

# Comparison of Atlanta Emission Inventory with Ambient Data Using Chemical Mass Balance Receptor Modeling

EPA/600/A-94/244

Teri L. Conner<sup>1</sup>, John F. Collins<sup>2</sup>, William A. Lonneman<sup>1</sup>, and Robert L. Seila<sup>1</sup>

<sup>1</sup>Atmospheric Research and Exposure Assessment Laboratory

U.S. Environmental Protection Agency

Research Triangle Park, North Carolina, 27711

<sup>2</sup>Civil Engineering Department, Environmental Engineering Program

University of Southern California

3620 South Vermont Avenue

Los Angeles, California, 90089-2531

## ABSTRACT

A comprehensive hydrocarbon database was obtained at 6 sites in the Atlanta metropolitan area during the summer of 1990. Samples were collected in stainless steel canisters and analyzed for 54 hydrocarbon species plus total non-methane organic compounds (TNMOC). The contributions of the major sources of TNMOC at each of the 6 sites were estimated through a procedure called Chemical Mass Balance (CMB) receptor modeling. Spatial variability of the source contributions is discussed. Results of the CMB analysis for one of the sites are compared with the emission inventory for Atlanta using several different approaches. The inventory highway mobile source estimate tends to be smaller than the minimum ambient data-derived highway mobile source estimate, and the inventory area plus point source estimate tends to be larger than the maximum ambient data-derived estimate for the data set examined. However, these source estimates are interdependent to some extent. Limitations of these comparisons are discussed.

## INTRODUCTION

During the summer of 1990, the U.S. Environmental Protection Agency conducted a 2-month, 6-site air quality monitoring study in the Atlanta metropolitan area referred to as the "1990 Atlanta Ozone Precursor Study". The goal of the study was to obtain hourly data for ozone and its precursors at multiple sites in an area which is out of compliance for the ozone NAAQS. Earlier reports on this work focussed on hydrocarbon data obtained from hourly on-site measurements made at one of the sites. In this report, we examine data obtained at all 6 sites, but on a less frequent basis, from samples collected in stainless steel canisters. The data set obtained provides an opportunity for reconciling ambient data with emission inventories. Previous efforts to compare ambient data with emission inventories have made use of a receptor modeling technique called Chemical Mass Balance (CMB) and other techniques employing the ambient data to determine the sources of the measured species.<sup>1-5</sup> Source contributions obtained in this way - from ambient data with no explicit use of meteorology or emission inventories - are compared in a relative sense with emissions inventories derived by traditional methods. In this effort, CMB is applied to the hydrocarbon data obtained at 6 sampling sites. Spatial variability of the results is examined, and new approaches for comparing CMB results with emission inventories are demonstrated for one of the sampling sites. While this analysis emphasizes results for motor vehicle-related sources, other sources are considered as needed for the CMB calculations.

## AMBIENT DATA

### Sample Collection and Analysis

Whole-air samples were collected in 6-liter Summa<sup>®</sup> polished stainless steel canisters. Sampling was conducted approximately every other day at each of 6 sampling sites throughout the Atlanta metropolitan area. Sites 2 and 3 were located near downtown Atlanta. Sites 4, 5, and 6

were located in a variety of surroundings away from the city. Site 4 was located about 5 km south of downtown Atlanta. Sites 5 and 6 were at typical downwind locations from Atlanta approximately 15 and 10 km, respectively. Site 1, located northwest of Atlanta, served as a background site. Start times for the 30-minute samples were rotated through the hours 0 (midnight), 8, 10, 12, 15, 18. A total of 163 samples were collected in this manner. Samples were analyzed by a well-characterized gas chromatograph with flame ionization detection (GC/FID) system operated by the Gas Kinetics and Photochemistry Research Branch of AREAL.<sup>6</sup> This is the same system used to analyze the mobile source profile samples prepared for this study.<sup>7</sup> Detailed descriptions of the sampling and analytical procedures are presented elsewhere.<sup>8</sup>

### **Data Screening**

The primary data screening tool used was to examine plots of each species versus every other species. This screening method has been demonstrated to be quite effective in identifying outliers and other problems with the data.<sup>5,9</sup> Particular attention was paid to those pairs of species which are expected to track each other very well. Notable outliers were observed for the butane and pentane isomers, resulting in elimination of the affected samples from further consideration. The problem was most severe at site 5, so this site was excluded from further analysis. The reason for the unusual butane and pentane compound concentrations is not known at this time. However, the objective of this exercise is to look at the usual situation, so screening out unusual data is appropriate in this case. Any offsets observed in comparing one species with another resulted in exclusion of the affected species from consideration as a fitting species in the CMB analysis. Species affected by offsets were 2-methylpentane and 2-methylhexane. One hundred of a possible 135 samples were retained (site 5 excluded) after data screening.

### **Data Averaging**

Chemical Mass Balance (CMB) calculations were performed on the average of samples from each site. Averages provide a snapshot of typical conditions at a sampling site and cancel out some of the random variations associated with individual samples. Only those samples which survived the screening process were included in the averages. Species with missing values were assigned a value of zero rather than excluding those samples from the average. This approach will tend to give a lower estimate of the species average concentration. In most, if not all cases, the missing values were assigned when concentrations were below detection; thus, assigning a value of zero to these species concentrations is a reasonable approach. It is unlikely such species will figure prominently into a CMB calculation, due to their high relative uncertainty.

## **CHEMICAL MASS BALANCE CALCULATION PROCEDURE**

### **Introduction**

The U.S. EPA/DRI Chemical Mass Balance model, version 7.0<sup>10,11</sup> was used to quantitatively apportion chemical species measured at the sampling sites to the major sources contributing to the total non-methane organic compounds, TNMOC (defined in this paper as the sum of all GC peaks excluding ethane) at those sites. The Chemical Mass Balance model, version 7 (CMB7) consists of an effective variance least squares solution to a set of linear equations which express each measured chemical species concentration as a linear sum of the contributions of each source to the chemical species. The effective variance solution gives the most weight to source and ambient measurements with the lowest relative uncertainty estimates. Source contributions are expressed as the product of the abundance of the species as emitted by the source and the total mass concentration contributed by the source. The set of abundances of all species as emitted by each source represents the "source profile" or "source fingerprint".

In practice, it is not possible to apportion mass to each individual contributing source. Individual sources may be too similar to one another, too numerous, or may not contribute significantly to the total mass loading. Sources are generally grouped together to represent a single

"source category" or "source type". Furthermore, these source types are distinguished chemically (e.g., gasoline headspace vapor) rather than by the mechanism of emission (e.g., fuel storage tank vapor displacement versus vehicle refueling vapor displacement.) This distinction will be important for comparison of CMB results with emission inventories.

Assumptions made in performing a chemical mass balance include: 1) the abundance of each species used in the fitting procedure is known for each source type, 2) all major sources of each species used in the fitting procedure must be included in the CMB, and 3) chemical species do not react with each other. The third assumption listed here is of particular concern for NMOC's, as virtually all of these species are reactive towards the OH radical to some degree. Thus, choice of fitting species must be restricted to the least reactive compounds. Other assumptions made for the CMB model are listed and explained elsewhere.<sup>10-12</sup>

The CMB7 model requires both ambient concentrations and uncertainties as input. Ambient measurement uncertainties were calculated from the following expression<sup>5</sup> applied to the average species concentrations at each site:

$$\Delta C_i(\text{ppbC}) = [(0.2)^2 + (0.05 * C_i)^2]^{1/2}$$

where  $C_i$  is the measured ambient concentration of species  $i$ . This method of estimating the uncertainty is used rather than calculating the standard deviation of the average, because the latter might give low-concentration species more weight in the CMB than the well-measured higher concentration species.

### Source Profiles

Local source profiles were developed for this study to represent roadway, whole gasoline, and gasoline headspace vapor.<sup>7</sup> The roadway profile (ROAD) was obtained from samples collected in an extended underpass in downtown Atlanta and captures the emissions of vehicles in motion (tailpipe plus running losses). This profile was not corrected for background concentrations and is dominated by emissions of vehicles in motion.<sup>7</sup> The whole gasoline (GAS) and gasoline headspace vapor (HS) profiles are each a composite of profiles of 3 octane grades from the 6 major vendors in the Atlanta area. The gasoline headspace vapor profiles used in this analysis were determined for 24°C. These mobile source related profiles were all analyzed on the same GC/FID system used to analyze the ambient samples.

In a previous analysis of Atlanta continuous monitoring GC data,<sup>5</sup> natural gas plus an additional propane-rich source were found to account for most of the propane and about 10% of the TNMOC (defined in that study as the sum of all GC peaks with ethane included). In that study, a propane-rich profile derived from the ambient data itself<sup>9</sup> was used to account for the excess propane. The abundance of propane in that profile was 45%, but it is not clear what made up the remaining 55%. In this study, a pure propane profile (PROPANE) was used to account for the propane not explained by the natural gas profile. Using a single-constituent profile in this way will underestimate the propane source if there are other species emitted by that source. However, this approach avoids the uncertainty associated with using a profile derived from a data set different from that to which the profile will be applied. The natural gas profile (NG) used in this exercise was obtained from existing measurements on Atlanta utility natural gas.<sup>3</sup>

For some sites and samples, toluene could not be fully explained by the mobile source profiles, suggesting the possibility of some solvent-type of activity. The existence of solvent-type activities is confirmed by the emission inventory.<sup>14</sup> Scheff et al. (1989)<sup>15</sup> compiled profiles for CMB apportionment of volatile organic compounds. Included in that compilation are profiles for surface coatings and other solvent-related activities. The autopaainting profile (AUTOCOAT) was found to best explain the data. The profiles are given as weight % while the other profiles used here are in ppbC %. The two profile representations are nearly the same, so no conversion was done for the AUTOCOAT profile.

All profiles are normalized to the TNMOC defined in this paper as excluding ethane. Ethane is excluded because it is excluded from the SIP guidance for emission inventories for ozone non-compliance.<sup>16</sup> The only profiles affected by this exclusion are roadway (ROAD) and natural gas (NG). The emission inventory also excludes many chlorinated compounds, but these are not expected to figure prominently into the sources apportioned here. The emission inventory includes formaldehyde and acetaldehyde which are, at best, only partially measured by the FID. These compounds could potentially figure prominently into the roadway emissions, but data from a previous roadside sampling study<sup>17</sup> suggests that these aldehydes are only a small part (< 2%) of the TNMOC.

### **Fitting Species**

One of the requirements of the CMB is that chemical compounds do not react with other species. In this exercise, fitting species will be those with a reactivity towards OH radical less than that of ethene. In a polluted atmosphere ( $\cdot\text{OH} = 5 \times 10^6$  molecules/cm<sup>3</sup>), this would correspond to a lifetime greater than 6.5 hours. In the canister data set, all alkenes, plus n-alkanes greater than or equal to C-8, xylenes, trimethylbenzenes, and methylheptanes are all excluded from being fitting species on the basis of reactivity. In addition, 2-methylpentane and 2-methylhexane were excluded from being fitting species because of offsets observed in the scatter plots. Methylcyclopentane, 3-methylhexane, methylcyclohexane, and n-heptane were eliminated from the fitting species because they were frequently missing or underpredicted. The list of measured species and their fitting species status is shown in Table 1.

## **CHEMICAL MASS BALANCE RESULTS AND DISCUSSION**

### **Results of Average - Fitting Diagnostics and Collinearities**

Table 2 shows the results of CMB7 calculations applied to the site averages as a percent of the total mass apportioned (sum of source estimates). Each calculation had at least 12 degrees of freedom, which should be ample to give statistically meaningful results. The goodness of fit indicators  $R^2$  and  $\text{Chi}^2$  were excellent. However, each fit was accompanied by one or more uncertainty/similarity clusters. These clusters are indicated in the CMB7 output display and are formed when the standard error of any of the source estimates involved is 50% or higher and there is excessive similarity among the source profiles, as indicated by an established diagnostic.<sup>12</sup> The standard error of the combined source contributions indicated in the cluster display may be smaller than the standard error of the source contribution estimate of any single source in the cluster if collinearity is the cause of the high standard error. Source combinations typically affected by collinearity were whole gas (GAS) and roadway (ROAD), and propane source (PROPANE) and natural gas (NG).

There are several ways of dealing with uncertainty/collinearity clusters in the CMB.<sup>12</sup> One approach is to use additional species in the fit. Unfortunately, species most likely to further distinguish ROAD and GAS from one another are fairly reactive (e.g., ethene, propene). However, ignoring the reactivity rules established for this exercise and including some of the species likely to be associated with exhaust did not eliminate the collinearity. Another approach would be to prepare a composite profile for the collinear sources. However, this would require having some independent estimate of the relative importance of the profiles involved.

A test can also be performed which may indicate if the resulting cluster is due to high profile uncertainty rather than true collinearity. The test consists of reducing uncertainties in the affected profile(s) and then rerunning the CMB. If the clusters are no longer listed (i.e., the standard error is reduced to less than 50%), then the apparent collinearity was likely due to profile uncertainties. This test was applied to the GAS/ROAD cluster for the average site 2 sample. Uncertainties in the GAS and ROAD profiles were arbitrarily reduced by half. When the source estimates were recomputed, the uncertainty/similarity clusters and high standard errors remained. Even when the GAS profile standard error was reduced to an unreasonably low level (0.00001, in terms of ratio of

species to total), the clusters and high uncertainties remained. It is therefore likely that a true collinearity as defined by CMB7 exists between the GAS and ROAD profiles and the two cannot be resolved reliably based on the collinearity criteria of the CMB7 model. Thus, the separate reporting of ROAD and GAS should be regarded with caution. The combined source contribution of ROAD and GAS as reported in the CMB7 cluster display are reported in Table 2, as well as their separate contributions. It is important to point out that the collinearity criteria of CMB7 are somewhat arbitrary and were originally established for typical particulate data. Furthermore, no ROAD plus GAS cluster was observed for site 3. The uncertainty estimated for GAS was below the 50% threshold for that site, but otherwise uncertainties for ROAD and GAS were similar to estimates at other sites.

The PROPANE and NG sources are not likely to be resolved with any standard remedies because of the small number of species associated with these sources. In this case, their source contributions are reported as the sum presented in the cluster display.

The surface coatings source (AUTOCOAT) was not collinear with other sources, but generally was small with a very large standard error, indicating it is probably not an important source at most sites. The standard error was less than 50% at sites 2 and 4, but greater than 50% at sites 1, 3, and 6. The AUTOCOAT profile was removed from the calculations for the latter sites without any significant effect on the other source estimates.

### Site Comparisons - Spatial Variability

Examination of the CMB results presented in Table 2 for site averages reveals both similarities and differences among the sites. The combined ROAD and GAS sources comprise by far the largest fraction of the total apportioned mass (83-91%). The HS accounts for only 4-15% of the total apportioned mass and was smallest at the downtown sites (sites 2 and 3). The combined fuel gases (PROPANE+NG) accounted for only a small fraction of the total (2-5%). AUTOCOAT was only found at sites 2 and 4, with site 2 having the largest percentage (5%). An automotive assembly plant, one of the largest VOC point sources in the Atlanta area, is located less than 10 km sites 2 and 4 (though not in the same direction from each site). The AUTOCOAT source impacts seen at these sites may also have contributions from one or more small but nearby sources.

ROAD and GAS were not resolvable based on the CMB7 criteria for resolving sources. However, it is worth noting their individual results, keeping in mind that the uncertainties are much larger than for the combined source estimate. While the sum of the ROAD and GAS percent varied little from site to site, the ratio of ROAD to GAS varied considerably (although uncertainties were sometimes quite large). The smallest ratios of ROAD to GAS (4-5) were observed at the downtown sites (sites 2 and 3), whereas ratios were much larger (10-13) at the outlying sites, with the largest ratio occurring at the background site (site 1). Whole gasoline emissions have several potentially significant sources, including spillage from refueling and fuel transfer and storage and hot soak evaporative emissions. Hot soak emissions are those evaporative emissions which escape from an automobile during the hour or so after it has been turned off but while it is still hot. These emissions are the most chemically similar to whole gasoline emissions. Differences in daytime traffic patterns in the vicinity of each site may explain the observed trend. Automobiles converge on the downtown area, mostly during the morning, and remain there only for the duration of the workday. One would expect a greater proportion of automobiles parked for a short time in the downtown area compared with the outlying areas where much of the traffic volume occurs (such as on the major interstate loop around the city). This distribution of automobile activity may explain the distribution of ROAD and GAS emissions suggested by the CMB7 results. Furthermore, automobiles are parked at residences in the outlying areas for longer periods of time and in greater numbers than in the downtown area. This longer period of inactivity would lead to a larger percentage of diurnal evaporative emissions and hence the larger percentage of gasoline headspace evaporative emissions in the outlying areas as suggested by the CMB results. Hourly results may shed some light on these hypotheses.

It is interesting to compare the percent of the TNMOC actually accounted for by the sources included in the CMB calculations, listed as SUM/TNMOC in Table 2 (recall that ethane must be subtracted from TNMOC to be consistent with the emission inventory). The smallest fraction of the TNMOC (49%) accounted for was found for site 1, the background site. This may be explained by the larger proportion of biogenic emissions likely to be found at the background site. Biogenic sources were not included in the CMB calculations. A much greater portion of the TNMOC (64-69%) was explained at the other sites with the exception of site 2, where only 54% of the TNMOC was accounted for. Hourly results may help explain these results.

## **COMPARISON OF CMB RESULTS WITH EMISSION INVENTORY**

### **Discussion of Limitations of Comparison and Possible Approaches**

The CMB model distinguishes among source types by their chemical differences, while emission inventories are generated based on a specific emission source or mechanism. Different emission mechanisms can produce emissions which are chemically similar and therefore would be treated as a single source type by the CMB model. Thus, the ability to directly compare CMB results with emission inventories is limited.

There are a number of different approaches that can be taken to maximize the information obtained from such comparisons. The inventory estimate for Highway Mobile Sources accounts for some, but not all, of the whole gas and headspace evaporative emissions. Therefore, a range of estimates for the Highway Mobile Sources should be calculated from the CMB estimates which would have a minimum equal to the ROAD estimate and a maximum equal to the sum of the ROAD, GAS, and HS estimates. One could get at the distribution of evaporative emissions possibly by using a typical MOBILE model<sup>18</sup> output to apportion the HS and GAS emissions that should be included in the Highway Mobile Source estimate. However, that approach would involve making some assumptions about the meaning of the ROAD estimate compared to the MOBILE model exhaust (plus running loss) estimate, the distribution of whole gasoline and gasoline vapor among the different evaporative emissions, and the accuracy of the ratio of evaporative to exhaust emissions, and would remove some of the independence of the ambient estimate.

The area and point sources are similarly difficult to deal with, in part because of the distribution of GAS and HS between mobile and non-mobile sources. A minimum estimate might equal the sum of NG, PROPANE, and AUTO COAT while a maximum estimate could equal the sum of these source estimates plus GAS and HS. This range could be compared to either area sources or the sum of area and point sources. The CMB may be missing or underestimating point and area sources, both because the CMB estimates are based only on FID measurements (and therefore may miss some chemical species important to architectural coatings, solvents, and other area sources) and because the impact of point sources on any particular sampling site is meteorology dependent.

Comparisons can include or exclude biogenic emissions. The ambient isoprene concentrations could be used to represent the lower-limit estimate of the biogenic fraction. Unfortunately, some suspiciously low values of isoprene are found at each site (as revealed by plots of isoprene versus 2,2-dimethylbutane). Another way to handle the biogenic fraction is to use the inventory value (16.8% biogenics) and then scale down the other source percent contributions estimated from the CMB so that the sum is 100%. This approach is probably closer to estimating the true biogenic emissions than using a minimum estimate and no worse than guessing species which should be included with biogenics and trying to estimate their reaction losses. Some of the species that should be included as biogenic emissions are aldehydes, which are not measured quantitatively by FID. Yet another approach would be to attribute all of the TNMOC unexplained by other sources to the maximum biogenic contribution. This approach could significantly overestimate the biogenic contribution, as species from other unapportioned sources plus secondary reaction products as well as biogenic species are all potential contributors to the unexplained TNMOC. A fourth alternative is to exclude biogenics altogether and renormalize the inventory



estimates to exclude biogenic emissions.

## Results and Discussion

The CMB results for the average site 2 sample (weekends excluded) were used to demonstrate the different approaches to comparing ambient data with emission inventories. A summary of these comparisons are presented in Table 3. The VOC emission inventory for Atlanta was obtained from the Georgia Department of Natural Resources.<sup>14</sup> The inventory was constructed to represent a typical ozone season 24-hour average weekday in 1990. The inventory developed for Fulton County, which encompasses most of downtown Atlanta, was judged to be the most appropriate for comparing with data from the site 2 site.<sup>5</sup> Inventory values are presented as a percent for all source categories and for source categories excluding biogenic sources (inventory estimates were normalized to sum to 100% in each case).

Starting with the comparisons which exclude biogenic emissions altogether, the inventory highway mobile source estimate is smaller than the minimum ambient data-derived highway mobile source estimate, and the inventory area plus point source estimate is larger than the maximum ambient data-derived estimate. A similar pattern is observed for the comparisons which accommodate biogenic emissions in a variety of ways, with the exception of the approach which sets the biogenic emissions equal to the total unexplained NMOC. In that comparison, the inventory-derived estimate for highway mobile sources exceeds the maximum ambient-derived estimate, and the difference between the inventory-derived and ambient-derived point plus area source estimate is even greater than for the other approaches. This observation is likely a direct result of attributing all the mass unexplained by the CMB to the biogenic fraction, an approach likely to overestimate the biogenic fraction while reducing the percent estimates of other sources because of renormalization. The inventory does not have a mechanism for determining whether mass has been unaccounted for, as does CMB.

The comparisons presented here are expressed in relative terms (percent of total emissions) rather than absolute terms (mass emitted per source category per unit time), resulting in a less sensitive comparison.<sup>19</sup> Nevertheless, the trend for most of the comparison approaches was that the inventory underestimates the highway mobile sources and overestimates the combined area and point sources. The comparison is complicated by the fact that the CMB results reflect mostly daytime (hours 8-18) conditions, while the inventory represents 24-hour average emissions.

## CONCLUSIONS

Data collected in Atlanta during the summer of 1990 are used to demonstrate some new approaches to comparing emission inventories with ambient data using CMB calculations. These new approaches make use of ranges of source estimates obtained from the ambient data. Ranges of source estimates can be used to deal with some of the inherent difficulties of comparing inventories with ambient data.

For most comparison approaches, the inventory highway mobile source estimate tended to be smaller than the minimum value estimated from the CMB results. However, the inventory estimate was derived from version 4.1 of the MOBILE model. Newer versions of that model may narrow the gap. Furthermore, the CMB estimate for mobile sources may include some non-automotive source emissions. An underestimation of the point plus area source category would also force the highway mobile sources to be overestimated (see below).

In the comparison in which the biogenic emissions were set equal to the total unexplained NMOC, the inventory highway mobile source estimate was larger than the CMB maximum estimate, most likely because this approach tends to overestimate biogenics, and thus underestimate the other sources on a percent basis. The results of Lewis et al., 1993<sup>5</sup> compared in the same manner (using maximum and minimum values and including unexplained TNMOC) yield an emission inventory estimate for highway mobile sources which falls in the middle of the minimum and maximum range of the CMB estimates. Differences between the Lewis et al., 1993 results and those of this analysis

may be due not only to differences in the dates and times the samples were collected but also the differences in the total NMOC reported by the different GC/FID systems used in the 2 studies.<sup>20</sup> The differences in the results of these 2 studies need to be investigated further.

The combination of point and area sources tended to be underestimated by the CMB compared with the emission inventory. Some chemical species which may be important components of solvents, coatings, and similar sources may not be measured adequately with a GC/FID and thus would not be included in estimates derived from the ambient data. Furthermore, chemical profiles used to represent the bulk of the point and area sources may be inadequate.

In general, comparisons of emission inventories with ambient data using CMB calculations are limited. Use of a more detailed highway mobile source inventory estimate in the comparisons may be helpful, as would uncertainty estimates for the emission inventory. The inventory used in these comparisons represents a 24-hour average of the entire summer, while the ambient data represented limited days and hours.

#### **ACKNOWLEDGEMENTS**

The authors thank Bobby Edmonds, John Wall, and Jerry Burger for collecting the canister samples used in this analysis.

#### **DISCLAIMER**

This paper has been reviewed in accordance with the U.S. Environmental Protection Agency's peer and administrative review policies and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendations for use.

#### **REFERENCES**

1. P.F. Nelson, S.M. Quigley, M.Y. Smith, Atmospheric Environment 1983 17, 439-449.
2. R.A. Harley, M.P. Hannigan, G.R. Cass, Environ. Sci. Technol. 1992 26, 2395-2408.
3. D.M. Kenski, R.A. Wadden, P.A. Scheff, W.A. Lonneman, "A receptor modeling approach to VOC emission inventory validation in five U.S. cities," in *Proceedings of the 86<sup>th</sup> Annual Meeting of the Air and Waste Management Association*, Denver, Colorado, 1993, Paper 93-WP-100.04.
4. P.A. Scheff and R.A. Wadden, Environ. Sci. Technol. 1993 27, 617-625.
5. C.W. Lewis, T.L. Conner, R.K. Stevens, J.F. Collins, R.C. Henry, in "Receptor modeling of volatile hydrocarbons measured in the 1990 Atlanta ozone precursor study," in *Proceedings of the 86<sup>th</sup> Annual Meeting of the Air and Waste Management Association*, Denver, CO, 1993, Paper 93-TP-58.04.
6. R.L. Seila, W.A. Lonneman, S.A. Meeks; *Determination of C<sub>2</sub> to C<sub>12</sub> Ambient Air Hydrocarbons in 39 U.S. Cities, From 1984 Through 1986*, EPA/600/3-89/058; U.S. Environmental Protection Agency: Research Triangle Park, 1989.
7. T.L. Conner, W.A. Lonneman, R.L. Seila "Transportation-related volatile hydrocarbon source profiles measured in Atlanta," accepted for publication in J. Air Waste Management Assoc.
8. L.J. Purdue et al.; *Atlanta Ozone Precursor Monitoring Study Data Report*, EPA-600/R-92/157; U.S. Environmental Protection Agency: Research Triangle Park, NC, 1992.



9. R.C. Henry, C.W. Lewis, J.F. Collins, Environ. Sci. Technol. 1994 28, 823-832.
10. J.G. Watson, N.F. Robinson, J.C. Chow, R.C. Henry, B.M. Kim, T.G. Pace, E.L. Meyer, Q. Nguyen, Environ. Software 1990 5, 38-49.
11. J.G. Watson et al.; *Receptor Model Technical Series, Volume III (1989 Revision): CMB7 User's Manual*, EPA-450/4-90/004; U.S. Environmental Protection Agency: Research Triangle Park, NC, 1990.
12. T.G. Pace and J.G. Watson; *Protocol for Applying and Validating the CMB Model*, EPA-450/4-87/010; U.S. Environmental Protection Agency: Research Triangle Park, NC, 1987.
13. W.E. Liss, W.H. Thrasher, G.F. Steinmetz et al., *Variability of Natural Gas Composition in Select Major Metropolitan Areas of the United States*, Report No. GRI-92/0123; Gas Research Institute: Chicago, IL, 1992, p. 58.
14. R. Methier, "Georgia's experience in preparing the 1990 baseline inventory," in *Proceedings of the EPA/A&WMA International Conference on Emission Inventory Issues*, 1992.
15. P.A. Scheff, R.A. Wadden, B.A. Bates, P.F. Aronian, IAPCA 1989 39, 469-478.
16. *Emission Inventory Requirements for Ozone State Implementation Plans*, EPA-450/4-91-010; United States Environmental Protection Agency: Research Triangle Park, NC, 1991.
17. R.B. Zweidinger, J.E. Sigsby, Jr., S.B. Tejada, F.D. Stump, D.L. Dropkin, W.D. Ray, J.W. Duncan, Environ. Sci. Technol. 1988 22, 956-962.
18. U.S. EPA, "User's guide to MOBILE5 (mobile source emission factor model)," U.S. Environmental Protection Agency, Office of Air and Radiation, Test and Evaluation Branch, 2625 Plymouth Road, Ann Arbor, MI, 48105, 1993.
19. C.W. Lewis, U.S. Environmental Protection Agency, Research Triangle Park, NC, personal communication, 1993.
20. J.H. Shreffler, JAWMA 1993 43, 1576-1584.

Table 1. Measured hydrocarbon species.

Hydrocarbon Species	Status	Hydrocarbon Species	Status
ethene	r	t-2-hexene	r
acetylene	*	c-2-hexene	r
ethane	n	methylcyclopentane	-
propene	r	2,4-dimethylpentane	*
propane	*	benzene	*
i-butane	*	cyclohexane	*
1-butene	r	2-methylhexane	o
n-butane	*	2,3-dimethylpentane	*
t-2-butene	r	3-methylhexane	-
c-2-butene	r	2,2,4-trimethylpentane	*
3-methyl-1-butene	r	n-heptane	-
i-pentane	*	methylcyclohexane	-
1-pentene	r	2,3,4-trimethylpentane	*
n-pentane	*	toluene	*
isoprene	x,r	2-methylheptane	r
t-2-pentene	r	3-methylheptane	r
c-2-pentene	r	n-octane	r
2-methyl-2-butene	r	ethylbenzene	*
2,2-dimethylbutane	x	m/p-xylene	r
cyclopentene	r	o-xylene	r
4-methyl-1-pentene	r	n-nonane	r
cyclopentane	-	i-propylbenzene	-
2,3-dimethylbutane	*	a-pinene	n,r
2-methylpentane	o	n-propylbenzene	*
3-methylpentane	*	1,3,5-trimethylbenzene	r
2-methyl-1-pentene	r	b-pinene	n,r
n-hexane	*	1,2,4-trimethylbenzene	r

KEY: x = unusual behavior in scatter plots

r = reactive

\* = fitting species

n = not available or not included in profiles

- = frequently underpredicted

o = offset

Table 2. Results of CMB analysis as percent of total mass apportioned (sum of source estimates), unless otherwise indicated.

Source Category	Site 1	Site 2 (weekends included)	Site 2 (weekends excluded)	Site 3	Site 4	Site 6
ROAD	79.1 ± 11.1%	69.3 ± 9.2%	72.7 ± 9.2%	75.0 ± 9.1%	78.1 ± 9.4%	75.1 ± 9.2%
GAS	6.3 ± 10.0%	16.9 ± 9.1%	14.0 ± 8.9%	16.1 ± 8.0%	7.6 ± 8.7%	7.6 ± 8.0%
ROAD+GAS	85.4 ± 4.4%	86.6 ± 4.9%	86.8 ± 4.6%	91.1 ± 12.1%	85.8 ± 4.5%	82.7 ± 3.6%
HS	9.4 ± 2.5%	4.3 ± 2.2%	4.0 ± 2.2%	6.4 ± 2.2%	10.7 ± 2.5%	14.7 ± 2.5%
PROP+NG	5.2 ± 0.7%	4.0 ± 0.6%	3.7 ± 0.6%	2.4 ± 0.4%	1.8 ± 0.4%	2.6 ± 0.4%
AUTOCOAT	0%	5.1 ± 1.1%	5.4 ± 1.1%	0%	1.8 ± 0.9%	0%
SUM	110.7 ± 5.7 ppbC	182.1 ± 10.0 ppbC	204.9 ± 10.8 ppbC	259.2 ± 31.9 ppbC	238.0 ± 12.4 ppbC	221.7 ± 9.7 ppbC
TNMOC	225.2 ± 19.5 ppbC	340.1 ± 19.7 ppbC	370.8 ± 18.8 ppbC	374.9 ± 19.6 ppbC	373.7 ± 19.6 ppbC	327.6 ± 19.5 ppbC
SUM/TNMOC	49.2 ± 5.0%	53.6 ± 4.3%	55.2 ± 4.0%	69.2 ± 9.3%	63.7 ± 4.7%	67.7 ± 5.0%
ROAD/GAS	12.6 ± 20.1	4.1 ± 2.3	5.2 ± 3.3	4.7 ± 2.4	10.3 ± 11.8	9.8 ± 10.4
Degrees of Freedom	13	12	12	13	12	13
R <sup>2</sup>	0.99	1.00	1.00	1.00	1.00	0.99
Chi <sup>2</sup>	0.69	0.32	0.30	0.26	0.44	0.77

Table 3. Comparison of emission inventory with relative source estimates derived from the site 2 ambient data (weekends excluded) using CMB calculations.

Summary of CMB Results*			Summary of Emission Inventory		
<b>Assume biogenic % = % reported in inventory (16.8%):</b>			<b>Including Biogenics:</b>		
	<b>MIN.</b>	<b>MAX.</b>			
<b>HIGHWAY MOBILE SOURCES:</b>	61%	75%			<b>HIGHWAY MOBILE SOURCES:</b> 56.0%
<b>POINT+AREA SOURCES:</b>	8%	23%	<b>POINT+AREA SOURCES:</b> 27.2%		
<b>Assume biogenic min. % = isoprene % (1.9%):</b>			<b>Excluding Biogenics:</b>		
	<b>MIN.</b>	<b>MAX.</b>			
<b>HIGHWAY MOBILE SOURCES:</b>	71%	89%			<b>HIGHWAY MOBILE SOURCES:</b> 67.3%
<b>POINT+AREA SOURCES:</b>	9%	27%	<b>POINT+AREA SOURCES:</b> 32.7%		
<b>Assume biogenic max. % = % unexplained (47.4%):</b>					
	<b>MIN.</b>	<b>MAX.</b>			
<b>HIGHWAY MOBILE SOURCES:</b>	41%	51%			
<b>POINT+AREA SOURCES:</b>	5%	15%			
<b>Assume no biogenics:</b>					
	<b>MIN.</b>	<b>MAX.</b>			
<b>HIGHWAY MOBILE SOURCES:</b>	73%	91%			
<b>POINT+AREA SOURCES:</b>	9%	27%			

\* Results presented as percent of total apportioned NMOC (sum of source estimates).

# **Comparison of Atlanta Emission Inventory with Ambient Data Using Chemical Mass Balance Receptor Modeling**

**Teri L. Conner  
John F. Collins  
William A. Lonneman  
Robert L. Seila**

**USEPA/AREAL  
USC  
USEPA/AREAL  
USEPA/AREAL**

## **STUDY OBJECTIVES**

- **Independent assessment of VOC source emissions**
- **Use ambient data and CMB receptor modeling**
- **Compare with emission inventory**
- **Emphasis on highway motor vehicle emissions**



## **What is Chemical Mass Balance Receptor Modeling?**

- **Ambient Concentrations = Sum of Contributions from Different Source Types**
- **Need Ambient Data and Chemical Profiles of Sources**

## **OUTDOOR AMBIENT SAMPLE COLLECTION**

- **Whole-air samples in stainless steel canisters**
- **6 sampling sites (urban, suburban, background)**
- **30 minute samples approx. every other day**
- **Start times rotated through hours 0, 8, 12, 15, 18**

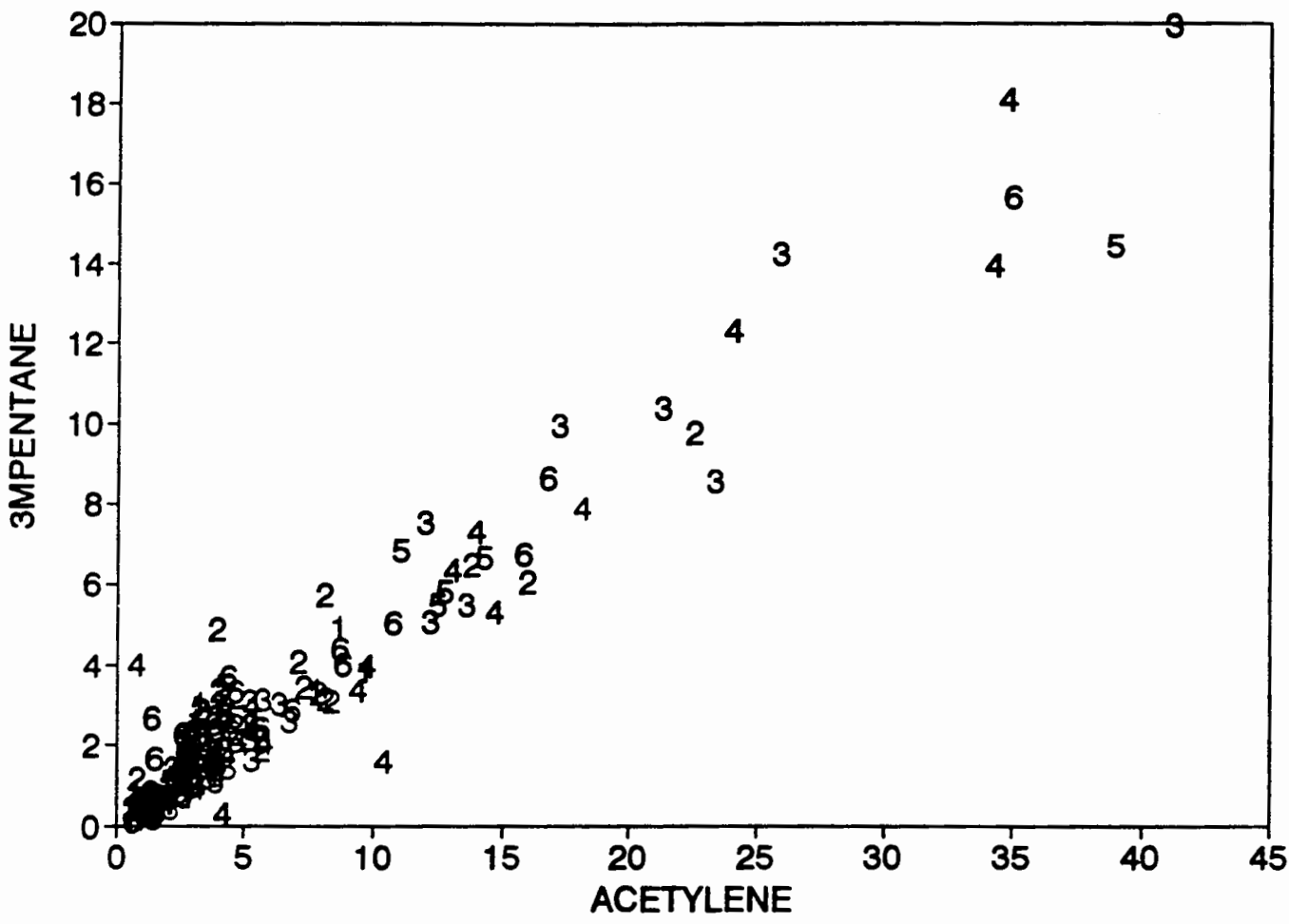
## **SAMPLE ANALYSIS**

- **$C_2 - C_{12} +$  non-methane hydrocarbons**
- **Two GC columns used to improve separation of  $C_2$ 's**
- **Flame-ionization detection**

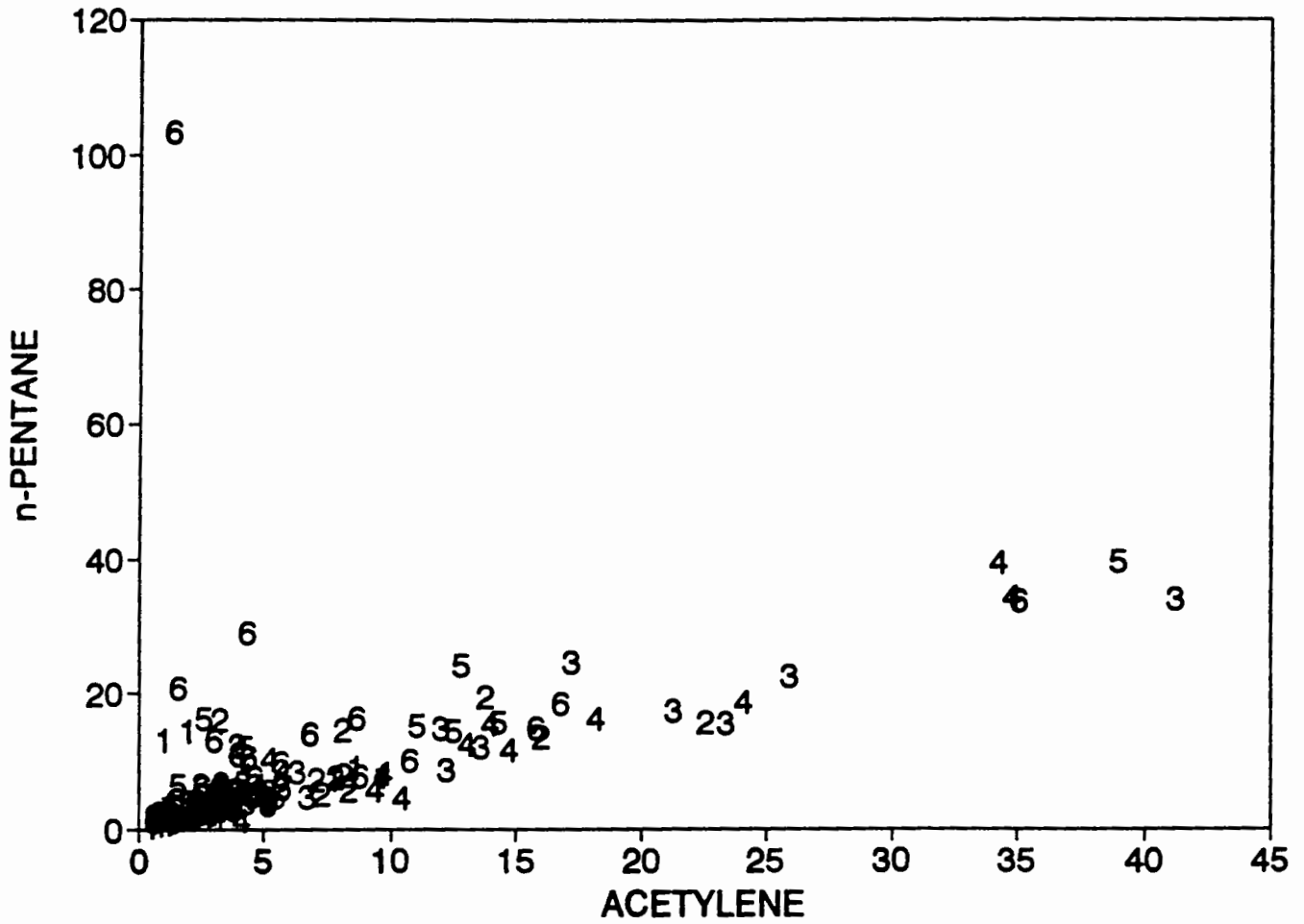
## **DATA SCREENING**

- **Necessary for any GC/FID data set**
- **Examine plots of each species vs other species**
- **Identify outliers and other problems**
- **Use procedure to guide selection of samples/species**

### CANISTER DATA



# CANISTER DATA





## **REASONS FOR EXCLUDING SPECIES FROM CMB CALCULATIONS**

- **Reactive**
- **Not available/not included in profiles**
- **Frequently underpredicted by CMB (missing source?)**
- **Offset or other unusual behavior in scatter plots**

**CMB FITTING SPECIES**  
**(18 out of possible 54 species)**

**acetylene**

**propane**

**i-butane**

**n-butane**

**i-pentane**

**n-pentane**

**2,3-dimethylbutane**

**3-methylpentane**

**n-hexane**

**2,4-dimethylpentane**

**benzene**

**cyclohexane**

**2,3-dimethylpentane**

**2,2,4-trimethylpentane**

**2,3,4-trimethylpentane**

**toluene**

**ethylbenzene**

**n-propylbenzene**

## **SOURCE PROFILES**

- **Derive from ambient data (SAFER)**
- **"Off-the-shelf" source measurements**
- **Concurrent, on-location source measurements**

## **PROFILES USED**

### **Measured:**

- **Emissions from a busy roadway**
- **Whole gasoline composite**
- **Headspace gasoline composite**

### **Literature:**

- **Auto painting profile**
- **Natural gas**

### **Other:**

- **Propane**

Table 2. Results of CMB analysis as percent of total mass apportioned (sum of source estimates), unless otherwise indicated.

Source Category	Site 1	Site 2 (weekends included)	Site 2 (weekends excluded)	Site 3	Site 4	Site 6
ROAD	79.1 ± 11.1%	69.3 ± 9.2%	72.7 ± 9.2%	75.0 ± 9.1%	78.1 ± 9.4%	75.1 ± 9.2%
GAS	6.3 ± 10.0%	16.9 ± 9.1%	14.0 ± 8.9%	16.1 ± 8.0%	7.6 ± 8.7%	7.6 ± 8.0%
ROAD+GAS	85.4 ± 4.4%	86.6 ± 4.9%	86.8 ± 4.6%	91.1 ± 12.1%	85.8 ± 4.5%	82.7 ± 3.6%
HS	9.4 ± 2.5%	4.3 ± 2.2%	4.0 ± 2.2%	6.4 ± 2.2%	10.7 ± 2.5%	14.7 ± 2.5%
PROP+NG	5.2 ± 0.7%	4.0 ± 0.6%	3.7 ± 0.6%	2.4 ± 0.4%	1.8 ± 0.4%	2.6 ± 0.4%
AUTOCOAT	0%	5.1 ± 1.1%	5.4 ± 1.1%	0%	1.8 ± 0.9%	0%
SUM	110.7 ± 5.7 ppbC	182.1 ± 10.0 ppbC	204.9 ± 10.8 ppbC	259.2 ± 31.9 ppbC	238.0 ± 12.4 ppbC	221.7 ± 9.7 ppbC
TNMOC	225.2 ± 19.5 ppbC	340.1 ± 19.7 ppbC	370.8 ± 18.8 ppbC	374.9 ± 19.6 ppbC	373.7 ± 19.6 ppbC	327.6 ± 19.5 ppbC
SUM/TNMOC	49.2 ± 5.0%	53.6 ± 4.3%	55.2 ± 4.0%	69.2 ± 9.3%	63.7 ± 4.7%	67.7 ± 5.0%
ROAD/GAS	12.6 ± 20.1	4.1 ± 2.3	5.2 ± 3.3	4.7 ± 2.4	10.3 ± 11.8	9.8 ± 10.4
Degrees of Freedom	13	12	12	13	12	13
R <sup>2</sup>	0.99	1.00	1.00	1.00	1.00	0.99
Chi <sup>2</sup>	0.69	0.32	0.30	0.26	0.44	0.77

## **PROBLEMS TO OVERCOME**

- **Compare ambient data and emission inventory on equal basis (e.g., species represented, definition of total emissions)**
- **Sources distinguished chemically by CMB vs mechanistically by inventory**
- **Representativeness of source profiles for CMB calculations**
- **Meet other CMB requirements**



# Source Profiles and MOBILE5 Outputs

---

**Reconcile ambient measurements with emissions inventories using CMB and profiles.**

---

- **MOBILE5 - distinguishes emissions by mechanism.**
- **CMB/profiles - distinguishes emissions by composition.**

## **COMPARISONS OF CMB RESULTS WITH EMISSION INVENTORY**

- **Calculate % of total mass apportioned (sum of source estimates) rather than percent of total NMOC**
- **Exclude ethane from TNMOC in CMB calculations and profile normalization**
- **Compare CMB with inventory on a relative basis**
- **Report minimum and maximum source estimates from CMB calculations**

## **1990 ATLANTA EMISSION INVENTORY**

	<b>Fulton Co.</b>	<b>13 Co.</b>
<b>Highway Mobile Sources</b>	<b>56.0%</b>	<b>39.8%</b>
<b>Area Sources</b>	<b>20.6%</b>	<b>19.4%</b>
<b>Point Sources</b>	<b>6.6%</b>	<b>3.9%</b>
<b>Biogenic Sources</b>	<b>16.8%</b>	<b>36.9%</b>

## **HIGHWAY MOBILE SOURCES**

- **Minimum = ROAD**
- **Maximum = ROAD + GAS + HS**

## **POINT + AREA SOURCES**

- **Minimum = NG + PROPANE + AUTOCOAT**
- **Maximum = NG + PROPANE + AUTOCOAT + GAS + HS**

## **BIOGENICS - DEAL WITH THEM INDIRECTLY**

- **Assume inventory % is correct; renormalize other source estimates accordingly**
- **Assume biogenic % = isoprene % (minimum estimate)**
- **Assume biogenic % = total unexplained NMOC (maximum estimate)**
- **Assume no biogenics; renormalize inventory to exclude biogenics**

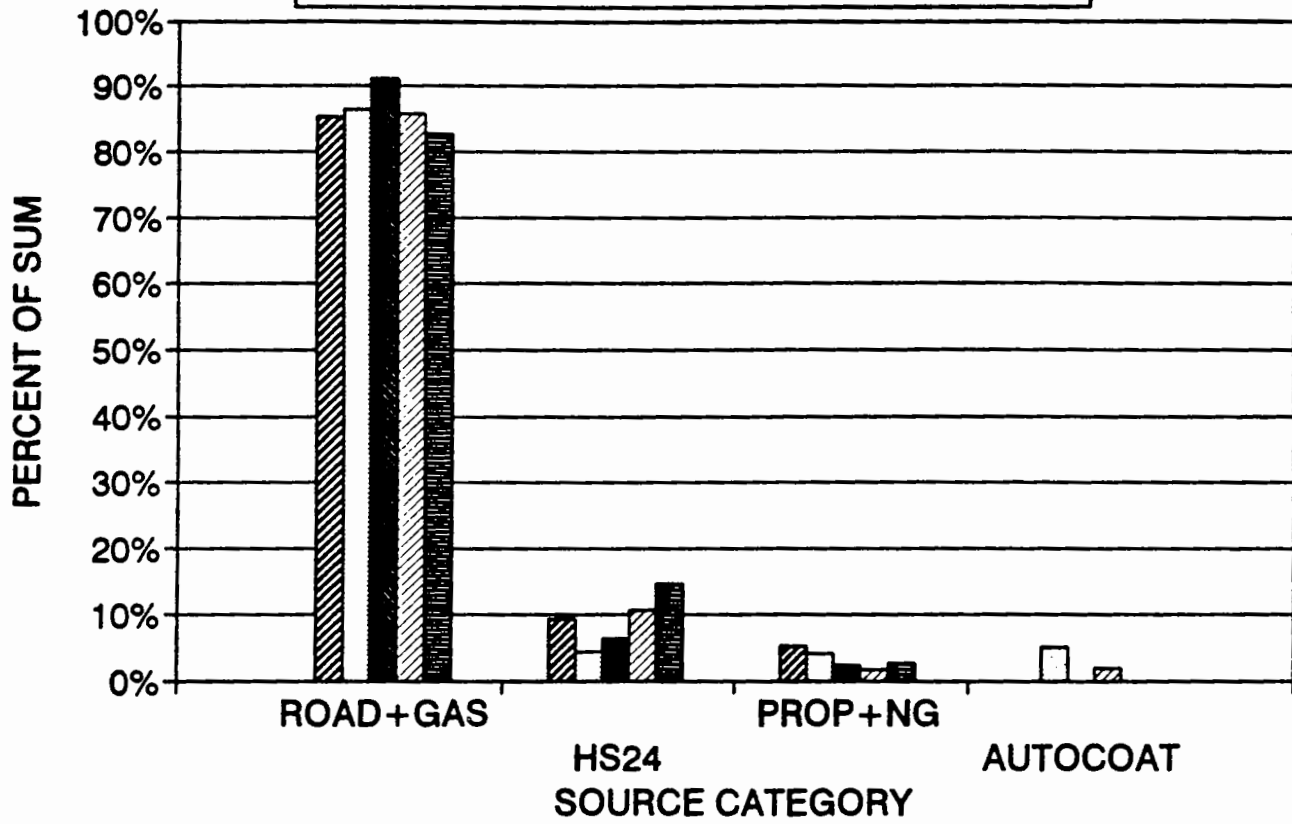


Table 3. Comparison of emission inventory with relative source estimates derived from the site 2 ambient data (weekends excluded) using CMB calculations.

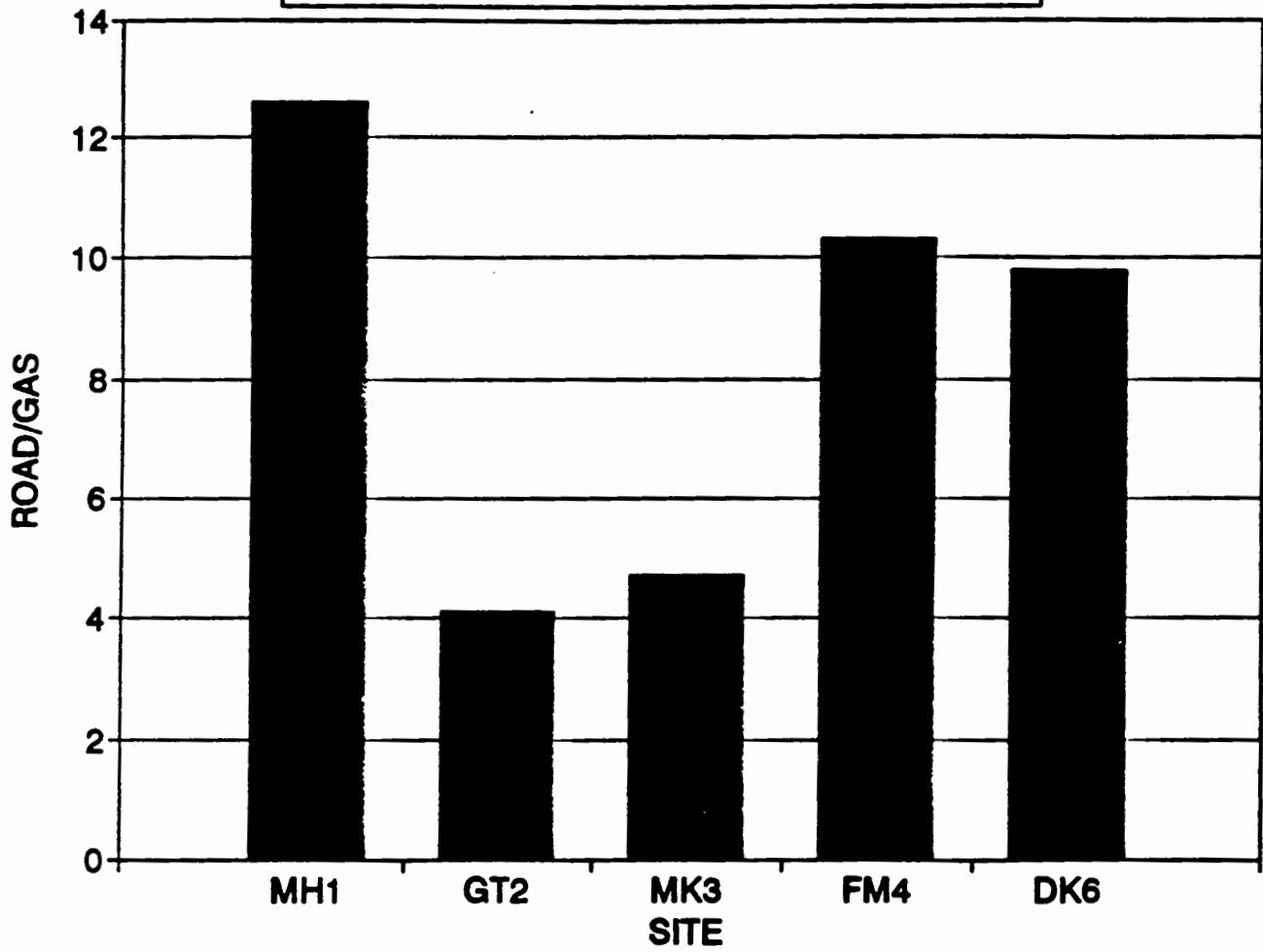
Summary of CMB Results*			Summary of Emission Inventory	
<u>Assume biogenic % = % reported in inventory (16.8%):</u>				
	<b>MIN.</b>	<b>MAX.</b>		
<b>HIGHWAY MOBILE SOURCES:</b>	61%	75%		
<b>POINT+AREA SOURCES:</b>	8%	23%	<u>Including Biogenics:</u>	
<u>Assume biogenic min. % = isoprene % (1.9%):</u>				
	<b>MIN.</b>	<b>MAX.</b>		
<b>HIGHWAY MOBILE SOURCES:</b>	71%	89%	<b>HIGHWAY MOBILE SOURCES:</b>	56.0%
<b>POINT+AREA SOURCES:</b>	9%	27%	<b>POINT+AREA SOURCES:</b>	27.2%
<u>Assume biogenic max. % = % unexplained (47.4%):</u>				
	<b>MIN.</b>	<b>MAX.</b>		
<b>HIGHWAY MOBILE SOURCES:</b>	41%	51%		
<b>POINT+AREA SOURCES:</b>	5%	15%	<u>Excluding Biogenics:</u>	
<u>Assume no biogenics:</u>				
	<b>MIN.</b>	<b>MAX.</b>		
<b>HIGHWAY MOBILE SOURCES:</b>	73%	91%	<b>HIGHWAY MOBILE SOURCES:</b>	67.3%
<b>POINT+AREA SOURCES:</b>	9%	27%	<b>POINT+AREA SOURCES:</b>	32.7%

\* Results presented as percent of total apportioned NMOC (sum of source estimates).

**CMB RESULTS FOR AVERAGE SAMPLES**



**CMB RESULTS FOR AVERAGE SAMPLES**



## **SUMMARY**

- **Method causes source %'s to be interdependent**
- **Representativeness of solvent-type source uncertain**
- **No breakout of components of highway mobile source**
- **Inventory and ambient data do not represent exactly the same dates and times**
- **Collinear sources in CMB**
- **No uncertainty estimates for emission inventory**

**TECHNICAL REPORT DATA**

1. REPORT NO. EPA/600/A-94/244		2.	3. RECIPI
4. TITLE AND SUBTITLE Comparison of Atlanta Emission Inventory with Ambient Data Using Chemical Mass Balance Receptor Modeling		5. REPORT DATE	
		6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) Teri L. Conner <sup>1</sup> , John F. Collins <sup>2</sup> , William A. Lonneman <sup>1</sup> , Robert L. Seila <sup>1</sup>		8. PERFORMING ORGANIZATION REPORT NO.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS <sup>1</sup> U.S. Environmental Protection Agency/AREAL MD-47 Research Triangle Park, NC, 27711  <sup>2</sup> University of Southern California Civil Engineering Dept., Env. Engineering Pgm. 3620 South Vermont Avenue Los Angeles, CA, 90089-2531		10. PROGRAM ELEMENT NO.	
		11. CONTRACT/GRANT NO.  CR-818410 (Coop. Agr. with USC)	
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency/AREAL MD-47 Research Triangle Park, NC, 27711		13. TYPE OF REPORT AND PERIOD COVERED Symposium Paper and Presentation	
		14. SPONSORING AGENCY CODE EPA/600/09	
15. SUPPLEMENTARY NOTES For presentation at EPA/A&WMA conference "The Emission Inventory: Applications and Improvement" to be held Nov. 1-3 in Raleigh, NC.			
16. ABSTRACT A comprehensive hydrocarbon database was obtained at 6 sites in the Atlanta metropolitan area during the summer of 1990. Samples were collected in stainless steel canisters and analyzed for 54 hydrocarbon species plus total non-methane organic compounds (TNMOC). The contributions of the major sources of TNMOC at each of the 6 sites were estimated through a procedure called Chemical Mass Balance (CMB) receptor modeling. Spatial variability of the source contributions is discussed. Results of the CMB analysis for one of the sites are compared with the emission inventory for Atlanta using several different approaches. The inventory highway mobile source estimate tends to be smaller than the <u>minimum</u> ambient data-derived highway mobile source estimate, and the inventory area plus point source estimate tends to be larger than the <u>maximum</u> ambient data-derived estimate for the data set examined. However, these source estimates are interdependent to some extent. Limitations of these comparisons are discussed.			
17. KEY WORDS AND DOCUMENT ANALYSIS			
a. DESCRIPTORS	b. IDENTIFIERS/ OPEN ENDED TERMS	c. COSATI	
18. DISTRIBUTION STATEMENT <u>Release to Public</u>	19. SECURITY CLASS ( <i>This Report</i> ) <u>Unclassified</u>	21. NO. OF PAGES	
	20. SECURITY CLASS ( <i>This Page</i> ) <u>Unclassified</u>	22. PRICE	