Data Reduction in Gas Chromatography

G.A. Eiceman
New Mexico State University, Las Cruces, USA
J. Gardea-Torresdey
University of Texas, El Paso, USA

The data obtained from a gas chromatographic measurement provide the user with details on both the quantitative and the qualitative composition of the sample. This information is displayed in the chromatogram from which information also can be gleaned concerning the performance and quality of the measurement. Quantitative determinations require attention to the relative response of the detector. The assignment of identity of individual constituents is possible in a formal sense only with authentic standards and mass spectra though retention times or indices can be indicators of identity. These aspects of data reduction have been understood for several decades and chemical information can be extracted from a chromatogram with ease and convenience.

In recent years, advanced chemometrics methods have been applied to chromatographic data with intrinsic advantages for probing patterns and trends too subtle or too complex for simple inspection of chromatograms. These methods are rooted in tools such as multivariate analysis and neural network analysis. Such methods are being developed in the area of natural products analysis.

1 INTRODUCTION

A chromatogram is the immediate result obtained from any analytical measurement by gas chromatography (GC) and discloses, for most detectors, all the chemical information about a sample that will be provided by the GC experiment. In addition, details about the overall performance of the chromatograph can be found in the chromatogram and diagnostics are best made when a known sample such as a standard mixture is analyzed. A chromatogram is a plot of detector response (usually in current or voltage) versus time (usually in minutes) following the introduction of a sample into the chromatograph. Most analyses by GC conclude with an understanding of the concentrations or amounts of one or more constituents in a sample and some indication of the chemical identity of such constituents. These results are readily available from the chromatogram and this accounts for the wide acceptance of GC as an analytical tool. At an elementary level, the information and the methods to extract this information have been unchanged during the last several decades and can be found in specific monographs on GC[1,2] or in general analytical chemistry texts.[3,4] These will be reviewed and illustrated concisely in the next section.

In addition to the routine use of chromatographic data to answer questions such as 'is there benzene in my sample and if so how much is there?', there may be interest in using the chromatogram for comparing with results from other samples. This might be motivated by interests in the origins of the sample or to learn if subtle changes in sample composition had occurred during storage, treatment, or processing of the sample. Normally, human skills of comparison are eclipsed when samples are highly complex or numerous. For example, observing the differences or similarities between chromatograms containing more than 20 peaks can be daunting especially when there may be only minor differences in concentrations of several components. Small differences in chromatograms from GC analysis of volatile organic compounds in blood or urine or breath might be essential in determining an illness, but easily overlooked without computation aids.

While comparisons of complex chromatograms are impractical through manual methods of data reduction, advanced computational methods can provide the necessary tools for such comparisons. These methods have been refined rapidly during the last few years with the wide

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

availability of affordable, powerful and small computers and the release of specialized, commercial software for advanced data reduction. In the third section below, two of the promising methods for advanced data reduction are described.

2 ESSENTIALS OF CHROMATOGRAPHIC DATA

2.1 General Facets of a Chromatogram

In Figure 1, a gas chromatogram is shown from the analysis of a standard mixture of normal alkanes by gas chromatography/mass spectrometry (GC/MS) and the results are useful for illustrating the information that can be gleaned from an analysis by GC. The important terms for the analyst include the number of constituents (seen ideally as separated or well-resolved peaks) and the relative abundances of each constituent (i.e. the relative heights or areas of each peak, as a rough approximation). Each peak can be defined by a few terms including: the retention time (t_r), peak area (or height), peak width (w), peak shape distortion or asymmetry factor (η), and the retention time for an unretained compound (t_m). These all assist in expressing the ultimate goal of a chromatographic experiment, namely, the satisfactory separation or resolution of constituents in a mixture. Other terms such as column length, carrier gas flow rate, inlet pressure, ambient pressure, and column temperature cannot be obtained from the chromatogram and are acquired from independent measurements. Such terms are especially useful when studying the details of chromatographic behavior but are not essential for analytical purposes. All these terms and their meaning have been standardized over the last 30 years[1–4] and are helpful in determining or documenting the composition of the sample or in exploring separation processes.

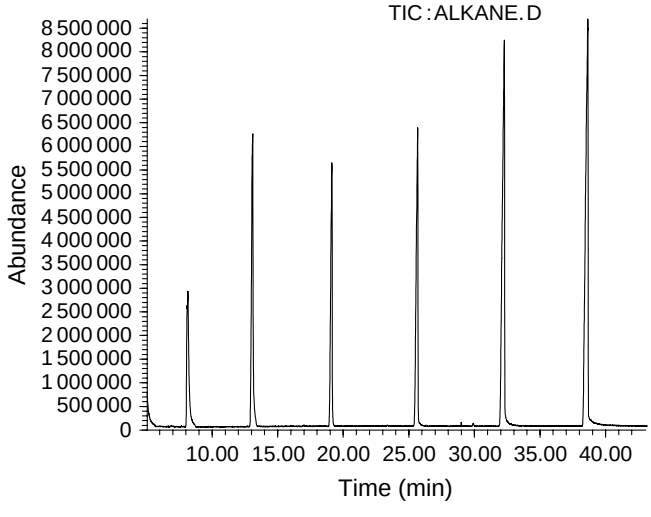

Figure 1 Gas chromatogram of *n*-alkanes taken with temperature programmed analysis. The analysis was made using a capillary column in a gas chromatograph mass spectrometer. Thus, the solvent peak was eluted and vented before data acquisition by the mass spectrometer was started. The chromatogram shows the general appearance of well-resolved peaks in a mixture.

The retention time is the timescale of the chromatographic measurement as shown in the *x*-axis of the chromatogram in Figure 1 and is measured from the point of injection (not seen in the figure). Each peak is assigned a retention time measured at the peak maximum for a substance and this is shown in Table 1. For example, the first peak has a retention time (t_r) of 8.154 min and the second peak shown has a t_r of 13.084 min and these supply information regarding the chemical identity of that peak as described below. In addition, the peak heights or areas contain information about the amounts of substances in the sample and this too is described below. A salient feature of the chromatogram before detailed discussion is the information so often sought: this sample shows the presence of six constituents and the amounts of each, as a first approximation, appear comparable in order of magnitude.

2.2 Quantitative Aspects of Gas Chromatography Data

The size of any peak is a measure of the amount of that chemical in the sample and there is a proportional relationship between peak area (or peak height) and amount of chemical (number of molecules, moles, mass, or concentration). However, all detectors in GC are operated on principles where there is some discrimination, based upon molecular structure, in the sensitivity of response (Δ response/Δ amount). Consequently, any direct comparison of peak heights of various substances without consideration of differences in detector response will be erroneous. Quantitative determinations for GC will be accurate only when the amount or concentration of each analyte is determined using individual calibration curves made for that specific analyte or when relative response factors can be used to normalize a response. Highly selective detectors such as electron capture detector (ECD)

Table 1 Results from analysis of chromatogram in Figure 1

Peak number (see Figure 1)	Retention time (min)	Width (min)	Area	Start time (min)	End time (min)
1	8.154	0.156	330 211 451	7.841	8.912
2	13.084	0.140	555 572 742	12.523	13.453
3	19.125	0.131	444 008 379	18.702	19.485
4	25.666	0.132	576 691 495	25.229	26.189
5	32.272	0.161	990 352 553	31.727	33.174
6	38.637	0.180	1 152 937 448	37.906	39.546

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

Table 2 Differences in response for various chemicals in two GC detectors. Absorption coefficients and normalization factors are measures of relative response for the same amount of chemical used in the original literature. Exact definitions can be found in refs. 5 and 6

ECD			
Chemical	Absorption coefficient	Chemical	Absorption coefficient
Methylene chloride	1	Chlorobenzene	1
1,3,5-Trichlorobenzene	60	*o*-Dichlorobenzene	42
Chloroform	500	*p*-Dichlorobenzene	11
Hexachlorobenzene	1100	Vinylchloride	0.1
FID			
Chemical	Normalization factor	Chemical	Normalization factor
n-Hexane	1.00	1-Butanol	0.66
Cyclopentane	0.98	Acetic acid	0.24
Benzene	0.97	Ethylacetate	0.38
2,2,4-Trimethylpentane	0.98	Acetone	0.49
1-Methyl-3-ethylbenzene	1.03	Ethyl phenyl ketone	0.80

FID, flame ionization detector.

and nitrogen-phosphorus detectors exhibit differences in response factors by orders of magnitude as shown in Table 2.[5] Other detectors such as the FID show a small range of responses within a chemical group such as alkanes though response factors may differ widely between chemical families.[6] In practice, this means that calibration curves should be made for each chemical. Creating individual calibration curves is not burdensome in GC methods since the separation capabilities allow standard solutions to be made as mixtures and calibrations to be made simultaneously for many constituents with a few analyses. The results in Figure 1 were taken from a mass spectrometer detector and the peaks are all alkanes with comparable response factors. Not all peaks have same intensity or height (or area), so differing concentrations between the constituents may be reasonably concluded (Table 1).

If chromatographic performance between samples and standards is identical or very similar, peak heights are suitable for quantitative work or peak areas may be used and are especially helpful when slight variations exist in chromatographic performance. These terms are generated automatically with modern gas chromatographs which are manufactured with data analysis options of varying degrees of sophistication. Difficulties arise in the automated determination of areas for partially resolved components. In the instance when best chromatographic performance is obtained and peaks are still not totally resolved, deconvolution of the chromatogram may be helpful. At present, commercial software packages allow deconvolution of peaks as shown in Figure 2 and these afford refined methods for obtaining peak areas.[7] The additional time and expense in data reduction for these deconvolutions is returned in analytical information that may not be obtainable otherwise.

2.3 Qualitative Aspects of Gas Chromatography Data

A second aspect of data reduction is the determination of the chemical identity for a peak and there is no true substitute for a mass spectrum of a peak as shown in Figure 1. The combination of prefractionation by GC with chemical identification by mass spectrometry (MS) explains, in part, the explosive growth of GC/MS instruments and methodologies. This principle about chemical identity cannot be diminished and is especially critical for samples where there is little or no understanding of possible constituents. In some instances where the number of constituents is limited and the sample is drawn from a well-characterized and fairly stable source, retention time of a peak (t_r) is a de facto measure of chemical identity. In these instances, this measure will suffice as all the proof necessary for an assignment of identity. In other instances, nothing short of a mass spectrum for a peak in a chromatogram and a confirmation of the analyte retention time with an authentic sample will be satisfactory for identifying a constituent in a chromatogram. Regardless, the retention time (t_r) is and has been an essential term in GC analyses and is best reported as the corrected retention time t'_r where the time is corrected and normalized to the time for an unretained solute (t_m) as shown in Equation (1). While selection of an unretained solute is more difficult than may be expected, methane is commonly used when column temperatures are high and the detector is a FID (hence the term t_m). In mass spectrometers, where

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

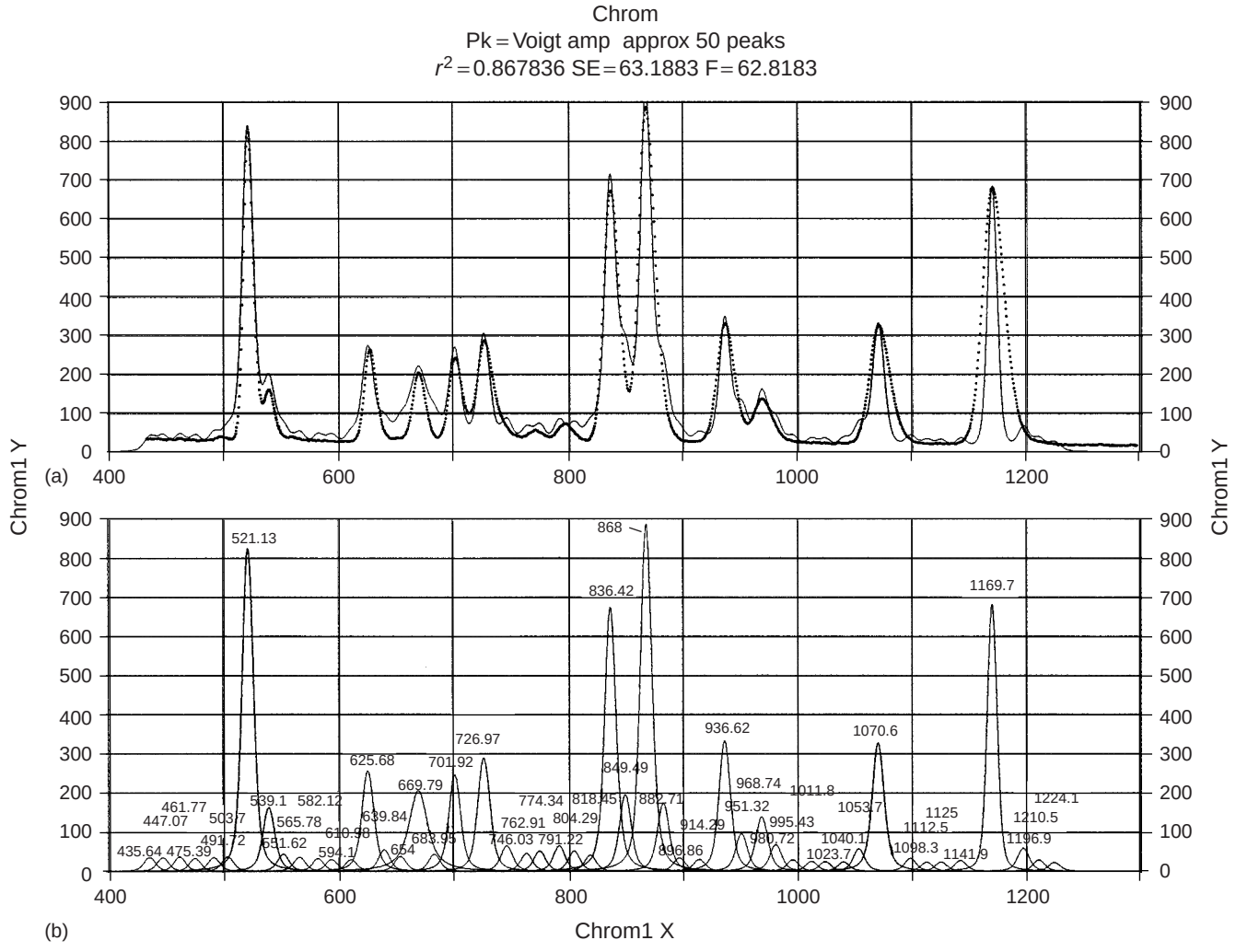

Figure 2 The use of deconvolution in obtaining peak areas for a chromatogram (a) in which there is incomplete chromatographic resolution. (b) The individual peaks as seen from deconvolution are shown. The results were generated using PeakFit 4.0.

the filament is usually started after the solvent or air peak has eluted, a reliable determination of t_m is inconvenient and must be made in a separate determination.

$$t'_r = t_r - t_m \quad (1)$$

The great facility of a corrected retention time is that there is a direct reference to the partition coefficient (K) of the analyte via the phase ratio for the column (β) as shown also in Equation (2). The value K is dependent upon stationary phase,

$$\frac{t'_r}{t_m} = \frac{K}{\beta} \quad (2)$$

chemical compound, and temperature. So, if stationary phase and temperature are controlled, K (and thus t'_r) will be characteristic for a given chemical.

Unfortunately, phase ratios are difficult to obtain or modify so practical chemical measurements require a means to compare results between laboratories. One means used in recent decades to allow comparison of chromatographic results on a given stationary phase is the use of chemical standards to calibrate retention times. A common standard comprised a homologous series of normal hydrocarbons and was called the Kovats index.[8,9] In this method, the retention of a chemical was expressed in carbon numbers, rather then retention time, by referencing retention for constituents to the retention of a set of *n*-alkane standards. Any differences in physical parameters of a column (flow, length, temperature, and phase ratio) would be compensated by this calibration of retention to chemical standards. Such methods were particularly useful since many chromatographers used self-made columns where experimental variables ranged widely. With the advent of commercial capillary columns manufactured with stringent quality control, retention times are matched via software to stored library values

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

and adjustments using carrier gas flows are made to meet standard t_r values. In this way, retention times are directly intercomparable between instruments or laboratory using retention timescales though retention on a column, in a sense, is still being calibrated against a chemical standard.

2.4 The Chromatogram as Diagnostic Tool

A final aspect of the chromatogram is that which provides a full measure of the performance of the gas chromatograph including injection port, column, and detector. A full discussion can be seen in sections on trouble shooting[1] and a brief discussion is given here.

Peaks that exhibit tailing are usually an indication of adsorptive interactions between the sample and surfaces in the inlet, column, or detector, poor flow connections, or overloading of the stationary phase. Active sites, that is highly polar or reactant surfaces in the inlet, column, or detector, are usually remedied through chemical deactivation of these sites. Alternatively, improved materials or new columns or improved designs may be needed.

Peaks that show front skewing are indicative of excessive sample loading on the column and the partition coefficient has been pushed into a region where the isotherm of C_l versus C_g is nonlinear where C_l and C_g are the concentrations of the sample in the liquid and gas phases respectively. This is shown in Figure 3.

When a hydrocarbon sample shows tailing on a fresh, nonpolar or semipolar column, the responsibility is usually poor connections or fitting and is shown in Figure 3. Drifting baselines and more are treated in a recent monograph that experimentalists will find useful.[1]

Band-broadening is a normal part of a chromatogram[10–12] and many of the advances in high resolution chromatography have occurred through improved management of band-broadening.[13] Additional band-broadening can arise through the injection of sample and other extra column contributions. These and baseline changes are treated in a series of practical guides.[14–16]

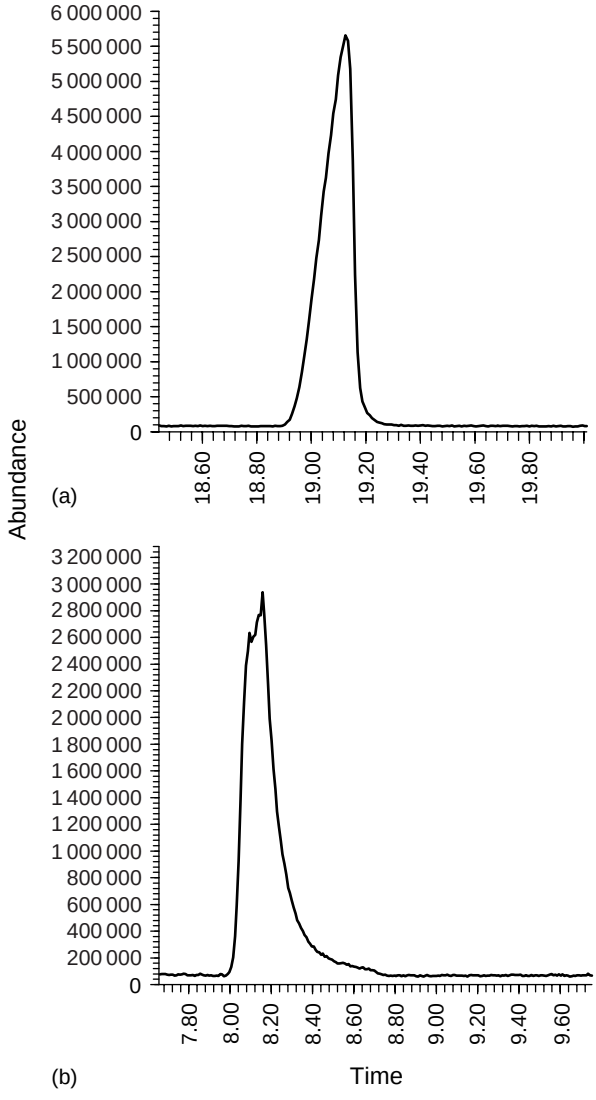

Figure 3 Two examples of peak distortion. (a) shows overloading of the liquid phase and (b) shows tailing indicative of a poor connection. The asymmetry factor is a ratio of distances from a center line.

3 ADVANCED DATA REDUCTION METHODS IN GAS CHROMATOGRAPHY

When chromatograms are simple and information regarding a single (or a few substances) is sought for a measurement, interpretation of the chromatogram can be uncomplicated and rapid. However, when chromatograms contain large numbers of peaks and the patterns defy simple interpretation, advanced methods of data reduction can be useful. There have been two approaches, broadly speaking, to coping with complex samples and these include principal component analysis or cluster analysis and neural network methods. Both methods are computationally intensive and have blossomed with the rise of personal computers in the 1980s and 1990s.

3.1 Multivariate Data Analysis in Gas Chromatography

In multivariate data analysis, data are manipulated so that subtle relationships between constituents in the data can be resolved or clarified. This is done by mathematical methods of recreating the reference points in the information contained in the data and the details of these methods can be found in the articles referenced below. The goal in multivariate analysis is to amplify similarities or differences in data for various samples and generally minimize the number of variables that

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

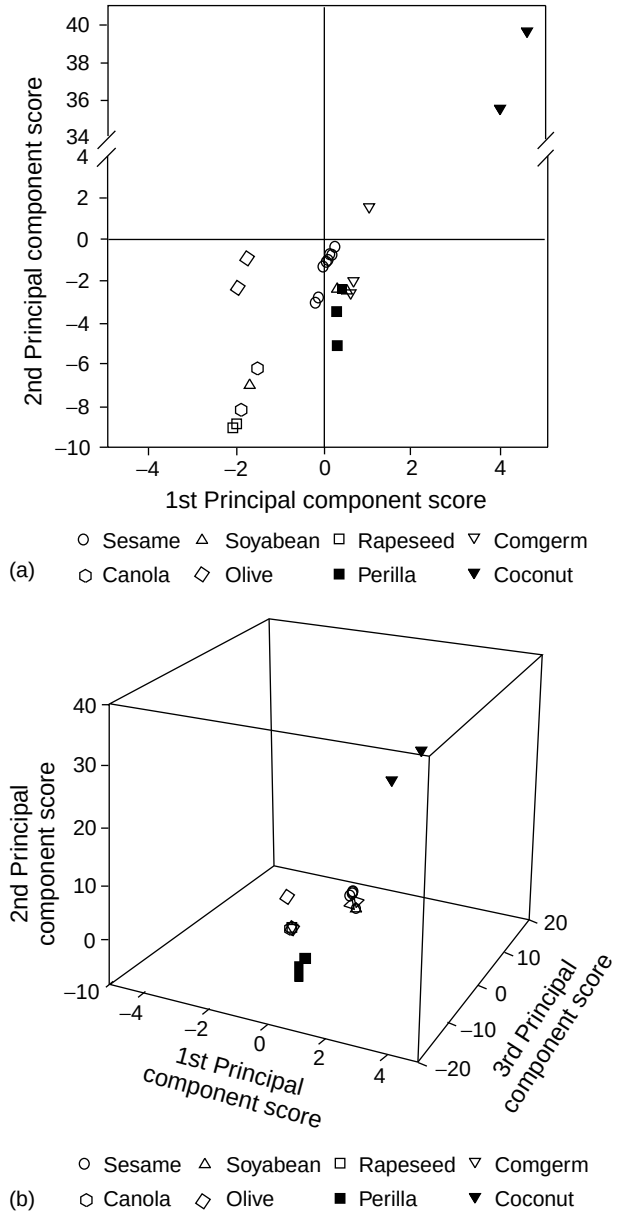

Figure 4 Principal components plot for fatty acid content of vegetable oils in 2-D (a) and 3-D (b) plots. (Drawn from the article by Lee et al.,[17] used with permission.)

describe a sample. The tools and rules to identify or create the terms for comparisons (principal components in principal component analysis) are beyond the scope of this article. It is sufficient to say that certain peaks or regions of a chromatogram may be compared with other peaks as a measure of similarities. When done for several such comparisons, computational methods allow rapid analysis of results. The end goal is the assignment of a sample as a single data point in two-dimensional (2-D) or three-dimensional (3-D) plots or in multidimensional space. The presence of other measurements in proximity to the sample constitutes evidence of a common or unique composition or identity.

An example of one use of these tools is principal component analysis of chromatographic data from characterizing fatty acids in vegetable oils.[17] Though the chromatograms were comparatively simple consisting of only ~10 fatty acids, the number of sample types were numerous including olive, coconut, sesame, and others. The results are shown in Figure 4 where the various types of samples are resolved in 2-D and 3-D.

Another example of a chemometric approach to processing gas chromatographic results can be seen in the highly complex patterns from pyrolysis gas chromatograms of polymers.[18] Polymers were identified or recognized in poor resolution GC through the presence of functional groups such as olefin, aromatic, amide bonds, carbonyls, and others. Such methods of data analysis may be expected to be beneficial in biomedical and medical applications of GC. One example of a successful application of chemometrics in biochemical measurements was demonstrated for urinary organic acid profiles and uterine cervical cancer.[19] In this work, a total of 50 organic acids from urine samples (0.25 mL) were resolved by GC and discriminant (cluster) analysis was performed on the GC data using 16 acids as the variables for discriminating between the two groupings. Discriminant analysis applied to these 16 variables correctly classified 26 urine samples into two separate clusters according to tumor types in the canonical plot.

Advanced data handling of gas chromatographic data has been used in microbiology where the volatile fatty acids in bacteria, long a reliable method to type bacteria through chemotaxonomy,[20] were characterized by GC. The data, when treated by chemometrics allowed subgrouping of 169 coagulase-negative staphylococci collected during an outbreak of nosocomial sepsis in a hematologic unit.[21] Their findings corroborated the results from conventional typing methods showing

Table 3 Examples of multivariate analysis methods and applications with GC data

Sample	Refs.
Foods and beverages	
Blueberries	22
Orange juice	23
Microbial defects in milk	24, 25
Fresh cabbage	26
Almond cultivars	27
Wines	28
Roots of echinacea	29
Non-foodstuffs	
Sulfur in coal by pyrolysis GC	30
Plastics in recycling	31

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

similarities in strain characteristics and, therefore, the same epidemiological origin.

Multivariate analysis methods then, are powerful tools to extract from gas chromatograms details that otherwise elude the unaided eye simply owing to the complexity or subtlety of the data. This is especially so with natural materials such as food odors or vapors. Several more examples are listed in Table 3.[22–31]

3.2 Neural Network Analysis in Gas Chromatography

Another approach to comparing complex chromatograms through what may be nonsubjective computational tools is that of neural network analysis. In all neural networks, the software is used to mimic the human ability to discern patterns. In contrast to humans, computers can operate without fatigue and can survey large amounts of data without the time restraints that humans experience. In neural network methods the network learns to recognize data from known samples. In practice, the chromatogram or the intensity versus retention time pattern of the chromatogram are entered into the input layer after which hidden neuron layers establish weight factors that form a correct answer for the known sample. Other known samples can be trained so that other answers are obtained and an overall weight factor for each neuron in each layer is created, providing that the network properly learns the differences between samples. Thus, unknown samples should be rapidly categorized when chromatographic results for an unknown sample are presented to the trained network. This is especially helpful when the chromatograms are complex and defy reasonable comparisons by humans.

Neural networks have been used in situations where the samples are complex and differences might be due to nonlinear causes. This is especially important in instances where uncertainty exists when two complex mixtures are combined and is illustrated nicely with the recognition of Arochlors in environmental samples.[32] In this approach, a simple architecture was developed in which ~10 peaks were used to classify samples on the basis of the level of Arochlor contamination in environmental samples.

In another example, neural networks were used to classify vegetable oils after GC analysis of samples for methyl ester derivatives of fatty acids[33] and to assign botanical origins of vinegar samples.[34] Polyalcohols, rather than volatile chemicals, were found useful in recognizing the country of origin of various vinegar samples. A similar use of neural networks and GC data from pyrolysis analysis of hashish was found successful in discriminating between hashish obtained from different sources.[35] Lastly, defects in milk were screened using neural network analysis with GC results.[36]

Unlike the traditional methods for interpreting chromatograms, these advanced methods can be expected to undergo further refinements in the years ahead. They do, however, establish a basis from which automated data reduction can be successfully completed and used as a powerful tool in the utilization of GC data.

4 FUTURE DEVELOPMENTS

The current standards for data processing in chromatographic systems afford a high level of convenience and precision in qualitative and quantitative determinations; these meet or exceed the analytical needs for resolving power of contemporary columns. Nevertheless, the technology is rooted in principles that have been long a part of GC separations and these may be expected to continue well into the future. In contrast, the methods necessary to compare analytical findings for a large number of samples are in nascent stages of application. The methods are developed but not well advanced in daily applications. These chemometric methods may be expected to undergo refinement and general adoption during the next decade. This may be especially so in view of the growing and now major emphasis placed on biomedical chemistry.

LIST OF SYMBOLS

t_r	retention time
η	asymmetry factor
t_m	retention time of unretained solute
K	partition coefficient
β	phase ratio for a column
C_g	concentration of sample in gas phase
C_l	concentration of sample in liquid phase
w	peak width

ABBREVIATIONS AND ACRONYMS

ECD	Electron Capture Detector
FID	Flame Ionization Detector
GC	Gas Chromatography
GC/MS	Gas Chromatography/Mass Spectrometry
MS	Mass Spectrometry
2-D	Two-dimensional
3-D	Three-dimensional

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503

RELATED ARTICLES

Chemometrics **(Volume 11)**
Clustering and Classification of Analytical Data • Multivariate Calibration of Analytical Data • Soft Modeling of Analytical Data

General Articles **(Volume 15)**
Quantitative Spectroscopic Calibration

REFERENCES

1. H.M. McNair, J.M. Miller, *Basic Gas Chromatography*, John Wiley & Sons, New York, 29–54, 126–141, 1998.
2. B.L. Karger, L.R. Snyder, C. Horvath, *An Introduction to Separation Science*, John Wiley & Sons, New York, 212–246, 1973.
3. D.C. Harris, *Quantitative Chemical Analysis*, 5th edition, W.H. Freeman & Co., New York, 675–712, 1999.
4. G.D. Christian, *Analytical Chemistry*, 5th edition, John Wiley & Sons, 505–514, 1997.
5. J.E. Lovelock, M.L. Gregory, in *Gas Chromatography*, eds. N. Brenner, J.E. Callen, M.D. Weiss, Academic Press, New York, 1962.
6. W.A. Dietz, 'Response Factors for Gas Chromatographic Analyses', *J. Gas Chrom.*, **5**(2), 68 (1967).
7. PEAKFIT 5.0
8. E. Kovats, 'Gas-chromatographische Charakterisierung Organischer Verbindungen, Tiel 1 Retention Sindices Aliphatischer Hologenide, Alkohole, Aldehyde and Ketone', *Helv. Chim. Acta.*, **206**, 1915 (1958).
9. L.S. Ettre, 'The Kovats Retention Index System', *Anal. Chem.*, **36**, 31A (1964).
10. J.J. van Deemter, F.J. Zuiderweg, A. Klinkenberg, 'Longitudinal Diffusion and Resistance to Mass Transfer as Causes of Non-ideality in Chromatography', *Chem Eng. Sci.*, **5**, 271 (1956).
11. H.W. Moody, 'Evaluation of the Parameters in the Van Deemter Equation', *J. Chem. Educ.*, (1982).
12. R.H. Perrett, J.H. Purnell, 'Contribution of Diffusion and Mass Transfer Processes to Efficiency of Gas Liquid Chromatography Columns', *Anal. Chem.*, **35**, 430 (1963).
13. K.J. Hyver, R. Sandra, *High Resolution Gas Chromatography*, 3rd edition, Hewlett-Packard Corporation, Part No. 5950-3562, 1989.
14. J. Hinshaw, 'GC Troubleshooting – A Systematic Approach', *LC-GC*, **5**, 790–794 (1987).
15. J. Hinshaw, 'GC Troubleshooting – Inlets', *LC-GC*, **5**, 954–960 (1987).
16. J. Hinshaw, 'GC Troubleshooting – Columns', *LC-GC*, **6**, 24–29, 228–231, 794–798.
17. D.-S. Lee, B.-S. Noh, S.-Y. Bae, K. Kim, 'Characterization of Fatty Acids Composition in Vegetable Oils by Gas Chromatography and Chemometrics', *Anal. Chim. Acta.*, **358**, 163–175 (1998).
18. E. Kullik, M. Kaljurand, M. Koel, 'Analysis of Pyrolysis Gas Chromatograms Using Pattern Recognition Techniques', *J. Chromatogr.*, **126**, 249–256 (1976).
19. K.R. Kim, H.G. Park, M.J. Paik, H.S. Ryu, K.S. Oh, S.W. Myung, H.M. Liebich, 'Gas Chromatographic Profiling and Pattern Recognition Analysis of Urinary Organic Acids from Uterine Myoma Patients and Cervical Cancer Patients', *J. Chromatogr., B*, **712**(1–2), 11–22 (1998).
20. G.A. Eiceman, W. Windig, A.P. Snyder, 'Gas Chromatography of Microorganisms', in *Gas Chromatography: Biochemical, Biomedical, and Clinical Applications*, ed. R.E. Clement, John Wiley & Sons, New York, 327–347, 1990.
21. P. Kotilainen, P. Huovinen, E. Eerola, 'Use of Gas-liquid-chromatography for Subgrouping Coagulase-negative Staphylococci During a Nosocomial Sepsis Outbreak', *Eur. J. Clin. Microbiol. Infect. Dis.*, **14**(5), 412–420 (1995).
22. D.W. Baloga, N. Vorsa, L. Lawter, 'Dynamic Headspace Gas Chromatography-Mass Spectrometry Analysis of Volatile Flavor Compounds from Wild Diploid Blueberry Species', *ACS Symp. Ser.*, **596**, 235–247 (1995).
23. P.E. Shaw, M.G. Moshonas, B.S. Buslig, 'Multivariate-analysis for Classification of Commercial Orange Juice Products by Volatile Constituents Using Headspace Gas-chromatography', *ACS Symp. Ser.*, **596**, 33–47 (1995).
24. Y. Horimoto, K. Lee, S. Nakai, 'Classification of Microbial Defects in Milk Using a Dynamic Headspace Gas Chromatograph and Computer-aided Data Processing. 1. Principal Component Similarity Analysis', *J. Agric. Food Chem.*, **45**(3), 733–742 (1997).
25. Y. Horimoto, K. Lee, S. Nakai, 'Classification of Microbial Defects in Milk Using a Dynamic Headspace Gas Chromatograph and Computer-aided Data Processing. 2. Artificial Neural Networks, Partial Least-squares Regression Analysis, and Principal Component Regression Analysis', *J. Agric. Food Chem.*, **45**(3), 743–747 (1997).
26. H. Siegl, A. Hanke, W.H. Schnitzler, 'Cultivar Differences of Volatile Aroma Compounds in Fresh Chinese Cabbage', *Gartenbauwissenshaft*, **62**(1), 17–22 (1997).
27. M.L.M. Carratala, C. Garcia-Lopez, V. Berenguerna-Varro, N. Graneteruel, 'New Contribution to the Chemometric Characterization of Almond Cultivars on the Basis of their Fatty Acid Profiles', *J. Agric. Food Chem.*, **45**(3), 743–747 (1997).
28. J. Weber, M. Beeg, C. Barzsch, K.H. Feller, D.D. Garcia, M. Reichenbacher, K. Danzer, 'Improvement of the Chemometric Variety Characterization of Wines by Improving the Detection Limit for Aroma Compounds', *HRC-J. High. Res. Chrom.*, **22**(6), 322–326 (1999).
29. D. Lienert, E. Anklam, U. Panne, 'Gas Chromatography Mass Spectral Analysis of Roots of Echinacea Species and

DOI: 10.1002/9780470027318.a5503

Classification by Multivariate Data Analysis', *Phytochem. Anal.*, **9**(2), 88–89 (1998).
30. P. Selsbo, P. Almen, I. Ericsson, 'Quantitative Analysis of Sulfur in Coal by Pyrolysis Gas Chromatography and Multivariate Data Analysis', *Energy Fuels*, **10**(3), 751–756 (1996).
31. P.J. Tandr, J.A. Butcher, H. Tao, P.D. Harrington, 'Analysis of Plastic Recycling Products by Expert-systems', *Anal. Chim. Acta.*, **312**(3), 231–244 (1995).
32. G.R. Magelssen, J.W. Elling, 'Chromatography Pattern Recognition of Aroclors Using Iterative Probabilistic Neural Networks', *J. Chromatogr., A*, **775**, 231–242 (1997).
33. R.A. Francelin, F.A.C. Gomide, F.M. Lancas, 'Use of Artificial Neural Networks for the Classification of Vegetable-oils after GC analysis', *Chromatographia*, **35**(3–4), 160–166 (1993).
34. V. Gerbi, G. Zeppa, R. Beltramo, A. Carnacini, A. Antonelli, 'Characterisation of White Vinegars of Different Sources with Artificial Neural Networks', *J. Sci. Food Agric.*, **78**(3), 417–422 (1998).
35. M. Hida, T. Mitsui, Y. Minami, Y. Fujimura, 'Classification of Hashish by Pyrolysis-gas Chromatography', *J. Anal. Appl. Pyrolysis*, **32**, 197–204 (1995).
36. Y. Horimoto, K. Lee, S. Nakai, 'Classification of Microbial Defects in Milk Using a Dynamic Headspace Gas Chromatograph and Computer-aided Data Processing. 2. Artificial Neural Networks, Partial Least-squares Regression Analysis, and Principal Component Regression Analysis', *J. Agric. Food Chem.*, **45**(3), 743–747 (1997).

Encyclopedia of Analytical Chemistry, Online © 2006 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Analytical Chemistry* in 2006 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470027318.a5503