery and, at best, an individual service station. It was also during the 1980s that two other researchers found that the relative concentrations of certain polycyclic aromatic hydrocarbons could be used to differentiate gasoline samples produced by different refineries.[4,5] The introduction of trace organic compounds during the refining and blending process in a refinery may introduce significant differences between batches of finished fuel. Two generic groups of trace organic compounds, polar compounds and polycyclic aromatic hydrocarbons (PAHs), were assessed for their potential to identify batch to batch variations in gasoline. It was also thought that the lower volatility of these compounds may allow them to be used in the comparison of evaporated gasoline samples.

One goal of the work described in this chapter was for the method(s) developed here to be directly transferable to a forensic laboratory. Gas chromatography is a common method used to separate and analyse complex organic mixtures. Gas chromatography is used extensively in the petroleum industry to obtain detailed information from both crude oils and refined products.[7,8,9] The use of gas chromatography for the forensic analysis of petroleum products and fire debris extracts is also well known.[10,11] Therefore, it was an aim of this thesis to use modern gas chromatographic techniques to characterise trace compounds present in gasoline.

5.1.1 Trace polar compounds in crude oils and petroleum products

Interest in detecting polar compounds in gasoline arose during the 1980s and lead to a number of articles being published on this topic. Using a liquid-liquid extraction (partitioning) method, Colgrove and Svec[12] attempted to extract polar compounds from gasoline prior to analysis by gas chromatography. They did not find any significant amounts of acids, phenols, amines, aldehydes, or ketones, however, their polar fraction did contain the alkylated homologues of benzene, indane, tetralin, and naphthalene. A similar pre-separation technique was used by Di Sanzo[13] to extract small concentrations of basic nitrogen compounds from gasoline. Dinh *et al.*[14] also developed a similar liquid extraction method and examined the nitrogen compound distribution of three middle distillate (unrefined) samples from different locations around the world. They found that pyridine compounds were present in large quantities relative to other nitrogen compounds.[14]

Youngless *et al.*[15] were able to improve the separation of trace polar compounds from gasoline using a silica gel column. They used gas chromatography to identify 62 polar compounds in gasoline, including alcohols, amines, phenols and pyridines. No derivatisation of the phenols was performed prior to injection of the extract onto the gas chromatograph. They found that the relative amounts of certain polar compounds (*e.g.* alkylphenols) differed between three different gasoline samples. Kanai *et al.*[16,17] used the method of Youngless *et al.*[15] to make simple comparisons of the polar compound content between different gasoline samples. Three different 87 octane unleaded gasoline samples were each found to contain a different combination of polar compounds, including different combinations of alkyl phenol compounds.[16] In one study, Kanai *et al.*[17] compared the polar compounds found in water that had been in contact with gasoline to the polar compounds originally present in the gasoline sample. They found that most polar compounds present in gasoline partitioned into the aqueous phase and could be recovered on a silica gel column. A total of thirty different polar compounds, including phenols, amines, and an indole, were extracted from the gasoline samples.[17]

A few articles have reported that lower molecular weight alkyl phenols (C_0- to C_4-) are present in varying relative concentrations in crude oil.[18,19,20] The solubility of alkyl phenols in seawater has been used to characterise marine oil spills.[21] It is proposed here that if polar compounds originally present in a crude oil survive the refining process, they may be suitable markers to differentiate between samples of gasoline.

5.1.2 *Trace polycyclic aromatic hydrocarbons (PAHs) in crude oils and petroleum products*

In 1977, Mach[22] reported the use of gas chromatography-mass spectrometry to examine polycyclic aromatic hydrocarbons (PAHs) in severely evaporated and burned samples of gasoline. Mach concluded that certain PAHs present in gasoline are unique to this product and may be used to determine the presence or absence of gasoline in a fire debris sample. Hennig[4] reported data for the higher boiling compounds found in three gasoline samples, each from a different refinery in Germany. Although not identified by Hennig, the higher boiling compounds in the published chromatograms appear to be the C_1- and C_2-naphthalenes, the relative amounts of which varied between refineries. Work done by Thatcher[5] on gasoline from three Australian refineries supported the results of Hennig. Thatcher found that in samples of regular and premium gasoline, the relative amounts of certain polycyclic aromatic hydrocarbons were dependent on which refinery made the gasoline. Thatcher found that, in particular, the relative ratios of naphthalene, 1,2-dimethylnaphthalene, and phenanthrene could be used to distinguish all six gasoline samples (three premium and three regular). The value of examining gasoline for the higher boiling compounds is that these compounds will remain even if the original sample has been evaporated, therefore, it may be possible to compare samples even after evaporation has occurred.

Extraction and detection of PAHs in refined petroleum products has been the subject of several recent articles.[23,24,25,26,27] Bundt *et al.*[23] developed a solid phase extraction (SPE) method to separate aliphatics, monoaromatics, diaromatics, polyaromatics and polar compounds from diesel fuel on a silica gel column. Song and Hatcher,[24] and later Lai and

Song,[25] reported a method for fractionating alkyl-substituted benzenes and polycyclic aromatic hydrocarbons from jet fuel on an alumina column. The use of PAHs to identify the source of a crude oil spill has been described by Wang and Fingas;[26] it is also the subject of a review.[27]

The relative stability of and, therefore, importance of using polycyclic aromatic hydrocarbon analytes to *differentiate* the source of a spilled crude oil or other heavy oil in the environment has been noted in the literature.[26,27] Alkylated PAHs such as naphthalene, phenanthrene, fluorene, chrysene, dibenzothiophene and their alkylated homologues can be used to fingerprint oils because they persist in the environment after the spill has occurred.[26] Radke *et al.*[28] examined the relative amounts of alkylated polycyclic aromatic hydrocarbons in six crude oil samples taken at different depths from five wells in the Handil oil field, Indonesia. They observed some differences in the relative ratios of the C_1-naphthalenes to the C_2-naphthalenes between different samples. If these differences were to persist through the refining process, then it may be possible to differentiate refined fuels based on the relative ratios of their polycyclic aromatic hydrocarbons.

Wang and Fingas[29] have noted that different crude oils have different distributions of PAHs and their alkylated homologues, including aromatic sulphur compounds such as the methyldibenzothiophenes. Aromatic sulphur compounds have received attention in crude oil comparisons because different unrefined oils tend to vary significantly in their sulphur content.[29,30] Differences in the relative amounts of thiophene, benzothiophene and their methylated homologues were found in four gasoline samples analysed by GC-AED (atomic emission detector).[31] No methyldibenzothiophenes were found in this study.[31] This study also reported that the *total* concentration of the highest boiling organic sulphur compounds detected in gasoline ranged from 0 to 126 ppm.[31] Di Sanzo *et al.*[32] used gas chromatography with sulphur chemiluminescence detection to analyse a large number of sulphur compounds in a gasoline stream. Coulombe[33] has described the detection of diphenyl disulfide and its homologues in evaporated gasoline and in fire debris extracts by full-scan GC-MS. These compounds were used as "chemical markers" to identify the presence or absence of gasoline in fire debris samples.[33] Chen and Lo[34] used flameless sulphur chemiluminescence detection to identify several thiophene and alkyldisulfide compounds in three gasoline samples (leaded, regular unleaded and premium unleaded) from the same service station. None of the gasoline samples contained diphenyl disulfide or any of the alkylbenzothiophenes.[34] Chen and Lo found that "the patterns of the sulphur compounds in the three samples are quite similar except that the concentration of dimethyl sulphide in the leaded gasoline is higher than the others."[34]

It is doubtful that aromatic sulphur compounds will be helpful in the characterisation of refined fuels because these compounds are intentionally being refined out of the finished petroleum product. In an effort to reduce motor vehicle emissions, different countries and regions around the world have different requirements for fuel quality. Part of this effort aims to decrease the allowable concentration of organic sulphur compounds in gasoline (Table 5.1). The World-Wide Fuel Charter is an attempt to harmonize fuel quality standards from around the world into four categories.[35]

The process of refining crude oil into gasoline not only intentionally removes certain undesirable materials such as the organo-sulphur compounds, but also removes most of the high molecular weight compounds. Refining processes significantly reduce the quantity of three-ring PAHs (*i.e.* fluorene and phenanthrene) and their alkylated homologues, and completely eliminate the four-ring PAHs (*i.e.* chrysene) from gasoline. It is, therefore, expected that the refining process will remove much of the chemical information that makes one crude oil different from another. Although finished petroleum products such as gasoline may retain trace amounts of two- and three-ring PAHs, the amount of information held by these compounds is expected to be reduced when compared with the crude oil from which it originated.

It was an aim of this chapter to evaluate the use of polar compounds and/or polycyclic aromatic hydrocarbons for characterising gasoline samples. To achieve this aim it was necessary to first isolate and identify these compounds using solid phase extraction. Once suitable target compounds were identified, the next step was to develop a method for the routine analysis of these compounds. The final step to achieve the goal was to use chemometric methods to explore the ability of this method to discriminate different samples of gasoline.

Table 5.1 Maximum sulphur concentration regulations for gasoline from selected industrialised countries and regions. World-Wide Fuel Charter shown for comparison.

Country	*Designation*	*Maximum sulphur (ppm)*	*Implementation date*
Australia[1]	-	150 - premium unleaded	1 January 2002
		500 - regular unleaded	1 January 2002
	-	150 – both grades	1 January 2005
Canada[2]	-	300	1 January 2004
	-	80	1 April 2005
Europe[3]	Euro II	150	1 January 2002
	Euro III	50	1 January 2005
	Euro IV	nil	-
USA – National[4]	-	120	1 January 2004
	-	30	1 January 2006
USA – California[5]	Phase 1	300	prior to 1 March 1996
	Phase 2	30	1 March 1996
	Phase 3	15	31 December 2002
World-Wide Fuel Charter[6]	Category 1	1000	-
	Category 2	200	-
	Category 3	30	-
	Category 4	nil	-

1. Fuel Standard (Petrol) Determination 2001. As amended made under section 21 of the Fuel Quality Standards Act 2000, No. 153. Commonwealth of Australia; 2000.
2. Government of Canada. Sulphur in Gasoline Regulations. Canada Gazette, Part II; 1999; 133(13):1464-1544 (June 23).
3. The Directive on the Quality of Petrol and Diesel Fuels. Official Journal of the European Communities L 350, 1998 December 28; 41:58–68. Directive 98/70/EC of the European Parliament and of the Council of 13 October 1998 relating to the quality of petrol and diesel fuels and amending Council Directive 93/12/EEC.
4. National Archives and Records Administration, Government of the United States of America. Control of air pollution from new motor vehicles: Tier 2 motor vehicle emissions standards and gasoline sulphur control requirements; final rule. Federal Register, 2000; 65(28): 6698-6870 (February 10).
5. California Code of Regulations: Title 13. Motor vehicles, Division 3. Air resources board; Chapter 5. Standards for motor vehicle fuels; Article 1. Standards for gasoline; Subarticle 1.
Gasoline standards that became applicable before 1996; Section 2252. Sulfur content of gasoline represented as unleaded sold before April 1, 1996; and, Subarticle 2. Standards for gasoline sold beginning March 1, 1996; Section 2262. The California reformulated gasoline Phase 2 and Phase 3 standards.
6. World-Wide Fuel Charter. Alliance of Automobile Manufacturers, Washington, DC; April, 2000.

5.2 Materials and methods

5.2.1 Samples

Thirty-five samples of gasoline were obtained from 24 service stations in metropolitan Sydney, Australia (Table 5.2) over a seven-month period (March to September, 2001). The majority of the samples (29) were collected over a period of seven weeks (March to April, 2001). Samples were collected in 125 mL amber, Boston round glass bottles, sealed with polypropylene caps fitted with teflon-faced foamed polyethylene liners (Wheaton) and stored in the dark at room temperature.

Table 5.2 Summary of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. Shaded and unshaded blocks indicate samples that were collected from the same service station on the same day. Sample numbers 1 through 6 are from the same service station.

No.	Sample	Date[†]	Brand	Grade[‡]
1	RU02	13.03	Shell	Regular
2	LR06	13.03	Shell	Lead rep
3	PU07	13.03	Shell	Premium
4	RU11	29.03	Shell	Regular
5	PU12	29.03	Shell	Premium
6	RU13	05.04	Shell	Regular
7	RU14	18.04	Caltex	Regular
8	RU15	18.04	Mobil	Regular
9	RU16	18.04	Caltex	Regular
10	PU17	18.04	Caltex	Premium
11	RU18	18.04	Shell	Regular
12	PU19	24.04	BP	Premium
13	RU20	24.04	BP	Regular
14	RU21	24.04	Caltex	Regular
15	RU22	24.04	Shell	Regular
16	RU23	24.04	BP	Regular
17	RU24	24.04	Mobil	Regular
18	RU25	24.04	Mobil	Regular
19	RU26	24.04	Caltex	Regular
20	RU27	24.04	Shell	Regular
21	PU28	24.04	Shell	Premium
22	RU29	24.04	BP	Regular
23	RU30	24.04	BP	Regular
24	PU31	24.04	Shell	Premium
25	RU32	24.04	Shell	Regular
26	RU33	24.04	Mobil	Regular
27	RU34	24.04	Mobil	Regular
28	RU35	24.04	Mobil	Regular
29	RU36	28.04	BP	Regular
30	PU37	14.08	BP	Premium
31	RU38	14.08	BP	Regular
32	PU39	12.09	Burmah	Premium
33	RU40	12.09	Burmah	Regular
34	PU41	12.09	Woolies	Premium
35	RU42	12.09	Woolies	Regular

[†] Date is expressed as day and month (dd.mm)

[‡] The three grades of unleaded gasoline are: regular, premium and lead replacement

5.2.2 Solid phase extraction (SPE)

Reagent grade petroleum spirit (bp 30–40 ^{o}C, Univar), spectrophotometric grade n-pentane (Aldrich), HLPC grade dichloromethane (BDH) and HPLC grade methanol (EM Science) were used without further treatment for SPE. The purity of each solvent used for SPE was checked by gas chromatography-mass spectrometry and all solvents were found to be free of contaminants that could potentially interfere with this study. Alumina micro-columns were made by sandwiching 1.0 gram of neutral alumina (Merck) between two plugs of silanised glass wool (Alltech) in a 14.6 cm (5 ¾ inch) Pasteur pipet (MapleLeaf). Alumina micro-columns were stored in an oven at 60 ^{o}C and were pre-heated to 110 ^{o}C for one hour prior to use.[36,37]

The procedure devised by Bundt *et al.*[23] for the SPE of diesel fuels on silica gel was modified for use with gasoline samples on alumina. While still warm, the column was conditioned by washing with 2 mL of petroleum spirit. A 200 μL aliquot of the gasoline sample was diluted with 0.8 mL petroleum spirit and the solution transferred to the head of the column. The aliphatic and monoaromatic compounds were eluted from the column with 5 mL of petroleum spirit followed by 4 mL of a solution of 5% dichloromethane in n-pentane (v/v). The first fraction, containing the polycyclic aromatic hydrocarbons, was eluted with 4 mL of a solution of 10% dichloromethane in n-pentane (v/v). The second fraction, containing the polar compounds, was eluted from the column with 2 mL of methanol. The volume of each fraction was reduced to approximately 0.25 mL under a gentle stream of dry nitrogen, then immediately transferred to a gas chromatography vial, fitted with a conical glass insert, and capped. Each sample was extracted in triplicate.

5.2.3 Gas chromatography-mass spectrometry

Gas chromatography was performed on a 30 m x 0.25 mm i.d. x 0.25 μm film HP5-ms capillary column using an Agilent 6890 gas chromatograph connected to an Agilent 5973 mass selective detector (MSD).

5.2.4 Analysis of extracts

Polar extracts (second fraction) were analysed by gas chromatography-mass spectrometry without chemical derivatisation of the phenolic compounds.[15,16,18] Chromatographic conditions were as follows: split injection (3:1) at 280 °C; temperature programmed from 40 °C (3 minutes isothermal) to 105 °C at 2 °C/minute, then to 290 °C at 20 °C/minute (held for 5 minutes); helium carrier gas held at a constant flow rate of 1.2 mL/minute. Polycyclic aromatic hydrocarbon extracts (first fraction) were analysed using the following chromatographic conditions: split injection (3:1) at 280 °C; temperature programmed from 40 °C (3 minutes isothermal) to 208 °C at 4 °C/minute, then to 290 °C at 25 °C/minute (held for 5 minutes); helium carrier gas held at a constant flow rate of 1.2 mL/minute. For the analysis of both polar and PAH extracts, the sample injection volume was 1 µL, and the detector was operated in full scan mode (30 to 350 daltons) with a sampling rate of 4.45 scans/second.

5.2.5 Selected ion monitoring (SIM) of polycyclic aromatic hydrocarbons

Samples were prepared for analysis using GC-MS (SIM) by diluting an aliquot of neat gasoline with dichloromethane (1:4, v/v). In all cases described below, five aliquots from each sample were analysed. The conditions set for the gas chromatograph and detector were modified according to the type of PAH compounds being detected. For the two-ring PAHs, the same gas chromatograph conditions used for the analysis of the PAH extracts (first fraction) were used here except that the split ratio was increased to 50:1. The detector was set to monitor the following ions: m/z^+ 127, 128, 141, 142, 155, and 156 (50 ms dwell time). A second method for the analysis of the two-ring PAHs by SIM was devised whereby the GC conditions described above were unaltered except that the split ratio was decreased to 15:1, and the detector was set to monitor the following ion groups: m/z^+ 127, 129, 139 and 143 (group 1; C_0- and C_1-naphthalenes); m/z^+ 128, 141, 142, 155, 156 and 157 (group 2; C_2-naphthalenes); m/z^+ 128, 141, 153, 155, 169, 170 and 171 (group 3; C_3-naphthalenes). Dwell times were 100 ms for group 1 and 50 ms for groups 2 and 3. For the three-ring PAHs the GC conditions remained the same except that the split was decreased to 2:1, and the detector was set to monitor the following ions: m/z$^+$ 178, 192 and 206. Ion fragments and their corresponding PAH isomer groups are given in Table 5.3.

Table 5.3 Selected ion fragments for each isomeric group of polycyclic aromatic hydrocarbons found in unleaded gasoline samples.

Two-ring	*m/z$^+$*
C_0-naphthalene	127, 128, 129
C_1-naphthalenes	139, 141, 142, 143
C_2-naphthalenes	127, 128, 141, 142, 155, 156, 157
C_3-naphthalenes	128, 141, 142, 155, 156, 169, 170, 171
Three-ring	
C_0-anthracene	178
C_0-phenanthrene	178
C_1-phenanthrenes	192
C_2-phenanthrenes	206

5.2.6 Retention index standards

Retention indices were calculated from the retention times of known standards. A standard solution of n-alkanes (n-C_5 to n-C_{22}), naphthalene, phenanthrene and chrysene was made up in dichloromethane. A 500 ppm solution (w/v) of each standard compound was made by diluting approximately 5 mg of each standard compound in 100 mL of HPLC grade dichloromethane (BDH). Standards were obtained from the following suppliers: n-C_5, n-C_{20} to n-C_{22}, and chrysene from Aldrich; n-C_6 to n-C_{19} from Polyscience; naphthalene from BDH; and, phenanthrene from Merck. Standards were used without further purification. All n-alkane standards used had a purity of 99% or greater, with the exception of n-C_{21} which had a purity of 98%. The polycyclic aromatic hydrocarbon standards had the following purities: naphthalene, 98%; phenanthrene, 96%; and, chrysene, 95%. The HPLC grade dichloromethane had a purity greater than 99.8%.

5.2.7 *Data analysis*

Peak area information was extracted from integrated chromatograms and written to an ASCII text file using a macro written specifically for this study using the instrument software (HP ChemStation, Agilent). The macro is reproduced in Appendix 5.1. Integration data was loaded and parsed into a spreadsheet (Microsoft Excel®). The data was normalized using a natural logarithmic transformation.[38] Principal component analysis, linear discriminant analysis, and cluster analysis was performed using Minitab® (version 13.1 for Windows).

5.2.8 *Quantitation of naphthalene, anthracene and phenanthrene*

The concentrations of naphthalene, anthracene and phenanthrene present in gasoline samples were determined by gas chromatography-mass spectrometry (GC-MS) operating in selected ion monitoring (SIM) mode. A calibration curve for each target compound was constructed by diluting aliquots of a prepared stock solution. The stock solution (approximately 1000 ppm) was made by dissolving an accurately weighed amount (approximately 142 g) each of naphthalene (BDH), anthracene (Hopkin & Williams), and phenanthrene (Merck) in HPLC grade dichloromethane (BDH). This solution was transferred to a 100 mL volumetric flask and the final volume made up with dichloromethane.

Aliquots of the prepared calibration mixtures were analysed by GC-MS without further dilution. Gasoline samples (Table 5.2) were diluted with dichloromethane (1:4 v/v) prior to analysis. Instrument conditions for GC-MS (SIM) analysis of naphthalene were as follows: split injection (50:1) at 280 ^{o}C; temperature programmed from 40 ^{o}C (3 minutes isothermal) to 148 ^{o}C at 4 ^{o}C/minute, then to 290 ^{o}C at 25 ^{o}C/minute (held for 5 minutes); helium carrier gas at a constant flow rate of 1.2 mL/minute. The detector was set to monitor the following ions: m/z^{+} 127, 128, 141, 142, 155, and 156 (50 ms dwell time). Instrument conditions for GC-MS (SIM) analysis of anthracene and phenanthrene were as follows: split injection (2:1) at 280 ^{o}C; temperature programmed from 40 ^{o}C (3 minutes isothermal) to 208 °C at 4 °C/minute, then to 290 °C at 25 °C/minute (held for 5 minutes); helium carrier gas at a constant flow rate of 1.2 mL/minute. The detector was set to monitor the following ions: m/z^+ 178, 192, and 206.

A five-point calibration curve was constructed for naphthalene and four-point calibration curves were constructed for anthracene and phenanthrene. Each of the three target compounds gave good linear response with no correlation coefficient falling below 0.991. The concentration of each target compound present in a gasoline sample was calculated from the calibration curve using its integrated peak area, taking into account the sample dilution with dichloromethane.

5.3 Polar compounds

5.3.1 Results

In order to evaluate the potential for using polar compounds for characterising the gasoline samples, it was necessary to first isolate and identify all of the polar compounds using solid phase extraction. A micro solid phase extraction method using activated alumina was developed to isolate polar compounds from a sample of gasoline prior to analysis by gas chromatography-mass spectrometry. For the thirty five gasoline samples extracted (in triplicate), visual inspection of the chromatograms did not reveal any discernable difference in the pattern of the polar compounds between samples. In some samples (samples 32 to 35, Table 5.2) the presence of the oxygenate ethanol was sufficient to cause the majority of the polar compounds to elute from the micro-column immediately after the introduction of the sample onto the head of the column. In these four samples, only a small concentration of the polar compounds was retained on the alumina column and subsequently recovered for analysis.

5.3.2 Discussion

5.3.2.1 Solid phase extraction of polar compounds

Solid phase extraction (SPE), or adsorption chromatography, is the use of a solid adsorbent to separate analytes of interest from a complex mixture. This technique has been successfully applied to the problem of separating a large variety of target analytes from a complex mixture in solution.[39,40]

Of the strongly polar adsorbents, silica gel and alumina are historically the most popular for solid phase extractions of complex mixtures.[40] Both of these adsorbents show similar preferential adsorption of polar organic compounds. One significant difference between these two adsorbents is the preference of alumina to adsorb acidic compounds such as alkyl phenols,[40,41] whilst silica has a preference to adsorb basic compounds such as aliphatic amines. Alumina is, therefore, generally considered to be basic in nature, while silica gel is considered to be acidic.[41] Snyder[41] also showed that for some weak organic bases such as aniline, silica gel does not necessarily exhibit greater adsorption than alumina.

There are only a few studies reported in the literature that describe the extraction of polar compounds from gasoline using an adsorbent. Each of these studies employed silica gel as the adsorbent. No studies were found where alumina was used to extract polar compounds for subsequent analysis by gas chromatography. Early studies, such as the work of Snyder and Buell,[42] did use an alumina adsorbent in column chromatography to effect a separation of a variety of compounds from petroleum fractions. Youngless *et al.*[15] removed polar compounds from hydrocarbon fuel using a silica gel column. They identified 62 compounds, the majority being polar compounds, including alcohols, amines, phenols and pyridines. Using methods similar to Youngless *et al.*,[15] Kanai *et al.*[16,17] also used a silica gel column to extract polar compounds from different gasoline samples. Unlike the results of Youngless *et al.*,[15] the majority of polar compounds recovered by Kanai *et al.*[16] were alkyl phenols. Bennett *et al.*[21] used a C_{18}-bonded silica adsorbent to extract C_0- to C_3-alkyl phenols from samples of crude oil. The alkyl phenols present in the crude oils examined by

Bennett *et al.*[21] were similar those found by Youngless *et al.*[15] and Kanai *et al.*[16,17] indicating that the alkyl phenols found in gasoline can come directly from crude oil. Although both silica gel and alumina would be considered good general adsorbents for polar molecules, it has been shown that alumina is the better choice for the extraction of polar molecules from a liquid organic matrix.[40,41,42] In the work presented in this chapter, alumina was chosen over silica gel for the extraction of polar compounds from gasoline.

5.3.2.2 Alumina activation

It is generally accepted that water molecules adsorbed onto alumina can be driven off at temperatures between 100 and 150 °C.[36] In a more recent study, Later *et al.*[37] examined the effect of oven temperature on the degree of alumina activation. Alumina activation is expressed as the weight percent of the amount of water adsorbed on the alumina.[37] The authors found that alumina stored in uncontrolled conditions in the laboratory could be re-activated using relatively mild conditions; heating alumina in an oven set to 150 °C yielded a water content of 1.0% to 1.5% within two hours, while heating at 100 °C yielded an alumina with a water content of 1.5% to 2.0% within two hours.[37] Later *et al.*[37] cautioned that these values were approximate because no attempt was made to control for the humidity levels in and around the oven. For the research described in this chapter, the alumina micro-columns were stored in an oven at 60 °C and were pre-heated to 110 °C for one hour prior to use (see section 5.2.2).

5.3.2.3 Analysis for polar compounds by gas chromatography

Chemical derivatisation has been applied to phenolic compounds prior to detection by gas chromatography.[21,43,44] The rationale behind derivatisation of phenols is primarily to improve quantitation by reducing or eliminating peak tailing and to assist in the resolution of isomers. However, it may not always be necessary to chemically derivatise polar compounds such as alkyl phenols prior to analysis by gas chromatography. A few researchers have analysed polar compound extracts obtained from gasoline,[15,16] and crude oil[18] without further treatment (*i.e.* with no chemical derivatisation). Ioppolo *et al.*[18]

studied the chromatographic behaviour and elution order of C_0- to C_3-phenols extracted from several crude oils on columns having different stationary phases. The authors concluded that peak tailing was minimal and that the majority of isomers could be separated, which allowed most compounds to be identified and quantitated.[18] The analysis for polar compounds extracted from the 35 gasoline samples in this work was made without derivatisation and good chromatographic separations were obtained.

The columns and temperature programs used for gas chromatographic separation of polar compounds have varied significantly over the years (Table 5.4). From the literature it can be seen that good, general use non-polar or weakly polar (*e.g.* 5% phenyl, 95% dimethylpolysiloxane) capillary columns have been employed for separation of alkylphenol and related types of polar compounds, including their silyl derivatives. Without derivatisation, some chromatographic peak tailing of the phenolic analytes is expected. However, separating phenolic compounds without derivatisation is not expected to affect the overall chromatographic pattern obtained from a sample, and so any differences in the relative amounts of different polar compounds between samples should still be evident. Column activity may play a role in the loss of trace polar compounds during analysis by gas chromatography.[45] Distortion of analyte peaks was not observed for the polar compounds extracted from the gasoline samples indicating that there was minimal column activity at the time of analysis. The affect of column activity on the results presented here is probably insignificant. An HP5-ms column was employed for all gas chromatographic separations in this work. This column has a stationary phase of 5% phenyl, 95% dimethylpolysiloxane but has a different cross-linked structure from the traditional columns of this type originally developed for flame ionisation detection (*e.g.* DB-5). The stationary phase of a DB-5 capillary column has the structure $\text{-[-O-Si(CH}_3)_2\text{-]-[-O-Si(C}_6\text{H}_5)_2\text{-]}_n\text{-}$, while the stationary phase of an HP5-ms capillary column, intended for use with mass spectrometry detectors, has the structure $\text{–[-O-Si(CH}_3)_2\text{-C}_6\text{H}_4\text{-Si(CH}_3)_2\text{-]-[-O-Si(C}_6\text{H}_5)_2\text{-]}_n\text{-}$.[46] The different structure of the MS column is intended to improve the signal to noise ratio so that it will work with the more sensitive GC-MS systems.

Table 5.4 Selected gas chromatographic conditions used to analyse for polar compounds.

Year	Column	Temperature program	Authors
1985	Non-bonded, non-cross linked capillary, DX-4 (15% dimethylpolysiloxane, 85% polyethylene glycol) 30 m x 0.33 mm i.d. 0.25 µm film thickness	50°C initial, hold 2 min 6°C/min to 230°C	Youngless *et al.*[1]
1989	Fused silica capillary, HP-PONA (100% dimethylpolysiloxane) 50 m x 0.20 mm i.d. 0.5 µm film thickness	35°C initial; 10°C/min to 250°C, hold 15 min	Kanai *et al.*[2]
1991	Fused silica capillary, HP-PONA (100% dimethylpolysiloxane) 50 m x 0.20 mm i.d. 0.5 µm film thickness	40°C initial; 15°C/min to 250°C, hold 15 min	Kanai *et al.*[3]
1992	Fused silica capillary, BP1 (100% dimethylpolysiloxane) 60 m x 0.2 mm i.d. film thickness not specified	40°C initial, hold 1 min 4°C/min to 280°C	Ioppolo *et al.*[4]
1994	fused silica capillary, Quadrex 100 m x 0.25 mm i.d. 0.5 µm film thickness	30°C initial, hold 2 min; 20°C/min to 70°C; 2°C/min to 320°C, hold 10 min	Green *et al.*[5]
1996	fused silica capillary, DB-5 (5% diphenyl, 95% dimethylpolysiloxane) 30 m x 0.32 mm i.d. 0.25 µm film thickness	35°C initial, hold 10 min; 2°C/min to 150°C; 8°C/min to 300°C, hold 20 min	Bennett *et al.*[6]

1. Youngless, T.L.; Swansiger, J.T.; Danner, D.A.; Greco, M. Mass spectrometry characterization of petroleum dyes, tracers, and additives. Analytical Chemistry, 1985; 57(9): 1894-1902.
2. Kanai, H.; Yazawa, L.; Maka, J. Baseline data on the quality of gasolines in Hawaii and their implication in preparing alcohol blends. Hawaii State Department of Agriculture, Honolulu, HI; 1989.
3. Kanai, H.; Inouye, V.; Goo, R; Yazawa, L.; Maka, J.; Chun, C. GC/MS analysis of polar compounds in "weathered" gasoline/water matrix as an aid in identifying gasoline. Analytical Letters, 1991; 24(1): 115-28.
4. Ioppolo, M.; Alexander, R.; Kagi, R.I. Identification and analysis of C_0-C_3 phenols in some Australian crude oils, Organic Geochemistry, 1992; 18(5): 603-609.
5. Green, J.B.; Yu, S.K.T.; Vrana, R.P. GC-MS analysis of phenolic compounds in fuels after conversion to trifluoroacetate esters. Journal of High Resolution Chromatography, 1994; 17(6): 439-451.
Gas chromatography conditions in:
Green, J.B.; Yu, S.K.T.; Vrana, R.P. GC-MS analysis of carboxylic acids in petroleum after esterification with fluoroalcohols. HRC-Journal of High Resolution Chromatography, 1994; 17(6): 427-438.
6. Bennett, B.; Fowler, B.F.J.; Larter, S.R. Determination of C_0-C_3 alkylphenols in crude oils and waters. Analytical Chemistry, 1996; 68(20): 3697-3702.

5.4 Polycyclic aromatic hydrocarbons (PAHs)

5.4.1 Results

5.4.1.1 Solid phase extraction of polycyclic aromatic hydrocarbons

It was necessary to isolate and attempt to identify the classes of polycyclic aromatic hydrocarbons present in the gasoline matrix prior to developing a method for analysing these analytes directly by gas chromatography-mass spectrometry with selected ion monitoring (SIM). Solid phase extraction was performed in triplicate on each of the 35 gasoline samples and the resulting PAH extracts were found to be reproducible for each sample. The PAH content in each of the 35 gasoline samples was found to be very consistent. Table 5.5 lists the isomeric groups of the two- and three-ring PAHs found in each of the 35 samples.

Table 5.5 Isomeric groups of polycyclic aromatic hydrocarbons identified in unleaded gasoline samples.

Two-ring	*Three-ring*	
C_0-naphthalene	C_0-fluorene	C_0-anthracene
C_1-naphthalenes	C_1-fluorenes	C_0-phenanthrene
C_2-naphthalenes	C_2-fluorenes	C_1-phenanthrenes
C_3-naphthalenes		C_2-phenanthrenes

5.4.1.2 Analysis for polycyclic aromatic hydrocarbons by gas chromatography

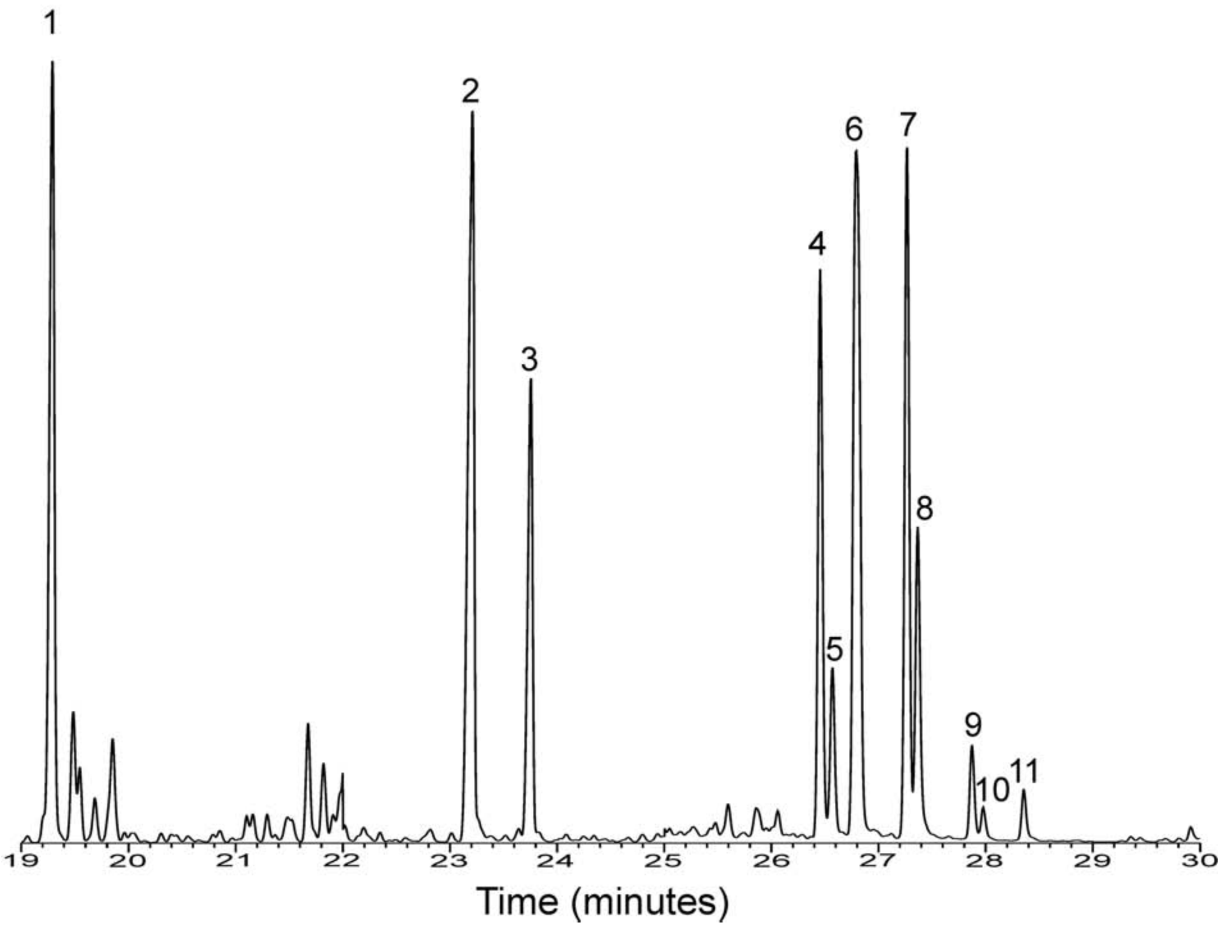
A visual inspection of the chromatogram for each polycyclic aromatic hydrocarbon (PAH) extract was made. Differences between gasoline samples were observed in the relative ratios of the C_0- to C_2-naphthalenes (two-ring PAHs), and in the relative amounts of anthracene and the C_0- to C_2-phenanthrenes (three-ring PAHs). Although the PAH extracts for any given sample were reproducible, it was found that the precision was not sufficient to give good clustering of individual samples after principal component analysis. In order to improve the analytical precision, GC-MS (SIM) was employed to analyse for the two- and three-ring PAHs directly. Although the precision between replicates was improved when analysing for the PAHs, the precision for the detection of three-ring PAHs was still unsatisfactory. The concentrations of anthracene and phenanthrene were determined by GC-MS (SIM) and found to be very low (Table 5.6). It is postulated that the low concentrations of these compounds resulted in a relatively large integration error, and so the desired analytical precision could not be obtained. The concentration of naphthalene was also determined for each sample (Table 5.6). It was found that the concentration of naphthalene (and by analogy, its methylated homologues) was significantly greater than that of the three-ring PAHs, and so attention was turned to these compounds.

Table 5.6 Concentrations of naphthalene, phenanthrene and anthracene in 35 unleaded gasoline samples.

	Concentration (ppm)		
	naphthalene	*phenanthrene*	*anthracene*
mean	1006	15	7
standard deviation	336	8	8
minimum	392	5	1
maximum	1880	42	41
number of chromatograms	175	175	175
number of gasoline samples	35	35	35

Although precision for the detection of C_0- to C_2-naphthalenes using solid phase extraction techniques was good, some samples had very low concentrations of the C_2-naphthalenes relative to the C_0- to C_1-naphthalenes which resulted in a lower precision for the integration results from these samples. A new approach was, therefore, attempted whereby the signal for the C_0- and C_1-naphthalenes was reduced by selectively monitoring the less abundant ions for these compounds, whilst continuing to monitor the most abundant ions for the C_2-naphthalenes. The split ratio could be decreased to allow more of the C_2-naphthalenes to be injected onto the column, and the signal for the C_0- and C_1-naphthalenes was reduced sufficiently to prevent the detector from overloading (Figure 5.1). Good precision in the integration results was obtained using this technique, however, care had to be taken to avoid overloading the column with naphthalene.

Figure 5.1 Chromatogram of C_0- to C_2-naphthalenes obtained by selected ion monitoring of the three isomer groups. GC-MS conditions given in *Materials and methods*, section 5.2.5. Peak assignments are given in Table 5.7.



The C_0- to C_2-naphthalenes were identified by comparison to NIST standard mass spectra and by retention indices reported in the literature.[25,47] Retention indices were calculated using both an n-alkane system (n-C_5 to n-C_{22})[48] and a system based on the retention of polycyclic aromatic hydrocarbon standards (naphthalene, phenanthrene and chrysene).[47] The retention indices for the C_0- to C_2-naphthalenes are shown in Table 5.7 and were calculated from 75 chromatograms representing 15 samples (*i.e.* five aliquots per sample).

Table 5.7 Retention indices for C_0- to C_2-naphthalenes by GC-MS (SIM).

		PAH retention index		n-alkane retention index	
No.	*Compound*	*Lee et al.*[1]	*This work*[†]	*Lai et al.*[2]	*This work*[†]
1	naphthalene	200.00	200.08 ± 0.012	1179.7	1179.6 ± 0.06
2	2-methylnaphthalene	218.14	221.22 ± 0.030	1290.5	1290.6 ± 0.16
3	1-methylnaphthalene	221.04	224.31 ± 0.021	1306.8	1307.0 ± 0.12
4	2-ethylnaphthalene	236.08	239.32 ± 0.003	1390.6	1390.0 ± 0.02
5	1-ethylnaphthalene	236.56	239.92 ± 0.002	1393.8	1393.4 ± 0.01
6	2,6-dimethylnaphthalene + 2,7-dimethylnaphthalene	237.58 + 237.71	241.13 ± 0.006	1400.9 + 1402.2	1400.2 ± 0.04
7	1,3-dimethylnaphthalene	240.25	243.75 ± 0.003	1416.5	1415.4 ± 0.02
8	1,6-dimethylnaphthalene + 1,7-dimethylnaphthalene	240.66 + 240.72	244.29 ± 0.003	1419.6	1418.7 ± 0.02
9	1,4-dimethylnaphthalene + 2,3-dimethylnaphthalene	243.55 + 243.57	247.02 ± 0.003	1436.3	1435.0 ± 0.02
10	1,5-dimethylnaphthalene	244.98	247.62 ± 0.002	1439.8	1438.4 ± 0.01
11	1,2-dimethylnaphthalene	246.49	249.70 ± 0.004	1452.2	1450.6 ± 0.02

† Retention indices calculated from retention index standards described in section 5.2.6.

1. Lee, M.L.; Vassilaros, D.L.; White, C.M.; Novotny, M. Retention indices for programmed-temperature capillary-column gas chromatography of polycyclic aromatic hydrocarbons. Analytical Chemistry, 1979; 51(6): 768-773.
2. Lai, W.C. and Song, C.S. Temperature programmed retention indices for g.c. and g.c.-m.s. analysis of coal- and petroleum-derived liquid fuels. Fuel, 1995; 74 (10): 1436-1451.

5.4.1.3 Data analysis by principal components

The data obtained from the C_0- to C_2-naphthalenes was examined using principal component analysis (PCA). The variation described by the first three principal components (PCs) is given in Table 5.8. The first two PCs represent 91.5% of the variation in the data. From a plot of the first two PC scores from the PCA results (Figure 5.2) it can be seen that the majority of the samples can be distinguished using this technique.

Figure 5.2 PC score plot of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. Data labels refer to samples described in Table 5.2.

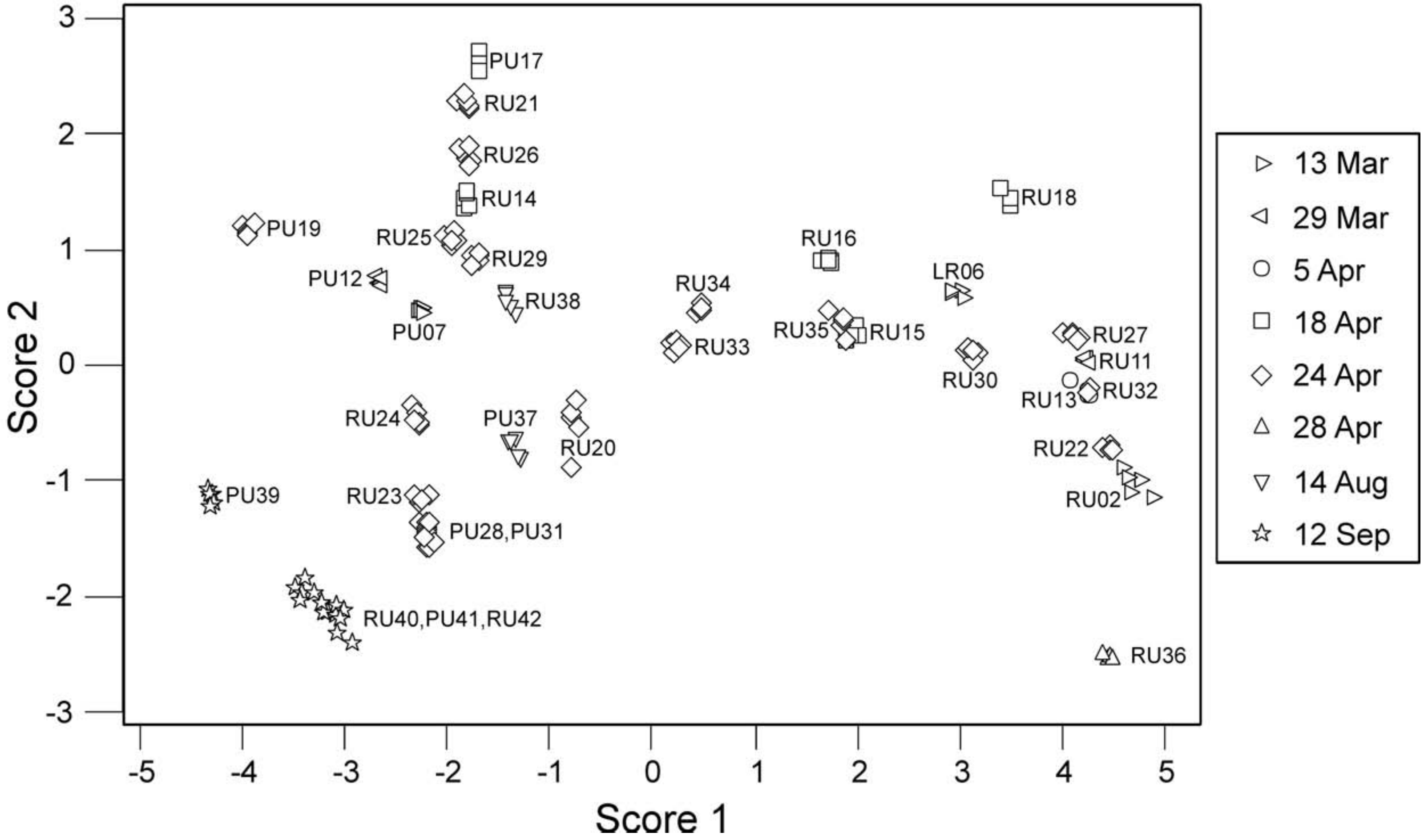


Table 5.8 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM).

	Eigenvalue	Proportion	Cumulative (%)
PC1	8.4765	0.771	77.1
PC2	1.5857	0.144	91.5
PC3	0.7925	0.072	98.7

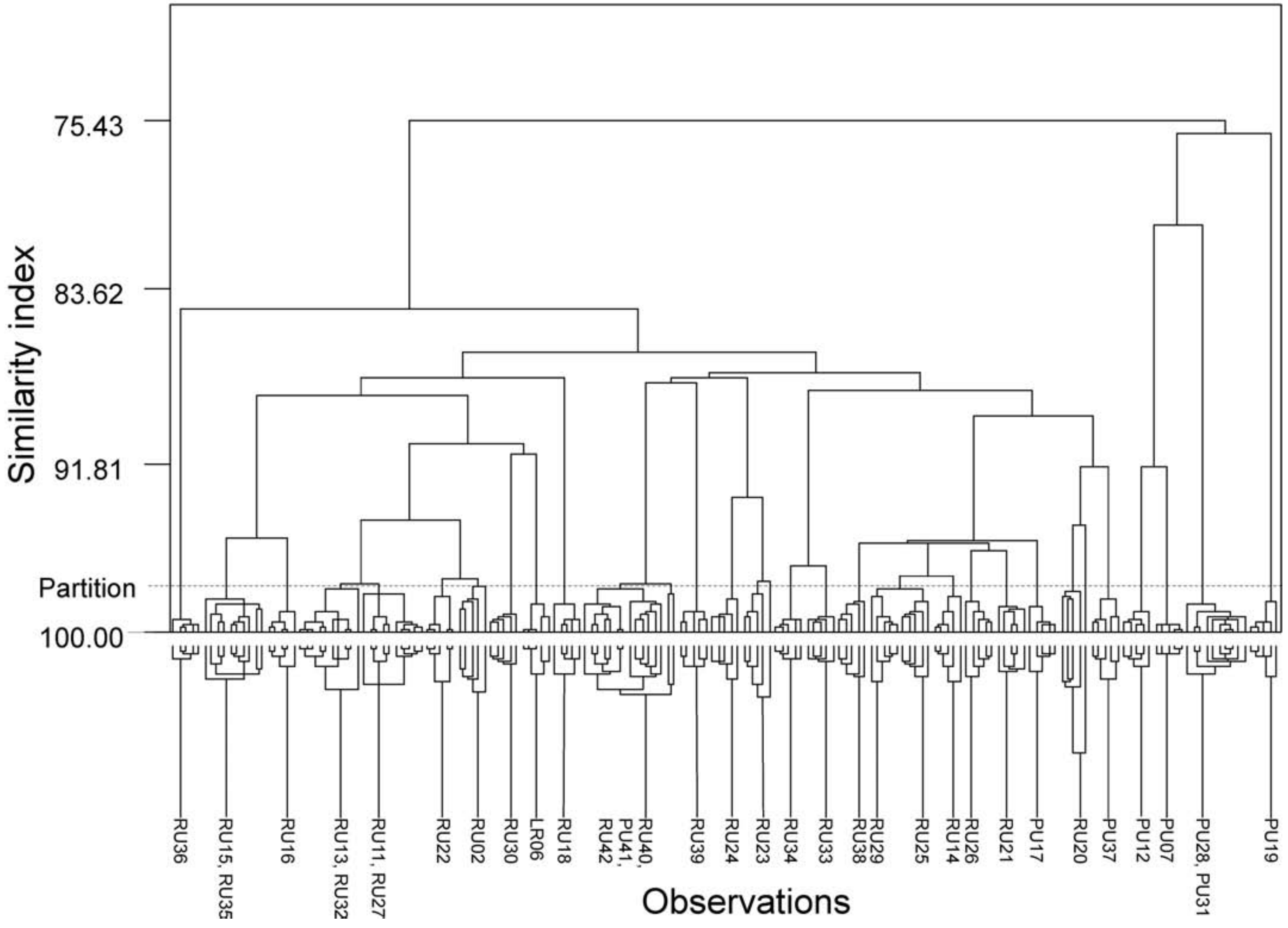
Linear discriminant analysis (LDA) with cross-validation using the first three PCs correctly classified 168 of the 175 chromatograms (96%) by sample. In practical terms, any overlap of one chromatogram (one aliquot) from one sample with any other chromatogram from another sample would cause these two gasoline samples to be indistinguishable. Therefore, the LDA results were re-examined and it was found that the 35 gasoline samples could be divided into 32 unique groups. Thirty of these groups each contained a single sample (*i.e.* five aliquots from one sample per group). The five remaining (misclassified) samples were divided between the two remaining groups. Two of these five samples formed a single group (samples 8 and 28; *i.e.* RU15 and RU35), while the other three samples formed another group (samples 33, 34 and 35; *i.e.* RU40, PU41, RU42) (Table 5.9).

Table 5.9 Results of LDA with cross-validation: number of aliquots from the five misclassified samples.

		True Group				
		RU15	RU35	RU40	PU41	RU42
Predicted Group	RU15	3	1	0	0	0
	RU35	2	4	0	0	0
	RU40	0	0	3	0	1
	PU41	0	0	0	5	1
	RU42	0	0	2	0	3

Cluster analysis was also attempted in order to sort the 175 chromatograms (*i.e.* 175 aliquots) into meaningful groups. The first three PCs were analysed using a "nearest neighbour" linkage between groups (clusters) based on Euclidean distance. It was found that a significant partition occurred around 35 groups (Figure 5.3). Cluster analysis divided the 35 samples into 29 unique groups. Twenty-four of these unique groups each contained

Figure 5.3 Dendogram resulting from cluster analysis performed on first three PCs using "nearest neighbour" linkage between groups (clusters) based on Euclidean distance. Data labels refer to samples described in Table 5.2.



a single sample. The remaining 11 samples were divided between the five remaining groups. The five groups containing more than one gasoline sample were as follows: RU11 and RU27; RU13 and RU32; RU15 and RU35; PU28 and PU31; and, RU40, PU41, and RU42 (Figure 5.3).

The PCA results demonstrate that gasoline samples collected on the same day from different retail stations can be distinguished from one another. For example, the five samples collected on April 18, 2001 (Table 5.2) were purchased from four different retail stations within a 400 meter radius. These five samples are highlighted in the PC score plot in Figure 5.4, where it can be seen that the five samples can be distinguished from one another. The 17 samples collected on April 24, 2001 (Table 5.2) were collected from 14 retail stations in a different part of Sydney from where the samples were collected on April 18. These 14 stations were all within a 7 km radius from each other. These samples are highlighted in the PC score plot in Figure 5.5, which shows good differentiation between most of the samples. Using the first three principal components, linear discriminant analysis correctly classified all 85 chromatograms (100%) by sample. The first two PC scores for all 35 gasoline samples were replotted by fuel grade in Figure 5.6. A close inspection of Figure 5.6 shows that all nine premium unleaded gasoline samples have a negative value for the first score. This indicates that the refining process imparted a particular naphthalene profile to these premium grade samples. There is, however, not enough information held in the naphthalene profile to allow the grade of the gasoline to be unequivocally assigned.

Cluster analysis was also applied to the PCA results to provide another, independent examination of the sample groupings. The appropriate number of groups (clusters) into which this particular data set may be divided is somewhere between 175 and one. The decision where to divide, or partition, the dendogram (Figure 5.3) is made by examining the similarity index. As the clustering proceeds from 175 groups (*i.e.* one aliquot per group) to one group containing all 175 aliquots the similarity index decreases from 100. Initially the similarity index will decrease slowly until a sudden, large decrease occurs. It is typically at this first, large decrease that the partition is made and the group membership of each aliquot

Figure 5.4 PC score plot of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. Five gasoline samples collected from four service stations on April 18, 2001 are hightlighted. Data labels refer to samples described in Table 5.2.

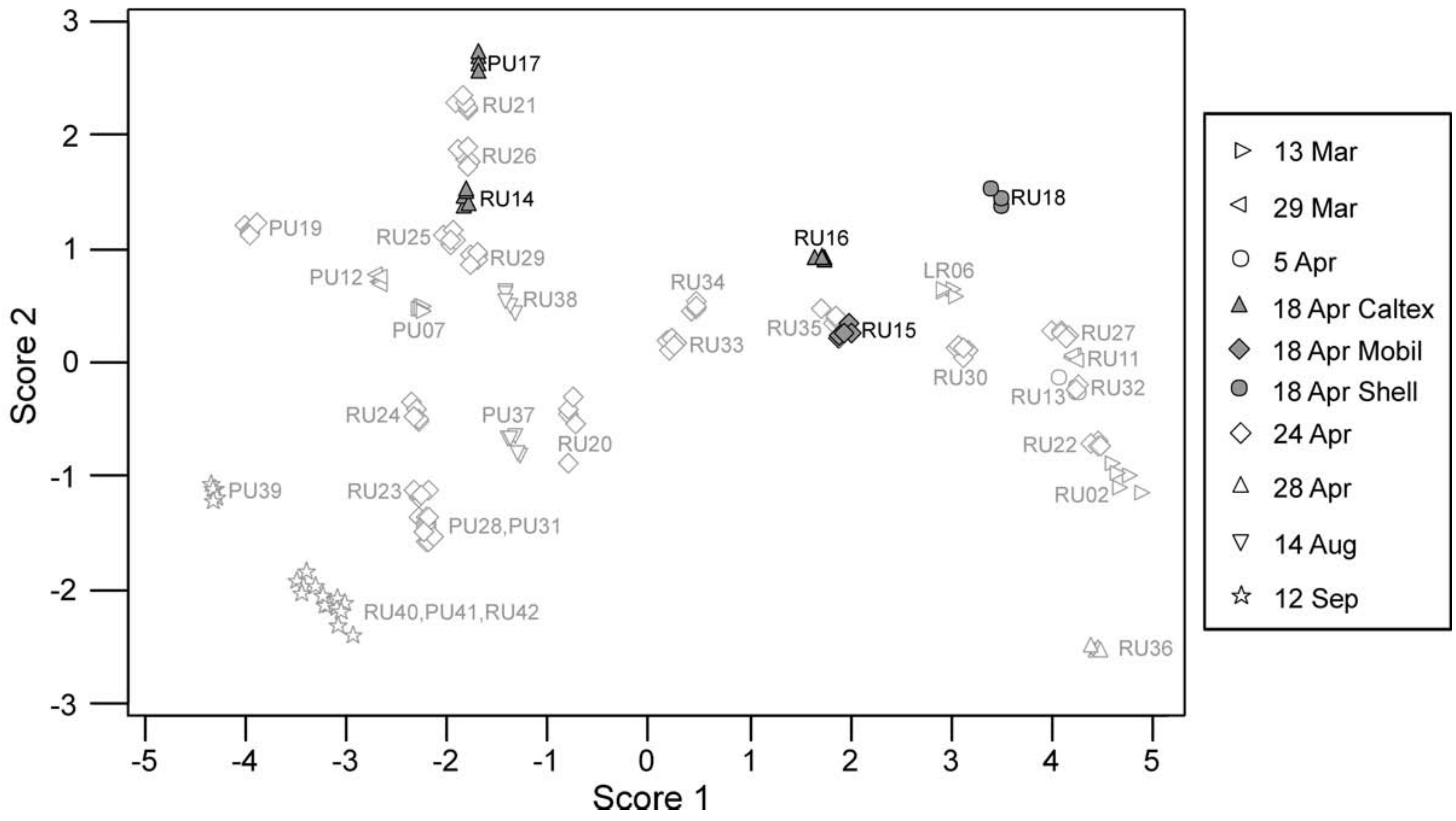
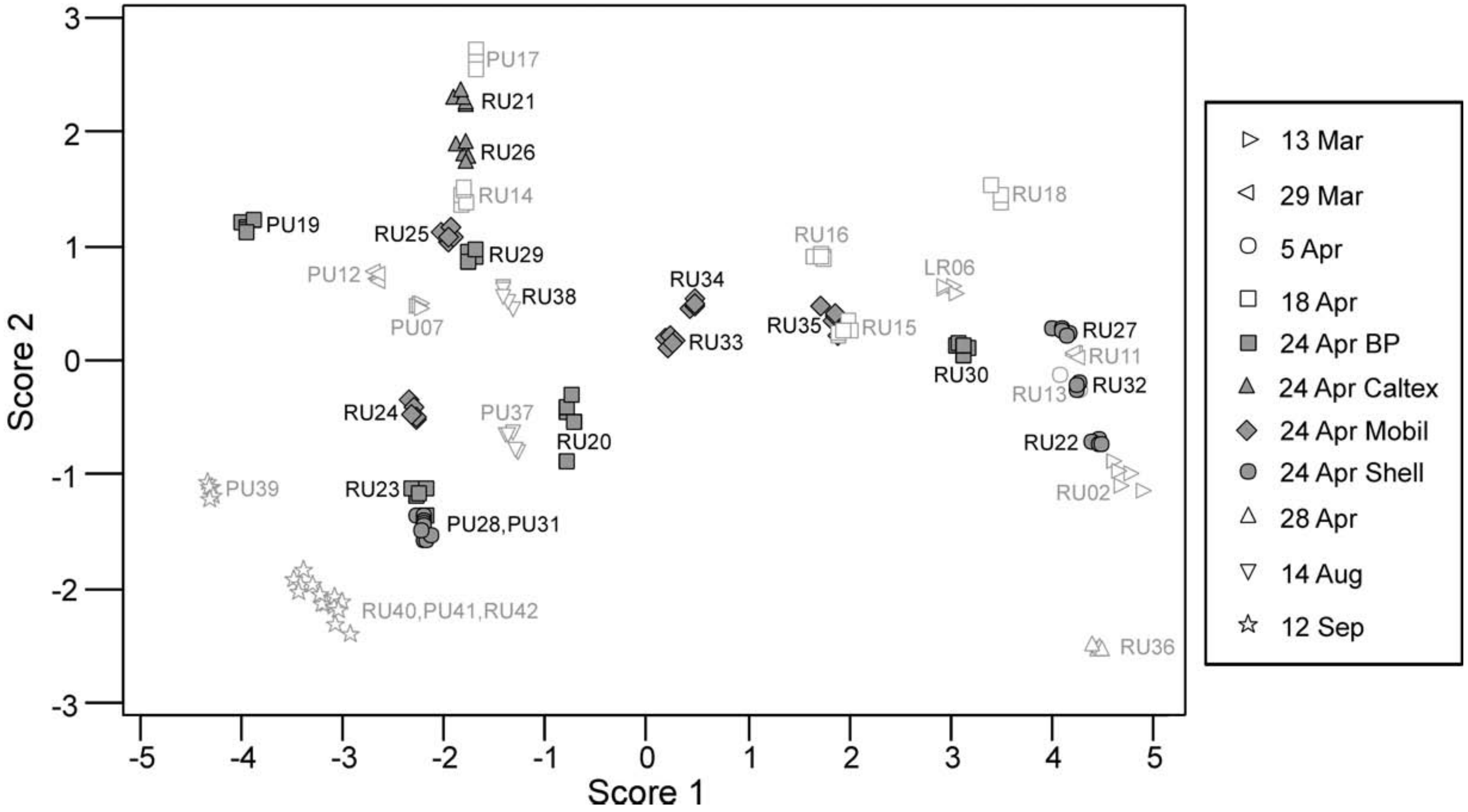


Figure 5.5 PC score plot of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. Seventeen gasoline samples collected from 14 service stations on April 18, 2001 are hightlighted. Data labels refer to samples described in Table 5.2.



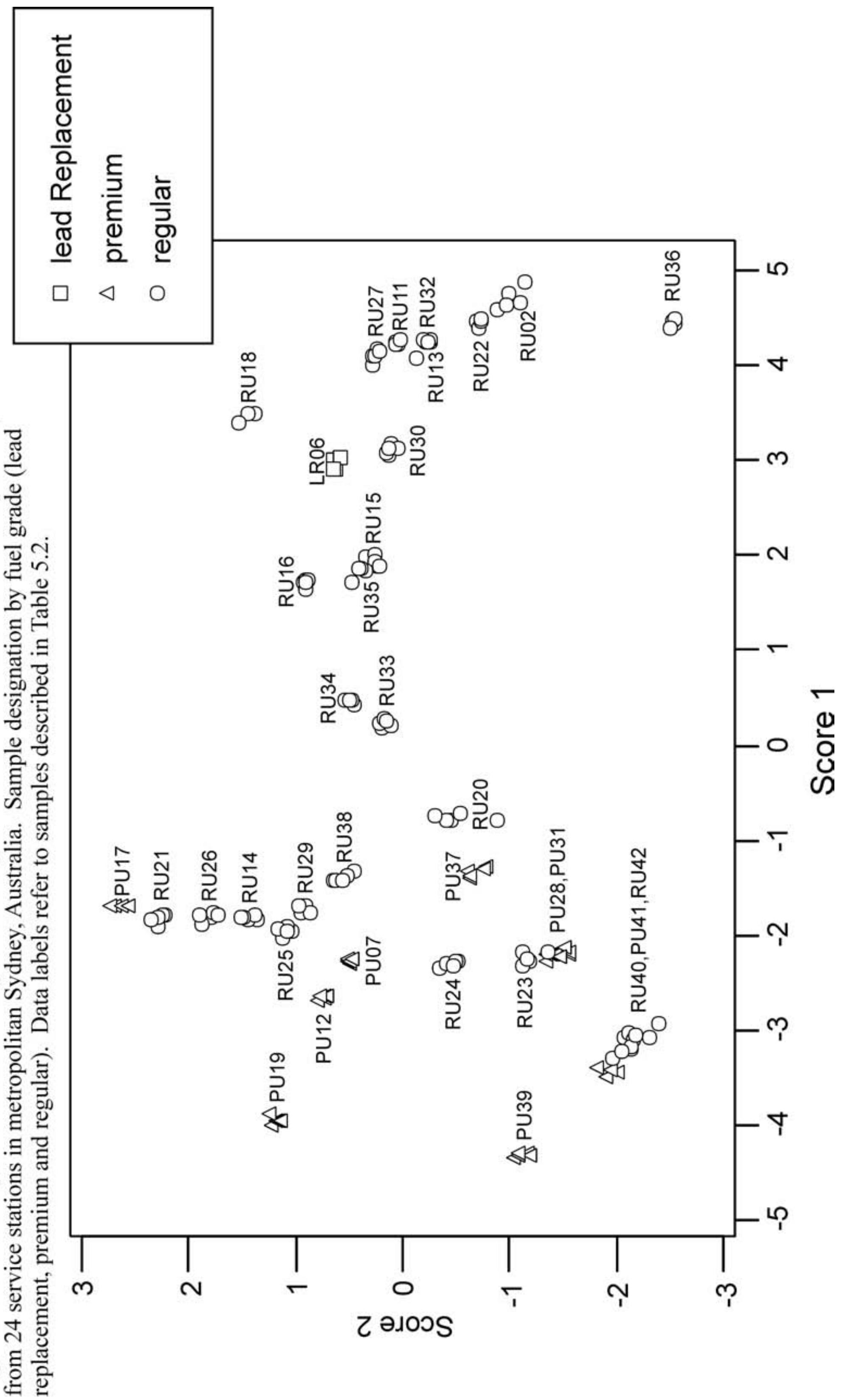


Figure 5.6 PC score plot of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. Sample designation by fuel grade (lead replacement, premium and regular). Data labels refer to samples described in Table 5.2.

is determined from this point. Cluster analysis gave very similar results to LDA in that the same two groups containing multiple samples were identified (RU15 and RU35; and RU40, PU41 and RU42). Cluster analysis was slightly less optimistic than LDA and identified three other clusters, each containing a pair of gasoline samples (RU11 and RU27; RU13 and RU32; and, PU28 and PU31). The five groups that contained more than one gasoline sample after cluster analysis have been highlighted in a PC score plot (Figure 5.7).

5.4.2 Discussion

5.4.2.1 Solid phase extraction of polycyclic aromatic hydrocarbons

It is well known that activated alumina is well suited to the adsorption and separation of aromatic hydrocarbons.[36] Snyder[49] compared the use of alumina and silica to separate different compound classes in petroleum. It was found that alumina was superior for the separation of monoaromatics (*e.g.* alkyl benzenes) from diaromatics (*e.g.* alkyl naphthalenes). In the second paper of this series, Snyder[50] demonstrated the utility of using alumina to separate napthalenes from phenanthrenes. The suitability of alumina for the separation of aromatic compounds is due, in part, to the planar or near planar nature of many aromatic and polycyclic aromatic hydrocarbons[36] which are retained by the planar geometry of the alumina crystal structure. Klemm *et al.*[51] developed the "Coplanarity Rule" after they observed that the more coplanar the PAH molecule, the greater its adsorption on alumina. Snyder[52] went on to examine the adsorption of a large number of fused polycyclic aromatic hydrocarbons on activated alumina. In this study, Snyder found that the free energy of adsorption of a given fused PAH is approximately proportional to the carbon number of the compound, suggesting that adsorption on alumina arises primarily from the orientation of these large, nearly planar molecules, with the crystal plane of the adsorbent.

More recently, Later *et al.*[53] described a procedure for the separation and identification of organic compounds in synthetic fuels using solid phase extraction over alumina and analysis by gas chromatography-mass spectrometry. Similar work was performed by Song

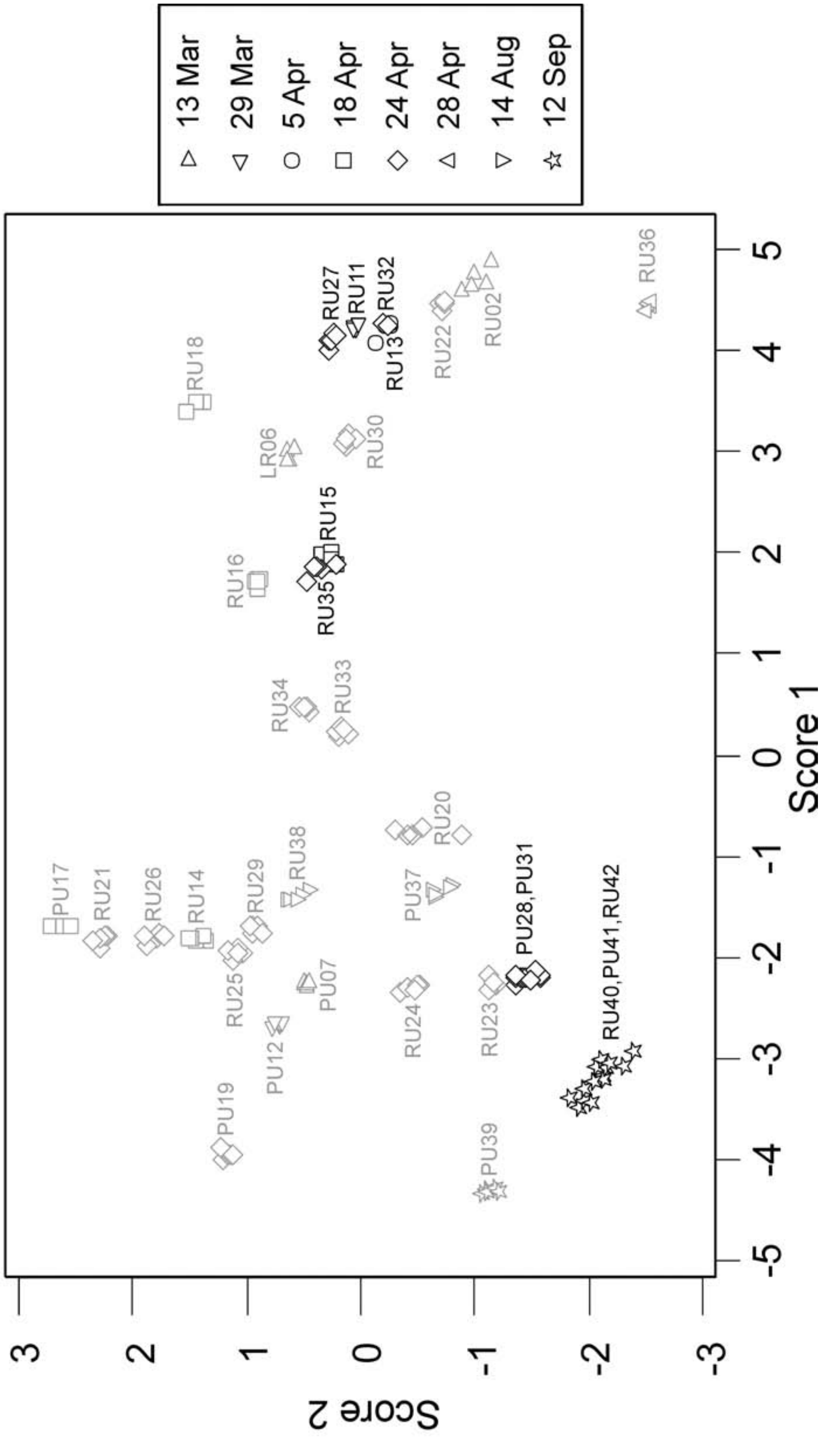


Figure 5.7 PC score plot of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. The five clusters that contain more than one gasoline sample are hightlighted. Clusters identified by cluster analysis (see text). Data labels refer to samples described in Table 5.2.

and Hatcher,[24] and Lai and Song[25] where alumina was used to separate different classes of aromatic compounds from jet fuels. It has also been shown, however, that silica gel can be used to separate aliphatics, monoaromatics, diaromatics, polyaromatics and polar compounds in diesel fuels.[23] The early work of Klemm *et al.*[51] and Snyder[49,50,52] has demonstrated that alumina is better suited than silica gel for solid phase extraction of polycyclic aromatic hydrocarbons, therefore, alumina was chosen for the work presented in this chapter.

5.4.2.2 Analysis for polycyclic aromatic hydrocarbons by gas chromatography

A 5% diphenyl, 95% dimethylpolysiloxane stationary phase has been widely used for efficient separation of PAHs by gas chromatography (Table 5.10).[25,26,47,53,54,55,56,57,58,59] Table 5.10 lists the conditions used by selected authors, over the past 25 years, for the analysis of PAHs by gas chromatography. It is interesting to note how temperature programs used to separate PAHs on a gas chromatograph have varied over the years (Table 5.10). Lai and Song[25] explored linear temperature programs in detail by comparing ramp rates of 2, 4 and 6 ^{o}C/minute (initial temperature 40 ^{o}C) on a 30 metre column. They found that different compounds either co-elute, or are separated depending on which ramp rate is used. Their results did not point to a single, best temperature program. Lai and Song[25] did find a large temperature dependence for the elution of two-, three- and four-ring polycyclic aromatic hydrocarbons. Overall they found little difference in the abilities of the different temperature programs to separate the compounds of interest and chose to only publish chromatograms obtained using the 4 ^{o}C/minute ramp rate. Selected ion monitoring (SIM) of PAHs for the identification and comparison of oil spills has been described.[26,27,57] Selected ion monitoring for specific analytes reduces the number of ions detected per scan resulting in an increase in sensitivity by approximately an order of magnitude.[26]

The analysis of PAHs by capillary gas chromatography has been studied extensively. Gas chromatography-mass spectrometry, including selected ion monitoring, is probably the most widely used and most useful method for separating and identifying polycyclic

Table 5.10 Selected gas chromatographic conditions used over the past 25 years to analyse for PAHs *(continued on next page).*

Year	Column (5% diphenyl, 95% dimethylpolysiloxane)	Temperature program	Authors
1976	glass capillary, SE-52 11 m x 0.26 mm i.d. film thickness not specified	70°C initial; 2°C/min to 230°C.	Lee *et al.*[1]
1979	glass capillary, SE-52 12 m x 0.30 mm i.d. 0.34 μm film thickness	50°C initial; 2°C/min to 250°C.	Lee *et al.*[2]
1981	fused silica capillary, SE-52 20 m x 0.30 mm i.d. 0.25 μm film thickness	50°C initial, hold 2 min; 3°C/min to 250°C, hold 5 min.	Later *et al.*[3]
1982	fused silica capillary, SE-52 25 m x 0.20 mm i.d. film thickness not specified	40°C initial, hold 2 min; 4°C/min to 260°C.	Lee *et al.*[4]
1985	fused silica capillary, SE-52 25 m 0.15 μm film thickness	50°C initial, hold 1 min; 8°C/min to 300°C.	Desideri *et al.*[5]
1985	fused silica capillary, DB-5 25m x 0.25 mm i.d. 0.25 μm film thickness	50°C initial, hold 2 min; 3°C/min to 300°C.	Wright *et al.*[6]

1. Lee, M.L.; Novotny, M.; Bartle, K.D. Gas chromatography/mass spectrometric and nuclear magnetic resonance determination of polynuclear aromatic hydrocarbons in airborne particles. Analytical Chemistry, 1976; 48(11): 1566-1572.
2. Lee, M.L.; Vassilaros, D.L.; White, C.M.; Novotny, M. Retention indices for programmed-temperature capillary-column gas chromatography of polycyclic aromatic hydrocarbons. Analytical Chemistry, 1979; 51(6): 768-773.
3. Later, D.W.; Lee, M.L.; Bartle, K.D.; Kong, R.C.; Vassilaros, D.L. Chemical class separation and characterization of organic compounds in synthetic fuels. Analytical Chemistry, 1981; 53: 1612-1620.
4. Lee, M.L.; Vassilaros, D.L.; Later, D.W. Capillary column chromatography of environmental polycyclic aromatic hydrocarbon compounds. International Journal of Environmental Chemistry, 1982; 11(3-4): 251-262.
5. Desideri, P.G.; Lepri, L.; Heimler, D.; Checchini, L.; Giannessi, S. Fingerprinting of crude oil spills. Journal of Chromatography, 1985; 322(1): 107-16.
6. Wright, C.W.; Later, D.W.; Wilson, B.W. Comparative chemical analysis of commercial creosotes and solvent refined coal-II material by high resolution gas chromatography. Journal of High Resolution Chromatography, Chromatography Communications, 1985; 8(6): 283-289.

Table 5.10 ***(continued)*** Selected gas chromatographic conditions used over the past 25 years to analyse for PAHs.

Year	Column (5% diphenyl, 95% dimethylpolysiloxane)	Temperature program	Authors
1991	Fused silica capillary, DB-5 30 m x 0.32 mm i.d. 0.25 µm film thickness	65°C initial, hold 5 min; 2°C/min to 300°C.	Bundt *et al.*[7]
1994	Fused silica capillary, HP-5 30 m x 0.25 mm i.d. 0.25 µm film thickness	90°C initial, hold 1 min; 25°C/min to 160°C; 8°C/min to 290°C, hold 15 min.	Wang *et al.*[8]
1995	fused silica capillary, HP-5 30 m x 0.25 mm i.d. 0.25 µm film thickness	50°C initial, hold 2 min; 6°C/min to 300°C, hold 16 min.	Wang and Fingas[9]
1995	fused silica capillary, DB-5 30 m x 0.25 mm i.d. 0.25 µm film thickness	40°C initial; 2, 4 or 6°C/min to 310°C.	Lai & Song[10]
2000	fused silica capillary, HP-5MS 30 m x 0.25 mm i.d. 0.25 µm film thickness	60°C initial; 10°C/min to 260°C; 1.5°C/min to 280°C; 40°C/min to 315, hold 9 min.	Zoccolillo *et al.*[11]

7. Bundt, J.; Herbel, W.; Steinhart, H.; Franke, S.; Francke, W. Structure-type separation of diesel fuels by solid phase extraction and identification of two- and three-ring aromatics by capillary gas chromatography-mass spectrometry. Journal of High Resolution Chromatography, 1991; 14: 91-98.
8. Wang, Z.; Fingas, M.; Li, K. Fractionation of a light crude oil and identification and quantitation of aliphatic, aromatic, and biomarker compounds by GC-FID and GC-MS, Part I. Journal of Chromatographic Science, 1994; 32: 361-366.
9. Wang, Z.; Fingas, M. Differentiation of the source of spilled oil and monitoring of the oil weathering process using gas chromatography-mass spectrometry. Journal of Chromatography A, 1995; 712: 321-343.
10. Lai, W.C. and Song, C.S. Temperature programmed retention indices for g.c. and g.c.-m.s. analysis of coal- and petroleum-derived liquid fuels. Fuel, 1995; 74(10): 1436-1451.
11. Zoccolillo, L.; Babi, D.; Felli, M. Evaluation of polycyclic aromatic hydrocarbons in gasoline by HPLC and GC-MS. Chromatographia, 2000; 52(5/6): 373-376.

---

aromatic hydrocarbons in crude oil and petroleum products. There is general agreement in the literature that the use of a slightly polar column (*i.e.* 5% phenyl, 95% dimethylsiloxane stationary phase), combined with a temperature program that uses a slow ramp rate, can be used to effect a good separation of a complex mixture of PAHs (Table 5.10).

5.4.2.3 Retention index systems for gas chromatography

A key component in data analysis is the consistent identification of each compound in the chromatogram. When compounds are consistently identified integration results can be extracted for the compounds of interest only. In gas chromatography, a retention index system is often employed to qualitatively identify compounds based on their retention behaviour relative to a series of known compounds. Kovats[60] originally developed the concept of the retention index in 1958 by comparing the retention behaviour of analytes to a series of n-alkanes in an isothermally programmed chromatogram. In 1963, Van den Dool and Kratz[48] applied Kovats' retention index system to temperature programmed chromatograms. In 1965, Kovats[61] described the use of a retention index system in more detail.

Just over 20 years after Kovats first described the retention index system, Lee *et al.*[47] reported that traditional retention indices employing a homologous series of n-alkanes had serious drawbacks when analysing for PAHs. Lee *et al.*[47] explored the effect of different film thicknesses on the reproducibility of retention indices for PAHs separated on a 5% diphenyl, 95% dimethylpolysiloxane stationary phase. Using the traditional n-alkane retention index system, they calculated a maximum variation of almost 20 index units for a PAH sample run on two columns having different film thicknesses. In response to this problem they developed a new retention index system based on selected PAH compounds. The standard index compounds used were naphthalene (200.00), phenanthrene (300.00), chrysene (400.00), and picene (500.00).[47] Lee *et al.*[47] used the equation published by Van den Dool and Kratz[48] to calculate their retention indices. This new retention index system gave a maximum variation of 0.14 index units for PAHs separated on two columns having different film thicknesses. Vassilaros *et al.*[62] explored the effects of initial column temperature, temperature program rate, and internal column diameter on the retention indices of 78 PAHs. They found the retention index system based on PAHs to be very stable and, therefore, a useful tool in the preliminary identification of PAHs from complex mixtures. This PAH retention index system has been used by Bundt *et al.*[23] and was the subject of a review.[63] The PAH retention index system was used in the work presented in this chapter and was found to give reproducible results for a large number of samples.

5.4.2.4 Data analysis by principal components

Once the data has been analysed by principal components, the individual samples can be grouped together (*i.e.* clustered) according to their chemical similarity. Grouping is done by comparing the new variables (*i.e.* PC scores) generated by PCA for each chromatogram (aliquot) and calculating the similarity of each chromatogram relative to all the others. One grouping method used was linear discriminant analysis (LDA). In linear discriminant analysis each group is defined prior to analysis. In this case 35 groups were constructed, one for each gasoline sample, and each group contained the five chromatograms for that sample. To classify a chromatogram as belonging to a particular gasoline sample (group) LDA examines the distance of the chromatogram from the centre (mean) of each pre-defined group. This distance, called the Mahalanobis distance, is the squared distance of the sample to the centre of a given group and is like a standard deviation in multidimensional space.[64] The chromatogram is assigned to a particular group for which the Mahalanobis distance is at a minimum. The shorter the Mahalanobis distance, the closer the sample is to the centre of the group. A cross-validation procedure is employed in order to minimise any bias when using LDA to classify samples. During cross-validation, the chromatogram being classified is excluded from the calculation of the group centroid prior to calculating the Mahalanobis distance of the chromatogram from each group. In this way the chromatogram being classified does not unnecessarily weight, or bias, the centroid to which it is being compared.

Another method for grouping similar chromatograms together is cluster analysis. Cluster analysis offers a very different approach from LDA. In LDA each sample group (cluster) is pre-defined and the object of the analysis is to compare the distance of a given chromatogram to each group and place the chromatogram into a particular group (classify it) based on which chromatogram-to-group distance is at a minimum. Cluster analysis takes the opposite approach to LDA in that the clustering of the samples is not pre-defined.

Cluster analysis begins by assuming that each chromatogram is its own separate group. The two groups (a pair of chromatograms) that are closest together are then joined into one group. The next step is to join the next pair of groups that are closest to one another. This step may join a third chromatogram to the first group containing two chromatograms, or it may be that two other groups, each containing one chromatogram, are joined to form a single group. The process continues until all chromatograms are joined into one group with the resulting graphical output being called a dendogram. There are several ways calculate which two groups are the closest. In this chapter, two different calculations were used for cluster analysis. In one case, a single linkage ("nearest neighbour") approach was used where the minimum Euclidean distance between a single chromatogram in one group and a single chromatogram in another group was calculated. In the other case a centroid linkage was used where the minimum Euclidean distance between the mean of one group and the mean of another group was calculated. Both approaches to cluster analysis placed the same gasoline samples into the same groups. The fact that cluster analysis requires the analyst to choose a partition point does introduce an element of subjectivity that is not found in linear discriminant analysis.

The work of Frank,[1] Chan[2] and Hirz[3] on differentiating gasoline samples based on their alkyl lead content, although out of date with the advent of unleaded gasoline, did demonstrate the value of examining gasoline for specific trace compounds. In a study of ten gasoline samples collected from the same city, Hennig[4] was able to differentiate all but two samples based on a comparison of the relative ratios of the higher boiling compounds (*i.e.* naphthalenes). Hennig[4] was also able to differentiate three samples, each from a different refinery, using this same comparison method. Similarly, Thatcher[5] found that it was possible to distinguish six gasoline samples, a regular and a premium each from three different refineries, based on the relative ratios of naphthalene, 1,2-dimethylnaphthalene and phenanthrene. More recently, Smallwood *et al.*[65] determined the $\delta^{13}C$ value for 16 compounds present in 19 unevaporated gasoline samples using gas chromatography-carbon isotope ratio mass spectrometry. Preliminary results suggested that it was possible to differentiate all of the samples, however, a sufficient number of replicate analyses was not always made for each sample. Smallwood *et al.*[65] noted that certain compounds, such as naphthalene and the C_1-naphthalenes, exhibited a wide range of carbon isotope values for the 19 samples studied. Given that seven samples were from the south central United States (four samples each from a different city in the state of Texas, and three samples from one city in the neighbouring state of Oklahoma) and twelve were from the eastern coast of the United States (state of New York), many of the samples must have originated from different refineries. In light of this information it is not too surprising that statistically significant differences were found between samples. The findings reported in this chapter confirm that trace compounds, and in particular the polycyclic aromatic hydrocarbons, can be used to differentiate gasoline samples. More importantly, the ability to differentiate all but two samples (RU15 and RU35, Table 5.9) out of 29 collected in a seven week period (March 13 to April 28, 2001; Table 5.2), with all samples being from service stations within a 7 km radius of each other in metropolitan Sydney, has demonstrated the value of characterising gasoline samples using the C_0- to C_2-naphthalenes.

Mann[66] compared twelve gasoline samples collected in one week (eight samples from metropolitan Seattle, four from outside the metropolitan area). The comparison relied on changes in the relative ratios of the more volatile, non-trace level organic components in the gasoline.[66] Mann was able to distinguish all twelve samples using this method. Mann[66] also collected ten samples of gasoline, one each from ten different service stations that each received a delivery of gasoline from the same distribution terminal on the same day. Mann found that all ten samples could be differentiated, however, the variation between these samples was small. Differences between gasoline samples collected from different service stations is likely a result of two processes: batch to batch differences that occur at the refinery as different petroleum streams are blended to create the final product, and mixing of different blends over time in the underground storage tank at the service station.[2,66] In this chapter it was found that five samples of gasoline collected from four service stations on one day (April 18, 2001) could be distinguished from one another, and 17 samples of gasoline collected from 14 service stations on another day (April 24, 2001) could also be distinguished from one another.

5.5 Conclusions

The use of trace compounds to differentiate one gasoline sample from another has been explored. An investigation to extract and identify trace polar compounds and polycyclic aromatic hydrocarbons in liquid gasoline samples has been conducted. In this study, a two-step process was undertaken to evaluate the potential of these two compound classes to differentiate samples of gasoline. In the first step, a novel extraction method using alumina was developed to separate trace amounts of both polar compounds and polycyclic aromatic hydrocarbons (PAHs) from the gasoline matrix. The intent of the separation by solid phase extraction prior to analysis by gas chromatography was to ensure that all potential analytes from each class were recovered. Both polar and PAH extracts were analysed by gas chromatography-mass spectrometry (GC-MS), and the polar extracts were analysed without further treatment (*i.e.* no chemical derivatisation of the phenolic compounds).

It was found that the polar compound composition did not vary significantly from one sample to another, and so these compounds were found to be unsuitable for use in differentiating gasoline samples. Solid phase extraction of PAHs from different gasoline samples revealed that PAH composition did vary from one gasoline sample to another. A GC-MS (SIM) method was developed to directly analyse for both two- and three-ring PAH compounds, however, precision in detecting lower concentrations of the three-ring PAHs was not as high as for two-ring PAHs. This resulted in the development of a one-step GC-MS (SIM) method that could determine precisely the relative amounts of the C_0- to C_2-naphthalenes in unevaporated gasoline. This study showed that, in conjunction with PCA and LDA, these two-ring polycyclic aromatic hydrocarbons can be used to distinguish between different gasoline samples.

5.6 References

1. Frank, H.A. Lead alkyl components as discriminating factors in the comparison of gasolines. Journal of the Forensic Science Society, 1980; 20(4): 285-292.

2. Chan, L. The determination of tetraalkyl lead compounds in petrol using combined gas chromatography atomic absorption spectrometry. Forensic Science International, 1981; 18: 57-62.

3. Hirz, R. Gasoline brand identification and individualization of gasoline lots. Journal of the Forensic Science Society, 1989; 29(2): 91-101.

4. Hennig, H.J. Möglichkeiten der differenzierung von Mineralölen, Gasölen und Vergaserkraftstoffen mit hilfe der gaschromatographie. Archive Fur Kriminologie, 1982; 170: 12-20.

5. Thatcher, P.J. The Identification of Petroleum Residues in Arsons. Thesis (Ph.D.) University of Melbourne, 1982.

6. Gibbs, L.M. Gasoline additives - when and why. SAE Paper 90204, 1990.

7. Mansfield, C.T.; Barman, B.N.; Thomas, J.V.; Mehrotra, A.K.; Philip, R.P. Petroleum and coal. Analytical Chemistry, 1997; 69(12): 59R-93R.

8. Mansfield, C.T.; Barman, B.N.; Thomas, J.V.; Mehrotra, A.K.; McCann, J.M. Petroleum and coal. Analytical Chemistry, 1999; 71(12): 81R-107R.

9. Barman, B.N.; Cebolla, V.L.; Mehrotra, A.K.; Mansfield, C.T. Petroleum and coal. Analytical Chemistry, 2001; 73(12): 2791-2804.

10. Tebbett, I. (Editor). Gas Chromatography in Forensic Science. Ellis Horwood, New York, 1992.

11. Bertsch, W. and Holzer, G. Analysis of accelerant by gas chromatography/mass spectrometry. In: Forensic Applications of Mass Spectrometry. Yinnon, J. (Editor), CRC Press: Boca Raton, Florida; 1995.

12. Colgrove, S.G. and Svec, H.J. Liquid-liquid fractionation of complex mixtures of organic components. Analytical Chemistry, 1981; 53: 1737-1742.

13. Di Sanzo, F.P. Characterization of basic nitrogen compounds in gasolines by GC-MS. Journal of High Resolution Chromatography and Chromatography Communications, 1981; 4(12): 649-651.

14. Dinh, H.T.; Mushrush, G.W.; Beal, E.J. Determination of nitrogen compound distribution from three source fuels. Petroleum Science & Technology, 1999; 17(3-4): 383-427.

15. Youngless, T.L.; Swansiger, J.T.; Danner, D.A.; Greco, M. Mass spectrometry characterization of petroleum dyes, tracers, and additives. Analytical Chemistry, 1985; 57(9): 1894-1902.

16. Kanai, H.; Yazawa, L.; Maka, J. Baseline data on the quality of gasolines in Hawaii and their implication in preparing alcohol blends. Hawaii State Department of Agriculture, Honolulu, HI, 1989.

17. Kanai, H.; Inouye, V.; Goo, R.; Yazawa, L.; Maka, J.; Chun, C. GC/MS analysis of polar compounds in "weathered" gasoline/water matrix as an aid in identifying gasoline. Analytical Letters 1991; 24(1): 115-28.

18. Ioppolo, M.; Alexander, R.; Kagi, R.I. Identification and analysis of C_0-C_3 phenols in some Australian crude oils, Organic Geochemistry, 1992; 18(5): 603-609.

19. Ioppolo-Armanios, M.; Alexander, R.; Kagi, R. Geosynthesis of organic compounds: I. Alkylphenols. Geochimica et Cosmochimica Acta, 1995; 59(14): 3017-3027.

20. Taylor, R.; Larter, S.; Jones, M.; Dale, J.; Horstad, I. The effect of oil-water-rock partitioning on the occurrence of alkylphenols in petroleum systems. Geochimica et Cosmochimica Acta, 1997; 61(9): 1899-1910.

21. Bennett, B.; Fowler, B.F.J.; Larter, S.R. Determination of C_0-C_3 alkylphenols in crude oils and waters. Analytical Chemistry, 1996; 68(20): 3697-3702.

22. Mach, M.H. Gas chromatography-mass spectrometry of simulated arson residue using gasoline as an accelerant. Journal of Forensic Sciences, 1977; 22: 348-357.

23. Bundt, J.; Herbel, W.; Steinhart, H.; Franke, S.; Francke, W. Structure-type separation of diesel fuels by solid phase extraction and identification of two- and three-ring aromatics by capillary gas chromatography-mass spectrometry. Journal of High Resolution Chromatography, 1991; 14: 91-98.

24. Song, C. and Hatcher, P.G. Compositional differences between coal- and petroleum-derived jet fuels. Preprints / Division of Petroleum Chemistry, American Chemical Society, 1992; 37: 529-539.

25. Lai, W.C. and Song, C.S. Temperature programmed retention indices for g.c. and g.c.-m.s. analysis of coal- and petroleum-derived liquid fuels. Fuel, 1995; 74(10): 1436-1451.

26. Wang, Z. and Fingas, M. Differentiation of the source of spilled oil and monitoring of the oil weathering process using gas chromatography-mass spectrometry. Journal of Chromatography A, 1995; 712: 321-343.

27. Wang, Z.D.; Fingas, M.; Page, D.S. Oil spill identification. Journal of Chromatography, 1999; 843(1-2): 369-411.

28. Radke, M.; Garrigues, P.; Willsch, H. Methylated dicyclic and tricyclic aromatic hydrocarbons in crude oils from the Handil field, Indonesia. Organic Geochemistry, 1990; 15(1): 17-34.

29. Wang, Z. and Fingas, M. Use of methyldibenzothiophenes as markers for differentiation and identification of crude and weathered oils. Environmental Science & Technology, 1995; 29(11): 2842-2849.

30. Bence, A.E.; Kvenvolden, K.A.; Kennicutt, M.C. Organic geochemistry applied to environmental assessments of Prince William Sound, Alaska, after the Exxon Valdez oil spill – a review. Organic Geochemistry, 1996; 24(1): 7-42.

31. Quimby, B.D.; Giarrocco, V.; Sullivan, J.J.; McCleary, K.A. Fast analysis of oxygen and sulfur compounds in gasoline by GC-AED. Journal of High Resolution Chromatography, 1992; 15: 705-709.

32. Di Sanzo, F.P.; Bray, W.; Chawla, B. Determination of the sulphur components of gasoline streams by capillary column gas chromatography with sulphur chemiluminescence detection. Journal of High Resolution Chromatography, 1994; 17(4): 255-258.

33. Coulombe, R. Chemical markers in weathered gasoline. Journal of Forensic Sciences, 1995; 40(5): 867-873.

34. Chen, Y.C. and Lo, J.G. Gas chromatography with flame ionization and flameless sulphur chemiluminescence detectors in series for dual channel detection of sulphur compounds. Chromatographia, 1996; 43(9-10): 522-526.

35. World-Wide Fuel Charter. Alliance of Automobile Manufacturers, Washington, D.C.; April, 2000.

36. Snyder, L.R. Principles of Adsorption Chromatography: The Separation of Nonionic Organic Compounds. New York: Marcel Dekker, Inc.; 1968.

37. Later, D.W.; Wilson, B.W.; Lee, M.L. Standardization of alumina and silica adsorbents used for chemical class separation of polycyclic aromatic compounds. Analytical Chemistry, 1985; 57: 2979-2984.

38. Rietjens, M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytical Chemistry, 1995; 316: 205-215.

39. Synder, L.R. Linear elution adsorption chromatography. V. Silica as adsorbent. Adsorbent standardization. Journal of Chromatography, 1963; 11(2): 195-227.

40. Snyder, L.R. Adsorption. In: Chromatography. Heftmann, E., Editor. Reinhold Publishing Corp., New York. 2nd edition; 1967.

41. Snyder, L.R. Linear elution adsorption chromatography. XII. Functional group adsorption energies on the metal oxide adsorbents. Journal of Chromatography, 1966; 23: 388-402.

42. Snyder, L.R. and Buell, B.E. Compound type separation and classification of petroleum by titration, ion exchange, and adsorption. An index of the acidity, basicity and adsorptivity on alumina of various petroleum compound types. Journal of Chemical and Engineering Data, 1966; 11: 545-553.

43. Green, J.B.; Yu, S.K.T.; Vrana, R.P. GC-MS analysis of phenolic compounds in fuels after conversion to trifluoroacetate esters. Journal of High Resolution Chromatography, 1994; 17(6): 439-451.

44. Li, D.; Park, J.; Oh, J-R. Silyl derivatization of alkylphenols, chlorophenols, and bisphenol A for Simultaneous GC/MS determination. Analytical Chemistry, 2001; 73: 3089-3095.

45. Bertsch, W.; Holzer, G.; Sellers, C.S. Laboratory Methods In: Chemical Analysis for the Arson Investigator and Attorney. Hüthig Buch Verlag GmbH, Heidelberg; 1993.

46. J&W Scientific. 1996/97 Catalogue and Technical Reference: GC, SPE, CE, LC. Folsom, CA, USA; 1996.

47. Lee, M.L.; Vassilaros, D.L.; White, C.M.; Novotny, M. Retention indices for programmed-temperature capillary-column gas chromatography of polycyclic aromatic hydrocarbons. Analytical Chemistry, 1979; 51(6): 768-773.

48. Van den Dool, H. and Kratz, P.D. A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. Journal of Chromatography, 1963; 11: 463-471.

49. Snyder, L.R. Applications of linear elution adsorption chromatography to the separation and analysis of petroleum. I. Compound class separation over alumina and silica. Analytical Chemistry, 1961; 33(11): 1527-1534.

50. Snyder, L.R. Applications of linear elution adsorption chromatography to the separation and analysis of petroleum. II. Compound class separations by a routine micro procedure. Determination of gasoline polyaromatics. Analytical Chemistry, 1961; 33(11): 1535-1538.

51. Klemm, L.H.; Reed, D.; Miller, L.A.; Ho, B.T. Chemical structure and chromatographic adsorptivity of aromatic hydrocarbons on alumina. Journal of Organic Chemistry, 1959; 24: 1468-1477.

52. Snyder, L.R. Adsorption from solution. I. Fused aromatic hydrocarbons on alumina. Journal of Physical Chemistry, 1963; 67: 234-240.

53. Later, D.W.; Lee, M.L.; Bartle, K.D.; Kong, R.C.; Vassilaros, D.L. Chemical class separation and characterization of organic compounds in synthetic fuels. Analytical Chemistry, 1981; 53: 1612-1620.

54. Lee, M.L.; Novotny, M.; Bartle, K.D. Gas chromatography/mass spectrometric and nuclear magnetic resonance determination of polynuclear aromatic hydrocarbons in airborne particles. Analytical Chemistry, 1976; 48(11): 1566-1572.

55. Lee, M.L.; Vassilaros, D.L.; Later, D.W. Capillary column chromatography of environmental polycyclic aromatic hydrocarbon compounds. International Journal of Environmental Chemistry, 1982; 11(3-4): 251-262.

56. Desideri, P.G.; Lepri, L.; Heimler, D.; Checchini, L.; Giannessi, S. Fingerprinting of crude oil spills. Journal of Chromatography, 1985; 322(1): 107-16.

57. Wang, Z.; Fingas, M.; Li, K. Fractionation of a light crude oil and identification and quantitation of aliphatic, aromatic, and biomarker compounds by GC-FID and GC-MS, Part I. Journal of Chromatographic Science, 1994; 32: 361-366.

58. Wang, Z.; Fingas, M.; Li, K. Fractionation of a light crude oil and identification and quantitation of aliphatic, aromatic, and biomarker compounds by GC-FID and GC-MS, Part II. Journal of Chromatographic Science, 1994; 32: 367-382.

59. Zoccolillo, L.; Babi, D.; Felli, M. Evaluation of polycyclic aromatic hydrocarbons in gasoline by HPLC and GC-MS. Chromatographia, 2000; 52(5/6): 373-376.

60. Kovats, E. Gas chromatographic characterization of organic compounds. I. Retention indexes of aliphatic halides, alcohols, aldehydes, and ketones. Helvetica Chimica Acta, 1958; 41: 1915-1932.

61. Kovats, E. Gas chromatographic characterization of organic substances in the retention index system. Advances in Chromatography, 1965; 1: 229-247.

62. Vassilaros, D.L.; Kong, R.C.; Later, D.W.; Lee, M.L. Linear retention index system for polycyclic aromatic hydrocarbons. Critical evaluation and additional indexes. Journal of Chromatography, 1982; 252: 1-20.

63. Castello, G. Retention index systems: alternatives to the n-alkanes as calibration standards. Journal of Chromatography A, 1999; 842: 51-64.

64. Garson, G.D. PA 765 Statnotes: An Online Textbook. North Carolina State University, USA. <http://www2.chass.ncsu.edu/garson/pa765/statnote.htm> Accessed August 12, 2002.

65. Smallwood, B.J.; Philip, R.P.; Allen, J.D. Stable carbon isotope composition of gasolines determined by isotope ratio monitoring gas chromatography mass spectrometry. Organic Geochemistry, 2002; 33: 149-159.

66. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography I: comparison methodology. Journal of Forensic Sciences, 1987; 32(3): 606-615.

Chapter 6. Study of unevaporated gasoline over time

6.1 Introduction

The significance of a forensic comparison between two gasoline samples is dependent on an understanding of how gasoline composition changes with time. The degree to which gasoline samples collected from a single service station change with time can provide a court of law with important information that will allow the court to assess the significance of the gasoline comparison. If, for example, it can be demonstrated that gasoline samples taken from a single service station can be differentiated from week to week then the court may place greater evidential weight on a comparison between two samples found to be similar. This fact was recognized by Mann.[1] Mann collected six samples from a single service station over a period of 28 days. The results of Mann's work showed that, as long as the distribution terminal received new shipments of gasoline from the refinery, differences between deliveries to a single service station did exist. To this author's knowledge, no other published work has examined the change in gasoline composition with time at a single service station.

In this chapter the work of Mann[1] is expanded and extended. The research presented here expands Mann's original study from one service station to three (each one representing a different brand of station), and extends the collection time from four weeks to 16 weeks. The aim of this chapter is to gain a better understanding of how the C_0- to C_2-naphthalene composition in gasoline changes with time at the service station and how it differs between stations over time.

6.2 Materials and methods

6.2.1 Samples

Ninety-six samples of gasoline were obtained from three service stations in metropolitan Sydney, Australia over a 16 week period (January 22 to May 7, 2002). The three service stations from which samples were collected, BP, Caltex, and Shell, represented the three major brands in Sydney. A deliberate choice was made to take samples only from branded service stations, that is, company owned or franchised stations, and not from independent stations or independent franchises such as Solo or Red Dot, nor from convenience store/grocery retailers such as 7-11 or Woolworth's ("Woolies"). This choice was made to ensure, as much as possible, that the gasoline delivered each week to the station was consistently coming from the same local refinery and through the same distribution network. Permission to obtain weekly samples for an extended period of time was obtained from each service station manager or owner.

Each week one regular unleaded and one premium unleaded gasoline sample was collected from each station (*i.e.* six samples per week). With the exception of the first week, all samples were collected on Tuesday. During the first week the samples from Shell were collected on Tuesday, the samples from Caltex on Wednesday, and the samples from BP on Thursday. The different collection days resulted from the Caltex and BP stations awaiting approval from their respective head offices before allowing sample collection. Tuesday was selected as sample collection day after consultation with the manager or owner of each service station because they all indicated that they routinely received deliveries on Saturday or Monday or both. Collecting samples on Tuesday provided the greatest assurance that at least one delivery had occurred between the current sample and the sample taken the previous week. It was important to consult with each station manager or owner prior to sample collection because it was found that at one station (Shell), all but one regular unleaded and one premium unleaded gasoline pump could be switched between more than one storage tank. In this situation it was decided to collect samples from the two pumps that could not be switched between storage tanks. As a precaution, each sample collected from each service station was always taken from the same pump over the 16 week period. Samples were collected in 125 mL amber, Boston round glass bottles sealed with polypropylene caps fitted with teflon-faced foamed polyethylene liners (Wheaton) and stored in the dark at room temperature.

6.2.2 Gas chromatography-mass spectrometry

Gas chromatography was performed on a 30 m x 0.25 mm i.d. x 0.25 μm film HP5-ms capillary column using an Agilent 6890 gas chromatograph connected to an Agilent 5973 mass selective detector (MSD). Chromatographic conditions were as follows: split injection (15:1) at 280 ^{o}C; temperature programmed from 40 ^{o}C (3 minutes isothermal) to 105 ^{o}C at 2 ^{o}C/minute, ramped to 290 ^{o}C at 20 ^{o}C/minute (held for 5 minutes); helium carrier gas held at a constant flow rate of 1.2 mL/minute. The sample injection volume was 1 μL.

Samples were prepared for GC-MS (SIM) by diluting an aliquot of neat gasoline with dichloromethane (1:4, v/v). The detector was set to monitor the following ion groups from the two-ring polycyclic aromatic hydrocarbons as follows (see section 5.2.5): m/z^+ 127 and 129 (group 1; C_0-naphthalene); m/z^+ 139 and 143 (group 2; C_1-naphthalenes); m/z^+ 128, 141, 142, 155, 156 and 157 (group 3; C_2-naphthalenes); and, m/z^+ 128, 141, 153, 155, 169, 170 and 171 (group 4; C_3-naphthalenes). Dwell times were 100 ms for groups 1 and 2, and 50 ms for groups 3 and 4. The PAH isomer group and its corresponding ion fragments are given in Table 6.1. In all cases five aliquots from each sample were analysed.

Table 6.1 Selected ion fragments for each isomeric group of the two-ring polycyclic aromatic hydrocarbons found in unleaded gasoline samples.

Two-ring PAHs	m/z^+
C_0-naphthalene	127, 128, 129
C_1-naphthalenes	139, 141, 142, 143
C_2-naphthalenes	127, 128, 141, 142, 155, 156, 157
C_3-naphthalenes	128, 141, 142, 155, 156, 169, 170, 171

6.2.3 Data analysis

Peak area information was extracted from the chromatograms using a macro. The macro is reproduced in Appendix 5.1. The integration results were written to an ASCII text file in a format suitable for export to a spreadsheet. Data was loaded and parsed into a spreadsheet (Microsoft Excel®). The data was normalised using a natural logarithmic transformation.[2] Principal component analysis and linear discriminate analysis with cross-validation were performed using Minitab® (version 13.1 for Windows).

6.3 Analysis of samples by brand and fuel type

A large volume of data was generated from the analysis of the 96 gasoline samples. The analysis of five aliquots for each of the 96 samples resulted in 480 chromatograms. Eleven naphthalene peaks (Table 5.7 and Figure 5.1) were extracted from each chromatogram resulting in an 11 by 480 data matrix. In order to simplify the initial exploration of the data and the reporting of results, the larger data set was divided according to brand (BP, Caltex or Shell) and fuel grade (regular unleaded or premium unleaded).

6.3.1 Results from the analysis of BP regular unleaded and premium unleaded gasoline

Sixteen premium unleaded and sixteen regular unleaded gasoline samples were collected from the BP service station. Principal component analysis (PCA) was performed on these 32 samples (5 aliquots per sample) and the results are given in Table 6.2 and Figure 6.1. The first three PCs explain 96.7% of the variation in the data set. In Figure 6.1 it can be seen that the regular and premium samples are randomly distributed with no clustering by fuel grade and no trend according to week.

Table 6.2 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): BP regular unleaded and premium unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	6.1436	0.559	55.9
PC2	3.4139	0.310	86.9
PC3	1.0752	0.098	96.7
PC4	0.1837	0.017	98.3
PC5	0.0952	0.009	99.2

Using the first three PC scores, linear discriminant analysis (with cross-validation) correctly classified all 160 chromatograms (100%) by sample. That is, there are 32 unique groups where each group contains only the five aliquots from one gasoline sample. This means that the variation between the five aliquots is smaller that the variation between samples (groups).

It can be seen in Figure 6.1 that some sample groups are very close to one another while other groups are well separated. In order to explore the data in more detail the data set was divided according to fuel grade and re-analysed by PCA. The PCA results for the 16 BP premium samples and the 16 BP regular samples are given in Tables 6.3 and 6.4, respectively. Principal components score plots are shown in Figure 6.2.

Table 6.3 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): BP premium unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	5.5605	0.505	50.5
PC2	3.8263	0.348	85.3
PC3	1.2974	0.118	97.1
PC4	0.1465	0.013	98.5
PC5	0.1218	0.011	99.6

Figure 6.1 PC score plot of BP premium unleaded and regular unleaded gasoline samples removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.

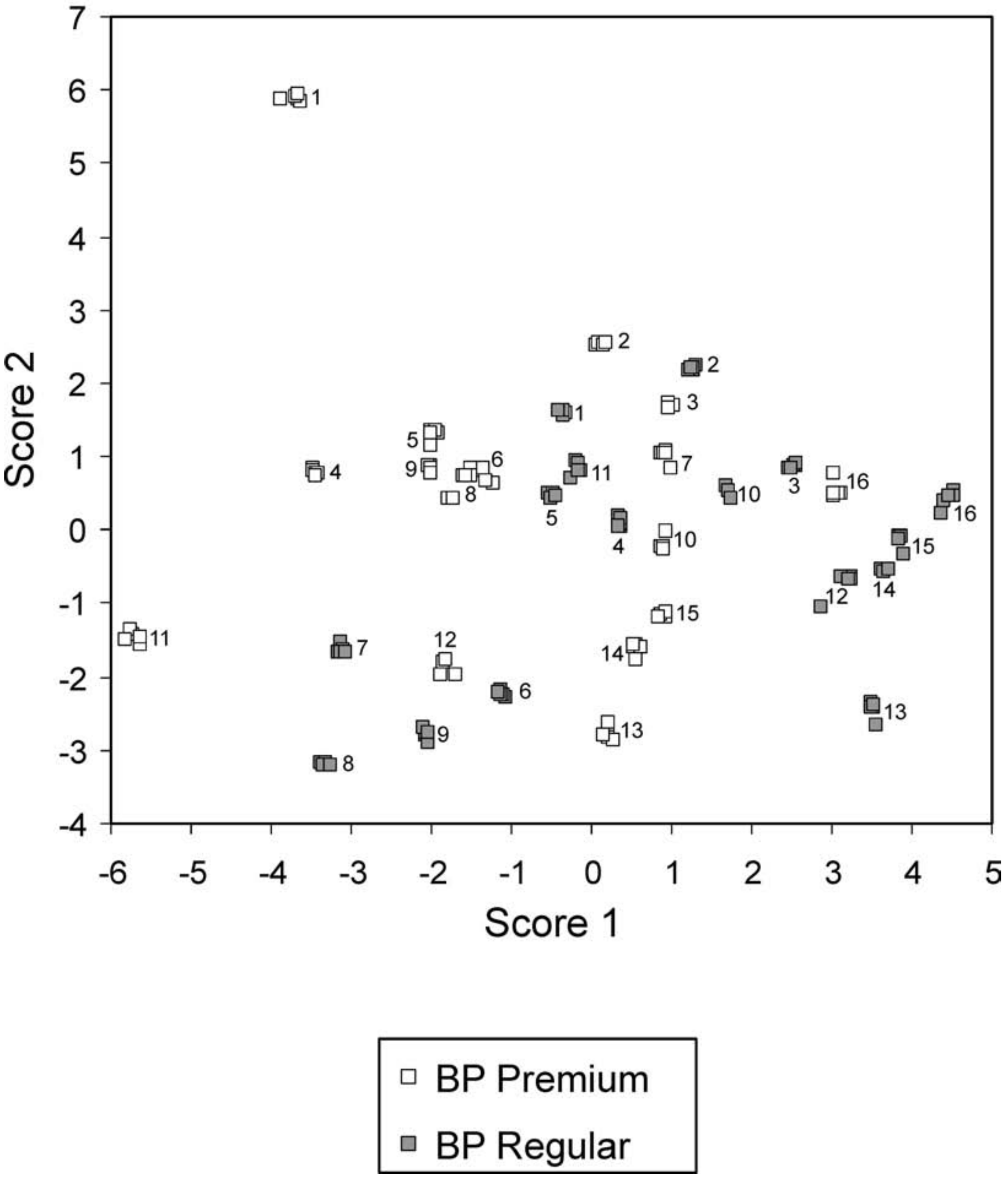


Figure 6.2 PC score plot of BP premium unleaded and regular unleaded gasoline samples. Each fuel type (regular or premium) was removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.

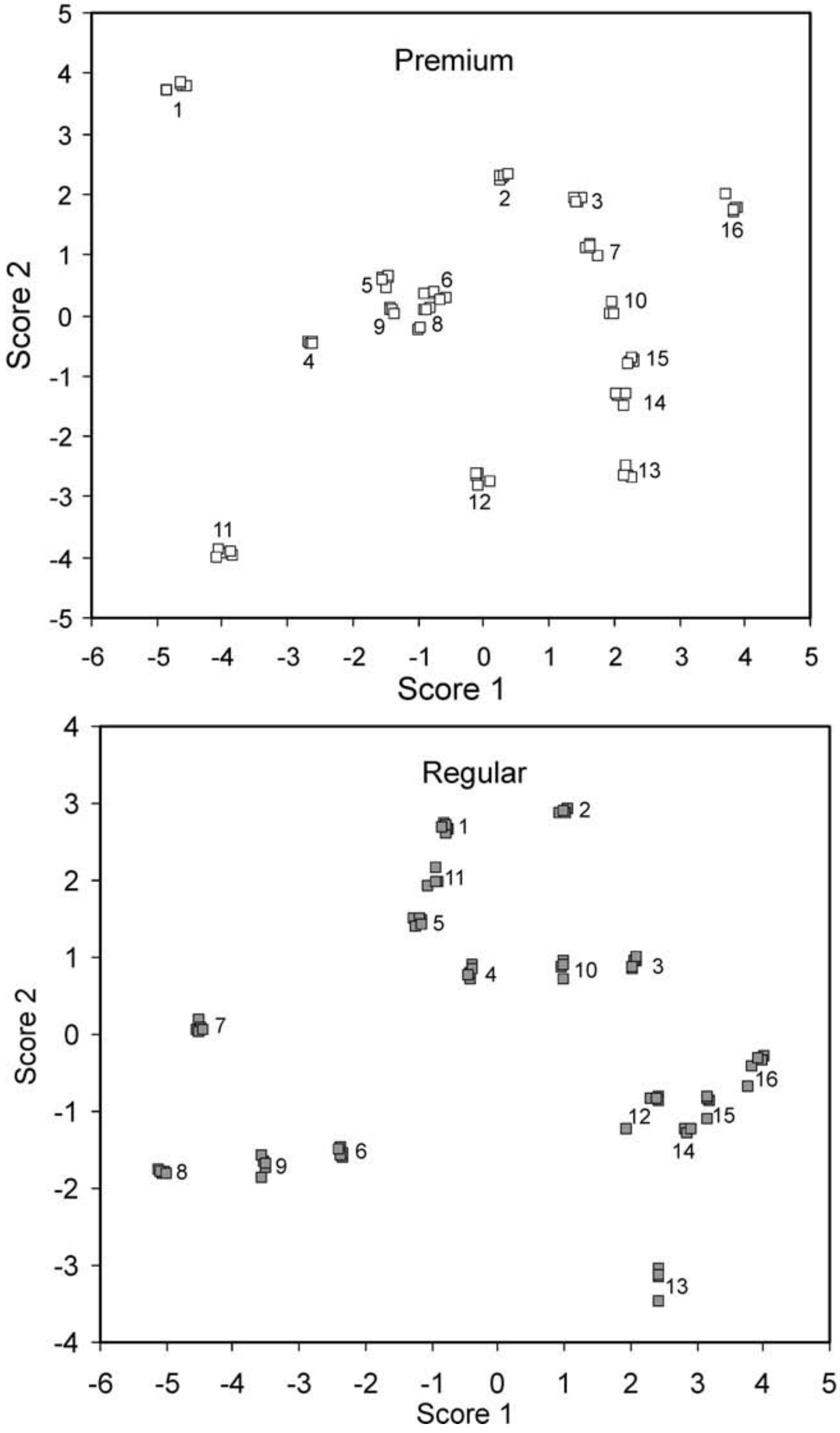


Table 6.4 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): BP regular unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	7.4733	0.679	67.9
PC2	2.9216	0.266	94.5
PC3	0.3621	0.033	97.8
PC4	0.1636	0.015	99.3
PC5	0.0342	0.003	99.6

The first two PCs for the regular samples account for nearly 10% more variation than do the first two PCs for the premium samples (94.5% versus 85.3%). However, the first three PCs for both the regular and premium samples account for nearly the same amount of variation in the data (>97%). By including the third PC, and calculating the third score, three-dimensional plots of the PCA results can be obtained for the premium gasoline samples (Figure 6.3) and the regular gasoline samples (Figure 6.4). Not surprisingly, linear discriminant analysis performed on the first three scores resulted in the correct classification of all premium aliquots and all regular aliquots.

Figure 6.3 PC score plots of BP premium unleaded gasoline samples. These samples were removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.

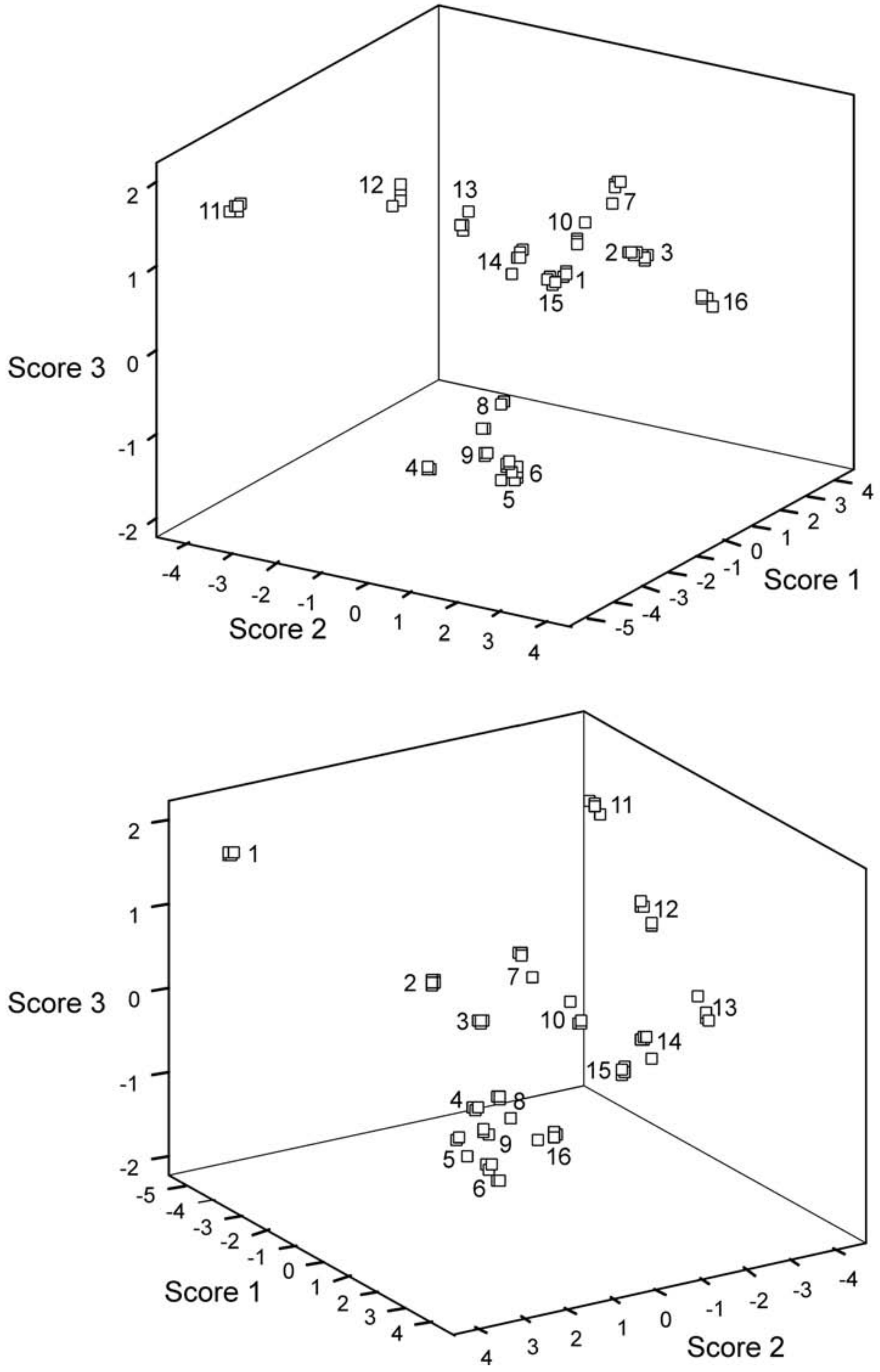
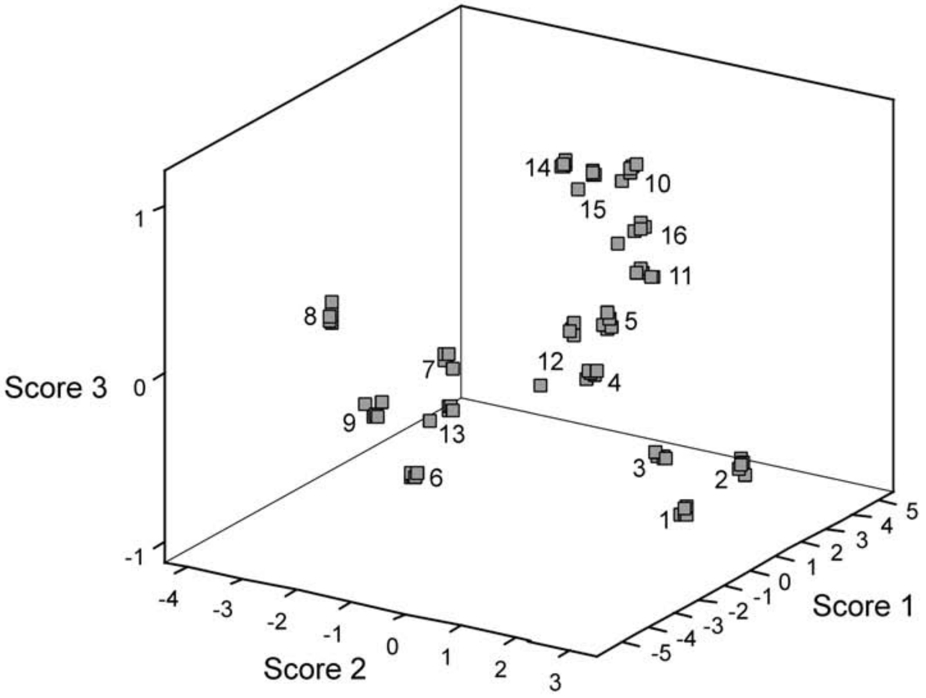
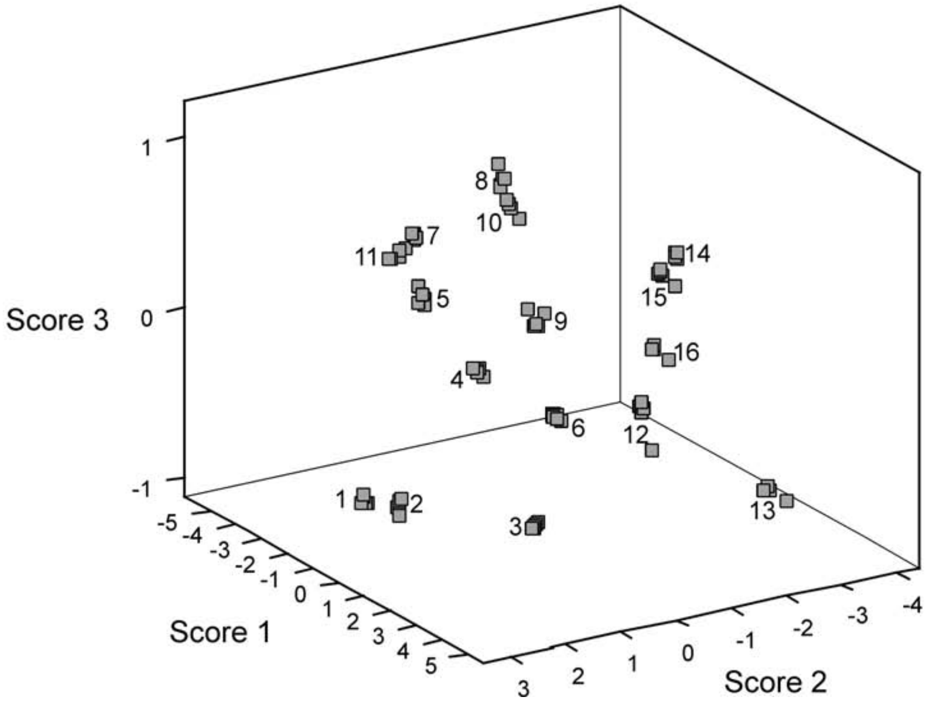


Figure 6.4 PC score plots of BP regular unleaded gasoline samples. These samples were removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.





6.3.2 Results from the analysis of Caltex regular unleaded and premium unleaded gasoline

Principal component analysis was performed on the 32 Caltex gasoline samples (16 premium unleaded and 16 regular unleaded) and the results are given in Table 6.5 and Figure 6.5. The first three PCs for the Caltex samples accounted for 97.8% of the variation in the data set, slightly more variation than the BP data set (96.7%, Table 6.2). However, the first two PCs for the Caltex data account for 9.7% more variation than the first two PCs calculated from the BP data (96.6% versus 86.9%). The first three PCs were retained for calculating the respective scores for the Caltex samples. In Figure 6.5 it can be seen that the regular and premium samples are randomly distributed with no clustering by fuel grade and no trend according to week.

Table 6.5 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): Caltex regular unleaded and premium unleaded samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	5.7860	0.526	52.6
PC2	4.8413	0.440	96.6
PC3	0.1272	0.012	97.8
PC4	0.0996	0.009	98.7
PC5	0.0615	0.006	99.2

Using the first three PC scores, linear discriminant analysis (with cross-validation) correctly classified 156 of the 160 chromatograms (97.5%) by sample. It was found that the 32 samples could be divided into 29 unique groups. Twenty-six of these groups each contained a single sample (*i.e.* five aliquots from one sample per group). The remaining six samples were divided between the three remaining groups as follows: Caltex premium weeks 1 and 2; Caltex regular week 5 and premium week 9; and, Caltex regular weeks 6 and 9. The number of misclassified aliquots are given in Table 6.6.

Figure 6.5 PC score plot of Caltex premium unleaded and regular unleaded gasoline samples removed from the larger data matrix and treated independently.

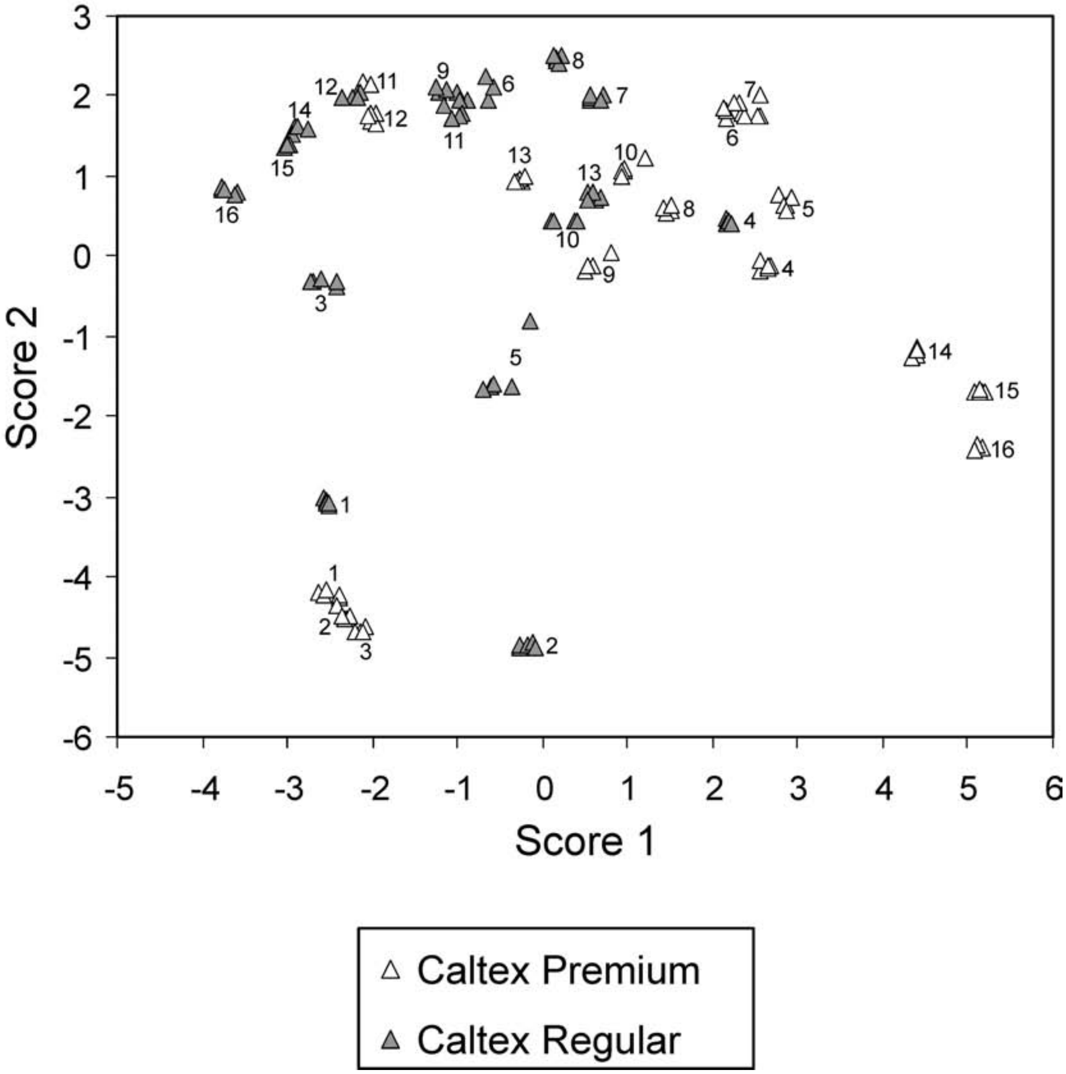


Table 6.6 Results of LDA with cross-validation: number of misclassified aliquots of Caltex premium unleaded and regular unleaded gasoline from the six misclassified samples.

		True Group			
		Caltex premium week 2	Caltex regular week 5	Caltex regular week 6	Caltex regular week 9
Predicted Group	Caltex premium week 1	1			
	Caltex premium week 9		1		
	Caltex regular week 9			1	
	Caltex regular week 6				1

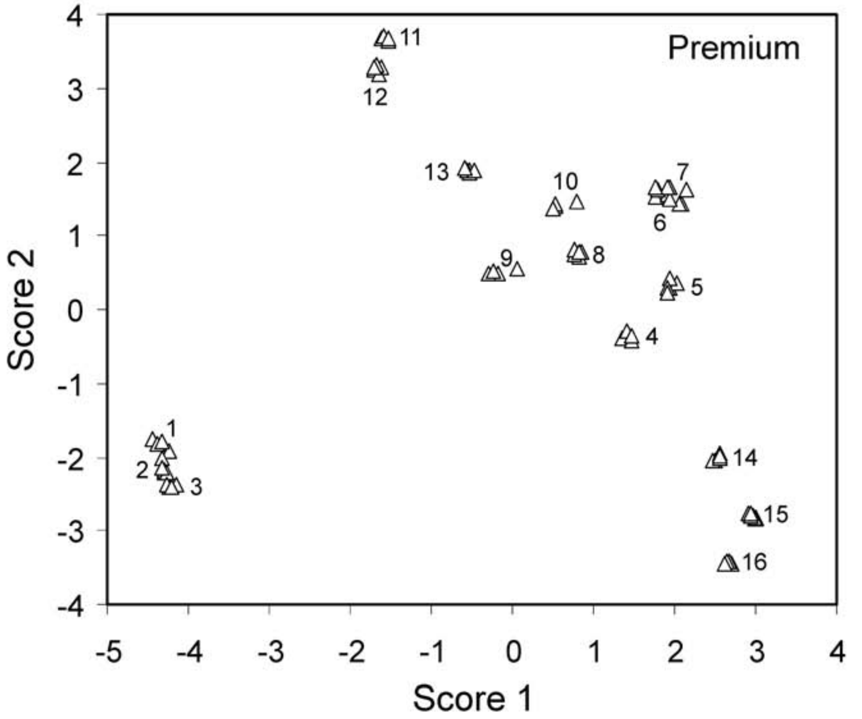
The Caltex data was explored in more detail by dividing it according to fuel grade. The PCA results for the 16 Caltex premium samples and the 16 Caltex regular samples are given in Tables 6.7 and 6.8, respectively; PC score plots are shown in Figure 6.6.

Table 6.7 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): Caltex premium unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	6.1414	0.558	55.8
PC2	4.6303	0.421	97.9
PC3	0.0891	0.008	98.7
PC4	0.0675	0.006	99.3
PC5	0.0333	0.003	99.7

Using the first three PC scores, linear discriminant analysis (with cross-validation) of the premium gasoline samples correctly classified all 80 chromatograms (100%) by sample. The three-dimensional plots of the first three scores for the premium samples is shown in Figure 6.7. Using the first three PC scores, linear discriminant analysis (with cross-validation) of the regular gasoline samples correctly classified 78 of the 80 chromatograms (97.5%) by sample. The 16 regular samples could be divided into 15 unique groups. Fourteen of these groups each contained a single sample. The remaining group contained two samples (weeks 6 and 9), the result of one aliquot from each sample being misclassified with the other (Table 6.9).

Figure 6.6 PC score plots of Caltex premium unleaded and regular unleaded gasoline samples. Each fuel type (regular or premium) was removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.



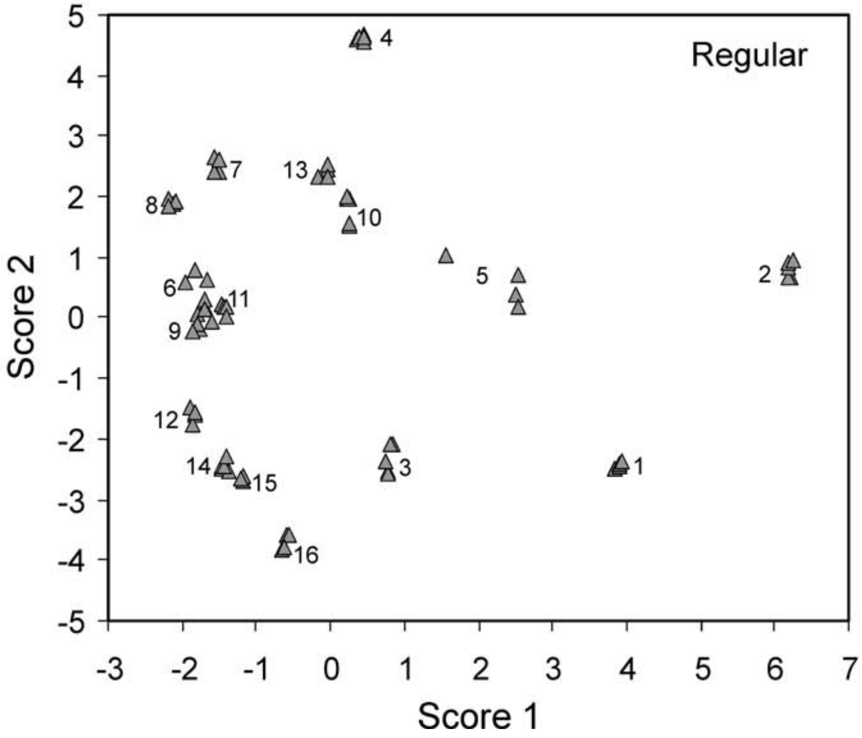


Figure 6.7 PC score plots of Caltex premium unleaded gasoline samples. These samples were removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.

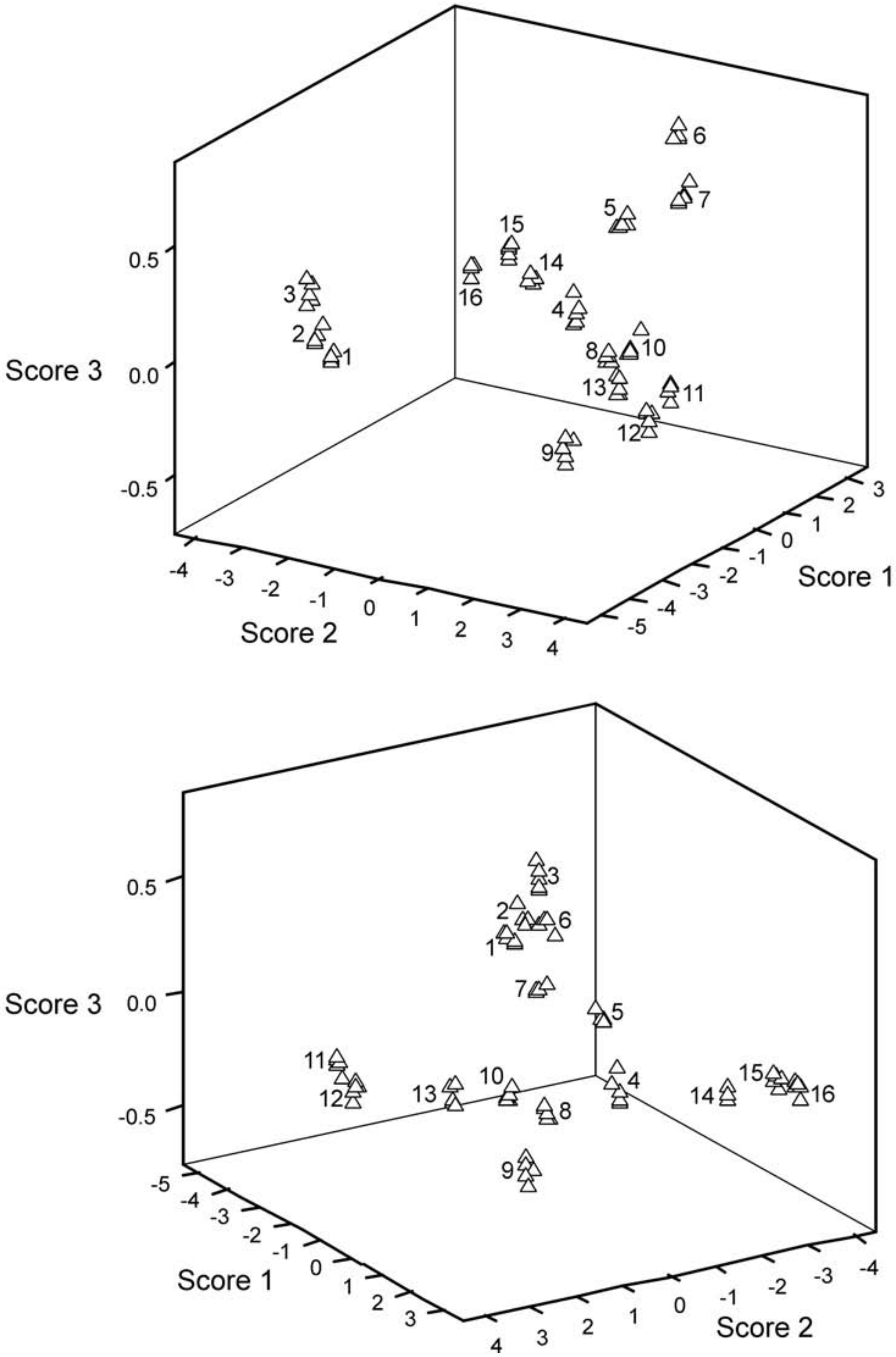


Table 6.8 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): Caltex regular unleaded gasoline samples collected over a 16 week period.

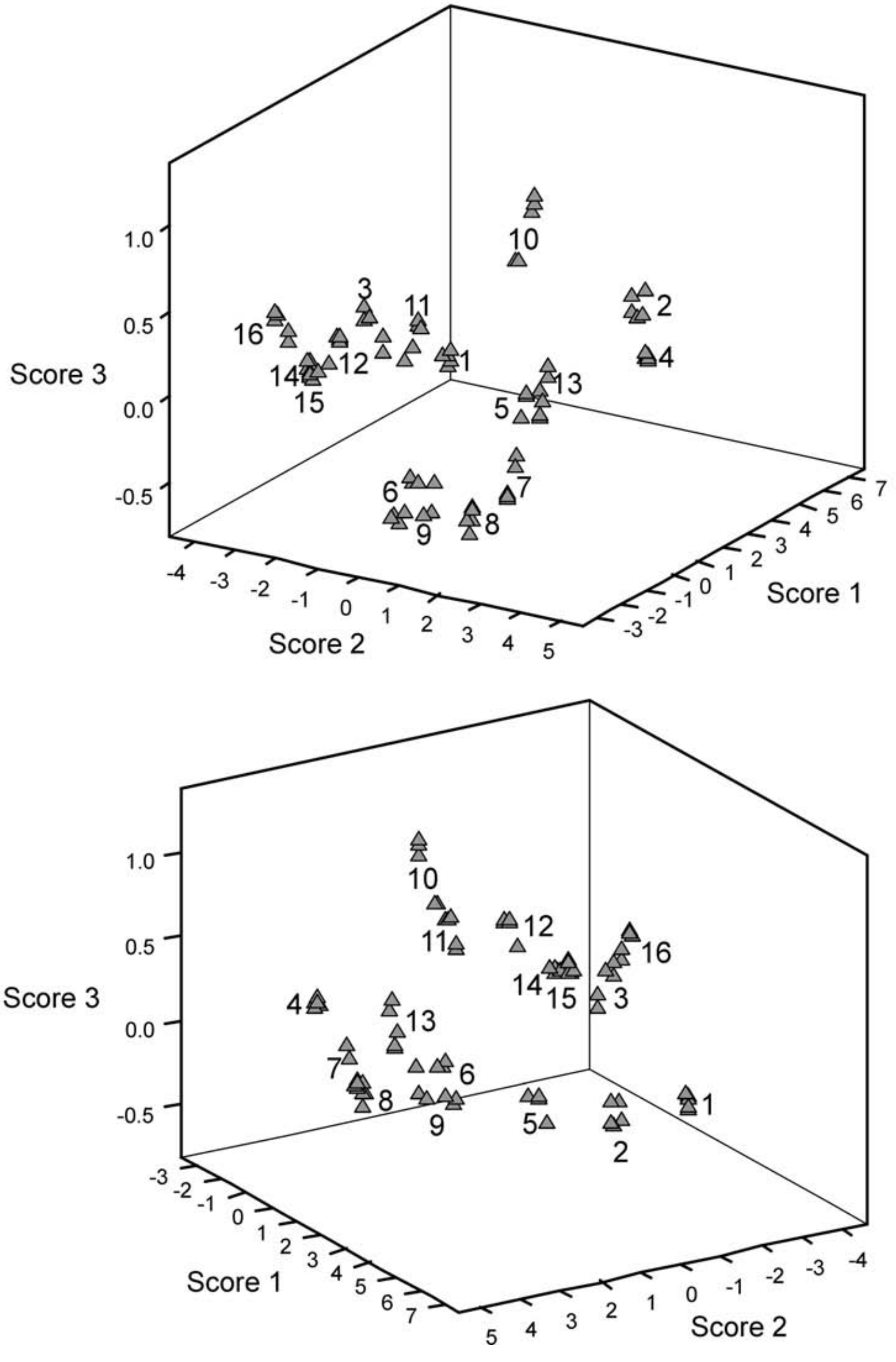
	Eigenvalue	Proportion	Cumulative (%)
PC1	5.2582	0.478	47.8
PC2	5.2054	0.473	95.1
PC3	0.2460	0.022	97.4
PC4	0.1671	0.015	98.9
PC5	0.0449	0.004	99.3

Table 6.9 Results of LDA with cross-validation: number of misclassified aliquots of Caltex regular unleaded gasoline from the two misclassified samples.

		True Group	
		Caltex regular week 6	Caltex regular week 9
Predicted Group	Caltex regular week 9	1	
	Caltex regular week 6		1

In Figure 6.6, the plot of the first two PC scores for the samples of Caltex regular gasoline shows weeks 6, 9 and 11 are all very close to one another. In the three-dimensional score plots of the first three PC scores (Figure 6.8) it can be seen along the axis of the third score that week 11 is well separated from weeks 6 and 9.

Figure 6.8 PC score plots of Caltex regular unleaded gasoline samples. These samples were removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.



6.3.3 Results from the analysis of Shell regular unleaded and premium unleaded gasoline

Principal component analysis was performed on the 32 Shell gasoline samples (16 premium unleaded and 16 regular unleaded) and the results are given in Table 6.10 and Figure 6.9. The first three PCs for the Shell samples explain 98.5% of the variation in the data set, slightly more variation than either the BP data set (96.7%, Table 6.2) or the Caltex data set (97.8%, Table 6.5). The first two PCs for the Shell data account for 96.4% of the variation in the data, which is approximately the same amount of variation accounted for in the Caltex data (96.6%, Table 6.5), but much more than the variation accounted for in the BP data (86.9%, Table 6.2).

Table 6.10 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): Shell regular unleaded and premium unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	8.4491	0.768	76.8
PC2	2.1598	0.196	96.4
PC3	0.2267	0.021	98.5
PC4	0.0926	0.008	99.3
PC5	0.0461	0.004	99.8

Using the first three PC scores, linear discriminant analysis (with cross-validation) correctly classified all 160 chromatograms (100%) by sample. That is, there are 32 unique groups where each group contains only the five aliquots from one gasoline sample. This means that the within sample variation is smaller than the variation between samples (groups). In the PC score plot (Figure 6.9) it can be seen that there is no trend according to week for either the premium or regular samples. Inspection of Figure 6.9 shows that the samples are clustered together by fuel grade. Furthermore, the premium samples are tightly clustered together. A closer look at the Shell data by fuel grade is required to understand the LDA results.

The data set was divided according to fuel grade. Principal component analysis was performed on the 16 Shell premium samples (Table 6.11) and on the 16 Shell regular samples (Table 6.12). The PCA results for both grades of fuel are plotted in Figure 6.10.

Figure 6.9 PC score plot of Shell premium unleaded and regular unleaded gasoline samples removed from the larger data matrix and treated independently.

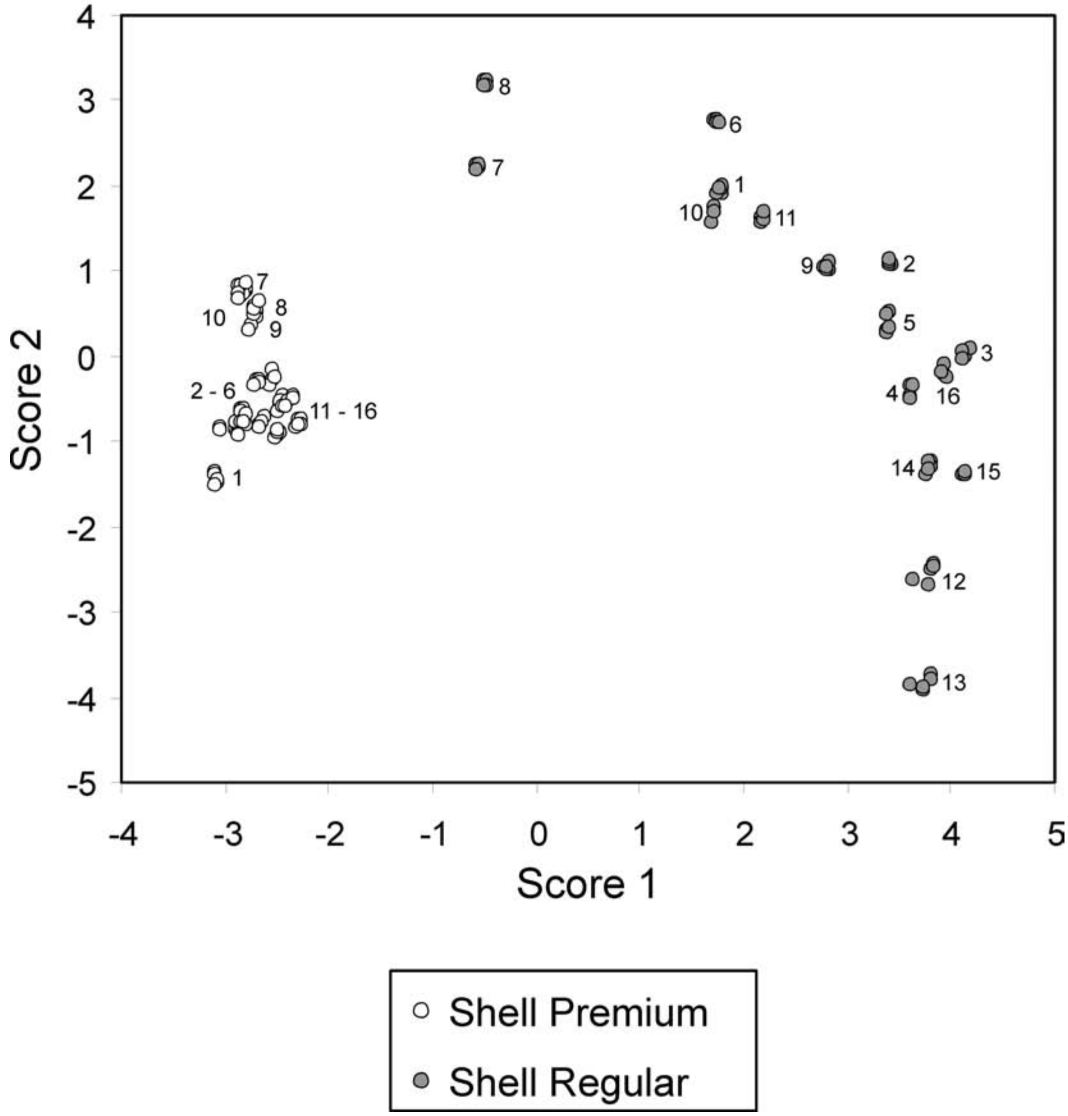
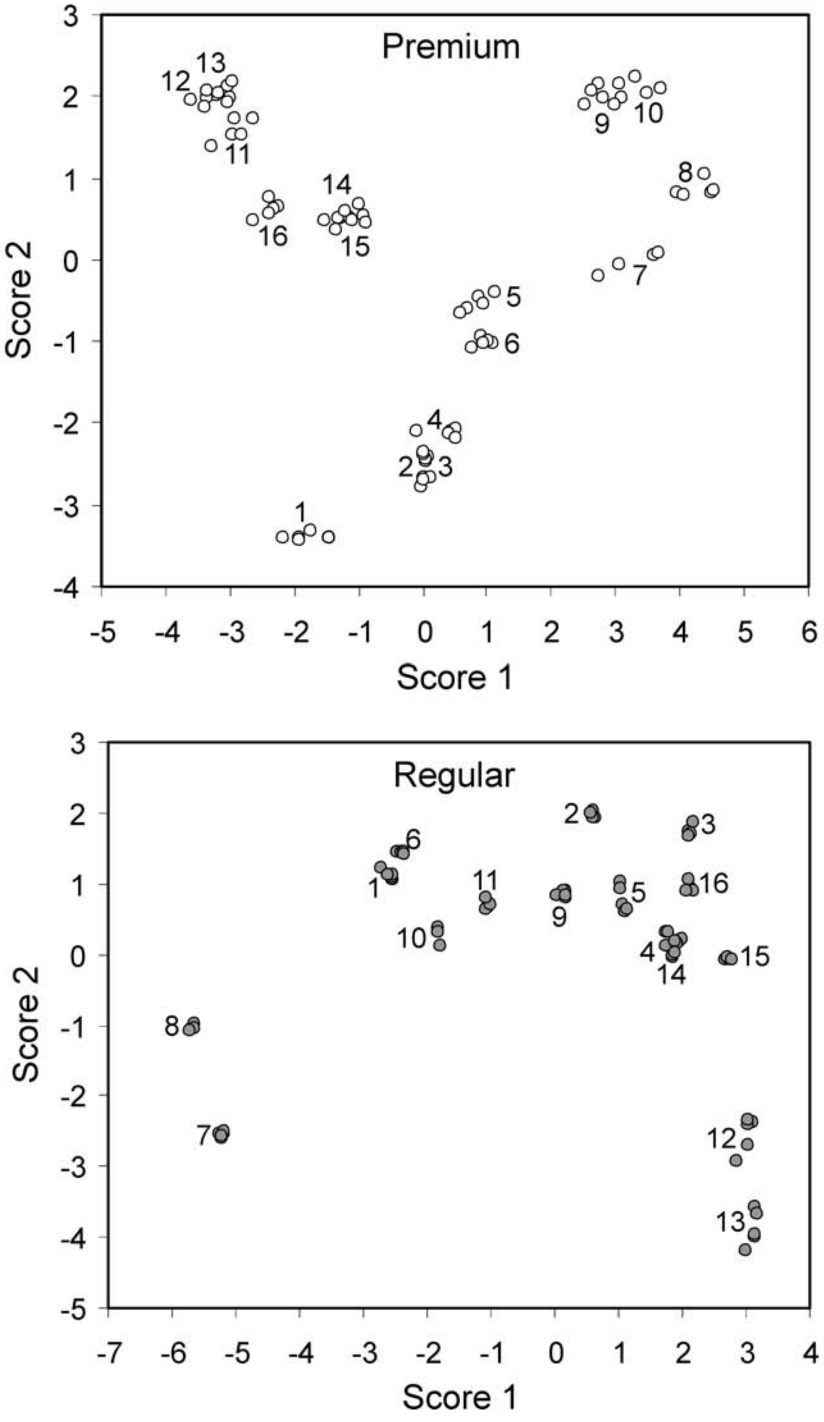


Figure 6.10 PC score plots of Shell premium unleaded and regular unleaded gasoline samples. Each fuel type (regular or premium) was removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.



The separation between premium gasoline samples becomes more apparent when the data for the two fuel grades has been separated (Figure 6.10, top). However, the third PC for the Shell premium gasoline samples (Table 6.11) accounted for a significant amount of variation in the data (12.3%). Inspection of the three-dimensional plots of the first three PC scores for the Shell premium samples (Figure 6.11) shows the separation of each sample into a unique group.

Table 6.11 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): Shell premium unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	5.7910	0.526	52.6
PC2	3.2375	0.294	82.1
PC3	1.3507	0.123	94.4
PC4	0.3415	0.031	97.5
PC5	0.1742	0.016	99.0

Table 6.12 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): Shell regular unleaded gasoline samples collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	7.5839	0.689	68.9
PC2	2.6814	0.244	93.3
PC3	0.3661	0.033	96.6
PC4	0.2649	0.024	99.1
PC5	0.0542	0.005	99.6

A larger proportion of the variation in the data is explained by the first two PCs for the Shell regular gasoline samples (93.3%, Table 6.12) compared with the first two PCs for the Shell premium samples (82.1%, Table 6.11). This explains why better separation is observed between most of the Shell regular gasoline samples (Figure 6.10, bottom) as compared to the premium samples (Figure 6.10, top). In the two-dimensional score plot of the regular gasoline samples (Figure 6.10, bottom), however, it is not possible to discern a difference between weeks 4 and 14. In plots of the first three scores for the Shell regular gasoline samples (Figure 6.12) it can be seen that aliquots from week 4 and week 14 are separated along the axis of the third PC score. Linear discriminant analysis performed on

Figure 6.11 PC score plots of Shell premium unleaded gasoline samples. These samples were removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.

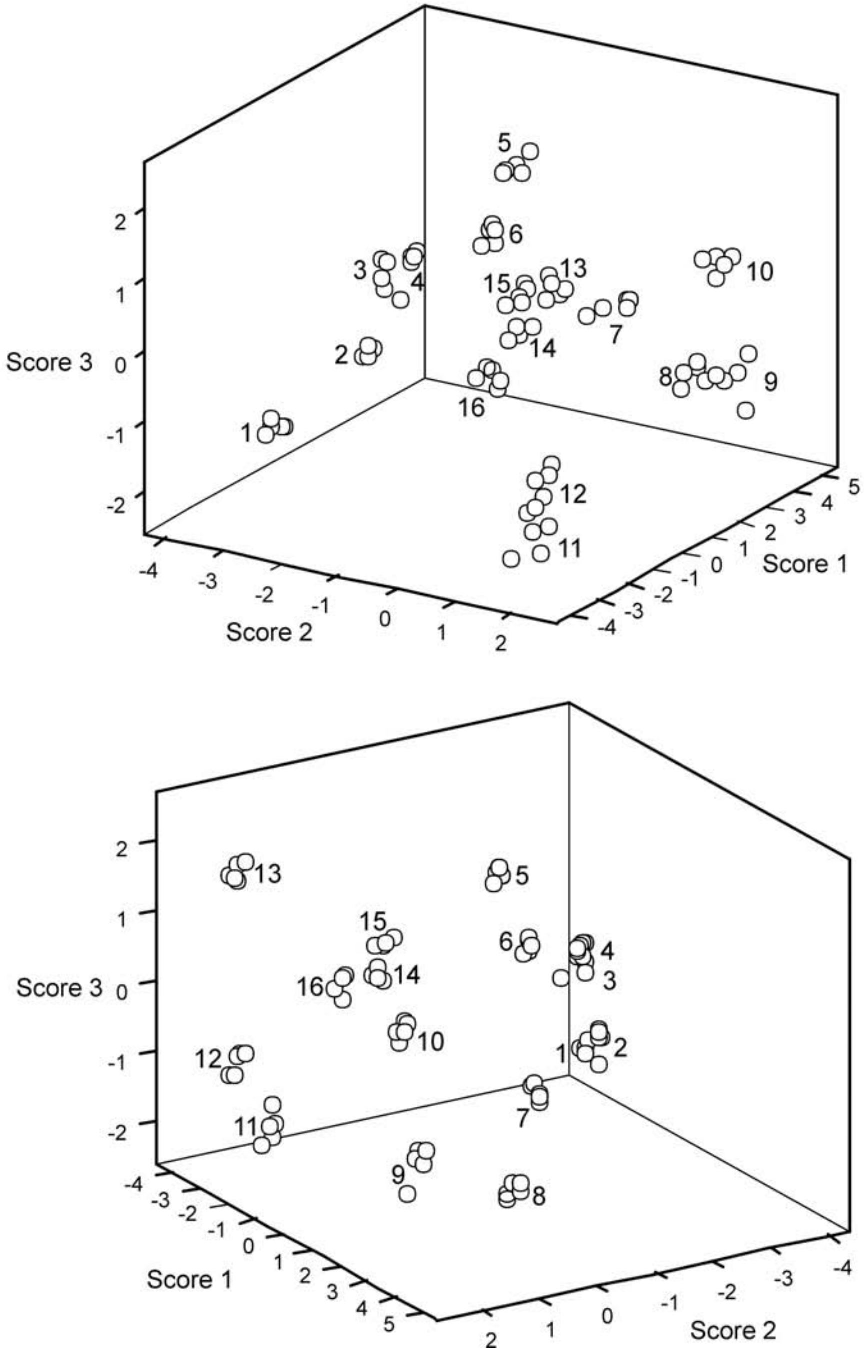
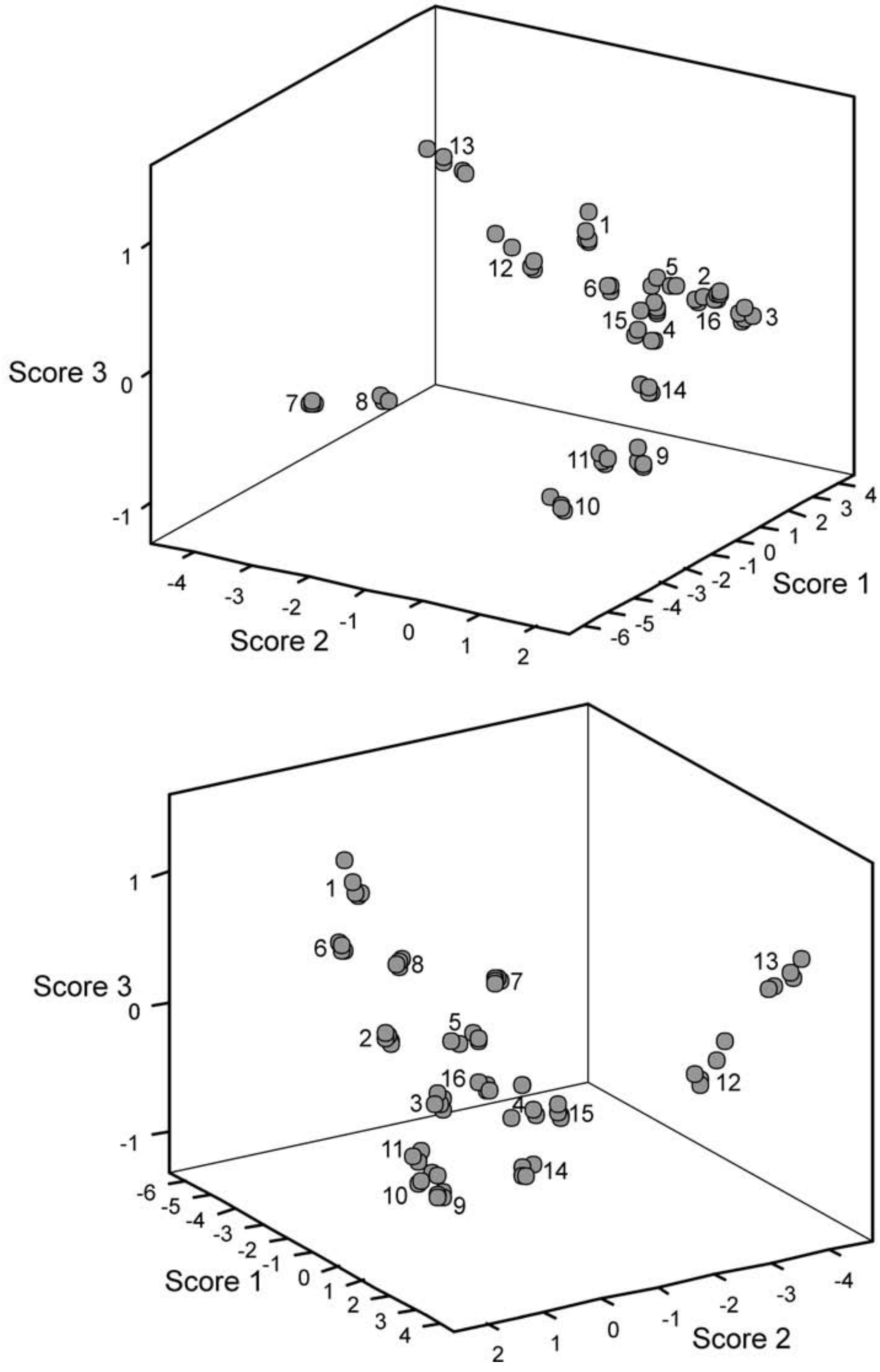


Figure 6.12 PC score plots of Shell regular unleaded gasoline samples. These samples were removed from the larger data matrix and treated independently. Sample week is indicated by the number next to each data cluster.



the first three PC scores resulted in the correct classification of all 16 premium aliquots (100%) and all 16 regular aliquots (100%) by sample.

6.3.4 Discussion

The comparison of 48 regular unleaded and 48 premium unleaded gasoline samples collected from three different service stations over a period of 16 weeks has not been reported before. Mann's collection of six samples of regular unleaded gasoline from a single service station over a period of 28 days is the most extensive study published in the scientific literature.[1] Forensic comparisons of large numbers of gasoline samples were published by Chan[3] in 1981 (46 samples) and Hirz[4] in 1989 (40 samples). Chan[3] was able to classify 26 of the 46 samples into 17 different groups based on the ratios of tetraethyl lead to tetramethyl lead. The remaining 20 samples only contained tetraethyl lead and so formed a separate group. Hirz[4] found that alkyl lead ratios in 23 premium gasoline samples were helpful in identifying the refinery from which a premium sample originated. Lead alkyls were not present in the 17 regular samples collected by Hirz.[4] Oxygenates, such as methanol, ethanol, *t*-butanol and MTBE, present in some of the samples collected by Hirz,[4] provided another point of comparison. Hirz[4] applied the method developed by Mann[1] and reported that a comparison of chromatographic profiles of the volatile compounds allowed all 40 unevaporated gasoline samples to be distinguished. This method involved a visual comparison between chromatograms and, therefore, had an element of subjectivity associated with it. The phasing out of leaded gasoline around the industrialised world has rendered the work on alkyl lead compounds by Chan[3] and Hirz[4] out of date.

For a given service station, monitoring the change in the C_0- to C_2-naphthalene profile from one week to the next provided a good basis for discriminating samples of gasoline. Analysis of the naphthalene profile by PCA and plotting the results of the first two scores often presented a good visual picture of the grouping of aliquots and the concomitant separation of samples. Although plots of the first two PC scores are attractive in their ability to give a simple visual explanation of the data, it must be remembered that the inclusion of other principal components may be necessary to correctly interpret the data. For example, the three-dimensional score plots of the Shell premium samples (Figure 6.11) gave a clearer representation of the uniqueness of each sample from one week to the next than either of the two-dimensional score plots (Figures 6.9 and 6.10, top). Another example is the incorporation of the third PC for the BP premium samples (Figure 6.3). In this case the third PC divided the 16 premium samples into two diagonal planes with samples from weeks 1, 2, 3, 7, 10, 11, 12, 13, 14, 15 and 16 on one plane and weeks 4, 5, 6, 8 and 9 on the other plane (Figure 6.3). A third example can be taken from the PCA of the Caltex premium samples. The plot of the first two PC scores for the Caltex premium gasoline samples (Figure 6.6, top) showed good separation for all samples except weeks 1, 2 and 3, and weeks 6 and 7. Plots of the first three PC scores (Figure 6.7) showed how samples from weeks 1, 2 and 3, as well as samples from weeks 6 and 7, are separated in the third dimension.

6.4 Analysis of all samples and weekly differences

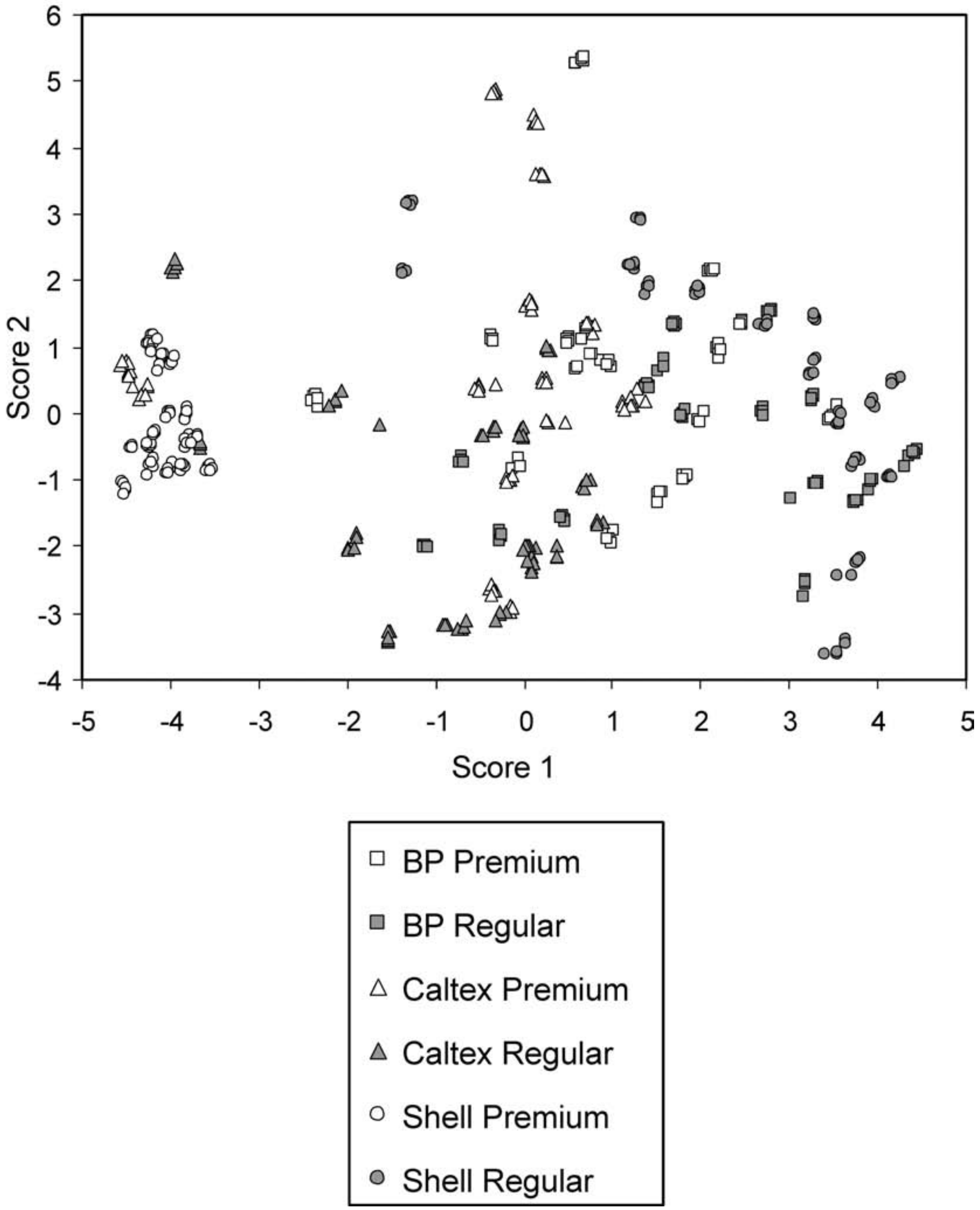
Trends between service stations over time were examined by analysis of all 96 gasoline samples (480 chromatograms); principal component analysis results are given in Table 6.13 and Figure 6.13.

Table 6.13 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): all brands (BP, Caltex, Shell) and both fuel types (regular unleaded and premium unleaded gasoline) collected over a 16 week period.

	Eigenvalue	Proportion	Cumulative (%)
PC1	6.9845	0.635	63.5
PC2	3.0091	0.274	90.9
PC3	0.5324	0.048	95.7
PC4	0.2767	0.025	98.2
PC5	0.1113	0.010	99.2

Linear discriminant analysis (with cross-validation) using the first three PCs correctly classified 473 of the 480 chromatograms (98.5%) by sample. In practical terms any overlap

Figure 6.13 PC score plot of BP, Caltex and Shell premium unleaded and regular unleaded gasoline samples collected over a 16 week period.



of one chromatogram (aliquot) from one sample with any other chromatogram from another sample would cause these two gasoline samples to be indistinguishable. Therefore, the LDA results were re-examined and it was found the 96 samples could be divided into 92 unique groups. Eighty-eight of these groups each contained a single sample (*i.e.* five aliquots from one sample per group). The eight remaining (misclassified) samples were divided between the four remaining groups. Each of these four groups contained a pair of samples as follows: Shell premium weeks 3 and 4; Caltex premium week 5 and BP premium week 5; Caltex premium week 11 and Caltex regular week 12; and, Shell regular week 4 and BP premium week 16 (Table 6.14).

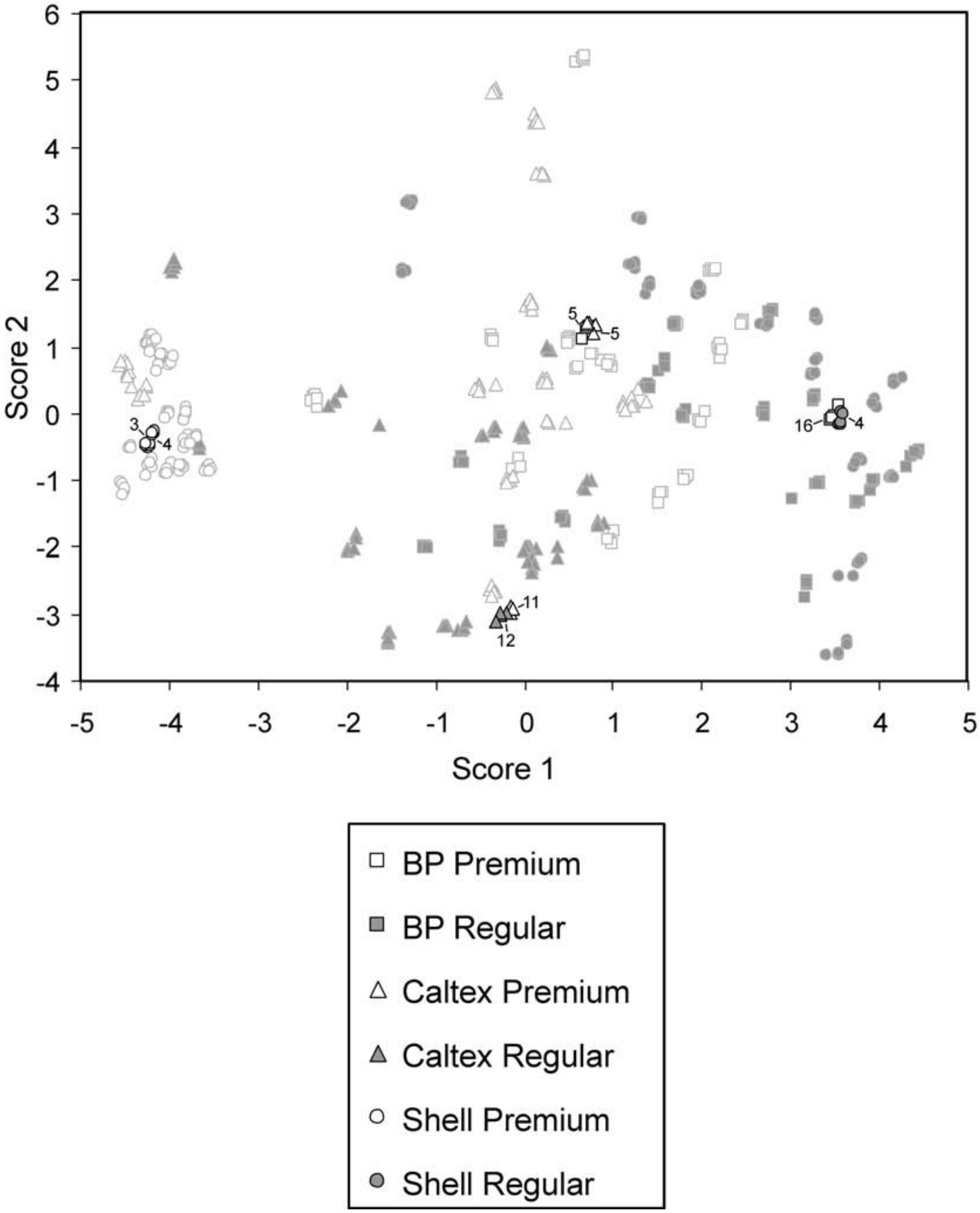
Table 6.14 Results of LDA with cross-validation: number of misclassified aliquots from the eight misclassified samples.

				True Group		
		Shell premium week 4	Caltex premium week 5	BP premium week 5	Caltex regular week 12	BP premium week 16
Predicted Group	Shell Premium week 3	1				
	BP premium week 5		1			
	Caltex premium week 5			2		
	Caltex premium week 11				2	
	Shell regular week 4					1

The pairs of misclassified samples are highlighted in Figure 6.14. Note that BP does not have its own refinery and must be supplied by the Caltex or the Shell refinery in Sydney, or purchase fuel from an offshore refinery (see Chapter 2). The similarity between the Caltex and BP premium unleaded samples collected on week 5 suggests that, if there was no mixing of the gasoline with fuel already held the service stations' underground storage tanks, the BP and Caltex service stations received deliveries of the same fuel from the same distribution terminal. It is also interesting to note that six of the eight misclassified samples are premium gasoline, indicating that there is less variation in the C_0- to C_2-naphthalene profile for premium unleaded gasoline than there is for regular unleaded gasoline.

Plots of the first two PC scores showing the similarities and differences between the six samples collected in a particular week are provided in Appendix 6.1 (Figures App 6.1.1 to App 6.1.8).

Figure 6.14 PC score plot of BP, Caltex and Shell premium unleaded and regular unleaded gasoline samples collected over a 16 week period. Four pairs of misclassified samples are highlighted and labeled by week. See Table 6.14 and the text for explanation of results. Sample weeks are indictated by the number next to each highlighted data cluster.



6.5 Conclusions

Principal component analysis, followed by linear discriminant analysis (with cross-validation), has been used to examine the C_0- to C_2-naphthalene profiles of 48 regular unleaded and 48 premium unleaded gasoline samples collected from three branded service stations over 16 consecutive weeks. Examination of the samples by brand showed that all 32 BP samples (16 regular and 16 premium) and all 32 Shell samples (16 regular and 16 premium) could be differentiated by PCA and LDA. The 32 Caltex samples (16 regular and 16 premium) were classified into 29 groups, 26 of which contained one sample. The remaining six samples formed the other three groups. The 96 gasoline samples from the three service stations were classified into 92 groups, 88 of which contained one sample. The remaining eight samples formed the other four groups. These results demonstrate that not only can the naphthalene profile be used to distinguish a large number of unevaporated gasoline samples, but also that a sample collected one week from one service station is likely to be different from a sample collected the next week from the same service station. Furthermore, it was shown that unevaporated gasoline samples collected on the same day of the week are likely to differ from one branded service station to the next. These results demonstrate that principal component analysis used in conjunction with linear discriminant analysis is a powerful tool for eliciting small but significant differences from chemical data.

6.6 References

1. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography. I: comparison methodology. Journal of Forensic Sciences, 1987; 32(3): 606-615.

2. Rietjens, M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytical Chemistry, 1995; 316: 205-215.

3. Chan, L. The determination of tetraalkyl lead compounds in petrol using combined gas chromatography atomic absorption spectrometry. Forensic Science International, 1981; 18: 57-62.

4. Hirz, R. Gasoline brand identification and individualization of gasoline lots. Journal of the Forensic Science Society, 1989; 29(2): 91-101.

Chapter 7. Comparison of unevaporated and evaporated gasoline samples

7.1 Introduction

Conditions at a crime scene before, during and after a fire are uncontrolled. It is, therefore, likely that the gasoline sample recovered from a fire scene will have a different level of evaporation than the gasoline sample seized from a suspect. Furthermore, control samples of gasoline obtained from local service stations by arson investigators will not be evaporated at all (if collected and stored appropriately). Having gasoline samples at different levels of evaporation will increase the complexity of the forensic comparison of those samples.

There are a number of studies in the scientific literature that have attempted to address the problem of comparing samples of unevaporated and evaporated gasoline.[1,2,3,4,5,6] Mann[1] used the chromatographic profiles of some of the more volatile compounds to differentiate unevaporated samples of gasoline. In a study of evaporated gasoline, Mann[2] observed that evaporated samples could be compared to less evaporated or unevaporated samples using the relative ratios of compounds that had similar vapour pressures. Although able to make limited comparisons between evaporated and unevaporated gasoline samples, Mann concluded that the method was most useful in "eliminating the possibility of common origin between two samples."[2] Hirz[3] compared chromatographic profiles of high boiling components of evaporated gasoline samples obtained from nine different refineries in Europe and was able to classify, in a limited way, which samples came from a particular refinery. Potter[4] used principal component analysis to explore how different isomeric groups could be used to compare a small number of evaporated and unevaporated of gasoline samples. Potter[4] was unable to satisfactorily compare a 50% evaporated gasoline sample with the original, unevaporated sample. It has been shown that three-dimensional fluorescence spectroscopy could not only distinguish different samples of gasoline,[5] but could also match a sample evaporated up to 50% back to the original, unevaporated gasoline sample.[6]

Environmental regulations have resulted in similar comparisons being attempted between evaporated and unevaporated crude oils. For example, Wang and Fingas[7] have discussed the use of isomer ratios to characterise and compare crude oil spills. Wang and Fingas[8] also studied and quantitated the changes that occur when a light crude oil is evaporated and even devised a “Weathering Index” to allow the degree of evaporation of a crude oil to be estimated. More recently, Smallwood *et al.*[9] used carbon isotope ratios for selected compounds in three gasoline samples to successfully compare an unevaporated sample to its original, unevaporated sample.

In this chapter, 35 samples of gasoline were compared at five different levels of evaporation (0%, 25%, 50%, 75% and 90%). These samples were compared within each level of evaporation to assess the relative degree of change in the C_0- to C_2-naphthalene profile. In addition, all 35 samples at each of the five levels of evaporations were compared altogether. This last comparison was made not only to assess the relative changes exhibited within a single sample of gasoline as it evaporates, but also to examine the ability of the method to differentiate a sample of gasoline at one level of evaporation from another sample of gasoline at a different level of evaporation.

7.2 Materials and methods

7.2.1 Unevaporated and evaporated samples

Thirty-five samples of gasoline were obtained from 24 service stations in metropolitan Sydney, Australia (Table 7.1) over a seven-month period (March to September, 2001). The majority of the samples (29) were collected over a period of seven weeks (March to April, 2001). Samples were collected in 125 mL amber, Boston round glass bottles sealed with polypropylene caps fitted with Teflon-faced foamed polyethylene liners (Wheaton) and stored in the dark at room temperature.

Table 7.1 Summary of 35 unleaded gasoline samples collected between March and September, 2001 from 24 service stations in metropolitan Sydney, Australia. Shaded and unshaded blocks indicate samples that were collected from the same service station on the same day. Sample numbers 1 through 6 are from the same service station.

No.	Sample	Date[†]	Brand	Grade[‡]
1	RU02	13.03	Shell	Regular
2	LR06	13.03	Shell	Lead rep
3	PU07	13.03	Shell	Premium
4	RU11	29.03	Shell	Regular
5	PU12	29.03	Shell	Premium
6	RU13	05.04	Shell	Regular
7	RU14	18.04	Caltex	Regular
8	RU15	18.04	Mobil	Regular
9	RU16	18.04	Caltex	Regular
10	PU17	18.04	Caltex	Premium
11	RU18	18.04	Shell	Regular
12	PU19	24.04	BP	Premium
13	RU20	24.04	BP	Regular
14	RU21	24.04	Caltex	Regular
15	RU22	24.04	Shell	Regular
16	RU23	24.04	BP	Regular
17	RU24	24.04	Mobil	Regular
18	RU25	24.04	Mobil	Regular
19	RU26	24.04	Caltex	Regular
20	RU27	24.04	Shell	Regular
21	PU28	24.04	Shell	Premium
22	RU29	24.04	BP	Regular
23	RU30	24.04	BP	Regular
24	PU31	24.04	Shell	Premium
25	RU32	24.04	Shell	Regular
26	RU33	24.04	Mobil	Regular
27	RU34	24.04	Mobil	Regular
28	RU35	24.04	Mobil	Regular
29	RU36	28.04	BP	Regular
30	PU37	14.08	BP	Premium
31	RU38	14.08	BP	Regular
32	PU39	12.09	Burmah	Premium
33	RU40	12.09	Burmah	Regular
34	PU41	12.09	Woolies	Premium
35	RU42	12.09	Woolies	Regular

[†] Date is expressed as day and month (dd.mm)
[‡] The three grades of unleaded gasoline are: regular, premium and lead replacement

Samples of evaporated gasoline were prepared from aliquots of each unevaporated gasoline sample listed in Table 7.1. The evaporation levels selected for this study (0%, 25%, 50%, 75% and 90% evaporated by weight) were chosen because they span a wide range of evaporation levels. Samples with target evaporation levels of 25% and 50% were prepared in 4 mL vials (Supelco), while samples with target evaporation levels of 75% and 90% were prepared in 7 mL vials (Supelco). All vials were clear glass and were sealed with Teflon-lined green phenolic caps (Supelco). A clean, unused vial was used for each sample at each target evaporation level. An evaporated sample was prepared in the following manner: the weight of an empty vial and its cap was determined using a precision analytical balance (Denver Instrument Co.); the vial was filled with an aliquot of unevaporated gasoline, capped and re-weighed; the total weight of gasoline in the vial was calculated; the target weight of the gasoline sample, vial and cap was then calculated based on the target level of evaporation; the weight of the sample, vial and cap was monitored over time while the sample was allowed to evaporate in a fume cupboard until it reached the pre-determined target weight for the desired evaporation level. The final weight of the sample, vial and cap was determined and the final weight and level of evaporation of the sample (by weight) was calculated. This procedure was repeated for each of the 35 gasoline samples at four of the five levels of evaporation (25%, 50%, 75% and 90% evaporated by weight). A total of 175 samples were prepared for analysis (35 samples each at 0%, 25%, 50%, 75% and 90% evaporation). The weights and calculations made for each sample at four of the five levels of evaporation (25%, 50%, 75% and 90%) are presented in tabular form in Appendix 7.1

7.2.2 Gas chromatography-mass spectrometry

Gas chromatography-mass spectrometry conditions are described in section 6.2.2. Samples were prepared for GC-MS (SIM) by diluting an aliquot of neat gasoline with dichloromethane (1:4, v/v). The detector was set to monitor the following ion groups from the two-ring PAHs as follows: m/z^+ 127 and 129 (group 1; C_0-naphthalene); m/z^+ 139 and 143 (group 2; C_1-naphthalenes); m/z^+ 128, 141, 142, 155, 156 and 157 (group 3; C_2-naphthalenes); and, m/z^+ 128, 141, 153, 155, 169, 170 and 171 (group 4; C_3-naphthalenes). Dwell times were 100 ms for groups 1 and 2, and 50 ms for groups 3 and 4. In all cases five aliquots from each sample at each evaporation level were analysed.

Some gasoline samples had a high naphthalene concentration. If left unchecked, the high naphthalene concentration would result in column overloading and/or detector saturation which, in turn, would result in poor peak shape for naphthalene with a concomitant error in peak area integration, particularly for the more evaporated samples (*i.e.* 90% evaporated). Each chromatogram was carefully inspected to ensure that good chromatography was obtained. Those few samples whose naphthalene concentration was found to be too high were analysed using the following conditions: sample dilution in dichloromethane was increased to 1:10 (v/v) and the split ratio of the gas chromatograph was increased to 60:1. This approach not only introduced a lower concentration of naphthalene onto the column (thus avoiding the problems described above), but also maintained a good chromatographic response for the C_1- and C_2-naphthalenes.

Total ion chromatograms (TICs) of unevaporated (0% evaporated) and evaporated gasoline (25%, 50%, 75% and 90% evaporated) were obtained for two samples, RU02 and PU12. Chromatographic conditions were as follows: split injection (15:1) at 280 ^{o}C; temperature programmed from 40 ^{o}C (3 minutes isothermal) to 60 ^{o}C at 5 ^{o}C/minute, then to 190 ^{o}C at 10 ^{o}C/minute, then to 290 ^{o}C at 20 ^{o}C/minute (held for 5 minutes); helium carrier gas at a constant flow rate of 1.2 mL/minute. The sample injection volume was 1 μL and the detector was operated in full scan mode (30 to 350 daltons) with a sampling rate of 4.45 scans/second.

7.2.3 Data analysis

Peak area information was extracted from the chromatograms using a macro. The macro is reproduced in Appendix 5.1. The integration results were written to an ASCII text file in a format suitable for export to a spreadsheet. Data was loaded and parsed into a spreadsheet (Microsoft Excel®). The data was normalised using a natural logarithmic transformation.[10] Principal component analysis and linear discriminate analysis was performed using Minitab® (version 13.1 for Windows). Error ellipses were calculated using the method described by Jackson.[11] Error ellipse calculations were performed using a Minitab macro modified for this work to allow user input of the F-statistic.[12] The macro is given in Appendix 3.1.

7.3 Unevaporated and evaporated gasoline samples

7.3.1 Overview of results for unevaporated and evaporated gasoline samples

For each of the 35 samples, the level of evaporation was brought as close to the target level as possible (Appendix 7.1). The level of evaporation achieved for the 35 samples at four of the five evaporation levels (25%, 50%, 75% and 90% evaporation) is summarised in Table 7.2. For each level of evaporation it was noted that not all samples evaporated at the same rate. It also became progressively easier to monitor and control evaporation rates as the target evaporation levels increased. The rate of evaporation was the most difficult to control for samples evaporated to 25% by weight and resulted in the greatest variation between samples, the rate of evaporation was slow for samples evaporated to 90% by weight which made it relatively easy to control the rate of evaporation and resulted in little variation between samples (Table 7.2). In all cases it was important to achieve, as much as possible, the same level of evaporation for all samples at a given target evaporation level so that any differences found between samples would be due solely to differences intrinsic to the sample and not due to differences in the levels of evaporation.

Table 7.2 Summary of level of evaporation achieved for 35 gasoline samples at each targeted evaporation level.

25% evaporated		*50% evaporated*	
mean	25.85%	mean	50.21%
sd	0.84%	sd	0.43%
95% ci	±0.29%	95% ci	±0.15%
min	24.90%	min	49.36%
max	27.69%	max	51.63%

75% evaporated		*90% evaporated*	
mean	75.07%	mean	90.02%
sd	0.14%	sd	0.13%
95% ci	±0.05%	95% ci	±0.04%
min	74.91%	min	89.85%
max	75.52%	max	90.47%

The analysis of five aliquots from each sample at each evaporation level resulted in 875 chromatograms (5 aliquots x 35 samples x 5 levels of evaporation {0%, 25%, 50%, 75% and 90%}). Eleven C_0- to C_2-naphthalene peaks were extracted from each chromatogram (Table 5.7 and Figure 5.1), therefore, the final data matrix was 11 by 875.

Total ion chromatograms for a sample of regular unleaded gasoline (RU02), Figure 7.1, and for a sample of premium unleaded gasoline (PU12), Figure 7.2, at each of the five

Figure 7.1 Regular unleaded gasoline sample RU02 total ion chromatograms (TICs) at five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight). Retention index standard shown in top chromatogram.

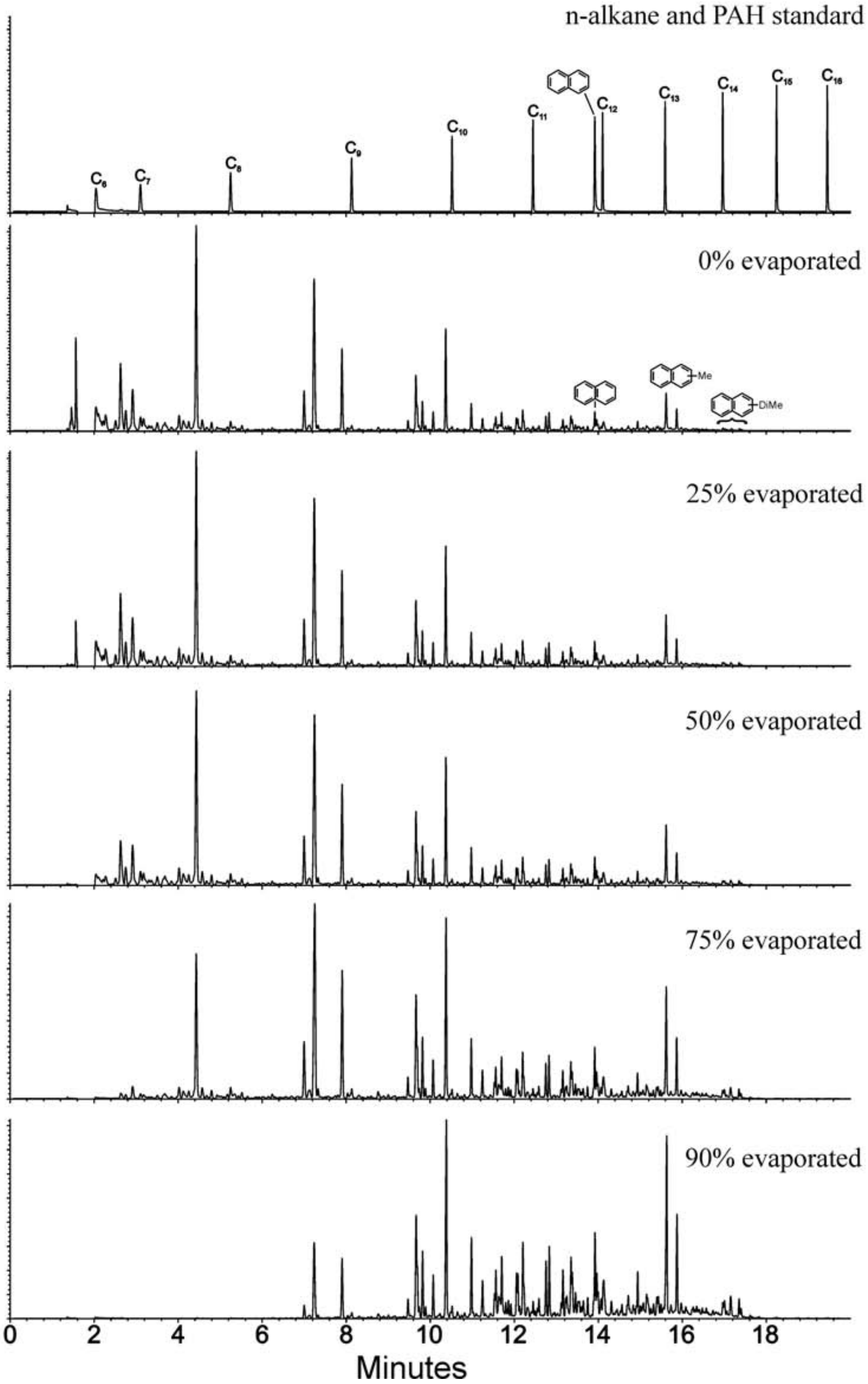
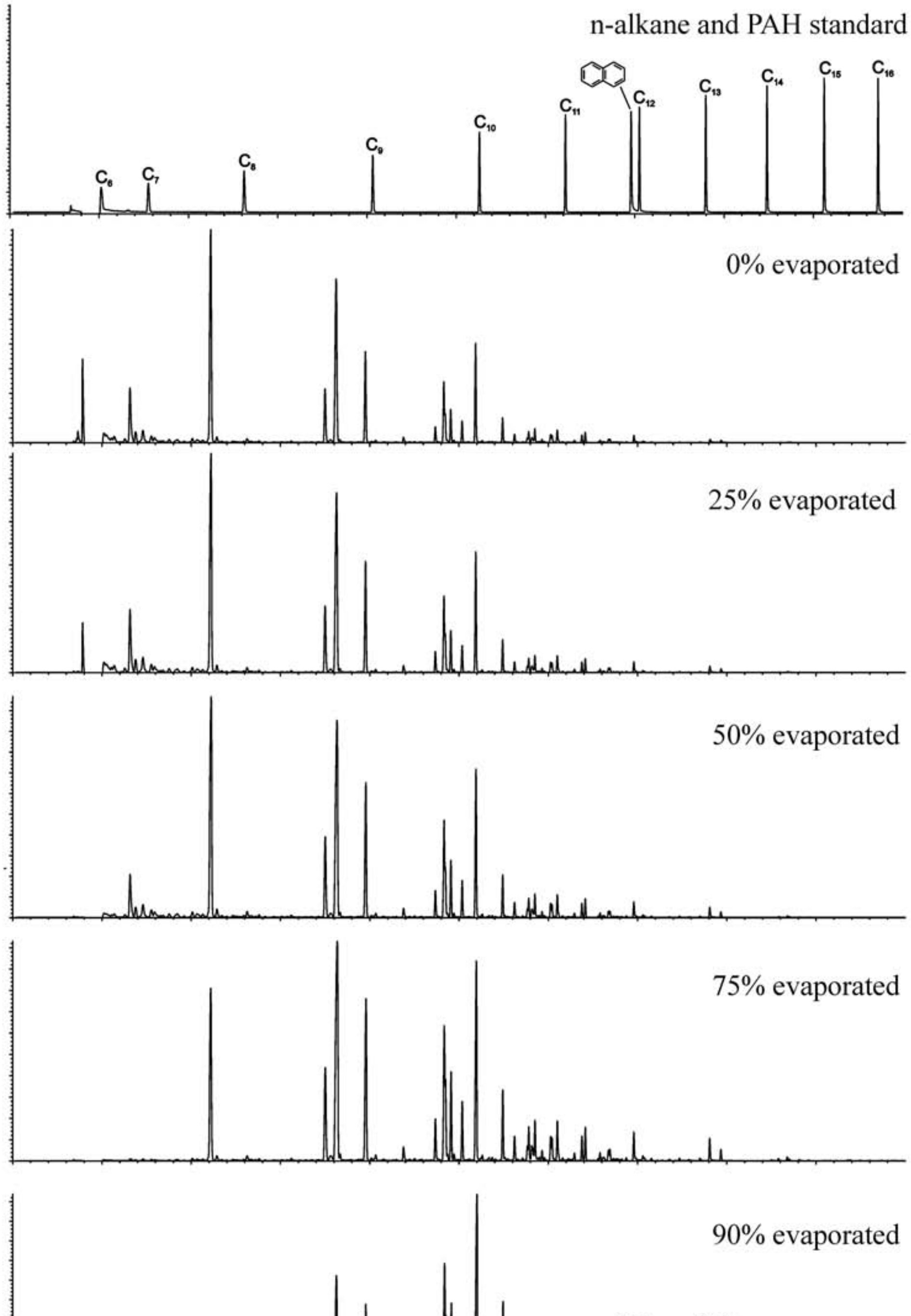


Figure 7.2 Premium unleaded gasoline sample PU12 total ion chromatograms (TICs) at five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight). Retention index standard shown in top chromatogram.



evaporation levels are shown. Chromatograms from selected ion monitoring of the C_0- to C_2-napthalenes for the regular gasoline sample (RU02) at each evaporation level are shown in Figure 7.3, and for the premium unleaded gasoline sample (PU12) in Figure 7.4.

The naphthalene profiles were extracted from the 175 chromatograms at each level of evaporation (0%, 25%, 50%, 75% and 90%) and analysed by principal components. The PCA results are summarised in Table 7.3. From the data in Table 7.3, it can be seen that the first three principal components explain most of the variation in the data set (>97%). The PCA results were consistent between each level of evaporation and when all evaporation levels were combined (Table 7.3). This similarity between evaporation levels may be seen by comparing plots of the first two PC scores. Plots of the first two principal component scores for the unevaporated (0%), 25%, 50%, 75% and 90% evaporated samples, are shown in Figures 7.5 to 7.9, respectively. The first two PCs account for more than 91% of the variation in each data set.

A PC score plot gives an illustration of chemical similarities or differences between samples.[13] Samples that are at opposite ends of a PC score plot are chemically more distinct than are samples that are very close to one another. From the PC score plots in Figures 7.5 to 7.9, it can be seen that the overall trends in chemical similarities or differences between the 35 samples remains reasonably intact from one level of evaporation to another. Good separation between the 35 samples at any given evaporation level is maintained because the between sample variation is greater than the within sample variation (*cf.* chromatograms in Figures 7.3 and 7.4).

The amount of variation explained by the principal components for all samples at all evaporation levels combined (n = 875) is similar to the PCA results at each of the five levels of evaporation (Table 7.3). When all samples at all evaporation levels are combined and analysed by principal components it is important to determine if the data for a given sample is homogeneous across all evaporation levels, or if it tends to cluster according to evaporation level (*i.e.* heterogeneous). It is also important to examine the size (*i.e.* the error) of the cluster formed by all evaporation levels for each sample. Seven of the 35 gasoline samples were selected to illustrate the homogeneity and range of cluster sizes

Figure 7.3 Regular unleaded gasoline sample RU02 at five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight). Chromatograms obtained from selected ion monitoring (SIM) of C_0- to C_2-naphthalenes. Retention index standard shown in top chromatogram.

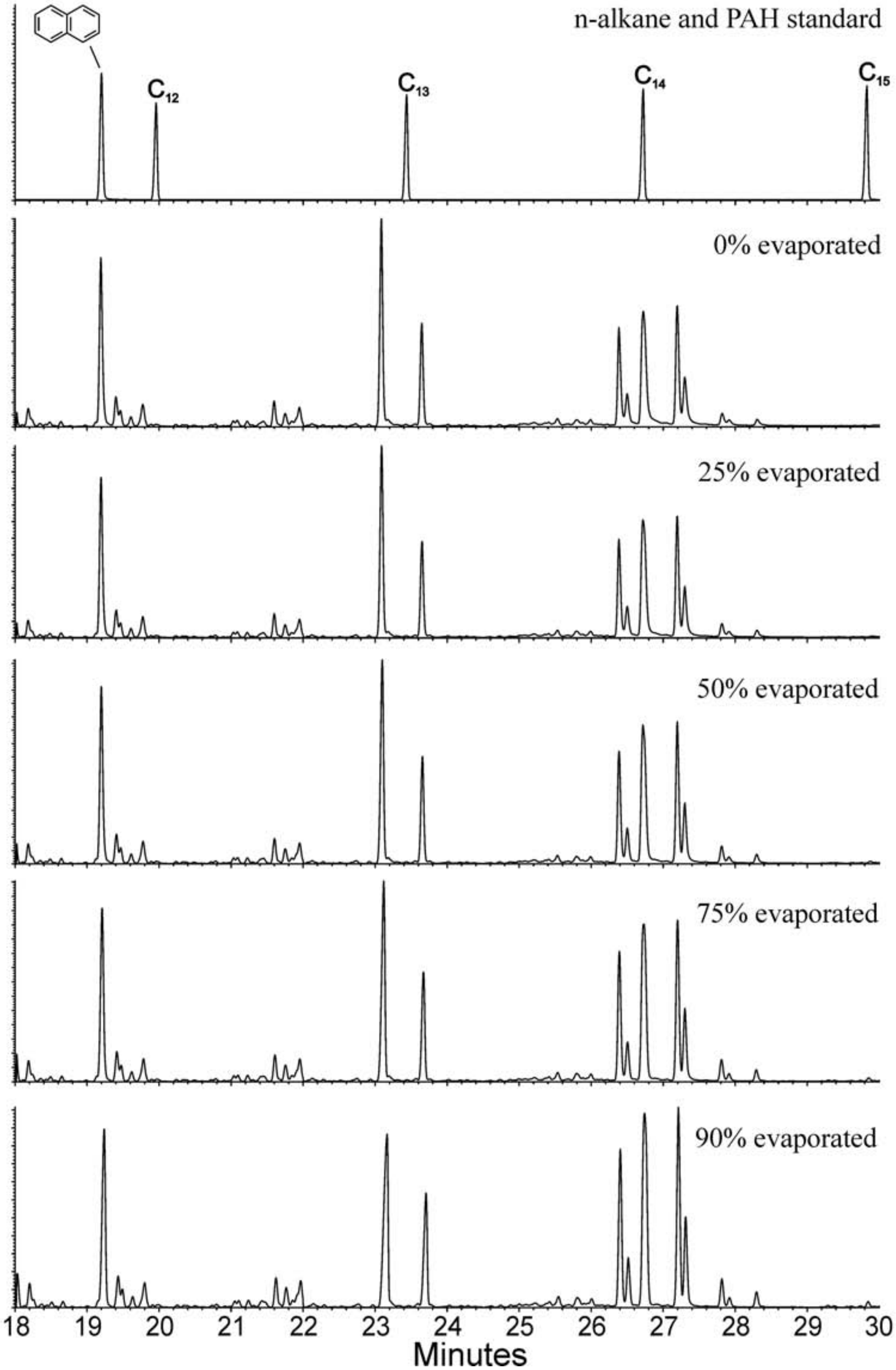


Figure 7.4 Premium unleaded gasoline sample PU12 at five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight). Chromatograms obtained from selected ion monitoring (SIM) of C_0- to C_2-naphthalenes. Retention index standard shown in top chromatogram.

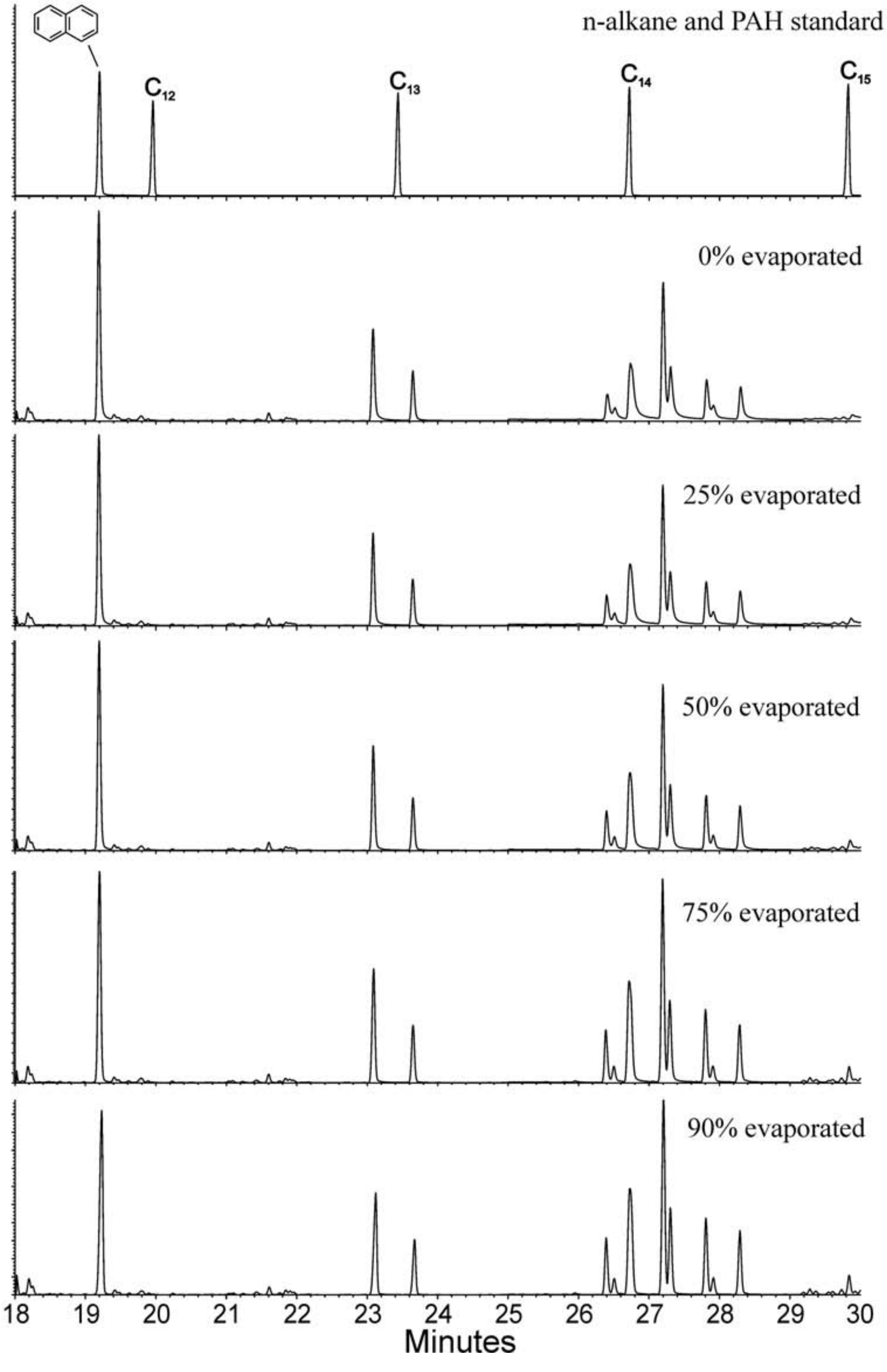


Table 7.3 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM)

		Eigenvalue	Proportion	Cumulative (%)
unevaporated samples (n = 175)	PC1	8.4765	0.771	77.1
	PC2	1.5857	0.144	91.5
	PC3	0.7925	0.072	98.7
	PC4	0.0714	0.006	99.3
	PC5	0.0466	0.004	99.8
25% evaporated samples (n = 175)	PC1	8.4997	0.773	77.3
	PC2	1.6404	0.149	92.2
	PC3	0.6276	0.057	97.9
	PC4	0.1104	0.010	98.9
	PC5	0.0754	0.007	99.6
50% evaporated samples (n = 175)	PC1	8.6219	0.784	78.4
	PC2	1.6912	0.154	93.8
	PC3	0.5002	0.045	98.3
	PC4	0.0998	0.009	99.2
	PC5	0.0449	0.004	99.6
75% evaporated samples (n = 175)	PC1	8.7081	0.792	79.2
	PC2	1.6102	0.146	93.8
	PC3	0.5183	0.047	98.5
	PC4	0.0973	0.009	99.4
	PC5	0.0265	0.002	99.6
90% evaporated samples (n = 175)	PC1	8.5133	0.774	77.4
	PC2	1.7306	0.157	93.1
	PC3	0.5850	0.053	98.4
	PC4	0.0721	0.007	99.1
	PC5	0.0400	0.004	99.5
all samples together (n = 875)	PC1	8.4432	0.768	76.8
	PC2	1.5885	0.144	91.2
	PC3	0.7099	0.065	97.7
	PC4	0.1202	0.011	98.7
	PC5	0.0778	0.007	99.5

observed for all the gasoline samples (Figure 7.10). The relative cluster sizes of these seven gasoline samples shown in Figure 7.10 ranged from small (PU17) to large (PU39). Each sample in Figure 7.10 is shown bounded by its 95% error ellipse (confidence

Figure 7.5 PC score plot of 175 chromatograms from 35 unevaporated gasoline samples (0% evaporated by weight).

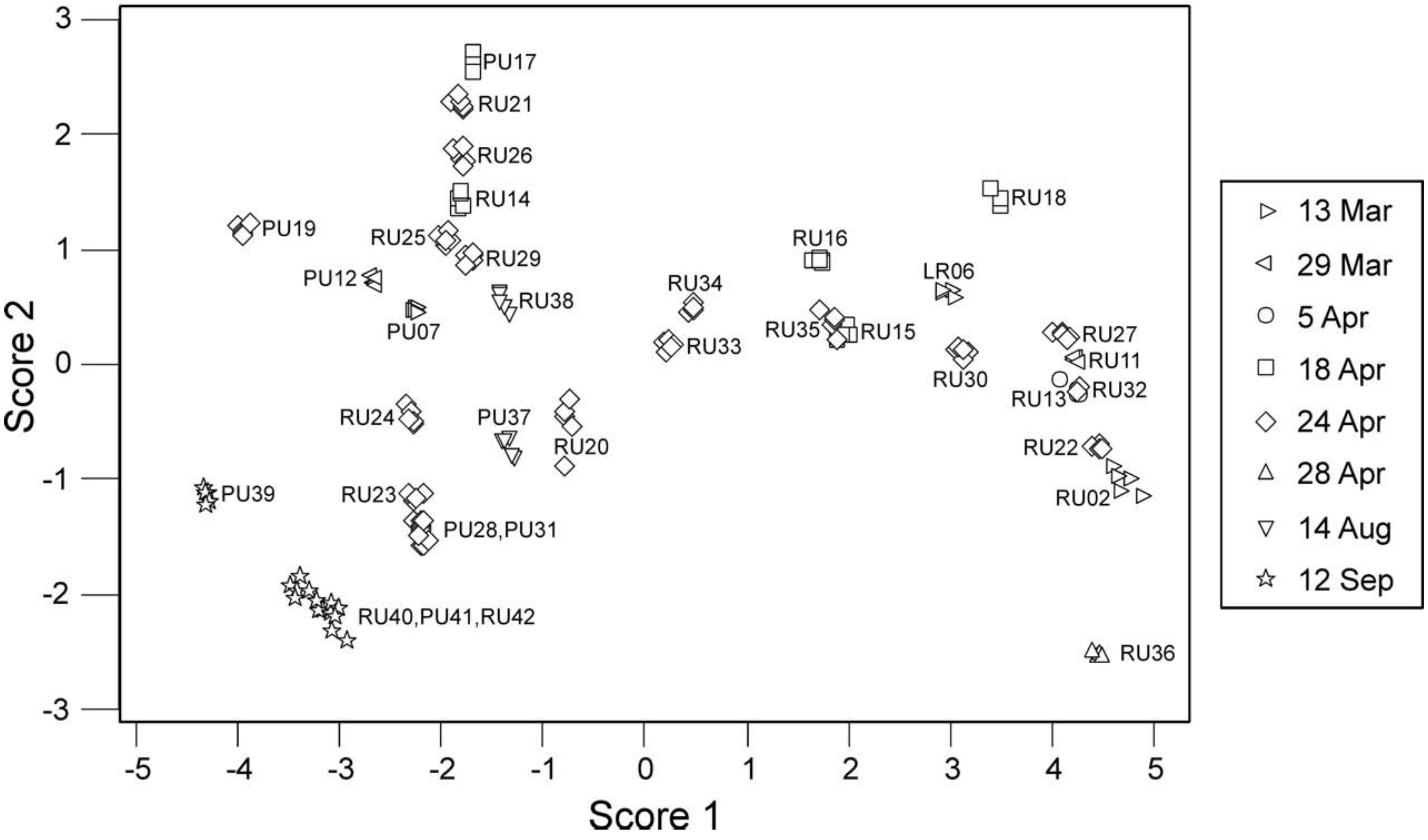


Figure 7.6 PC score plot of 175 chromatograms from 35 gasoline samples evaporated 25% by weight.

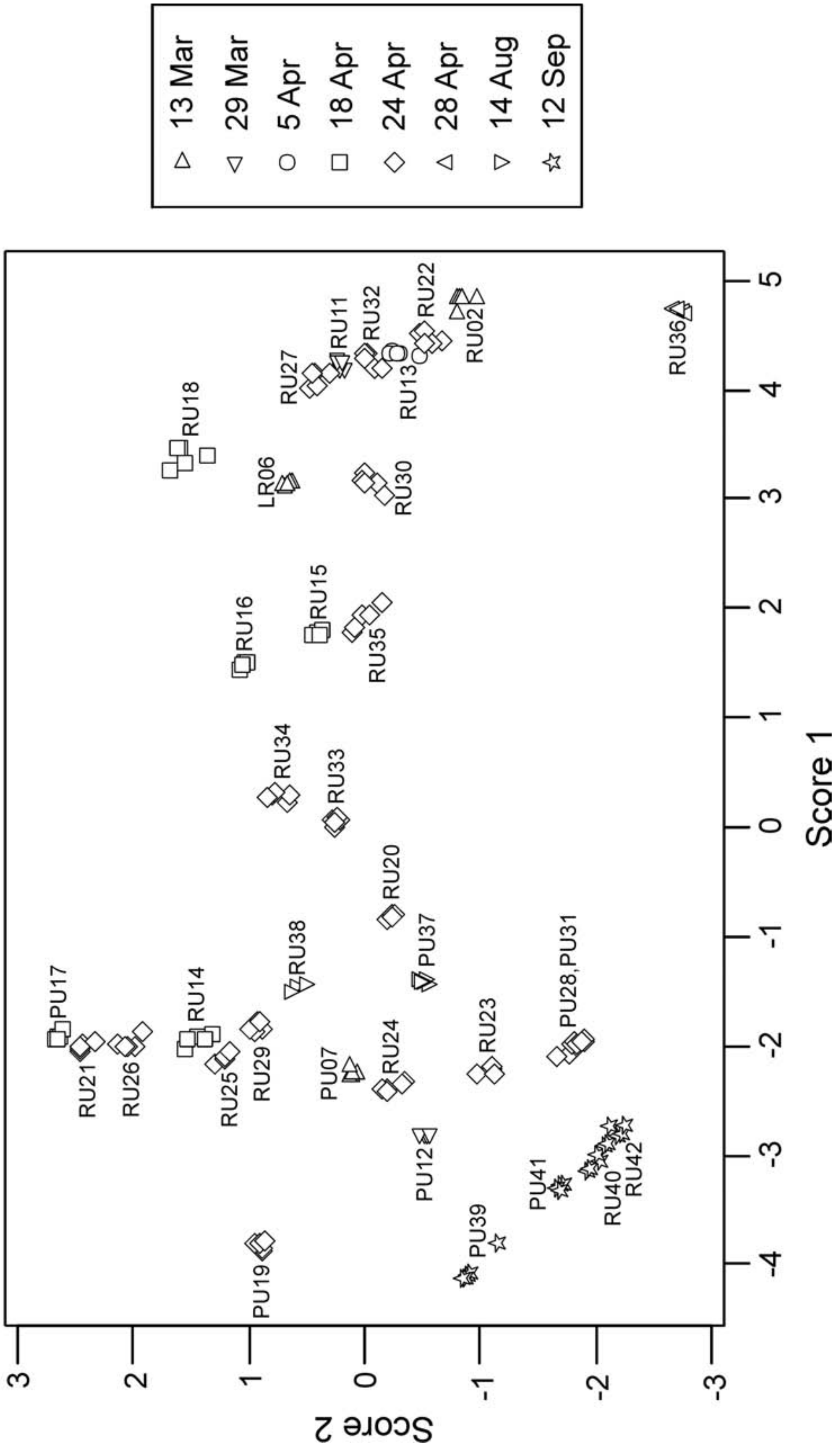


Figure 7.7 PC score plot of 175 chromatograms from 35 gasoline samples evaporated 50% by weight.

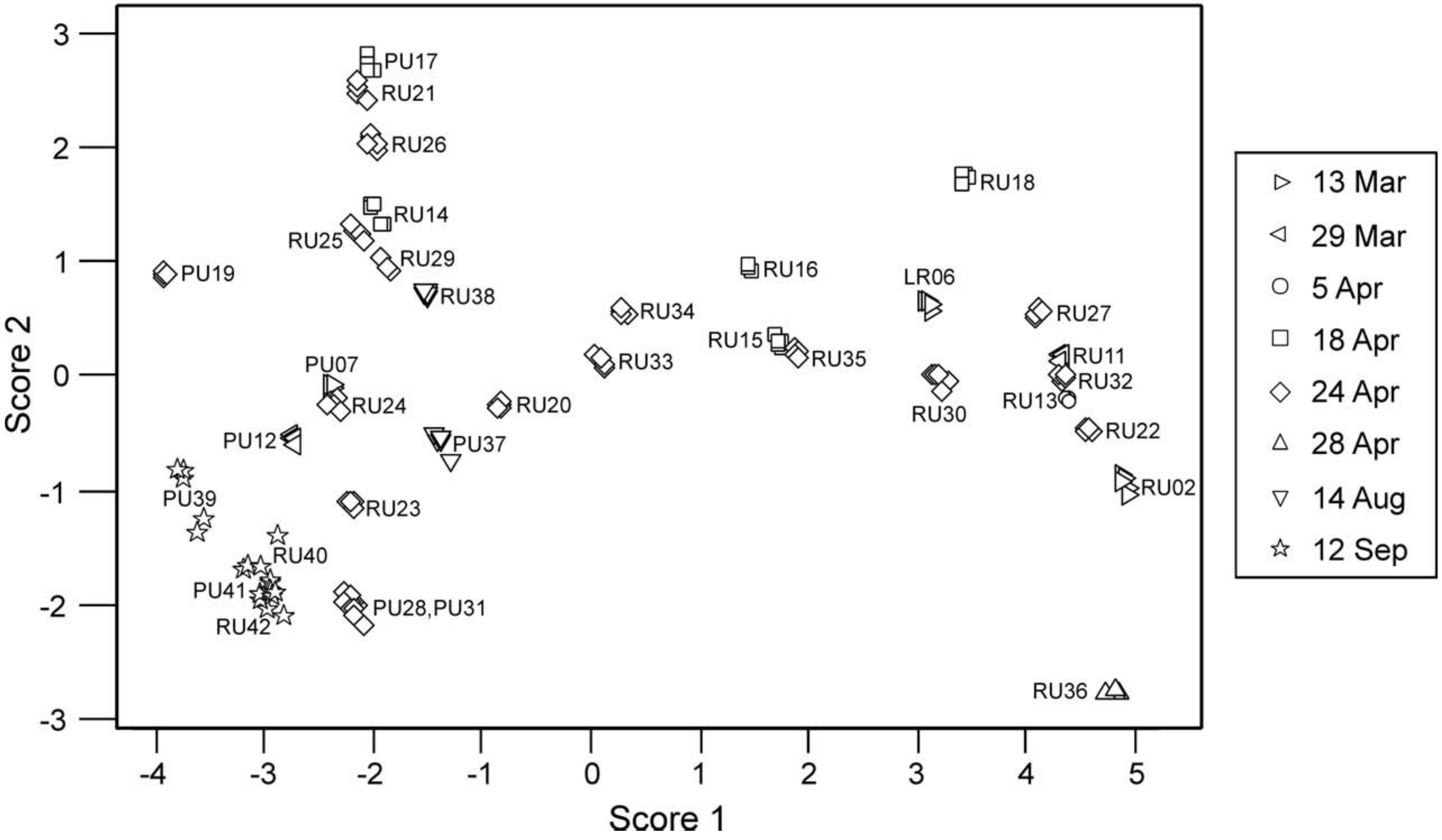


Figure 7.8 PC score plot of 175 chromatograms from 35 gasoline samples evaporated 75% by weight.

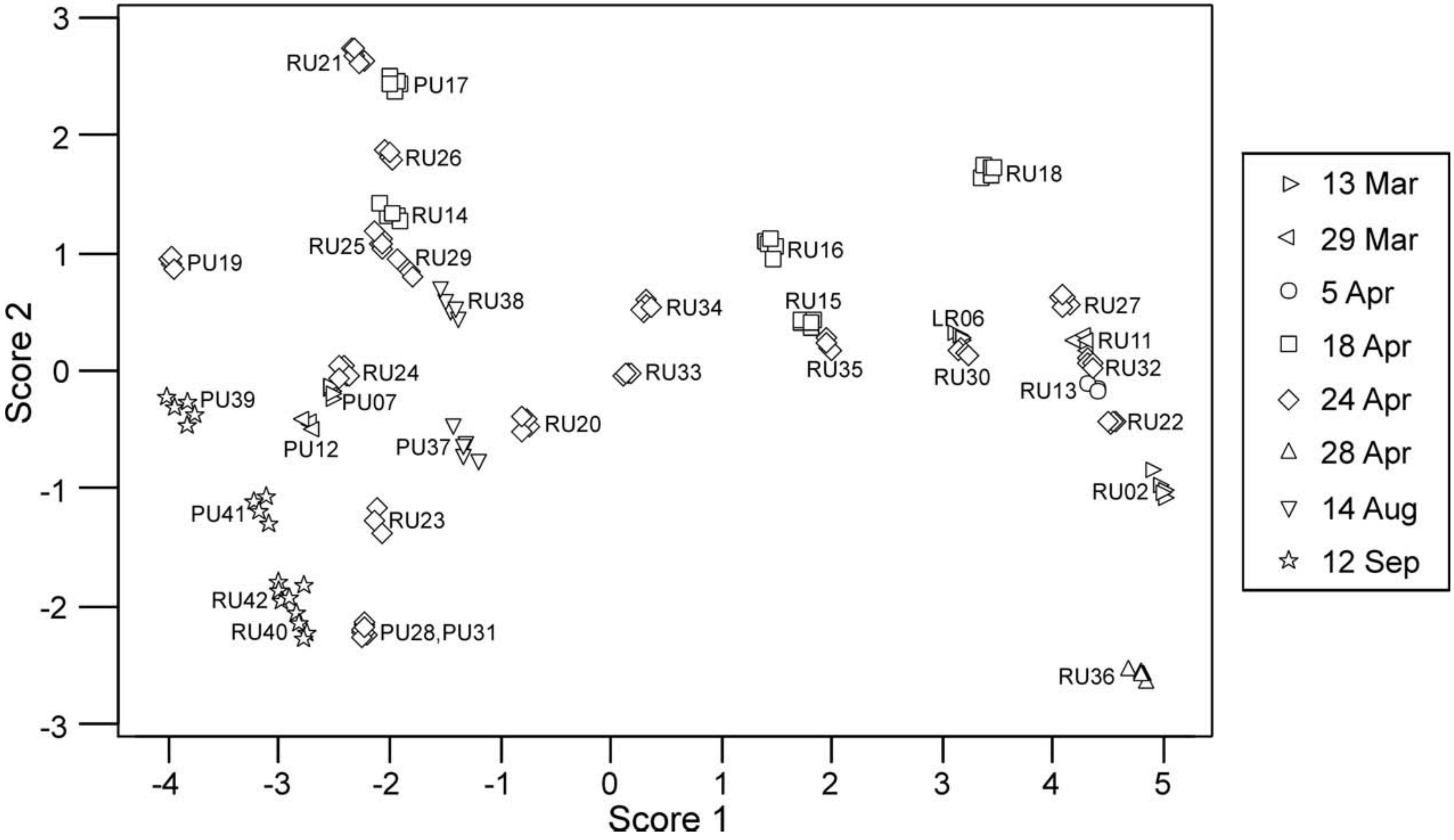
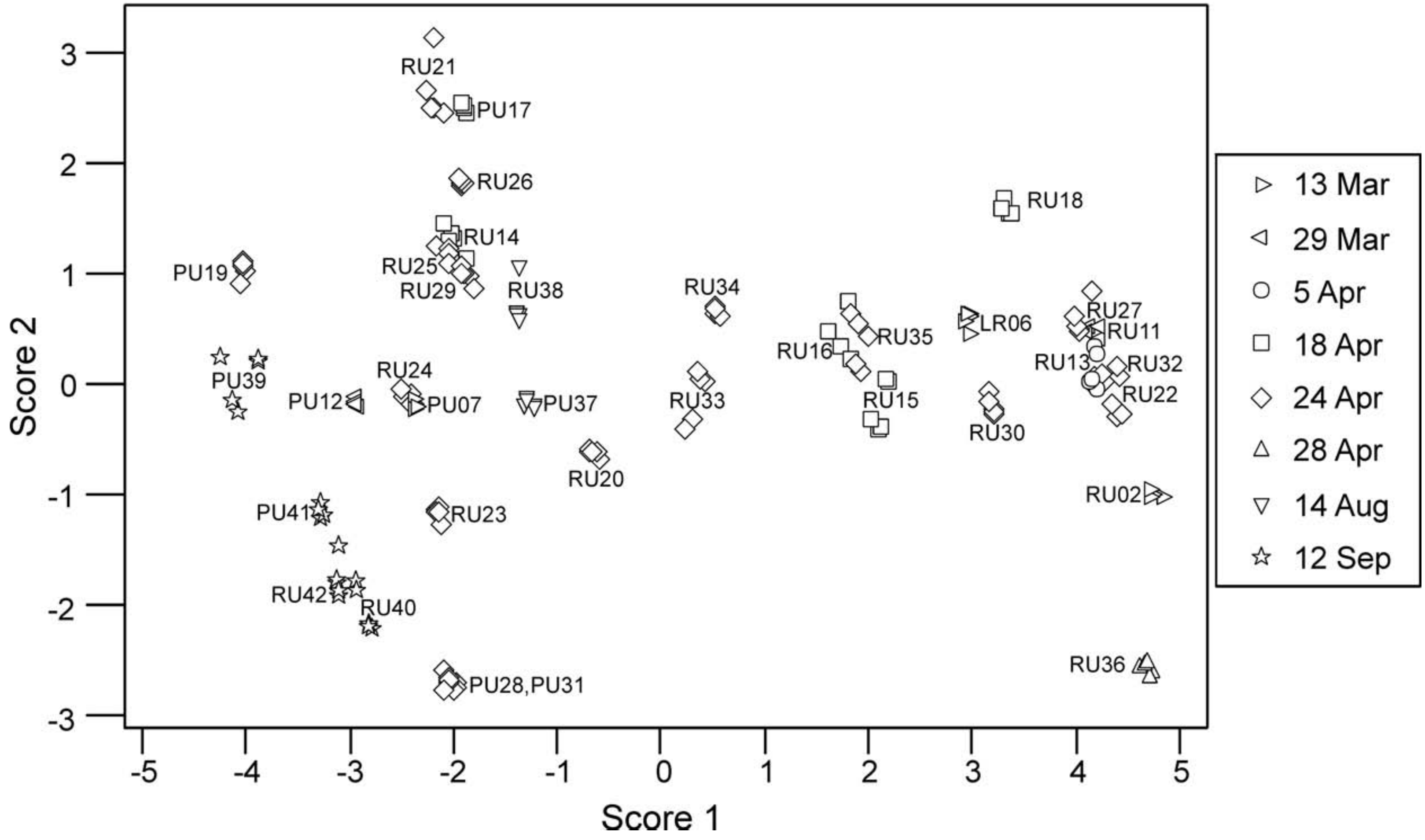


Figure 7.9 PC score plot of 175 chromatograms from 35 gasoline samples evaporated 90% by weight.



envelope). Each of these gasoline samples has been re-plotted in Figures 7.11 to 7.14 to illustrate the scatter of the individual data points within each sample (*i.e.* five aliquots at each of the five levels of evaporation). Figure 7.11 shows the scatter of the PC score data for two small sample clusters, samples PU17 and RU34 as defined by the confidence envelopes. Figure 7.12 illustrates the scatter of the PC score data for samples RU25 and RU26. These two samples exhibit a relatively high degree of data homogeneity between evaporation levels. Figures 7.13 and 7.14 show PC score data for samples RU13, RU36 and PU39 which illustrate heterogeneity within a sample due to clustering of aliquots by evaporation level. The samples in Figures 7.13 and 7.14 also provide examples of data scatter that result in large sample clusters.

Linear discriminant analysis (with cross-validation) was performed on the first three PC scores obtained at each evaporation level. The number of aliquots misclassified by LDA is shown for each evaporation level in Table 7.4. The grouping of samples based on the overlapping of data points (*i.e.* aliquots) from one sample onto another is very consistent for the 25%, 50% and 75% evaporated samples (Table 7.4). A difference in sample grouping is observed for the unevaporated samples compared with the 25%, 50% and 75% evaporated samples. A large difference in sample grouping is observed for the 90% evaporated samples compared to the less evaporated samples. This suggests that at 90% evaporation, information in the C_0- to C_2-naphthalene profile has been lost resulting in fewer differences between samples.

Figure 7.10 PC score plot of selected samples illustrating small scatter of data points (PU17, RU25, RU26, RU34), homogeneity of evaporation levels within a sample (RU25, RU26), heterogeneity of evaporation levels within a sample (RU13, RU36), and large scatter of data points (RU13, PU39). See text and figures of individual samples for more detail. 95% error ellipses shown.

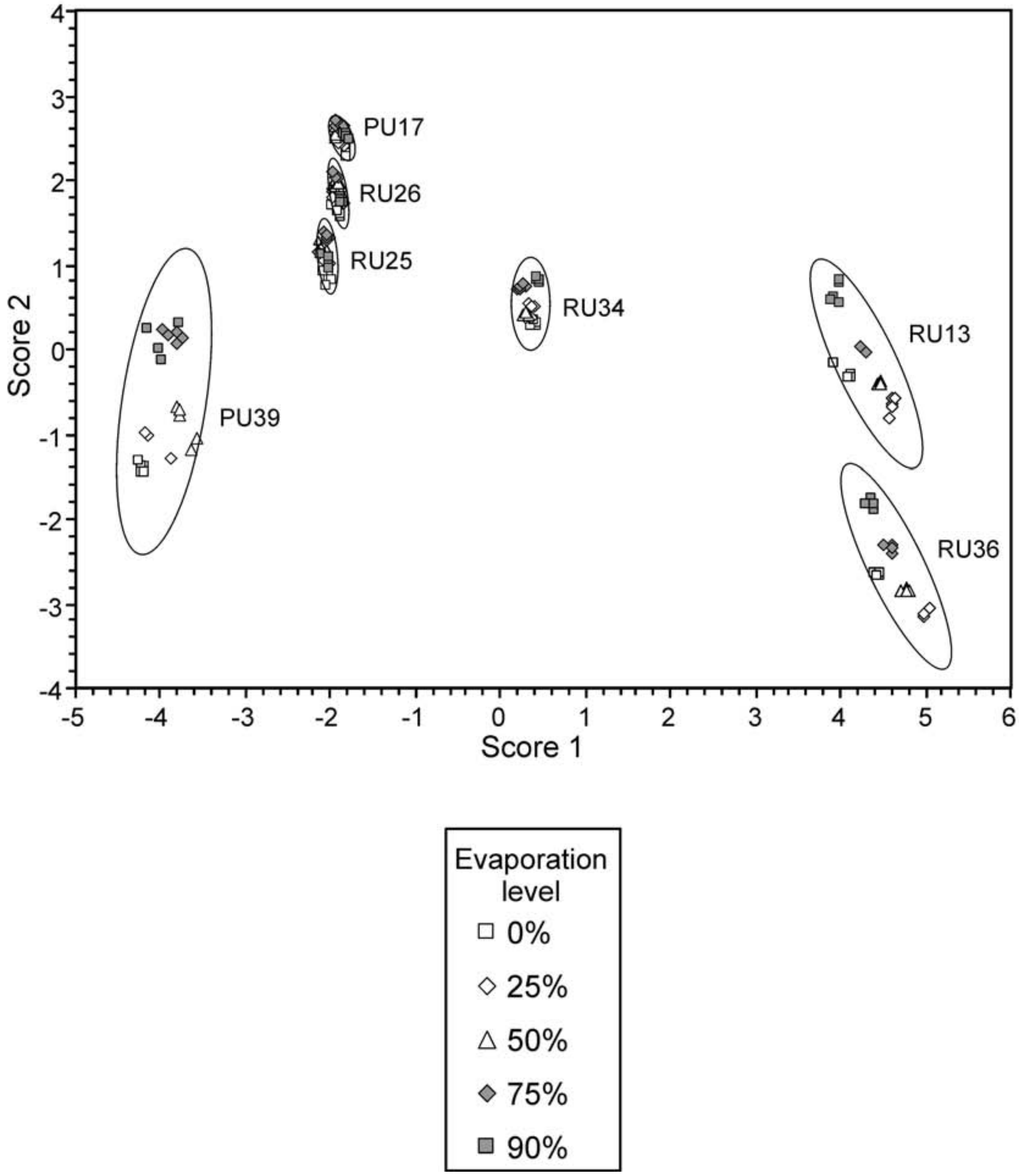


Figure 7.11 PC score plots of unevaporated (0%) and evaporated (25%, 50%, 75% and 90% by weight) samples PU17 and RU34. Plots illustrate a typical small scatter of data points. 95% error ellipses shown.

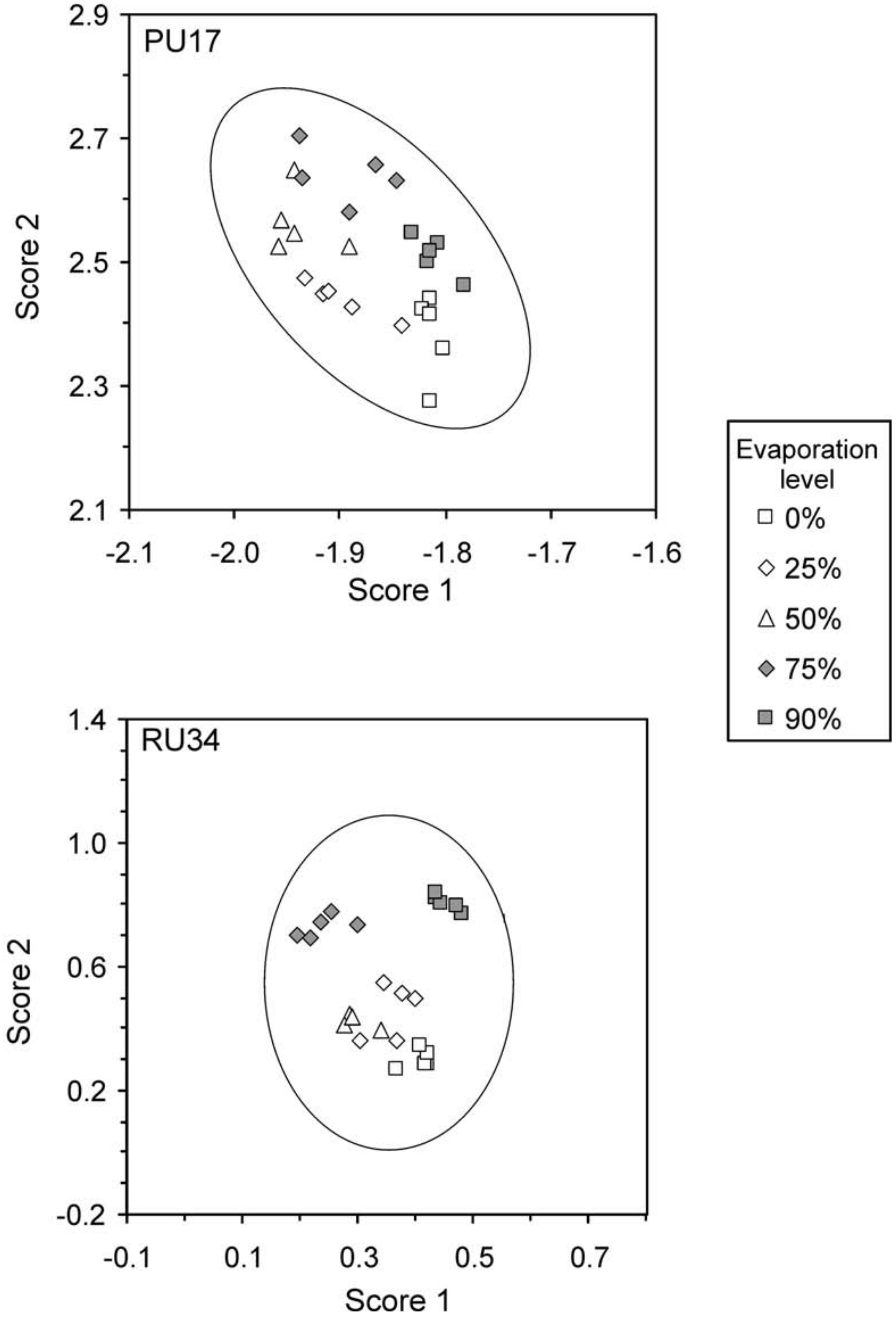


Figure 7.12 PC score plot of unevaporated (0%) and evaporated (25%, 50%, 75% and 90% by weight) samples RU25 and RU26. Plots illustrate a typical small scatter of data points and homogeneity of evaporation levels within a sample. 95% error ellipses shown.

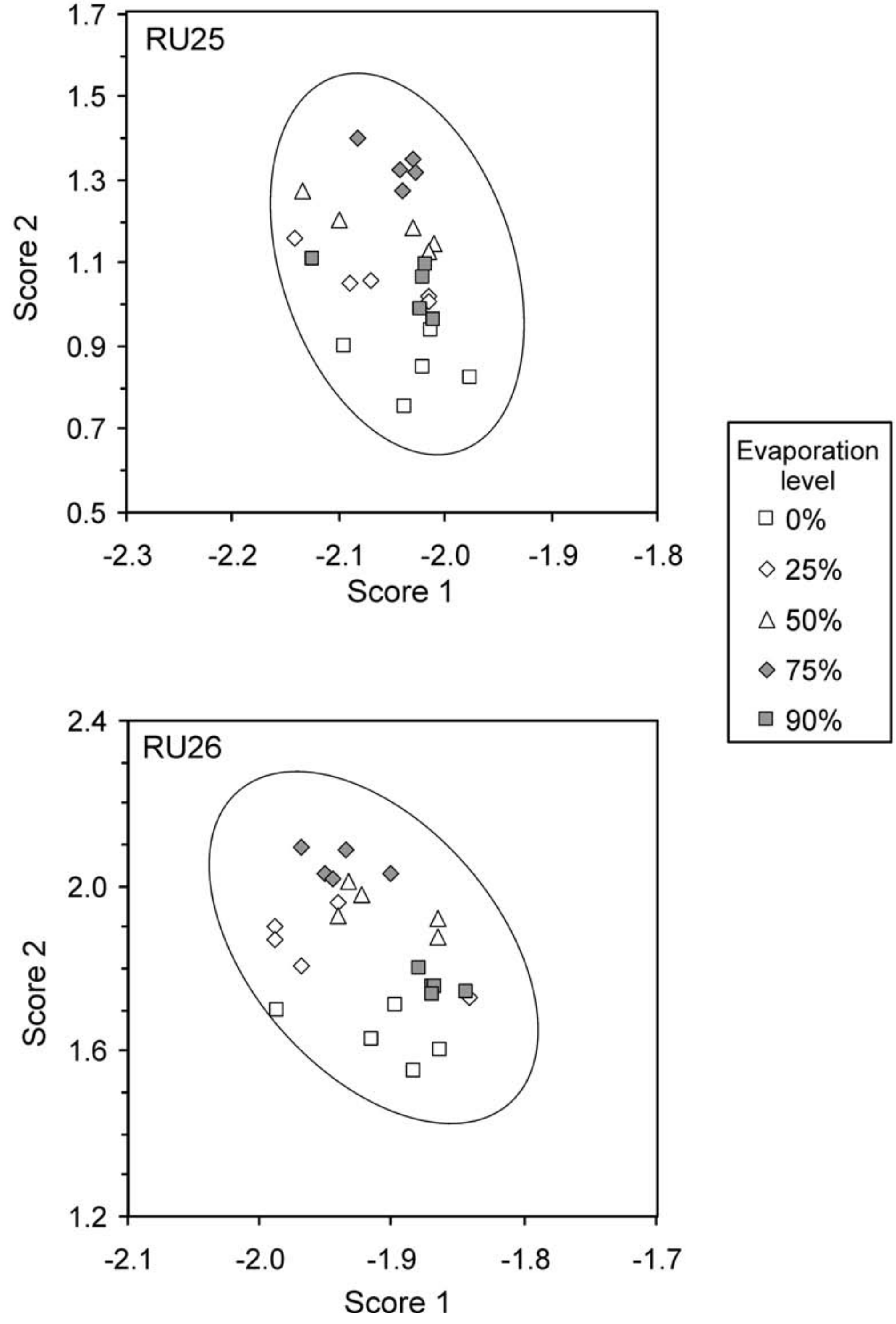


Figure 7.13 PC score plot of unevaporated (0%) and evaporated (25%, 50%, 75% and 90% by weight) samples RU13 and RU36. Plots illustrate heterogeneity of evaporation levels within a sample. 95% error ellipses shown.

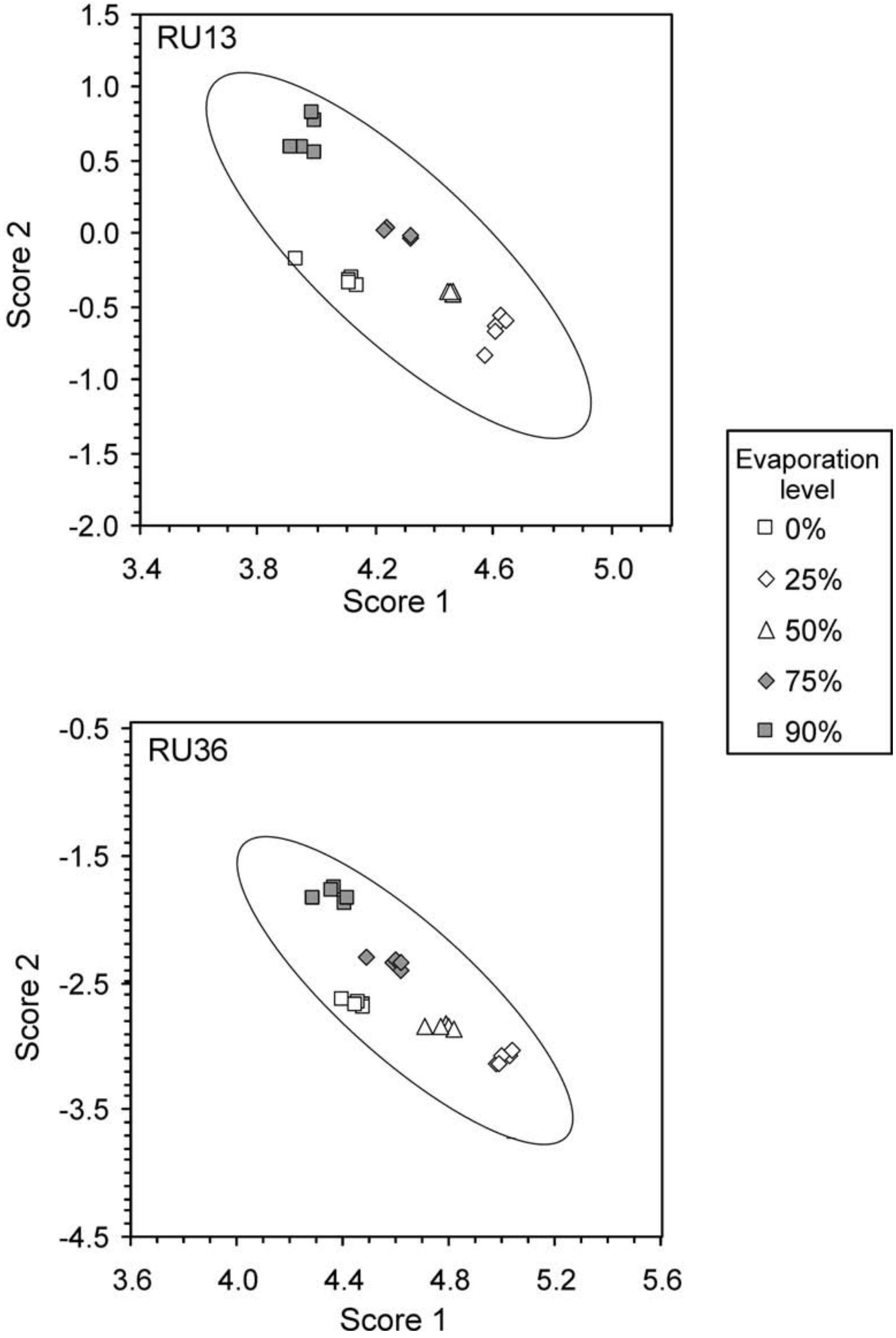


Figure 7.14 PC score plots of unevaporated (0%) and evaporated (25%, 50%, 75% and 90% by weight) samples RU13 and PU39. Plots illustrate a typical large scatter of data points. 95% error ellipses shown.

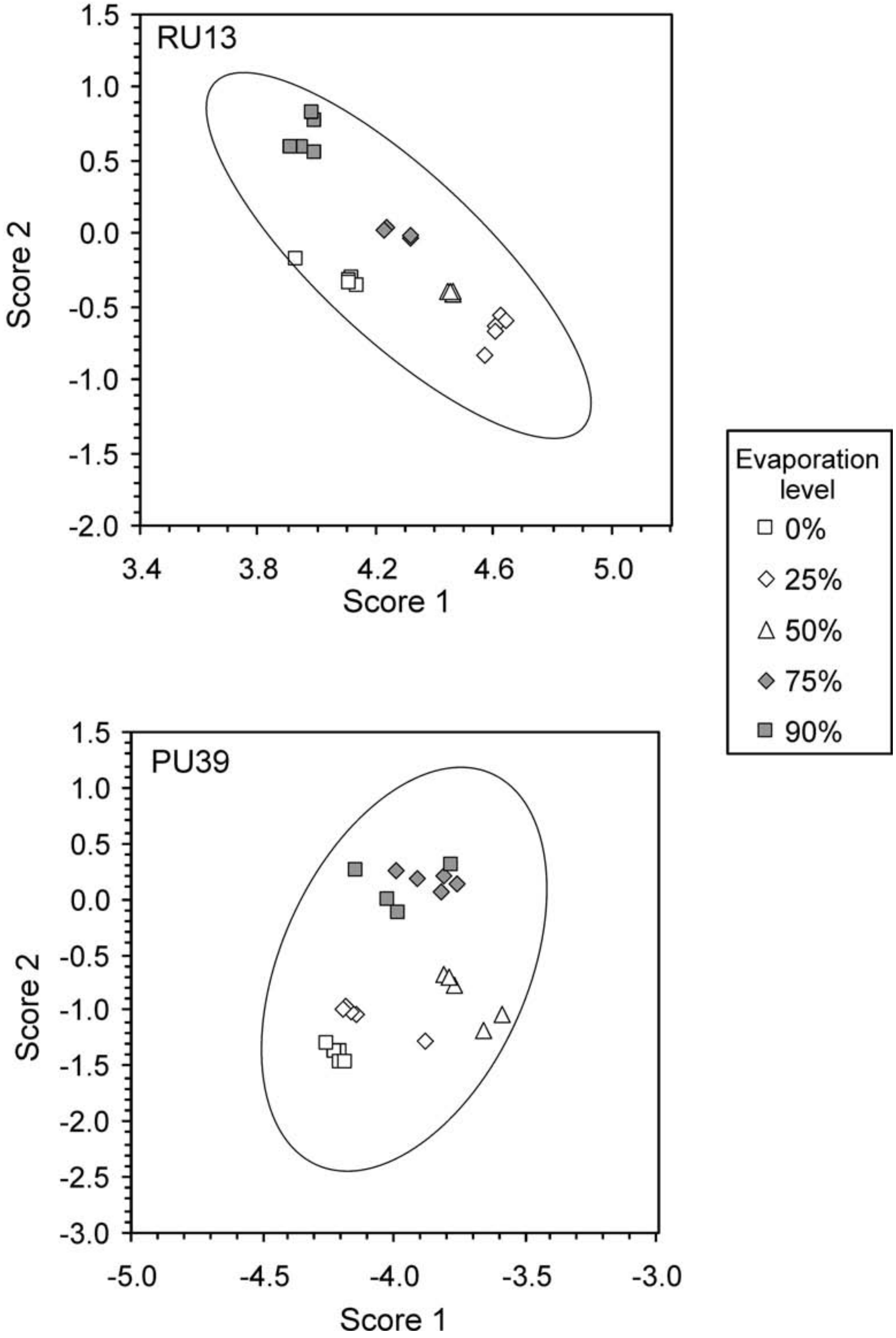


Table 7.4 Results of LDA with cross-validaton: number of aliquots from misclassified gasoline samples at difference evaporation levels.

unevaporated gasoline

		True Group				
		RU15	RU35	RU40	PU41	RU42
Predicted Group	RU15	3	1	0	0	0
	RU35	2	4	0	0	0
	RU40	0	0	3	0	1
	PU41	0	0	0	5	1
	RU42	0	0	2	0	3

25% evaporated gasoline

		True Group			
		PU28	PU31	RU40	RU42
Predicted Group	PU28	4	1	0	0
	PU31	1	4	0	0
	RU40	0	0	4	1
	RU42	0	0	1	4

50% evaporated gasoline

		True Group			
		PU28	PU31	RU40	PU41
Predicted Group	PU28	5	2	0	0
	PU31	0	3	0	0
	RU40	0	0	4	0
	PU41	0	0	1	5

75% evaporated gasoline

		True Group			
		PU28	PU31	RU40	RU42
Predicted Group	PU28	4	0	0	0
	PU31	1	5	0	0
	RU40	0	0	4	0
	RU42	0	0	1	5

90% evaporated gasoline

		True Group							
		PU28	PU31	RU14	RU29	RU11	RU13	RU27	RU32
Predicted Group	PU28	4	0	0	0	0	0	0	0
	PU31	1	5	0	0	0	0	0	0
	RU14	0	0	4	0	0	0	0	0
	RU29	0	0	1	5	0	0	0	0
	RU11	0	0	0	0	5	1	1	0
	RU13	0	0	0	0	0	2	0	1
	RU27	0	0	0	0	0	0	4	0
	RU32	0	0	0	0	0	2	0	4

7.3.2 Detailed examination of results for gasoline samples collected on two different days

The use of C_0- to C_2-naphthalenes to discriminate gasoline samples, irrespective of their level of evaporation, can be explored in more detail by examining samples taken on the same day from service stations in the same geographical area. On April 18, 2001 five samples of gasoline were collected from four service stations in one Sydney, Australia suburb (Table 7.1). The distance between the stations that were furthest apart was approximately 800 metres. The PCA results are given in Table 7.5. The plot of the first two PC scores for the 25 chromatograms from these five samples at five different levels of evaporation is shown in Figure 7.15. The five gasoline samples form five groups, one sample per group. Using the first three PCs, linear discriminant analysis (with cross-validation) showed that samples RU15 and RU16 do not have any overlapping aliquots and so these two samples each form their own group.

Table 7.5 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): five gasoline samples collected on April 18, 2001 in one Sydney, Australia suburb.

	Eigenvalue	Proportion	Cumulative (%)
PC1	9.6006	0.873	87.3
PC2	0.9559	0.087	96.0
PC3	0.3143	0.029	98.8
PC4	0.0859	0.008	99.6
PC5	0.0152	0.001	99.7

On April 24, 2001, 17 gasoline samples were collected from 14 different service stations within a 7 km radius of each other in south central Sydney (Table 7.1). These 17 gasoline samples were analysed at the five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight) resulting in 425 chromatograms. The PCA results are given in Table 7.6 and in a plot of the first two PC scores in Figure 7.16. Linear discriminant analysis (with cross-validation) showed that the 17 gasoline samples formed 11 groups. Six of these groups each contained a single sample (*i.e.* 25 aliquots from one sample per group), while the 11 remaining (misclassified) samples were divided between another five groups (Table 7.7).

Figure 7.15 PC score plot of five gasoline samples at five evaporation levels (25%, 50%, 75% and 90% by weight) collected from four service stations in one Sydney, Australia suburb on April 18, 2001. Regular unleaded sample RU16 and premium unleaded sample PU17 are from the same station. 95% error ellipses are shown.

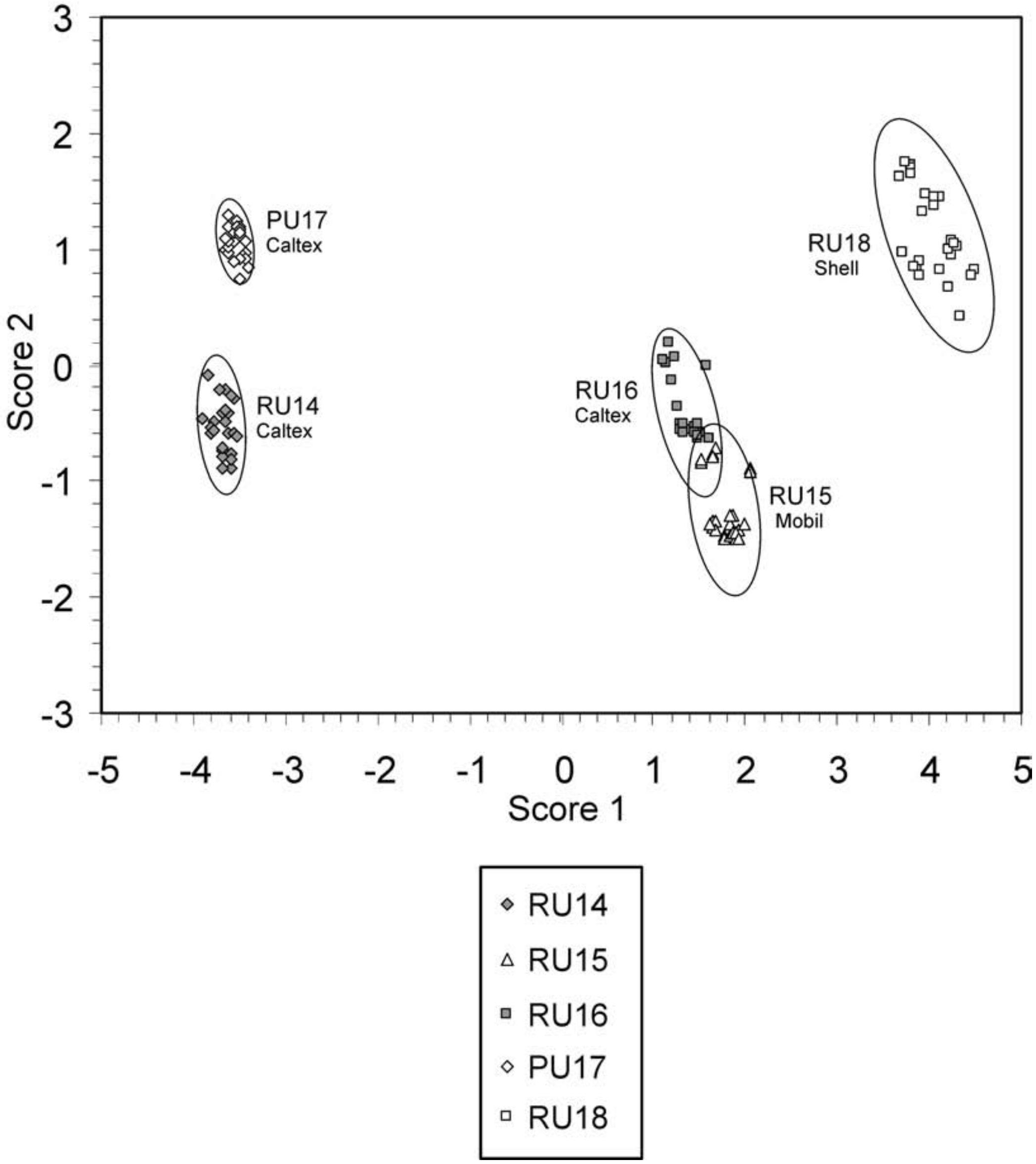


Table 7.6 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): seventeen gasoline samples collected on April 24, 2001 in south central Sydney.

	Eigenvalue	Proportion	Cumulative (%)
PC1	8.4529	0.768	76.8
PC2	1.7602	0.160	92.8
PC3	0.5505	0.050	97.9
PC4	0.1215	0.011	99.0
PC5	0.0750	0.007	99.6

Table 7.7 Results of LDA with cross-validation: number of aliquots at five evaporation levels (0%, 25%, 50%, 75% and 90% by weight) from the eleven misclassified gasoline samples.

		True Group					
		RU21	RU26	RU25	RU29	PU28	PU31
Predicted Group	RU21	24	1	0	0	0	0
	RU26	1	24	0	0	0	0
	RU25	0	0	25	1	0	0
	RU29	0	0	0	24	0	0
	PU28	0	0	0	0	14	11
	PU31	0	0	0	0	11	14

		True Group				
		RU22	RU27	RU32	RU33	RU34
Predicted Group	RU22	20	0	0	0	0
	RU27	2	22	5	0	0
	RU32	3	3	20	0	0
	RU33				23	0
	RU34				2	25

Linear discriminant analysis did not indicate any misclassification of aliquots (chromatograms) between samples RU24 and RU23, nor did it indicate any misclassification between sample RU23 and the group of two samples, PU28 and PU31; results that are not readily apparent from the PC score plot in Figure 7.16. These four samples form three distinct groups: RU23; RU24; and, PU28 and PU31. When the third principal component is included, as it was for LDA, the separation of these samples in the third dimension can be seen (Figure 7.17).

Figure 7.16 PC score plot of 17 gasoline samples at five levels of evaporation (0%, 25%, 50%, 75% and 90% by weight) collected from 14 service stations in south central Sydney on April 24, 2001. Premium and regular samples collected from the same service stations are as follows: PU19 and RU20; PU31 and RU32; and, RU27 and PU28. 95% error ellipses are shown.

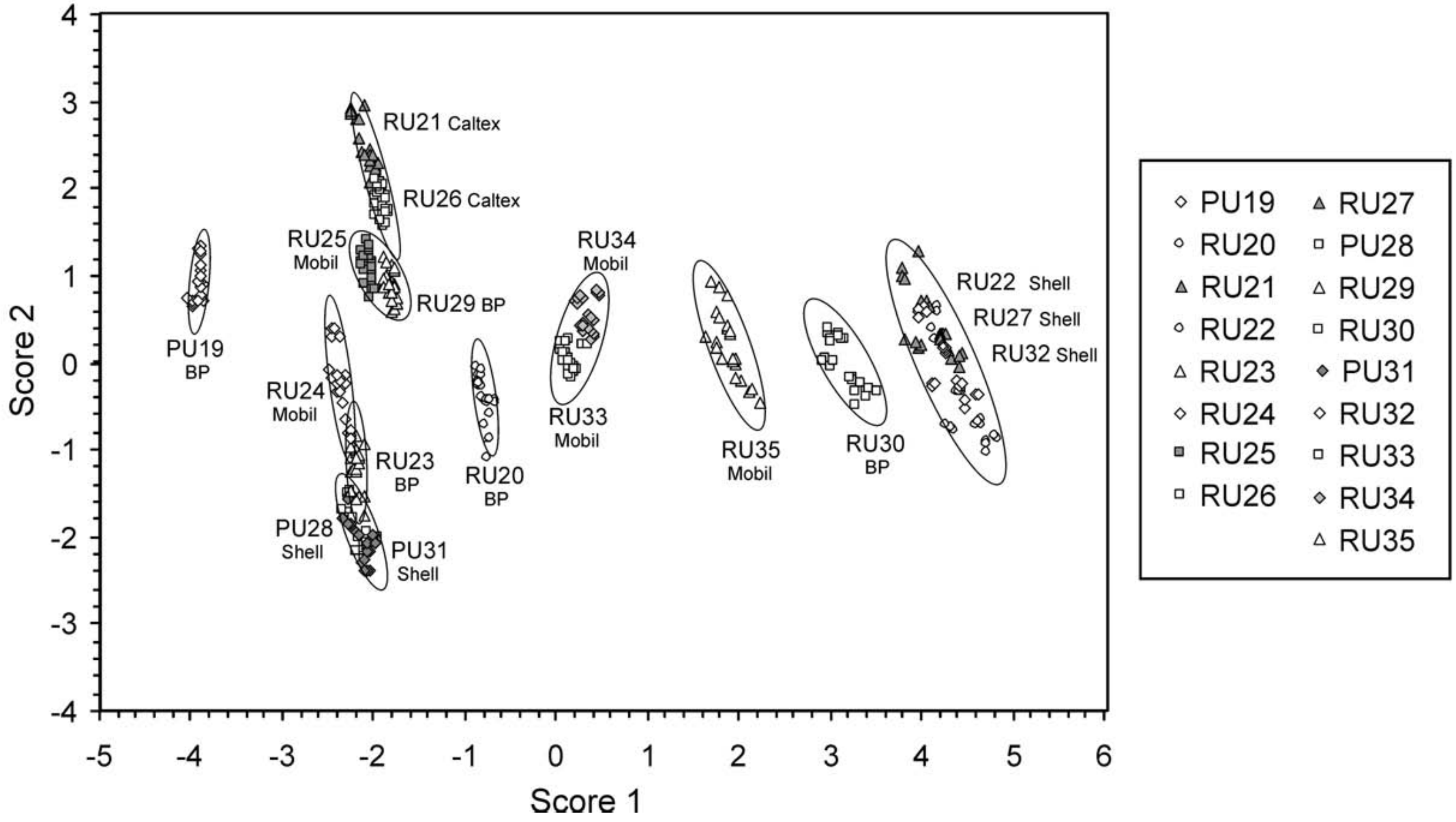
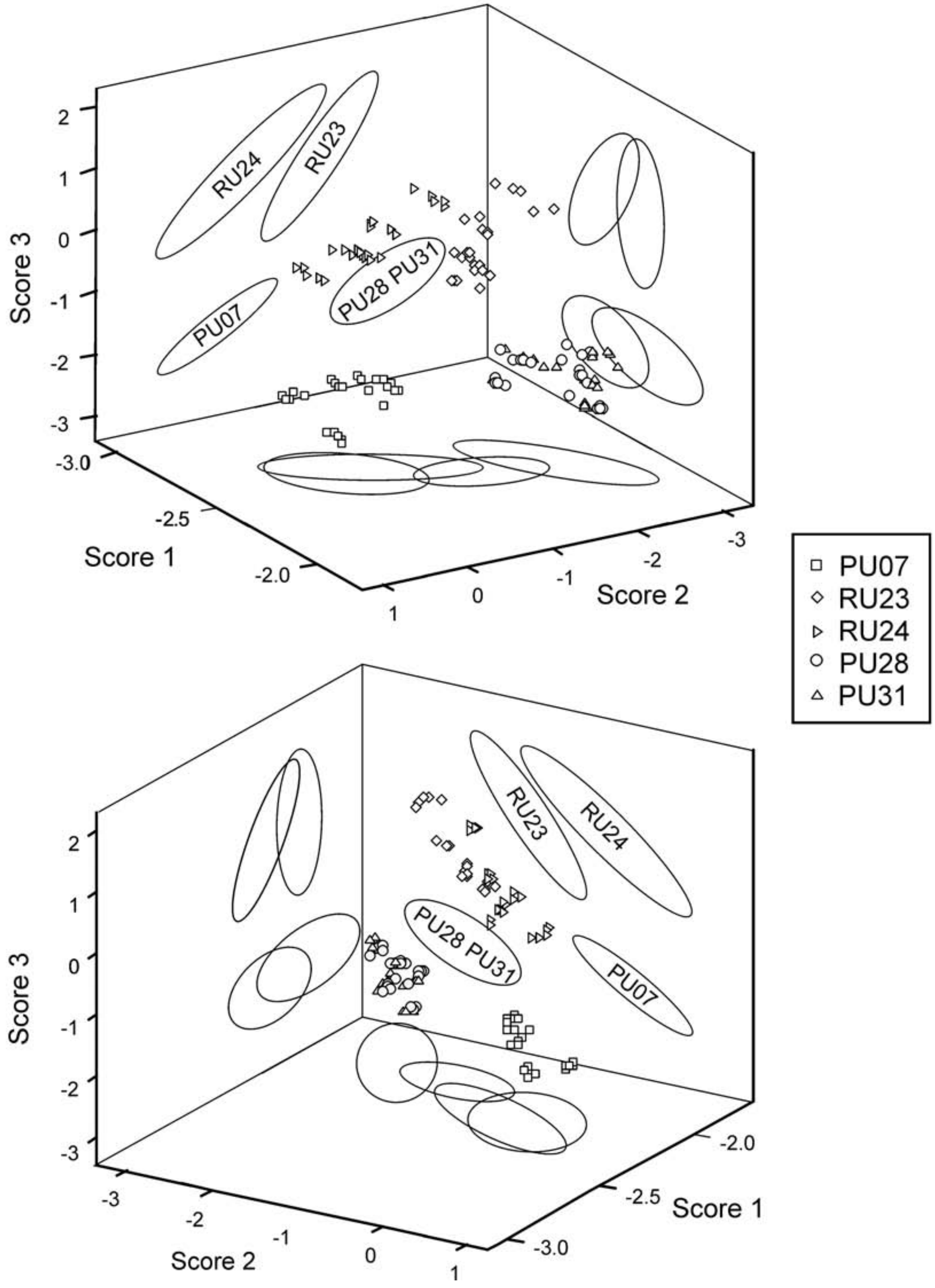


Figure 7.17 PC score plots of PU07, RU23, RU24, PU28 and PU31 demonstrating the separation of groups in the third dimension (score 3). 95% error ellipses are projected onto the back planes of the plots.



7.3.3 Examination of all gasoline samples at all evaporation levels

Linear discriminant analysis (with cross-validation) was performed on the first three PC scores for all 35 gasoline samples at all five evaporation levels (*n* = 875). The 35 gasoline samples formed 18 groups. Eleven of these groups each contained a single sample (*i.e.* 25 aliquots from one sample per group) while the remaining 24 samples were divided between the remaining seven groups (Table 7.8).

Linear discriminant analysis was used in conjunction with visual inspection of three-dimensional plots of the first three PC scores to demonstrate group membership for selected samples. The 95% error ellipses (confidence envelopes) were calculated for these samples and projected onto the plots to aid in the demonstration. The three-dimensional PC score plots of one group of two samples, RU33 and RU34, is shown in Figure 7.18. In the top orientation of this plot, it can be seen that three aliquots of sample RU33 are well separated from the cluster of sample RU33 data points and are very close to the cluster of sample RU34 data points. These three data points are three aliquots of 90% evaporated sample RU33. In this case, the grouping together of samples RU33 and RU34 by LDA (Table 7.8, top) is readily observed in the three-dimensional PC score plots. The clustering of samples RU02, RU11, RU13, RU22, RU27 and RU32 into a single group was also indicated by LDA (Table 7.8, bottom). The three-dimensional PC score plots for these samples (Figure 7.19) includes sample RU18 for comparison. Although sample RU02 is on the periphery of the cluster formed by samples RU11, RU13, RU22, RU27 and RU32, it can be seen why LDA included sample RU02 within this cluster (Figure 7.19). Samples LR06 and RU30 were also grouped together by LDA (Table 7.8, top) and the PC score plots in Figure 7.20 demonstrate how similar these two samples are in three-dimensional space. Linear discriminant analysis indicated that samples PU39, RU40, PU41 and RU42 belonged to the same group (Table 7.8, middle). For comparison purposes, sample PU12 was included in the PC score plots for samples PU39, RU40, PU41 and RU42 (Figure 7.21) because it is chemically similar to the other four samples. Although there are three groups visible in Figure 7.21, two data points from sample PU39 were assigned by LDA as being closer to

Table 7.8 Results of LDA with cross-validation: number of aliquots at five evaporation levels (0%, 25%, 50%, 75% and 90% by weight) from the 24 misclassified samples.

	True Group								
Predicted Group	LR06	RU30	PU28	PU31	RU15	RU16	RU35	RU33	RU34
LR06	24	5	0	0	0	0	0	0	0
RU30	1	20	0	0	0	0	0	0	0
PU28	0	0	14	11	0	0	0	0	0
PU31	0	0	11	14	0	0	0	0	0
RU15	0	0	0	0	18	1	7	0	0
RU16	0	0	0	0	2	24	2	0	0
RU35	0	0	0	0	5	0	16	0	0
RU33	0	0	0	0	0	0	0	24	0
RU34	0	0	0	0	0	0	0	1	25

	True Group								
Predicted Group	RU14	RU21	RU25	RU26	RU29	PU39	RU40	PU41	RU42
RU14	20	24	1	1	1	0	0	0	0
RU21	0	0	0	2	0	0	0	0	0
RU25	3	0	24	0	0	0	0	0	0
RU26	0	1	0	22	0	0	0	0	0
RU29	2	0	0	0	24	0	0	0	0
PU39	0	0	0	0	0	23	0	0	4
RU40	0	0	0	0	0	0	19	1	8
PU41	0	0	0	0	0	2	1	19	0
RU42	0	0	0	0	0	0	5	5	18

	True Group					
Predicted Group	RU02	RU11	RU13	RU22	RU27	RU32
RU02	21	0	0	0	0	0
RU11	0	9	1	1	4	5
RU13	0	4	10	3	19	5
RU22	4	0	10	18	0	0
RU27	0	5	4	1	0	3
RU32	0	7	0	1	2	12

sample PU41 (Table 7.8, middle). Linear Discriminant analysis examines each data point separately and assigns a data point to a group based on the shortest Mahalanobis distance. Thus, the large scatter in the data for sample PU39 caused this sample to be grouped with samples RU40, PU41 and RU42. Samples RU15, RU16 and RU35 formed another group, as did samples PU28 and PU31 (Table 7.8, top). The intermingling of data points from one sample with data points from another sample within these two groups is significant and so no three-dimensional PC score plots for either of these two groups have been reproduced here. A PC score plot showing the 18 groups that were identified based on linear discriminant analysis of the 875 chromatograms from the 35 gasoline samples at each of the five evaporation levels (0%, 25%, 50%, 75% and 90% by weight) is shown in Figure 7.22.

Figure 7.18 PC score plots of samples RU33 and RU34 demonstrating that these two samples form one group. 95% error ellipses are projected onto the back planes of the plots.

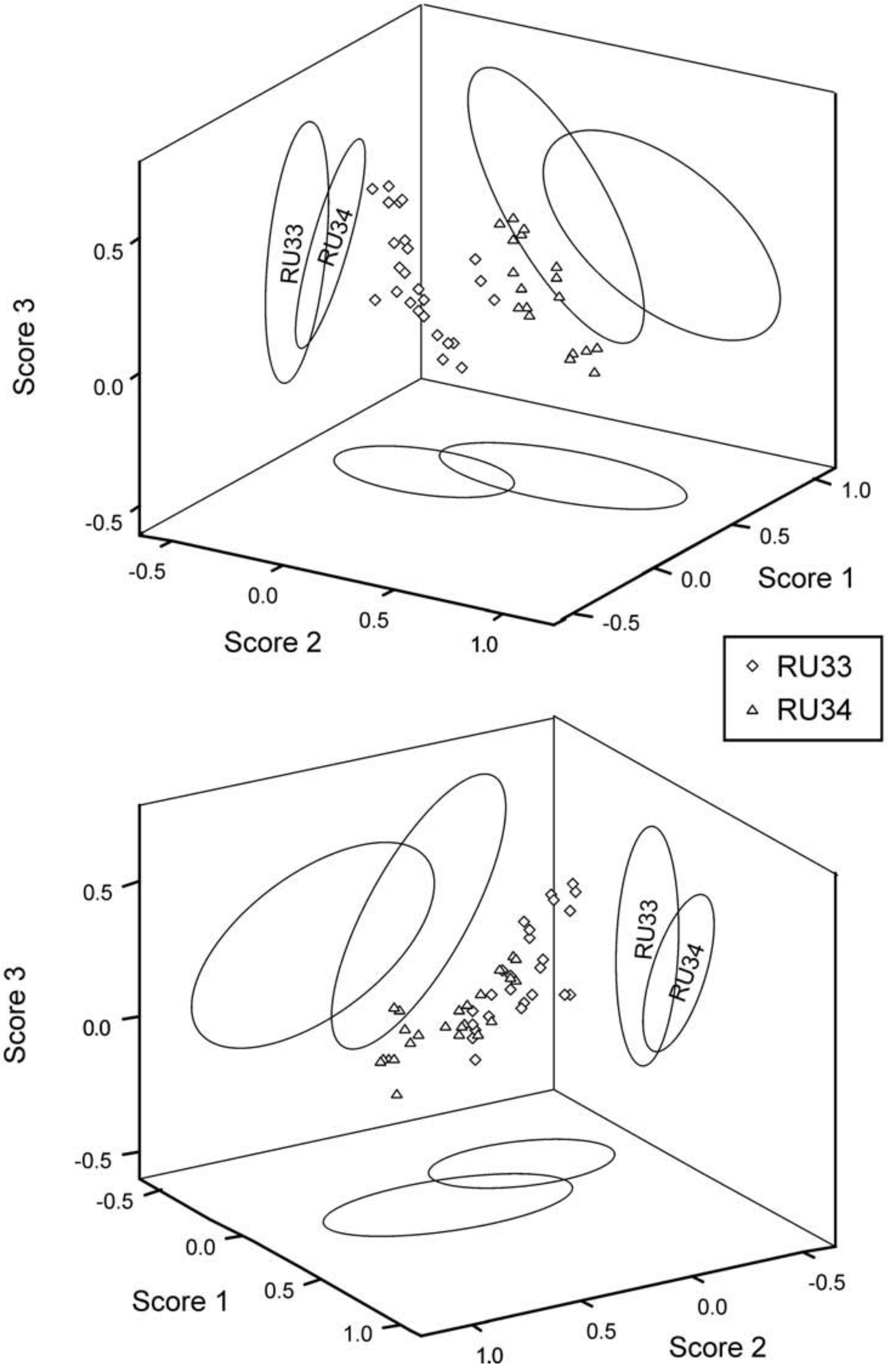


Figure 7.19 PC score plots of RU02, RU11, RU13, RU22, RU27 and RU32 demonstrating that these six samples form one group. Sample RU18 has been included for comparison. 95% error ellipses are projected onto the back planes of the plot.

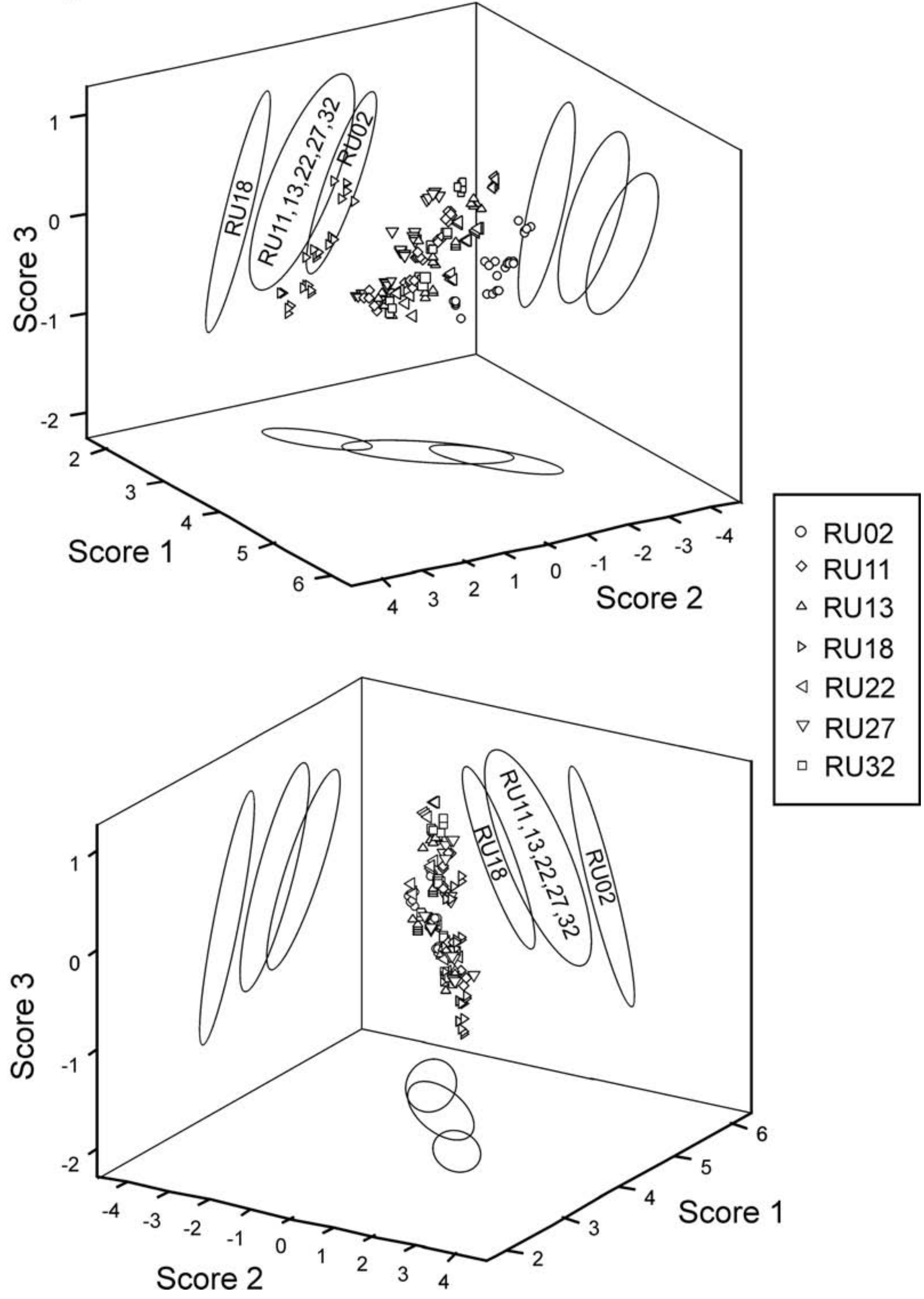


Figure 7.20 PC score plots of samples LR06 and RU30 demonstrating that these two samples form one group. 95% error ellipses are projected onto the back planes of the plots.

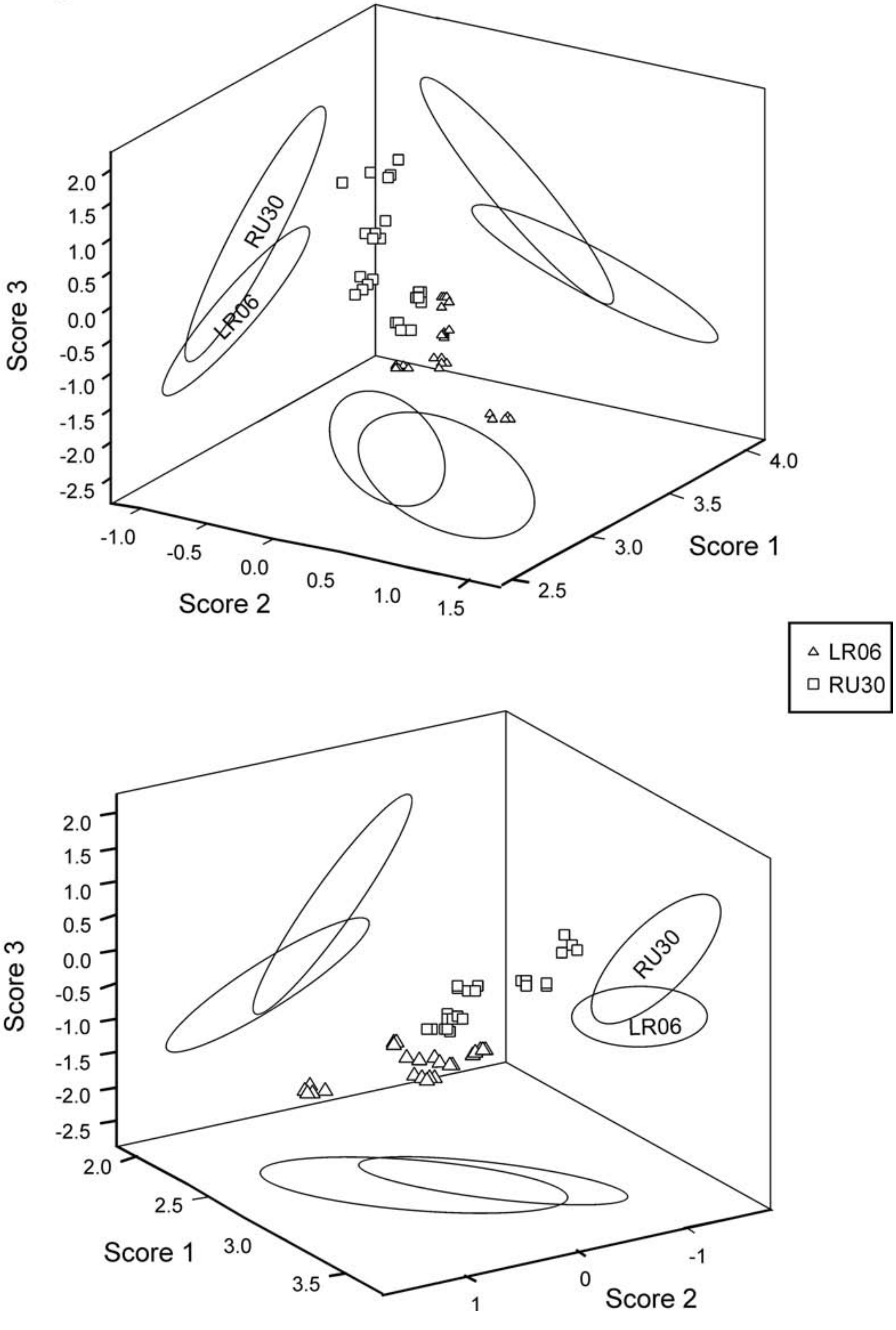


Figure 7.21 PC score plots of samples PU39, RU40, PU41, RU42 demonstrating that these four samples form one group. Sample PU12 has been included for comparison. 95% error ellipses are projected onto the back planes of the plots.

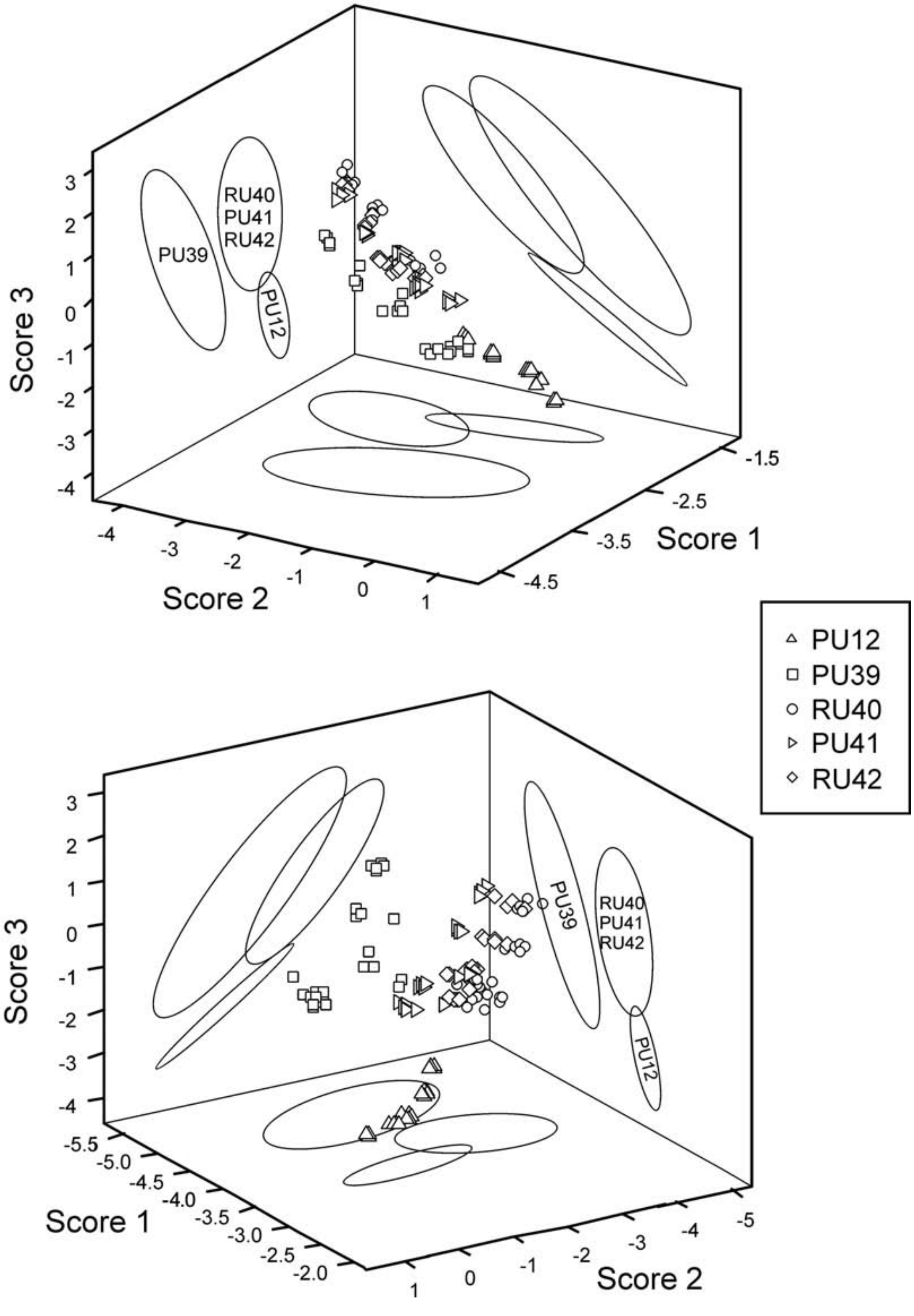
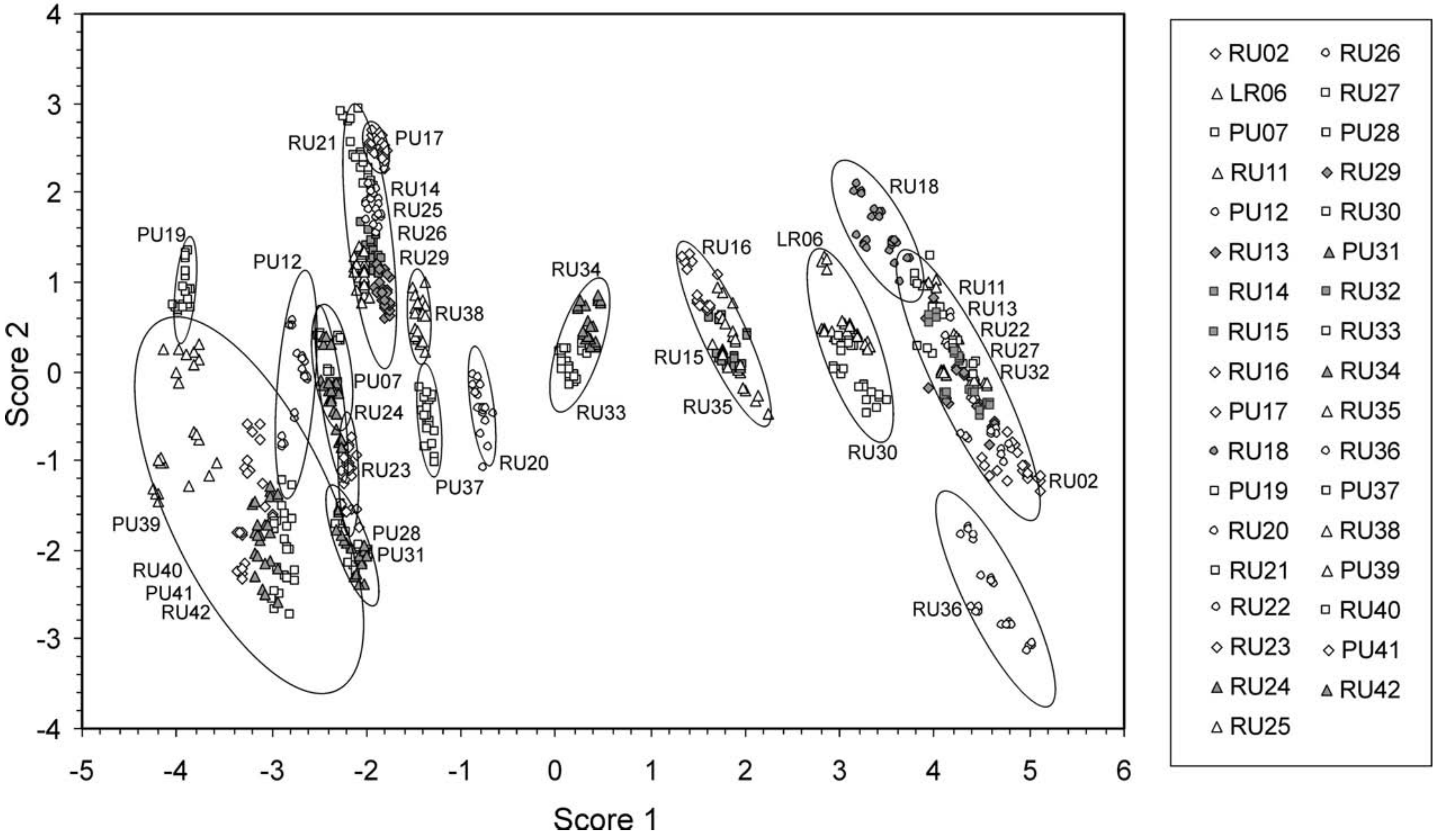


Figure 7.22 PC score plot of 35 gasoline samples at five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight). 95% error ellipses are shown. The 35 samples were divided into 18 unique groups by LDA. See text for explanation of groups.



7.4 Discussion

Sheff and Siegel[6] compared unevaporated, 50% evaporated and "100%" evaporated regular leaded, regular unleaded and premium unleaded gasoline samples using three-dimensional fluorescence spectroscopy. They found that there was little difference between fluorescence spectra of a partially evaporated gasoline sample (*i.e.* 50% evaporated) and the original, unevaporated sample. As the amount of evaporation in a sample increased so that no more liquid could be removed from the sample (*i.e.* "100%" evaporated), the fluorescence spectrum changed dramatically and comparison between an unevaporated and evaporated sample was no longer possible.[6] Differences in fluorescence between samples were caused by variations in the polycyclic aromatic hydrocarbon content.[5] Thus, the lack of change in the fluorescence plots obtained by Sheff and Siegel[6] indicated that the relative concentrations of the PAHs did not vary significantly between unevaporated and partially evaporated gasoline samples. This finding supports the use of the C_0- to C_2-naphthalene profile obtained by gas chromatography-mass spectrometry to compare unevaporated and evaporated gasoline samples.

In a comparison of unevaporated and 94% to 96% evaporated (by weight) leaded and unleaded gasoline samples, Hirz observed that "[t]he analysis of weathered samples showed very consistent relative concentrations of the high boiling components in gasoline from any one refinery."[3] Hirz did not identify which of the high boiling compounds were used for the comparison, however, Hirz was able to assign all samples to their respective refineries. All of the samples examined in this chapter, except samples PU39, RU40, PU41 and RU42, likely came from one of the two refineries operating in Sydney (Caltex Kurnell and Shell Clyde). Samples PU39, RU40, PU41 and RU42, purchased from Burmah and Woolworths (Table 7.1), are believed to have come from a refinery in Singapore.[14,15]

Potter[4] compared the C_8-alkanes profiles of six unleaded gasoline samples using principal component analysis. Potter had initially chosen several isomeric groups for study including the C_2-benzenes, the C_1-naphthalenes, the C_3-benzenes, and the C_8-alkanes. This choice was based on the rationale that the profile of an isomeric group would exhibit only small differences between an unevaporated and an evaporated sample.[4] Chemical isomers do have similar vapour pressures and so will evaporate at approximately the same rate. Mann[2] also used this concept in the comparison of evaporated and unevaporated gasoline samples. Wang and Fingas[7] examined the relative ratios of selected isomers, including the naphthalenes, in order to compare crude oil samples taken from environmental spills.

Smallwood *et al.* [9] compared the carbon isotope ratios of selected compounds from evaporated gasoline samples to those in their unevaporated parent sample. They analysed three samples from one city in the state of Oklahoma and found that at least 67% of the compounds examined in the evaporated gasoline samples remained "isotopically similar" to the corresponding compounds in the unevaporated samples. Smallwood *et al.* measured evaporation by time and not by weight or volume of sample lost. This is a serious shortcoming of the work of Smallwood *et al.*[9] not only because it made it difficult to compare samples reported in their investigation, but also because it prevented a comparison with the results presented in this thesis. The work presented in this chapter showed that all gasolines *do not* evaporate at the same rate even when placed together in the same fume cupboard at the same time. Preliminary results from Smallwood *et al.*[9] suggested that carbon isotope ratios for the C_0- and C_1-naphthalenes did not vary significantly between the evaporated samples and their original, unevaporated samples. This result is in agreement with the findings for the C_0- to C_2-naphthalenes presented in this chapter.

The fact that grouping of samples by LDA was found to be consistent for the 25%, 50% and 75% evaporated samples (Table 7.4) is a significant finding. The significance of this result is that one would be more likely to perform a forensic comparison on samples at these levels of evaporation than on samples that are either unevaporated or 90% evaporated. Evaporation of forensic samples to between 25% and 75% by weight is realistic because gasoline samples recovered from the scene or seized from the suspect are usually exposed to conditions that allow some evaporation to occur. A thorough forensic investigation would also include a number of control samples taken from local service stations and/or other sources of gasoline available to the suspect. The PCA results for five gasoline samples collected from four service stations on the same day (Figure 7.15) demonstrated that all five samples could be distinguished, irrespective of their level of evaporation. In a second study of 17 gasoline samples collected from 14 service stations on the same day (Figure 7.16), PCA and LDA results showed that the samples, irrespective of their levels of evaporation, could be divided into 11 groups. While it is always best to compare gasoline samples that have similar levels of evaporation, these results showed that a meaningful comparison between samples is possible even when the samples have not undergone exactly the same amount of evaporation. A comparison between a large number of gasoline samples collected over a seven month period can also be successfully made. The comparison of 35 gasoline samples at five different levels of evaporation (0%, 25%, 50%, 75% and 90% by weight), showed that these samples could be divided by LDA into 18 different groups based on their C_0- to C_2-naphthalene profile alone (Table 7.8 and Figure 7.22). Inspection of the three-dimensional score plots for some of these 18 groups (Figures 7.17 to 7.21) demonstrates that LDA is able to elicit true statistical differences between groups, even in a large, complex data set.

7.5 Conclusions

The value of using the C_0- to C_2-naphthalene profiles to compare unevaporated and evaporated gasoline samples has been demonstrated. The division of gasoline samples into unique groups is possible even when the samples span a wide range of evaporation levels (0%, 25%, 50%, 75% and 90% by weight). The results presented in this chapter show that a forensic comparison of gasoline samples collected on the same day can be made even if the samples do not have exactly the same level of evaporation. Although possible, an attempt to compare an unevaporated gasoline sample to a 90% evaporated gasoline sample is not recommended because the within group variation for a given sample at these two levels of evaporation can become large. In a forensic comparison, it would be prudent to compare samples that have similar levels of evaporation.

7.6 References

1. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography I: comparison methodology. Journal of Forensic Sciences, 1987; 32(3): 606-615.

2. Mann, D.C. Comparison of automotive gasolines using capillary gas chromatography II: limitations of automotive gasoline comparisons in casework. Journal of Forensic Sciences, 1987; 32(3): 616-628.

3. Hirz, R. Gasoline brand identification and individualization of gasoline lots. Journal of the Forensic Science Society, 1989; 29(2): 91-101.

4. Potter, T. Fingerprinting petroleum products: unleaded gasoline. In: Petroleum Contaminated Soils: Volume 3. Lewis Publishers, Chelsea, MI; 1990.

5. Alexander, J.; Mashak, G.; Kapitan, N.; Siegel, J.A. Fluorescence of petroleum products II. Three-dimensional fluorescence plots of gasolines. Journal of Forensic Sciences, 1987; 32(1): 72-86.

6. Sheff, L.M. and Siegel, J.A. Fluorescence of petroleum products V: three-dimensional fluorescence spectroscopy and capillary gas chromatography of neat and evaporated gasoline samples. Journal of Forensic Sciences, 1994; 39(5): 1201-1214.

7. Wang, Z. and Fingas, M. Differentiation of the source of spilled oil and monitoring of the oil weathering process using gas chromatography-mass spectrometry. Journal of Chromatography A, 1995; 712: 321-343.

8. Wang Z.D. and Fingas, M. Study of the effects of weathering on the chemical composition of a light crude oil using GC/MS GC/FID. Journal of Microcolumn Separations, 1995; 7(6): 617-639.

9. Smallwood, B.J.; Philp, R.P.; Allen, J.D. Stable carbon isotope composition of gasolines determined by isotope ratio monitoring gas chromatography mass spectrometry. Organic Geochemistry, 2002; 33(2): 149-159.

10. Rietjens, M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytical Chemistry, 1995; 316: 205-215.

11. Jackson, J.E. Principal components and factor analysis: part I. Journal of Quality Technology, 1980; 12(4): 201-213.

12. Lindsay, W.S. Bivariate control chart: uses Hotelling's T in combination with principal components analysis. January 1, 1984. <http://www.minitab.com/support/macros/index.asp?cat=QC_DOE#17> Accessed May 8, 2002.

13. Lavine, B.K. Clustering and classification of analytical data. In: Encyclopedia of Analytical Chemistry, Meyers, R.A., Editor. John Wiley & Sons, Chichester, UK; 2000.

14. Wilson, Nigel. Woolies Wants Change to Fuel Rules. The Australian, The Nation, Page 2; Tuesday, May 8, 2001.

15. Mr. Len Daly, Commercial Development Manager, Royal Vopak (Australia) Limited. Personal communication.

Chapter 8. Analysis and comparison of gasoline samples from New Zealand

8.1 Introduction

8.1.1 Overview of petroleum refining and distribution in New Zealand

The New Zealand gasoline market is different from the market in Australia. New Zealand has only one refinery which is located north of Auckland at Marsden Point. This refinery, which is jointly owned by the four major oil companies (BP, Caltex, Mobil and Shell), is incorporated as the New Zealand Refining Company Ltd. (NZRC). Ships deliver refined products from this refinery to twelve seaboard distribution terminals throughout New Zealand.[1] The city of Auckland is supplied by a 170 km long pipeline that connects the refinery to the Wiri terminal located in south Auckland.[2] In 2001, approximately 93% of the crude oil and other feedstock processed at the Marsden Point refinery was imported,[3] with the balance supplied from local oil basins. The capacity of the refinery is 95,000 barrels per stream day (15.1 megalitres per day). This refinery currently produces approximately 90% of the diesel fuel and 65% of the gasoline consumed in New Zealand.[3] In 2000, approximately 75% of all gasoline sold in New Zealand was regular unleaded (91 octane) and 25% was premium unleaded (96 or 98 octane).[4] No leaded gasoline is produced or sold in New Zealand.[5] Most of the gasoline and diesel fuel imported by New Zealand comes from the Asia-Pacific region with the majority of imports originating from Australia, Singapore and Korea.[4] Gasoline and diesel fuel are also imported on a less frequent basis from the USA and the Middle East.[4]

8.1.2 General comparison of the Auckland, New Zealand and Sydney, Australia gasoline markets

The vast majority of the gasoline sold in Auckland, New Zealand comes directly from the Marsden Point refinery via pipeline.[6] Only a small portion of the gasoline sold in Auckland is imported.[4] Gull Petroleum, an independent fuel distributor and retailer in New Zealand, is known to import gasoline.[4] In contrast, most of the gasoline sold in Sydney, Australia is produced at the Caltex refinery, Kurnell and at the Shell refinery, Clyde. An independent distribution terminal owned by Trafigura/Vopak supplies a small portion of the Sydney gasoline market with imported gasoline. The Caltex refinery at Kurnell has been estimated to supply approximately 85% of the gasoline consumed in Sydney.[7] The Caltex Kurnell refinery imports approximately 70% of its crude oil and refinery feed stock[7] compared to the Marsden Point refinery which imports approximately 93%.[2] During the 1990s, Australia exported more gasoline to New Zealand than to any other nation. Gasoline exports from Australia to New Zealand represented 52% to 65% of all the gasoline Australia exported to other nations (Table 8.1).

Table 8.1 Volume of gasoline exported from Australia to New Zealand. Volumes are in megalites.

	Year				
	1994[1]	*1995*[2]	*1996*[3]	*1997*[4]	*1998*[5]
Australian exports to New Zealand	444.8	513.1	696.5	773.9	881.6
total Australian exports	787.2	870.7	1,161.9	1,466.1	1,364.6
percentage of Australian gasoline exported to New Zealand	56.50%	58.93%	59.94%	52.79%	64.61%

1. Oil and Australia statistical review 1995. Petroleum Gazette. 1995; 2: 14.
2. Oil and Australia statistical review 1996. Petroleum Gazette. 1996; 3: 14.
3. Oil and Australia statistical review 1997. Petroleum Gazette. 1997; 4: 14.
4. Australian petroleum in facts and figures statistical review 1998. Petroleum Gazette. 1998; 3: 40.
5. Australian petroleum in facts and figures statistical review 1999. Petroleum Gazette. 1999; 3: 44.

The data in Table 8.2 shows that the volume of imported gasoline consumed in New Zealand between 1999 and 2000 was typically between 27% and 40%. In contrast with the New Zealand gasoline market, Australia did not import any gasoline from New Zealand in 1994,[8] 1995,[9] or 1998[10] and imported the relatively small quantities of 7.1 and 12.4 megalitres in 1996[11] and 1997,[12] respectively (*cf*. Table 8.1.).

Table 8.2 Volume of gasoline imported from all sources by New Zealand. Volumes are in megalitres.

	Year		
	1999[1]	*2000*[2]	*2001*[3]
gasoline imported by New Zealand†	810	1,197.5	961.2
gasoline consumed in New Zealand†	2,970	3,010.5	2,929.5
percentage of gasoline imported byNew Zealand	27.27%	39.78%	32.81%

†Volumes originally given in tonnes. There are approximately 1350 litres of gasoline per tonne. Institute of Petroleum <http://www.petroleum.co.uk> Petroleum Average Conversion Factors. Accessed September 13, 2002.

1. Explore New Zealand petroleum 2000. Crown Minerals, Ministry of Economic Development. Wellington, New Zealand, 2000.
2. Explore New Zealand petroleum 2001. Crown Minerals, Ministry of Economic Development. Wellington, New Zealand, 2001.
3. Explore New Zealand petroleum 2002. Crown Minerals, Ministry of Economic Development. Wellington, New Zealand, 2002.

Compared to the total amount of gasoline consumed, Australia tended to import a relatively small quantity of gasoline from around the world (2.8% to 4.2% between 1994 and 1998, Table 8.3) compared with New Zealand (27.3% to 39.8% between 1999 and 2001, Table 8.2).

Table 8.3 Volume of gasoline imported from all sources by Australia. Volumes are in megalitres.

			Year		
	1994[1]	*1995*[2]	*1996*[3]	*1997*[4]	*1998*[5]
gasoline imported by Australia	627.1	546.4	647.0	754.4	506.0
gasoline consumed in Australia	17,664.0	17,750.0	17,967.0	17,901.2	18,033.4
Percentage of gasoline imported by Australia	3.55%	3.08%	3.60%	4.21%	2.81%

1. Oil and Australia statistical review 1995. Petroleum Gazette. 1995; 2: 16.
2. Oil and Australia statistical review 1996. Petroleum Gazette. 1996; 3: 12, 16.
3. Oil and Australia statistical review 1997. Petroleum Gazette. 1997; 4: 12, 16.
4. Australian petroleum in facts and figures statistical review 1998. Petroleum Gazette. 1998; 3: 36, 42.
5. Australian petroleum in facts and figures statistical review 1999. Petroleum Gazette. 1999; 3: 40, 46.

Note that the *Petroleum Gazette* ceased publication in 1999.

The aim of this chapter is to apply the method developed for the discrimination of gasoline samples based on their C_0- to C_2-naphthalene profile, to a gasoline market that is different from that of Sydney, Australia. The close proximity of New Zealand to Australia provided the opportunity to collect and analyse gasoline samples as well as to compare them to those obtained in Sydney.

8.2 Materials and methods

8.2.1 Samples

Two collections, totalling 28 samples of gasoline, were made at service stations in metropolitan Auckland, New Zealand in 2002 (Table 8.4). All samples were collected by Mr. Kevan Walsh, Institute of Environmental Sciences and Research (ESR), Auckland according to the sampling protocol in Chapter 4 and Appendix 4.1. The first collection of 14 samples occurred in the summer (February, 2002) and the second collection of 14 samples was made in the winter (August, 2002). There are six service station brands in Auckland. Four of the brands (Caltex, Challenge, Gull and Shell) sold only regular

Table 8.4 Summary of 28 unleaded gasoline samples collected in Auckland, New Zealand during February and August, 2002.

Summer samples					*Winter samples*				
No	Sample	Date†	Brand	Grade‡	No	Sample	Date†	Brand	Grade‡
1	RU343	13.02	Mobil	Regular	15	PU357	12.08	Caltex	Premium
2	PU344	13.02	Challenge	Premium	16	RU358	12.08	Caltex	Regular
3	RU345	13.02	BP	Regular	17	PU359	12.08	Shell	Premium
4	PU346	13.02	Mobil	Premium	18	RU360	12.08	Shell	Regular
5	RU347	13.02	Gull	Regular	19	PU361	12.08	Gull	Premium
6	PU348	13.02	Caltex	Premium	20	RU362	12.08	Gull	Regular
7	PU349	13.02	BP	Premium+	21	PU363	12.08	Mobil	Premium
8	PU350	14.02	Mobil	Premium+	22	RU364	12.08	Mobil	Regular
9	PU351	13.02	Gull	Premium	23	PU365	12.08	Challenge	Premium
10	RU352	13.02	Caltex	Regular	24	RU366	12.08	Challenge	Regular
11	PU353	13.02	Shell	Premium	25	PU367	20.08	BP	Premium
12	PU354	14.02	BP	Premium	26	PU368	12.08	BP	Premium+
13	RU355	13.02	Challenge	Regular	27	RU369	12.08	BP	Regular
14	RU356	13.02	Shell	Regular	28	PU370	12.08	Mobil	Premium+

† Date is expressed as day and month (dd.mm)

‡ The three grades of unleaded gasoline are: Regular (91 octane), Premium (96 octane) and Premium+ (98 octane).

(91 octane) unleaded and premium (96 octane) unleaded gasoline. The other two brands (Mobil and BP) sold regular (91 octane) unleaded, premium (96 octane) unleaded and premium plus (98 octane) unleaded gasoline. The samples were collected from twelve different service stations. In February 2002 (summer), samples of each fuel grade of gasoline available were collected from six service stations, one station for each of the six brands. In August 2002 (winter), samples of each fuel grade of gasoline available were collected from six service stations, one station for each of the six brands. The locations of the six service stations from which the winter samples were collected were different from those used to collect the summer samples. Samples were collected in 125 mL amber, Boston round glass bottles sealed with polypropylene caps fitted with Teflon-faced foamed polyethylene liners (Wheaton).

8.2.2 Gas chromatography-mass spectrometry

The gas chromatography-mass spectrometry conditions described in chapter 6 were used to analyse the samples for their C_0- to C_2-naphthalene profiles. Total ion chromatograms (TICs) of unevaporated and evaporated gasoline were obtained for the six Mobil samples (RU343, PU346, PU350, PU363, RU364, PU370). Chromatographic conditions were as follows: split injection (15:1) at 280 ^{o}C; temperature programmed from 40 ^{o}C (2 minutes isothermal) to 200 ^{o}C at 10 ^{o}C/minute, then ramped to 290 ^{o}C at 20 ^{o}C/minute (held for 5 minutes); helium carrier was gas at a constant flow rate of 1.2 mL/minute. The sample injection volume was 1 μL, and the detector was operated in full scan mode (30 to 350 daltons) with a sampling rate of 4.45 scans/second.

8.2.3 Retention index standards

Retention indices were calculated from the retention times of known standards. A standard solution of n-alkanes (n-C_5 to n-C_{22}), naphthalene, phenanthrene and chrysene, was made up in dichloromethane. A 500 ppm solution (w/v) of each standard compound was made-up by diluting approximately 5 mg of each standard compound into 100 mL of dichloromethane (HPLC grade, BDH). Standards were obtained from the following suppliers: n-C_5, n-C_{20} to n-C_{22} and chrysene from Aldrich; n-C_6 to n-C_{19} from Polyscience; naphthalene from BDH; and, phenanthrene from Merck.

Standards were used without further purification. All n-alkane standards used had a purity of 99% or greater, except n-C_{21} which had a purity of 98%. The polycyclic aromatic hydrocarbon standards had the following purities: naphthalene, 98%; phenanthrene, >96%; and, chrysene 95%. HPLC grade dichloromethane (diluent) had a purity of >99.8%.

8.2.4 Data analysis

Peak area information was extracted from the chromatograms using a macro. The macro is reproduced in Appendix 5.1. The integration results were written to an ASCII text file in a format suitable for export to a spreadsheet. Data was loaded and parsed into a spreadsheet (Microsoft Excel®). The data was normalised using a natural logarithmic transformation.[13] Principal component analysis and linear discriminate analysis with cross-validation was performed using Minitab® (version 13.1 for Windows).

8.3 New Zealand gasoline samples

8.3.1 Results

All 28 gasoline samples collected in New Zealand were analysed by principal component analysis (PCA) and the results are given in Table 8.5. The first two PCs account for 69.0% of the variation in the data. A plot of the first two PC scores is shown in Figure 8.1.

Linear discriminant analysis (LDA) with cross-validation using the first four PCs correctly classified 126 of the 140 chromatograms (90%) by sample. However, close inspection of the LDA results revealed a difference between the classification of the winter samples compared with the summer gasoline samples. Linear discriminant analysis correctly

Table 8.5 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM): 28 New Zealand gasoline samples.

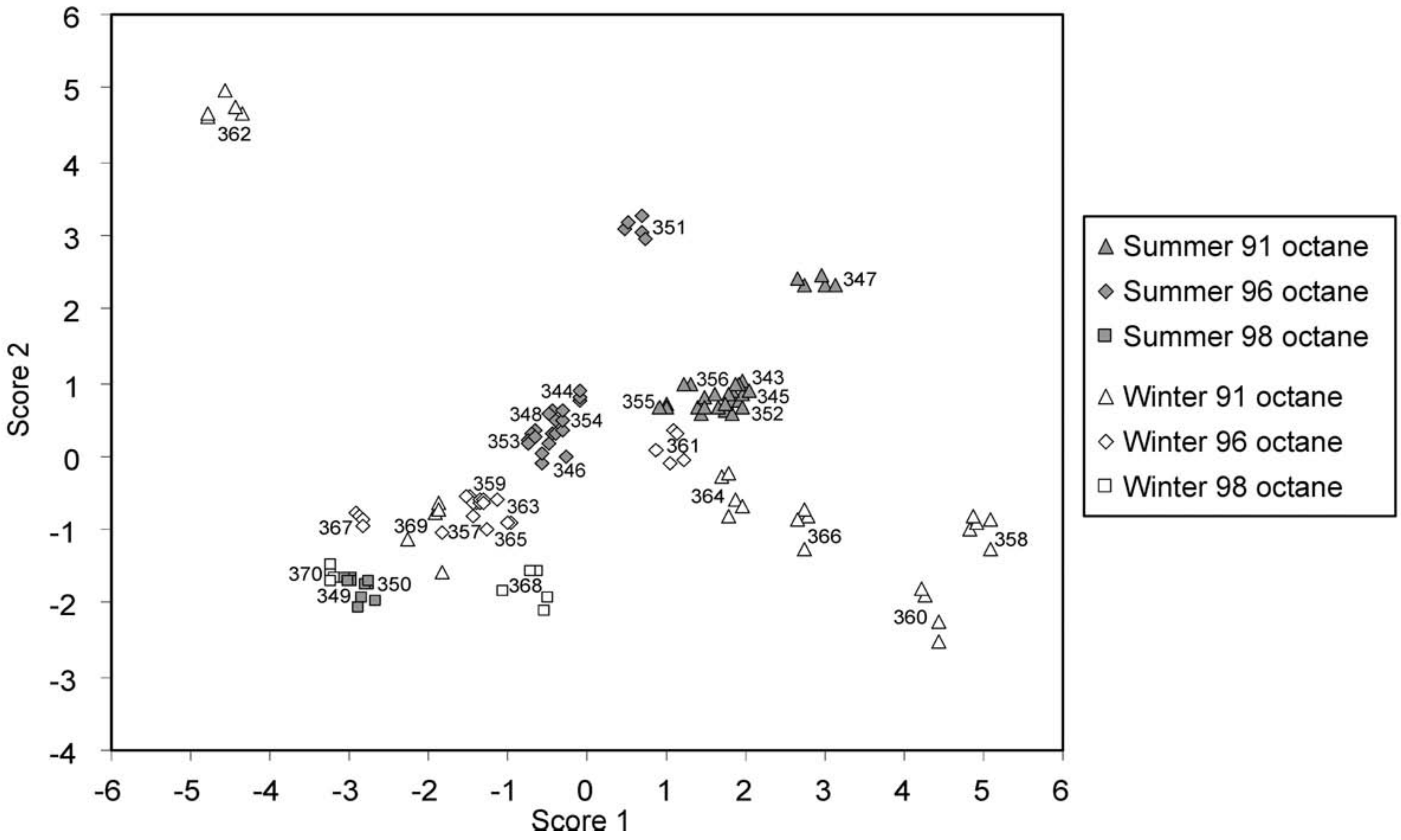
	Eigenvalue	Proportion	Cumulative (%)
PC1	5.2556	0.478	47.8
PC2	2.3392	0.213	69.0
PC3	2.0663	0.188	87.8
PC4	0.9472	0.086	96.4
PC5	0.2123	0.019	98.4

classified all 70 winter gasoline chromatograms (100%) by sample, and correctly classified 54 of the 70 (77%) summer gasoline chromatograms by sample. In practical terms, any overlap of one chromatogram (one aliquot) from one sample with any other chromatogram from another sample, would cause these two gasoline samples to be indistinguishable. The LDA results were, therefore, re-examined and it was found that the 14 summer gasoline samples formed seven groups. Five of these groups each contained a single sample (*i.e.* five aliquots from one sample per group). The remaining nine samples were divided between the two remaining groups (Table 8.6).

Table 8.6 Results of LDA with cross-validation: number of aliquots from the nine misclassified summer samples.

		True Group								
		PU344	PU348	PU353	PU354	RU343	RU345	RU352	RU355	RU356
Predicted Group	PU344	3	0	0	2	0	0	0	0	0
	PU348	0	3	1	1	0	0	0	0	0
	PU353	0	2	4	0	0	0	0	0	0
	PU354	2	0	0	2	0	0	0	0	0
	RU343	0	0	0	0	0	0	0	0	0
	RU345	0	0	0	0	2	4	0	0	0
	RU352	0	0	0	0	1	1	5	0	0
	RU355	0	0	0	0	1	0	0	5	0
	RU356	0	0	0	0	1	0	0	0	5

Figure 8.1 PC score plot of 28 samples collected in Auckland, New Zealand. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample identification.



8.3.2 Discussion

There is more variation in the C_0- to C_2- naphthalene profile between the samples of winter gasoline than between the samples of summer gasoline. Analysis of the PC score data by LDA showed that each of the 14 winter gasoline samples formed their own group (*i.e.* 100% classification by sample). Analysis of the 14 summer gasoline samples showed that five samples formed distinct groups and the remaining nine samples were divided between two other groups (Table 8.6). The summer samples also tended to cluster together according to their fuel grade (octane rating). Two exceptions to this trend are two samples from Gull, samples RU347 and PU351. Grouping of samples by fuel grade is not observed for the winter samples.

Three of the four Gull samples (RU347, PU351 and RU362) are removed from the other samples along the axis of the second PC score (Figure 8.1). The segregation of these samples indicates a significant chemical difference in the C_0- to C_2-naphthalenes between these three samples as well as between these samples and all the others. When Gull Petroleum first entered the New Zealand market in 1998, all gasoline sold at Gull service stations was imported.[14] By 2001, however, Gull was able to purchase gasoline from the New Zealand Refinery Corporation's (NZRC) Marsden Point refinery.[14] In the New Zealand market, Gull is still considered to be primarily an importer of gasoline and diesel fuel.[15]

When Challenge and Gull first entered the New Zealand motor fuel market in the 1990s, they were both considered to be independent operators and, therefore, imported all of their fuel stocks.[4] With the sale of Challenge to Caltex, Gull Petroleum is now the only truly independent gasoline retailer in New Zealand. Gull Petroleum currently supplies gasoline to the retailer Woolworths.[16] Taking into consideration that Caltex owns (and therefore controls) Challenge service stations, it is not surprising that fewer differences were observed between the gasoline samples from Challenge and the four major oil companies (BP, Caltex, Mobil, and Shell) than were observed between the gasoline samples from Gull and the other five brands.

In Auckland, four brands of service station (Caltex, Challenge, Gull and Shell) sell two grades of gasoline (91 and 96 octane) and two brands of service station (BP and Mobil) sell three grades of gasoline (91, 96 and 98 octane). Differences in the C_0- to C_2-naphthalene profiles between these three grades can be seen in the separation of these samples in a PC score plot. The differences between fuel grades for each brand are highlighted in Figures 8.2 to 8.7. The 12 different gasoline samples representing the three different fuel grades from BP and Mobil service stations are summarised in Table 8.7. Total ion chromatograms of the Mobil samples collected in the summer and the winter are shown in Figures 8.8 and 8.9, respectively. The selected ion monitored chromatograms showing the C_0- to C_2-naphthalenes for the same Mobil samples collected in the summer and the winter are shown in Figures 8.10 and 8.11, respectively. As the octane of the gasoline increases, the size of the C_2-naphthalene peaks increase relative to the C_0- and C_1-naphthalene peaks. This change in the C_2-naphthalene profile explains the separation between the different fuel grades observed in the PC score plots (Figures 8.2 to 8.7).

Table 8.7 Summary of the 12 samples of the three grades of gasoline obtained from BP and Mobil service stations in Auckland, New Zealand.

Summer Samples				*Winter Samples*			
Sample	Date†	Brand	Grade‡	Sample	Date†	Brand	Grade‡
RU345	13.02	BP	Regular	RU369	12.08	BP	Regular
PU354	14.02	BP	Premium	PU367	20.08	BP	Premium
PU349	13.02	BP	Premium+	PU368	12.08	BP	Premium+
RU343	13.02	Mobil	Regular	PU363	12.08	Mobil	Premium
PU346	13.02	Mobil	Premium	RU364	12.08	Mobil	Regular
PU350	14.02	Mobil	Premium+	PU370	12.08	Mobil	Premium+

† Date is expressed as day and month (dd.mm)

‡ The three grades of unleaded gasoline are: Regular (91 octane), Premium (96 octane) and Premium+ (98 octane).

Figure 8.2 PC score plot of 28 samples collected in Auckland, New Zealand. BP samples are highlighted. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample information.

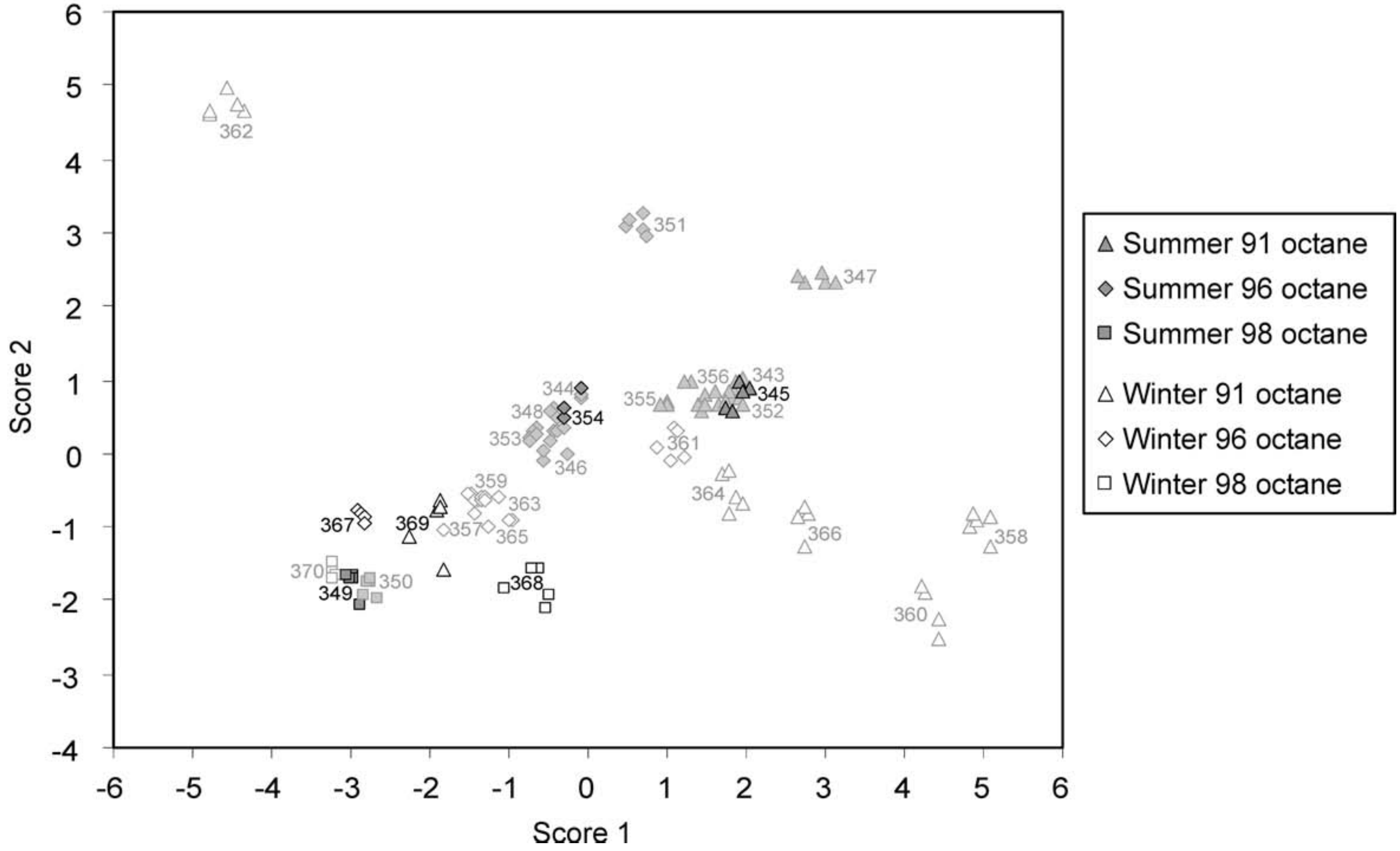


Figure 8.3 PC score plot of 28 samples collected in Auckland, New Zealand. Caltex samples are highlighted. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample information.

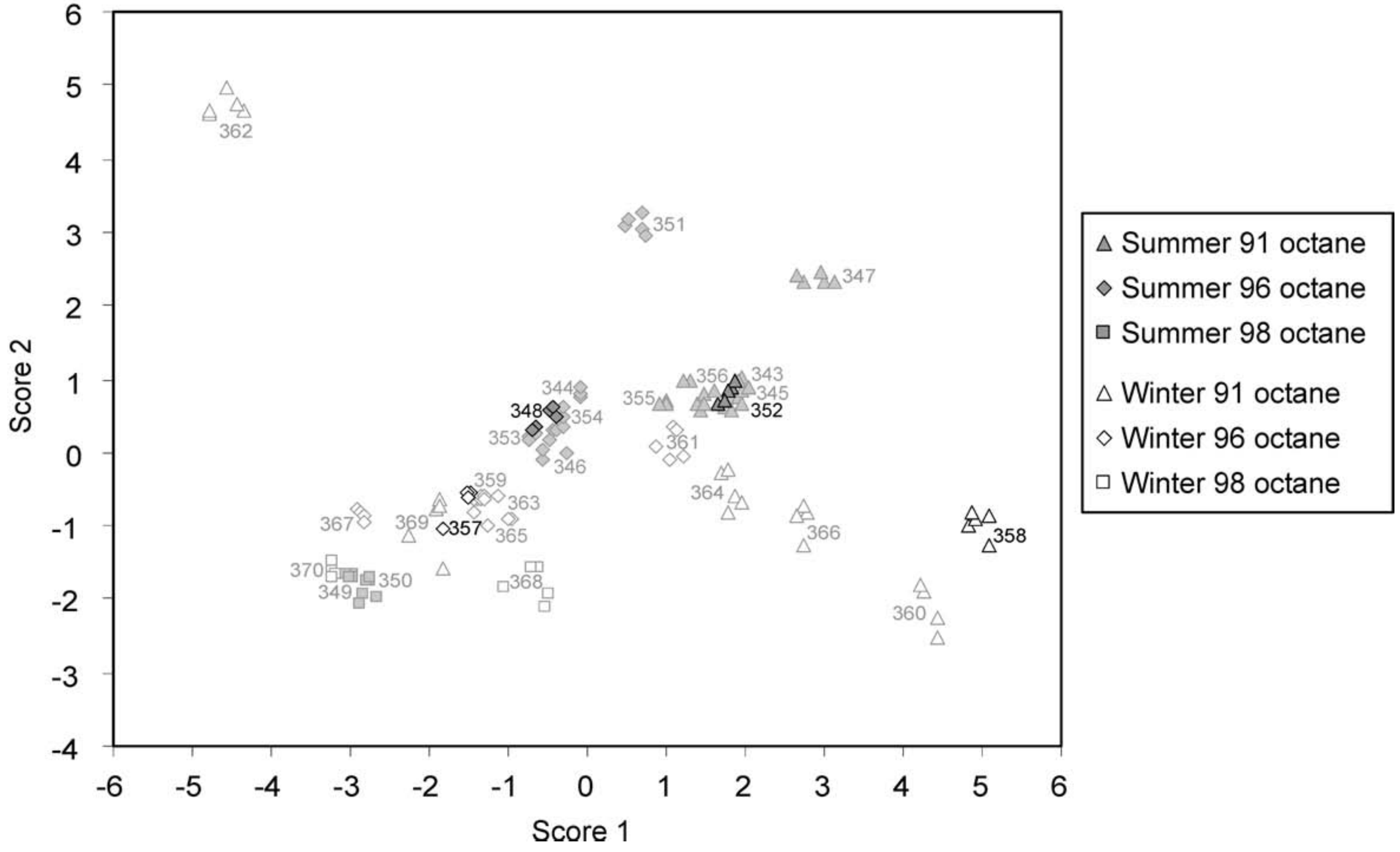


Figure 8.4 PC score plot of 28 samples collected in Auckland, New Zealand. Challenge samples are highlighted. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample information.

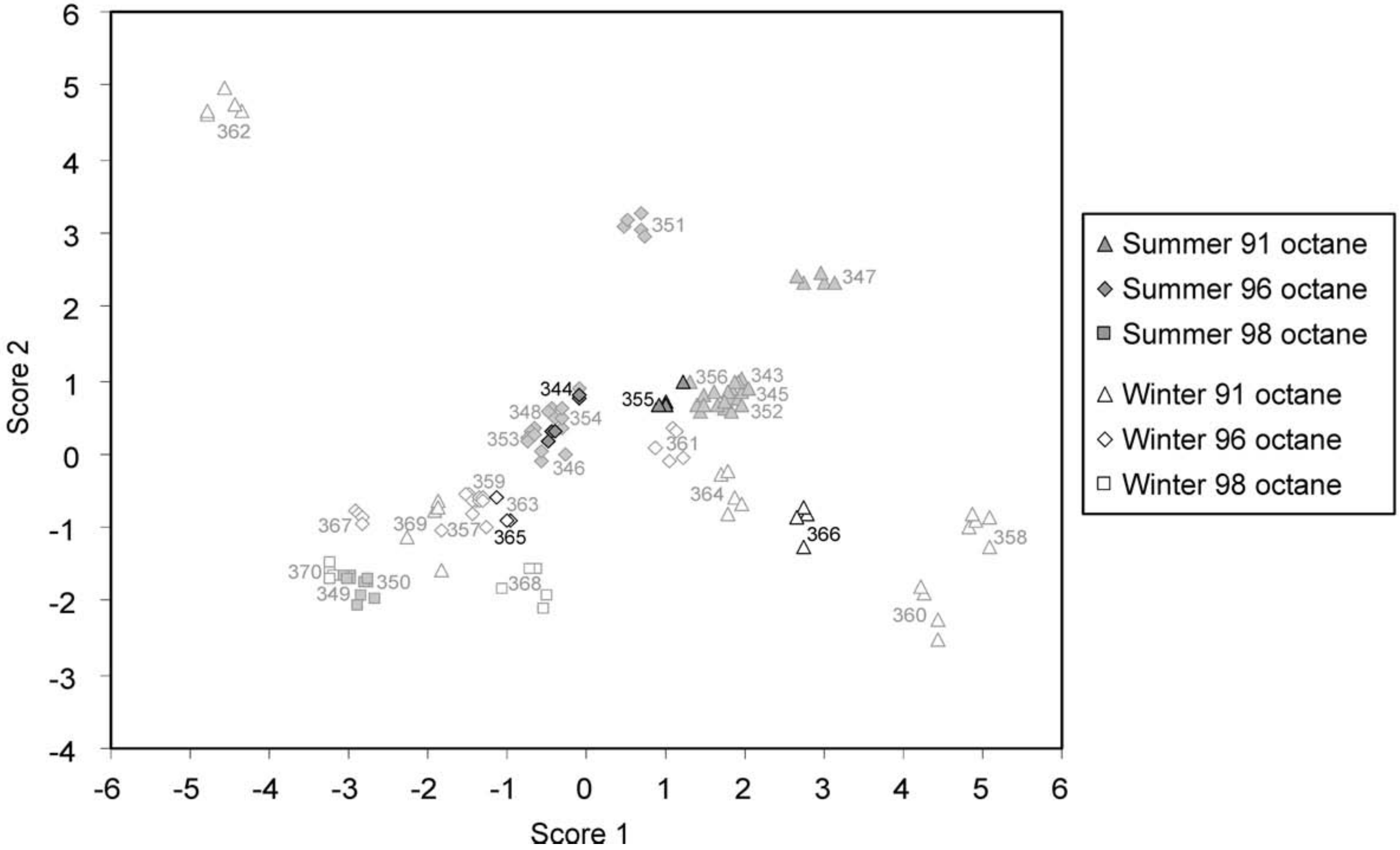


Figure 8.5 PC score plot of 28 samples collected in Auckland, New Zealand. Gull samples are highlighted. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample information.

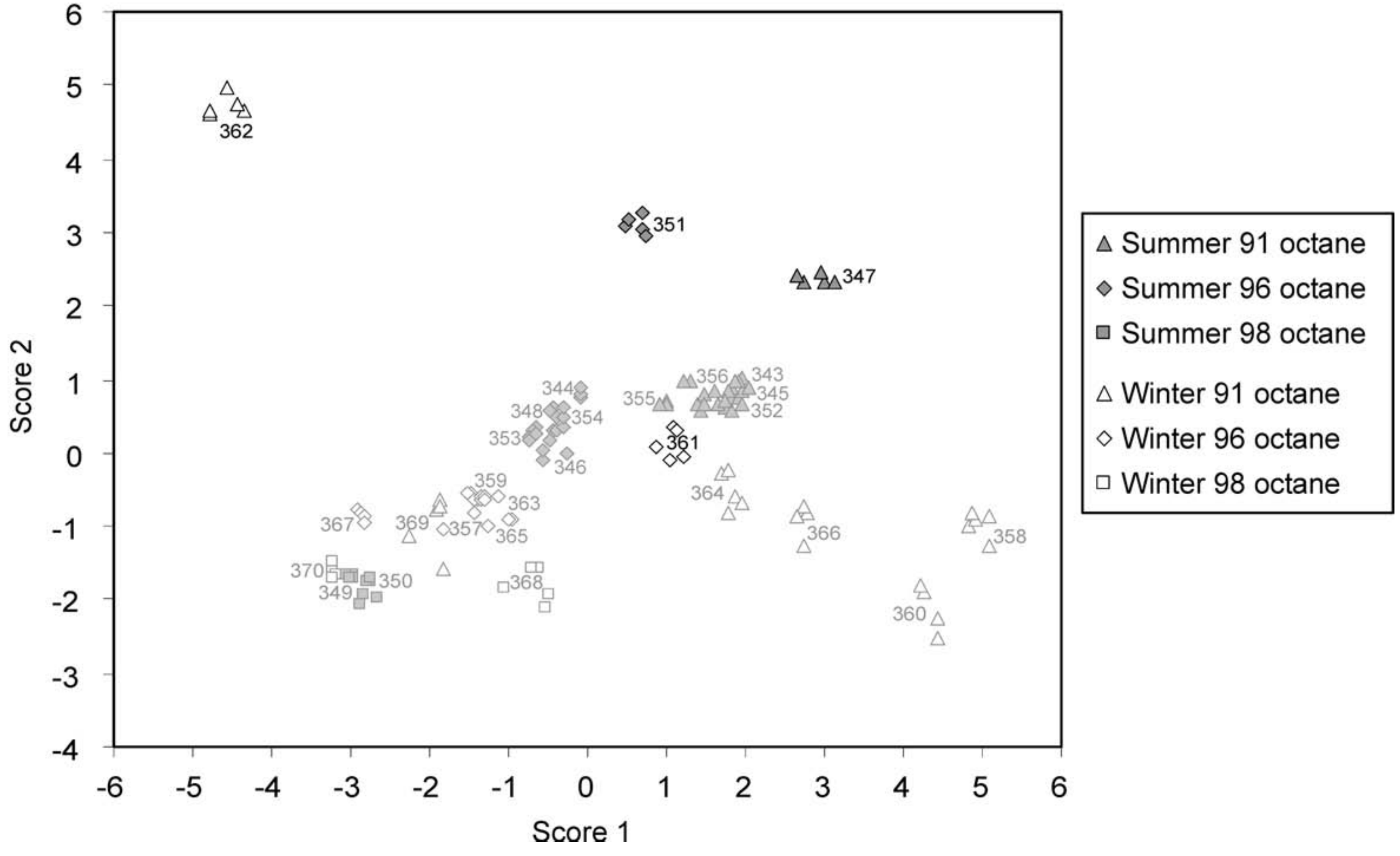


Figure 8.6 PC score plot of 28 samples collected in Auckland, New Zealand. Mobil samples are highlighted. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample information.

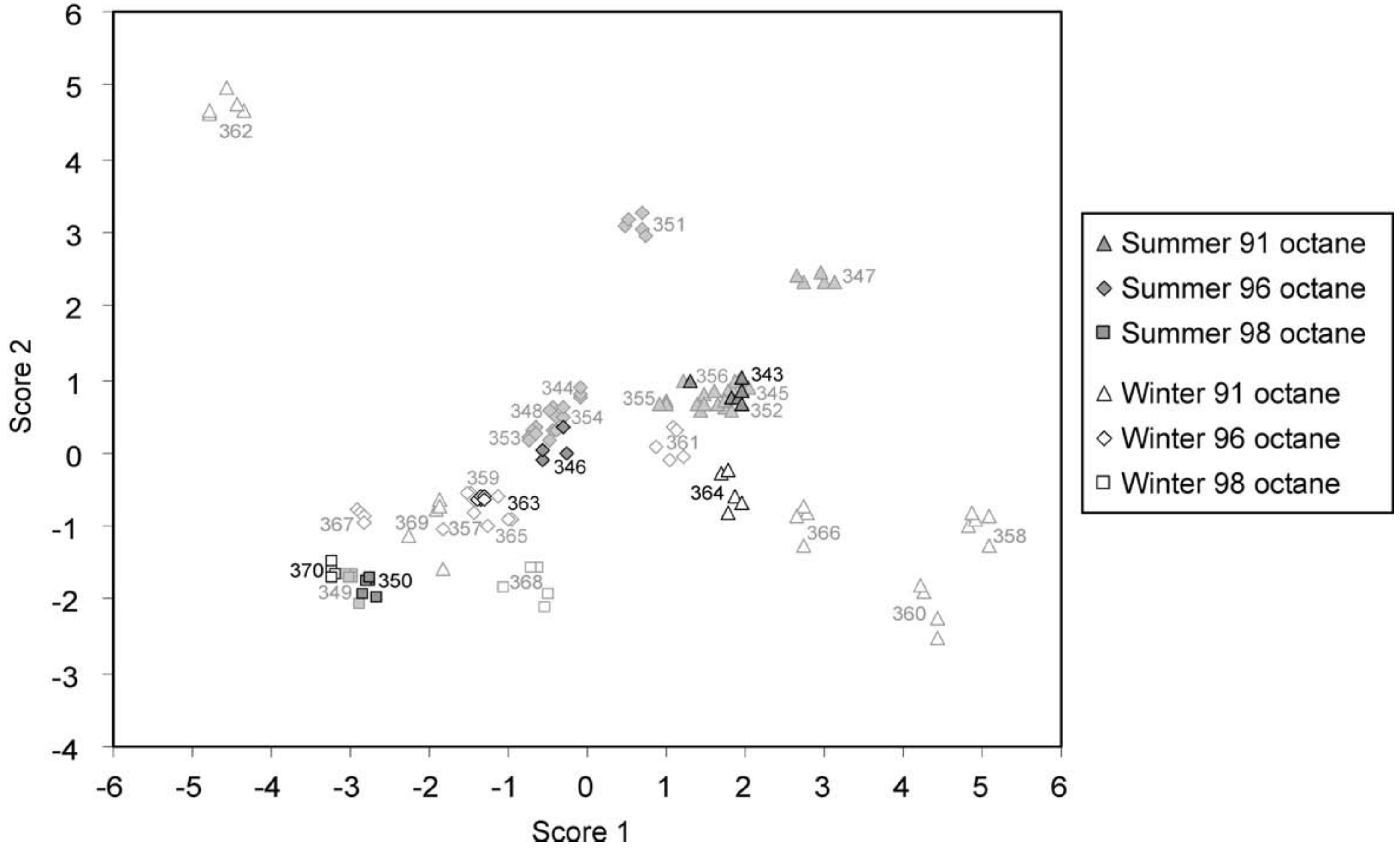


Figure 8.7 PC score plot of 28 samples collected in Auckland, New Zealand. Shell samples are highlighted. Summer samples were collected in February, 2002 and winter samples were collected in August, 2002. Refer to Table 8.4 for sample information.

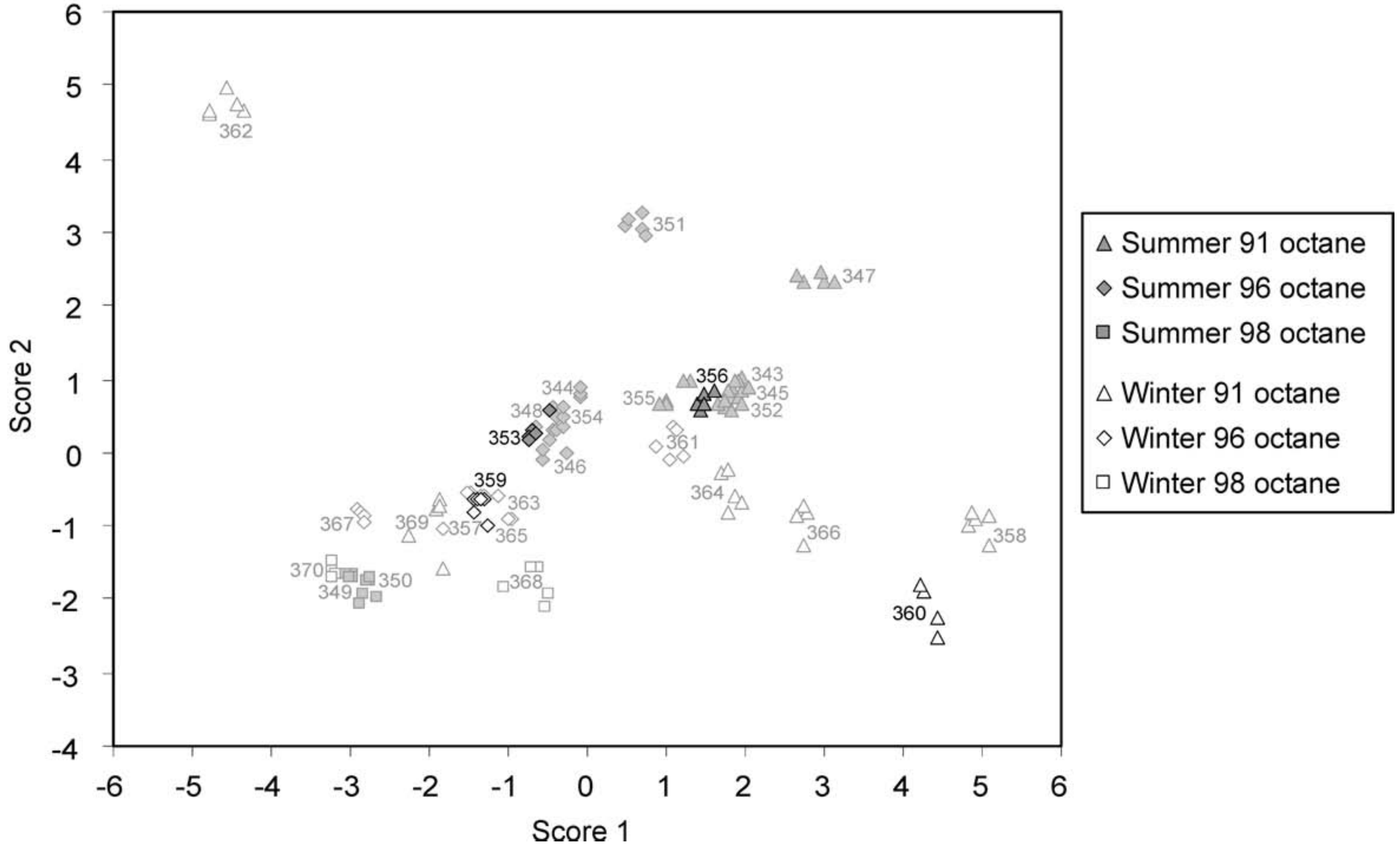


Figure 8.8 Total ion chromatograms of Mobil summer gasoline samples collected in Auckland, New Zealand in February, 2002.

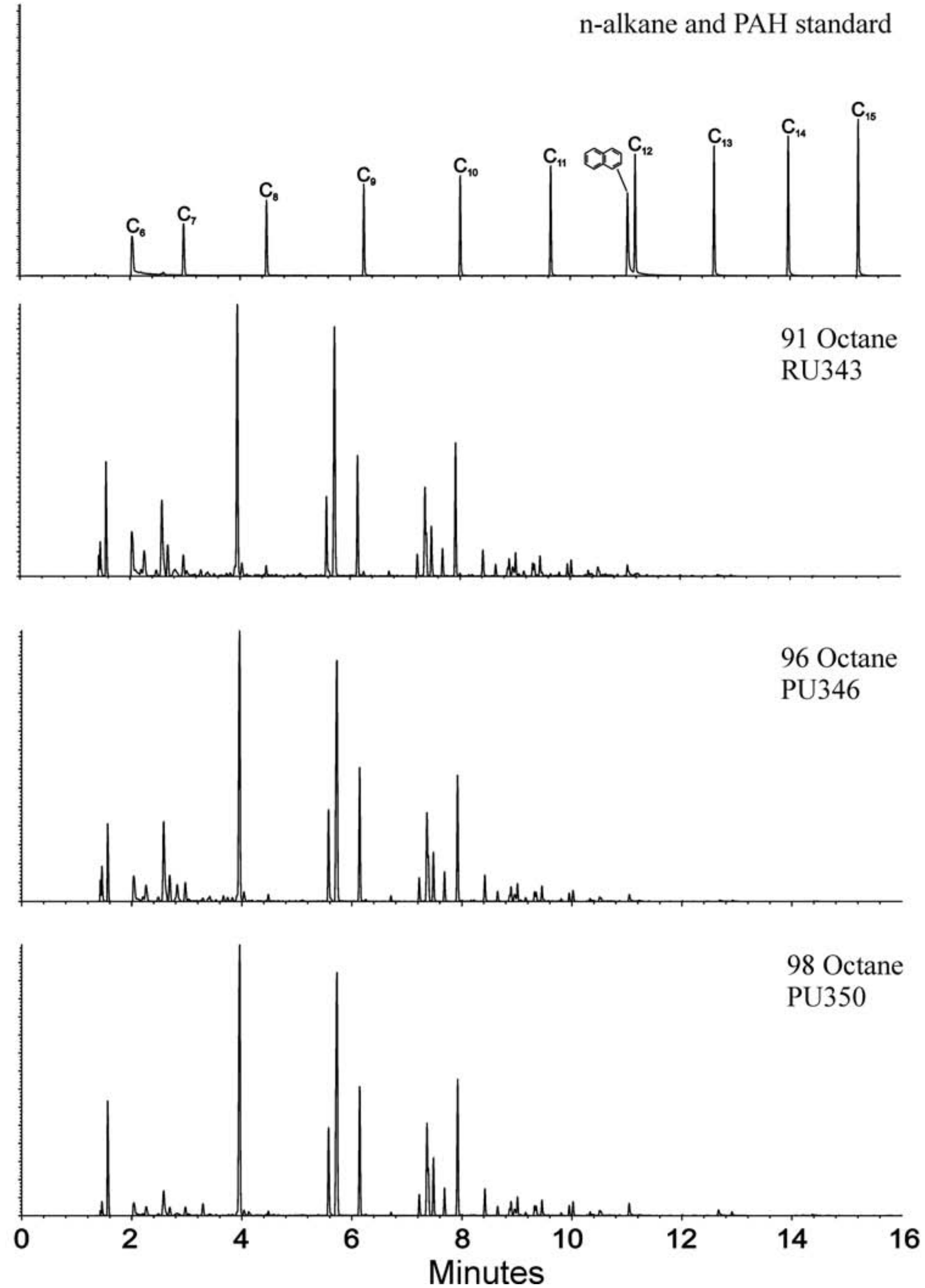


Figure 8.9 Total ion chromatograms of Mobil winter gasoline samples collected in Auckland, New Zealand in August, 2002.

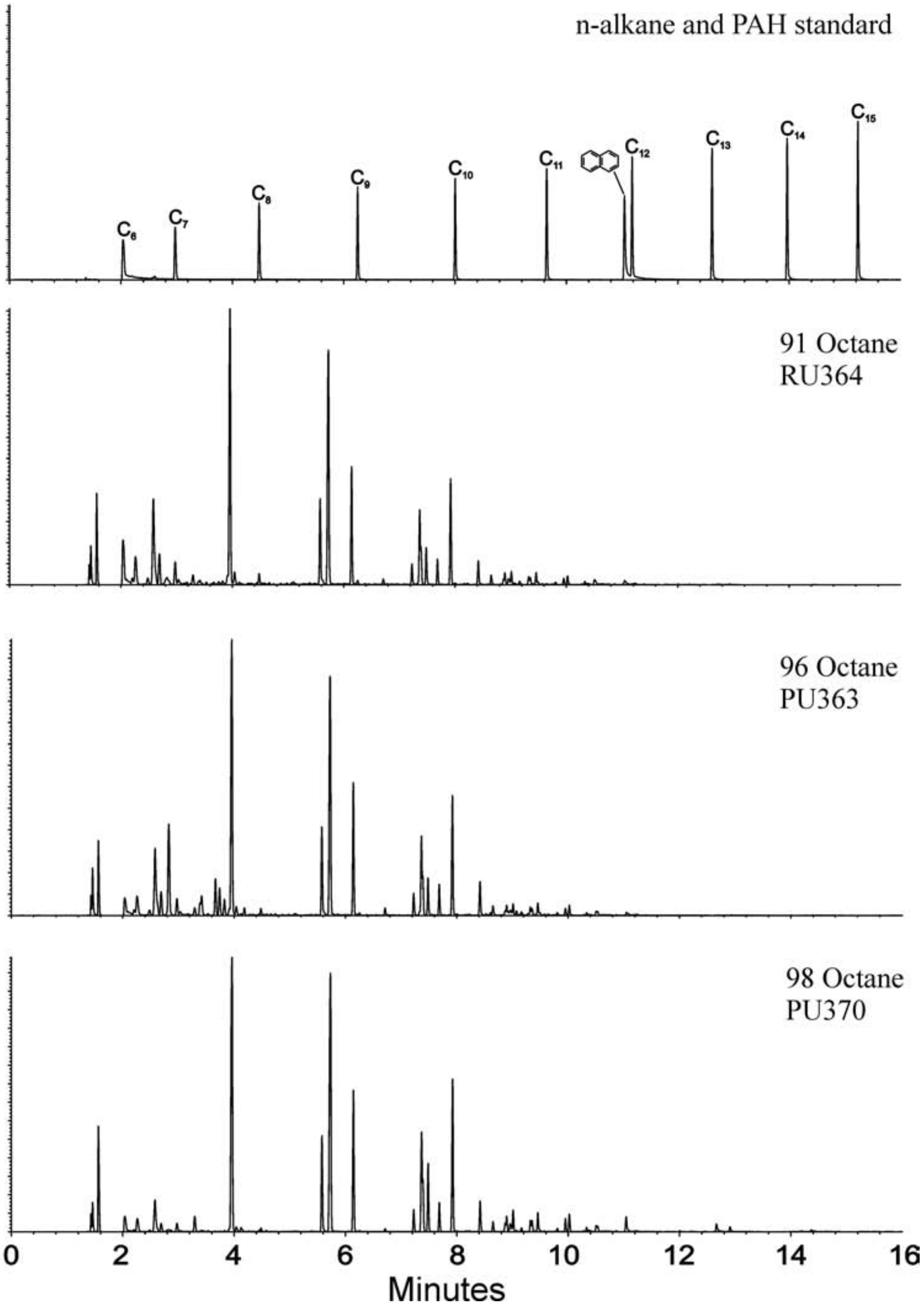


Figure 8.10 Mobil summer gasoline samples collected in Auckland, New Zealand in February, 2002. Chromatograms obtained from selected ion monitoring (SIM) of C_0- to C_2-naphthalenes.

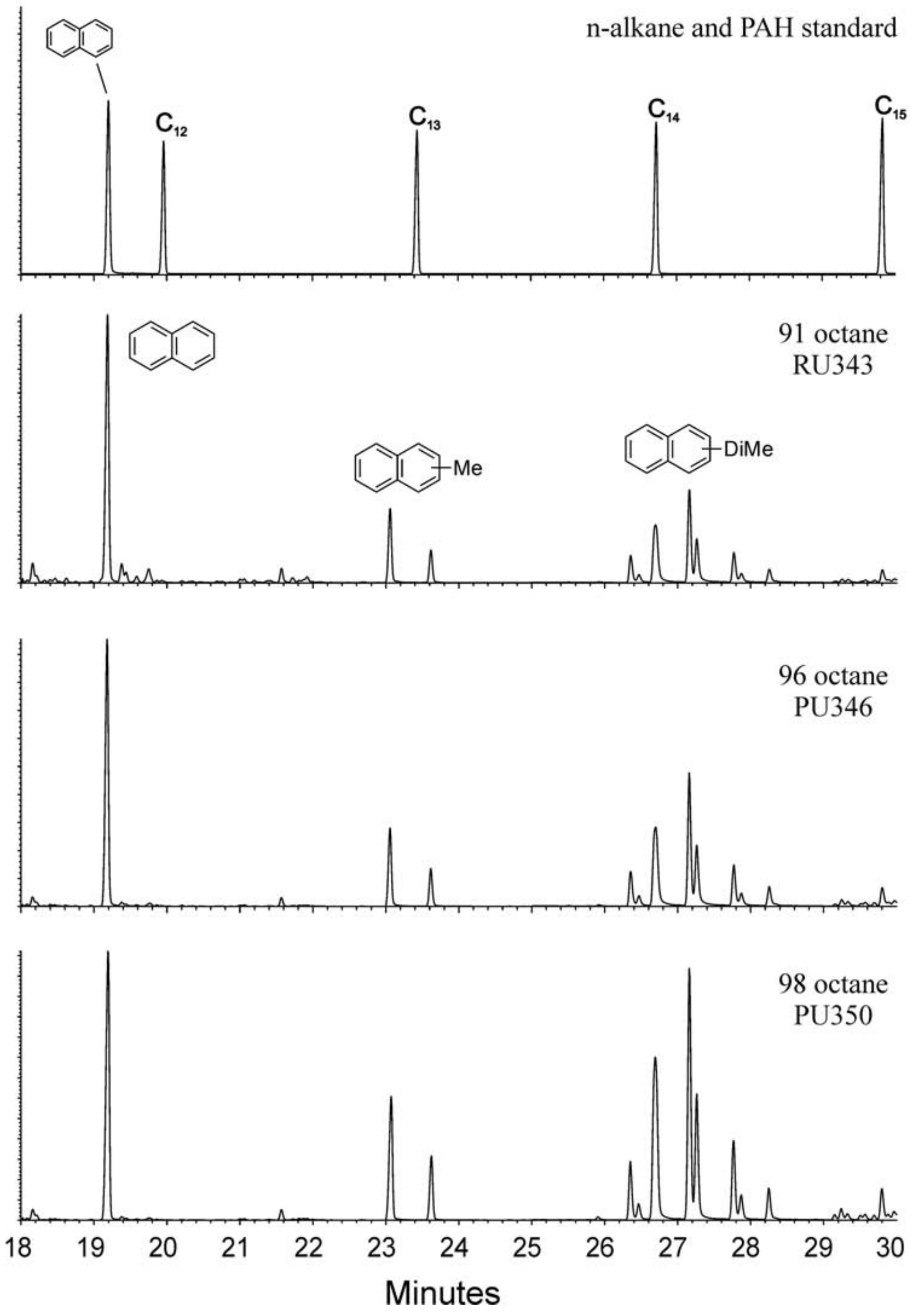
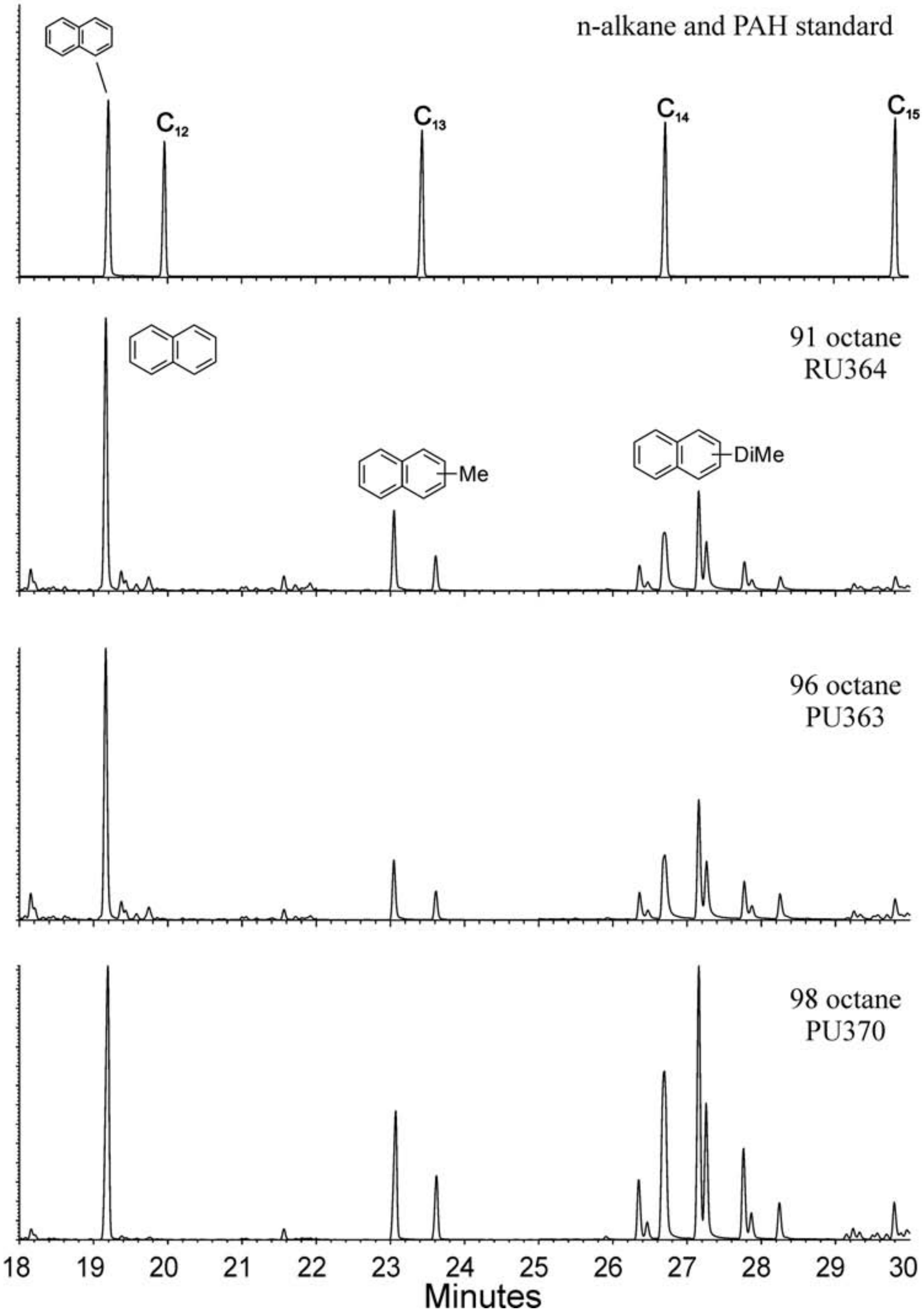


Figure 8.11 Mobil winter gasoline samples collected in Auckland, New Zealand in August, 2002. Chromatograms obtained from selected ion monitoring (SIM) of C_0- to C_2-naphthalenes.



8.4 Comparison of Australian and New Zealand gasoline samples

8.4.1 Results

The collection of gasoline samples in the same month (February, 2002) from both Auckland, New Zealand and Sydney, Australia provided the opportunity to compare gasoline from these two different markets. Twenty-four gasoline samples were collected from three different service stations in Sydney in February, 2002. The details of this sample collection are given in Chapter 6 and the February samples are summarised in Table 8.8. Fourteen samples were collected in February, 2002 in Auckland (Table 8.4). Five aliquots of each sample were analysed. Chromatographic data from these 38 samples (190 chromatograms) were combined and the resulting data matrix analysed by principal components.

Table 8.8 Summary of 24 unleaded gasoline samples collected from three service stations in Sydney, Australia during February, 2002.

No	Sample	Date†	Brand	Grade‡	No	Sample	Date†	Brand	Grade‡
1	RU69	05.02	Caltex	Regular	13	RU81	19.02	Caltex	Regular
2	PU70	05.02	Caltex	Premium	14	PU82	19.02	Caltex	Premium
3	RU71	05.02	BP	Regular	15	RU83	19.02	BP	Regular
4	PU72	05.02	BP	Premium	16	PU84	19.02	BP	Premium
5	RU73	05.02	Shell	Regular	17	RU85	19.02	Shell	Regular
6	PU74	05.02	Shell	Premium	18	PU86	19.02	Shell	Premium
7	RU75	12.02	Caltex	Regular	19	RU87	26.02	Caltex	Regular
8	PU76	12.02	Caltex	Premium	20	PU88	26.02	Caltex	Premium
9	RU77	12.02	BP	Regular	21	RU89	26.02	BP	Regular
10	PU78	12.02	BP	Premium	22	PU90	26.02	BP	Premium
11	RU79	12.02	Shell	Regular	23	RU91	26.02	Shell	Regular
12	PU80	12.02	Shell	Premium	24	PU92	26.02	Shell	Premium

† Date is expressed as day and month (dd.mm)
‡ The two grades of unleaded gasoline are: Regular (91 octane) and Premium (96 octane).

The variation in the data described by the first five PCs is given in Table 8.9. The first two PCs accounted for 91.1% of the variation in the data. A plot of the first two PC scores is shown in Figure 8.12. The sign of the first PC score is positive for the majority (17 out of 24, 70.8%) of the Australian samples and negative for all of the New Zealand samples.

Table 8.9 Results of PCA for C_0- to C_2-naphthalenes by GC-MS (SIM) for Australian and New Zealand gasoline samples collected in February, 2002.

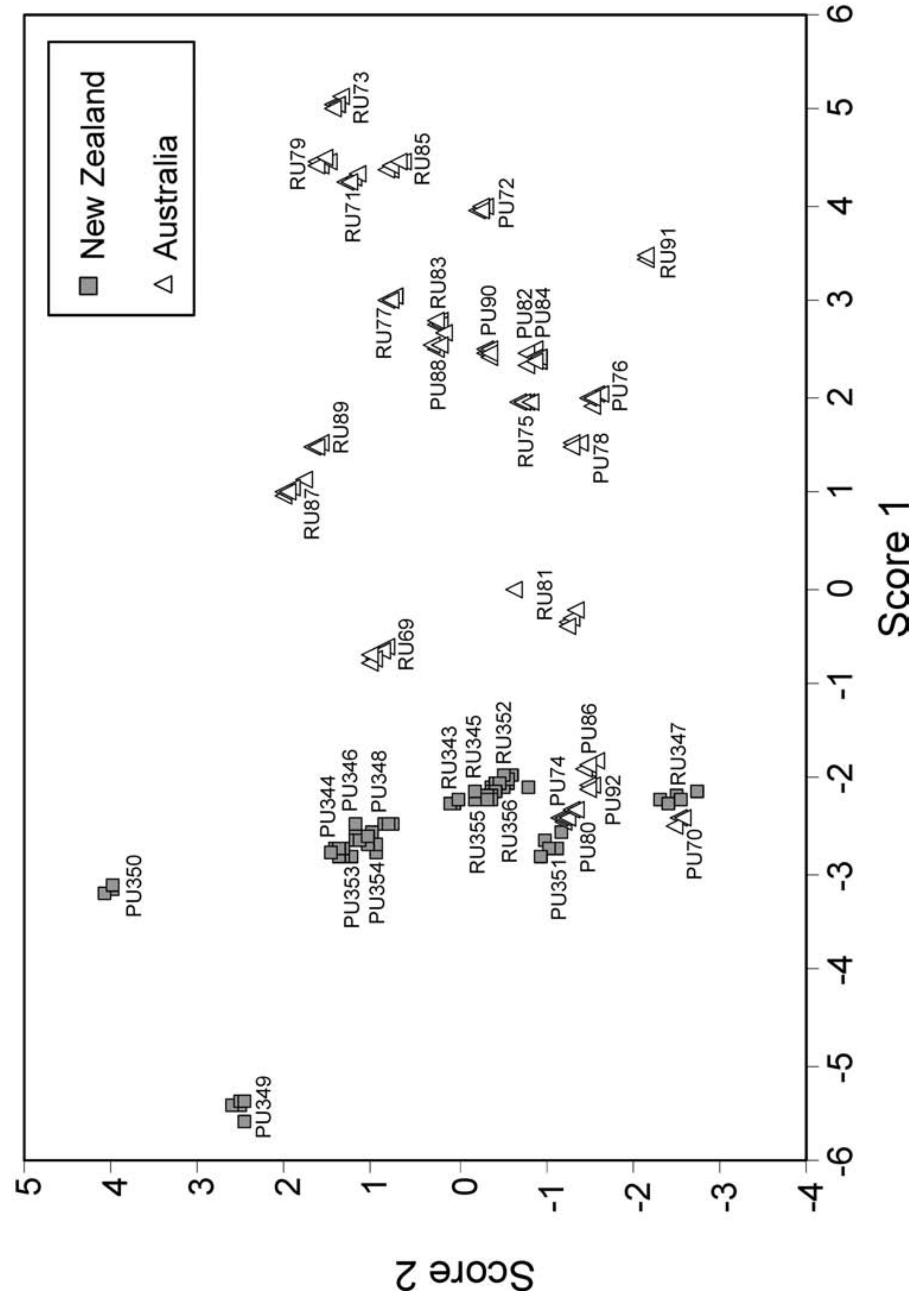
	Eigenvalue	Proportion	Cumulative (%)
PC1	7.9364	0.721	72.1
PC2	2.0857	0.190	91.1
PC3	0.5497	0.050	96.1
PC4	0.2251	0.020	98.2
PC5	0.1131	0.010	99.2

8.4.2 Discussion

The majority of the samples (19 out of 24, 79.2%) collected in Sydney, Australia are significantly different from the gasoline samples collected in Auckland, New Zealand during the same month. The New Zealand samples with the greatest chemical similarity to the Australian samples are both from Gull (RU347 and PU351). The Australian sample PU70 from Caltex is the most similar in its C_0- to C_2-naphthalene profile to the Gull sample RU347. Four other Australian samples, PU74, PU80, PU86 and PU92 from Shell are similar to Gull sample PU351. When Gull Petroleum first entered the New Zealand market in 1998, all the gasoline sold at Gull service stations was imported and Gull continues to import most of the gasoline sold at its service stations.[14] While these results suggest that these two Gull samples have been imported from Australia, it is not possible to know for certain.

Overall the C_0- to C_2-naphthalene profile for samples of Auckland gasoline is significantly different from the profile for samples of Sydney gasoline. Despite the fact that in 2001 approximately 33% of New Zealand's gasoline was imported (Table 8.2), historically much of it from Australia, there is a difference between most of the samples based on their country of origin. In drawing this conclusion, it is assumed that the gasoline in the majority of samples collected in Auckland originated via pipeline from the NZRC,[6] and the gasoline in the majority of samples collected in Sydney came from either the Caltex Kurnell or the Shell Clyde refinery. Like the New Zealand market, the gasoline market in Sydney is complex. For example, the Shell refinery at Clyde can receive product from the Caltex

Figure 8.12 PC score plot of 14 gasoline samples collected in Auckland, New Zealand and 24 gasoline samples collected in Sydney, Australia during February, 2002. Sample information is in Tables 8.4 and 8.8.



Kurnell refinery (via a pipeline that runs through the Caltex Banksmeadow distribution terminal[16]) thus creating the possibility of Shell occasionally distributing the same, or similar, product as Caltex.

8.5 Conclusions

The method developed for the discrimination of gasoline samples based on their C_0- to C_2-naphthalene profile was applied to samples collected in Auckland, New Zealand. The Auckland gasoline market is different from the one in Sydney in several aspects, including the number of refineries supplying the market and the relative amount of gasoline that is imported. The majority of the gasoline sold in Auckland, however, comes directly from the Mardsen Point refinery, the only refinery in the country. The fact that the distribution terminals supplying gasoline to the Auckland service stations receive fuel directly from the refinery via pipeline may have contributed to the lack of differences observed in the C_0- to C_2-naphthalene profile for the summer gasoline samples. A larger variation between samples was observed for the gasoline samples collected in winter. Three of the four gasoline samples that were collected at Gull service stations were readily distinguished from the other New Zealand samples suggesting that they may have been imported from overseas refineries. A comparison between samples collected in Auckland, New Zealand and Sydney, Australia in February, 2002 showed that most samples could be grouped based on the country of origin, suggesting that the C_0- to C_2-naphthalene profile is dependant on the refinery at which the gasoline was made.

8.6 References

1. ACIL Economic and Policy Pty. Ltd. Barriers to entry to the New Zealand downstream oil market: a report to New Zealand Ministry of Commerce. Canberra, ACT, Australia; 1997.

2. The New Zealand Refining Company Ltd. Education. <http://www.nzrc.co.nz/> Accessed September 13, 2002.

3. Explore New Zealand petroleum 2002. Crown Minerals, Ministry of Economic Development. Wellington, New Zealand; 2002.

4. Petrol and diesel: delivering quality. A review of the petroleum products specifications regulation. Resources and Networks Branch, Ministry of Economic Development. Wellington, New Zealand. August 2001.

5. Explore New Zealand petroleum 2001. Crown Minerals, Ministry of Economic Development. Wellington, New Zealand; 2001.

6. The New Zealand Refining Company Ltd. Submission to: Energy and Efficiency and Conservation Authority. Draft National Energy Efficiency and Conservation Strategy, June 12, 2001. <www.http://eeca.govt.nz/subpdf/NZRC.pdf> Accessed November 15, 2001.

7. Mr. Paul Seage, Chief Chemist, Caltex Kurnell Refinery. Personal Communication.

8. Oil and Australia statistical review 1995. Petroleum Gazette, 1995; 2: 16.

9. Oil and Australia statistical review 1996. Petroleum Gazette, 1996; 3: 16.

10. Australian petroleum in facts and figures statistical review 1999. Petroleum Gazette, 1999; 3: 46.

11. Oil and Australia statistical review 1997. Petroleum Gazette, 1997; 4: 16.

12. Australian petroleum in facts and figures statistical review 1998. Petroleum Gazette, 1998; 3: 42.

13. Rietjens, M. Reduction of error propagation due to normalization: effect of error propagation and closure on spurious correlations. Analytical Chemistry, 1995; 316: 205-215.

14. Pickford, M and Wheeler, C. The petrol industry: deregulation, entry and competition. NZ Trade Consortium Working Paper No. 12. 2001.

15. Belgrave, M.J.; Berry, M.N.; Rebstock, P.R. Decision No. 434: Caltex New Zealand Limted and Challenge Petroleum Limited, Commerce Commission, Auckland, New Zealand June 28, 2001.

16. Woolworths and Gull team up. Retail Today. May 2001.

17. Mr. John Morgan, Director of Operations, Banksmeadow Distribution Terminal, Caltex Australia Limited. Personal Communication.

Chapter 9. Final conclusions

Arson is a destructive crime and Australian society incurs a high cost each year in lost property, in interruptions to business, and in fire fighting and fire prevention costs. Arson is also a difficult crime to prosecute, often requiring strong physical evidence to successfully link a suspect to a crime scene. The most common liquid accelerant used in arson is gasoline. The ability to uniquely characterise gasoline would allow a comparison to be made between a gasoline sample obtained from a suspect and a gasoline sample recovered from the fire scene.

In order to develop a method for differentiating one gasoline sample from another, one must understand the crude oil refining process as well as the gasoline distribution network from refinery to consumer. Differences between crude oils, blending at the refinery, mixing of the finished gasoline in storage tanks at a bulk distribution terminal and in an underground storage tank at a service station, as well as the gasoline turn over rate at a service station, may all lead to differences between gasoline samples obtained from different service stations or from the same service station at different times. A comparison method must also be able to take into account the inevitable evaporation of the sample(s) that will occur upon transport, storage and handling by the suspect and the investigators. In light of this information, the higher boiling compounds present in gasoline were investigated to determine their utility in characterising both unevaporated and evaporated gasoline.

The use of higher boiling trace polar compounds and polycyclic aromatic hydrocarbons (PAHs) in liquid gasoline to differentiate one sample from another was explored. Methods were developed to extract and identify these trace compounds. It was found that polar compounds did not vary significantly from one sample to another and so these compounds were not suitable for differentiating gasoline samples. Solid phase extraction of polycyclic aromatic hydrocarbons from different gasoline samples revealed that PAH composition did vary from one gasoline sample to another. A GC-MS (SIM) method was developed to directly analyse for both two- and three-ring PAHs, however, precision in detecting lower concentrations of three-ring PAHs was not as high as for two-ring PAHs. This resulted in the development of a one-step GC-MS (SIM) method that allowed for precise determination of C_0- to C_2-naphthalenes in unevaporated gasoline. It has been shown that 35 unevaporated gasoline samples could be classified into 32 unique groups using principal component analysis and linear discriminant analysis (with cross-validation). Furthermore, a group of 17 samples collected from 14 service stations on the same day could all be distinguished from one another thus demonstrating the usefulness of this method for discriminating unevaporated gasoline samples.

A second study involved the comparison of 96 unevaporated gasoline samples (48 regular unleaded and 48 premium unleaded) collected from three different branded service stations over 16 consecutive weeks. Using the methods described here, all 32 BP samples (16 regular and 16 premium) and all 32 Shell samples (16 regular and 16 premium) could be differentiated using PCA and LDA. The 32 Caltex samples (16 regular and 16 premium) could be classified into 29 unique groups. This study showed that a sample collected from a service station one week is likely to be different from a sample collected from the same station the following week. These results also demonstrated that unevaporated gasoline samples collected in the same week are likely to differ from one branded service station to the next. Another study compared 14 summer and 14 winter gasoline samples collected in Auckland, New Zealand. Using the methods described here the 14 summer samples could be classified into seven unique groups and all 14 winter samples could be differentiated. Results from a comparison of gasoline samples collected in Auckland, New Zealand and Sydney, Australia in the same month demonstrated that there were often significant differences between gasoline samples supplied by the New Zealand refinery and samples supplied by the Australian refinery.

The method developed here for the discrimination of unevaporated gasoline samples based on their C_0- to C_2-naphthalene profiles was applied to samples of evaporated gasoline. It was found that the within group variation for a sample of gasoline over a wide range of evaporation levels (0%, 25%, 50%, 75% and 90% by weight) was sufficiently small to allow many gasoline samples to be differentiated from one another. It is recommended, however, that a forensic comparison of gasoline should be attempted with samples that have a similar level of evaporation. One value of the method developed here is that a valid comparison between gasoline samples is possible even when the samples have not undergone exactly the same amount of evaporation. The author cautions against the long term storage of results (*i.e.* creation of a database) for future comparison with other samples because this study has not attempted to explore the validity of this approach. Furthermore, the author recommends that a comparison be made not only between a sample recovered from a suspect and a sample recovered from a fire, but comparison should also be made with samples collected from other sources. These other sources might include any of the following:

- a service station;
- an acquaintance's garage/shed; or,
- an employer's fuel storage depot;

that is close to any of these or the fire scene:

- the suspect's place of work;
- the suspect's residence; or,
- along a route frequently travelled by the suspect.

Samples collected from other sources should be obtained as close to the time of the incident as practicable.

The research presented here describes a method for the discrimination of gasoline samples based on their C_0- to C_2-naphthalene profiles. This method uses, without modification, a standard gas chromatograph-mass spectrometer like the ones currently used for fire debris analysis by forensic laboratories around the world. The effect of evaporation on a sample, and the speed, precision, and simplicity of the analytical method has been demonstrated.

The method developed here may be extended and applied to the forensic comparison of other petroleum products in the future. Polycyclic aromatic hydrocarbons (not necessarily the naphthalenes) may be used to characterise other products such as aviation gasoline, medium petroleum distillates (*e.g.* paint thinners), kerosene and jet fuels, diesel fuels and heating oils. The application of other analytical techniques such as gas chromatography-carbon isotope ratio mass spectrometry may be attempted to further refine and extend the results described in this work. Future improvements in the analysis and characterisation of liquid petroleum products may also encourage the development of new methods to improve the recovery of petroleum products from fire debris with a view to making a direct comparison between a liquid sample and an extract from fire debris.

Appendix 3.1

Macro used to extract selected aromatic ions from chromatograms of gasoline.

File name: *ms~kovats~model01.mac*

```
!This macro "ms~kovats~model01.mac" is designed to extract aromatics
!from the full scan TIC of evaporated gasolines.
!The integrated peaks are tabulated in the file: <myfile.txt>
!Results are tabulated in the following order:
!peak number; retention time; Kovats index; area area; baseline peak height
!
!(c) Mark Sandercock, University of Technology, Sydney, Australia 2002FEB22
!
!*We*all,*like*sheep,*have*gone*astry,*each*of*us*has*turned*to*his*own*wa
!*                                                                       y
!* Instructions for use:                                                 ;
!* On command line execute                                               *
!* OPEN "C:\MSDChem\1\Data\myfile.txt" for output as #67                 A
!* macro "ms~kovats~model01.mac",go                                      n
!* After macro is complete CLOSE #67                                     d
!*                                                                       *
!******(VIN)*6:35*haiasI**.lla*su*fo*ytiuqini*eht*mih*no*dial*sah*droL*eht
!

NAME DO_HEADER
       !This sub-routine will construct a file header for "myfile.txt",
       !note "/" is a carriage return
       PRINT USING #67,"40%/",_Datafile$
RETURN

NAME DO_EXTRACTION
       !Integrate the peaks in the SIM ion chromatogram in register X
       !Can check the above integrator parameters by entering TABULATE RESULTS
       !on the command line.
       !The following integration parameters have been selected to optimize
       !integration
       INITTHRESH 14       !allowed values are -12 through 25, system default=18
       INITWIDTH 0.019     !allowed values are 0-100, 0.019 is a typical peak
                           !width
       AREAREJECT 0,0      !set reject area to 0 starting at time zero
       INTEGRATE X,0:11.5  !integrates EIC; number of peaks found equals !variable
                           npeaks

       !Set values of all local variables
       counter = 1
       ppeaks = npeaks
       totalarea = 0
       kindex = 0
       image$ = "  3# 3#.3# 5# 9# 9#/"  !image$ sets the fields for the PRINT
                                        !USING command

       !Kovats retention index table of n-alkanes
       !These values are adjusted based on current carbon ladder results
       C6 = 2.05    !Retention time of C6
       C7 = 3.12    !Retention time of C7
       C8 = 5.28    !Retention time of C8
       C9 = 8.16    !Retention time of C9
       C10 = 10.55  !Retention time of C10
       C11 = 12.47  !Retention time of C11
       C12 = 14.13  !Retention time of C12
```

```
      C13 = 15.62  !Retention time of C13

      !Tabulate the desired results by placing selected integration results
      !into a text file for printing or importing:  Set-up the column headings
      !for the integration results for the aromatics; sum peak areas for each
      !extracted ion
             WHILE (counter <= ppeaks) DO
                    PEAKVARS counter,top,X

      !Kovats retention index routine: kovats index calculated as local
      !variable kindex
                    IF ret_time >= C6 AND ret_time <C7 THEN
                          kindex = 100*((ret_time)-(C6))/((C7)-(C6))+600
                    ENDIF

                    IF ret_time >= C7 AND ret_time <C8 THEN
                          kindex = 100*((ret_time)-(C7))/((C8)-(C7))+700
                    ENDIF

                    IF ret_time >= C8 AND ret_time <C9 THEN
                          kindex = 100*((ret_time)-(C8))/((C9)-(C8))+800
                    ENDIF

                    IF ret_time >= C9 AND ret_time <C10 THEN
                          kindex = 100*((ret_time)-(C9))/((C10)-(C9))+900
                    ENDIF

                    IF ret_time >= C10 AND ret_time <C11 THEN
                          kindex = 100*((ret_time)-(C10))/((C11)-(C10))+1000
                    ENDIF

                    IF ret_time >= C11 AND ret_time <C12 THEN
                          kindex = 100*((ret_time)-(C11))/((C12)-(C11))+1100
                    ENDIF

                    IF ret_time >= C12 AND ret_time <C13 THEN
                          kindex = 100*((ret_time)-(C12))/((C13)-(C12))+1200
                    ENDIF
             !End Kovats retention index routine

                    PRINT USING #67,image$,counter,ret_time,kindex,peak_area,
                    bslpeakheight
                    !Write peak to file
                    !PEAKVARS Sets variables with information from the
                    !integrated chromatogram in register X; each peak is
                    !designated by "counter" (1,2,...npeaks)
                    counter = counter + 1
             ENDWHILE

      !Need to write dummy data to fill <myfile.txt.> with the same number of
      !line entries per sample (ie number of peaks found will vary from sample
      !to sample; this is important for parsing data in spreadsheet using a
      !macro
             WHILE (counter <= 49) DO
                    PRINT USING #67,image$,counter,99.999,9999,9999,0
                    !Write dummy data to file
                    counter = counter +1
             ENDWHILE

      PRINT "Writing to destination file: ",file_counter," of 100 completed"
      !Command line comment
RETURN
```

```
NAME EIC_INT07.MAC

!This is the main macro
!remove  all  data  from  working  registers  on  the  stack  to  prevent  unexpected
!addition from previous analysis

       !Clear old files from stack
       REMOVE X,Y,Z,T,R0,R1,R2,R3,R4,R5,R6
       file_counter = 1
       XLOW = 0
       XHIGH = 30

       WHILE (file_counter <= 100) DO
       !Counter needs to be set to number of files to extract
       !Asks for file to be worked on
       FILE
       TIC,R0

       !Extract and sums ions: 78, 91, 105, 119 and molecular ions 106,120 and
       !134

       PRINT "Extracting Aromatic Ions" !Command line comment

       EIC XLOW:XHIGH,78
       EIC XLOW:XHIGH,77   !include this ion for benzene (m/z78=100%, m/z77=28%)
       ADD X,Y
       EIC XLOW:XHIGH,91
       ADD X,Y
       EIC XLOW:XHIGH,92   !include this ion for toluene (m/z91=100%, m/z92=62%)
       ADD X,Y
       EIC XLOW:XHIGH,105
       ADD X,Y
       EIC XLOW:XHIGH,119
       ADD X,Y
       EIC XLOW:XHIGH,133
       ADD X,Y
       EIC XLOW:XHIGH,106
       ADD X,Y
       EIC XLOW:XHIGH,120
       ADD X,Y
       EIC XLOW:XHIGH,134
       ADD X,Y

       LOCAL counter,ppeaks,totalarea,image$
       !set local variables for use in this macro

       !OPEN "D:\GCMS\Data\Mark_S\myfile.txt" for output as #67
       DO_HEADER     !Call subroutine DO_HEADER
       DO_EXTRACTION !Call subroutine DO_EXTRACTION

       file_counter = file_counter + 1
       ENDWHILE

       PRINT "Finished File",_Datafile$,"....Close #67 to end macro session"
       !Command line comment
RETURN
```

Appendix 3.2

Minitab macro for the calculation of error ellipses. Macro was modified to allow user input of the F-statistic.
Source: Lindsay, W.S. Bivariate Control Chart: Uses Hotelling's T in combination with principal components analysis. January 1, 1984. < http://www.minitab.com/support/macros/index.asp?cat=QC_DOE#17> Accessed May 8, 2002.

Macro file name: *minitab_ellipse3.mtb*

```
# TITLE: Bivariate Control Chart
# DESCRIPTION: Uses Hotelling's T in combination with principal
# components analysis.
# AUTHOR: Wayne S. Lindsay, 1/1/1984.
# USE: Use with releases: 9.x or later
#
# URL http://www.minitab.com/support/macros/index.asp?cat=QC_DOE#17
# Accessed on May 8, 2002
#
# Macro written for Minitab (R), version 13 for Windows
#
# Before running this macro, input data must be in C1 and C2
# (of equal length) and K11 must contain the proper F value
# I.E. F(2,N-2,alpha) where N is the number of data points
# and alpha is significance level desired for chart.
#
# This procedure uses Hotelling's T in combination
# with principal component analysis as per the article:
# Jackson, J.E. Principal components and factor analysis: part I
# Journal of Quality Technology, 1980; 12(4): 201-213
#
#
# Macro Modified by P.M.L. Sandercock (see below)
# University of Technology, Sydney
# May 7, 2002
#
# Modification made was to allow user input of F statistic in K11
# from keyboard.  F-statistic based on 95% confidence interval
# automatically calculated and displayed on command line for user
# input.
#
# Before executing the macro:
# load X,Y data into columns C1,C2
# display command line by enabling commands from the <Editor> menu
# Select the macro from the <File> menu, chosing <Other files>,
# <Run an Exec...>
#
# Values for the plotting of the ellipsoid are found at the end of
# columns C5 and C6 (variable number 2).  These may be copied and
# pasted into a spreadsheet and used for plotting purposes outside
  Minitab.
#
# CODE
NOECHO      #turn off echo
#
```

```
COUNT C1 K2
let k21 = k2 - 2
invcdf 0.95;
F 2 k21 0.
NOTE Input the F statistic value.
NOTE I.e. enter the value for x (above) from the Inverse Cumulative
Distribution Function
set 'terminal' c9;      #input value from keyboard (terminal) of
                        #Rho (one value/observation expected)
nobs 1.                 #nobs is number of observations expected
                        #from input command
let k11 = c9(1)         #i.e. k1 = value of Rho inputted above (Column
C1, cell 1)
#
LET K10=2*((K2-1)/(K2-2))*K11
LET K12=SQRT(K10)       #Hotelling’s T
LET C11=C1-MEAN(C1)
LET C12=C2-MEAN(C2)
LET K1=1/(K2-1)
COPY C11 C12 M1
TRANS M1 M2
MULT M2 M1 M3
MULT K1 M3 M3           #Sample covariance matrix
EIGE M3 C10 M4
SQRT C10 C10
COPY M4 C3 C4           #Principal components
LET K4=C10(1)*C3(1)
LET K5=C10(1)*C3(2)     #These are the SQRT(Eigenvalue) = Eigenvector
LET K6=C10(2)*C4(1)
LET K7=C10(2)*C4(2)
#
# NOW GET SQRTS OF T**2-1 AND T**2-4
#
LET K13=SQRT(K10-1)
LET K14=-1*K13
LET K15=SQRT (K10-4)    #Does not always return a valid number
                        #(i.e. sqrt(neg))
LET K16=-1*K15
LET K17=-1*K12
#
#CREATE AND PLOT ELLIPSE
#
STACK 0 0 K12 K17 1 -1 1 -1 K13 K14 K13 K14 2 -2 2 -2 &
K15 K16 K15 K16 C14
STACK K12 K17 0 0 K13 K13 K14 K14 1 1 -1 -1 K15 K15 K16 &
K16 2 2 -2 -2 C13
LET K8=MEAN(C1)
LET K9=MEAN(C2)
LET C5=K8+(C13*K4)+(C14*K6)
LET C6=K9+(C13*K5)+(C14*K7)
stack c2 c6 c6;
subs c7.
stack c1 c5 c5
PLOT C6*C5;             #Data and control ellipse
symbol c7.
#endofmacro
```

Appendix 4.1

The following two pages show the data collection sheet that was completed with each sample collected. Fuel distributor information was strictly voluntary and many service station employees either did not know the information or were reluctant to provide it.

UNIVERSITY OF
TECHNOLOGY SYDNEY

Eric Du Pasquier & Mark Sandercock
Department of Chemistry, Materials & Forensic Science
PO Box 123, Broadway, NSW 2007
Phone: 9514 1743 Fax: 9514 1460

Reference Number:	

Unleaded Petrol Sample Information

Collection Details	Collection Date: (yyyy/mmm/dd)	
	Collection Time: (24 hour clock)	
	Collected By: (print name)	

Petrol Station Information

Name of Petrol Station:
☐ Ampol ☐ BP ☐ Burmah
☐ Caltex ☐ Mobil ☐ Shell
☐ Other:
(specify)

Street Address:

Suburb/Town:

State: ☐ NSW ☐ NT ☐ QLD
☐ SA ☐ VIC ☐ WA ☐ ACT

Post Code:

↓ Unleaded Petrol ONLY ↓

Petrol Sample	Pump Octane Rating: (check one)	☐ Regular	☐ Premium

Fuel Distributor Information

Fuel Distributor	Delivery Schedule	Regular	Premium (Super etc.)
	Last Delivery Date: (yyyy/mmm/dd)		
	Volume Delivered: (Litres)		
	Delivery Frequency: (Days)		

Sample Collection Protocol:

Please read these instructions prior to sample collection.

1. It is not necessary to discharge petrol from the pump prior to sample collection.

    A. IF petrol was pumped into a vehicle tank by the collector, or was pumped by another customer immediately prior to the collection, then indicate the amount of petrol pumped in **Collection Details** box "Amount pumped prior to sample collection."

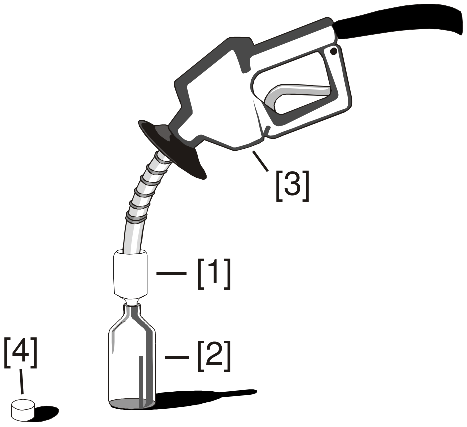
    B. IF the petrol station pump was idle prior to collection then mark zero in **Collection Details** box "Amount pumped prior to sample collection."

2. Place a new disposable plastic funnel **[1]** into mouth of 120 millilitre amber glass collection bottle **[2]**.

3. Pull the pump nozzle trigger **[3]** slowly and gently. Fill the bottle with approximately 100 millilitres (0.1 litres) of petrol and cap **[4]** immediately. Ensure that cap is tight.

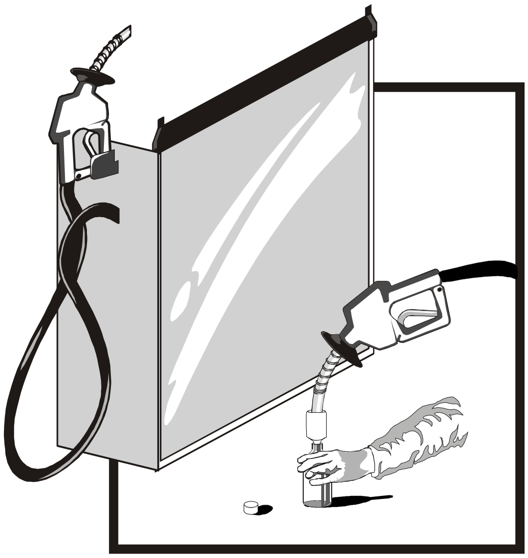
4. Discard plastic funnel in rubbish bin. Do not reuse funnel.

5. Write the reference number provided on the bottle cap **[4]** in the top right section of the form marked **Reference Number**.



6. Complete the form obtaining all information requested. This information is voluntary so if the attendant is unable/unwilling to divulge the **Fuel Distributor Information** then please obtain a contact number for follow-up with the petrol station owner/supervisor.

7. Store petrol sample in a cool, dark place with good ventilation. Do not expose to direct sunlight and do not leave sample in boot unless transporting it. Do not transport or store sample in vehicle interior at any time.



Appendix 5.1

Macro written for the extraction of GC-MS (SIM) data using the HP ChemStation.

File name: *ms~kovats~simnaph.mac*

```
!This macro "ms~kovats~simnaph.mac" is designed to extract individual peaks
!from the chromatogram of C0- to C2-naphthalenes extracts obtained via SIM.
!The integrated peaks are tabulated in the file: <myfile.txt>
!Results are tabulated in the following order:
!peak number; retention time; kovats index; area area; baseline peak height
!
!(c) Mark Sandercock, University of Technology, Sydney, Australia 2002FEB22
!
!*What*good*is*it*for*a*man*to*gain*the*whole*world,*yet*forfeit*his**soul
!*                                                                       ?
!* Instructions for use:                                                 *
!* Make changes to RT values of aromatic ring standards for Kovats       *
!* On command line execute                                               F
!* OPEN "C:\MSDChem\1\Data\myfile.txt" for output as #67                 o
!* macro "ms~kovats~simnaph.mac",go                                      r
!* After macro is complete CLOSE #67                                     *
!*                                                                       w
!************(VIN)*63:8*kraM**?luos*sih*rof*egnahcxe*ni*evig*nam*a*nac*tah
!

NAME DO_HEADER
       !This sub-routine will construct a file header for "myfile.txt",
       !note "/" is a carriage return
       PRINT USING #67,"40%/",_Datafile$
RETURN

NAME DO_EXTRACTION
       !Integrate the peaks in the SIM ion chromatogram in register X
       !Can check the above integrator parameters by entering TABULATE RESULTS
       !on the command line.

       !The following integration parameters have been selected to optimize
       !integration
       INITTHRESH 12       !allowed values are -12 through 25, system default=18
       INITWIDTH 0.019     !allowed values are 0-100, 0.019 is a typical peak
                           !width
       AREAREJECT 0,0      !set reject area to 0 starting at time zero
       INTEGRATE X,18:29   !integrates EIC; number of peaks found equals
                           !variable npeaks

       !Set values of all local variables
       counter = 1
       ppeaks = npeaks
       totalarea = 0
       kindex = 0
       image$ = "   3# 3#.3# 5# 9# 9#/"
       !image$ sets the fields for the print using command to use

       !Kovats retention index table of n-alkanes
       !These values must be changed to reflect results of current standard
       onering = 5.00      !Retention time of Benzene assigned default value of
                           !5.00 minutes
       tworing = 19.20      !Retention time of naphthalene  (2-ring)
       threering = 37.50    !Retention time of phenanthrene (3-ring)
       fourring = 48.57    !Retention time of chrysene     (4-ring)
```

```
       !Tabulate the desired results by placing selected integration results
       !into a text file for printing or importing:  Set-up the column headings
       !for the integration results for the aromatics; sum peak areas for each
       !extracted ion
       WHILE (counter <= ppeaks) DO
                    PEAKVARS counter,top,X

       !Kovats retention index routine: kovats index calculated as local
       !variable kindex
       IF ret_time >= onering AND ret_time <tworing THEN
       kindex = 1000*((ret_time)-(onering))/((tworing)-(onering))+1000
       ENDIF

       IF ret_time >= tworing AND ret_time <threering THEN
       kindex = 1000*((ret_time)-(tworing))/((threering)-(tworing))+2000
       ENDIF

       IF ret_time >= threering AND ret_time <fourring THEN
       kindex = 1000*((ret_time)-(threering))/((fourring)-(threering))+3000
       ENDIF
       !End Kovats retention index routine

       PRINT USING #67,image$,counter,ret_time,kindex,peak_area,bslpeakheight
       !Write peak to file
       !PEAKVARS Sets variables with information from the integrated
       !chromatogram in register X; each peak is designated by "counter"
       !(1,2,...npeaks)
       counter = counter + 1
       ENDWHILE

       !Need to write dummy data to fill <myfile.txt.> with the same number of
       !line entries per sample (ie number of peaks found will vary from sample
       !to sample; this is important for parsing data in spreadsheet using a
       !macro
              WHILE (counter <= 99) DO
                     PRINT USING #67,image$,counter,99.999,9999,9999,0
                     !Write dummy data to file
                     counter = counter +1
              ENDWHILE

       PRINT "Writing to destination file: ",file_counter," of 175 completed"
       !Command line comment
RETURN


NAME EIC_INT07.MAC

!This is the main macro
       !remove all data from working registers on the stack to prevent
       !unexpected ADDition
       !from previous analysis
       REMOVE X,Y,Z,T,R0,R1,R2,R3,R4,R5,R6
       file_counter = 1


       WHILE (file_counter <= 175) DO   !Counter needs to be set to number of
                                        !files to extract
       !Asks for file to be worked on
       FILE
       TIC    !Places TIC in Register X

       !Set values of global variables
```

```
       !XLOW=32            !start time for scanning chromatogram
       !XHIGH=48           !end time for scanning chromatogram
       !MzRANGE = -0.3:0.7 !Sets m/z range of extracted ions

       LOCAL counter,ppeaks,totalarea,image$  !set local variables for use in
                                               !this macro

       !OPEN "D:\GCMS\Data\Mark_S\myfile.txt" for output as #67
       DO_HEADER     !Call subroutine DO_HEADER
       DO_EXTRACTION !Call subroutine DO_EXTRACTION

       file_counter = file_counter + 1
       ENDWHILE

       PRINT "Finished File",_Datafile$,"....Close #67 to end macro session"
RETURN
```

Appendix 6.1

The following pages contain PC score plots of BP premium unleaded and regular unleaded, Caltex premium unleaded and regular unleaded, and Shell premium unleaded and regular unleaded gasoline for each of the 16 weeks that samples were collected.

Figure App 6.1.1 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 1 and 2 of a 16 consecutive week collection.

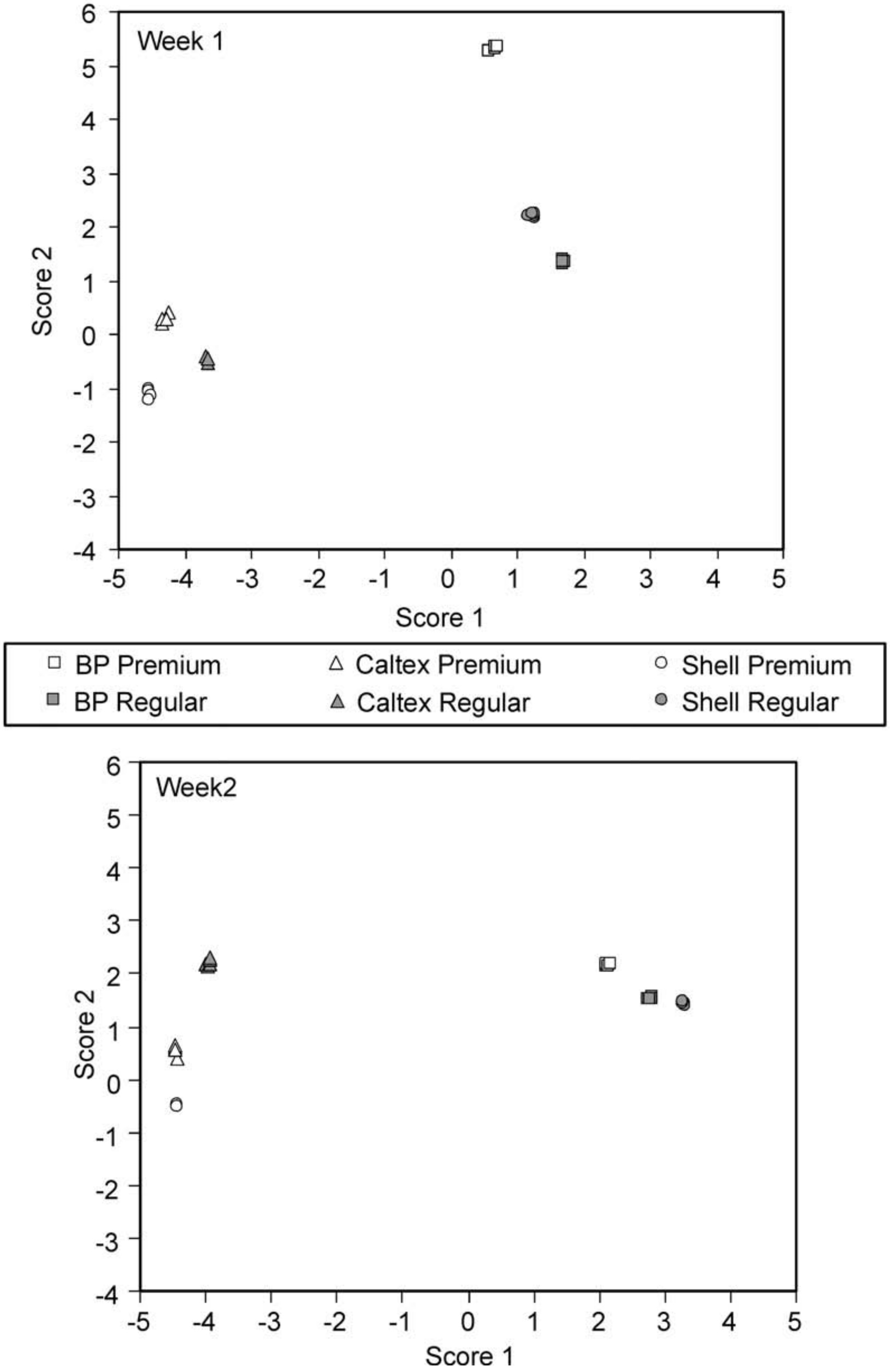


Figure App 6.1.2 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 3 and 4 of a 16 consecutive week collection.

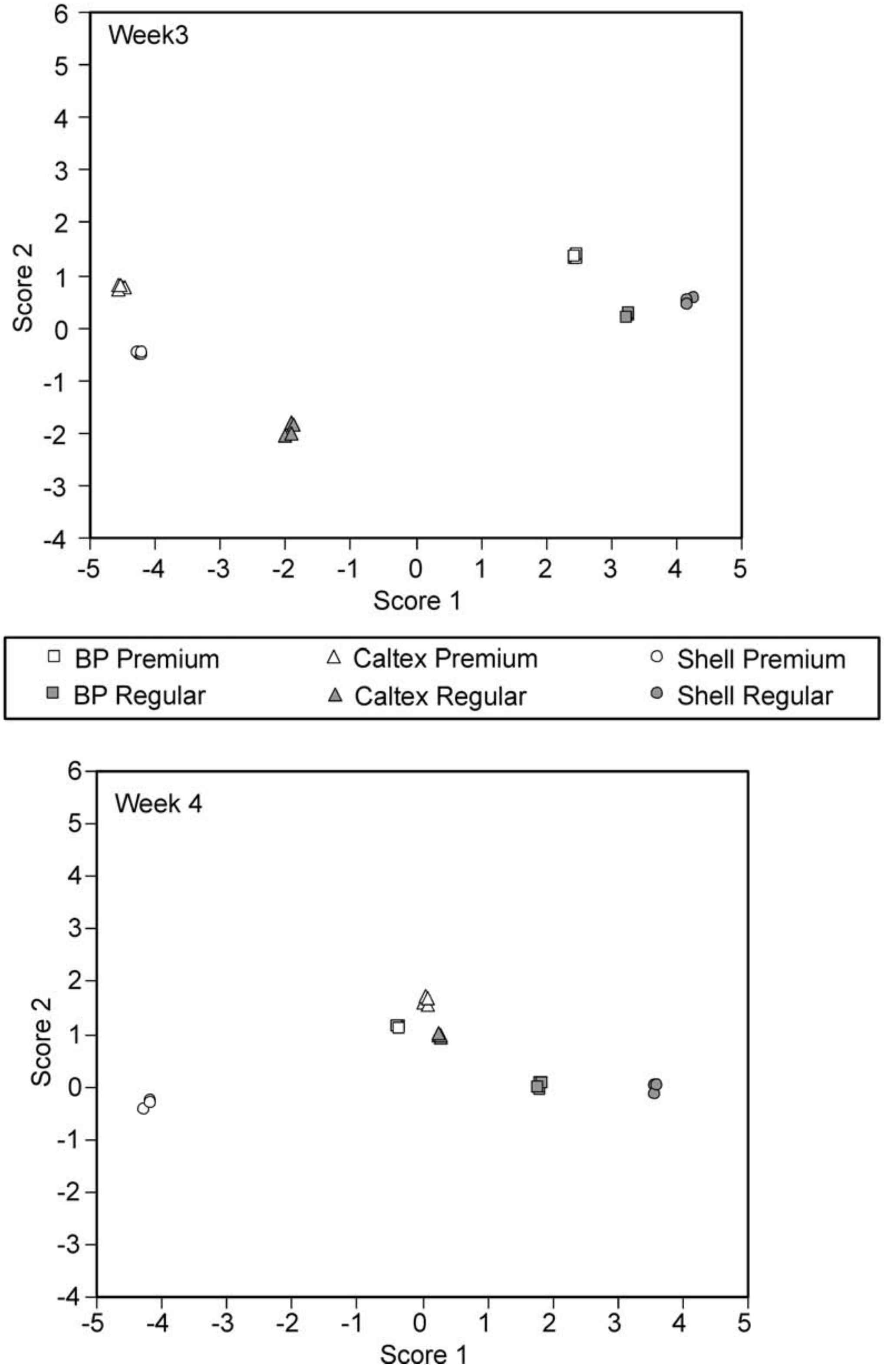


Figure App 6.1.3 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 5 and 6 of a 16 consecutive week collection.

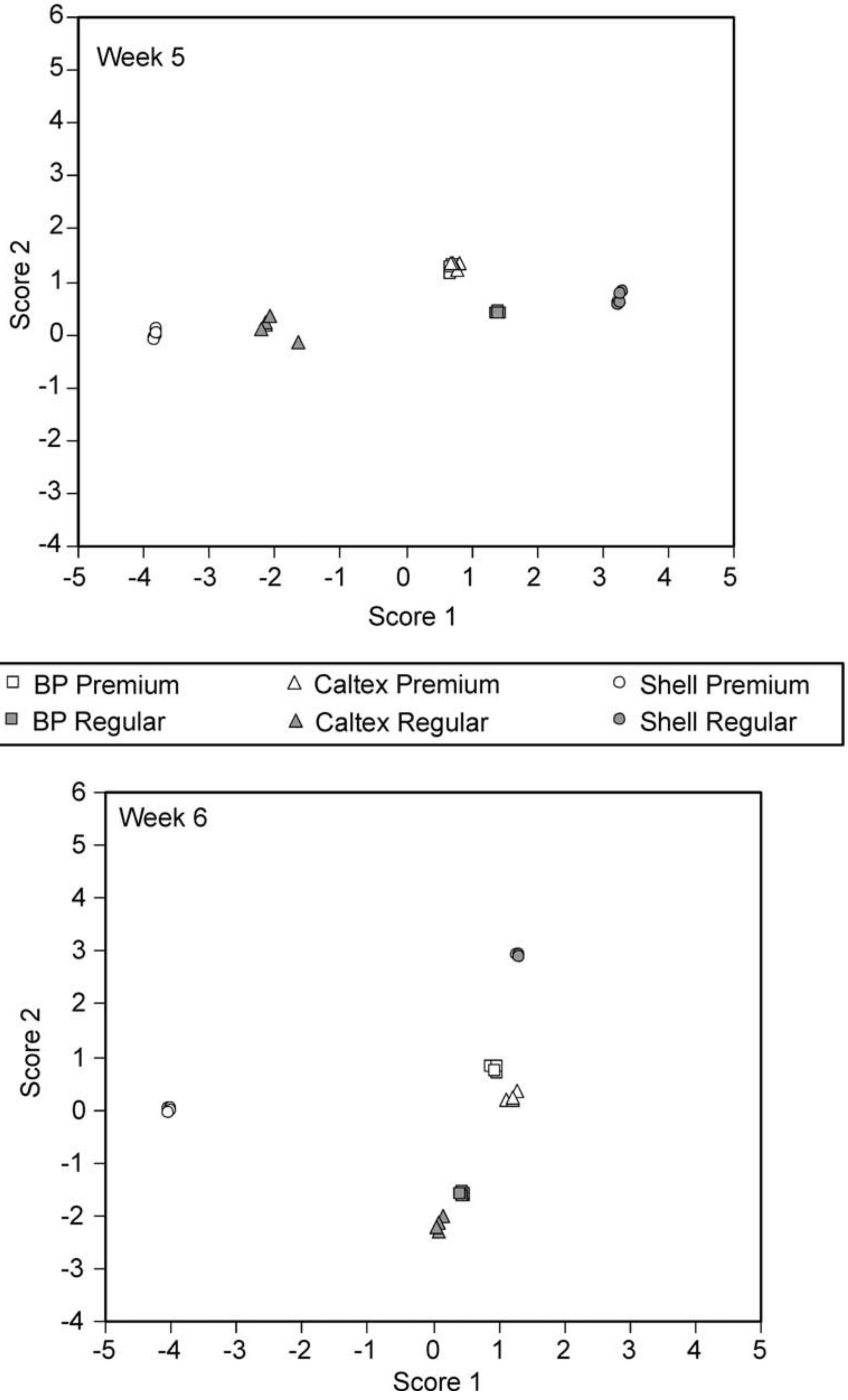


Figure App 6.1.4 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 7 and 8 of a 16 consecutive week collection.

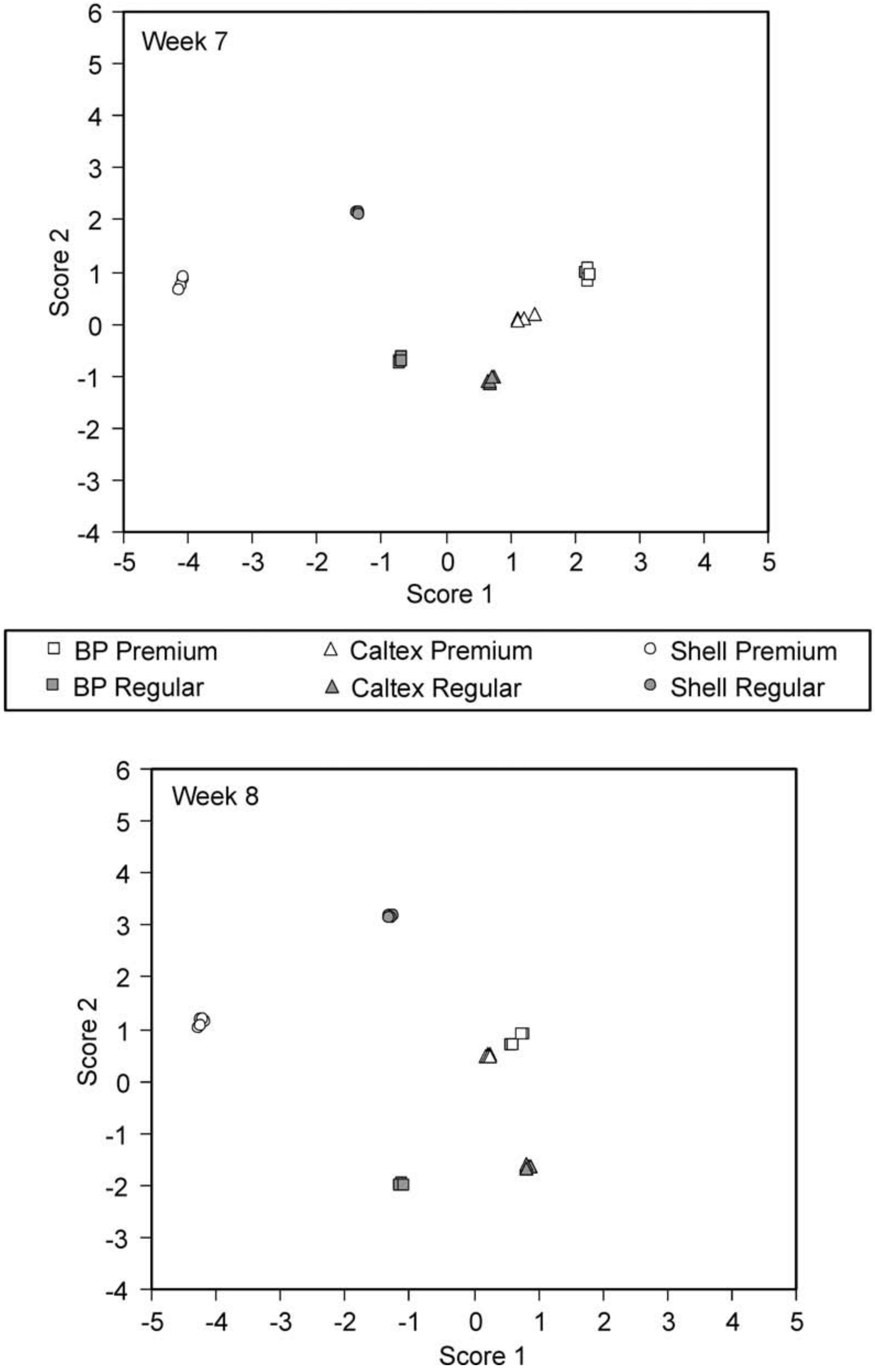


Figure App 6.1.5 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 9 and 10 of a 16 consecutive week collection.

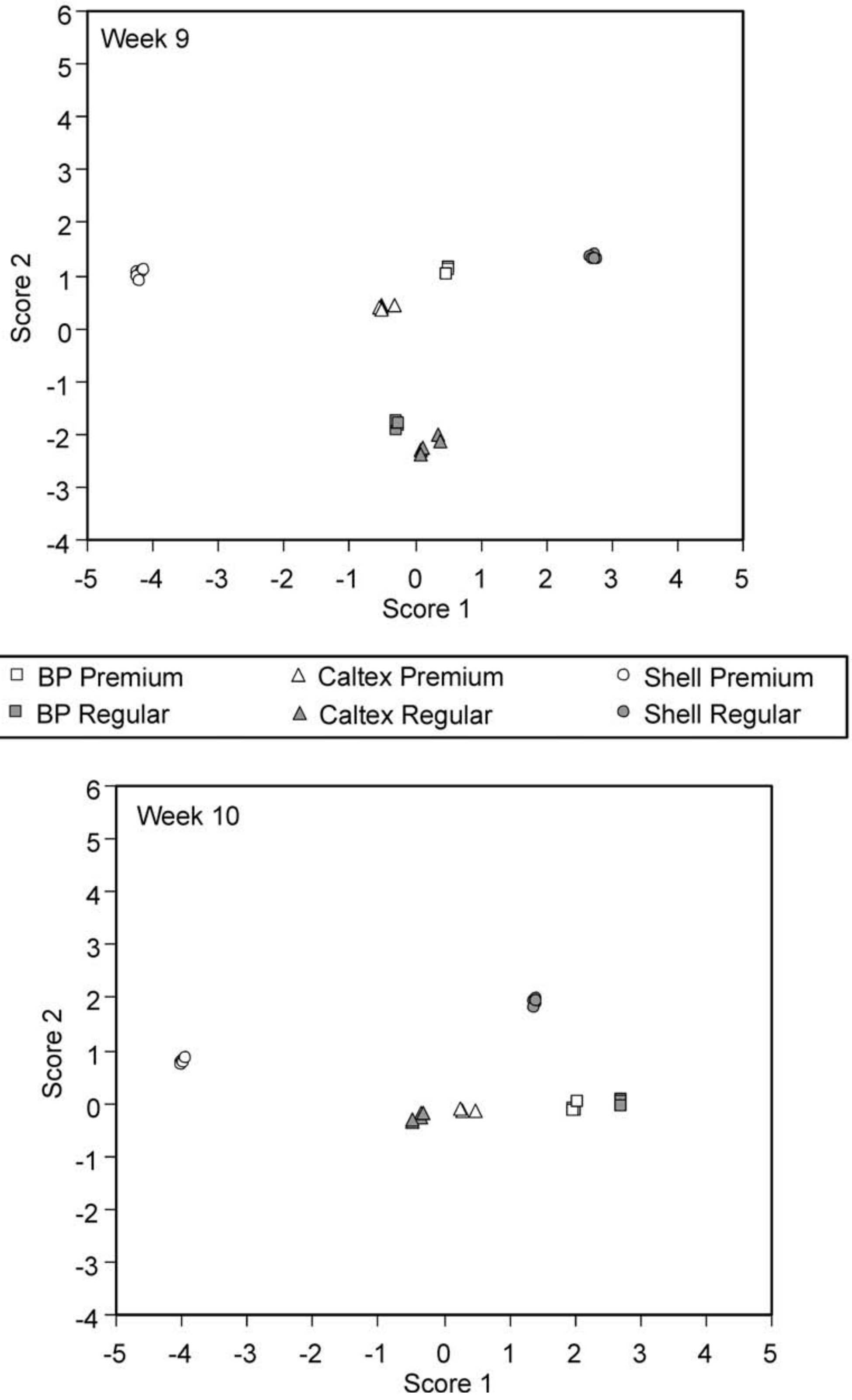


Figure App 6.1.6 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 11 and 12 of a 16 consecutive week collection.

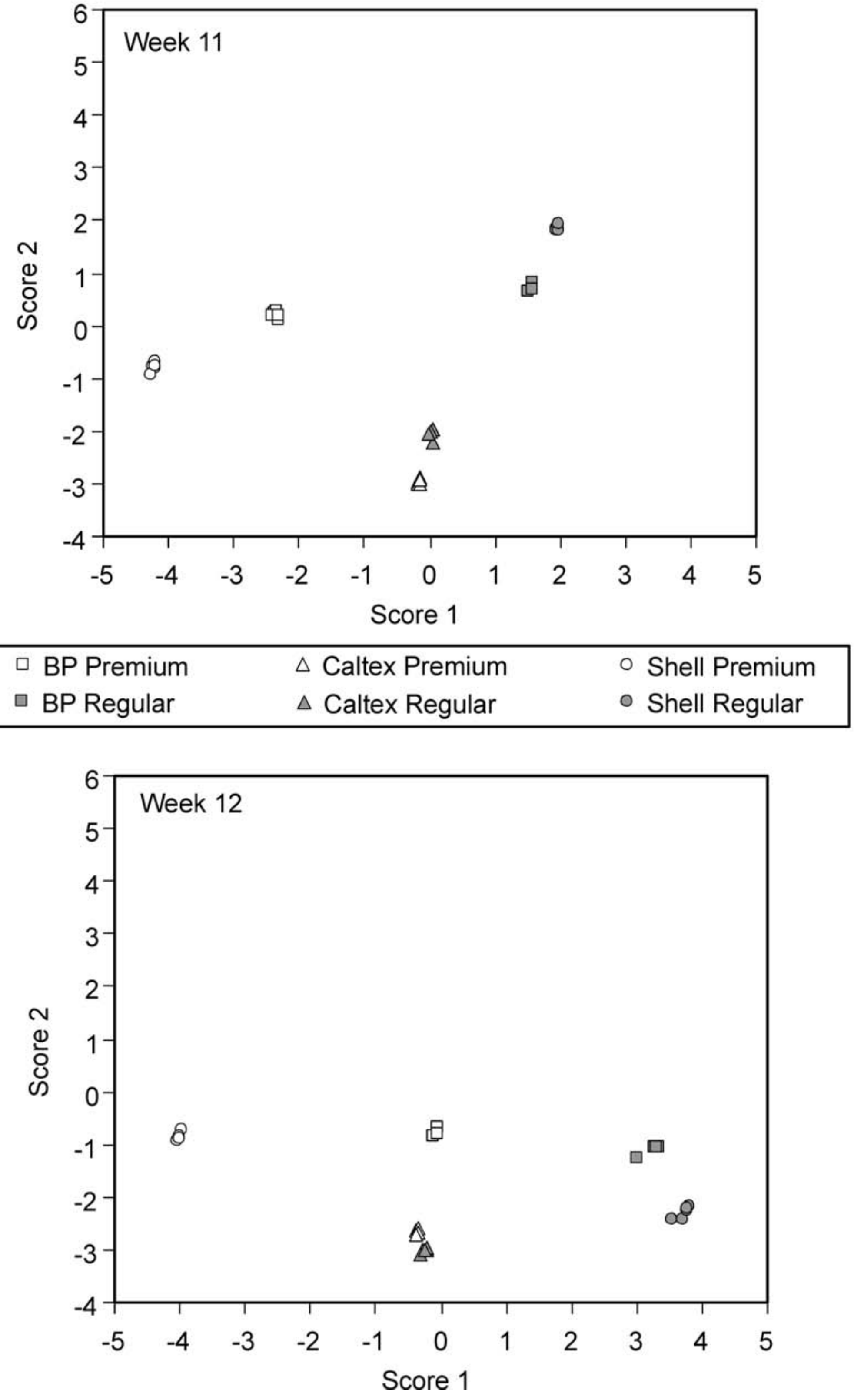


Figure App 6.1.7 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 13 and 14 of a 16 consecutive week collection.

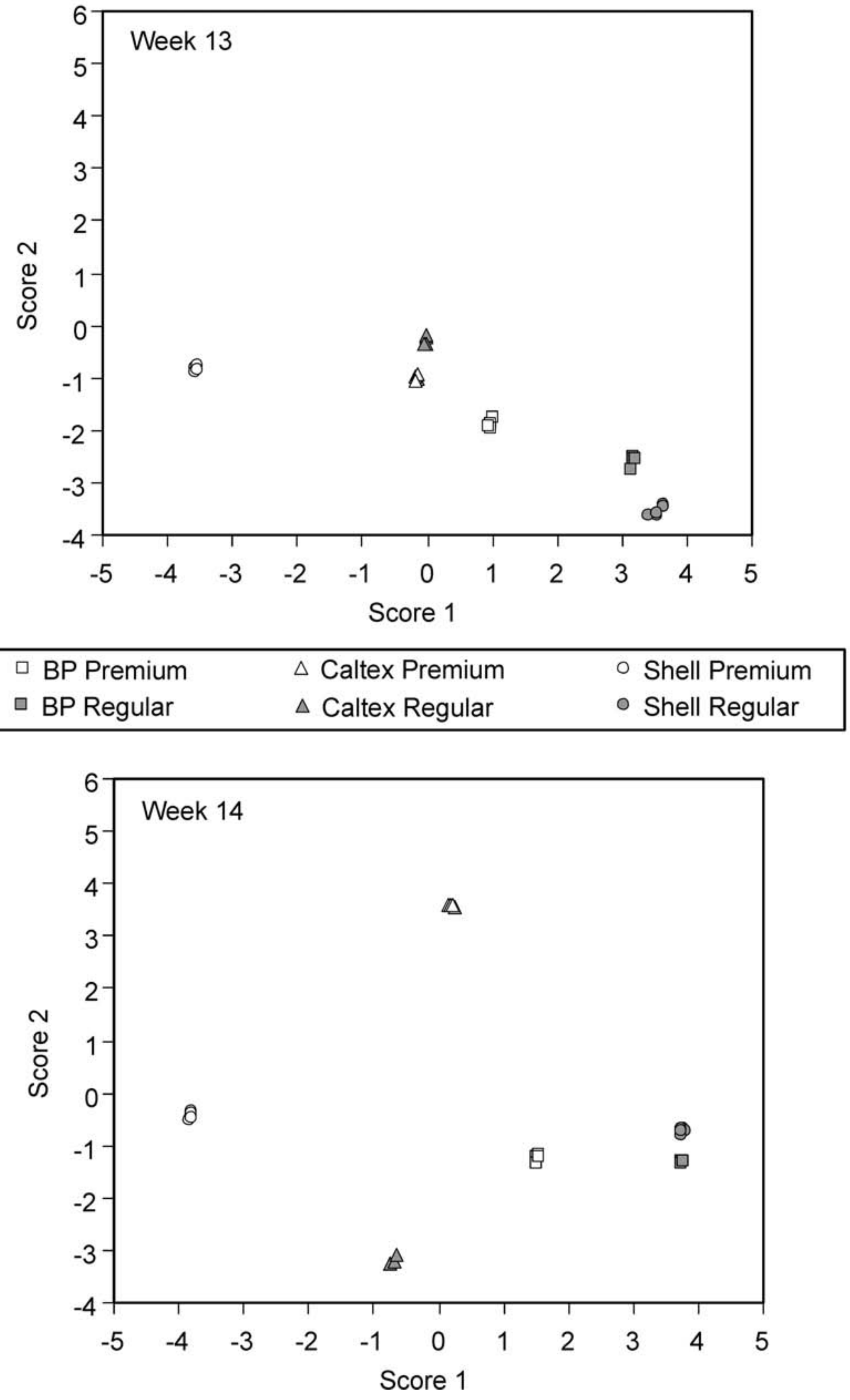
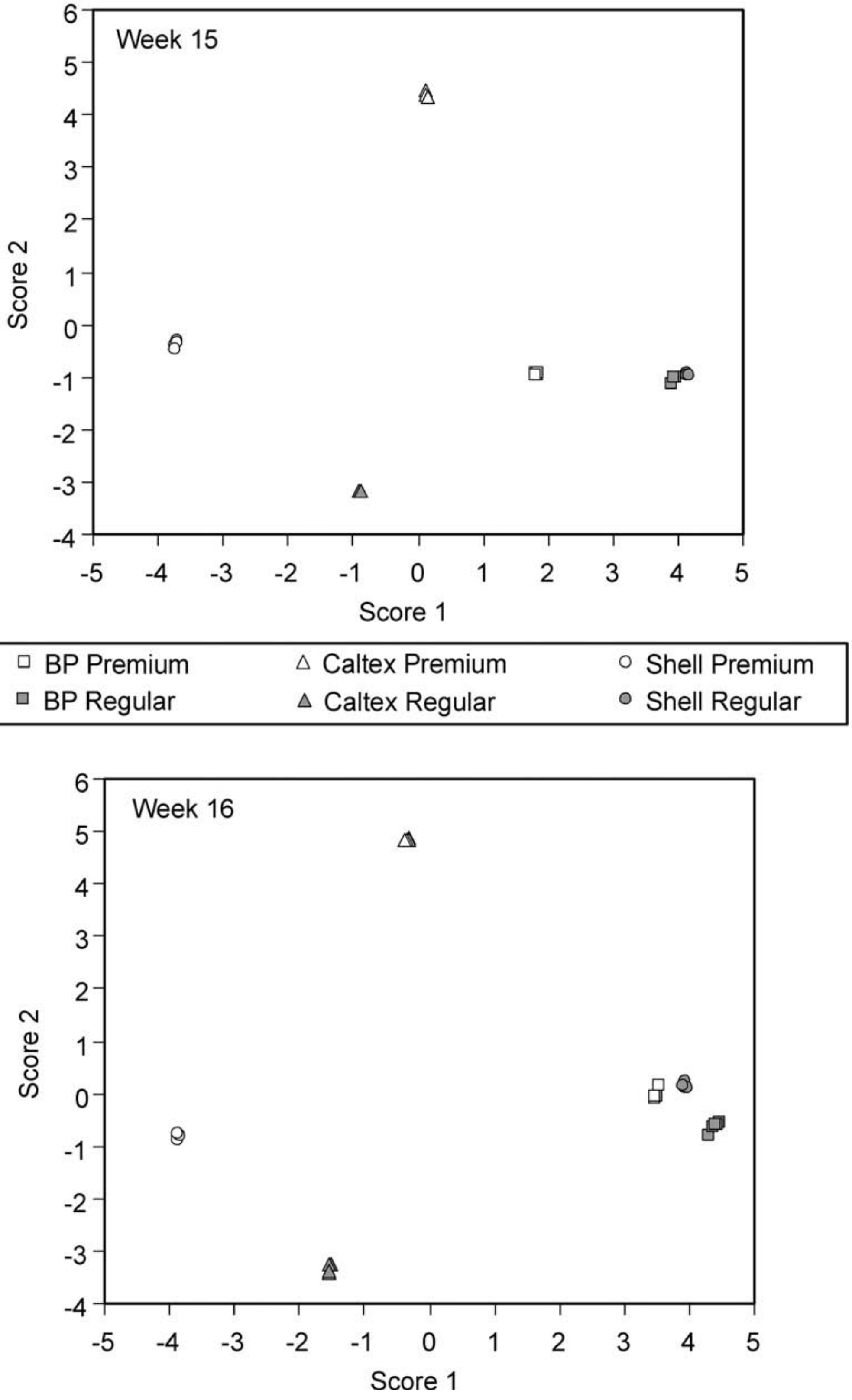


Figure App 6.1.8 PCA score plot of BP, Caltex and Shell premium and regular unleaded gasolinesamples showing only samples collected on weeks 15 and 16 of a 16 consecutive week collection.



Appendix 7.1

Tables of evaporation data for 35 gasoline samples.

25% evaporated samples

	Weight (grams)						
Sample	Empty vial & cap	Vial, cap & sample	Initial weight sample	Vial, cap & sample (target)	Vial, cap & sample (actual)	Final weight sample	Evaporation level (%)
RU2	6.3346	9.5799	3.2453	8.7686	8.7545	2.4199	25.43
LR6	6.2388	9.5249	3.2861	8.7034	8.7039	2.4651	24.98
PU7	6.1985	9.4920	3.2935	8.6686	8.6692	2.4707	24.98
RU11	6.2394	9.4202	3.1808	8.6250	8.5640	2.3246	26.92
PU12	6.2423	9.4966	3.2543	8.6830	8.6707	2.4284	25.38
RU13	6.3454	9.4684	3.1230	8.6877	8.6725	2.3271	25.49
RU14	6.1999	9.4755	3.2756	8.6566	8.5992	2.3993	26.75
RU15	6.2353	9.4124	3.1771	8.6181	8.5912	2.3559	25.85
RU16	6.2537	9.4545	3.2008	8.6543	8.6517	2.3980	25.08
PU17	6.2681	9.5600	3.2919	8.7370	8.7367	2.4686	25.01
RU18	6.2339	9.5328	3.2989	8.7081	8.7000	2.4661	25.24
PU19	6.2568	9.5739	3.3171	8.7446	8.7340	2.4772	25.32
RU20	6.2558	9.4187	3.1629	8.6280	8.5871	2.3313	26.29
RU21	6.3311	9.4999	3.1688	8.7077	8.6460	2.3149	26.95
RU22	6.2195	9.4293	3.2098	8.6269	8.5715	2.3520	26.72
RU23	6.2463	9.4416	3.1953	8.6428	8.6352	2.3889	25.24
RU24	6.2073	9.4147	3.2074	8.6129	8.5885	2.3812	25.76
RU25	6.1921	9.3583	3.1662	8.5668	8.4964	2.3043	27.22
RU26	6.2175	9.4562	3.2387	8.6465	8.5995	2.3820	26.45
RU27	6.2128	9.4355	3.2227	8.6298	8.6097	2.3969	25.62
PU28	6.2363	9.5615	3.3252	8.7302	8.6881	2.4518	26.27
RU29	6.2605	9.4161	3.1556	8.6272	8.6068	2.3463	25.65
RU30	6.2364	9.4712	3.2348	8.6625	8.6169	2.3805	26.41
PU31	6.3096	9.4856	3.1760	8.6916	8.6850	2.3754	25.21
RU32	6.1826	9.3394	3.1568	8.5502	8.4730	2.2904	27.45
RU33	6.2224	9.4019	3.1795	8.6070	8.5471	2.3247	26.88
RU34	6.2313	9.3933	3.1620	8.6028	8.5178	2.2865	27.69
RU35	6.2724	9.4947	3.2223	8.6891	8.6271	2.3547	26.92
RU36	6.2205	9.4536	3.2331	8.6453	8.6390	2.4185	25.20
PU37	6.2877	9.4048	3.1171	8.6255	8.6221	2.3344	25.11
RU38	6.2595	9.5085	3.2490	8.6963	8.6882	2.4287	25.25
PU39	6.2328	9.3679	3.1351	8.5841	8.5872	2.3544	24.90
RU40	6.3268	9.4505	3.1237	8.6696	8.6696	2.3428	25.01
PU41	6.2540	9.4127	3.1587	8.6230	8.6230	2.3690	25.17
RU42	6.2213	9.3033	3.0820	8.5328	8.5328	2.3115	24.97

50% evaporated samples

	Weight (grams)						
Sample	Empty vial & cap	Vial, cap & sample	Initial weight sample	Vial, cap & sample (target)	Vial, cap & sample (actual)	Final weight sample	Evaporation level (%)
RU2	6.2230	9.4048	3.1818	7.8139	7.7997	1.5767	50.45
LR6	6.2621	9.5076	3.2455	7.8849	7.8895	1.6274	49.86
PU7	6.2691	9.6288	3.3597	7.9490	7.9406	1.6715	50.25
RU11	6.2508	9.4066	3.1558	7.8287	7.7772	1.5264	51.63
PU12	6.2492	9.6012	3.3520	7.9252	7.9360	1.6868	49.68
RU13	6.2148	9.3062	3.0914	7.7605	7.7495	1.5347	50.36
RU14	6.2395	9.4185	3.1790	7.8290	7.8213	1.5818	50.24
RU15	6.2037	9.4441	3.2404	7.8239	7.8065	1.6028	50.54
RU16	6.2533	9.4646	3.2113	7.8590	7.8552	1.6019	50.12
PU17	6.2830	9.8631	3.5801	8.0731	8.0741	1.7911	49.97
RU18	6.1928	9.4436	3.2508	7.8182	7.8163	1.6235	50.06
PU19	6.1910	9.4778	3.2868	7.8344	7.8350	1.6440	49.98
RU20	6.2478	9.5418	3.2940	7.8948	7.8907	1.6429	50.12
RU21	6.2463	9.5011	3.2548	7.8737	7.8684	1.6221	50.16
RU22	6.2389	9.4405	3.2016	7.8397	7.8221	1.5832	50.55
RU23	6.2444	9.5419	3.2975	7.8932	7.9033	1.6589	49.69
RU24	6.2025	9.4284	3.2259	7.8155	7.8362	1.6337	49.36
RU25	6.2122	9.4416	3.2294	7.8269	7.7893	1.5771	51.16
RU26	6.2263	9.4771	3.2508	7.8517	7.8448	1.6185	50.21
RU27	6.2899	9.5611	3.2712	7.9255	7.9134	1.6235	50.37
PU28	6.2088	9.5248	3.3160	7.8668	7.8668	1.6580	50.00
RU29	6.1974	9.3863	3.1889	7.7919	7.7795	1.5821	50.39
RU30	6.2209	9.4906	3.2697	7.8558	7.8392	1.6183	50.51
PU31	6.2625	9.7369	3.4744	7.9997	8.0168	1.7543	49.51
RU32	6.2937	9.5599	3.2662	7.9268	7.9059	1.6122	50.64
RU33	6.2769	9.4957	3.2188	7.8863	7.8738	1.5969	50.39
RU34	6.2288	9.5376	3.3088	7.8832	7.8752	1.6464	50.24
RU35	6.2888	9.4782	3.1894	7.8835	7.8768	1.5880	50.21
RU36	6.1969	9.4016	3.2047	7.7993	7.8083	1.6114	49.72
PU37	6.2656	9.5204	3.2548	7.8930	7.9032	1.6376	49.69
RU38	6.2383	9.4037	3.1654	7.8210	7.8072	1.5689	50.44
PU39	6.2044	9.4837	3.2793	7.8441	7.8349	1.6305	50.28
RU40	6.2270	9.4186	3.1916	7.8228	7.8177	1.5907	50.16
PU41	6.2016	9.3910	3.1894	7.7963	7.7931	1.5915	50.10
RU42	6.2639	9.3649	3.1010	7.8144	7.8064	1.5425	50.26

75% evaporated samples

Sample	*Weight (grams)*						Evaporation level (%)
	Empty vial & cap	Vial, cap & sample	Initial weight sample	Vial, cap & sample (target)	Vial, cap & sample (actual)	Final weight sample	
RU2	9.5954	15.8573	6.2619	11.1609	11.1606	1.5652	75.00
LR6	9.6699	16.2763	6.6064	11.3215	11.3220	1.6521	74.99
PU7	9.5289	15.9213	6.3924	11.1270	11.1259	1.5970	75.02
RU11	9.6246	15.7704	6.1458	11.1611	11.1543	1.5297	75.11
PU12	9.6274	16.2446	6.6172	11.2817	11.2737	1.6463	75.12
RU13	9.5583	15.7822	6.2239	11.1143	11.1152	1.5569	74.99
RU14	9.5891	15.9771	6.3880	11.1861	11.1788	1.5897	75.11
RU15	9.5322	15.7232	6.1910	11.0800	11.0833	1.5511	74.95
RU16	9.6618	15.8896	6.2278	11.2188	11.2113	1.5495	75.12
PU17	9.5346	15.7376	6.2030	11.0854	11.0845	1.5499	75.01
RU18	9.5587	15.7866	6.2279	11.1157	11.1169	1.5582	74.98
PU19	9.6180	15.9556	6.3376	11.2024	11.1864	1.5684	75.25
RU20	9.4819	15.6363	6.1544	11.0205	11.0250	1.5431	74.93
RU21	9.6645	15.7407	6.0762	11.1836	11.1805	1.5160	75.05
RU22	9.6372	15.8723	6.2351	11.1960	11.1945	1.5573	75.02
RU23	9.5488	15.8064	6.2576	11.1132	11.0954	1.5466	75.28
RU24	9.5339	15.7173	6.1834	11.0798	11.0810	1.5471	74.98
RU25	9.5952	15.7859	6.1907	11.1429	11.1449	1.5497	74.97
RU26	9.6172	15.8489	6.2317	11.1751	11.1757	1.5585	74.99
RU27	9.5190	15.6541	6.1351	11.0528	11.0544	1.5354	74.97
PU28	9.4914	15.8211	6.3297	11.0738	11.0794	1.5880	74.91
RU29	9.6078	15.6955	6.0877	11.1297	11.1222	1.5144	75.12
RU30	9.5512	15.6952	6.1440	11.0872	11.0738	1.5226	75.22
PU31	9.5472	15.8288	6.2816	11.1176	11.1122	1.5650	75.09
RU32	9.5247	15.6482	6.1235	11.0556	11.0557	1.5310	75.00
RU33	9.6161	15.8658	6.2497	11.1785	11.1709	1.5548	75.12
RU34	9.5289	15.6196	6.0907	11.0516	11.0573	1.5284	74.91
RU35	9.6006	15.8539	6.2533	11.1639	11.1622	1.5616	75.03
RU36	9.5696	15.8679	6.2983	11.1442	11.1438	1.5742	75.01
PU37	9.5827	15.8038	6.2211	11.1380	11.1352	1.5525	75.04
RU38	9.5218	15.5989	6.0771	11.0411	11.0466	1.5248	74.91
PU39	9.5232	15.5879	6.0647	11.0394	11.0132	1.4900	75.43
RU40	9.4928	15.5052	6.0124	10.9959	10.9984	1.5056	74.96
PU41	9.5410	15.6783	6.1373	11.0753	11.0432	1.5022	75.52
RU42	9.5237	15.6271	6.1034	11.0496	11.0308	1.5071	75.31

90% evaporated samples

Sample	*Weight (grams)*						Evaporation level (%)
	Empty vial & cap	Vial, cap & sample	Initial weight sample	Vial, cap & sample (target)	Vial, cap & sample (actual)	Final weight sample	
RU2	9.6177	15.7173	6.0996	10.2277	10.2265	0.6088	90.02
LR6	9.6283	16.1783	6.5500	10.2833	10.2862	0.6579	89.96
PU7	9.5385	16.0268	6.4883	10.1873	10.1970	0.6585	89.85
RU11	9.5722	15.7832	6.2110	10.1933	10.1904	0.6182	90.05
PU12	9.5668	16.1291	6.5623	10.2230	10.2249	0.6581	89.97
RU13	9.5672	15.7894	6.2222	10.1894	10.1896	0.6224	90.00
RU14	9.5157	15.6687	6.1530	10.1310	10.1269	0.6112	90.07
RU15	9.5770	15.7484	6.1714	10.1941	10.1746	0.5976	90.32
RU16	9.5370	15.7086	6.1716	10.1542	10.1594	0.6224	89.92
PU17	9.5481	15.8303	6.2822	10.1763	10.1797	0.6316	89.95
RU18	9.6745	16.0488	6.3743	10.3119	10.3168	0.6423	89.92
PU19	9.5810	15.8822	6.3012	10.2111	10.2124	0.6314	89.98
RU20	9.5387	15.7737	6.2350	10.1622	10.1637	0.6250	89.98
RU21	9.4558	15.5989	6.1431	10.0701	10.0710	0.6152	89.99
RU22	9.6316	15.8749	6.2433	10.2559	10.2443	0.6127	90.19
RU23	9.6003	15.9954	6.3951	10.2398	10.2246	0.6243	90.24
RU24	9.5063	15.6935	6.1872	10.1250	10.1290	0.6227	89.94
RU25	9.5569	15.8851	6.3282	10.1897	10.1602	0.6033	90.47
RU26	9.5555	15.7215	6.1660	10.1721	10.1756	0.6201	89.94
RU27	9.6593	16.1589	6.4996	10.3093	10.3141	0.6548	89.93
PU28	9.5353	15.9308	6.3955	10.1749	10.1778	0.6425	89.95
RU29	9.4889	15.6261	6.1372	10.1026	10.1037	0.6148	89.98
RU30	9.5207	15.7487	6.2280	10.1435	10.1484	0.6277	89.92
PU31	9.5031	15.8852	6.3821	10.1413	10.1478	0.6447	89.90
RU32	9.5592	15.7897	6.2305	10.1823	10.1874	0.6282	89.92
RU33	9.6213	15.9695	6.3482	10.2561	10.2521	0.6308	90.06
RU34	9.5005	15.6698	6.1693	10.1174	10.1221	0.6216	89.92
RU35	9.5372	15.7794	6.2422	10.1614	10.1536	0.6164	90.13
RU36	9.6645	16.0665	6.4020	10.3047	10.2998	0.6353	90.08
PU37	9.5505	15.7203	6.1698	10.1675	10.1700	0.6195	89.96
RU38	9.5316	15.7538	6.2222	10.1538	10.1403	0.6087	90.22
PU39	9.5326	15.6371	6.1045	10.1431	10.1406	0.6080	90.04
RU40	9.5400	15.7198	6.1798	10.1580	10.1621	0.6221	89.93
PU41	9.5855	15.7893	6.2038	10.2059	10.2089	0.6234	89.95
RU42	9.5287	15.6698	6.1411	10.1428	10.1443	0.6156	89.98