

Reconciliation of the VOC and NO_x Emission Inventory with Ambient Data in the Houston, Texas Region

Steven G. Brown, Stephen B. Reid, Paul T. Roberts, Martin P. Buhr, Tami H. Funk
Sonoma Technology, Inc., 1360 Redwood Way, Suite C, Petaluma, CA 94954
steveb@sonomatech.com

Eugene Kim, Phillip K. Hopke
Clarkson University, Potsdam, NY

ABSTRACT

This study, sponsored by the Houston Advanced Research Center (HARC) in support of the Texas Commission on Environmental Quality (TCEQ), was designed to use innovative methods for evaluating and reconciling the emission inventory for the Houston-Galveston area (HGA). The objective of this project was to evaluate the emissions inventory for large point sources in the HGA for 2000 by comparing emissions estimates to ambient data collected during 2000 and 2001, and to make recommendations on possible improvements to the point source inventory. This project focused on sources of non-methane organic compounds (NMOC, often called volatile organic compounds, VOC) that are most likely to contribute to ozone formation, and on oxides of nitrogen (NO_x) in the HGA. In particular, the project focused on the 12 highly reactive VOCs or groups of compounds that TCEQ has targeted because of their significant contributions to ozone formation. These compounds and groups include the following: acetaldehyde; formaldehyde; ethylene; propylene; 1,3-butadiene; all butenes (butylenes); isoprene; all pentenes; toluene; all xylenes; all ethyltoluenes; and all trimethylbenzenes. However, due to limitations in the measurement systems, we did not address acetaldehyde and formaldehyde, and ethyltoluenes and trimethylbenzenes were often lumped together.

The overall results of this study indicate that the NMOC and NO_x point source emission inventory appears to be underestimated and should be evaluated and corrected as necessary. In addition, estimates for non-road sources appear to be high in some areas, particularly for lawn and garden equipment and non-road equipment for Ellington Field. Overall, the speciation profiles used to prepare the emission inventory correlate relatively well with the weight percentages of organic species (and species groups) detected by the ambient monitoring system. However, propylene, n-butane, and ethane were underrepresented and benzene and C-5 paraffins were overrepresented in the emission inventory. The point source emission inventory should be examined to identify the sources that contribute the most mass for the compounds listed above. Both the chemical speciation profiles used to speciate these identified sources, and the speciation profiles used to speciate the point source inventory for refineries and chemical manufacturing should be assessed.

INTRODUCTION

Emission inventories are used to support air quality modeling and develop regulatory control strategies. Because of the complexities involved in developing emission inventories and the implications of errors in the inventory on air quality model performance and control strategy assessment, it is important to evaluate the accuracy and representativeness of any inventory that is intended for use in air quality modeling. There are several techniques used to evaluate emissions data, including bottom-up evaluations that begin with emissions activity data and estimate the emissions accordingly; and top-down evaluations that compare emission estimates to ambient air quality data or use ambient data to estimate emissions profiles. This project focused on top-down evaluation methods using existing surface and aloft (via aircraft) ambient data. The result is an evaluation independent of the inventory itself, as well as an evaluation independent of the modeling that will later use the inventory to predict ozone concentrations and sensitivities.

The technical approach used to evaluate the emission inventory involved a multi-task effort that utilized the extensive collection of surface and aircraft data available for 2000 and 2001 in the HGA. We used three unique but synergistic approaches to perform a top-down emission inventory analysis for emissions estimates in the HSC area. The three analysis tasks included (1) performing source apportionment analyses using positive matrix factorization (PMF) and comparing source contribution profiles with emissions profiles, (2) performing emissions and air quality data reconciliation using surface data, and (3) performing emissions and air quality data reconciliation using aircraft data. The results of each of the three tasks were combined and synthesized to develop recommendations for improving emissions estimates.

Figure 1 depicts the greater Houston-Galveston area including the air quality monitoring sites near the Houston Ship Channel (HSC) where ambient surface and aircraft data were collected and analyzed for this study. The HSC is a heavily industrialized area influenced by large emissions sources, such as petroleum refineries and chemical manufacturing facilities. Figure 2 illustrates the industrial nature of this region and depicts clusters of emissions sources by Standard Industrial Classification (SIC). The VOC and NO_x emission inventory for the HGA is summarized in Table 1.

Figure 1. Map of the greater Houston-Galveston area, including the air quality monitoring sites near the Houston Ship Channel (HSC).

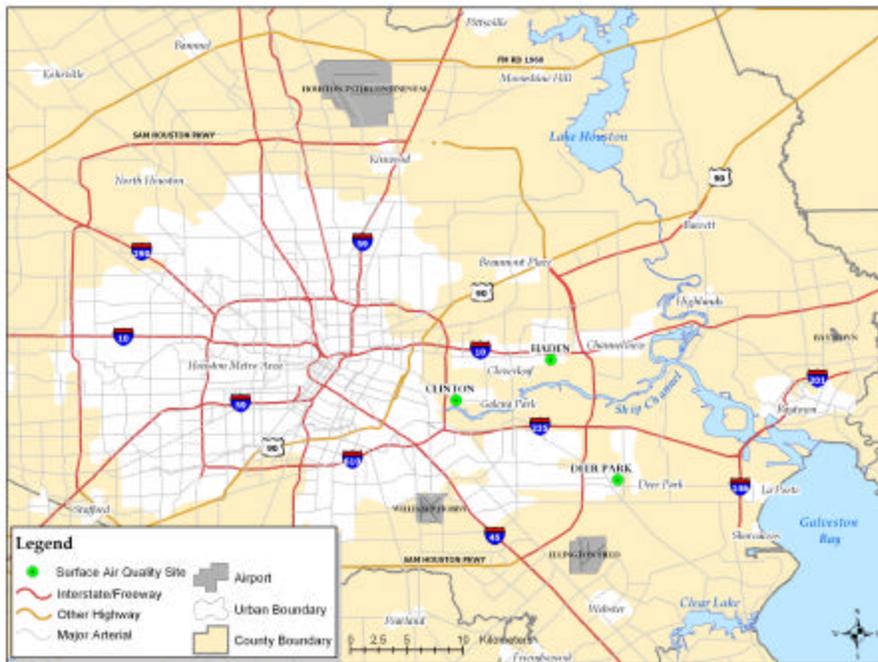


Figure 2. Individual point sources in the Houston Ship Channel area, by industry type. (One facility may consist of many points; several examples are shown.)

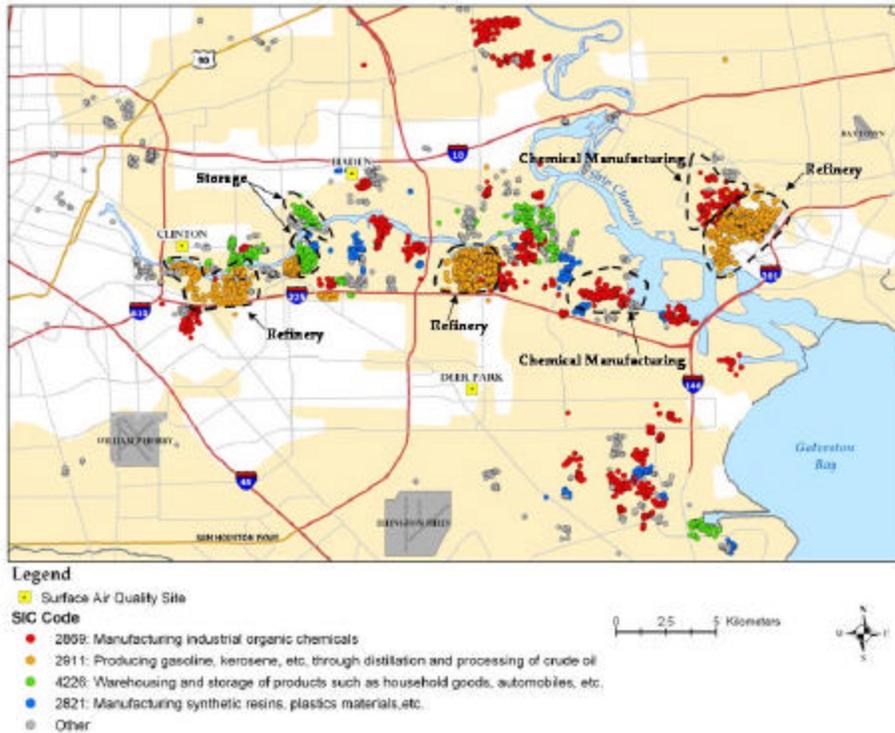


Table 1. Summary of the daily emission inventory (tons/day) for the HGA for August 29, 2000 provided by TCEQ.

Source Category	VOC	NO _x
On-road Mobile	43	77
Non-road Mobile	88	87
Area	45	10
Point	111	195
Total	288	369

TECHNICAL APPROACH

Overview

The data required to implement this three-task approach included (1) an hourly, gridded, and speciated hydrocarbon and NO_x emission inventory, (2) speciated surface hydrocarbon and NO_x data, (3) speciated aloft hydrocarbon and NO_x data, and (4) wind speed and wind direction data. The analysis focused on the abundance of ambient NMOC and NO_x data that have been newly collected in the HSC area during recent monitoring and analysis efforts. This section discusses the technical approach and a summary of findings from the source apportionment analyses, the emissions and air quality data reconciliation using surface data, and the emissions and air quality data reconciliation using aircraft data.

Source Apportionment Analyses Using Positive Matrix Factorization

Hourly speciated VOC data were collected at several sites using automatic gas chromatographs (auto-GCs) in 2001. The auto-GCs record hourly concentrations of nearly 60 hydrocarbons. Other air quality measurements (such as ozone and NO_x) and meteorological data are collocated at these sites. Exploratory source apportionment of hourly VOC data from three sites around the HSC (Clinton Drive, Deer Park, and Haden Rd. as shown in Figure 1) was performed to isolate emission factors and determine if the emission inventory is representative of ambient data. The factor analysis-based receptor model Positive Matrix Factorization (PMF)¹⁻⁴ was used to identify likely emissions sources.

Receptor modeling is a mathematical procedure for identifying and quantifying the sources of ambient air contaminants at a receptor, primarily on the basis of ambient concentration measurements at that receptor (also called source apportionment). Multivariate receptor models require the input of data from multiple samples and extract the source apportionment information from all of the sample data simultaneously. The benefit for the extra complexity of these models is that they estimate not only the source contributions to ambient concentrations but also the source compositions (i.e., chemical profiles or fingerprints). Receptor models such as PMF provide factors that can be related to emission source types and estimate the quantitative contribution of each factor in every sample. Thus, the variation of source strength by time of day, day of week, and wind direction can be explored. These source apportionment results can reveal the types of emission sources, their relative strengths, temporal behavior, and the directional influence.

For application of PMF, data collected at Deer Park, Haden Rd., and Clinton Drive were prepared. The data preparation included data validation, treatment of missing and below detection data, and development of uncertainty files for PMF. To focus on the period when point source emissions are fresh and undiluted (i.e., mixing heights are low) and predominantly industrial (since mobile and biogenic sources are nominal in the middle of the night) and to minimize the influence of photochemistry, data used were limited to nighttime hours. Data used were also limited to the ozone season (summer), because it is the period for which the emission inventory is available, and one of the general goals is to identify sources that contribute to ozone formation in the Houston area. Additionally, the Conditional Probability Function (CPF)⁵⁻⁷ was used to determine the predominant wind direction and source area of each factor.

Key Findings of the Source Apportionment Task

PMF was successfully applied to hourly speciated hydrocarbon data from three sites in the Houston area using nighttime data from summer 2001. Seven factors were isolated at Deer Park and Haden Rd., and eleven factors were identified at Clinton Drive. Mass was well-reconstructed, suggesting that the identified factors accurately represented the ambient data at the three sites. Table 2 summarizes the types of factors identified at each site and the percentage of the total VOC mass they accounted for. The high number of factors found at Clinton Drive is likely due to the better spatial segregation of sources around Clinton Drive and the proximity of the site to major sources. Figures 3 and 4 show examples from Clinton Drive of the median composition of VOC mass from the PMF factors and the relative distribution of mass between industrial and mobile sources in the emission inventory. The figures show that industrial sources appear to be underestimated or that mobile sources are overestimated in the emission inventory.

Table 2. Factor types identified by PMF at each site, and the percentage of VOC mass accounted for.

Factor ID	% VOC at Deer Park	% VOC at Haden Rd.	% VOC at Clinton Drive	Comments
Accumulation	15	8	26	Clinton Drive includes some C-4 to C-5 paraffins, increasing its mass contribution.
Industrial C-4 to C-5 olefins	19	15	12	Deer Park factor also contains C-4 to C-5 paraffins, increasing its mass contribution.
Evaporative emissions	46	29	6	Not distributed similarly at each site.
Isoprene	2		1	Isoprene at Haden Rd. was not separated but grouped with C-5 to C-8 alkanes.
Petrochemical	8	27	11	Haden Rd. includes some C-4 to C-5 paraffins, increasing its mass contribution.
Mobile	1	6	2	Percentages seem low. Evaporative at Haden Rd., mobile mix at Deer Park, and aromatics at Clinton Drive account for other mobile sources.
C-5 to C-8 alkanes		10	9	Includes isoprene at Haden Rd.
Mobile/industrial mix	15	5	17	Composition is different at each site (two identified at Clinton Drive).
Solvent/paint			13	Some likely mobile influence as well.
Diesel			3	Proximity to freeway, shipping and rail enables resolution of this factor at Clinton Drive.

Figure 3. Median composition of TNMOC by PMF factors at Clinton Drive.

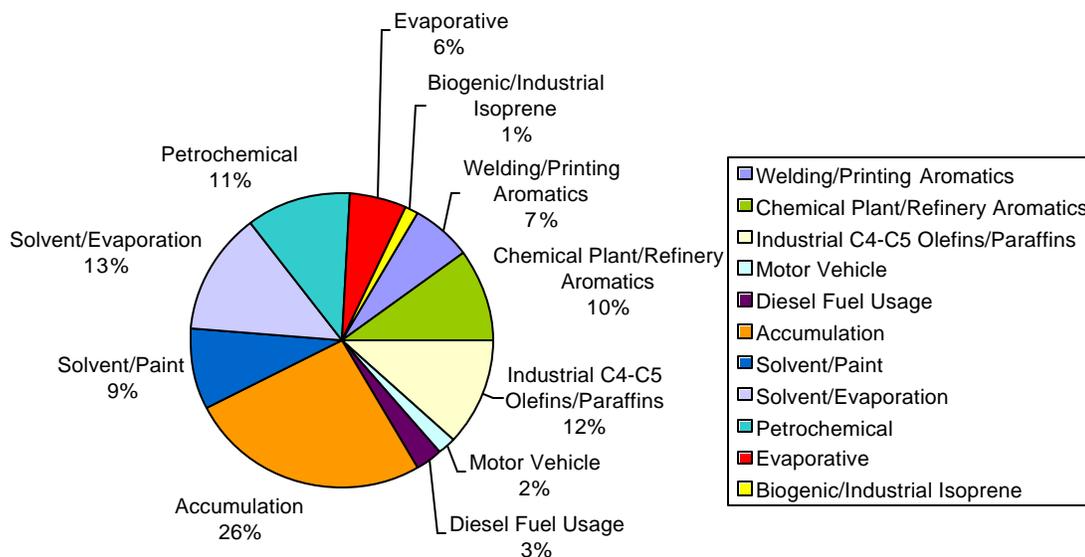
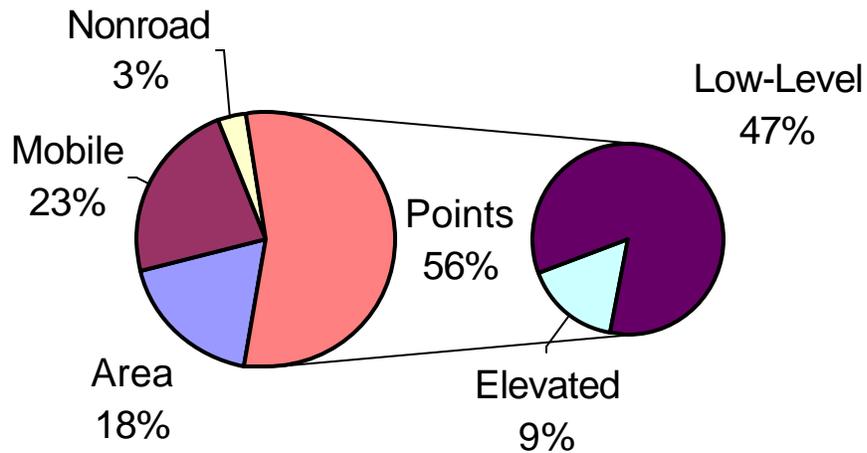


Figure 4. Relative distribution of VOC sources in 11 x 11 grid (2-km x 2-km grid cells) around Clinton Drive during 0500-0700 CST for 2001 including elevated sources.



Overall findings include:

- Industrial sources appear to dominate the VOC composition at all three sites, though not in the same proportion as suggested by the emission inventory. Both mobile and industrial factors were identified, and industrial factors generally made up 90% of the median VOC mass at each site. This is consistent with the emission inventory mass composition around Deer Park, but the emission inventory split around Haden Rd. was 85 industrial/15 mobile, suggesting that mobile sources may be overestimated by a factor of 1.5. The emission inventory split around Clinton Drive was 75/25, suggesting that the industrial portion of the emission inventory is underestimated or that mobile sources are overestimated by a factor of 2.
- Propene appears to be significantly underestimated around Deer Park. The mean ambient composition excluding ethane and propane showed propene at 16% (weight percent) while the emission inventory showed propene at only 5%. Propene was predominantly in the petrochemical factor identified by PMF, suggesting that these specific source types are underestimated. Toluene (6% ambient versus 4% emission inventory weight percent) was also underestimated.
- Propene, toluene, and C3/C4/C5 alkylbenzenes appear to be underestimated by a factor of 2 in the emission inventory around Haden Rd. These compounds were mostly in industrial factors, suggesting that these source types are underestimated in the emission inventory.
- The light olefins (C2-C5) around Clinton Drive appear to be underestimated to the south of Clinton Drive. Analysis suggests that sources of these reactive compounds to the East and South should be of similar magnitude (as they are also a similar distance), but the emission inventory shows significant sources only to the East.
- Mobile sources are spatially distributed and are, therefore, difficult for PMF to completely separate out from point sources in the direction of major freeways. The “mobile” factors at each site accounted for 6% or less of the mass but when combined with fractions of other factors (evaporative, aromatic mixes), the total influence from mobile sources is about 10% of the total mass at the three sites.
- Industrial factors were consistent with the spatial distribution of similar industrial emissions according to the emission inventory. Analyses using wind direction data helped identify the likely source areas by identifying the directions where each factor had the highest concentration.

Sources of C-2 to C-5 olefins, C-2 to C-8 paraffins, and alkyl benzenes were all consistent between PMF results and the emission inventory, although the magnitude was not examined.

- Similar factors were found at each site, indicating that some sources impact the entire area near the HSC. Wind direction analyses at each site showed these factors were all strongest from the same area. The following factors were observed at all three sites:
 - accumulation and natural gas usage characterized by ethane and propane;
 - industrial activity using butenes and pentenes;
 - evaporative emissions containing butanes and pentanes;
 - an isoprene factor predominantly from biogenic sources but also from industrial activity;
 - petrochemical production characterized by ethene and propene; and
 - mobile sources identified by acetylene, benzene, toluene, and xylenes.

Emission Inventory Reconciliation Using Surface Data

Top-down comparative analyses of emissions data with surface air quality data are routinely termed “emissions reconciliation”. An emissions reconciliation is a selective, quantitative comparison of inventory- and ambient-derived molar pollutant ratios (e.g, NMOC/NO_x) and NMOC speciation profiles. Findings often point toward weaknesses or omissions in the emission inventory, which can be iteratively remedied until the inventory data and ambient data reconcile with one another.⁸⁻¹⁰

Emissions reconciliation requires careful selection of data sets to minimize the effects of two critical confounding factors: (1) transport of aged pollutants from distant sources and (2) transformation of fresh pollutants due to photochemical reactions. When significant, these factors invalidate any direct comparison of emissions data with ambient data. However, it is possible to minimize these factors by applying emissions reconciliation techniques only to geographic areas with high emission densities and time periods when chemical reaction rates are low. Thus, one of the most basic principles of emissions reconciliation is to carefully select and analyze emissions and ambient data sets that represent morning periods in urban or industrial areas. Under these conditions, emission rates are high, mixing depths are low, and long-range transport and chemical reactions are minimal. Ambient data collected during the summer from 0500 to 0900 CST (includes samples collected at begin times 0500, 0600, 0700, 0800, and 0900 CST) at Clinton Drive (2000 and 2001), Deer Park (2000), and Haden Rd. (2000) were used for this emission inventory reconciliation.

To assess the sensitivity of NMOC/NO_x ratios to transport, wind analyses were performed to calculate air parcel travel distance during the 0500-0900 CST time period at each site. Because of various factors that may affect differences between ambient- and emission inventory-derived ratios and chemical species compositions, several comparisons were made for different spatial extents. Comparisons between NMOC/NO_x ratios were calculated for the area immediately surrounding the ambient site as well as a broader region defined by directional wind quadrants.

Understanding the uncertainties associated with comparisons of ambient- and emission-inventory-derived pollutant ratios is essential to assess the suitability of top-down evaluation analyses. There are three general categories of uncertainty issues associated with top-down emissions reconciliation analyses: (1) accuracy of the emission inventory, (2) accuracy of the ambient concentration measurements, and (3) suitability of comparisons. Uncertainties associated with the accuracy of the emission inventory generally stem from emissions estimation techniques (emission factors and activity data) and emissions processing techniques (spatial, temporal, and chemical speciation of the inventory). Uncertainties associated with ambient measurements include the influence of instrument detection limits, precision of measurements, and sampling and handling losses. Uncertainties associated with the comparison of ambient and emissions inventory data arise from the

spatial and temporal matching of ambient and emissions inventory data, meteorological factors, and atmospheric reactions. The methodology and data processing steps used for this analysis was designed to minimize the uncertainties associated with the ratio comparisons.

Preparing data for use in top-down emissions reconciliation involves several steps including:

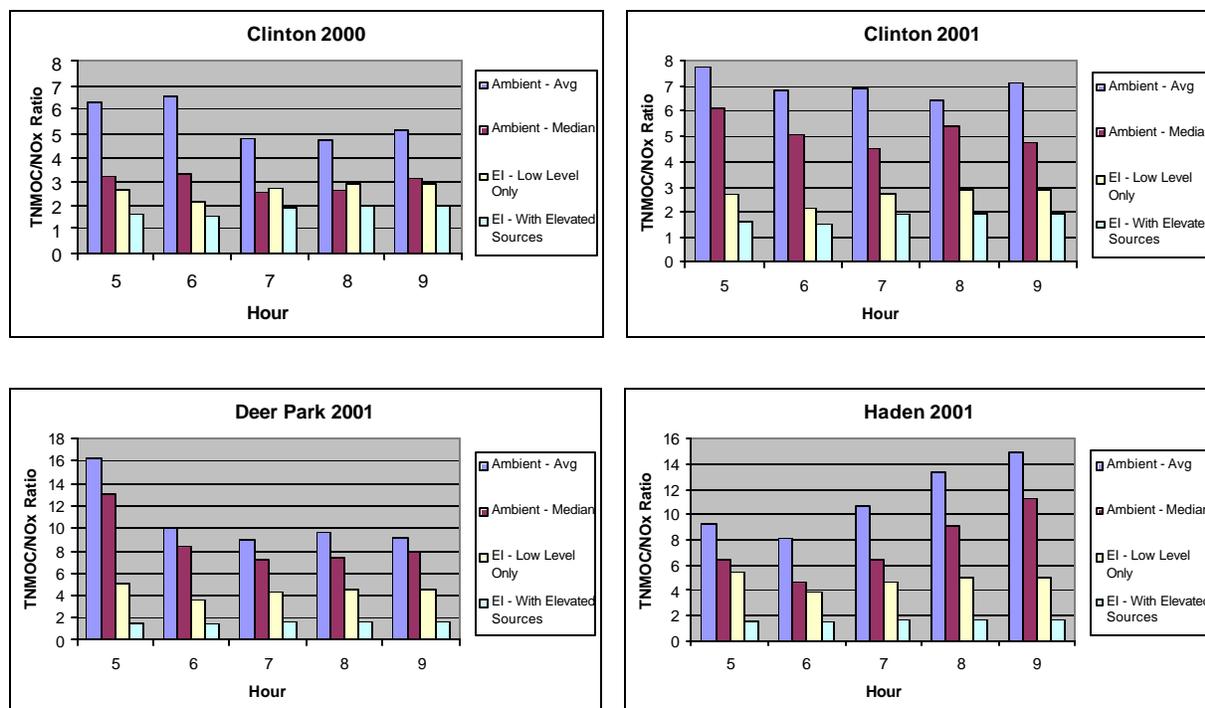
- 1) Ambient data must first be quality-assured to ensure a certain level of confidence in the data. Quality assurance was performed as part of the source apportionment work.
- 2) Ambient data must then be segregated by time period to include data for morning sampling periods and weekdays only.
- 3) Wind analyses must be performed to determine average wind speeds and predominant wind directions for the analysis time period. The meteorology determines how large a grid extent (and how many grid cells) to include for the ratio analyses. The meteorology also indicates the predominant wind directions and, consequently, the sources that are likely to most impact the ambient monitor.
- 4) The chemical species in the emission inventory must be matched to the ambient data. The emission inventory contains many more organic and inorganic chemical species than are measured by the ambient hydrocarbon monitoring systems. To properly match the emissions data to the ambient data, the chemical species measured in the ambient air must be matched to the same species reported in the emission inventory. The emission inventory must then be converted from a mass basis to a molar basis so that the emission inventory data reported in mass can be compared to the ambient data reported in ppbC.

Key Findings of the Surface Emission Inventory Reconciliation

Sufficient time resolved specified ambient measurements and highly resolved speciated emissions were available to draw the following conclusions:

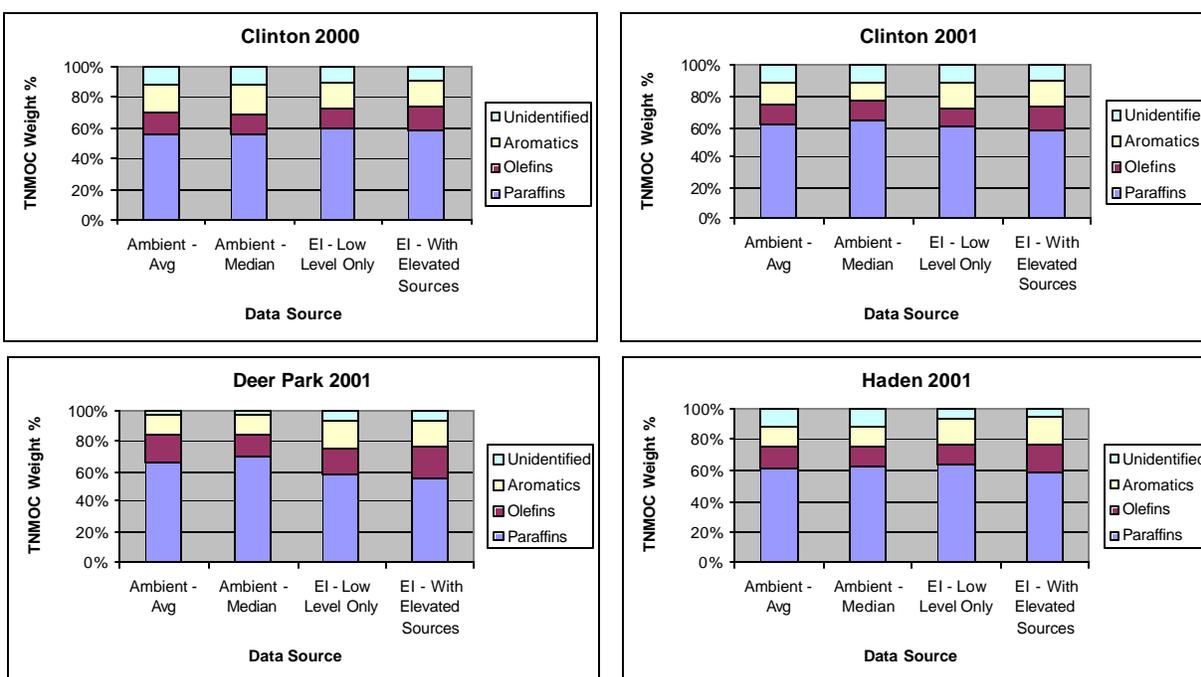
- Overall, the comparisons for the median ambient and emission inventory ratios including elevated sources indicate that the ambient ratios are consistently higher than emission inventory ratios at all three sites. At the Clinton Drive site, ambient ratios are about a factor of 2 to 5 higher than emission inventory ratios; at the Deer Park site, ambient ratios are about a factor of 4 to 12 higher than emission inventory ratios; at the Haden Rd. site, ambient ratios are about a factor of 5 higher than emission inventory ratios. These findings indicate that overall, the emission inventory is substantially underrepresenting the relative amounts of TNMOC and NO_x that are observed in the ambient air during the 0500 to 0900 CST time period. The results of the ratio comparisons by site are presented in Figure 5.
- In general, the TNMOC/NO_x ratios by site show the greatest agreement between the ambient data and emission inventory at the Clinton Drive site. For the entire 0500 to 0900 CST period, the median ambient ratio is about two times higher than the emission inventory ratio (with elevated sources included) for Clinton Drive in 2000 and about three times higher in 2001. However, for Deer Park and Haden Rd., the median ambient ratio is about five times higher than the emission inventory ratio. On an hour-to-hour basis, there tends to be better agreement between the ambient and emission inventory ratios after 0700 CST, when a significant increase in non-road, mobile TNMOC emissions is observed in the emission inventory data. These findings indicate that the TNMOC/NO_x ratios for Deer Park and Haden Rd. show better agreement when the ambient data is influenced more by mobile sources than industrial sources.

Figure 5. Hourly TNMOC/NO_x ratios for Clinton Drive (2000 and 2001), Deer Park, and Haden Rd. sites.



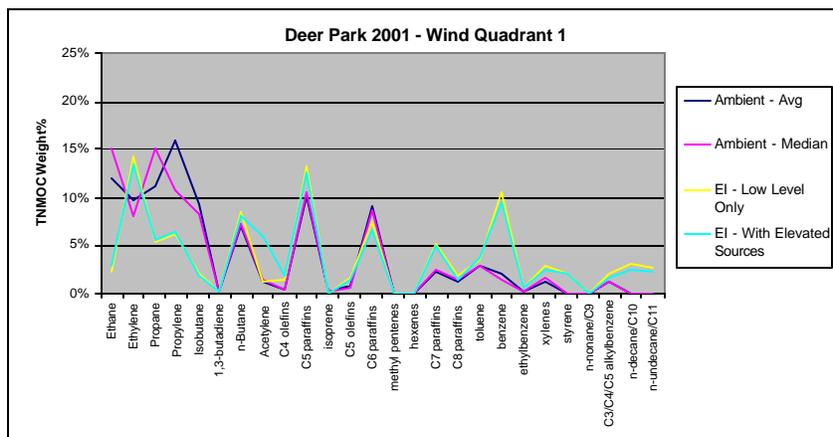
- Analysis of the total relative compositions of the three species groups by site indicate that there is generally good agreement between the ambient data and emission inventory, with paraffins consistently making up about 60% of total TNMOC. Figure 6 shows the relative compositions of the three species groups by site.

Figure 6. Relative composition of major species groups by site.



- For comparisons at all sites, the paraffin/NO_x ratios show the largest discrepancies—the inventory ratios are about 3 to 8 times lower than observed ratios; however, the olefin/NO_x and aromatic/NO_x ratios are also substantially higher (about 2 to 5 times higher) in the ambient median data compared to the emission inventory data including elevated sources. Moreover, the relative amounts of olefins and aromatics in the inventory have a greater impact on the ozone formation potential represented by the inventory. Because ambient levels of these species groups are much higher than reported in the inventory, it is likely that the inventory is underrepresenting the reactive potential of emissions in the HGA, particularly when the spatial distribution of these data is considered.
- At the Deer Park site, when winds are from the northeast, the ambient-derived data show a spike in the fraction of propylene, whereas the emission inventory-derived data show an elevated fraction of benzene. The emission inventory-derived compositions of benzene are also significantly higher than the ambient-derived compositions at the Haden Rd. site when winds are from the east. Examination of pre-gridded point source records indicates that the sources of these emissions at both sites are chemical manufacturing operations at Shell Oil Company and Rohm & Haas facilities northeast of Deer Park. It appears that the chemical speciation profiles used to speciate the point source inventory overrepresent the relative amount of benzene (by about a factor of 2 to 5) and underrepresent the relative amount of propylene emitted by these sources (by about a factor of 2). Figure 7 shows comparison between the 0500 to 0900 CST ambient- and emission inventory-derived relative compositions of individual species for the northeast wind quadrant for Deer Park in 2001.

Figure 7. Comparison between the 0500 to 0900 CST ambient- and emission inventory-derived relative compositions of individual species for the northeast wind quadrant for Deer Park in 2001.



- For individual aromatic species, benzene is consistently overrepresented in the emissions inventory; toluene is underrepresented in the emissions inventory at Clinton Drive and to the southwest of Deer Park and the northwest at Haden Rd., but fairly consistent in other directions at Deer Park; and the C-3, C-4, C-5 alkylbenzene group is underrepresented to the northwest and southeast at Clinton Drive and to the southeast at Deer Park, and significantly underrepresented to the northwest at Haden Rd.
- The emission inventory-derived compositions of C-5 paraffins tend to be higher than the ambient-derived compositions. At the Clinton Drive site, the difference is less pronounced with winds from the west and than when winds were from the east, suggesting that point source speciation profiles may be overestimating paraffins.
- When winds are from the southeast, the ambient-derived data at the Clinton Drive site show a spike in the fraction of n-butane for 2000. This wind quadrant contains large petroleum and chemical manufacturing facilities, and it is likely that the chemical speciation profiles used to

speciate the point source emission inventory are underrepresenting the relative mass of n-butane emitted by these industrial sources.

Emission Inventory Reconciliation Using Aircraft Data

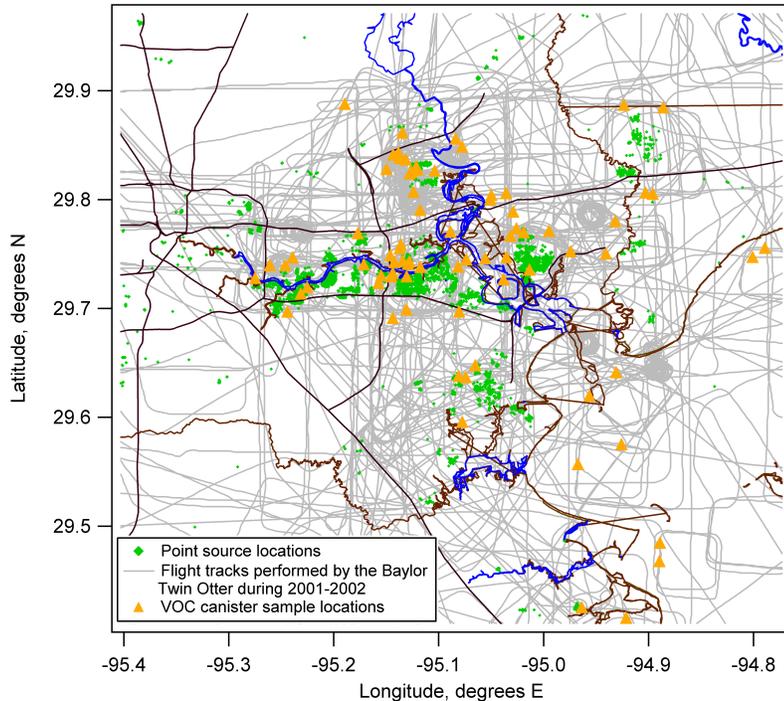
The analysis of ambient aloft pollutant ratios used to evaluate emission ratios is limited by differing chemical and physical loss rates for the species included in the ratio. This limitation may be addressed by using a photochemical clock (e.g., NO_x/NO_y ratio) to extrapolate the measured ratio to the emission ratio. The work described here applies this method to the large set of continuous measurements of NO_x ($\text{NO} + \text{NO}_2$) and NO_y and numerous canister-based VOC analyses that were collected aboard the Baylor Twin Otter aircraft in the HGA in 2001 and 2002. TCEQ is currently evaluating the ratios of a variety of important olefin, aldehyde, and aromatic hydrocarbon species to NO_x from these measurements. In this work we selected the subset of these measurements collected in isolated plumes from petrochemical industry sources and extended the VOC/NO_x ratio analysis to include an analysis of photochemical age, taken as the NO_x/NO_y ratio, for each relevant sample. This analysis helped us gain a better understanding of the change in the observed VOC/NO_x ratio with photochemical processing time. More significantly, the aircraft-collected data allowed us to evaluate the VOC/NO_x ratios from individual point sources

Use of the aircraft data for this analysis is important because of the ability to select plumes that are isolated from other surface sources. The plumes from many of the facilities sampled were intercepted at points both near the facility where the pollutant ratios should be most similar to the actual emission ratios, and at points further downwind where photochemical transformation altered the observed ratio. The canister samples collected during the flights allowed examination of a variety of VOC/NO_y ratios and comparison to the expected emission ratios of VOC/NO_x . The NO_y measurements were used to calculate the observed ratios since NO_y is conserved on the timescales of the transport observed relative to NO_x which is rapidly oxidized.

Comparison between the aircraft-collected VOC samples and the emission inventory data was accomplished by examining the flight track and wind direction associated with each VOC canister sample. Based on this spatial examination, a group of the point sources were selected as having a high probability of influencing the air parcel sampled into the canister. The emission rates of VOC and NO_x for the sources identified were then compiled en suite into the expected VOC_i/NO_x emission ratios. Continuous NO_x and NO_y data collected aboard the aircraft was then integrated over the sampling period of the VOC canister (typically 10-20 seconds) and the observed VOC_i/NO_y emission ratios were calculated.

The ambient data used for this analysis included VOC canister grab samples and continuous NO , NO_2 , NO_y , and meteorological data collected aboard the Baylor University Twin Otter during their 2001 and 2002 measurement campaigns. In total there were 130 ambient VOC samples collected during this period. Figure 8 shows the flight tracks performed during these sampling periods and the location of the VOC canister samples.

Figure 8. Spatial plot of the point sources, flight tracks, and VOC canister sampling locations in the Houston Ship Channel region used in this analysis.



Key Findings of the Aloft Emission Inventory Reconciliation

- Analysis of the data collected by Baylor University during the 2001 and 2002 sampling seasons resulted in isolating and evaluating 9 major point source facilities in the Houston region using 59 VOC canister samples.
- Comparison of the observed $\Sigma \text{VOC}/\text{NO}_y$ ratios to the expected $\Sigma \text{VOC}/\text{NO}_x$ ratios indicated that the emission ratios from the facilities examined except for Oil Tanking Houston may be underestimated by factors ranging from 2 to 43. The observed and expected $\Sigma \text{VOC}/\text{NO}_x$ ratios for Oil Tanking Houston were in good agreement. Table 3 summarizes the results of the point source and aloft sample comparisons for $\Sigma \text{VOC}/\text{NO}_y$. The standard deviations reported for the emission inventory values reflect differences in the individual point sources selected as having influence on a given canister sample. The standard deviations reported for the observed ratio reflect the variation in concentrations observed among a set of canister samples associated with a given source.
- Comparison of the observed $\Sigma \text{Alkenes}/\text{NO}_y$ ratios to the expected $\Sigma \text{Alkenes}/\text{NO}_x$ ratios indicated that the emission ratios from all of the facilities examined may be underestimated by factors ranging from 3 to 64.
- Comparison of the observed $\Sigma \text{Aromatics}/\text{NO}_y$ ratios to the expected $\Sigma \text{Aromatics}/\text{NO}_x$ ratios indicated that the emission ratios from the facilities examined except for Oil Tanking Houston and Shell Oil Company may be underestimated by factors ranging from 3 to 20. The observed and expected $\Sigma \text{Aromatics}/\text{NO}_y$ for Oil Tanking Houston and Shell Oil Company were in good agreement.
- The method used to evaluate the observed VOC/NO_y ratios as a function of the NO_x/NO_y “photochemical clock” did not provide compelling reason to use the extrapolated value as an indication of the emission ratio at the source. However, examination of the data presented in this way for several of the sources does suggest that if samples are collected too close to a facility the resulting emission ratio may be biased high because inadequate mixing of the VOC and NO_x sources has taken place.

- Uncertainty with respect to the upwind concentration of VOCs was a limiting factor in this analysis. At one extreme, this uncertainty sets the observed VOC/NO_y ratios as an upper limit for the target facility.

Table 3. Expected and observed total VOC-to-NO_x ratios for the facilities evaluated.

Principal source	Emission Inventory SVOC/NO _x	Observed SVOC/NO _y	Ratio of observed to expected emissions
Chevron Phillips Chemical Co LP - 0145	Average= 0.588 Std. Deviation=0.004	Average= 14.7 Std. Deviation= 9.0	25.0
The Dow Chemical Co - 0041	Average= 0.131 Std. Deviation= 0.013	Average= 3.8 Std. Deviation= 3.8	29.0
Sterling Chemicals Inc - 0010	Average= 0.865 Std. Deviation=0.000	Average= 7.1 Std. Deviation= 4.2	8.2
Shell Oil Co - 0039	Average=1.123 Std. Deviation= 0.191	Average= 7.2 Std. Deviation= 6.1	6.4
Oil Tanking Houston Inc - 0277	Average= 19.154 Std. Deviation=0.020	Average= 9.8 Std. Deviation=10.7	0.5
Hoechst Celanese Chemical Group, Inc. - 0003	Average= 0.198 Std. Deviation= 0.015	Average= 8.5 Std. Deviation= 7.9	42.9
EXXON MOBIL Chemical Co – 0014	Average=0.896 Std. Deviation=0.340	Average= 8.8 Std. Deviation= 5.1	9.8
Equistar Chemicals LP - 0075	Average=0.782 Std. Deviation= 0.162	Average= 5.2 Std. Deviation= 3.5	6.6
BP Solvay Polyethylene N. America - 0004	Average= 2.548 Std. Deviation=0.937	Average= 6.1 Std. Deviation= 2.7	2.4

INTEGRATED RESULTS

Integrated Results from All Three Analysis Tasks

The three-pronged analysis approach provides added value in that the findings from each individual task discussed above can be integrated to develop corroborative results. The following findings are supported by results from at least two or more detailed investigations and typically by results from more than one of the three analysis approaches that were used.

- Total VOC as identified using VOC/NO_x ratios for multiple sources, was typically underrepresented in the emissions inventory by a factor of 2 to 10. For single sources, total VOC, as identified using VOC/NO_x ratios, was typically underrepresented in the emissions inventory by a factor of 2 to 43.
- Most of the underrepresentation seems to be in the industrial facilities segment of the inventory.
- Olefins, and specifically propylene, are underrepresented in the emissions inventory by a factor of 2 to 5. For single sources, total alkenes, as identified using VOC/NO_x ratios, were typically underrepresented in the emissions inventory by a factor of 3 to 64.
- Total aromatics were also underrepresented in the emissions inventory by a factor of 2 to 5. For single sources, total aromatics were typically underrepresented in the emissions inventory by a factor of 1.4 to 20. For individual aromatic species, benzene is consistently overrepresented in

the emissions inventory; toluene is underrepresented in the emissions inventory around Clinton Drive and to the southwest of Deer Park and the northwest of Haden Rd., but fairly consistent in other directions around Deer Park; and the C-3, C-4, C-5 alkylbenzene group is underrepresented to the northwest and southeast of Clinton Drive and to the southeast of Deer Park, and significantly underrepresented to the northwest of Haden Rd.

- Total paraffins were also underrepresented in the emissions inventory by factor of 3 to 8. For individual paraffin species, ethane and propane are consistently underrepresented in the emission inventory, especially at the Deer Park site, while the C-5 paraffins tended to be overrepresented in the emission inventory. For single sources, total alkanes were typically underrepresented in the emissions inventory by a factor of 2 to 41. However, because of the low reactivity of these species, these discrepancies in the emissions inventories are less important than ones with more-reactive species.

RECOMMENDATIONS

This work has identified specific improvements that could be made to the emissions inventory:

- Non-road emissions appear to be high in general, particularly TNMOC emissions occurring from 0700 to 0900 CST around the Clinton Drive site, where non-road TNMOC emissions are approximately the same as point source TNMOC emissions. Lawn and garden equipment activity is the main contributor to non-road TNMOC emissions for Harris County as a whole. Emissions of NO_x from activities at Ellington Field (south of the Deer Park site) appear to be overestimated and should be re-evaluated.
- Since TNMOC/NO_x ratios show the greatest discrepancy between the ambient- and emission inventory-derived values at sites (and wind quadrants) dominated by industrial sources, it appears that point source TNMOC emissions are underestimated, NO_x emissions are overestimated, or both. However, given the prevalence of refineries and chemical manufacturing facilities in the region and the difficulties inherent in estimating NMOC emissions, it seems more likely that NMOC emissions are underestimated. The point source NMOC inventory should be evaluated and corrected as necessary.
- Overall, the speciation profiles used to prepare the emission inventory track relatively well with the weight percentages of organic species (and species groups) detected by the ambient monitoring system. However, propylene, n-butane, and ethane were underrepresented and benzene and C-5 paraffins were overrepresented in the emission inventory.
- The point source emission inventory should be assessed to determine which sources contribute the most mass for the compounds listed above. The chemical speciation profiles used to speciate these sources should be assessed for the compounds listed above. Also, speciation profiles used to speciate emissions from refineries and chemical manufacturing should be assessed.

Additional recommendations for future emissions inventory reconciliation include the following:

- This same emission inventory reconciliation analysis should be performed for more sites within the Houston/Galveston area.
- Run PMF on data limited to the morning hours used in the emission inventory analysis (0500-0900 CST) to better examine the temporal variations of the ambient data, and to investigate how accurate the temporal profiles of specific emission inventory types (such as off-road mobile) are.
- The recommended changes to the emission inventory should be implemented and the emission inventory reconciliation should be performed again based on the updated inventory.
- Similar ambient-to-emission inventory comparisons using aloft data should be performed with data collected in 2003. In addition, future data collection campaigns using aircraft should

include an upwind canister for each set of downwind samples collected. The resulting data would significantly reduce the uncertainty that exists in the current data with respect to the upwind concentration of the VOC species.

REFERENCES

1. Hopke, P. K. "A guide to positive matrix factorization"; Prepared for Positive Matrix Factorization program, Potsdam, NY, by the Department of Chemistry, Clarkson University, Potsdam, NY. 2003.
2. Paatero, P.; Tapper, U. "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values", *Environmetrics* 1994, 5, 111-126.
3. Paatero, P. "User's guide for positive matrix factorization programs PMF2 and PMF3, part 1: tutorial"; Prepared by University of Helsinki, Finland. 2000.
4. Poirot, R. L.; Wishinski, P. R.; Hopke, P. K.; Polissar, A. V. "Comparative application of multiple receptor methods to identify aerosol sources in northern Vermont", *Environ. Sci. Technol.* 2001, 35(23), 4622-4636.
5. Ashbaugh, L. L.; Malm, W. C.; Sader, W. Z. "A residence time probability analysis of sulfur concentrations at Grand Canyon National Park", *Atmos. Environ.* 1985, 19(8), 1263-1270.
6. Kim, E.; Hopke, P. K.; Larson, T. V.; Covert, D. S. "Analysis of ambient particle size distributions using UNMIX and positive matrix factorization", *Environ. Sci. Technol.* 2003, 38, 202-209.
7. Kim, E.; Hopke, P. K.; Edgerton, E. S. "Source identification of Atlanta aerosol by positive matrix factorization", *J. Air & Waste Manage. Assoc.* 2003, 53, 731-739.
8. Haste, T. L.; Chinkin, L. R.; Kumar, N.; Lurmann, F. W.; Hurwitt, S. B. "Use of ambient data collected during IMS95 to evaluate a regional emission inventory for the San Joaquin Valley"; STI-997211-1800-FR, Final report prepared for the San Joaquin Valleywide Air Pollution Study Agency, c/o the California Air Resources Board, Sacramento, CA, by Sonoma Technology, Inc., Petaluma, CA. 1998.
9. Korc, M. E.; Jones, C. M.; Chinkin, L. R.; Main, H. H.; Roberts, P. T.; Blanchard, C. "Use of PAMS data to evaluate the Texas coast emission inventory"; Work assignment 2-95, EPA Contract No. 68D30020, STI-94520-1558-FR, Final report prepared for U.S. Environmental Protection Agency, Research Triangle Park, NC, by Sonoma Technology, Inc., Santa Rosa, CA. 1995.
10. Fujita, E. M.; Croes, B. E.; Bennett, C. L.; Lawson, D. R.; Lurmann, F. W.; Main, H. H. "Comparison of emission inventory and ambient concentration ratios of CO, NMOG, and NO_x in California's South Coast Air Basin", *J. Air & Waste Manag. Assoc.* 1992, 42, 264-276.

ACKNOWLEDGMENTS

This research was sponsored by the Houston Advanced Research Center (HARC). We thank the current HARC contract manager, Dr. Jay Olaguer, and Ms. Anne Brun, who was the contract manager for the early part of the project. We acknowledge the hard work and assistance given by numerous persons responsible for both surface and aloft air quality data collection, validation, and processing; and for emissions inventory compilation. We would also like to thank the Texas Commission on Environmental Quality (TCEQ) for providing the emission inventory data and support.

KEY WORDS

Emission inventory
Reconciliation
Source Apportionment
PMF
Aircraft
TNMOC
VOC
NO_x
Houston
Texas