Identification of atmospheric organic sources using the carbon hollow tube-gas chromatography method and factor analysis

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 1989

Impact Factor: 5.64 · DOI: 10.1021/ac00183a011

3 AUTHORS, INCLUDING:

George P Cobb

Baylor University

167 PUBLICATIONS **4,583** CITATIONS

SEE PROFILE

(30) Appelqvist, R.; Marko-Varga, G.; Gorton, L.; Torstensson, A.; Johansson, G. *Anal. Chim. Acta* **1985**, *169*, 237-247.
(31) Appelqvist, R. Doctoral Thesis, Lund, 1987.
(32) Hoffmann, E.; Marko-Varga, G.; Csiky, I.; Jönsson, J. Å. *Int. J. Environ. Anal. Chem.* **1986**, *25*, 161-171.
(33) Aspinall, G. O. *The Polysaccharides*; Academic *Press*: New York, 1985; Vol. 3.
(34) Gorton, L. *J. Chem. Soc., Faraday Trans.* **1986**, *82*, 1245-1258.
(35) Marko-Varga, G., Lund, unpublished results, Jan 1988.
(36) Leyden, D. E.; Collins, W. *Silylated Surfaces*; Gordon and Breach Publishers, Inc.: New York, 1980.
(37) Emnéus, J.; Appelqvist, R.; Marko-Varga, G.; Gorton, L.; Johansson, G. *Anal. Chim. Acta* **1986**, *180*, 3-8.
(38) Kissinger, P. T.; Heineman, W. R. *Laboratory Techniques in Electroanalytical Chemistry*; Marcel Dekker: New York, 1984.
(39) Urbaniczky, C.; Lundström, K. *J. Electroanal. Chem.* **1984**, *176*, 169-182.
(40) Stevens, T. S. *Ind. Res. Dev.* **1983**, *25*, 96.
(41) Marko-Varga, G.; Csiky, I.; Jönsson, J. Å. *Anal. Chem.* **1984**, *56*, 2066-2069.
(42) Csiky, I., Fermenta Products AB, Strängnäs, Sweden, personal communications, March 1988.
(43) Murray, R. W.; Ewing, A. G.; Durst, R. A. *Anal. Chem.* **1987**, *59*, 379A-390A.

RECEIVED for review September 23, 1988. Accepted January 11, 1989. Financial support from the Swedish Board for Technical Development (STUF) is gratefully acknowledged.

Identification of Atmospheric Organic Sources Using the Carbon Hollow Tube–Gas Chromatography Method and Factor Analysis

G. P. Cobb*[1] and R. S. Braman

Department of Chemistry, University of South Florida, Tampa, Florida 33620

R. A. Gilbert

Chemical Engineering Department, University of South Florida, Tampa, Florida 33620

Atmospheric organics were sampled and analyzed by using the carbon hollow tube–gas chromatography method. Chromatograms from spice mixtures, cigarettes, and ambient air were analyzed. Principal factor analysis of row order chromatographic data produces factors which are eigenchromatograms of the components in the samples. Component sources are identified from the eigenchromatograms in all experiments and the individual eigenchromatogram corresponding to a particular source is determined in most cases. Organic sources in ambient air and in cigarettes are identified with 87% certainty. Analysis of clove cigarettes allows the determination of the relative amount of clove in different cigarettes. A new nondestructive quality control method using the hollow tube–gas chromatography analysis is discussed.

INTRODUCTION

The carbon hollow tube–gas chromatography (CHT–GC) method has proven applicable in studies of atmospheric organics (*1, 2*). Complex systems may be sampled with the CHT and the components separated and quantitated by using GC. Dozens of peaks are routinely obtained in chromatograms of atmospheric samples (*2–4*). Data of this complexity presents special data interpretation problems. The number of peaks present in environmental samples makes visual peak matching to determine sources contributing to organic patterns difficult at best. A statistical method is needed for pattern recognition in these complex systems.

Chemometrics, one of the most rapidly growing fields of chemistry, encompasses the various methods of complex data interpretation. One versatile chemometric technique, factor analysis (FA), provides insight into the chemical nature of our environment. Psychologists developed FA during the period between 1900 and 1905 (*3*). Their original goal was to determine the number of factors that contribute to an observed response (*5*). The technique also has the advantage of removing bias from the observations (*3, 5, 6*). Until the advent of modern computers, the laborious matrix diagonalization and inversions required for FA restricted useful applications. With the current accessibility of computers, FA can be applied to large data sets within seconds.

Chemists have used FA to study various physical properties of systems (*6–8*), model reaction intermediates (*9, 10*), correlate toxicity of chemicals based on structure (*11*), and identify molecules by using chromatographic or spectroscopic data (*4, 6, 12–14*). Biennial reviews of chemometric studies have appeared since 1980 (*15–18*).

This paper describes uses of principal factor analysis (PFA) for identifying the contribution of specific, complex organic sources to an air sample. The computations involved are centered on five steps: [1] matrix arrangement of data, [2] data preprocessing, [3] data matrix deconvolution into weighting and factor matricies, [4] error removal, [5] target testing. The reader is referred elsewhere (*6*) for basic details of PFA.

One primary attribute of PFA is extraction of factors (eigenvectors) in order of importance to the data set. The first factor calculated for a data set will describe more of the data than will any other single factor. If these factors were assigned to specific organic sources, individual components of the sample could be identified. Unfortunately, eigenvectors do not directly correspond to unique molecules represented by the data. Despite this fact, PFA of CHT–GC samples can lead to identification of characteristic patterns from volatile organic sources.

In order to give eigenvalues physical significance, a different view of factor determination is useful. The $n \times m$ data matrix (**D**) to be analyzed is constructed of n CHT sample chromatograms with peak areas (retention areas) at m retention times. In all cases discussed, the number of retention areas is greater than the number of experiments. Analysis of **D** constructed

[1] Present address: Institute of Wildlife Toxicology, Western Washington University, Bellingham, WA 98225.

0003-2700/89/0361-0838$01.50/0

from such data will produce a large covariance matrix (C) in route to factor determination. Since a larger C increases PFA calculation time, other investigators analyze data matrices in the form of $\mathbf{D}^T$. In that way the size of C is minimized; however, the resulting eigenvectors have no physical significance. By contrast, analysis of the original $n \times m$ experimental data increases PFA calculation time but now each eigenvector extracted from the data matrix is a chromatogram (eigenchromatogram) of a source contributing to each experiment represented in the data matrix.

The goal of the data manipulation is to produce two matrices, scalar multiples (S) and eigenvectors (F), such that

$$\mathbf{D} = \mathbf{S}\,\mathbf{F} \tag{1}$$

In the course of the PFA, the eigenvector matrix (F) and the associated eigenvalues are determined by using the covarience matrix of D. After determining F, S may be calculated with

$$\mathbf{S} = \mathbf{F}^T\mathbf{D} \tag{2}$$

In the CHT–GC application, each row of F contains an eigenchromatogram having m normalized peak areas. The elements of S reflect the contribution of each eigenchromatogram to the chromatograms in D. Thus, scalar multiple, s_{ij}, is the weighting of the jth eigenchromatogram (factor) in the ith experiment

$$(d_{i,1}, \ldots, d_{i,j}, \ldots d_{i,m}) = s_{i,1}(f_{1,1}, \ldots, f_{1,j}, \ldots, f_{1,m}) + \ldots + s_{i,j}(f_{j,1}, \ldots, f_{j,j}, \ldots, f_{j,m}) + \ldots + s_{i,n}(f_{n,1}, \ldots, f_{n,j}, \ldots f_{n,m}) \tag{3}$$

More generally, a chromatogram is represented as a row in D and is the product of a set of scalars, a row in S, with n eigenchromatograms ⟨comm A⟩ arranged as rows in matrix F

$$(\mathbf{D})_i = (\mathbf{S})_i\mathbf{F} \tag{4}$$

The set of scalars is important because it determines the relative amount each eigenchromatogram that contributes to the chromatogram.

Error is inherent in experimental data. Thus in the eigenanalysis, error vectors are determined along with the main factors (*3*). This fact ensures a number of eigenvectors equal to the smallest dimension of D. For example, PFA of a 6 × 7 data matrix produces an F with six factors, each contributing to the seven retention areas in each chromatogram in D. The first (largest) factors represent eigenchromatograms of real atmospheric components while the last eigenvector(s) are error. In many instances, some of the six vectors in F simply reproduce error in D. These error vectors should be distinguished from the eigenchromatograms at some point in the analysis.

Malanowski's factor indicator (IND) function (*3*, *6*, *19*) is one method of identifying error vectors. The IND function shows a minimum at the last significant eigenvector. All factors with eigenvalues smaller than the indicated vector are considered to represent error. Another error identification procedure involves determining the proportion of the data matrix represented by each factor. All of the eigenvectors collectively account for 100% of the data matrix. The first PFA factor calculated has the largest eigenvalue and accounts for more of the data than any other factor in the set. If each eigenvalue is divided by the sum of all the eigenvalues, then the resulting decimals indicate the percentage of the data matrix each factor represents. The smallest eigenvectors that collectively account for a portion of the data less than or equal to the known experimental error are also said to be error vectors. If these error vectors are removed from F to form an eigenchromatogram matrix, F#, then S must be correspondingly reduced in size to S# and

$$\mathbf{D}\# = \mathbf{S}\#\,\mathbf{F}\# \tag{5}$$

If the error vectors are the only factors removed, the reproduced data (D#) will contain less experimental error than did the raw data.

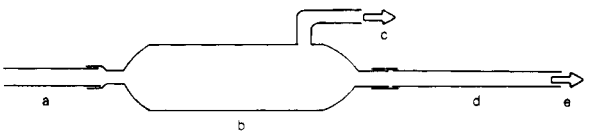

Figure 1. Cigarette sampling apparatus: (a) cigarette, (b) mixing chamber, (c) to main pump, (d) CHT, (e) to sampling pump.

Once the error factors have been found and removed, chromatograms of organic sources may be checked against F# to determine the contributions of suspected sources to the complex data set. This procedure is known as target testing and provides a powerful tool for identifying constituents in a data set (*4*, *6*, *7*, *10*, *11*, *13*).

EXPERIMENTAL SECTION

Air Sampling. Atmospheric samples containing organics from spices, cigarettes, or organic solvents were studied. In the first set of experiments, spice samples were placed in a chamber and 200 mL/min of helium was passed through it. Each chamber contained 1 g of spice(s). Four spices—cinnamon, clove, ginger, and nutmeg—were mixed in various ratios—11:1.2:74:14, 20:8:52:20, 20:16:33:31, and 25:25:25:25. The organics in the carrier were collected in a CHT. The CHT–GC analysis was performed on each sample. Clove, ginger, and nutmeg standards were also sampled individually for use in target tests. In the second set of experiments, air was drawn through five commercial brands of unburned cigarettes at a rate of 0.5 L/min for 2 min. Three individual clove cigarettes were also analyzed in this way for comparison purposes. Each type of cigarette was then burned and the smoke drawn into a 1-L mixing chamber (Figure 1) at 0.5 L/min. During the burning process, samples were taken from the chamber at a rate of 50 mL/min for 0.25 min. A clove standard was also analyzed. In the final set of experiments, the air inside a solvent cabinet was sampled for 5 min. Air samples were also taken for 30–100 min at varying distances from the cabinet.

After loading, tube contents were thermally desorbed into a HP-3790A capillary column gas chromatograph for analysis. This procedure has been described previously (*1*, *2*). All GC analyses were carried out with 0.52 mm × 30 m SE-30 column at a flow rate of 1.1 mL/min. The normal temperature program was 10 min at 30 °C, 4 °C/min to 250, hold 2 min. The spice samples were also analyzed with no initial hold at 30 °C, and a 4-min initial hold was used in the solvent cabinet experiment.

Factor Analysis. Computations were carried out with a Basic V program written for a Prime 750 computer. The program was checked for proper function by reproducing the results of previously published data analysis (*6*). Eigenvectors, eigenvalues, scalar multiples and reconstructed data were all calculated initially. Data may also be reconstructed from a subset of eigenvectors. This reduces the error contribution in reconstructed D#. Test vectors may be analyzed by using any number of the eigenvectors. Any test vector may also be matched with the experimental data to find significant contributions from external sources or error.

In the analysis of laboratory air, a data filter was used to reduce the number of elements in the data matrix. Five submatricies were generated from the original 6 × 82 CHT–GC data matrix. These submatricies were obtained by increasing the threshold for peak recognition. As the threshold was raised, zero was assigned as a replacement for any retention area removed from the data matrix. If a zero was assigned to the same retention area in five of the six chromatograms, then that data column was removed from the data matrix. This action prevented the sixth chromatogram from automatically being classified as an unique member of the data set.

DISCUSSION

Spice Study. Chromatographic analyses of spice samples are shown in Figure 2. Standards (Figure 2b,c,d) show marked differences in chromatographic patterns. PFA of spice data produces four factors, but the relative size of the first three factors (Table I) indicates their predominance over the

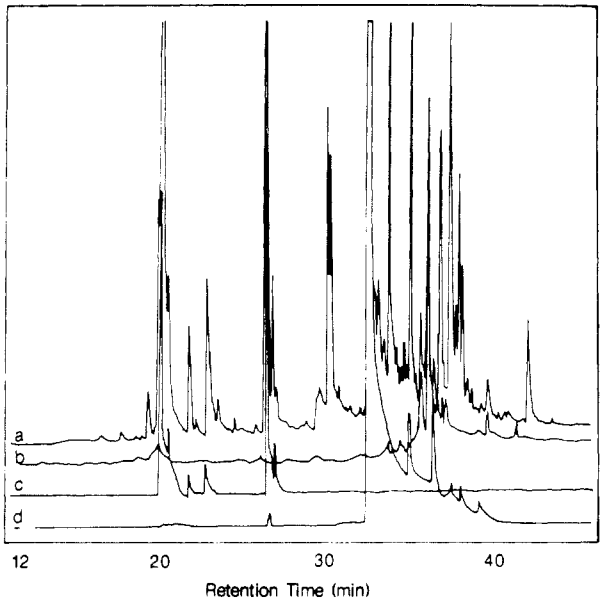

Figure 2. Spice chromatograms: (a) mixture of nutmeg, ginger, clove, and cinnamon, (b) ginger, (c) nutmeg, (d) clove.

Table I. Eigenvalues of Spice Factors

factor	eigenvalue
1	2938
2	332
3	117
4	9

Table II. Percent Error in Spice Target Tests

spice	long analysis	short analysis
nutmeg	49.0	42.6
ginger	83.1	35.4
clove	46.5	28.3

fourth and suggests that the fourth vector is due to error. The largest elements of the principal factor are found in both nutmeg and ginger.

The fact that both nutmeg and ginger are represented in the most important eigenchromatogram suggests that these spices will not be well represented in the remaining factors. Since the fourth factor represents experimental error and should not be considered significant, only the middle two eigenchromatograms might be assigned to individual spices. Target testing was used to aid in this assignment. The results (Table II, long analysis) indicate a large error in the ginger target test. Also the other two spice chromatograms show nearly 50% error when tested in the same data space. This indicates that none of the spices fit the data space well. This is puzzling considering the spice mixtures all contained each of the test spices. The inconsistency could have developed due to uncertainty in assigning retention times to chromatographic peaks. Long elution times increase retention time errors. Therefore, mistakes may be made when assigning retention times to slowly, closely eluting components.

To improve the reliability in retention times and reduce the chance of peak misassignment, the chromatographic analysis time was shortened by 25%. PFA of these chromatograms followed by target tests with the standard chromatograms produced three good fits (Table II, short analysis). Although these conditions reduced the number of resolvable peaks, the decrease also reduced the chance of misassignment and simplified retention time assignments. Improved precision in retention time assignments improves the reliability of the

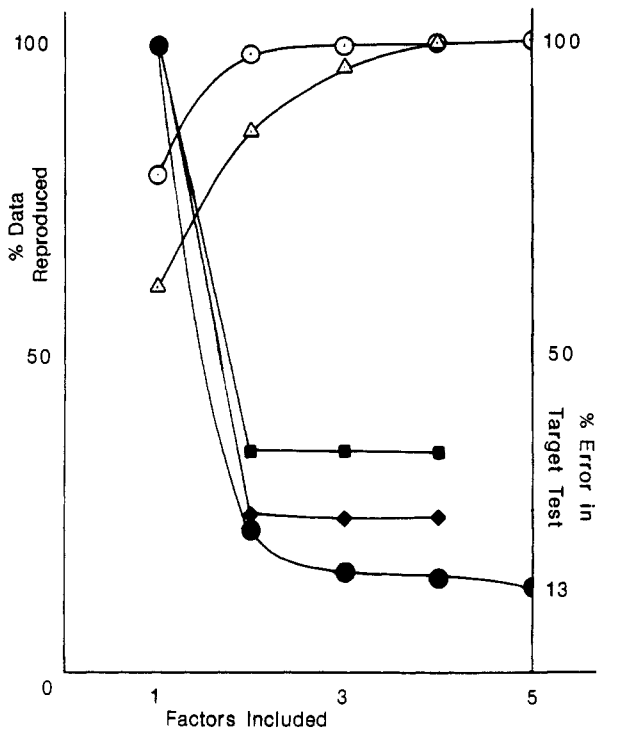

Figure 3. Cigarette data reproduction and target test error: data reproduction (○) cigarette, (△) burned cigarette; target test errors (●) cigarette test in smoked cigarette factor space, (■) clove test in smoked cigarette factor space, (◆) clove target test into cigarette factor space.

data interpretation. Fortunately, PFA can account for the presence of more than one component in a chromatographic peak if the relative percentages of these components are varied from analysis to analysis. Since this variation existed in the spice experiment, the PFA short analysis isolated the sources of the spice mixtures with greater certainty from data collected in a shorter period of time!

Cigarette Study. In a similar experiment, five unburned cigarettes, one of which contained clove, were analyzed. Dried tobacco contains few volatile organics, thus lowering the number of observable peaks. Six principle peaks from each sample were used to construct **D**. The upper portion of Figure 3 shows the percentage of the data described by increasing the number of factors considered. For all unburned cigarettes, three factors account for over 99% of the data. Therefore, only these three eigenchromatograms can describe sources of volatile organics.

A target test analysis was used to associate the second eigenchromatogram with the clove component of a clove cigarette. The error associated with this clove test vector fit into the unburned cigarette factor space was 13%. Additional evidence that the second eigenchromatogram represents clove in the cigarette was obtained from a graph of test vector error versus factors removed (Figure 3). Removal of the eigenchromatogram which represents the source being tested will drastically increase the error in test vector fit. Removing the smallest three factors from the target space only generated a small increase in clove test vector error. In contrast, removal of the second factor causes near 100% error in the test vector (Figure 3). Therefore, the second factor describes the clove in the cigarette and the scalars associated with that vector reflect the amount of clove in the original cigarette chromatograms.

The success in matching the clove factor to an eigenchromatogram suggests an interesting quality assurance procedure. To simulate a quality assurance process, chromatograms of three clove cigarettes were analyzed. Each chromatogram was

Table III. Scalar Multiples for Factors in Cigarettes

cigarette type	1	2	3	4	5
		Experiment 1			
clove	1.967	9.801	−0.246	−0.014	−0.003
Canadian	9.990	−0.330	0.038	0.277	0.209
U.S. regular	9.989	−0.926	−0.938	−0.646	0.006
U.S. low tar	9.875	−0.885	−1.259	0.414	−0.079
unfiltered	9.749	0.197	2.239	−0.044	−0.079
		Experiment 2			
clove	2.658	9.636	−0.191	−0.011	−0.003
Canadian	9.983	−0.502	0.054	0.283	0.207
U.S. regular	9.885	−0.983	−0.978	−0.649	−0.008
U.S. low tar	9.865	−0.963	−1.280	0.417	−0.138
unfiltered	9.739	−0.142	2.286	−0.051	−0.079
		Experiment 3			
clove	2.447	9.692	−0.297	−0.015	−0.004
Canadian	9.988	−0.404	0.013	0.266	0.188
U.S. regular	9.879	−1.119	−0.893	−0.647	0.002
U.S. low tar	9.858	−1.115	−1.213	0.417	−0.123
unfiltered	9.759	0.241	2.190	−0.036	−0.069

Table IV. Scalar Ratios for Clove Content in Three Clove Cigarettes[a]

experiment	ratio
1	4.98
2	3.63
3	3.96

[a] Mean ± standard deviation was 4.19 ± 0.7.

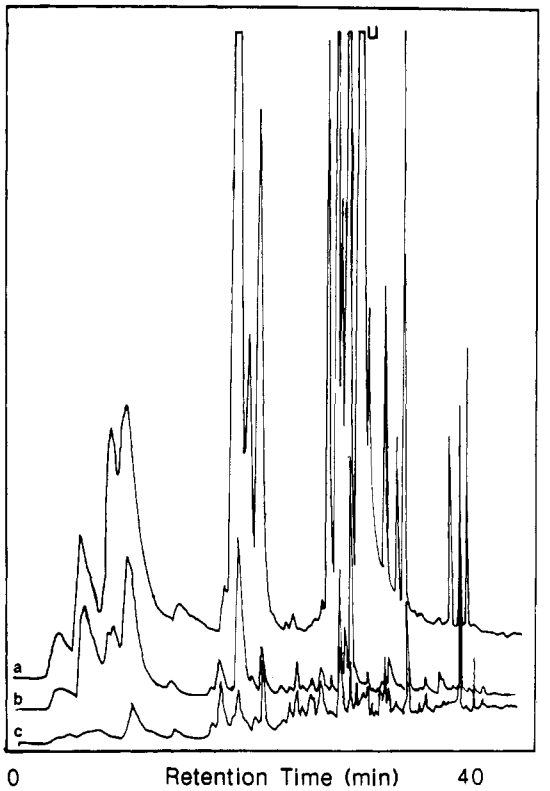

Figure 4. Laboratory air chromatograms: (a) solvent cabinet, (b and c) lab air, (u) unique analyte.

used, in turn, as the first row of **D**. In each analysis, the balance of **D** was composed of the same four regular cigarette chromatograms. Therefore, the four regular cigarettes need only be analyzed once. Five eigenchromatograms resulted from each PFA. The clove chromatogram target tested into each data space with an 85 ± 2% confidence. In each PFA, the scalar multiple of the second eigenchromatogram reflects the relative amount of clove in the cigarettes (Table III). For the three clove cigarette samples, scalar values ($s_{1,2}$) differed by a maximum of 0.86%, and the three corresponding ratios, $s_{1,2}/s_{1,1}$, differed by a maximum of 17% (Table IV). Since the first eigenchromatogram corresponded to the tobacco volatiles in the sample, these $s_{1,2}/s_{1,1}$ values indicated the clove/tobacco ratio in the three individual clove cigarettes.

The CHT–GC analysis of clove cigarettes followed by PFA could easily be adapted for nondestructive quality assurance. In this example, the CHT–GC–PFA was capable of detecting deviations in the clove concentration and variations in the clove/tobacco ratio in cigarettes. In general, the analysis of any sample for its characteristic volatile organics would be facile. When the pattern of any component in a sample can be associated with its eigenchromatogram, quantitation of that organic pattern would quantitate the source of the pattern. This happens because the eigenchromatogram represents the volatile organics produced by the component being studied and the scalar of this eigenchromatogram will describe the relative amount of that component in each sample. Replacing a sample chromatogram by a standard chromatogram would result in a scalar value calibrated for the desired component.

Identification of cigarette components in cigarette smoke was also accomplished. Four different cigarettes were burned. Once again one of these was a clove cigarette. The largest of the four eigenchromatograms was assigned to the tobacco smoke. The chromatograms of clove and the unburned clove cigarette were both used as test vectors. Target tests in the factor space that described the burned cigarette indicated clove as a factor with 62% certainty. The clove cigarette had 86% certainty of being a factor. The target test analysis summarized in Figure 3 identifies the second factor as the unsmoked clove cigarette.

It is interesting that the identification of volatile components in a cigarette and the identification of components from cigarette smoke have the same certainty, 87%. Also noteworthy is that the certainty of clove being a factor describing the cigarette smoke data space is the product of [1] the certainty of clove as a factor in a clove cigarette (85%) and [2] the certainty of clove cigarette as a factor in the cigarette smoke (86%). This calculation predicts that the clove should fit the smoked cigarette factor space at 73% confidence. The relative difference in calculated and observed fits is 13%.

Laboratory Air Quality Study. CHT–GC is also a valuable tool for laboratory air quality analysis. A large laboratory space that contained a solvent cabinet was used in this study. Samples were taken at varying distances from the cabinet, a complex point source. After each sample was analyzed (Figure 4), a PFA was performed. Variations in component concentrations were assured by the wide range of diffusion rates for components in the cabinet. Each of the six raw chromatograms contain 82 integral peaks. Only the largest peaks from each analysis were considered. The number of elements in **D** was incrementally decreased with the data filter routine. Initial data filtering resulted in a 6 × 29 data submatrix. Uniqueness testing and further data filtering produced four additional submatricies. Results of the PFA changed with each compression of the data set. The top four curves in Figure 5 show the data reproduction error for each of the final four submatricies versus the number of factors used. As the number of peaks in the data set is increased, the relative size of the smaller eigenvalues increased.

The actual number of error eigenvectors in each of the four submatricies was estimated by evaluating the information obtained from four different error determination methods: [1] the IND function, [2] the relative magnitude of the eigenvalues, [3] comparison of reproduction errors to a known value for experimental error, [4] comparison of reproduction errors among different size data submatrices. The IND

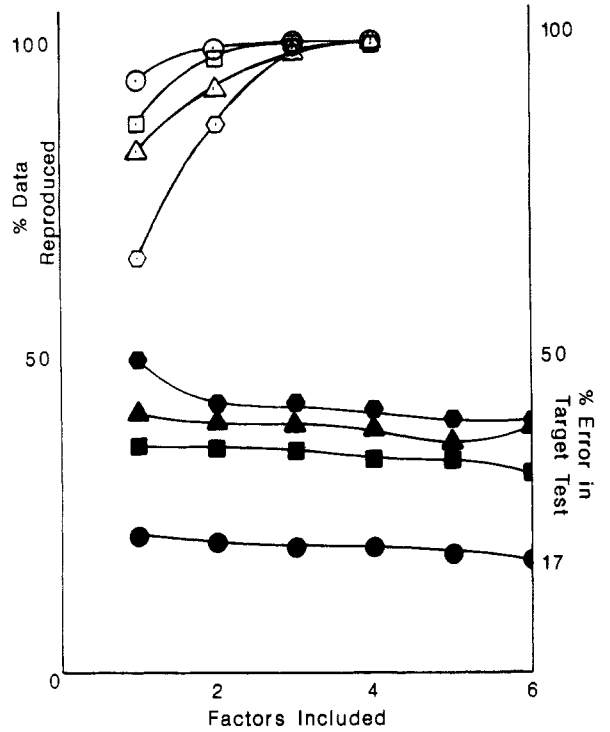

Figure 5. Laboratory air data reproduction and target test error: (○) 11 peaks considered; (□) 16 peaks considered; (△) 20 peaks considered; (○) 27 peaks considered; solvent cabinet target tests in solid symbols.

function analysis predicts four eigenchromatograms. Unfortunately, the IND function is not well understood (*6*) and an analysis of relative eigenvalue magnitudes indicates that each of the data submatricies required their first three factors to achieve 98% reproduction. This suggests a maximum of three data eigenchromatograms and three remaining error vectors. Knowledge of the precision of the CHT–GC method (*1*) provides additional insight to the exact number of error vectors. The CHT–GC precision is ±10% and the smallest four factors describe less than 10% of the data in three of the submatrices and 13% of the data in the other. This analysis assigns these four small factors as error vectors. Finally, a comparison of data reproduction error among the room air submatricies in Figure 5 also substantiates the presence of two eigenchromatograms with physical significance. Three factors are required to reproduce 98% of the data in each case. The submatricies, from which the plotted eigenvalues were extracted, differed only in the number of retention areas in each data vector. The data filtering method ensured that retention areas removed from the data matrix were the smallest in the matrix. Since error in the data sets must decrease as the smaller peaks are deleted (this is true because there are fewer retention areas contributing error and because relatively uncertainty in the area of a small peak is large), the only way in which these four data submatrices can be reproduced to the same level of accuracy is to include error in the reproduction. In this case, convergence occurs at three factors and the third factor is concluded to be an error vector.

With the number of important eigenvectors for the lab air space data established, their physical significances needed to be determined. Factors describing the submatricies of the lab air data matrix were used to monitor the change in the solvent cabinet test vector error as factors were removed from the test space. For each submatrix, the solvent cabinet test vector was adequately described by the first eigenchromatogram (Figure 5). For the three smallest submatricies, deleting all but one eigenvector used to form the transform matrix did not increase the test vector error. Therefore, the first factor must be the eigenchromatogram of the organics that entered lab air from the solvent cabinet. When target testing the solvent cabinet chromatogram in the space described by the eigenchromatograms of the largest submatrix, the 6 × 27 matrix, an 8% error was obtained upon discarding factor 2. This error increase is attributed to the large number of zeros dispersed throughout the data matrix because of the uncertainty in retention time assignments. Consider air samples taken at different points on a line radiating from an organic source. If the chromatogram of one of the samples has a zero response at retention time t_r, where other samples have a response, the value at retention time t_r in the first eigenchromatogram will be reduced by virtue of that zero. Therefore, the value at t_r in the second eigenchromatogram would now have to account for the residual t_r retention areas from the other chromatograms. An abundance of such zeros can spoil a target test.

Although the target test of the solvent cabinet chromatogram with the first eigenchromatogram from each of the four submatrices was ultimately accomplished, the test of the original solvent cabinet vector with the original 6 × 29 data submatrix was unsuccessful. The test failure was attributed to the existence of retention areas present in the solvent cabinet chromatogram but not found in any of the room air sample chromatograms. Uniqueness testing was used to determine which retention areas were specific to the solvent cabinet chromatogram. This procedure involved target testing vectors of unit length that contain a single entry. The unit vectors are constructed by replacing the retention area at the t_r in question with a one. Uniqueness of the retention areas tested is assured when the magnitude of the predicted retention values is large compared to the other areas in the predicted vector (*6*). In this study, the uniqueness test showed one peak (Figure 4u), which was responsible for 70% of the error in the original target test, as well as one insignificant peak to be unique to the solvent cabinet. Deletion of these two retention areas from consideration in the PFA and additional filtering produced the four submatrices used to obtain the successful results shown in Figure 5.

Further examination of the solvent cabinet chromatogram suggests an explanation for the absence of the significant unique peak from the room air. Since that component has a long retention time, it must have a low volatility and should diffuse slowly in air. Also, the component is shown to be polar by its tailing GC peak. Adsorption of polar components on the cabinet interior and any surface outside the cabinet is likely. Finally, there is the possibility of reaction outside the cabinet. Both atmospheric and surface reaction would decrease the concentration of analyte in the room air. The effluent from a NO_x analyzer was vented 5 ft from the solvent cabinet. The ozone released could react with this organic molecule to remove it from the room air. This also may explain some of the spurious data obtained when components were present at low concentration.

SUMMARY

The three studies show the usefulness of PFA in identifying complex sources of atmospheric components. The spice study illustrated that PFA analysis could determine a chromatographic fingerprint of the volatile spice component in mixtures even when the chromatography conditions are not optimal. The cigarette study indicated the possibility of using the CHT–GC–PFA method for quality-control applications. Percentages of solids, liquid, or gaseous mixtures in a product can be determined by the quantity of characteristic volatiles released from the product. If the analysis of a solid or liquid is carried out prior to packaging, no loss of product would result from this analysis. The cigarette and solvent cabinet studies indicate the methods applicability to solving indoor pollution problems. Problem sources could be identified and

emissions stopped, reduced, or vented. A study of cigarette patterns indoors could also prove useful in discussing secondary exposure to cigarette smoke.

Row order data entry produces easily interpretable eigenchromatograms. These abstract chromatograms often correspond to chromatograms of volatile organics from a particular source. With further refinement and establishment of a library of volatile patterns from common organic sources, eigenchromatograms could be matched to chromatograms of the indexed sources. This approach could be invaluable in off-site monitoring of industrial atmospheric effluents in areas where there are several major sources of atmospheric organics.

ACKNOWLEDGMENT

The intricacies of the original computer program were explained by Mark Dale. His program was the backbone of the program used in this study. Timely suggestions for improved data analysis were provided by Dr. J. A. Llewellyn.

LITERATURE CITED

(1) Cobb, G. P.; Hua, L. M.; Braman, R. S. *Anal. Chem.* **1986**, *58*(11), 2213-2217.
(2) Cobb, G. P.; Braman, R. S. *Measurement of Toxic Air Pollutants*; USEPA/APCA: Raleigh, NC, 1986; pp 314-325.
(3) Harman, H. H. *Modern Factor Analysis*; The University of Chicago Press: Chicago, IL, 1976.
(4) Seivers, R. *Measurement of Toxic Air Pollutants*; USEPA/APCA: Raleigh, NC, 1986; pp 287-303.
(5) Amick, D. A.; Walberg, H. J. *Introduction to Multivarient Analysis*; McCutchan Publishing Co.: Berkley, CA, 1975.
(6) Malanowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; John Wiley & Sons, Inc.: New York, 1980.
(7) Coomans, D.; Broeckaert, I. *Potential Pattern Recognition in Chemical and Medicinal Decision Making*; John Wiley & Sons, Inc.: New York, 1986.
(8) Windig, W.; Jakab, E. *Anal. Chem.* **1987**, *59*(2), 317-323.
(9) Gilbert, R. A.; Llewellyn, J. A.; Swartz, W. E.; Palmer, J. W. *Appl. Spectrosc.* **1982**, *36*(4), 428-430.
(10) Gilbert, R. A.; Llewellyn, J. A.; Swartz, W. E. *Appl. Spectrosc.* **1985**, *39*(2), 316-320.
(11) Jurs, P. C. *Science* **1986**, *232*, 1219-1224.
(12) McConnell, M. L.; Rhodes, G.; Watson, U.; Novotny, M. *J. Chromatogr.* **1979**, *162*, 495-506.
(13) Lavine, B.; Carlson, D. *Anal. Chem.* **1987**, *59*(6), 468A-470A.
(14) Lohniger, H.; Varmuza, K. *Anal. Chem.* **1987**, *59*(2), 236-244.
(15) Ramos, L. S. *Anal. Chem.* **1986**, *58*(5), 294R-315R.
(16) Delaney, M. F. *Anal. Chem.* **1984**, *56*(5), 261R-277R.
(17) Ildiko, E. F.; Kowalski, B. R. *Anal. Chem.* **1982**, *54*(5), 232R-243R.
(18) Kowalski, B. R. *Anal. Chem.* **1980**, *54*(5), 112R-122R.
(19) Uguagliati, P.; Benedetti, A. *Comput. Chem.* **1984**, *8*(3), 161-168.

RECEIVED for review August 25, 1988. Accepted January 19, 1989.

Fundamental Factors in the Performance of Diffusive Samplers

Daniel P. Adley

Liberty Mutual Insurance Company, 600 Grant Street, Pittsburgh, Pennsylvania 15219

Dwight W. Underhill*

School of Public Health, University of South Carolina, Columbia, South Carolina 29208

The performance of diffusive samplers can be determined from dimensionless parameters characterizing the sampling time and the geometry of the diffusive sampler. Should the sorption isotherm contain adjustable parameters, the number of dimensionless parameters increases accordingly. For linear and irreversible isotherms, sampler performance was calculated in closed form, whereas for the Dubinin–Radushkevich isotherm (which describes the uptake of many compounds on activated carbon), sampler performance was determined by using the Crank–Nicolson implicit numerical procedure. For all three isotherms, generalized plots of sampling efficiency were developed as a function of sampling time and sampler geometry. With the linear isotherm, a very rapid loss of sampling efficiency occurs, whereas for both the Dubinin–Radushkevich and the irreversible isotherms, high sampling efficiencies can be maintained until a sizable fraction of the adsorbent is saturated. These calculations show the performance that can be expected from a diffusive sampler and, in particular, explain the high performance seen with diffusive samplers containing an activated carbon adsorbent.

INTRODUCTION

Figure 1 gives a schematic diagram of a common type of diffusive sampler. While this sampler is in use, the analyte diffuses across an internal air gap to a sorbent, which serves to retain it. The windshield at the front of the sampler serves as a barrier to convective air movement, which otherwise would disrupt the rate-limiting aspect of the diffusion across the air gap in the diffusive sampler. The impervious barrier supporting the sorbent prevents loss of analyte through the back of the sampler. Ideally a diffusive sampler would, until the sorbent is exhausted, have an uptake rate exactly proportional to the instantaneous ambient concentration of analyte. This ideal response is usually only approximated because (1) for many analytes there is no useful sorbent for which the sorption process is irreversible and therefore loss of some sample by desorption and reverse diffusion occurs and (2) as the more exposed layers of sorbent become saturated, the distance that the analyte must travel before reaching fresh sorbent increases, decreasing the sampling rate of the analyte. Any realistic model of a passive sampler must include both effects.

THEORY

Equations for Mass Transfer in a Diffusive Sampler. The model developed here includes these factors: (1) the external concentration of analyte, (2) the width of the air gap, (3) the diffusion coefficient of the analyte in air, (4) the porosity, (5) tortuosity, and (6) thickness of the sorbent pad, and (7) the sorption isotherm. In the following analysis, these factors are combined into a minimum of two dimensionless parameters, i.e. one dimensionless parameter describing the sampling time and another the geometry of the sampler and, should the sorption isotherm have adjustable parameters, additional nondimensional parameters for each of these.

0003-2700/89/0361-0843$01.50/0