Journal of Exposure Analysis and Environmental Epidemiology **(2001) 11, 295–307**
© 2001 Nature Publishing Group All rights reserved 1053-4245/01/$17.00
www.nature.com/jea

Source apportionment of exposure to toxic volatile organic compounds using positive matrix factorization

MELISSA J. ANDERSON,[a] SHELLY L. MILLER[b] AND JANA B. MILFORD[b]

[a]*Department of Civil, Architectural and Environmental Engineering, The University of Colorado at Boulder, Boulder, Colorado*
[b]*Department of Mechanical Engineering, The University of Colorado at Boulder, Boulder, Colorado*

Data from the Total Exposure Assessment Methodology studies, conducted from 1980 to 1987 in New Jersey (NJ) and California (CA), and the 1990 California Indoor Exposure study were analyzed using positive matrix factorization, a receptor-oriented source apportionment model. Personal exposure and outdoor concentrations of 14 and 17 toxic volatile organic compounds (VOCs) were studied from the NJ and CA data, respectively. Analyzing both the personal exposure and outdoor concentrations made it possible to compare toxic VOCs in outdoor air and exposure resulting from personal activities. Regression analyses of the measured concentrations versus the factor scores were performed to determine the relative contribution of each factor to total exposure concentrations. Activity patterns of the NJ and CA participants were examined to determine whether reported exposures to specific sources correspond to higher estimated contributions from the factor identified with that source. For a subset of VOCs, a preliminary analysis to determine irritancy-based contributions of factors to exposures was carried out. Major source types of toxic VOCs in both NJ and CA appear to be aromatic sources resembling automobile exhaust, gasoline vapor, or environmental tobacco smoke for personal exposures, and automobile exhaust or gasoline vapors for outdoor concentrations. *Journal of Exposure Analysis and Environmental Epidemiology* (2001) **11,** 295–307.

Keywords: *factor analysis, personal exposure, positive matrix factorization, receptor modeling, source apportionment, volatile organic compounds.*

Introduction

Hazardous air pollutants (HAPs) are generally defined as pollutants that are known or suspected to cause cancer or other serious health effects, or to cause harm to the environment (OAQPS, 1998). The 1990 Clean Air Act Amendments (CAAA), Section 112, seek to reduce human exposure to HAPs by defining a statutory list of these compounds. More than 80% of the compounds on the federal HAPs list are toxic volatile organic compounds (VOCs). The main pathways for exposure to toxic VOCs include breathing contaminated air, ingesting contaminated water or soil, and dermal contact. This study focuses on identifying sources of inhalation exposure to toxic VOCs, most of which are on the current HAPs list.

The majority of toxic VOCs in ambient air originate from sources that emit to the outdoors, such as drycleaners, power plants, and vehicle emissions (UATW, 2000). It is unclear, however, whether these sources are the predominant contributors to human exposure. Although not major sources of emissions, personal activities and indoor sources may be the dominant sources of exposure for many compounds (Wallace, 1991). Some examples of sources of personal exposure to toxic VOCs include household cleaners, vehicle exhaust, gasoline vapors, drycleaned clothes, and environmental tobacco smoke (ETS).

With concentration data, receptor modeling may be used to identify source types that contribute significantly to measured concentrations. Although receptor modeling has been widely used for estimating source contributions for particulate matter (PM) air pollution, relatively few receptor modeling studies have been conducted for VOCs. The studies that have been done have focused primarily on estimating source contributions to VOCs in outdoor air rather than to personal exposure. Early examples of receptor modeling applied to VOCs include apportionment of ambient VOCs (Lin and Milford, 1994; Mukund et al., 1996), and the use of acetylene (Lonneman et al., 1974; Whitby and Altwicker, 1978) and aromatic organic compounds (Singh et al., 1985; Edgerton et al., 1989) as indicators of motor vehicle emissions.

The primary goals of this study were to determine sources contributing to personal exposure and outdoor concentrations, and the relative contribution of each source to total concentrations. Personal exposure and outdoor concentration data for residents in Elizabeth and Bayonne, NJ, and Los Angeles (L.A.), Pittsburg, and Antioch, CA from the United States Environmental Protection Agency's (US-EPA) Total Exposure Assessment Methodology (TEAM)

1. Address all correspondence to: Dr. Jana B. Milford, Campus Box 427, The University of Colorado at Boulder, Boulder, CO 80309. Tel.: +1-303-492-5542. Fax: +1-303-492-3498. E-mail: jana.milford@colorado.edu
Received 16 October 2000; accepted 19 March 2001.

and California Air Resources Board (CARB) California Indoor Exposure studies were analyzed using positive matrix factorization (PMF; Paatero and Tapper, 1994).

Methods

Data

The NJ TEAM data used were accessed from the Total Human Exposure Risk database and Advanced Simulation Environment (THERdbASE), version 1.2 (Pandian et al., 1989; NERL, 2000). This database includes data from the main TEAM studies, which measured 24-h exposures of 600 people in NJ, CA, North Dakota (ND) and North Carolina (NC) to various toxic chemicals in air and drinking water. Monitored compounds were selected based on their toxicity, carcinogenicity, and amenability to collection on Tenax sorbent (Wallace, 1987; Wallace et al., 1984). The CA TEAM and CARB data were accessed from the Californian Exposures Database (CED) (Clayton and Perritt, 1993). CARB studies were carried out using the same general procedures as the TEAM studies. A key feature of these studies was the use of a probability-based sampling design to represent the exposures of large populations in the various cities.

Exposures were measured using personal monitors, which collected two 12-h samples, representing overnight and daytime exposures. Concurrent outdoor samples were collected in two 12-h samples from the backyards of a subset of the study homes. Air samples were analyzed by gas chromatography/mass spectrometry. Both studies used time-weighted average exposure concentrations as the measure of exposure. Two questionnaires were administered to participants. At the beginning of each study, household questionnaires were administered regarding age, gender, occupation, smoking status, and customary activities of the participants. Immediately following each study, participants filled out a 24-h activity diary to establish an activity pattern for each subject. These questionnaires provided information needed to identify the likely sources and human activities contributing to exposures.

Models

PMF, a multivariate technique, was used for source apportionment (Paatero and Tapper, 1994). Previous PMF applications include identifying sources of bulk wet deposition concentrations of strong acids in Finland (Antilla et al., 1995), and more recently, sources of PM from the USEPA Particle TEAM study (Yakovleva et al., 1999) and PM in Phoenix urban aerosol (Ramadan et al., 2000). PMF incorporates error estimates of the data to solve matrix factorization of a linear model as a constrained, weighted least-squares problem. These error estimates account for sampling errors, detection limits, missing data observations, and outliers. The input data must be finite, positive numbers. One portion of the model solution is a matrix of factors. These factors, which are roughly interpreted as source profiles, represent the relative amounts of each compound in each source. Because source profiles are spatially and temporally variable, the factor profiles presented are a qualitative guide to the types of sources to which people could be exposed and have been qualitatively related to the best available profiles in the literature. Each factor is constrained to be nonnegative. This requirement decreases the rotational freedom used to produce meaningful factors, and oftentimes, the result is fully unique with no rotational freedom (Paatero, 1998).

PMF was applied to 24-h personal and outdoor concentration data obtained by averaging overnight and daytime samples, as well as to separate daytime and overnight personal concentration data, to investigate potential sources of exposure to toxic VOCs. The compounds included in this study are listed in Table 1. Due to insufficient data collected for participants in ND, NC, and Woodland, CA, data from these locations were not used in this study.

Two-way PMF seeks to solve the matrix equation

$$X_{ij} = \sum_{h=1}^{p} G_{ih} F_{hj} + E_{ij} \quad i = 1, 2, \ldots, n; j = 1, 2, \ldots, m \tag{1}$$

where X_{ij} is an element of an n by m matrix of observed data with n chemical species and m study participants. G_{ih} is an element of an n by p unknown factor matrix, which is interpreted as a matrix of source profiles. F_{hj} is an element of a p by m matrix of unknown factor scores, whose relative magnitudes indicate how much the hth factor contributes to the exposure of the jth participant. E_{ij} is the residual error. The number of factors, p, is chosen by the user. Given p, PMF solves this matrix problem by minimizing the sum of squares, Q:

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{E_{ij}}{S_{ij}} \right)^2 \tag{2}$$

where S_{ij} is the standard deviation representing the uncertainty in the observation X_{ij}. In this study, the X_{ij} are assumed to be lognormally distributed and the values of S_{ij} are iteratively refined so that the solution to Equation 1 approximates a maximum likelihood solution. Further details in the error model specification are given below and by Paatero (1997).

In this study, two-way PMF was applied separately to the personal exposure and outdoor concentration data from NJ and CA.

Three-way PMF seeks to solve a similar matrix equation for an extended set of observations that includes multiple

Table 1. Compounds investigated in this study.

Compound (ABBREV)	Study[a]	IF[c]	Compound (ABBREV)	Study[a]	IF[c]
Benzene (BNZ)[b]	N, C	0.47	*n*-Octane (OCT)	C	NA
Carbon tetrachloride (TET)[b]	N, C	NA	α-Pinene (PIN)	C	4.2
Chlorobenzene (CBZ)[b]	N, C	4.3	Styrene (STR)[b]	N, C	7.8
Chloroform (CFM)[b]	N, C	NA	Tetrachloroethane (PRC)[b]	N, C	NA
n-Decane (DEC)	C	NA	1,1,1-Trichloroethylene (TCA)	N, C	NA
n-Dodecane (DOD)	C	NA	Trichloroethylene (TCE)[b]	N, C	NA
1,2-Dichlorobenzene (ODB)	N	25	*n*-Undecane (UND)	C	NA
1,4-Dichlorobenzene (PDB)[b]	N	25	*o*-Xylene (OXY)[b]	N, C	3.1
1,2-Dichloroethane (DCA)	N, C	NA	*m*-, *p*-Xylenes (MPX)[b]	N, C	3.4
Ethylbenzene (EBZ)[b]	N, C	1.6			

[a]N, included in NJ analysis; C, included in CA analysis.
[b]On current federal HAPs list.
[c]IF, irritancy factor (Ten Brinke, 1995; Kasanen et al., 1998). These values are used to calculate irritancy contributions of factors, which are presented in the Irritancy Contributions section in the text. NA indicates that the IF was not available for that compound.

sample modes, such as corresponding personal and outdoor samples. The three-way factor model can be written as

$$X_{ijk} = \sum_{h=1}^{p} A_{jh} B_{ih} C_{kh} + E_{ijk}$$
$$i = 1, 2, \ldots, n; j = 1, 2, \ldots, m; k = 1, 2, \ldots, q \qquad (3)$$

where i indexes the chemical species, j indexes the study participant, k indexes the sample mode, and h indexes the factors. A_{jh} is an element of the m by p matrix of unknown factor scores, B_{ih} is an element of the n by p unknown factor matrix, and C_{kh} is an element of the q by p matrix that weights the factor contributions between sample modes. Q is calculated as shown in Equation 2.

Three-way PMF was applied here to the subset of TEAM and CARB study participants for which both personal and outdoor concentrations were measured, as well as to daytime and overnight concentrations for personal exposures.

Diagnostic Tools

The Q value (Equation 2), used to estimate the best number of factors, is an indicator of the goodness of fit and a measure of how close the modeled values are to the observed values. Because more of the variation in the data can be explained with more factors, the Q value will generally continue to decrease as the number of factors increases. The additional factors, however, may not be physically meaningful and may just be explaining noise in the data. Thus, other diagnostic measures, as well as judgment of the user, are required to select an appropriate number of factors. Additional diagnostics used include residuals for individual species and for total VOC concentrations.

In the presence of outliers in the data, it can be difficult to determine if the reported Q value is too high (Paatero, 1998). Therefore, the distribution of the scaled residuals for individual compounds may be used as a diagnostic measure as well. Ideally, the residuals should be centered around zero, forming a sharp peak at zero.

A linear regression analysis of the factor scores versus the sum of the measured VOC concentrations was used to help ascertain the optimal number of factors. In addition, the regression was used to determine the relative contribution of each factor to exposure. Multiplying the factor scores (unitless) by the linear regression coefficients (concentration units) gives a total modeled VOC concentration from each factor. Summing these modeled concentrations over all factors gives a total modeled concentration that allows the user to compare the model-predicted total VOC concentration to the measured total VOC concentration for each observation. In addition, the percent contribution of each factor to total VOCs can be determined by dividing the modeled concentration for each factor by the total modeled concentration. The number of factors was selected, in part, to avoid negative regression coefficients. R-squared values and the ratios of modeled to measured concentrations were examined to determine how well the regression model fit the measured data for total VOCs.

Treatment of Missing Data and Uncertainties

Compounds were included in the PMF analysis only if more than 60% of all study participants had values for that compound. Values reported as below the detection limit were included in the analysis to create the largest concentration matrix possible, but were assigned large errors and thus given low weight. Similarly, participants were included only if they had data for at least 60% of the compounds. In this case, missing values for a given compound were filled in using the median value for that compound across all participants. Tables 2 and 3 show that relatively few data points were missing after compounds and participants not meeting the specified criteria were excluded.

Table 2. Summary of the NJ concentration data used in the three-way and two-way PMF models.[a]

Compound	Three-way PMF				Two-way PMF			
	No. data points		GM (GSD), $\mu g/m^3$		No. data points		GM (GSD), $\mu g/m^3$	
	Personal	Outdoor	Personal	Outdoor	Personal	Outdoor	Personal	Outdoor
BNZ	143	122	9.1 (4.3)	6.1 (3.9)	506	135	12 (4.1)	5.8 (4.1)
TET	143	142	1.1 (2.0)	0.87 (2.1)	505	156	1.4 (2.6)	0.87 (2.1)
CBZ	144	142	0.37 (2.2)	0.23 (3.2)	507	155	0.33 (2.2)	0.22 (3.3)
CFM	144	121	2.6 (3.8)	0.85 (6.3)	506	133	2.9 (3.7)	0.88 (6.3)
PDB	144	140	4.9 (5.6)	0.87 (2.4)	506	153	5.4 (5.7)	0.88 (2.4)
ODB	144	140	0.36 (2.9)	0.22 (2.0)	507	153	0.35 (2.8)	0.21 (2.0)
DCA	144	142	0.46 (2.7)	0.22 (2.4)	507	155	0.54 (2.7)	0.22 (2.4)
EBZ	143	140	7.1 (3.1)	2.5 (3.0)	506	152	7.4 (2.9)	2.5 (3.0)
STR	143	142	1.8 (3.0)	0.5 (2.8)	507	155	2.0 (2.7)	0.5 (2.8)
PRC	144	142	6.2 (3.3)	2.9 (3.9)	508	155	7.8 (3.2)	2.8 (4.0)
TCA	143	126	15 (3.5)	4.3 (3.7)	501	138	16 (3.9)	4.4 (3.7)
TCE	144	142	2.9 (2.9)	1.6 (4.7)	508	155	3.1 (3.4)	1.6 (4.7)
OXY	143	141	7.0 (2.8)	2.6 (2.8)	506	154	6.8 (2.6)	2.6 (2.8)
MPX	143	140	18 (2.9)	7.4 (3.0)	507	153	18.5 (2.8)	7.5 (3.0)

[a]The original TEAM documentation (Wallace, 1987) states that 50 out of 157 canisters from the NJ summer analysis in 1982 were contaminated, of which nine were discarded. The concentrations from the remaining 41 canisters were corrected for high background levels. These data were included in the analysis because background corrections were made to the data before they were reported and the TherDBase database does not flag which concentration values from 1982 in NJ are suspect.

One limitation of PMF is its inability to extract factors that fit widely varying exposure concentrations, which is apparent when comparing modeled concentrations from the regression analysis to measured concentrations. When the measured concentrations analyzed include both very high and very low values, the modeled concentrations are generally much higher or lower than the true values. When the measured concentrations are restricted to a smaller range, the modeled concentrations reflect measured concentrations much more closely. Because of this limitation, only participants with total exposure concentrations of less than 2000 $\mu g/m^3$ were included in the NJ personal exposure data analysis, excluding 4% of the participants. This cutoff concentration was chosen because it was an obvious divider

Table 3. Summary of the CA concentration data used in the three-way and two-way PMF models.

Compound	Three-way PMF				Two-way PMF			
	No. data points		GM (GSD), $\mu g/m^3$		No. data points		GM (GSD), $\mu g/m^3$	
	Personal	Outdoor	Personal	Outdoor	Personal	Outdoor	Personal	Outdoor
BNZ	110	117	9.8 (2.2)	4.9 (2.4)	296	130	9.7 (2.1)	4.8 (3.4)
TET	113	118	0.65 (1.6)	0.64 (1.4)	299	131	0.73 (1.7)	0.65 (1.4)
CBZ	105	111	0.057 (2.2)	0.046 (2.0)	287	123	0.063 (2.2)	0.046 (2.0)
CFM	116	116	0.82 (3.3)	0.29 (3.6)	298	129	0.70 (4.1)	0.28 (3.5)
DCA	104	116	0.12 (2.3)	0.067 (2.3)	283	128	0.15 (2.5)	0.068 (2.3)
EBZ	118	119	5.0 (2.3)	2.5 (2.8)	304	132	5.1 (2.6)	2.4 (2.8)
OCT	116	116	3.6 (2.3)	1.3 (2.5)	302	127	3.3 (2.3)	1.3 (2.4)
STR	117	119	1.8 (3.0)	0.75 (3.9)	301	132	1.6 (2.9)	0.74 (3.8)
PRC	108	119	5.0 (3.3)	2.2 (3.4)	291	132	4.9 (3.3)	2.3 (3.3)
TCA	119	118	18 (3.2)	7.5 (2.7)	305	131	18 (3.4)	7.5 (2.6)
TCE	115	113	0.94 (5.3)	0.15 (2.8)	298	125	1.1 (6.2)	0.15 (2.8)
OXY	119	118	6.6 (2.5)	3.5 (2.9)	305	131	6.5 (2.7)	3.5 (2.8)
MPX	119	118	19 (2.2)	9.6 (2.7)	305	131	17 (2.3)	9.5 (2.6)
DEC	111	104	2.6 (3.2)	1.1 (2.4)	295	116	2.2 (3.2)	1.1 (2.3)
DOD	106	111	1.6 (3.4)	0.35 (2.4)	288	123	1.3 (3.1)	0.34 (2.4)
UND	112	114	2.8 (3.2)	0.73 (2.8)	295	126	2.3 (3.1)	0.72 (2.8)
PIN	116	113	2.4 (2.9)	0.31 (2.8)	301	125	2.2 (3.0)	0.51 (0.60)

between the high and low concentrations. No outdoor concentrations were greater than 2000 $\mu g/m^3$. In the CA studies, only one participant had a total exposure concentration greater than 2000 $\mu g/m^3$. Therefore, all participants were included in the analysis for CA.

All data were treated as coming from lognormal distributions, as indicated by examination of histogram plots, means, and medians for the concentration data. The error model used by PMF for lognormally distributed data is:

$$S_{ij} = \left[T_{ij}^2 + \frac{1}{2}V_{ij}^2|Y_{ij}|\left(|Y_{ij}| + |X_{ij}|\right)\right]^{\frac{1}{2}} \quad (4)$$

where Y_{ij} is the fitted concentration value, represents normally distributed measurement error and V_{ij} is associated with inherent randomness in the data. For this study, T_{ij} was determined for each observation using the following formula:

$$\begin{aligned} &\text{if } X_{ij} > 0, T_{ij} = \sqrt{\left[(0.5 \times \mathrm{DL}_i)^2 + (0.1 \times X_{ij})^2\right]} \\ &\text{if } X_{ij} \text{ is missing, } T_{ij} = 2M_i \end{aligned} \quad (5)$$

where DL_i is the detection limit and M_i is the median value for the compound of interest across all observations, as given by Pandian et al. (1989) and Clayton and Perritt (1993). A constant value of 0.1 was used for V_{ij}. For the three-way analysis, S_{ij} is replaced by S_{ijk}, T_{ij} by T_{ijk}, etc. It was found that changing the error estimates had a negligible effect on the factors obtained by the model.

Results

Geometric means (GM) and geometric standard deviations (GSD) of the personal exposure and outdoor concentration data used in the PMF modeling are presented in Table 2 for the NJ data and Table 3 for the CA data. NJ and CA average personal exposure concentrations for all compounds are higher than outdoor concentrations.

Standard deviations of the factors are plotted as error bars on all source apportionment results. These values are based on a global least squares fit, in which all three matrices, **A**, **B** and **C**, are determined simultaneously (Paatero, 1998). In general, the standard deviation matrices represent both individual random uncertainty in the factor elements and uncertainty due to factor rotation. If there is negligible rotational freedom, however, the error estimates in **A**, **B**, and **C** reflect only the random uncertainty in the factor elements, and the values in the standard deviation matrices will be small (Paatero, 1998). Although error bars are plotted for each compound in the factor profiles, some errors are small enough such that they are not visible, suggesting negligible rotational freedom of the factors.

NJ Factorization

Factorization was performed using from four to nine factors for both the two-way and three-way PMF data sets. Six factors and five factors were chosen as the optimal number of factors for the personal and outdoor two-way PMF data sets, respectively. Eight factors were chosen for the three-way PMF model. Adding more factors slightly improved the Q values and R-squared values, but investigation of the resulting factor profiles revealed that the additional factors were not easily interpreted and were likely just explaining noise in the data. Factor profiles and modal associations for the three-way model are shown in Figure 1. The right-hand side of each bar graph shows the relative contribution of each factor to personal exposure (PER) or outdoor (OUT) concentrations, and also to personal daytime (PER DT) versus personal overnight (PER ON) exposures. In all factor profiles, the height of the bars can only be compared within a given factor, and not across all factors, because the sum of chemical contributions in a factor is normalized to one.

The factors obtained from the three-way PMF analysis generally matched the factors obtained from the two-way PMF analysis. Both analyses showed similar contributions of factors to total concentrations. This occurs despite the fact that the three-way model was applied to only a subset of the participants included in the two-way models. Factor 6p, however, did not have an equivalent three-way factor, nor was there an equivalent factor to the three-way Factor 3 in the two-way outdoor model. This is likely due to the use of different data sets in the two-way and three-way models.

In addition to 24-h personal and outdoor data, 12-h daytime and overnight personal observations were modeled using three-way PMF. Daytime sources were generally the same as overnight sources, though the sources contributed differently to exposure. Modal associations for daytime and overnight personal factors are shown in Figure 1.

Figure 2 shows examples of the residuals for benzene and TCE for the personal and outdoor modes of the three-way PMF model. If the majority of the scaled residuals (E_{ij}/S_{ij}) are between ± 2 standard deviations (σ), the fit of the model is acceptable (Paatero, 1998). The percentages of residuals within $\pm 2\sigma$ are shown in Table 4.

To estimate the contributions of factors to total personal exposure and outdoor concentrations (summed over all VOCs), the observations of these values were regressed against the factor scores produced by the PMF model. Figure 3 shows an example of the factor scores for each NJ participant for Factor 1 in the three-way model. The regression analysis results, showing the average source contributions to personal exposure and outdoor concentrations for both the two-way and three-way NJ PMF data sets, are presented in Table 5. For example, Factor 3

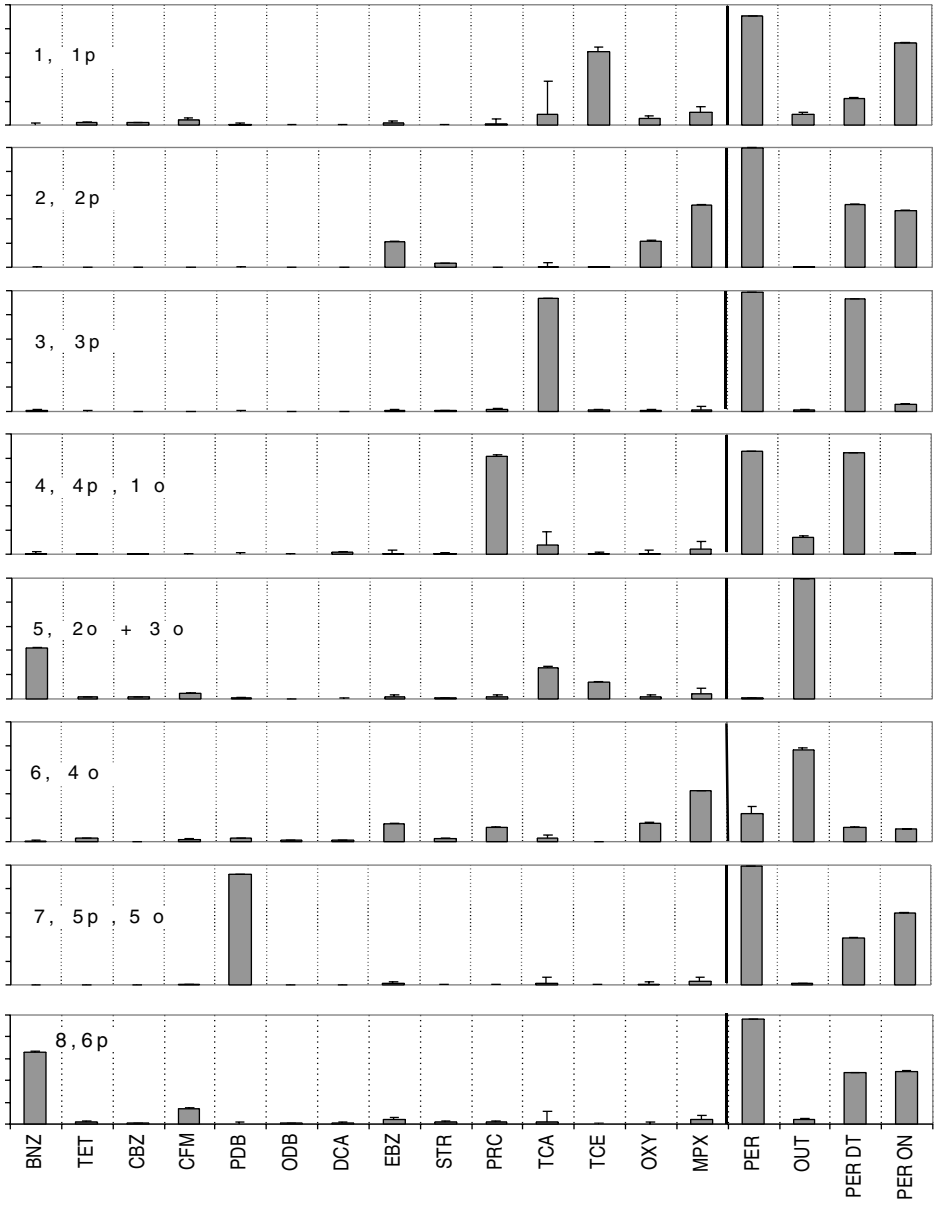

Figure 1. Source apportionment for the NJ three-way PMF model. The *y* axes represent normalized concentrations with arbitrary units. The labels on the left side of each plot indicate which factors were also seen in the two-way personal (p) or outdoor (o) models.

contributed, on average, 22% to participants' total 24-h personal VOC exposure, whereas Factor 4 contributed approximately half as much (11%) to total exposure. Averaged across all of the study participants, Factors 1 and 2 contribute most to total personal exposures, whereas Factors 5 and 6 contribute most to total outdoor concentrations of the 14 VOCs studied.

The measured concentrations were well predicted by the regression. The personal modeled concentrations were, on average, slightly overpredicted at 110±40% and 105±12% of measured concentrations for the three-way and two-way models, respectively. The predicted outdoor concentrations were approximately 100±32% and 97±22% of measured concentrations for the three-way and two-way models, respectively.

CA Factorization

Factorization was attempted for the CA data using from 4 to 11 factors for both the two-way and three-way PMF data sets. Eight factors and three factors were chosen as the optimal number of factors for the personal and outdoor two-way PMF data sets, respectively. Nine factors were chosen for the three-way PMF model. Factor profiles and modal associations for the three-way model are shown in

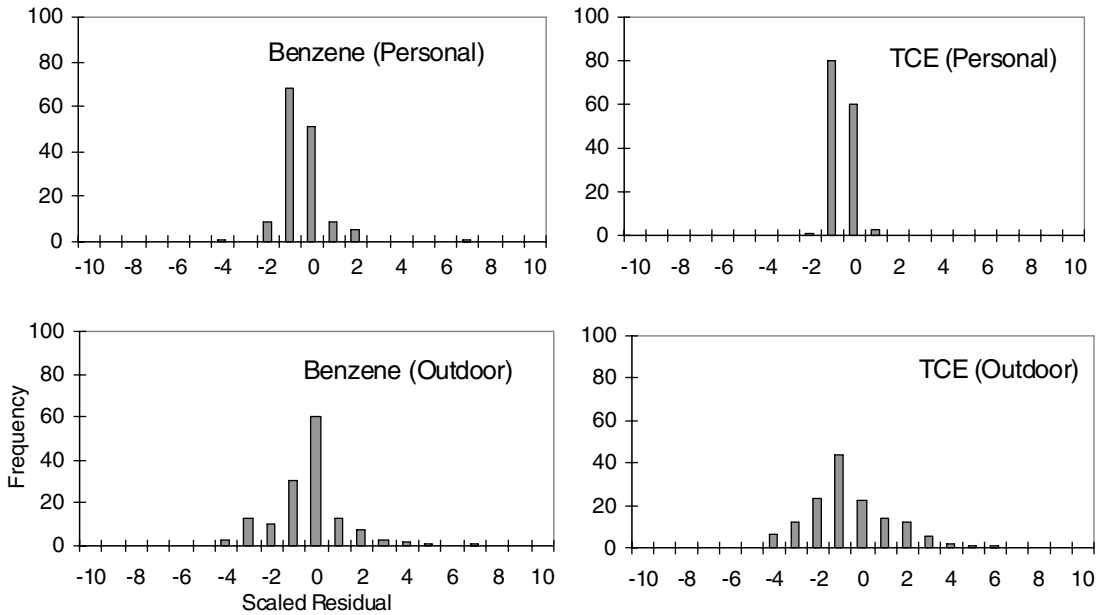

Figure 2. Examples of frequency distributions of residuals scaled by standard deviations for the NJ three-way PMF solutions for personal and outdoor concentrations of benzene and TCE.

Figure 4. As with the NJ data, the factors obtained from three-way PMF generally matched the factors obtained from the two-way PMF analysis. Factor 1 from the three-way model, however, did not appear in the personal two-way model. The percentages of residuals within $\pm 2\sigma$ are shown in Table 4.

Table 4. Percentages of NJ and CA residuals within $\pm 2\sigma$.

Compound	NJ[a]				CA[a]			
	Three-way		Two-way		Three-way		Two-way	
	Personal (%)	Outdoor (%)	Personal (%)	Outdoor (%)	Personal (%)	Outdoor (%)	Personal (%)	Outdoor
BNZ	95	79	88	100	75	78	70	80
TET	92	97	82	96	91	79	74	79
CBZ	92	92	91	95	94	98	93	97
CFM	73	83	69	82	55	74	100	83
PDB	99	88	100	100	NA	NA	NA	NA
ODB	86	98	86	97	NA	NA	NA	NA
DCA	92	98	94	99	82	93	75	92
EBZ	92	99	94	99	86	75	91	76
OCT	NA	NA	NA	NA	58	74	69	78
STR	70	90	63	92	74	67	63	64
PRC	97	67	100	100	76	73	100	74
TCA	100	90	100	94	94	82	100	83
TCE	100	72	100	95	100	96	100	95
OXY	94	98	96	99	95	94	95	95
MPX	93	98	96	99	87	95	91	95
DEC	NA	NA	NA	NA	77	81	84	81
DOD	NA	NA	NA	NA	63	92	68	91
UND	NA	NA	NA	NA	92	80	85	80
PIN	NA	NA	NA	NA	99	84	100	83

[a]NA, not applicable; compound was not included in study.

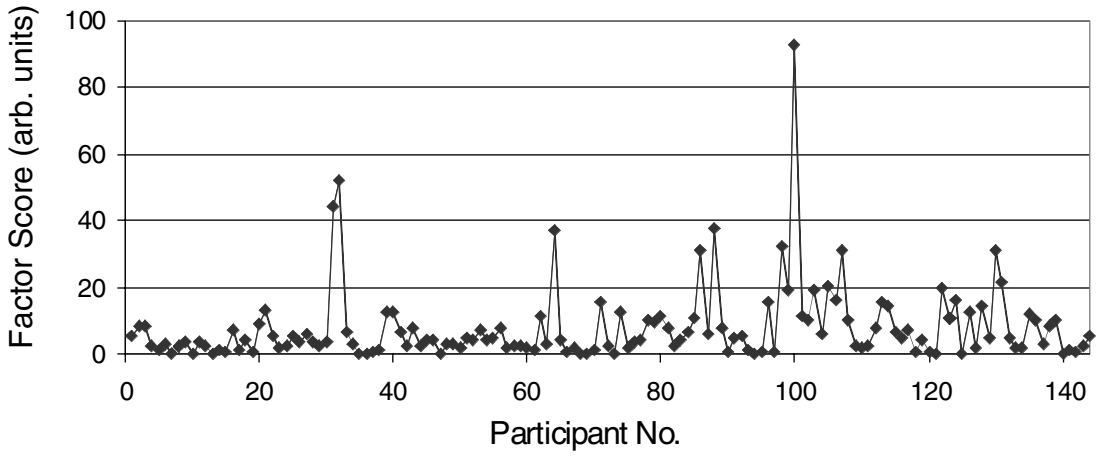

Figure 3. Factor scores for Factor 1 (personal TCE) for each participant included in the NJ three-way model.

Daytime and overnight personal observations were modeled using three-way PMF for the CA data as well. Modal associations for the daytime and overnight personal factors are shown in Figure 4. Factor 6 from the three-way model did not have an equivalent daytime/overnight factor.

The regression analysis results for the two-way and three-way CA PMF models are presented in Table 6. Averaged across all of the study participants, Factors 1, 5, and 9 contribute most to personal exposure and Factor 7 contributes most to outdoor concentrations of the 17 VOCs studied.

The measured concentrations were well predicted by the regression. The personal modeled concentrations were, on average, slightly overpredicted at 120±41% and 100±11% of measured concentrations for the three-way and two-way models, respectively. The predicted outdoor concentrations were approximately 104±18% and 102±13% of measured concentrations for the three-way and two-way models, respectively.

Discussion

Interpretation of Factor Profiles

Factors identified in three-way PMF were interpreted as 10 different source types contributing to NJ and CA VOC concentrations that were reflected in either one or both of the data sets. The interpretation was made on the basis of qualitative comparison to source or exposure profiles reported in the literature, as detailed below. Note that precise matches were not expected due to variability in the reported profiles and the fact that VOCs undergo chemical degradation from the time they are emitted. The sources are: contaminated water (TCE); solvents (TCA); gasoline vapors, automobile exhaust, or ETS (aromatics,

Table 5. Average and standard deviations of source contributions in NJ for 24-h personal and outdoor data sets analyzed using three-way and two-way PMF, and 12-h personal daytime and overnight data sets analyzed using three-way PMF.[a]

Three-way (n=144/528)[b]					Two-way (n=508/528)[b]		Two-way (n=158/528)[b]	
Factor	Personal			Outdoor	Factor	Personal	Factor	Outdoor
	24-h	12-h (daytime)	12-h (overnight)	24-h		24-h		24-h
1	27±21%	15±15%	23±14%	NS	1p	5±8%	–	–
2	27±23%	28±19%	40±19%	NS	2p	32±19%	–	–
3	22±19%	12±14%	7±11%	6±11%	3p	20±18%	–	–
4	11±13%	11±13%	NS	NS	4p	9±10%	1o	14±13%
5	NS	–	–	48±25%	–	–	2o, 3o	29±20%(2o), 23±21%(3o)
6	NS	–	–	46±23%	–	–	4o	33±22%
7	13±21%	9±17%	11±20%	NS	5p	13±21%	5o	NS
8	NS	26±18%	20±15%	NS	6p	21±14%		–

[a]NS, not significant; "–" indicates that factor did not appear in mode.
[b]n indicates the number of participants included in the analysis out of the total number possible.

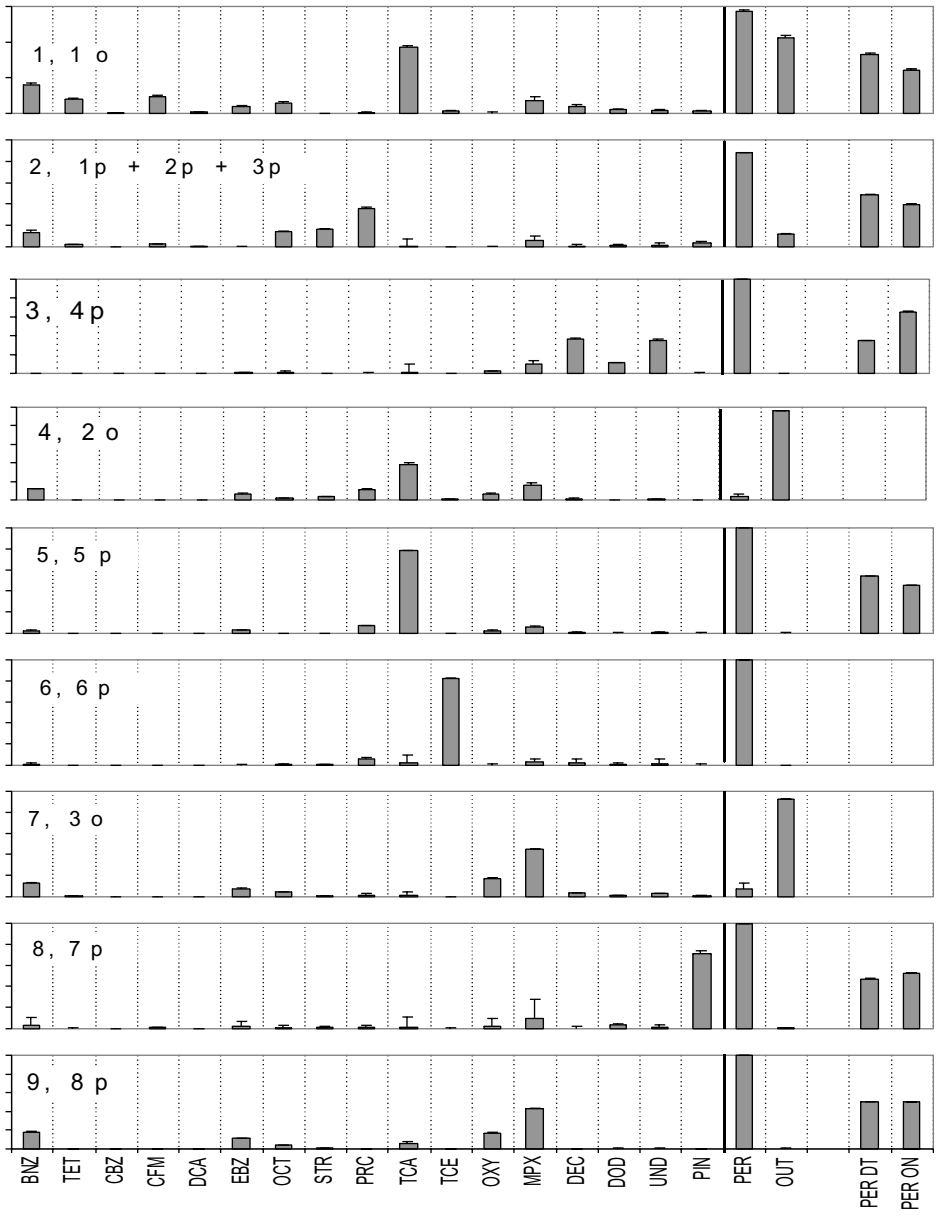

Figure 4. Source apportionment for the CA three-way PMF model. The *y* axes represent normalized concentrations with arbitrary units.

including benzene, ethylbenzene, and xylenes); wastewater treatment plant emissions (chloroform, TCA, TCE, benzene, xylenes); drycleaning chemicals (PERC); deodorizers or mothballs (*p*-DCB); consumer products (PERC, styrene, benzene, *n*-octane); building materials (alkanes and xylenes); background ambient concentrations (TCA, benzene, ethylbenzene, PERC, xylenes); and cleaners (α-pinene). These sources were, in general, confirmed by two-way PMF results. A summary of the factorization results for the NJ and CA PMF modeling is presented in Table 7.

The contaminated water factor, comprised of mainly TCE, may be a reflection of exposure to contaminated drinking and showering water or solvents (McKone, 1987; Wallace et al., 1989). The main exposure route to TCE appears to be from personal activities, as evidenced by Factors 1 and 6 from the NJ and CA three-way models, and the corresponding Factors 1p and 6p from the NJ and CA personal two-way models.

The TCA source likely reflects exposure to solvents, household cleaners, and household pesticides (Wallace et al., 1989). The main exposure route to TCA, as evidenced by NJ Factors 3 and 3p, and CA Factors 5 and 5p, is due to personal activities as well.

The aromatics factor, resembling ETS (Daisey et al., 1994), automobile exhaust or gasoline vapors (Rappaport

Table 6. Average and standard deviations of source contributions in CA for 24-h personal and outdoor data sets analyzed using three-way and two-way PMF, and 12-h personal daytime and overnight data sets analyzed using three-way PMF.[a]

Three-way (n=119/447)[b]					Two-way (n=305/447)[b]		Two-way (n=132/447)[b]	
Factor	Personal			Outdoor	Factor	Personal[c]	Factor	Outdoor[c]
	24-h	12-h (daytime)	12-h (overnight)	24-h		24-h		24-h
1	30±26%	NS	NS	25±21%	–	–	1o	23±22%
2	22±15%	12±11%	13±11%	NS	1p, 2p, 3p	8±8%, 11±6%, 3±6%	–	–
3	5±6%	4±11%	5±9%	NS	4p	11±13%	–	–
4	NS	–	–	35±21%	–	–	2o	35±21%
5	31±31%	31±20%	27±15%	NS	5p	18±15%	–	–
6	NS	–	–	NS	6p	4±9%	–	–
7	NS	–	–	41±19%	–	–	3o	42±20%
8	NS	NS	7±8%	NS	7p	5±7%	–	–
9	32±24%	52±20%	48±23%	NS	8p	40±17%	–	–

[a]NS, not significant; "–" indicates that factor did not appear in mode.
[b]n indicates the number of participants included in the analysis out of the total number possible.
[c]Source contributions of a few factors varied significantly by location and season. The source contribution from Factor 2p was slightly higher in Pittsburg/Antioch (14±7%, n=67) than in L.A. (11±6%, n=238). The contribution from 5p was slightly higher in L.A. (19±16%) than in Pittsburg/Antioch (13±11%). The differences between the contributions of all factors to outdoor concentrations in Pittsburg/Antioch and L.A. were statistically different, but data were only available from nine outdoor samples for Pittsburg/Antioch. The source contribution from Factor 4p was slightly higher in the summer (14±16%, n=89) than in the winter (10±11%, n=149). The source contribution from Factor 1o was approximately twice as high in the summer (26±19%, n=61) as in the winter (14±14%, n=62), whereas the contribution from 2o was approximately twice as high in the winter (43±23%) as in the summer (28±15%). The relative contributions from all other factors were not statistically different at the α=0.05 level, based on season or on location.

et al., 1987; Scheff et al., 1991; MacIntosh et al., 1995; Wallace, 1996), is present as both a personal exposure and outdoor source in NJ and CA. Exposure to ETS could result from smoking or being proximal to a smoker. Personal exposure to gasoline or exhaust could result from activities such as sitting in traffic, pumping gasoline, or from being exposed to outdoor air with automobile exhaust or gasoline vapors that has penetrated indoors. NJ Factors 2 and 2p, and 6 and 4o consist mainly of ethylbenzene and xylenes. Benzene, a major component of gasoline, automobile exhaust and ETS, is not included in these factors. It is, however, the dominant component in other NJ personal and outdoor factors (Factors 8 and 6p, and Factors 5 and 2o, respectively). This suggests that benzene is present in

Table 7. Summary of factorization results for NJ and CA two-way and three-way PMF modeling. Only the dominant components for each factor are listed, although others are present in most factors.[a]

Factors	Sources	NJ				CA			
		Three-way PMF		Two-way PMF		Three-way PMF		Two-way PMF	
		Personal	Outdoor	Personal	Outdoor	Personal	Outdoor	Personal	Outdoor
TCE	Contaminated water	1	–	1p	NA	6	–	6p	NA
TCA	Solvents	3	–	3p	NA	5	–	5p	NA
BNZ,[b] EBZ, xylenes	ETS,[c] automobile exhaust, gasoline vapors	2, 8	6	2p, 6p	4o	9	7	8p	3o
CFM, TCA, TCE, xylenes, BNZ	Wastewater treatment	–	5	NA	2o, 3o	1	1	NA	1o
PRC	Drycleaners	4	4	4p	1o	NA	NA	NA	NA
PDB	Mothballs, deodorizers	7	–	5p	5o	NA	NA	NA	NA
BNZ, OCT, STR, PRC	Consumer products	NA	NA	NA	NA	2	–	1p, 2p, 3p	NA
DEC, DOD, UND, MPX	Building materials, carpeting	NA	NA	NA	NA	3	–	4p	NA
TCA, BNZ, EBZ, PRC, xylenes	Ambient air	NA	NA	NA	NA	–	4	NA	2o
PIN	Cleaners	NA	NA	NA	NA	8	–	7p	NA
Q value		8620	8620	16308	2191	13585	13585	14351	6170
R^2 value		0.77	0.88	0.97	0.92	0.88	0.98	0.98	0.98
Total number of factors		8	8	6	5	9	9	8	3

[a]NA, not applicable (factor was not in PMF results), "–" indicates that factor did not appear in mode.
[b]In NJ, benzene appears in a separate factor from ethylbenzene and xylenes.
[c]ETS is likely only a personal source.

widely varied concentrations in several sources. PMF, to account for the variability in the data, might make benzene a separate source. As shown with CA Factors 9 and 8p, and Factors 7 and 3o, an aromatics factor consisting of benzene, ethylbenzene, and xylenes was a personal and outdoor source in the CA data as well.

The outdoor source comprised of chloroform, TCA, xylenes, and TCE resembles wastewater treatment plant emissions (Scheff et al., 1989). This source is evident with NJ Factor 5, which appears to be a combination of Factors 2o and 3o in the two-way model, and CA Factor 1, which affects both personal exposure and outdoor concentrations. This factor did not appear in the CA two-way personal model, but was present in the two-way outdoor results, as shown by Factor 1o.

The PERC source could result from exposure to drycleaned clothes, or being inside or living near a drycleaning shop (Wallace, 1989; Wallace et al., 1989). This source was only extracted from the NJ data, although PERC was measured in CA too. As shown in NJ Factors 4, 4p and 1o, the PERC source affects both personal exposure and outdoor concentrations.

The *p*-DCB source is most likely a reflection of exposure to mothballs and room deodorizers (Wallace, 1989). Because this compound was not included in the CA modeling due to a limited number of measurements, the *p*-DCB source appeared only in the NJ data. NJ Factors 7 and 5p show that *p*-DCB primarily affects personal exposures. A *p*-DCB source was present in the NJ two-way PMF outdoor results, although it was not significant to outdoor concentrations.

The source resembling consumer products (Samfield, 1992; Sheldon et al., 1992), such as adhesives and aerosols, is comprised of PERC, styrene, benzene, and *n*-octane. This source, appearing only in the CA results, was reflected in CA Factor 2, a personal exposure factor that appears to be a combination of Factors 1p, 2p and 3p from the two-way model.

The building material and carpeting source (Mølhave, 1982; Wallace, 1987), composed of higher alkanes (*n*-decane, *n*-undecane, *n*-dodecane) and xylenes, was extracted from only the CA data as well, as shown in CA Factors 3 and 4p. These alkanes were not measured in the NJ study.

As shown in Factors 4 and 2o, the outdoor source comprised of TCA, benzene, ethylbenzene, PERC and xylenes is present only in the CA results. This factor resembles background outdoor concentrations of these compounds (Rosenbaum et al., 1999).

The last source, extracted only from the CA data, is comprised of α-pinene, which is most likely a reflection of exposure to deodorizers and cleaners (Apte and Daisey, 1999). Factors 8 and 7p show that α-pinene primarily affects personal exposures. This compound was not measured in the NJ study.

Activity Patterns

Activity pattern data from questionnaires administered in the CARB and TEAM studies were analyzed along with the personal factor scores to determine if factors with high scores were reflected in the reported activities. For example, we expected that participants who reported having smoked or pumped gasoline should have higher scores for factors dominated by aromatics than those who did not report being exposed to smoke or gasoline. Although many activities were analyzed, only a subset is discussed here.

Participants in NJ who reported visiting the drycleaners (4 out of 321 participants) had factor scores for the personal PERC factor that were approximately twice as high as those who did not report visiting the drycleaners. It should be noted, however, that the sample size for NJ is likely not large enough to produce conclusive evidence as to exposure from being at the drycleaners. In addition, significant exposure can arise from wearing or being around drycleaned clothes. Therefore, the highest exposures may have occurred outside of the drycleaners (Wallace, 1989).

Three hundred five out of 321 NJ participants and 282 out of 311 CA participants reported drinking tap water, as opposed to bottled water. These participants had approximately four times higher factor scores for the TCE factor (Factor 1p in NJ and Factor 6p in CA), which is thought to reflect exposure to contaminated water. CA participants who reported having showered or bathed (204 out of 234 participants) had two times higher factor scores for Factor 6p. This activity was not reported for the NJ participants.

CA participants who reported having used paint or solvents (16 out of 94 participants) had approximately five times higher factor scores for the 1,1,1-TCA factor (Factor 5p) and seven times higher factor scores for the alkanes factor (Factor 4p). This activity was not reported for the NJ participants.

The NJ participants who reported having smoked or lived with smokers in NJ (224 out of 321 participants) had approximately 85% higher factor scores for the aromatics factor (Factor 2p) and 40% higher factor scores for the benzene factor (Factor 6p) than people who neither smoked nor lived with smokers. Those who reported having pumped gasoline during the study (22 out of 321 participants) showed 80% higher factor scores than those who did not pump gasoline for Factor 2p and 20% higher factor scores for Factor 6p. The fact that higher factor scores for the aromatics factors are linked to both ETS and pumping gasoline may be evidence of PMF having difficulty separating sources with similar profiles.

The CA participants who reported having smoked or lived with smokers (145 out of 305 participants) did not have elevated factor scores for the aromatics factor compared to those who did not smoke or live with smokers. This suggests that either ETS was not strongly reflected as a source of exposure in the PMF factors or that most

participants received similar levels of exposure to ETS, whether they reported exposure to ETS or not. A study published in 1996 by Pirkle et al. (1996) indicates widespread exposure to ETS in the U.S. Of American nonsmokers tested, 88% showed evidence of exposure to ETS over the previous 1 to 2 days.

Forty-eight out of 305 CA participants reported having pumped gasoline and 50 out of 93 reported having an attached garage, but neither group showed an increase in the factor scores, on average, for the aromatics factor (Factor 8p). Exposure to gasoline or automobile exhaust, therefore, may not have been a strong personal source of exposure for CA participants or, as with ETS, it may have been a common source of exposure that most participants received.

Irritancy Contributions

Tables 5 and 6 show mass-based contributions of various sources to exposures. Because this research focuses on exposure to toxic VOCs, a preliminary analysis using irritancy-based contributions of sources to exposures was carried out. Mass fractions of each factor were scaled by unitless irritancy factors (IFs) to determine contributions of each source to total irritancy for a given participant. IFs are calculated using RD_{50} values (Ten Brinke, 1995; Kasanen et al., 1998), or the concentration of airborne irritant causing a 50% decrease in the respiratory rate of an animal. The RD_{50} for toluene is divided by the RD_{50} for the compound of interest (Ten Brinke, 1995). Compounds for which IFs were available (Ten Brinke, 1995; Kasanen et al., 1998) and that were included in this calculation are indicated in Table 1. Using these IFs, the greatest irritancy-weighted contribution to exposure in NJ is from Factor 5p (*p*-DCB), which contributes an average of 48% of irritancy. In CA, the personal aromatics factor resembling either ETS, gasoline vapors or automobile exhaust (Factor 8p) contributes the greatest amount to irritancy, on average, with 66% of the total contribution. A factor with a similar profile (Factor 2p) is the second highest contributor to irritancy in NJ, averaging 37% of the total contribution, followed by the benzene source (Factor 6p), with 12% of irritancy. The second highest contributor to irritancy in CA is α-pinene (Factor 7p), with 12% of irritancy, followed by the consumer products factor (Factor 2p), with 11%. All other source contributions to total irritancy were small (0.5–7%) compared to these factors.

IFs were available for only 8 out of 14 and 8 out of 17 compounds studied in the NJ and CA analyses, respectively. Therefore, this analysis serves primarily to illustrate alternate methods of analyzing source apportionment results. IFs are only one of several methods of scaling source contributions to analyze different aspects or consequences of exposure to chemicals. For example, cancer slope factors or reference doses could be used to analyze chronic cancer risk or acute risk, respectively.

Conclusions

The main sources of personal exposure to toxic VOCs in NJ and CA in the mid to late 1980s appear to be ETS and/or automobile exhaust or gasoline vapors, followed by TCA sources, such as solvents and cleaning agents. Good agreement was found between reported activity patterns and factor scores for pumping gasoline, being exposed to ETS, drinking tap water, and visiting the drycleaners in NJ, and drinking tap water, showering, and using paints and solvents in CA.

We are continuing this study by analyzing a synthetic data set with known sources and contributions to test the robustness of PMF as a source-apportionment technique. We will also be analyzing the NJ, CA and synthetic data sets with the principal components analysis (PCA), chemical mass balance (CMB), and GRACE/SAFER (Henry et al., 1994) models. The results from all models will be compared.

Acknowledgments

This research was supported by U.S. EPA grant number R82-6788-010. The PMF model was used under a licensing agreement with Pentti Paatero of the University of Helsinki. We are grateful to Eileen Daly and Bill Oliver for their assistance with the statistical analysis of the data.

Although the research described in this article has been funded by the U.S. EPA, it has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References

Antilla P., Paatero P., Tapper U., and Järvinen O. Source identification of bulk wet deposition in Finland by positive matrix factorization. *Atmos Environ* 1995: 29: 1705–1718.

Apte M.G., and Daisey J.M. VOCs and "Sick Building Syndrome": application of a new statistical approach for SBS research to U.S. EPA base study data. *Indoor Air* 1999: 177–122.

Clayton C.A., and Perritt R.L. Data base development and data analysis for California Indoor Exposure Studies. California Air Resources Board, 1993, pp. A133–187.

Daisey J.M., Mahanama K.R.R., and Hodgson A.T. Toxic volatile organic compounds in environmental tobacco smoke: emissions factors for modeling exposures of California populations. California Air Resources Board, 1994, pp. A133–186.

Edgerton S.A., Holdren M.W., Smith D.L., and Shah J.J. Inter-urban comparison of ambient volatile organic compounds concentrations in U.S. cities. *J. Air Waste Manage Assoc* 1989: 39: 729–732.

Henry R.C., Lewis C.W., and Collins J.F. Vehicle-related hydrocarbon source compositions from ambient data: the GRACE/SAFER method. *Environ Sci Technol* 1994: 28: 823–832.

Kasanen J.P., Pasanen A.L., Pasanen P., Liesivuori J., Kosma V.M., and Alarie Y. Stereospecificity of the sensory irritation receptor for nonreactive chemicals illustrated by pinene enantiomers. *Arch Toxicol* 1998: 72: 514–523.

Lin C., and Milford J.B. Decay-adjusted chemical mass balance receptor modeling for volatile organic compounds. *Atmos Environ* 1994: 28: 3261–3276.

Lonneman W.A., Kopczynski S.L., Darley P.E., and Sutterfield F.D. Hydrocarbon composition of urban air pollution. *Environ Sci Technol* 1974: 8: 229–236.

MacIntosh D.L., Xue J., Ozkaynak H., Spengler J.D., and Ryan P.B. A population-based exposure model for benzene. *J Exposure Anal Environ Epidemiol* 1995: 5: 375–403.

McKone T. Human exposure to volatile organic compounds in household tap water: the indoor inhalation pathway. *Environ Sci Technol* 1987: 21: 1194–1201.

Mølhave L. Indoor air pollution due to organic gases and vapours of solvents in building materials. *Environ Int* 1982: 8: 117–127.

Mukund R., Kelly T.J., and Spicer C.W. Source attribution of ambient air toxic and other VOCs in Columbus, Ohio. *Atmos Environ* 1996: 30: 3457–3470.

National Exposure Research Laboratory (NERL), USEPA. THERdbASE Exposure Assessment Software. http://www.epa.gov/nerl/heasd/therd-home.htm (accessed July 5, 2000).

Office of Air Quality Planning and Standards (OAQPS), USEPA. Taking toxics out of the air: progress in setting "maximum achievable control technology" standards under the Clean Air Act. EPA/451/K-98-001, 1998.

Paatero P. Least squares formulation of robust non-negative factor analysis. *Chemom Intell Lab Syst* 1997: 37: 15–35.

Paatero P. User's guide for positive matrix factorization programs PMF2 and PMF3. 1998. ftp://rock.helsinki.fi/pub/misc/pmf/.

Paatero P., and Tapper U. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 1994: 5: 111–126.

Pandian M.D., Bradford J., and Behar J.V. THERDBASE: total human exposure relational database. Proceedings of the EPA/AWMA Specialty Conference: Total Exposure Assessment Methodology, A New Horizon. Air & Waste Management Association, Pittsburgh, PA, 1989, pp. 204–209.

Pirkle J.L., Flegal K.M., Bernett J.T., Brody D.J., Etzel R.A., and Maurer K.R. Exposure of the U.S. population to environmental tobacco smoke. *J Am Med Assoc* 1996: 275: 1233–1240.

Ramadan Z., Song X.-H., and Hopke P.K. Identification of source of Phoenix aerosol by positive matrix factorization. *J Air Waste Manage Assoc* 2000: 50: 1308–1320.

Rappaport S.M., Selvin S., and Waters M.A. Exposures to hydrocarbon components of gasoline in the petroleum industry. *Appl Ind Hyg* 1987: 2: 148–154.

Rosenbaum A.S., Axelrad D.A., Woodruff T.J., Wei Y., Ligocki M.P., and Cohen J.P. National estimates of outdoor air toxics concentrations. *J Air Waste Manage Assoc* 1999: 49: 1138–1152.

Samfield M.M. Indoor air quality data base for organic compounds. EPA-600-R-92-025, 1992.

Scheff P.A., Porter J.A., and Doskey P.V. Improvement of VOC source fingerprints for vehicles and refineries. Paper No. 91–79.5, Air & Waste Management Association, 1991.

Scheff P.A., Wadden R.A., Bates B.A., and Aronian P.F. Source fingerprints for receptor modeling of volatile organics. *J Air Pollut Control Assoc* 1989: 39: 469–478.

Sheldon L., Clayton A., Jones B., Keever J., Perritt R., Smith D., Whitaker D., and Whitmore R. Indoor pollutant concentrations and exposures. California Air Resources Board (CARB), A833-156 1992, pp. 9.39–9.47.

Singh H.B., Salas L.J., Cantrell B.K., and Redmond R.M. Distribution of aromatic hydrocarbons in the ambient air. *Atmos Environ* 1985: 19: 1911–1919.

Ten Brinke J. Development of new VOC exposure metrics and their relationship to "sick building syndrome" symptoms. Lawrence Berkeley Laboratory, LBL-37652, 1995.

Unified Air Toxics Website (UATW), USEPA. Unified air toxics website: the pollutants. http://www.epa.gov/ttn/uatw/ (accessed June 30, 2000).

Wallace L.A. The Total Exposure Assessment Methodology (TEAM) study: summary and analysis: vol. I. EPA/600/6-87/002a, 1987.

Wallace L.A. The Total Exposure Assessment Methodology (TEAM) study: an analysis of exposures, sources and risks associated with four volatile organic chemicals. *J Am Coll Toxicol* 1989: 8: 883–895.

Wallace L.A. Comparison of risk from outdoor and indoor exposure to toxic chemicals. *Environ Health Perspect* 1991: 95: 7–13.

Wallace L.A. Environmental exposure to benzene: an update. *Environ Health Perspect* 1996: 104: 1129–1135.

Wallace L.A., Pellizari E.D., Hartwell T., Rosenzweig M., Erickson M., Sparacino C., and Zelon H. Personal exposure to volatile organic compounds: I. Direct measurement in breathing-zone air, drinking water, food, and exhaled breath. *Environ Res* 1984: 35: 293–319.

Wallace L.A., Pellizari E.D., Hartwell T.D., Davis V., Michael L.C., and Whitmore R.W. The influence of personal activities on exposure to volatile organic compounds. *Environ Res* 1989: 50: 37–55.

Whitby R.A., and Altwicker E.R. Acetylene in the atmosphere: sources, representative ambient concentrations, and ratios to other hydrocarbons. *Atmos Environ* 1978: 12: 1289–1296.

Yakovleva E., Hopke P.K., and Wallace L. Receptor modeling assessment of Particle Total Exposure Assessment Methodology data. *Environ Sci Technol* 1999: 33: 3645–3652.