Anal. Chem. **2001,** *73,* 596−605

Application of Quantitative Chemometric Analysis Techniques to Direct Sampling Mass Spectrometry

William P. Gardner,[†,‡] Ronald E. Shaffer,[§,⊥] James E. Girard,[†] and John H. Callahan*[,§]

American University, 4400 Massachusetts Avenue, Washington, D.C. 20016; Naval Research Laboratory, Code 6115, 4555 Overlook Avenue, Washington, D.C. 20375, and Geo-Centers, Inc. at NRL

This paper explores the use of direct sampling mass spectrometry coupled with multivariate chemometric analysis techniques for the analysis of sample mixtures containing analytes with similar mass spectra. Water samples containing varying mixtures of toluene, ethyl benzene, and cumene were analyzed by purge-and-trap/direct sampling mass spectrometry. Multivariate calibration models were built using partial least-squares regression (PLS), trilinear partial least-squares regression (tri-PLS), and parallel factor analysis (PARAFAC), with the latter two methods taking advantage of the differences in the temporal profiles of the analytes. The prediction errors for each model were compared to those obtained with simple univariate regression. Multivariate quantitative methods were found to be superior to univariate regression when a unique ion for quantitation could not be found. For prediction samples that contained unmodeled, interfering compounds, PARAFAC outperformed the other analysis methods. The uniqueness of the PARAFAC model allows for estimation of the mass spectra of the interfering compounds, which can be subsequently identified via visual inspection or a library search.

In recent years, considerable attention has been focused on developing analytical instrumentation and methodology for the rapid determination of volatile and semivolatile organic compounds in air, soil, and water samples. One method that has received attention is direct sampling mass spectrometry (DSMS), due to fast sampling times and detection limits in the mid part-per-trillion to low part-per-billion range.[1] DSMS involves introducing the analytes from a sample directly into a mass spectrometer, without preliminary chromatographic separation of the analytes. Typically, the analytes are introduced into the mass spectrometer via a capillary restrictor or membrane interface (i.e., membrane-introduction mass spectrometry), and sample preparation is minimal or not required prior to analysis. This characteristic, when combined with the elimination of long chromatographic separations, results in sample analysis times of near-real time to a few minutes.

The capillary restrictor interface allows for direct atmospheric sampling by letting the vacuum of the mass spectrometer directly pull air samples into the mass spectrometer for analysis. The transport time through a 25-cm by 50−100-μm i.d. capillary restrictor is on the order of 100 ms, which allows for nearly instantaneous detection of analytes in air samples.[1] Extraction of analytes from water and soil matrixes is achieved by using traditional techniques such as needle sparging, aerosol stripping, thermal desorption, and purge-and-trap (P&T),[2] with the extracted analytes directed into the capillary restrictor for analysis.[3] Several interesting applications of DSMS with capillary restrictors have been used for environmental applications in recent years. Volatile organic compounds (VOCs) and semi-volatile organic compounds (SVOCs) have been monitored in flue gas by using direct atmospheric sampling and sorbent-trap/thermal-desorption DSMS.[4] Other researchers have demonstrated in situ sparging of VOCs from groundwater samples using a site characterization and analysis penetrometer system (SCAPS) in which the sparged analytes are transported to the surface for analysis by an ion-trap mass spectrometer.[5] Another SCAPS system was used for in situ thermal desorption of VOCs from soil over 55 feet below the surface, with the desorbed analytes being trapped above the surface prior to analysis by thermal-desorption DSMS.[2]

Membrane inlets for DSMS, more commonly referred to as membrane introduction mass spectrometry (MIMS), have received considerable attention in the past decade, due to the simplicity and sensitivity for analyzing VOCs in water and air samples.[6−8] MIMS involves coupling an air or liquid sample stream with a mass spectrometer via a thin sheet or capillary membrane. The membrane acts as an interface between the vacuum of a mass spectrometer and the sample stream, and as the sample stream flows across the membrane, pervaporation causes the dissolved

* Corresponding author. Phone: 202-767-0719. Fax: 202-404-8119. E-mail: callahan@ccsalpha3.nrl.navy.mil.

[†] American University.

[‡] Geo-Centers Inc.

[§] Naval Research Laboratory.

[⊥] Present address: General Electric Company, Corporate Research and Development, Characterization and Environmental Technology Laboratory, 1 Research Circle, Niskayuna, NY 12309.

(1) Wise, M. B.; Thompson, C. V.; Merriweather, R.; Guerin, M. R. *Field Anal. Chem. Technol.* **1997**, *1*, 251−76.

(2) Myers, K. F.; Karn, R. A.; Eng, D. Y.; Hewitt, A. D.; Strong, A. B.; Brannon, J. M. *Field Anal. Chem. Technol.* **1998**, *2*, 163−71.

(3) Wise, M. B.; Guerin, M. R. *Anal. Chem.* **1997**, *69*, 26A−32A.

(4) Hart, K. J.; Dindal, A. B.; Smith, R. R. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 352−60.

(5) Davis, W. M.; Wise, M. B.; Furey, J. S.; Thompson, C. V. *Field Anal. Chem. Technol.* **1998**, *2*, 89−96.

(6) Harland, B. J.; Nicholson, P. J. D.; Gillings, E. *Water Res.* **1987**, *21*, 107−13.

(7) Bier, M. E.; Cooks, R. G. *Anal. Chem.* **1987**, *59*, 597−601.

(8) Ketola, R. A.; Ojala, M.; Sorsa, H.; Kotiaho, T.; Kostiainen, R. K. *Anal. Chim. Acta* **1997**, *349*, 359−65.

10.1021/ac000690p
Published on Web 12/19/2000

analytes to selectively diffuse through the membrane into the mass spectrometer.[9] Pervaporation of the analytes into the mass spectrometer occurs quickly, and detection limits in the part-per-billion, part-per-trillion,[10] and part-per-quadrillion[11] range are obtainable. Some applications of MIMS appearing in the literature include using an advanced mobile analytical laboratory (AMAL) for on-site analysis of BTEX in groundwater,[12] automated analysis of trihalomethanes and other VOCs in water,[13] on-line monitoring of the reactions of epichlorohydrin in water by using MIMS equipped with a liquid membrane,[14] and on-line chemical process monitoring of a photolysis reaction.[15] The use of DSMS, with capillary and membrane interfaces, has received substantial attention for chemical process monitoring applications, and two informative reviews have recently been published.[16,17]

The primary disadvantage in using direct sampling mass spectrometric methods results when analyzing sample mixtures in which the analytes have similar mass spectra. Because there is no preliminary chromatographic separation and the analytes are introduced nearly simultaneously, the resulting mass spectral data is the sum of the mass spectra of each ionizable analyte in the mixture. The most common quantitation method for DSMS analysis is univariate least-squares regression (uLS) of the response of individual mass channels or sums of mass channels that are characteristic for the analytes of interest. To prevent quantitation errors due to spectral contributions from species other than the analyte of interest, this requires each of the analytes being quantitated to have a unique fragment ion or molecular ion that is not shared with other analytes in the sample. This requirement for unique ions may be feasible for certain analyte mixtures, but if the DSMS data contains spectrally similar analytes with overlapping mass spectra, the chemical source of the ions will be unclear. This ultimately results in limitations and errors in quantitative data analysis. In one study, Virkki et al. used univariate least-squares regression analysis to quantify benzene (m/z 78), toluene (m/z 92), xylenes (m/z 106) and trichloroethene (m/z 96, 98, 100) during on-site MIMS analysis of groundwater.[12] One limitation to this approach is that high concentrations of xylenes with respect to toluene will interfere with the quantitation of toluene due to the isotopic contribution at m/z 92 from the primary fragment ion of xylenes (m/z 91). Another example of a problem associated with using uLS is demonstrated by the in situ thermal desorption of VOCs by Myers et al.[2] In this paper, the authors used m/z 91 as the quantitation ion for toluene, ethyl benzene, and xylenes, which resulted in the inability to differentiate among the three analytes. During quantitation, the analytes were grouped together and the reported concentration was a total of the three.

One approach to improved quantitation during DSMS analysis is to use the full mass spectrum and multivariate data analysis methods to exploit the subtle differences in the analyte spectra. Chemometric methods, specifically multivariate calibration, have been used extensively for quantifying mixtures of spectrally similar analytes in optical spectroscopy[18,19] and is also being used for chemical sensor-based applications.[20] With respect to direct-inlet mass spectrometric techniques, most applications of multivariate calibration have been employed for the analysis of pyrolysis mass spectrometry data.[21−23] One example of a multivariate calibration methodology for DSMS has been demonstrated by Ketola et al., who recently developed an approach called nonlinear asymmetric error function-based least mean square (NALMS) for the MIMS analysis of multicomponent mixtures.[24] The NALMS approach assumes the intensity of a m/z channel is a linear function of the concentration of analytes contributing to that particular m/z, and it is a modification of the general deconvolution method. The NALMS algorithm has been successfully applied to several practical applications, including a completely automated system for the analysis of industrial wastewater via MIMS.[25]

A more popular multivariate regression technique, partial least-squares regression (PLS),[26] was applied by Ohorodnik et al. to analyze a mixture of benzene, toluene, ethyl benzene, and *p*-xylene via MIMS.[27] Unlike the Myers study in which toluene, ethyl benzene, and xylenes could not be quantitated separately using uLS of m/z 91,[2] PLS data analysis of the full spectral data allowed for individual quantification of each of the analytes with good precision. The researchers also demonstrated that PLS regression was superior to univariate regression analysis of a mixture containing the isomers ethyl benzene and *p*-xylene, which have nearly identical spectra. The PLS regression algorithm has recently been incorporated into the data analysis software of a commercially available MIMS based instrument (MS-200, Kore Technology, Cambridge, U.K.)[28] and a headspace-based DSMS instrument (HP 4440A chemical sensor, Agilent Technologies, Palo Alto, CA).

Despite the advantages offered by PLS, there are limitations to its application. In cases in which a complex mixture contains multiple analytes with essentially identical spectra, PLS may be subject to problems during quantitation. One method for decreasing these problems is to add another level of dimensionality to the data. For example, by exploiting differences in the time

(9) Srinivasan, N.; Johnson, R. C.; Kasthurikrishnan, N.; Wong, P.; Cooks, R. G. *Anal. Chim. Acta* **1997**, *350*, 257−71.
(10) Bauer, S.; Solyom, D. *Anal. Chem.* **1994**, *66*, 4422−31.
(11) Soni, M.; Bauer, S.; Amy, J. W.; Wong, P.; Cooks, R. G. *Anal. Chem.* **1995**, *67*, 1409−12.
(12) Virkki, V. T.; Ketola, R. A.; Ojala, M.; Kotiaho, T.; Komppa, V.; Grove, A.; Facchetti, S. *Anal. Chem.* **1995**, *67*, 1421−25.
(13) Lopez-Avila, V.; Benedicto, J.; Prest, H.; Bauer, S. *Am. Lab.* **1999**, *31*, 32−37.
(14) Johnson, R. C.; Koch, K.; Cooks, R. G. *Ind. Eng. Chem. Res.* **1999**, *38*, 343−51.
(15) Wong, P.; Srinivasan, N.; Kasthurikrishnan, N.; Cooks, R. G.; Pincock, J. A.; Grossert, J. S. *J. Org. Chem.* **1996**, *61*, 6627−32.
(16) Cook, K. D.; Bennett, K. H.; Haddix, M. L. *Ind. Eng. Chem. Res.* **1999**, *38*, 1192−204.
(17) Workman, J.; Veltkamp, D. J.; Doherty, S.; Anderson, B. B.; Creasy, K. E.; Koch, M.; Tatera, J. F.; Robinson, A. L.; Bond, L.; Burgess, L. W.; Bokerman, G. N.; Ullman, A. H.; Darsey, G. P.; Mozayeni, F.; Bamberger, J. A.; Greenwood, M. S. *Anal. Chem.* **1999**, *71*, 121R-80R.
(18) Brown, S. D.; Sum, S. T.; Despagne, F.; Lavine, B. K. *Anal. Chem.* **1996**, *68*, 21R-61R.
(19) Lavine, B. K. *Anal. Chem.* **1998**, *70*, 209R-28R.
(20) Janata, J.; Josowicz, M.; Vanysek, P.; DeVaney, D. M. *Anal. Chem.* **1998**, *70*, 179R-208R.
(21) Goodacre, R.; Neal, M. J.; Kell, D. B. *Anal. Chem.* **1994**, *66*, 1070−85.
(22) Tas, A. C.; van der Greef, J. *Mass Spectrom. Rev.* **1994**, *13*, 155−81.
(23) Goodacre, R. *Appl. Spectrosc.* **1997**, *51*, 1144−53.
(24) Ketola, R. A.; Ojala, M.; Komppa, V.; Kotiaho, T.; Juujärvi, J.; Heikkonen, J. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 654−62.
(25) Kotiaho T, Kostiainen R.; Ketola, R. A.; Mansikka, T.; Mattila, I.; Komppa, V.; Honkanen, T.; Wickstrom, K.; Waldvogel, J.; Pilvio, O. *Proc. Control Qual.* **1998**, *11*, 71−78.
(26) Martens, H.; Næs, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.
(27) Ohorodnik, S. K.; Shaffer, R. E.; Callahan, J. H.; Rose-Pehrsson, S. L. *Anal. Chem.* **1997**, *69*, 4721−27.
(28) White, A. J.; Blamire, M. G.; Corlett, C. A.; Griffiths, B. W.; Martin, D. M.; Spencer, S. B.; Mullock, S. J. *Rev. Sci. Instr.* **1998**, *69*, 565−71.

response of the analytes, the subtle variations in the spectra can be enhanced. Some direct sampling methods, such as MIMS and purge-and-trap MS (P&T/MS), can be configured so that slight temporal differences exist between the analytes entering the instrument. For example, in MIMS, slight differences in pervaporation are caused by chemical and physical interactions between the analytes and membrane material. Overney and Enke developed a data analysis strategy that utilizes the time-dependent pervaporation of analytes in MIMS.[29] The method is based on the periodic change between exposing sample and background solutions to the membrane, and a custom software program was developed that used linear equations to resolve the analytes in the sample on the basis of differences in pervaporation. Although the differences in time profiles are usually subtle, a dramatic example of differences in pervaporation through zeolite membranes has recently been presented.[30] Time differences in the P&T/MS temporal profiles can be attributed to differences in boiling points and physical interactions with the trapping media and transfer lines.

A potentially powerful approach to analyzing time-dependent DSMS data is through the use of multiway or second-order methods, such as trilinear PLS[31] (tri-PLS) and parallel factor analysis (PARAFAC).[32,33] Second-order instrumentation produces a matrix of data per sample. Used in this context, *order* refers to its tensor mathematical form (e.g., a scalar value is a zero-order tensor, a vector is a first-order tensor, and a matrix is a second-order tensor, etc.). Second-order chemometric methods are designed specifically to exploit the data generated by second-order instrumentation. For example, a sample data matrix can be constructed using the mass spectrum versus retention time during GC/MS and LC/MS analysis. This differs from first-order methods, such as PLS and NALMS, which use data vectors (i.e., a single mass spectrum per sample) at a specific retention time for quantitation. The application of second-order calibration methods to DSMS data is directly analogous to their application to the deconvolution and quantitation of coeluted analytes in hyphenated chromatography methods, for example, high performance liquid chromatography with UV/vis diode array detection (HPLC/DAD). Improved quantification can be observed by using second-order methods, even in the case of crude or poor chromatographic separation. To date, no applications of second-order methods to DSMS have been reported in the literature.

For DSMS to mature as an analytical method, its use in routine on-line, at-line, or automated scenarios is necessary. In these types of applications, the role of the expert analyst is minimized and most of the data processing steps are performed without user intervention. One often-overlooked aspect of blindly applying chemometric methods for complex samples (e.g., environmental samples) is the treatment of unanticipated or unmodeled interferences. This topic has been covered in great detail elsewhere,[32] with a complete discussion of the mathematics behind zero-, first-, and second-order calibration, but will be covered here briefly for completeness. Univariate approaches are biased by the presence of interferences, unless the m/z channel is unique to the analyte, which is rare for many applications. For autonomous use, univariate approaches can only be used in well-controlled environments. First-order chemometric methods such as PLS can handle interferences provided they are known in advance and incorporated into model development. First-order methods can be used to detect any unanticipated interference but quantification is often biased. Autonomous use of PLS models is done routinely in chemical process monitoring applications, but care must be exercised for complex environments. NALMS has some robustness against unexpected interferences, because they will be accumulated in the error term of the fitting function. Second-order chemometric algorithms such as PARAFAC have the potential to eliminate this worry completely, because the occurrence of interferences that are not included in the calibration set does not bias the measurement if certain mathematical requirements are met. The ability of second-order chemometric methods to quantify in the presence of unexpected interferences is commonly referred to as the "second-order advantage."[32]

In the work presented here, purge-and-trap/direct inlet mass spectrometry was used to analyze mixtures of toluene, ethyl benzene, and cumene in water. These three analytes were chosen for use in the test mixtures due to their overlapping mass spectra and the subtle differences in temporal profiles when analyzed by P&T/MS. The overlapping mass spectra of the analytes is desired to evaluate the performance of univariate and multivariate quantitation methods, and the difference in temporal profiles will allow for the evaluation of second-order calibration methods. The goal of this study is to further our understanding of how differing calibration approaches affect the operational performance of DSMS methods for multicomponent analysis. Four data analysis algorithms (uLS, PLS, tri-PLS, and PARAFAC) were used to quantify the three analytes in calibration and predictions samples. Critical assessments of each algorithm will be based on (1) prediction performance, and (2) the ability to handle prediction samples that also contain previously unmodeled contaminants (e.g., unknown interferences). The calibration approaches studied in this work cover a broad range of complexity: zero-order (uLS), first-order (PLS), and second-order (tri-PLS and PARAFAC). The ability of the second-order PARAFAC and tri-PLS approaches to exploit the subtle temporal differences found in the P&T/MS data will be critically evaluated and compared to the simplicity of the uLS and bilinear PLS approaches. In addition to the predictive ability of each approach, the use of PARAFAC for obtaining pure component spectra of analytes and interferences will be discussed.

EXPERIMENTAL SECTION

Reagents. Stock standard solutions were prepared by diluting toluene (Fisher Scientific, Fair Lawn, NJ), ethyl benzene, and cumene with methanol (Aldrich, Milwaukee, WI). The stock standards were stored at −10 °C prior to use. Immediately prior to analysis by purge and trap, duplicate water samples were prepared by diluting aliquots of the stock standards to 25 mL with 18 MΩ water from a MilliQ+ system (Millipore, Bedford, MA) in a volumetric flask. Two 5-mL aliquots of each sample were transferred to 5-mL gas-tight syringes for injection into the purge-and-trap instrument.

Mass Spectrometer. The experiments were performed on a Finnigan GCQ ion-trap mass spectrometer (Finnigan Corp., San

(29) Overney, F. L.; Enke, C. G. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 93−100.
(30) Bennett, K. H.; Cook, K. D.; Falconer, J. L.; Noble, R. D. *Anal. Chem.* **1999**, *71*, 1016−20.
(31) Bro, R. *J. Chemometrics* **1996**, *10*, 47−61.
(32) Booksh, K. S.; Kowalski, B. R. *Anal. Chem.* **1994**, *66*, 782A-91A.
(33) Smilde, A. K. *Chemom. Intell. Lab. Syst.* **1992**, *5*, 143−57.

Jose, CA) operated in electron impact (70 eV) and full-scan mode (m/z 46−150). The GCQ had been previously modified by replacing the 100 L/s diffusion pump with a 200 L/s turbomolecular pump. The GCQ was controlled using Custom Tune instrument control software (version 1.0 beta, Build 8; Finnigan Corp, Austin, TX), which allows the user to have manual control over all instrumental parameters. Automatic gain control (AGC) was turned off, and the ion injection time was set to 10.0 milliseconds. The scan speed was approximately 1 scan per second. The ion source and GCQ transfer line were held at 200 °C. A fused-silica capillary restrictor (30 cm × 75 μm i.d.) was used to interface the P&T instrument with the mass spectrometer in order to limit helium flow into the MS and to maintain sufficient vacuum ($\sim 10^{-5}$ Torr). The capillary restrictor was installed in the transfer line of the MS in the same manner as one would install a GC column, which resulted in approximately 5 cm of the restrictor extending out of the MS transfer line. The restrictor was coupled to the transfer line of the P&T instrument via a $^1/_{16}$-in. stainless steel union. To heat the portion of the restrictor which extended out of the MS transfer line, it was inserted into an approximately 4-in. section of 1-in. i.d. copper tubing, which was wrapped in heating tape then aluminum foil. A rheostat was used to adjust the temperature inside the copper tubing to 150 °C, as measured by a thermocouple.

Purge and Trap. All purge-and-trap experiments were performed using an OI Analytical (College Station, TX) model 4560 instrument with a frit sparger. The trap was OI model 9, which contains sections of Tenax, silica gel, and charcoal. Because the pressure of helium into the P&T also controls the flow through the capillary restrictor, it was determined that a head pressure of 30 psi resulted in an estimated flow of 3−4 mL/min into the MS. The helium purge pressure was set to 20 psi, and the transfer line and valve body were maintained at 150 °C. Water samples (5 mL) were analyzed using the following program: (1) sample purged for 10 min with helium, (2) desorb trap at 190 °C for 4 min, and (3) trap baked at 200 °C for 5 min. The scan number in which the sampling valve of the P&T was actuated (at the beginning of trap desorption) was noted to record the injection time of each sample into the mass spectrometer. The water management feature of the 4560 P&T was used for all analysis to reduce the amount of water introduced into the mass spectrometer. The 4560 P&T allows the user to program the instrument to preheat the trap to the desorb temperature prior to the switching valve's directing the sample into the mass spectrometer. This feature was not used in order to maximize the time separation of the analytes during desorption from the trap, due to differences in boiling point. The sparging vessel was flushed with 6 mL of 18 MΩ water between samples.

Data Analysis. The mass spectrometer data was translated from binary to ASCII format prior to use for data analysis. All data analysis was performed using routines written in MATLAB, version 5.3 (Mathworks, Inc., Natick, MA). The MATLAB routines used for PLS regression and PARAFAC were provided in the PLS_Toolbox, version 2.0.1c (Eigenvector Technologies, Inc., Manson, WA). The MATLAB functions for performing tri-PLS were adapted from the multiway toolbox from Rasmus Bro.[29] Mass spectral library searches were performed using the NIST Mass Spectral Search program, version 1.6d.

Table 1. Calibration and Prediction Sets

	Concentration (ppb)		
calibration sample	toluene	ethyl benzene	cumene
1	0.863	2.000	0.858
2	1.500	0.200	0.199
3	2.000	0.863	0.858
4	0.863	0.863	1.990
5	0.863	0.863	0.000
6	1.500	1.500	0.199
7	0.200	0.200	0.199
8	0.863	0.000	0.858
9	1.500	1.500	1.490
10	1.500	0.200	1.490
11	0.200	0.200	1.490
12	0.200	1.500	1.490
13	0.000	0.863	0.858
14	0.863	0.863	0.858
15	0.200	1.500	0.199
prediction sample			
1	0.863	0.000	0.000
2	0.000	0.000	0.000
3	0.000	0.863	0.000
4	0.000	0.000	0.858
5	1.010	0.604	0.399
6[a]	1.010	0.604	0.399
7[b]	0.863	0.863	0.858
8[c]	1.200	1.200	1.190
9	0.401	2.000	0.980

[a] Also contains 1.775 ppb chloroform and 1.17 ppb pyridine. [b] Also contains 2.4 ppb benzene. [c] Also contains 0.6 ppb benzene.

DATA ANALYSIS METHODS

Mixtures with varying concentrations (0.2−2.0 parts-per-billion) of toluene, ethyl benzene, and cumene were analyzed by P&T/direct sampling mass spectrometry, and calibration and prediction data sets were constructed. The concentrations of the three analytes in the calibration samples were determined using a central composite design with 15 samples and 5 calibration levels. Nine samples were analyzed to construct the prediction set. Three of the nine prediction samples had contaminants added in order to explore the performance of each quantitation algorithm when the sample contains species that were not present during model building. The contaminating species were pyridine, chloroform, and benzene. Table 1 lists the concentrations of each analyte in the calibration and prediction samples. Duplicate analysis was performed for each sample, except for sample 2 in the prediction set, which is a blank.

Univariate Least-Squares Regression. The instrument response for the molecular ion of each analyte, that is, m/z 92 for toluene, m/z 106 for ethyl benzene and m/z 120 for cumene, were regressed against the known concentrations in the calibration set using the simple least-squares regression model. The response for each sample was determined by averaging 16 scans, starting 100 scans (approximately 100 seconds) after injection, which covered the maximum spectral response in the time profile. Using an average of multiple scans for quantitation was determined to be necessary to reduce the noise, which is a characteristic of the mass spectrometer that was used for the experiments.

To evaluate the performance of the regression models for calibration and prediction, the root-mean-squared error in calibra-

tion (RMSEC) and the root-mean-squared error in prediction (RMSEP) were calculated for each analyte. The root-mean-square error calculation is given by

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{c}_i - c_i)^2}{n}} \quad (1)$$

where n is the number of samples, c_i is the actual concentration in sample i, and $\hat{c}_i$ is the predicted concentration. The samples in the calibration set were used to calculate the RMSEC, and the samples in the prediction set were used to calculate the RMSEP.

Partial Least Squares Regression. The average of 16 spectra was used for each sample, as in the uLS analysis, and the full mass spectrum of each sample was used to construct an $n \times 105$ data matrix, where 105 is the number of ions in the mass spectra from m/z 46 to 150. The resulting data matrix was mean-centered before performing multivariate calibration. The principles of (bilinear) PLS have been described in detail elsewhere,[26] and will not be discussed here. A single PLS-1 model was built for each analyte. The optimum number of PLS factors (maximum of 14) to use for each calibration model was determined by performing leave-one-out cross-validation and finding the number of PLS factors that minimized the root-mean-square error of cross validation (RMSECV).[34] RMSECV was calculated using eq 1, in which i is the sample removed from the calibration set and predicted using a model constructed with the remaining n-1 samples.

Trilinear Partial Least-Squares Regression. To introduce the time dimension into the data analysis, a data matrix for each sample was constructed by extracting 200 mass spectra starting 40 scans after the start of trap desorption. The data matrix for each sample was combined to form a data cube with the dimensions of $200 \times 105 \times n$. Before analysis using the tri-PLS algorithm, the data cube was processed by unfolding on the third dimension (sample), then it was mean-centered. The tri-PLS algorithm is an extension of the PLS algorithm from first- to second-order data, and a detailed description of the algorithm is given by Bro.[31] Like the bilinear PLS algorithm, tri-PLS involves data reduction followed by construction of a regression model. The data reduction step seeks to find a set of triads that maximizes the covariance between the analyte concentrations and the time-dependent mass spectral data. The optimum number of factors (maximum of 14) for each calibration model was determined using leave-one-out cross-validation.

Parallel Factor Analysis. The $200 \times 105 \times n$ data cube constructed for the tri-PLS analysis was used for analysis by PARAFAC. Preprocessing of the data cube was not performed in this experiment. PARAFAC uses iterative procedures for the decomposition of higher order data, in that an $I \times J \times K$ data cube, x, is decomposed via the model[35,36]

$$x_{i,j,k} = \sum_{n=1}^{N} a_{i,n} b_{j,n} c_{k,n} + e_{i,j,k} \quad (2)$$

where a, b, and c are the estimated loading matrixes for factor n, in time, mass spectra, and relative concentration, respectively. The sum of squares of the residuals in the calibration model is given by e. Alternating least squares is used to optimize the model by adjusting the loading matrixes until e is minimized. Both PARAFAC and tri-PLS utilize the trilinear model but differ in the optimization criteria. Tri-PLS builds the optimal calibration model by finding covariance between the data cube and the known analyte concentrations, while PARAFAC reduces the error in the fit of the experimental data.

Each factor in a PARAFAC model represents a distinct component (e.g., analyte, background, interference) that contributes to the spectral information contained in the data cube, and the number of factors, N, must be chosen correctly for optimum performance of the model.[35] In this work, the number of factors, N, was varied from 4 to 8, and separate PARAFAC models were constructed at each N. The PARAFAC model that resulted in the best estimation of the mass spectra and time profiles for the analytes (by visual inspection) and minimum RMSEC values was chosen as the final model.

The most salient property of the PARAFAC model is that it results in a unique solution for certain types of second-order analytical data such as GC/MS, HPLC/DAD, and fluorescence EEM. PARAFAC models can produce pure component profiles (a and b in eq 2) and can estimate analyte concentrations even in the presence of unknown interferences. These benefits arise from the rank additivity and bilinearity of second-order analytical data. The addition of each new component (analyte or interfering agent) in the prediction sample simply increases the number of necessary components (N in eq 2) by one (i.e., increases the rank of the data cube by one) and does not change the pure component profiles for the other components. Thus, the PARAFAC model can handle several unanticipated interferences in a given sample, depending on the noise level of the measurement and the degree of overlap between analytes and interferences. A complete mathematical derivation of the second-order advantage of PARAFAC and related methods are given elsewhere.[32,37−39]

Prediction of the analyte concentration of new samples is done by augmenting the data cube of calibration samples with the sample to be predicted. Each time a new prediction sample is acquired, a new data cube is created by removing the previous prediction sample and replacing it with the current prediction sample (i.e., the calibration samples in the data cube are constant). The PARAFAC algorithm simultaneously models both the calibration and prediction samples. This is in sharp contrast to the other calibration methods, in which a mathematical model is developed once using only the calibration set and then subsequently applied to all future samples. As shown in eq 2, PARAFAC is an unsupervised quantitation method in that the known concentrations of the analytes in the calibration samples are not used during PARAFAC model optimization. Once the final solution is found

(34) Wise, B. M.; Gallagher, N. B. *PLS_Toolbox Chemometrics Tutorial, version 2.0*; Eigenvector Research: Manson, WA, 1998.
(35) Booksh, K. S.; Muroski, A. R.; Myrick, M. L. *Anal. Chem.* **1996**, *68*, 3539−44.
(36) Bro, R. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 149−71.
(37) Sanchez, E.; Kowalski, B. R. *J. Chemom.* **1990**, *4*, 29−45.
(38) Sanchez, E.; Kowalski, B. R. *J. Chemom.* **1988**, *2*, 265−80.
(39) Bro, R. "Multi-Way Analysis in the Food Industry: Models, Algorithms, and Applications", Doctoral Thesis, University of Amsterdam, Denmark, 1998.

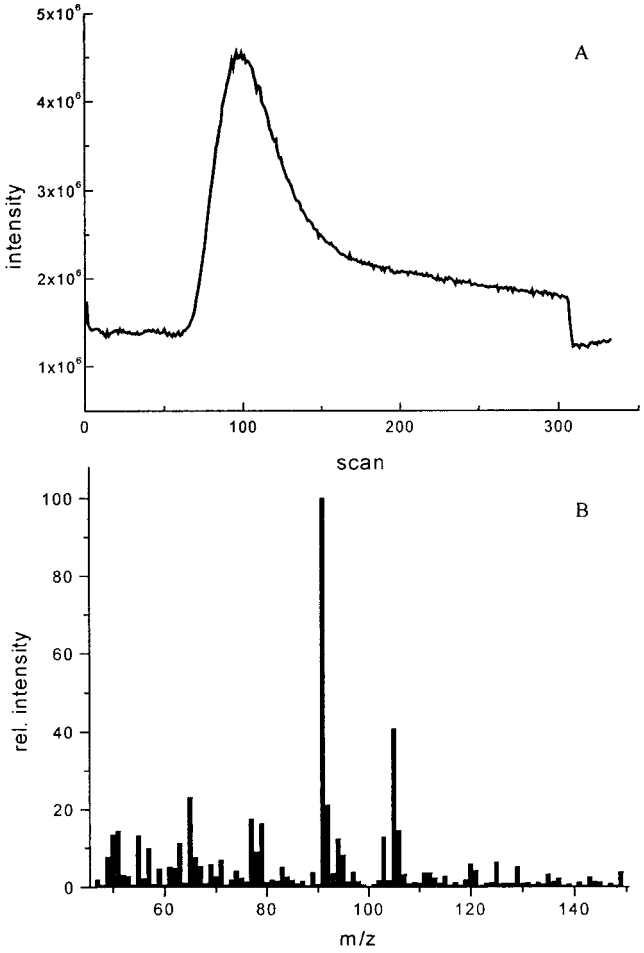

Figure 1. Total ion current profile during P&T/MS analysis (A) and a mass spectrum of 0.86 ppb mixture of toluene, ethyl benzene, and cumene (B). Scan zero in plot A corresponds to the beginning of desorption and actuating of the valve in the P&T.

(i.e., the model has been optimized), the estimated concentrations for each sample (c in eq 2) are relative to the other samples, and scaling must be performed to determine actual concentrations. This scaling is performed through the use of a simple linear regression model that is built using the known analyte concentrations from the calibration samples in the data cube and the estimated relative analyte concentrations of the calibration samples from the PARAFAC model. The estimated concentrations of the analytes in the prediction sample can then be predicted using this model.

RESULTS AND DISCUSSION

A typical total ion current profile and mass spectrum resulting from the P&T/MS analysis of a three-component mixture is given in Figure 1. The tailing of the peak in the total ion current profile is not due to tailing of the three analytes, but it is actually due to a background component that begins to elute from the P&T around scan 90. The mass spectrum of the late-eluting background component contained mainly m/z 95, and the chemical source of the background was not investigated. In looking at the mixture mass spectrum in Figure 1b and the pure component mass spectra in Figure 2a−c, the difficulty in analyzing this three-component mixture becomes clear. The spectral complexity is due to the overlapping mass spectra of the analytes, for example, many spectral channels have a response that results from two or three of the analytes. Due to the structural similarity between toluene and ethyl benzene, it is not surprising that the mass spectra for the two analytes are highly overlapping, excluding the stronger intensity of m/z 92 for toluene and the spectral signals between m/z 102 and 107 for ethyl benzene. With respect to ethyl benzene and cumene, both analytes have overlapping spectra in the ranges of m/z 102−106 and m/z 77−79. A fragment ion at m/z 91 and a cluster of ions centered at m/z 50 are common to all three analytes.

Table 2. Prediction Results for Different Quantitative Analysis Techniques

analyte	measure	uLS	PLS	tri-PLS	PARAFAC
toluene	RMSEP[a]	0.216	0.645	0.298	0.112
	RMSEP[b]	0.226	0.121	0.110	0.113
ethyl benzene	RMSEP[a]	0.342	0.180	0.358	0.239
	RMSEP[b]	0.364	0.079	0.104	0.174
cumene	RMSEP[a]	0.106	0.189	0.103	0.102
	RMSEP[b]	0.094	0.112	0.097	0.078

[a] Root-mean-square error of prediction (ppb) using all 9 prediction samples. [b] Root-mean-square error of prediction (ppb) excluding prediction samples 7 and 8.

Due to the overlapping mass spectra, the chemical source of many ions will be unclear when analyzed with univariate techniques. In performing uLS (zero-order calibration) analysis using the molecular ions for the three analytes, overlapping spectral contributions exist for toluene and ethyl benzene. The molecular ion for toluene (m/z 92) overlaps with the isotopic ^{13}C contributions from the m/z 91 fragment ions of ethyl benzene and cumene, and the molecular ion for ethyl benzene (m/z 106) overlaps with the isotopic contribution from the m/z 105 fragment ion of cumene. Of course, the molecular ion for cumene (m/z 120) does not overlap with the other analytes. Therefore, univariate least-squares analysis using molecular ions is expected to perform well for cumene and poorly for toluene and ethyl benzene. The results for uLS analysis of the prediction data set are given in Table 2. In comparing the results for toluene and ethyl benzene to the results for cumene, it is obvious that the overlapping isotopic peaks at m/z 92 and m/z 106 have a negative effect on the univariate quantitation of toluene and ethyl benzene. The approximately 2-fold and 3-fold higher RMSEP errors for toluene and ethyl benzene, respectively, clearly demonstrate that univariate least-squares regression is not an ideal quantitation method when the mixture contains analytes with overlapping mass spectra.

To reduce the negative effects of overlapping mass spectra during quantitation, multivariate data analysis techniques are necessary. PLS regression using the full mass spectrum data was used to analyze the calibration and prediction data sets, and the results are given in Table 2. The number of PLS factors used for each of the calibration models were 5, 4, and 5 for toluene, ethyl benzene and cumene, respectively. The RMSECV statistics for the analytes were 0.075, 0.060, and 0.057 ppb for toluene, ethyl benzene, and cumene, respectively.

In evaluating the RMSEP statistics for PLS calibration, at first one may think that univariate regression does a better job of sample prediction for toluene and cumene. Close investigation into

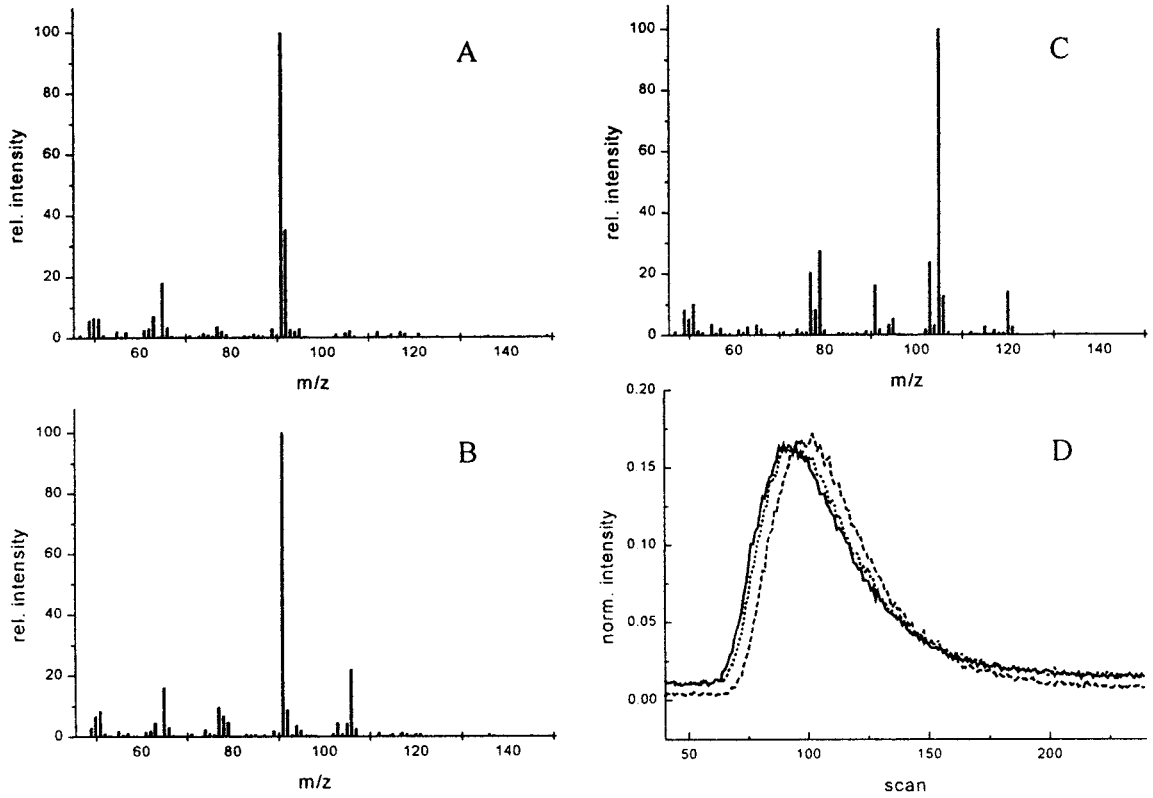

Figure 2. Normalized, background-subtracted P&T/MS mass spectra of 1.5 ppb toluene (A), ethyl benzene (B), and cumene (C). Plot D illustrates the time profile differences between the three analytes. The solid line is toluene (*m*/*z* 91), the dotted line is ethyl benzene (*m*/*z* 91), and the dashed line is cumene (*m*/*z* 105).

the prediction results for each sample clarifies the source of the problem with multivariate PLS analysis. Two of the samples (7 and 8) in the prediction set contain benzene, in addition to the three analytes of interest, which was not included in model building. The poor toluene and cumene prediction results for these two samples caused the poor overall RMSEP statistics. Because benzene does not overlap with the molecular ions of the calibrated analytes, uLS was unaffected by its presence in the prediction samples. On the contrary, the multivariate nature of PLS led to the poor results in prediction for samples 7 and 8.

The rationale for these results can be found by close examination of the regression coefficients and latent vectors for the PLS models. The first loading vector for each model is nearly identical to the mass spectrum of the analyte being calibrated. The remaining loading vectors are very similar to difference spectra between the calibrated analytes and the other analytes in the mixture. Simply stated, the first PLS factor models the analyte being calibrated and the remaining factors subtract the spectral contribution from the other analytes and background. In looking at the PLS regression vectors, the PLS model for toluene has strongly negative regression coefficients at *m*/*z* 78 and 77 which overlap with the molecular and fragment ions for benzene. The probable cause of these regression coefficients is the modeling of minor features in the spectra for ethyl benzene and cumene. The spectral response for benzene in samples 7 and 8 and the negative regression coefficients led to biased concentration estimates. The predicted concentrations for toluene in duplicate analysis of samples 7 and 8 were −0.991, −0.929, 0.932, and 0.817, respectively. The dramatic effect of the negative regression coefficient is apparent in sample 7, which contained a high concentration of benzene (2.4 ppb) relative to the three analytes of interest (0.86 ppb). The low concentration of benzene in sample 8 (0.6 ppb) relative to the three analytes of interest (1.2 ppb) results in a less pronounced reduction in predicted concentration.

A similar trend in regression coefficients is observed for the cumene model, except that the regression coefficients for *m*/*z* 77 and 78 are smaller than for toluene. The prediction results for the ethyl benzene model in these two samples were overestimated, which was caused by positive regression coefficients for *m*/*z* 77 and 78. To determine the influence of samples 7 and 8 on the RMSEP statistic, the two samples were removed from the prediction set, and the RMSEP statistics were recalculated. The recalculated RMSEP results show that the overall prediction of toluene and ethyl benzene in the remaining samples was improved over uLS. The recalculated RMSEP result for cumene was only slightly higher than what was obtained with uLS.

It is interesting to note that prediction sample 6 also contained chloroform and pyridine, and the PLS prediction results were not biased by their presence. In looking at the regression vectors for the three analytes, none of the analytes had significant regression coefficients at *m*/*z* 83/85 and *m*/*z* 79, the primary ions for chloroform and pyridine, respectively. Thus, the presence of chloroform and pyridine in sample 6 did not adversely affect the prediction result.

To overcome the prediction bias caused by interfering species, the analyst typically recalibrates the system using calibration samples that also contain the interfering species. In doing so, the spectral contributions from the interfering species are modeled

by the PLS factors. This recalibration process can be time-consuming and requires correct identification of the interfering species, which may be difficult.[26] However, this is an area of ongoing research interest, and recently developed methods may make this process easier.[40] It should also be noted that recalibration is possible only if the analyst knows that the prediction samples contain unmodeled interferences. One procedure for spotting potentially difficult prediction samples is to utilize statistical outlier detection strategies such as the T^2 and Q residual statistics, which are commonly found in commercial chemometric software packages.[41] The T^2 statistic is a measure of the distance from a sample and the multivariate mean of the calibration samples. The Q residual is a measure of the variation of a sample outside the data space defined by the PLS model. Using 99% confidence intervals for the T^2 and Q residual statistics, the spectra for prediction samples 6, 7, and 8 were identified as outliers for the toluene model. The ethyl benzene PLS model flagged spectra from prediction sample 7 as outliers, while the cumene model determined that the spectra from prediction samples 6 and 7 were outliers. Once a spectrum is flagged as an outlier, the presence of unmodeled interferences can be verified and actions taken. Regardless, the ability to alert the operator to changes in sample composition is a powerful advantage of first-order calibration methods such as PLS, as compared to zero-order calibration algorithms.

Although the use of PLS led to improvements in prediction (excluding prediction samples 7 and 8) in comparison to uLS, further improvements may be possible by increasing the dimensionality of the data. Because it was known that the three analytes of interest have slightly different time profiles in the P&T/MS data, the use of second-order calibration techniques was investigated to determine if improvements in quantitation could be achieved. Figure 2d shows the differences in the time profiles for the three analytes: approximately 1−2 s between toluene and ethyl benzene and 7 s between toluene and cumene. The source of the time separation between the three analytes is believed to be due to differences in desorption times from the trap and migration rates along the transfer line and capillary restrictor. With respect to desorption times from the trap, the differences are due to the boiling points of the analytes (toluene = 110, ethyl benzene = 136, and cumene = 152 °C) and the differences in the physical and chemical interactions between the analytes and the trapping media. As for the migration of the analytes, a limited chromatographic separation likely occurs due to differences in the physical interactions between the analytes and the fused silica interior walls of the transfer line and capillary restrictor.

Tri-PLS models were built using 6, 7, and 4 PLS factors for the toluene, ethyl benzene, and cumene models, respectively. The RMSECV statistics were 0.078, 0.069, and 0.067 ppb for toluene, ethyl benzene, and cumene, respectively. The RMSEP results for the tri-PLS analysis are given in Table 2.

In evaluating the RMSEP values for the tri-PLS predictions, it is apparent that tri-PLS is also subject to increased errors when the prediction samples contain unexpected interfering species, for example, benzene in samples 7 and 8. Similarly to the bilinear PLS algorithm, toluene and ethyl benzene models have significant regression coefficients that overlap with the response from benzene, which leads to the poor prediction performance for samples 7 and 8. The tri-PLS predicted concentrations for toluene in samples 7 and 8 were 0.051, 0.059, 1.141, and 0.991 ppb, respectively. The presence of benzene is less detrimental to prediction performance in sample 8, due to the low benzene concentration with respect to toluene. A similar trend is seen with the ethyl benzene results, except that the ethyl benzene results were overestimated. If the RMSEP values are recalculated removing samples 7 and 8, the results improve dramatically, with prediction errors for all 3 analytes around ∼0.1 ppb, which is acceptable for this application. Similarly to PLS, outlier detection methods may be used with tri-PLS to alert the operator to changes in the sample composition and actions taken to identify the interfering species.

Regarding the improved prediction performance for ethyl benzene using PLS, further examination reveals that caution should be used when attempting to interpret the significance of small differences in SEP. As stated earlier, the optimal model size for PLS and tri-PLS was chosen by using the well-known leave-one-out/cross-validation (LOOCV) method. Further inspection reveals that lower SEP values can be found at slightly smaller model sizes for both PLS and tri-PLS. For example, the minimal SEP value for the ethyl benzene model is 0.08 and 0.07 for PLS and tri-PLS, respectively. Thus, in this case, PLS and tri-PLS perform similarly and the slight (apparent) improvement on the part of PLS is due to the ability of LOOCV to choose a better model size. If a larger calibration set were available, the likelihood of this occurring would decrease dramatically. Inspection of the SEP values of toluene and cumene for PLS and tri-PLS using smaller model sizes reveals a similar trend. Tri-PLS does not appear to provide a significant advantage over PLS for this application.

The slight differences in the P&T/MS time profiles for the three analytes were also utilized for quantitative analysis using PARAFAC. An important feature results from the uniqueness of the PARAFAC model. The model loadings are the estimated time profiles and mass spectra for each component in the mixture. Inspection of the loadings in these dimensions allows for the analyst to identify the time profile and mass spectrum of each analyte without having prior knowledge of their characteristics. In building a model for the 15 calibration samples only, a 6-factor model was determined to provide the best results. The PARAFAC-estimated time profiles for the six components in the mixture is shown in Figure 3. The estimated time profiles for toluene, ethyl benzene, and cumene in Figure 3A match well with the actual time profiles shown in Figure 2d. The evolving background component, mentioned earlier, consists mainly of m/z 95 and is represented by the dashed line in Figure 3b. The solid and dotted lines in Figure 3b represent additional background components in the data. Although the PARAFAC-estimated mass spectra of the three calibrated analytes are not shown, visual comparison of the estimated spectra and the actual mass spectra in Figure 2 revealed that they were nearly identical. To obtain quantitative estimates of the quality of the estimated spectra, a small mass spectral library was constructed in the NIST mass spectral database using the actual mass spectra of the pure components in Figure 2. The PARAFAC-estimated spectra were purity-searched

(40) Haaland, D. M. *Appl. Spectrosc.* **2000**, *54*, 246−54.
(41) Wise, B. M.; Gallagher, N. B. *PLS_Toolbox, 2.0*, ed.; Eigenvector Research Inc: Manson, WA, 1998.

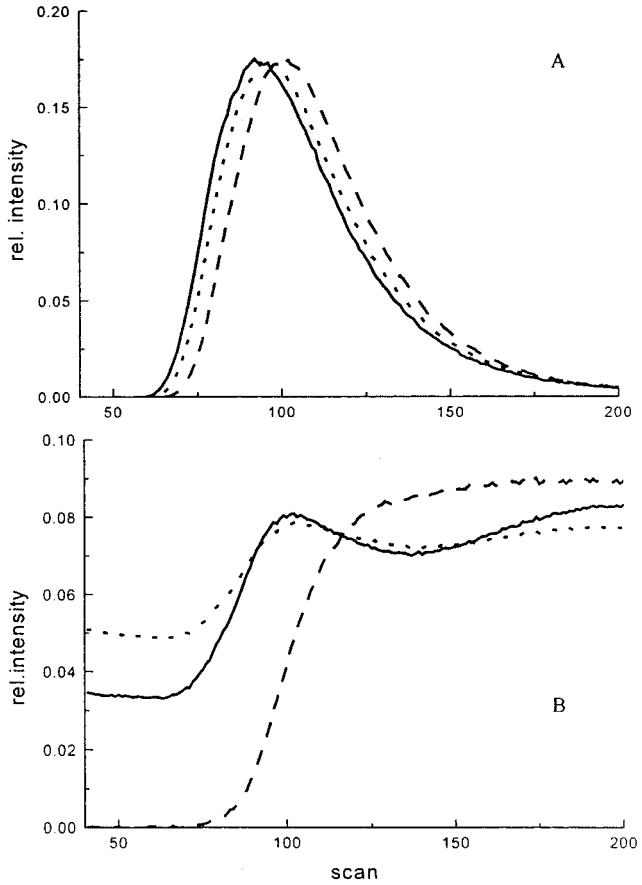

Figure 3. PARAFAC-estimated time profiles for calibration data set. Plot A illustrates the time profiles for the analytes toluene (solid), ethyl benzene (dotted), and cumene (dashed). Plot B illustrates the time profiles for three PARAFAC-determined background components in the P&T/MS data.

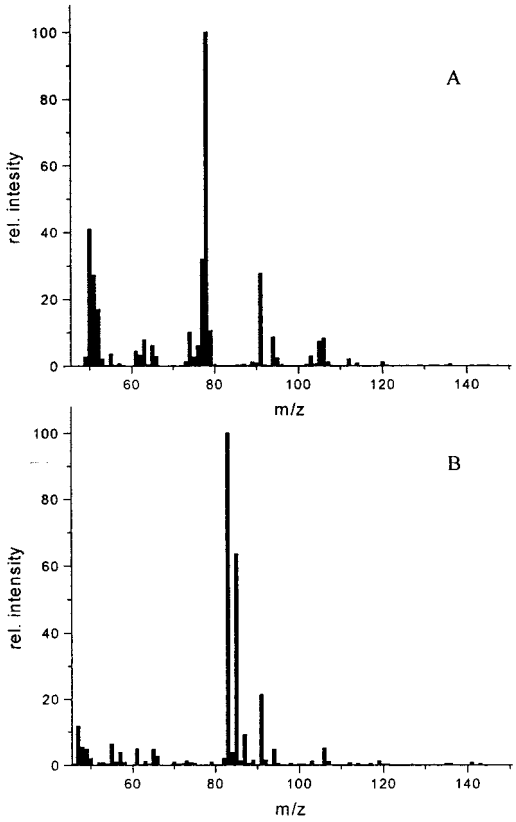

Figure 4. PARAFAC-estimated mass spectra for benzene in prediction sample 7 (A) and chloroform in prediction sample 6 (B).

against the actual mass spectra, and the match factors were 805, 905, and 869 for toluene, ethyl benzene and cumene, respectively (perfect match = 1000). Similar match factors were obtained when the actual and PARAFAC-estimated spectra were searched against the NIST mass spectral library. This demonstrates that the PARAFAC-estimated mass spectra are very useful for analyte identification purposes, and further automation of the model-building and -evaluation process is possible.

In performing quantitative analysis of the prediction samples, nine different PARAFAC models are required (i.e., one model per prediction sample). In looking at the RMSEP statistics of the full prediction set for PARAFAC in Table 2, there is an overall improvement in the prediction performance over the other multivariate methods when the prediction samples contained unmodeled interferences. The better prediction performance of PLS for ethyl benzene is the only exception. The improved prediction performance of PARAFAC over the other multivariate methods is mainly due to the ability of PARAFAC to quantify the prediction samples while interfering species are present (i.e., the second-order advantage). In contrast, when the prediction samples did not contain unmodeled interferences, PLS and tri-PLS provided similar or improved prediction results. The improved performance of PLS and tri-PLS for these cases is most likely due to PARAFAC not using the known analyte concentrations in the standards during model-building.

In building the models for quantitation of the prediction samples, the optimum number of factors was determined to be 6 for samples 1−5 and 9, and 7 factors for samples 6, 7, and 8. It was not surprising that prediction samples 7 and 8 were optimized with 7 factors in the model, because the presence of benzene in these samples required an additional factor. It was somewhat surprising that 8 factors was not the optimum for sample 6, because that prediction sample contained two additional analytes. Inspection of the mass spectral loadings showed that chloroform was successfully modeled as an individual factor, but pyridine was not modeled by any of the factors. Further investigation of the data for prediction sample 6 revealed that the response for pyridine was very small in comparison to the other analytes. The low sensitivity for pyridine was likely due to an active site in the P&T/MS system, which resulted in poor transfer from the P&T to the ion source of the mass spectrometer. The benzene and chloroform interfering species were successfully modeled in samples 6, 7, and 8, and the PARAFAC-estimated mass spectra for the two analytes are shown in Figure 4. Visual inspection of the two spectra reveals that they are in good agreement with known spectra, except for the inclusion of responses at m/z 91, 95, and 102−106. The source of these spectral responses is most likely the other analytes and background components in the data. PARAFAC was unable to obtain better estimates of the spectra because the interfering species were only present in one sample out of 16 (15 calibration samples). The estimated mass spectra for benzene and chloroform were reverse-searched against the NIST mass spectral database and the match factors against the NIST library spectra were 891 (hit 1) and 896 (hit 2), respectively. This demonstrates a powerful feature of PARAFAC in that the quantitation of

prediction samples in the presence of interfering species is possible, and the estimated mass spectral loadings for the interfering species can be used for subsequent interference identification.

CONCLUSIONS

Regardless of which calibration method was employed, the P&T method of DSMS was able to achieve excellent sensitivity (low ppb) for three environmentally important analytes. The time required for analysis was not optimized in this work, so further improvements can be expected. Among the various DSMS approaches, MIMS and P&T afford the capability of creating differences in the temporal profiles of the analytes entering the spectrometer. In studying the four different calibration algorithms, it became clear that each approach has unique advantages and, when used in the proper scenario, will perform as required. The zero-order uLS method must be used in carefully guarded situations in which no interferences overlap with the quantitation ion of the analyte. The first-order PLS regression models provided excellent accuracy as long as the samples being predicted do not have interferences present that were not included in the calibration set. For process monitoring or automated analysis applications, outlier detection methods must be used. In a few of the cases studied here, the PLS model was able to predict analyte concentration in the presence of unexpected interferences, but only when the PLS model had negligible regression coefficients at the location of the molecular and fragment ions of the interference. The additional use of the time profiles and the second-order calibration method, tri-PLS, resulted in no improvements in prediction results in comparison to the bilinear (first-order) PLS algorithm. Similarly to PLS, tri-PLS was only able to accurately predict analyte concentration in the presence of unexpected interferences when the model had negligible regression coefficients that overlapped with the mass spectrum and time profile of the interferences. The other second-order method studied, PARAFAC, was able to accurately estimate analyte concentration in the presence of unexpected interferences. In addition, the uniqueness feature of the PARAFAC method allowed for the possible identification of the interference when it was used in conjunction with a spectral database. In general, PLS and tri-PLS provided the best results during prediction when the samples did not contain unmodeled interferences, and PARAFAC performed better than the other methods during sample prediction when unmodeled interferences were present.

With respect to the quantitative and qualitative attributes, PARAFAC outperformed the other calibration methods studied. It was able to take advantage of the subtle temporal variations in the P&T/MS data. The only disadvantage of the PARAFAC method is the computational effort required. Because the entire data cube must be modeled each time a prediction is performed, PARAFAC model-building takes on the order of several minutes, even for fast personal computers. This is in contrast to the much faster PLS and tri-PLS algorithms, in which models are built once using the calibration data and require only simple calculations for prediction. Thus, until speed improvements in the PARAFAC method are found, PLS or tri-PLS must be the method of choice for process monitoring or automated analysis applications in which speed is critical. However, one can envision hybrid approaches that utilize PLS for routine samples, but default to PARAFAC for quantification and identification of the new interference when more complicated samples are encountered.

Further work in our laboratories will focus on the development of novel DSMS instrumentation to improve the temporal variations between environmentally important analytes. Improved separation will further improve tri-PLS and PARAFAC model performance. We also intend to explore the possibilities of using variable selection methods to expedite the PARAFAC calculations. Future work will also be performed to compare the quantitative performance of the NALMS and other related fitting methods to the algorithms employed here.

ACKNOWLEDGMENT

W.P.G. thanks the National Science Foundation for financial assistance in the form of a NSF GRT-traineeship. This project was financially supported by the Office of Naval Research under Contract No. N00014-00-WX20319. The authors thank Dr. Rasmus Bro for helpful comments and suggestions regarding parallel factor analysis.

Received for review June 15, 2000. Accepted November 1, 2000.

AC000690P