Practical Analysis of Flavor and Fragrance Materials

Practical Analysis of Flavor and Fragrance Materials

Edited by

Kevin Goodner and Russell Rouseff

A John Wiley & Sons, Ltd., Publication

This edition first published 2011

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

Practical analysis of flavor and fragrance materials / edited by Kevin Goodner, Russell Rouseff.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4051-3916-8 (cloth)
1. Flavor – Analysis. 2. Flavoring essences – Analysis. 3. Flavor – Biotechnology. I. Goodner, Kevin. II. Rouseff, Russell.
TP418.P73 2011
664′.07–dc23

2011013226

A catalogue record for this book is available from the British Library.

Print ISBN: 978-1-405-13916-8; ePDF ISBN: 978-1-444-34314-4;
oBook ISBN: 978-1-444-34313-7; ePub ISBN: 978-1-119-97521-2; mobi ISBN: 978-1-119-97522-9

Typeset in 10.5/13pt Sabon by Laserwords Private Limited, Chennai, India

Contents

Preface

PRACTICAL ANALYSIS OF FLAVOR AND FRAGRANCE MATERIALS

Flavor is one of the most important factors in consumer purchases and long term consumption. However, flavor is not easily quantified as the factors that impact flavor are almost always trace components. So from a chemical point of view, flavor analysis is essentially trace organic analysis. The human factor is essential to understanding flavor because humans have different genetic and cultural backgrounds which will alter their perception of flavor. Therefore all flavor analysis should be guided by human sensory panels. For too many years the study of flavor was conducted by analytical chemists who measured what they could measure using traditional analytical techniques rather than quantifying those trace impact compounds which should be measured. For many years the use of human assessors (sensory analysis) was conducted without interest in determining what was producing flavor changes in products being evaluated. Because sensory panels are impractical for routine quality control purposes, most food and fragrance manufacturers have chosen a middle ground where sensory panel data is used to guide chemists as to which compounds should be monitored to maintain quality or a specific sensory profile.

This book is an attempt to demonstrate how to develop this hybrid approach to flavor analysis. The few books that exist for flavor analysis have exclusively detailed either chemical analysis with sensory input or exclusively sensory analysis without regard to chemical composition.

This book is aimed at the practical side of analytical analyses. We attempt to produce a book as a reference book or as a primer for analytical chemists who are starting out in the flavor and fragrance industry with useful chapters on some of the major topics that someone new to the industry might encounter, including some of the basic tests one might see in the labs such as °Brix, water activity, turbidity, and similar tests.

David Rowe summarized much of the descriptive information from his recent book on *Chemistry and Technology of Flavour and Fragrance* into the first chapter. Sample preparation techniques are described by Russell Bazemore in the next chapter. It provides a detailed description of classic and cutting edge sampling techniques that ultimately determine the success of any flavor analysis. Traditional analytical techniques that have been used to measure the quality of raw flavor materials and finished products are presented next. Gas chromatography-mass spectrometry is included in this chapter as it is the most common technique employed by flavor chemists.

Gas chromatography-olfactometry, GC-O, is a hybrid technique employing the separation power of high resolution gas chromatography with the particular selectivity and sensitivity of human olfaction. This chapter written by Kanjana Mahattanatawee and Russell Rouseff, covers the hardware, software, and various techniques used for GC-O, along with selected applications, and benefits.

Vanessa Kinton wrote the next chapter on multivariate techniques which are commonly used for data analysis. This chapter describes the mathematical background and theory behind these techniques. The focus is to provide a basic understanding of the theory behind these mathematical approaches knowing that in practice the procedures are handled as a ''black box''. These techniques are used extensively in many areas of analysis (electronic nose, MS chemsensor, sensory analysis, etc.) and this chapter provides the basics while the other chapters provide the application examples.

Chapters 5 and 6, by Marion Bonnefille and Ray Marsili respectively, employ many of the multivariate data treatments for two very different sensor types. Chapter 5 concerns the metal oxide based electronic nose while chapter 6 is on the MS-based chemical sensor. Although both techniques employ pattern recognition software from instrumental sensors to mimic human olfaction, they differ profoundly in the types and number of sensors used to obtain the data arrays.

The chapter on sensory analysis by Carlos Margaria and Anne Plotto is likely to be an area in which most chemists have little familiarity. This chapter provides a wealth of practical information about conducting

sensory panels both trained and untrained with many anecdotes from their own experience.

The last chapter describes the ever changing regulations that affect flavor analysis in the industry and is written by Robert Kryger. This is an extremely important issue that is rarely taught in schools or universities. He discusses many of the basic terms and regulations as well as some of the complications in interpreting these regulations which vary from country to country.

The editors hope that this compilation will benefit those scientists beginning their careers in the area of flavor. Finally, and most importantly, we wish to thank each contributor for their time and efforts they put into their respective chapters. This book was a long time in the making and we are most appreciative of individual authors for their dedication and expertise in making this book possible.

About the Editors

Kevin L. Goodner received both a B.S.Ch. in Chemistry and a B.S. in Mathematics from the University of Memphis in 1992 and a Ph.d. in Analytical Chemistry from the University of Florida working with Fourier Transform Mass Spectrometry. His focus changed to flavor chemistry after a 1.5 year post-doctorial position at the University of Florida with Russell Rouseff. Kevin then worked for 9 years at the USDA Citrus and Subtropical Products Laboratory researching flavor and quality aspects of many products. In January of 2008, Kevin switched to industry at Sensus, LLC working on tea, coffee, and other products where he currently is the Director of Research and Development. Kevin has over 50 peer-reviewed publications.

Russell Rouseff is a professor at the University of Florida's Citrus Research and Education Center specializing in flavor and color chemistry. He has 35 years experience in the Florida citrus industry, first with the Department of Citrus and then the University. He has written or edited five books, 37 book chapters and over 108 referred journal articles. He has mentored scores of domestic and international students, 8 post docs and numerous visiting scientists. He has worked with aroma volatiles in fruits, coffee, wine, flowers and foliage and bitter non-volatiles. He is a Fellow of the American Chemical Society's Agricultural and Food Chemistry Division, recipient of the IFT's Citrus Products Division Research and Development Award and received the American Chemical Society's Award for the Advancement of Food and Agricultural chemistry in 2009. Hobbies include salt water reef aquariums, tennis and motor cycles.

List of Contributors

Russell Bazemore, Volatile Analysis Corporation, USA

Marion Bonnefille, Alpha MOS, Toulouse, France

Kevin Goodner, Sensus, LLC, Hamilton, USA

Vanessa Kinton, Alcohol and Tobacco Tax and Trade Bureau (TTB), Ammendale, USA

Robert A. Kryger, Citrus Resources LLC, Lakeland, USA

Kanjana Mahattanatawee, Department of Food Technology, Siam University, Thailand

Carlos Margaria, US Distilled Products, Princeton, USA

Ray Marsili, Marsili Consulting Group,Rockford, USA

Anne Plotto, USDA, ARS, Citrus and Subtropical Products Laboratory, Winterhaven, USA

Russell Rouseff, IFAS, Citrus Research and Education Center, University of Florida, USA

David Rowe, Riverside Aromatics Ltd , Poole, UK

1

Overview of Flavor and Fragrance Materials

David Rowe
Riverside Aromatics Ltd, Poole, UK

The nature of this chapter must be that of an overview as the alternative would be a multivolume series! The difficulty is not a shortage of material but rather a surfeit, and a second issue is how to give a rational coverage; should the materials be classified by chemistry, by odor or by application? The approach here is a combination of all three, and is based in part on a précis of *The Chemistry and Technology of Flavours and Fragrances* [1].

There is, of course, a massive overlap between flavor and fragrance; for example, *cis*-3-hexenol, discussed below, has a 'green', cut-grass odor, and hence contributes freshness to both flavors and fragrances. The division between the two Fs is itself not always a natural one!

1.1 FLAVOR AROMA CHEMICALS

1.1.1 Nature Identical

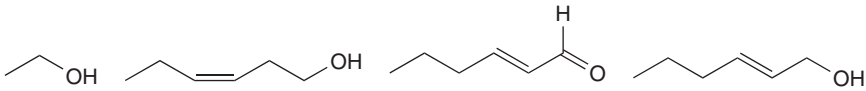
The vast majority of the aroma chemicals used in flavor are nature identical (NI), that is, they have been identified as occurring in foodstuffs in the

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

human food chain. This is a key method of identifying the most important components which create a flavor, and until recently, there were also regulatory implications. European Council Directive 88/388/EEC defined these as ''flavouring substances identical to natural substances'', with the alternative being ''artificial flavouring substances'', with the latter leading to the stigma of ''artificial flavors''. The newest regulations, REGULATION (EC) No 1334/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, no longer differentiates between Nature Identical and artificial, but the concept is still important – as a guide to flavorists, knowing a material is NI is important, and it can be especially so the context of ''from the named food'' type of flavours. Regulation 1334/2008 now only differentiates between ''flavouring substances'' and ''natural flavouring substances'', which harmonizes to an extent with the USA, where the NI classification has never been used. Even there, though, the NI concept has value, as materials have to be on the FEMA GRAS list, that is they are ''Generally Recognized As Safe'', and the vast majority of such substances are found in Nature.

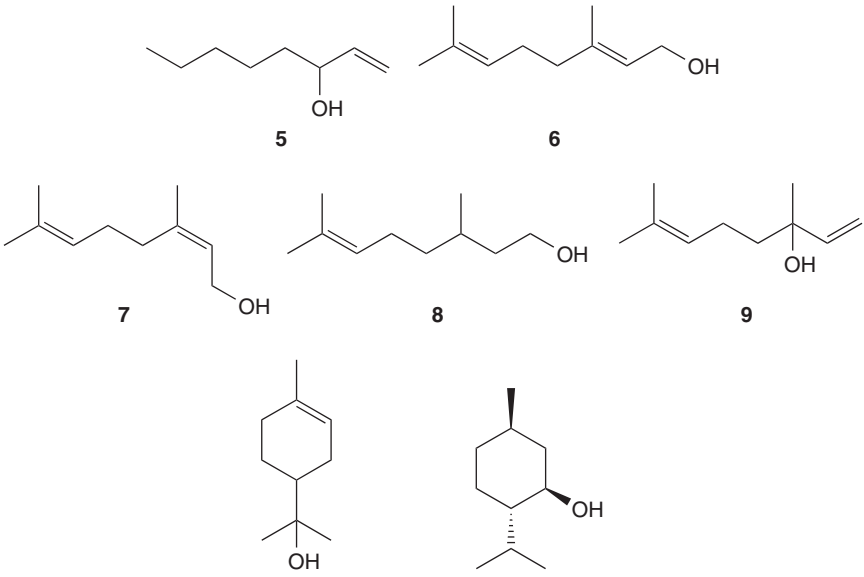
1.1.1.1 Alcohols

It should be noted that ethanol **1** itself is a flavor component of 'alcoholic drinks' as anyone tasting alcohol-free drinks will report! In fact it may considered as a solvent (especially in fragrances), as a flavour substance (FEMA 2419) or an additive (E1510)! *cis*-3-Hexenol **2**, mentioned above, is produced in nature as a 'wound chemical', that is, when plant tissue is damaged, ingressing oxygen is 'mopped up' by reaction with linoleic acid, which generates the unstable *cis*-3-hexenal, which is enzymatically reduced to the alcohol. Also formed are *trans*-2-hexenal **3**, which has a harsher, more acrid greenness and *trans*-2-hexenol **4**, which is rather sweeter:



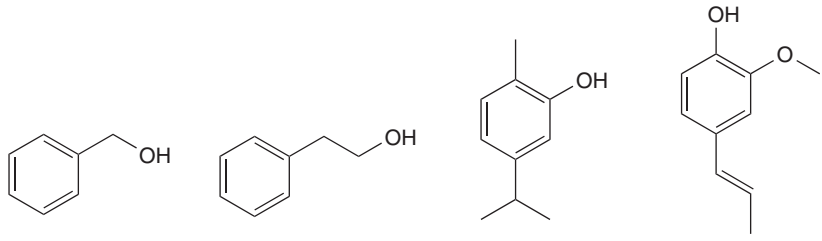
1 **2** **3** **4**

1-Octen-3-ol, 'mushroom alcohol' **5**, has the earthy note characteristic of mushrooms. The 'terpenoid' alcohols, C10 derivatives, include geraniol **6** and its isomer nerol **7**, citronellol **8** and linalool **9** [2]. Cyclic terpenoid alcohols include α-terpineol **10** and menthol **11**:



5 **6** **7** **8** **9** **10** **11**

Benzyl alcohol **12** has relatively little odor and is more commonly used as a solvent in flavors; phenethyl alcohol **13** is a component of rose oil and has a pleasant rose-like aroma. Two important phenols are thymol **14** and eugenol **15**, which are also major components of thyme and clove oils respectively:

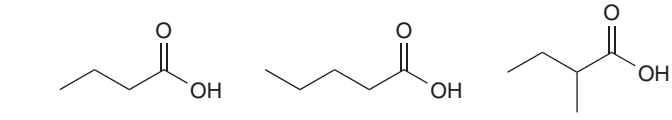


12 **13** **14** **15**

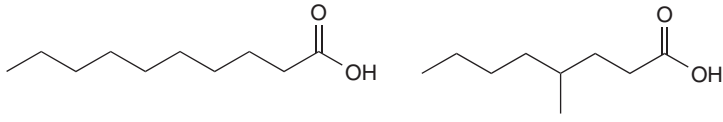
1.1.1.2 Acids

Simple acids contribute sharp notes which often become fruity on dilution. Butyric acid **16** is indisputably 'baby vomit' in high concentration; valeric acid **17** is cheesy, whereas 2-methylbutyric acid **18** is fruitier. Longer chain acids such as decanoic **19** are fatty and are important in

dairy flavors. 4-Methyloctanoic acid **20** has the sharp fatty character of roasted lamb:



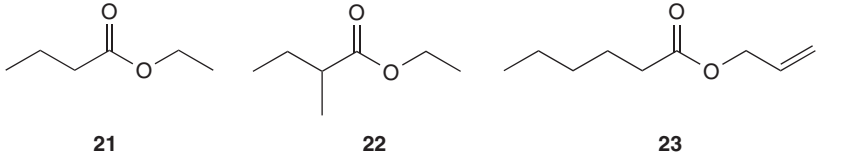
16 **17** **18**



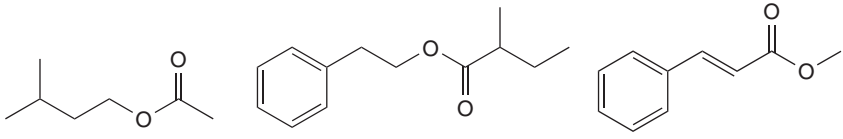
19 **20**

1.1.1.3 Esters

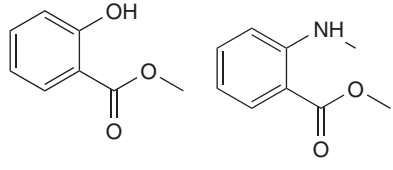
Numerous esters are used in flavors, so it is almost a case of any flavor alcohol combined with any flavor acid! Important simple esters include the fruity ethyl butyrate **21** and 2-methylbutyrate **22**; allyl hexanoate **23** has a familiar pineapple aroma and isoamyl acetate **24** is 'pear drops'. Phenethyl 2-methylbutyrate **25** is 'rose bud ester' and the warm sweet aroma of methyl cinnamate **26** makes it valuable in strawberry flavors. Methyl salicylate is the main component of wintergreen oil **27** and methyl N-methylanthranilate **28** is found in mandarin, which differentiates this from the other citrus oils:



21 **22** **23**



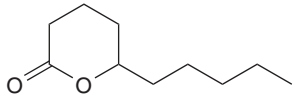
24 **25** **26**



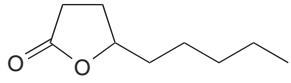
27 **28**

1.1.1.4 Lactones

These cyclic esters are usually found as gamma-lactones (five-membered rings) and delta-lactones (six-membered). Like their acyclic cousins they are used in fruit flavors and also for dairy, especially the delta-lactones such as delta-decalactone **29**. Gamma-nonalactone **30**, also misleadingly known as Aldehyde C18, has a powerful coconut odor:



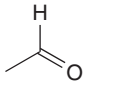
29



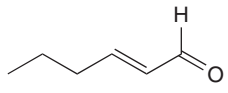
30

1.1.1.5 Aldehydes

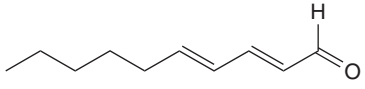
Acetaldehyde **31** is ubiquitous in fruit aromas, though its volatility (b.p. 19 °C) makes it difficult and dangerous to handle as a pure aroma chemical. Unsaturated aldehydes such as the previously mentioned *trans*-2-hexenal (leaf aldehyde) **3** are very important. *trans*-2-*trans*-4-Decadienal **32** is intensely 'fatty-citrus'; *trans*-2-*cis*-6-nonadienal **33** is 'violet leaf aldehyde'. 'Citral', a mixture of the isomers geranial **34** and neral **35**, is intensely lemon; it is a key flavor component of lemon and to a lesser extent other citrus oils:



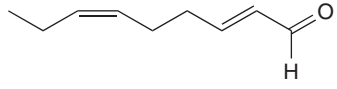
31



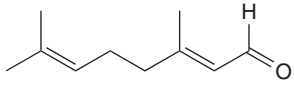
3



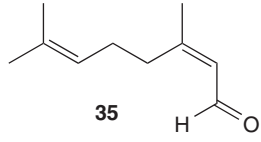
32



33



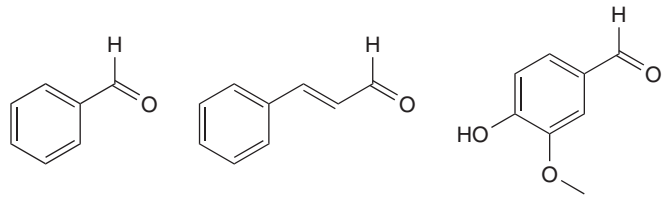
34



35

Benzaldehyde **36** is widely used in fruit flavors, especially for cherry, though in fact it is not a key component of cherries. Cinnamaldehyde **37** is found in cassia and cinnamon oils. The most important aromatic

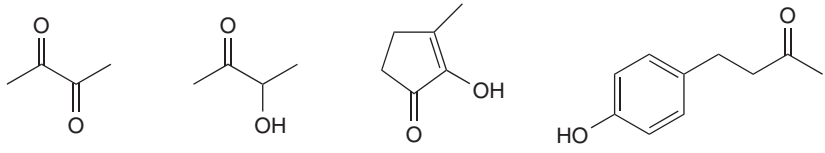
aldehyde, and one of the most significant of all aroma chemicals, is vanillin **38**:



36 **37** **38**

1.1.1.6 Ketones

The C4 ketones diacetyl **39** and acetoin **40** are used in butter-type flavors for margarines and other dairy products and hence are used in very large quantities. The former is very volatile and is believed to have led to respiratory damage amongst people exposed to large quantities of its vapor. The cyclic diketone 'maple lactone' **41** occurs as the enolic methylcyclopentenolone (MCP) and has the characteristic sweet, caramel odour of maple syrup. Raspberry ketone **42** is unusual in the bizarre world of flavor and fragrance trade names in that it is actually found in raspberries, tastes of raspberries and is a ketone!



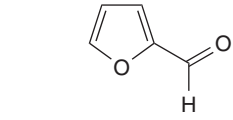
39 **40** **41** **42**

1.1.2 Heterocycles [3]

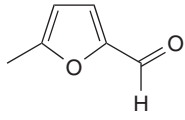
1.1.2.1 Oxygen-containing

The importance of materials containing the five-membered furan ring cannot be overstated [4]. Furfural **43** is formed by the Maillard reaction from pentoses in the cooking process, and 5-methylfurfural **44** from hexoses similarly. The latter has an almond, 'marzipan' aroma similar to benzaldehyde but with more naturalistic character. Methyl tetrahydrofuranone, 'coffee furanone' **45**, is sweet and caramelic, but the most important flavor furan must be 2,5-dimethyl-4-hydroxy-3[2H]-furanone

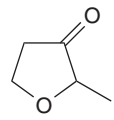
46, an aroma chemical of many names, including strawberry furanone, and pineapple ketone. This has sweet, fruit and caramel notes, making it of obvious importance in fruit flavors, but it is also important in meat flavors, where it seems to function as a flavor enhancer. Its homologue Soy Furanone **47** is also very sweet, whereas its isomer 'Sotolone', or fenugreek lactone **48**, has an intense fenugreek tonality, becoming more caramel-like in high dilution. The saturated furan Theaspiran **49** is found in black tea and a number of fruits:



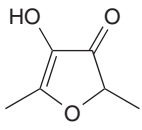
43



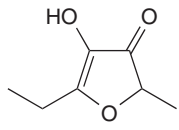
44



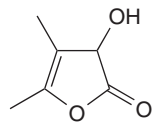
45



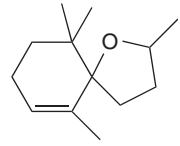
46



47

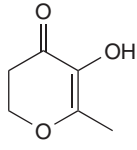


48

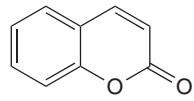


49

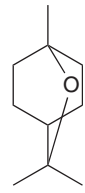
The most important pyrans must be Maltol **50** and Coumarin **51**. The former is another caramel compound, with the latter having sweet and spicy notes. The saturated furan 1,8-cineole, or eucalyptol **52**, is the main component of eucalyptus oil as well as being widespread in other oils such as lavender, distilled lime and rosemary:



50



51

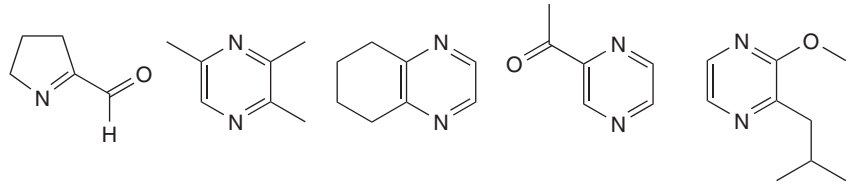


52

1.1.2.2 Nitrogen-containing

The pyrrole group is relatively unimportant in flavors, though mention should be made of 2-acetyldihydropyrrole **53**, which has the 'Holy Grail'

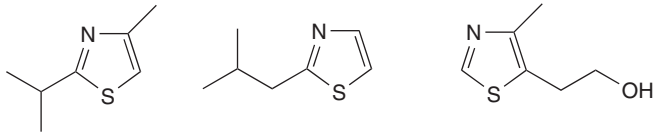
aroma of freshly baked bread but is too unstable for commercial use. The most important nitrogenous heterocycles are pyrazines, which are readily formed in the Maillard reaction from amino acids and sugars; simple alkyl pyrazines such as 2,3,5-trimethylpyrazine **54** have roasted, cocoa-like notes making them important for chocolate and roasted notes. Tetrahydroquinoxaline **55** has particularly noticeable roasted notes. 2-Acetylpyrazine **56** has very pervasive roasted, biscuit notes. The alkoxyalkylpyrazines are also found in fresh fruits and vegetables, the intensely odorous 2-methoxy-3-isobutylpyrazine **57** often being known as 'Bell Pepper Pyrazine'.



53 **54** **55** **56** **57**

1.1.2.3 Sulfur-containing

Whilst a few simple thiophenes are used, the most important sulfur heterocycles are thiazoles, especially 2-isopropyl-4-methylthiazole **58** and 2-isobutylthiazole **59**, which have peach/tropical and tomato vine character respectively. 4-Methyl-5-thiazoleethanol, Sulfurol **60**, is widely used in dairy and savory flavors.



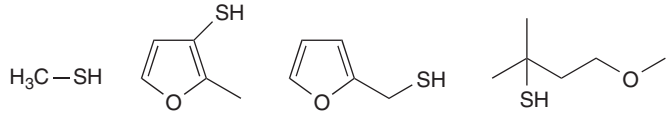
58 **59** **60**

1.1.3 Sulfur Compounds [5]

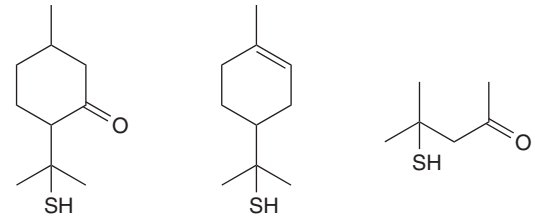
The importance of sulfur compounds reflects their highly odorous character; the most odorous compounds known are sulfur compounds, with odor thresholds down to the 10^{-4} parts per billion level. They are the single largest group of 'High Impact Aroma Chemicals', materials which provide 'character impact' even at very low levels [6].

1.1.3.1 Mercaptans

These are generally the most odorous of the most odorous, as it were, the *capo di capi* of the flavor industry. Methyl mercaptan **61** is widespread in meat aromas, as is 2-methyl-3-furanthiol (MFT) **62**; the latter is especially important in beef. Furfuryl mercaptan **63** is a character impact aroma chemical of roasted coffee. The latter two are Maillard reaction products formed from cysteine and pentoses. 'Fruity' mercaptans include the blackcurrant/cassis materials **64** and thiomenthone **65**, and p-menthene-8-thiol, the Grapefruit Mercaptan **66**. The accurately named Cat Ketone, 4-mercapto-4-methyl-2-pentanone **67**, is also found in grapefruit and wines.



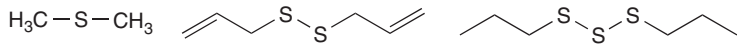
61 **62** **63** **64**



65 **66** **67**

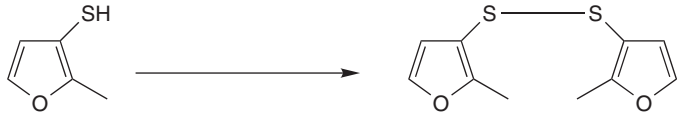
1.1.3.2 Sulfides

The simplest sulfide, dimethyl sulfide (DMS) **68** has a vegetable, sweetcorn odor; sulfides are less odorous than mercaptans, and hence a key aspect of quality is the need to remove all traces of mercaptans; impure DMS is quite repellent. Propyl and allyl sulfides are perhaps the commonest, especially as di- and higher sulfides; allyl disulfide **69** is the major component of garlic oil, with the remainder being mostly higher sulfides. Propyl compounds such as dipropyl trisulfide **70** are found in onion; ethyl compounds are found in Durian fruit, and to human noses other than those raised with the fruit, are at best unpleasant and sewer-like:



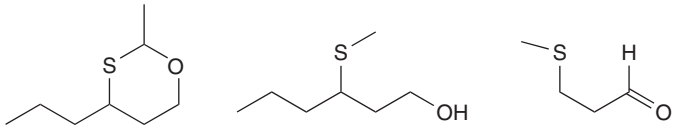
68 **69** **70**

Some mercaptans oxidize very easily to form disulfides, such as the formation of bis(2-methyl-3-furyl) disulfide **71** from MFT **62**:



62 **71**

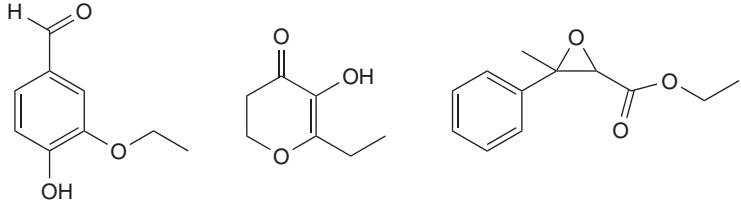
There are a number of fruity sulfides, often derived in some way from C6 units with an oxygen atom in the 3-position relative to the sulfur; such a grouping is found in the 'tropicals' Tropathiane **72** and 3-methylthiohexanol **73** as well as the potato-like methional **74**:



72 **73** **74**

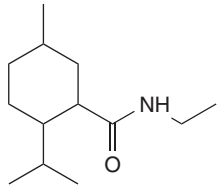
1.2 FLAVOR SYNTHETICS

There are still a number of important flavor materials which have, to date, not been found in nature. They are often used because they have properties which suitable NI materials lack; for example the so-called Ethyl vanillin **75** has a lower odor threshold than vanillin and is more soluble in organic solvents, making it more suitable for use in oil-based flavors, and 'Ethyl maltol' **76** is more powerful than maltol. Several glycidate esters are used, such as ethyl 3-methyl-3-phenylglycidate **77**, so-called Aldehyde C16, which has a powerful strawberry aroma and is used in flavours as well as fragrances.

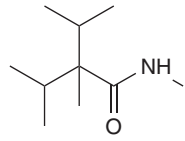


75 **76** **77**

Synthetics have proved especially valuable in the area of what might be termed 'sensates', the molecules of taste and sensation [7]. For example, the carboxamides **78** and **79** are both cooling agents which are longer-lasting than menthol **81**:



78



79

Over the years a number of 'synthetics' have been found in nature, changing their status, such as the previously mentioned cat ketone and allyl hexanoate, and other materials are surely 'synthetic' simply because they are still hiding in foodstuffs! In addition, the "new" European regulations, EC 1334/2008, removes the "artificial" classification, and gives, potentially, a new lease of life to these materials.

1.3 NATURAL AROMA CHEMICALS [8]

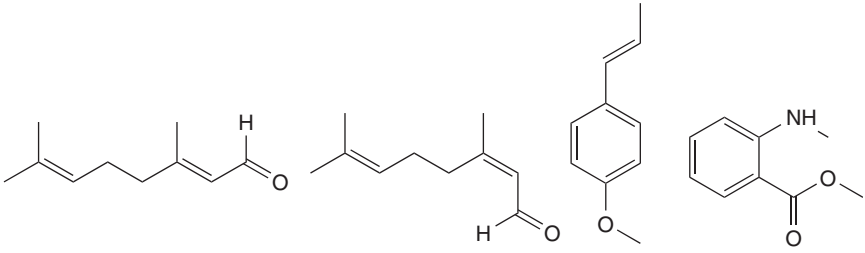
The seemingly innocent term 'natural' is, in fact, a more troublesome one than it seems. In essence 'natural' materials are those which are:

(a) obtained by physical means from materials in the human food chain, that is, isolates;
(b) obtained by biological conversions of natural materials, that is, biotechnology; or
(c) obtained by reacting natural materials together in the absence of chemical reagents or catalysts, that is, cooking chemistry or soft chemistry.

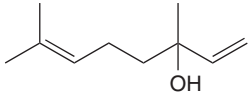
These definitions are enshrined in US (CFR 21, 101.22 (a) (3)) and European (REGULATION (EC) No 1334/2008) regulations. As far as aroma chemicals are concerning, "Natural" is a marketing conceit; the marketing departments of flavour and food companies, the supermarkets and other major retailers are unlikely to reverse their policies of promoting their subliminal (and sometimes not so subliminal) formula of Natural = Healthy, especially with the so-called "Clean Label" concept. The importance of the regulations is that they set the criteria which enable a material to be called "Natural".

1.3.1 Isolates

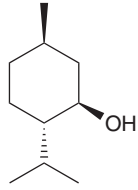
A number of essential oils consisting of high levels of valuable components, make their direct isolation by physical means commercially viable. Examples include citral **34, 35** from litsea cubeba oil, anethole **80** from star anise oil, methyl N-methyl anthranilate **28** from mandarin petitgrain oil, linalool **9** from ho wood oil, and L-menthol **81** from mint oils. In the latter case, isolation is by cooling the crude mint oil to deposit the familiar large 'bright crystals' of commerce:



34 **35** **80** **28**

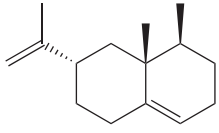


9

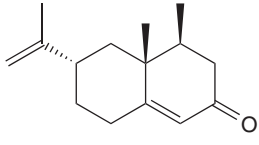


81

If the value of the end product is sufficiently high and the raw material is cheap enough, then isolation is viable even if the component is at a low level, for example valencene **82** from orange oils and nootketone **83** from grapefruit oil:



82

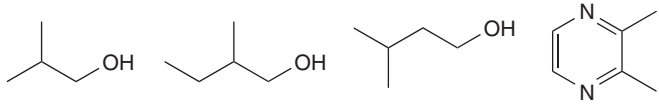


83

1.3.2 Biotechnology

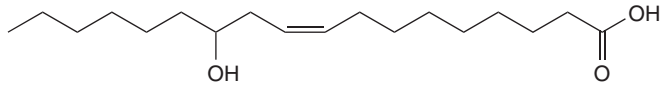
This very modern term actually covers one of our species' oldest hobbies – brewing! Alcoholic fermentation, as well as forming ethanol,

produces other alcohols such as isobutyl **84** and isoamyl **85, 86.** The latter is the main component of fusel oil, the residue after the distillation of liquors such as brandy, which can also be a source of pyrazines such as **87.**

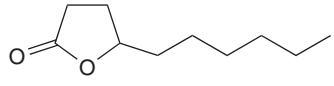


84 **85** **86** **87**

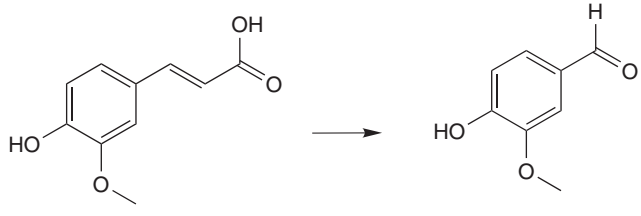
Acetobacter species, unwanted contaminents in brewing, produce carboxylic acids. Gamma-lactones such as γ-decalactone **88** can be produced from ricinoleic acid **89,** a component of castor oil, and vanillin **38** can be obtained from ferulic acid **90,** a by-product of cereal production:



89



88



90 **38**

1.3.3 'Soft Chemistry'

This is best illustrated by the formation of esters by heating an alcohol with an acid; if the alcohol has a high boiling point this often takes place rapidly even though no catalyst is permitted. Another important

example is the synthesis of 2,5-dimethyl-4-hydroxy-3[2H]-furanone **46** from the hexose rhamnose **91**:

91 **46**

This is the area of greatest contention in the natural aroma chemicals arena. EC 1334/2008 defines the techniques that may be used as "listing of traditional food preparation processes in Annex II, and does not involve, inter alia, the use of singlet oxygen, ozone, inorganic catalysts, metal catalysts, organometallic reagents and/or UV radiation." but it is often easier to define what cannot be done than what can be: for example, when is a solvent also a reagent? All solvents interact with their solutes – they would not act a solvents otherwise!

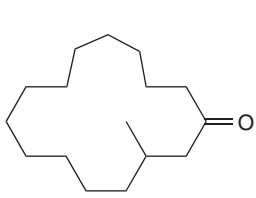
1.4 FRAGRANCE AROMA CHEMICALS [9–11]

As noted above, there are many aroma chemicals whose use overlaps both flavor and fragrance; many esters, aldehydes, heterocycles and indeed anything other than the most savory of flavor chemicals, tend to have uses in the fragrance sector. However, the freedom to move away from naturally occurring materials opens up a range of what we might call designer synthetics for the key notes of fragrance.

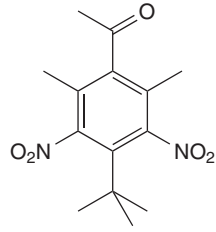
1.4.1 Musks [12]

The main odiferous component of natural musk is the macrocycle Muscone **92**. The scarcity of the natural material and the difficulty of synthesizing large carbocycles has driven chemists to develop synthetics since the late nineteenth century. The first artificial musks were the nitro musks such as Musk Ketone **93**, discovered serendipitously during explosives research. Discoloration problems and toxicity issues have restricted the use of the nitromusks and from the 1950s the so-called polycyclic musks were introduced. These have the advantages of stability, especially in household use, as well as ease of synthesis. Galaxolide **94** is the most important of these, with other related materials including Tonalid (Fixolide) **95** and Celestolide **96**. The stability and hydrophobicity of

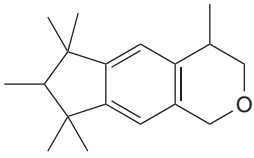
these materials has led to high usage in fabric softeners and detergents as well as fine fragrances.



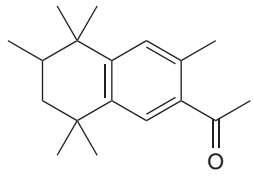
92



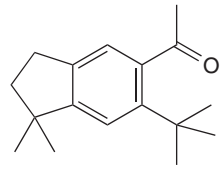
93



94

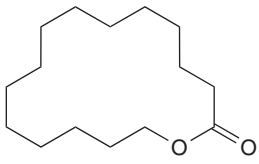


95

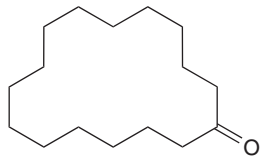


96

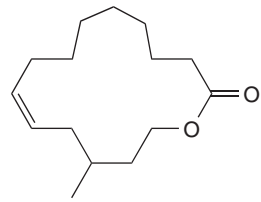
Advances in synthetic methods have increased the availability of macrocyclic musks. The lactone ring was relatively easy to form, and 15-pentadecanolide **97** has been sold under many names for many years; the corresponding ketone **98** is now available. Modern musk structures such as Nirvanolide **99** combine natural musk character with low odor thresholds:



97



98

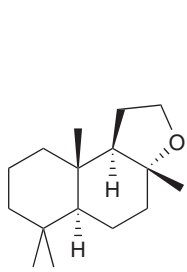


99

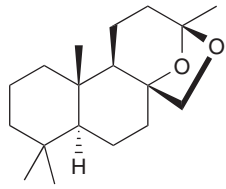
1.4.2 Amber

'Amber' materials are so-called due to their resemblance to *Ambergris*, a material formed in the stomachs of whales, probably as a pathological response to damage by shelly parts of plankton. This was formerly available as a by-product of the 'whaling industry', but now is only

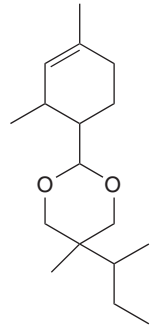
found occasionally on beaches (a 15 kg piece was found on a beach in Australia in 2006), making it a rare and expensive material. Ambergris consists mostly of steroidal materials, and this structure forms the basis of classic ambers such as Ambroxan **100** and Amberketal **101**. More recently amber molecules lacking this steroidal structure have been produced, such as Karanal **102** and Spirambrene **103**:



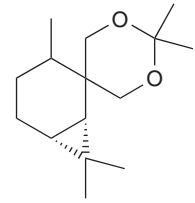
100



101



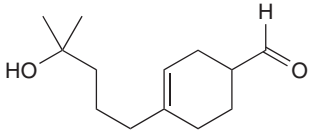
102



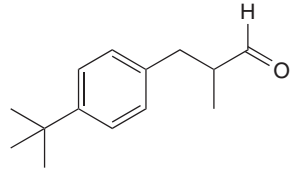
103

1.4.3 Florals

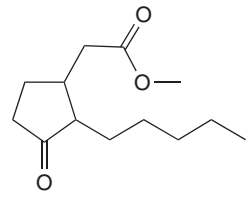
Lily of the valley, or 'muguet', materials include the aldehydes Lyral **104** and Lilial **105**. Methyl dihydrojasmonate **106** is a powerful jasmine molecule, first used in Dior's famous 'Eau Sauvage'. For 'cheap and cheerful' jasmine-type notes for household fragrances simpler materials such as α-hexylcinnamaldehyde **107** can be used:



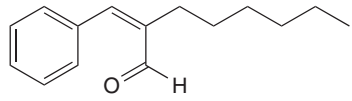
104



105

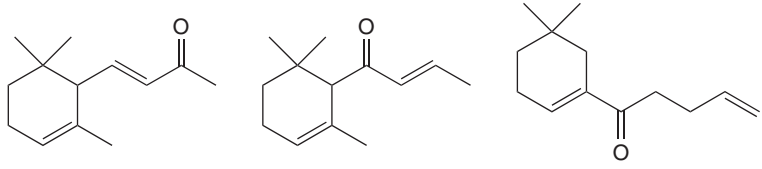


106



107

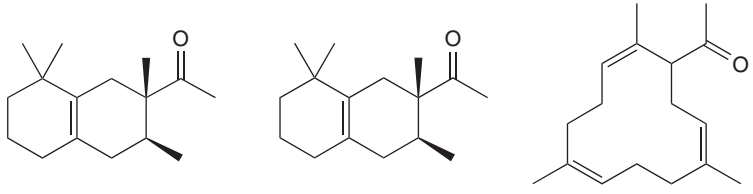
The ionones, such as α-ionone **108**, were amongst the first synthetics to be used in perfumery; these violet-type materials have since been found in nature, for example in raspberry, and are also used in flavors. The damascones, such as α-damascone **109**, were identified in the oil from rosa damascena; the synthetic analogue Dynascone **110** is used in the popular perfume 'Cool Water' by Davidoff:



108 **109** **110**

1.4.4 'Woodies'

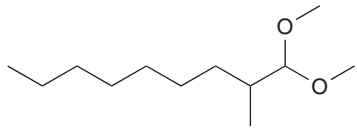
Materials associated with this note include Iso E Super **111** and Georgywood **112**; the macrocyclic ketone Trimofix **113** combines woody and amber notes:



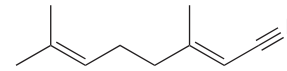
111 **112** **113**

1.4.5 Acetals and Nitriles

Another advantage in the use of synthetics is that stability issues can be addressed; aldehydes are prone to oxidation and Aldol-type condensation reactions, especially under the harsh conditions required for household fragrances. Acetals, such as 2-methylundecanal dimethyl acetal (Aldehyde C12 MNA DMA!) **114** and nitriles such as geranonitrile **115** have similar character to their 'parent' aldehydes but are much less prone to these damaging reactions:



114



115

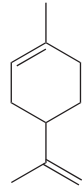
1.5 MATERIALS OF NATURAL ORIGIN

1.5.1 Essential Oils

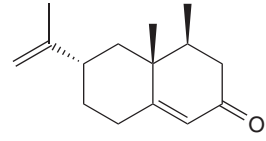
Naturally occurring materials were, of course, the foundation of both the flavor and fragrance industry and organic chemistry. The so-called essential oils are those materials obtained from plant materials by simple physical means, especially cold pressing and steam distillation.

1.5.1.1 Cold-pressing – Citrus Oils

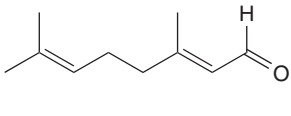
The peels of the citrus family – which includes orange, lemon, lime, bergamot, grapefruit, tangerine, and mandarin – contain glands which release oils when crushed. Cold-pressing the peel gives a mixture of water and oil, which is simply separated. The main component of all these oils is the monoterpene hydrocarbon limonene **116**, typically about 95 % in orange and grapefruit, slightly lower in lemon and lime. The character of the oil is determined by 'trace' components, which are also considered to be 'markers' for the quality of the oil, for example Nootketone **83** in grapefruit (0.1–0.4 %), citral **34, 35** in lemon (1–2 %), methyl N-methyl anthranilate **28** (0.3–0.6 %) in mandarin; it should be emphasized, however, that the quality of the oils is the sum of their parts, not a marker plus carrier!



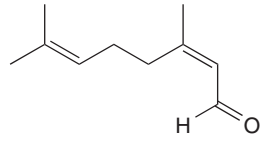
116



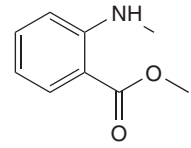
83



34

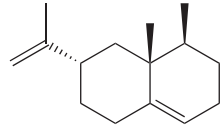


35



28

Citrus essence oils, or phase oils, are by-products of the juice market; concentration of the juice to reduce transport costs leads to oil separating out which can be sold or further processed. Whilst the essence oils are similar in components to their 'parent', there are some differences in that the essence oil is nearer to being an extract of the juice. The most important is orange essence oil; this contains more volatiles such as acetaldehyde and ethyl butyrate and a higher level of the sesquiterpene valencene **82** than cold-pressed orange oil.



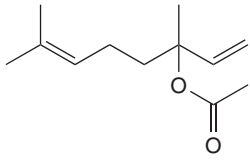
82

The high levels of hydrocarbons in the citrus oils leads to poor solubility in water, a particular problem for their use in soft drinks. This can be overcome, in part, by folding; the oil is 'folded' by distilling off the more volatile monoterpene hydrocarbons, in effect concentrating the more valuable flavor components and removing the most hydrophobic components. The distillates also have value as solvents, diluents and, since the terpenes do carry over some odiferous components, as flavor or fragrance ingredients in their own right; for example in creating a tenfold orange oil, 9 kg of orange terpenes are produced per kg of 'Orange Oil 10X'.

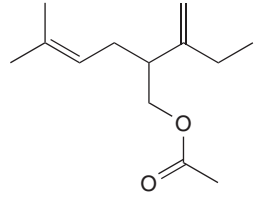
1.5.1.2 Steam-distilled Oils

Most plant materials contain much less volatile oil than the citrus fruits, often less than 1 %, and these are simply steam distilled to obtain the oil; examples include cassia, cinnamon, mint, rose and lavender. This process

can, of course, lead to changes in composition due to thermal decomposition, oxidation and hydrolysis; by tradition, the best lavender oil is produced at high altitude, as the lower boiling point of water leads to less hydrolysis of esters such as linalyl **117** and lavandulyl **118** acetates.

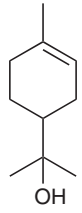


117

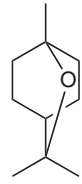


118

Chemical changes can also be beneficial. The most important lime oil is "Distilled Lime"; this is not redistilled cold-pressed oil but oil which is steam distilled from the macerated fruit. The high acidity of the juice leads to hydration of the hydrocarbons, giving high levels of α-terpineol **10** and 1,8-cineole **52**. This oil has a fresh, juicy aroma which contrasts with the waxy, floral odour of the cold-pressed oil.



10



52

1.5.1.3 A Note on 'Adulteration'

Historically, especially prior to the routine availability of gas chromatography, the addition of cheap materials to 'cut' or adulterate oils, was more widespread. It is still a problem; without the use of sophisticated isotopic analysis, the addition of a quantity of a synthetic material, for example, cinnamaldehdye to cinnamon oil, is impossible to detect, let alone prove. Similarly, the addition of a low percentage of terpenes to a citrus oil which is already over 90 % terpenes, is not going to be detected easily. Ultimately the answer has to lie in the trust in the supplier, as applied with a modicum of common sense. Customers are themselves

part of the problem – the desire for lower priced raw materials is itself a driving force in adulteration. Put simply, if you wish to buy an oil at 50 % of the market price, then do not be surprised if the product you are buying is only 50 % oil!

1.5.2 Absolutes and Other Extracts

This final section in fact covers the oldest approach to flavor and fragrance materials – the extraction of aromatic materials into organic solvents. Extraction of plant material with a nonpolar solvent such as hexane, followed by removal of the solvent, yields a concrete. As the name implies, this is often solid or semisolid due to the presence of plant waxes as well as pigments and nonvolatiles. Extraction into ethanol followed by solvent removal gives an absolute, which is more manageable and the more usual item of commerce. This is a very labour-intensive process and hence absolutes are more expensive than most oils, and are more associated with higher-value materials such as violet flower absolute and orange flower absolute. Oleoresins are more associated with spice oils such as ginger and garlic and are a solution of an absolute-type extract in an essential or solvent such as vegetable oil or propylene glycol.

ACKNOWLEDGMENTS

Many thanks to the contributors to *The Chemistry and Technology of Flavours and Fragrances*, whose work I have cheerfully plundered for this chapter!

REFERENCES

1. Rowe, D. J. (ed.) (2004) *The Chemistry and Technology of Flavours and Fragrances*, Blackwell, Oxford, UK.
2. Teisseire, P. J. (1994) *Chemistry of Fragrant Substances*, VCH, New York, USA.
3. Zviely, M. (2004) Aroma chemicals II: heterocycles, in *The Chemistry and Technology of Flavours and Fragrances*, (ed. D. J. Rowe), Blackwell, Oxford, UK, pp. 85–115.
4. Rowe, D. (2004) Fun with Furans, *Chemistry and Biodiversity*, **1**, 2034–2041.
5. Jameson, S. B. (2004) Aroma chemicals III: sulfur compounds, in *The Chemistry and Technology of Flavours and Fragrances*, (ed. D. J. Rowe), Blackwell, Oxford, UK, pp. 116–142.

6. Rowe, D. J. (2002) High impact aroma chemicals, in *Advances in Flavours and Fragrances: From the Sensation to the Synthesis* (ed. K. Swift), Royal Society of Chemistry, Cambridge, UK, pp. 202–236.
7. Dewis, M. L. (2004) Molecules of taste and sensation, in *The Chemistry and Technology of Flavours and Fragrances*, (ed. D. J. Rowe), Blackwell, Oxford, UK, pp. 199–243.
8. Margetts, J. (2004) Aroma chemicals V: natural aroma chemicals, in *The Chemistry and Technology of Flavours and Fragrances*, (ed. D. J. Rowe), Blackwell, Oxford, UK, pp. 169–198.
9. Frater, G., Bajgrowicz, J. A. and Kraft, P. (1998) Fragrance chemistry. *Tetrahedron*, **54**, 7633–7703.
10. Kraft, P., Bajgrowicz, J. A., Denis, C. and Frater, G. (2000) *Angew. Chem. Int. Ed.*, **39**, 2980–3010.
11. Sell, C. (1999) Ingredients for the modern perfume industry, in *The Chemistry of Fragrances* (eds D. Pybus and C. Sell), Royal Society of Chemistry, Cambridge, UK, pp. 51–124.
12. Kraft, P. (2004) Aroma chemicals IV: musks, in *The Chemistry and Technology of Flavours and Fragrances*, (ed. D. J. Rowe), Blackwell, Oxford, UK, pp. 143–168.

2

Sample Preparation

Russell Bazemore
Volatile Analysis Corporation, USA

2.1 INTRODUCTION

While injecting samples 'neat' (i.e., directly into the GC with no sample preparation) is a common practice for liquid samples in the flavor and fragrance industry, the practice is straightforward and requires little to no explanation or discussion. Therefore, preparation of samples by extracting volatiles for subsequent introduction into a gas chromatograph is the focus of this chapter. New or relatively new extraction techniques including solid phase microextraction (SPME), stir bar sorptive extraction (SBSE), polydimethylsiloxane (PDMS) foam, and Microvial are discussed in addition to tried and true technologies including static/dynamic headspace, and the solvent extraction methods MIXXOR, soxhlet, and Solvent Assisted Flavor Evaporation (SAFE).

The use of sample prep techniques that are rapid, precise, and may be automated have become popular and necessary in industry and academia due to an increasing number of projects managed and sample analyses required. A large portion of this review is devoted to products that incorporate polydimethylsiloxane (PDMS), a polymeric material proven to be useful for volatile extraction and fast, relatively simple

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

procedural techniques that lend themselves to automation for multi-sample analysis. A few basic points regarding PDMS are helpful in understanding its unique characteristics and capabilities.

2.2 PDMS

Siloxanes contain a Si-O-Si backbone, and those with organic groups attached to silica are called polyorganosiloxanes with the structure indicated in Figure 2.1. PDMS is a versatile substance. It is a primary component in Silly Putty, a substance well known to many baby-boomers as a malleable childhood toy. It is also a component in silicone grease and lubricants, defoaming agents, cosmetics, hair conditioner and a filler fluid in breast implants [1]. PDMS exhibits unique flow (rheological) properties. If placed on a given surface for several hours it will flow to cover the surface and will also cover any imperfections on the surface [2]. Polymer length and/or branches or cross links dictate viscoelasticity but in general at high temperatures PDMS resembles a very viscous liquid and at low temperatures it resembles an elastic solid. It has been used extensively as a stationary phase in gas chromatography (GC), and can be used over a broad temperature range (−20 °C to 320 °C, [3]). It is considered to be inert, nontoxic and nonflammable.

An important characteristic of PDMS is that it is hydrophobic. It does not bind water appreciably while it does extract other volatile components present in a sample matrix (immersed in a liquid or from headspace) by absorption into the polymer liquid phase. It does not require the use of solvents. These are the principal reasons why it has become popular for extracting volatiles and semivolatiles from foods, beverages, and biological materials.

The octanol–water partition coefficient (K_{O-W}) is the ratio of a compound's concentration in octanol and concentration in water at equilibrium at a specified temperature. The logarithmic ratio of the concentrations of solute in solvent is the definition of log *P*. PDMS extraction capacities can be predicted based on octanol-water partition

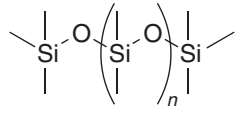


Figure 2.1 Structure of PDMS.

coefficients. Generally lipophillic substances have larger K_{O-W} values and are more readily absorbed by PDMS [4, 5].

When PDMS is exposed to organic solvents, especially pentane and xylenes, they diffuse into the polymer causing it to swell [2]. For this reason manufacturers of the two most utilized PDMS products in sample preparation chemistry, SPME (Supelco, Bellefonte, PA, USA) and Twister (Gerstel GmbH, Mülheim an der Ruhr, Germany) recommend aqueous extractions when directly immersing the SPME fiber or Twister. When conducting headspace analysis of most organic solvents, swelling still occurs and chromatograms display large solvent peaks, thus undermining a main advantage of PDMS extraction.

2.3 STATIC HEADSPACE EXTRACTION

Headspace refers to the gas phase located above a liquid and solid phase present in a sealed vessel. Volatility, or rate of evaporation of a given component, is governed by both Henry's law constant (gas to liquid distribution constant) and vapor pressure. Volatile compounds partition into the gas phase due to a number of factors, all interrelated in their effects, including propensity for solubility in water (hydrophilic or hydrophobic), polarity, ionic nature of analyte and solvent, molecular weight, and temperature. In static headspace extraction, the liquid and/or solid sample is placed in a sealed vial with an inert septum, usually Teflon (inert to prevent volatiles from sticking to the surface via *adsorption*, or being *absorbed* into the septum material). Volatile components are allowed to reach equilibrium between liquid/solid and gas phases. An aliquot of headspace is removed, normally by gas-tight syringe, and injected directly into a gas chromatograph injection port. A good rule of thumb to optimize the rate and concentration of compounds in headspace is to fill the vial to two-thirds full with sample thus leaving a third for headspace. Too large a volume of headspace requires more time to reach equilibrium conditions. Too small a volume and there may be insufficient to extract for analysis and wasted sample.

2.3.1 Advantages and Disadvantages

Static headspace extraction is simple, inexpensive, and lends itself to automation. It provides a true representation of volatile compounds responsible for aroma because it reflects natural headspace concentrations. However, this method is notorious for poor precision due to

nonequilibrium conditions during extraction or selective changes in volatile concentrations associated with serial extractions. It is also difficult to detect potentially important components because the volatiles are not concentrated. Only the volatiles present in greatest quantities will be detected by direct headspace sampling [5].

2.4 DYNAMIC HEADSPACE EXTRACTION

Dynamic extraction, also called purge and trap, differs from static extraction in that volatile compounds are continuously swept from the headspace into a trap by flow of inert carrier gas, nitrogen or helium (and sometimes air with the risk of oxidation). Figure 2.2 is a diagram of a typical purge and trap apparatus. Once trapping is complete the volatiles are released for chromatographic analysis. Rapid heating is the most efficient method for release of components from a trap. Inert gas then carries the desorbed volatiles onto the GC column to form a tight band for optimal chromatography. Other methods release volatiles by solvent desorption followed by solvent evaporation under a stream of inert gas, and then injection into a gas chromatograph.

Traps may contain one or combinations of substances. Included in the list is activated carbon, and modified carbon products with uniform pore sizes. Tenax (2,6-diphenylene-oxide polymer) is commonly used in this application due to its stability when heated (upper limit of 350 °C) and hydrophobic properties. PDMS foam is a newer product suitable for purge and trap and it will be discussed further in Section 2.6.

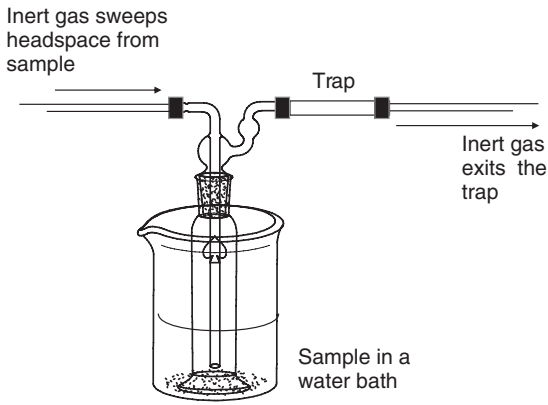


Figure 2.2 Diagram of a typical dynamic headspace extraction or purge and trap apparatus.

Cryocooling with liquid nitrogen (b.p. −196 °C) to condense volatiles onto glass wool and/or adsorbent material enhances extraction capacity. Following completion of extraction, traps are rapidly heated to desorb volatiles, as previously described.

2.4.1 Advantages

Purge and trap allows for concentrating samples and therefore this method has lower limits of detection. Solid samples may also be analyzed because the confounding equilibration variable associated with solid samples is not a concern. There is a continuous introduction and departure of carrier gas into and out of a vessel. Cryocooling increases the number of highly volatile, lower molecular weight compounds extracted.

2.4.2 Disadvantages

Disadvantages of purge and trap include Tenax's low surface area and a low adsorption capacity. The method has a propensity for extracting nonpolar components and a low affinity for polar compounds. Reineccius [5] provides a good discussion of Tenax's adsorption capabilities and shortcomings. Due to ice buildup, aqueous samples are problematic for cryocooled traps. For this reason a means for removing water vapor from the carrier gas stream must be introduced prior to the cold trap.

2.5 SOLID PHASE MICROEXTRACTION (SPME)

Arthur and Pawliszyn [6] first utilized PDMS coated on a fused silica fiber to extract analytes from an aqueous media in a process called solid phase microextraction. Diagrams of manual and auto sampler SPME fiber assemblies are indicated in Figure 2.3 and the procedure is indicated in Figure 2.4. This method was commercialized and made available by Supelco Corp. (Bellefonte, PA, USA).

2.5.1 Research

Since Pawliszyn's initial work he and many other researchers have successfully utilized SPME to investigate a plethora of topics associated with volatile and semivolatile components. Among these are Marsili [7, 8] who, in addition to a broad range of flavor work in foods and beverages,

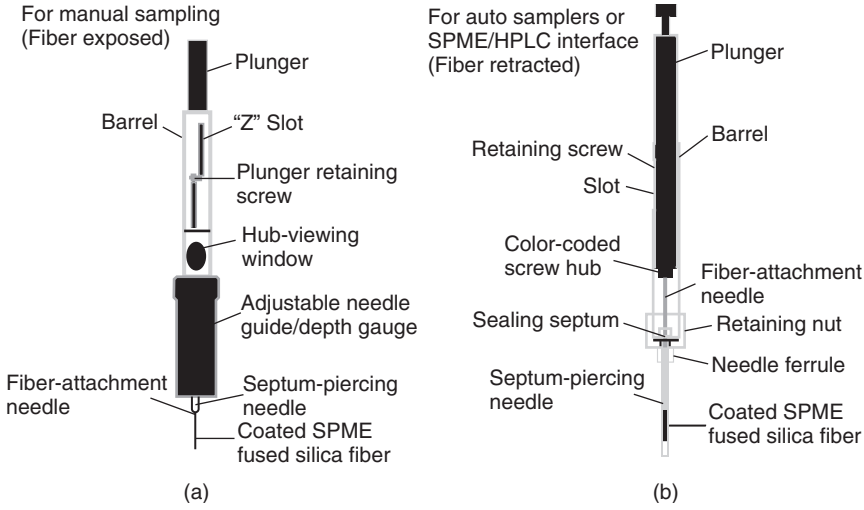


Figure 2.3 Diagrams of (a) manual and (b) auto sampler solid phase microextraction devices.

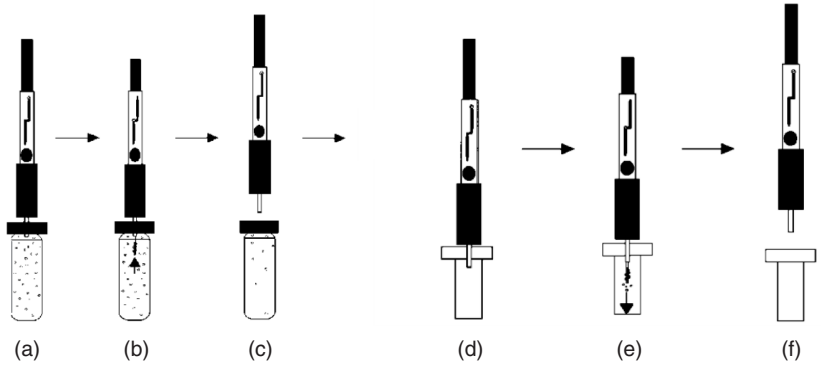


Figure 2.4 SPME diagram: (a) protective cover piercing septum; (b) fiber exposure and insertion into sample; (c) fiber withdrawn into cover and removed from sample; (d) cover inserted into GC injection port; (e) fiber exposed; (f) fiber withdrawn and removed from injection port.

has extensively investigated dairy flavors. Rouseff [9, 10] utilized SPME and published comprehensively on a diverse list of food and flavor topics with a special emphasis in citrus. Wright [11–13] identified key aroma impact components and off odors in a variety of foods, beverages and agricultural products.

2.5.2 Practical

As well as the original PDMS fiber coating, a variety of coatings is now available that adds to SPME extraction capabilities and includes Carboxen, divinyl benzene, polyacrylate, and carbowax (polyethylene glycol; PEG).

Coatings may be combined or available separately, or both. The manufacturer recommends 85 μm Carboxen PDMS for extracting components with 125 mw (molecular weight) or less. For larger components of interest, polarity plays a larger role in extraction efficacy. Polyethylene glycol (PEG), polyacrylate (PA), and DVB/Carboxen fibers are recommended for polar analytes [14].

Figure 2.5 shows area response by fiber type and Figure 2.6 shows analyte concentration vs. response for a Carboxen–PDMS fiber. Fibers with different film thickness are also available. A thicker film coating allows for more analyte loading into the liquid polymer coating (absorption), and by theory more analyte for analysis and detection. For example there are three fibers available with different PDMS film thicknesses: 100 μm, 30 μm, and 7 μm. The larger 100 μm film will absorb more analyte, but may also release the analyte less efficiently, especially larger molecular weight semivolatiles. This may result in carryover where peaks are seen on subsequent analyses. A 30 μm coating is

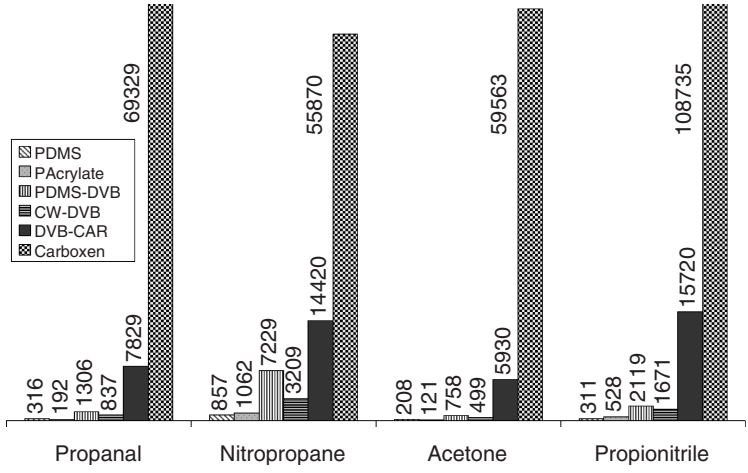


Figure 2.5 Area response by fiber type. PDMS is polydimethylsiloxane; PAcrylate is polyacrylate; PDMS-DVB is polydimethylsiloxane-divinylbenzene; CW-DVB is carbowax-divinylbenzene. *Source*: Supelco Corp.

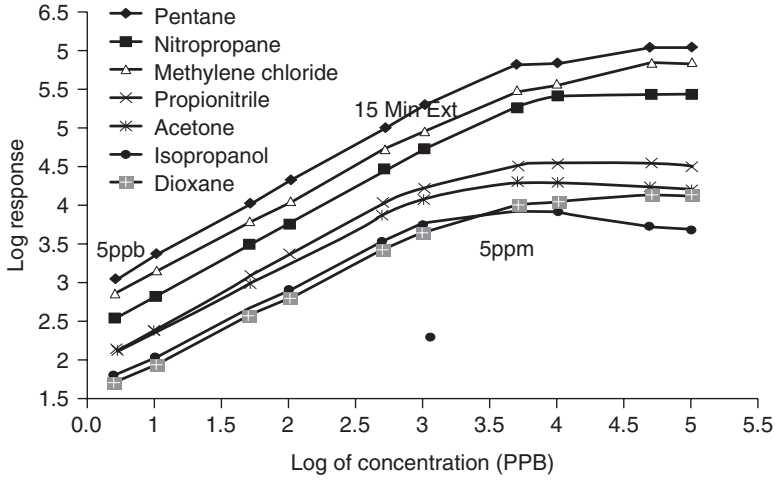


Figure 2.6 Carboxen–PDMS SPME concentration vs. analyte response. *Source*: Supelco Corp.

intended for analysis of semivolatiles and the 7 μm fiber is intended for immersion and extraction of semivolatiles with molecular weights of 250 or more.

In addition to the original fused silica fiber, a new metal alloy fiber is available. The advantage of this new core material is its improved durability and usefulness in extractions associated with the popular CTC robotic auto sampler that has a capability to agitate samples while exposing the fiber to sample or sample headspace (CTC, Zwingen, Switzerland).

Fibers also come in a variety of needle gauges. The smaller 24 gauge fiber is recommended for extractions and injections utilizing the manual holder and silica GC inlet septa. This is to decrease septa coring. The slightly larger 23 gauge fiber is recommended for use with the Merlin Microseal (Merlin Instruments, Half Moon Bay, CA, USA) septum that has a longer life (should last for 1 year or 25 000 injections) and does not contaminate injection port liners with shaved silica particles. The larger size is needed here to ensure no leakage into the injection port.

Another important item to consider when conducting extractions with SPME and optimizing chromatography is the GC injection port liner. A 0.75 mm injection port liner delivers superior chromatographic results compared with the normal 2 mm internal diameter liner when using

SPME. Peak shape and less tailing are noticeably improved due to less band broadening.

SPME fibers must first be conditioned before initial use based on the manufacturer's specific conditioning recommendations (a GC injection port or similar device created for fiber conditioning). This step is important because residual artifacts from manufacturing and the environment are baked off and could otherwise obscure important chromatographic peaks. Additionally fibers benefit from brief bake-outs prior to subsequent daily use to remove residual volatiles and volatiles collected from the laboratory atmosphere. It is good practice to conduct blank analyses with SPME periodically, especially if a series of extractions are conducted with an autosampler, to ensure fibers have not collected artifacts or environmental contaminants.

Unlike conventional static headspace extraction where volatile partitioning between the sample and sample headspace is a limiting factor, successful SPME results depend on two separate partitioning coefficients. The first is the same as with convention static extraction, that is, partitioning between sample and sample headspace. This partitioning has a practical application in determining how much time is required for components of interest to pass from the liquid or solid sample into the sample headspace, or how long should the sample be placed in a sealed container prior to extraction. Thus, unless equilibrium and exposure times are carefully duplicated, headspace concentrations will not be the same from one analysis to the next and precision will be unacceptable.

If sample components are not allowed sufficient time to reach equilibrium between the two phases (sample and headspace) the resultant analyte headspace concentration will be lower. However, most procedures do not wait for equilibrium conditions, requiring precise control of exposure time to maintain reasonable reproducibility. It should be noted there are ways of increasing headspace analyte concentrations including increasing sample temperature and changing the ionic strength of aqueous sample (add salt). Also by ensuring headspace volume is approximately one third of the volume of the sample, a maximum headspace concentration will result [15].

The second partitioning effect between sample and fiber may be considered in terms of how long the fiber is exposed to the sample. Another factor which is important in complex samples is the competitive absorption between headspace analytes. As briefly discussed previously, headspace components absorb into PDMS, or adsorb to particles embedded in PDMS, for example, Carboxen, divinyl benzene.

2.5.3 Advantages

Because of the dual equilibrium associated with SPME functionality, semivolatiles are more easily extracted than with conventional headspace methods. This is because once most semivolatile analytes are present in the headspace they are readily extracted by a PDMS fiber. These larger analytes are retained in the fiber coating where a concentration effect occurs that improves recovery (compared with static and dynamic headspace analyses). Also, unlike traditional headspace methods that may require heating samples above 100 °C, SPME will commonly extract these analytes at lower temperatures usually ranging between 45–70 °C [16]. SPME is very suitable for comparing samples of the same product, determining if a product has been adulterated, or determining a source of a flavoring or fragrance.

Once familiar with the basic tenets of SPME methodology, a researcher will appreciate the capacity for rapid sample analysis, simple automation, and minimal cost. Because there is no need for organic solvents, the hassle and expense associated with hazardous wastes are minimized. The opportunity for good accuracy and precision are easily attainable. Additionally, by employing innovation, unconventional volatile extractions are possible. An example of thisis novel work which was conducted by Payne *et al.*, [17] who utilized SPME to extract volatiles directly from the oral cavity.

2.5.4 Disadvantages

SPME may not be an answer to all questions regarding volatile profiles. A researcher requires an accurate picture of what is occurring regarding propensities for extractions of certain functional groups vs. others. PDMS has a greater affinity for nonpolar volatiles [4], although with the addition of Carboxen, and other fiber coating ingredients and film thicknesses, this problem is improved. Based on the discussion above, the extraction of components by PDMS may follow a model predicted by log *P*, or the octanol–water partition coefficient. Additionally extraction efficiency is determined by the phase ratio (β) defined by:

$$\beta = (V \text{ aqueous phase}/V \text{ PDMS phase})$$

Given a 100 μm PDMS coated fiber has approximately 0.5 μL of PDMS, variable extraction for nonpolar volatiles may result due to competition between the coated fiber, the aqueous phase, and the glass walls

of the sample vessel [4, 5]. Such competition on an absorbent PDMS fiber is more common if one analyte is at a much higher concentration and saturates the absorbent phase. Adsorbent type fibers are more prone to competition due to limited pore sites. An analyte with higher affinity can displace an analyte with less affinity. This problem can be greatly reduced by shortening extraction times. It is difficult to quantify complex mixtures using SPME, since it is not an exhaustive extraction technique. If quantification is required for samples with a variety of analytes, SPME is probably not the best extraction tool. However, it can be an ideal tool for quantifying a specific analyte, or a small group of analytes, especially if an internal standard is utilized.

As previously stated PDMS may result in sample carryover especially with larger semivolatiles. The durability of fibers as well as precision were both once a problem but technical advances in these areas have resulted in major improvements.

2.6 STIR BAR SORPTIVE EXTRACTION

Stir bar sorptive extraction (SBSE) was developed by Baltussen *et al.* [18] and is similar to the technology of the SPME fiber that incorporates a coating of PDMS (only). GERSTEL Twister is the commercial product that utilizes this technology and is manufactured and available from Gerstel GmbH. (Gerstel, Mülheim an der Ruhr, Germany). Twister is made of an inner, magnetized, metallic bar. A glass layer covers the metal bar and PDMS coating covers the glass (Figure 2.7). Currently there are four different sizes available that range from 10 mm length and 500 μm phase thickness with 24 μL PDMS up to a length of 20 mm, 1000 μm phase thickness and 126 μL of PDMS. The larger size of SBSE allows for a greater amount of PDMS coating and more efficient stirring of larger volumes. Twister has approximately 50–250 times more extraction phase than the comparable 100 μm PDMS coated SPME fiber and solute detection limits of less than 1 ng/L (ppt) have been reported [19].

2.6.1 Research

SBSE usefulness has proved diverse and capable for extracting volatiles and semivolatiles from aqueous media, foods, and beverages. A few of the many studies include those by Novotny's group who utilizes SBSE for extraction of aqueous media in chemical ecology and biological media [17, 18, 20]. Demyttenaere [21] measured malt whiskey, Kishimoto

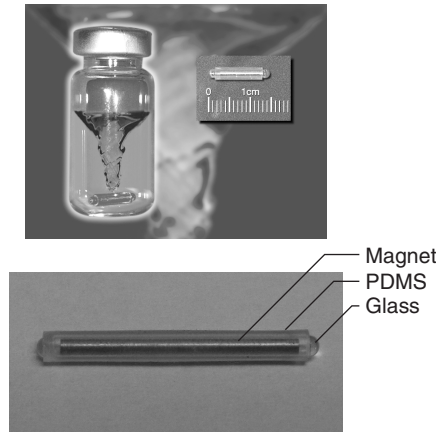


Figure 2.7 Twister consists of a magnetic stir bar coated with a layer of glass and polydimethylsiloxane. *Source*: Gerstel Inc.

measured terpenoids in beer [22], and León [23] extracted semivolatiles from water. Gurbuz and Rouseff [24] successfully utilized SBSE to extract wine volatiles for analysis by gas chromatography olfactometry. David and Sandra [25] have an excellent Stir Bar Sorptive Extraction review article currently in press.

2.6.2 Practical

Twister may be added directly into a liquid sample or it may be utilized for headspace extraction (sometimes referred to as HSSE). In the latter an open glass insert is available from Gerstel to suspend the Twister in the headspace above the sample. At a pinch a paperclip inserted through a septum will also work. The magnetic attraction between the stir bar and paper clip will hold the stir bar suspended above the sample.

Twister, like SPME, must be preconditioned prior to use and this is accomplished by heating as per the manufacturer's guidelines. Following an extraction, Twister is manually placed into a TDU tube, or glass desorption liner. The desorption process differs from that associated with SPME where the polymeric coated fiber is inserted into a hot GC injection port. Besides the obvious that Twister is too large for desorption in a heated injection port, the thicker PDMS coating on Twister requires additional time for desorption and requires refocusing prior to transferring to the analytical column. Gerstel has automated this process (after placing Twister into a desorption liner and putting the liner in an auto sampler storage tray). If using the MPS2 auto sampler

(Gerstel) desorption liners containing Twisters are placed in the MPS2's storage tray fitted with cylindrical glass tubes to house up to 98 Twisters (two trays or up to 196 liners may be continuously analyzed). The MPS2 auto sampler transfers liners from the storage tray to the Twister Desorption Unit (TDU) located above and attached to a cooled injection system (CIS). The CIS is located on the GC in the same position as standard on-column or split/splitless injectors. The TDU may be cooled below ambient temperature by integration with a Peltier cooling device (utilizes a combination of water–ethanol) and then heated to desorb volatiles concentrated in Twister (Twister is located inside the heated TDU). The TDU is attached to a liquid nitrogen cooled injection system (CIS) that contains a glass liner (with glass wool insert) and functions to cryorefocus (trap) components desorbed from Twister in the TDU. Following this period of cryotrapping volatiles that are transferred from the TDU, the CIS has programmable heating to thermally desorb comp onents (components that have been cryofocused in the liner located inside the CIS) by rapid vaporization. Components are then transferred in either the split or splitless mode onto the analytical column for chromatographic separation. The TDU may be cooled again by the Peltier device to decrease time between multiple sample analyses.

2.6.3 Advantages

SBSE has established a good track record of performance. Twisters are durable and have a useful life of approximately 50–300 extractions depending upon extraction media (E. Pfannkoch, personal communication). As previously discussed, SBSE has more PDMS coating and thus a larger capacity for volatile extraction vs. PDMS SPME. Loughrin [26] measured extraction of volatile components from waste water and found the performance of the PDMS stir bars compared favorably with that of the polar SPME fibers at octanol–water ($\log k_{ow}$) values > 1.50. Bicchi [27] reported both direct (placed in liquid) and headspace SBSE extractions yielded greater volatile recoveries from Arabica coffee than obtained by SPME. Also important is the fact that this system may be automated with the Gerstel MPS2 and up to 196 samples may be analyzed continuously.

2.6.4 Disadvantages

Currently SBSE (Twister) products are available only with PDMS coating. An adsorption material that would aid in binding highly volatile

polar components would be useful. Carryover may be a concern with the thicker PDMS coating. Bake-out times must be sufficient to desorb all volatiles/semivolatiles and occasional blank runs should be conducted to ensure there is no carryover from previous analyses. Equipment necessary to utilize Twister to its fullest, including Gerstel's TDU, CIS, and MPS2 requires some capital investment.

2.7 PDMS FOAM AND MICROVIAL

Two novel and relatively new extraction methods from Gerstel, PDMS foam and Microvial, provide increased capabilities and will be briefly discussed below.

2.7.1 PDMS Foam

In addition to PDMS on a fiber (SPME) and stir bar (SBSE), it is also available as a foam plug loaded into a TDU tube (available through Gerstel). Each tube contains 85 mg +/− 2.7 mg of PDMS foam. This method offers the largest PDMS mass and surface area available and can be used to trap volatile components in a dynamic headspace extraction process. Desorption is conducted utilizing the same equipment as with Twister (Gerstel's MPS2, thermal desorption unit and cooled injection sytem). In samples with a lot of water, vapor may condense inside the TDU tube during the purge and trap process. Formation of an ice plug in the cryocooled liner may result. A nice feature of this system is the capability to purge tubes with nitrogen and vent water vapor prior to cryorefocusing absorbed components. Marsili and Laskonis [28] constructed a dynamic extraction apparatus (Figure 2.8) that recirculated air through a PDMS foam tube. Results after 5 minutes of recirculation are displayed in Figure 2.9. They concluded this extraction technique provided the accuracy (as measured by standard calibration curves) and precision necessary for most analytes studied.

2.7.2 Microvial

This procedure refers to a novel technique that evolved from problems associated with direct liquid injection of samples contaminated with nonvolatile components (a process sure to decrease column life). A disposable glass microvial capable of holding up to approximately 200 μL of liquid (more liquid than in a normal liquid injection) is

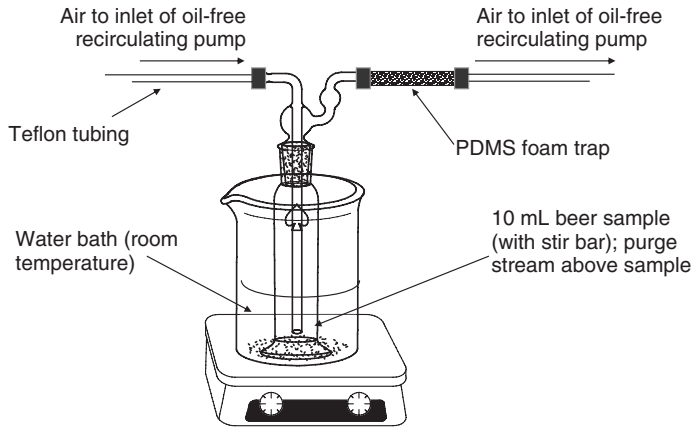


Figure 2.8 Recirculating dynamic headspace extraction apparatus with PDMS foam as the trap. *Source*: Marsilli Consulting Group.

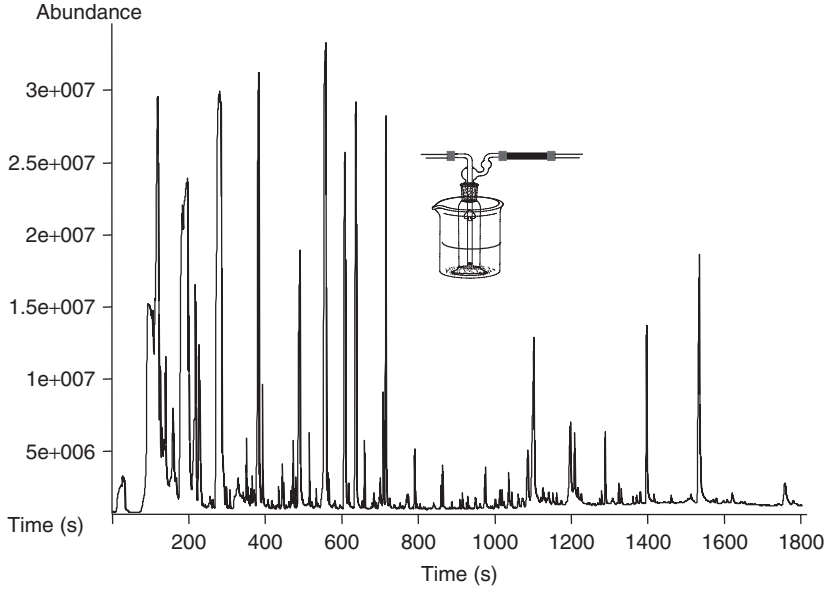


Figure 2.9 Results from PDMS foam extraction of beer headspace utilizing recirculating extraction apparatus. *Source*: Marsilli Consulting & Gerstel Corp.

placed inside a TDU tube (Figure 2.10). Both are placed into the thermal desorption unit. As discussed with PDMS foam, purge and programmable heating capabilities of the thermal desorption unit allow for evaporation and venting of solvent/water and cryorefocusing volatiles

onto an inlet liner. The volatiles trapped in the liner are then vaporized by rapid heating and transferred onto the column without problems related to nonvolatiles and an overload of solvent. A big advantage of this method is efficiency and the means to analyze multiple samples.

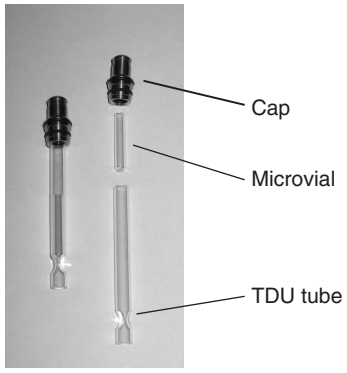


Figure 2.10 Microvial constituents including: (a) TDU tube, (b) microvial, and (c) metal cap. The cap with visible O-rings functions to provide an airtight seal so that the TDU can maintain system pressure. It also functions as a means for the auto sampler to grip and transfer the tube between the storage tray and TDU. *Source*: Gerstel Inc.

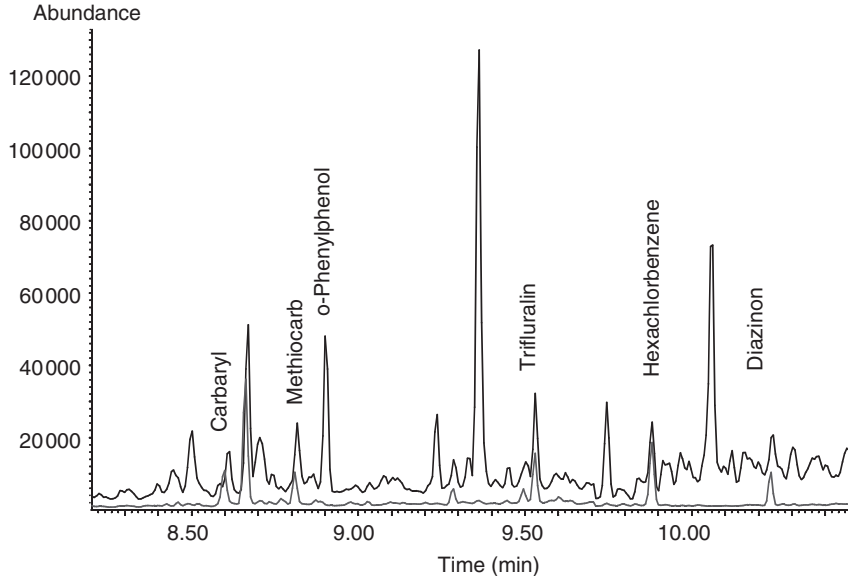


Figure 2.11 Microvial analysis. Overlay of 10 ppb pesticide in standard and 10 ppb pesticide in carrot juice extract. *Source*: Gerstel Inc.

Gerstel has automated it entirely from injection of a dirty sample into the microvial and transfer to the thermal desorption unit to desorption, analysis, and removal of microvial.

Pfannkoch, Whitecavage, and Stuff [29] analyzed by microvial 10 μL of carrot extract spiked with 10 ppb pesticides and the same concentration of pesticide in solvent (Figure 2.11). The peak sizes of pesticides obtained from the complex carrot juice sample were similar in size and shape to those from the much cleaner standard sample. After analysis, the microvial and residual nonvolatile material from the extract were disposed of, minimizing contamination of the inlet.

2.8 SOLVENT EXTRACTION

In its simplest form solvent extraction is conducted by the addition of a liquid or solid sample to a liquid solvent, agitation to ensure mixture, and removal of the solvent plus that which dissolved in it. However, many factors affect solubility and include lipophilicity, polarity, temperature, pH, molecular weight, pressure, mixing thoroughness, and time. A brief overview of some liquid extraction methods follow.

2.8.1 MIXXOR

A separatory funnel is probably the best known solvent extraction tool. The sample is added to immiscible liquids, shaken, and the appropriate layer with solute is removed. MIXXOR (NBS Systems, Haifa, Israel) improves separation due to vigorous blending and corresponding mass transfer (Figure 2.12). Samples processed with MIXXOR or a similar device, and centrifuged to separate layers, have provided successful results. Parliament [30] reported success utilizing a similar apparatus and Jella *et al*. [31] followed this method to extract flavor components from grapefruit juice.

2.8.2 Soxhlet Extraction

A liquid extraction process for utilizing solvent to extract solid samples involves the Soxhlet extractor (Figure 2.13). A container (thimble) open at the top, made of permeable, filter paper-like cellulose, is placed inside the soxhlet extractor and positioned directly beneath a condenser. Solvent is added to a round bottom flask and heated with a heating mantle.

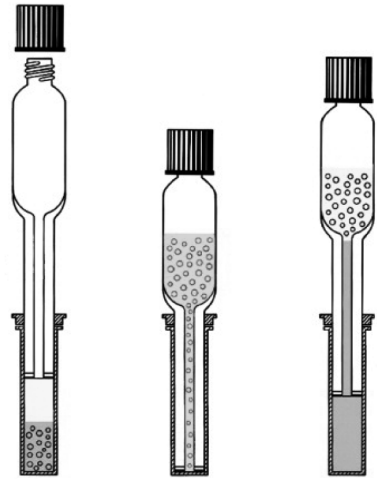


Figure 2.12 A MIXXOR apparatus takes the separatory funnel concept a step further in efficiency. *Source*: Sigma Aldrich.

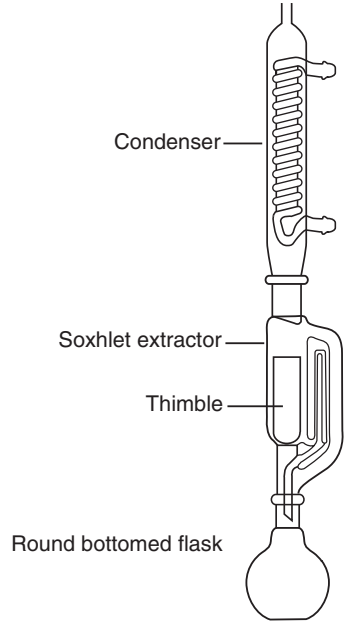


Figure 2.13 Soxhlet extraction apparatus.

Solvent evaporates from the round bottomed flask, rises through the extractor, condenses in the condenser, drips back into the extractor, and gradually immerses the contents of the thimble. As the level of solvent reaches a critical volume in the extractor, the apparatus design is such that all of the solvent in the extractor drains back into the round bottomed flask leaving behind solid material in the thimble. This

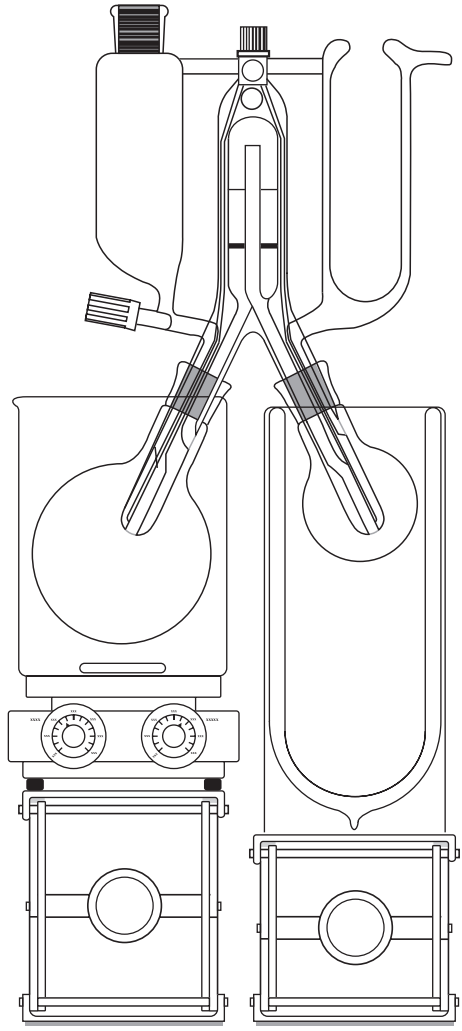


Figure 2.14 Solvent Assisted Flavor Evaporation (SAFE) apparatus developed by Engel *et al*. [32] improved upon pre-existing simultaneous distillation and solvent extraction techniques.

process is repeated many times over several hours to provide multiple extractions of the solid material present in the thimble. Ideally suited to remove lipid material from a solid, depending on the solvent used this method may also be used to extract other components.

2.8.3 Solvent Assisted Flavor Evaporation (SAFE)

Engel *et al.* [32] described Solvent Assisted Flavor Evaporation (Figure 2.14), a method that combined vacuum distillation, cold trapping, and in some cases solvent extraction (solid samples such as popcorn, coffee or bread crust required solvent extraction). This method has proven useful and a step above previous simultaneous distillation extraction (SDE) procedures. What distinguishes this method is it is faster, safer, and easier to maintain temperature controls than previous methods that employed similar high vacuum, cold trapping conditions. Additionally it yields an extract that minimizes formation of off-flavor artifacts and provides a close approximation to authentic flavor. In addition to Engel's original work, another good review of this technique is found in Werkoff *et al.* [33].

2.9 SUMMARY

Due to a greater requirement for rapid sample screening, sample prep techniques for subsequent GC separation have necessarily become automated to allow multisample analysis. Tools that lend themselves to automation such as SPME, Twister, and dynamic headspace extraction traps will be utilized more frequently in the coming years as demand for basic chemical information continues to increase. The novice scientist must be able to utilize such techniques in order to be an asset in industry and academia. The informed scientist must also understand fundamental extraction methodology because principles of future techniques that meet the rapid and automated criteria will be based on tried and true methods including static, dynamic, solvent, and distillation extractions.

REFERENCES

1. Micro Total Analysis Systems, Silly Putty, and Fluorous Peptides http://www.fluorous.com/journal/?p=86.
2. Lee, J. N., Park, C., and Whitesides, G. M. (2003) Solvent compatibility of poly(dimethylsiloxane)-based microfluidic devices. *Anal. Chem.*, 75, 6544.

3. Lotters, J. C., Olthuis, W., Veltink, P. H., and Bergveld, P. (1997) The mechanical properties of the rubber elastic polymer polydimethylsiloxane for sensor applications. *J. Micromech. Microeng.*, **7**, 145–147.
4. David, E.T., Tienpont, B., Sandra, P. (2003) Stir-bar sorptive extraction of trace organic compounds from aqueous matrices. *LC/GC*, **21**, 109.
5. Reineccius, G. (2006) *Flavor Chemistry and Technology*. CRC Press, Boca Raton, FL, USA, pp. 33–72.
6. Arthur, C. L., and Pawliszyn, J. (1990) Solid phase microextraction with thermal desorption using fused silica optical fibers. *Anal.Chem.*, **62**, 2145.
7. Marsili, R.T. (1999) SPME-MS-MVA as an electronic nose for the study of off-flavors in milk. *J. Agric. Food Chem.*, **47**, 648–654.
8. Marsili, R. (1999) Comparison of solid phase microextraction and dynamic headspace methods for the gas chromatographic – mass spectrometric analysis of light induced lipid oxidation products in milk. *J.Chrom. Sci.*, **37**, 17–23.
9. Rouseff, R. L. and Cadwallader, K. R. (2001) *Headspace Analysis of Foods and Flavors: Theory and Practice*. Kleur Publishing, New York, USA, pp. 212.
10. Rouseff, R., Bazemore, R.; Goodner, K., and Naim, M. (2000) *GC-Olfactometry with Solid Phase Microextraction of Aroma Volatiles from Heated and Unheated Orange Juice. In Headspace Analysis of Foods and Flavors: Theory and Practice* (eds R. L. Rouseff and K. R. Cadwallader), Plenum, New York, USA.
11. Koziel, J., Cai, L., Wright, D., and Hoff, S. (2007) Solid phase microextraction as a novel air sampling technology for improved, GC-olfactometry based, assessment of livestock odors. In press *J. Chrom. Sci.*
12. Wright, D. *et al.* (2005) Multidimensional gas chromatography-olfactometry for the identification and prioritization of malodors from confined animal feeding operations *J. Agric. Food Chem.*, **53**, 8663–8672.
13. Wright, D.W. *et al.* (2004) Multidimensional gas chromatography-olfactometry based investigations of odor quality issues in packaging and consumer products. Proceedings of TAPPI Conference, Indianapolis, Indiana, USA. TAPPI.
14. Shirey, R. (1999) Sample prep made easy. *SPME Newsletter Winter 1999/2000*.
15. Anonymous (1998) Solid phase microextraction: theory and optimization of conditions. *Supleco Bulletin*, **923**, 1–8.
16. Zhang, Z. and Pawliszyn, J. (1993) Headspace solid phase microextraction. *Anal.Chem.*, **65**, 1843–52.
17. Payne, R., Labows, J., and Liu, X. (2000) Released oral malodors measured by solid phase microextraction/gas chromatography mass spectrometry (HS-SPME-GC-MS). Proceedings of ACS Flavor Release No. 0841236925. American Chemical Society, Washington, DC, USA.
18. Baltussen, E., Sandra, P., David, F., and Cramers, C. (1999) Stir bar sorptive extraction (SBSE), a novel extraction technique for aqueous samples: Theory and principles. *J. Microcol.*, **11**, 737–747.
19. Soini, H. A. *et al.* (2005) Stir bar sorptive extraction: a new quantitative and comprehensive sampling technique for determination of chemical signal profiles from biological media *J. Chem. Ecol.*, **31**, 377–392.
20. Soini, H. *et al.* (2005) Comparative investigation of the volatile urinary profiles in different *Phodopus* hamster species. *J. Chem. Ecol.*, **31**, 1125–1143.
21. Demyttenaere, J. C. R., Moriña, R., and Sandra, P. (2003) Analysis of volatiles of malt whisky by solid-phase microextraction and stir bar sorptive extraction. *J. Chrom. A.*, **985**, 221–232.

22. Kishimoto, T., Wanikawa, A., Kagami, N., and Kawatsura, K. (2005) Analysis of hop-derived terpenoids in beer and evaluation of their behavior using the stir bar-sorptive extraction method with GC-MS. *Agric. Food Chem*., **53**, 4701–4707.
23. León, V. M. *et al*. (2003) Analysis of 35 priority semivolatile compounds in water by stir bar sorptive extraction-thermal desorption-gas chromatography-mass spectrometry I. Method optimization. *J. Chrom. A*., **999**, 91–101.
24. Gurbuz, O. and Rouseff, R. (2005) Analysis of aroma active compounds in wine using stir bar sorptive extraction with GC-olfactometry. Proceedings of the 56th Citrus Processor's Meeting, October 20, 2005, Lake Alfred, FL, USA.
25. David, F. and Sandra, P. (2007). Stir bar sorptive extraction for trace analysis. *J. Chrom. A*., **1152**, 54–69.
26. Loughrin, J. (2006) Comparison of solid-phase microextraction and stir bar sorptive extraction for the quantification of malodors in wastewater. *J. Agric. Food Chem*., **54**, 3237–3241.
27. Bicchi, C., Iori, C., Rubiolo, P., and Sandra, P. (2002) Headspace sorptive extraction (HSSE), stir bar sorptive extraction (SBSE), and solid phase microextraction (SPME) applied to the analysis of roasted arabica coffee and coffee brew. *J. Agric. Food Chem*., **50**, 449–459.
28. Marsili, R. and Laskonis, L. (2006) Evaluation of PDMS-based extractiontechniques and GC-TOFMS for the analysis of off-flavor chemicals in beer. Gerstel application note 10/2006.
29. Pfannkoch, E. and Whitecavage, J. (2006) Elimination of non-volatile sample matrix components after GC injection using a thermal desorber and microvial inserts. Gerstel application note 4/2006.
30. Parliament, T. H. (1986) A new technique for GLC sample preparation using a novel extraction device. *Perfum. Flavor*., **11**, 1–8.
31. Jella, P., Rouseff, R., Goodner, K., and Widmer, W. (1998). Determination of key flavor components in methylene chloride extracts from processed grapefruit juice. *J. Agric. Food Chem*., **46**, 242–247.
32. Engel, W., Bahr, W., and Schieberle, P. (1999) Solvent assisted flavor evaporation – a new and versatile technique for the careful and direct isolation of aroma compounds from complex food matrices. *Eur. Food Res. Technol*,, **209**, 237–241.
33. Werkhoff, P., Brennecke, S., Bretschneider, W., and Bertram, H.J. (2002) In *Flavor: Fragrance, and Odor Analysis* (ed. R. Marsili), Marcel Dekker, New York, USA, p. 139.

3

Traditional Flavor and Fragrance Analysis of Raw Materials and Finished Products

Russell Rouseff
IFAS, Citrus Research and Education Center, University of Florida, Lake Alfred, USA

Kevin Goodner
Sensus, LLC, Hamilton, USA

3.1 OVERVIEW

Most flavor samples have a biological origin and thus typically consist of complex chemical mixtures with a wide range of chemical and physical properties which require separation. In addition, flavor active components are typically trace components which require that the compounds of interest need to be extracted and concentrated in order to simplify the analysis and to raise the concentration of the components of interest above instrumental detection limits. Specific details of this

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

critical step have been presented in the proceeding chapter. However, in discussing analytical procedures, it cannot be overemphasized that the most sophisticated analytical procedures cannot compensate for an inadequate extraction. The objective of the analysis process is to separate an extract into its constituent chemicals to determine either their nature (qualitative analysis) or their amounts (quantitative analysis).

In the flavor industry the acquired data is used by various groups within a company to satisfy specific needs. For example, purchasing decisions are predicated on a material's ability to meet certain analytical specifications that may include standards of identity, authenticity, chemical purity, and overall component profile. The manufacturing sector relies upon analytical information to control processes and assess equipment effectiveness. Research groups require the ability to identify and quantify unique chemicals in order to measure the completeness of an anticipated reaction. Quality control groups would be most interested in determining the major flavor impact compounds in order to maintain a consistent flavor profile as well as identifying off-flavors that occasionally appear. Figure 3.1 illustrates how analytical chemistry associates with nearly every facet of a flavor and fragrance company.

The abundance, complexity, and application of analytical methodology have significantly increased with the advancement of instrumental technologies with the associated computer hardware and software. Almost all modern instruments employ computers to control the instrumental processes and record data thus allowing for unattended analysis and automation. Modern analytical laboratories are typically equipped with a wide array of instrumentation capable of trace analysis.

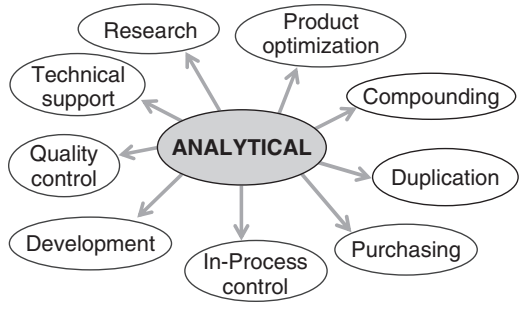


Figure 3.1 Reliance on analytical support throughout the various departments of a flavor and fragrance company.

One of the major goals for flavor and fragrance companies is to produce high quality products with consistent sensory properties. Analytical tools have altered the once exclusively intuitive flavor/fragrance creation processes towards one of more calculated science based upon chemical composition. The increased ability to determine a product's complete composition analytically has motivated flavorists and perfumers to increase the sophistication of their formulations to achieve differentiation and thwart duplication from the competition. This is typically achieved by adding potent, high sensory-impact components that are difficult to quantify and identify chemically.

At one time sensory evaluation was the exclusive means of assessing the quality and consistency of a given flavor or fragrance and it still is a major assessment tool. Today, sensory analysis has been augmented with detailed analytical information which is particularly useful when determining subtle differences in sensory profiles. It was the only available tool prior to the development of analytical instrumentation. Sensory analysis still remains the final checkpoint for a developer or quality control analyst; however, analytical equipment provides objective, repeatable, and reliable assessment.

Traditional flavor and fragrance analysis involves the evaluation of physical, chemical, and sensory characteristics. This chapter covers the analytical aspects, so sensory techniques will be discussed later in the book. Prior to gas and liquid chromatography, classical methods focused on basic physical attributes such as color and clarity. More objective methods evolved to include specific gravity, optical rotation, refractive index, and so on. These tools were employed to determine the authenticity of essential oils and other flavoring materials. Today, methods are specifically tailored to identify and quantify individual chemicals within every imaginable matrix (i.e., living material, extract, headspace, finished flavor and fragrances, store-bought beverage, etc.). This shift from physical to chemical attribute evaluation has been made possible because of greater instrumental selectivity for specific components, improved detector sensitivity, and computer controlled instrumental automation.

3.2 PHYSICAL ATTRIBUTE EVALUATION

There are a number of physical characteristics of products that are important to the food and flavor industry. Generally, product appearance, moisture content, and concentration are some of the most important and most often measured.

3.2.1 Color – Optical Methods

One of the primary methods consumers use to evaluate food and beverages is their appearance. Therefore, the color of the final product – and, by inference, its ingredients – is a critical attribute to determine. There are two generally used measurements for determining color: absorbance of light at a specific wavelength or by measuring the color using the LAB system (discussed below).

Specific wavelengths are sometimes used to describe the intensity of color of a product. For example, caramel color manufacturers Sethness Products and D.D.W. Williamson, two leading caramel color manufactures, each use the absorbance of a 0.1 % weight/volume solution but at differing wavelengths. Sethness Products describes their products based on the absorbance at 560 nm while D.D.W. Williamson uses 610 nm. Both companies use the hue index to describe the redness of their products. The hue index is 10 times the logarithm of the absorbance at 510 nm divided by the absorbance at 610 nm [1].

$$\text{Hue} = 101\text{Log}\left(\frac{A_{510}}{A_{610}}\right)$$

These two values – intensity and hue – are the main appearance characteristics used in the caramel color industry.

Another example is in the brewing industry. One of the main methods is the Standard Reference Method, which is defined by

$$\text{SRM} = 12.7 * D * A_{430}$$

where D is a dilution factor. This method yields numbers generally between 2 and 70 which would correspond to a light yellow colored beverage to a very dark brew. These two industries are just examples of some of the methods that specific products use involving absorption.

While some companies and researchers use specific wavelengths, others use the LAB color space. The Hunter 1948 LAB color space is designed to approximate human vision using the variables L, a, and b. However, the Hunter 1948 method has largely been replaced with the CIE 1976 L^*, a^*, b^* space where CIE is the abbreviation for the International Commission on Illumination [2]. In general, the L^* value matches the human perception of lightness ($L^* = 0$ is black and $L^* = 100$ is white), a^* is a measure of the red/green of a sample (i.e., negative a^* values mean green and positive values red), and b^* is a measure of the yellow/blue

of a sample (i.e., negative b^* values mean blue and positive values mean yellow).

A common colorimeter used by academics and the food and flavor industry is the Minolta CR-400 series colorimeter. This instrument is a portable, handheld unit that can determine L^*, a^*, b^* values of products by using reflectance or transmittance with an adapter. Given the history of the LAB system, it is no surprise that Hunter Lab systems are also popular colorimeters. These systems and others can provide analytical data on the color of ingredients or final products.

One thing to keep in mind, whether using a spectrometer to measure specific wavelengths or a colorimeter to obtain L^*, a^*, and b^* values, these numbers are useful for quality control/quality assurance, research and development, and product development, but do not indicate consumer acceptance. Sensory testing (Chapter 8) is critical to determine consumer acceptance.

3.2.2 Turbidity

While color is an obviously critical component in food products, another important factor in beverages is turbidity (i.e., clarity) [3]. The desired clarity of a product varies depending on the product and even varies within a product category. For example, some fruit juices are clarified and the consumer has come to expect that (e.g., apple and grape) while others are not (e.g., orange juice). Other examples include tea and coffee, where different brands have different turbidity levels ranging from clear to highly turbid. Turbidity is measured by how much a sample will scatter light passing through it. Measuring turbidity in this manner uses a nephelometer which reports nephelometric turbidity units (NTU). Turbidity tests are generally performed on relatively inexpensive equipment and only take a couple of minutes to obtain results.

3.2.3 Water Activity

Water activity (A_w) is an important physical characteristic of a product. Water activity (A_w) is a measure of the energy of the water in a sample and it has no units. It is often described as measuring the 'available' water in a system. This description is used to explain the situation where two products have the same water content, but different water activity, or different moisture content and the same water activity. For example, pasta with a moisture content of 12 % could have a A_w of

0.50 while rolled oats with 10 % moisture could have a A_w of 0.70. One of the reasons water activity is so important in the food industry is that microorganisms will only grow if the water activity is acceptable. For example, many microorganisms cannot grow if the water activity is below 0.9, most molds require a water activity above 0.8, and a water activity below 0.6 will inhibit all microbial growth [4]. There are two main methods for determining water activity, the chilled mirror method and the capacitance method. The chilled mirror method uses a mirror which is chilled until dew is formed and detected with an optical sensor. The capacitance instruments use two charged plates with a polymer membrane between them. Instruments vary in price but generally start around $2000 and can determine the water activity of a sample in a few minutes.

3.2.4 Moisture Content

The moisture content of samples is another commonly measured parameter. There are several ways to determine this, but loss on drying (or evaporative residue) is probably the most common. This can be done in an oven or using bench top instrumentation. Generally the oven is cheaper and can dry more samples at once, but it takes longer to get results and is more manual, meaning a technician has to weight the sample before and after drying and perform the calculations. Additionally, oven drying can be problematic for some samples if they decompose or char because if there is only one oven but differing conditions are needed for various samples. These problems can be ameliorated by using a bench top automated moisture analyzer. Typical instruments today can have multiple methods for different types of samples and each method can be optimized to minimize drying time. Additionally, some units will use preliminary data to extrapolate to an end point, further decreasing analysis time and preventing charring or decomposition. Bench top moisture analyzers cost somewhere between $2000 and $20 000 depending on features such as balance accuracy/precision or using microwave energy to speed the heating.

3.2.4.1 Karl Fischer Method

While loss on drying is the most common and one of the most accepted methods for determining moisture, there are other methods. The only other primary moisture determination method is the Karl Fischer titration. The Karl Fischer titration is named after the German chemist Karl

Fischer who developed the method in 1935 [5]. The advantage of the Karl Fischer method is that it is specific for water while loss on drying will measure the loss of any volatile compounds (e.g., flavor volatiles). The Karl Fischer method is also used for very low moisture content samples as it is the most precise, accurate, and sensitive method. The titration uses an alcohol, base, sulfur dioxide, and iodine. The reactions that occur in the sample are listed below.

$$CH_3OH + SO_2 + C_3H_4N_2 \rightarrow (C_3H_4N_2 \cdot H) \cdot SO_3CH_3$$

$$(C_3H_4N_2 \cdot H) \cdot SO_3CH_3 + 2 \cdot C_3H_4N_2 + I_2 + H_2O \rightarrow (C_3H_4N_2 \cdot H) \cdot SO_4CH_3 + 2(C_3H_4N_2 \cdot H)I$$

In current laboratories the Karl Fischer method is most commonly performed with an automated system which is available from several manufacturers for approximately $10 000.

3.2.4.2 Secondary Moisture Determination Methods

While loss on drying and the Karl Fischer titration are considered primary moisture analyzers, there are several so-called secondary moisture determination methods. One method is to measure the water activity and determine the moisture content. This method is considered secondary as it can only be used once a relationship has been generated between water activity and moisture content, as determined by a primary method, and water activity. This relationship between moisture content and water activity has to be determined for every product. A new product would have to have the relationship between water activity and moisture determined. While that is a negative, such instruments determine water activity and moisture content in a single analysis. Another secondary method is by utilizing the near infrared (NIR) spectrum. Much like using water activity to measure moisture content, a calibration curve has to be generated for each product using the NIR spectra and moisture content from a primary moisture analysis (loss on drying or Karl Fischer). While the instrumentation is generally considerably more expensive than the other methods, results are available in seconds.

3.2.5 Optical Rotation

Optical rotation is a parameter which is commonly used in the flavor industry as a measure of identity. If plane polarized light is passed

through solutions, it can become rotated either clockwise or counter-clockwise [6]. Measuring this optical rotation can help determine the concentration and identity of a compound in solution. For this reason, it is a fairly common analysis of flavors to ensure that the concentration and identity are within the specifications provided by the manufacturer. Instruments to measure the optical rotation are generally available for around $10 000 for a bench top model.

3.2.6 Specific Gravity

Specific gravity, or relative density, is the ratio of the density of an item to that of a reference material. Most often, the reference material is water. The temperature at which the density measurement is determined must be specified for both materials. A common temperature for the reference material of water is 4 °C as the density of water at that temperature is 1.00 g/mL and therefore the density of the sample is equal to the specific gravity. In this manner, many people treat specific gravity to be the same as density (due to the convenient choosing of the reference). Density and/or specific gravity are both properties that are used to determine the purity of a product. There are many ways to determine specific gravity and/or density. The simplest and probably least accurate is to use a graduated cylinder to measure a volume of material which divided into the weight will provide the density of liquid. A more accurate method, but not without issues, particularly for viscous fluids is to use a volumetric flask in the same matter as the graduated cylinder. An item of specialized glassware called a pycnometer is a small vessel with a lid that has a small hole in it such that the vessel can be filled and excess material is eliminated through the small hole. The weights of the sample and water are then divided to determine the specific gravity. There is also instrumentation that can determine density and specific gravity. There are handheld units that cost a few thousand dollars and bench top models which are more precise and accurate but cost considerably more.

3.2.7 Refractive Index

Refractive index is a measure of how much the speed of light is reduced in an item. As light passes from one medium to another, the light will change direction. The greater the difference in velocity of light through the two media, the greater the observed angle [7]. Refractive index

is measured using a refractometer which can vary between handheld and bench top using either transmittance or reflectance measurement. As refractive index is an inherent property of a substance, it is commonly used to confirm the purity of samples, particularly in a QA/QC (quality assurance/quality control) type of environment. One particularly common use of refractive index is to measure the concentration of aqueous solutions. The refractive index of a solution will change as the concentration of the solute changes. This relationship can be measured and used to determine the concentration from the refractive index. Modern refractometers will automatically apply the relationship and displace the concentration of many analytes. Some examples include: ethanol content, high fructose corn syrup (HFCS), salinity, and most commonly, °Brix (sucrose concentration). This last relationship is used by many in the industry for many types of products – some of them have no sucrose at all. For example, many people measure tea and coffee solids as °Brix. In a sucrose solution, a 50 °Brix solution is 50 % solids and 50 % water. For other materials, such as tea and coffee, that is not the case. For example, a 50 °Brix tea or coffee concentrate will have approximately 40 % solids and 60 % water. However, while the measurement is not accurate, it is consistent and has been used for many years in the tea and coffee industry and therefore is not likely to be replaced.

3.2.8 Sugars/Soluble Solids

When dealing with sugar solutions, °B (degrees Brix) is used to describe the amount of dissolved sucrose. This is common in fruit juices, wines, and some other products [8]. °Brix is generally determined by measuring a physical property and relating that to the percentage of sucrose in a solution. This physical property can be various properties such as density, specific gravity, refractive index, or infrared vibrational wavelengths. This means that any solution can be analyzed to determine a °B measurement, even if little to no sucrose is in solution. While this might at first seem to be a totally erroneous method, it is not necessarily so. Consider that as the concentration of a solute (or in the case of complex mixture, solutes) increase, the refractive index will change. This will provide an apparent °B reading. This °B reading can either be used directly or related to actual concentration of solutes (which can be determined by drying or percent moisture analyses). However, using refractive index to determine °B is predicated upon one important assumption – that all solutes are actually in solution (and

thus affect refractive index) and not suspended (and thus do not affect refractive index). Additionally, there is the danger of suspended solids settling on the instrument's prism and interfering with the measurement. Density and specific gravity might solve the problems of suspended solids settling on the prism and possibly of suspended solids being correctly accounted for. There is the very real possibility that the density would be different if the solutes are suspended as opposed to being in solution. However, for many compounds the effects are very small and likely within experimental error. Compounds that dramatically affect density do so generally by affecting the volume of the solution. While this does occur it is usually the exception and not the rule. Infrared is specific for sugar vibration bands and therefore would not work for other solutes.

3.2.9 Viscosity

While many physical properties are useful for describing a product, possibly one of the more important for processing is viscosity. Many pumps and other equipment will not work if the viscosity is too high. Viscosity is a measure of a product's fluidity [9]. The lower the viscosity, the easier the product will flow. Viscosity can be measured with a viscometer which can range from glassware to bench top units. Viscosity is often reported in units of centipoise. However, this explanation of viscosity is really only skimming the surface of this vast topic. Whole books have been written solely on viscosity and one should explore those if viscosity is important to their product.

3.3 INSTRUMENTAL ANALYSIS

There are a wide range of instrumental techniques available to today's analytical flavor chemist. The technique employed will depend primarily on what information is desired and what techniques are available. Pure compounds will be treated differently than complex mixtures. In addition the physical state, volatility, solubility, and sample size are additional features which may exclude some techniques. Gas and liquid chromatography will be the major separation techniques discussed in this chapter. Major identification techniques will include mass spectrometry and nuclear magnetic resonance (NMR) as they are the most commonly employed techniques.

3.3.1 Separation Techniques

3.3.1.1 Gas Chromatography (GC)

Although there are many types of chromatography, high resolution capillary gas chromatography is by far the major technique employed to separate flavor materials which can be volatilized. It is not uncommon to have hundreds of components in a wide range of concentrations in a flavor sample. Good separation is difficult but required to obtain an accurate identification. Coeluting substances add extraneous data which will reduce identification confidence or in extreme cases result in a false identification.

3.3.1.2 GC Retention Data

The time it takes for a specific volatile to elute from the end of a GC column after injection is a characteristic of the volatile and its interaction with the column stationary phase. This time is also a function of several GC parameters such as carrier gas flow rate, column length, and oven temperature. It is called either elution or retention time, with the latter term being more widely used. Although this time is characteristic of each volatile, retention time is not a unique measurement and many compounds will potentially share the same retention time. For this reason, it should be kept in mind that GC is primarily a separation technique and not an identification technique. Although chromatographic retention time has been used as an identification technique in earlier literature, it is no longer considered acceptable practice. To reduce the number of potential volatiles that will have similar retention times, a second retention time is determined with a column stationary phase different from the first. If retention times for both a standard and unknown match from two distinctly different column stationary phases, then it is highly likely they are identical, but not an absolute identification.

3.3.1.3 Standardized Retention Index Systems

Since there are so many variables in GC, an absolute time measurement is of limited value as column conditions tend to drift with time and use and they are therefore difficult to duplicate. To allow for a more uniform system of reporting retention behavior, systems based on relative retention of the compound of interest compared to a series of standards

have been developed. In this way data from columns of different lengths and carrier gas flow rates could be compared. However, it needs to be emphasized that each system is valid only for the stationary phase from which the data is obtained and cannot be directly used for columns with different stationary phases. One of the first to be developed was one popularized by Ervin Kováts in the late 1950s and was based on a homologous series of linear alkanes [10]. This system ran a series of linear alkanes from C5–C25 where each alkane was given a value equal to the number of carbon atoms times 100. Each new volatile was given a value by interpolating between the two nearest alkane values. However, this system is based on isothermal temperature systems which are rarely employed today because of the wide range of volatilities in most flavor samples. This approach has been modified to allow for the use of linear temperature gradients [11]. The term 'Kovats Index' has been largely supplanted by the term 'retention index' or 'linear retention index' (LRI) for values obtained under thermal gradient conditions. An example is shown in Figure 3.2 in the case of octanol separated on a DB-5 column. In this situation octanol elutes between decane (C10) and undecance (C11). Its LRI value can be calculated using the retention times from the chromatographic report as shown in the general equation:

$$\mathrm{LRI} = 100\left(\frac{Rt(\text{peak of interest}) - Rt(\text{preceding alkane})}{Rt(\text{following alkane}) - Rt(\text{preceding alkane})} + \#C(\text{preceeding alkane})\right)$$

In the case of octanol shown in Figure 3.2:

$$\mathrm{LRI} = 100\left(\frac{9.22 - 7.77}{9.82 - 7.77} + 10\right) = 1071$$

3.3.1.4 GC Injection

GC injections are generally of either solvent or solventless types. Earlier chromatographic studies primarily employed solvent based liquid injections as solvents were typically used to extract the compounds of interest from the matrix and/or then concentrated by carefully removing the solvent to concentrate the volatiles of interest. Injections of liquid extracts have the advantage of knowing exactly how much material is put onto the column (if in the preferred splitless mode). In the case of standards, both internal and external standardization procedures can be

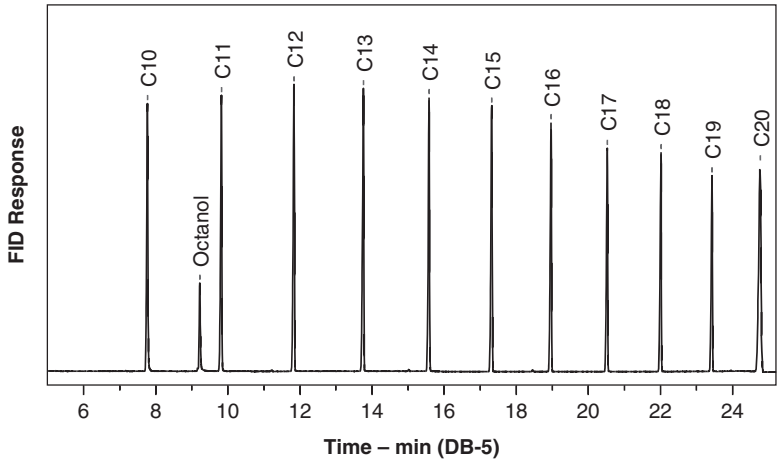


Figure 3.2(a) Mixture of linear alkane standards and standard octanol separated on a 30 m 5 % phenyl methylsiloxane column, where C10 = decane, C11 = undecane, C12 = dodecane, and so on.

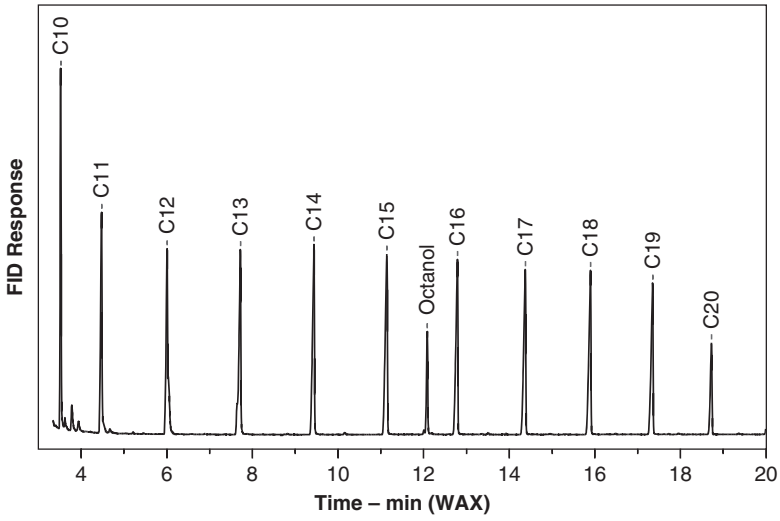


Figure 3.2(b) Homologous series of alkane standards with octanol, separated on a 30 m wax column.

used. Liquid injection is also simpler, more dependable, and less expensive to automate than solventless techniques. The major disadvantage of liquid injection is that it often introduces material onto the column that may be of limited volatility and over many injections will degrade

column performance. One way to minimize this problem is to clean up the extract using either distillation/condensation or classic column chromatography prior to concentration and injection. In addition, the solvent chosen must have a boiling point at least 20 °C lower than the most volatile component of interest. The most commonly employed solvents are either pentane or ethyl ether or a mixture of the two as they have relatively low boiling points and are still liquids at room temperature. A major disadvantage to solvent based injection is the large solvent peak which coelutes with those volatiles that are only slightly retained by the column; rendering analysis of these compounds virtually impossible.

Solventless techniques include fixed volume headspace injection, thermal desorption purge and trap and solid phase micro extraction (SPME) developed by Janusz Pawliszyn [12]. Of these SPME has become increasingly popular as standard GC injectors can be employed and the injection apparatus is simple and relatively inexpensive compared to the alternatives. SPME has been well described in the previous chapter on sample preparation and additional details can be found there. All solventless injection systems allow for the detection of highly volatile components normally blocked out by the solvent peak. The disadvantage of SPME is that the fibers are fragile and easily broken. In addition, quantification can be difficult as the amount sorbed on the fiber is time, temperature, and matrix dependent. Adding known amounts of standards to the matrix is the recommended way to quantify specific volatiles in samples with complex matrices.

3.3.1.5 GC Columns (Stationary Phases)

There are a wide range of column types available for capillary gas chromatographic separations. A few of the most common are listed in Table 3.1. The 5 % phenyl methylpolysiloxane columns are probably the most widely employed phase. More standard retention index values are listed for this stationary phase than any other. It is a rugged column that has slightly better separation power than the previously popular 100 % methylpolysiloxane. Another reason it is widely used is that it produces reasonably reproducible retention index values so that researchers can compare their data with that found in the literature with a fair degree of confidence.

Wax (polyethylene glycol)/Free Fatty Acid columns produce the best separations for oxygenated (more polar) volatiles. Flavor chemists are typically more interested in oxygenated volatiles than hydrocarbon

Table 3.1 Comparison of common stationary phases used in capillary gas chromatography.

Stationary phase	Temperature range (°C)	Recommended uses	Trade names
Methylpolysiloxane	50–325	Nonpolar compounds – low selectivity, volatiles separated by increasing boiling point. Good thermal stability	DB-1, SE-30, OV-1, OV-101, HP-1, RTx-1
5% Phenyl methyl-polysiloxane	50–325	Most widely used phase. Best for nonpolars, similar to methyl polysiloxane, but more selectivity due to phenyl content. Rugged column, good thermal stability	DB-5, SE-54, OV-23, HP-5, RTx-5,
50% Phenyl methyl-polysiloxane	40–325	Best for high boiling flavones, coumarins, and steroids. Additional selectivity due to higher phenyl content. Excellent thermal stability	DB-17, OV-17, HP-17, RTx-17,
Polyethylene glycol	20–260	Best for polar compounds containing oxygen. Susceptible to oxygen degradation.	Carbowax 20M, DB-Wax, Stabilwax, BP-20, HP-20M, AT-Wax
Free fatty acid	20–260	Especially useful for volatile fatty acids and fatty acid methyl esters (FAMEs). Good thermal stability but susceptible to oxygen degradation	OV-351, HP-FFAP, SP-1000, AT-1000,
Porous silica or modified alumina or porous DVB*	−80–300	Best for gases and low molecular weight hydrocarbons	Gas Pro, GC-PLOT, HP-PLOT

Note: DVB* = divinyl benzene polymer.

terpenes as the oxygenated species are generally more aroma active. Unfortunately, the retention index values from these column types tend to be much more dependent on coating thickness, oven gradient rates and injection history than those columns which employ methyl polysiloxane

such as DB-1 or DB-5. Therefore, although these columns are commonly employed, their retention index values cannot be as readily compared with those that can be found in the literature.

Another column type of interest to flavor analysts when it might be necessary to separate highly volatile aroma active volatiles like hydrogen sulfide or methanethiol is the porous layer open tubular column typically called a PLOT column. As indicated in Table 3.1, the material employed for this type of column can vary widely. PLOT columns are columns with thick interior wall coatings of porous material designed to retain highly volatile materials not normally retained by the major column types. The coating consists of particulate materials which offer a range of unique selectivities for volatiles which are gases at room temperature. They are, however, highly retentive and should not be routinely used for separating samples which contain compounds of low volatility. Even though these columns have a high temperature range to allow elution of highly retained material, the chance of contamination or increased background levels due to previous injections is fairly high. These columns are best used for headspace analyses.

An example PLOT separation is shown in Figure 3.3 for a refinery gas sample. Few if any refinery gases are aroma active, but the chromatogram illustrates the ability of the PLOT column to separate highly volatile components. The PLOT column can separate highly volatile hydrogen sulfide, methane thiol and carbon disulfide.

3.3.1.6 GC Detectors

Two basic detector types are commonly employed in chromatographic systems, general mass detectors which respond to most or all volatiles and selective detectors which are specific for a desired subset of volatiles. General mass detectors include the common FID detector and the total ion current (TIC), MS detector (MS-scan) as listed in Table 3.2. General mass detectors are necessary to determine column loading and to help quantify the volatiles present in highest concentrations. Selective detectors are employed when complex samples are analyzed and coelution is a major problem. In these cases the detector should respond only to the compounds of interest and it would not be necessary to chromatographically separate all interferences from the compounds of interest unless the interfering volatiles are in sufficiently high concentration to prevent the selective detector from operating properly. Examples would include fluorescence quenching for optical systems such as the PFPD or

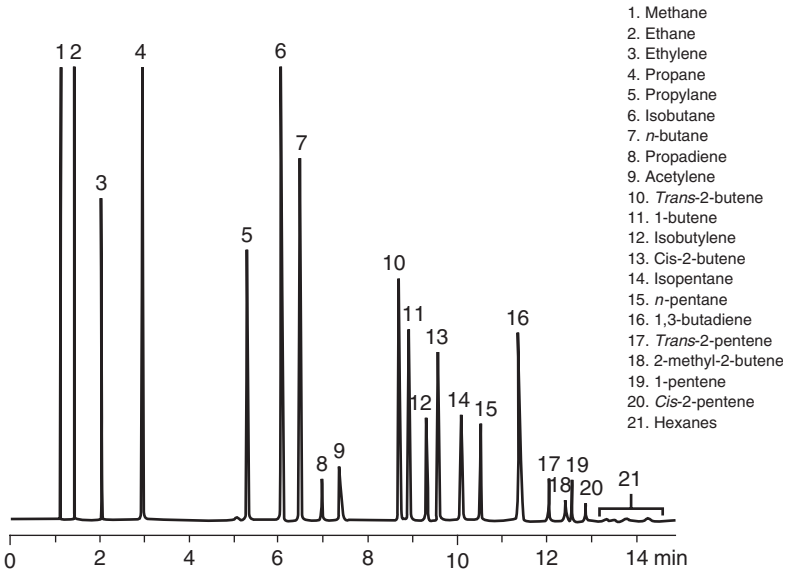


Figure 3.3 Separation of refinery gases using a PLOT column. Column = 50 m, 0.53 mm id, 10 μm Rt®-Alumina BOND/Na_2SO_4, hydrogen carrier gas at 8.0 psi. FID detection. Oven temperature program = 45 °C (hold for 1 min) then to 200 °C at 10 °C/min. Courtesy: Restek.

ion–molecule interactions with MS in the SIM mode. Selective detectors often have greater sensitivity (lower detection limits) than many general mass detectors. The detectors listed in Table 3.2 are far from inclusive but contain the detectors most commonly used in flavor analysis. Even though these detectors are ranked in terms of diminishing linear range, the sensitivity (minimum detection limits) is equally or more important. The human nose is a highly sensitive detector whose detection limits for some volatiles are difficult to match with instruments. In some cases the human nose is more sensitive than the best analytical detectors. Examples include the sulfur containing methional (cooked potato) and the chlorinated phenol, 2,4,6-Trichlorophenol, TCA (musty) among others.

Shown in Figure 3.4 is the comparison of a sulfur specific detector to that of a more general carbon PFPD detector. Three of the more important coffee volatiles are separated and quantified using the sulfur specific detector including the coffee character impact compound 2-furfuryl thiol (A). The upper trace of this chromatogram is specific for sulfur volatiles. The lower trace is a more general detection of carbon volatiles.

Table 3.2 Comparison and characteristics of common GC detectors sorted by linear dynamic range.

Name	Type	Selective for:	Detection limits	Linear range
FID	General mass	Volatiles that ionize in air-H_2 flame	5 pg C/s	$>10^7$
NPD	Selective	N, P	0.4 pg N/s 0.2 pg P/s	10^6
PFPD	Selective	Primarily P,S	0.1 pg S/s	10^5
MS-scan	General mass	Most organic volatiles	10 ng	10^5
MS-SIM	Selective	Specific m/z ions	10 pg	10^5
ECD	Selective	Primarily halogens	~0.1 pg Cl/s	10^4
FPD	Selective	Primarily P,S	20 pg S/s 0.9 pg P/s	10^3
AES	Selective	Most elements	0.1–1 ng element specific	10^3

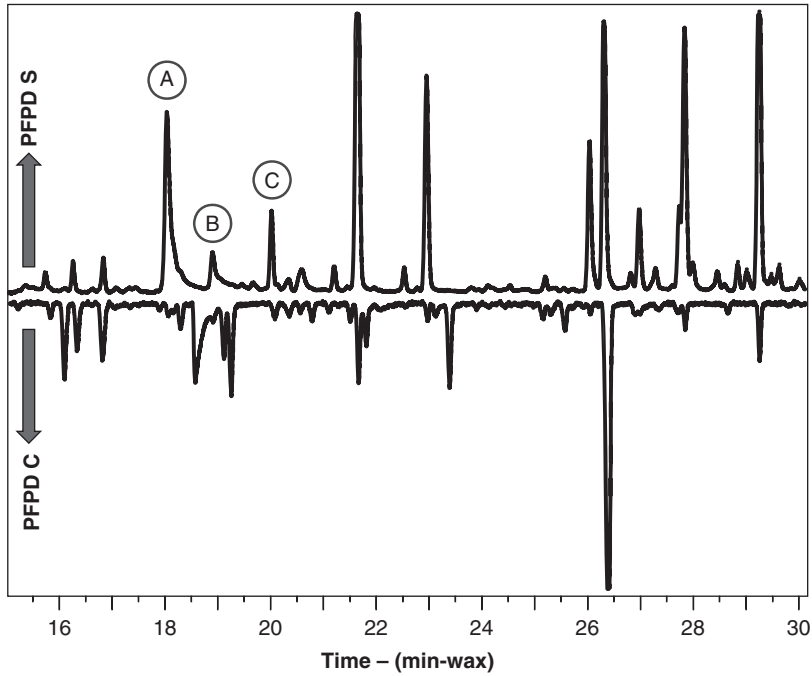


Figure 3.4 Aroma volatiles from ground coffee comparing the response from a sulfur specific detector to that of a carbon specific detector: A = 2-furfuryl thiol, B = methional, C = furfuryl methylsulfide 30 m × 0.32 mm id wax column, oven program = 60–180 °C at 3 °/min and with a 5 min hold at 180 °C.

3.3.2 Identification Techniques

3.3.2.1 Retention Index Approach

The time it takes for each volatile to pass through columns with various chromatographic stationary phases is a characteristic of that substance and the stationary phase. Unfortunately it is not a unique characteristic as many compounds can share very similar or identical retention values. Determining this time as a retention index value does not reduce the number of compounds that share this same value. The general rule of thumb for identification is to have two independent tests confirm the identity of an unknown. Therefore a single retention index match cannot be used for identification. The second piece of independent information would be matching the unknown with a standard on another column of very different chromatographic properties. Typically, the match is made with something like a DB-5 column and a wax column as shown in Figures 3.2(a) and (b) and for the case of octanol. In Figure 3.2(a), octanol elutes between decane (C10) and undecane (C11) on the nonpolar DB-5 column with an LRI value of 1071. When the somewhat polar octanol is chromatographed on a polar wax column, it is more strongly retained and elutes between C15 and C16 with an LRI value of 1565. If the unknown of interest also had an LRI value of something like 1070, it would indicate that the unknown peak could be octanol, but certainly does not prove it, as many other compounds have similar values. However, if the unknown also had a wax LRI of 1565, then it was quite likely to be octanol as few if any other compounds would possess similar values on both dissimilar chromatographic columns.

It is always recommended to match retention index values of unknowns with standards run under identical conditions. However, it is impossible for most flavor analytical chemists to have at their disposal, much less test, the thousands of volatiles known to exist. This is where standard tables of compounds with their LRI values become extremely useful as they can narrow the potential list of standards to be tested. It should be pointed out that the values found in these tables are similar, but rarely agree perfectly. For example, the sample octanol had a value of 1071 on a DB-5 column. This can be compared with values of 1066 and 1078 from the Reading University site, a value of 1070 from the Pherobase site and 1068 for the University of Florida site. Again the DB-5 column value is chosen to make an initial tentative identification as LRI variability for it is much less than that of wax. Standard octanol

Table 3.3 LRI data bases.

Books
Adams, R.P. (1995) *Identification of Essential Oil Components By Gas Chromatography/Mass Spectroscopy*, Allured Publishing Corporation, Carol Stream, IL, USA.
Jennings, W. and Shibamoto, T. (1980) *Quantitative Analysis of Flavor and Fragrance Volatile by Glass Capillary Column Gas Chromatography*, Academic Press, New York, USA.
Sadtler Research Laboratories (1984) *The Sadtler Standard Gas Chromatography Retention Index Library*, Sadtler Research Laboratories, Philadelphia, PA, USA.

Websites	
Name	Location
Reading University – LRI and Odour Database	http://www.odour.org.uk/information.html
Cornell – Flavornet	http://www.flavornet.org/flavornet.html
The Pherobase	http://www.pherobase.com/database/kovats/kovats-index.php
University of Florida – Citrus Database	http://www.crec.ifas.ufl.edu/rouseff

can then be compared on the same wax column for possible identity confirmation.

LRI data bases can also be found in books but these are costly and in many cases out of print or have used chromatographic stationary phases which are no longer commonly used. Data bases are free, readily available, and much faster at locating data. These books are listed as possible references as well in Table 3.3.

3.3.2.2 GC–MS

Gas chromatography-mass spectrometry is an instrumental combination of the separation power of high resolution capillary gas chromatography with the identification power of the mass spectrometer. The combination is a powerful tool in the analysis of complex flavor mixtures and an essential piece of instrumentation in any modern flavor analysis lab. It works on the principle that volatiles are fragmented into ions of predictable size and frequency. The weakest chemical bonds holding the molecule together will be the place at which the molecule is most frequently fragmented and ions form. The most common fragmentation uses electron impact (EI) at about 70 eV. This high energy stream of

electrons fragments the volatiles as they elute from the end of the capillary GC. The ions formed are then focused and sent to a mass analyzer that sorts the ions in terms of their mass to charge ratio, abbreviated m/z. The numbers of ions at each mass are counted and displayed as a bar chart diagram for each mass. The resulting fragmentation pattern is called a mass spectrum and is characteristic of each molecule. The mass spectrum can be compared with other mass spectra either self-determined or from commercial sources such as John Wiley & Sons, Inc. and NIST (National Institute of Standards and Technology). Unfortunately, these libraries were created for environmental, hospital, and forensic labs and the vast majority of their contents are of little interest to flavor analysts. John Wiley & Sons, Inc. has recently introduced a flavor and fragrance library that has addressed this problem. In addition, Allured Publishing has a small essential oil library that is based primarily on ion trap MS.

Modern computer based pattern recognition programs match the spectrum obtained with those found in their data base and lists the identifications of those with the highest match. Misidentifications are still possible as some compounds (especially terpenes) have very similar spectra. Therefore, retention index values are also employed to confirm the identity of the chromatographic peak of interest. The Allured essential oil MS library includes DB-5 retention index (LRI) values for each volatile. Beginning in 2005, NIST libraries have begun to include retention index values for some but not all of the volatiles in its libraries for several of the major chromatographic stationary phases. The matching of MS fragmentation pattern and LRI values with standards are considered solid identification confirmation. If only library spectra and literature LRI values are matched the identification is generally considered tentative.

3.3.2.3 MS/MS

Shown in Figure 3.5 is a portion of a TIC chromatogram from a grapefruit juice extract. It illustrates the complexity and wide concentration ranges found in many flavor samples. Of interest is the small peak labelled 46. This peak is due to vanillin which produces a strong aroma peak (in GC/O) even though its TIC peak is very small and poorly resolved.

Generally, when increased sensitivity and selectivity are necessary, the MS is switched to the selected ion (SIM) mode. Instead of scanning from the customary 25–300 m/z range the MS is set at only one or two masses. In this way a hundred- to thousandfold increase in sensitivity is achieved as the MS collects all the ions from the masses selected instead

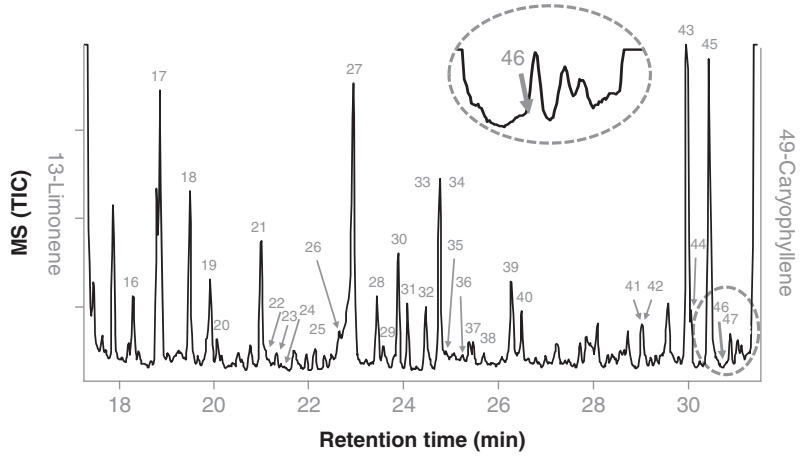


Figure 3.5 GC/MS TIC chromatogram of a grapefruit juice extract in chromatographic region between limonene and caryophyllene. The small shoulder peak labelled 46 is due to vanillin and is magnified in the inset for greater clarity.

of just a few in the brief time ion signals are collected in the scan mode. However, it is critical that the masses selected be characteristic and as unique as possible as the trade off for increased sensitivity is the loss of full spectrum identification. If at least two masses are collected then the ratios of these masses in the unknown can be compared to that of the standard as at least a partial confirmation that the ions collected at the appropriate retention time are only due to those of the compound of interest.

Although not as widely available (due to increased instrument costs) is the use of GC/MS/MS. In this situation the first MS acts as a filter to allow only a single mass (often the molecular ion, M^+) of interest to proceed to the second MS. It is ionized prior to entering the second MS and a full mass spectrum of the compound of interest can be obtained. Shown in Figure 3.6 is an example for the case of vanillin in a grapefruit juice extract. In comparing Figures 3.5 and 3.6, it should be noted that the use of MS/MS detection allows for shorter run times as the chromatographic separation need not be as good because the first MS filters out all but the mass of the compound of interest. In the case of vanillin, the first MS was set at 152 (the molecular ion). Then only the fragments from this mass are recorded in the MS/MS mode.

It can be seen that the baseline drops sharply due to the filtering effect of the first MS. The peak at 8.3 min in Figure 3.6 was confirmed as vanillin by comparing its spectra with that of standard vanillin shown

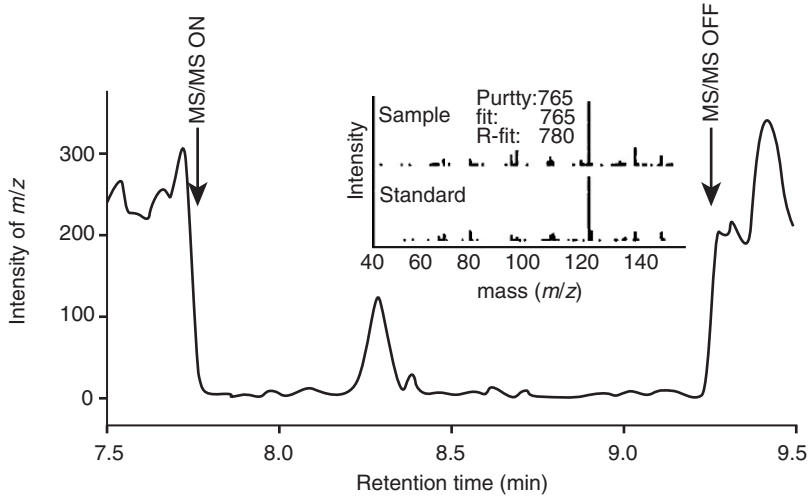


Figure 3.6 Selective MS/MS detection and quantitation of vanillin in a grapefruit juice using m/z 123 which is the major ion in vanillin. From [13].

in the insert. GC/MS/MS is a very powerful separation technique that allows for shorter analysis time with increased sensitivity and selectivity. The only drawbacks are higher instrument costs and MS/MS spectra are sufficiently different from standard GC/MS spectra that all confirming spectra must be self-generated from standards.

REFERENCES

1. Linner, R. T. (1970) Caramel color: a new method of determining its color hue and tinctorial power, in *Proceedings of the Society of Soft Drink Technologists Annual Meeting*, pp. 63–72.
2. Schanda, J. (2007) *Colorimetry*, Wiley & Sons, Inc., New York, USA.
3. Nielsen, D.M. and Nielsen, G.L. (2006) Ground-water sampling, in *Practical Handbook of Environmental Site Characterization and Ground-Water Monitoring* (ed. D. M. Nielsen), Taylor and Francis Group, CRC Press. Boca Raton. USA, pp. 959–1112.
4. Fennema, O. R. (ed.) (1985) *Food Chemistry*, 2nd edn revised and expanded, Marcell Dekker, Inc., New York, USA, pp. 46–50.
5. Scholz, E. (1984) *Karl Fischer titration: determination of water (chemical laboratory practice)*, Springer, Germany.
6. Hecht, E. (1998) *Optics*, 3rd edn, Addison-Wesley, New York, USA.
7. Born, M. and Wolf, E. (2000) *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th edn, Cambridge University Press, Cambridge, UK.

8. Boulton, R., Singleton, V., Bisson, L., and Kunkee, R. (1996) *Principles and Practices of Winemaking*, Chapman & Hall, UK.
9. Symon, K. (1971) *Mechanics*, 3rd edn, Addison-Wesley, New York, USA.
10. Kovats, E. (1958) Gas-chromatographische charakterisierung organischer verbindungen.1. Retentionsindices aliphatischer halogenide, alkohole, aldehyde und ketone. *Helv. Chim. Acta*, **41**, 1915–1932.
11. Vandendool, H. and Kratz, P. D. (1963) A generalization of retention index system including linear temperature programmed gas-liquid partition chromatography. *J. Chromatogr.*, **11**, 463.
12. Arthur, C. L. and Pawliszyn, J. (1990) Solid phase microextraction with thermal desorption using fused silica optical fibers. *Anal. Chem.*, **62**, 2145–2148.
13. Goodner, K. L., Jella, P., and Rouseff, R. L. (2000) Determination of vanillin in orange, grapefruit, tangerine, lemon, and lime juices using gc-olfactometry and gc-ms/ms. *J. Agric. Food Chem.*, **48**, 2882–2886.

4

Gas Chromatography/ Olfactometry (GC/O)

Kanjana Mahattanatawee
Department of Food Technology, Siam University, Thailand

Russell Rouseff
IFAS, Citrus Research and Education Center, University of Florida, USA

4.1 INTRODUCTION

Gas chromatography/olfactometry (GC/O) is a hybrid technique that combines the separating power of gas chromatography (GC) with the specific selectivity and sensitivity of the human nose. Gas chromatography is a relatively mature science and dozens of excellent books have been written about this topic during the last 50 years. This chapter will deal primarily with the olfactometry aspects of GC/O as it would be beyond the scope of this chapter to also discuss the many aspects of gas chromatography. GC/O is a method to determine directly which compounds in a complex mixture of volatiles have aroma activity. Although chromatographers had noted specific aromas eluting during GC runs, it

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

was not until 1964 that it was proposed that a dedicated human evaluate the eluting aroma compounds during a GC separation [1]. Although GC/O can determine which volatiles have aroma activity and which do not, it should be kept in mind that GC/O results do not provide an indication of how these aroma-active compounds will interact with each other or with the food matrix. Preliminary identifications of these aroma-active compounds are usually achieved by comparing retention behavior and sensory descriptors from 2–3 dissimilar capillary GC column types. Tentative identifications can be confirmed by matching the retention times and sensory characteristics of authentic standards with that from the tentatively identified volatile. Whenever possible, mass spectrometry is also employed as an identification confirmation tool. Minimum requirements for the proper identification of GC/O aroma volatiles have recently been published [2] and selected aspects will be discussed in more detail in later sections of this chapter.

Traditional chromatograms of food volatiles employing either FID or MS detection will indicate which volatiles are present in highest concentration, as both FID and MS detectors are general mass detectors. Unfortunately FID or TIC chromatograms usually do not represent the aroma profile of the food because most aroma active compounds exist as potent, low concentration volatiles which produce little FID or MS response while other compounds exist in high concentrations with little or no aroma activity. The human nose is a highly selective and highly sensitive detector which has a theoretical odor detection limit of about 10^{-19} moles [3]. Therefore it is the detector of choice when determining aroma activity. Since the limit of detection of the human nose is lower than most instrumental detectors for many compounds, it is often necessary to concentrate the sample in order to identify specific volatiles using instrumental techniques.

4.2 ODOR ASSESSORS' SELECTION AND TRAINING

Although many of the early GC/O studies were conducted using a single trained odor evaluator (sometimes called a sniffer), two to three trained sniffers are now considered an absolute minimum. Multiple sniffers are needed to compensate for individual threshold differences and to eliminate or at least minimize problems due to specific anosmias. Unfortunately, there is no commonly accepted training procedure for GC/O sniff panel members and there are conflicting reports about the benefits of assessor training [4–6]. Friedrich and coworkers [7] have

suggested that a standard set of 40 odorants would be sufficient to cover all aroma categories. Normally a set of standards containing 10 to 20 of the anticipated aroma-active compounds is used for training. Panelists would typically be expected to evaluate this set of standards repeatedly until they could consistently respond to the aroma compounds of interest. Panelists who cannot perform consistently (reproducibly) or cannot detect certain of the training compounds should not be used. Training normally takes at least 3–6 weeks unless the sniffers have had previous sensory or GC/O experience. However, there are some forms of GC/O such as the frequency of detection method that utilizes 10–12 untrained sniffers. The implicit assumption of this technique is that there will be a normal distribution of aroma thresholds within the panel and they are familiar with whatever technique is used to indicate an aroma response.

4.3 SENSORY VOCABULARY

The description of aroma volatiles as they elute from the GC column by sensory assessors can be a useful piece of information in the identification of individual aroma volatiles. Usually investigators will borrow heavily from the vocabulary established from sensory aroma profiling studies. A panelist's initial description of aroma compounds is based upon the assessor's experience. Compounds are typically described in terms of previously experienced foods or other volatile substances.

There have been limited attempts to standardize aroma vocabulary. It has been complicated by the fact that there are often disagreements in terms of what something smells like because of genetic differences in olfactory receptors. In terms of practical GC/O, two basic approaches have been developed. One is a fixed choice procedure in which a list of descriptors is developed from several initial GC/O runs and all subsequent responses must be selected from that list. The other option is to allow free choice descriptions. This option allows for more flexibility in description and takes into consideration perception variations among sensory assessors. This second option also does not require as much training as a fixed list because it does not require agreement in terms of aroma perceived. However, the interpretation of the final data is more difficult because assessors will use different descriptors for the same aroma. Often the only way to resolve the discrepancy in description is to rerun a standard (providing identification can be made) and ask the assessor if that is the same aroma that they observed in the sample.

With the fixed choice option data interpretation is less ambiguous as the assessor is forced to choose a particular descriptor. This forced choice requirement can sometimes be confusing to assessors when dealing with complex samples and aroma components which elute rapidly and in close proximity to each other. Assessors can sometimes be searching amongst the list of descriptors and miss the next aroma component.

4.4 GC/OLFACTOMETERS (SNIFFERS)

There are a number of olfactometer designs which have been developed to evaluate aroma volatiles as they elute from the end of a GC column. Some units are sold separately (such as the ones from Gerstel (Baltimore, MD, USA), SGE (Austin, TX, USA) and Brechbühler (Houston, TX, USA)). Others are sold only as part of a complete instrument (such as the ones from Datu, Geneva, NY, USA, and Microanalytics, Round Rock, TX, USA). They all share a single feature. That is, they all allow for the provision of humidified air to be added to the effluent stream from the GC column. Some designs also allow for the addition of a neutral make-up gas to change the overall velocity of the gas mixture as it impinges on the assessor's nose. One study reported that as overall gas velocity increased aroma detection accuracy increased [8]. However, this study employed a particular olfactometer design which may or may not be applicable to other olfactometers. The purpose of the humidified air is to cool the heated gases (as much as 250 °C) as they elute from the GC column oven and to prevent dehydration of the assessor's nasal passages during sniffing. The relative amount of this cooled humidified air varies considerably among the different designs. Another difference between olfactometer designs is how close to the nose the humidified air is added to the GC column effluent. In some cases, the humidified air is added directly into the sniffing cone only a few centimeters from the assessor's nose. Other designs add the humidified air more than 50 cm ahead of the assessor's nose. A final and potentially more important design factor is the distance between the end of the column and the assessor's nose. This distance will impact on how much mixing the GC effluent will undergo before it is assessed. In some designs this distance is 2 cm. Other designs use the base of the FID detector to introduce the effluent into the fast moving stream of humidified air in a 1 cm diameter tube about 80 cm long.

An important factor to consider in the use of GC/olfactometers is the relative comfort of the assessor during the sensory evaluation. A general

rule of thumb is to adjust the chromatographic conditions such that the evaluation time would not exceed 30 minutes. Some designs require the assessor to stand or assume positions close to the GC that are uncomfortable or may distract from the assessors ability to focus on aroma evaluation simply because of physical discomfort. Ideally, the assessors should be seated in a comfortable position so that they can be alert and focus their attention entirely towards aroma detection and description.

4.5 PRACTICAL CONSIDERATIONS

There are a few practical factors well known to experienced workers in the area of GC/olfactometry, but not immediately obvious to workers just entering this field. First of all the GC/O unit should be situated in a low traffic area to minimize assessor distraction. It should also be situated away from any food preparation/cooking areas such as microwave ovens or kitchens. The instrument should be located in a separate room equipped with charcoal filter ventilation and at positive pressure compared to the surrounding rooms. Finally, assessors should not use colognes, scented deodorants or hair sprays on the days they will be assessing GC peak odors. The assessors should not have eaten or have drunk a flavored beverage for at least 1 hour before sniffing. Many of these later suggestions are identical with common sensory practices.

4.6 TYPES OF GC/OLFACTOMETRY

Four different GC/O techniques have been developed to identify potent odorants in foods. They include: dilution analysis, time-intensity methods, detection frequency methods, and posterior intensity methods.

4.6.1 Dilution Analysis

This is the oldest and most popular GC/O method. It determines the relative strength of the aroma-active volatiles based on stepwise dilution, usually as a series of 1:2 or 1:3 dilutions. Starting with the most concentrated sample and proceeding to successively more dilute samples, each dilution is sniffed until no odor is detected. However, the diluted samples should be presented in randomized order to avoid bias introduced by knowledge of the samples. The most diluted sample in which an

odor-active compound can detected is called a dilution value or dilution factor value (FD value). For example, if the original aroma extract was diluted as a series of 1:2 (one part in two parts of solvent), and the odor of interest was perceived until the sixth dilution, then the dilution value was $2^5 = 32$. All forms of dilution analysis require the aroma volatiles to be concentrated in some way. Solvent extraction is the most common procedure used to extract and concentrate aroma volatiles for this type of GC/O.

Two slightly different dilution techniques have been developed by research groups in America and Germany: charm analysis by Acree and coworkers [9] and aroma extract dilution analysis (AEDA) by Grosch and coworkers [10, 11]. These methods are based on odor detection thresholds rather than psychological estimations of stimulus intensity at super threshold levels. The difference between the two methods is that charm analysis integrates and sums the dilution values and odor response durations, whereas AEDA determines only maximum dilution value. Charm analysis requires specific software and computer availability where as AEDA can be done with only pen and paper, and is perhaps the reason why it is more widely used.

An AEDA aromagram can be obtained by plotting maximum dilution value on the y axis and retention time (or retention index) on the x axis. An example of an AEDA aromagram is presented in Figure 4.1 demonstrating the aroma components in an extract from freshly prepared popcorn [12]. This method has been employed to determine the potent odor-active compounds in many food products including (but not limited to): grapefruit juice [13], green and black tea [14], cheese [15], and coffee [16].

In charm analysis, when the sniffer detects an odor in the olfactometer air stream, they press the mouse button and hold it for the duration a particular odor is perceived. When the odor is no longer perceived the sniffer releases the mouse button and indicates the odor character from a predetermined list of sensory descriptors. Sniffers must be trained and must have developed and mastered a suitable vocabulary or descriptor list.

Times and durations of the individual sniffs are combined and graphed to yield an aromagram with peaks and integrated peak areas (charm values), which are used to quantify potency. A charm value can be calculated according to the formula $c = d^{n-1}$ where n is the number of coincident responses and d is the dilution factor [17]; Figure 4.2 shows an idealized construction of a charm aromagram. Charm aromagrams are generated when the time-dilution strength blocks produced from all the samples in a dilution series are added together to produce a

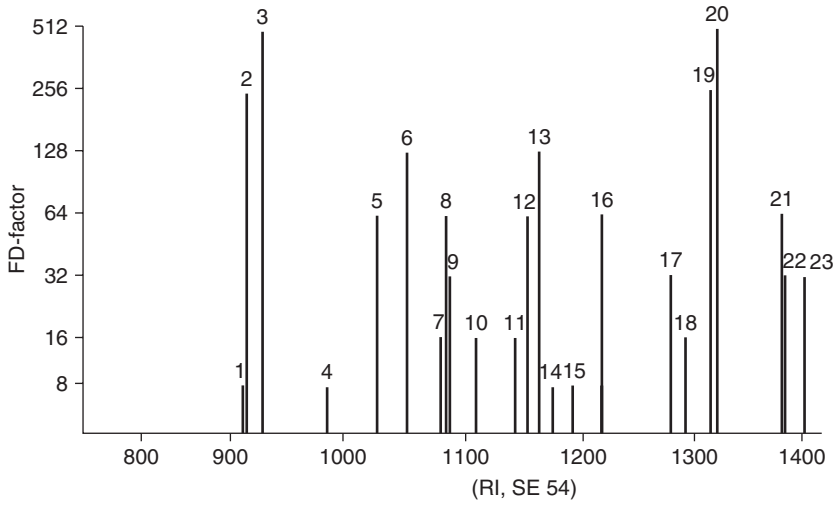


Figure 4.1 Dilution factor (FD) chromatogram of the odorants of an extract from freshly prepared popcorn where 1 = methional, 2 = 2-furfurylthiol, 3 = 2-acetyl-1-pyrroline, 4 = 1-octen-2-one, 5 = 2-propionyl-1-pyrroline, 6 = 2-acetyltetrahydropyridine, 7 = Furaneol, 8 = 2, 5-dimethyl-3-ethylpyrazine, 9 = 2-methoxyphenol, 10 = unknown, 11 = 2-acetyltetrahydropyridine, 12 = (Z)-2-nonenal, 13 = 2, 3-diethyl-5-methylpyrazine, 14 = unknown, 15 = 2, 4-nonadienal, 16 = (E,E)-2,4-nonadienal, 17 = unknown, 18 = (E,Z)-2,4-decadienal, 19 = 4-vinyl-2-methoxyphenol, 20 = (E,E)-2,4-decadienal, 21 = 4, 5-epoxy-(E)-2-decenal, 22 = (E)-β-damascenone, 23 = vanillin [12].

coincident response chromatogram (see Figure 4.2). To make the peak areas proportional to the amount of compound present, the cumulative responses must be transformed using the algorithm shown in Equations (4.1) and (4.2):

$$\mathrm{d}v = F^{n-1}\mathrm{d}i \tag{4.1}$$

$$\text{charm} = \int_{\text{peak}} \mathrm{d}v \tag{4.2}$$

where dv is the dilution value (or number of dilutions made before a particular dilution becomes odorless), F is the dilution factor that was used to prepare the dilution series, n is the number of dilutions in the series that yielded a square block at a particular chromatographic elution time, and di is the time or retention index over which the responses are determined. The software program then combines the data from several sniffs using Equation (4.1) to produce a chromatogram. The

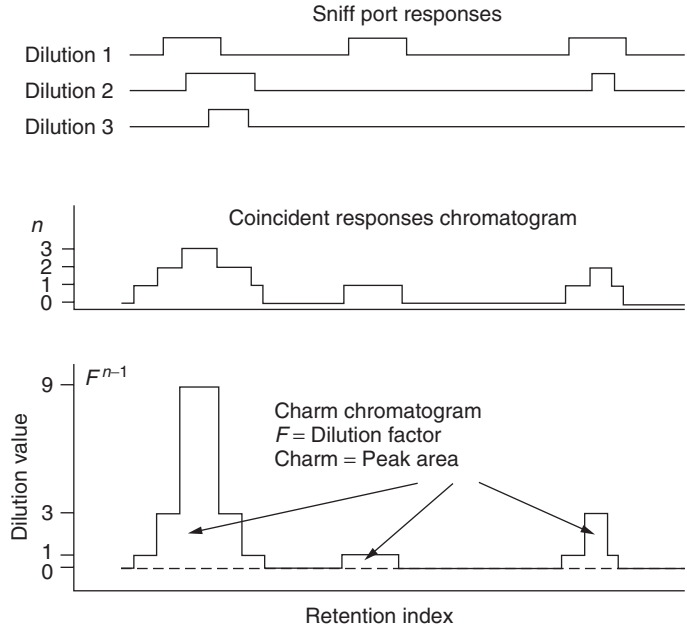


Figure 4.2 Construction of a charm chromatogram [17].

chromatogram is integrated in the region of a peak using Equation (4.2) to yield the charm value for each aroma-active peak.

Charm analysis has been employed for the determination of odor-active compounds in a wide range of foods, beverages, and spices such as: coffee [18, 19], stored boiled potatoes [20], coriander [21], beer [22], and citrus peel oil [23].

4.6.2 Time Intensity

McDaniel and coworkers [24–26] developed a time-intensity GC/O method called OSME (derived from the Greek word for smell) which was based on psychophysical laws. Psychophysical smell behaviors are most commonly represented by Stevens' Law [27, 28]. Stevens established that the odor intensity (I) of a compound increases with its concentration (C) that exceeds its threshold raised to a power n:

$$I = k(C - T)^n \tag{4.3}$$

where T is the compound's threshold value and k is the constant of proportionality. Furthermore, under Stevens' Law, two different

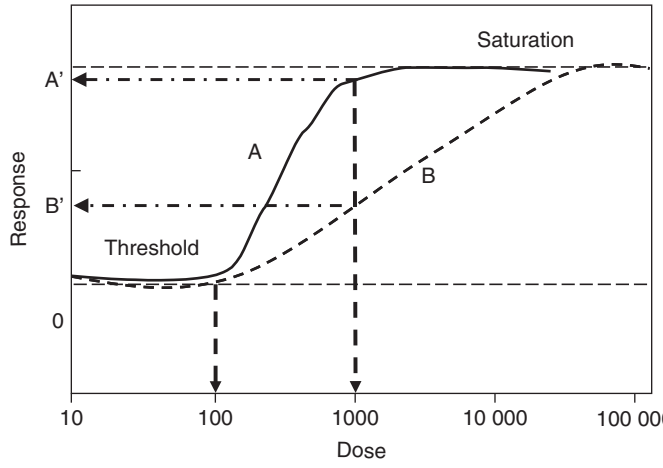


Figure 4.3 Stevens' Law dose response curves for two substances with the same threshold but different dose response behaviors. At 10× concentrations the responses A′ and B′ are appreciably different even though they would have identical odor activity values (OAVs).

compounds at the same concentration (C), and possessing very close threshold values (T), but showing different exponents (n), may provide different individual odor intensities (I) (Figure 4.3) and consequently provide different individual odor contributions to the intensity and quality of a flavor system.

The OSME method differs from charm analysis and AEDA in that only a single concentration of the extract is evaluated. No dilution series are evaluated; hence OSME is not based on odor detection thresholds but on perceived odor intensity. OSME or time–intensity methods are based on estimation of the odor intensity with time for each compound detected at a sniffing port. The panelist moves a variable resistor as a function of the intensity of perception and at the same time describes the odor of the eluting compound. The trained sniffer rates the intensity of an eluting compound using an electronic time–intensity scaling device (15 cm scale, 0 = none, 7 = moderate, and 15 = extreme) coupled with computer data-handling software which provides an FID-style aromagram called an osmegram (Figure 4.4).

In this study [26] a trained panel of four assessors was used to determine relationships between odor intensities and concentration in a series of aroma-active standards. Both the maximum odor intensity of the compounds and the area under the odor intensity peak showed significant correlations with the physical concentration of the compounds in the GC effluent.

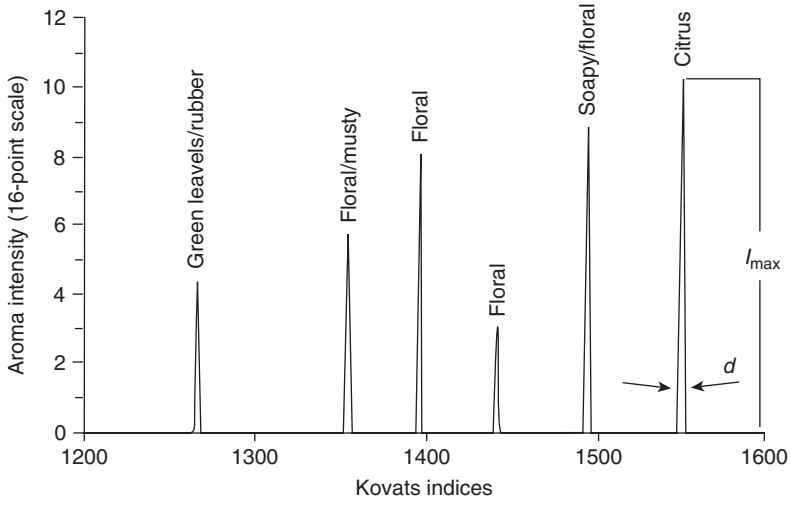


Figure 4.4 Osmegram of aroma active standards [26].

Etievant *et al*. [29] reported a cross-modality matching method with the finger span (GC/O/FSCM) based on the same principle. They described a prototype for the precise measurement and acquisition of the distance between the thumb and another finger during analysis (Figure 4.5). A four-member panel was able to determine intensity

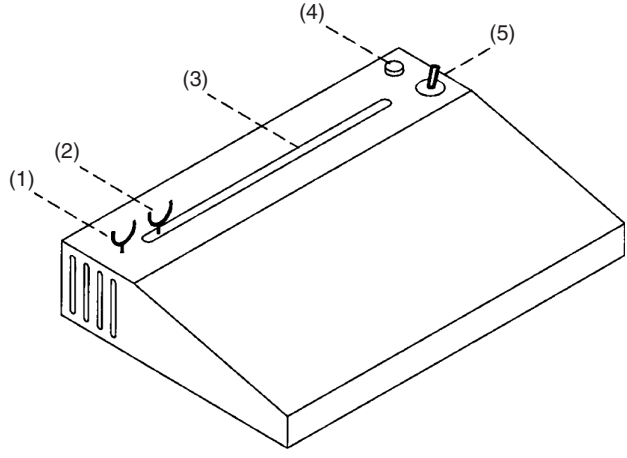


Figure 4.5 Finger span prototype used to measure distance between finger: (1) fixed ring for the thumb; (2) mobile ring for the major or the index finger connected to a 195 mm long rheostat; (3) cursor track; (4) signal lamp; (5) on/off switch [29].

characteristics of 11 standard solutions with good precision. However, individual panelist reproducibility varied considerably. This variation between assessors demonstrated the importance of the using a panel rather than a single assessor to evaluate aroma active materials.

Time intensity GC/O is not as popular as GC/O dilution techniques because it requires additional hardware (intensity transponder) and a second data channel with corresponding software. It has been used (for example) in hop oil [30], blackberries [31], wines [25], and extensively used in citrus aroma [32–35]. Using chromatographic software such as ChromPerfect allows for the rapid visualization of individual FID and GC/O responses. An example is shown in Figure 4.6, where the GC/O response has been inverted in the so called 'fishbone' fashion for improved clarity.

4.6.3 Detection Frequency

The detection frequency method uses the number of assessors who perceived an odor at the sniff port from a single concentration sample. However, a larger group (six to 12) of untrained assessors is employed to determine odor perception, rather than estimating intensity or noting

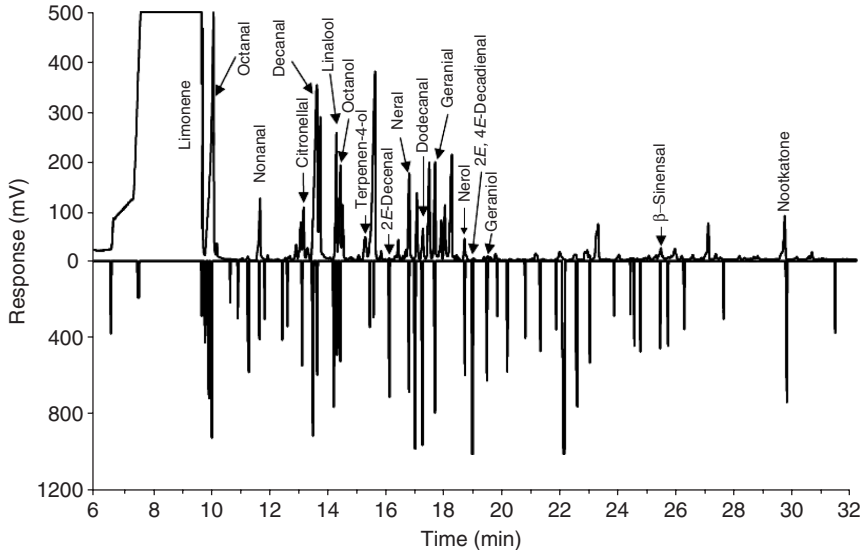


Figure 4.6 GC/FID chromatogram (top) and time-intensity aromagram (inverted, bottom) of a grapefruit oil, separation performed on a DB-Wax column [34].

aroma activity in successive dilutions. The percentage of assessors detecting an odor-active compound was reported to correspond with odor intensity [36–39]. The method was first proposed by Linssen *et al.* [36] using a group of 10 assessors to overcome the limitation of a small number of assessors and only one dilution level was used. The aim of the study was only to identify the volatile compounds that were responsible for the taint found in water packed in LDPE/aluminum/cardboard laminate packaging.

Van Ruth *et al.* [40] evaluated odor released from rehydrated French beans. Ten assessors detected and rated perceived intensities of the eluting compounds at the sniffing port on a nine-point intensity interval scale (1 = extremely weak; 9 = extremely strong). Tenax TA tubes without adsorbed volatile compounds were used as dummy samples for determining the signal-to-noise level of the group of assessors. GC sniffing of dummy samples showed that detection of an odor at the sniffing port by one or two of the 10 assessors can be considered as 'noise'. The number of assessors perceiving an odor correlated significantly with odor intensity scores (Spearman's ranked correlation test, $P < 0.05$) indicating that the number of assessors is a sufficient measure for odor intensity. Furthermore, significant correlations between the number of assessors perceiving odor active compounds correlated well with the intensity scores of attributes in sensory analysis [40–42]. An example of a sniffing chromatogram by detection frequency is shown in Figure 4.7.

Pollien *et al.* [38] have reported a similar technique based on the frequency of detection with six untrained assessors. Elution of odor-active compounds was recorded by pressing a button during the whole sensory impression. The square signal was registered by an HP Pascal workstation. When peak recognitions were needed, the assessor recorded the corresponding odor descriptors on a tape recorder. The six individual aromagrams of a given sample were summed to one chromatogram and normalized with homemade software, yielding an averaged aromagram. Peak heights (percentage detection frequency of an odor by panelists) and areas (expressing the detection frequencies and the detection time duration) are called NIF and SNIF (nasal impact frequency and surface of nasal impact frequency, respectively) (Figure 4.8). It was claimed that the method provided both repeatability and reproducibility. The authors had set an experiment for the number of assessors to establish an aromagram. Details about the procedures may be found in the literature cited above. These results suggest that one or two untrained panelists cannot provide a reliable aroma profile. The first assessor may detect or miss a peak above his/her own detection threshold. Even when the

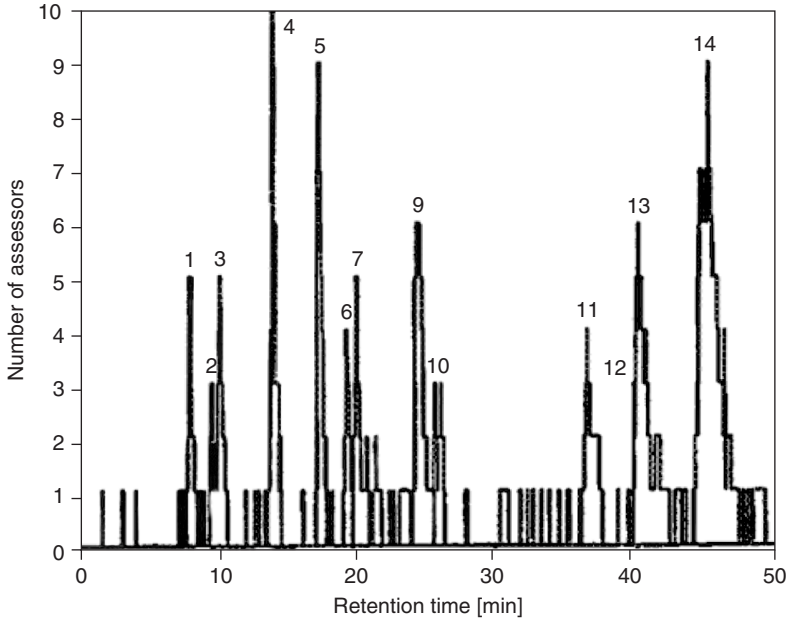


Figure 4.7 Sniffing chromatogram of volatile compounds of diced French beans, rehydrated for 5 min in a closed flask; numbers on the chromatogram refer to identified compounds (data not shown) [40].

second assessor correctly detects the same peak, a 50 % difference from the asymptotic result can be observed. As a result, the reasonable number of untrained assessors should be eight to 10 panelists [38].

There have been subsequent attempts to quantify odorants using the GC/SNIF method where detection frequency is correlated with odorant concentration. It has been suggested that quantification is possible when the odorants have unimodal detection thresholds as a function of the logarithm of their concentration. However, the primary assumption in SNIF quantification is that the panel has a normal distribution of aroma thresholds amongst the assessors for the compounds of interest. This procedure can also be applied to odorants whose concentration is below the sensitivity of physical detectors or when the odorant coelutes in an inseparable mixture of volatiles which have high odor thresholds [39, 43]. The method has been used for analyzing odor active compounds in foods such as: French beans, bell peppers and leeks [44], yogurt [45] mineral water in polyethylene laminated packages [36], black currant [46], and orange juice [39].

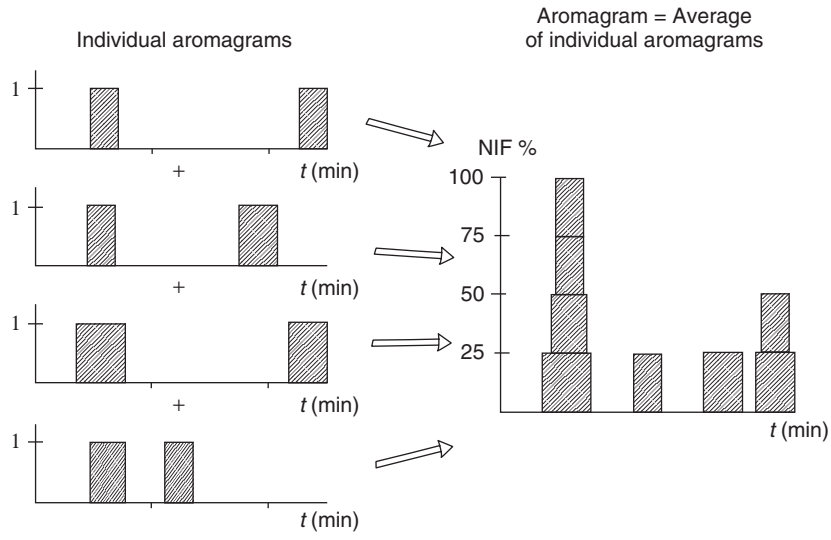


Figure 4.8 Data treatment procedure using the responses from four panelists [38].

4.6.4 Posterior Intensity Method

The posterior intensity method involves the estimation of the odor intensity of a compound in the GC effluent by trained assessors. The perceived intensity is scored on a scale after a peak has eluted from the GC column. The method has been called gas chromatography/sniffing port analysis (GC/SP) [40], odor-port evaluation or GC/sniffing [47], GC odor profiling [48], posterior intensity [49], or perceived intensity method [50]. They are based on the same principle but differ in the number of assessors [47–49] and intensity scale (3, 5, and 9 point intensity [40, 47, 51] respectively). The posterior intensity method is actually a hybrid method employing aspects of time intensity as well as detection frequency procedures. This method is similar to the detection frequency method in that a relatively large group of assessors is used, but differs in that assessors must estimate intensity rather than simply indicating aroma activity. It is similar to OSME/time intensity in that panelists are required to estimate aroma intensity but do not do so continuously, only after an aroma active peak elutes. Therefore, responses can be recorded with pen and paper and a second computer data channel is not required. The disadvantage is that they are also required to record the time a peak elutes (something done automatically when computer software is employed).

To validate the posterior intensity method, Van Ruth *et al.* [49] examined the relationships between perceived GC/O odor intensities with perceived intensities of the same compounds in sensory headspace analysis. The correlation between GC/O posterior intensity scores and sensory odor intensity scores were highly correlated. The method has been applied in studies on Cheddar cheese [47] rehydrated French beans [40], orange juice [48], and wines [52].

4.7 SAMPLE INTRODUCTION

One of the most overlooked and undoubtedly one of the most important aspects of GC/O is the process of collecting and introducing into the GC a representative sample of the product being evaluated. The entire second chapter in this book is devoted to this topic. In the case of dilution analysis, the method of sample preparation will almost always be solvent extraction. Different extracting solvents will produce different aroma profiles because all solvents selectively extract some aroma components more efficiently than others. Therefore the choice of the extracting solvent will greatly influence the aroma components that are introduced into the GC. Traditionally polar solvents such as diethyl ether and ethyl acetate will extract more of the polar aroma components than will nonpolar solvents such as pentane. A serious limitation of this approach is that solvent extractions do not allow for the analysis of highly volatile components as these elute at the same time as the solvent. Another problem with solvent extraction is that it extracts materials that are semivolatile or nonvolatile which may degrade long-term chromatographic performance or produce thermal degradation artifacts in the hot GC injection port.

Headspace analysis is the other option that is commonly employed in GC/O. Classic static headspace analyzes the volatiles in a fixed volume of headspace gas and has been used on a wide range of samples with some degree of success. The primary limitation with this technique is that the headspace often contains a large amount of water and relatively small amounts of the components of interest. Therefore the sensitivity from the direct injection of headspace gas is fairly low. Dynamic headspace can be used in those situations where the interest is primarily in the most volatile compounds in the sample. In passing a fixed amount of the gas over the surface of the sample increased amounts of the more volatile compounds can be trapped and later eluted into the GC.

An increasingly popular method of sample concentration and introduction is the use of solid phase microextraction commonly known

as a SPME. Headspace volatiles are concentrated on the thin layer of absorbent on the silica fiber. The amounts of volatiles collected will be dependent on the temperature of the sample and headspace as well as the exposure time. Different absorbent materials will selectively concentrate headspace volatiles in a manner similar to solvents of different polarity in solvent extraction. Therefore the choice of solid phase in adsorbent will greatly influence the relative amounts of volatiles observed. The guiding principle should always be to choose a fiber that will produce the most representative sample of the headspace gases.

4.8 IDENTIFICATION OF AROMA-ACTIVE PEAKS

One of the factors that differentiate GC/O from the earlier odor-activity value (OAV) approach is that all aroma-active components could be detected and their intensities recorded without actually having an analytical method to measure them. Therefore, the identity of each volatile measured was known. With GC/O, only the aroma intensity or dilution strength is known along with sensory descriptor. Therefore, the major challenge in GC/O is to identify correctly the volatiles responsible for producing the aroma activity. Correct identification of aroma-active peaks can be extremely challenging as a given food product will contain hundreds of volatiles. Correct identification is further complicated by the fact that food volatiles will differ in concentration by several magnitudes. The disadvantage of the GC/O approach is that aroma-active peaks must be identified using a variety of techniques and usually confirmed with some analytical method. Tentative identifications can be based on limited GC/O data. However, firm identifications require sensory and analytical confirmations from at least two independent methods.

4.8.1 Standardized Retention Index Values

One of the primary ways that aroma-active peaks are identified (at least in a preliminary fashion) is by using either the alkane- or ester-based standardized retention index systems. The reporting of retention times is unacceptable in that these values have little value outside of the reporting lab as they are subject to flow rate, oven temperature program, and column length differences. The alkane-based system is more commonly known as the linear retention index or LRI system. In this system a progressive series of linear alkane's is used to standardize all the aroma volatiles. In practice, this requires a sample containing a

series of linear alkanes from C 5 to C 20 (sometimes as high as C 25) to be injected under the same conditions as samples will be run under.

4.8.2 Aroma Description Matching

Tentative identifications can often be achieved by matching the retention values on two columns and aroma descriptions from standardized tables found on the web. However, the lack of a standardized vocabulary makes this process difficult as the same compound may be described in a wide range of terms usually based on prior experience.

4.8.3 MS Identifications

Mass spectrometry (MS) is a mature technology that has been coupled to gas chromatography for identification purposes. Today these instruments are tuned and calibrated under computer control so that they can now be used by any scientists whereas before a dedicated specialist was required to operate them. Identifications are based on the fragmentation patterns of the compounds as they elute from the end of the GC column. The fragmentation patterns are highly reproducible and characteristic of each compound. Libraries containing over 100 000 volatiles can be purchased from NIST or John Wiley & Sons, Inc. Compounds are identified by matching the mass of fragments and relative intensity with those in the library using one of several computer pattern matching programs. This has allowed chromatographers to identify components separated by the GC column rapidly. However, there has been an overreliance on the computer-derived matching identification. Problems occur when several compounds produce similar fragmentation patterns, for example, terpenes. Other problems occur when the fragmentation pattern of the compound of interest is not present in the library. When this happens the computer will identify the compound incorrectly with the closest match it can find in its library. The need to couple MS pattern matching identification with some independent data such as retention index values has been discussed [2]. The most recent version of the NIST library has included retention index values for many compounds to help users select the correct identification from the list of matches proposed by the computer.

The use of MS to identify aroma compounds can be extremely challenging when the aroma volatile is extremely potent and coelutes with an inactive volatile present in much greater concentrations. In this case

the MS will correctly identify the volatile present in highest concentration, but it may not be the aroma-active compound. Often this false identification can be ruled out from aroma descriptions or more reliably from retention characteristics on two or more columns. In the final analysis, overreliance on MS identifications should be avoided and correctly considered as only one piece of information which must be confirmed with additional independent data.

4.8.4 Use of Authentic Standards

After a preliminary identification has been made by matching several pieces of independent data from the literature or internal databases, the identification should be confirmed using authentic standards. The standards should be run under the same conditions as the samples and the data compared. LRI values on polar and nonpolar columns should match within 1 %. Mass spectrometry fragmentation patterns should be essentially identical except for compounds found in low concentration producing signal to noise ratios less than 10:1. The matching of sensory description of the standard compared with that found in the sample at the same retention time is also highly desirable, but in order to evaluate properly the concentration of the standard should be matched closely with that found in the sample.

4.9 CONCLUSION

GC/O is a powerful technique that has become a staple tool in most flavor chemists' laboratory. The results from GC/O should not be overinterpreted. It should be remembered that this technique determines the relative aroma activity and aroma character of component volatiles as individual components. Therefore, the comparison of sensory data, where volatiles are evaluated simultaneously in mixtures, with GC/O data – where components are determined individually – should be done with some caution.

Finally, it should be recognized that each of the many forms of GC/O have their specific strengths and limitations. All GC/O data will be influenced by the manner by which volatiles are extracted and concentrated, and probably represent the largest source of variation. Another factor to consider is how concentrated should the sample extract be. There is always the possibility of overconcentrating sample volatiles. Time-intensity GC–O has the advantage of sensing aroma

volatiles at supra threshold levels but is limited by the wide range that various assessors estimate aroma intensity. This can be minimized by normalizing each assessor's data. Dilution methods are based on threshold information with the assumption that the dose response slopes for all volatiles are identical or very similar. There are many cases where this is not true. On the other hand, the data based on threshold characterization tends to be more reproducible than techniques which require magnitude estimation.

REFERENCES

1. Fuller, G. H., Tisserand, G. A., and Steltenk. R. (1964) Gas chromatograph with human sensor – perfumer model. *Ann. N. Y. Acad. Sci.*, **116**, 711–714.
2. Molyneux, R. J., and Schieberle, P. (2007) Compound identification: a Journal of Agricultural and Food Chemistry perspective. *J. Agric. Food Chem.*, **55**, 4625–4629.
3. Reineccius, G. (1994) *Source Book of Flavors*, 2nd edn, Springer, New York, USA p. 24.
4. Hulin-Bertaud, S., O'Riordan, P. J., and Delahunty, C. M. (2001) Gas chromatography/olfactometry panel training: a time-intensity approach for odor evaluation, in *Food Flavors and Chemistry* (eds A. M. Spanier, F. Shahidi, T. H. Parliment, and C.-T. Ho) Royal Society of Chemistry, London, UK, Vol. 274, pp. 394–403.
5. Callement, G. *et al.* (2001) Odor intensity measurements in gas chromatography-olfactometry using cross modality matching: evaluation of training effects. *ACS Symp. Ser.*, **782**, (Gas Chromatography-Olfactometry), 172–186.
6. Van Ruth, S. M. (2001) Methods for gas chromatography-olfactometry: a review. *Biomolecular Engineering*, **17**, 121–128.
7. Friedrich, J. E., Acree, T. E., and Lavin, E. H. (2001) Selecting standards for gas chromatography-olfactometry. *ACS Symp. Ser.*, **782**, (Gas Chromatography-Olfactometry), 148–155.
8. Hanaoka, K., Sieffermann, J.-M., and Giampaoli, P. (2000) Effects of the sniffing port air makeup in gas chromatography-olfactometry. *J. Agric. Food Chem.*, **48**, 2368–2371.
9. Acree, T. E., Barnard, J., and Cunningham, D. G. (1984) A procedure for the sensory analysis of gas chromatographic effluents. *Food Chem.*, **14**, 273–286.
10. Schieberle, P. and Grosch, W. (1987) Evaluation of the flavor of wheat and rye 111–113.
11. Ullrich, F. and Grosch, W. (1987) Identification of the most intense volatile flavor compounds formed during autoxidation of linoleic acid. *Z. Lebensm.-Unters. Forsch.*, **184**, 277–282.
12. Schieberle, P. (1991) Primary odorants in popcorn. *J. Agric. Food Chem.*, **39**, 1141–1144.
13. Buettner, A. and Schieberle, P. (1999) Characterization of the most odor-active volatiles in fresh, hand-squeezed juice of grapefruit (citrus paradisi macfayden). *J. Agric. Food Chem.*, **47**, 5189–5193.

14. Guth, H. and Grosch, W. (1993) Identification of potent odorants in static headspace samples of green and black tea powders on the basis of aroma extract dilution analysis (AEDA). *Flav. Fragr. J.*, **8**, 173–178.
15. Qian, M. and Reineccius, G. (2003) Static headspace and aroma extract dilution analysis of Parmigiano Reggiano cheese. *J. Food Sci.*, **68**, 794–798.
16. Blank, I., Sen, A., and Grosch, W. (1992) Potent odorants of the roasted powder and brew of Arabica coffee. *Z. Lebensm.-Unters. Forsch.*, **195**, 239–245.
17. Acree, T. E. (1993) Bioassays for flavor, in *Flavor Science: Sensible Prinicples and Techniques* (eds T. E. Acree and R. Teranishi) American Chemical Society, Washington, DC, USA, pp. 1–20.
18. Sagara, Y. *et al.* (2005) Characteristic evaluation for volatile components of soluble coffee depending on freeze-drying conditions. *Drying Technology*, **23**, 2185–2196.
19. Deibler, K. D., Acree, T. E., and Lavin, E. H. (1998) Aroma analysis of coffee brew by gas chromatography-olfactometry, in *Food Flavors: Formation, Analysis, and Packaging Influences* (eds Contis, E. T. *et al.*), Elsevier, New York, USA, Vol. 40, pp. 69–78.
20. Jensen, K., Acree, T. E., and Poll, L. (2000) Identification of odour-active aroma compounds in stored boiled potatoes, in *Frontiers of Flavour Science* (eds P. Schieberle, K.-H. Engel), Deutsche Forschungsanstalt für Lebensmittelchemie: Munich, pp. 65–68.
21. Eyres, G. *et al.* (2005) Identification of character-impact odorants in coriander and wild coriander leaves using gas chromatography-olfactometry (GCO) and comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry (GC ∗ GC-TOFMS). *J. Sep. Sci.*, **28**, 1061–1074.
22. Murakami, A. A. *et al.* (2003) Investigation of beer flavor by gas chromatography-olfactometry. *J. Am. Soc. Brewing Chem.*, **61**, 23–32.
23. Gaffney, B. M., Haverkotte, M., Jacobs, B., and Costa, L. (1996) Charm analysis of two Citrus sinensis peel oil volatiles. *Perfumer & Flavorist*, **21**, 1-2, 4–5.
24. McDaniel, M. R. *et al.* (1990) Pinot noir aroma: a sensory/gas chromatographic approach. *Dev. Food Sci.*, **24**, (Flavors Off-Flavors '89), 23–36.
25. Miranda-Lopez, R., Libbey, L. M., Watson, B. T., and McDaniel, M. R. (1992) Odor analysis of Pinot noir wines from grapes of different maturities by a gas chromatography-olfactometry technique (Osme). *J. Food Sci.*, **57**, 985–993, 1019.
26. da Silva, M. A. A. P., Lundahl, D. S., and McDaniel, M. R. (1994) The capability and psychophysics of Osme: a new GC-olfactometry technique, in *Trends in Food Science* (ed. H. M. a. D. G. V. d. Heij) Amsterdam, Holland Elsevier Science B.V. Vol. 35, pp. 191–209.
27. Stevens, S. S. (1957) On the psychophysical law. *Psychological Review* **64**, 153–181.
28. Stevens, S. S. (1961) To honor Fechner and repeal his law. *Science*, **133**, 80–86.
29. Etievant, P. X. *et al.* (1999) Odor intensity evaluation in gas chromatography-olfactometry by finger span method. *J. Agric. Food Chem.*, **47**, 1673–1680.
30. Sanchez, N. B. *et al.* (1992) Sensory and analytical evaluation of hop oil oxygenated fractions. *Dev. Food Sci.*, **29**, (Food Sci. Hum. Nutr.), 371–402.
31. Klesk, K. and Qian, M. (2003) Preliminary aroma comparison of Marion (Rubus spp. hyb) and Evergreen (R. laciniatus L.) blackberries by dynamic headspace/OSME technique. *J. Food Sci.*, **68**, 697–700.

32. Bazemore, R., Goodner, K., and Rouseff, R. (1999) Volatiles from unpasteurized and excessively heated orange juice analyzed with solid phase microextraction and GC-olfactometry. *J. Food Sci.*, **64**, 800–803.
33. Goodner, K. L., Jella, P., and Rouseff, R. L. (2000) Determination of vanillin in orange, grapefruit, tangerine, lemon, and lime juices using GC-olfactometry and GC-MS/MS. *J. Agric. Food Chem.*, **48**, 2882–2886.
34. Lin, J. and Rouseff, R. L. (2001) Characterization of aroma-impact compounds in cold-pressed grapefruit oil using time-intensity GC-olfactometry and GC-MS. *Flav. Fragr. J.*, **16**, 457–463.
35. Mahattanatawee, K., Rouseff, R., Valim, M. F., and Naim, M. (2005) Identification and aroma impact of norisoprenoids in orange juice. *J. Agric. Food Chem.*, **53**, 393–397.
36. Linssen, J. P. H., Janssens, J. L. G. M., Roozen, J. P., and Posthumus, M. A. (1993) Combined gas chromatography and sniffing port analysis of volatile compounds of mineral water packed in polyethylene laminated packages. *Food Chem.*, **46**, 367–371.
37. vanRuth, S. M., Roozen, J. P., and Cozijinsen, J. L. (1996) Gas chromatography/sniffing port analysis evaluated for aroma release from rehydrated French beans (Phaseolus vulgaris). *Food Chem.*, **56**, 343–346.
38. Pollien, P. *et al.* (1997) Hyphenated headspace-gas chromatography-sniffing technique: screening of impact odorants and quantitative aromagram comparisons. *J. Agric. Food Chem.*, **45**, 2630–2637.
39. Bezman, Y., Rouseff, R. L., and Naim, M. (2001) 2-Methyl-3-furanthiol and methional are possible off-flavors in stored orange juice: aroma-similarity, NIF/SNIF GC-O, and GC analyses. *J. Agric. Food Chem.*, **49**, 5425–5432.
40. Van Ruth, S. M., Roozen, J. P., Hollmann, M. E., and Posthumus, M. A. (1996) Instrumental and sensory analysis of the flavor of French beans (Phaseolus vulgaris) after different rehydration conditions. *Z. Lebensm.-Unters. Forsch.*, **203**, 7–13.
41. Van Ruth, S. M., Roozen, J. P., Cozijnsen, J. L., and Posthumus, M. A. (1995) Volatile compounds of rehydrated French beans, bell peppers and leeks. Part II. Gas chromatography/sniffing port analysis and sensory evaluation. *Food Chem.*, **54**, 1–7.
42. Van Ruth, S. M., Roozen, J. P., and Posthumus, M. A. (1995) Instrumental and sensory evaluation of the flavor of dried French beans (Phaseolus vulgaris) influenced by storage conditions. *J. Sci. Food Agric.* **69**, 393–401.
43. Pollien, P., Fay, L. B., Baumgartner, M., and Chaintreau, A. (1999) First attempt of odorant quantitation using gas chromatography-olfactometry. *Anal. Chem.*, **71**, 5391–5397.
44. Van Ruth, S. M., Roozen, J. P., and Cozijnsen, J. L. (1995) Volatile compounds of rehydrated French beans, bell peppers and leeks. Part 1. Flavor release in the mouth and in three mouth model systems. *Food Chem.*, **53**, 15–22.
45. Ott, A., Fay, L. B., and Chaintreau, A. (1997) Determination and origin of the aroma impact compounds of yogurt flavor. *J. Agric. Food Chem.*, **45**, 850–858.
46. Varming, C., Petersen, M. A., and Poll, L. (2004) Comparison of isolation methods for the determination of important aroma compounds in black currant (Ribes nigrum L.) juice, using nasal impact frequency profiling. *J. Agric. Food Chem.*, **52**, 1647–1652.

47. Arora, G., Cormier, F., and Lee, B. (1995) Analysis of odor-active volatiles in Cheddar cheese headspace by multidimensional GC/MS/sniffing. *J. Agric. Food Chem.*, **43**, 748–752.
48. Tonder, D., Petersen, M. A., Poll, L., and Olsen, C. E. (1998) Discrimination between freshly made and stored reconstituted orange juice using GC odor profiling and aroma values. *Food Chem.*, **61**, 223–229.
49. Van Ruth, S. M. and O'Connor, C. H. (2001) Evaluation of three gas chromatography-olfactometry methods: comparison of odor intensity-concentration relationships of eight volatile compounds with sensory headspace data. *Food Chem.*, **74**, 341–347.
50. Van Ruth, S. M. (2004) Evaluation of two gas chromatography-olfactometry methods: the detection frequency and perceived intensity method. *Journal of Chromatography, A*, **1054**, 33–37.
51. Pet'ka, J., Ferreira, V., and Cacho, J. (2005) Posterior evaluation of odour intensity in gas chromatography-olfactometry: comparison of methods for calculation of panel intensity and their consequences. *Flav. Fragr. J.*, **20**, 278–287.
52. Cullere, L., Escudero, A., Cacho, J., and Ferreira, V. (2004) Gas chromatography-olfactometry and chemical quantitative study of the aroma of six premium quality Spanish aged red wines. *J. Agric. Food Chem.*, **52**, 1653–1660.

5

Multivariate Techniques

Vanessa Kinton
Alcohol and Tobacco Tax and Trade Bureau (TTB), Ammendale, USA

5.1 INTRODUCTION

Nowadays, some commercially available instruments have chemometric software bundled with their data analysis software. Chemometrics is not typically taught in undergraduate chemical degrees and can intimidate the novice analyst, but chemometric analysis can be a powerful tool for analyzing complex data. The International Chemometrics Society defines chemometrics as 'the chemical discipline that uses mathematical and statistical methods, (a) to design or select optimal measurement procedures and experiments; and (b) to provide maximum chemical information by analyzing chemical data' [1]. The usage of chemometrics has increased due to the availability of analytical instruments coupled with fast computers, such as spectrometers and chromatographs that can easily generate megabytes of data. Whereas years ago the computer time required for multivariate analysis of such large data-sets was prohibitive, it is now possible to perform complex statistical analysis and obtain insightful information by applying different multivariate tools in a few minutes.

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

Chemometrics can play an important role in the development of an analytical method by assisting the analyst in selecting what variables of an analytical signal are relevant for measurement. Output from instruments that produce sufficiently complex multidimensional signals is often difficult to interpret without the assistance of automated data processing. Some mathematical manipulations of data (averaging, smoothing, spectral analysis via the Fourier transform, etc.) have become common in the software provided with some instruments, but the use of multivariate statistics for recognizing patterns in data and classification of samples by their spectra is less common.

Quantitative and qualitative analysis of laboratory data have been largely based on univariate data-sets in which a single variable is used for prediction, quantification, or classification. Choosing the specific wavelengths or retention times for an analyte of interest can be a difficult task, especially in the presence of complex matrices. For example, even when the absorption or emission characteristic of an analyte occurs at a specific wavelength, impurities may be present which compromise the signal of interest. Multivariate techniques allow the analyst to treat the entire spectrum or chromatogram as a pattern, regardless of dimensionality, and to model specific variability due to a chosen analyte in the presence of other substances [2].

In some other instances, the analyst may be interested in comparing an adulterated sample to a reference or a new prototype flavor found in nature to a new formulation developed in the laboratory. This comparison may involve multiple analytical techniques and the data can be very complex. This task, however, can be simplified with the use of chemometrics. For example, comparison of multiple chromatograms can be reduced to a single plot. This chapter describes the basics of multivariate analysis and serves as an introduction to chemometric techniques.

In order to perform a multivariate analysis, the analyst should be familiar with some basic notation described in the literature. For example, matrices are usually denoted with upper case bold letters such as:

$$\mathbf{X}_{n\times p} = \begin{bmatrix} x_{11} & \cdots & \cdots & x_{1j} & \cdots & \cdots & x_{1p} \\ \vdots & & & \vdots & & & \vdots \\ x_{i1} & \cdots & \cdots & x_{ij} & \cdots & \cdots & x_{ip} \\ \vdots & & & \vdots & & & \vdots \\ x_{n1} & \cdots & \cdots & x_{nj} & \cdots & \cdots & x_{np} \end{bmatrix} \tag{5.1}$$

where each element x_{ij} represents a single object obtained for sample i at variable j. The objects will always be arranged in n rows, where

n is the total number of objects (e.g., spectra or chromatograms). The variables (measurements) will be in p columns, where p is the total number of variables or features (chromatographic peaks, wavelengths, etc.). The use of spreadsheets, such as Excel, makes it easier for the user to understand the concept of a matrix; for an example of a data matrix see Section 5.6.1, Table 5.1.

The transpose of a matrix will be denoted by $\mathbf{X}'$ and its inverse by $\mathbf{X}^{-1}$. Vectors will be considered column vectors and will be represented with lower case bold letters, for example, $\mathbf{x}$. A row vector is the transpose of the column vector and will be represented by $\mathbf{x}'$.

Table 5.1 Example of a data matrix.

	46	47	48	...	147	148	149
Sample A-lot 60	200428	5256	878	...	90514	45304	29760
Sample D-lot 16	198508	2237	666	...	143773	75155	39880
Sample D-lot 16	171930	685	1077	...	143161	71811	38764
Sample C-lot 25	197901	4310	454	...	15650	2494	517
Sample B-lot 13	233630	4764	659	...	64011	33155	17499
Sample C-lot 26	198790	4503	252	...	15785	2125	587
Sample C-lot 55	198754	3764	330	...	15842	2245	486
Sample A-lot 25	200667	7910	913	...	78827	39446	24862
Sample D-lot 11	163494	164	1160	...	143163	70707	38233
Sample B-lot 07	140608	2731	3187	...	57771	23885	9080
Sample A-lot 56	222919	3833	637	...	90806	46469	29158
Sample B-lot 31	156297	3035	2722	...	58711	25344	10438
Sample C-lot 09	201904	4114	52	...	15984	2460	640
Sample D-lot 12	188802	1727	911	...	143158	74020	39827
Sample D-lot 11	215380	3279	500	...	143770	77364	40943
Sample A-lot 16	202342	7188	886	...	81546	40887	25946
Sample B-lot 20	221278	3825	337	...	61270	30149	15320
Sample C-lot 03	203905	4016	230	...	16151	2443	701
Sample A-lot 34	204017	6466	860	...	84266	42329	27030
Sample B-lot 43	202252	4157	1589	...	62131	30237	14784
Sample B-lot 21	187675	3642	1792	...	60591	28262	13153
Sample B-lot 28	171986	3338	2257	...	59651	26803	11795
Sample C-lot 05	200235	3135	548	...	18463	2233	1011
Sample A-lot 51	205692	5744	833	...	86985	43770	28114
Sample C-lot 15	226108	4546	894	...	12886	2492	480
Sample C-lot 25	210537	3639	348	...	19080	2628	1441
Sample C-lot 16	199902	4212	253	...	15817	2477	578
Sample D-lot 10	180366	1206	994	...	143160	72916	39296
Sample D-lot 12	203134	2790	1069	...	141922	76168	41848
Sample D-lot 11	155058	357	1243	...	143164	69602	37701
Sample A-lot 55	203778	3812	825	...	95953	48187	31928

One of the major concerns in experimental design is choosing an adequate number of samples in relation to the number of variables. Some of the multivariate analyses that are described here, have their roots in social sciences such as psychology or sociology where many objects or subjects are measured with very few variables. This is one of the biggest differences between chemical data where measurements on one sample can result in hundreds of variables; for example, a mass spectrum could contain ion abundance measurements for 300 different m/z values for a single compound, IR spectroscopy might have response measured at 4000 wavelengths, and a chromatographic analysis, run for an hour at 10 Hz, would have 30,000 data points.

Bender *et al.* [3] and Kowalski *et al.* [4] have studied the difficulty of finding the adequate number of samples in relation to the number of variables. They measured the reliability of data-sets by calculating the ratio of samples to variables ($R = n/p$). A ratio of $R > 3$ is adequate for pattern recognition analysis. This ratio is difficult to achieve with hyphenated methods such as GC/MS and a central problem of applying chemometrics to analytical data is that of reducing the number of variables to a more manageable level. Goodner *et al.* [5] recommended using as few variables as possible to develop chemometric models. In their work, they suggest a ratio of 6:1 to 10:1 samples to variables.

Caution should be taken when selecting the samples for any training data-set; the samples should be independent of each other. Particular attention should be placed on obtaining samples that represent the entire population, for example, samples from different lots of samples, samples obtained with different raw materials, samples from different crop years, and so on. Consideration should also be given to collection of sufficient data to use in validating any multivariate models made from the training set.

Once the data has been tabulated as in Equation (5.1), the next step is to examine the raw data and preprocess it. The data should be then examined visually, preferably in both tabular and graphical form. Viewing plots of the raw data may expose unusual values which may derive from transposition errors, equipment malfunction, or simply aberrant samples. Preprocessing is defined as 'any manipulation of the data prior to the primary analysis' [6]. Preprocessing can be performed on the samples (rows of $\mathbf{X}$) or on the variables (columns of $\mathbf{X}$).

Examples of sample-preprocessing techniques include normalization, weighting, smoothing, and baseline corrections. These techniques place samples on the same scale (normalization), increase the influence of some samples over others (weighting), reduce the amount of noise

(smoothing), or reduce systematic variation (baseline correction). For instance, taking second derivatives of spectra removes any baseline offset and facilitates fair comparison of samples [2].

Normalization removes possible variation due to sample size; it is necessary when the detector signal is a function of sample mass as for most gas chromatographs detectors. Normalization to unit area is accomplished by dividing each element in the rows of **X** by the '1-norm' (in chromatography, known as area % normalization):

$$\text{1-norm} = \sum_{j=1}^{p} |x_j| \tag{5.2}$$

Normalization to unit length is accomplished by dividing every element in the rows of **X** by the '2-norm':

$$\text{2-norm} = \sqrt{\sum_{j=1}^{p} x_j^2} \tag{5.3}$$

Johansson *et al.* [7] proposed selective normalization as a way to evade the problems occurring from closure of the data. Sahota and Morgan [8] discuss selective normalization for chromatographic data and demonstrate the effects of normalization on correlated variables.

Variable-preprocessing tools include mean centering and variable weighting. Mean centering is achieved by subtracting the mean of the variable vector from all the columns of **X**. Mean centering accounts for the intercept in a calibration model and is usually recommended prior to performing principal component analysis (PCA). Variable weighting tools include variable selection, variable scaling, and autoscaling. Variable selection is a vast field and many different algorithms have been developed for the selection of optimal discriminating variables; sometimes this is done to maintain an adequate sample to variable ratio. Variable scaling is also used to remove differences in units between variables, which can be accomplished by dividing each element of **X** (x_{ij}) by the standard deviation of that variable (s_j) [2].

Application of both variable-mean-centering and scaling is known as autoscaling. Autoscaling both accounts for the intercept in a calibration model and removes scaling differences between variables. The elements of the autoscaled data matrix **Z** are obtained by:

$$z_{ij} = \frac{x_{ij} - \overline{x}_j}{s_j} \tag{5.4}$$

where

$$\overline{x}_j = \frac{1}{n}\left(\sum_{i=1}^{n} x_{ij}\right) \tag{5.5}$$

is the mean for variable j, and

$$s_j = \left\{ \frac{\left[\sum_{i=1}^{n} (x_{ij} - \overline{x}_j)^2\right]}{(n-1)} \right\}^{\frac{1}{2}} \tag{5.6}$$

is the standard deviation of variable j. Autoscaling is also commonly referred to as 'standardizing' each variable. Once autoscaling has been performed, each variable (column) in the transformed data matrix will have a mean of zero and unit variance.

The selection of which preprocessing technique to use is important because the magnitude of the variance of a variable will determine its importance in any final model, that is, any approximation of the original data set. Sanchez *et al.* [9] investigated the effects of eight different preprocessing techniques prior to PCA on chromatograms and spectra obtained with high-performance liquid chromatography coupled to diode-array detection. The three best results were obtained (a) without any preprocessing, (b) with selective normalization, and (c) with log transformations.

There are many more preprocessing algorithms, some more applicable to specific techniques, such as first and second derivatives for spectrometry and smoothing transforms based on Savitzki–Golay polynomial filter. Something to keep in mind is that transmittance does not vary linearly with concentration; therefore transmittance values should be converted to absorbance if a regression model is needed.

Some software companies make a clear distinction when the preprocessing is applied to the rows (independent variables) or columns (specific to a set of samples) of the data-set, then the order on how they are applied is also defined. It is important for the user of any chemometric software package to be well informed on how the preprocessing is carried out; since there is no universal approach to carrying out preprocessing, different packages do them in different ways.

In summary:

- mean centering shifts the origin without altering inter-sample relationships, usually performed on data prior to a principal component analysis;

- autoscaling involves sequential application of mean centering, then variance scaling. Autoscaled variables have mean 0 and variance of 1.
- derivatives/smooth are based on the Savitzki–Golay polynomial filter; these are ways to remove baseline features from spectroscopic or chromatographic data.
- rule of thumb: transform as little as possible and avoid random transforms.

Pattern recognition has been defined by Albano *et al*. [10] as: 'a methodology for finding rules of classification, that is, given a number of classes, each of which is defined by a set of objects (the training reference sets) and the values of *M* measurements made on each of the objects, rules are defined that make it possible to classify new objects (the test set) on the basis of the same *M* measurements made on these new objects.' Two different categories exist for pattern recognition: unsupervised learning and supervised learning.

Unsupervised learning refers to analysis of data in an exploratory mode. The real identity of the sample is not included in the model development. The goal of unsupervised learning is to find relations and similarities between samples. For instance, given a number of chromatograms one might need to group them according to how similar they are. Examples of unsupervised learning techniques include principal component analysis (PCA) and hierarchical cluster analysis (HCA).

Models developed by *supervised learning* techniques take into account the identity of the samples being investigated. The data matrix is increased with one more column (see Section 5.6.4, Figure 5.6). Optimal models are then built based on misclassification rates of classification algorithms. The lower the misclassification rate, the better the model. Examples of supervised techniques include *k*-nearest neighbors (*k*-NN), soft independent modeling of class analogy (SIMCA), and linear discriminant analysis (LDA).

5.2 HIERARCHICAL CLUSTER ANALYSIS (HCA)

This is an unsupervised technique that displays in a dendogram the distances between samples. A dendogram is a two-dimensional plot which shows the distances between samples (see Section 5.6.3, Figure 5.2). When the distances are small, the samples are similar, when the distances between samples are large, samples are dissimilar. Since there are no pre-existing categories assigned to the samples, this type of analysis emphasizes the natural grouping in the data-set.

The most commonly used multivariate distance is the Euclidean distance, which is defined as:

$$d_{ab}^2 = \sum_j^m (x_{aj} - x_{bj})^2 \tag{5.7}$$

where d_{ab} is the multivariate distance between samples a and b, and x_{aj} and x_{bj} are the sample vectors.

In a HCA the distances between each pair of samples are calculated and samples are linked according to form clusters. The similarity measure is calculated as:

$$\text{similarity}_{ab} = 1 - \frac{d_{ab}}{d_{\max}} \tag{5.8}$$

where $d_{\max}$ is the largest distance in the data-set. Similarity = 1 for identical samples and similarity = 0 for the most dissimilar samples.

There are several ways to link samples: *single link* links to the nearest neighbor (that is, the sample with the smallest distance), *complete link* links to the farthest neighbor and *center* links to the centroid, median, or group average. When working with very different groups, any linkage will work; for poorly separated groups a centroid-based method (for example, group average or incremental) is recommended.

5.3 PRINCIPAL COMPONENT ANALYSIS (PCA)

Pearson first described principal component analysis 100 years ago [11]. At that time, he did not propose a practical method for the solution of problems with more than two variables. Hotelling later expanded Pearson's idea to a more practical method for calculation of principal components in 1933 [12]. Recent advances in technology; specifically data storage, retrieval, and increased processor speed make PCA a practical technique, even for large data-sets.

Principal component analysis (PCA) builds linear multivariate models for complex data-sets [13–15]. PCA facilitates the simultaneous interpretation of the entire data-set by finding linear combinations of variables (or principal components, PCs) that successively account for the maximum variability of the original data-set. One of the goals of PCA is to exclude PCs that correspond to noise and keep only PCs associated with systematic variance, thus reducing the dimensionality of the data.

Normally, PCA is performed on the covariance matrix. For an $\mathbf{X}_{n\times p}$ the covariance matrix, **C**, is defined as:

$$\mathbf{C}_{n\times n} = (\mathbf{X}_{n\times p}\mathbf{X}'_{n\times p})/(n-1) \tag{5.9}$$

PCA is efficiently computed using singular value decomposition (SVD) [16]. The basis of SVD is that a matrix $\mathbf{X}_{n\times p}$ can be expressed as the product of three matrices:

$$\mathbf{X}_{n\times p} = \mathbf{U}_{n\times r}\mathbf{S}_{r\times r}\mathbf{V}'_{r\times p} \tag{5.10}$$

where the matrix **S** is diagonal and the s_k elements are the square roots of the non-zero eigenvalues of the square matrix $\boldsymbol{XX}'$. The columns of **U** are the eigenvectors of $\boldsymbol{XX}'$ and the rows of $\mathbf{V}'$ are the eigenvectors of $\boldsymbol{X}'\boldsymbol{X}$. Wold *et al*. [14] have recommended PCA as 'an initial step to any multivariate analysis to obtain a first look at the structure of the data, to help identify outliers, delineate classes. . .'.

When a PCA is carried out, several diagnostic plots are obtained. These plots provide information on variables, the samples, and the model. The natural grouping between samples is the scores plot. Because PCA concentrates information related to variance in the first principal components, sample similarities and differences are emphasized in plots of the sample positions on these early factors. Sample points in these plots are called scores plots, and these scores plots are key to visualizing the sample relationships.

Variables that are important in the matrix decomposition are more heavily "loaded" in the early principal components, thus these vectors are often referred to as loadings. The amount of variance captured during the decomposition, known as eigenvalues, will decrease as more components are captured. When these are plotted as a function of factor number (called a scree plot), the pattern of eigenvalues can be used to decide the number of PCs, that is, which are relevant (information) and which are irrelevant (noise).

In summary, exploratory analysis, PCA, and HCA provide tools for observing natural differences among samples, for detecting unusual samples (outliers) and discriminating important variables by inspection of their PC loadings.

5.4 CLASSIFICATION MODELS

These techniques are a two step process: build and validate. The goal of this type of analysis is to compare new samples against a previously-analyzed data-set. These types of methods are supervised methods, a class is assigned to each sample in the training set (the training set is used to build the model), then a class is predicted and assigned to samples with unknown classes. There are several types of classification models, but two commonly used are *K*NN and SIMCA.

5.4.1 *k*-Nearest Neighbors (*k*-NN)

This is a classification method based on a distance comparison between samples. Multidimensional distances between samples are calculated, and then the predicted class of unknown samples is determined based on the identity of those samples closest to the unknown. The unknown identity is found by the number of samples closest to the unknown, and these are referred as 'votes'. The number of optimal votes is the '*k*' in the model's name. It is important to keep in mind that the number of maximum votes (*k*) is restricted by the category in a data-set with the fewest number of samples.

This model is based on the multivariate distance, as shown in the HCA section; the Euclidean distance is the distance between objects in a multidimensional space. The unknowns are classified to one and only one class and the quality of the assignment is unspecified. The unknowns are classified based on their proximity to samples already placed in categories. The predicted class of the unknown depends on the class of its *k* nearest neighbors. 'Each of the *k* closest training samples votes once for its class, the unknown is classified with the class with most votes.

The number of neighbors can range from one to one less than the training set. If during the prediction of an unknown sample, two classes received equal votes, the tie is broken by calculating the accumulative distances. The class with the smallest accumulative distance is chosen. Rule of thumb: while running *k*-NN, choose the maximum number of neighbors less than twice the size of the smallest category.

The goodness value is used to measure the quality of the prediction for unknown samples and it is defined as:

$$g_i = \frac{d_i - \overline{d}_q}{sd(d_q)} \tag{5.11}$$

- it is a value similar to the *t* value in statistics; it indicates the number of standard deviations units the unknown is from the average class distance;
- negative values are normal; this means the sample belongs to the class with a high degree of confidence;
- large values indicate that the sample has a poor fit.

5.4.2 Soft Independent Modeling of Class Analogy (SIMCA)

This is another classification method that is based on principal component analysis. The basis of this model is to create principal components

for each separate class present in the data-set, followed by the selection of relevant principal components for each class, better known as 'rank'. Once the rank has been determined for each class, 'multidimensional boxes' are constructed around each class. These boxes then define boundary regions for each class. Just like any other chemometric model, the training set is used to create it and the testing set validates/rejects its prediction ability.

Another important advantage of using SIMCA as a classification model is its ability of determining if a sample does belong to any predefined category but SIMCA also determines if a sample does not belong to any class at all. *k*-NN will predict a class for an unknown sample regardless if the prediction is reasonable or not.

5.5 PRINCIPAL COMPONENT REGRESSION

Prediction of the quantitative composition of mixtures requires calibration experiments that relate the instrumental measurements to the components of interest. The initial step for any calibration requires measurements for reference samples, also commonly referred to as the training or calibration set from which the calibration model is developed. In the second step, measurements for a training sample are obtained and related to the calibration model developed from the calibration data. This prediction step is normally repeated many times with new test samples.

Principal component regression (PCR) [17, 18] is a multivariate calibration technique that consists of two steps. First, PCA is performed on the predictor matrix (**X**); secondly, the number and identity of the PCs to be included in the model is chosen. Selection of PCs based solely on the order of variance typically results in poor PCR models [17, 18]. It is much better to select PCs on the basis of their correlations with **y** (predicted vector). If there is high correlation between a PC and the dependent variable vector **y**, then the PC is a good predictor and should be included in the regression. Regression coefficients are calculated in PCR with:

$$\mathbf{b}_{p\times 1} = \mathbf{V}_{p\times k}\mathbf{S}^{-1}_{k\times k}\mathbf{U}'_{k\times n}\mathbf{y}_{n\times 1} \tag{5.12}$$

To validate a PCR model the data is split into two sets: training and testing. The model is built with the training set, and the testing set is used for prediction. Cross-validation (CV) is a diagnostic approach widely used in chemical applications because of the limited size of data-sets. CV makes use of all the cases and it gives more useful information regarding

the stability of the tree-structure. Parsimonious treatment of the data-set with cross-validation (CV) is used when there are not enough samples for separate training and testing sets. In this case, a portion of the data is left out at a time and the prediction is obtained using a model developed with remaining data. In v-fold cross-validation the original sample is divided at random into v subsets, each containing the same number of cases. When a single sample is left out this approach is also called jackknifing.

CV proceeds as follows: the data is divided into v random portions and then the first portion is reserved for testing. A model is obtained with the remaining v-1 (training set) portions and tested using the first portion (test set) that was left aside. In the next step, the data is rotated and a second portion is reserved for testing. A second model is then obtained excluding the second portion and including the first portion. Then the second portion is used to measure the error rate of the model. This process is repeated until all the portions of the data have been rotated through testing and a cross-validated error of prediction is reported.

5.6 EXAMPLE OF DATA ANALYSIS FOR CLASSIFICATION MODELS

5.6.1 Tabulating Data

The first step to any multivariate analysis is arranging the data in a spreadsheet. This data matrix will have the variables arranged in columns and the samples in rows. If the data set is chromatographic, variables can be area peaks at different retention times, if the data is spectra, then the variables can be absorbance at different wavelengths. If the detector is a mass spectrometer, the data can be arranged as variables being ion counts. It is important to have a fast method that arranges the data in a spreadsheet and some companies specialize in macros that create this data matrix. For this example, a data matrix with four sample types (A, B, C and D) and 105 variables was chosen. Notice that even though, the data has only four samples, the samples have many replicas. An important factor to keep in mind is that the chemometric model will be as good as the data collected. The samples should be independent samples that represent their population variance. For example, samples from different lots, samples collected at different

times, samples using different raw materials. The replicas should not be the same exact sample from a single lot which has been run several times; this will lead to not capturing the variance of the category that is being modeled. Table 5.1 shows the example data.

5.6.2 Examining Data

This step should be done without any preprocessing and with preprocessing. There are several ways of preprocessing a data matrix, on their variables, on the samples or on any combination of both:

- on variables – columns:
  - mean centering,
  - range scale, variance scale, autoscale:
- on samples (transforms) – rows:
  - normalization,
  - weighting,
  - smoothing, and
  - baseline corrections.

Once the data is in a tabulated format, the user must find multivariate software to analyze it and/or create models. For this example, we will use Pirouette [19]. The data can then be imported from a variety of formats, such as Excel, text file, and so on. A preliminary analysis should be carried out that determines if any sample looks extremely different or if there are errors in variables. Examination of the data can be done with plots using two variables (biplots).

It can be seen by examining the bi-plot in Figure 5.1 that there is a sample that appears different than the others. This sample could be an outlier. If the task of comparing bi-plots is too complex (too many variables), you can move into multivariate exploratory analysis.

5.6.3 Multivariate Exploratory Analysis

For this analysis, the samples are explored without a prior assignment of classes. The goal is to determine if they are natural groupings or if there is any correlation or structures among samples and variables. Figure 5.2 shows an HCA dendrogram. In this example, there are four clusters well separated. This indicates that the four categories, A, B, C,

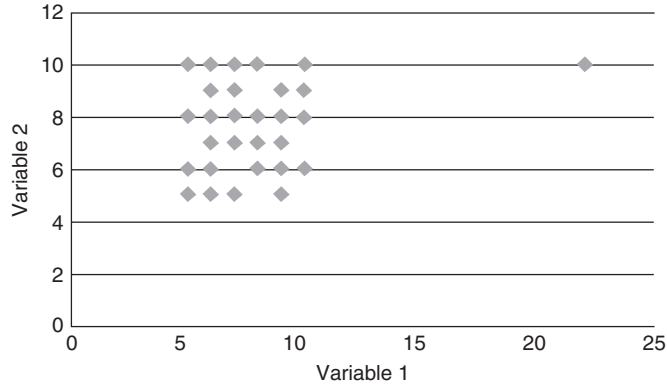


Figure 5.1 Biplot of variables 1 and 2.

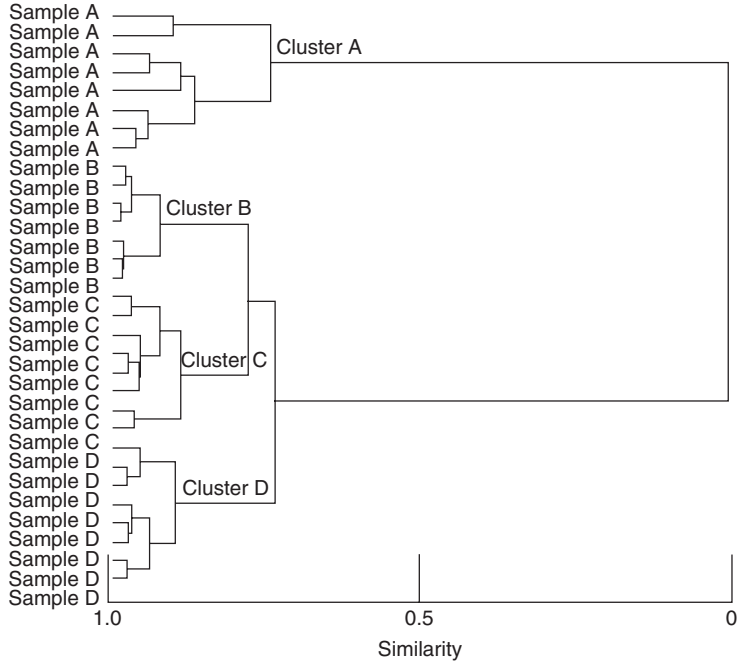


Figure 5.2 HCA of data-set with four categories.

and D have inheriting natural separation and this set could be used for a classification model.

The most commonly used multivariate analysis is PCA. There are a lot of diagnostic plots when this model is created. An example is shown in Figure 5.3 shows a scree plot (see section 5.4). The selection of the

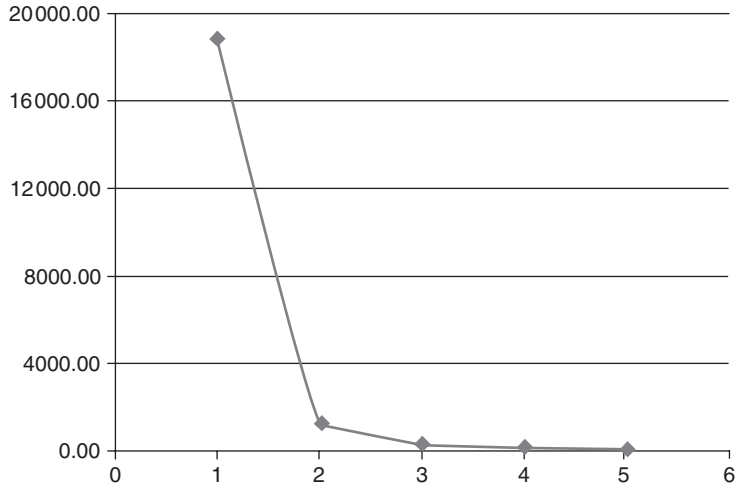


Figure 5.3 PCA scree plot.

number of PCs to keep is important as discussed previously. In this example, it appears that three PCs capture enough variance.

Another PCA plot, and probably the one most commonly recognized, is the scores plot. An example is shown in Figure 5.4. This plot has axes created in the order of amount of variance in the data; each PC is orthogonal to the other. In Figure 5.4, the projection of samples into the first three PCs is displayed. It is important to report which PCs are displayed (in this case, the first three) and the amount of variance captured (in this case 98.65%).

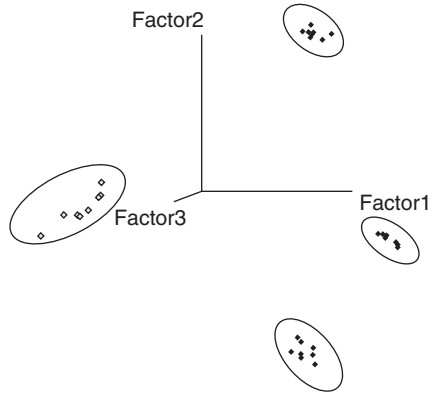


Figure 5.4 PCA scores plots, ellipsoids do not represent confidence regions (98.65% variance captured within the first three principal components).

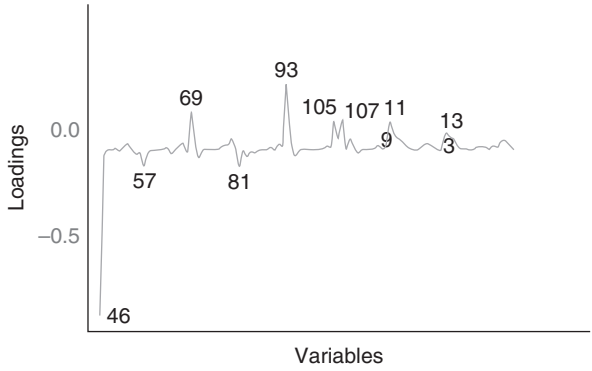


Figure 5.5 PCA Loadings plots.

In some research projects, the chemometric analysis goes as far as showing a scores plot (finding natural grouping). This is a problem if the data is to be used as a classification tool. It can also be a problem if the variance of the population was not properly captured. If for example, all samples of one category were analyzed on a specific day, then a PC may have captured variance due to the day and not due to the sample type. Randomization of the order in which samples are run for a training set is highly recommended to avoid capturing systematic errors.

Another important plot that should be examined while carrying out a PCA analysis is the loadings plot for all the chosen PCs. Figure 5.5 shows the loadings for the first principal component. This plot shows what variables were important for the creation of the first PC. In this figure, variables 46 and 93 appear to be important. If these were ion masses, then maybe a compound with those ions could be differentiating the samples. The loadings plots for PC 2 and PC 3 are not shown here but should also be examined since the scree plot indicated three PCs were adequate to keep for this set.

5.6.4 Creation of a Classification Model with a Training Set and Validation with a Testing Set

For this type of analysis, we need to add one more column to the data matrix: each sample class assignment. The training set becomes as shown in Figure 5.6.

For *k*-NN the number of nearest neighbors (*k*) is chosen with the training set; *k* is made as large as possible but the number of misses is kept small. Once this is done the testing data-set is evaluated. As seen

Variable	46	47	48	···	147	148	149	
Sample C	226108	4546	894	···	12886	2492	480	
Sample A	203778	3812	825	···	95953	48187	31928	
Sample B	202252	4517	1589	···	62131	30237	14784	
Sample C	203134	2790	1069	···	141922	76168	41848	
Sample D	210537	3639	348	···	19080	2628	1441	
Sample A	222919	3833	637	···	90806	46469	29158	
Sample B	233630	4764	659	···	64011	33155	17499	
Sample C	215380	3279	500	···	143770	77364	40943	
Sample D	198754	3764	330	···	15842	2245	486	
Sample A	200428	5256	878	···	90514	45304	29760	
Sample B	221278	3825	337	···	61270	30149	15320	
Sample C	198508	2237	666	···	143773	75155	39880	
Sample D	200235	3135	548	···	18463	2233	1011	

class
3
1
2
3
4
1
2
3
4
1
2
3
4

31 ´ 105 *31 ´ 1*

Figure 5.6 Data matrix before classification models, with class column added.

Table 5.2 *k*-NN classification of unknown samples using goodness value.

	Class 1	Class 2	Class 3	Class 4
Unknown 1	−0.58756	47.93282	61.98627	31.5728
Unknown 2	51.801	0.632427	51.19658	31.9879
Unknown 3	178.6284	169.7851	289.4688	1.003831
Unknown 4	−0.52808	46.51608	63.17764	31.07885
Unknown 5	50.2593	0.347203	56.22467	30.8833
Unknown 6	46.96932	32.74984	−0.3749	36.09097
Unknown 7	163.0246	155.9081	268.1997	−1.22517
Unknown 8	−0.38146	47.05276	67.22506	29.8066
Unknown 9	53.20723	0.563078	47.55336	30.69936
Unknown 10	46.73367	33.31105	−0.11794	35.9984
Unknown 11	160.5863	153.6007	265.2962	−1.92559
Unknown 12	−0.07051	47.53406	67.21384	30.61489
Unknown 13	50.82423	−1.22794	50.84089	29.97689

in Table 5.2, goodness values can be negative. This means the sample belongs to the class with a high degree of confidence; in Table 5.2 the class with the lowest goodness value is underlined. The large goodness values indicate that the sample has a poor fit.

SIMCA models develop PCA models for each category. A multidimensional 'box' is created for each class and the unknown samples are classified according to which box (if any) they belong to. Prediction of unknowns: combination of how far the sample is from (a) the PCA model (i.e., residual), and (b) its projection is from the SIMCA box boundary. The boundaries depend on the number of samples of the training set. (Rule of thumb: have at least 10 independent replicas per class.) These boundaries also depend on the critical value selected

(probability cut-off); $p = 0.95$ will predict differently from $p = 0.999$. The box drawn has straight edges but in reality the region in space is curved due to the nature of the F test.

In order to optimize the model, the optimal number of principal components for each category must be selected. The number of principal components can be different for each class. Class separation can be improved by selecting variables and examining total modeling power and discrimination power. Another important diagnostic to examine for a SIMCA model is the interclass distance. This measurement indicates how well the classes are separated from each other. As a good rule of thumb, interclass distances greater than three are considered well separated. For the samples displayed in Table 5.3 these distances indicate good separation between samples.

SIMCA develops principal component models for each category of the training set. The bounding ellipses in Figure 5.7 form a 95 % confidence interval for the distribution of these categories. In this case,

Table 5.3 Interclass distances for a SIMCA model.

	Interclass distances			
	CS1@4	CS2@2	CS3@2	CS4@2
CS1	0	18.99593	21.8742	17.42387
CS2	18.99593	0	11.42804	23.16223
CS3	21.87419	11.42804	0	22.96419
CS4	17.42387	23.16223	22.9642	0

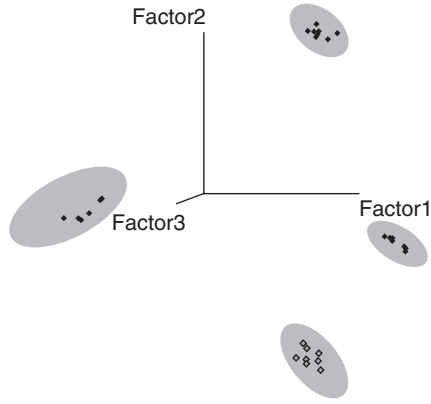


Figure 5.7 SIMCA scores plot, ellipsoids represent 95 % confidence regions.

the projection of the variables indicates good clustering between samples without overlap. Another indication of a good SIMCA model is the interclass distances between samples.

Once classification models have been created and validated, they should be tested periodically. For example, some of the samples used in the creation or validation of the model should be saved. Each time a sequence of unknown samples is run, some of the saved samples should be run alongside. If the saved samples predict accordingly then the model is still valid, if not close inspection to the model should be carried out to update or re-create the model. This chapter represents a small introduction in chemometrics. For comprehensive information regarding multivariate analysis, please see reference 5. Beeve, Pell and Seasholtz [5] advocate six habits for a good chemometrician. These six habits include not only examining data, preprocessing data, estimating models as shown in this chapter but also include validating the model, using the model for prediction and validation of predictions.

REFERENCES

1. Frank, I. E. and Kowalski, B. R. (1982) Chemometrics. *Anal. Chem*., **54**, 232R–243R.
2. Kinton, V. (2001) Chemometric techniques for modeling and classification of composition and identity in multivariate analytical chemical data. Ph. D. Thesis, University of South Carolina, Columbia, SC, USA.
3. Bender, C. F., Shepherd, H. D., and Kowalski, B. R. (1973) Alternate representation of mass spectra for the spectral identification problem (pattern recognition). *Anal. Chem*., **45**, 617–618.
4. Kowalski, B. R. and Bender, C. F. (1972) Pattern recognition. A powerful approach to interpreting chemical data. *J. Am. Chem. Soc*., **94**, 5632–5639.
5. Beebe, K. R., Pell, R. J., and Seasholtz, M. B. (1998) *Chemometrics: A Practical Guide*, John Wiley & Sons, Inc., New York, USA.
6. Goodner, K. L., Dreher, J. G., and Rouseff, R. L. (2001) The dangers of creating false classifications due to noise in electronic nose and similar multivariate analysis. *Sens. Actu. B*, **80**, 261–266.
7. Johansson, E., Wold, S., and Sjödin, K. (1984) Minimizing effects of closure on analytical chemistry. *Anal. Chem*., **56**, 1685–1688.
8. Sahota, R. S. and Morgan, S. L. (1992) Recognition of chemical markers in chromatographic data by an individual feature reliability approach to classification. *Anal. Chem*., **64**, 2383–2392.
9. Sanchez, F. C., Lewi, P. J., and Massart, D. L. (1994) Effect of different preprocessing methods for principal component analysis applied to the composition of mixtures: detection of impurities in HPLC-DAD. *Chemomet. Intell. Lab. Sys*., **25**, 157–177.
10. Albano, C. *et al*. (1978) Four levels of pattern recognition. *Anal. Chim. Acta*, **103**, 429–443.

11. Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. *Phil. Mag*., **2**, 559–572.
12. Hotelling, H. (1933) Analysis of a complex of statistical variables into principal components. *J. Ed. Psych*., **24**, 417–441 and 498–520.
13. Jackson, J. E. (1980) Principal components and factor analysis: Part I-principal components. *J. Qual. Tech*., **12**, 201–213.
14. Wold, S., Esbensen, K., and Geladi, P. (1987) Principal component analysis. *Chemomet. Intell. Lab. Sys*., **2**, 37–52.
15. Jollife, I. T. (1986) *Principal Component Analysis*, Springer Verlag, New York, USA.
16. Press, W. H., Flannery, R. P., Teukolsky, S. A., and Vetterling, W. T. (1986) In *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, pp 60–72.
17. Egan, W. J., Brewer, W. E., and Morgan, S. L. (1999) Measurement of carboxyhemoglobin in forensic blood samples using UV-visible spectrometry and improved principal component regression, *Appl. Spectrosc*., **53**, 218–225.
18. Martens H. and Næs, T. (1989) *Multivariate Calibration*, John Wiley & Sons, Inc., New York, USA.
19. Pirouette (2007) Version 3.12, Infometrix, Seattle, USA.

6

Electronic Nose Technology and Applications

Marion Bonnefille
Alpha MOS, Toulouse, France

6.1 INTRODUCTION

Over the last decade, 'electronic sensing' – or 'e-sensing' – technologies have undergone important developments from a technical and commercial point of view. The expression 'electronic sensing' refers to the capability of reproducing human senses using sensor arrays and pattern recognition systems. The first sense to be reproduced by an instrument was hearing and devices that could be called 'electronic ears' were developed for industrial purposes. In recent years, 'electronic eyes' have also experienced a number of developments for biometric applications such as iris recognition for safety purposes or for industrial routine quality control.

For the last 15 years, research has been conducted to develop technologies, commonly referred to as electronic noses, which could detect and recognize odors and flavors. The stages of the recognition process are similar to human olfaction and are performed for identification, comparison, quantification, and other applications. However, hedonic

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

evaluation is a specificity of the human nose given that it is related to subjective opinions. These electronic nose devices have undergone much development and are now used to fulfill industrial needs in research and development departments, quality control for flavor and fragrance, food and beverages, packaging, pharmaceutical, cosmetic and perfumes, and chemical companies. More recently they can also address public concerns in terms of olfactive nuisance monitoring with networks of on-field devices.

Most electronic noses use sensor-arrays that react to volatile compounds on contact: the adsorption of volatile compounds on the sensor surface causes a physical change of the sensor. A specific response is recorded by the electronic interface transforming the signal into a digital value. Recorded data are then computed based on statistical models. The more commonly used sensors include metal oxide semiconductors (MOS), conducting polymers (CP), quartz crystal microbalance, surface acoustic wave (SAW), and field effect transistors (MOSFET).

Recently, other types of electronic noses have been developed that utilize mass spectrometry or ultra fast gas chromatography. However, this chapter will only focus on sensor-array based technologies and will present:

- sampling systems,
- detection technologies,
- data treatment tools,
- application range,
- case studies.

6.2 HUMAN SMELL AND ELECTRONIC NOSES

The sense of smell in humans is one of the most important senses and the one that historically is useful for identification and protection. The human process of olfaction includes:

- the perception of a flavor by odor receptors,
- the conversion of the perception into a signal,
- the transduction of the signal to the limbic system of the brain, that is, the area in which emotional responses occur and which is associated both with memory and with physiological responses independent of conscious decisions, and
- passing the signal on to the neocortex for labelling (recognition, etc.).

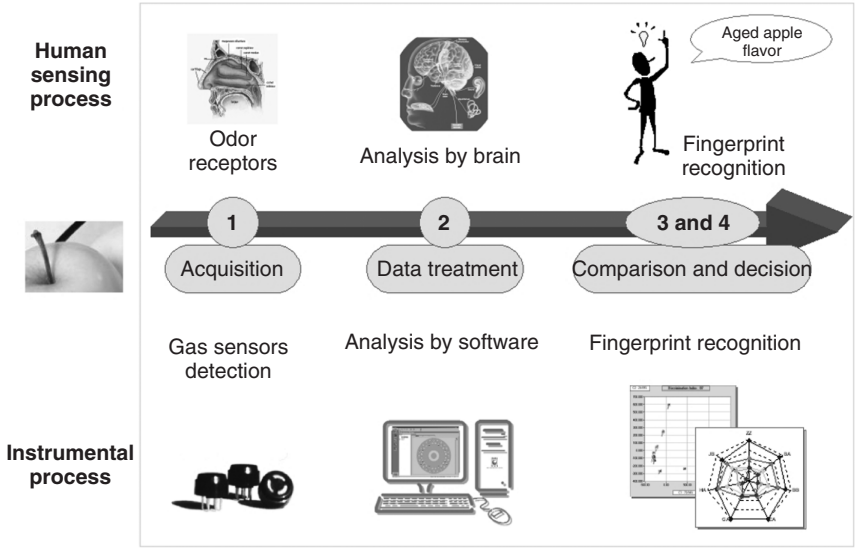


Figure 6.1 Comparison of human sensing and instrumental sensing of aroma volatiles.

The electronic nose was developed in order to mimic human olfaction which functions in an unseparated mode: that is, an odor/flavor is perceived as a global fingerprint (see Figure 6.1).

6.3 TECHNIQUES TO ANALYZE ODORS/FLAVORS

In industry, aroma assessment is usually performed by human sensory analysis or by gas chromatography (GC, GC/MS). The latter technique gives information about volatile organic compounds but the correlation between analytical results and global odor perception is not direct due to potential interactions between several odorous components. This is why electronic noses are being used more widely in industrial applications.

6.3.1 Sensory Panel

Companies look to non-trained sensory panels for subjective evaluations (for instance consumer tests). To evaluate the conformity of a manufacture's product, or to assess the quality or origin, they often have to constantly train panelists and establish standardized quotation scales or criteria. Trained sensory panels can detect aroma differences in a sample,

from major differences to minor ones. Sensory evaluation is certainly the most widespread among odor/flavor assessment techniques. More and more sensory panel analyses are being allocated to new product development (research and development) or to starting (as a reference) and ending (end-user evaluation) points of a production process.

The two chief advantages of human sensory panels are the human sensitivity to some molecules (concentrations in the sub-ppb range) and the selectivity of the human nose. It is the most common method used by consumers for the final assessment of a product. However, sensory panels present some major drawbacks for industrial purposes:

- limited availability,
- time consumption and costs,
- human fatigue which limits the number of analyses per day,
- related subjectivity and variability entailing variability in scientific evaluation,
- reluctance in risking the testing of unpleasant or noxious products.

6.3.2 GC and GC/MS

Analytical techniques such as gas chromatography (GC) and mass spectroscopy (MS) are also commonly used by chemists to identify compounds and their concentrations in an odor/flavor. Given that they are separation techniques, the results are not directly linked to information produced by human sensory panels and this correlation is difficult to achieve. They are perfectly suitable for the analysis of a single pure molecule but it is much more difficult to interpret results for odor/flavors that consist of complex mixtures of compounds.

6.3.3 GC/Olfactometry

A method combining GC with human evaluation (GC/olfactometry) is sometimes used as a complement: for example, an analyst can smell individual compounds of a mixture, after separation by a GC-column. By assessing the type and intensity of each component, this method gives detailed information on the composition of an odor. Nevertheless, it is not a global analysis that will enable comparison between various flavors. In addition it is a very lengthy technique to implement and therefore it is not suitable for routine quality control tests.

6.3.4 Electronic Nose

As a first step, an electronic nose needs to be trained with qualified samples so as to build a database of reference. Then the instrument can recognize new samples by comparing volatile compound fingerprints to those contained in its database. Thus they can perform qualitative or quantitative analysis.

The electronic nose presents various advantages for flavor measurement:

- no or very little sample preparation,
- consistency and reproducibility of the measurement,
- rapidity of obtaining results,
- high throughput of analyses,
- permanent availability,
- nondestructive and global analysis of an odor/aroma,
- fingerprint results as human assessment,
- results correlated with human perception due to multivariate data treatment.

These advantages explain why electronic noses are used not only in research and development laboratories (as a fast and extensive screening technique) but also at the production stage (for rapid quality control).

In the flavor and fragrance industry and related areas, the electronic nose proves to be a powerful tool which broadens aroma analysis capabilities. However, the electronic nose cannot replace the human nose for subjective analysis (for example, consumer preference tests or qualification of a new product never analyzed before). More generally, electronic nose analyzers can be applied to all types of volatile compounds, for example, to monitor batch to batch variation, raw material variability, and samples that are in conformity/out of specification for whatever reason.

6.3.5 Electronic Nose Technology and Instrumentation

6.3.5.1 Architecture

Gardner and Bartlett (1993) defined the electronic nose as 'an instrument including a set of electronic chemical sensors with a cross selectivity,

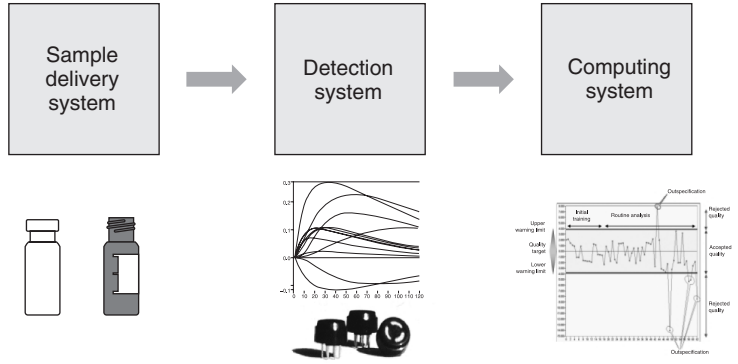


Figure 6.2 Block diagram of components of the process involved in electronic nose operation.

and a fitted pattern recognition system capable of recognizing simple or complex odors'. Figure 6.2 illustrates the three major parts that comprise an electronic nose:

- a sample delivery system,
- a detection system,
- a computing system.

Sample delivery system This enables the generation of the headspace (volatile compounds) of a sample, which is the fraction analyzed. The system then injects this headspace into the detection system of the electronic nose. The sample delivery system is essential to guarantee constant operating conditions.

Detection system This consists of a sensor set, which is the 'reactive' part of the instrument. When in contact with volatile compounds, the sensors react, which means they experience a change of electrical properties. Each sensor is sensitive to all volatile molecules but each in their specific way.

Computing system This works to combine the responses of all of the sensors, which represents the input for the data treatment. This part of the instrument performs global fingerprint analysis and provides results and representations that can be easily interpreted. Moreover, the electronic nose results can be correlated to those obtained from other techniques (sensory panel, GC, GC/MS).

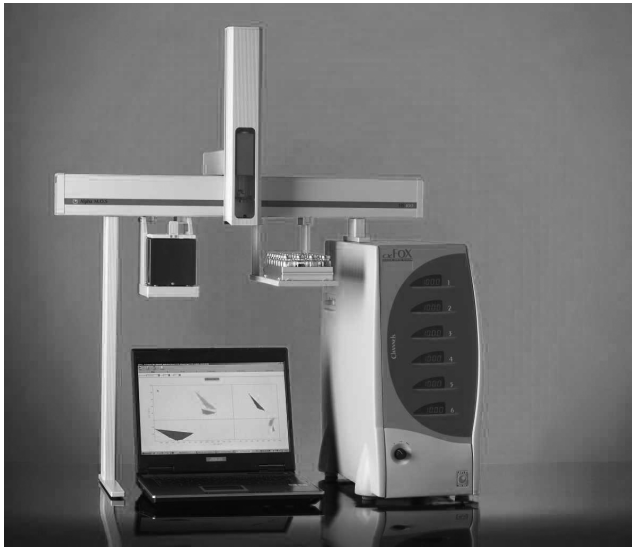


Figure 6.3 Photo of the Alpha MOS Fox electronic nose with autosampler. *Source*: Alpha MOS, France.

Figure 6.3 represents a typical system commercially available on the market and consisting of: an autosampler, a sensor array based electronic nose including 18 metal oxide sensors (the FOX from Alpha MOS, France) and a computer.

6.3.5.2 Air Generator

In the early days of electronic noses, ambient air changes during the day rapidly generated variations in sensor response over time. To overcome these variations linked to external conditions, electronic noses are now connected to a source of pure and constant air.

In order to operate properly and in optimum conditions, it is recommended that electronic nose instruments are fuelled with a constant and high quality pure air as the carrier gas with the specifications:

- $H_2O < 5$ ppm
- $C_nH_m < 5$ ppm
- $O_2 + N_2 > 99.95\ \%$
- $O_2 = 20\ \% +/- 1\ \%$.

The source of pure air consists of a total organic compounds (TOC) generator which produces purified air from compressed air (from an air

compressor or compressed air delivery network). The air generator uses a combination of filtration, combustion, and pressure swing adsorption to remove hydrocarbons, CO_2, and water from the compressed air.

6.3.5.3 Sampling

Electronic nose instruments are used for volatile organic compound and odor analyses. Sample preparation with an aim to generate headspace remains a key parameter to provide consistent and accurate results. Since the 1990s, several new headspace sampling techniques have been developed to adapt analytical instrument capabilities to all kinds of samples and extend their performance in terms of detection limits.

Automated or semi-automated sampling devices At first, analyses were run by presenting the sample at the 'entrance' to the electronic nose. Since the reproducibility of the analyses was not conclusive, several developments have been implemented. For example Alpha MOS Company introduced a 'measurement chamber' to isolate the sample in a vial with air flowing directly on the content through valves. The vial was heated to generate a headspace, and air was sent to the sample manually after triggering data recording. In order to improve synchronization of air flow and data recording, a four-way valve was used. In the meantime, a mass flow controller (MFC) was installed to monitor a constant gas vector flow.

Between the end of 1995 and the beginning of 1996, autosamplers were connected to the instruments. An autosampler consists of:

- a tray to store several vials containing samples,
- a multiposition oven with orbital stirring, which proves to be efficient to guarantee homogeneity,
- a heated syringe block (various volumes available).

The multiposition oven allows the heating and stirring of several samples simultaneously so as to overlap preparation times and thus reduce the overall analysis sequence time. At the present time the autosampler and its operating parameters can be piloted directly from the electronic nose software.

The use of autosamplers provides automation and facilitates analyses while allowing a better understanding of issues related to sampling. Moreover, these devices guarantee the reproducibility of: heating/stirring times, stirring speed, heating accuracy, and injected volume.

Instead of the autosampler, other sampling devices with a lower degree of automation were developed to meet the needs of users who perform fewer analyses and who do not require a high throughput:

- the 2T station allows the automation of the heating time of various samples simultaneously and ensures reproducible conditions (time, temperature);
- the Matrix system also performs automated heating of samples and further enables semiautomated SPME extraction of headspace.

SPME Among the different headspace sampling methods, solid phase microextraction (SPME) is very innovative and completely fulfills electronic nose requirements. In SPME, a coated fused silica fiber is introduced into the sample, and volatile organic compounds (VOCs) are adsorbed onto the coating. Then the analytes are desorbed from the fiber to the analyzer by a heated injection port. By changing the coating type or thickness, the selectivity can be altered in favor of more or less volatility. This property can be monitored to selectively adsorb sample components, as desired.

Headspace SPME is ideal for:

- minimizing the matrix background during an analysis,
- revealing additional VOCs that have been obscured in a direct headspace technique,
- improving the quantity of compounds to be analyzed and thus increasing the sensitivity, and
- increasing the selectivity in headspace sampling.

Thermodesorption devices Thermodesorption devices are aimed at performing enrichment of volatile or semivolatile compounds during headspace analysis. Their working principle is based on dynamic extraction of headspace and automated thermodesorption in the instrument. The sample headspace is repeatedly pumped by the syringe. Sensitivity can be improved by increasing the number of pumping strokes. During thermal desorption into the instrument, the adsorbing phase is rapidly flash heated in the injection port.

Among recent sample concentration techniques, the solid phase dynamic extraction (SPDE) from Chromtech, Germany, consists of a syringe comprising a needle with an inner coating of adsorbing material. Another device commercially available, the ITEX from CTC, Switzerland) uses a syringe including a microtrap (Tenax or activated charcoal) placed between the syringe and the needle.

A dynamic thermodesorption device, the TDAS-2000 from Chromtech, Germany) was developed for fully automated thermal desorption to analyze organic compounds that vaporize or semivaporize in solid, liquid, or gaseous samples. It assumes that the volatiles have been previously trapped (during a dynamic extraction called purge) on a sorbent in a tube. The desorption is performed at high temperature. This technique allows the concentration of medium and low volatile compounds and their detection with the analytical system. This dynamic extraction is more efficient than a static one. Moreover this module can be mounted on commercially available autosamplers (such as CTC) so that the transport of the sample is over short distances, which yields results which are highly reproducible and even more sensitive. There are no critical parts involved, such as heated transfer lines and/or switching valves, so errors related to sample contamination and low reproducibility are eliminated.

Other sampling options Specific modules can be used to achieve appropriate sampling:

- *For temperature-sensitive samples* that need to be stored at a low temperature, a Peltier cooler can be mounted on the autosampler to cool down samples on their tray.
- *For on-line analysis* a flow cell can be used to aspirate the sample headspace at chosen intervals or to spike a liquid or gas stream at certain time intervals with a reagent or standard. The main applications concern on-line monitoring of processes (fermentation, chemical reactions, a pilot plant, and cooking processes) and environment analysis (waste or drinking water lines).
- *To automate the weighing-in of samples* the Balance Pal option automatically weighs samples and transfers them to the detector, while transmitting the weight data to the computer.
- *For large volume samples* that cannot be introduced into a standard vial (10–20 mL) the Baker oven from Chromtech, Germany) can contain samples up to 750 mL. This module is designed to generate headspace from large volume samples.
- *Measurement in microbial applications*: Petri dishes with a specific sealing system were developed to directly sample and analyze volatile compounds emanating from cultures. This module can be directly fixed on the headspace sampler. Petri dish sampling can be temperature controlled using a cooling device. Once the inoculation of organisms is performed in the Petri dish, the growth rate can be directly monitored via the electronic nose measurement.

6.3.5.4 Detection Technologies

Metal oxide sensors (MOS) As shown in Figure 6.4, metal oxide semiconducting sensors are generally made of:

- a ceramic substrate,
- metal electrodes to measure conductance,
- a heating element (wire) to activate reactions and to allow the elimination of contaminants on the sensor surface, and
- a coating consisting of a semiconducting metal oxide layer.

The metal oxide coatings used are mainly of two types:

- *n*-type (n = negative electron) oxides that include zinc oxide, tin dioxide, titanium dioxide, or iron (III) oxide; they are more sensitive to oxidizing compounds because the excitation of these sensors results in an excess amount of electrons in its conduction band.
- *p*-type (p = positive hole) oxides such as nickel oxide or cobalt oxide; they are more reactive to reducing compounds since they develop an electron deficiency under excitation.

Metal oxide sensors are known to have a low sensitivity to moisture and are quite sturdy. The typical operating temperature range is from 400 °C to 600 °C. Volatile compounds are adsorbed onto the surface of the semiconductors thus generating a change in the surface electrical resistance which is also a function of the gas concentration. Figure 6.5 describes the simplified mechanism commonly proposed, for instance by Morrison and Kohl, to explain sensor/gas interaction.

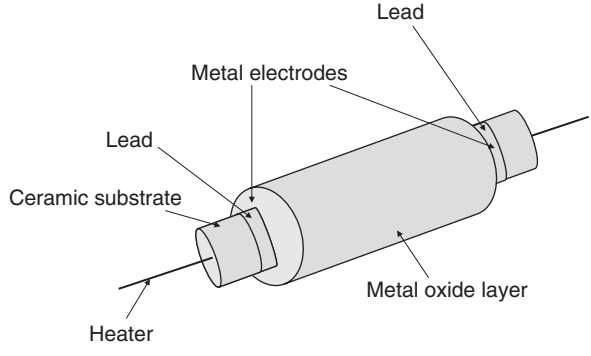


Figure 6.4 Diagram of a metal oxide sensor.

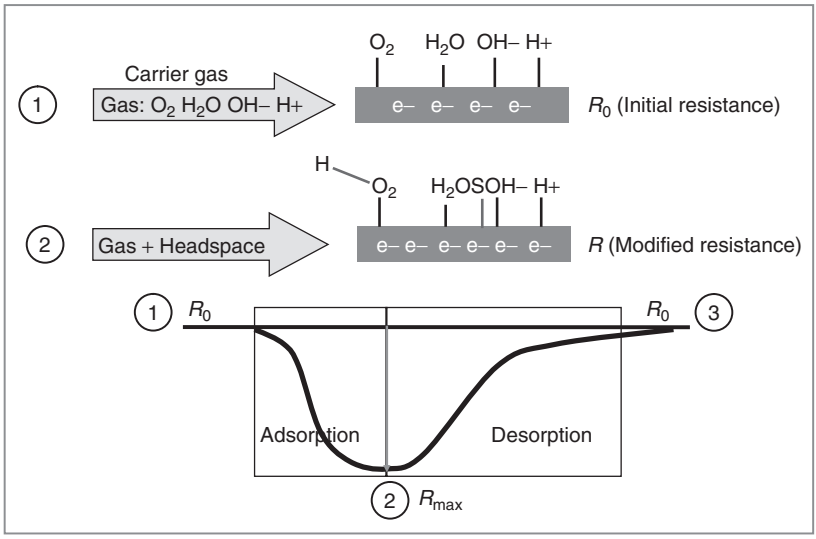


Figure 6.5 Representation of the resistance change which occurs on a metal oxide sensor when a volatile is absorbed on its surface.

- At equilibrium, under reproducible and constant carrier gas flow (usually pure air produced by a TOC generator), the sensor has a stable resistance R_0.
- When sample headspace is injected with the carrier gas, the volatile compounds that form the headspace are adsorbed on the sensor surface and react with the oxygen contained in the carrier gas. The sensor resistance is thus modified up to an optimum value R_{max}. These data (maximum resistances) are recorded over time by the system for further treatment. The presence of oxygen is absolutely necessary in this mechanism. Then desorption of volatile compounds occurs and the sensor recovers its inital resistance R_0, since the process is completely reversible.

Sensitivity (See Table 6.1)

To determine sensor sensitivity, various formulas are proposed in the literature. Generally, they imply values measured with air and with gas. The main parameters that influence the sensitivity of metal oxide sensors are:

- the characteristics of the sensor material,
- the operating temperature of the sensor,
- the ambient conditions (humidity, temperature), and
- the composition and concentration of the gas.

Table 6.1 Some detection thresholds found in the literature.

Compound	Detection threshold	Operating conditions
H_2S	0.1 ppm	SnO_2 sensors Operating temperature: 300 °C
Methanol, ethanol propanol and acetone	10 ppb to 100 ppb	Commercial sensors (Figaro) TGS 812, TGS 824, and TGS 800; static injection in a volume of 20 L
NO	1 ppm	SnO_2 sensors Measuring cell: 1700 cm^3 Range of temperature: 25–300 °C
H_2	0.1 ppm	MOS-Pd sensors coupled with a Pt strand (catalytic action)

The most sophisticated electronic noses have included devices that control these parameters and thus achieve stable operating conditions (chambers with thermotats to control sensor temperature, a source of pure air to limit moisture and contamination, etc.).

Conducting polymers sensors Conducting polymer sensors are composed of an organic conducting polymer film laid down on a silicon or carbon substrate including an electrode, as shown on Figure 6.6. The polymer is obtained through electrochemical polymerization of an aromatic monomer solution (example: pyrrole, thiophene, anilline, indole . . .) with an electrolyte in a solvent. They can be used at room temperatures but not much higher. They react to many volatile compounds but they are more sensitive to moisture (water decreases resistance).

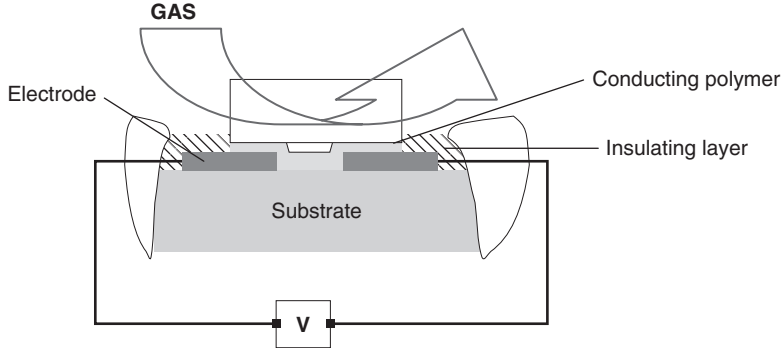


Figure 6.6 Diagram of a conducting polymer sensor.

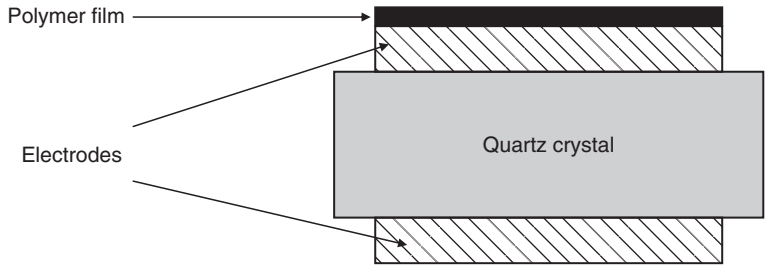


Figure 6.7 Diagram of a quartz microbalance sensor.

When exposed to various volatile compounds, the polymer film conductivity is modified and electrical resistance variation is measured; conduction is achieved by electrons and not by ions. These changes are reversible at room temperature. For practical purposes, either the current intensity is measured under constant voltage or voltage is measured under constant intensity.

Quartz microbalance sensors (QMB) Two types of sensors from this family of piezoelectric crystal sensors exist: those vibrating on a bulk acoustic mode (quartz microbalance) and those vibrating on a surface acoustic mode (SAW). As shown in Figure 6.7 these sensors consist of a quartz crystal on which are placed metal electrodes (aluminum or gold) coated with a polymer film. Usually the polymer films are hydrophobic (polyethylene, polystyrene, etc.). Acoustic waves are generated by an oscillating electrical field induction in the device.

Surface acoustic wave (SAW) sensors These sensors use surface acoustic wave transmission, which is a normal and periodic alteration of the surface under alternative voltage. The solid is usually a piezoelectric crystal and the wave is induced by microelectrodes. In both types of sensors, the volatile compounds adsorbed on the film lead to a change in mass that modifies the wave spreading. The measurable parameter is the oscillating frequency variation. Table 6.2 presents the comparison of the characteristics of the three types of sensors previously described.

Metal oxide semiconductors field effect transistor (MOSFET) sensors As shown in Figure 6.8, MOSFET sensors consist of three layers including a silicon semiconductor, a silicon oxide insulator, and a catalytic metal. The latter is commonly called the gate and is usually made of palladium, platinum, iridium, or rhodium.

Table 6.2 Comparison of metal oxide, conducting polymers & quartz microbalance sensors.

	Metal oxide sensors	Conducting polymers	Quartz microbalance
Sensitivity	+++++	+	+++
Selectivity	+++	++++	+++++
Stability	+++++	+	+++++
Sample humidity effect	Low	Severe	Low to mild
Reaction mode	Oxidation	Polarity	Polar or nonpolar
Measurement function	Resistance change	Conductivity change	Mass vs. frequency
Ruggedness	+++++	++	+++
Sample injection temperature	<200 °C	<45 °C	<120 °C
Resistance to poisoning/ damage	+++++	++	++++

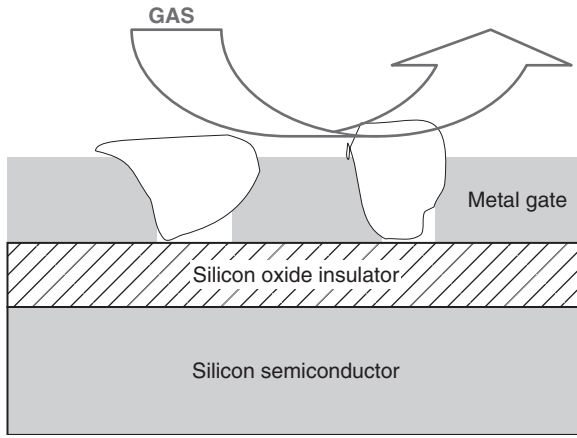


Figure 6.8 Diagram of a MOSFET sensor.

In the MOSFET sensor, the gate and the drain are connected and the sensor operates as a two-terminal device. The field-effect transistor (FET) controls the current between these two points and operates by the effects of an electric field on the flow of electrons through a semiconductor. When a voltage is applied to the gate, current flows within the sensor in a channel from the source to the drain. The silicon oxide insulator keeps current from flowing between the gate and the channel. The gate terminal generates an electric field that controls the current.

As for MOS sensors, MOSFET sensors can include two types of metal oxide semiconductor: *n*-type (conducting through electrons) and *p*-type (conducting through 'holes'). The principle is based on the field effect generated by gases in metal oxide semiconductor field effect transistor (MOSFET) devices with catalytic metal as a gate. The interaction of volatile compounds with the catalytic metal gate induces charges or dipoles, and modifies the gate voltage. The voltage at a constant current is recorded.

Photo ionization detectors (PID) The PID detector consists of a special UV lamp mounted on a low volume flow-through cell with a thermostat. PIDs rely on ionization as the basis of detection. When volatile compounds absorb the energy from the UV lamp, molecules become 'excited' and are ionized. Ions are then collected and produce a current proportional to the number of ionized molecules. Gases are injected through a filter to remove particles. All gases with an ionization potential lower than the photon energy produced by the lamp can be ionized and detected. The PID is therefore considered as a nonspecific gas sensor without interference from the permanent gases of air.

Electrochemical cell As shown in Figure 6.9, electrochemical cells are miniature fuel cells. The simplest form of electrochemical cell consists of two electrodes – sensing and counter – separated by a thin layer of electrolyte. This is enclosed in a plastic housing that has a small capillary

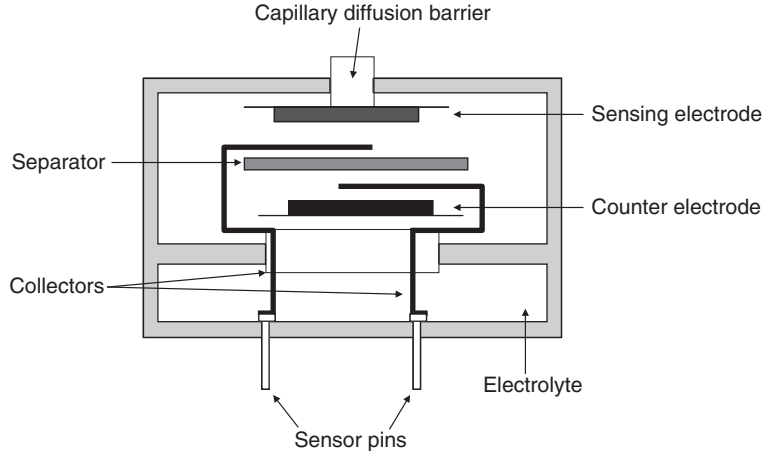


Figure 6.9 Diagram of an electrochemical cell sensor.

to allow gas entry to the sensing electrode and includes pins which are electrically attached to both electrodes and allow easy external interface. These pins may be connected to a simple resistor circuit which allows the voltage to drop resulting from any current flow to be measured. Gas diffusing into the sensor is either oxidized or reduced at the sensing electrode and, coupled with a corresponding (but converse) counter reaction at the other electrode, a current is generated through the external circuit. Since the rate of gas entry into the sensor is controlled by the capillary diffusion barrier, the current generated is proportional to the concentration of gas present outside the sensor and gives a direct measure of the toxic gas present.

The reactions that occur at the electrodes in a carbon monoxide sensor are:

$$\text{sensing electrode: } CO + H_2O \rightarrow CO_2 + 2H^+ + 2e^-$$

$$\text{counter electrode: } {}^1\!/_2O_2 + 2H^+ + 2e^- \rightarrow H_2O$$

$$\text{overall reaction: } CO + {}^1\!/_2O_2 \rightarrow CO_2.$$

Similar reactions take place for all other toxic gases that are capable of being electrochemically oxidized or reduced. In theory, any gas that is capable of being electrochemically oxidized or reduced can be detected. Commercially available sensors are mainly designed to detect toxic gases (CO, ozone, chlorine, H_2S, NH_3, HCl, etc.). Some sensors allow the detection of organic compounds such as mercaptanes or tetrahydrothiophene.

6.3.6 Data Treatment Tools

Each analysis with an electronic nose begins with a training step which consists of analyzing known samples, namely those that were previously assessed by another technique (sensory panel, GC-MS, etc.) to be either qualified or scored. The objective of this first step is to 'train' the electronic nose in the same way as the human nose learns odors. In a second step the analyzer can recognize samples.

Statistical analysis is useful to compute, interpret, and understand the sensor responses of an electronic nose and to carry out their discriminatory power. Differentiation, identification and characterization of samples require the use of multivariate factor analysis. Factor analysis is a type of multivariate analysis that relates to the internal relations of a set of variables.

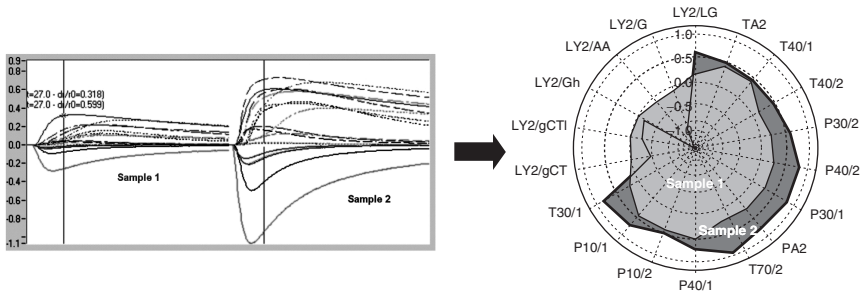


Figure 6.10 Time intensity response plot of the multiple sensors from an electronic nose along with a corresponding 'radar' plot of the maximum responses from each sensor.

Several algorithms and methods have been especially optimized for Electronic Nose applications. Comparing the fingerprint of two samples obtained with one or several sensors (radar plot) is relatively easy. However, when more than two samples are analyzed with several sensors, the data interpretation is much more complex and becomes impossible to perform. At this step, one must use more complex data analysis techniques to study the data-sets as a whole. These tools are known as multivariate statistic algorithms (see Chapter 5) allowing the determination of which differences between samples are important, to identify unknown samples, in blind samples sensory descriptor intensity or quantify substance concentration in an unknown samples.

The most commonly used tools will be described:

- qualitative analysis for identification or recognition: radar plot, Principal Component Analysis (PCA), and discriminant factorial analysis (DFA);
- qualitative analysis for quality control: soft independent modeling by class analogy (SIMCA), and statistical quality chart (SQC);
- quantitative analysis: partial least square (PLS) and odor unit model.

Radar Plot As a basis for sample comparison, a radar plot representation can be used, as shown on Figure 6.10. It consists of representing the optimum of each sensor response on an axis. On this radar plot, the blue and the red zones correspond to two different samples.

Principal component analysis (PCA) Principal component analysis (PCA) is used to explore the data and to assess discrimination performance (i.e., the capability to determine which of the differences are

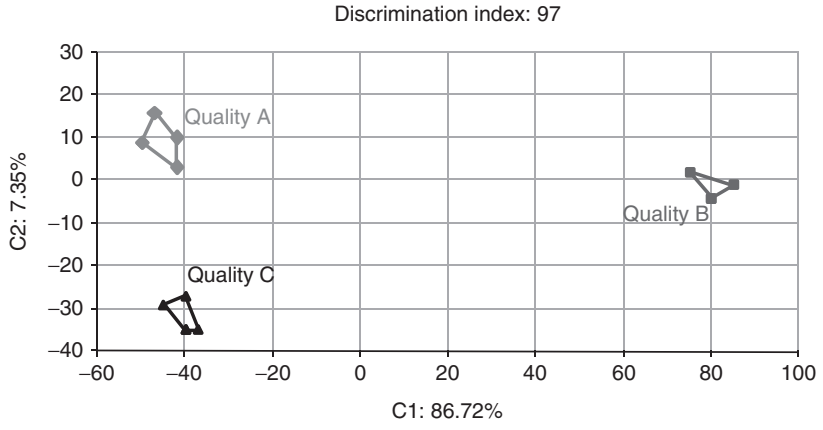


Figure 6.11 Discriminate analysis plot of three quality levels of the same flavor.

important and to which degree). An example of a representation is given in Figure 6.11. In this case, three qualities (good, bad, and medium) of a same flavor were analyzed. The graphic representation gives the plan (two axes) that shows the best discrimination. The percentage mentioned below each axis represents the part of the information brought by this axis: in this example, the horizontal axis accounts for the main part of the information (86.7 %). The discrimination index is calculated by dividing the surface of groups by the surface intergroups. It gives the discrimination quality based on an indication of the surface between groups. When groups overlap each other, the discrimination index is negative. When groups are distinct, the discrimination index is positive; and the higher this value, the better the discrimination.

For the electronic nose, the PCA can be used to assess:

- the discrimination and similarities between different samples and groups,
- the repeatability of the analytical method, and
- the detection of outliers.

It can be used for qualitative analysis, correlation with sensory panel or other technology as a preliminary study but also to perform product matching (i.e. comparison between different companies and competitive brands or formulations).

Discriminant factorial analysis (DFA) Discriminant factorial analysis (DFA), as shown in Figure 6.12, is used to identify unknown samples

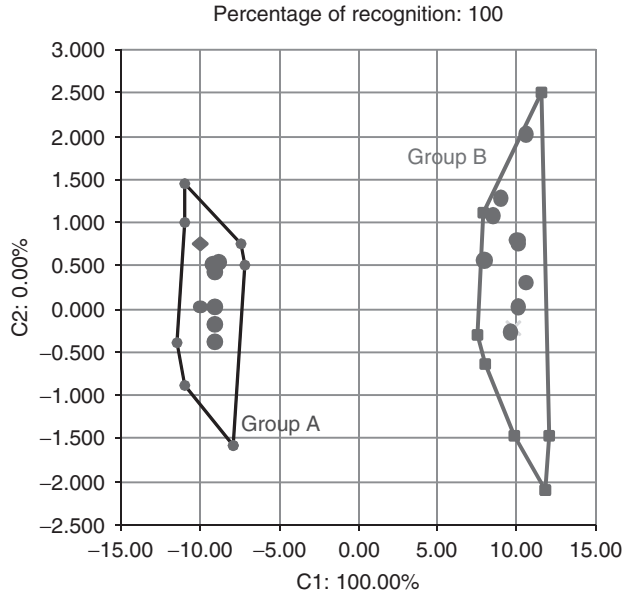


Figure 6.12 Discriminate factor analysis plot of two training groups used to place unknowns into one of these groups.

into one of the training groups. To elaborate, a reliable model for new samples identification – the electronic nose – must be trained with a training sample set representative of all occurrences. For this purpose, the user must correctly identify the characteristics (e.g. quality groups, origins, and ages) of each sample so that the sample training set can be divided into several groups. In the example described in Figure 6.12, two groups of samples are differentiated (A and B). In order to validate the model, a percentage of recognition is calculated this way: a sample is isolated and projected on the model set-up with all other samples to check the identification performed by the instrument. This step is repeated for each sample. The model is considered as valid if the percentage of recognition is higher than 90 %.

Qualitative analysis for quality control Various tools are suitable for quality control purposes including the soft independent modeling by class analogy (SIMCA) and the statistical quality chart (SQC).

Soft independent modeling by class analogy (SIMCA)
SIMCA is used to compare unknown samples to a reference. As can be seen in Figure 6.13, the model is built by taking into account one group considered as the 'gold reference'. The model is going to identify

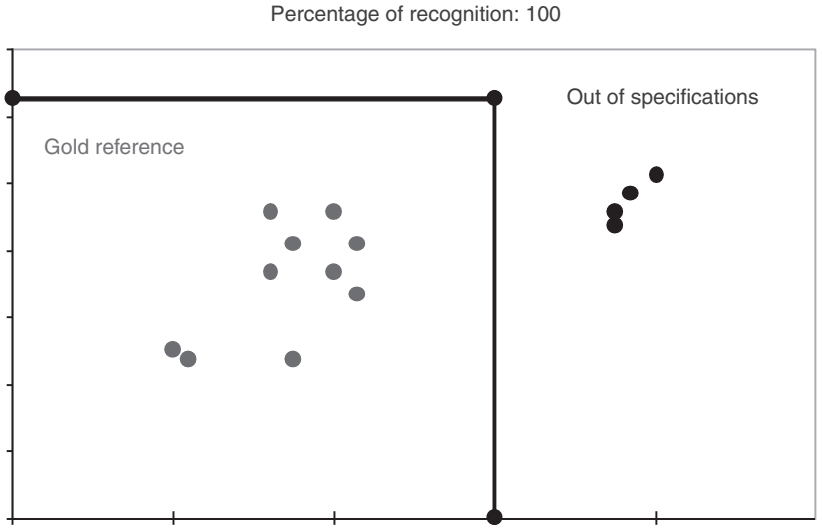


Figure 6.13 A SIMCA model of a defined group is to determine if unknown samples belong to this group.

unknown samples as belonging or not belonging to the one and only group previously defined. The advantage of the SIMCA model is that it is not necessary to collect samples from different groups but only from the group of interest.

Statistical quality chart (SQC)

This model is a quality control monitoring tool. These charts allow one to follow the quality of a product with an acceptable variability of this quality. As shown on Figure 6.14, the data represented are olfactory distances from the target product, concentrations, or sensory panel scores. In a qualitative application, a model is built which relates the sensor data to a quality. For this purpose, a training phase relates the variability of the product to the sensor data. At the conclusion of this step, a chart is constructed with two areas: conforming and nonconforming. In this chart, the upper and lower limits (horizontal lines) define the conformity area (acceptability). If a blind sample is mapped into this area, it will be recognized as conforming. Otherwise (i.e. if the unknown sample is mapped outside this area), it will be recognized as nonconforming.

Quantitative analysis

Partial least square (PLS)

This method is used to correlate the electronic nose measurement with quantitative sample characteristics (substance concentration,

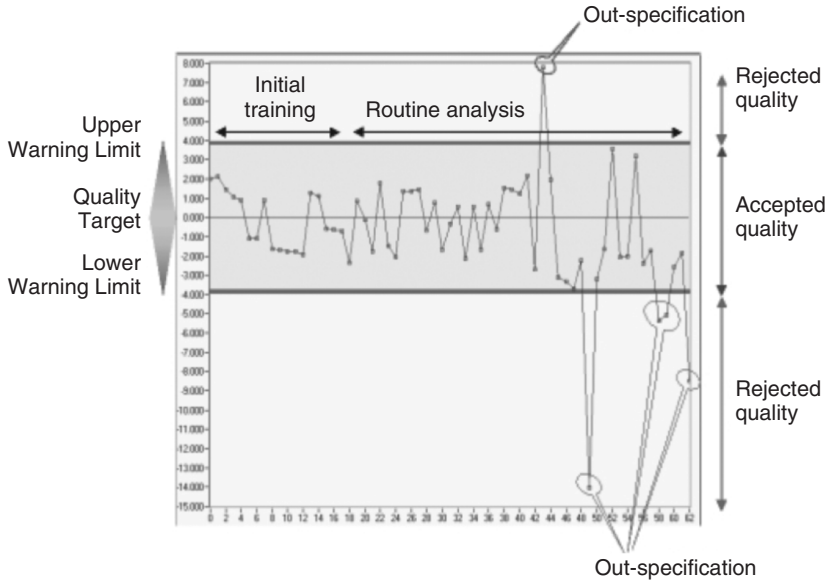


Figure 6.14 A statistical quality chart (SQC) illustrating acceptable and unacceptable sample responses.

descriptor intensity, etc). Once the model has been built based on known samples, unknown samples can be projected on the curve for value prediction, as shown on Figure 6.15.

The correlation coefficient gives an indication on the performance and accuracy of the model: for concentrations, a coefficient higher than 0.90 indicates that the model is valid, whereas for panel scores, the model is valid with a coefficient >0.80. There are mainly two types of quantitative applications:

– determination of the concentration of specific compounds within the product (basic taste substance, off-taste substance, bitterness unit, etc.) and
– sensory panel scores.

Multiband SQC

Very similar to the SQC seen previously, this pattern can integrate various groups of a product (various qualities, scores, etc.). Thus the model built this way allows one to predict a quantitative value linked to samples. The example described in Figure 6.16 highlights three categories of products ranked based on a sensory panel score (0, 1 or 2). The model was set up with the samples assessed by

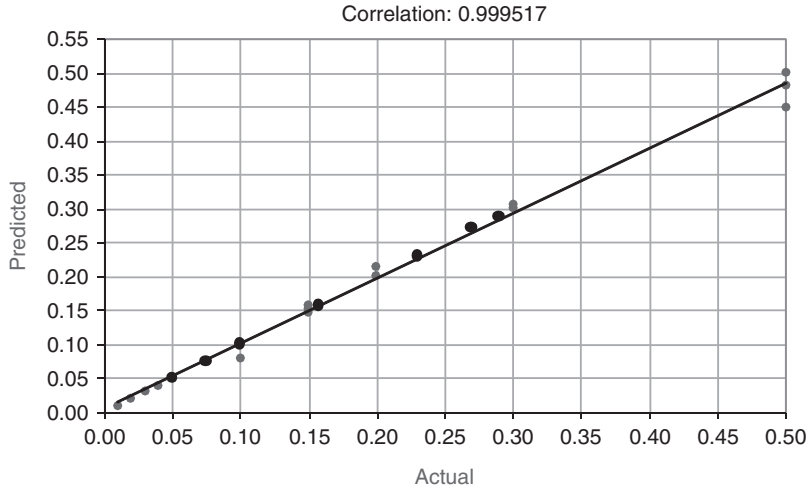


Figure 6.15 Partial least squares analysis showing both training set and unknown samples.

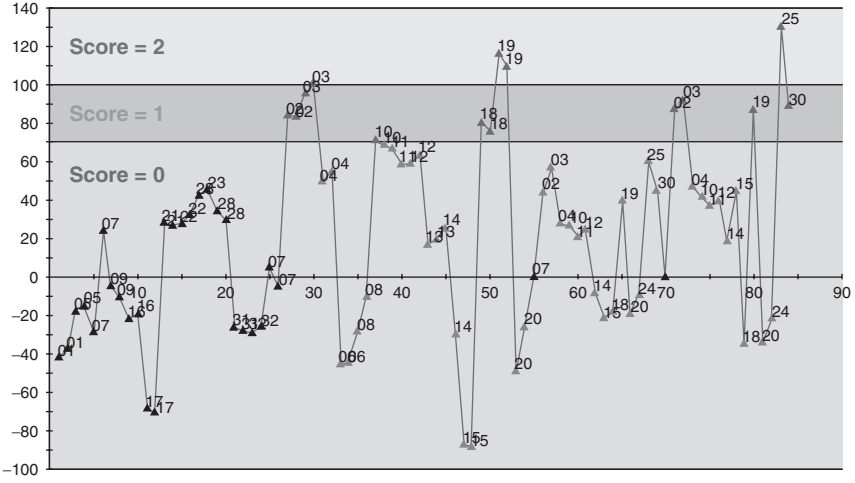


Figure 6.16 A multiband SQC showing various categories of product based on a sensory panel score (0, 1, or 2).

panelists. Then new samples can be projected on the model and their scoring determined as the panel would do.

Odor unit

Another pattern allows the characterization of attributes (flavor, fragrance, stench, etc.) of an unknown odor by measuring its intensity.

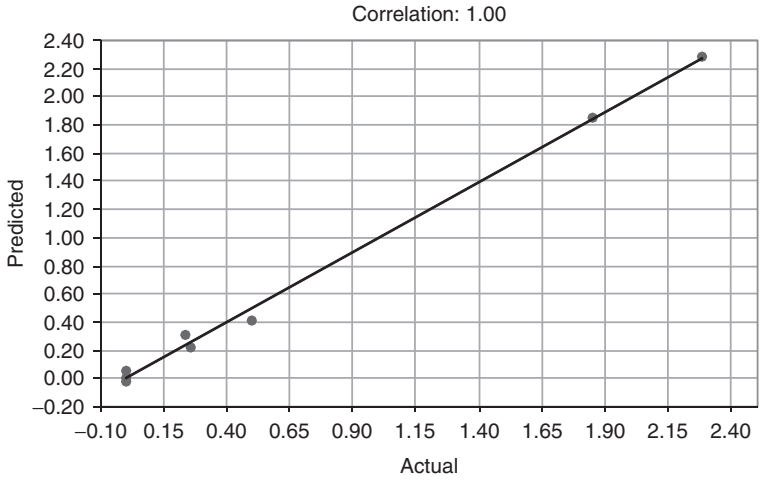


Figure 6.17 An illustration of an application using the multiparametric function to obtain an objective measurement.

Currently, the evaluation of odors is performed by the olfactory sense of human being using an olfactometry method or olfactive evaluation by a sensory panel. With the odor measurement and the multiparametric function, it is now easy to obtain an objective measurement and to give an intensity value following the method, as illustrated in Figure 6.17.

6.4 THE MAIN CRITICISMS DIRECTED AT THE ELECTRONIC NOSE

In the early years of electronic noses, users complained about the following.

- The lack of reproducible results was a matter of concern. This issue had various origins:
  - The variability of the carrier gas has an impact on the quality of the results [1, 2]. Interfacing the electronic nose with a TOC generator continuously supplying pure air of a high and constant quality is recommended. Variability of the carrier gas is therefore eliminated resulting in more reproducible data.
- The variability of the headspace generation step can be explained by the inconsistent quantity of sample analyzed. Today, however, the quantity/volume of sample placed in the analysis vial is accurately

weighed or measured. Moreover, headspace injections, specifically when done manually, were not always similar. That is why autosamplers are now used to assure reproducible and reliable operating conditions for headspace generation and injection.

- Sensor drift was often reported. Drift can be due to degradation and/or response drift over time [3–5]. Some users have established calibration methods [4] consisting of analyzing reference compounds regularly so as to check measurement reproducibility and detect potential drift. Nowadays, some manufacturers of electronic nose instruments have included a periodical, automated diagnostic of sensor sensitivity, performed with reference chemical products. Additionally, an automated mathematical adjustment can be employed. Consequently, the proper functioning of sensor responses is repeatedly checked and assured.
- Sensitivity to environmental conditions (humidity, temperature): in the past, the influence of moisture and temperature on sensors was reported as a major drawback [6, 7]. Since that time, several technical improvements have been developed to overcome this issue. A humidity sensor has been included in some instruments to track the level of moisture during the analysis. The use of a TOC generator can guarantee a source of dry air (less than 5 ppm of humidity) as well. As for temperature, sensor chambers now have thermostats to maintain a constant temperature of the sensors.*Sensors lifespan*: this lifetime depends on the sensor type; 18–24 months for MOS and between 6-9 months for conducting polymers and QMB. Presently, this lifespan is continuously increased by sensor manufacturers due to production process improvements.*Non-specificity of sensors*: some studies [3] mention the lack of specificity of sensors to volatile compounds as a restraint. Indeed, sensors are purposely cross-selective, just like human odor receptors, in order to react to the widest range of components. However, some instruments now integrate specific detectors or sensors in order to detect one target compound. Moreover, sensors can be selected by prior tests, based on the user's applications and objectives with a view to optimizing results.
- *Long database set-up and method development* [7]: due to the difficulty of gathering a representative set of samples qualified by another method (sensory panel, GC, GC/MS), the database set up can take a lot of time. New mathematical models recently developed, especially for quality control applications, require fewer samples to establish a reference. Besides, manufacturers now tend

to elaborate and recommend standard operating procedures that define the optimum number of samples needed and thus save time.
- *Software not easy to use* [8]: software development is part of ongoing technical improvements. Presently, electronic nose software includes tools that were created to facilitate data treatment and that are accessible to nonskilled users.

6.5 MARKET AND APPLICATIONS

6.5.1 Application Range

Electronic nose instruments are used by research and development laboratories, quality control laboratories and process and production departments for various purposes:

- in research and development laboratories for:
  - formulation or reformulation of products,
  - benchmarking with competitive products,
  - shelf life and stability studies,
  - selection of raw materials,
  - packaging interaction effects, and
  - simplification of consumer preference test.
- in quality control laboratories for in line quality control such as:
  - conformity of raw materials, intermediate and final products,
  - batch to batch consistency,
  - detection of contamination, spoilage, adulteration,
  - origin or vendor selection, and
  - monitoring of storage conditions.
- in process and production departments for:
  - managing raw material variability,
  - comparison with a reference product,
  - measurement and comparison of the effects of manufacturing process on products,
  - following-up cleaning in place process efficiency,
  - scale-up monitoring, and
  - cleaning in place monitoring.
    - The fields of applications include: Flavors and fragrances: selection of the desired fragrance [9], checking the quality of products [10] or raw materials [11]. Food: fish freshness monitoring [12, 13], aroma characterization of cooked meats [14, 15], meat aging [16] and quality [17–19], differentiation

of fruits based on maturity [20, 21] or variety [1], shelf life of vegetables [22] and milk [23], quality [24] and deterioration [25, 26] of cheese, quality control [4, 5, 8, 27] and classification [28] of edible oil. Beverages: aroma analysis of tea [6, 29] and coffee [3, 30], optimization of beer ageing [31], influence of the production process on cider quality [32], differentiation [2, 33], authentification of fruit juice [34] and correlation of results with sensory evaluation [35]. Aroma analysis of tobacco [36]. Packaging: odor analysis of polymer pellets [37].

- Pharmaceutical industry: assessment of unpleasant odors in coated tablets [38], differentiation and quantification of various flavors in oral formulations [39].
- Nutraceuticals: evaluation of nutritive drinks palatability [40], recognition of ginseng origin [41].
- Cosmetics and perfumes: identification and quantification of a raw material in a composition [42] or recognition of various perfumes [43], applications such as substantivity, olfactory intensity and conformity, product efficacy to mask unpleasant odors, selection of raw materials based on quality [44].
- Chemicals: replication and quantification of a perfume in cleaning products [45], assessment of candle fragrance [46], odor analysis of polymer textile films and foams in the automobile industry [47], oil odor evaluation [48], shelf life of chemicals [49].
- Environment: odor analysis of wastewater [50], quality monitoring of potable water [51], detection of contamination in fish [52, 53], odor intensity measurement in the air [54], air quality monitoring [55].

Several case studies will be detailed below.

Case Study 1 – Perfumery compound detection in a fragrance [42]: this study, performed by IFF (Argentina) and the Asociacion Argentina de Quimocos Cosmeticos (Argentina), compares the performance of three methods (the electronic nose, GC/MS and sensory panel) for the detection of a compound in a composition.

Case Study 2 Cosmetic natural raw materials [11]: characterization of volatile constituents of benzoin gum.

This case investigated the use of an electronic nose to check the quality of natural raw materials based on various parameters (origin, harvesting year, grade) and was conducted by Charabot (France) and Université de Nice (France).

Case Study 3 Home care products [45]: identification and quantification using an electronic nose in the perfumed cleaner industry.

This study conducted by Promer (France) is aimed at identifying among various perfumes, the one that better matches target expectations, then to quantify the perfume concentration in the formulation during the adding process.

Case Study 4 Pharmaceutical products [39]: flavor analysis in liquid oral formulations.

This first study describes the use of an electronic nose in the pharmaceutical formulation department of Merck & Co. Inc (USA) to differentiate various flavors used in medicines, to recognize them once introduced in formulations and to evaluate the quantity of flavor in the formulation.

6.5.2 Perfumery Compound Detection in a Fragrance

The identification of the notes composing a fragrance is a common issue in the perfume industry. Usually, this evaluation is done by humans. More recently, the detection of a compound (the mangone) in a fragrance was conducted by three methods in order to compare their performance [42]: an electronic nose including an array of 11 tin-oxide based sensors, a GC/MS system and a sensory panel of 20 trained panelists. The mangone was selected for its low threshold value of sensory detection and because it belongs to a different olfaction family (citrus, grapefruit) with respect to the compounds of the fragrance (green and pinecone notes). Successive dilutions of the mangone in the fragrance were tested. The range of dilution assessed was between 10^{-1} and 10^{-4} %.

The sensory panel evaluation consisted of a triangle test: each panelist had to smell three samples, two samples of pure fragrance and one sample containing mangone. The panelist had to indicate which of the three samples was different. An individual identification was considered as positive when the panelist had successfully recognized the different sample. For the global panel of 20 panelists, identification was deemed positive when 70 % of panelists had passed the test.

Results

Between a concentration of 1 and 10^{-1}, the three techniques allowed the presence of mangone to be detected in the fragrance (Table 6.3). Starting from lower concentrations (10^{-2} %), the electronic nose was the only method that could pick-up the mangone presence.

Table 6.3 Comparison of GC/MS, sensory panel and electronic nose for the detection of mangone in a fragrance.

Mangone concentration*	GC/MS**	Sensory panel** (% of successful identification in the panel)	Electronic nose**
1	Yes	Yes (100 %)	Yes
10^{-1}	Yes	Yes (89 %)	Yes
10^{-2}	No	No (37 %)	Yes
10^{-3}	No	No (45 %)	Yes
10^{-4}	No	No (30 %)	Yes

*These values correspond to a percentage (weight/weight) of mangone in the fragrance.
**Yes = the GC/MS is able to distinguish the presence of mangone in the fragrance.

Conclusion

For the mangone compound diluted in a fragrance, the electronic nose had a lower threshold of detection than GC/MS and the trained sensory panel.

6.5.3 Cosmetic Natural Raw Materials: Characterization of Volatile Constituents of Benzoin Gum

Benzoin gums are commonly used in the formulation of luxury flavors and fragrances [11].

The quality and consistency of raw materials are difficult to check given that they can be produced by craft methods, graded locally by each producer and traded by brokers. Moreover, their prices are defined according to the claimed quality. So it is crucial both for users and producers to objectively assess the quality of raw materials.

Several gums have been analyzed with an electronic nose using 18 metal oxide sensors (FOX 4000, Alpha MOS, France) to determine the quality and the olfactive features of Benzoin gums. The selected factors being the harvest year, the country of origin, the grade of fragrance, the method can also be used to detect counterfeited Benzoin gums.

Samples and operating conditions are detailed within Tables 6.4 and 6.5.

Results

- ***Differentiation of gum origin***
  A clear discrimination between Siam (red) and Sumatra (blue) gums could be observed (radar plot Figure 6.18). In the same way, Siam, Sumatra and false gums were clearly separated (discriminant

Table 6.4 FOX E-Nose operating conditions for benzoin gum analysis.

Sensor array	18 sensors
Quantity of sample	0.2 g
Headspace generation	10 min at 60 °C
Injection volume	500 μL
Acquisition time	120 s

Table 6.5 Benzoin gum samples.

Origin	Grades	Number of samples
Siam	2 – 3 – 5	13
Sumatra	A – B – C – D	17
False	None	3
Unknown	?	23

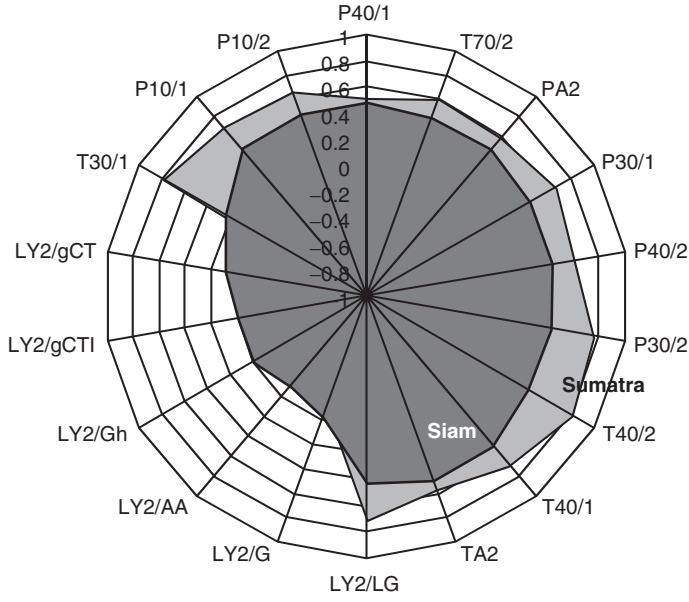


Figure 6.18 Radar plot of sensor responses for two Benzoin gums of different origins.

factorial analysis, Figure 6.19.). Samples for which origin were successfully identified with a high percentage of recognition (99 %).

- ***Differentiation of the harvesting year***
  A clear differentiation based on harvesting year was achieved (PCA Figure 6.20).

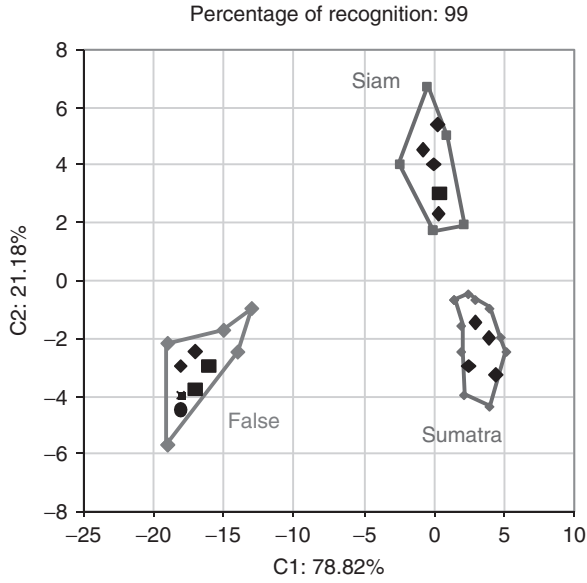


Figure 6.19 Discriminant function analysis of the same Benzoin gums shown in Figure 6.18, along with some false gums.

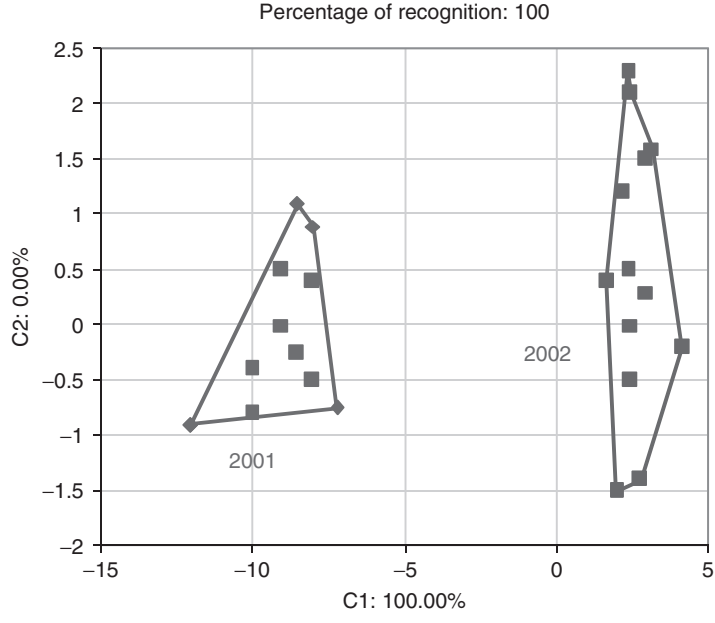


Figure 6.20 Principle components analysis of Benzoin gums of different harvest years.

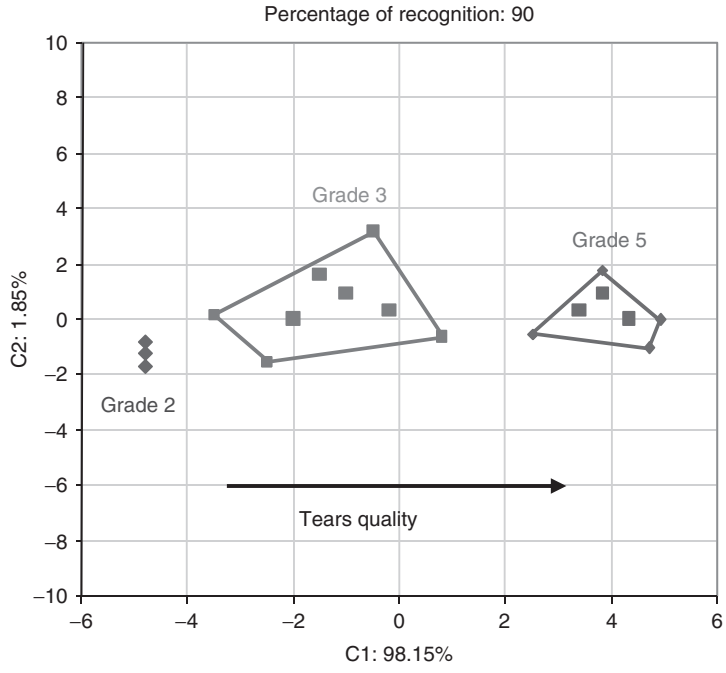


Figure 6.21 Discriminant function analysis of three varying qualities of Benzoin gums.

- *Differentiation of qualities*
  Siam benzoin gums (13 samples): the three qualities of Siam benzoin gums (13 samples) are clearly separated (DFA Figure 6.21).

Conclusions

The FOX electronic nose successfully and rapidly discriminated benzoin gums based on various criteria. It offered an alternative rapid technology to conventional techniques such as chromatographic methods and sensorial analysis for raw material quality control and prediction. The electronic nose is a fast and objective tool for the flavor and fragrance industry for checking raw material quality rapidly.

6.5.4 Home Care Products: Identification and Quantification Using an Electronic Nose in the Perfumed Cleaner Industry

The improvement of analytical techniques to analyze perfumed products is becoming crucial for industries [45]. Traditional analyses rely

Table 6.6 Experimental conditions for the analysis of Rose perfume.

Sample preparation	
Quantity of sample in the vial	1 mL
Total volume of the vial	10 mL
Headspace generation	
Headspace generation time	10 min
Headspace generation temperature	40 °C
Agitation speed	500 rpm
Headspace injection	
Injected volume	2000 μL
Injection speed	2000 μL/s
Syringe temperature	45 °C
Acquisition parameters	
Acquisition time	120 s
Time between 2 injections	18 min

on sensory panel evaluation and gas chromatography analyses of formulations. For the development of routine control tests, electronic nose instruments – a simpler and faster to use technique – become of interest.

The use of an MOS-based electronic nose for qualitative and quantitative analysis of perfumed cleaner products is reported here with the identification of the closest product for a target reference and determination of the level of perfume added onto one product.

Experimental conditions are detailed within Table 6.6.

Results

- *Identification of the best matching perfume*
  The objective was to determine which perfume in a set of three (elaborated by various suppliers, noted RO10, RO15 and RO20) could be used to copy an existing ROSE perfume.
  The electronic nose was used to compare product smell-prints and find which one was the most similar to the reference ROSE perfume. All samples were separated (PCA Figure 6.22). The group distance enabled the measurement of the proximity between samples.
  A model was set up to monitor the quality of rose perfumes (SQC Figure 6.23). In this SQC model, rose perfume was used to establish an area corresponding to acceptable natural variations (e.g. change of flavor quality). Visually, the model built presented an acceptable bandwidth defined between two limits; all samples outside of this band were considered out of specifications.

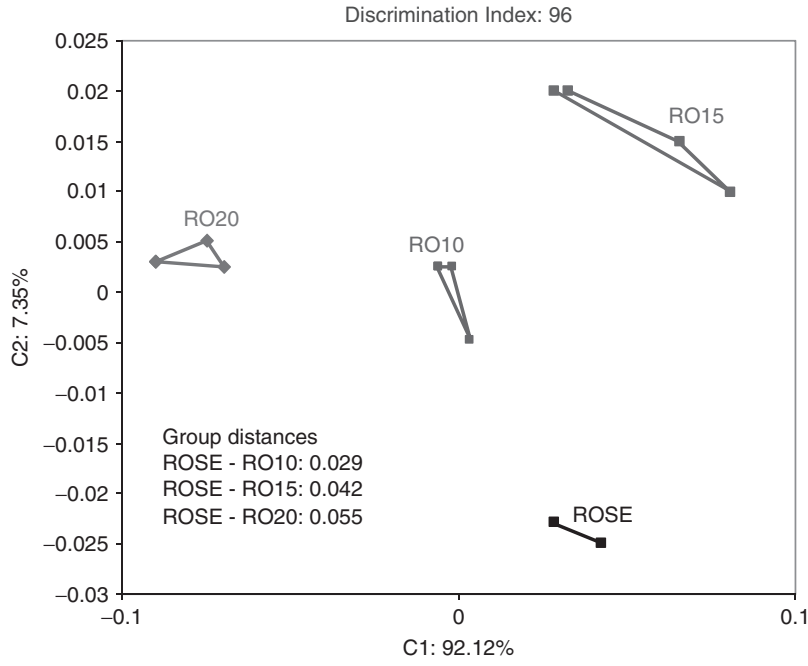


Figure 6.22 Principle components analysis of three cleaning products with a rose aroma along with a reference sample.

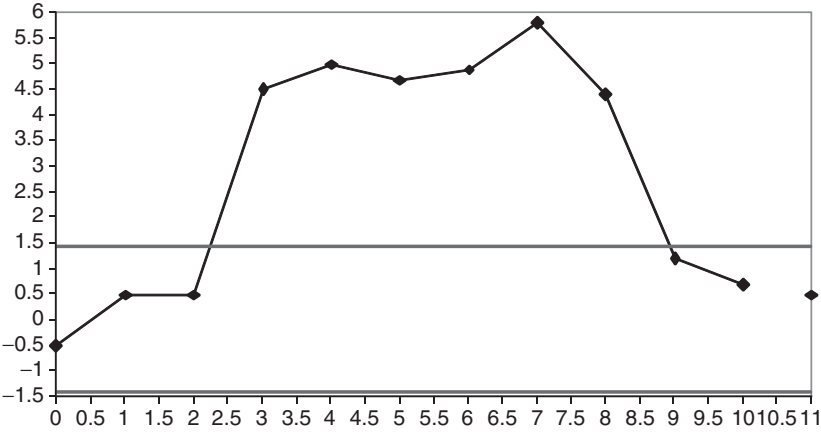


Figure 6.23 A statistical quality chart using the reference sample and illustrating how two of the cleaning products were determined to be outside of the acceptable natural variations of the rose aroma.

- *Quantification of perfume concentration*
  The objective was to determine the level of perfume in an unknown sample perfume MNE using a set of eight prepared calibration formulations containing different concentrations (0.6–4 %). Results showed (PCA Figure 6.24):
  – a good separation and reproducibility and
  – a distribution from right to left with increasing level of perfume.

  In order to quantify the concentration of rose perfume in the coming batches, a quantification model was set up (PLS Figure 6.25). This model showed:
  – a high correlation coefficient of 0.998 and
  – an MNE sample projected as an unknown onto the model gave a perfume concentration of 3.49 %.

Conclusions

For perfume replication, the results obtained were consistent with sensory evaluation. The electronic nose offered a very promising application in product placement and selection of matching formulations. For quantification, after building a calibration model, the system could accurately predict the level of perfume introduced into solutions. The electronic nose could be used successfully for routine monitoring of process variations of added level of perfume.

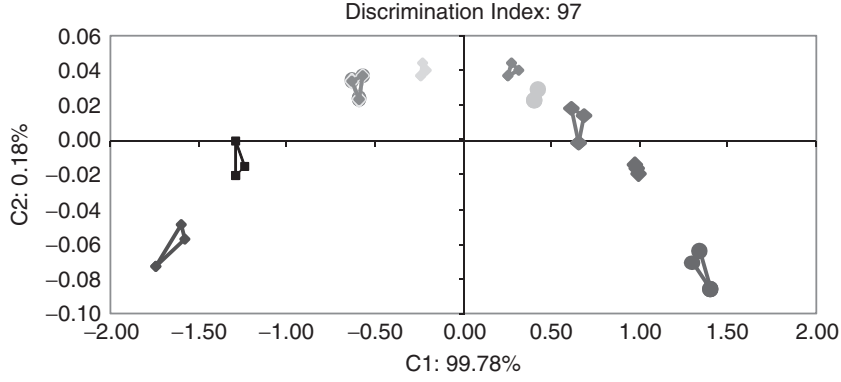


Figure 6.24 Principle components analysis of an unknown sample (MNE) along with a set of eight known calibration formulations (0.6–4.0%).

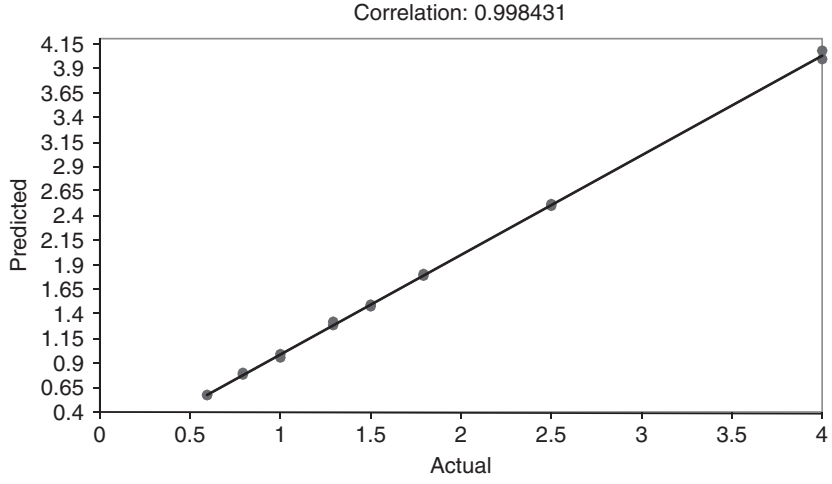


Figure 6.25 Partial least squares analysis of the same unknown samples as in Figure 6.24.

6.5.5 Pharmaceutical Products: Flavor Analysis in Liquid Oral Formulations

Flavors are commonly used in the food and beverage industries, personal care and pharmaceutical products [39]. It is highly desirable to qualitatively and quantitatively analyze different flavors for formulation development, stability and quality control purposes.

An electronic nose using 18 metal oxide sensors (FOX 4000, Alpha MOS, France) was used to complement human sensory perception and other traditional analytical methods for consistent qualitative and quantitative analysis of flavors in formulations. This was done via a qualitative study of six various flavors and a quantification of various flavor concentrations. Samples and experimental conditions are detailed in Tables 6.7 and 6.8.

Results

- ***Differentiation of flavor samples (Figure 6.26)***
  The differentiation of the six flavors was achieved with a high level of discrimination. As the reproducibility of the instrument was high (measurement precision RSD = 1.3 % and method precision RSD = 0.5 %), results showed a low sample variation except for

Table 6.7 Liquid oral formulation samples.

Analyzes	Sample name	Sample description
Qualitative analysis		
Batch to batch analysis	A1, A2, A3 and A5	Placebo flavored with raspberry (various lots of raw materials)
Aging analysis	A4	Sample A1 stored at room temperature for 8 months
Training standard solutions A5 to A11	A6 to A10	Placebo flavored with cherry, strawberry, red berry, pineapple, and orange respectively
	A11	Placebo without any flavor
Three unknown samples (prepared by an analyst and analyzed by another one without flavor knowledge)	Unknown 1 and 2	4 mg/mL raspberry flavor and red berry flavor
	Unknown 3	Mixture: 2 mg/mL strawberry + 2 mg/mL raspberry flavor
Quantitative analysis		
Training standards with raspberry flavor from manufacturer's lot 4	B1	1.01 mg/mL raspberry flavor from manufacturer's lot 4
	B2	2.01 mg/mL raspberry flavor from manufacturer's lot 4
	B3	3.00 mg/mL raspberry flavor from manufacturer's lot 4
	B4	4.04 mg/mL raspberry flavor from manufacturer's lot 4
	B5	5.02 mg/mL raspberry flavor from manufacturer's lot 4

Table 6.8 E-Nose operating conditions for the analysis of liquid oral formulations.

Sensor array	18 sensors
Sample volume	1 mL
Headspace generation	240 s at 40 °C
Injection volume	2 mL
Acquisition time	120 s

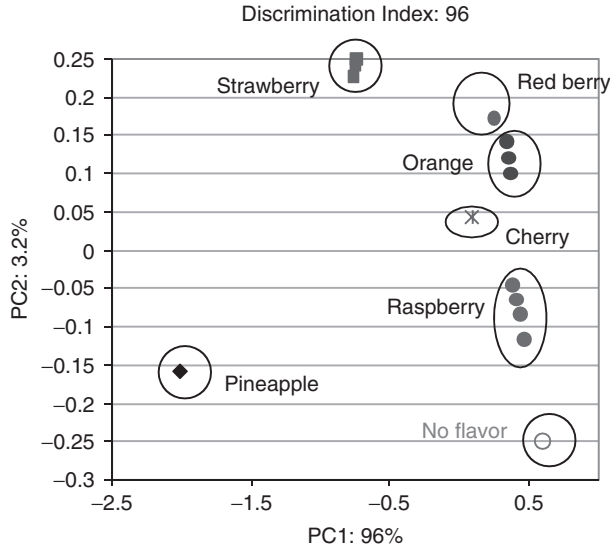


Figure 6.26 Principle components analysis of six flavors along with a sample with no flavor added.

raspberry. Therefore, further analysis of the raspberry flavor was done in order to detect a possible batch to batch variation.

- ***Correlation with sensory panel analyses (Figure 6.27)***
  In order to check the correlation of the electronic nose measurement with human sensory evaluation, a model was built using solutions of known flavor composition. Then analyses of new samples were proceeded with an aim to identify them. The results showed that:
  - unknown samples 1 (raspberry) and 2 (red berry) were correctly identified;
  - unknown sample 3 (a mixture of raspberry and strawberry) was identified as containing strawberry flavor. When evaluated by the human nose, this sample gave the perception of strawberry flavor, though it did not smell exactly the same as the samples containing only strawberry flavor.

  Therefore, the result from the electronic nose correlated with the human nose assessment, indicating that the discrimination ability of the instrument was comparable with the human nose.
- ***Batch to batch variation analysis – discrimination of fresh and aged samples (Figure 6.28)***
  In order to understand the variation in the raspberry flavor and study the effects of storage conditions on the flavor stability, five

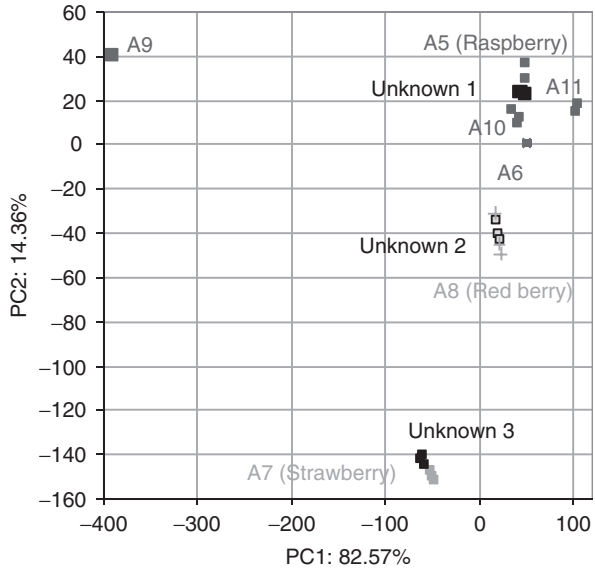


Figure 6.27 Principle components analysis of three flavors along with three unknowns that were correctly identified.

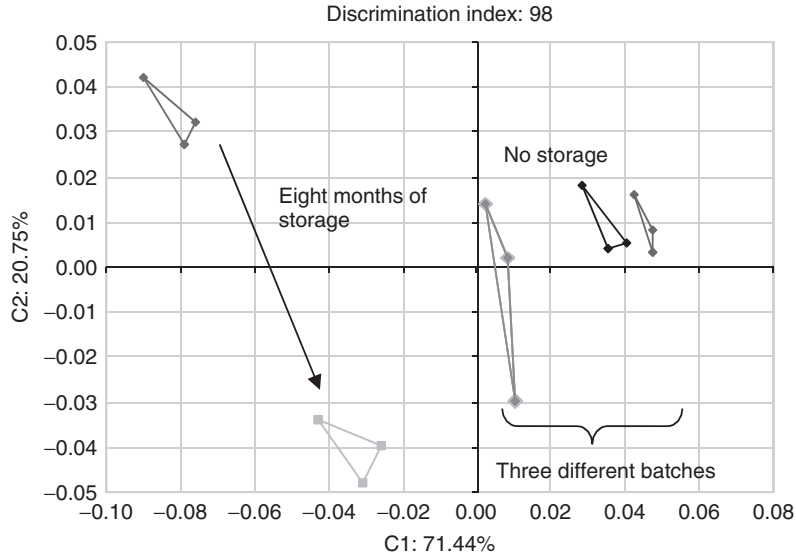


Figure 6.28 Principle components analysis of four flavors and one stored for 8 months.

placebos containing different lots of fresh or aged raspberry flavors were analyzed. Samples A1 and A4 (A4 = sample A3 stored at room temperature for 8 months) had significantly different fingerprints compared with the other three lots of raspberry flavors (samples A2, A3, and A5). Samples A2, A3, and A5 showed a very small batch to batch variation.

A GC analysis showed that sample A1 had a different peak profile compared with the other three lots (samples A2, A3, and A5). This means that the electronic nose results correlated with GC results. Sample A4 containing the same lot of sample A3, but with 8 months storage in ambient conditions was discriminated by electronic nose. The human nose evaluation of these two samples verified that the raspberry flavor in the aged sample was weaker. Therefore, the electronic nose, in the same way as the human nose, could pick up differences between fresh and aged samples. This study proved that storage time and temperature might have an impact on the quality of the flavors in formulations.

- ***Predicting model for the concentration of raspberry flavor (Figure 6.29)***

  A calibration curve was generated from formulations with a known concentration of raspberry flavor (PLS Figure 6.29). The *x*-axis

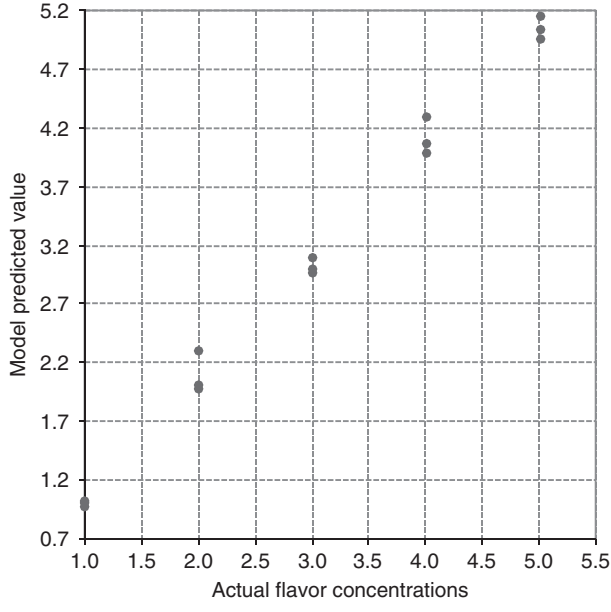


Figure 6.29 Partial least squares calibration curve of raspberry flavor concentration.

represents actual flavor concentrations of the standards input to the model, and the *y*-axis corresponds to the prediction value produced by the model.

The coefficient of determination obtained ($R = 0.9954$) indicated that this calibration model could be used to reliably predict raspberry flavor concentration in unknown samples.

Conclusions

The electronic nose could qualitatively distinguish and recognize various unknown flavors and lots, differentiate fresh from aged flavor samples, and quantify flavor amounts.

This analyzer is a rapid tool with an adequate selectivity and sensitivity to perform flavor identification in edible products but also to assay the flavor concentration during release testing, to check the quality of raw material according to batches and suppliers, and to monitor flavor stability during shelf life. It is therefore an efficient tool to build a database of flavors used for drug formulation and thus speed up the formulation process.

REFERENCES

1. Hirschfelder, M., Hanrieder, D., and Ulrich, D. (1998) Discrimination of strawberry varieties by a gas sensor array in correlation with human sensory evaluation. *Seminars in Food Analysis*, **3**, 27–36.
2. Goodner, K.L. and Baldwin, E.A. (2000) The use of an electronic nose to differentiate NFC (not from concentrate) orange juices. *Proc. Fla State Hort*., **113**, 304–306.
3. Gretsch, C., Toury, A., Estebaranz, R., and Liardon, R. (1998) Sensitivity of metal oxide sensors towards coffee aroma. *Seminars in Food Analysis*, **37–42**, 5–13.
4. Jones, H. and Engelen-Cornax, J. (1998) Quality control of edible oil using an electronic nose. *Seminars in Food Analysis*, **3**, 15–25.
5. Garcia Gonzalez, D. and Aparicio, R. (2002) Detection of defective virgin olive oils by metal oxide sensors. *European Food Research and Technology*, **215**, 118–123.
6. Dutta, R. *et al*. (2003) Electronic nose based tea quality standardization. *Neural Networks*, **16**, 847–853.
7. Carrasco, A. and Saby, C. (1998) Discrimination of Yves Saint Laurent perfumes by an electronic nose. *Flavour and Fragance Journal*, **13**, 335–348.
8. Shiers, V. and Adechy, M. (1998) Use of multi-sensor array devices to attempt to predict shelf-lives of edible oils. *Seminars in Food Analysis*, **3**, 43–52.
9. Fukai, S. and Abe, Y. (2002) Discrimination of lily fragrance by use of an electronic nose. Acta horticulturae (International Society for Horticultural Science) Proceedings of the 20th International Eucarpia Symposium, p. 572.
10. Strassburger, K. (1998) Electronic nose technology in the flavor industry: moving from R&D to the production floor. *Seminars in Food Analysis*, **3**, 5–13.

11. Fernandez, X. *et al.* (2006) Volatile constituents of benzoin gums: Siam and Sumatra, Part 3. Fast Characterization with an Electronic Nose. *Flavour Fragr. J.*, **21**, 439–446.
12. Haugen, J.E. *et al.* (2005) Rapid control of smoked Atlantic Salmon (salmo salar) by electronic nose: correlation with classical methods. *Sensors & Actuators B*, **116**(1–2), 72–77.
13. Olafsdottir, G. *et al.* Prediction of microbial and sensory quality of cold smoked atlantic salmon (Salmo salar) by Electronic Nose. *J. Food Sci.*, **70**, S563–574.
14. Taurino, A.M. *et al.* (2003) Analysis of dry salami by means of an electronic nose and correlation with microbiological methods. *Sensors & Actuators B*, **95**, 123–131.
15. Otero, L. *et al.* (2003) Detection of Iberian Ham aroma by a semiconductor multi-sensorial system. *Meat Science*, **65**, 1175–1185.
16. Boothe, D.H. and Arnold, J.W. (2002) Electronic nose analysis of volatile compounds from poultry meat samples, fresh and after refrigerated storage. *J. Sci. Food Agric.*, **82**, 315–322.
17. González-Martín, I. *et al.* (2000) Differentiation of products derived from Iberian breed swine by electronic olfactometry. *Anal. Chim. Acta*, **424**, 279–287.
18. Suwansri, S., Pohlman, F.W., Meullenet, J-F. and McEylea, K. S. (2001) Application of an electronic nose for discrimination of ground beef quality through simulated retail display. *IFT*, Annual meeting – New Orleans, LA, presentation 30C-11, June 23–27, 2001.
19. Sullivan, M.G., Byrne, D.V. and Jensen, M.T. (2003) A comparison of warmed-over flavour in pork by sensory analysis, GC/MS and the electronic nose. *Meat Science*, **65**, 1125–1138.
20. Saevels, S. *et al.* (2003) Electronic nose as a non-destructive tool to evaluate the optimal harvest date of apples. *Postharvest Biology and Technology*, **30**(1), 3–14.
21. Supriyadi *et al.* (2004) Maturity discrimination of snake fruit (Salacca edulis Reinw.) cv. Pondoh based on volatiles analysis using an electronic nose device equipped with a sensor array and fingerprint mass spectrometry. *Flavor and Fragrance Journal*, **19**, 44–50.
22. Benedetti, S., Toppini, P.M. and Riva, M. (2003) Shelf life of fresh cut vegetables as measured by an electronic nose: preliminary study. *J. Food Sci.*, **13**, 201–212.
23. Labrèche, S., Bazzo, S., Cadé, S. and Chanié, E. (2003) Shelf life determination by electronic nose: application to milk. *Sensors & Actuators B, Chemicals, SNB7947*, **106**, 199–206.
24. Decker, M., Trihaas, J. and Nielsen, P.V. (2003) White mould cheese: new tools for objective quality evaluation. *Technical University of Denmark*, Denmark (2003) pp210.
25. Chung, H., Partridge, J. and Harte B. (2003) Evaluation of light deterioration in cheddar cheese using the olfactory sensing technique. *IFT Annual meeting – Chicago*, IL, presentation 71-4, July 12–16, 2003.
26. Drake, M.A., Gerard, P.D., Kleinhenz, J.P. and Harper, W.J. (2003) Application of an electronic nose to correlate with descriptive sensory analysis of aged cheddar cheese. *Lebensm. Wiss. U. Technol.*, **36**, 13–20.
27. Cerrato Oliveros, M. C. *et al.* (2002) Electronic nose based on metal oxide semiconductor sensors as a fast alternative for the detection of adulteration of virgin olive oils. *Anal. Chim. Acta*, **459**, 219–228.

28. Gonzalez, Y., Perez, J.L., Moreno, B. and Garcia, C. (1999) Classification of vegetable oils by linear discriminate analysis of electronic nose. *Anal. Chim. Acta*, **384**, 83–94.
29. A.S.M. Ou, L.S. Chen, Aroma quality analysis of Taïwan local tea by sensory evaluation and electronic nose. *IFT Annual meeting – New Orleans*, LA, presentation 73G-11, June 23–27, 2001.
30. Costa Freitas, A.M., Parreira, C. and Vilas-Boas, L. (2001) The use of an electronic aroma sensing device to assess coffee differentiation – comparison with SPME GC-MS aroma pattern. *J. Food Comp. Anal.*, **14**, 513–522.
31. McKellar, R. C. *et al.* (2002) Use of the electronic nose and gas chromatography-mass spectrometry to determine the optimum time of aging of beer. *Master Brewers Association of the Americas Technical Quarterly*, **39**, 99–105
32. Rye, G.G. and Mercer, D.G. (2003) Changes in headspace volatile attributes of apple cider resulting from thermal processing and storage. *Food Research International*, **36**, 167174.
33. Goodner, K.L. and Baldwin E.A. (2001) The comparison of an electronic nose and gas chromatograph for differentiating NFC orange juices. *Proc. Fla. State Hort. Soc.*, **114**, 158–160.
34. Steine, C., Beaucousin, F., Siv, C., and Peiffer, G. (2001) Potential of semi-conductor sensor arrays for the origin authentification of pure Valencia orange juices. *J. Agric. Food Chem.*, **49**, 3151–3160.
35. Bleibaum, R.N. *et al.* (2002) Comparison of sensory and consumer results with electronic nose and tongue sensors for apple juices. *J. Food Qual. Pref.*, **13**, 409–422.
36. Robie, D.J. (1997) Comparison of two electronic noses: three tobacco applications. *Seminars in Food Analysis*, 247–256.
37. Das, R. (2004) Development of electronic nose method for evaluation of HDPE data, correlated with organoleptic testing. *Packaging Technology and Science* **20**(2) (2007) 125–136.
38. Ohmori, S., Ohno, Y., Makino, T. and Kashihara, T. (2005) Application of an electronic nose system for evaluation of unpleasant odor in coated tablet. *J. Pharm. Biopharm.*, **59**, 289–297.
39. Zhu, L., Seburg, R.A. and Tsai, E. (2004) Flavor analysis in a pharmaceutical oral solution formulation using an electronic nose. *J. Pharm. Biomed. Anal.*, **34**, 453–461.
40. Kataoka, M. *et al.* (2005) Evaluation of the taste and smell of bottled nutritive drinks. *Int. J. Pharm.*, **305**, 13–21.
41. Sohn, H.J. *et al.* (2002) Discrimination of Ginseng habitat by using instrumental analysis techniques. Volume 106, Issue 1, 29 April 2005, Pages 7–12.
42. Branca, A. *et al.* (2003) Electronic nose based discrimination of a perfumery compound in a fragrance. *Sensors & Actuators B*, **92**, 222–227.
43. Ormancey, X., Puech, S., Bernard, A. and Coutière, D. (1998) New applications in olfactometry. *Seminars in Food Analysis*, **3**, 77–84.
44. Ormancey, X., Puech, S. and Coutière, D. (2000) Substantivity of fragrances on cloth. *Perfumer & Flavorist*, **25**, 24–29.
45. Poprawski, J., Boilot, P., and Tetelin, F. (2005) Counterfeiting and quantification using an electronic nose in the perfumed cleaner industry. *Sensors & Actuators B*, **116**, 156–160.

46. International Group Incorporation (IGI) (2004) Candle new waxes – investigation of the effects of each blend colour, fragance throw, wick choice, burn properties. *NCA White Paper*.
47. Boilot, P. *et al.* (2003) Sensor array system in the automobile industry: study on polymer textile films and foams, Pittcon Conference and Exposition, March 10–15 Orlando, Florida (2003).
48. Janata, J., Josowicz, M., and Sepcic, K. (2003) Sensing array for oil diagnostics, International Symposium on Olfaction and Electronic Nose, Riga, Latvia June 26–28, (2003).
49. Hansen, W., Van Der Bol, H., and Wieslemann, S. (2003) Combined use of electronic nose with purge and trap GC–MS in chemical shelf-life study, Pittcon Conference and Exposition, March 10–15 Orlando, Florida (2003).
50. Dewettinck, T., Van Hege, K., and Verstraete, W. (2001) The electronic nose as a rapid sensor for volatile compounds in treated domestic wastewater. *Water Research*, **35**, 2475–2483.
51. Gardner, J. W., Woo Shin, H., Hines, E. L, and Dow, C. S. (2000) An electronic nose system to monitoring the quality of potable water correlated with the quality assessments by an expert assessor/inspector. *Sensors and Actuators B*, **69**, 336–341.
52. York, R. K., Mitchell, S., and Cosham, C. (2002) Feasibility study on electronic nose evaluations of quality and tainting in fish correlated with the quality assessments by an expert assessor/inspector, International Symposium on Olfaction and Electronic Nose Rome, Italy, Sept 29 – Oct 2 (2002)
53. Bencsath, F. A. *et al.* (2000) Analysis of petroleum tainted seafood with electronic nose and gas chromatography/mass spectrometry. FDA Science Forum, Washington D.C., February 14–15, (2000).
54. Hudon, G. and Guy, C. (2000) Measurement of odor intensity by an electronic nose correlated with the quality assessments by an expert assessor/inspector. *Journal of the Air Waste Management Association*, **50**, 1750–1758.
55. Prica, M. (2001) Electronic nose studies of air quality monitoring in Australia. Field Screening Europe 2001, The 2nd International conference on strategies and techniques for investigation and monitoring of contaminated sites, Karlsruhe, Germany, May 14–16, 2001.

7

MS/Nose Instrumentation as a Rapid QC Analytical Tool

Ray Marsili
Marsili Consulting Group, Rockford, USA

7.1 INTRODUCTION

A mass spectrometer-based instrument specifically designed for rapid headspace quality control applications was first introduced commercially as the HP 4440A in 1998 by Hewlett Packard Co. (now known as Agilent Technologies). MS e-noses, like solid state e-nose instruments, were initially aimed at a wide variety of flavor and fragrance applications that seek to augment or even replace time-consuming and expensive human sensory evaluations. However, their usefulness extends into many other important industrial quality control applications.

Agilent Technologies, Inc. transferred training, support and further development of the HP4440A to Gerstel GmbH & Co. in 2002. The HP 4440A was renamed ChemSensor by Gerstel. While any GC/MS system can be used as an MS e-nose, the ChemSensor incorporates Pirouette multivariate analysis software, in addition to sophisticated macros and algorithms that automatically transfer the copious amounts of mass intensity data into a Pirouette-compatible spreadsheet format.

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

The ChemSensor was developed as a collaborative effort between Agilent, Gerstel, and Infometrics (Bothell, WA, USA), a recognized leader in multivariate statistical analysis software. ChemSensor software provides a variety of algorithms and macros that allow seamless data transfer and manipulation which greatly simplifies the calibration and prediction processing. A further advantage is that the ChemSensor software is embedded in the Agilent ChemStation program. The ChemSensor MS-based e-nose instrument can also be configured as a conventional GC/MS. New Agilent MS upgrades, such as improvements in MS sensitivity and inertness, can be incorporated into current versions of the ChemSensor.

Table 7.1 Advantages and disadvantages of solid state and MS e-nose instruments.

Technology	Average throughput	Primary advantages	Primary disadvantages
Solid state sensors	Moderate (5–15 min/ sample)	• Responds to a variety of analytes • Minimal sample preparation • Can be miniaturized • Easy on-line implementation	• Sensor poisoning • Sensor overloading • No structural information • Time-consuming calibration • Alcohol & water interfere • Short and long term drift due to changes in relative humidity, etc.
Headspace MS (GERSTEL ChemSensor)	Moderate (3–5 min/ sample)	• Responds to all volatiles • Minimal sample preparation • Fast method development • Structural information • Same instrument can be used as a GC-MS to determine specific chemical(s) responsible for out-of-spec samples	• Vacuum pump required • Subject to tuning inconsistencies • Can't distinguish optical isomers

In some respects, the ChemSensor is similar to the solid state electronic based e-noses that employ conducting polymer sensors, metal oxide sensors, surface acoustic wave (SAW) sensors, quartz microbalance (QMB) sensors, and so on. However, the differences between the MS-based approach and the solid state sensors are so significant that the MS approach probably should not be considered as a simple subset of e-nose instrument types. The ChemSensor MS instrument has been shown to overcome many of the limitations of the competing e-nose technologies. Table 7.1 compares the advantages and disadvantages of solid state sensor e-noses and the ChemSensor.

7.2 OPERATING PRINCIPLE

Figure 7.1 illustrates the difference between data collected using conventional GC/MS and the ChemSensor. It shows a conventional GC-MS chromatogram of a milk sample contaminated with 1300 ppm Matrixx, a sanitizer used in milk processing lines after cleaning. Milk contaminated with Matrixx has an oxidized-type flavor, which is difficult to distinguish by taste from light-abused milk. The active ingredients in Matrixx are peroxyacetic acid and octanoic acid. Oleic acid in milk can be oxidized to heptanal by hydrogen peroxide in Matrixx. The presence of three key chemicals – octanoic acid, acetic acid (a decomposition product of peroxyacetic acid) and heptanal – is a good indication that a potential off-flavor problem from sanitizer contamination is possible. While conventional GC/MS testing can be used to confirm Matrixx contamination of processed milk, the problem with this approach is that the chromatographic time (at least 20 minutes) and subsequent chromatogram interpretation are time-consuming and not conducive for use as a rapid screening test to monitor production samples.

A more practical approach for efficient screening of a large number of milk samples for sanitizer contamination is based on the application of the headspace ChemSensor. In this approach, the analytical column (normally 30 m long) is replaced by a retention gap, an uncoated fused silica capillary column approximately 1 m in length. Analysis time is reduced from 20 minutes to less than 2 minutes. What is sacrificed, however, is the ability to identify individual chemical components. By analyzing numerous control (normal-tasting) milk samples and several samples intentionally contaminated with sanitizer, multivariate analysis techniques can then be applied to 'train' the chemometrics software to distinguish mass intensity patterns that correspond to Matrixx-contaminated samples from mass intensity patterns that correspond

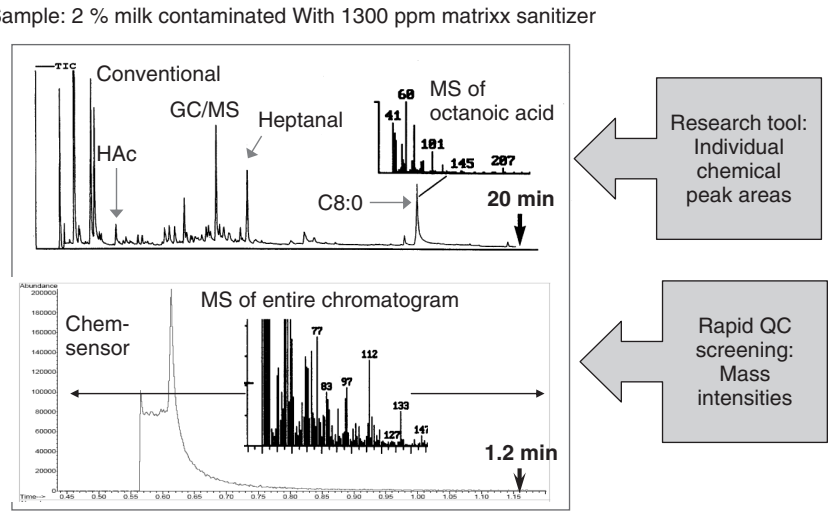


Figure 7.1 Conventional GC/MS vs. ChemSensor.

to control samples. Chemometrics software can be used to develop a classification model. Once created, the model can then be used to make class predictions (i.e., contaminated with sanitizer or not contaminated with santizer) of unknown milk samples. Once properly calibrated, the entire process from sample testing to sample classification is completed automatically at a rate of approximately one sample every two to three minutes. Several applications illustrating the usefulness of the MS-based e-noses have been published [1–12].

With solid state e-noses, the number and types of sensors must be specified prior to purchase. When the MS is used as the chemical sensor, preconfiguration is unnecessary. With the Gerstel ChemSensor, for example, the MS can scan from 2 to 1050 amu using the Agilent 5975 MSD. This is analogous to using 1049 solid state sensors. The user selects the scan range based on the sample matrix, expected analytes, and so on. Typically, a scan range of *m/z* 50–300 is a good starting point.

Mass intensity data generated for samples are converted into spreadsheet format by Gerstel macros. Data can then be transformed to useful visual formats using multivariate analysis algorithms with Pirouette software (Infometrix).

Whether using solid state or MS-based e-noses, the first step in method development is to calibrate the instrument with known samples belonging to pre-established classifications such as geographic areas, shelf life

age, flavor scores, and so on. If the goal is to classify coffee samples according to their geographic origin, several 'known' coffee samples (e.g., 5–10) from each geographical region must first be analyzed. The coffees in each group should span the range of chemical variation that would normally be encountered for each specific group. Once calibration has been completed, the MS intensities of unknown samples at each selected *m/z* serves as a fingerprint to compare to profiles obtained for the calibration samples. In effect, the approach is similar to creating a conventional GC/MS library for specific chemicals to which subsequent unknown chromatographic peaks in sample chromatograms can be compared. In normal GC/MS testing, the mass spectrum of a peak is compared to library mass spectra in order to identify the chemical. In the case of MS e-noses, the MS profile of all volatile chemicals in the sample is used for characterization – not just the MS profile of one chemical. In fact, standard GC/MS library tools have been used to identify products instead of individual chemicals [13]. Chromatographic peak resolution is unnecessary and sacrificed for speed of analysis.

Results for calibration samples (e.g., in the coffee example, coffees from various known geographic regions) should be examined using the exploratory techniques of principal component analysis (PCA) and/or hierarchical cluster analysis (HCA), as has been previously described in Chapter 5 in this book. Samples from the same class should cluster together in PCA plots. If not, preprocessing of data should be conducted. If this fails to provide class clustering, inspection of PCA loadings, residuals, and outlier diagnostic plots should be scrutinized to determine which *m/z* values and/or samples could be excluded to improve class clustering.

If the ultimate goal is to determine the classification of an unknown sample – for example, if an unknown coffee sample is Guatemalan, Brazilian, Jamaican, and so on – then mass intensity profiles of unknowns can be compared to the calibration samples. Classification of unknowns can be made by creating models using the multivariate algorithms of *k*-nearest neighbor (*k*-NN) or soft independent modeling of class analogy (SIMCA). Once the models are created, they can be used to make class predictions of unknown samples.

If the goal of multivariate analysis studies is to predict flavor scores, shelf life or some other discrete continuous property, then principal component regression (PCR) or partial least squares (PLS) can be used to prepare models and predict the discrete continuous property of unknowns.

7.3 ADVANTAGES OF MS OVER SOLID STATE SENSORS

Solid state e-noses are subject to interference from water and alcohol in samples. This problem not only reduces sensitivity to more important differentiating chemical components but also lowers sample throughput because of the time required to restabilize sensors. Another serious problem is that solid state sensors can be poisoned by sulfur compounds and other strongly adsorbing components.

Even though a solvent peak or a volatile present in high concentration in the sample may generate a large signal that saturates the MS detector at one or several *m/z* settings, the performance at other *m/z* settings is largely unaffected. An example of this is the analysis of alcoholic beverages by static headspace. Ethanol is present at relatively high concentrations compared with other flavor-active components in the sample. Overloading solid state detectors with ethanol can be problematic for most solid state sensors. However, alcoholic beverage applications with the ChemSensor are not impacted by this problem. (See classification of whiskey samples by brand type using the ChemSensor presented in Section 7.9.)

It is not uncommon for solid state e-noses to demonstrate long- and short-term drift due to changes in relative humidity. In addition, individual sensors have to be replaced periodically. In contrast, changes in the external environment have minimal effects on mass spectrometers. High stability is critical for any analytical technique that relies on standard library searching.

It should also be noted that MS sensitivity is affected by the amount of sample analyzed, the headspace vial temperature, and the scan range. The sensitivity of solid state sensors is determined by their type, the flow rate of sample vapors over the sensor, the analyte, and the temperature.

7.4 USING OTHER SAMPLE PREPARATION MODES

Static headspace is a good sample preparation choice for many applications. It offers advantages over other sample preparation techniques commonly used prior to GC analysis for flavor and fragrances. It is highly reproducible; whereas solid phase microextraction (SPME) may suffer from fiber-to-fiber variations, static headspace syringes and other types of automated instrumentation used for static headspace GC analysis are more consistent. As a solventless technique, headspace sampling

has the significant advantages of reducing artifact peaks and minimizing background contamination. Unfortunately, static headspace also suffers from significant disadvantages. It is not particularly sensitive for some types of analytes, and it does not work well for high boiling point compounds. In many cases, SPME and other sample preparation techniques are a better choice. Selection of optimum sample preparation/extraction techniques can be key factors to success with e-nose applications

7.5 TECHNIQUES FOR IMPROVING RELIABILITY AND LONG-TERM STABILITY

7.5.1 Calibration Transfer Algorithms

A major benefit of using the MS based nose is its fast analysis time. Success with e-nose instruments requires training the sensor with standard samples. The mass spectra from the standards are used to create chemometric models. Classification of unknown samples is obtained by comparing their mass spectra to those in the models.

The accuracy of predictions is dependent on the quality of the model, and reliable models are created with numerous replicates [14]. Construction of calibration models is time-consuming and therefore costly. If the MS sensor is disturbed – for example, if the filament is changed or MS maintenance is required – recalibration of the instrument may be necessary to compensate for the new instrument conditions. An alternative to recalibration is the use of calibration transfer algorithms.

The minor differences in MS fingerprints caused by replacing filaments, retuning, and cleaning can significantly reduce the reliability of the chemometric models. To compensate for these minor differences in MS profiles, two alternatives can be considered. One option is to create a new model with new data; this can be time-consuming. Another alternative is to use a computational adjustment that compensates for instrument differences. This second approach is known as transfer of calibration (TOC).

The TOC algorithms available with commercial units adjust profiles obtained with the new set of instrument conditions to look like those collected before any parameter changes occurred. In a recent study [15], a series of food samples and individual compounds were used to monitor the reliability of the calibration transfer algorithms for a period of 10 weeks. The instrument was disrupted in three ways: (a) filaments were replaced, (b) tuning algorithms were changed, and

(c) preventive maintenance was conducted. Instrument drift over the 10 week period was also investigated in a separate system and different types of calibration transfer algorithms were examined. For this study the TOC method known as direct standardization was used. Direct TOC relates variables in the data collected with initial instrument settings to each corresponding variable measured after slightly different instrument conditions have been made. Calibration transfer algorithms can be performed for either quantitative (e.g., PLS or PCR) or qualitative models (*k*-NN and SIMCA). In this study, the efficiency of the calibration transfer was investigated using SIMCA and PLS models.

Results of this study showed that successful use of calibration transfer is highly dependent on the type and number of transfer samples and the model type. Overall *k*-NN models appear to be more robust than SIMCA models. The best accuracy was obtained using *k*-NN models and SIMCA models with calibration transfer. Improved accuracy can be obtained if the calibration transfer algorithm is based on reanalyzing one or more samples from each class observed in the original chemometric calibration.

7.5.2 Internal Standards

Internal standards can be applied to MS e-nose applications as a way to improve precision and accuracy of results – especially when testing is conducted over an extended period of time (e.g., several weeks or longer). There are several ways that internal standards can be incorporated into MS e-nose studies. One approach is to add a chemical (the internal standard) that contains a significant *m/z* peak which is not present in chemical components in samples. In a study designed to predict the shelf life of processed milk, chlorobenzene was added as an internal standard to all milk samples [16]. All mass intensities in the samples were divided by the mass intensity of the *m/z* 112 peak, the molecular ion for chlorobenzene. Results based on these normalized mass intensities improved accuracy for long-term shelf life prediction of milk samples compared with shelf life prediction results obtained without the internal standard.

Another approach to applying an internal standard is utilizing fast GC or partial GC separation of analytes. An internal standard, which elutes early or late in the chromatogram, can be added to all the samples. The peak area or the mass intensity of the molecular ion (or some other ion with a large abundance) can be used to normalize mass intensities obtained for the remainder of the chromatogram.

7.6 TWO INSTRUMENTS IN ONE

Since commercial e-nose instruments are relatively expensive, careful consideration should be given to which instrument technologies afford the most flexibility, long-term usefulness, and greatest chance for success. The ChemSensor allows for both conventional GC/MS testing for QC and research applications, as well as e-nose type work. A typical analytical capillary column used for normal GC/MS testing can be left in place for MS e-nose applications. In the MS e-nose mode, the column is operated isothermally at elevated temperatures so sample analytes can be eluted from the column in 4 minutes or less. There is no need to switch the analyte column with a meter-long retention gap as was commonly done in the past. This saves considerable down time; there is no need to pump down the MS since switching from an analytical column to a retention gap and back again is avoided.

Furthermore, information learned in conventional GC/MS experiments can be applied to assist in the development of robust, accurate e-nose methods with the same instrumentation. For example, if one is differentiating between multiple groups and known chemical compounds are already known which are unique to the groups, then this information can be used to ensure that ions from these compounds are used in the resulting models to ensure causative models and not just correlative models. Using GC/MS experiments to optimize e-nose methods is ideal. Another advantage of this configuration is samples diagnosed as out-of-specification (i.e., not belonging to a desired or expected classification) can easily be reanalyzed by conventional GC/MS on the same instrument to verify results and to better understand which specific chemical(s) may be causing the out-of-specification classification. Configuring a flexible system capable of performing both e-nose type work and conventional GC/MS analyses offers many advantages.

7.7 APPLICATION EXAMPLES

MS e-noses have been used in a wide variety of applications from many industries. A few examples include:

- quality monitoring of goods received and shipped,
- quality, odor, and product consistency issues,
- monitoring surfactants and odor compounds,
- testing product quality and contamination,

- product shelf life and residual solvent analysis associated with packaging materials, and
- out-gassing of automobile interior products.

The following examples illustrate how PCA clustering of similar samples types can be made with the MS e-nose approach.

7.8 CLASSIFICATION OF COFFEE SAMPLES BY GEOGRAPHIC ORIGIN

Figure 7.2 shows classification clusters of four different coffee types (Guatemalan, decaffeinated Guatemalan, Sumatran, and decaffeinated Sumatran) generated from static headspace GC-MS testing using the Gerstel ChemSensor. The independent variables considered were mass intensities for 120 different atomic mass units (*m/z* 51–170). Inspection of Figure 7.2 shows one of the decaffeinated Guatemalan samples is potentially an outlier. A two-dimensional plot of sample residuals vs. Mahalanobis distances provides a good indication of samples that are possible outliers and should be excluded from data analysis (Figure 7.3). Figure 7.4 shows a three-dimensional PCA plot of the four classes of coffee samples after the decaffeinated Guatemalan outlier was excluded from the data-set. With excellent clustering of similar sample types, these results could be used to create a *k*-NN model for class prediction. The model could then be used to analyze coffee samples of unknown origin in order to determine their proper geographic class origin. This example

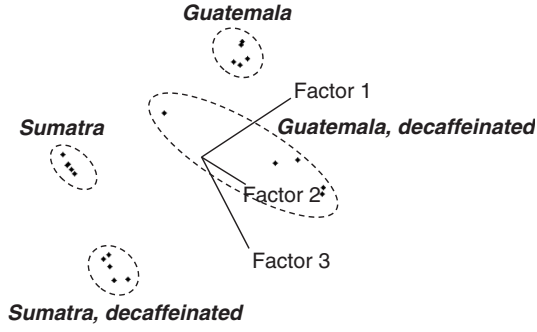


Figure 7.2 Three-dimensional PCA plot of four classes of coffees (five samples for each class) based on ChemSensor static headspace testing using mass intensity results from *m/z* 51 to 170.

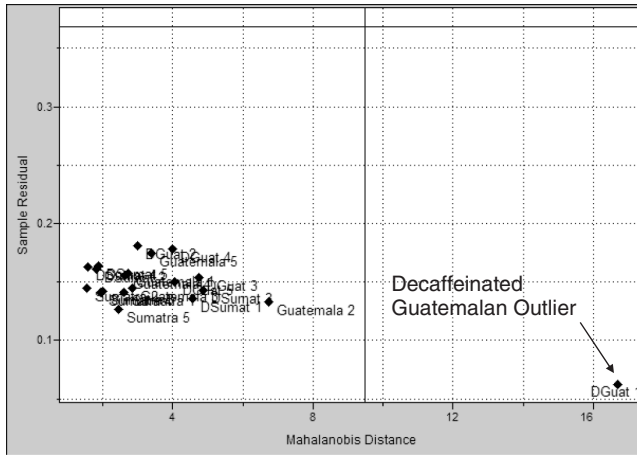


Figure 7.3 Determination of decaffeinated Guatelmalan outlier sample from Mahalanobis outlier diagnostics.

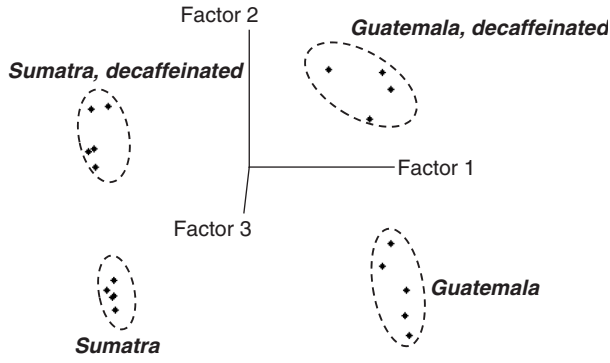


Figure 7.4 Three-dimensional PCA plot of four classes of coffees based on ChemSensor static headspace testing using mass intensity results from 51 to 170 amu after deletion of decaffeinated Guatemalan outlier.

illustrates the diagnostic power of chemometric software to determine outliers.

'Loadings' plots can be used to determine the independent variables most influential in clustering of sample classes (Figure 7.5). In this case, the loadings plots for Factor 1 show that mass intensities for *m/z* 52, *m/z* 60, and *m/z* 79 were most influential, and for Factor 2 the most important masses for clustering were the same three masses plus *m/z* 81, *m/z* 95, and *m/z* 98. These results could indicate that acetic acid (*m/z* 60) and pyridine (*m/z* 79), which are known to be important flavor

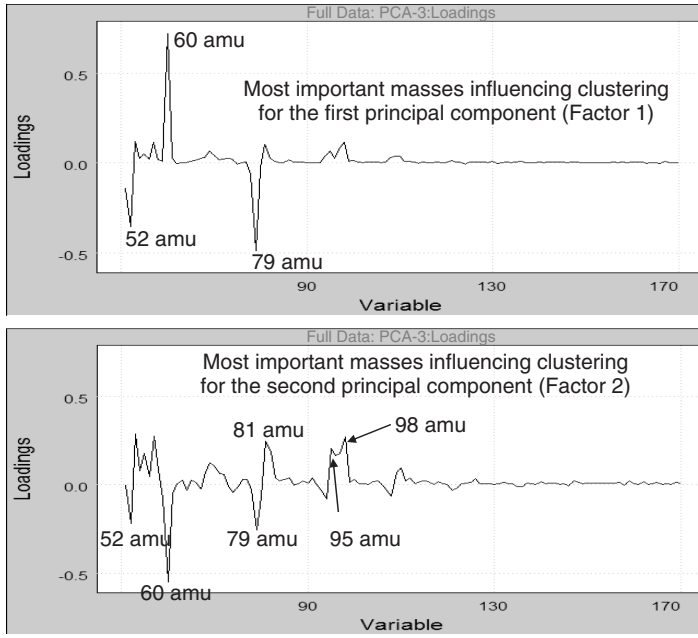


Figure 7.5 Examination of PCA loadings plots for coffee samples to determine masses most influential to clustering.

contributors to some coffees, are likely to be strongly responsible for differentiating classes for these four types of coffees. A follow-up GC/MS study done with the same GC/MS instrument which was used for the e-nose work could confirm the suspicion that acetic acid and pyridine are important class influencers.

Information revealed in the loadings plots could be used to facilitate efficient identification of the important constituents that define class clustering. Examining classical chromatographic separations of sample constituents with extraction ion or SIM plots based on key masses indicated in the loadings plots could help identify the important constitutents.

7.9 CLASSIFICATION OF WHISKEY SAMPLES BY BRAND

In this application, three samples each of four different brands of whiskey were analyzed with an MS e-nose. In addition, two 'blind' samples of whiskey were submitted. The goal of this work was to classify the

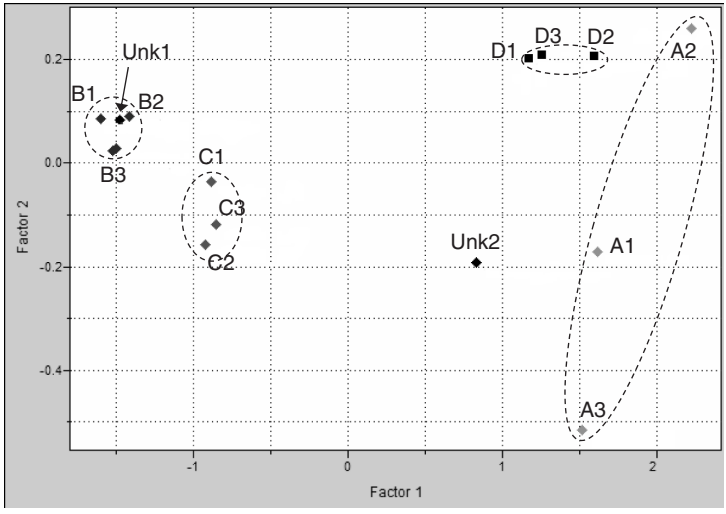


Figure 7.6 Two-dimensional PCA plot of four classes (brands) of whiskey (three samples for each brand) based on ChemSensor static headspace testing using mass intensity results from *m/z* 51 to 170.

unknown samples according to type of brand. Since a relatively small number of samples was submitted, brand predictions of blind samples were made based on the *k*-nearest neighbor (*k*-NN) algorithm. The PCA multivariate analysis plot (Figure 7.6) showed the four brands of whiskey samples formed four distinct clusters, thus permitting accurate brand classification of the 'blind' whiskey samples submitted for analysis.

First, 5 mL samples were incubated for 20 min and heated to 60 °C. A 1 mL aliquot of equilibrated headspace gas was injected with the heated (75 °C) 2 mL gas-tight syringe using the MPS2 robotic sampler. Mass intensities from *m/z* 46–157 were measured. The following masses were excluded: *m/z* 64, 66, 80, 82, 90, 92, 93, 98, 100, 106–114, 118–126, 128, 130–154, and 156. Samples were injected into a 30 m × 0.25 mm i.d. × 0.25 μm DB5 column heated to 220 °C. The GC run time was 3 min.

Based on *k*-NN, one of the blind whiskey samples was classified as belonging to Brand A and the other blind whiskey sample was classified as belonging to Brand B. This agrees with estimates of brand predictions based on a simple visual inspection of the 2D-PCA plot in Figure 7.6. The company that submitted the sample set for analysis confirmed that the blind samples were properly identified by brand with the ChemSensor. The company was unable to assign proper brand identity for these

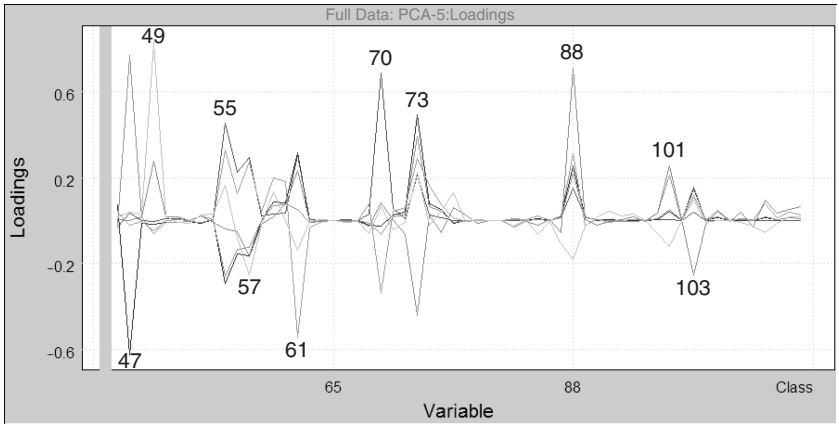


Figure 7.7 Examination of PCA loadings plots for whiskey samples indicates masses most influential to brand clustering for the first six factors.

blind unknowns by simple visual inspection of their traditional GC/MS chromatograms with no statistical treatments.

Inspection of the loadings plot in Figure 7.7 reveals the masses that are most important for modeling for the first six factors. Masses 88, 101, and 103 are a good indication that esters are important influencers for clustering. It should be noted that static headspace is not a good technique for measuring free fatty acids (FFAs). FFAs may be important for differentiating brands of whiskey and should be extracted by another analytical sample preparation technique (e.g., SPME with the DVB/PDMS fiber). A rapid static headspace GC/MS e-nose test was developed for the rapid identification of whiskey samples by brand in less than 2 days of experimentation using a 4 min test procedure.

7.10 FUTURE DIRECTIONS: PARTNERING MS/NOSE WITH GC/MS

GC/MS is an excellent tool for food research applications but unfortunately is not well suited to routine QC applications. While it can provide a good deal of important chemical information about raw materials and finished products, analysis time can be lengthy. Conversely, speed of analysis and ease of interpretation of results make the MS e-nose well suited to many routine QC applications. Unlike GC/MS testing, however, MS e-nose instruments do not normally provide details about the

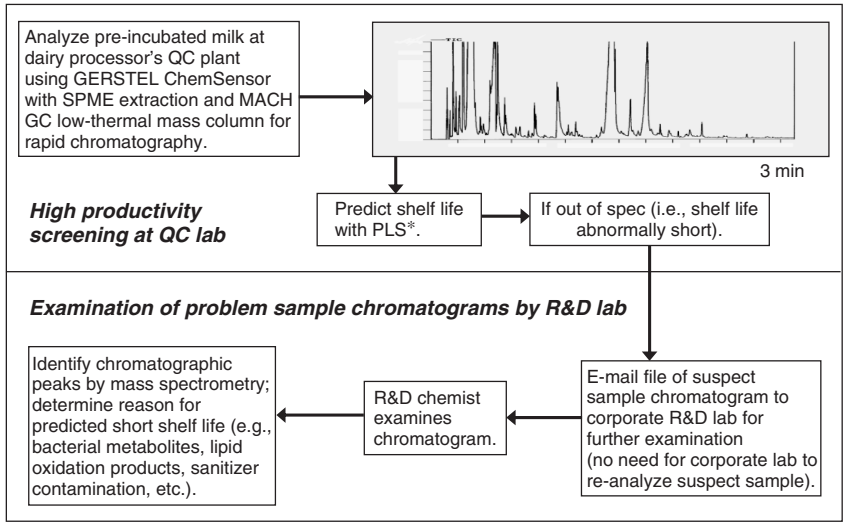


Figure 7.8 Partnering MS-nose and GC/MS technologies to predict the shelf life of processed milk and the reason for premature off-flavor development.

amounts and kinds of specific chemicals that are different from sample to sample.

There is a good way to team the rapid prediction benefits of the ChemSensor with the superior ability of GC/MS to provide details about specific chemicals. A flow chart explaining the strategy appears in Figure 7.8. Previous research has shown that it is feasible to predict the shelf life of processed milk accurately using a ChemSensor [9]. Once appropriate robust PLS models are developed for predicting shelf life, a dairy QC lab can test incoming processed milk samples using a ChemSensor and a fast GC method. With this approach, analysis can be completed by the QC lab in 3–5 min per sample. Mass spectra intensity data can be imported into Pirouette and a rapid prediction of the milk shelf life can be made with the PLS models. If a sample is flagged as being out-of-specification with an unusually short predicted shelf life, the actual chromatogram for the suspect sample can be further scrutinized to determine the cause of premature off-flavor development. The file for the GC/MS chromatogram can be emailed to a corporate analytical research lab. The corporate lab could have the staff expertise and time to examine the chromatogram carefully in order to identify specific chemicals responsible for the off-flavor and perhaps deduce the mechanism of off-flavor development. This can all be accomplished by

performing only one GC/MS analysis of the sample at the QC lab. This strategy is an ideal way of overcoming the limitations of both MS e-nose and GC/MS approaches to real world problem solving in an industrial laboratory environment. This strategy is an ideal way to overcome the inability of the MS e-nose to identify specific chemical constituents and the unsuitability of conventional GC/MS to provide rapid screening of a large number of samples due to the slow speed of generating and interpreting chromatograms.

7.11 CONCLUSION

MS e-noses offer the following advantages:

- less drift (proven bench-top MS technology),
- fast (2–5 min per sample),
- able to tolerate water,
- results not hampered by presence of alcohols, sulfur-containing chemicals, or other polar compounds,
- scan range determines the number of sensors,
- linear to 10^4,
- no poisoning of sensors,
- correlation with GC/MS,
- capable of identifying ions that differentiate two samples, and
- can be combined with GC/MS to help elucidate the actual mechanism of off-flavor development in foods and beverages.

One of the most important benefits of using MS as a sensor array in an e-nose instrument is that it can be used to determine not only if a test sample is different from a standard sample but also why it is different. The ability to use MS e-nose instruments as both a rapid screening productivity tool and a research tool to provide more details about specific chemical components in samples is an appealing combination.

REFERENCES

1. Pena, F., Cardenas, S., Gallego, M., and Valcarcel, M. (2002) Characterization of olive oil classes using a ChemSensor and pattern recognition techniques. *Journal of the American Oil Chemists Society*, **79**, 1103–1108.
2. Marsili, R.T. (2002) Combining mass spectrometry and multivariate analysis to make a reliable and versatile electronic nose, *Flavor, Fragrance and Odor Analysis* (ed. R. Marsili), Marcel Dekker, New York, USA p. 349.

3. Kinton, V.R., Collins, R.J., Kolahgar, B., and Goodner, K.L. (2003) Fast analysis of beverages using a mass spectral based chemical sensor, *Gerstel Application Note 4/2003*.
4. Kinton, V.R., Whitecavage, J.A., Heiden, A.C., and Gil, C. (2004) Use of a mass spectral based chemical sensor to discriminate food and beverage samples: olive oils and wine as examples, *Gerstel Application Note 1/2004*.
5. Kinton, V.R., Pfannkoch, E.A., Mabud, M.A., and Dugar, S.M. (2003) Wine discrimination using a mass spectral based chemical sensor, *Gerstel Application Note 2/2003*.
6. Kolahgar, B. and Heiden, A.C. (2002) Discrimination of different beer sorts and monitoring of the effect of aging by determination of flavor constituents using SPME and a chemical sensor, *Gerstel Application Note 11/2002*.
7. Marsili, R.T. (2001) SPME-MS-MVA as a rapid technique for assessing oxidation off-flavors in foods, *Headspace Analysis of Foods and Flavors: Theory and Practice/Advances in Experimental Medicine and Biology, Volume 488* (eds R.L. Rouseff and K.R. Cadwallader), Kluwer Academic/Plenum Publishers, New York, USA, p. 56.
8. Pavon, P.J. et al. (2003) A method for the detection of hydrocarbon pollution in soils by headspace mass spectrometry and pattern recognition techniques. *Anal. Chem.*, **75**, 2034–2041.
9. Lorenzo, M.I. et al. (2002) Application of headspace-mass spectrometry for differentiating sources of olive oil. *Anal. Bioanal. Chem.*, **374**, 1205–1211.
10. Pena, F., Cardenas, S., Gallego, M., and Valcarcel, M. (2003) Direct sampling of orujo oil for determining residual hexane by using a ChemSensor. *Journal of the American Oil Chemists Society*, **80**, 613–618.
11. Vinaixa, M. et al. (2004) Early detection of fungal growth in bakery products by use of an electronic nose based on mass spectrometry. *J. Agric. and Food Chemistry*, **52**, 6068–6074.
12. Saevels, S. et al. (2004) An electronic nose and a mass spectrometry-based electronic nose for assessing apple quality during shelf life. *Postharvest Bio. Tech.*, **31**, 9–19.
13. Goodner, K.L. and Kinton, V.R. (2006) Using a mass spectrometer library matching system to identify food products. *Proc. Fla. State Hort. Soc.*, **119**, 383–387.
14. Goodner, K.L., Dreher, J.G., and Rouseff, R.L. (2001) The dangers of creating false classifications due to noise in electronic nose and similar multivariate analyses. *Sensors and Actuators B*, **80**, 261–266.
15. Heiden, A.C. et al. (2003) Use of calibration transfer algorithms on a mass spectrometry based chemical sensor – preliminary results. *Gerstel Application Note 3/2003*.
16. Marsili, R.T. (2000) Shelf-life prediction of processed milk by solid-phase microextraction, mass spectrometry and multivariate analysis. *J. Agric. and Food Chemistry*, **48**, 3470.

8

Sensory Analysis

Carlos Margaria
US Distilled Products, Princeton, USA

Anne Plotto
USDA ARS Citrus and Subtropical Laboratory, Winterhaven, USA

8.1 INTRODUCTION

This chapter is intended to serve as a reference for the inexperienced sensory analyst. Our goal is to help reduce to a minimum the number of things that could go wrong and help make the sensory analysis experience easier and the experimentation more accurate and enjoyable. This chapter will concentrate on the everyday problems that the novice investigator will face, and hopefully will insure that the data gathered will be transformed into useful information, regardless of the experimental design or the technique selected. We will leave the in-depth exploration of sensory science to the excellent books already available from many renowned authors [1–4] and will concentrate on the different stages of sensory analysis experimentation, such as panel selection, panel organization, available sensory techniques, and expression of results.

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

8.2 THE PURPOSE OF SENSORY ANALYSIS

Sensory evaluation represents the only moment during research and development where the experimenter is in touch with the final judge: the consumer and his/her perception. Sensory analysis has unique advantages over other methods. 'Is my product "*good*"?' is a question that no analytical technique could answer. It requires knowledge of psychology, physiology, and statistical methods. Appropriate experimental design and thoughtfully designed questionnaires will help find out the details of why is it '*good*', what '*good*' means, is it '*good*' for whom, and so on. Sensory science presents unique challenges that might make people who are unfamiliar with the concept think that it is not as accurate or reliable as other disciplines. Sensory science is accurate and reliable if conducted properly and this chapter aims to help the experimenter use appropriate methodology and techniques to ensure reliable and accurate results.

Product development, quality control, product matching, shelf life studies, product reformulation, and product acceptability are the main areas where sensory analysis is of central importance. Each of the following areas or scenarios will have a slightly different approach. The following paragraphs will briefly discuss these scenarios.

(1) In new product development, the analysis will involve developing 'from scratch' a product that matches some given guidelines, often generated by the marketing department based on an identified niche or opportunity. This case gives the most freedom to experiment with different ideas and interpretations in order to fulfill the specified need, but it is probably the most challenging of all scenarios, since there is no tangible product to target. It is important for the researcher to keep the marketing department involved in order to monitor that the progress of the prototype corresponds to their expectations, and to adjust the prototype according to the final goal, while recording the findings throughout the whole process. After the prototype(s) starts to take shape, the product developer should consider the possibilities and hurdles that were not evident in the original idea, and address them to develop a successful final product. In addition, packaging can be an integral part of the characteristics of the product, since it provides not only physical protection and attractive graphics and labels, but also convenience. Packaging extends shelf-life in a way that is intimately related to the physical-chemical-sensorial characteristics of the product (examples are with ready-to-eat products, gas barrier films, modified atmosphere packaging, etc.).

The other possible scenarios where sensory analysis is used share one important characteristic: there is already at least one existing product that needs to be matched, improved, or optimized.

(2) The quality control scenario is arguably the easiest one, where sensory analysis is used to verify that the current product matches the characteristics of the original formulation within acceptable margins.
(3) In shelf life studies, a given product or prototype is tested against itself; the only variable should be storage conditions. In these cases, different ingredients such as antioxidants, colorants, and preservatives can be tested to determine their effect on the overall shelf life of the product. Also, as discussed later, microbial evaluation is a main factor to be considered in sensory analysis.
(4) In product reformulation, the most common cause for such a request is that one of the ingredients has been discontinued, or its cost increased dramatically. In both cases, the 'problem' ingredient needs to be replaced by a similar one to yield a product that matches the characteristics of the original formula.

Learn from others' experience:

In some cases it could be difficult to believe that just a price change could require a reformulation of a product, considering that it might also imply label, packaging, or process changes. This was the case when supplies of vanilla extracts dropped from 2002 to 2005 due to a natural disaster. In this case a hurricane in Madagascar, an island nation in Southeast Africa, which is the most important producer of vanilla beans, resulted in a threefold increase in the price of vanilla extract. This situation forced most vanilla extract users to find alternatives that would allow them to continue producing products at a reasonable cost.

(5) Product acceptability deals with an existing prototype that needs testing in a simulated commercial situation to determine whether it is acceptable to consumers, or if further development is needed, to adjust the product to the market preferences, or even to find a niche market for that specific product. This latter situation should be explored by the marketing department, prior to any product development, using techniques such as focus groups to determine the acceptance of a new concept. In some cases, when a company

has an existing successful product in a different market (such as in a different country or region) and wants to introduce it to a new market, the potential customers need to be identified.

There are two different approaches an experimenter can use to answer the ultimate question: (a) sensory analysis by means of a trained or expert panel, or (b) consumer testing. The researcher must bear in mind that the information that one or the other panel techniques provide is essentially different, and one cannot be substituted by the other. For instance, a typical questionnaire for a trained panel might include questions such as: is this product different from the standard/target? How different is it? How would you rate this product's intensity of sweet/sour/bitter/salty? All this information would be invaluable for product development, fine tuning of prototypes, matching existing products, and for reformulations. A trained (or expert) panelist has the tools necessary to effectively communicate specific traits of a certain product. The panel leader through training sessions, lexicon development, panel discussions and evaluation of standards provides those tools [2, 5]. Once trained, a panelist should not be asked consumer preference questions because his/her judgment would then not be objective due to bias from knowing too much about the product [1]. In addition, trained panelists are screened and selected to discriminate small sensory differences between samples, and are not at all representative of an untrained consumer, who will be selected for usage of the target product. However, in product development, preference tests can be performed with in-house personnel, provided respondents are not closely involved with the product of interest, and provided enough panelists are available for statistical validity. These in-house panels may be valuable in pilot studies to fine tune a questionnaire before a large consumer study is conducted [6]. In a consumer study, panelists should not be asked to rate the intensity of an attribute, since they have not been trained to recognize a specific descriptor or for the use of a scale.

For consumers, the set of questions should aim at exploring their emotional response to the product: do you like this product? Which of these two products do you prefer? Would you buy it? What would make this product more appealing to you? The objective of these types of tests is to measure the degree of liking, preference, or purchase intent of a product.

The rationale behind this separation is simple: consumers always know what they like, but not necessarily why they like it or how to express their reasons verbally. There are also many psychological reasons behind product selection and preference, but psychological analysis belongs to market research and other disciplines [6].

8.3 FLAVOR PERCEPTION

In an analytical laboratory, where chemical, microbiological, or physical analyses are performed, data acquisition can be very straightforward. Such analyses involve equipment that is readily available, and easily calibrated as needed. Results are reliable and easy to interpret (e.g., pH measurements, refractometry, acid level by titration, chromatograms, etc.). In the case of a sensory panel, the 'instrument' that would 'measure' the samples and yield data, is a group of humans (or animals if pet food is tested) made up either of trained members or consumers. Panels are comprised of people who respond differently to stimuli, have different perception thresholds, and different histories of food products experience. In addition, the way the sample is presented can strongly influence the results. For example, colors strongly influence the perception of flavors [2, 7]. A simple exercise in an introductory sensory science course consisted of adding flavors to colored jell-O, and asking panelists to identify the flavor. A green jell-O with lemon flavor will be perceived as having mint or lime flavor, or red jell-O will be described as having cherry flavor, even if strawberry was the flavor added. The same aroma-active volatile compound can be perceived with a much lower intensity, or not perceived at all under a certain threshold concentration, if it is presented in a matrix that binds the compound [8]. It can even be given different descriptors depending on the concentration in the food [9]. Finally, people can represent a large variation in the perception of odor and taste due to their genetic make up. For instance, partial anosmia, where people can perceive all odorants except one, has been extensively described by Amoore [10–12]. In our laboratory, we have shown that 50% of the population expressed partial anosmia to β-ionone [13]. Variations in the perception of odorous compounds between individuals can be at the olfactory bulb, olfactory receptor level, due to enzymes in the nasal tissue [14] or in the oral cavity [15]. It has recently been discovered that the enzymes can change one odorous molecule into another and, therefore, change its quality. Not all individuals have the same enzymatic activities and receptors, and thus the same molecules can be perceived quite differently, based on biochemistry and physiology. All this makes the work of the sensory scientist and panel leader challenging, and even with the best panelists and training practices, human subjects are not exactly like instruments. That is why it is so important that the panel leader keeps the panel trained, so that they can provide reliable measurements.

Sensory evaluation may not be able to analyze all differences at the molecular level, but that does not prevent it from yielding excellent

results. Sensory analysis should be viewed as a multicomponent approach to a problem that presents synergistic interactions. Perception is not the effect of one factor only. Before even putting something into our mouths, there is an avalanche of external sensations already bombarding our brain: color, visual texture, tactile texture, temperature, presentation, packaging, environment, and first odor impression. In addition, the physiological state of the subject affects how a food is smelled and tasted. By the time the food enters the mouth, there is already a lot of information that will bias the second set of sensations – flavor, mouthfeel, sound while chewing – which will confirm or not the expectations derived from the first set of impressions. The experimental design should consider all these components and be able to isolate the variable(s) of interest, keeping variation of other components at a minimum by standardizing conditions. For example, red lighting should be used to mask color if color might induce a bias, and tests for odors should be performed in booths equipped with positive air pressure and air filtration to minimize the influence of outside environmental odors, resulting in odor cross contamination.

Flavor-matrix interactions are of central importance, but usually forgotten and poorly understood due to the huge number of variables that could be involved, so that matrix interaction is not easily categorized. Macromolecules, gums, pH, chemical composition, chelating agents, particle size, and so on, can modify the original flavor perception, enhancing, suppressing or – in extreme cases – changing its profile.

Most processed foods these days include one or more added flavors, which provide a characteristic profile for each specific food matrix. This is where the genius of flavorists shines, since they are able to find the balance of flavoring substances that result in the desired flavor perception for a specific food item. This flavoring, once combined with the food matrix, should be able to produce adequate odor and flavor release in spite of interactions with macromolecules (which could mask their perception) or substances like ethanol, sugar, and acids (which could enhance it).

8.4 SENSORY ANALYSIS TECHNIQUES

The literature provides different classifications for sensory tests such as discrimination tests, time intensity methods, acceptability methods, descriptive methods, and so on [1, 2, 4, 6]. There is a point that should be emphasized here, to make sure that the nature of the problem and the

limitations of the different approaches and techniques are understood. These different techniques are all based on statistical models, which would give better or worse predictability and errors according to the power of the different tests, the statistical distribution involved, and the number of subjects evaluated. For instance, the widely used nine-point hedonic scale, a 15 cm intensity scale (with or without anchor points), or an intensity rating from 0–100, yield results that have been proved to be statistically sound and are widely accepted. The novice experimenter should be aware of these conventions and use them wisely, not being afraid to test alternatives which might yield better results and could be more appropriate from a practical standpoint [1, 16].

Another important point is: what question do you want to answer? Often, an experimenter or product developer would come to a sensory analyst and say: 'I want to know if there is any difference between my experimental products'. What type of difference does this person want to know? 'Can a consumer detect if there is any difference between the two products?', 'The products are different, but how large is the difference, and how can the difference be qualified or quantified?', or 'Would a consumer prefer one product over the other?'. The first two questions do not address quality on an emotional level (preference, liking), while the third question does. Different techniques can answer each of these general questions. The three major classes of sensory tests are difference tests, descriptive tests, and acceptance tests, each of which can be used to answer the above three questions, respectively. It is important to know in which case to use these different techniques, and the limitations and advantages of each.

8.4.1 Overall Difference Tests

These tests determine whether a general difference is perceived between two and up to five samples. If the difference between two products is obvious, such tests should not be performed as they would be a waste of time and resources. Difference tests can also be used to determine whether two samples using different ingredients are similar enough to be used interchangeably, for instance, if two flavors can be used in a product, the cheaper or easier to source would be preferred from a manufacturer's standpoint. Before performing a difference or similarity tests, one should determine the parameters for sensitivity, α, β, and p_d. The α-risk, the most commonly used in difference testing, is "the probability of concluding that a perceptible difference exists when there is actually no difference" [2]. The β-risk, used in similarity tests, is the probability

to conclude that there is no perceptible difference, when actually there is, and p_d is the proportion of distinguishers. The α-risk (β-risk) is often used at a confidence level of 5% or 1% (0.05–0.01), and it indicates strong evidence that a difference (similarity) is apparent. A higher risk (α of 0.05 to 0.10) only indicates moderate evidence [2]. Meilgaard *et al.* [2] have developed a spreadsheet application to help researchers selecting values for α, β, and p_d depending on the desired sensitivity and available resources, mainly the number of panelists. For example, a lower risk might be considered if one wants to determine the storage life of a product, perhaps limited by the appearance of an off-flavor.

8.4.1.1 Triangle Test

This is one of the most known and used tests. Here, panelists are presented with three coded samples; they are told that two samples are the same, and one is different, and they are asked to choose the different or odd sample, based on visual, tactile, odor, or taste cues. Correct answers are tallied and compared to the values in the appropriate standards table (Table 8.1) for interpretation of the results.

There are six possible presentations of the samples: AAB, ABA, BAA, ABB, BAB, and BBA. Combinations should be presented in a random but balanced order, that is, A or B presented first an equal amount of times. For that reason, it is preferred to have a multiple of six as the number of panelists. As little as 20, and up to 40 panelists are acceptable for a difference test using the triangle method; however, more (50 to 100 panelists) are necessary for similarity testing [2].

Some space might be provided on the ballot sheet for panelists' comments (they may write down why they thought there was a difference), but affective questions should not be asked, as selecting the odd sample might bias the answer to which sample is preferred. Because there is a 33% chance of guessing the correct answer, this test is supposed to be more robust statistically than other difference tests. However, if the products tested tend to have carry-over, a lingering effect, or involve sensory fatigue or adaptation, other tests should be considered.

A similar test is the two out of five test. In this test, two coded samples of A and three coded samples of B (or two Bs and three As) are presented, and panelists are asked to select the two samples that are different from the other three. Because of the higher number of samples, this test is mostly recommended when there is no sensory fatigue, usually for visual or texture (feel) attributes. This test is statistically more robust than the triangle test and any other difference test, because the chance of guessing correctly two out of five samples is only one out of 10 (10%). Therefore,

Table 8.1 Number of correct responses needed for significance in a triangle test.

	α						α					
n	0.20	0.10	0.05	0.01	0.001	*n*	0.20	0.10	0.05	0.01	0.001
6	4	5	5	6	...	32	14	15	16	18	20
7	4	5	5	6	7	33	14	15	17	18	21
8	5	5	6	7	8	34	15	16	17	19	21
9	5	6	6	7	8	35	15	16	17	19	22
10	6	6	7	8	9	36	15	17	18	20	22
11	6	7	7	8	10	37	16	17	18	20	22
12	6	7	8	9	10	38	16	17	19	21	23
13	7	8	8	9	11	39	16	18	19	21	23
14	7	8	9	10	11	40	17	18	19	21	24
15	8	8	9	10	12	41	17	19	20	22	24
16	8	9	9	11	12	42	18	19	20	22	25
17	8	9	10	11	13	43	18	19	20	23	25
18	9	10	10	12	13	44	18	20	21	23	26
19	9	10	11	12	14	45	19	20	21	24	26
20	9	10	11	13	14	46	19	20	22	24	27
21	10	11	12	13	15	47	19	21	22	24	27
22	10	11	12	14	15	48	20	21	22	25	27
23	11	12	12	14	16	54	22	23	25	27	30
24	11	12	13	15	16	60	24	26	27	30	33
25	11	12	13	15	17	66	26	28	29	32	35
26	12	13	14	15	17	72	28	30	32	34	38
27	12	13	14	16	18	78	30	32	34	37	40
28	12	14	15	16	18	84	33	35	36	39	43
29	13	14	15	17	19	90	35	37	38	42	45
30	13	14	15	17	19	96	37	39	41	44	48
31	14	15	16	18	20	102	39	41	43	46	50

Note 1: Entries are the minimum number of correct responses required for significance at the stated α-level (i.e., column) for the corresponding number of assessors, *n* (i.e., row). Reject the assumption of 'no difference' if the number of correct responses is greater than or equal to the tabled value.

Note 2: For values of *n* not in the table, compute the missing entry as follows: minimum number of responses (x) = nearest whole number greater than $x = (n/3) + z\sqrt{2n/9}$, where z varies with the significance level as follows: 0.84 for $\alpha = 0.20$; 1.28 for $\alpha = 0.10$; 1.64 for $\alpha = 0.05$; 2.33 for $\alpha = 0.01$; 3.10 for $\alpha = 0.001$.

a smaller number of panelists can be used (10 to 20) but 20 or a multiple of 20 is better, since the number of possible permutations of order of AABBB is 10, or of BBAAA is 10, for a total of 20. Meilgaard *et al.* [2] recommend using this test with trained panelists, since the number of samples presented requires memorization and possibly fatigue.

Both tests are used for testing product differences or similarities resulting from ingredient change, storage, or processing technique. In addition, triangle or two out of five tests are recommended for panelist selection prior to a descriptive panel, as well as panelist monitoring. A triangle test with products similar to those that will be used for a descriptive panel will determine the panelists' ability to discriminate between given differences [2]. The two out of five test is recommended when sensory fatigue effects are small.

8.4.1.2 Duo–Trio Test

In the Duo–Trio test, panelists are presented with a reference sample, and two coded samples, one of which matches the reference sample. Panelists are asked to select which coded sample matches the reference sample. The reference sample can always be sample A in the case of the 'constant reference mode', or it can be alternatively A and B in the case of the 'balanced reference mode'. In the constant reference mode, samples presented to panelists will be either A(ref)AB or A(ref)BA; in the balance reference mode, samples presented to the panelists will be A(ref)AB, B(ref)BA, A(ref)BA, or B(ref)AB. The correct answers are counted and compared to a standard table for one-tailed (or one-sided) tests (Table 8.2). The minimum number of panelists should be 20, but the discrimination power is improved when more panelists can participate. As in the triangle test, this test is exclusively used to determine if there is a general difference between two products. Affective questions should not be asked.

This test is statistically less robust than the triangle test because the chance of guessing the correct answer is 50%. However, because a reference is presented, the task is easier to perform: it is easily understood and the reference provides an anchor when looking for a difference. Like the triangle test, three samples must be tasted, and there also can be a problem if there is a strong carry-over effect from sample to sample.

8.4.1.3 Simple Difference Test

This test is to be used when the products have too much carry-over and may confuse the subjects in triple or multiple comparisons as described above. Panelists are presented with two samples, A and B, or either A and A or B and B to account for the "placebo effect". Panelists are asked whether the samples are the same or different. Results are analyzed by comparing the number of 'different' responses given when two samples of the same product were presented, to the number of 'different' responses given to the A/B pair, using the χ^2-test.

Table 8.2 Number of correct responses needed for significance in a Duo–Trio, or one-sided directional difference test.

	α						α					
n	0.20	0.10	0.05	0.01	0.001	*n*	0.20	0.10	0.05	0.01	0.001
5	4	5	5	...	...	33	20	21	22	24	26
6	5	6	6	...	...	34	20	22	23	25	27
7	6	6	7	7	...	35	21	22	23	25	27
8	6	7	7	8	...	36	22	23	24	26	28
9	7	7	8	9	...	40	24	25	26	28	31
10	7	8	9	10	10	44	26	27	28	31	33
11	8	9	9	10	11	48	28	29	31	33	36
12	8	9	10	11	12	52	30	32	33	35	38
13	9	10	10	12	13	56	32	34	35	38	40
14	10	10	11	12	13	60	34	36	37	40	43
15	10	11	12	13	14	64	36	38	40	42	45
16	11	12	12	14	15	68	38	40	42	45	48
17	11	12	13	14	16	72	41	42	44	47	50
18	12	13	13	15	16	76	43	45	46	49	52
19	12	13	14	15	17	80	45	47	48	51	55
20	13	14	15	16	18	84	47	49	51	54	57
21	13	14	15	17	18	88	49	51	53	56	59
22	14	15	16	17	19	92	51	53	55	58	62
23	15	16	16	18	20	96	53	55	57	60	64
24	15	16	17	19	20	100	55	57	59	63	66
25	16	17	18	19	21	104	57	60	61	65	69
26	16	17	18	20	22	108	59	62	64	67	71
27	17	18	19	20	22	112	61	64	66	69	73
28	17	18	19	21	23	116	64	66	68	71	76
29	18	19	20	22	24	122	67	69	71	75	79
30	18	20	20	22	24	128	70	72	74	78	82
31	19	20	21	23	25	134	73	75	78	81	86
32	19	21	22	24	26	140	76	79	81	85	89

Note 1: Entries are the minimum number of correct responses required for significance at the stated α-level (i.e., column) for the corresponding number of assessors, *n* (i.e., row). Reject the assumption of 'no difference' if the number of correct responses is greater than or equal to the tabled value.

Note 2: For values of *n* not in the table, compute the missing entry as follows: minimum number of responses (x) = nearest whole number greater than $x = (n/2) + z\sqrt{n}/4$, where z varies with the significance level as follow: 0.84 for $\alpha = 0.20$; 1.28 for $\alpha = 0.10$; 1.64 for $\alpha = 0.05$; 2.33 for $\alpha = 0.01$; 3.10 for $\alpha = 0.001$. This calculation is an approximation. See Meilgaard *et al*. [2] for a more complete calculation.

As many as 200 panelists can be used for this test, and they should be untrained. A variation of the test is to present more than one pair to one panelist. The data should then be analyzed using the McNemar test [2]. In their second and third editions, Meilgaard *et al*. [2, 17] give practical examples for each test, pointing out the differences in each situation, and going through the complete scenario, with ballot sheets, worksheets, and data analysis.

8.4.2 Single Attribute Difference Tests

8.4.2.1 Difference from Control

In addition, to test whether a difference exists between a control and one or more products, this test is used to measure the size of that difference. The objective may be to measure an overall difference (i.e., how different is the overall flavor?), or it may be specific (i.e., how different is the product for off-flavor?)

Panelists are presented with a control, and two or more test samples, and are asked to rate the difference between the control and the sample(s) on a given scale. A coded control is included in the presentation to measure the placebo effect. The mean difference from control for each test product will be compared to the difference from control values obtained with the blind control, and analyzed by analysis of the variance if more than two samples are tested, or the paired *t*-test if only one sample is compared with the control.

This test assumes that panelists are trained, or are otherwise familiar with the scale. It is easier if the scale is at least anchored at both ends, or otherwise has verbal cues at each point. Anchor words can be 'no difference' to 'very large difference', or 'extremely different'. Common scales are the 10 point (0 to 9), 16 point (0 to 15), or magnitude scale (0 to 100). Panelists should be aware of the format of the test, and know that at least one blind control is included in the test comparison. Twenty to 50 presentations (20 panelists can take the test twice) should be performed to determine a reliable degree of difference. If some difference is perceived between two products, the next step may be a consumer panel to determine which product is preferred.

8.4.2.2 Paired Comparison Test

This test is a directed test, that is, the experimenter wants to know if a sample is sweeter, more bitter, or has a different flavor than the control

intense for the defined attribute. Usually, a 'no difference' response is not allowed. However, under some test situations, an experimenter has to allow the 'no difference' option. Whichever way the data are handled, the sensory analyst must be aware of the consequences on the outcome (i.e. increase or decrease the power of the test) [2]. The experimenter also has to decide whether the test is one-sided (i.e., sample A is sweeter than B), or two-sided (sample A is different from sample B for sweetness) when results are tabulated and compared to the critical number of correct responses. For a complete discussion on directionality of hypothesis, refer to Meilgaard *et al.* [2].

8.4.2.3 Ranking Tests

Ranking tests are used to determine the difference between samples for one attribute, but the difference is not quantified. In the simple ranking test, panelists are asked to rank the samples for the attribute of interest (e.g., flavor, sweetness, off-flavor, preference). If more than one attribute is to be tested, it is better to present as many sets of samples as there are attributes to be tested, each set with its own three-digit codes. Otherwise, if two attributes are to be tested on the same set of samples, the response to one attribute is likely to influence the response to the next question. For instance, if panelists are asked to rank five samples of beer based on the intensity of bitterness, and the next question is to rank the samples by decreasing order of preference, the answer to the second question will be biased by the answer to the first question. For that reason, it is also recommended to always ask a preference question before asking a question on the level of intensity of a specific attribute.

Sixteen [2] or 30 [6] panelists are recommended to give reliable results. Panelists do not require training but instructions should be clear and, if necessary, a short session should be organized to discuss attributes. Ranks given by panelists are added for each sample, and the significance of the test is calculated using the Friedman–type statistic test for ranking data, with the nonparametric analog to Fisher's LSD for ranked sums [2].

When two samples are clearly different from one another, panelists will easily assign them rank 1 and rank 2. If, however, the difference is not so obvious, panelists should be asked to make their best guess. The simple ranking test is well adapted for three to six samples. It was shown to be comparable to a conventional descriptive test for a corn product [18].

A variation of the simple ranking test is the pair-wise comparison ranking test. In this test, samples are presented as pairs, one pair at a

time, and panelists are asked 'which sample is sweeter?' (more bitter, preferred, etc.). By presenting one pair at a time, panelists experience less confusion and less fatigue if samples have a lingering effect. They are well adapted to compare three to six samples. For example, with five samples, the pairs AB, AC, AD, BC, BD, and CD are presented in a completely randomized and balanced order. The Friedman statistical test is also used with the pair-wise comparison test.

8.4.3 Descriptive Tests

Conventional profiling is generally used by companies to understand the specifics of a product: appearance, aroma, flavor, aftertaste, texture, mouth-feel, and so on. In descriptive tests, panelists are trained to identify specific attributes describing a sample, and then to rate these attributes using a determined scale. A descriptive profile for a product can help with the understanding of the drivers of consumer preference for that product. Ultimately, results of descriptive analysis can be used in product development to correlate the intensity of desirable or undesirable attributes to instrumental measurements. Different methods for descriptive analysis have been developed in the last 50 to 60 years and the choice of one method over another depends on panel leader preferences and resources. Descriptive methods are all developed in two phases: the first phase consists of finding a set of attributes to describe the product, and the second phase is where panelists rate the intensity of the attributes found in the first phase [6].

In the *flavor profile* method, developed by Arthur D. Little Inc. in the late 1940s, five to six strictly selected and highly trained panelists develop the terminology to describe a product and rate 'character notes' on a seven-point intensity scale. Panelists develop their own reference standards and anchor points on the intensity scale. Panelists finally rate the samples and discuss the ratings to arrive at a *consensus profile*. The flavor profile is considered to be a qualitative descriptive analysis [1, 2], and because of the consensus, data cannot be analyzed statistically, and do not account for panelists' variability. This method has been criticized as prone to bias if the panel leader or one panelist has a strong personality and a tendency to impose their view on the other panelist members.

The *quantitative descriptive method* (QDA®) was developed by the Tragon Corporation in 1974 in collaboration with the Department of Food Science at the University of California, USA, [2]. This technique has been widely adopted by the sensory community, and the original

published technique has been adapted and modified to suit particular user requirements [6]. The QDA method relies heavily on statistical analysis to evaluate panelists' performances (usually 10 to 16 qualified panelists) and to calculate product differences. Similar to the flavor profile, panelists are trained with references, but the panel leader acts as a facilitator rather than an instructor. There should be a good communication between the panel leader and the project director for the generation of descriptive terms, which can be adjusted during the course of training. During training, panelists also discuss the reference standards, and the leader must make sure all panelists understand the terminology and use of standards. Once trained, panelists evaluate the samples one at a time in individual booths, and do not discuss their results after evaluation. The rating scale is usually a 15 cm line scale with anchor words at the end, but other types of scales can be used [6]. Panelists' responses are analyzed by analysis of variance (ANOVA), multivariate analysis of variance (MANOVA), and multiple comparisons for means separation [19]. The use of multivariate analysis, such as principle components analysis (PCA), is now a common practice to present descriptive profile results in a perceptual map, with each principal component (PC) representing a linear combination of attributes explaining most of the variation in each dimension. According to Meilgaard *et al*. [2], the QDA method is the closest to the ideal of treating human subjects as calibrated instruments and, as such, results of QDA can be correlated with instrumental data from product analyses using the partial least square (PLS) statistical method [20].

The *Spectrum*^{TM} descriptive analysis method was developed by G.V. Civile in the 1970s at Sensory Spectrum Inc. This method distinguishes itself from the preceding ones by using a predetermined universal standardized lexicon of descriptors and reference standards. The scale is a standardized 16-point intensity scale (0 = none, 15 = extreme), and it is considered to be an absolute scale. For example, intensities of saltiness and sweetness are considered equal when they are given the same score on this scale. Because all the procedures and instructions are standardized, it is theorized that results produced from different trained panels are comparable [1, 2].

Unlike the preceding methods, where panelists are trained to use the same terminology, the *free choice profiling* method allows panelists to use their own vocabulary and to develop their own questionnaire [21, 22]. The advantages of this method are that it does not require as much intensive training, and it accounts for panelists' differences

in perception and previous experience. Panelists should, nevertheless, be consistent in using and rating their own attributes and, as in the preceding methods, presentations are made in duplicate or triplicate. Also, the basic rules about product coding and balanced design still apply. The data are analyzed by general Procrustes analysis (GPA), a multivariate technique which adjusts for panelists' use of different parts of the scale by normalizing and centering the data. GPA also finds the best combination of variables (product descriptors) that explain differences between products. As in PCA, samples are mapped in the dimensions that explain most of the variation. However, because individual panelists may have diverse backgrounds, it may be difficult to extract a common understanding between attributes.

Other techniques requiring less stringent panelist training have been investigated, but the experimenter must understand their limitations. For instance, the ranking test described in the preceding section can be used for a simple profiling of a product with few attributes and minimal sampling by using predetermined attributes [18], or allowing panelists to develop their own vocabulary such as in the free choice profiling [23]. This is called 'flash profiling', and is an attractive alternative for companies when a rapid comparison is needed between products.

8.4.4 Affective Tests

Affective tests – acceptance or preference tests – can be performed in close collaboration with the marketing department, or can be done as 'premarket' studies, using a determined segment of the population that is preferably representative of the anticipated consumers of the product. Untrained panelists are preferably recruited with a questionnaire that will select for purchasing and eating habits, income level, age, and so on. The number of panelists will depend on the desired level of significance for the test and willingness to take an α-risk (Type I error) or β-risk (Type II error) (see Section 8.4.1). For example, Hough and collaborators suggested that 84 to 138 panelists were necessary to obtain results with $\alpha = 0.05$, and β ranging from 0.05 to 0.20, with an average root mean square error divided by the scale length (RMSL) of 0.23 [24]. They came up with these values by analyzing data from 108 consumer acceptance studies conducted in five countries.

Techniques used in affective tests include measurement of the 'liking' of a product (absolute rating), or comparison of preference between

products (relative rating). Absolute ratings are usually measured with the nine point hedonic scale, which is widely accepted. Other scales have been used, such as a seven point 'excellent' scale (excellent, very good, good, fair, poor, very poor, terrible), or facial hedonic scales [2]. Numbers, one to nine, are assigned to each point of the scale, and although the scale is categorical, it is assumed that the intervals are equal, and data are treated like continuous numbers, and analyzed with parametric statistics (ANOVA). Other scales are categorical such as the 'just right' scale (for example 'much too sweet', 'very sweet', 'just right', 'not quite sweet enough', 'not at all sweet enough'). In such scales, the percentage of answers in each category is summarized and the distribution response analyzed with the χ^2 test [2].

Ranking tests (simple or pairwise ranking) (see Section 8.4.2.3) are well adapted for acceptance studies. Unlike with the use of hedonic scales, the distance of liking between two products is not known: for example, if the average ratings of product A and B are 8.8 (like very much) and 5.0 (neither like nor dislike) on a nine point hedonic scale, a ranking test will only give the same products ranks of 1 and 2, respectively. Likewise, ranking tests indicate which product(s) consumers prefer, but do not indicate if they like it or not (for example, they might prefer A over B, but dislike both products).

8.5 PREPARATION AND PLANNING

Preparation and planning and then, more preparation and planning are the key to success.

8.5.1 Experimental Design

There is no universal design that fits every situation, so keeping anxiety levels low will help select appropriate tests for the different aspects under study, minimizing the risk of overextending the possibilities of a given design. Be aware of limitations from the number of samples and number of available panelists. A good experimental design is very important, and it is also very important to know how much information a specific design can provide and how much is needed. Do not ask too many questions, and keep questions phrased in a simple and clear manner.

Learn from other's experience:

The authors have had the opportunity of participating in some research projects from the very beginning and from different roles: the 'hard science', the sensory science, and the administrative. One of the most important lessons learned is that the quality of the final product has to be guarded at every step. Nothing is unimportant, and this goes beyond sensory techniques. The best trained panel, with the greatest number of members, and the best statistical software to analyze the data cannot make up for poor experimental design, mislabeled samples or too few samples due to budget constraints to obtain reliable results.

For example, when testing for the best quality ham for a certain market, which involves different pork breeds, production practices, diet, and so on, it would be worthwhile to go to the production facility, to get to know the people rearing the animals, talk to the investigator who is responsible there, see with your own eyes how easy or difficult it is for them to dose the right amount of feed and how accurate that really is. It is advantageous to learn how they keep track of their animals and the chance of one of them losing a tag; whether the assignment of animals to the different treatments was truly random as you had designed for the analysis with your team? How is animal stress handled? Is it an issue? How do they handle movement within the farm and transport to the slaughterhouse? Are measurements taken with care and in their full extent? How much does the controlled experimental environment make your experiment different from the normal procedures, and should you write recommendations for future practices? All these experiences could also be a lot of fun, and they give the experimenter the chance to leave the laboratory and experience some fresh air, but most importantly, they will give the researcher a better understanding of what can and cannot be done, and the effort (cost) involved, so that the experimental design can be improved, or optimized. The counter experience is also very productive: inviting production people over to the laboratory and showing them how the samples are processed, so that they will understand why all the requests in the experimental design are important and how not following procedures will significantly affect the final result.

8.5.2 Environment

One of the first things to take into consideration is the facility where the analysis will take place. In the case of an 'in house' test (usually a trained or expert panel), a place that provides a quiet environment which allows panelists to concentrate on their job is required. It should be well illuminated, protected from foreign odors and noises, ideally, isolated from the building 'traffic', but not too remote, have entrance and exit doors, a preparation area, and a serving (with isolated booths) and discussion area. Ideally, serving will be through some sort of opening in the wall that separates the preparation area from the tasting booths, and it should have filtered air and positive pressure. For a more complete description refer to [2].

Another important consideration is that all attending panelists should not wear any perfumed substances (lipstick, after-shave, perfume, etc.) since these will interfere with the analysis. The behavior during the test should be quiet, avoiding any unnecessary conversation, exclamation, noises, or gestures that might interfere with other assessors' performance.

8.5.3 Sample Preparation

Food safety for panelists is paramount, so every precaution should be taken to assure that the samples presented to the panel follow the most rigorous standards of food safety and cleanliness, including cooking to safe temperatures and appropriate storage, to avoid any source of contamination that could compromise the panelists' health. This is a point where no compromises can be made. Consider the need for profuse cleaning of all surfaces and utensils, use of hair and beard nets, gloves, and correct storage and maintenance temperatures for all food items and ingredients. In shelf life studies, the sensory evaluation should always come after the food safety member of the group (microbiologist) has analyzed the samples and given the "green light", ensuring that the product is safe for panelists' consumption. If there is no microbiologist on the team, samples should be sent out for microbiological analysis before using them for panel testing.

Once the location and panel have been selected, the logistics of sample preparation and presentation begins. Samples should be appropriately prepared; they should be presented in a neat, organized way. Plastic disposable cups (with or without lids), or paper plates are usually very convenient. Each sample should bear a legible, randomized three-digit number that does not give any clue to the taster about any particular

sample or order. If the samples consist of a product that is usually consumed refrigerated (e.g., yogurt, orange juice) or hot (e.g., hot cocoa, meat), those samples should reflect that and be presented at the intended consumption temperature. Furthermore, most care should be taken to have all the samples at the same temperature. This in itself could prove to be a challenge for larger studies.

Everything else should be standardized and, if this is not possible, the design should take variations into account in the randomization pattern including the time of day, the day of the week, temperature, presentation, environment, labels, colors, sizes, and illumination. Intimate mechanisms of flavor perception and liking are not fully understood so the only way to reduce variation is to be extremely cautious and inflexible with standardization. This will prevent contaminants and off-flavors from packaging, cork stoppers/lids, containers, handler's lotions, contaminated cups, and so on, from invalidating results.

An unexpected source of variability could be process practices, such as purification methods designed to strip undesired flavoring substances from the base matrix. These processes could also contaminate the product with trace amounts of chemicals that, unless specifically tested for, could go undetected and influence the overall final flavor profile. Some of these contamination sources could be filtration pads, ion exchangers, and carbon filters that are designed to strip undesirable amounts of certain compounds, but could also alter sensory quality by contributing ions or removing desirable components. A similar situation could appear when the industrial treatment, designed to modify some of the properties of the raw material, uses chemicals that will remain in the final product (e.g., during some meat tenderization techniques, calcium salts are injected to optimize the extent and duration of the process, but beyond a certain concentration, a 'metallic' off-flavor could appear in the final product).

8.6 PANEL SELECTION

Sensory evaluation is the ultimate test for the panel leader's commitment to quality: every detail should receive attention since panels have enough intrinsic variability without introducing more due to careless procedures. Part of the variability in panels is arguably a lack of knowledge of the underlying mechanisms of perception and its communication. Therefore, standardizing every procedure that can be controlled is important to minimize noise and other unwanted variability. As with any analytical

test, there could be many sources of interference, and it is the challenge for the researcher to avoid and minimize them. The 'instrument' (i.e., the panelists), the main tool the researcher uses to obtain raw data, talks back, has mood changes, varies in response over time, or simply does not show up for the experiment. Extra measurements must be taken to assure the reliability of the data. As in any other procedure, the final results and conclusions can only be as good as the data. For all these reasons, listed here are some of the important things that should be addressed before, during, and after any sensory test.

8.6.1 Trained Panels

Panels can be tuned and calibrated to become a trained panel. Panelists should be selected from a group of available people based on their ability to taste, smell, and discriminate different intensities for the four(five) basic tastes in a set of standardized tests and/or selected aromas/odors that will be encountered in the product. People that do not perform satisfactorily, whose schedule will not allow participation at each session, or whose personality is disruptive or overpowering should be screened out and diplomatically dismissed [2]. After selecting a total of 12–18 members, so that at any given time there will be eight to12 participants available, the panel should receive training on basic tastes and the different aspects of the product(s) to be tested in order to make sure that all the assessors understand what it is expected of them and that they are able to communicate their results in a homogeneous and clear way. This includes discussions about how to express results, the development of (or the adoption of) a lexicon (a lexicon is a collection of words used to describe flavor traits (see [5]).

Trained panelists are a renewable but limited and very expensive resource. They are difficult to find and train and very easy to lose, so the panel leader should take good care of them. Panelist selection and training takes months, and if the pool of people from which to draw panelists is fixed or restrained by any special consideration (e.g., when testing alcoholic beverages, subjects such as nondrinkers, pregnant women, and personnel driving and/or operating machinery are to be excluded, and this can reduce significantly the number of potential panelists available) there might not be enough members to perform a meaningful analysis.

This brings up a difference that should be noted: trained panels are obviously composed of trained panelists. Consumer panels, on the other hand, are composed of consumers with no sensory training. Using the

same people for consumer evaluation over and over again should be avoided because it defeats the basic principle in that the consumers become somewhat trained or, at least, they can no longer be considered naïve, and this could bias their response and invalidate the results. For example, in-house panelists should not be used for consumer panels because they might know the product too well or have a preconceived preference, and therefore could bias the results.

8.6.2 Consumer Panels

Another important consideration is the human aspect of panelists, so make sure to address any questions that might arise from objections to the sample to be tasted, its ingredients or form of preparation. For example, some people philosophically object to genetically modified organisms, or have cultural or religious backgrounds that require them to avoid certain foods, specific production techniques, or ingredients (e.g., kosher or halal ingredients or preparation, or certain cultures and religions that avoid meat of animals considered 'impure').

Learn from other's experience

One of the authors of this chapter helped run a consumer panel evaluating qualities of tomatoes at The University of Georgia in Athens. Unfortunately, for different reasons that were beyond the control of the main investigator, the experiment date fell during Ramadan, and many of the students available on campus were Muslims who were fasting during the day, and it was really difficult to find enough panelists. On the other hand, some people from the available group insisted in adding salt to the tomatoes ('I don't eat ***my*** *tomatoes without salt'). In such cases, the researchers were forced to go along, sacrifice a sample with their best smile and discretely mark the ballot sheet for later destruction (one of the attributes being tested was tomatoes' saltiness…).*

Allergies and intolerances are two other important aspects which should be taken into account. If the product being tested or any of its components can produce allergens, it is important to take the corresponding measures to guarantee that a sample that could generate allergic reactions or intolerance is never served to a sensitive person.

Since pregnancy could alter flavor perception, pregnant women should not be called to participate in the panel, unless the product is specifically targeted for that market.

In certain institutions, usually universities, there are regulations and procedures concerning the use of human subjects. It is usually a requirement that the test is preapproved (demonstrate the need and the innocuousness of the test), and that all the participants sign consent forms. In any case the researchers must be aware of any legal aspects that might apply to their workplace.

Recruitment of future panel participants can be done at random for the most part, but when relevant for specific applications, it is advisable to take into account the potential influence of ethnicity, cultural background, religious/philosophical beliefs, familiarity with the product, and so on, since these factors could mean the difference between failure or success. The senses, and in particular the sense of smell, is a powerful trigger of memories which are connected to personal and cultural backgrounds and personal history. In the case of a consumer panel, the participants could also be chosen for their relevance to the product: if the goal is to test acceptability of a new yogurt, a reasonable place to start would be a grocery store, near the dairy aisle, inviting people that are dairy consumers to participate. This segmentation of the population would help eliminate 'noise' in the data that might mask any useful results (most of the time it is not a good idea to ask a consumer about liking/intention of purchase if they do not like that specific type of product and/or they are not willing to consume it). This practice should be aligned with the goal of the project, and if the goal is to test only yogurt consumers, the results would be only valid for the subpopulation of yogurt consumers, and cannot be extrapolated to a wider group.

8.7 CONDUCTING A PANEL

Always treat panelists respectfully, make them feel that their opinion is very important, and assure them that there are no right and wrong answers but that the results are just a matter of perception. Panelists who are rewarded with positive feedback will agree to participate in more tests. After the test is completed, if paper ballots are used, review each ballot page as it is submitted to make sure all the boxes are checked, comments are written, and demographics are answered. In cases where computer programs are used instead of paper ballots, the software should prevent submission of incomplete ballots.

By the day of the actual test, every aspect should have been addressed. No improvisation or last minute changes should be necessary if the planning was complete and correct. In cases of complex or large panels, it is a good idea to recruit helpers and schedule a rehearsal to make sure that everything runs smoothly. This will provide those helpers with the opportunity to understand their roles better, and to comprehend the panel mechanism and panelist treatment. Have enough consumables to cover any unforeseen accidents (spills, leaks, etc.). Accidents do occur, so having extra unlabeled cups and labels ready will save time and trouble.

Make sure that all the required elements are available and organized in a way that makes sense to everybody helping to run the panel. Allow extra time for sample preparation in cases where samples need to be prepared prior to panel (for example, cooked meat, etc.). In many cases, it is best to present samples in small disposable plastic cups (2–4 oz, 60–120 ml) covered with lids.

8.8 EXPRESSION OF RESULTS

The researcher should plan in advance what he/she intends to do and how he/she will handle the results. For example, what to do in the case of conflicting results? What would be the next step? Redesign? Include market segmentation? Increase the number of panelists? Is this feasible (from a practical and financial standpoint)?

This goes back to the planning stage: was there a line for comments in the ballots? Did the questions require open or closed answers? Most open answers come back unanswered, and closed comments are likely to suffer from bias from the investigator, or panelist misunderstanding, but are also more likely to be answered. Again, the experimenter should be aware that the way questions are phrased could affect the results.

For closed answer ballots (where a selection of possible answers is provided, like a multiple choice) it is strongly recommended that each possible answer is given much thought and that an experienced investigator reviews the questionnaire according to a pre-established lexicon. Sometimes, running a test panel and recording the spontaneous answers is warranted, and these answers can be used as a base for formulating the closed answers offered on the official panel test. A focus group is another option.

Depending on the sensory technique used, the statistics behind it will differ, and each situation will require a different minimum number of samples tested and/or participating assessors to yield a statistically

significant result, as discussed in previous sections. In sensory evaluation, there is another aspect that does not apply to most 'traditional' analytical techniques: apart from being statistically significant, the results must be biologically significant.

To illustrate this, let us assume that the experimenter wants to compare meat tenderness between two different bovine cross-breeds. Instrumental data (Warner Bratzler shear force method) analysis of the meat found statistically significant differences at 0.5 lb (227 g). When running a sensory panel on the same material, panelists were able to detect differences 1 lb (454 g) or larger. Which values should be considered as a relevant result? Does the instrumental value have the most relevance? In other words, is a difference of 0.5 lb (227 g) important when no one is likely to detect it? Maybe the 1 lb (454 g) level is more relevant to determine whether consumers are able to tell the two cross-breeds apart based on meat tenderness. The same would apply to results from fruit and fruit products, where a change in 1°Brix, a common unit used to measure soluble solids, mostly sugars in fruit, is not usually enough to correlate with a change in perceived sweetness intensity.

As with any other measurement, sensory evaluation results will always be a statistical value: nine out of 10 panelists thought that the new product was great and might buy it, 34% rated tenderness of the calcium-treated steak at 9; 27% as 7; 21% as 6; 10% as 4 and 8% as 1. What does one do with these results? Now is the time when the researcher uses the secret weapon: common sense; but remember, before making any decisions, drawing conclusions, even before taking any measurements, the researcher has to know the desired precision and exactitude of the test, and most importantly, the questions to be answered.

8.9 CONCLUSIONS

Taste panels can present logistic headaches, are time consuming, can strain personal relationships, and the results obtained are restricted and not always straightforward. So why bother? The answer is because sensory analyses give you the 'magical' answer, something no instrument invented so far can do, and that is give an idea (probability) of how well accepted a new product will be (intent of purchase), or how will consumers like it in comparison with the competitors' (consumer panels), or how the intensity of the flavor, texture, overall sensation of the product compares with the competition or your desired 'gold standard'.

There is a saying that 'the nose knows'. In this book, Chapter 3 shows that a human subject, smelling compounds as they exit the column

in a gas chromatograph, can detect compounds that the instrument cannot detect. Therefore, results from a trained panel can give a complementary picture to instrumental data that helps understand a product in terms of human perception. Consumer panels on the other hand, provide affective responses, an emotional picture which no instrument can provide.

REFERENCES

1. Lawless, H. T. and Heymann, H. (1999) *Sensory valuation of Food. Principles and Practices*, Aspen Publishers Inc., Gaithersburg, USA, 827 pp.
2. Meilgaard, M., Civille, G.V., and Carr, B.T. (2007) *Sensory Evaluation Techniques*, 4th edn, CRC Press, Boca Raton, USA.
3. Resurreccion, A.V.A. (1998) *Consumer Sensory Testing for Product Development*, Aspen Publishers Inc., Gaithersburg, USA, 254pp.
4. Stone, H. and Sidel, J. (2004) *Sensory Evaluation Practices*, 3rd edn, Academic Press, New York, USA.
5. Drake, M.A. and Civile, G.V. (2003) Flavor lexicons, in *Comprehensive Reviews in Food Science and Food Safety*, Vol. 2, pp. 33–40.
6. Carpenter, R.P., Lyon, D. H., and Hasdell, T. A. (2000) *Guidelines for Sensory Analysis in Food Product Development and Quality Control*, 2nd edn, Aspen Publishers Inc., Gaithersburg, USA, 210 pp.
7. Demattè, M.L., Sanabria, D., and Spence, C. (2006) Cross-modal associations between odors and colors. *Chemical Senses*, **31**, 531–538.
8. Plotto, A. *et al.* (2004) Odour and flavour thresholds for key aroma components in an orange juice matrix: terpenes and aldehydes. *Flavour and Fragrance J.*, 19, 491–498.
9. Laing, D.G., Legha, P.K., Jinks, A.L., and Hutchinson, I. (2003) Relationship between molecular structure, concentration and odor qualities of oxygenated aliphatic molecules. *Chemical Senses*. **28**, 57–69.
10. Amoore, J.E. (1969) A plan to identify most of the primary odors, in *Olfaction and Taste III* (ed. C. Pfaff-Mann); Rockefeller University Press, New York, USA, pp. 158–171.
11. Amoore, J.E. (1971) Olfactory genetics and anosmias, in *Handbook of Sensory Physiology, Vol 4, Pt.1*, pp 245–256.
12. Amoore, J.E. (1991) Specific anosmias, in *Smell and Taste in Health and Disease* (eds T.V. Getchell *et al.*) Raven Press, New York, USA, pp. 655–664.
13. Plotto, A., Barnes, K.W., and Goodner, K.L. (2006) Specific anosmia observed for β-ionone, but not for α-ionone: significance for flavor research. *J. Food Sci.*, **71**, S1–S6.
14. Schilling, B. (2006) Odorant metabolism in the human nose, in AChems, 28th Annual Meeting, April 26–30, 2006, Sarasota, Florida, USA. Book of abstracts.
15. Buettner, A. (2006) Flavor metabolism in the oral cavity, in AChems, 28th Annual Meeting, April 26–30, 2006, Sarasota, Florida, USA. Book of abstracts.
16. Margaria. C.A. (2001) Flavor quality models for consumer acceptability using partial least squares regression.University of Georgia, Athens, USA, Master's Thesis.

17. Meilgaard, M., Civille, G.V., and Carr, B.T. (1991) *Sensory Evaluation Techniques*, 3rd edn, CRC Press, Boca Raton, USA.
18. Rodrigue, N., Guillet, M., Fortin, J., and Martin, J.-F. (2000) Comparing information obtained from ranking and descriptive tests of four sweet corn products. *Food Qual. Pref.*, **11**, 47–54.
19. Lea, P., Næs, T., and Rødbotten, M. (1998) *Analysis of Variance for Sensory Data*, John Wiley & Sons, Inc., New York, USA, 102 pp.
20. Martens, M. and Martens, H. (1986) Partial least square regression, in *Statistical Procedures in Food Research* (ed. J.R. Piggott), Elsevier Applied Science Publishers Ltd., UK, pp. 293–359.
21. Arnold, G.M. and Williams, A.A. (1986) The use of generalized Procrustes technique in sensory analysis, in *Statistical Procedures in Food Research*, (ed. J.R. Piggot), Elsevier, London, UK, pp. 233–253.
22. Williams, A.A. and Arnold, G.M. (1984) A new approach to the sensory analysis of foods and beverages, in *Proceedings of the 4th Weurman Flavor Research Symposium* (ed. J. Adda), Elsevier Science Publishers, New York, USA, pp. 35–50.
23. Delarue, J. and Sieffermann, J.-M. (2004) Sensory mapping using Flash profile. Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. *Food Qual. Pref.*, **15**, 383–392.
24. Hough, G. *et al.* (2006) Number of consumers necessary for sensory acceptability tests. *Food Qual. Pref.*, **17**, 522–526.

9

Regulatory Issues and Flavors Analysis

Robert A. Kryger
Molecular Separation Specialists, Lakeland, USA

9.1 INTRODUCTION

Flavor mixtures, as routinely manufactured today for addition to food or beverage products, usually contain dozens of ingredients and thousands of individual chemicals. Despite the complexity, technology for flavor manufacturing is relatively simple – only a formula for the mixture, access to the raw materials, and blending equipment are required. Traditionally, the flavor manufacturer has tried to protect his craft by keeping his product formulations secret.

The flavor manufacturers' economic interest in their products is balanced by the desire of food manufacturers – and consumers – for the assurance of safe, wholesome foods. How do they ensure the safety of the flavors they purchase and consume? If a food manufacturer or consumer wants a flavor produced with only natural ingredients, instead of cheaper artificial ingredients, how does the flavor purchaser ensure compliance? How do flavor manufacturers compete fairly with each other, if some may be tempted to 'cheat' and use much less expensive

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.

'unallowed' ingredients? For these reasons, among others, governments and other regulatory agencies have established regulatory guidelines for flavors.

Within the context of the flavor producers' desire to 'protect their formula' and the purchaser/consumer/regulatory agency's desire to 'open up the formula', the technology for chemical and sensory analysis of flavors has undergone revolutionary development. In the early twentieth century, flavor analysis was limited to simple chemical tests and human sensory analysis of the mixtures, yielding only hints of the total formulation. Today, with modern chromatographic techniques routinely available, nearly the entire chemical content can be elucidated with substantial precision. This allows for relatively easy reverse engineering of flavor mixtures.

Despite these developments, the information that can be gleaned analytically is helpful, but not complete. There are still many limitations that complicate the regulation of flavor products. For example, the use of authentic vanilla bean extract in a finished beverage may be mandated, but what exactly constitutes the analytical test for 'authentic vanilla extract' when this natural extract is manufactured from a plant product which is grown in different parts of the world, fermented according to various traditional and modern techniques, and extracted according to various techniques?

This chapter focuses on the regulatory framework that governs flavors which has developed as a result of these tensions between the flavor suppliers, food manufacturers, and consumers in the context of substantial developments in the technology of flavor analysis. The next section gives an overview of the flavor regulation framework, while the following section focuses on specific regulatory issues where analytical techniques and limitations play a fundamental role.

9.2 REGULATORY OVERVIEW

9.2.1 History

Historically, food and beverage manufacturers combined spices, essential oils, and other ingredients with other bulk ingredients to manufacture their food products. The 'flavor' never existed as a separate entity, but was incorporated into the overall product formula. The food manufacturer controlled the sourcing of all of the ingredients. The 'flavor ingredients' were largely natural spices and extracts – for example, salt, vinegar, citrus oils, cinnamon, vanilla, and so on. As consumer tastes

expanded and world trade routes developed, the number of different flavor ingredients increased. Often, however, these ingredients were expensive, grown in remote locations, and subject to frequent supply disruptions. Regulatory interest was in ensuring a safe food product. Since the food manufacturer had complete knowledge of the raw materials, consumers and potential regulators could easily identify the responsible party in the event of a 'problem'.

With the advent of modern organic chemistry in the nineteenth and early twentieth centuries, some key chemicals responsible for the flavor of natural products were identified. For example, vanillin as the main contributor to vanilla flavor was first isolated from vanilla pods by Gobley in 1858, and citral as the key lemon flavor was isolated by Semuler and Tiemann in 1896. This knowledge allowed for flavor to be quantified in some sense, at least by a sophisticated party. Chemists also started discovering methods to synthesize important flavor chemicals. One example is vanillin by Tiemann in 1874. Synthetic aroma chemicals revolutionized the manufacturing of flavors, reducing the need for natural products which may be expensive, available in limited quantity, and in uncertain quality. Furthermore, the effort to source flavor ingredients reliably was often inefficient at the food manufacturers' level.

With time, the expertise to manufacture aroma chemicals, source a large variety of natural extracts and essential oils, and consistently combine these ingredients into a large variety of different 'flavor profiles' resulted in the modern flavor industry. Now, the flavor manufacturer specializes in producing many different flavors for sale to other food and beverage manufacturers. The modern food manufacturer purchases a flavor package from a flavor company, designed to conform to the flavor, labelling, and regulatory requirements of the finished product. This flavor mixture is then combined with the other food ingredients, packaged, and distributed for purchase by the consumer.

Concomitant with the development of the modern flavor industry, government interest in the regulation of the production of food substantially increased. In the US, this was motivated by exposés such as *The Jungle* by Upton Sinclair (1906) detailing horror stories regarding the meat production industry. Eventually, this led to the creation of the Food and Drug Administration (FDA) and the Federal Food, Drug and Cosmetic Act of 1938; this has subsequently been amended. The regulatory model here was to treat all substances, including flavors, which were added to food products as food ingredients subject to regulatory constraints. However, the regulation respected the flavor industry desire to prevent disclosure of the exact flavor formulation. While the food

manufacturers might not be aware of all of the ingredients in a flavor package they purchased, the flavor manufacturers were responsible for conforming to the applicable regulations. This model of protecting the flavor formulation as a trade secret, but requiring the flavor mixture to conform to specific regulations spread worldwide as the 'standard' regulatory model for flavors.

In the intervening years, as additional safety, regulatory, and product disclosure concerns have developed, the specific flavor regulations worldwide have been modified and/or expanded to address these concerns. In broad strokes, the major regulatory issues can be broken down into three main categories: safety, consumer product labeling, and fair trade regulations. Each of these three is treated below.

9.2.2 Safety Regulations

Perhaps the primary concern of any consumer eating a food product is whether the product is safe to eat. Safety concerns can arise from toxic ingredients, with either acute or long-term toxicity, or inadvertent contaminants like heavy metals or pesticides. Regulations designed to ensure that a flavor mixture is safe to consume have generally developed around the concept of creating 'positive lists' of allowed safe ingredients. Examples include the FEMA/GRAS list of 'generally regarded as safe' flavor ingredients, or the EU list mandated by Regulation 2232/96/EC. There are several advantages of a positive list regulatory approach. First, the flavor manufacturer is not required to disclose the ingredients used, only ensure compliance with the positive list. Also, at least conceptually, noncompliance can be checked by analytical tests for components not on the list. Table 9.1 contains a list of sources of important positive ingredient lists for flavor products from around the world.

Positive lists also have several disadvantages. In particular, flavor innovation can be limited by the difficulty associated with adding new ingredients to approved lists. The modification of regulatory approved lists is usually a very time-consuming and expensive process. This

Table 9.1 Some key positive ingredient lists for flavors.

EU	– Regulation of the EU Parliament, Positive List of Flavouring Substances [1]
Japan	– Administration of Food Safety, List of Approved Food Additives [2]
USA	– FEMA/GRAS List [3]

problem is magnified if many different lists are in use worldwide. Protection for proprietary or 'trade-secret' ingredients can be difficult to maintain with these types of lists. Furthermore, the lists themselves can be ambiguous. For example, a natural essential oil which contains hundreds of constituent chemicals may be on the approved list. Concentrated fractions of this essential oil, which may contain high levels of particular constituents, are presumably also included. However, the constituent chemicals themselves may not all be included on the approved list. At what point does the concentrated fraction become a completely different item, warranting a separate entry within the positive list?

In some cases, specific 'negative lists' exist regarding prohibited or limited in use flavor ingredients. Usually, this applies to an ingredient which was used at one time, but was discovered to be unsafe for one reason or another. An example is the US Code of Federal Regulations (CFR) regulation concerning coumarin (21 CFR 189.130). Coumarin was used in artificial vanilla formulations prior to being banned in the 1940s. The US regulations in 21 CFR 189 list a number of prohibited ingredients for foods and flavors. Another example is Annex II of the EU Flavouring Directive 88/388/EEC which lists substances which cannot be added to flavors.

A second source of safety concerns regards avoiding inadvertent contaminants in the flavor product – either chemical or microbiological in nature. Here, food ingredient regulations apply to the flavor ingredients themselves. For example, limits on heavy metal contaminants in natural extracts which are used in flavor formulations apply. The flavor manufacturers have to ensure the safety/regulatory conformity of the raw materials they use in their products. Other contaminants of concern include pesticides and other agricultural residues, naturally occurring toxins, and inadvertent allergen contaminants. In Table 9.2, a number of these contaminants of interest in flavor manufacturing are listed. More generally, flavors must be manufactured under good manufacturing procedure (GMP) conditions appropriate to all food ingredients. These requirements address sanitation and cleanliness concerns regarding the production environment among other factors. Table 9.3 lists some of the important manufacturing regulations worldwide governing food flavors.

A third category of safety regulations derive from the fact that flavors are fundamentally chemical mixtures. Workplace safety requirements regarding chemical contact, MSDS, inhalation, and safe-handling requirements apply based upon the individual chemicals within the mixtures. Also, all chemical registration and transportation regulations for

Table 9.2 Some contaminants of regulatory concern in flavoring products.

Allergens:
Peanuts
Tree nuts
Milk
Egg
Soybean
Fish
Crustacea
Wheat
GMO proteins
Heavy metals (Pb, As, Cd, Hg)
Natural toxins:
Aflatoxin
Patulin
Pesticides
Solvent Residues

Table 9.3 Some key manufacturing regulations regarding food/flavoring products.

Canada	– Food and Drugs Regulations [4]
EU	– Regulation on the Hygiene of Foodstuffs [5]
UN/FAO/WHO	– Recommended International Code of Practice General Principles of Food Hygiene [6]
	– General Standard for Food Additives [7]
US	– Federal Food, Drug, and Cosmetic Act [8]
	– GMP Regulations for Food Manufacturing [9]

chemical mixtures apply. Since flavor mixtures may be flammable with relatively low flashpoints, hazardous chemical handling and transport regulations can be critical. Further discussion of these generic chemical regulations, while very important, will not be covered in this chapter.

9.2.3 Product Labelling Regulations

Secondary to safety concerns, consumers have an increasing desire to know the contents and health impact of the food they consume. Consequently, regulations have developed which require the disclosure of ingredient statements and nutritional content of manufactured food products. As discussed above, this desire has to be balanced against the manufactures' desire to protect their formulation as a trade secret.

For the most part, the flavor industry has managed to limit the disclosure of the flavor mixture ingredients to the consumer to a generic ingredient statement such as 'flavor added'. While disclosing no information to the consumer on the contents of the flavor mixture, the consumer is aware that a separate ingredient – a flavor mixture – was added to the product by the food manufacturer.

Early on, consumer interest in 'natural' products drove the requirement that 'natural flavors' should be distinguished from 'artificial flavors'. In this way, a food or beverage manufacturer could include the statement 'natural flavors added' on the package contents to differentiate their product. Consequently, regulations to define a 'natural' flavor were developed. While conceptually easy, the exact definition of a natural flavor can be tricky in practice. For example, a flavor made entirely from botanical extracts – such as an orange flavor derived from orange peel oil – might be obviously natural. If we further add a flavor chemical isolated in nearly pure form from a natural extract – say adding trans-2-hexenal isolated by distillation from mint oil – this may also be 'natural'. However, what if the trans-2-hexenal is generated by a microbial fermentation process, or perhaps by a genetically modified microbe? At what point does the trans-2-hexenal no longer qualify for the description'natural'? A similar example involves the chemical modification of natural raw materials – for example, catalytic esterification of natural alcohols and acids. Is the resulting ester still 'natural'? Because of the importance of 'natural' products to the consumer, regulations governing the definition of a 'natural flavor' have been very important.

The specific definitions of natural and artificial flavors can be complex especially with different definitions of natural in different regulatory regions. Most of the differences concern which specific processes are allowed during the manufacture of natural flavors. Of course, the specific language of any regulation often allows for 'gray areas' of interpretation. Also, some regulatory regions, such as the EU, allow for hybrid flavor types such as nature identical (NI). An NI flavor chemically mimics a natural flavor but may use artificial ingredients (see Chapter 1, Section 1.5). Table 9.4 lists some of the important flavor regulations defining flavor types.

Beyond natural and artificial flavor labeling, consumer demand is driving the development of new labeling categories such as 'GMO free', 'organic', or even 'country of origin'. An example of this latter regulation is the requirement in the US that juice products must be labeled with the country of origin of the juice/juice concentrates. For each new labeling category, specific regulations need to be developed which define the

Table 9.4 Some regulations regarding flavor types.

Australia/New Zealand	– User Guide to Flavourings and Flavor Enhances [10]
EU	– Flavour Directive [11]
Japan	– List of Plant or Animal Sources for Natural Flavourings [12]
UN/FAO/WHO	– General Requirements for Natural Flavourings [13]
USA	– Flavor Labeling Regulations [14]

concept as applied to flavor blends. Examples include the National Organic Program, administered by the US Department of Agriculture, which defines what flavor ingredients are allowed in organic foods and the EU Regulation (EC) 50/2000 governing the labeling of foods with GMO derived flavors.

A final interesting class of labeling regulations regards religious acceptability regulations – kosher or halal labeled products. While only a small fraction of the US population is Jewish, a large fraction of the manufactured food in the US is labeled 'kosher'. Consumers have traditionally valued the kosher label as a sign of quality. The regulations regarding 'kosher' status are entirely controlled by private organizations. Kosher approved ingredients and manufacturing techniques are required to manufacture flavors for use in kosher foods. A similar situation is true for halal flavors as defined by the Islamic community.

9.2.4 Fair Trade/Conformity with Established Standards

A third category of regulations affecting flavors, ultimately tied to product labeling, is what can be categorized as fair trade regulations. Broadly speaking, these apply to requirements on how a flavor is formulated for a product so that the flavor and resultant product competes fairly in the marketplace. Usually, these regulations are written to apply to the finished consumer product, and flavors are treated within the context of an 'ingredient' in the finished product. However, some regulations deal particularly with the manufacturing and sale of flavors themselves.

'Standard of identity' regulations apply to a number of food products worldwide. For example, there are regulations in the US Code of Federal Regulations and an EU Directive regarding the definition of orange juice. These definitions, differing somewhat from each other, establish what products can be marketed in their respective regions as 'orange juice'. Among other proscriptions, the regulations define which flavoring ingredients can be added to orange juice, or juice concentrate which is marketed as orange juice. Another example includes

Table 9.5 Some examples of standards of identity for food products containing flavorings.

Alcoholic Beverages:	
Canada	– Food and Drug Regulations [16]
USA	– Code of Federal Regulations [17]
Fruit Juices:	
EU	– Fruit Juice Directive [18]
USA	– Code of Federal Regulations [19]
Vanilla Extracts:	
Canada	– Food and Drug Regulations [20]
USA	– Code of Federal Regulations [21]

the US regulations regarding vanilla extracts. The CFR designates quite explicitly the ingredients and manufacturing process for any product sold as a 'vanilla extract' (see Table 9.5).

A slightly different type of regulation applies to flavors allowed for use in alcoholic beverages in the US. These flavors must have Alcohol and Tobacco Tax and Trade Bureau (TTF) approval and consequently must meet special regulations designed around concerns regarding alcohol-tax issues and fair trade in alcoholic beverages [15]. These regulations govern the use of ethanol as a solvent in flavor blends as well as limit the types of ingredients that can be used. There is even a separate definition of 'natural' applicable for flavors in alcoholic beverages (see Chapter 1, Section 1.5).

There are several 'standard of identity' regulations worldwide covering thousands of products. Some prohibit the addition of any flavoring agents while others establish limits on the use of flavors or ingredients present in the flavors. Because of the huge variety of requirements, flavor manufacturers have to treat these products on a case-by-case basis. Table 9.5 lists some of the important standard of identity regulations worldwide.

9.2.5 Flavor Types

In response to the variety of regulations governing flavors, some broad categories of flavors have developed to ease flavor commercial exchange and regulation:

- *Natural and artificial flavors (N and A)* comprises the broadest category of flavors. N and A flavors contain a mixture of natural and artificial ingredients. Of course, the flavor mixture must meet

specific regulations as required by its final use market. Thus, an N and A flavor suitable for use in USA may not be suitable in Japan.

- *Natural flavor* is a very broad category. A natural flavor must be composed entirely of naturally-occurring ingredients. Ingredients in these flavors can be isolated by physical means – for example, extraction or distillation – but only limited chemical modifications are allowed. In this latter category, fermentation or roasting processes are usually considered 'natural'. Several subcategories of natural flavors exist driven by labeling requirements on specific end use consumer products. A *natural type flavor* must be all natural, but may not contain any ingredients from the named source. Thus a natural apple type flavor might contain no ingredients derived from apples. A *natural flavor 'with other natural ingredients (WONF)'* is a category of natural flavors where the flavor must contain some ingredients derived from the named flavor type. Thus a natural orange flavor WONF must contain flavor ingredients from the orange fruit, but other natural ingredients are also allowed. A *from the named fruit (FTNF)* natural flavor contains ingredients derived entirely from the named substance.
- *Nature identical flavor (NI)* is a specific category of flavors created by regulation in some markets like Europe. An NI flavor is a midpoint between a 'natural' and 'artificial' flavor. In simplest terms, an NI flavor is made from natural and artificial ingredients, but the only artificial ingredients allowed are those chemically found in the corresponding natural flavor. For example, one important flavor contributor to apples is ethyl-2-methyl butyrate (E2MB), but this is an expensive ingredient from natural sources. In an NI flavor, synthetic E2MB could be used.
- *TTF natural flavor* is a category specific for use in alcoholic beverages within the US. TTF flavors must be approved by the Alcohol and Tobacco Tax and Trade Bureau, but their regulations allow for a flavoring containing up to 0.1 % artificial top notes to be used in alcoholic beverages as a 'natural flavor'. This allowance can substantially reduce the cost of manufacturing this type of flavor product.

Other examples arise in response to marketplace and regulatory agency demands. *Organic* and *GMOfree* flavors are two relatively new categories. These flavors must conform to regulations governing organic and GMO free foods. Other important categories of flavors such as kosher and halal have been discussed above.

9.2.6 Governing Authorities

In principle, each country has its own governing authorities that regulate food and consequently flavor products. However, in practice, the largest manufactured food consumer markets – the USA and EU – have developed much of the regulatory framework that has served as the model worldwide. Japan, which is an increasingly important flavor market, has also developed flavor regulations which are different in some key points from the EU and US. Finally, the UN has an ongoing program to develop worldwide standards for food production and trade which has lead to a number of flavor-related requirements. Table 9.6 shows some of the most important national and supranational government regulatory agencies. As is illustrated by the case of the USA, because flavors are simultaneously food ingredients, chemical mixtures, and often utilize agricultural ingredients, sometimes different regulatory agencies within a single government are involved in regulating related aspects of flavor products.

In order to influence these regulatory agencies most effectively, the flavor industry and major flavor consumer industries often utilize trade organizations. Table 9.7 shows a list of the major industry and trade groups that represent the flavor industry. In addition to the flavor manufacturers, flavor consuming industries obviously also have strong interests in flavor regulations. For instance, the beverage industry is a good example of an industry which purchases large quantities of flavoring products and whose products are strongly influenced by flavor

Table 9.6 Some major regulatory agencies for flavors.

Australia/New Zealand	– Food Standards Australia New Zealand (FSANZ)
Canada	– Health Canada
Japan	– Ministry of Health, Labour and Welfare
EU	– European Food Safety Authority (EFSA) and the Scientific Committee Panel on Food Additives, Flavourings, Processing Aids and Materials in Contact with Food (AFC Panel)
	– European Commission Scientific Committee on Food (SCF)
UN	– WHO/FAO - Codex Alimentarius Commission (Codex) and the Joint Expert Committee on Food Additives (JECFA)
USA	– Food and Drug Agency (FDA)
	– US Department of Agriculture (USDA)
	– Alcohol and Tobacco Tax and Trade Bureau (TTF)

Table 9.7 Some major industry/trade regulatory groups for flavors or foods which use flavors.

European Flavor and Fragrance Association (EFFA)
Flavor and Extract Manufacturer's Association (FEMA)
International Federation of Essential Oils and Aroma Trades (IFEAT)
International Organization of the Flavor Industry (IOFI)
Japan Flavor and Fragrance Materials Association (JFFMA)
American Beverage Association
Distilled Spirits Council of the USA
EU Soft Drinks Association
European Fruit Juice Association (AIJN)
Shutzgemeinshaft der Fruchtsaft-Industrie (SGF)

regulations. Some of the key beverage trade organizations which seek to influence flavor regulations are also included in Table 9.7.

Not shown are the private, independent organizations which seek to certify flavor manufacturers for compliance with quality systems like ISO, or verify kosher, halal or organic status. The status of these organizations is somewhat fluid, primarily derived from the approval given, tacitly or explicitly, by major flavor customers.

9.2.7 Role of Flavor Analysis in Regulatory Conformance

The increasing sophistication and sensitivity of chemical analysis techniques has given food manufacturers who purchase flavors, regulators, and other interested parties powerful new tools to verify regulatory compliance. However, there are significant 'gray areas' that complicate the 'black and white' picture of verification by analysis. For example, regulations governing the process of manufacturing flavors are very difficult to confirm by product analysis; so, while a microbiological screen can verify that a particular flavor is clean, it cannot demonstrate that the flavor was manufactured using GMP processes. Perhaps more obviously, there will never be an analytical test for a religious designation like kosher. However, analysis can sometime disclose clues about the manufacturing process. For example, the presence of trace levels of a hydrocarbon solvent, like hexane, can indicate that an extraction was performed during manufacturing.

A secondary issue with many flavor analysis methods is that there are few 'official' analytical methods for flavor analysis. While an aroma analysis generally consists of a preconcentration step, followed by GC/MS, a large variety of preconcentration steps are possible choices. The specific

preconcentration step can introduce different recoveries for different components as well as introduce artifacts. Few proprietary methods used by companies or regulatory agencies are subject to the reproducibility and interlab robustness testing that are routine for methods such as AOAC International methods. The result can be differing analyses and interpretations of similar tests done on identical materials at different labs. Many times, there is no independent standard to reference against.

For these reasons, among others, regulatory compliance will never be completely verified by chemical analysis. However, the continuously improving analytical situation has certainly shed light on the use and manufacturing of flavor materials.

9.3 SPECIFIC REGULATORY ISSUES

This section is devoted to a discussion of a number of specific regulatory issues where chemical analysis plays an important role. Almost always this situation involves a verification process that a particular regulatory requirement is being met. Sometimes this could be from the point of view of a regulating agency, trying to confirm compliance. Often, flavor manufacturers are screening their own raw materials and finished products to ensure conformance. While the specific analytical issues are often unique to each problem, the varied cases presented here are intended to provide a good overview by example.

When attempting to bring analytical chemistry techniques to regulatory compliance, it is important to recognize three key steps in the process. First, the analytical strategy must be determined. This may be as simple as choosing a well-established analytical method or as complex as developing a list of analytes that need to be measured. Second, the measurements themselves need to be completed with suitable accuracy. Finally, the interpretation of the results and the corresponding regulations must be completed. The interpretation step, as discussed below, can be difficult.

9.3.1 Identifying the Presence of 'Forbidden' Substances

Perhaps the most straightforward application of analytical chemistry techniques to regulatory issues is in identifying the presence of substances which are prohibited or only allowed in limited amounts. A specific example would be the heavy metal content in flavors. The EU regulation

on flavors (88/388/EEC) specifically limits lead (Pb) content in flavoring products to 10 mg/kg. In the case of a well-defined chemical substance, a suitable test method is usually easy to develop with the right equipment. Sometimes a standard method is available, having been developed for other food products. For more difficult measurements, industry groups sometimes collaborate to develop suitable test methods [22].

However, complications can arise for a number of reasons. For example, a regulated substance may not be well-defined chemically. Allergens provide a good example of this problem. Allergen ingredients which are not declared on the label are prohibited in food blends. Accidental allergen contamination, due to process equipment or raw material contamination, is another concern. However, a test for 'peanut content' is not really a well-defined request to an analytical chemist. With many different peanut byproducts, what exactly constitutes a 'peanut' from a chemical perspective?

Test sensitivity is another potential complication. Many pesticides are approved for use on specific food products only. Approved pesticide usage rates may lead to the presence of pesticide residues on food products in the part per million (ppm) to part per billion (ppb) range. If a sample is tested with a method sensitive to 0.1 ppb, and the result is clean, it would seem to affirm that no unapproved pesticide was used. However, if a pesticide test is much more sensitive than that, how does one interpret a positive result of 0.002 ppb for an 'unapproved' pesticide? The presence of such a low level could be the result of inadvertent exposure during the growing and/or production of the food product. Determining if this is a regulatory concern is a difficult question.

9.3.1.1 Heavy Metals such as Pb, As, Hg, and Cd

Heavy metal content of food and flavoring products is a relatively straightforward example of a limited contaminant. Heavy metals are limited in food products due to health concerns associated with metal accumulation. The allowed limits are generally well defined either in the flavoring products themselves or in the finished food products. Furthermore, the technology for heavy metal testing is well established [23] and the interpretation of test results is straightforward

9.3.1.2 Pesticides

Pesticides, fungicides, miticides, and related products used to improve agricultural yields are very well defined from a regulatory aspect because

these products must satisfy a very detailed government registration and approval process prior to use in most of the world. As part of this process, limits are established for the use of these chemicals on various agricultural products and tests for the residues of these chemicals are established. In fact, for most pesticide-type products, regulatory/agency approved analysis methods exist because analysis methods must be submitted with the pesticide registration [24]. As discussed above, the most difficult regulatory questions concern the interpretation of very low pesticide levels, especially if a particular pesticide is 'not allowed', as well as the applicability of regulatory pesticide limits on agricultural byproducts. For example, whole fruit pesticide limits exist in the US on citrus fruits [25]. How do you interpret those limits with regard to fruit byproducts such as peel comminute and/or peel oils?

9.3.1.3 Environmental Toxins

The presence of certain naturally occurring environmental toxins – like patulin in apple juice, aflatoxin in cereals, or spice and nut products – is a more complex example. Obvious natural toxins (such as the alkaloids in certain mushrooms) were identified a long time ago by their acute symptoms after consumption. New toxins, or suspected toxins, are constantly being brought to regulators' attention by health and/or food researchers based upon long-term health concerns such as carcinogenicity. These toxins often occur at very low levels and regulators can feel pressure to act prior to the development of robust analytical methods. Some examples include patulin in apple products and acrylamide in fried foods [26].

Patulin testing in apple products is an illustrative example. Patulin is a naturallyoccurring toxin generated by molds which can grow on fruit surfaces like apples [27]. Patulin is suspected of causing carcinogenic or mutagenic toxicity in humans [28]. When contaminated apples are processed the toxin can contaminate the resulting juice and juice by-products. Any of these products which are incorporated into a flavoring mixture are a potential source of patulin. Regulatory limits for patulin are in the neighborhood of 25–50 ppb [29]. The development of a robust test by HPLC for patulin was complicated because preconcentration is required and patulin can be unstable under some common conditions. Concentrated efforts by industry groups and others have lead to the development of acceptable methods [22]. However, prior to these developments, it was difficult to obtain consistent results from different laboratories.

9.3.1.4 Allergen Testing

Allergic reactions to specific food products in some individuals have been known for a long time. Regulations that require the label declaration of common allergens have been implemented to address this safety issue. However, the ability to test for allergen content in food or flavor products has been difficult. Our knowledge of the specific agents that cause food allergies is still limited. Also, the concentration levels at which food allergens are dangerous in food is not well established. Both factors likely vary among affected individuals. Also complicating the regulatory picture is whether components of an allergenic material are equally dangerous. Does the oil from soybeans contribute to the allergenic reaction of soybean? Research seems to indicate that most food allergies result from proteins. While there is little economic reason not to declare known allergens on the label, accidental contaminations due to carry-over on the processing equipment used to manufacture the final product or any ingredient is a big concern. Recently, methods that detect specific proteins from known allergens have been developed to test products and process equipment rapidly [30]. Unfortunately, these tests are available for only a few allergens and not enough testing has been done to establish their complete effectiveness.

9.3.2 Testing Whether a Product is 'Natural' or Meets a 'Standard of Identity'

A more challenging question for the analytical chemical chemist is to determine if a 'natural flavor' is truly natural. In this case, one is usually confronted with determining whether a complex mixture of chemicals contains only ingredients which conform to the applicable regulations concerning natural ingredients and that the ingredients were processed in an acceptable fashion.

A similar problem confronts the question of whether a flavor meets the 'standard of identity' for a particular end use. For example, FTNF flavors which are suitable for use without label declaration in products like fruit juices must be made only from ingredients derived from the named fruit. Determining whether a complex apple flavor is derived only from apple by-products can be a challenging task.

In practice, the motivation to adulterate is almost always economic and the testing can be limited to ingredients where there is a large difference in cost between natural and artificial sources. For example, flavor chemicals in the 'green note' class, such as hexanal, trans-2-hexenal

and cis-3-hexenal, are used frequently in natural fruit flavors. Prior to the development of viable biofermentation routes to synthesize 'natural' versions of these compounds [31], they were isolated from plant extracts such as mint oils, usually at very low concentration. Collection and purification was expensive, leading to prices for these natural chemicals at thousands of dollars per kg. At the same time, artificial green notes derived from petrochemicals were available at a fraction of the cost. Naturalness verification for green note compounds was routine.

In general three main approaches are used. First, the identification of compounds which are not found in the natural products is often a simple marker of an adulterated flavor. For example, until recently a key sulfur note from grapefruit – 1-p-methene-8-thiol – was not available except as a synthetic chemical or in very dilute form within grapefruit oils. This chemical is relatively easy to detect by GC/MS and the presence of a significant quantity of this grapefruit thiol was a marker for synthetic adulteration. Another example is the use of synthetic cooling agents in mint flavors. Some major flavor companies have developed very intense 'cooling agents' that are much stronger than natural menthol. Examples include Frescolat by Symrise and TK-10 by Takasago [32]. Many of these compounds are not naturally occurring and therefore are an easy marker for an artificial ingredient. One complication to keep in mind is the presence of trace levels of some synthetic chemicals – such as solvents like hexane – can be introduced as a result of processing aids during the production process. Processing aids are perfectly acceptable in the manufacturing of natural flavors.

A second main strategy is the identification of trace compounds that should or should not be present based upon the available sources of natural products. In this case, detailed knowledge on the source of the natural flavoring ingredients is needed by the analyst. If an important flavor ingredient is only available from a limited number of natural sources, then knowledge of the usual chemical composition of the natural source is very helpful. For example, an FTNF apple flavor suitable for use in apple juice must be made from ingredients entirely derived from apple. One important ester present in apple aroma, a key flavor ingredient of FTNF flavors, is ethyl-2-methyl butyrate (E2MB). In natural apple, the level of E2MB is much less than the other volatile ethyl esters. If this ratio is found to be unusually high in an apple flavor, it is an indication that E2MB from a nonapple source has likely beenadded. Natural vanilla extract is another example. Trace levels of p-hydroxybenzoic acid and vanillic acid in the right ratios with vanillin serve as indicators of natural vanilla extract [33].

Finally, the most sophisticated test strategy is a comparison of physical properties of individual molecules which differ depending upon the natural or artificial source. A simple example is in the case of flavor components which are chiral – coming in both a left- and right- handed geometric isomer. Often, the naturally occurring source favors one form over the other, while synthetic versions are equally distributed in both forms ('racemic'). This was recognized early on in the case of citrus oils which are mostly composed of the limonene in the right-handed or (+)-limonene form. An old test for adulteration with less expensive (-)-limonene was to measure the optical rotation of the oil. Chiral molecules rotate polarized light differently depending upon the enantiomer. Measuring the rotation factor for a citrus oil provided a fast estimate of the (+)-limonene content. Currently, the development of chiral columns for GC analysis and the sophistication of GC equipment allow the measurement of the chiral ratio of individual molecules. Therefore, the enantiomeric ratio of the E2MB within an apple flavor can be used to determine if the source is natural or artificial. This type of measurement can also distinguish between different natural sources of compounds. For example, we find the enantiomeric ratio of a compound like β-pinene differs depending upon the source, even for closely related plants. In the case of citrus oils, β-pinene in lemon oil has 4–7 % in the (+) isomer, while mandarin oil is around 98 % (+) isomer [34]. As this type of equipment becomes more readily available, tables of typical enantiomeric ratios for important chiral compounds from various natural sources are being published.

Natural and synthesized compounds can also differ due to the levels of certain isotopes found in the compound. For example, naturally occurring flavor molecules, often derived from plants, incorporate ^{14}C at levels associated with the amount of ^{14}C present during plant metabolism. However, since the half-life of ^{14}C is around 5700 years, the same molecules derived from petroleum by-products have much less ^{14}C present due to the age of the petroleum. Carbon-14 testing has been heavily utilized to distinguish natural and artificial flavor molecules. Carbon-14 was one of the earliest isotopes tested because of the ease it can be counted using scintillation detectors since it is naturally radioactive. Other isotopes can also be used. For example, ^{13}C, ^{3}H (tritium) or ^{15}N vary in some molecules depending upon the different natural sources [35]. Testing of nonradioactive isotopes is more difficult, usually involving some sort of high-resolution mass spectrometry with or without chemical derivitization. Coupling high-resolution mass spectrometry with gas chromatography allows online determination of isotopes for

many individual flavor components nearly simultaneously. Accelerator mass spectrometry (AMS) has also been applied to this problem. Many examples are available in the scientific literature. However, the equipment and expertise necessary to do these types of measurements are very expensive and limited to a few laboratories. Few companies and other organizations can maintain this type of equipment internally. Often, these measurements are outsourced to specialized laboratories.

Even more recently, the development of SNIF–NMR [36], a method that utilizes NMR technology to determine the isotope ratio for specific locations within a given molecule, gives even more sensitivity. For example, ethanol contains two carbon molecules. A mass spectrometry based measurement of the ^{13}C-to-^{12}C ratio on ethanol will average over both carbons in the molecule. However, NMR can be used to measure this ratio on a specific carbon – say the one attached to the OH group. This more detailed information can shed even more light on the source of important molecules. However, the equipment cost is at least an order of magnitude higher, requiring very specialized laboratories to perform. One example of SNIF–NMR applied to flavor ingredients is vanillin [37].

Clearly, naturalness testing generally requires substantial knowledge on the part of the analyst regarding the type of contaminants and adulterants to expect. Furthermore, each flavor type has different issues. It is difficult for regulatory agencies to possess this detailed knowledge. Practically speaking, this type of information usually resides only in specialty laboratories and within the flavor industry itself.

9.3.3 Testing for Other Regulatory Compliance Requirements

Confirming regulatory compliance with other flavor categories – like GMO-free or country of origin labeling–involve many of the same issues as discussed above. GMO verification involves measuring trace levels of particular proteins found only in the GMO raw material. For most flavor formulations, any protein content is incidental to the formulation so that testing is not likely to be very effective. Country of origin testing, of theoretical interest for some labeling requirements, is practically of little interest for flavors which are generally used in small quantities in the finished product. Even more difficult cases involve verification of organic or kosher status for flavors. Since noncompliance with these requirements does not necessarily change the chemical makeup of the flavor, verification by analysis is difficult. At best, one can look for forbidden substances in the product. Practically speaking, verification is often done by outside groups auditing the production process.

To conclude, flavor manufacturers today has a broad array of regulations and requirements that must be met in the production and sale of their products. These requirements influence the ingredients, the process, the distribution, and the final use of flavor mixtures. It is of interest to the manufacturers, competing manufacturers, regulatory agencies, and the consumer that these regulations and requirements are met. In this quest, analytical chemistry as applied to flavor mixtures plays an important role. While flavors are mostly composed of volatile chemicals, suitable for gas chromatographic analysis, a much broader range of analytical tools must be applied to this problem. A knowledgeable analytical chemist is also indispensable. Continuing research in the tools and methods of flavor analysis will continue to open opportunities in this field.

REFERENCES

1. Regulation (EC) No 2232/96 of the European Parliament and of the Council of 28 October 1996 laying down a Community procedure for flavouring substances used or intended for use in or on foodstuffs. OJ L 299, 23.11.1996, pp. 1–4, http://ec.europa.eu/food/food/chemicalsafety/flavouring/index_en.htm (accessed 31 August 2006).
2. Administration of Food Safety (2003) http://www.ffcr.or.jp/zaidan/FFCRHome.nsf (accessed 31 August 2006); Japanese Food Chemical Research Foundation (2006) http://www.ffcr.or.jp/zaidan/FFCRHome.nsf/pages/list-desin.add-x (accessed 27 April 2011).
3. Substances Generally Regarded as Safe (2005) 21 CFR Part 182 with periodic updates published in the journal *Food Technology*. Most recent update *Food Technology*, **59**, 24.
4. Health Canada (2006) The Food and Drugs Act – Part B, Division 1 and Division 10, http://laws-lois.justice.gc.ca/eng/acts/F-27/ (accessed 27 Apr 2011).
5. Regulation (EC) No 852/2004 of the European Parliament and of the Council of 29 April 2004 on the hygiene of foodstuffs. CS.ES Chapter 13, Volume 34, pp. 319–337.
6. Codex Alimentarius Commission (2003) Recommended International Code of Practice General Principles of Food Hygene. CAC/RCP 1-1969, Rev. 4.
7. Codex Alimentarius Commission (2005) General Standard for Food Additives. CAC/STAN 192-1995, Rev. 6.
8. Food and Drug Administration (2004) Federal Food, Drug, and Cosmetic Act. http://www.fda.gov/opacom/laws/fdcact/fdctoc.htm (accessed 31 August 2006).
9. Current Good Manufacturing Practice in Manufacturing, Packing, or Holding Human Food. 21 CFR Part 110. Revised 2010.
10. Food Standards Australia and New Zealand (FSANZ) (2002) *User Guide to Flavorings and Flavor Enhancers*, http://www.foodstandards.gov.au/thecode/assistanceforindustry/userguides/index.cfm (accessed 31 August 2006).

11. Council Directive 88/388/EEC of 22 June 1988 on the approximation of the laws of the Member States relating to flavourings for use in foodstuffs and to source materials for their production. OJ L 184, 15.7.1988, pp. 61–66.
12. Japanese Food Chemical Research Foundation (2006) *List of Plant or Animal Sources of Natural Flavorings*, http://www.ffcr.or.jp/zaidan/FFCRHome.nsf/pages/list-nat.flavors (accessed 27 Apr 2011).
13. Codex Alimentarius Commission (1987) General requirements for natural flavourings. *CAC/GL 29-1987*.
14. Food Labelling Foods: Labelling of Spices, Flavorings, Colorings and Chemical Preservatives. 21 CFR Part 101.22. Revised 2010.
15. Drawback on Taxpaid Distilled Spirits Used in Manufacturing Nonbeverage Products. 27 CFR Part 17 (1996) *Alcohol and Tobacco Trade and Tax Bureau. Industry Circulars and Rulings for Manufacturers Non-Beverage Products*, http://www.ttb.gov/industrial/mnbp.shtml (accessed 31 August 2006).
16. Health Canada (2006) *The Food and Drugs Act - Part B, Division 2*, http://laws-lois.justice.gc.ca/eng/acts/F-27 (accessed 27 Apr 2011).
17. Distilled Spirits – Standards of Identity, *27 CFR Part 5.22*. Wines – Standards of Identity. *27 CFR Part 4.21*. Use of Ingredients Containing Alcohol in Malt Beverages; Processing of Malt Beverages. *27 CFR Part 7.11*.
18. Council Directive 2001/112/EC of 20 December 2001 relating to fruit juices and certain similar products intended for human consumption. *OJ L 10, 12.1.2002*, pp. 58–66.
19. Canned Fruit Juices. *21 CFR Part 146*.
20. Health Canada (2006) *The Food and Drugs Act - Part B, Division 10*. http://laws-lois.justice.gc.ca/eng/acts/F-27 (accessed 27 Apr 2011).
21. Food Dressings and Flavorings. *21 CFR Part 169*.
22. AOAC Official Method 995.10 (1996) Patulin in apple juice. *J. AOAC*, **79**, 451.
23. US Pharmacopeia (2003) *Food Chemical Codex*, 5th edn, National Academies Press.
24. US Environmental Protection Agency (2006) *Residue Analytical Methods*. http://www.epa.gov/pesticides/science/index.htm (accessed 31 August 2006).
25. Tolerances and Exemptions from Tolerances for Pesticide Chemicals in Food. 40 CFR Part 180.
26. US Food and Drug Administration. (2004) *FDA Action Plan for Acrylamide in Food*, http://www.fda.gov/Food/FoodSafety/FoodContaminantsAdulteration/ChemicalContaminants/Acrylamide/ucm053519.htm (accessed 27 Apr 2011).
27. Harrison, M. A. (1989) Presence and stability of patulin in apple products: a review. *J. Food Safety*, **9**, 147–153.
28. IARC (1986) Patulin. *IARC Monog. Eval. Carcinog. Risk Chem. Humans*, **40**, 8398; and WHO (1990) Patulin. *WHO Food Addit. Ser*., **26**, 143–165.
29. US Food and Drug Administration (2000) Patulin in apple juice, apple juice concentrates and apple juice products. *Fed. Register*, **65**, 37791–37792.
30. Health Canada (2006) *Allergen detection methods – the compendium of food allergen methodologies*, http://www.hc-sc.gc.ca/fn-an/res-rech/analy-meth/allergen/index_eng.php (accessed 27 Apr 2011).
31. Fabre, C. and Goma, G. (1999) A review of the production of green notes. *Perfumer and Flavorist*, **24**, 1.
32. Erman, M. (2004) Progress in physiological cooling agents. *Perfumer and Flavorist*, **29**, 34.

33. IOFI (2000) Information Letter 1271 – Authenticity of Natural Vanilla Products.
34. Dugo, G. *et al*. (2001) Enantiomeric distribution of volatile components of citrus oils by MDGC. *Perfumer and Flavorist*, **26**, 20.
35. Schmidt, C.O. *et al*. (2001) Stable isotope analysis of flavor compounds. *Perfumer and Flavorist*, **26**, 3.
36. Site-Specific Natural Isotope Fractionation – Nuclear Magnetic Resonance. Trademark of Eurofins Laboratories, Nantes, France. *Eurfins. SNIF-NMR – A Unique Method to Prove Authentic Origin*, http://www.eurofins.com/food-testing/food-analyses/snif-nmr/en (accessed 31 August 2006).
37. Tenailleau, E. J., Lancelin, P., Robins, R. J., and Akoka, S. (2004) Authentication of the origin of vanillin using quantitative natural abundance 13C NMR. *J. Ag. Food Chem*., **52**, 7782–7787.

Index

Practical Analysis of Flavor and Fragrance Materials, First Edition.
Edited by Kevin Goodner and Russell Rouseff.