Source characteristics of oxygenated volatile organic compounds and hydrogen cyanide

Changsub Shim[1], Yuhang Wang[1], Hanwant B. Singh[2], Donald R. Blake[3], and Alex B. Guenther[4]

[1] Department of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia
[2] NASA Ames Research Center, Moffett Field, California
[3] Department of Chemistry, University of California, Irvine, California
[4] National Center for Atmospheric Research, Boulder, Colorado

Abstract. Trace gas measurements from TRACE-P, PEM-Tropics B, and INTEX-NA experiments are analyzed to examine the major source factors contributing to the observed variabilities of oxygenated volatile organic compounds and cyanides. The positive matrix factorization (PMF) is applied to coincidental measurements of 11 chemicals including CH_3OH, CH_3COCH_3, CH_3CHO, C_2H_2, C_2H_6, i-C_5H_{12}, CO, CH_3Cl, and $CHBr_3$. Measurements of HCN and CH_3CN are available for TRACE-P and INTEX-NA. We identify major source contributions from terrestrial biosphere, biomass burning, industry/ubran regions, and ocean. While source separation is fairly clear for TRACE-P and PEM-Tropics B data, PMF cannot resolve source contributions in INTEX-NA dataset due to active mixing processes over North America in summer and broad collocations of different types of sources in the region. Spatial and backtrajectory characteristics of these factors are examined. Large biomass burning contributions are found via long-range transport of aged air masses at high altitude. The biogenic contributions appear to result from dispersed continental sources. Industrial, urban, and oceanic concentrations clearly result from surface sources and transport. Based on TRACE-P and PEM-Tropics B data, we find that the terrestrial biogenic factor explains 80–88% of the CH_3OH variability, 20–40% of CH_3COCH_3, 7–35% of CH_3CHO, and 41% of HCN. The biogenic contribution to HCN is much larger than previous global estimates. Cyanogenesis in plants is likely a major emission process for HCN, not fully accounted for previously. Larger contributions than previous global estimations to CH_3COCH_3 and CH_3CHO by biomass burning and industry/urban sources likely reflect significant secondary productions from VOC oxidation. No evidence was found for large emissions of CH_3COCH_3 from the ocean. The oceanic CH_3CHO contribution shows large regional variations.

1. Introduction

Oxygenated volatile organic compounds (OVOCs) are a critical component of tropospheric photochemistry [e.g., *Singh et al.*, 2004]. Previous studies found significant emissions of OVOCs from the biosphere, which affect not only trace gas composition and chemistry but also aerosols and subsequently the radiative budget of the atmosphere [*e.g., Fuentes et al.,* 2000; *Guenther et al.*, 2000; *Fall,* 2003]. There have been many efforts to estimate the global budgets of OVOCs [*Guenther et al.*, 1995; *Galbally and Kristine*, 2002; *Heikes et al.,* 2002; *Jacob et al.,* 2002; *Singh et al.*, 2004]. Despite these efforts, there are still large uncertainties and many unknowns.

Hydrogen cyanide (HCN) is an often used indicator for biomass burning [e.g., *Rinsland et al*., 1998]. Additionally, biogenic processes are known to produce HCN and CH_3COCH_3 simultaneously [*Conn et al.,* 1991]. A specific metabolic process is cyanogenesis in plants (e.g. food crops, clovers, and eucalyptus leaves), which is part of plant self-defensive activity to fend off herbivores during growing seasons [*Fall et al.,* 2003; *Gleadow and Woodrow*, 2000]. Global estimate of the biogenic HCN production is highly uncertain partly because this complicated biochemical process is not well understood over the diverse terrestrial ecosystems. Furthermore, direct HCN measurements alone cannot be used to identify biogenic sources because of its long lifetime.

In order to identify and quantify major source contributions to OVOCs and cyanides, we analyze aircraft measurements made during the TRACE-P, PEM-Tropics B, and INTEX-NA experiments [*Raper et al.,* 2001; *Singh et al*., 2002; *Jacob et al.*, 2003]. We apply factor analysis using the positive matrix factorization (PMF) method to analyze these measurements. We show the characteristics of each source factor and calculate their contributions to the OVOCs and cyanides. Finally, we compare these results with the global budgets estimated by the previous studies.

2. Methodology

We choose a suite of chemicals, including methanol (CH_3OH), acetone (CH_3COCH_3), acetaldehyde (CH_3CHO), ethyne (C_2H_2), ethane (C_2H_6), i-pentane (i-C_5H_{12}), carbon monoxide (CO), methyl chloride (CH_3Cl), and bromoform ($CHBr_3$). The suite of chemicals is chosen to resolve the sources for OVOCs and cyanides; the limited tracer set is not meant to resolve different industrial/urban emission sources. Methanol has large terrestrial biogenic emissions [*Fall and Benson*, 1996; *Galbally and Kristine*, 2002]. Major CH_3COCH_3 sources are from the terrestrial biosphere [*Pouton*, 1990; *Jacob et al.*, 2002; *Singh et al*., 2004], biomass burning [*Jost et al.*, 2003], and the ocean [*Jacob et al.*, 2002]. CH_3CHO has known biogenic and oceanic sources [*Harry and Kimmerer*, 1991; *Singh et al.*, 2004], although the uncertainty in the budget is large. C_2H_2 and CO are good tracers for combustion. C_2H_6 is a good tracer for the usage of fossil fuel. Isopentane is a unique tracer for gasoline evaporation and gasoline powered motor vehicle emissions and hence the emissions from urban regions in general. It is not included in the PEM-Tropics B analysis because its concentrations in the remote tropical Pacific are generally below the detection limit. CH_3Cl is a typical tracer for biomass burning and biogenic emissions [*Yoshida et al.*, 2004 and references therein]. $CHBr_3$ is a tracer for ocean emissions [*Singh et al.*, 2004]. In addition, HCN and CH_3CN, available in TRACE-P and INTEX-NA, are included. They are generally considered biomass burning tracers [*de Gouw et al*., 2003; *Li et al.*, 2003; *Singh et al.*, 2003a].

To investigate the major sources contributing to the variability of these tracers, we use the PMF method, which has been applied previously to identify source factors of aerosols and trace gases [e.g., *Paatero*, 1997; *Wang et al.,* 2003; *Liu et al.,* 2005]. The factors are obtained from examining the covariance structure of the datasets [*Paatero*, 1997]. In the analysis, we selected only coincident measurements of the chemicals. Missing data are not used in order to reduce the uncertainty in the analysis.

The data matrix X of m measurements by n tracers can be decomposed by PMF for p factors as

$$X = GF + E \tag{1}$$

or

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \tag{2}$$

$$i = 1, \ldots, m;\ j = 1, \ldots, n;\ k = 1, \ldots, p.$$

where the m by p matrix G is the mass contributions of k^{th} factor to i^{th} sample (factor score), the p by n matrix F is the gravimetric average contributions of k^{th} factor to j^{th} chemical species (factor loadings), and the m by n matrix E is the error. We also use the explained variation (EV),

$$\mathrm{EV}_{kj} = \sum_{i=1}^{m} | g_{ik} f_{kj} | / \left[\sum_{i=1}^{m} \left(\sum_{k=1}^{p} | g_{ik} f_{kj} | + | e_{ij} | \right) \right] \tag{3}$$

to define the relative contributions of each factor to chemical species since the mixing ratios of different compounds are directly comparable. The analysis approach is the same as used by *Wang et al.* [2003].

As commonly found in principal component and factor analyses, the minimization of the error matrix does not land to a unique decomposition of matrix X to matrices G and F. Matrix rotation is one way to explore the constraints of the covariance structure. As in the work by *Wang et al.* [2003], the results presented here are not sensitive to rotation. For factor analysis, a key parameter is the choice of factor number. By evaluating the error matrix E [*Paatero*, 1997;

Lee et al., 1999; *Paatero et al.*, 2002], a range of factor numbers is determined. We then inspect the PMF results and determine the number of factors (with necessary rotation) that leads to clear separation of tracer signals. Following this procedure, we determined 6 factors for TRACE-P and 5 factors for PEM-Tropics B, which provide most physically meaningful results.

In the work by *Wang et al.* [2003], PMF analysis was used to identify air mass characteristics. That is the first objective of this work. The second objective is to investigate if the tracer signals in factor contributions can be used to gauge our understanding of emissions of these trace gases. In this case, the non-negativity in factor contributions calculated using an alternating least-square method in PMF [*Paatero,* 1997] offers a major advantage over less advanced techniques because source contributions by definition can only be positive. In order to attempt the second objective, the inversion of sources using factor analysis makes the assumption that factor profiles represent a linear combination of source contributions. The assumption is obviously false for chemicals with large secondary sources. It is also false if the chemicals have short lifetimes in the troposphere. The influence of transport is difficult to quantify in factor analysis. We note here two specific problems arising from transport. First, the anti-correlation driven by source location difference (for example, the often observed anticorrelation between ozone transported from the stratosphere and CO emitted from the surface), which tends to manifest in principal component analysis, does not pose a problem in PMF analysis. Second, transport processes do not “create” air masses with clean tracer characteristics; mixing during transport reduces source signals instead. As a result, the source characters cannot always be resolved by factor analysis. In this work, we show that the tracer source characteristics can be resolved for TRACE-P and PEM-Tropics B. However, transport/mixing processes during

INTEX-NA results in mixed source characteristics. PMF cannot clearly resolve the source contributions in that dataset.

Despite all the assumptions, we will attempt to compare source contributions inferred from PMF analysis of the two datasets with current source estimates. The empirical approach is advantageous over 3-D model studies when the source types, distributions, and strengths for the chemicals of interests are not well understood. Furthermore the PMF results are not affected by uncertainties in 3-D model chemistry and transport. Figure 1 shows the measurement regions during TRACE-P, PEM-Tropics B, and INTEX-NA experiments.

3. Results

3.1 TRACE-P

The TRACE-P experiment was conducted to investigate the Asian outflow to the Pacific during springtime (March – April, 2001) [*Jacob et al.,* 2003]. There are 386 coincidental measurements over the region of 15°−45°N and 114°E−124°W. 50% of the selected data is distributed below 3km. Figure 2 shows the EV profiles of 6 factors, and Figure 3 shows the absolute variability explained by each factor.

We label each factor based on its key chemical signatures. The “terrestrial biogenic” factor is characterized with large signals in CH_3OH (80%, 862 pptv) and CH_3COCH_3 (40%, 215 pptv). Terrestrial biosphere is the largest known source for CH_3OH [e.g., *Heikes et al.*, 2002] and a substantial source for CH_3COCH_3 [e.g., *Jacob et al.*, 2002]. HCN is often thought to originate from biomass burning. However, the terrestrial biogenic factor also explains a large fraction of the HCN variability (41%, 72 pptv). Cyanogenesis is likely to be a major biogenic process emitting HCN and its by-product, CH_3COCH_3 (40%, 215 pptv), from a variety of plants [*Pouton*, 1990; *Vetter,* 2000; *Fall,* 2003]. The biogenic factor explains 30% of CH_3CN (27 pptv). However, this fraction is not so significant considering its large measurement uncertainty of 35 pptv (Table 1). The biogenic CH_3CHO variability (35%, 79 pptv) is much greater than its measurement uncertainty (Table 1), providing support for its biogenic production [*Harry and Kimmerer*, 1991; *de Gouw et al.*, 1999; *Karl et al.*, 2002].

The "biomass burning" factor is loaded with largest signals of CH_3CN (69%, 63 pptv) and CH_3Cl (61%, 40pptv), two chemicals with large biomass burning sources [e.g., *Singh et al*., 2003a; *Yoshida et al.,* 2004]. HCN, known to have a large biomass burning source [e.g., *Li et al.*, 2000], also has a large signal (30%, 53 pptv). The signals of CH_3COCH_3 (19%, 104 pptv) and CO (23%, 22 ppbv) are also significant in this factor.

Two key combustion tracers, C_2H_2 (82%, 392 pptv) and CO (46%, 44 ppbv), appear in the "combustion" factor. Both gases are emitted from biomass burning and motor vehicle combustion. The similar correlations appear in the covariance structure of the dataset, resulting in one factor that represents a mixture of two sources. The source ambiguity, however, does not affect the gases of interest in this study.

The industry and urban pollution contributions are split into two factors. The first factor has significant signals of CH_3CHO (32%, 73 pptv), C_2H_2 (12%, 58 pptv), and i-C_5H_{12} (91%, 37 pptv) (industry/urban 1), while the second factor shows notable variabilities of C_2H_6 (78%, 888 pptv), CH_3COCH_3 (30%, 163 pptv), and CO (20%, 19 ppbv) (industry/urban 2). The first factor is more characteristic of urban emissions while the second is more of industrial emissions. The CH_3CHO variability explained by the industry/urban 1 factor is much higher than the primary biomass burning (5%) or industrial (< 1%) CH_3CHO source estimated by *Singh et al.* [2004], which likely reflects secondary production of CH_3CHO from VOC oxidation. The variability of

CH_3Cl (28%, 18 pptv) in the industry/urban 2 factor is much higher than the global industrial source fraction of (4%) [*McCulloch et al.*, 1999], reflecting the influence by biofuel combustion in relatively fresh industrial plumes sampled during TRACE-P. The most prominent component in the "oceanic" factor is $CHBr_3$ (84%, 0.75 pptv). This factor also explains 32% of CH_3CHO (72 pptv), supporting the notion of substantial CH_3CHO oceanic emissions [*Singh et al.*, 2003b].

Previous studies suggested large sources of CH_3Cl from the terrestrial biosphere [*Yoshida et al.*, 2004 and references therein]. However, the CH_3Cl variability explained by the terrestrial biogenic factor is negligible. The fairly small variability of the biogenic CH_3Cl likely reflects its dispersed terrestrial sources in the tropics [*Yoshida et al.*, 2004] and mixing during long-range transport. Considering the small contribution to CH_3Cl by its largest known source from the tropical terrestrial ecosystems to this factor, the much larger HCN signal here in turn likely indicates that measured enhancements originate from the forests at mid latitudes, not in the tropics. Figure 4 shows the kinematic 5-day back-trajectories of biogenic and biomass burning factors calculated with the European Centre for Medium-Range Weather Forecasts (ECMWF) meteorological fields [*Fuelberg et al*, 2003]. The biomass burning factor is associated with transport from tropical regions, while the terrestrial biogenic factor is related to transport from dispersed continental locations.

To further investigate the characteristics of the biomass burning factor, we examine how the factor correlates with latitude, altitude, and C_2H_6/C_3H_8 ratio (Figure 5). Higher C_2H_6/C_3H_8 ratios indicate more aged air masses [*Wang and Zeng*, 2004]. The correlation is strongest with C_2H_6/C_3H_8 ratio (r = 0.73), indicating aged biomass burning air masses. The positive correlation with altitude (r = 0.59) and negative correlation with latitude (r = -0.31) indicate transport of aged biomass burning air masses from tropical regions at high altitude after initial plume rise driven by combustion heating, consistent with backtrajectories (Figure 4).

The significant positive correlations of biomass burning factor with altitude and C_2H_6/C_3H_8 ratio are unique (Table 2). The terrestrial biogenic factor has no apparent correlation with latitude or altitude, implying a geographically dispersed distant source. The lack of correlation with C_2H_6/C_3H_8 ratio reflects in part that the biogenic sources are not collocated with alkane sources. The two industry/urban factors show significant anitcorrelations with altitude and C_2H_6/C_3H_8 ratio, representing relatively fresh plumes from low-altitude sources. Between them, industry/urban 2 factor with more chemical signatures of industrial than urban emissions have stronger anticorrlations with altitude and C_2H_6/C_3H_8 ratio. It also has clear latitude dependence. These characteristics are consistent with more industrial activities in the northeast of China and more dispersed urban distributions in the region. The oceanic factor has large anticorrelations with altitude and C_2H_6/C_3H_8 ratio, consistent with characteristics of fresh surface emissions. The anticorrelation with C_2H_6/C_3H_8 ratio reflects in part a positive correlation of the hydrocarbon ratio with altitude.

3.2 PEM-Tropics B

The PEM-Tropics B campaign was conducted during the wet season over the tropical Pacific (March – April, 1999) [*Raper et al.,* 2001]. 581 coincident measurements of 8 chemicals over the region of 36°S−35°N and 107°W−162°E are analyzed. 63% of the selected data is distributed above 5 km. We find no significant bias in the spatial distributions of selected dataset compared to the entire dataset.

Five factors are resolved by PMF (section 2). Figure 6 shows the EV profiles and Figure

7 shows the absolute chemical variability. Compared to TRACE-P (Figs. 2 and 3), the tracer variability during PEM-Tropics B is much lower, leading to more ambiguity in interpretation of the analysis results because observed variabilities were at times smaller than measurement uncertainties (Table 1). As one could expect, the tracer variabilities of industy/urban related factors in PEM-Tropics B are much smaller than those in TRACE-P, which reflects the sampling of photochemically aged air plumes during PEM-Tropics B.

As previously, we label each factor based on its key chemical signatures. The "terrestrial biogenic" factor is characterized by the large variability of CH_3OH (88%, 653 pptv). The signal of CH_3COCH_3 is below the measurement uncertainty. The "biomass burning" factor has large signals of CH_3COCH_3 (55%, 98 pptv), CH_3CHO (64%, 30 pptv), CO (52%, 11 ppbv), and CH_3Cl (82%, 44 pptv). The "combustion/industry" factors reflect a mix of biomass burning, automobile combustion, industry, and urban sources. There are substantial signals of C_2H_2 (84%, 28 pptv) and C_2H_6 (9%, 21 pptv) in the combustion/industry 1 factor, and large signals of C_2H_6 (76%, 188 pptv) and CO (34%, 8 ppbv) in the combustion/industry 2 factor. The CH_3Cl fraction (13%, 7 pptv) in the latter factor is below the measurement uncertainty. The "oceanic" factor explains 94% of $CHBr_3$ variability (0.55 pptv) and does not show significant variability in the other chemicals.

We analyze the correlations between factors and latitude, altitude, and C_2H_6/C_3H_8 ratio to help identify the source characteristics (Table 3). As for TRACE-P, the terrestrial biogenic factor shows no significant correlations reflecting geographically dispersed sources. The biomass burning factor is anticorrelated with latitude ($r = -0.51$), implying a southern hemispheric origin particularly from Africa and Australia (Figure 8). As in TRACE-P, it is positively correlated with altitude and C_2H_6/C_3H_8 ratio, indicating long-range transport of aged pollutants after initial rise due to combustion heating. The combustion/industry 1 and combustion/industry 2 factors show significant positive correlations with latitude and negative correlations with C_2H_6/C_3H_8 ratio, implying most long-range transport of industrial/urban/combustion emissions from the northern hemisphere [*Wang et al*., 2001]. As in Trace-P, the oceanic factor is the only one with significant anticorrelation with altitude. It also has a week anticorrelation with C_2H_6/C_3H_8 ratio, representing mixing of relatively fresh emissions from the ocean with transport from C_2H_6 and C_3H_8 from northern industrial regions.

3.3 INTEX-NA

The INTEX-NA experiment is conducted to investigate the formation and transport of tracers and aerosols and inflow and outflow of pollutants over North America during the summer (July – August, 2004) [*Singh et al*., 2002]. Two characteristics separate INTEX-NA from TRACE-P and PEM-Tropics B. First, mixing processes are much more active over the continent in summer than spring (TRACE-P) or over the tropical Pacific (PEM-Tropics B). Second, North American urban and industrial sources are more dispersed than over East Asia, where major industrial and urban emissions are densely located near coastal regions. As a result, we find that source-separation is much more ambiguous in INTEX-NA dataset than the other two. In the selected dataset, 35% of the data is measured below 2 km; 50% is measured above 5 km. We find 5 factors in the PMF analysis. Figure 9 shows the EV profiles. The factors are named after the tracer that has largest EV. The C_2H_2 factor is a combustion factor with a large CH_3OH signal (> 25%), representing either a large industrial related CH_3OH source over North America or a mixture of combustion and biogenic emissions because we cannot define a clear biogenic factor. The CH_3Cl factor with large signals of biomass burning (CH_3Cl, CO, HCN, and CH_3CN) and biogenic tracers (CH_3OH, CH_3COCH_3, and HCN) likely represent a mixture of the two. The CH_3Br, C_2H_6, and i-C_5H_{12} factors are dominated by a single species. Photochemical oxidation and mixing during INTEX-NA (summer) are considerably faster than TRACE-P (spring). Assuming similar emissions, we would expect lower concentrations in INTEX-NA than TRACE-P. Inspection of vertical profiles of trace gases we selected (not shown) shows that is true for most species. However, concentrations of CH_3OH, acetone, and HCN are much higher in INTEX-NA than TRACE-P. We postulate that these enhancements are due at least in part to larger biogenic emissions in summer. However, the analysis method used in this study cannot be applied to evaluate our hypothesis.

4. Comparisons to global estimations

Despite the limitations of PMF methods (section 2), it is still instructive to compare our results with previous global budget estimations. Additional precautions must also be heeded to prevent oversimplifications. First, our results are drawn from only two aircraft missions. Second, the factor contributions do not necessarily translate into source strengths. Third, it is generally hard to distinguish primary sources from secondary chemical production in the factor analysis. However, the following assessment points to the directions of needed research when large discrepancies are found between previous emission estimations and our results. Table 4 summarizes the contributions of various factors to OVOCs and HCN for TRACE-P and PEM-Tropics B; previous global estimates are also listed for comparison purposes.

The 80–88% variability of CH_3OH in the terrestrial biogenic factor from this study is consistent with the estimated global biogenic fractions of 62–81% [*Heikes et al.,* 2002; *Singh et* *al.,* 2004; *Jacob et al*., 2005]. For CH_3COCH_3, the biogenic fractions of this study (20 − 40 %) are also comparable with the global estimates (27−53%) [*Jacob et al.,* 2002; *Singh et al.,* 2004].

Large terrestrial biogenic HCN variability (41%) is found during TRACE-P; it is considerably higher than the global biogenic fraction (~18%) estimated by *Li et al.* [2003, Table 4]. The cyanogenic process in plants is known to emit both CH_3COCH_3 and HCN. It appears to be a key contributor to HCN emissions [*Fall*, 2003] from geographically dispersed source regions. Contribution of biomass burning to HCN (~30%) in TRACE-P is significantly smaller than the global estimates (> 80%). The industry/urban/ocean factor contributions are uncertain since they are less than the measurement uncertainty of HCN (~ 25%).

The CH_3COCH_3 variabilities explained by industry/urban and biomass burning factors (8 −30% and 19 − 55%, respectively) are much higher than the estimated primary sources (1 − 4% and 4 − 10%, respectively). The total difference between PMF factor contributions and previously estimated primary sources, however, are comparable to the 30 – 50% contribution from VOC oxidation [*Jacob et al*., 2002; *Singh et al*., 2004], indicating secondary productions in those factors. There is no clear evidence for significant oceanic CH_3COCH_3 emissions, which is consistent with the recent measurements by *Marandino et al.* [2005].

The terrestrial biogenic CH_3CHO variability is quite different between TRACE-P and PEM-Tropics B (35% and 7%, respectively), likely reflecting relatively fast photochemical loss of CH_3CHO or its precursors during long-range transport to the remote PEM-Tropics B region. The corresponding global fraction estimated by *Singh et al.* [2004] is 21%. We find large contributions from industry/urban 1 (32%) for TRACE-P and from biomass burning (64%) for PEM-Tropics B. Given the relatively short lifetime of CH_3CHO (~ 1 day), our results indicate that secondary production from VOCs, likely produced in combustion processes, is a major source of observed CH_3CHO. The oceanic CH_3CHO variability is significant during TRACE-P (32%) but not PEM-Tropics B, indicating the possibility of regional variations in oceanic CH_3CHO emissions. In comparison, the estimated global oceanic contribution is much larger at 63% [*Singh et al.*, 2004]. In general, the interpretation of PMF results of short-lived species such as CH_3CHO is more ambiguous. More detailed and quantitative analysis will require 3-D chemical transport model simulations that explicitly treat the emissions, secondary production, oxidation, and transport of CH_3CHO.

5. Conclusions

We apply an advanced factor analysis method, PMF, to analyze the source contributions to OVOCs and cyanides measured in three aircraft experiments. In order to identify contributions by emissions from the terrestrial biosphere, biomass burning, industry/urban regions, and ocean, we select a suite of trace chemicals including CH_3OH, CH_3COCH_3, CH_3CHO, C_2H_2, C_2H_6, i-C_5H_{12}, CO, CH_3Cl, and $CHBr_3$. Two cyanide compounds, HCN and CH_3CN are available and included in TRACE-P and INTEX-NA analyses. While feasible for TRACE-P and PEM-Tropics B measurements, PMF analysis cannot clearly separate source contributions in INTEX-NA measurements. More active transport/mixing processes during summer over North America and broad collocations of different types of sources in the region are likely the reasons.

We resolve 6 factors for TRACE-P and 5 factors for PEM-Tropics B. In TRACE-P, the terrestrial biogenic factor is characterized by significant variability of CH_3OH (80%, 862 pptv) and CH_3COCH_3 (40%, 215 pptv). This factor also shows large fraction of HCN variability (41%, 72 pptv). Backtrajectory analysis and a lack of correlation of this factor with altitude, latitude, or

C_2H_6/C_3H_8 ratio imply widely dispersed continental sources. The significant variabilities of HCN and CH_3COCH_3 in this factor likely reflect emissions from cyanogenesis in diverse plants.

The biomass burning factor is characterized by large variabilities of CH_3CN (69%, 63 pptv), CH_3Cl (61%, 40 pptv), HCN (30%, 53 pptv), and CO (23%, 22 ppbv). It is the only factor that positively correlates with altitude and C_2H_6/C_3H_8 ratio, indicating long-range transport of aged biomass burning air masses at high altitude. Backtrajectories and a negative correlation with latitude indicate that it originates from subtropical/tropical regions.

The combustion factor, characterized by variabilities of C_2H_2 (82%, 392 pptv) and CO (46%, 44 ppbv), reflects a mixture of biomass burning, industrial, and urban sources. Significant negative correlations of this factor with altitude and C_2H_6/C_3H_8 ratio indicate relatively fresh surface emissions. There are two industry/urban factors. Industry/urban 1 factor, characterized by large variability of $i\text{-}C_5H_{12}$ (91%, 37 pptv), is affected more by urban emissions. The second, characterized by a strong signal of C_2H_6 (78%, 888 pptv), is affected more by industrial emissions. As a result, the second factor has a stronger dependence on latitude likely due to the high density of industries over the northeast of China. Both factors are negatively correlated with altitude and C_2H_6/C_3H_8 ratio, implying surface sources and transport.

The oceanic factor is characterized by largest variability of $CHBr_3$ (84%, 0.75 pptv). It is clearly a surface source. The substantial CH_3CHO variability (32%, 72 pptv) in this factor implies a significant oceanic contribution to CH_3CHO.

Compared with TRACE-P, factor contributions in PEM-Tropics B are smaller because of transport over longer distance to the region. The terrestrial biogenic factor is similarly characterized by large variability of CH_3OH (88%, 653 pptv). The biomass burning factor has large variabilities of CH_3Cl (82%, 44 pptv), CH_3CHO (64%, 30 pptv), CH_3COCH_3 (55%, 98 pptv), and CO (52%, 11 ppbv). The factor correlations with altitude, latitude, and C_2H_6/C_3H_8 ratio for these factors are similar to TRACE-P, indicating widely dispersed biogenic sources and long-range transport of aged biomass burning air masses at high altitude.

The two combustion/industry factors show large signals of C_2H_2 (84%, 28 pptv, combustion/industry 1), C_2H_6 (76%, 188 pptv, combustion/industry 2), and CO (34%, 8 ppbv, combustion/industry 2). Both factors are positively correlated with latitude and negatively correlated with C_2H_6/C_3H_8 ratio, representing long-range transport of industrial/urban/combustion emissions from the northern hemisphere. The oceanic factor characterized by $CHBr_3$ from surface emissions does not have a significant CH_3CHO signal.

We compare the source contributions from two aircraft field experiments to previous global estimations; identifying large discrepancies points to the directions of needed research. First, the terrestrial biogenic contribution of CH_3OH in this study is fairly consistent with the global biogenic emission estimates (80 – 88% and 62 – 81%, respectively). Second, the terrestrial biogenic contribution of HCN in TRACE-P (41%) is substantially higher than the global emission estimates (0 – 18%), which suggests that the cyanogenesis in plants from widely dispersed regions is likely to be a major source of HCN in addition to biomass burning. Third, the biogenic contribution to CH_3COCH_3 variability is comparable with the global emission estimates (20 – 40% and 27 – 53%, respectively). However, there are much larger CH_3COCH_3 industry/urban and biomass burning contributions (8 – 30% and 19 – 55%, respectively) in this study than previous global emission (1 – 4% and 4 – 10%, respectively), reflecting the importance of secondary productions. We do not find large oceanic contributions to CH_3COCH_3. Fourth, considering its relatively short lifetime, the large contributions to CH_3CHO variability from industry/urban 1 (32%) for TRACE-P and from biomass burning (64%) for PEM-Tropics B imply that secondary production from combustion/industrial VOCs is likely an important source. The oceanic CH_3CHO contribution (10 – 32%) shows considerable regional dependence and it is lower than previous global emission estimates (63%). Based on this analysis, three specific areas emerge for future research: the global sources of CH_3COCH_3 and HCN from the cyanogenic process in plants, the VOC precursors to sustain CH_3COCH_3 and CH_3CHO concentrations in industry/urban and biomass burning plumes, and the magnitudes and regional variability of oceanic CH_3CHO emissions.

Acknowledgements. The work was supported by the National Science Foundation Atmospheric Chemistry Program and NASA's Global Tropospheric Experiment Program.

References

Conn E E. (1991), The metabolism of a natural product –lessons learned from cyanogenic glycosides, *Planata Medicda, 57*, s1-s9.

de Gouw, J. A. et al. (1999), Emissions of volatile organic compounds from cut grass and clover are enhanced during the drying process, *Geophys. Res. Lett, 26*, 811−814.

de Gouw, J. A., C. Warneke, D. D. Parrish, J. S. Holloway, M. Trainer, and F. C. Fehsenfeld (2003), Emission source and ocean uptake of acetonitrile (CH3CN) in the atmosphere, *J. Geosphys. Res, 108(D11)*, 4329, doi: 10.1029/2002JD002897, 2003.

Fall, R. (2003), Abundant oxygenates in the atmosphere: A biochemical perspective, *Chem. Rev., 103*, 4,941−4, 951.

Fuentes, J. D. et al. (2000), Biogenic hydrocarbons in the atmospheric boundary layer: A review, *Bull. Amer. Meteor. Soc., 81*, 1,537−1,575.

Fuelberg, H. E., R. E. Newell, D. J. Westberg, J. C. Maloney, J. R. Hannan, B. D. Martin, M. A. Avery, and Y. Zhu (2001), A meteorological overview of the second Pacific exploratory in the tropics, *J. Geophys. Res., 106 (D23)*, 32427-32443.

Fuelberg, H. E., C. M. Kiley, J. R. Hannan, D. J. Westberg, M. A. Avery, and R. E. Newell (2003), Meteorological conditions and transport pathways during the Transport and Chemical Evolution over the Pacific (TRACE-P) experiment, *J. Geophys. Res., 108(D20),* 8782, doi:10.1029/2002JD003092.

Fukui, Y., and P. V. Doskey (1998), Air-surface exchange of nonmethane organic compounds at a grassland site: Seasonal variations and stressed emissions, *J. Geophys. Res.,* 103, 13,153−13,168.

Galbally, I. E., and W. Kirstine (2002), The production of methanol by flowering plants and the global cycle of methanol, *J. Atmos. Chem., 43,* 195−229.

Gleadow and Woodrow (2000), Temporal and spatial variation in cyanogenic glycosides in Eucalyptus cladocalyx, *Tree Physiology, 20,* 591-598.

Guenther, A. et al. (1995), A global-model of natural volatile organic-compound emissions, *J. Geophys. Res., 100,* 8,873−8.892.

Guenther, A. et al. (2000), Natural emissions of non-methane volatile organic compounds; carbon monoxide, and oxides of nitrogen from North America, *Atmos. Environ., 34*, 2,205−2,230.

Harry, D. E., and T. W. Kimmerer (1991), Molecular genetics and physiology of alcohol dehydrogenase in woody plants, *Forest Ecol. Manag., 43,* 251−272.

Heikes, B. G. et al. (2002), Atmospheric methanol budget and ocean implication, *Global Biochem. Cycles,* 16(4), 1133, doi:10.1029/2002GB001895.

Jacob, D. J. et al. (2003), Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft mission: Design, execution, and first results, *J. Geophys. Res., 108* (D13), 1 –19.

Jacob, D. J. et al. (2002), Atmospheric budget of acetone, *J. Geophys. Res., 107* (D12), 10.1029/2001JD000694.

Jacob, D. J. et al. (2005), Global budget of methanol: Constraints from atmospheric observations, *J. Geophys. Res., 110*, D08303, doi:10.1029/2004JD005172.

Jost, C. et al. (2003), Trace gas chemistry in a young biomass burning plume over Namibia: Observations and model simulations, *J. Geosphys. Res., 108*(D13), 8482, doi:10.1029/2002JD002431.

Karl, T. et al. (2002), Transient releases of acetaldehyde from tree leaves – products of a pyruvate overflow mechanism?, *Plant Cell Environ., 25,* 1,121−1,131.

Li, Q. B. et al. (2000), Atmospheric Hydrogen Cyanide (HCN): Biomass burning source, Ocean sink?, *Geophys. Res. Lett.*, *27,* 357-360.

Li, Q. B. et al. (2003), A global three-dimensional model analysis of the atmospheric budget of HCN and CH_3CN: Constraints from aircraft and ground measurements, *J. Geophys. Res.*, *108* (D21), 8827, doi:10.1029/2002JD003075.

Liu, W. et al. (2005), Atmospheric aerosol over two urban-rural pairs in the southestern United States: Chemical composition and sources, *Atmos. Environ.*, *39,* 4,453−4,710.

Marandino, C. A., et al. (2005), Oceanic uptake and the global atmospheric acetone budget, *Geophys. Res. Lett., 32,* L15806, doi:10.1029/2005GL023285.

McCulloch, A. et al. (1999), Global emissions of hydrogen chloride and chloromethane from coal combustion, incineration, and industrial activities: Reactive chlorine emissions inventory, *J. Geophys. Res., 104*, 8,391−8,403.

Paatero, P. (1997), Least squares formulation of robust non-negative factor analysis, *Chemometr. Intell. Lab. Syst., 37*, 23−35.

Paatero, P. et al. (2002), Understanding and controlling rotations in factor analytic model, *Chemometr. Intell. Lab. Syst., 60*, 253−264.

Pouton, J. E. (1990), Cyanogenesis in plants, *Plant Physiol., 94*, 401−405.

Raper, J. L. et al. (2001), Pacific Exploratory Mission in the tropical Pacific: PEM-Tropics B, March−April 1999, *J. Geophys. Res., 106*, 32,401−32,425.

Rinsland, C. P. et al. (1998), ATMOS/ATLAS 3 infrared profile measurements of trace gases in the November 1994 tropical and subtropical upper troposphere, *J. Quant. Spectrosc. Radiat. Transfer, 60*, 891–901.

Singh, H, D. Jacob, L. Pfister, and J. Crawford (2002), INTEX-NA: Intercontinental Chemical Transport Experiment - North America, from http://cloud1.arc.nasa.gov/docs/intex-na/white_paper.pdf.

Singh, H. et al. (2003a), In situ measurements of HCN and CH_3CN over the Pacific Ocean: Sources, sinks, and budgets, *J. Geophys. Res.*, *108* (D20), 8795, doi: 10.1029/2002JD003006

Singh, H. et al. (2003b), Oxygenated volatile organic chemicals in the ocean: Inferences and implications based on atmosphere observations and air-sea exchange models, *Geophys. Res. Lett.*, *30*, 16, 1862, doi: 10. 1029/2003GL017933.

Singh, H. et al. (2004), Analysis of the atmospheric distribution, sources, and sinks of oxygenated volatile organic chemicals based on measurements over the Pacific during TRACE-P, *J. Geophys. Res.*, *109*, D15S07, doi:10.1029/2003JD003883.

Tyndall, G. S., J. J. Orlando, T. J. Wallington, and M. D. Hurley (2001), Products of the chlorine-atom- and hydroxyl-radical-initiated oxidation of CH_3CN, *J. Phys. Chem, A, 105*, 5380-5384.

Vetter, J. (2000), Pant cyanogenic glycosides, *Toxicon*, *38*, 11−36.

Wang, Y., et al. (2001), Factors Controlling Tropospheric O_3, OH, NO_x, and SO_2 over the Tropical Pacific during PEM-Tropics B, *J. Geophys. Res.*, *106*, 32,733-32,748.

Wang, Y. et al. (2003), Intercontinental transport of pollution manifested in the variability and seasonal trend of springtime O_3 at northern middle and high latitudes, *J. Geophys. Res.*, *108* (D21), 46 83, doi: 10.1029/2003JD003592.

Wang, Y., and T. Zeng (2004), On tracer correlations in the troposphere: The case of ethane and propane, *J. Geophys. Res.*, *109*, D24306, doi:10.1029/2004JD005023.

Yoshida, Y. et al. (2004), A three-dimensional global model study of atmospheric methyl chloride budget and distributions, *J. Geophys. Res., 109*, D24309, doi: 24310.21029/2004JD004951.

Table 1. Relative and average absolute measurement uncertainties

	CH_3OH	CH_3COCH_3	CH_3CHO	HCN	CH_3CN	C_2H_2	C_2H_6	C_5H_{12}	CO	CH_3Cl	$CHBr_3$
Relative uncertainty	10%	10%	10%	~25%	~25%	1%	1%	1%	7%	2%	5%
TRACE-P[1]	110	90	30	50	35	5	15	2	7	12	0.03
PEM-Tropics B[1]	90	40	10	-	-	3	3	-	4	12	0.03

[1] Unit is pptv except for CO (ppbv).

Table 2. Correlations (R) between factor scores and latitudes, altitudes, and C_2H_6/C_3H_8 ratio for TRACE-P.

Factors	Latitude	Altitude	C_2H_6/C_3H_8 ratio
Terrestrial biogenic	0.06	0.01	0.08
Biomass burning	-0.31	0.59	0.73
Combustion	0.1	-0.21	-0.37
Industry/urban 1	0.04	-0.27	-0.39
Industry/urban 2	0.29	-0.53	-0.78
Oceanic	0.1	-0.57	-0.55

Table 3. Correlations (R) between factor scores and latitudes, altitudes, and C_2H_6/C_3H_8 ratio for PEM-Tropics B.

Factors	Latitude	Altitude	C_2H_6/C_3H_8 ratio
Terrestrial biogenic	0.01	-0.09	-0.14
Biomass burning	-0.51	0.16	0.45
Combustion/industry 1	0.6	-0.1	-0.54
Combustion/industry 2	0.68	-0.07	-0.57
Oceanic	0.13	-0.42	-0.17

Table 4. Comparison factor contributions with global estimates[1] of OVOCs (%)

	CH3OH		CH_3COCH_3		CH3CHO		HCN	
	Global estimates	PMF	Global estimates	PMF	Global estimates	PMF	Global estimates	PMF
Biogenic	60 – 80	80 – 88	27 – 53	20 – 40	17.5	7 – 35	0 – 18	41
BB	5 – 9	3 – 16	4 – 10	19 – 55	5	1 – 68	80 – 98	30
Industry/urban	2 – 4	3 – 5	1 – 4	8 – 30	0.5	1 – 28	0 – 2	10
Ocean	0	0.5 – 3	0 – 28	2 – 4	62.5	6 – 36	0	2
Others[2]	15 – 33	-	30 – 50	-	15	-	-	17

[1] Global estimates are compiled from Heikes et al. [2002], Li et al. [2003], Singh et al. [2004], and Jacob et al. [2002, 2005]. The contributions less than measurement uncertainty in PMF are expressed in Italic face.

[2] Unexplained fraction by PMF.

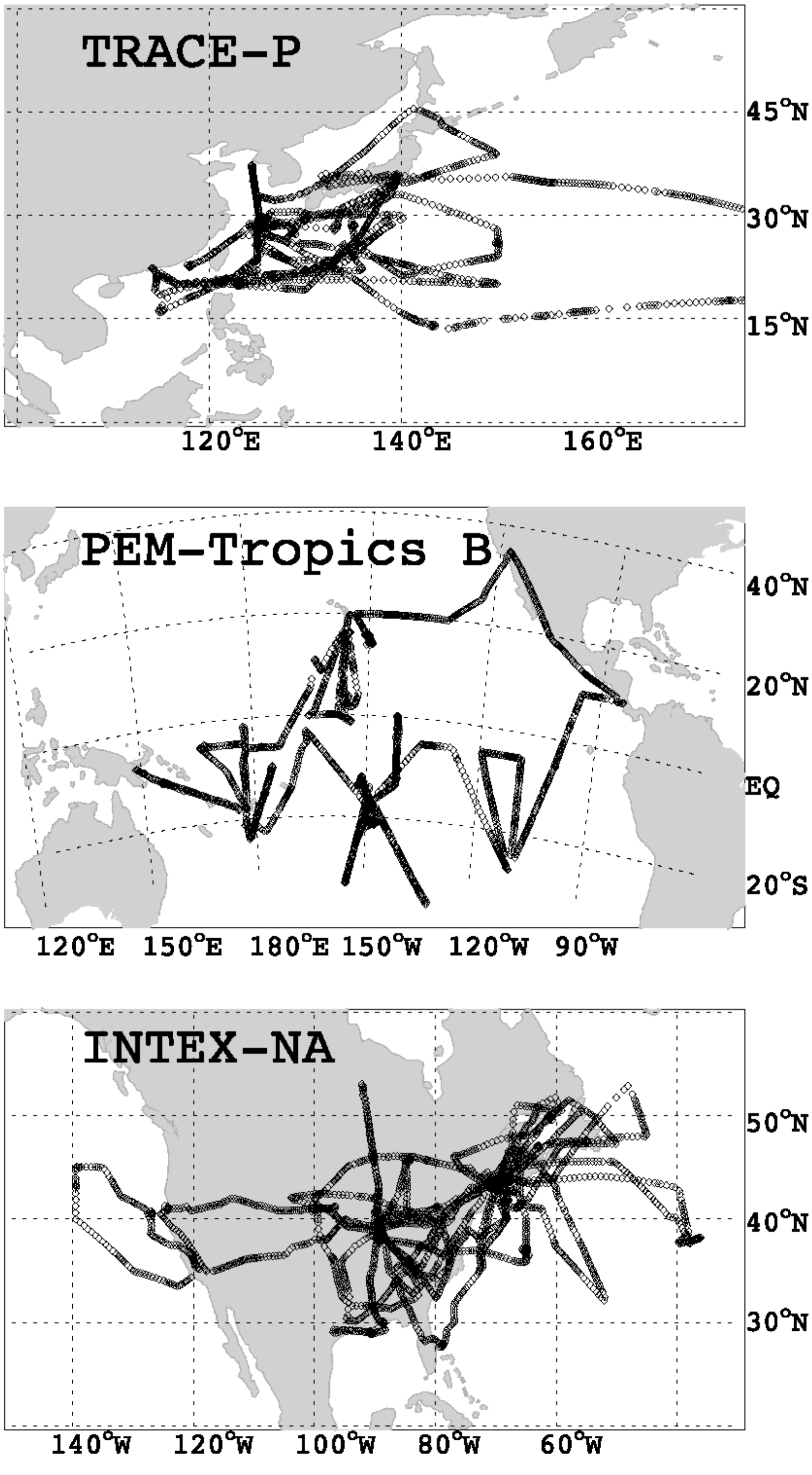


Figure 1. Location of aircraft measurements during TRACE-P (top), PEM-Tropics B (middle), and INTEX-NA (bottom).

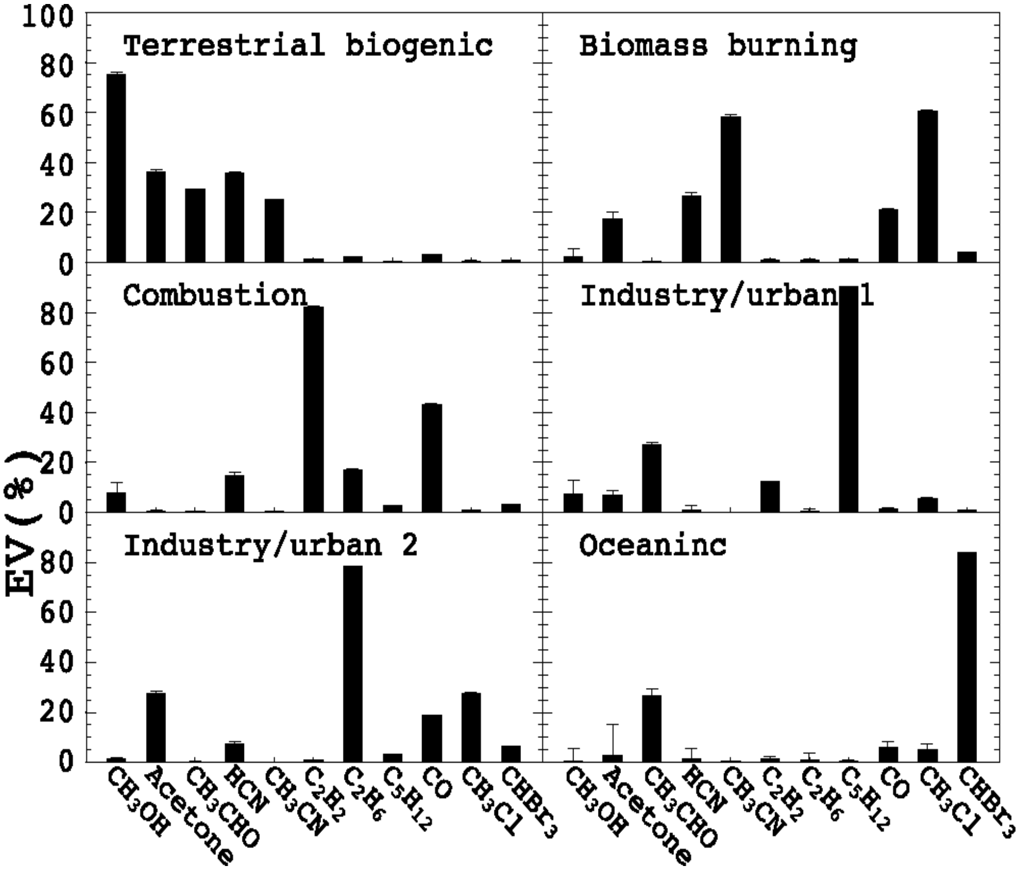


Figure 2. Explained variations of chemical compounds in the 6 factors for TRACE-P. The factors are named based on the chemical characteristics.

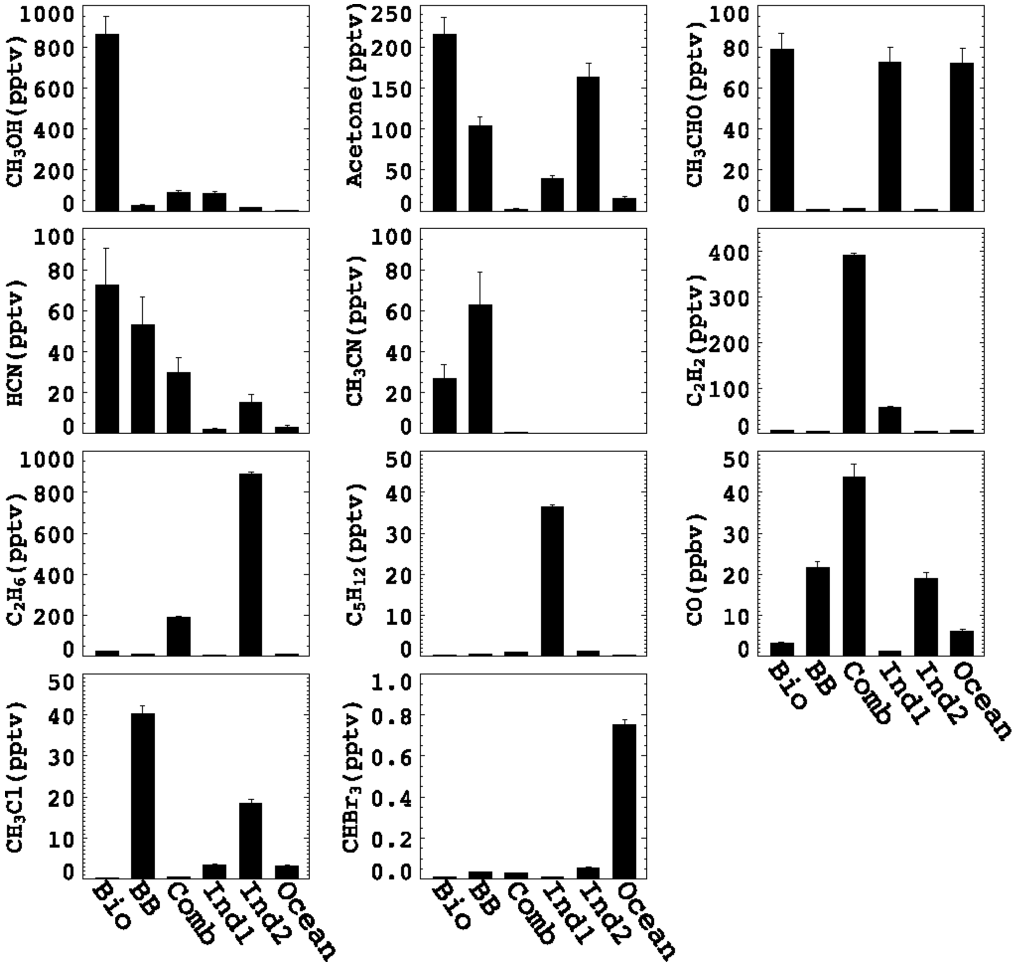


Figure 3. The variabilities in mixing ratio explained by each factor for TRACE-P. “Bio” denotes terrestrial biogenic; “BB” denotes biomass burning; “Comb” denotes combustion; and “Ind” denotes industry/urban. The vertical bars show the uncertainties of the PMF analysis.

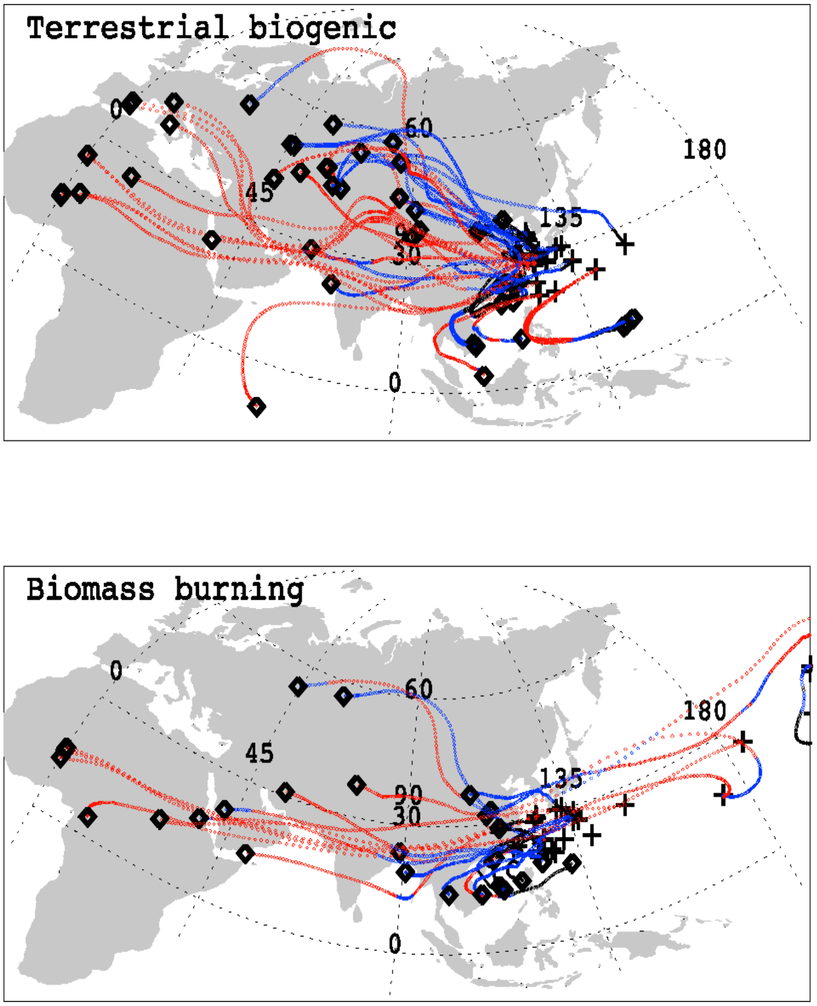


Figure 4. Five day back trajectories of data points that rank in the top 10% by factor score for the terrestrial biogenic (top) and biomass burning (bottom) factors in TRACE-P. The crosses show the locations of measurements. The back trajectory locations are represented by dots with colors indicating heights (black (< 2 km), blue (2 – 6 km), and red (> 6 km)) and the end point is marked by a diamond.

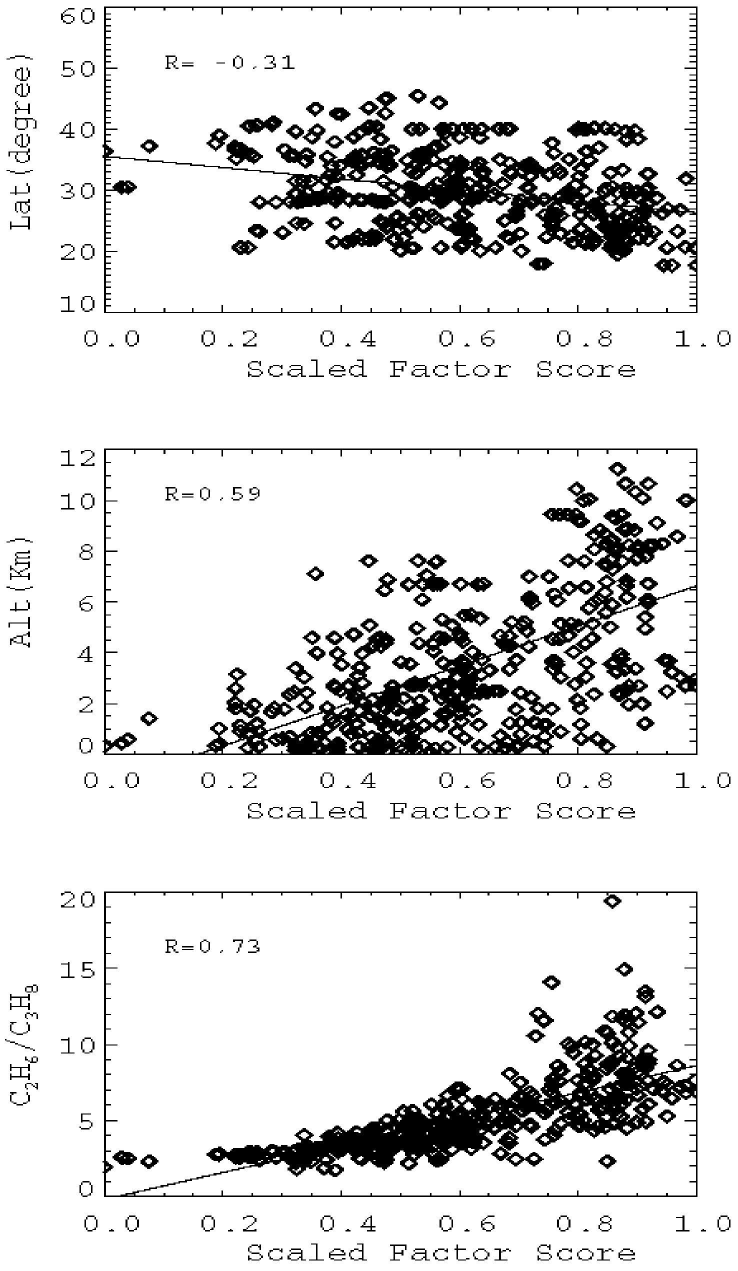


Figure 5. Factor score correlation with latitude (upper), altitude (middle) and C_2H_6 / C_3H_8 ratio (bottom) for the biomass burning factor in TRACE-P.

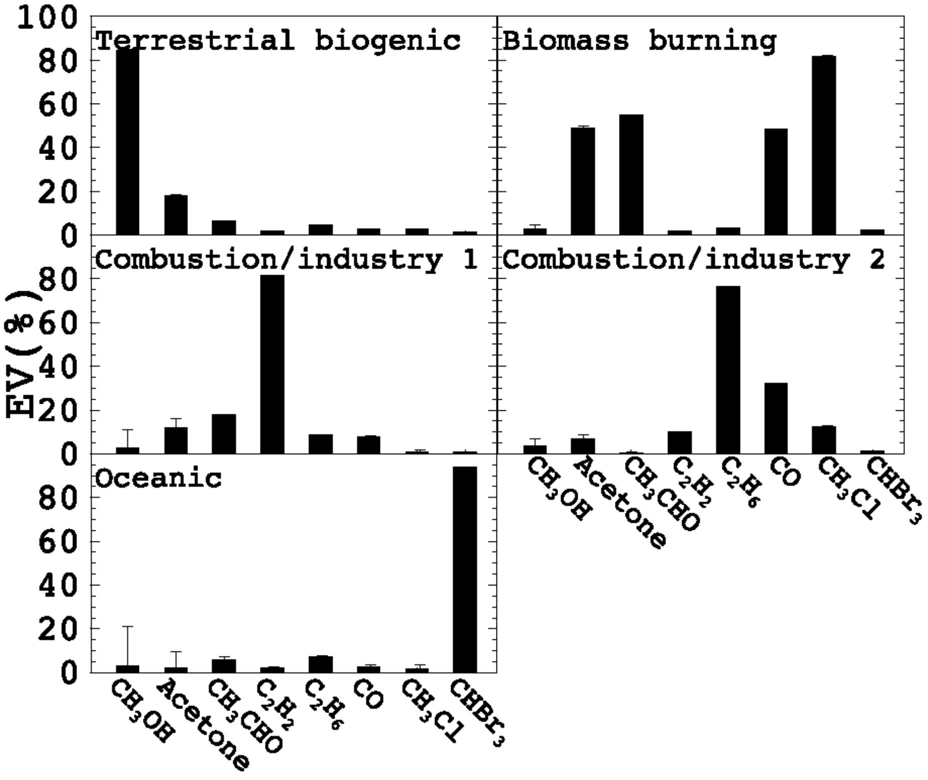
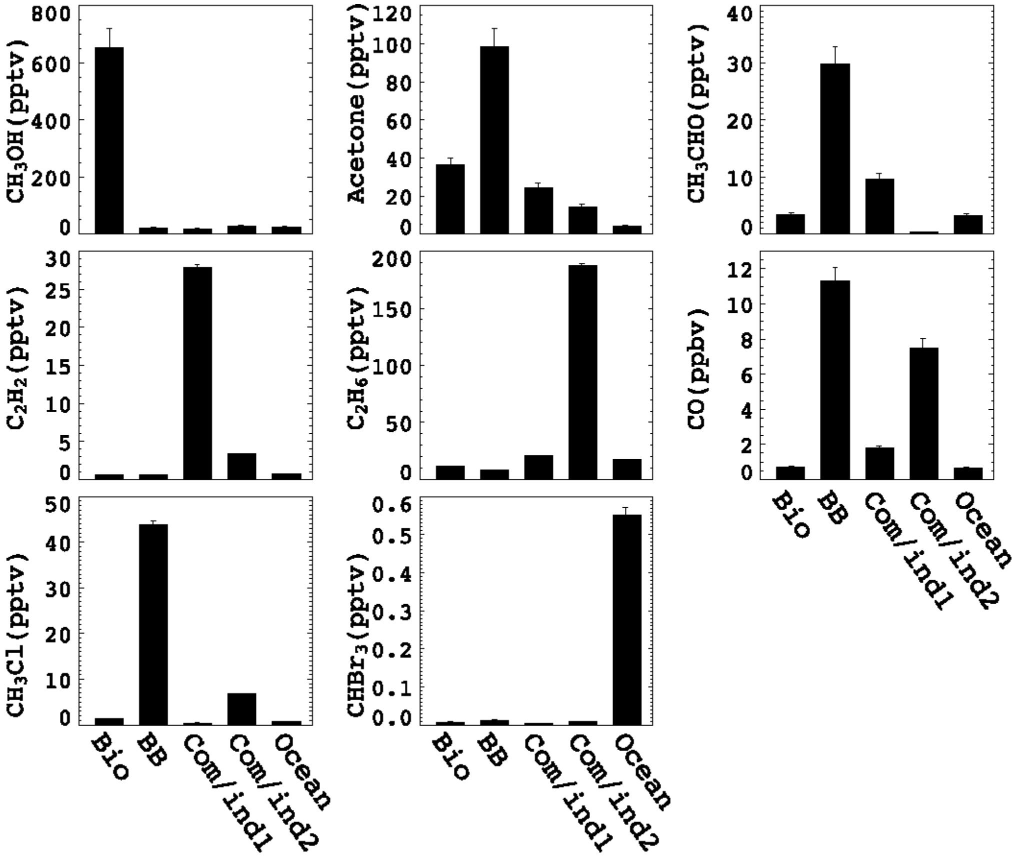


Figure 6. Same as Figure 1 but for PEM-Tropics B.



Figrure 7. Same as Figure 2 but for PEM-Tropics B. “Com/ind1” and “Com/ind2” denote combustion/industry 1 and 2, respectively.

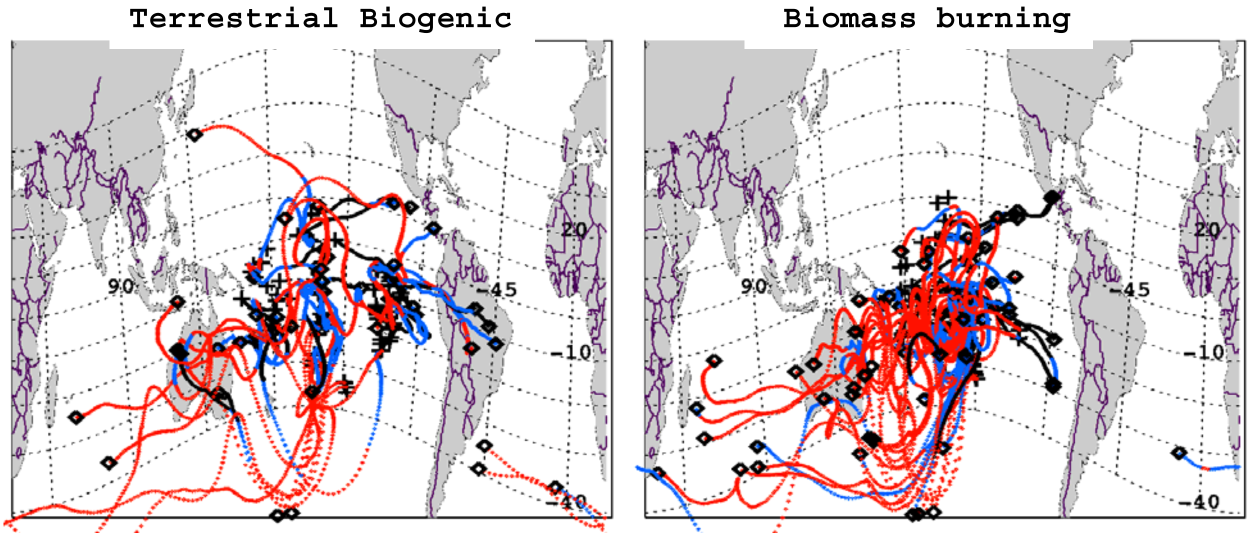


Figure 8. Same as figure 4 but 10 day back-trajectories for PEM-Tropics B.

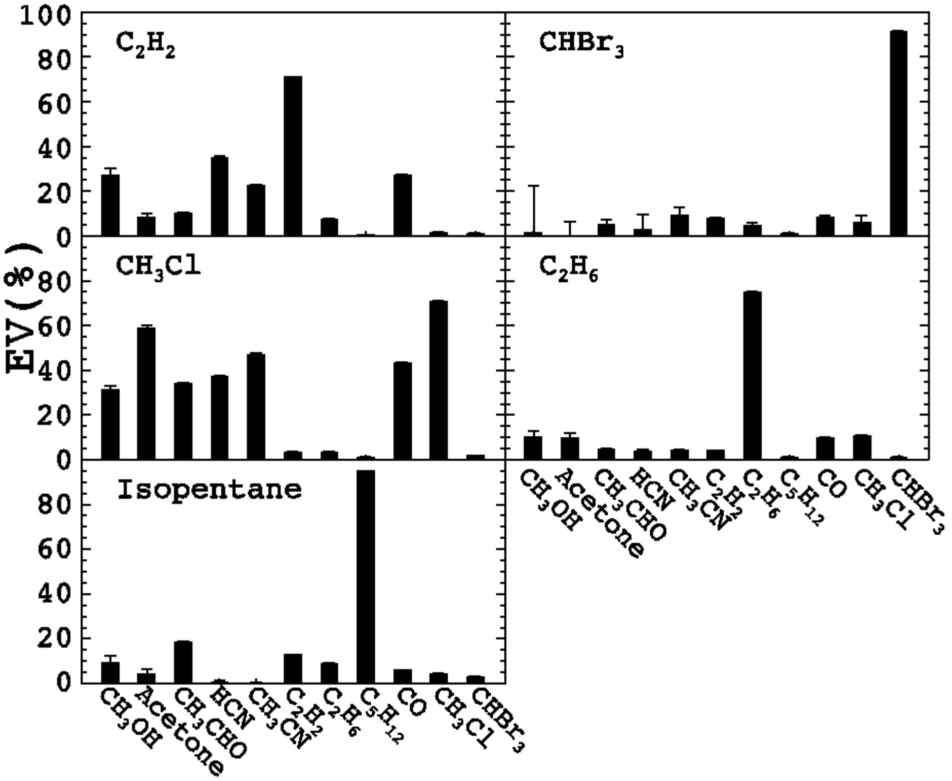


Figure 9. Same as Figure 1 but for INTEX-NA.