

This discussion paper is/has been under review for the journal Atmospheric Chemistry and Physics (ACP). Please refer to the corresponding final paper in ACP if available.

# Can Positive Matrix Factorization identify sources of organic trace gases at the continental GAW site Hohenpeissenberg?

M. Leuchner<sup>1,4</sup>, S. Gubo<sup>1</sup>, C. Schunk<sup>1</sup>, C. Wastl<sup>1,\*</sup>, M. Kirchner<sup>2</sup>, A. Menzel<sup>1,4</sup>,  
and C. Plass-Dülmer<sup>3</sup>

<sup>1</sup>Fachgebiet für Ökologiklimatologie, Technische Universität München,  
Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising, Germany

<sup>2</sup>Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH),  
Cooperation Group of Comprehensive Molecular Analytics, Ingolstädter Landstraße 1, 85764  
Neuherberg, Germany

<sup>3</sup>Meteorologisches Observatorium Hohenpeissenberg, Deutscher Wetterdienst,  
Albin-Schwaiger-Weg 10, 82383 Hohenpeissenberg, Germany

<sup>4</sup>Institute for Advanced Study, Technische Universität München, Lichtenbergstraße 2a, 85748  
Garching, Germany

\*now at: Zentralanstalt für Meteorologie und Geodynamik, Hohe Warte 38, 1190 Wien, Austria

Can PMF identify  
sources of organic  
trace gases at a GAW  
site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Received: 24 February 2014 – Accepted: 7 March 2014 – Published: 25 March 2014

Correspondence to: M. Leuchner (leuchner@wzw.tum.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

ACPD

14, 8143–8183, 2014

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Abstract

From the rural Global Atmosphere Watch (GAW) site Hohenpeissenberg in the pre-alpine area of Southern Germany, a dataset of 24 C<sub>2</sub>–C<sub>8</sub> non-methane hydrocarbons over a period of seven years was analyzed. Receptor modeling was performed by Positive Matrix Factorization (PMF) and the resulting factors were compared to literature source profiles. Photochemical aging during transport to the relatively remote site violates the PMF prerequisite of mass conservation from source to receptor. However, previous studies showed plausible results with this method at remote sites; the applicability and restrictions of the PMF model to such a remote dataset and the influence of photochemical processing on the interpretability of the results are discussed. A six factor solution showed a high stability and the most plausible results. In addition to biogenic sources and remote sources of very stable compounds – reflecting the continental background – four additional anthropogenic factors were resolved that could be divided into two short- and two long-lived patterns from evaporative sources and incomplete combustion processes, respectively. A method to increase the uncertainty for each individual compound by including photochemical reactivity did not improve the results, but decreased the stability of the model output. The contribution of the different source categories at the site over the entire period was, in decreasing order: remote sources, long-lived evaporative sources, residential heating and long-lived combustion sources, short-lived evaporative sources, short-lived combustion sources, and biogenic sources. Despite a low overall impact, biogenic sources played an important role during summer, in particular in terms of reactivity.

## 1 Introduction

Tropospheric ozone is an environmental pollutant that causes adverse effects to vegetation, e.g. by reducing and altering physiological processes and plant growth (Matussek et al., 2010; Nunn et al., 2005), and to humans, where respiratory diseases

ACPD

14, 8143–8183, 2014

### Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

can be linked to ozone. In addition to these effects, ozone has been one of the most important greenhouse gases since the beginning of industrialization with a large impact on radiative forcing (Gauss et al., 2003). Atmospheric background concentrations of tropospheric ozone are expected to increase in the 21st century (Vingarzan, 2004).

5 In contrast to other greenhouse gases, such as carbon dioxide, methane, and nitrous oxide, it is not emitted directly, but produced in the atmosphere by photochemical processes from precursor substances. The main drivers for the production of ozone, besides nitrogen oxides ( $\text{NO} + \text{NO}_2$ ) and carbon monoxide (CO), are volatile organic compounds (VOC) and in particular non-methane hydrocarbons (NMHC) (Atkinson, 2000).  
10 NMHC are not only important for the photochemical formation of tropospheric ozone, but also for other secondary air pollutants such as peroxy-carboxylic nitric anhydrides, formaldehyde (Rappenglück et al., 2010), and secondary organic aerosols. In addition, many NMHC species act directly as air toxics or hazardous air pollutants. Thus, monitoring and modeling of the spatiotemporal distribution of these species is important for  
15 mitigation strategies concerning radiative forcing as well as for human health.

The most important sources of NMHC are combustion of fossil fuels from road traffic and industrial processes, handling and evaporation of fuels, solvents, and gases, as well as plant emissions during photosynthesis and stress defense, amongst others. Stevenson et al. (2005) expected a strong increase of biogenic NMHC emissions  
20 caused by temperature stress, with ethylene, isoprene, and monoterpenes as the most prominent compounds. On a global scale biogenic emissions dominate total VOC emissions, while in urban areas they generally only play a minor but non-negligible role.

Reliable long-term scientific data of NMHC are gathered by several groups and various networks. Within the framework of the World Meteorological Organization (WMO), the Global Atmosphere Watch (GAW) program has been developed to achieve global  
25 measurements of the chemical composition of the atmosphere with high data quality (WMO, 2007).

In order to quantify impacts of biogenic and anthropogenic origin on photochemical production of ozone, aerosols, and other compounds, an apportionment into specific

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

source categories is necessary (Badol et al., 2008). Several receptor models such as principal component analysis/absolute principal component scores (PCA/APCS) (e.g. Chan and Mozurkewich, 2007; Guo et al., 2004, 2006), chemical mass balance (CMB) (e.g. Badol et al., 2008; Na and Kim, 2007), or UNMIX (e.g. Jorquera and Rappenglück, 2004; Olson et al., 2007) were used for source apportionment. In particular Positive Matrix Factorization (PMF) (e.g. Lingwall and Christensen, 2007; Paatero, 1997, 1999), a multivariate mathematical receptor model, has been shown to be quite reliable at identifying and quantifying source categories. But most studies investigating NMHC composition were concentrated in urban metropolitan areas with mainly anthropogenic emissions (e.g. Brown et al., 2007; Leuchner and Rappenglück 2010; Na and Kim, 2007).

Yuan et al. (2012) stressed the importance of different reactivity of the NMHC compounds and the impact of photochemical aging on the interpretability of the resolved source profiles that has not been considered in most of the studies applying PMF. The impact of photochemical processing increases with longer transport times from source to receptor which emphasizes the importance of considering this factor at remote sites. Only a few studies have applied PMF receptor modeling at remote sites with a focus on the global or continental background. Lanz et al. (2009) and Sauvage et al. (2009) used PMF analysis for reactive species such as NMHC at remote sites in Switzerland and France despite the PMF assumption of mass conservation from source to measurement site.

In the current work, NMHC data from the GAW global site Hohenpeissenberg, Southern Germany, was used to quantify the impact of different source categories as well as their seasonal contributions at this rural site with PMF analysis. One main objective of this study was to test and discuss whether PMF as a statistical tool is actually capable of reliably identifying the main sources of reactive trace gases at the remote site despite the mass conservation assumption. Limitations of the PMF model at remote sites will be discussed along with the results of the PMF model.

## 2 Methods

### 2.1 Experimental setup

The GAW Observatory Hohenpeissenberg is located about 70 km southeast of Munich (47°48' North, 11°02' East) at 980 m a.s.l. on top of a hill which is about 300 m above the surrounding countryside (approx, 70 % pasture and 30 % forest). Sample air was routinely measured daily at 01:00 CET (41 % of data) and 13:00 CET (48 %), however, 11 % of the data were measured at other times of the day. At 13:00 CET, the site was generally in a vertically fully developed mixed-layer and local emissions may have affected measurements only low to moderately. For this analysis only the 13:00 CET data were used to minimize the influence of local sources and shallow boundary layer conditions during nighttime as well as to ensure the homogeneity of the dataset.

C<sub>2</sub>–C<sub>8</sub>-NMHC were measured with an on-line GC-FID system. It consisted of a 3600 CX Varian gas chromatograph combined with a flame ionization detection (FID) system until January 2008 and was then replaced by a Varian CP-3800 GC-FID. The air intake was 17 m above the ground, and 2 m above a flat roof which was about the same height as the nearby forest canopy (> 10 m distance). The intake was a downwards-facing glass-funnel connected to a permanently flushed glass-manifold (375 L min<sup>-1</sup>, 8 m length, 4 cm I.D.). The GC sampling unit was connected to a port on the manifold via a 1/16" Sulfinert line (Restek, length: 2 m, 0.96 mm I.D., 50 mL min<sup>-1</sup>) such that an overall residence time in the lines of 3.5 s was achieved. After the port, the sample gas passed a filter for aerosol and ozone removal (PTFE filter holder: 25 mm I.D., Metron Technology, PTFE coated glass fiber filter, Fiberfilm, Pall Life Sciences, impregnated with sodium thiosulfate (Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>), backed by a PTFE-membrane filter, 20–30 μm pores, Metron Technology), and further downstream a stainless steel-screen (10 μm pores, VICI). A custom-built sampling and gas flow system was used comprising a moisture trap at 228 K (0.5 m 1/8" Sulfinert, Restek), a VOC trap using cryo-adsorption on glass beads (87 K adsorption, 403 K desorption) (SPT-type by Varian, installed in custom built LN<sub>2</sub> dewar), a sample volume determination by measuring

### Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

the pressure increase in an evacuated reference volume, and the corresponding Valco (VICI) switching valves mounted in a temperature controlled compartment at 293 K. After sampling for 20 min, the cryo-trap was dry-purged by helium at  $15 \text{ mL min}^{-1}$  for 15 min. The adsorbed NMHC were then thermally desorbed in a helium carrier gas flow at  $5 \text{ mL min}^{-1}$  injected and separated on a PLOT column ( $\text{Al}_2\text{O}_3/\text{KCl}$ ,  $50 \text{ m} \times 0.53 \text{ mm}$  I.D., Chrompack, the Netherlands). After an initial isothermal phase (313 K for 2 min), the GC column was heated in two phases, first to 345 K ( $4 \text{ K min}^{-1}$ ), then with a rate of  $6 \text{ K min}^{-1}$  to 473 K. This temperature was kept for 33.7 min. The end of the column extended to the FID system, where the separated compounds were identified. This system was regularly checked and calibrated with helium (zero gas), calibration gas by NPL (certified mixtures of a few ppb in nitrogen of some 30 NMHC), and different reference gases holding synthetic and whole air mixtures in pressurized cylinders (see metadata at WDCGG, 2013).

A more detailed description of measurement, integration, and error assessment was published by Plass-Dülmer et al. (2002). For this study mixing ratios of 24 substances over the course of seven years from 2003 to 2009, measured daily, were used. An individual uncertainty for each compound and each measurement comprising systematic uncertainty contributions and random factors was estimated and assigned to each value. It considers blank values, peak integration errors (including insufficient chromatographic separation) and detection limit, calibration uncertainties and random fluctuations in the system response (Plass-Dülmer et al., 2002).

## 2.2 Positive Matrix Factorization (PMF) model description

Next to Chemical Mass Balance (CMB), UNMIX, and PCA (principal components analysis), Positive Matrix Factorization (PMF) has become an accepted and regularly used tool for receptor modeling. In this study, factor analysis was conducted with PMF 3.0 (US-EPA, 2011). Positive Matrix Factorization determines the number of source factors  $\rho$ , a species profile  $f$  for each factor, and the amount  $g$  that each factor contributes to

each sample.

$$\mathbf{X}_{ij} = \sum_{k=1}^p \mathbf{g}_{ik} \times \mathbf{f}_{kj} \quad (1)$$

In PMF Eq. (1) is solved by decomposing the matrix  $\mathbf{X}_{ij}$  of measurement data, with  $i$  the number of samples and  $j$  the number of the different chemical species, into two matrices, factor contributions, and factor profiles. Both factor contributions and factor profiles can then be analyzed.

There are natural and logical physical conditions for such a model: the original data have to be reproduced by the model, the predicted source compositions and contributions must be non-negative, and the sum of the predicted mass contributions for each source must be less than or equal to the total measured mass for each substance (Hopke, 2003).

The multilinear engine ME-2 (Paatero, 1999) gives PMF the ability to solve multilinear problems and implement constraints like the replacement of missing values and individual weighting of data points by associating an uncertainty value  $U_{ij}$  to each point. The object function (Eq. 2) is then minimized using these uncertainties (Norris et al., 2008).

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left[ \frac{\mathbf{X}_{ij} - \sum_{k=1}^p \mathbf{g}_{ik} \times \mathbf{f}_{kj}}{U_{ij}} \right]^2 \quad (2)$$

PMF 3.0 also gives the opportunity to test the stability and uncertainty of the computed solutions by using a bootstrap technique. It also provides a tool called Fpeak to control the rotations of the different factors (Norris et al., 2008).

Several studies have compared PMF, CMB, PCA, and UNMIX (Anderson et al., 2002; Miller et al., 2002; Paatero and Tapper, 1994; Willis, 2000) and found limitations and advantages of the different models. Some of the advantages of PMF are the good performance not only with simulated data, the non-negativity constraint, and the possibility of individual treatment for single data points.

Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Various studies have shown that PMF provides physically reasonable results for source identification of NMHC in environments located in proximity to the sources (e.g. Buzcu and Fraser, 2006; Lanz et al., 2008; Leuchner and Rappenglück, 2010) or particulate matter (e.g. Santoso et al., 2008; Tauler et al., 2009; Yue et al., 2008). But as receptor models are only mathematical models, they do not use pollutant emissions, chemical transformation mechanisms, or meteorological data to identify and quantify the sources at a receptor location. It is difficult to use PMF for data from remote sites, because with chemical reactions during the transport of air masses (Atkinson and Arey, 2003) and the effects of mixing (Parrish et al., 2007) the presumed mass conservation from source to measurement site, necessary for using receptor models (Hopke, 2003), is not given. Despite these limitations, PMF has been used for remote VOC data (Lanz et al., 2009; Sauvage et al., 2009) and obtained reasonable results. However, it needs to be considered that not only emission profiles but also their different aging determine the factor solutions of PMF.

### 2.3 Data treatment

A total of 2335 valid day (13:00 CET) measurements of 24 substances per measurement were available for the investigated time span of which 345 values of individual substances were missing (0.6% of the entire dataset). There are different ways to treat these data. PMF 3.0 provides the option to exclude the entire sample. In that case a loss of 15% of the sample data would have occurred. To avoid such a high loss of data, the missing values were replaced by the respective species geometric mean and the corresponding uncertainties were set to four times the geometric mean according to Sauvage et al. (2009). Five different treatments of missing value replacement and uncertainty assessment were performed as described below and shown in Table 1.

Values greater than zero, but below the specific detection limit, were replaced by half the detection limit. The uncertainty for zero values and values below the detection limit was set to the specific detection limit (Sauvage et al., 2009). Fpeak values, indicating

the degree of rotation of the solutions, were varied between 5.0 and -5.0 in steps of 0.1.

The remote character of the research site at Hohenpeissenberg implies that only a few substances were emitted locally, but rather transported from their sources. During the transport time photochemical reactions occur and the original emission pattern is altered, due to different photochemical reactivity of the compounds (Atkinson, 2000, 2008). However, PMF assumes inertness of the substances; it cannot integrate reactivity into the model. Sauvage et al. (2009) proposed a method to integrate photochemical reactivity into the overall uncertainty to compensate this limitation of PMF. For each compound a potential error  $E_{j(\text{reactivity})}$  was computed with Eq. (3) assuming pseudo first-order reaction kinetics and photochemical reactivity mainly driven by OH radical reactions.

$$E_j(\text{reactivity}) = X_{ij}(1 - e^{-k_j[\text{OH}]\tau}) \quad (3)$$

$k_j$  is the second-order rate constant of the reaction between the substance  $j$  and OH (Atkinson and Arey, 2003),  $\tau$  is the source-receptor time of transport. [OH] is the seasonally and spatially averaged OH concentration published by Spivakovsky et al. (2000). The overall uncertainty  $s_{ij}$  for the PMF modeling was then computed following the ISO guide rule for uncertainty (ISO 13005, according to Sauvage et al., 2009) (Eq. 4) and used for PMF computations.

$$s_{ij} = 2\sqrt{\left(\frac{U_{ij}(\text{measure})}{2}\right)^2 + \left(\frac{E_j(\text{reactivity})}{\sqrt{3}}\right)^2} \quad (4)$$

This interpretation, however, needs some more modification. The PMF results in factors which correspond to the prevailing factors determined at the receptor site and which are different from pure emission profiles. Instead, they may be seen as aged source profiles with an aging corresponding to some mean transport time from many diverse source locations to the receptor site. Accordingly, the resulting factors are uncertain

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



due to variations in these  $[\text{OH}]$  and  $\tau$ , or with the simplification of known OH, the uncertainty is due to variations of the transport times around a characteristic value. Our dataset was tested with transport time variations of 0, 12, 24, 36, and 72 h. It should be pointed out that this approach reduces the impact of the shorter lived NMHC on the achieved results and that obtained factors have high uncertainty in these short lived compounds. Accordingly, comparison with emission profiles should be performed with caution, especially for the short-lived compounds. The applicability of this method is discussed later.

The impact of different treatments of missing values, of values below the detection limit, and of zero values on the overall results of the PMF analysis was assessed by utilizing five methods of value replacement. In this dataset the number of these values was rather low at 2.3% of all data. Table 1 shows the five different treatments with missing values (treatments 1–5), values below the detection limits (3–5), and zero values (4) replaced by the species median (1) and geometric mean (2–5). The respective uncertainties were fitted accordingly in treatments 1–4 and with additional 20% to *n*-hexane in treatment 5, due to peak overlaps with an unknown substance in the 2003 and 2004 data. Statistical differences in the treatments were tested with a Levene's test for variances. For the test, data were linearly transformed by normalization with the respective arithmetic mean value of the different treatments.

There are different guidelines to help determine the number of factors that best model the measured reality. Mathematical variables like  $Q$  values or the distribution of residuals and stability of the solution can be taken into account, but interpretation of the computed factors by the analyst is a crucial part of selecting the most appropriate solution (Hopke, 2003; Norris et al., 2008). Comparing computed  $Q$  values as a function of the number of factors to theoretical  $Q$  values (approximately the number of data points) seems to work only for certain kinds of weighted uncertainties (Hopke, 2003).

PMF solutions with 2 to 20 factors were calculated, but only the four most plausible solutions (five to eight factors) were compared in this work. Selection criteria were mathematical indicators such as the  $Q$  value, residual distribution, explained variance,

as well as the plausible explanation of the source categories by expert knowledge of the authors.

At the GAW site Hohenpeissenberg other trace gases have also been measured (Gilge et al., 2010). Resolved factors can be compared to these independent measurements to verify the apportionment of the source when the hydrocarbons and the trace gas are emitted from the same source or at the same time but by a different source. Thus, the contributions of factors resolved by PMF were correlated to the additional trace gases NO, NO<sub>2</sub>, SO<sub>2</sub>, CO, and NO<sub>y</sub>. Since sometimes correlations can also be coincidental, Lanz et al. (2008) suggested that these correlations can add evidence to the source apportionment of the factors, but should not be used as the only basis for the attribution of sources.

Carbon monoxide (CO) is a product of incomplete combustion, its main sources are traffic emissions, but it is emitted by all incomplete combustion processes. The nitrogen oxides NO and NO<sub>2</sub> are mainly emitted by combustion engines, NO<sub>y</sub> is the sum of oxidized nitrogen species containing NO, NO<sub>2</sub>, PAN and alkyl nitrates, HNO<sub>3</sub> and others. The main sources of sulfur dioxide (SO<sub>2</sub>) emissions are power plants burning fossil fuels for energy generation but also households to a minor extent (UBA, 2011). Only if a resolved source profile correlated with one or more of the other trace gases, it was mentioned in the text.

PMF 3.0 provides a bootstrap function that selects blocks of input samples and creates new input files from them. These files have the same dimensions as the original input files. Then PMF is run and the resulting factors are mapped to the base factor they correlate with best. In this study, 200 bootstrap runs were performed on the base of the final solution with six factor profiles.

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[⏪](#)

[⏩](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



## 3 Results and discussion

### 3.1 Determination of the number of factors

The calculated  $Q$  values of four different factor solutions with five to eight factors and other diagnostic parameters are shown in Table 2. The values decreased with increasing number of factors, due to a better explanation of the variability of the measured NMHC mixing ratios by a higher number of factors, and lower global minima of the object functions. All scaled residuals should be within  $\pm 3\sigma$  (Willis, 2000) and symmetrically distributed. With this dataset all solutions from five to eight factors had normally distributed residuals. The number of residuals beyond three standard deviations decreased for an increasing number of factors (Table 2) down to four for 20 factors. An  $F_{\text{peak}}$  value of 0.0 showed the lowest  $Q$  values for all solutions displayed in Table 2, additional rotation of the factors did not improve the results. Coefficients of determination of modeled to measured mixing ratios and the mean ratio were high for all solutions and increasing with the number of resolved factors from 0.86 ( $R^2$ ) and 0.97 (ratio), respectively, with five to 0.89 and 0.98, respectively, with eight factors.

An indication for an appropriate number of factors is the stability of the factors after performing the analysis at least five times with the same parameters but randomized starting points. No multiple solutions should be found (Hopke, 2003). In this analysis the tested five, six, seven, and eight factor solutions were very stable and showed the same distribution of factors for all five computations (not shown here).

On the basis of these statistical indicators no final decision about the optimal number of factors could be made. Thus, it remained a process based on plausibility arguments mainly by checking the resulting factors vs. reasonable source or aged source profiles (cf. Sect. 3.2), also in comparison to previous studies (Lanz et al., 2008, 2009; Sauvage et al., 2009). Careful consideration led to the decision of choosing six factors as the most reasonable.

## 3.2 Source apportionment

In the following, the six factor solution (Fig. 1) will be discussed and compared to alternative solutions with five, seven, and eight factors. In addition, the apportionment to source categories and the interpretation of chemical aging for the respective factors will be performed by comparison to source profiles throughout the literature. Because of the complexity of the atmospheric system with transport, mixing, and chemistry, each individual factor can not be attributed exclusively to one source category. The factors should be seen as more or less aged profiles originating from different groups of sources (Sauvage et al., 2009). For a better attribution to source categories, the annual courses (Fig. 2) including the winter–summer amplitude of the retrieved profiles were compared to the seasonal change of OH at the site. Based on the continuous OH-radical measurements at Hohenpeissenberg, this effect accounted for an amplitude of about 7, e.g. short lived compounds emitted at constant rates throughout the year and with fairly homogenous spatial distribution, like vehicular exhaust, would be factors of about 7 higher in winter than in summer. However, there are effects that might potentially disturb this rather simple relation. The generally lower height of the mixed layer in winter might lead to less dilution and consequently higher winter–summer ratios. This, however, is valid only for those measurements that were performed within the mixed layer. On the other hand, in winter during daytime the site was frequently above the mixed layer and thus decoupled from the emissions at ground. Both effects are expected to partly compensate each other. In addition to a change of composition patterns during transport, emission profiles in literature vary due to measurement uncertainty associated with the use of different techniques, e.g. on-line or canister samples, the number of substances measured, the experimental set-up and associated conditions, and the published units. These variations made direct quantitative comparisons of the profiles difficult, but allowed a qualitative assessment for identification of possible source profiles.

Figure 1 shows the absolute and relative contributions of substances to the six factor solution. Absolute values are the mixing ratios of each substance that PMF apportioned to each factor. Relative contributions are the fractions of each substance contributing to each factor, therefore, the sum of all factors for each substance is 1.

### 3.2.1 Biogenic sources

Factor I explained 34 pptv of the measured isoprene which was the only biogenic NMHC measured. On the absolute scale ethane (49 pptv), ethene (21 pptv) and isopentane (14 pptv) also contributed considerably to this factor, but due to the small total amount of isoprene, 94 % of it was explained by this factor in contrast to only about 4 % of the total amount of ethane and ethene. Since isoprene is the only substance in this study that is almost exclusively produced by plants, this factor is apportioned to biogenic sources. The high reactivity and short life-time of isoprene in the atmosphere exclude distant sources. Small amounts of ethene found in this factor might also be of biogenic origin, since ethene is an important plant hormone (Fall, 1999). Factor I was the only factor with a distinct maximum in summer (Fig. 2). Since isoprene emissions depend on the photosynthesis of plants, temperature, and solar radiation (Fuentes and Wang, 1999), the maximal source strength of the highly reactive diene and thus the maximal mixing ratios could be found in July despite higher photochemical reactivity and higher mixing during this time of year. The ethane contribution to this factor might have derived from biomass burning (Stein and Rudolph, 2007) with a maximum during the summer.

Similar profiles for this factor, including amounts of ethane, *n*-pentane and isopentane were also found by Sauvage et al. (2009) for three remote sites in France. The alkane contributions in this factor could be attributed to artifacts from the PMF model or to temperature related emissions like evaporation from fuel.

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



### 3.2.2 Short-lived incomplete combustion sources

Factors II and III both showed a large contribution of short-lived substances. Factor II contained rather short-lived (< 1 day) alkenes such as ethene (133 pptv, 27%), propene (46 pptv, 79%), 1,3-butadiene (6 pptv, 90%), and some butenes, typical for incomplete combustion processes. The winter–summer amplitude (Fig. 2) of about 9 was quite consistent with the seasonal OH change of 7 indicating constant emissions throughout the year with possibly larger fetch areas for the short-lived compounds during winter due to the longer life-time. Due to these constant emissions and the following comparison to emission profiles, this factor was referred to as short-lived combustion sources mainly from vehicular exhaust.

Literature profiles of vehicle emissions show similarly high contributions of ethene, propene, isopentane, benzene, toluene and xylenes (e.g. Friedrich and Obermeier, 1999; Hellen et al., 2003; Sagebiel et al., 1996; Thijssse et al., 1999). Sauvage et al. (2009) provided a vehicle exhaust factor very similar for one of the French remote sites.

Factor II correlated well with NO<sub>2</sub> ( $r = 0.84$ ), NO (0.75), NO<sub>y</sub> (0.73), and CO (0.66), all associated with traffic emissions.

### 3.2.3 Short-lived evaporative sources

Factor III explained large parts of the measured toluene (63 pptv, 55%), ethylbenzene (11 pptv, 60%), and *m* + *p*-xylene (23 pptv, 65%). Alkanes such as pentanes, *n*-hexane, and *n*-heptane also contributed to this factor. The life-times of the compounds were in the short- to medium-lived range of 1–4 days. The aromatics were the only substances measured in this study that are found in evaporative solvent emission profiles, e.g. in paint or wood coating. The winter–summer amplitude of about 3.6 (Fig. 2) indicated stronger emissions during the summer, typical for evaporative sources dependent on temperature such as solvents and fuel evaporation. This factor is thus further referred to as short-lived evaporative sources, additionally containing

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



some medium-lived combustion compounds, which is shown also by its correlation with combustion tracers NO<sub>2</sub> ( $r = 0.74$ ), NO (0.74), and NO<sub>y</sub> (0.75).

### 3.2.4 Residential heating and long-lived incomplete combustion sources

In comparison to Factors II and III, the compounds in Factors IV and V exhibited longer atmospheric life-times. Large parts of the alkynes acetylene (216 pptv, 49 %) and propyne (8 pptv, 56 %) were explained in Factor IV that also contained ethene (287 pptv, 58 %), benzene (65 pptv, 48 %) and toluene (23 pptv, 21 %) that are typical compounds emitted by residential heating and other incomplete combustion processes, in particular from road traffic. The major compounds resembled in this factor had atmospheric life-times of 1–10 days. Only a few source profiles for domestic combustion emissions were found in literature. A wood burning profile by Friedrich and Obermeier (1999), adjusted to the substances measured for this study, showed a similar composition to Factor IV, as does a “residential heating” profile by Klemp et al. (2002) and Mannschreck et al. (2002). Ethene and acetylene contributions of the literature profiles differed from those found in Factor IV. These differences indicated that not only wood burning but also other sources contributed to this factor. Similar profiles to Factor IV were found by Lanz et al. (2008) and Sauvage et al. (2009), who attributed the French profiles to hot water generation and building heating by burning fossil fuels and wood. The benzene/toluene ratio for those French sites was 2, for Mannschreck et al. (2002) 3.2, and for Factor IV 2.7. The winter–summer ratio was very high with an amplitude of 15.8 (Fig. 2) indicating a much higher source strength in winter than in summer according to a higher heating demand during the cold season. The shorter-lived compounds such as ethene and some of the aromatics possibly had larger fetch areas during winter due to the longer life-times at lower temperatures contributing to this large winter–summer spread. Factor IV correlated well with the combustion tracer CO ( $r = 0.87$ ). A correlation to SO<sub>2</sub> (0.62) was found for this factor, which is reasonable, since household emissions contribute to SO<sub>2</sub> in the atmosphere (UBA, 2011).

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



### 3.2.5 Long-lived evaporative sources

In Factor V alkanes from evaporation were the predominant substances, with ethane (331 pptv, 22 %), propane (306 pptv, 54 %), isobutane (75 pptv, 59 %), *n*-butane (128 pptv, 61 %), *n*-hexane (9 pptv, 47 %), isopentane (37 pptv, 26 %), *n*-pentane (38 pptv, 50 %), and small amounts of acetylene and other compounds contributing to it. These were typical evaporative sources, but longer-lived (> 4 days) than in Factor III (< 4 days). Literature emission profiles for crude oil production, natural gas, gasoline, and LPG evaporation (e.g. Friedrich and Obermeier, 1999; Lanz et al., 2008; Leuchner and Rappenglück, 2010; Mannschreck et al., 2002; Thijsse et al., 1999; Watson et al., 2001) closely resembled this factor. Although emission profiles and PMF results of urban datasets presented high isopentane/*n*-pentane ratios, Sauvage et al. (2009) found ratios close to 1 for remote sites in France for evaporative sources, similar to our study. Factor V can be viewed as an aged combined profile of evaporative losses of gasoline, LPG, and parts of natural gas with atmospheric life-times of 4–60 days (C<sub>2</sub>–C<sub>4</sub> alkanes). The winter–summer amplitude had a value of 7 (Fig. 2) which coincided with the seasonal OH course. That indicated rather constant sources throughout the year such as leakage rather than temperature dependent evaporation. The correlation with CO (*r* = 0.80) indicated that this factor may not solely consist of evaporated compounds, but also of long-lived combustion products, e.g. from natural gas burning.

### 3.2.6 Remote sources

Factor VI can be apportioned to remote sources showing the continental background. It explained most of the measured ethane (881 pptv, 59 %) and quite large amounts of benzene (42 pptv, 31 %). Ethane is the most abundant and longest living compound measured in this study and appears as an indicator for aged air masses. This factor was largely driven by ethane with substantial contributions of biomass burning which were higher in summer, accordingly the winter–summer amplitude was only 3.1 (Fig. 2). Propane, acetylene, and benzene also are very stable substances (life-times 7–10

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



days) that underline the background character of this factor. Hellen et al. (2003) provided a measured profile attributed to distant sources in Finland that showed a similar composition. Sauvage et al. (2009) identified a remote source factor with high ethane loadings and amounts of propane in France and Lanz et al. (2008) found an ethane factor. These findings led to the conclusion that this factor represented the European NMHC background. Factor VI did not correlate well with any of the additional trace gases.

### 3.2.7 Seasonality and total contributions

The winter–summer amplitude was already used for better attribution of source categories to the individual factors in the previous sections (Fig. 2). The total monthly variation of factorial NMHC mixing ratios is summed in Fig. 3. The maximum was encountered in February (7.1 ppbv), the minimum in July (1.8 ppbv). During fall and winter, chemical reaction rates decreased due to lower OH concentrations as a consequence of lower UV light available and lower temperatures. Thus, the NMHC were not depleted as rapidly as during the summer months.

Within this dataset, biogenic sources only contributed to the measured and modeled NMHC in summer, but are responsible for 20 % of the total modeled amount of hydrocarbons in July. Due to the high reactivity of isoprene, this overall small factor plays an important role in atmospheric processes in summer. All anthropogenic factors (Factors II–VI) were more or less consistent with the annual variation of OH, depending on the seasonal source strengths of the individual factors as discussed in the previous chapters. Due to massive summer depletion, the short-lived combustion (II) and long-lived evaporative (V) compounds had factor 7–9 higher mixing ratio contributions in winter than in summer, although their atmospheric turn-over remained the same all over the year. The remote factor VI containing only long-lived hydrocarbons dominated throughout the year with more than 50 % abundance in May. The substances contributing to the short-lived Factor II were expected to be enhanced in winter due to cold start emissions of vehicles (Borbon et al., 2004) and higher use of heating functions in vehicles

(winter/summer = 9). Contrary, short lived evaporative emissions were driven by temperature and much enhanced in summer (Factor III). The residential heating Factor IV showed the most pronounced maximum in winter as expected. The annual course of Factor V showed fairly constant emissions such as leakage from natural gas, LPG, and vehicular carburetors, rather than temperature dependent evaporation.

The absolute contributions of Factor VI had their minimum in July/August and the maximum in March, however, later than e.g. Factors IV and V. The annual mixing ratio variability followed the OH abundance with a lag caused by the very slow reaction rate constant of ethane (Plass-Dülmer et al., 2002).

Over the entire seven year period remote sources contributed most to the NMHC volume (Fig. 4a) and mass at the remote site (31.3 vol%, 25.8 mass%) followed by long-lived evaporative sources (24.6 vol%, 25.4 mass%). Residential heating/long-lived combustion sources accounted for 23.4 vol% (21.0 mass%), short-lived evaporative sources for 11.3 vol% (17.8 mass%), and short-lived combustion sources for 6.0 vol% (5.9 mass%), respectively. The smallest fraction was attributed to biogenic sources (3.5 vol%, 4.0 mass%), but mainly emitted in summer, being an important precursor to ozone formation. Weighted with the corresponding OH reaction rate constants, the mean OH-reactivity of the factors was as follows (Fig. 4b): biogenic – 24.2 %, short-lived combustion – 22.5 %, residential heating – 20.6 %, short-lived evaporative – 15.1 %, remote sources – 9.0 %, long-lived evaporative – 8.5 %. Despite low volume and mass fractions the biogenic isoprene factor has the highest contribution for the total reactivity. The impact is even higher during maximum emissions in the summer.

### 3.3 Comparison to alternative factor solutions

In Fig. 5 the additional factors that were resolved in the seven and eight factor solutions and one different factor for the five factor solution are shown. In general, all factors of the six factor solution could also be found in these alternative computations. For the five factor computation considerable differences to the solution above could be found in Factors II, IV, and V. Factor II also contained large amounts of ethane, smaller amounts

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of acetylene, benzene, and toluene. The longer-lived Factors IV and V were combined to one factor. The seven and eight factor solutions provided an additional common factor that contained more than one third of the modeled isobutene and considerable amounts of 2-methylpentane, 3-methylpentane, benzene, toluene and some other substances. Thus, the amounts of these substances in other factors, especially Factor III, decreased accordingly, compared to the six factor computation. Ethane in Factor V only occurred for the six factor solution; it was distributed in other factors, e.g. Factor VI, for the seven and eight factor solutions. Isopentane and *n*-pentane accounted for the second additional factor in the eight factor solution, thus, their contribution to Factor III was much lower than in the six factor solution.

The five factor solution combined the two long-lived Factors IV and V from the six factor solution that should appear separately when compared e.g. to Sauvage et al. (2009) as well as to the source apportionment in Sect. 3.2. The seven and eight factor solutions showed an additional factor with comparatively small contributions of many substances that could not be apportioned to proper source categories. Since the statistical values of the PMF analysis did not show a clear indication for the ideal number of factors, the decision was made to choose six factors as the most reasonable and explainable solution that PMF calculated.

### 3.4 Effects of different data treatment methods

Five methods to deal with missing values, values below the detection limit and uncertainty values due to interference with other substances or other measurement problems were applied to the six factor solution to analyze the effects of different treatments to the overall solution (cf. Table 1). Figure 6 shows the results of the five different scenarios displayed on a relative scale. Only the results for Factor V are shown, since within this factor the largest differences between the computed results of different datasets can be found.

No significant statistical difference for any profile was found for any of the five treatments with a Levene's test for variances. Thus, in this study different treatment of these

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



values had only a very small quantitative impact on the final result. An additional reason for the differences between the computations could be different starting points in PMF that led to slightly different resulting numbers, but had no influence on the overall stable result.

### 5 3.5 Inclusion of a rough proxy for profile distortion due to photochemical aging

Based on the rate constants for the respective substances (Atkinson, 2000), a reactivity error value depending on different transport or reaction time was calculated for every sample and every species (cf. Sect. 2.3) following the approach by Sauvage et al. (2009). Reactivity error contributions were calculated for up to 72 h. However, it should be pointed out that the variation of the transport times around characteristic values should be considered in relation to the respective factor and the lifetimes of compounds present in the factor. Accordingly, for Factors II and III with compound lifetimes of hours to few days, it does not make sense to vary the transport times by more than 12–24 h, otherwise compounds vanish. Furthermore, it should be considered that a factor is distorted due to aging depending on the different lifetimes of the VOC compounds present in the respective factor. But the factors determined by PMF in this paper have been demonstrated to represent compounds with similar lifetimes. Thus, the distorting effect is rather small, and in fact the uncertainty due to aging should be scaled to lifetime differences rather than absolute lifetimes. Nevertheless, we want to use and discuss the Sauvage et al. (2009) approach, and Table 3 shows the average values of the calculated uncertainties for modeled reaction times of 0 and 36 h exemplarily. Uncertainties for compounds with low reactivity such as some alkanes and aromatic species remained at relatively low levels; ethane as the most stable substance changed its uncertainty from 7.1 to 8.0 %, benzene from 8.6 to 16.3 %. A stronger influence of reactivity existed for alkenes, e.g. for 1-butene from 29.3 to 92.4 % or 1,3-butadiene from 43.6 to 108.9 %. The uncertainty for the highly reactive isoprene rose from 33.3 % to 101.7 % after 12 h (not shown) and to 115.8 % after 36 h of assumed reaction time. These high uncertainties made proper modeling in PMF impossible and for

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the reasons discussed before this demonstrates the limits of this approach. To partly overcome this, the isoprene uncertainty due to reactivity was not considered as the biogenic factor mainly depends on isoprene alone.

The attempt to compensate the mass conservation assumption of the PMF model by changing the uncertainty values of the reactive species had an impact on the resulting factors. The model reacted to the change of uncertainty data by changing the quantitative composition of each factor; the  $Q$  values for the model decreased with increasing uncertainties. The largest change of substance contributions to each factor was found between zero and 12 h where the uncertainties change the most. Despite the sometimes large changes of some substance contributions, especially ethane and ethene, the relative profiles of the factors remained more or less recognizable for the different tested times. This demonstrates the robustness of the results even though this approach is rather crude for the reasons discussed above. The absolute and relative factor profiles of the PMF solution with six factors and 36 h reaction time for all substances except isoprene can be seen in Fig. 7 to illustrate the changes in the solutions with integrated reactivity. However, due to the discussion above the apparently small deviations to the original solution without introducing this enhanced uncertainty due to reactivity were not further discussed.

In terms of model stability, bootstrap runs for the 0 h solution were very stable with only small variations in the contributions of the substances to the factors. Small variability could only be seen for Factors II and III in regards to some alkenes. Bootstrap runs including reaction times of 12–72 h were still stable for most factors, in particular for longer-lived factors (Factor IV, VI). Factor V also exhibited a quite stable behavior with only some variability of ethane. The bootstrap process attributed factor profiles with various different contributions of aromatic compounds, *n*-heptane, and, to some extent, methylpentanes and hexane, to both Factors II and III. For Factor II the median of the contributing percentages of *n*-heptane, toluene, ethylbenzene, and xylenes of the 200 bootstrap profiles was much higher than the base run value. The stabilities from the bootstrap runs with the inclusion of the photochemical aging showed a tendency

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



that factors with high relative contributions of long-lived substances are more stable than those with high contributions of substances with higher rate constants such as ethene, propene, and toluene, as expected.

In addition to the less stable bootstrap runs, transport times from major source areas such as the surrounding cities of Munich, Augsburg, and Weilheim were on average shorter than 12 h. Additionally, the patterns of the factors were shifted towards longer-lived compounds, reducing the influence of species with higher reactivity in a way that the new factors do not well resemble the original emission profiles at the sources. Aging can thus better be taken into account when the profiles are interpreted rather as aged profiles than source emission profiles unaltered by higher uncertainties due to photochemical reactivity. In conclusion, an application of the method proposed by Sauvage et al. (2009) seems not suitable to better interpret the results at this site.

## 4 Conclusions and outlook

Positive Matrix Factorization (PMF) was used on this dataset of NMHC measurements at the remote GAW site Hohenpeissenberg. For the determination of the number of factors, statistics on the provided results were not decisive. Interpretability of the computed factors was the most important parameter in the analysis and interpretation of the results. Treatment of missing values and uncertainties had no substantial influence on the solutions. It could be interesting for other datasets to evaluate the number of missing values that the model can compensate for before an effect on the result is seen.

The stability of the modeled factors depended on the individual uncertainties of the contributing substances. Especially short-lived Factors II (incomplete combustion) and III (evaporative) showed lower stability in the bootstrap runs performed by PMF. These two factors were also the ones that were attributed to shorter-lived categories containing compounds with high reactivity and thus stronger depletion from the sources to the receptor site. The other factors could be attributed to biogenic sources,

### Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Can PMF identify sources of organic trace gases at a GAW site?**

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

residential heating including long-lived incomplete combustion, long-lived evaporative sources (e.g. from gas leakage), and remote sources that reflect the continental background. The measured anthropogenic NMHC have maxima in winter that were enhanced mainly due to lower corresponding OH concentrations. The overall influence of biogenic isoprene sources is non-negligible. In terms of reactivity it showed the highest reactivity of all factors and was much higher than by its volume or mass contribution. When considering the seasonal variation with maxima in summer and during the day, while anthropogenic sources had their minima, the impact to the reactivity at certain times can be dominant. Contributions by other biogenic emissions such as monoterpenes and sesquiterpenes that were not considered in this study can even increase the biogenic impact.

The factor composition profiles were generally in agreement with source profiles from literature and calculated results from PMF studies in urban areas. Most of the literature source profiles are rather old, with large uncertainties and only low substance resolution, thus the comparisons to these possibly outdated profiles leave a lot of room for interpretation (Theloke and Friedrich, 2007). Improvement and updates of emission profiles, particularly in regard to new legal limitations of emissions, seem necessary.

Since Yuan et al. (2012) emphasized the non-negligible influence of photochemical aging on PMF results at urban sites, the pronounced inclusion of aging into the interpretation of the factors (in this case Factors II–V) rather than a typical source apportionment seems necessary for remote sites that cannot fulfill the assumption of mass conservation during transport. Factors II and IV were thus attributed to incomplete combustion originating mainly from vehicular exhaust and residential heating (lead components: light alkenes, alkynes, benzene), and Factors III and V to evaporative losses from fuel (including fuel not burned e.g. during cold-start), natural gas and solvents as they predominantly contain alkanes and aromatics. Differences between Factors II and IV, and between III and V, respectively, were the different life-times of the compounds of less than 1 day (II) and more than 1 day (IV), and less than 4 days (III) and more than 4 days (V). In view of interpreting patterns in the atmosphere, the

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

regional background is resembled by the longer-lived patterns (IV and V) and fresh, local impacts by the shorter lived patterns (II and III), the superposition reflects the different degrees of aging and mixing. This interpretation with respect to only two lumped source categories is plausible in the light of mainly two factors: measurements were made at times of well mixed conditions (13:00 CET) and local, more pronounced impacts of individual sources were damped. Secondly, emissions are usually correlated with population density for major source types especially in the rural surroundings of Hohenpeissenberg with no major local sources, e.g. the closest motorway is some 30 km east. Thus, signatures of different individual sources were not expected to show up in different air masses in a pronounced way.

The factors resolved for the Hohenpeissenberg dataset were very similar to those found by Sauvage et al. (2009) for remote sites in France; PMF seems to be able to calculate reasonable results for reactive species without including reactivity into the uncertainty for the PMF model.

Using another receptor model like CMB or UNMIX on this dataset could further confirm the resolved factors, as could expanding the dataset to include more species including aerosols.

*Acknowledgements.* We gratefully acknowledge the technical work by K. Michl and E. Tensing in set-up of instruments, measurements, and data evaluation. We further thank S. Gilge and the team at the GAW station Hohenpeissenberg for non-organic trace gas data and T. Sparks for editing the language of our manuscript. This work was supported by the German Research Foundation (DFG) and the Technische Universität München within the funding program Open Access Publishing.

## References

- Anderson, M. J., Daly, E. P., Miller, S. L., and Milford, J. B.: Source apportionment of exposures to volatile organic compounds: II. Application of receptor models to TEAM study data, *Atmos. Environ.*, 36, 3643–3658, 2002.
- Atkinson, R.: Atmospheric chemistry of VOCs and NO<sub>x</sub>, *Atmos. Environ.*, 34, 2063–2101, 2000.

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Atkinson, R.: Our present understanding of the gas-phase atmospheric degradation of VOCs, in: *Simulation and Assessment of Chemical Processes in a Multiphase Environment*, edited by: Barnes, I. and Kharytonov, M. M., 1–19, Springer, Netherlands, Dordrecht, 2008.

Atkinson, R. and Arey, J.: Atmospheric degradation of volatile organic compounds, *Chem. Rev.*, 103, 4605–4638, 2003.

Badol, C., Locoge, N., and Galloo, J. C.: Using a source–receptor approach to characterise VOC behaviour in a French urban area influenced by industrial emissions, Part II: Source contribution assessment using the chemical mass balance (CMB) model, *Sci. Total. Environ.*, 389, 429–440, 2008.

Borbon, A., Coddeville, P., Locoge, N., and Galloo, J. C.: Characterising sources and sinks of rural VOC in eastern France, *Chemosphere*, 57, 931–942, 2004.

Brown, S. G., Frankel, A., and Hafner, H. R.: Source apportionment of VOCs in the Los Angeles area using positive matrix factorization, *Atmos. Environ.*, 41, 227–237, 2007.

Buzcu, B. and Fraser, M. P.: Source identification and apportionment of volatile organic compounds in Houston, TX, *Atmos. Environ.*, 40, 2385–2400, 2006.

Chan, T. W. and Mozurkewich, M.: Application of absolute principal component analysis to size distribution data: identification of particle origins, *Atmos. Chem. Phys.*, 7, 887–897, doi:10.5194/acp-7-887-2007, 2007.

Fall, R.: Biogenic emissions of volatile organic compounds from higher plants, in: *Reactive Hydrocarbons in the Atmosphere*, edited by: Hewitt, N., Academic Press, London, 41–95, 1999.

Friedrich, R. and Obermeier, A.: Anthropogenic emissions of volatile organic compounds in: *Reactive Hydrocarbons in the Atmosphere*, edited by: Hewitt, N., Academic Press, London, 1–39, 1999.

Fuentes, J. D. and Wang, D.: On the seasonality of isoprene emissions from a mixed temperate forest, *Ecol. Appl.*, 9, 1118–1131, 1999.

Guo, H., Wang, T., and Louie, P. K. K.: Source apportionment of ambient non-methane hydrocarbons in Hong Kong: application of a principle component analysis/absolute principle component scores (PCA/APCS) receptor model, *Environ. Pollut.*, 129, 489–498, 2004.

Guo, H., Wang, T., Blake, D. R., Simpson, I. J., Kwok, Y. H., and Li, Y. S.: Regional and local contributions to ambient non-methane volatile organic compounds at a polluted rural/coastal site in Pearl River Delta, China, *Atmos. Environ.*, 40, 2345–2359, 2006.

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)



[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)

- Gauss, M., Myhre, G., Pitari, G., Prather, M. J., Isaksen, I. S. A., Bernsten, T. K., Brasseur, G. P., Dentener, F. J., Derwent, R. G., Hauglustaine, D. A., Horowitz, L. W., Jacob, D. J., Johnson, M., Law, K. S., Mickley, L. J., Müller, J. F., Plantevin, P. H., Pyle, J. A., Rogers, H. L., Stevenson, D. S., Sundet, J. K., van Weele, M., and Wild, O.: Radiative forcing in the 21st century due to ozone changes in the troposphere and the lower stratosphere, *J. Geophys. Res.*, 108, 4292, doi:10.1029/2002JD002624, 2003.
- Gilge, S., Plass-Duelmer, C., Fricke, W., Kaiser, A., Ries, L., Buchmann, B., and Steinbacher, M.: Ozone, carbon monoxide and nitrogen oxides time series at four alpine GAW mountain stations in central Europe, *Atmos. Chem. Phys.*, 10, 12295–12316, doi:10.5194/acp-10-12295-2010, 2010.
- Hellen, H., Hakola, H., and Laurila, T.: Determination of source contributions of NMHCs in Helsinki (60° N, 25° E) using chemical mass balance and the Unmix multivariate receptor models, *Atmos. Environ.*, 37, 1413–1424, 2003.
- Hopke, P. K.: Recent developments in receptor modeling, *J. Chemometr.*, 17, 255–265, 2003.
- Jorquera, H. and Rappenglück, B.: Receptor modeling of ambient VOC at Santiago, Chile, *Atmos. Environ.*, 38, 4243–4263, 2004.
- Klemp, A., Mannschreck, K., Pätz, H. W., Habram, M., Matuska, P., and Slemr, F.: Determination of anthropogenic emission ratios in the Augsburg area from concentration ratios: results from long-term measurements, *Atmos. Environ.*, 36, 61–80, 2002.
- Lanz, V. A., Hueglin, C., Buchmann, B., Hill, M., Locher, R., Staehelin, J., and Reimann, S.: Receptor modeling of C<sub>2</sub>–C<sub>7</sub> hydrocarbon sources at an urban background site in Zurich, Switzerland: changes between 1993–1994 and 2005–2006, *Atmos. Chem. Phys.*, 8, 2313–2332, doi:10.5194/acp-8-2313-2008, 2008.
- Lanz, V. A., Henne, S., Staehelin, J., Hueglin, C., Vollmer, M. K., Steinbacher, M., Buchmann, B., and Reimann, S.: Statistical analysis of anthropogenic non-methane VOC variability at a European background location (Jungfrauoch, Switzerland), *Atmos. Chem. Phys.*, 9, 3445–3459, doi:10.5194/acp-9-3445-2009, 2009.
- Lingwall, J. W. and Christensen, W. F.: Pollution source apportionment using a priori information and positive matrix factorization, *Chemometr. Intell. Lab.*, 87, 281–294, 2007.
- Leuchner, M. and Rappenglück, B.: VOC source–receptor relationships in Houston during TexAQS-II, *Atmos. Environ.*, 44, 4056–4067, 2010.
- Mannschreck, K., Klemp, A., Kley, D., Friedrich, R., Kühlwein, J., Wickert, B., Matuska, P., Habram, M., and Slemr, F.: Evaluation of an emission inventory by comparisons of mod-

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

elled and measured emission ratios of individual HCs, CO and NO<sub>x</sub>, Atmos. Environ., 36, 81–94, 2002.

Matyssek, R., Wieser, G., Ceulemans, R., Rennenberg, H., Pretzsch, H., Haberer, K., Löw, M., Nunn, A. J., Werner, H., Wipfler, P., Oßwald, W., Nikolova, P., Hanke, D. E., Kraigher, H., Tausz, M., Bahnweg, G., Kitao, M., Dieler, J., Sandermann, H., Herbinger, K., Grebenc, T., Blumenröther, M., Deckmyn, G., Grams, T. E. E., Heerd, C., Leuchner, M., Fabian, P., and Häberle, K. H.: Enhanced ozone strongly reduces carbon sink strength of adult beech (*Fagus sylvatica*) – resume from the free-air fumigation study at Kranzberg Forest, Environ. Pollut., 158, 2527–2532, 2010.

Miller, S. L., Anderson, M. J., Daly, E. P., and Milford, J. B.: Source apportionment of exposures to volatile organic compounds: I. Evaluation of receptor models using simulated exposure data, Atmos. Environ., 36, 3629–3641, 2002.

Na, K. and Kim, Y. P.: Chemical mass balance receptor model applied to ambient C2–C9 VOC concentration in Seoul, Korea: effect of chemical reaction losses, Atmos. Environ., 41, 6715–6728, 2007.

Norris, G., Vedantham, R., Wade, K., Brown, S., Prouty, J., and Foley, C.: EPA Positive Matrix Factorization (PMF) 3.0 Fundamentals & User Guide, US Environmental Protection Agency, available at: [http://www.epa.gov/heads/documents/EPA\\_PMF\\_3.0\\_User\\_Guide.pdf](http://www.epa.gov/heads/documents/EPA_PMF_3.0_User_Guide.pdf) (last access: 13 March 2014), 2008.

Nunn, A. J., Kozovits, A. R., Reiter, I. M., Heerd, C., Leuchner, M., Lütz, C., Liu, X., Löw, M., Winkler, J. B., Grams, T. E. E., Häberle, K. H., Werner, H., Fabian, P., Rennenberg, H., and Matyssek, R.: Comparison of ozone uptake and sensitivity between a phytotron study with young beech and a field experiment with adult beech (*Fagus sylvatica*), Environ. Pollut., 137, 494–506, 2005.

Olson, D. A., Norris, G. A., Seila, R. L., Landis, M. S., and Vette, A. F.: Chemical characterization of volatile organic compounds near the World Trade Center: ambient concentrations and source apportionment, Atmos. Environ., 41, 5673–5683, 2007.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemometr. Intell. Lab., 37, 23–35, 1997.

Paatero, P.: The multilinear-engine – a table-driven, least squares program for solving multilinear problems, including the *n*-way parallel factor analysis model, J. Comput. Graph. Stat., 1, 854–888, 1999.

---

**Can PMF identify sources of organic trace gases at a GAW site?**M. Leuchner et al.

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Paatero, P. and Tapper, U.: Analysis of positive matrix factorization: a nonnegative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5, 111–126, 1994.
- Parrish, D. D., Stohl, A., Forster, C., Atlas, E. L., Blake, D. R., Goldan, P. D., Kuster, W. C., and de Gouw, J. A.: Effects of mixing on evolution of hydrocarbon ratios in the troposphere, *J. Geophys. Res.-Atmos.*, 112, D10S34, doi:10.1029/2006JD007583, 2007.
- Plass-Dülmer, C., Michl, K., Ruf, R., and Berresheim, H.: C-2–C-8 hydrocarbon measurement and quality control procedures at the Global Atmosphere Watch Observatory Hohenpeisenberg, *J. Chromatogr. A*, 953, 175–197, 2002.
- Rappenglück, B., Dasgupta, P. K., Leuchner, M., Li, Q., and Luke, W.: Formaldehyde and its relation to CO, PAN, and SO<sub>2</sub> in the Houston-Galveston airshed, *Atmos. Chem. Phys.*, 10, 2413–2424, doi:10.5194/acp-10-2413-2010, 2010.
- Sagebiel, J. C., Zielinska, B., Pierson, W. R., and Gertler, A. W.: Real-world emissions and calculated reactivities of organic species from motor vehicles, *Atmos. Environ.*, 30, 2287–2296, 1996.
- Santoso, M., Hopke, P. K., Hidayat, A., and Diah, D. L.: Sources identification of the atmospheric aerosol at urban and suburban sites in Indonesia by positive matrix factorization, *Sci. Total Environ.*, 397, 229–237, 2008.
- Sauvage, S., Plaisance, H., Locoge, N., Wroblewski, A., Coddeville, P., and Galloo, J. C.: Long term measurement and source apportionment of non-methane hydrocarbons in three French rural areas, *Atmos. Environ.*, 43, 2430–2441, 2009.
- Spivakovsky, C. M., Logan, J. A., Montzka, S. A., Balkanski, Y. J., Foreman-Fowler, M., Jones, D. B. A., Horowitz, L. W., Fusco, A. C., Brenninkmeijer, C. A. M., Prather, M. J., Wofsy, S. C., and McElroy, M. B.: Three-dimensional climatological distribution of tropospheric OH: update and evaluation, *J. Geophys. Res.-Atmos.*, 105, 8931–8980, 2000.
- Stein, O. and Rudolph, J.: Modeling and interpretation of stable carbon isotope ratios of ethane in global chemical transport models, *J. Geophys. Res.*, 112, D14308, 1–18, doi:10.1029/2006JD008062, 2007.
- Stevenson, D. S., Doherty, R., Sanderson, M., Johnson, C. E., Collins, W. J., and Derwent, R. G.: Impacts of climate change and variability on tropospheric ozone and its precursors, *Faraday Discuss.*, 130, 1–17, 2005.
- Tauler, R., Viana, M., Querol, X., Alastuey, A., Flight, R. M., Wentzell, P. D., and Hopke, P. K.: Comparison of the results obtained by four receptor modelling methods in aerosol source apportionment studies, *Atmos. Environ.*, 43, 3989–3997, 2009.

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

- Theloke, J. and Friedrich, R.: Compilation of a database on the composition of anthropogenic VOC emissions for atmospheric modeling in Europe, *Atmos. Environ.*, 41, 4148–4160, 2007.
- Thisse, T. R., van Oss, R. F., and Lenschow, P.: Determination of source contributions to ambient volatile organic compound concentrations in Berlin, *JAPCA J. Air Waste Ma.*, 49, 1394–1404, 1999.
- UBA: Nationale Trendtabellen für die deutsche Berichterstattung atmosphärischer Emissionen seit 1990 (Stand: 15. April 2011), available at: <http://www.umweltbundesamt-daten-zur-umwelt.de/umweltdaten/public/document/downloadPrint.do?ident=20681> (last access: 13 March 2014), 2011.
- US-EPA: Research in Action – EPA Positive Matrix Factorization (PMF) Model, available at: <http://www.epa.gov/heasd/research/pmf.html> (last access: 6 August 2013), 2011.
- Vingarzan, R.: A review of surface ozone background levels and trends, *Atmos. Environ.*, 38, 3431–3442, 2004.
- Watson, J. G., Chow, J. C., and Fujita, E. M.: Review of volatile organic compound source apportionment by chemical mass balance, *Atmos. Environ.*, 35, 1567–1584, 2001.
- WDCGG: World Data Centre for Greenhouse Gases: Hohenpeissenberg – DWD, available at: <http://ds.data.jma.go.jp/gmd/wdcgg/cgi-bin/wdcgg/accessdata.cgi?cntry=Germany&index=HPB647N00-DWD&para=VOCs&param=200612120859&select=parameter&parac=processing> (last access: 13 March 2014), 2013.
- Willis, R. D.: Final Report, Workshop on UNMIX and PMF as applied to PM<sub>2.5</sub>, Research Triangle Park, NC, 14–16 February 2000, 2000.
- WMO: GAW Report No. 172, WMO Global Atmosphere Watch (GAW) Strategic Plan: 2008–2015, 2007.
- Yuan, B., Shao, M., de Gouw, J., Parrish, D. D., Lu, S., Wang, M., Zeng, L., Zhang, Q., Song, Y., Zhang, J., and Hu, M.: Volatile organic compounds (VOCs) in urban air: how chemistry affects the interpretation of positive matrix factorization (PMF) analysis, *J. Geophys. Res.*, 117, D24302, 1–17, doi:10.1029/2012JD018236, 2012.
- Yue, W., Stolzel, M., Cyrus, J., Pitz, M., Heinrich, J., Kreyling, W. G., Wichmann, H. E., Peters, A., Wang, S., and Hopke, P. K.: Source apportionment of ambient fine particle size distribution using positive matrix factorization in Erfurt, Germany, *Sci. Total. Environ.*, 398, 133–144, 2008.

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Table 1.** Different treatments of missing values, values below the detection limit, and zero values.

#	Treatment
1	missing values replaced by species median; corresponding uncertainties set to four times the species median
2	missing values replaced by geometric mean; corresponding uncertainties set to four times the geometric mean
3	additional to 2: values below detection limit (dl) set to half the detection limit; corresponding uncertainties set to detection limit
4	additional to 3: zero values set to half the detection limit
5	additional to 3: uncertainty of <i>n</i> -hexane increased by 20%



**Can PMF identify sources of organic trace gases at a GAW site?**

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Table 2.** Mathematical diagnostics for the results of PMF computations for different numbers of factors.

	number of factors ( <i>k</i> )			
	5	6	7	8
number of samples	2335	2335	2335	2335
number of species	24	24	24	24
<i>Q</i> value from PMF	150 678	117 832	94 444	75 804
<i>F</i> <sub>peak</sub>	0.0	0.0	0.0	0.0
mean ratio NMHC modeled/NMHC measured	0.97	0.97	0.98	0.98
mean coefficient of determination ( $R^2$ )	0.86	0.88	0.89	0.89
number of species unexplained > 25 %	5	3	3	3
scaled residuals beyond 3σ	4354	2937	1971	1391

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

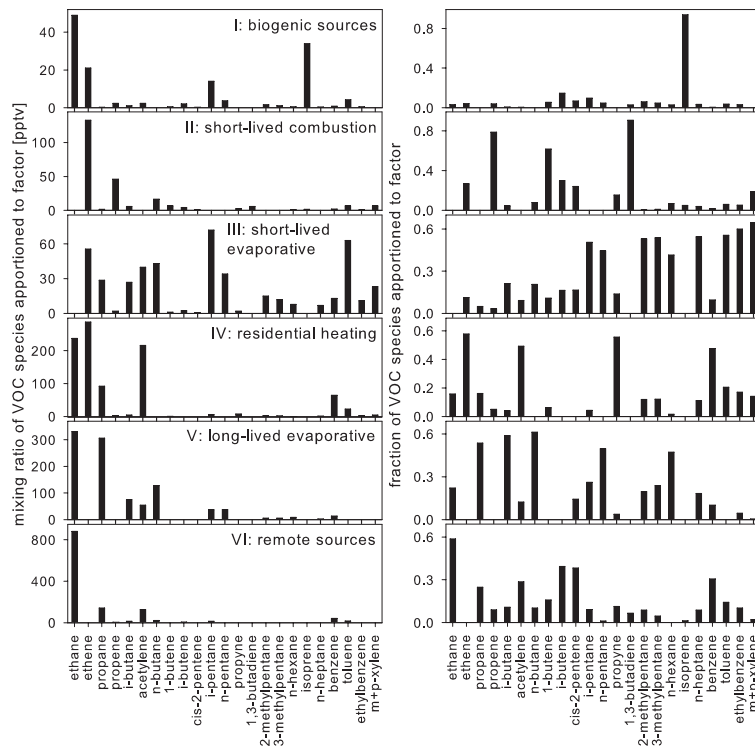
Interactive Discussion

**Table 3.** Mean values of calculated uncertainties for assumed time of reaction with OH of 0 h and 36 h.

Compound	Average uncertainty [%]	
	0 h	36 h
ethane	7.1	8.0
ethene	11.0	56.3
propane	6.2	13.5
propene	12.6	82.7
isobutane	6.9	21.6
acetylene	21.0	25.2
<i>n</i> -butane	26.0	39.2
1-butene	29.3	92.4
isobutene	16.9	96.2
<i>cis</i> -2-pentene	39.1	105.3
isopentane	11.2	33.8
<i>n</i> -pentane	6.8	32.4
propyne	33.5	55.1
1,3-butadiene	43.6	108.9
2-methylpentane	15.0	42.8
3-methylpentane	25.7	48.9
<i>n</i> -hexane	18.3	45.3
isoprene	33.3	115.8
<i>n</i> -heptane	12.9	48.4
benzene	8.6	16.3
toluene	10.9	44.8
ethylbenzene	30.7	59.5
<i>m</i> + <i>p</i> -xylene	25.2	72.8

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.

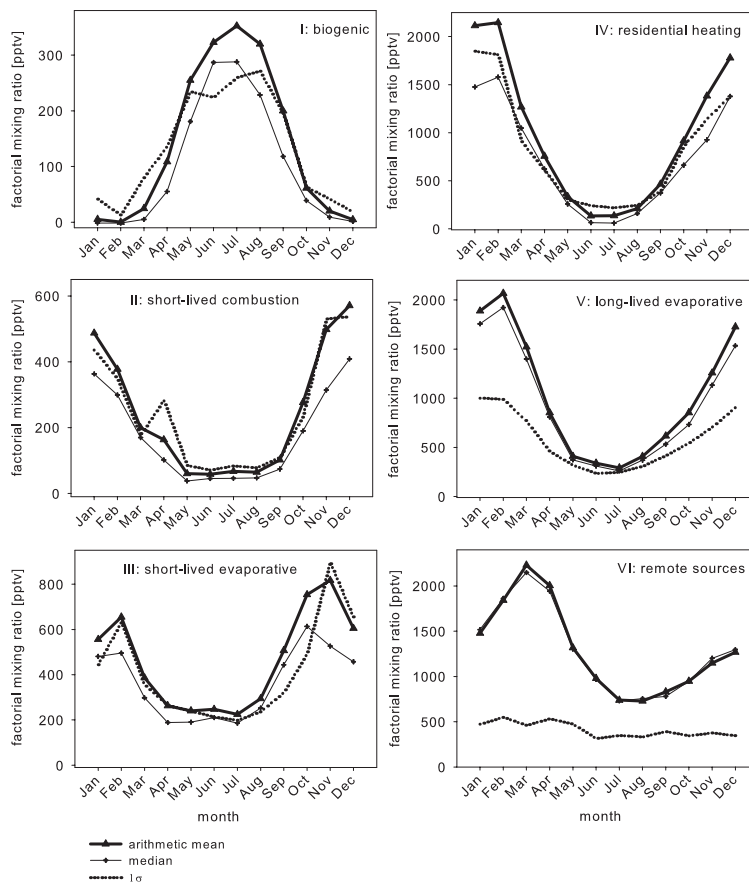


**Fig. 1.** Source composition profiles for the six factor solution calculated by PMF, left: mixing ratio of each species apporportioned to each factor [pptv], right: relative fraction of each species to the overall species mixing ratio. Note that the scales for each subplot are different due to large variations in absolute mixing ratios.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.



**Fig. 2.** Annual pattern of the mean, median, and standard deviation of the factorial mixing ratio [pptv] of the respective source category.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

⏪ ⏩

⏴ ⏵

Back Close

Full Screen / Esc

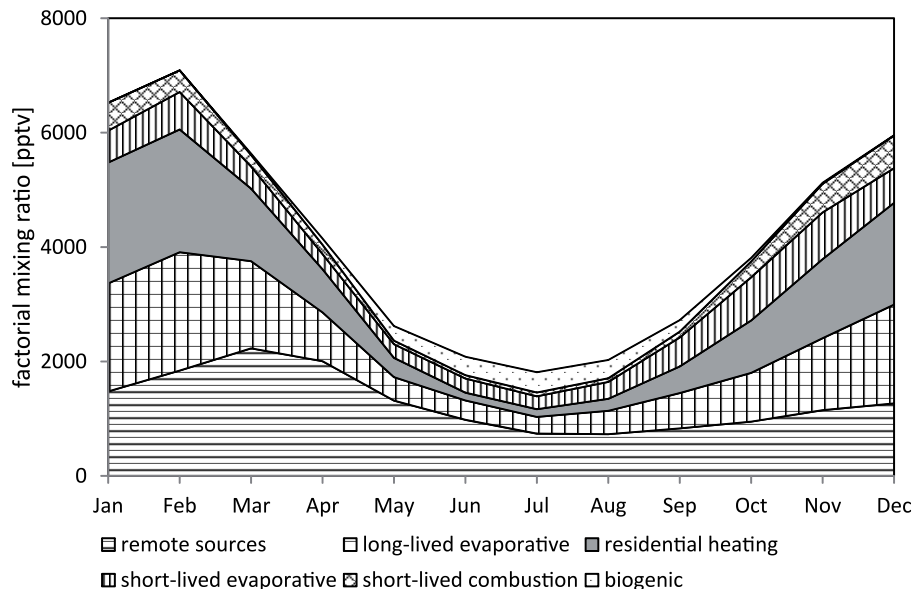
Printer-friendly Version

Interactive Discussion



**Can PMF identify sources of organic trace gases at a GAW site?**

M. Leuchner et al.



**Fig. 3.** Mean monthly variability of the contributions from the six factors.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

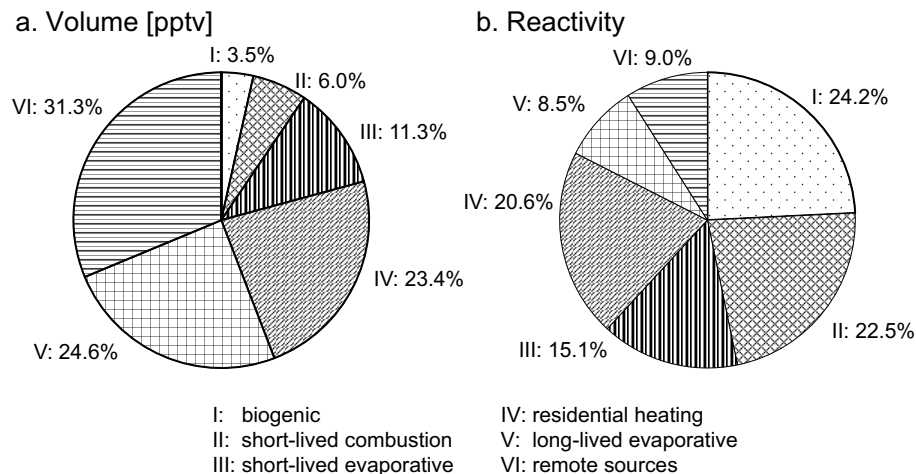
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Can PMF identify sources of organic trace gases at a GAW site?

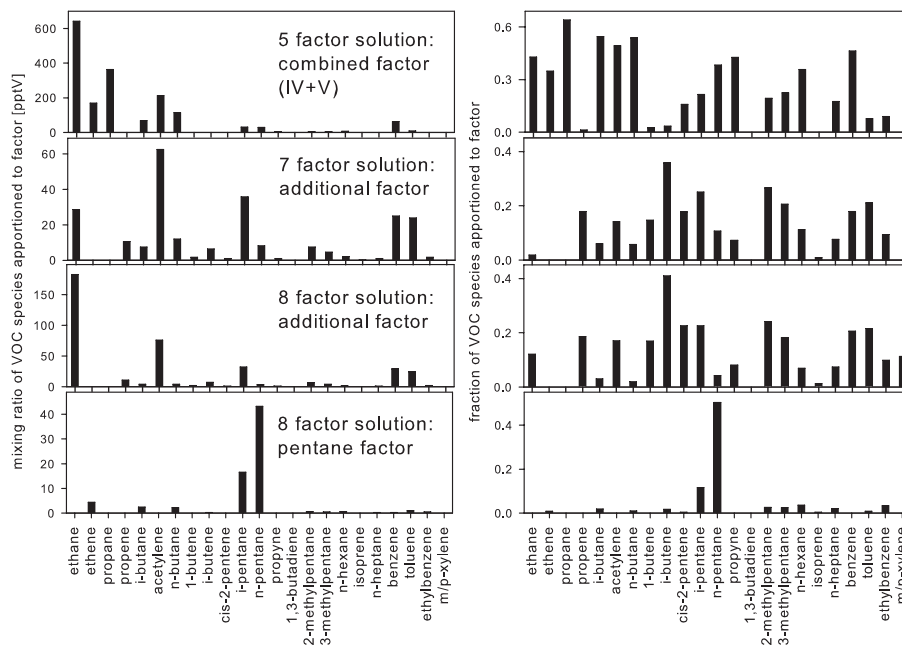
M. Leuchner et al.



**Fig. 4.** Contributions of individual factors, left: to the total amount of modeled NMHC [pptv], right: to the total OH reactivity of modeled NMHC [ $\text{pptcm}^3 \text{molec}^{-1} \text{s}^{-1}$ ].

Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.



**Fig. 5.** Source composition profiles of alternative solutions calculated by PMF, left: mixing ratio of each species apporportioned to each factor [pptv], right: relative fraction of each species to the overall species mixing ratio.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

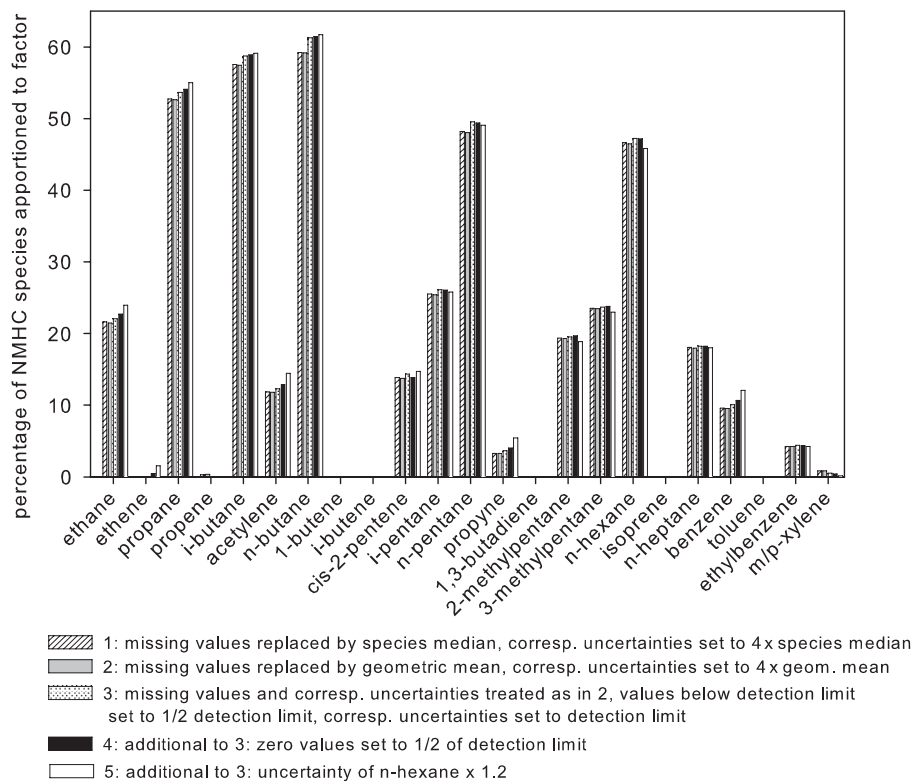
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.



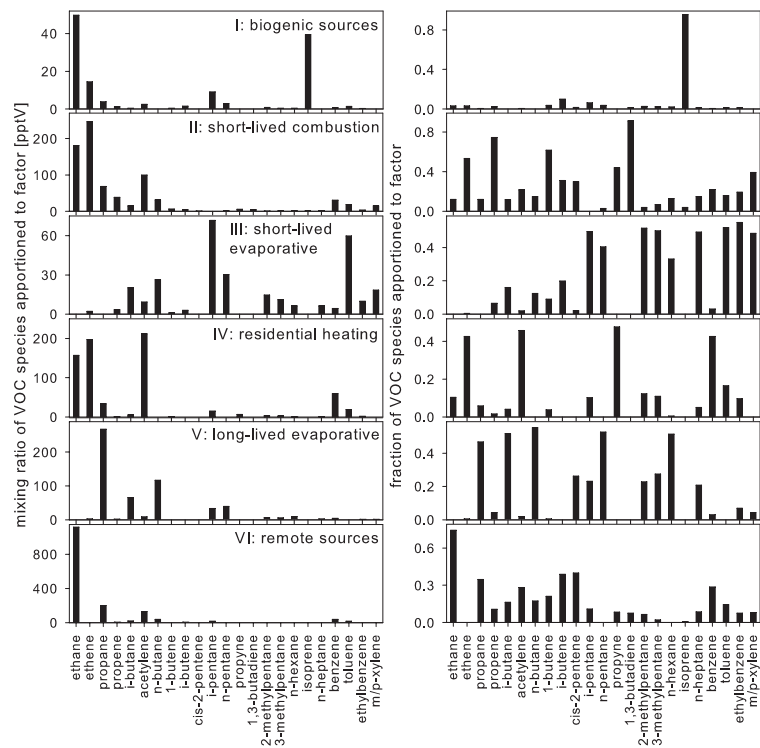
**Fig. 6.** Composition profiles of Factor V, showing the relative contribution of each substance to the factor, for five different treatments of missing and small values (see Table 1).

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



## Can PMF identify sources of organic trace gases at a GAW site?

M. Leuchner et al.



**Fig. 7.** Source composition profiles for the six factor solution calculated by PMF, assumed reaction time: 36 h, for isoprene: 0 h. left: mixing ratio of each species apportioned to each factor [pptV], right: relative fraction of each species to the overall species mixing ratio.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion