least an order of magnitude less (*8*).

A study that supports the contention that a concentration-dependent conformational change occurs in the range of Table IV was reported by Saar and Weber (*15*). The degree of cadmium(II) complexation by a FA similar to the Contech material was found to be a decreasing function of sample concentration in solutions of 0.1 M KNO_3; this was attributed to an expanded conformation in dilute (<70 mg/L) solution although the change may have been induced by the metal ion. Langford et al. (*9*) observed that light scattering from dilute aqueous solutions of Contech FA at low ionic strength is a linear function of sample concentration. This implies an absence of conformational change, which is consistent with the disaggregated state reported here for FA at 150 mg/L in the absence of background electrolyte. In contrast, Lapen and Seitz reported the absence of a conformational change over the concentration range of interest; however, the technique employed, steady-state fluorescence polarization, is relatively insensitive to small changes in molecular size. Finally, it should be noted that the solubility of FA in glycerol/water mixtures is less than that in 100% water. Therefore, the concentration-induced aggregation reported here may be absent or observed only at higher concentrations in pure water.

CONCLUSION

The results reported here show that the molecular size and shape of the fluorescent components in aqueous solutions of a fulvic acid are variable properties dependent on the solution environment. In particular, disaggregation and molecular expansion are favored in dilute solution at high pH. Contraction of the molecular units and aggregation among them are promoted at higher concentration, or when the sample is protonated, in a medium of high ionic strength. Such behavior is consistent with that of a flexible polyelectrolyte (*14*). It must be emphasized that due to the negative correlation of fluorescence with particle size, these results may or may not be germane to the larger molecular weight fractions of FA. However, previous studies (*9*, *10*, *16*) indicate that copper(II) complexation and acidic functional group content are decreasing functions of particle size; thus, fluorescence spectroscopy may be a selective probe of the more reactive fractions of the FA mixture.

LITERATURE CITED

(1) Hayes, M. H. B.; Swift, R. S. *The Chemistry of Soil Constituents*; Greenland, D. J., Hayes, M. H. B., Eds.; Wiley: London, 1983; Chapter 3.
(2) Gamble, D. S.; Schnitzer, M. *Trace Metal and Metal-Organic Interactions in Natural Water*; Singer, P. S., Ed.; Ann Arbor Science: Ann Arbor, MI, 1973; Chapter 9.
(3) Bresnahan, W. T.; Burcar, P. J.; Weber, J. H. *Anal. Chem.* **1978**, *50*, 1675-1679.
(4) Roemelt, P. M.; Seitz, W. R. *Environ. Sci. Technol.* **1982**, *16*, 613-616.
(5) Rook, J. J. *Environ. Sci. Technol.* **1977**, *11*, 478-482.
(6) Zepp, R. G.; Baughman, G. L.; Schlotzhauer, P. F. *Chemosphere* **1981**, *10*, 109-117.
(7) Ghosh, K.; Schnitzer, M. *Soil Sci.* **1980**, *129*, 266-276.
(8) Thurman, E. M.; Wershaw, R. L.; Malcolm, R. L.; Pinckney, D. J. *Org. Geochem.* **1982**, *4*, 27-35.
(9) Langford, C. H.; Gamble, D. S.; Underdown, A. W.; Lee, S. *Aquatic and Terrestrial Humic Substances*; Christman, R. F., Gjessing, E. T., Eds.; Ann Arbor Science: Ann Arbor, MI, 1983; Chapter 11.
(10) Underdown, A. W.; Langford, C. H.; Gamble, D. S. *Environ. Sci. Technol.* **1985**, *19*, 132-136.
(11) Lapen, A. J.; Seitz, W. R. *Anal. Chim. Acta* **1982**, *134*, 31-38.
(12) Azumi, T.; McGlynn, S. P. *J. Chem. Phys.* **1962**, *37*, 2413-2420.
(13) Wahl, Ph. *Biophys. Chem.* **1979**, *10*, 91-104.
(14) Tanford, C. *Physical Chemistry of Macromolecules*; Wiley: New York, 1961; Chapter 7.
(15) Saar, R. A.; Weber, J. H. *Can. J. Chem.* **1979**, *57*, 1263-1268.
(16) Saar, R. A.; Weber, J. H. *Anal. Chem.* **1980**, *52*, 2095-2100.

RECEIVED for review December 19, 1985. Accepted April 3, 1986. This work was supported, in part, by a grant from the Research Council of Duke University and by the National Science Foundation under Grant CHE-8119600.

Infrared Screening Technique for Automated Identification of Bulk Organic Mixtures

Mark A. Puskar and Steven P. Levine*

*Department of Environmental and Industrial Health, School of Public Health, The University of Michigan,
Ann Arbor, Michigan 48109*

Stephen R. Lowry

Nicolet Instrument Corporation, 5225 Verona Road, Madison, Wisconsin 53711

A program for automated waste mixture identification (PAWMI) was developed as a preliminary screening tool for unknown organics handled on remedial action sites. PAWMI has three subroutines: the interferogram processing and peak selection subroutine (PUSHSUB), the interpretation subroutine (PAIRS), and final processing subroutine to subtract spectral similarity (PAIRSPLUS). The method can be used to interpret spectra generated with both transmission and internal reflection techniques. The qualitative limit of detection of minor components in mixtures tested with PAWMI ranged from 0.4% to 15%, depending on the spectral similarity of components of the mixtures. The sensitivity of the method was 77%. False positives represented 13% of the total positive results reported.

Material analysis during hazardous waste remediations is accomplished in on-site laboratories by compatibility testing (*1*). This method classifies wastes into broad groups such as organics, halogenated organics, acids, bases, etc. This approach prevents catastrophic incidents that might be caused by rapid mixing of incompatible compounds. Generally, compatibility testing procedures separate the organic wastes into only four categories: halogenated organics, nonhalogenated organics, polar organics, and PCB containing (>50 ppm). Thus, compatibility testing procedures offer little useful information on organic chemical identification that can be used in protecting people working on or living near the site.

Compatibility testing is not as extensive as the characterization of each individual container of waste by standard laboratory testing procedures (e.g., RCRA series, gas chromatography/mass spectroscopy (GC/MS), emission spectroscopy). GC/MS offers the ability to identify many volatile and semivolatile organic substances; however recent findings (*2*) have documented major classes of organic compounds that

0003-2700/86/0358-1981$01.50/0

Table I. 62 Compounds Selected for Training Set

acetone
anthracene
benzaldehyde
benzene
2-butanone
n-butyl acetate
chlorobenzene
1-chlorodecane
chloroform
o-chlorotoluene
m-cresol
o-cresol
p-cresol
cyclohexanone
dibutyl phthalate
o-dichlorobenzene
m-dichlorobenzene
1,1-dichloroethane
1,1-dichloroethene
trans-1,2-dichloroethylene
dichloromethane
2,4-dichlorophenol
1,2-dichloropropane
1,3-dichloropropane
1,3-dichloropropene
2,3-dichloro-1-propene
2,3-dichlorotoluene
dicyclopentadiene
diethyl phthalate
2,4-dimethylphenol
dioctyl phthalate
ethanol
ethoxyethyl acetate
ethyl acetate
ethylbenzene
2-ethylphenol
m-ethyltoluene
p-ethyltoluene
1,3-hexachlorobutadiene
hexachlorocyclopentadiene
n-hexane
2-hexanone
methanol
3-methylpentane
4-methyl-2-pentanone
pentachloroethane
phenol
1-propanol
2-propanol
n-propyl acetate
styrene
1,2,3,4-tetrachlorobenzene
1,1,2,2-tetrachloroethane
toluene
1,2,4-trichlorobenzene
1,1,1-trichloroethane
1,1,2-trichloroethane
1,2,3-trichloropropane
1,2,4-trimethylbenzene
m-xylene
o-xylene
xylene-mixed

are not detected by this method in hazardous waste samples. Also, complete chemical characterization of each waste container would be far too costly and time-consuming for the purpose of hazardous waste cleanup. A technique is needed that has the speed advantages of compatibility testing (100 samples/day) but, like GC/MS, the ability to identify many major components of the waste mixtures.

The sensitivity and rapidity of Fourier transform infrared (FT-IR) spectrometry meets many of the requirements of a near real-time screening technique for organic hazardous waste samples (*1*). A primary area where the advantages of qualitative FT-IR spectrometry can be realized is direct identification of major components of hazardous waste mixtures. Direct identification techniques usually perform very little (drying the sample) or no preseparation on the sample prior to spectral interpretation. This saves the time involved in extraction and chromatography and also enables nonchromatographable compounds to be identified in the unknown waste.

This paper describes a direct infrared spectral mixture interpretation technique, PAWMI (program for automated waste mixture interpretation). PAWMI is an automated rule based artificial intelligence program designed to identify the major components in a complex mixture by interpreting the infrared spectrum of the mixture. Presently, the method has been developed to identify 62 of the most commonly identified organic compounds on hazardous waste sites (*3, 4*). Table I lists these compounds. The program was developed to be easily adapted for adding additional components to the intrepretation list, thus giving it the ability to be routinely adapted to site-specific compounds. PAWMI was designed to be automatic so it could be operated on-site (in a mobile laboratory) during a remedial action project, by trained personnel. Other applications of the PAWMI technique would include screening incoming organic waste at RCRA disposal facilities or interpreting gas-phase spectra obtained during industrial hygiene air monitoring.

New rules were written for PAIRS (*5–11*) that adapted it from a program that identifies functional groups of pure compounds into a rule-based artificial intelligence program designed to identify the major components in an unknown mixture, by interpreting the infrared spectrum of the mixture. This new use of PAIRS is described in detail elsewhere (*12*).

When this technique was tested on spectra of mixtures frequently found on hazardous waste sites (*12*), significant success was documented for two through five components. However, two problems were identified as areas for further research before this technique could be applied to actual unknown hazardous waste mixtures. The first problem was that a technique was needed that would be able to automatically select peaks in spectra, especially where the base lines are nonlinear. Broad peak envelopes and nonlinear base lines are common in hazardous waste samples due to the presence of water and polymeric mixtures. The second problem, false positive results, occurs because many of the compounds of interest on remedial action sites are structurally similar molecules (e.g., toluene, ethylbenzene, and xylene mixed).

In this paper, an automated technique is described for limiting the effect of each of these problems: PUSHSUB for peak selection and PAIRSPLUS for spectral similarity. The PAWMI software package was evaluated for the interpretation of an array of mixtures prepared to estimate the methods qualitative limit of detection and to determine if both transmission and internal reflection spectra may be successfully interpreted.

EXPERIMENTAL SECTION

Both transmission and internal reflection spectra were generated. Thin film transmission spectra where generated by placing a drop of the sample between two 13 × 2 mm KBr crystals. A double-pass fixed angle nine penetration internal reflection plate apparatus (Harrick PLC-S1M) was used. Cadmium telluride (CdTe) was selected as the prism material because of its transparency range to 400 cm^{-1}. The major advantage of this apparatus design is that the unit, when inverted and mounted into the spectrometer, could easily be dipped into a vial containing the sample. Also, it could be easily cleaned between samples. With transmission spectroscopy, the sample holder must be disassembled and cleaned between samples.

The interpretation rules were written for each of the compounds of interest using CONCISE (*7, 8*) and were compiled prior to mixture interpretation. The length of time required to interpret a spectrum through the 62 rules (which when printed out are over 200 pages long) is 60 s. This entire process, including rule writing, compiling, and spectral interpretation, was performed with a Nicolet 1280 computer. Spectra were generated with a Nicolet 20-SX optical bench with a background and sample signal averaging of 128 scans. The number of data points collected was 16 384 and the instrument resolution was 2 cm^{-1}. All solvents were Aldrich Spectrometric Grade or equivalent. Mixtures were prepared on a weight basis.

RESULTS AND DISCUSSION

PAWMI has three main subroutines: the interferogram processing and peak selection subroutine (PUSHSUB), the interpretation subroutine (PAIRS), and final processing subroutine to subtract spectral similarity (PAIRSPLUS). Figure 1 is a flow chart of the PAWMI process, where the logic of each of the three major subroutines is diagrammed.

PUSHSUB. To automate PAWMI, a peak selection subroutine PUSHSUB, was developed that does not require the operator to set a peak selection threshold. The previously used routine, PEAK PICKER, has a linear threshold value that is set manually by the operator. Only peaks that maximize above this threshold line are selected by the program for intrepretation. Using PEAK PICKER, the operator makes the decision between actual peaks and noise. However, if the operator sets the threshold too high, important peaks that are needed by the interpreter to identify components would be lost, producing false negative situations. If the threshold is

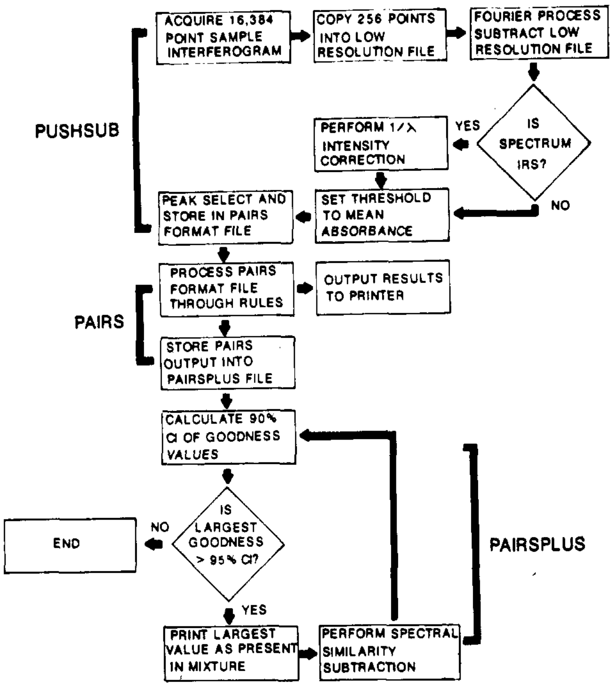

Figure 1. Flow chart of the PAWMI process, showing the logic of each of the three major subroutines.

set too low, noise would be selected as actual peaks and interpreted, producing false positives. Although the Nicolet linear threshold method works well when the spectrum has a flat linear base line, it fails when the spectrum has a nonlinear base line. This frequently occurs when dealing with actual hazardous waste sample spectra, as illustrated in Figure 2A, which is an infrared transmission spectrum of a unknown organic hazardous waste sample, collected from an EPA hazardous waste remedial action site. The sample was bright purple, and like many hazardous waste samples, water was a component of the mixture.

PUSHSUB selects peaks by transforming the first 256 data points right of the centerburst from the original 16 384 data point sample interferogram into a threshold curve, Figure 2B (low resolution spectrum), which, when subtracted from the original sample, creates a spectrum with a straight base line for peak selection, Figure 2C. PUSHSUB automatically calculates the threshold value from this file. The program has the ability to process both transmission spectra and internal reflectance spectra. If a transmission spectrum is being processed, the threshold value is directly set to the mean absorbance of the data points in the subtracted file. If an internal reflectance spectrum is processed, a 1/wavelength intensity correction (*13*) is first applied to the spectrum before the mean absorbance is calculated. In either case, the resultant value is stored in the threshold register, and peak selection is conducted by PEAK PICKER. PUSHSUB stores the peak file in a format that can be used by the interpreter program (PAIRS).

PAIRS. The mixture interpretation software uses peak location information and is based on a three-level filter algorithm designed to compensate for potential peak shifts in the mixture spectrum (*12*). If a peak falls in a relatively wide frequency window assigned to a certain compound, a percentage of the overall "goodness" value will be added to the total. "Goodness" is a measurement of closeness of match between the neat spectrum used in rule writing and the unknown spectrum. The goodness scale ranges from 0.01 for a complete mismatch to 0.99 for a complete match. As the window size is reduced, the amount of "goodness" is increased for each level where the peak in the unknown spectrum still meets the criteria.

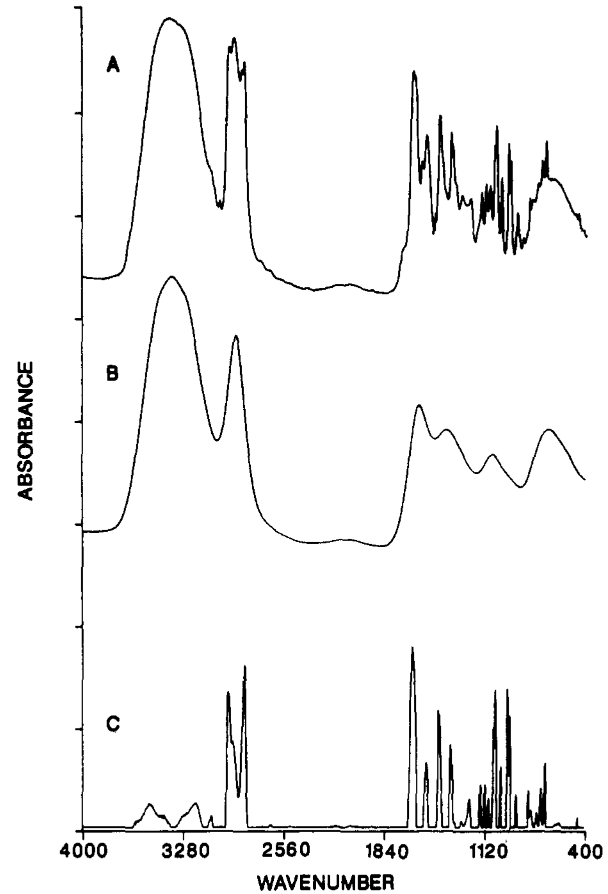

Figure 2. Comparison of neat hazardous waste spectrum (A), threshold curve (B), and subtraction (C) used in peak selection.

PAIRSPLUS. PAIRSPLUS was developed to limit the effect of spectral similarity. Statistical studies have been performed (*12*) to evaluate the quality of the final goodness value reported by PAIRS for actual compound assignment. Based on these results, a goodness value greater than 0.59 out of a possible 0.99 generally indicated the presence of the compound in the unknown spectrum. However, if a large goodness value (greater than 0.85) is returned by the interpreter for a compound that has many spectrally similar compounds in the interpretation set, goodness values greater than 0.59 may be returned for spectrally similar compounds. Because these compounds are not actually in the sample, but are interpreted to be there, they are considered false positives. For example, if a neat spectrum of ethylbenzene is interpreted by PAIRS, 0.99 will be returned for ethylbenzene but 0.67 will also be returned for toluene. So, unless this effect is corrected for, each time ethylbenzene is identified in a sample (true positive) with a large goodness value, toluene will also be identified (false positive).

By adapting a subroutine of the PAIRS program, two files were created. The first file contains a 62 by 62 array of the goodness value returned by the interpreter when each of the neat compounds was interpreted by each of the 62 rules. Table II is a portion of that array showing the goodness values returned by the interpreter for neat spectra of toluene, chlorobenzene, 1,1,1-trichloroethane, and benzene. When toluene is interpreted by the rules, 0.99 goodness is returned for toluene, 0.23 goodness is returned for chlorobenzene, 0.30 goodness is returned for 1,1,1-trichloroethane, and 0.23 is returned for benzene. Although 0.30 goodness is returned by the interpreter for 1,1,1-trichloroethane when toluene is interpreted, only 0.01 goodness is returned for toluene when 1,1,1-trichloroethane is interpreted. The difference in goodness between the two interpretations is due to the differences in number of peaks selected for rule-writing and interpretation.

Table II. Portion of PAIRSPLUS Spectral Similarity Array, Showing Goodness Values Reported for Four Compounds When Neat Spectra Are Interpreted by the Rules

	pairs results				
compound processed	toluene	chlorobenzene	1,1,1-TCE	benzene	total goodness
toluene	[0.99]	0.23 (2)	0.30 (0)	0.23 (1)	0.76
chlorobenzene	0.24 (0)	[0.99]	0.44 (2)	0.28 (1)	0.96
1,1,1-TCE	0.01 (3)	0.24 (4)	[0.99]	0.01 (0)	0.26
benzene	0.04 (2)	0.10 (0)	0.01 (0)	[0.99]	0.15
total additnl compds identified by PAIRSPLUS (KBr)	(5)	(6)	(2)	(2)	

Table III. Top Ten Compounds Identified and Goodness Values Reported by PAIRS When a Four-Component Mixture of 45.5% 1,1,1-Trichloroethane, 45.5% Benzene, 4.5% Toluene, and 4.5% Chlorobenzene was Interpreted

	compound name	goodness value
1	1,1,1-trichloroethane	0.99
2	benzene	0.99
3	chlorobenzene	0.46
4	anthracene	0.36
5	ethoxyethyl acetate	0.35
6	toluene	0.32
7	1,1-dichloroethane	0.27
8	*n*-butyl acetate	0.26
9	*o*-chlorotoluene	0.23
10	1,2-dichloropropane	0.23

The ten largest peaks in toluene were selected for rule-writing (the criteria are given elsewhere (*12*)). However, when a neat spectrum of toluene is processed by PUSHSUB, 23 peaks are selected for interpretation. Only six peaks where selected for rule-writing for 1,1,1-trichloroethane, and when a neat spectrum of 1,1,1-trichloroethane is processed by PUSHSUB, only eight peaks are selected for interpretation. The difference in number of peaks selected accounts for the differences in goodness values reported.

The second file created by PAIRS contains the 62 goodness values generated by the rules during spectral interpretation. PAIRSPLUS accesses both the array and interpretation results and subtracts the percentage of spectral similarity corresponding to the compound with the largest goodness value from all the remaining compounds' goodness values. A statistical check is then conducted on the remaining compounds to determine if another compound should be reported as present in the unknown sample. This is accomplished by calculating the mean and standard deviation of the remaining 61 goodness values. If the next largest goodness value, in the 61 remaining, is greater than two times the standard deviation above the mean (95% confidence interval), it is reported, the array is accessed, and the percentage of spectral similarity corresponding to its goodness value is subtracted from all the remaining compound's goodness values. This is repeated until the statistical check determines there are no goodness values greater than the 95% confidence interval. At that point, the program terminates.

The power of this routine is demonstrated when an actual component of a mixture is present, but at a low concentration. The goodness value corresponding to it is below 0.60 and, according to the cutoff level set for PAIRS, it is not considered present in the mixture. However, once the compounds in the interpretation set that have spectral similarity with the major components are removed by PAIRSPLUS, the compound is identified as being present in the mixture. An example of this is when the four-component mixture of 45.5% 1,1,1-trichloroethane, 45.5% benzene, 4.5% toluene, and 4.5% chlorobenzene is processed by PAWMI. Table III lists the top ten compounds identified by the interpreter with their corresponding goodness values. Goodness values of 0.99 are returned for both 1,1,1-trichloroethane and benzene. Chlorobenzene has the next highest goodness, reported at 0.46. Toluene, which is the seventh compound listed in Table III, has a goodness value of 0.32 reported when the mixture is interpreted. With a 0.60 cutoff for compounds being present in the mixture, only 1,1,1-trichloroethane and benzene would be identified in the mixture by PAIRS.

Table IV. Results of PAIRSPLUS Processing of PAIRS Results Listed in Table III

	compound name	goodness value	results of remaining compounds		
			mean	σ	95% CI
1	1,1,1-Trichloroethane	0.99	0.14	0.19	0.51
2	Benzene	0.98	0.05	0.16	0.37
3	Toluene	0.28	0.00	0.09	0.18
			−0.02	0.08	0.14

Table IV lists the results when the PAIRS results from Table III are processed by PAIRSPLUS. For each compound listed, its goodness value, the mean of the remaining goodness values, the standard deviation (σ) of the remaining goodness values, and the corresponding 95% confidence interval are listed. PAIRSPLUS identified 1,1,1-trichloroethane, benzene, and toluene in the mixture, but missed chlorobenzene (a false negative). The reason chlorobenzene was missed, even though it had the third largest goodness value, and toluene was identified, even though it had the seventh largest goodness value, is a function of the spectral similarity between the four compounds. From Table II, when 1,1,1-trichloroethane is processed by PAIRS, 0.24 goodness is returned for chlorobenzene. So when the mixture is interpreted and 0.99 goodness is returned for 1,1,1-trichloroethane, PAIRSPLUS subtracts 99% of the spectral similarity for chlorobenzene (0.24) from the chlorobenzene goodness total. So when the next largest goodness value, benzene at 0.98 goodness, is processed, the chlorobenzene goodness total has already been lowered to 0.23. Benzene lowers the chlorobenzene score another 0.10 due to the spectral similarity between benzene and chlorobenzene. By the third pass through the PAIRS results, toluene has the next highest goodness value of 0.28 (note from Table II that toluene has low spectral similarity when 1,1,1-trichloroethane and benzene are processed) and the goodness value for chlorobenzene has dropped to 0.13. After 28% of toluene corresponding spectral similarity is removed from the remaining 59 compounds, the chlorobenzene goodness has dropped to 0.07, and the 95% confidence interval calculated is 0.14 goodness. No additional compounds have goodness values above the confidence interval and the process stops.

Qualitative Limits of Detection. To estimate PAWMI's ability to identify minor components in mixtures, an array of 66 four-component mixtures were prepared using toluene, chlorobenzene, 1,1,1-trichloroethane, and benzene as an example of compounds in a hazardous waste mixture. These four compounds were selected because they are four of the most commonly identified organic compounds on remedial action sites and have typical spectral similarity features.

Four concentration ranges where prepared for each compound in the mixtures. The approximate concentrations were

Table V. Results of PAIRS' and PAIRSPLUS' Ability To Identify Toluene in 16 Four-Component Mixtures, Generated by Both Transmission (KBr) and Internal Reflection (ATR) Techniques, as the Concentration of Chlorobenzene in the Mixtures Decreases

	% chlorobenzene	% toluene	toluene interpretation results: KBr PAIRS	KBr PLUS	ATR PAIRS	ATR PLUS
1.	25	25	0.74	yes	0.74	yes
	25	14	0.64	yes	0.64	yes
	25	3	0.24	no	0.16	no
	25	0.3	0.16	no	0.16	no
			13%[a]		13%[a]	
2.	14	29	0.74	yes	0.74	yes
	14	17	0.72	yes	0.64	yes
	14	4	0.22	no	0.24	no
	14	0.4	0.12	no	0.14	no
			14%[a]		16%[a]	
3.	3	32	0.72	yes	0.74	yes
	3	19	0.72	yes	0.74	yes
	3	5	0.32	yes	0.34	yes
	3	0.5	0.12	no	0.14	no
			15%[a]		14%[a]	
4.	0.3	33	0.72	yes	0.74	yes
	0.3	20	0.72	yes	0.74	yes
	0.3	5	0.32	yes	0.34	yes
	0.3	0.5	0.02	no	0.04	no
			16%[a]		15%[a]	

[a] Percent toluene when goodness = 0.60.

30%, 18%, 4%, and 0.4% by weight. Each compound was prepared at the same concentration in a group of mixtures where two of the other components were at the same concentration, while the final compound's concentration was changing. A mixture was prepared and interpreted for each possible combination. This was done to generate data on how the qualitative limits of detection of a minor component in a mixture was affected by the presence of other components in that mixture.

An example of this is shown in Table V, which tabulates the results of PAIRS' and PAIRSPLUS' ability to identify toluene in 16 four-component mixtures, generated by both transmission and internal reflection techniques, as the concentration of chlorobenzene in the mixture decreases. The data are presented in four groups of four mixtures. In the first group of four mixtures, the chlorobenzene mass ratio is 1 and concentration is 25% in the mixtures as the toluene concentration lowers from 25% in the first mixture to 0.3% in the fourth. This logic continues through the remaining groups, so that in the fourth group of mixtures, the chlorobenzene mass ratio is now 0.01 and the concentration is 0.3% in the mixtures as the toluene concentration lowers from 33% in the first mixture to 0.5% in the fourth.

The interpretation results for toluene are presented for PAIRS and PAIRSPLUS. By use of the 0.60 cutoff for PAIRS interpretation, toluene was identified in the four groups of four-component mixtures down to the 15% concentration range in both transmission and internl reflection spectra. The PAIRS limit of detection is estimated for toluene in the mixture group by calculating the expected concentration for a goodness of 0.60 (using a linear regression equation). Note that the limit of detection determined for the transmission spectra of toluene, using the linear regression method, increases from 13% in mixture group I to 16% in mixture group III. This increase is due to the changing slope of the linear regression line, caused by the decreasing spectral similarity effect of chlorobenzene on toluene, thereby lowering toluene goodness.

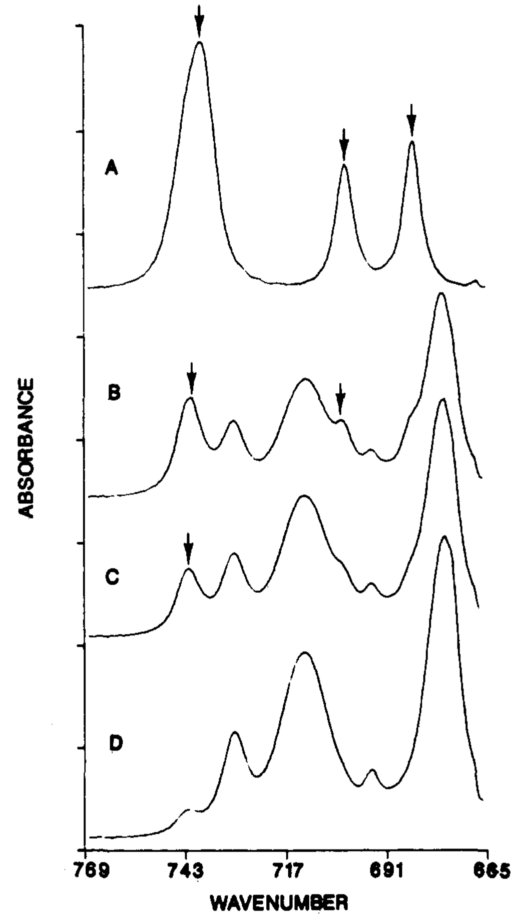

Figure 3. Comparison of chlorobenzene peaks in neat spectrum (A), 17% chlorobenzene in four-component mixture (B), 4% chlorobenzene in similar mixture (C), and 0.4% chlorobenzene in similar mixture (D).

Generally, running the results of the PAIRS process through PAIRSPLUS lowered the limit of detection of the method down to the 5% level. The data in Table V show that results obtained with PAIRSPLUS are not only dependent on the concentration of the component of interest in the mixture but are also dependent on the concentration of other components in the mixture. For example, once the concentration of chlorobenzene in the samples was lowered from 14% in mixture group II to 3% in mixture group III, PAIRSPLUS was able to identify toluene at the 5% level. The reason for this decrease in limit of detection is the reduction of spectrally similar compounds, as was explained earlier.

The data in Table V was also generated for toluene as 1,1,1-trichloroethane concentration changed while benzene and chlorobenzene remained constant and for toluene as benzene concentration changed while 1,1,1-trichloroethane and chlorobenzene remained constant. The entire process was also repeated for determining chlorobenzene, 1,1,1-trichloroethane, and benzene limits of detection while one of the other compound's concentration changed and the other two compounds' concentrations remained constant. A total of 12 Table Vs were generated, but the remaining 11 are not presented to save space.

The PAIRS method identified all minor components down to the 15% level in the four-component mixtures. However, at the 5% level the majority of goodness values had dropped below the 0.60 cutoff and PAIRS interpretation was ineffective. One of the reasons for this is shown in Figure 3. Figure 3A is an absorption spectrum of chlorobenzene in the range of 665–769 cm^{-1}. The three marked peaks are used by the PAIRS rules to identify the presence of chlorobenzene in mixtures. Figure 3B is a portion of a four-component mixture spectrum where the toluene concentration is 17%, 1,1,1-trichloroethane and benzene concentration are 33%, and the chlorobenzene concentration is 17%. At a concentration of 17% in the mixture, chlorobenzene peak number three has become a shoulder of a larger peak at 675 cm^{-1}. PUSHSUB is unable to select shoulders as peaks, so the peak is lost during peak selection and its corresponding goodness values are lost during interpretation. However, the goodness value calculated for chlorobenzene in this mixture was still 0.74, and chlorobenzene is considered a true positive in the mixture. In Figure 3C, the concentration of chlorobenzene has dropped to 4%. At this concentration, chlorobenzene peak 2 is lost in a shoulder of a broad peak at 714 cm^{-1}, but peak 1 is still selected. The goodness value calculated for chlorobenzene in the mixture is 0.46, which is considered a false negative. Figure 3D is a portion of a similar four-component mixture spectrum, the only difference being that the chlorobenzene concentration is now at 0.4% in the mixture. Although peak 1 can be still be seen in Figure 3D, the peak selection routine, PUSHSUB, does not select it. The goodness value calculated for chlorobenzene in the mixture is 0.36.

PAIRS identified 1,1,1-trichloroethane down to the 5% level in mixtures where either toluene, chlorobenzene, or benzene concentrations were reduced, and down to the 0.5% level in mixtures where the concentration of both toluene and chlorobenzene were at the 33% level. The reason for the lowering of the detection limit of PAIRS for 1,1,1-trichloroethane is an artifact, due to spectral similarity with both toluene and chlorobenzene.

Table II shows that a goodness of 0.30 is assigned to 1,1,1-trichloroethane when toluene is present and a goodness of 0.44 when chlorobenzene is present. Thus, the goodness assigned to 1,1,1-trichloroethane when a mixture of toluene and chlorobenzene is present should be at least 0.44 but could be as high as 0.74. In actuality, when a 1:1 mixture of toluene and chlorobenzene is interpreted, 0.64 goodness is returned for 1,1,1-trichloroethane. So, by strictly using PAIRS results, 1,1,1-trichloroethane will be identified in any mixture where toluene and chlorobenzene are in significant concentrations, whether it is there or not. In the mixtures described above, it is actually there and is considered a true positive. However if it is was not actually present in the toluene–chlorobenzene mixture, it would still be identified and would be considered a false positive.

As is shown in Table V, PAIRSPLUS has the ability to identify minor components of mixtures, in specific cases, below the 15% level of PAIRS down to the 5% and 0.5% levels. In the 36 spectra interpreted where a minor component was at the 5% level and one of the other components was at or below the 15% level, the minor component was identified in 44% of the transmission spectra and 55% of the internal reflection spectra. In the 24 spectra where a minor component was at the 5% level and one of the other components was at or below the 5% level, the minor component was identified in 63% of the transmission spectra and 83% of the internal reflection spectra. The reason why PAIRSPLUS only lowers the methods limit of detection for specific mixtures is explained by spectral similarity.

In Table II, next to each PAIRS result, in parentheses, is listed the number of additional compounds identified by PAIRSPLUS for the compound processed, when the concentration of the compound listed under the PAIRS Results heading was lowered in the group of mixtures. For example, Table V list the data generated on toluene interpretation as the chlorobenzene concentration changes in the group of mixtures. For the transmission spectra, toluene was correctly identified in 8 of the 16 mixtures by PAIRS. PAIRSPLUS identified toluene correctly in 10 of the 16 mixtures. The net difference between the PAIRS and PAIRSPLUS interpretation for toluene as chlorobenzene concentration changes is 2. Thus, in Table II under the spectral similarity goodness value for chlorobenzene when toluene is processed, it is a 2. This same format was repeated for the other 11 combinations.

Table VI. Comparison of PAIRSPLUS Results between Spectra Generated Using Transmission vs. Internal Reflection Techniques on the Total Array of Four-Component Mixtures

	transmission spectra (KBr)	internal reflection spectra (ATR)
true positives	200	208
false positives	28	32
false negatives	64	60
true negatives	3800	3854
total	4092	4154
sensitivity	76	78
specificity	99.3	99.2

When total goodness for a specific compound processed is compared to the total additional compounds identified by PAIRSPLUS when that compound's concentration is lowered in the mixtures, a linear relationship is observed with a correlation coefficient of 0.993. This means that when chlorobenzene, the compound with the most spectral similarity to the other three compounds (0.96), has its concentration lowered in mixtures containing the other three components, the greatest number of addition minor components (*6*) are identified by PAIRSPLUS, rather than by PAIRS. Also, when benzene, which has the least spectral similarity to the other three compounds (0.15), has its relative concentration lowered in mixtures containing the other three components, the least number of addition components (2) are identified by PAIRSPLUS, rather than by PAIRS. Therefore, PAIRSPLUS' ability to identify components in mixtures at the 5% level or lower is spectrally dependent.

In the array of mixtures interpreted during this study, internal reflection spectra where correctly interpreted at approximately the same percentage as transmission spectra. This is worth discussing because: (1) a slight difference in peak locations exists between transmission and internal reflection spectra, (2) the rules were written using peak locations from transmission spectra, and (3) the transmission spectra are of a higher quality and resolution than the internal reflection spectra.

Table VI compares PAIRSPLUS results between spectra generated using transmission (KBr plates) vs. internal reflection techniques (nine reflection ATR) on the total array of four-component mixtures. The sensitivity of PAIRSPLUS was 76% for transmission spectra and 78% for internal reflection spectra. Sensitivity is defined as the percentage of compounds actually present in the mixtures identified by the technique. The specificity of PAIRSPLUS using transmission spectra was 99.3% and for internal reflection spectra was 99.2%. Specificity is defined as the percentage of compounds not actually present in the mixtures and identified as such.

In comparing the ability of PAIRSPLUS to correctly identify components in mixtures when using internal reflection vs. transmission spectra, the two spectral generation techniques were comparable through the entire array of mixtures, except for two mixture groups where the results obtained were far superior when interpreting the internal reflection spectra. These two groups are responsible for the higher total results compiled for internal reflection spectra and the reason for each is discussed below.

The first mixture group where minor components were interpreted in internal reflection spectra at a far superior rate

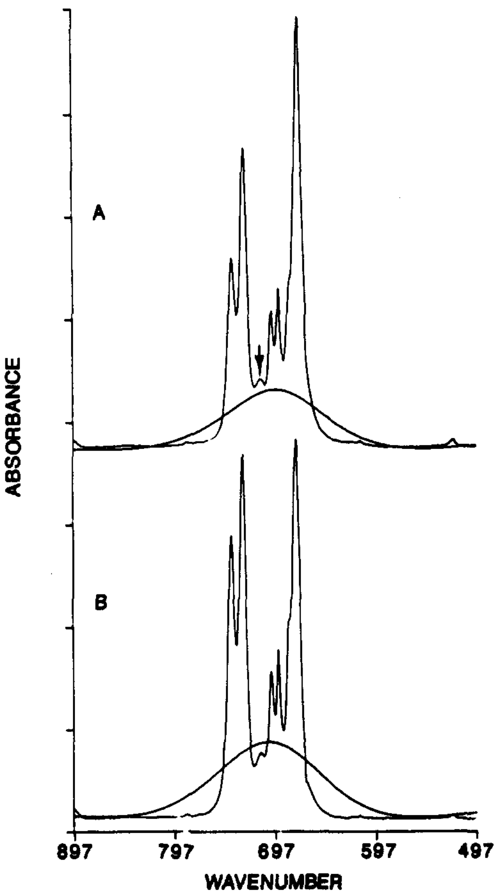

Figure 4. Comparison of internal reflection (A) and transmission (B) spectra of four-component mixture. Threshold curves used in peak selection are overlaid.

over the transmission spectra was the group of mixtures where 1,1,1-trichloroethane was being monitored as the concentration of benzene was being lowered. For example, in a mixture where the concentration of toluene and chlorobenzene was 38%, benzene concentration was 19%, and 1,1,1-trichloroethane concentration was 4%, a goodness value of 0.80 was returned by PAIRS for 1,1,1-trichloroethane and it was identified by PAIRSPLUS in the internal reflection spectra. However, with the transmission spectra, only 0.63 goodness was returned for 1,1,1-trichloroethane by PAIRS and it was not identified by PAIRSPLUS. This was also true in the other mixtures in this group where the concentrations were similar for the other three components, but benzene concentration was at 4% and 0.4%. The difference between the goodness values returned by the interpreter for 1,1,1-trichloroethane is attributed to the loss of one additional peak in the transmission spectrum when compared to the internal reflection spectrum.

Figure 4 compares the transmission and internal reflection spectra of the mixture between 497 cm^{-1} and 897 cm^{-1}, with their corresponding threshold curves used in peak selection. Figure 4A shows the internal reflection spectrum. The arrow points to a 1,1,1-trichloroethane peak at 713 cm^{-1}. Because the penetration depth, in internal reflection spectroscopy, is proportional to wavelength, the effective thickness of the sample also increases with wavelength. This is the reason that, for internal reflectance spectra of bulk materials, the absorption bands at the longer wavelengths are relatively stronger in intensity; i.e., two bands that are equal in strength in transmission spectra will have unequal strengths in ATR spectra. Thus, the shorter wavenumber bands in Figure 4A appear to be stronger than the same peaks in Figure 4B. The same is true for the low-resolution spectrum which is generated from the sample interferogram and subtracted from the sample during PUSHSUB. This intensity difference causes the peak, at 713 cm^{-1}, to be above the threshold curve and thus selected for interpretation, in Figure 4A. Figure 4B is the transmission spectrum of the same mixture. The peak intensities are not a function of wavelength and the peak intensities are significantly different than the peak intensities in Figure 4A. Although the 1,1,1-trichloroethane peak is visible at 713 cm^{-1}, it is below the threshold curve and is not selected during peak selection. Thus, the difference between the transmission and internal reflection spectra interpretation, for this example, is a peak selection problem caused by the intensity differences between transmission and internal reflection spectra.

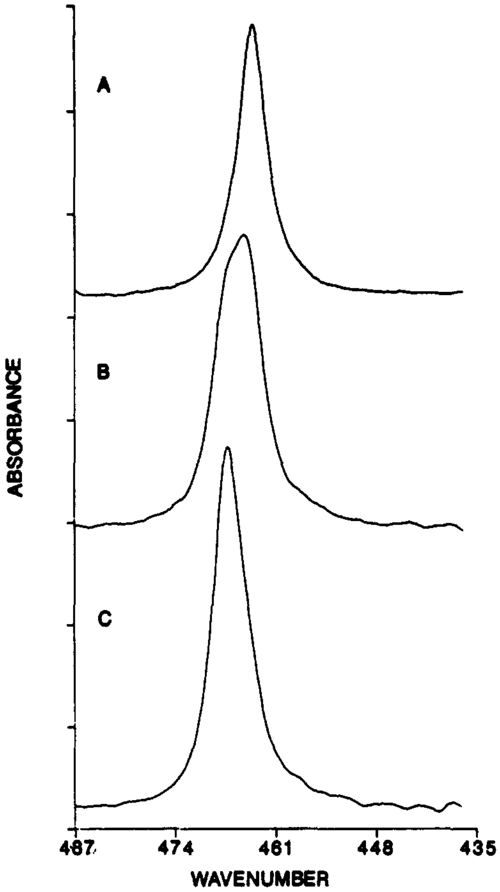

Figure 5. Comparison of toluene peak at 464 cm^{-1} between neat spectrum peak used in rule writing (A) and four-component mixture spectra generated with either internal reflection (B) or transmission (C) spectroscopy.

The second mixture group where minor components were interpreted in internal reflection spectra at a far superior rate than the transmission spectra, was with a group of four mixtures where the concentrations of chlorobenzene and 1,1,1-trichloroethane where either 45% or 49.5% and the toluene and benzene concentrations where either 5% or 0.5%. The reason for higher goodness values being reported for toluene in the internal reflection spectra can be attributed to greater shifting of the toluene peak at 464 cm^{-1} in the transmission spectrum than what was observed in the internal reflection spectrum. Figure 5 compares the toluene peak at 464 cm^{-1} generated with transmission vs. internal reflection techniques. Figure 5A is the toluene peak at 464 cm^{-1} in a neat spectrum, which was used for rule writing. Figure 5B is the internal reflection spectrum peak of a mixture which contains 45% chlorobenzene, 45% 1,1,1-trichloroethane, 5% benzene, and 5% toluene. The location of this peak is 465 cm^{-1}. Figure 5C is the transmission spectrum peak of the same mixture. Its location is 469 cm^{-1}. To receive the total goodness assigned to this toluene peak during interpretation, a peak

Table VII. Results of PAIRS Interpretation for Xylene-Mixed When Neat Compound and Mixture Spectra Are Processed

compound interpreted	goodness value reported for xylene-mixed
toluene	0.39
chlorobenzene	0.09
1,1,1-trichloroethane	0.02
benzene	0.01
toluene/chlorobenzene	0.58
toluene/1,1,1-trichloroethane	0.41
toluene/benzene	0.42

must be found in the mixture between 467 cm^{-1} and 461 cm^{-1} (*12*). The internal reflection spectrum peak does fall within this window and thus will receive the entire goodness value assigned to that peak. However, the transmission spectrum peak does not fall within the window and will only receive partial goodness. This difference in goodness will account for the reason why, when the interpretation results are processed by PAIRSPLUS, toluene is identified in the internal reflection spectrum but not in the transmission spectrum. Thus, for this example, the reason for the difference between the transmission and internal reflection spectral interpretation was a peak-shifting problem.

Table VI summarizes and compares PAIRSPLUS results between spectra generated using transmission vs. internal reflection techniques on the total array of four-component mixtures. A total of 60 false positives were reported, which represents 13% of the total positive results reported. However, 85% of the 60 false positives reported by PAIRSPLUS were the last and lowest concentration compound reported.

Forty-five percent of the false positives reported were for the single component, xylene-mixed. Xylene-mixed is a mixture spectrum of the ortho and meta xylene isomers. It was selected for rule writing because this mixture is very commonly used in industry and is found frequently on remedial action sites. When the mixtures where xylene-mixed was returned as a false positive were compared, it was observed that these mixtures contained both toluene and chlorobenzene at significant concentrations. The concentration of the two components together averaged 74% in the mixtures. Thus, an artifact existed in mixtures that had a majority of toluene and chlorobenzene.

Table VII lists the goodness values reported for xylene-mixed when the four-mixture components are interpreted separately and when a two-component mixture of toluene and each of the other three components (1:1) w/w are interpreted. The false positive occurs because the mixture spectrum of toluene and chlorobenzene has more spectral similarity with the xylene-mixed rules than the total spectral similarity of each of the two components interpreted independently. When a neat spectrum of toluene is interpreted by the rules, 0.39 goodness is returned for xylene-mixed. When chlorobenzene is interpreted, a goodness of 0.09 is returned for xylene-mixed. However, when a mixture of toluene and chlorobenzene is interpreted, a goodness of 0.58 is returned for xylene-mixed, which is 0.10 goodness values greater than the two independent scores combined. This extra 0.10 goodness value for xylene-mixed becomes statistically significant after the toluene and chlorobenzene are identified and processed by PAIRSPLUS. Note that the spectral similarity of the toluene/1,1,1-trichloroethane and toluene/benzene mixtures are approximately the additive spectral similarity of the compounds processed independently.

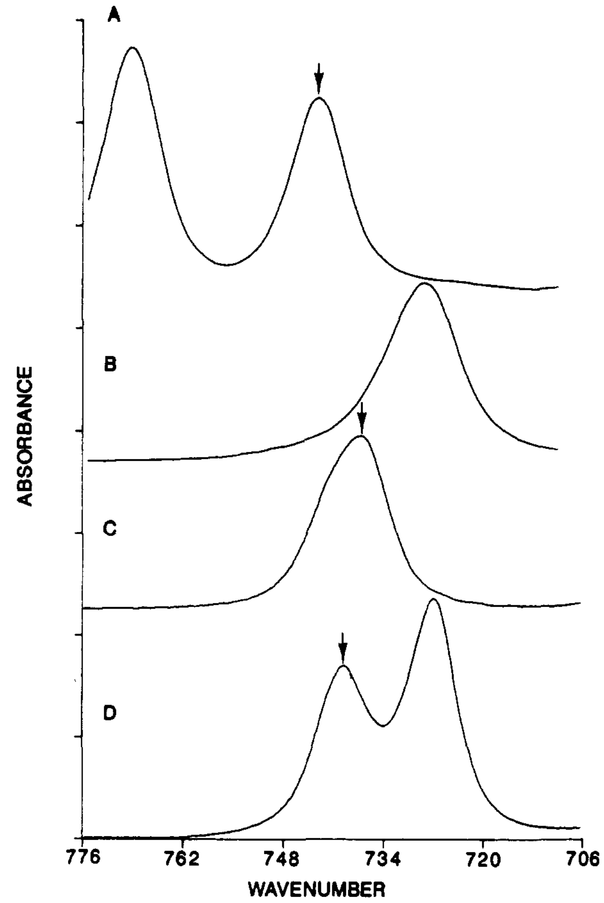

Figure 6. Comparison of xylene-mixed peak location at 743 cm^{-1} between neat spectrum used for rule writing (A), neat toluene spectrum (B), neat chlorobenzene spectrum (C), and toluene/chlorobenzene mixture spectrum (D).

Peak shifting is responsible for 0.58 goodness being reported for xylene-mixed when a mixture containing toluene and chlorobenzene is interpreted, while a total of 0.48 is returned when the compounds are interpreted independently. In two unique locations of the infrared spectrum of the mixture, peak shifting causes a peak to be located in a tighter window corresponding to the presence of xylene-mixed. One of the locations, as an example, is shown in Figure 6. Figure 6A is a portion of the xylene-mixed spectrum between 706 cm^{-1} and 776 cm^{-1}. Both peaks shown are used by the rules to determine if xylene-mixed is in an unknown spectrum. The right peak is at 743 cm^{-1}. Figure 6B is the same portion of a toluene spectrum. Note that toluene has no spectral similarity with the two xylene-mixed peaks shown. Figure 6C is the same portion of a chlorobenzene spectrum. The chlorobenzene peak at 738 cm^{-1} is only 5 cm^{-1} away from the xylene-mixed peak and the xylene-mixed total goodness will receive 0.05 goodness from that peak when a spectrum of chlorobenzene is processed through the rules. Figure 6D is the same portion of a toluene/chlorobenzene mixture (1:1 w/w) spectrum. The chlorobenzene peak in the mixture spectrum has shifted from 738 cm^{-1} to 741 cm^{-1}. This 3-cm^{-1} shift has caused the peak to move into the tightest window assigned for the 743-cm^{-1} peak in xylene-mixed. Thus, when this mixture spectrum is interpreted, the total goodness for xylene-mixed will receive 0.10 goodness from the shifted peak. This is a 0.05 goodness increase over the two compounds independently interpreted. Peak shifting also occurs with a toluene peak in the spectrum at 2863 cm^{-1}, which adds the other 0.05 goodness to the xylene-mixed goodness. Thus, peak shifting is the reason for xylene-mixed being considered a false positive when a mixture contains large concentrations of toluene and chlorobenzene.

CONCLUSIONS

Significant improvements were made on PAIRS' ability to identify components of mixture spectra. The peak selection software, PUSHSUB, is a improvement over PEAK PICKER alone on spectra with nonlinear base lines. In addition, PAIRSPLUS was able to lower the qualitative limit of detection of the PAIRS interpretation in specific mixtures where spectral similarity effects could be subtracted.

To understand PAWMI's ability to identify major and minor components in infrared spectra of mixtures, an array of mixtures were prepared, both transmission and internal reflection spectra were generated, and the spectra were interpreted. The qualitative limit of detection of minor components in mixtures ranged from 0.4% to 15%, depending on the components of the mixture. This was not only dependent on the concentration of the component of interest in the mixture but also on the concentration of other spectrally similar components in the mixture.

In the array of mixtures interpreted during this study, internal reflection spectra where correctly interpreted at a slightly higher percentage than transmission spectra. The reasons were: intensity differences between transmission and internal reflection spectra causing additional peaks to be selected by PUSHSUB, resulting in higher goodness value assigned for minor components in the internal reflection spectra; greater peak shifting in transmission mixture spectra, causing lower total goodness values being reported by PAIRS and the minor components not being identified by PAIRSPLUS.

False positives represented 13% of the total positive results reported by PAWMI. This occurred when the spectral similarity of a mixture to the false positive component was greater than the additive spectral similarity of each component of the mixture. Many of these false positives could be eliminated by placing additional constraints on the PAIRSPLUS processing routine. However, a number of true positives would also be lost. Reducing the number of false positives is an area of future research.

ACKNOWLEDGMENT

The authors wish to thank Arthur Palmer for his help with the programming and Greg Kinnes for his help in preparing the mixtures.

LITERATURE CITED

(1) Puskar, M. A.; Levine, S. P.; Turpin, R. In *Protecting Personnel at Hazardous Waste Sites*; Levine, S. P., Martin, W. F., Eds.; Butterworths/Ann Arbor: Woburn, MA, 1985; Chapter 6.
(2) Gurka, D. F. *Project Summary*: *Interlaboratory Comparison Study*: *Methods for Volatile and Semivolatile Compounds*; Environmental Monitoring Systems Laboratory: Las Vegas, NV, June 1984; EPA-600/S4-84-027.
(3) Hallstedt, P. A.; Levine, S. P.; Puskar, M. A. *J. Hazard. Waste Hazard. Mater.*, in press.
(4) Eckel, W. P.; Trees, D. P.; Kovell, S. P. "Distribution and Concentration of Chemicals and Toxic Materials Found at Hazardous Waste Dump Sites"; Proceedings of the National Conference on Hazardous Waste and Environmental Emergencies, May 1985.
(5) Woodruff, H. B.; Munk, M. E. *J. Org. Chem.* **1977**, *42*, 1761–1767.
(6) Woodruff, H. B.; Munk, M. E. *Anal. Chim. Acta* **1977**, *95*, 13–23.
(7) Woodruff, H. B.; Smith, G. M. *Anal. Chem.* **1980**, *52*, 2321–2327.
(8) Woodruff, H. B.; Smith, G. M. *Anal. Chim. Acta* **1981**, *133*, 545–553.
(9) Tomellini, S. A.; Saperstein, D. D.; Stevenson, J. M.; Smith, G. M.; Woodruff, H. B. *Anal. Chem.* **1981**, *53*, 2367–2369.
(10) Tomellini, S. A.; Stevenson, J. M.; Woodruff, H. B. *Anal. Chem.* **1984**, *56*, 67–70.
(11) Tomellini, S. A.; Hartwick, R. A.; Stevenson, J. M.; Woodruff, H. B. *Anal. Chim. Acta* **1984**, *162*, 227–240.
(12) Puskar, M. A.; Levine, S. P.; Lowry, S. R. *Anal. Chem.* **1986**, *58*, 1156–1162.
(13) Harrick, N. J. *Internal Reflection Spectroscopy*; Wiley: New York, 1967; 2nd Printing, Harrick Scientific Corp.: Ossining, NY, 1979.

RECEIVED for review January 6, 1986. Accepted April 18, 1986. This work was supported by Grant 1-R01-OHO2066-01 from the National Institute for Occupational Safety and Health of Centers for Disease Control.

Spatial and Temporal Distributions of Particulates Formed from Metallic Surfaces by Laser Vaporization

Carmen W. Huie and Edward S. Yeung*

Ames Laboratory—USDOE and Department of Chemistry, Iowa State University, Ames, Iowa 50011

By observation of scattered radiation at 90° relative to a probe laser beam, the particulate matter that is formed above metallic surfaces following laser vaporization can be monitored. An acoustooptic deflector scans the direction of the probe beam so that spatially and temporally resolved distributions can be recorded from a single transient event. Distinct scattering maps are obtained as a function of laser power, surface property, absorption characteristics, and material volatility. The results can be used to understand the laser-surface interactions, so that atom formation can be better controlled for analytical applications.

The use of lasers to vaporize, dissociate, excite, or ionize species on solid surfaces has the potential of becoming a powerful analytical tool. Since laser beams can be focused to very small sizes, local concentrations of materials on surfaces can be probed. The material vaporized is usually quite well confined, so that analytical sensitivity is enhanced. In general, either optical spectroscopy (neutral species) (*1*) or mass spectrometry (ions) (*2, 3*) can be used to probe the laser-vaporized material. The main problem is the lack of reproducibility in the generation of atoms or ions, in part due to the pulse-to-pulse fluctuations in the laser intensity, and in part due to variations in the surface properties of the sample. To turn the semiquantitative results obtained so far into reliable quantitative results, one must try to obtain a better understanding of the laser-surface interaction. Progress has been made by obtaining the spatial and temporal distribution of atoms (*4*), molecules (*5*), translational temperature (*6*), and vibrational temperature (*7*) in the laser-generated plumes. It may then be possible to use thermodynamic information to model the vaporization–dissociation–excitation–ionization process for various laser beam and surface properties. The one type of data that has so far been neglected is that of particle formation from the laser vaporization process. This is the first step in the chain for atom production. One must be able to volatilize these particles completely before even worrying about dissociation. Current successes in atomic spectroscopy in flames, inductively coupled plasma, and graphite furnaces have benefited greatly from earlier studies of nebulization processes and the design of sample introduction techniques to optimize nebulization (*8*). The same