

Sources of Bay Area Fine Particles

Technical Report

Prepared by: David Fairley, Statistician

*Contributions & review from: Saffet Tanrikulu, Research & Modeling Section Manager
Philip Martien, Senior Advanced Project Advisor*

Approved by: Henry Hilken, Director of Planning, Rules and Research Division

April 2008

**Bay Area Air Quality Management District
Planning, Rules and Research Division
939 Ellis Street
San Francisco, CA 94109**

April 2008

DISCLAIMER

This report presents an analysis of available data to provide information on Bay Area PM_{2.5} sources for the District's activities to reduce exposure to fine particulate matter. Data were obtained from the field program of the California Regional Particulate Air Quality Study (CRPAQS) and from routine air monitoring stations. To estimate the contributions from various sources, a Chemical Mass Balance (CMB) analysis was conducted. Because of limitations in the data and uncertainties in the CMB analysis, some of the findings presented in this report should be viewed as preliminary. Certain assumptions, as described in the document, were made to aid the analysis. While the District continues making additional routine particulate matter measurements, analyses similar to those presented in this report will be conducted and the findings of this report will be updated as appropriate.

CONTENTS

CONTENTS.....	1
FIGURES.....	2
TABLES.....	2
SUMMARY.....	3
Methods.....	3
Results and Key Findings.....	4
1. INTRODUCTION.....	6
2. DATA.....	6
2.1 Ambient Data.....	7
2.2 Source Profiles.....	7
2.3 Carbon-14 Analysis.....	8
3. SUMMMARY STATISTICS.....	9
3.1 Total Mass Comparisons.....	9
3.2 Seasonal Distribution of PM _{2.5}	9
3.3 Speciation of PM _{2.5}	10
3.4 Carbon-14 Analysis Results.....	13
4. SOURCE APPORTIONMENT METHODOLOGY.....	14
4.1 CMB Modeling.....	14
4.2 Incorporating Carbon-14 measurements.....	16
5. SOURCE APPORTIONMENT RESULTS.....	17
5.1 Annual Summary.....	17
5.2 Peak Summary.....	19
5.3 Analysis by Season.....	21
5.4 Comparisons with the 1994 Winter PM ₁₀ Field Study.....	24
6. REFINING SOURCE APPORTIONMENT WITH EMISSIONS INVENTORY ESTIMATES.....	26
7. KEY FINDINGS AND ON-GOING WORK.....	32
7.1 Key Findings.....	32
7.2 Ongoing work.....	32
8. REFERENCES.....	34
Appendix A. Adjustment of OC and EC data.....	A1
Appendix B. Comparison of mass measurements.....	B1
Appendix C. Chemical species used in the CMB model.....	C1
Appendix D. Determining weights applied to the fits.....	D1
Appendix E. Initial CMB results and source category selection.....	E1
Appendix F. Modification of CMB results Based on Carbon-14 Analysis.....	F1

FIGURES

Figure 3.1. Quarterly averaged PM _{2.5} for special study sites.	10
Figure 3.2. Mean mass contribution of PM _{2.5} species	11
Figure 3.3. Boxplots of most abundant species by site.....	12
Figure 3.4. Annual percentages of contemporary carbon from Bay Area sites for 1998 and 2004.....	13
Figure 5.1. Estimated annual source contributions to Bay Area ambient PM _{2.5}	17
Figure 5.2. Estimated annual percent contributions from various source categories	18
Figure 5.3. Estimated source contributions to peak Bay Area ambient PM _{2.5}	19
Figure 5.4. Estimated annual percentage contributions to peak PM _{2.5}	20
Figure 5.5a. PM _{2.5} source contributions by season for Bethel Island	21
Figure 5.5b. PM _{2.5} source contributions by season for Livermore	22
Figure 5.5c. PM _{2.5} source contributions by season for San Francisco.....	22
Figure 5.5d. PM _{2.5} source contributions by season for San Jose 2000	23
Figure 5.5e. PM _{2.5} source contributions by season for San Jose 2001	23
Figure 5.5f. PM _{2.5} source contributions by season for Point Reyes.....	24
Figure 5.6. Comparison with percent source attributions from the winter 1993-94 field study.....	26
Figure 6.1. Estimated contributions to annual Bay Area PM _{2.5} from various sources	29
Figure 6.2. Estimated contributions to peak Bay Area PM _{2.5} from various sources	31
Figure A1. Comparisons of EC/TC ratios.....	A1
Figure B1. Comparison of PM _{2.5} mass measured by CARB and BAAQMD for Livermore, 2000.....	B1
Figure E1. Comparison of model uncertainties with total uncertainty	E3

TABLES

Table 2.1. Source profiles used in CMB analysis.....	8
Table 5.1. CMB results for annual PM _{2.5} samples.....	18
Table 5.2. CMB results for peak PM _{2.5} samples.....	20
Table 5.3. CMB results for aggregated wintertime high PM ₁₀ samples, 11/93-1/94.....	25
Table 6.1. Emissions inventory estimates of carbonaceous and secondary precursors ...	28
Table 6.2. Percent estimated total annual contribution from various sources	29
Table 6.3. Emissions inventory estimates of winter carbonaceous and secondary precursors.....	30
Table 6.4. Estimated percentage total peak contribution from various sources	30
Table A1. Mean and median EC/TC ratios.....	A2
Table B1. Same-day comparisons of BAAQMD, CARB and EPA PM _{2.5} measurements	B2
Table C1. Chemical species – frequency above the limits of detection	C1

SUMMARY

This study uses Bay Area ambient particulate matter (PM) measurements and emissions data to determine the major PM sources and their approximate contributions to Bay Area PM concentrations.

The focus of this study is a chemical mass balance (CMB) analysis where a computer model is used to apportion ambient PM collected on filters to a set of source categories. The filters contain particles less than or equal to 2.5 microns in diameter (PM_{2.5}) collected over 24-hour periods at monitoring sites around the Bay Area. Each filter was analyzed for a range of chemical species. The same species were measured in special studies of emissions from various sources, such as motor vehicles and wood burning. The CMB model finds the mix of these source measurements that best matches the ambient sample, chemical species by chemical species.

CMB analysis was limited to identifying source *categories*, such as fossil fuel combustion. In order to make finer distinctions, the CMB results were combined with the Bay Area Air Quality Management District's emissions inventory.

Methods

The ambient data were obtained from four different studies or agencies: 1) the California Regional Particulate Air Quality Study (CRPAQS), 2) the Speciated Trends Network (STN), 3) the IMPROVE network, and 4) Bay Area Air Quality Management District (BAAQMD) routine measurements. The source profile data were obtained from sample collections of CRPAQS, the California Air Resources Board (ARB) and BAAQMD.

Speciated ambient data were collected from five monitoring sites – Bethel Island, Livermore, San Francisco, San Jose, and Point Reyes. The data for the first three sites were collected as part of the CRPAQS. They cover a 14 month period, extending from December 1999 through January 2001. The Point Reyes (IMPROVE) and San Jose (STN) data for 2000 were obtained to supplement the CRPAQS measurements. San Jose data from 2001 were also added to the analysis to investigate how source contributions vary from year to year.

Source categories used in the CMB analysis initially included fossil fuel combustion, vegetative burning, commercial cooking, tire and brake wear, geological dust, marine air, ammonium nitrate, and ammonium sulfate. CMB analysis is limited in that the contributions from certain sources may occur at levels below the model's ability to detect; and certain combinations of sources may be effectively impossible to differentiate if their chemical profiles are too similar.

In the initial CMB analysis, it was found that PM from tire and brake wear did not exceed the limits of detection, and commercial cooking could not be differentiated from vegetative burning or fossil fuel combustion. These sources were omitted from subsequent analyses.

Considerable uncertainty remained in the apportionment of two source categories – vegetative burning and fossil fuel combustion. To better apportion these categories, carbon-14 was measured on PM₁₀ filters. This analysis provides a reliable estimate of the amount of "new carbon", from vegetative burning and cooking, relative to "fossil carbon", from burning fossil fuels. The results of this analysis were used to adjust the CMB estimates of the vegetative burning and fossil fuel combustion apportionment.

The CMB analysis produced estimates of the contributions from source *categories*, but could not differentiate between wood smoke and cooking or among on-road, off-road, refinery, and power plant emissions. To better differentiate among individual sources, the CMB results were combined with emissions inventory estimates for the Bay Area. The combination apportions not only directly emitted PM_{2.5} but also the ammonium nitrate and ammonium sulfate with the assumption that these are proportional to the emissions of NO_x and SO₂ respectively.

The California and federal governments each have established two PM standards – a 24-hour standard to protect against short-term PM exposure, and an annual standard to protect against long-term exposure. The CMB results are presented in terms of peak and annual average PM to correspond with these standards.

Results and Key Findings

This study found that the three source categories – vegetative burning, fossil fuel combustion and ammonium nitrate – are the largest contributors to annual average PM_{2.5}. Marine air and ammonium sulfate are substantial contributors. Geological dust is a relatively minor contributor.

The CMB results were also summarized for peak PM_{2.5} – the average of the 10 highest PM_{2.5} samples for each site. For peak PM_{2.5}, the top three source categories – vegetative burning, fossil fuel combustion and ammonium nitrate – comprise more than 90% of the total. Marine air and ammonium sulfate make up a much smaller percentage of the total; and geological dust is negligible.

The combined CMB and emissions inventory analysis showed that vegetative burning was the largest single source, contributing about 24% of annual PM_{2.5} and 33% of peak PM_{2.5}. On-road vehicles were the second largest source of PM_{2.5}. Off-road vehicles, commercial cooking, ships, refining and marine air were also found to be large contributors.

Additional findings are listed below:

- Most anthropogenic PM₁₀ and PM_{2.5} derive from burning wood or fossil fuels.
- Geological dust is a small contributor to PM₁₀ and a negligible contributor to PM_{2.5}. Tire/break wear is also a negligible PM_{2.5} source.

- Peak PM occurs largely in winter. Reasons include a high rate of ammonium nitrate formation and more wood burning, both related to low temperatures. Analysis of wintertime meteorological data shows that periods of low winds are conducive to the buildup of PM_{2.5}.
- Ammonium nitrate contributes almost 40% to peak PM_{2.5} and 20% to annual PM_{2.5} on average.
- Carbonaceous PM, that is, PM directly emitted from burning, accounts for roughly half of peak PM_{2.5} and annual PM_{2.5}.
- Ammonium sulfate contributes about 10% to annual PM_{2.5}, but only 5% to peak PM, on average.

1. INTRODUCTION

The Bay Area Air Quality Management District (BAAQMD or District) is responsible for assuring clean air in the nine-county San Francisco Bay Area. The Bay Area occasionally exceeds health-based ambient air quality standards for particulate matter.

Airborne particulate matter (PM) has serious adverse effects on health. High levels of PM, especially PM_{2.5}, are correlated with exacerbations of respiratory problems such as asthma, increases in emergency room visits, and increases in respiratory and cardiac related deaths. Such effects have been noted in the San Francisco Bay Area.

In December 2006, the U.S. EPA adopted a new ambient air quality standard for fine particulate matter, or PM_{2.5}, of 35 µg/m³. Monitoring data indicates that the Bay Area may be designated non-attainment for the new standard. The Bay Area does not meet the more stringent California PM_{2.5} and PM₁₀ standards.

In order to reduce the Bay Area's PM levels most effectively, it is necessary to know what the major sources are and their approximate contributions to the total PM. Many sources contribute to PM. Sources of direct, or *primary*, emissions include on-road and off-road vehicles, power plants, refineries, wood burning, cooking, and dust from roads, fields, construction, and farming.

PM also forms from chemical processes in the atmosphere. Earlier studies have shown that these chemically formed particulates, or *secondary particulates*, constitute a sizeable fraction of the Bay Area's PM. The major secondary components are ammonium nitrate, formed from transformation of NO_x and ammonia, and ammonium sulfate, formed from transformation of sulfur dioxide and ammonia. There are also secondary organic molecules formed from chemical transformation of gaseous organic molecules in the atmosphere.

This study analyzes speciated PM_{2.5} collected at several Bay Area sites using a computer model (chemical mass balance, or CMB) to estimate the contributions of various sources to total PM. The analysis is summarized both annually and for peak PM to correspond with the annual and 24-hour PM standards. The results are studied together with BAAQMD emissions inventory estimates to provide a finer breakdown of source contributions. Carbon-14 analyses were also used to help distinguish between PM contributed by wood burning and fossil fuel combustion.

2. DATA

In order to conduct CMB analysis, both ambient and source profile data were needed. The ambient data were obtained from four different studies or agencies: 1) the California Regional Particulate Air Quality Study (CRPAQS), 2) the Speciated Trends Network, 3) the IMPROVE network, and 4) Bay Area Air Quality Management District (BAAQMD) routine measurements. The source profile data were obtained from CRPAQS, and from ARB and BAAQMD sample collections.

2.1 Ambient Data

CRPAQS included speciated analysis of filters containing PM_{2.5} samples collected over 24-hour periods from a variety of central California sites, including three in the Bay Area: Bethel Island (BI), Livermore (LI) and San Francisco (SF). The speciated analysis included measurement of elements, using X-Ray Fluorescence; ions, using ion chromatography; and elemental and organic carbons, using the IMPROVE methodology. Samples were collected on a 1-in-6 day schedule from December 1999 through January 2001.

The Speciated Trends Network (EPA 1999) collected speciated PM_{2.5} at San Jose – 4th St. on a 1-in-3 day schedule. PM_{2.5} data were obtained for years 2000 (SJ0) and 2001 (SJ1). The abbreviation SJ will be used when referring to both years. The 2001 data are used to investigate changes in source contributions from 2000 to 2001 at this site.

The BAAQMD routine monitoring network collected gravimetric PM_{2.5} data at Livermore, San Francisco and San Jose contemporaneously with these other networks.

The IMPROVE network, which collects PM_{2.5} data from national parks around the United States, provided data at Point Reyes (PR). These data were collected on a 1-in-3 day schedule. Most of the time, the PR site samples air coming onshore from the Pacific Ocean so typically its samples are representative of background PM.

Generally, the species measurements seemed consistent among the labs conducting data analyses as confirmed by an inter-lab comparison (EPA/NAREL 2002). One exception is the measurement of carbon. There are two distinct approaches for carbon measurements: the NIOSH and IMPROVE methods. Although the total carbon measurements are similar, the NIOSH method attributes a much higher percentage of carbon to OC than the IMPROVE method (Chow et al., 2001). The source profiles as well as the CRPAQS and Point Reyes samples used the IMPROVE method, whereas the EPA/RTI samples were analyzed by the NIOSH method. In order to make the two sets of measurements more comparable, a conversion from NIOSH to IMPROVE was made as presented in Appendix A.

2.2 Source Profiles

CRPAQS source profile data included a range of samples: geological, woodsmoke, motor vehicle exhaust, commercial cooking, and tire and brake wear. These profiles were combined with PM₁₀ source profiles available from a previous District CMB analysis conducted in 1994. A subset of these profiles was used for CMB analysis, listed in Table 2.1.

Table 2.1. Source profiles used in CMB analysis

Source Abbrev.	Agency ^a	Description
SJ4PVRD	BAAQMD/DRI	Paved road dust from San Jose
BYRDC	BAAQMD/DRI	Paved road dust from selected sites around the Bay Area
ARB_DUST	CRPAQS	Paved road dust from around central California
AMSUL		Ammonium sulfate
AMNIT		Ammonium nitrate
NASUL		Sodium sulfate
NANIT		Sodium nitrate
MARINE		Marine air with intermediate aging
MAR0		Fully aged marine air
MAR100		Fresh marine air
GUNPOWDR		Gun powder from fireworks
BYTUN4	BAAQMD/DRI	Composite motor vehicle exhaust profile
BYWS4	BAAQMD/DRI	Composite wood burning
WBOakEuc	CRPAQS	Oak & Eucalyptus wood burning
BurnWdAg	CRPAQS	Composite wood burning
WBOak	CRPAQS	Oak burning profile
WBEucal	CRPAQS	Eucalyptus burning
COOK	CRPAQS	Composite meat cooking
CAMV	CRPAQS	CRPACS combined motor vehicle
TireBrke	CRPAQS	Tire and Brake wear

^a Agency that collected the data: BAAQMD samples were collected in 1993, DRI (Desert Research Institute). No agency indicates that the source has a known chemical composition.

2.3 Carbon-14 Analysis

Carbonaceous material mostly from the combustion of fossil fuels and vegetative materials represents a large fraction of the Bay Area's PM_{2.5}. However, the contributions of the various carbonaceous sources are difficult to distinguish based on the set of measurements made. Carbon-14 analysis offers a way to clearly distinguish two categories of these compounds.

Carbon-14 is a naturally occurring radioactive isotope of carbon that is constantly being replenished in the atmosphere, representing about one part in a trillion. Living things contain this fraction of C-14. On the other hand, C-14 has a half-life of about 5,000 years, so that fossil fuels contain virtually none. Thus, the fraction of contemporary C-14 in the PM allows us to differentiate between sources with contemporary concentrations of C-14 (vegetative smoke, cooking emissions, and secondary biogenic PM), and sources with fossil concentrations (diesel, gasoline, and natural gas).

Estimates of annual fractions of contemporary and fossil carbon for each site were obtained as follows. PM₁₀ filters were used because they are quartz and because they contain a much greater mass of material to analyze. For each of the District's PM₁₀ sites, a representative set of filters from 2004 was selected, and a composite sample created. For selected sites, composites of 1998 filters were also selected. The District contracted

with the University of Arizona for a C-14 analysis of these composite filters. Section 3.4 contains the results.

3. SUMMARY STATISTICS

Before conducting the CMB analysis, some of the co-located measurements by BAAQMD, the U.S. EPA and CRPAQS at Livermore, San Jose, and San Francisco were compared for consistency among different measurement methods. Available ambient data were summarized to better understand the seasonal distribution of PM_{2.5}, its composition, and the locations of peak concentrations.

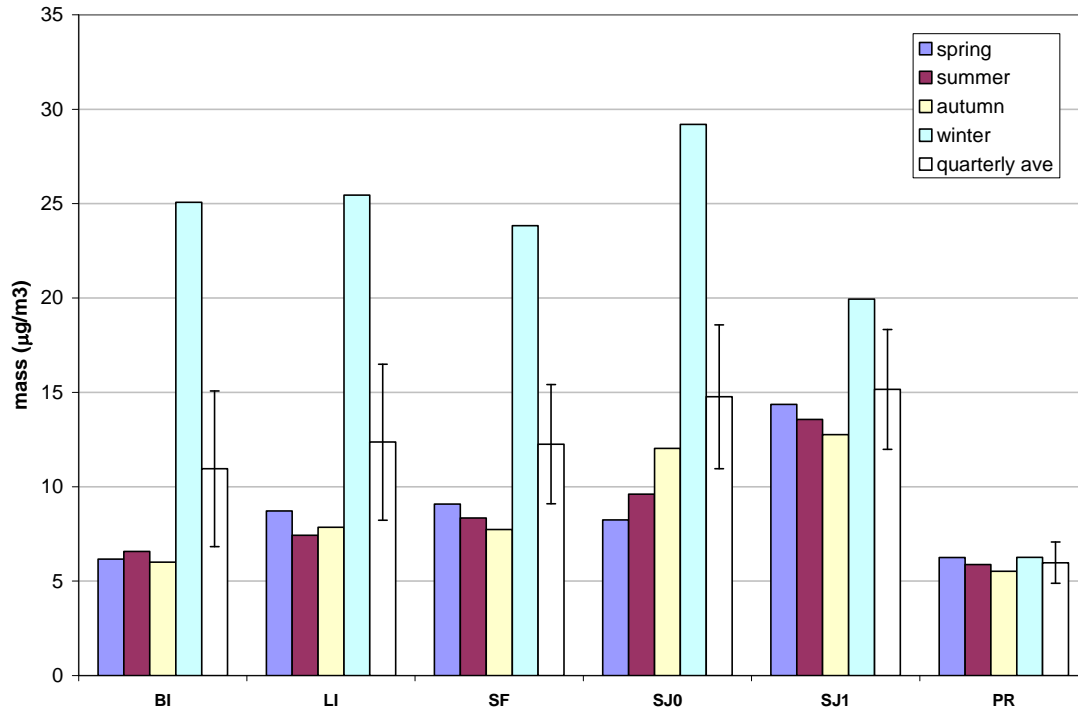
3.1 Total Mass Comparisons

A comparison of CRPAQS, the U.S. EPA and BAAQMD total mass measurements shows a high degree of correlation for all 3 sites (LI, SJ and SF), but a substantial difference in magnitude for the LI and SF sites. In particular, the CRPAQS measurements average about 70-75% of the District measurements. The masses were also compared with the sums of the individual chemical species. The sums of species are expected to be smaller than the total because not all species are included. Yet the CRPAQS masses were often less than the sum of species, whereas the BAAQMD masses were generally larger (see Appendix B). Because of the large discrepancies in the LI and SF measurements, their total mass measurements, along with measurements at Bethel Island, were adjusted to be consistent with BAAQMD measurements. For San Jose, measured for EPA, there was a slight difference in masses, with the EPA measurements averaging about 105% of the District measurements. These measurements were not adjusted.

3.2 Seasonal Distribution of PM_{2.5}

The seasonal distribution of PM_{2.5} was studied to better understand when the highest concentrations occur in the different regions of the District. The seasons were defined to correspond to the PM_{2.5} season, namely for this study "winter" is defined to include November, December, and January. The other quarters follow from the definition of the winter quarter: "spring" is February, March and April; "summer" is May, June and July; and "autumn" is August, September and October.

Figure 3.1. Quarterly averaged PM_{2.5} for special study sites



Note: Data from 2000 except for 2001 for SJ1. Quarters are spring = Feb-Apr, summer = May-July, autumn=Aug-Oct, winter=Nov-Jan. These quarters were chosen to correspond to the PM season rather than the standard definition of seasons. White bars are averages of the 4 quarters. Uncertainty bars represent 90% confidence intervals for the annual mean.

Figure 3.1 shows quarterly averages and the average of the four quarters. The figure shows a clear seasonal pattern, with the winter quarterly averages far exceeding those of the other quarters, except at Point Reyes. During the non-winter quarters, the 2000 average values are not much higher than those of the Point Reyes site. Note the large difference between 2000 and 2001 for San Jose, which is likely due to meteorology. In particular, the winter quarter decreases from 2000 to 2001 whereas the other quarters increase.

The white bars show the means of the 4 quarters. Also shown are 90% confidence intervals for the means. Except for the Point Reyes background site, the confidence intervals for all the other sites straddle the national annual standard of 15 µg/m³ and the state annual standard of 12 µg/m³. Thus, the Bay Area is on the borderline for these standards.

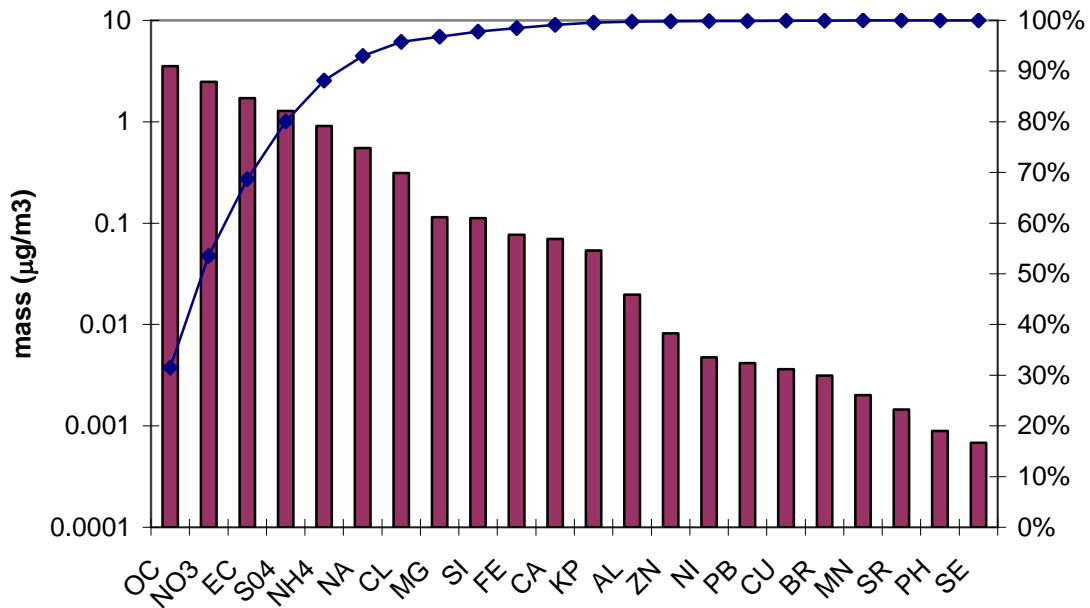
3.3 Speciation of PM_{2.5}

Many PM_{2.5} species were measured, but for a number of the elements, the concentrations rarely if ever exceed the limits of detection. Appendix C has a list of all species and the number of times the measured concentrations exceeded one and two standard errors. Those species whose concentrations exceeded one standard error in less than 10 samples

were eliminated from further analysis because they would likely detract from the capacity of the CMB model to differentiate between sources.

Figure 3.2 shows the annual means of the remaining species, averaged across all sites, in decreasing order of magnitude. Also shown is the cumulative fraction of total mass. Organic carbon (OC) and Nitrate (NO₃) account for over 50% of the mass. Total carbon, that is, OC plus elemental carbon (EC), and NO₃ along with sulfate (SO₄) and ammonium (NH₄), account for almost 90% of the total. Thus, most Bay Area PM is either carbon or secondary PM.

Figure 3.2. Mean annual mass contribution of PM_{2.5} species



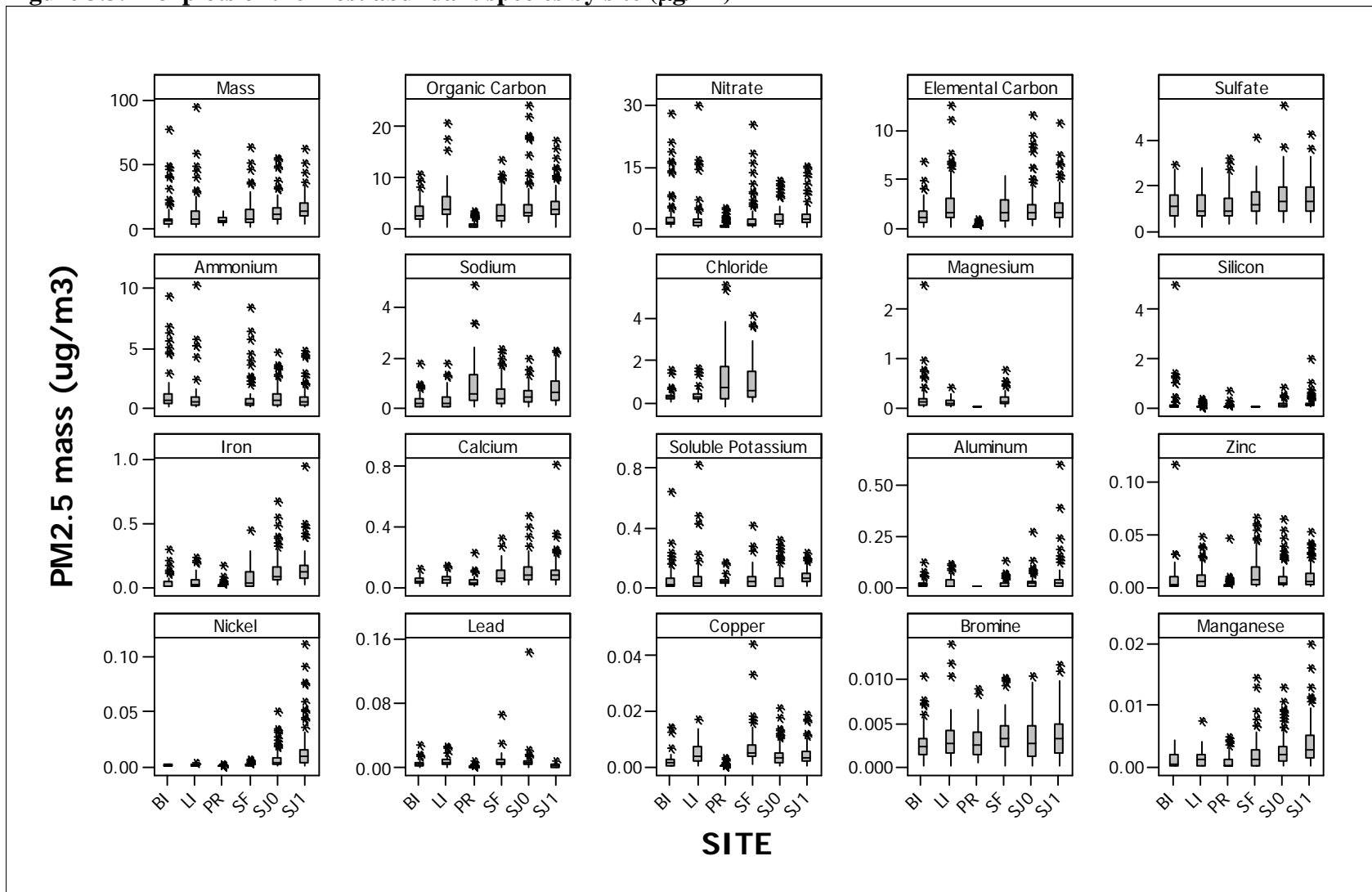
Note: Masses averaged across all sites. Bars indicate individual species corresponding to left-hand axis. The cumulative amount is shown by the line and corresponds to the right-hand axis. Appendix C provides a list of compounds corresponding to the abbreviations on the horizontal axis.

Sodium and chlorine are the next greatest contributors, both present in marine air. Sodium is greater than chlorine because some chlorine is replaced by nitrate as the marine air mixes with air containing NO_x. Elements associated with soil – notably silicon and aluminum – are present, but in small quantities, indicating that geological dust is not a large component of Bay Area PM_{2.5}.

Figure 3.3 shows boxplots of mass and the major species. Each box represents the range from the 25th to 75th percentile, with the horizontal line in the middle representing the median. The vertical lines above and below the box extend to the 95th and 5th percentiles. Asterisks beyond these lines represent extreme or outlying values of the distribution.

The figure shows large contributions of OC and to a lesser extent EC at every site except Point Reyes. Chloride and sodium are larger at Point Reyes than other Bay Area sites,

Figure 3.3. Boxplots of the most abundant species by site ($\mu\text{g}/\text{m}^3$)

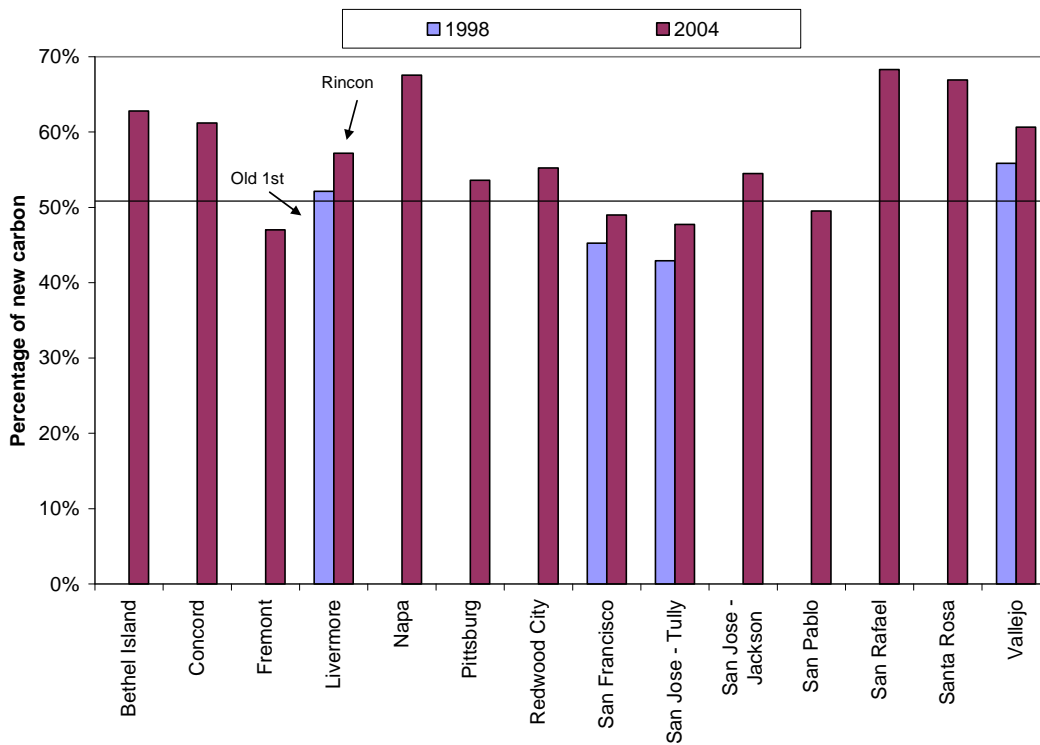


clearly a result of the site's exposure to the sea breeze. The San Francisco site, which is also exposed to air off the ocean and bay, also has elevated levels of these species. The patterns for nitrate and sulfate are different, with nitrate occasionally showing very large values, whereas sulfate has essentially no outliers. Also, Point Reyes nitrate values are much lower than for the other sites, but its sulfate values are comparable, perhaps because sulfate is a component of marine air. Aluminum and silicon are elevated on a few occasions, but even for these, the concentrations are not that high, indicating that occasionally windblown dust contributes somewhat to PM_{2.5} at these sites, but not in very large quantities.

3.4 Carbon-14 Analysis Results

Figure 3.4 shows the fraction of new carbon for each of the samples analyzed by the University of Arizona. For 2004, the percentages range from 47% for Fremont to 68% for San Rafael. From 1998 to 2004, the percents of new carbon increased approximately 5%. At 10 of the 14 sites, there was more new carbon than fossil carbon (Fairley, 2005). Standard errors of these percentages are all less than 0.5%.

Figure 3.4. Annual percentages of new carbon from Bay Area sites for 1998 and 2004



4. SOURCE APPORTIONMENT METHODOLOGY

Chemical Mass Balance (CMB) modeling is a standard methodology that uses chemical measurements of ambient PM on filters from air quality monitors for determining categories of PM sources. It is excellent at differentiating among sources that have chemical signatures that are clearly defined and measured.¹ However, CMB is limited when source signatures are too similar, nor can it attribute secondary PM such as ammonium nitrate to the original sources. Thus, the CMB analysis results were modified using other data sources.

First, with these measurements, CMB could not well differentiate among various carbonaceous sources (wood burning, diesel, gasoline, cooking, etc.) To augment the CMB analysis, measurements of Carbon-14 were made. This yielded the fraction of carbonaceous PM from “new” sources (wood, cooking, secondary organic) and “fossil” sources (diesel, gasoline, natural gas).

Second, to further identify sources and incorporate source contributions to secondary PM, the results of the CMB/Carbon-14 analysis were merged with information from the District’s emissions inventory.

4.1 CMB Modeling

CMB modeling provides a way of estimating the amount that various sources contribute to ambient PM concentrations. The CMB model is fit using a computer program whose inputs are source profiles and ambient PM samples that have been analyzed for a variety of chemical components. For each ambient sample, the CMB model finds the mix of sources whose combined amounts of chemical components best approximates those on the sample. In other words, the output of the CMB model includes estimates of the amounts (concentrations) from the various sources on each ambient filter.

The PM that deposits on an ambient filter sample comes from a wide variety of sources, only a few for which source profiles have been developed. Even those sources, like wood smoke, auto exhaust or geological dust, exhibit infinite variations in the relative amounts of various constituents. To some extent, this variation is accounted for in the model, which incorporates the variability recorded in the source profiles. Yet, these variations in the source profiles can affect the quality of the CMB results.

Species whose concentrations were below the limits of detection were not used for the CMB fit. Also, several species were measured both as ions and elements: potassium, chloride/chlorine, and sodium. Only one of the forms was used to avoid double counting. Similarly, sulfur was not fit because it duplicates sulfate. The table in Appendix C shows

¹ For example, approximately 40% of Bay Area geological dust is aluminum and silicon, whereas other potential sources – wood smoke, gasoline and diesel exhaust, cooking, tire and brake wear – have virtually none. Thus, the presence of aluminum and silicon indicates geological dust, and their absence indicates an absence of geological dust.

which species were used for fitting. Not all the species were measured at every site. For example, ammonium was not measured at Point Reyes and therefore could not be used in the CMB analysis for that site.

CMB version 7 was used in this analysis (Watson *et al.* 1990). In previous applications of the CMB model, the approach used was to find the "best" fit using a variety of criteria – low chi-squared value, high R^2 , all positive coefficients, all statistically significant coefficients, and lack of identifiability problems.² The weakness of this approach is that there may be more than one reasonable fit to a set of data.

The approach taken here was to find a weighted average of fits, weighting by the relative likelihood of the fit.³ This process took several iterations to improve the likelihood function so that it better matched the actual distribution.⁴

The method that was ultimately used was to define the likelihood as the product of two p-values:⁵ the p-value associated with the chi-square statistic for goodness of fit of individual chemical species, and the p-value for the difference between estimated and measured mass. The p-values were set to zero if any of the estimated source coefficients were negative. These likelihoods provided the weights applied to different fits. (Appendix D provides the details.)

When the CMB model was applied using all source profiles, the results showed large uncertainties for some of the source categories. This indicated that the data should be reanalyzed with fewer categories. Details of the analysis with all source profiles and the rationale for dropping specific source profiles are explained in Appendix E. The profiles included in the analysis were: ammonium sulfate, ammonium nitrate, marine, road-dust, fossil fuel, and wood smoke. Dropped from the analysis were the commercial cooking and tire/brake wear profiles.

² A program was written by District staff to automate the application of the CMB model. Specifically, CMBRUNS.EXE is a program that generates a file with keystrokes that operate the CMB model, allowing it to be run in batch mode and to try a variety of fits. In particular, fits were made for every combination of the following sources: 1) marine/mar0/mar100/none, 2) bytun4/camv/none, 3) byws4/WBOakEu/WBOak/none, 4) amm. nitrate/none, 5) amm. sulfate/none, 6) cook/none, 7) BYRDC/ARB_DUST/none, and 8) TireBreak/none. The "none" option permitted running the model without this source, which is desirable because CMB7 estimated contributions can be statistically insignificant or even negative. This creates a total of 2,304 fits for each ambient sample.

³ If the weights were statistical likelihoods, then this approach would have a Bayesian interpretation – attempting to approximate the mean of a posterior distribution for the model. This approach was tried but it had the weakness that the models themselves were uncertain. In particular, this approach led to situations where model A might produce estimated species contributions closer to the measured than for model B, yet model B would have a higher likelihood because its estimated uncertainties were less.

⁴ For example, initially, the difference between the measured and calculated mass was assumed to have a Gaussian distribution. But it was found that, for some samples, the two values differed dramatically – many standard deviations apart, i.e., the Gaussian provides a poor fit and leads to unrealistic results. The likelihood was modified to minimize the effect of the Gaussian term in these cases. See Appendix D for details.

⁵ p-values can be produced for statistical tests. A p-value represents the probability of seeing something at least as extreme as what was recorded given the null hypothesis were true. In this case, the null hypothesis is that the model is satisfactory. Small p-values indicate that the model is inadequate.

4.2 Incorporating Carbon-14 measurements

In theory, Carbon-14 measurements can be incorporated directly into CMB analysis, as simply another component in the array of chemical measurements made. But this can only be done if the C-14 measurements are made on the same air sample as the other measurements. For this study, we have C-14 measurements from PM₁₀ filters for 1998 and 2004. (The PM_{2.5} filters analyzed in the rest of the study were unavailable.) Also the C-14 measurements represent only an annual average. However, because C-14 measurements yield a clear-cut division between new and fossil sources, these measurements were used to modify the CMB results for the “vegetative burning” and “fossil” categories. (See Appendix F for a more detailed rationale for the adjustment.)

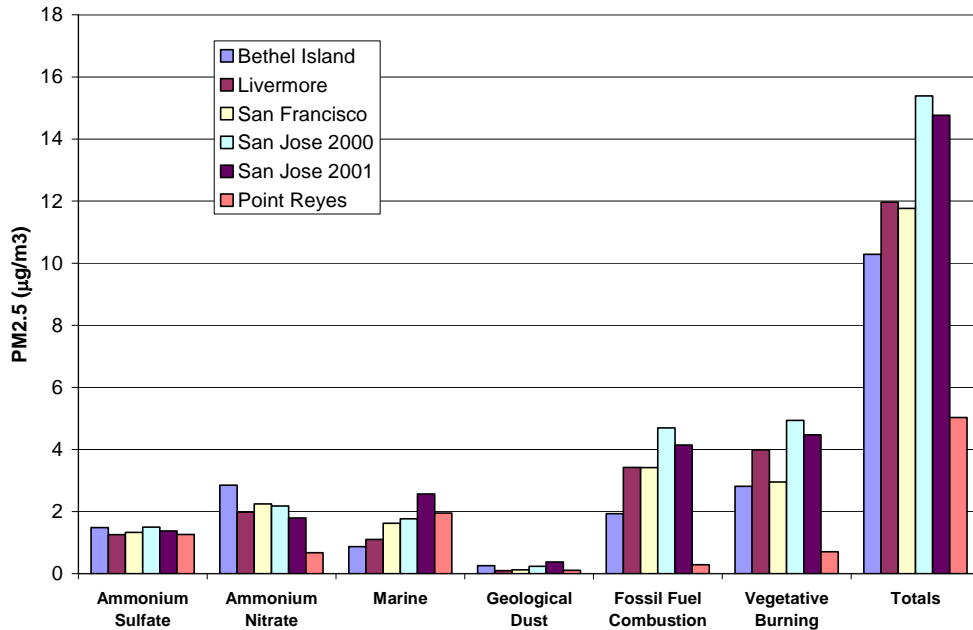
5. SOURCE APPORTIONMENT RESULTS

This section summarizes the results of the Carbon-14 adjusted CMB analysis and relates these results to the annual and 24-hour PM standards. The annual standard is based on the average of the 4 quarterly averages, so the quarterly averaged results are presented. The 24-hour standard relates to peak PM_{2.5} values, so the results for the samples with the 10 highest measured masses are presented.

5.1 Annual Summary

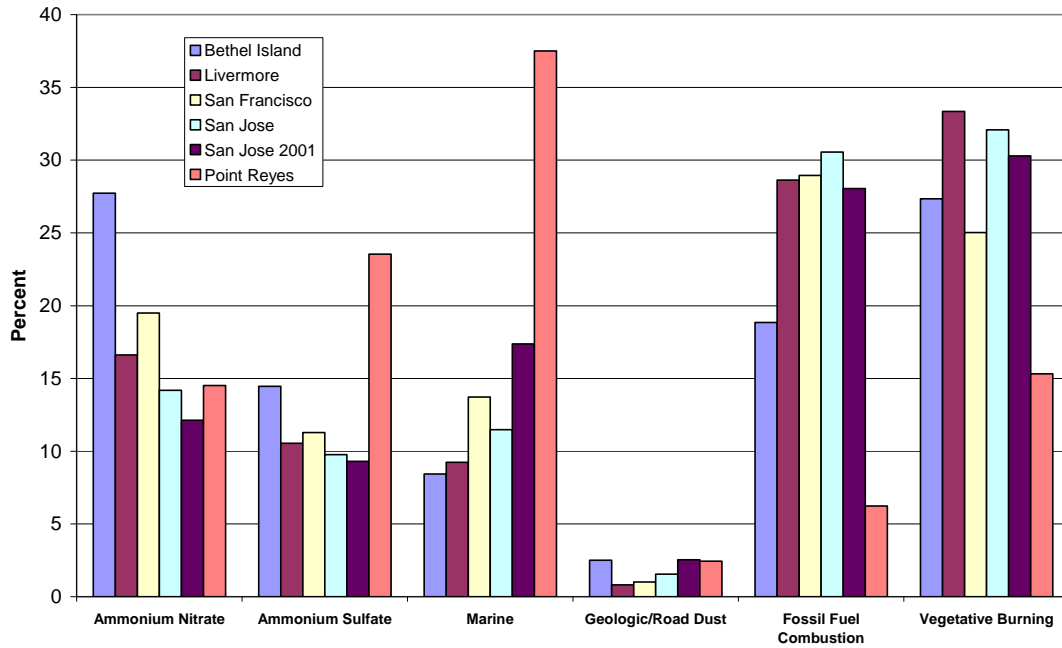
Figure 5.1 shows the estimated annual mass contributions from different source categories for the various sites. The major categories at the non-background sites are direct, combustion-related, largely carbonaceous sources – fossil fuel and wood burning; and secondary, combustion-related sources – ammonium nitrate and ammonium sulfate.

Figure 5.1. Estimated annual source contributions to Bay Area ambient PM_{2.5}



Note: Estimated annual source contributions to Bay Area ambient PM_{2.5} for 2000 and San Jose 2001. Values are quarterly averaged means of individual CMB results, except that the vegetative burning and fossil fuel totals have been adjusted based on Carbon-14 measurements. Totals are sums of individual source contributions.

Figure 5.2. Estimated annual percent contributions from various source categories



Note: Estimated annual percentage contributions to Bay Area ambient PM_{2.5} for 2000 and San Jose 2001. Values are percentages of total quarterly averaged PM_{2.5} from individual CMB results with fossil fuel combustion and vegetative burning adjusted for Carbon-14 results.

Table 5.1. CMB results for annual PM_{2.5} samples^a

Conc. (µg/m ³)	Ammonium Sulfate	Ammonium Nitrate	Marine	Geological Dust	Fossil Fuel	Vegetative Burning	Gun Powder	Est. Total
Bethel Island	1.5	2.9	0.9	0.3	1.9	2.8	0.1	10.3
Livermore	1.3	2.0	1.1	0.1	3.4	4.0	0.1	12.0
San Francisco	1.3	2.3	1.6	0.1	3.4	3.0	0.1	11.8
San Jose 2000	1.5	2.2	1.8	0.2	4.7	4.9	0.1	15.4
San Jose 2001	1.4	1.8	2.6	0.4	4.1	4.5	0.0	14.8
Point Reyes	1.3	0.7	2.0	0.1	0.3	0.7	0.0	5.0
Percentages^b								
Bethel Island	14.5	27.7	8.4	2.5	18.8	27.3	0.7	100
Livermore	10.5	16.6	9.2	0.8	28.6	33.3	0.8	100
San Francisco	11.3	19.1	13.8	1.0	29.1	25.1	0.5	100
San Jose 2000	9.8	14.2	11.5	1.5	30.6	32.1	0.4	100
San Jose 2001	9.3	12.1	17.4	2.5	28.1	30.3	0.2	100
Point Reyes	25.2	13.4	38.9	2.2	5.8	14.1	0.4	100
4 site average ^c	11.3	18.8	10.9	1.4	27.3	29.7	0.6	100

^a Average of quarterly averages.

^b Percentages of estimated mass.

^c BI, SF LI, and SJ 2000.

For the urban sites, Livermore, San Francisco and San Jose, the two carbonaceous source categories – fossil fuel and vegetative burning – constitute more than half the total. The contributions from these two categories is approximately the same. (see Figure 5.2 and

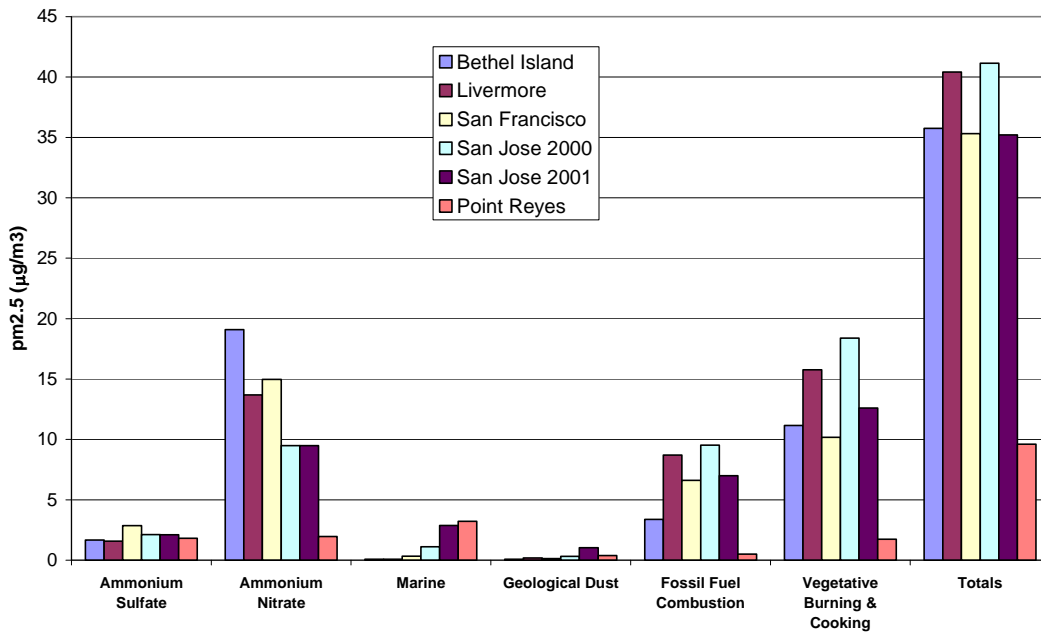
Table 5.1). For Bethel Island, secondary PM constitutes a similar percentage to its carbonaceous fraction, that is, ammonium nitrate plus ammonium sulfate constitutes the same percentage as vegetative burning plus fossil fuel combustion. For Point Reyes, the largest category is marine air.

Ammonium nitrate is a large source at every site. Ammonium sulfate and marine air are also important categories at all sites. The combined direct and secondary combustion⁶ sources constitute over 80% of the total for all 2000 sites except Point Reyes, and 79.8% for San Jose 2001. At no site does geological dust represent more than 4% of the total.

5.2 Peak Summary

Figure 5.3 shows the mass contributions for the 10 days at each site with the highest measured PM_{2.5} masses, and Figure 5.4 shows the percent contributions.

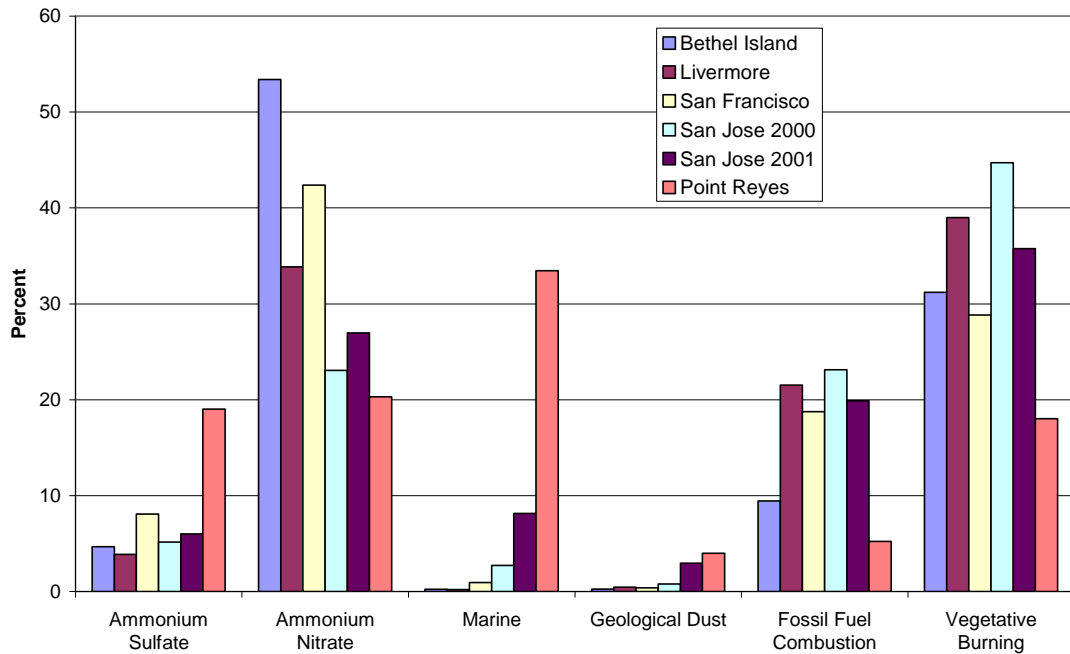
Figure 5.3. Estimated source contributions to peak Bay Area ambient PM_{2.5}



Note: Estimated source contributions to peak Bay Area ambient PM_{2.5} for 2000 and San Jose 2001. Values are averages of CMB and C-14 estimated source contributions from 10 days with highest PM at each site. Totals are sums of individual source contributions.

⁶ The District emissions inventory shows that most of the emissions of the precursors of sulfate and nitrate, namely SO₂ and NO_x, derive from combustion. Available evidence suggests that reducing these precursors results in reductions in ammonium sulfate and ammonium nitrate concentrations. Thus, it seems reasonable to consider these secondary sources as combustion-related.

Figure 5.4. Estimated annual percentage contributions to peak PM_{2.5}



Note: Estimated annual percentage contributions from various source categories to the 10 highest PM_{2.5} days, based on CMB analysis adjusted for Carbon-14. The values shown are the masses from individual source categories as a percentage of the total estimated mass.

Table 5.2. CMB results for peak PM_{2.5} samples^a

Conc. (µg/m ³)	Ammonium Sulfate	Ammonium Nitrate	Marine	Geological Dust	Fossil Fuel	Vegetative Burning	Gun Powder	Est. Total
Bethel Island	1.7	19.1	0.1	0.1	3.4	11.2	0.3	35.7
Livermore	1.6	13.7	0.1	0.2	8.7	15.8	0.4	40.4
San Francisco	2.9	15.0	0.3	0.1	6.6	10.2	0.2	35.3
San Jose 2000	2.1	9.5	1.1	0.3	9.5	18.4	0.2	41.1
San Jose 2001	2.1	9.5	2.9	1.0	7.0	12.6	0.1	35.2
Point Reyes	1.8	2.0	3.2	0.4	0.5	1.7	0.0	9.6
Percentages^b								
Bethel Island	4.7	53.4	0.2	0.3	9.5	31.2	0.8	100
Livermore	3.9	33.9	0.2	0.4	21.5	39.0	1.1	100
San Francisco	8.1	42.4	0.9	0.4	18.8	28.8	0.6	100
San Jose 2000	5.2	23.1	2.7	0.8	23.1	44.7	0.4	100
San Jose 2001	6.0	27.0	8.1	2.9	19.9	35.7	0.3	100
Point Reyes	18.9	20.2	33.4	4.0	5.2	17.9	0.4	100
<i>4 site average^c</i>	5.4	38.2	1.0	0.4	18.2	36.0	0.7	100.0

^a Average of results for 10 highest PM_{2.5} measurements at each site.

^b Percentages of estimated mass.

^c BI, SF LI, and SJ 2000.

The pattern of source contributions to peak PM_{2.5} is different from that of annual PM_{2.5}. In particular, ammonium nitrate and vegetative burning represent larger fractions.

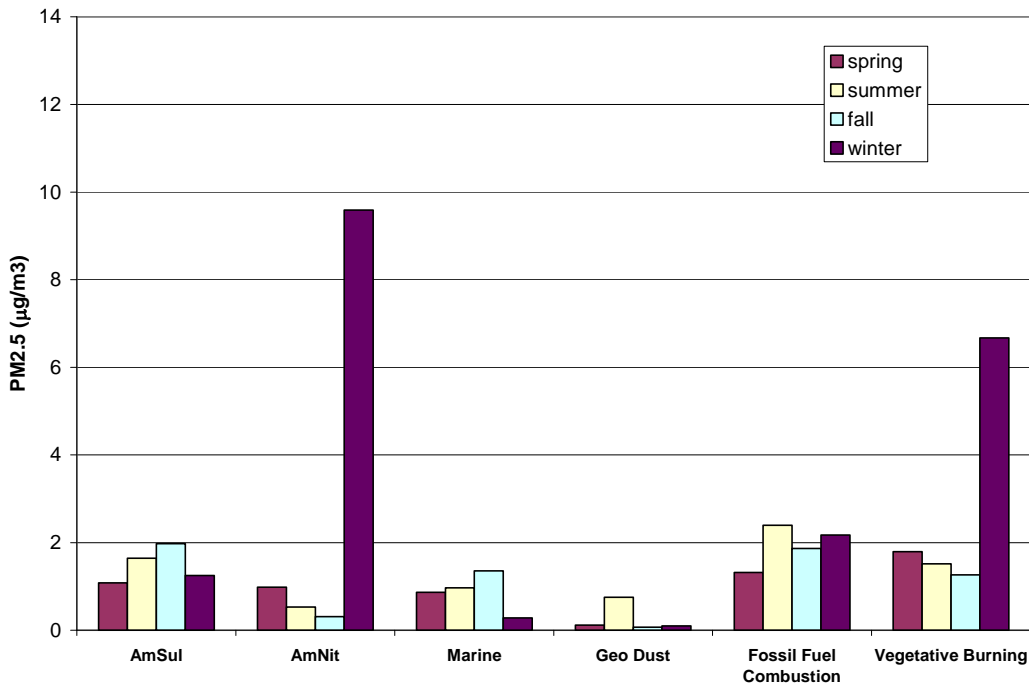
Ammonium nitrate represents over 50% for Bethel Island, over 40% for San Francisco, and over 30% for Livermore. Vegetative burning is also a greater factor, representing over 28% for all other sites except Point Reyes. The percent contribution of fossil fuel is smaller for peak PM_{2.5} than for annual PM_{2.5}, although is still a large contributor. Ammonium sulfate, marine and geological dust are also smaller contributors to peak PM_{2.5} than to the annual totals. Note that the total peak concentrations are between 35 - 40 µg/m³ at all sites except Point Reyes.

5.3 Analysis by Season

Figures 5.5a-5.5f show source contributions by season for each site. As expected, the largest contributions occur in the winter quarter. Ammonium nitrate and either fossil fuel combustion or vegetative burning, or both, provide the dominant contribution to total PM_{2.5} at every site, except Point Reyes.

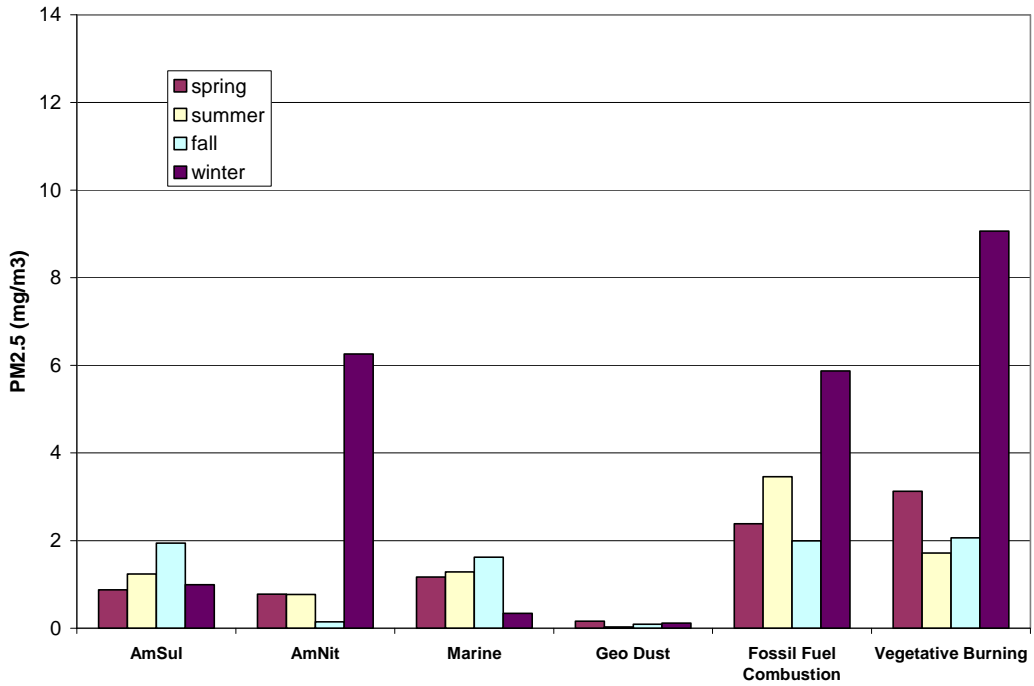
For other seasons, fossil fuel and vegetative burning are large contributors and ammonium sulfate forms as large or larger contributor than ammonium nitrate in almost every instance. The fossil fuel category is generally as large or larger than “vegetative burning,” but often the two are comparable. Note that the “vegetative” category includes cooking, secondary organics, planned burns, and accidental fires. This may account for the similarity.

Figure 5.5a. PM_{2.5} source contributions by season for Bethel Island



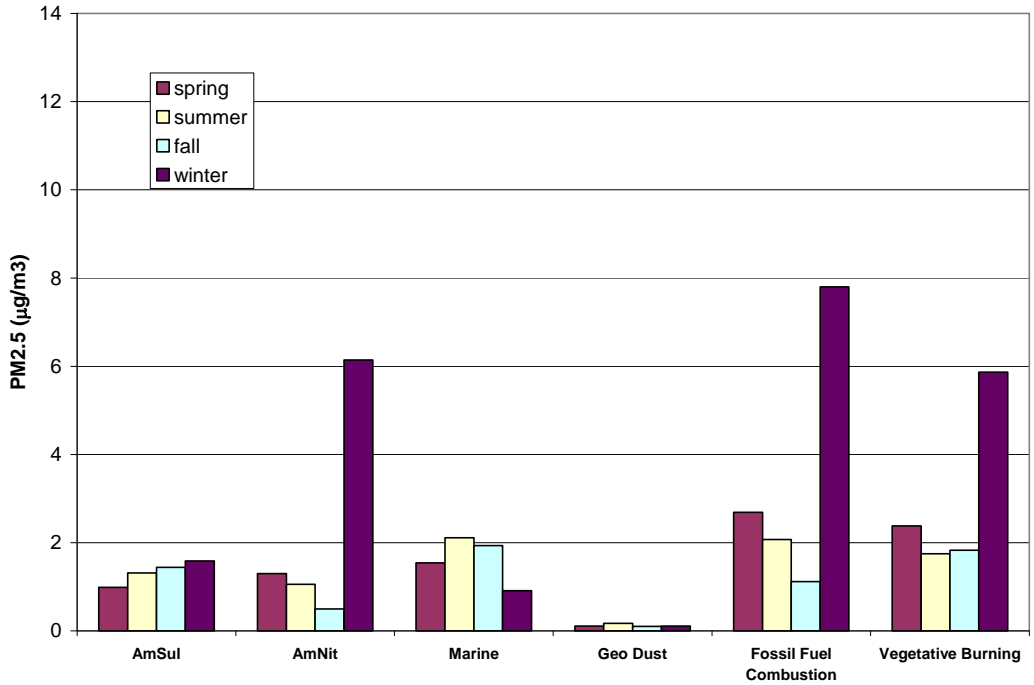
Note: Bethel Island PM_{2.5} source contributions by season. Values represent averages for the 4 pm seasons.

Figure 5.5b. PM_{2.5} source contributions by season for Livermore



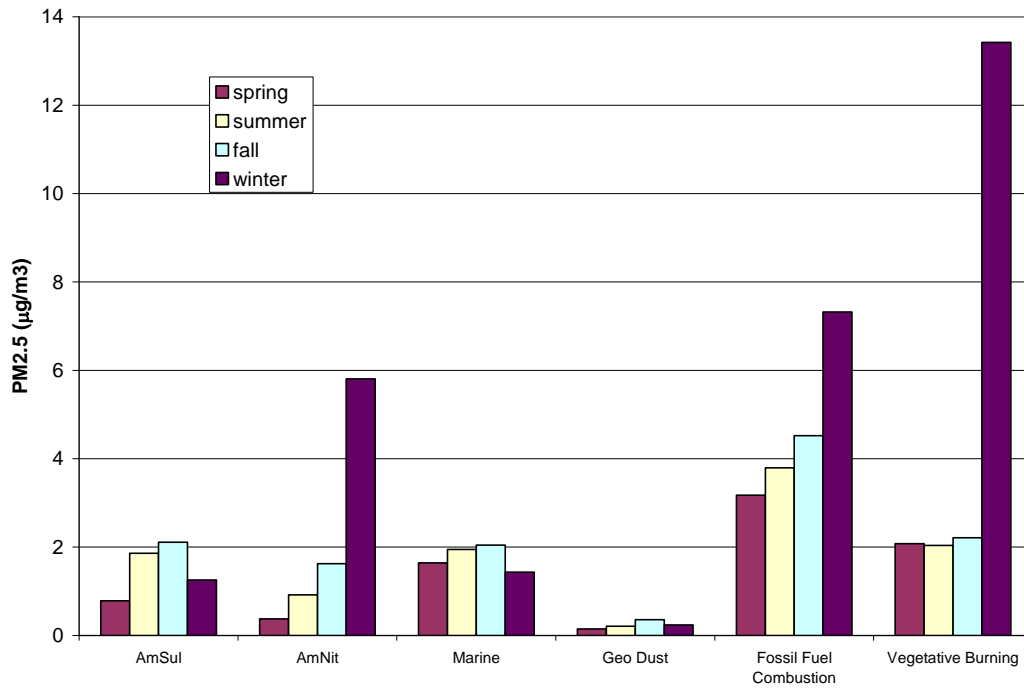
Note: Livermore PM_{2.5} source contributions by season. Values represent averages for the 4 pm seasons.

Figure 5.5c. PM_{2.5} source contributions by season for San Francisco



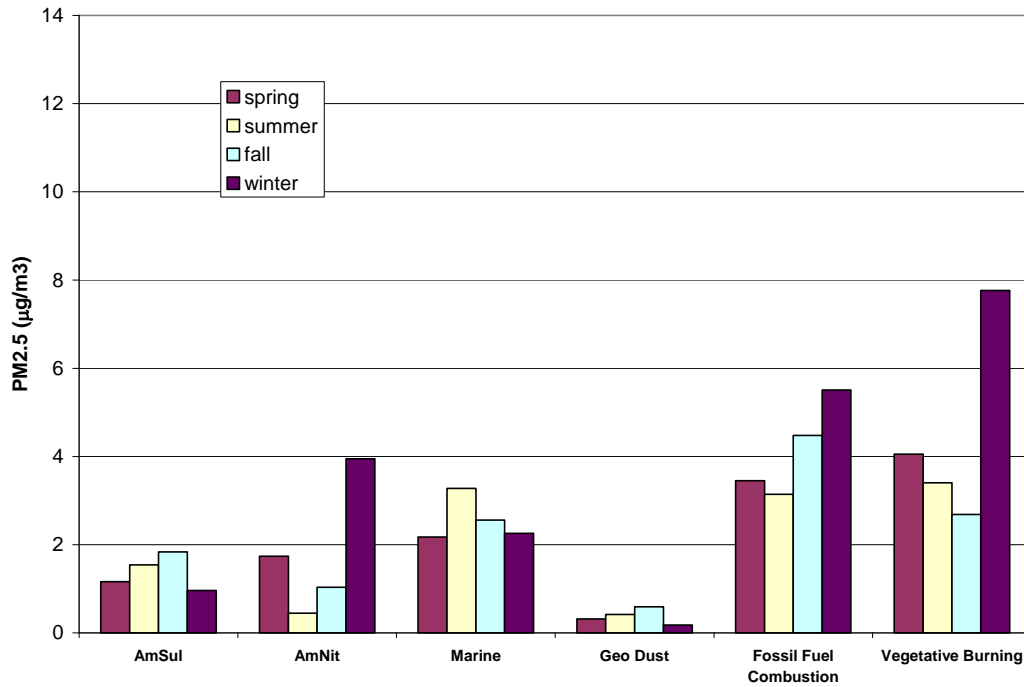
Note: San Francisco PM_{2.5} source contributions by season. Values represent averages for the 4 pm seasons.

Figure 5.5d. PM_{2.5} source contributions by season for San Jose 2000



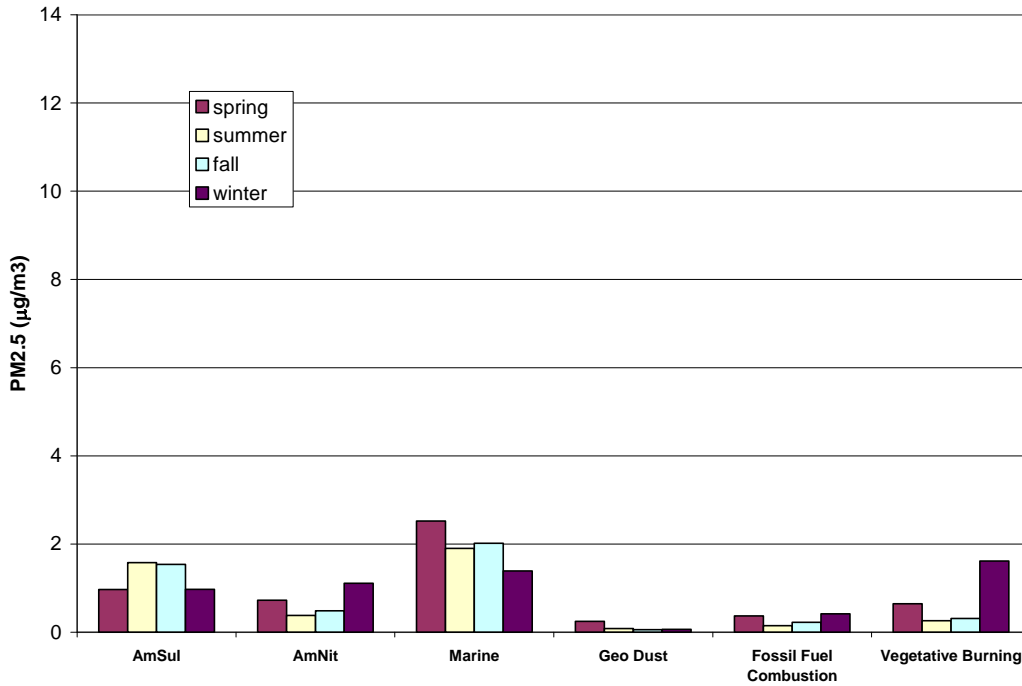
Note: San Jose 2000 PM_{2.5} source contributions by season. Values represent averages for the 4 pm seasons.

Figure 5.5e. PM_{2.5} source contributions by season for San Jose 2001



Note: San Jose 2001 PM_{2.5} source contributions by season. Values represent averages for the 4 pm seasons.

Figure 5.5f. PM_{2.5} source contributions by season for Point Reyes



Note: Point Reyes PM_{2.5} source contributions by season. Values represent averages for the 4 pm seasons.

5.4 Comparisons with the 1994 Winter PM₁₀ Field Study

As an external check on the findings of the current study, comparison was made to the results of a previous PM study. During the winter of 1993-94, PM₁₀ samples were collected from the four sites analyzed in the present study (BI, LI, SF, and SJ). Samples with PM₁₀ greater than 50 µg/m³ were aggregated and analyzed for C-14 for each site. Aggregated profiles for wood smoke, auto exhaust, and geological dust were also analyzed for C-14. A CMB analysis was then performed on the aggregated samples. Table 5.3 shows the results.

**Table 5.3. CMB results^a for aggregated wintertime high PM₁₀ samples, 11/93-1/94.
Source: Fairley (1995)**

Conc. (µg/m ³) (std. errors)	Marine	Ammonium Sulfate	Ammonium Nitrate	Geological Dust ^b	Fossil Fuel ^c	Wood/plant burning ^d	Est. Total	Meas. Total
Bethel Island	0.5 (0.4)	3.0 (0.2)	30.4 (0.6)	5.6 (0.7)	3.5 (0.9)	18.6 (2.3)	61.7 (2.4)	63.2 (1.1)
San Francisco	2.1 (0.4)	4.3 (0.2)	23.8 (0.5)	10.2 (1.2)	10.7 (1.8)	13.7 (2.0)	64.9 (2.4)	64.5 (1.0)
Livermore ^e	1.1 (0.4)	2.6 (0.3)	27.3 (0.5)	7.5 (0.7)	6.5 (2.2)	30.0 (2.7)	75.0 (2.8)	74.2 (1.0)
San Jose ^e	2.2 (0.5)	2.3 (0.3)	19.1 (0.6)	11.7 (1.2)	14.9 (2.0)	36.2 (3.5)	86.3 (3.7)	80.9 (1.1)
Percentages^f								
Bethel Island	0.9	4.9	49.3	9.1	5.7	30.1	100	102.4
San Francisco	3.2	6.7	36.7	15.8	16.5	21.2	100	99.4
Livermore ^e	1.4	3.5	36.4	10.0	8.7	40.0	100	98.9
San Jose ^e	2.6	2.6	22.1	13.5	17.3	41.9	100	93.7
<i>4 site average</i>	<i>2.1</i>	<i>4.5</i>	<i>34.9</i>	<i>12.2</i>	<i>12.4</i>	<i>34.2</i>	<i>100</i>	<i>98.2</i>

^a CMB estimates based on combined samples for days where PM₁₀ > 50 µg/m³, using data collected 11/93 - 1/94. This analysis includes old and new carbon as species based on the University of Arizona C-14 analysis.

^b Geological dust profile collected from various sites around the Bay Area, on and off road, mainly near the 4 sites.

^c Fossil Fuel profile collected in Caldecott Tunnel, adjusted to eliminate other sources. PM₁₀ attributed to this profile is mainly auto exhaust, but the CMB model probably attributes PM₁₀ from other fossil fuel burning (such as natural gas burning) to this source also.

^d Wood/plant burning profile collected in three San Jose back yards of homes burning wood in their fireplaces. PM₁₀ attributed to this profile is mainly from woodburning, but there may be some from agricultural burning and other sources also.

^e The 2 samples for San Jose and Livermore are averaged.

^f Percentages of estimated mass.

Figure 5.6 compares the percent of peak PM₁₀ attributed to different source categories for the winter of 1993-94 with the percent peak PM_{2.5} for different source categories for the present study.

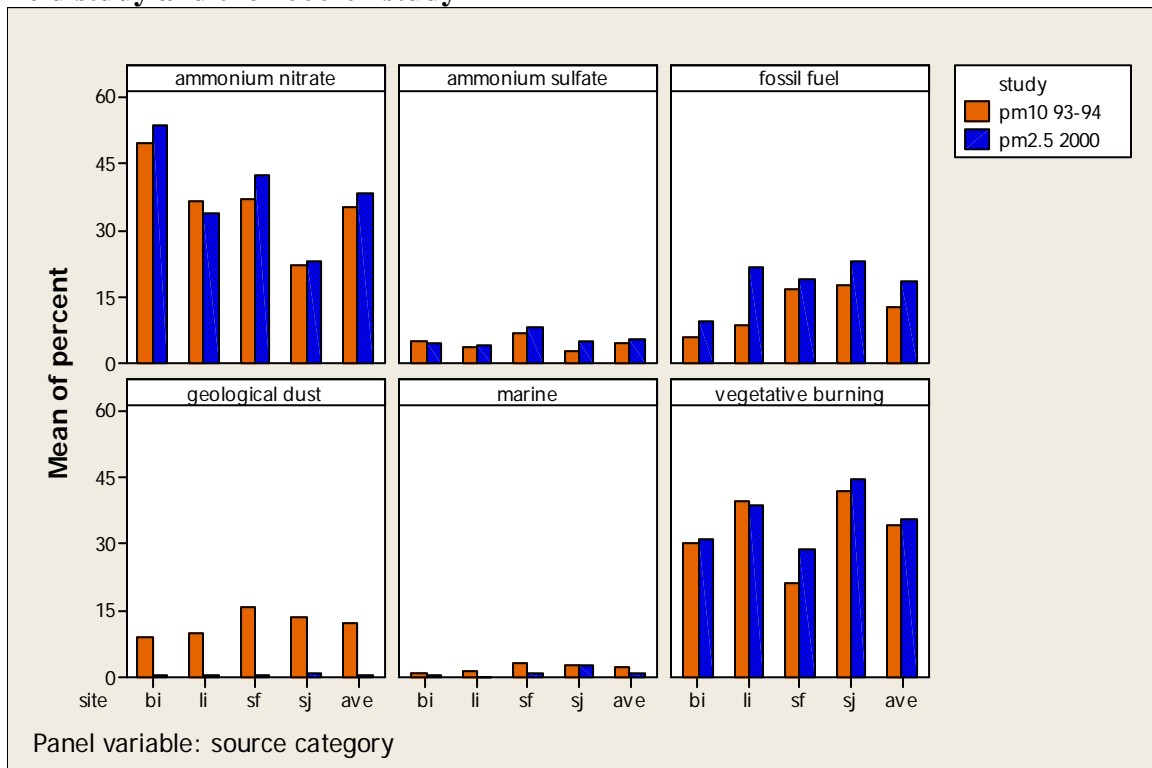
Certain differences are expected because the comparison is between PM₁₀ in the earlier study and PM_{2.5} in the present study. One key difference is that most airborne geological dust is coarse – greater than 2.5 microns in diameter, so that a much higher percentage would be expected in PM₁₀ than PM_{2.5}. Indeed, that appears to be the case: geological dust represents 12.2% of PM₁₀ on average in the earlier study, and only 0.5% of PM_{2.5} in the present study. Marine air also contains a large coarse fraction. The earlier study has somewhat higher marine percentages overall (2.1% vs. 1.0%) and for every site except SJ.

For ammonium nitrate and ammonium sulfate, the results seem consistent between studies. The overall percentages are somewhat higher in the present study; this is expected, because most of this secondary PM is PM_{2.5}. The relationship of the sites is reasonably consistent, with BI registering the highest ammonium nitrate in both studies, SJ the lowest, and SF the highest ammonium sulfate. Note that there is an increase in the

percent contribution of ammonium sulfate, also secondary PM, at all sites, particularly at San Jose.

Figure 5.6 also shows larger contributions from fossil fuel and vegetative combustion in the PM2.5 study. The overall percent from fossil fuel has increased at every site, with an average increase from 12% to 18%. The greatest increase is at Livermore, which experienced significant traffic increases between the two study periods. The vegetative burning percentages increased at every site except for Livermore (possibly influenced by the relocation of the monitor and increase in traffic between the two study periods). The relative order has remained the same, however, with SF having the lowest percentages to San Jose having the most.

Figure 5.6. Comparison with percent source attributions from the winter 1993-94 field study and the 2000-01 study



6. REFINING SOURCE APPORTIONMENT WITH EMISSIONS INVENTORY ESTIMATES

CMB analysis with currently available data can only provide accurate distinctions between source categories, not specific sources. In an attempt to refine the source apportionment, emissions inventory estimates were combined with the CMB results. Specifically, the BAAQMD emissions inventory provides estimates of emissions within source categories. For the "vegetative burning" category, the emissions inventory includes domestic burning (wood burning), commercial cooking, accidental fires, tobacco smoke, and waste burning. The major emissions inventory categories within the "fossil fuel" category are on-road vehicles, off-road vehicles, aircraft, refineries, and power

generation.⁷ The emissions inventory is the product of years of effort by the District along with ARB and the U.S. EPA. Emissions estimates are based on carefully analyzed methods, and have been tested against emissions measurements and ambient concentration data.

CMB analysis and the emissions inventory provide complementary information about PM sources. CMB analysis uses measured PM concentrations, so that it reflects what is actually in the air. But the measurements are specific to a particular location and can only distinguish among a limited number of sources. The emissions inventory is based on engineering calculations for a wide range of sources, and represents totals across the entire Bay Area, but it doesn't reflect transport, deposition, or chemical transformations in the atmosphere.

In this section, CMB is used to provide estimates of the relative contribution of major primary and secondary source categories. The emissions inventory is used to apportion the PM in these major categories to individual sources.

The emissions inventory provides estimated annual average emissions, appropriate for apportioning annual average PM_{2.5} sources. Peak PM_{2.5}, however, occurs almost exclusively in winter, making the winter emissions inventory more appropriate for this analysis.

Table 6.1 shows estimated relative contributions of various sources to annual emissions based on estimated annual 2000 emissions from the base year 2002 BAAQMD emissions inventory. The table presents four components of the emissions inventory: PM_{2.5} from combustion, PM_{2.5} from geological sources (such as road dust), NO_x, and SO₂. The latter two are precursors to ammonium nitrate and ammonium sulfate, respectively. The "vegetative burning" category is broken into cooking and "vegetative fires", which includes residential wood burning, wildfires, planned burns and cigarette smoke. Information in the table was used to apportion the CMB source categories into more specific sources.

⁷ In some of these, various subcategories have been combined. For example, "refineries" includes "basic refining processes," "flares," "other refining processes," and "oil refineries external combustion."

Table 6.1. Emissions inventory estimates of annual carbonaceous and secondary precursors (% of emission inventory total)

	PM2.5- combustion	NOx	SO ₂	PM2.5- geological
Vegetative Fires	44	1	1	0
Commercial Cooking	11	0	0	0
Fossil Fuel Combustors:				
On-road	10	57	3	72
Off-road	14	20	0	9
Trains	1	2	1	0
Aircraft	2	3	1	0
Ships	4	3	10	0
Refining	6	4	65	0
Power Generation	2	2	1	0
Domestic fossil heat	2	3	0	0
Other fossil	4	4	4	0
Fossil Fuel Total:	45	99	88	82

Table 6.1 shows that the emissions inventory attributes 44% of combustion PM_{2.5} to vegetative fires, 11% to commercial cooking, and 45% to fossil fuel burning. The table also lists specific fossil fuel sources including on-road, off-road, trains, aircraft, ships, refining, and power generation. Particles from these sources are likely to be attributed to the "fossil" CMB source category.

Note that the percentage of directly emitted PM_{2.5} from combustion estimated from vegetative fires alone is 44%, nearly as large as all fossil fuel emissions. Combined with cooking, the total is 55%.

Table 6.2 combines the information in Table 6.1 with the 4-site totals of Table 5.1. The percentages in the Table 5.1 are assumed to provide correct totals for each source category⁸, and Table 6.1 is used to apportion PM to specific sources from within each source category. For example, Table 5.1 shows that CMB attributes 18.8% of annual PM_{2.5} to ammonium nitrate. This ammonium nitrate is assumed to derive from NOx. Table 6.1 shows that 57% of the Bay Area's annual NOx emissions derive from on-road vehicles. Therefore, we assume that 57% of the 18.8% or 11% of the Bay Area's annual PM_{2.5} derives from ammonium nitrate originating from on-road motor vehicles.

The numbers in Table 6.2 represent the percent of the total PM attributable to specific sources within each category. The numbers in the right-hand column are the percent of total PM_{2.5} attributable to that specific source so that, combining both direct emissions and secondary contributions, vegetative fires are estimated to contribute 24% of the total annual PM_{2.5}, commercial cooking 6%, on-road 18% and so on. Figure 6.1 summarizes the results.

⁸ In other words, the column totals in Table 6.2 are equal to the 4-site percentages at the bottom of Table 5.1.

Table 6.2. Percent estimated total annual contribution from various sources.

Source	Source Category:						Totals*
	Marine	Ammonium Sulfate	Ammonium Nitrate	Geological	Fossil	Vegetative	
Vegetative fires		0	0	0		24	24
Commercial cooking		0	0	0		6	6
On-road		0	11	1	6		18
Off-road		0	4	0	9		12
Trains		0	0	0	0		1
Aircraft		0	1	0	1		2
Ships		1	1	0	2		4
Refining		7	1	0	3		12
Power generation		0	0	0	2		2
Domestic fossil		0	0	0	1		2
Marine	11						11
Other		2	1	0	2		5
Totals*	11	11	19	1	27	30	99

* Totals may differ due to rounding.

Figure 6.1. Estimated contributions to annual Bay Area PM_{2.5} from various sources

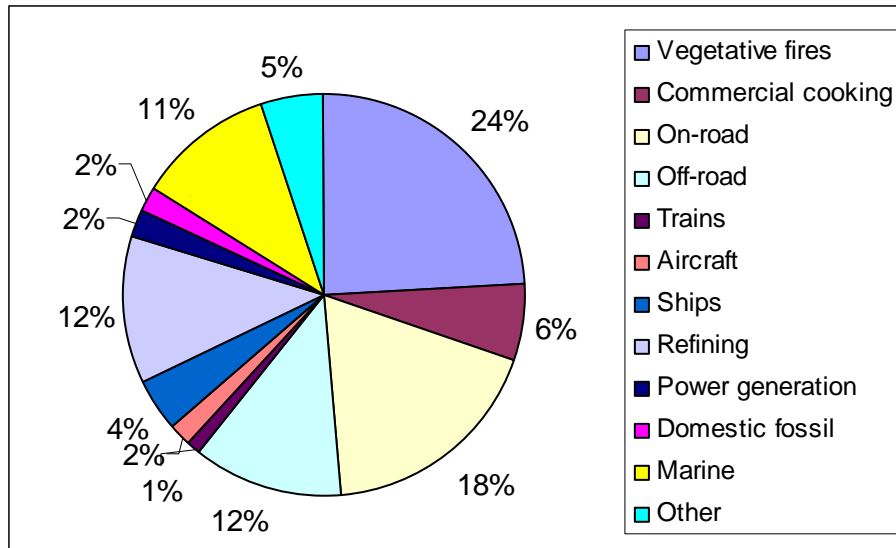


Table 6.3 shows winter emissions inventory percentages for carbonaceous and secondary precursors. The inventory attributes 38% of combustion PM_{2.5} to fossil fuels and 62% [56%+6% = 62%] to vegetative fires and cooking. This is close to the percentages estimated based on the carbon-14 adjusted CMB analysis shown in Table 5.2, where fossil fuels were estimated to contribute 18.2% and vegetative burning 36.0%. Thus, the CMB analysis attributes 34% [100%*18.2/(18.2+36.0) = 34%] of direct combustion-related PM_{2.5} to fossil fuels, and 66% [(100-34)% = 66%] to vegetative burning.

Table 6.3. Emissions inventory estimates of winter carbonaceous and secondary precursors (% of emission inventory total)

	PM2.5- combustion	NOx	SO ₂	PM2.5- geological
Vegetative fires	56	1	1	0
Commercial cooking	6	0	0	0
Fossil Fuel Combustors:				
On-road	9	60	3	72
Off-road	11	17	0	5
Trains	0	2	1	
Aircraft	2	3	1	
Ships	3	3	10	
Refining	4	4	65	
Power generation	2	2	1	
Domestic fossil heat	3	4	1	0
Other Fossil	3	4	4	0
Fossil Fuel Total:	38	99	88	77

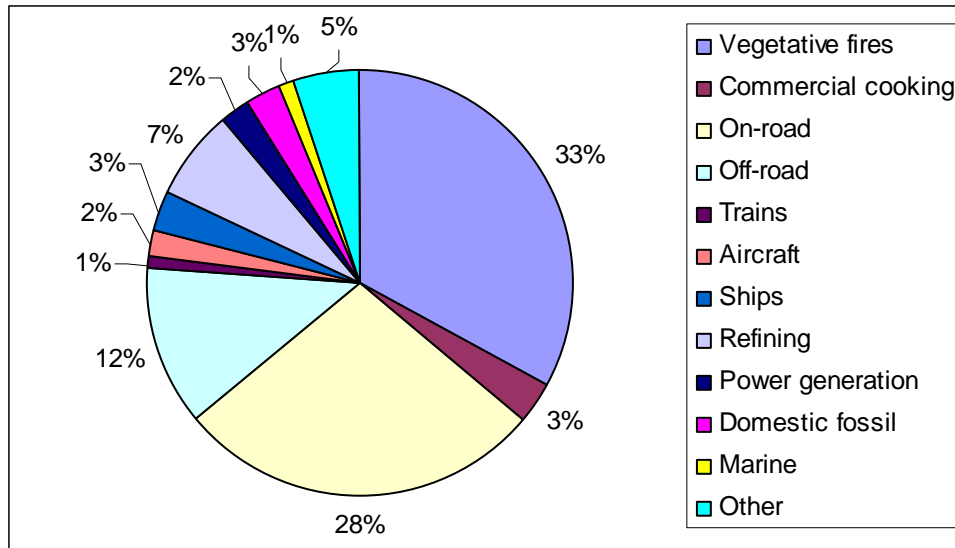
As with Table 6.2, Table 6.4 contains estimates of contributions from individual sources to total PM_{2.5}, but for winter days with high PM_{2.5}. As expected, vegetative fires (almost entirely domestic wood burning) constitutes a larger percentage than in Table 6.2, and is the largest single source. Interestingly, on-road vehicles also contribute more of the total, due to the large amount of ammonium nitrate in wintertime PM_{2.5}. These two sources contribute about half of winter peak PM_{2.5}. Other large sources include off-road vehicles, ships, and commercial cooking. Figure 6.2 summarizes the results.

Table 6.4. Estimated percentage total peak contribution from various sources.

Source	Source Category:						Totals*
	Marine	Ammonium Sulfate	Ammonium Nitrate	Geological	Fossil	Vegetative	
Vegetative fires		0	0	0	0	33	33
Commercial cooking		0	0	0	0	3	3
On-road		0	23	0	4	0	28
Off-road		0	7	0	5	0	12
Trains		0	1	0	0	0	1
Aircraft		0	1	0	1	0	2
Ships		1	1	0	1	0	3
Refining		4	2	0	2	0	7
Power generation		0	1	0	1	0	2
Domestic fossil		0	1	0	1	0	3
Marine	1					0	1
Other		1	2		2	0	5
Totals*	1	5	38	0	19	36	99

* Totals may differ due to rounding.

Figure 6.2. Estimated contributions to peak Bay Area PM_{2.5} from various sources



7. KEY FINDINGS AND ON-GOING WORK

The currently available data, coupled with the above analysis, makes it possible to draw some conclusions about sources of fine PM in the Bay Area.

7.1 Key Findings

- Most anthropogenic PM₁₀ and PM_{2.5} derive from burning wood or fossil fuels.
- Geological dust is a small contributor to PM₁₀ and a negligible contributor to PM_{2.5}. Tire/break wear is also a negligible PM_{2.5} source.
- Peak PM occurs largely in winter. Reasons include a high rate of ammonium nitrate formation and more wood burning, both related to low temperatures. Analysis of wintertime meteorological data shows that periods of low winds are conducive to the buildup of PM_{2.5}.
- Residential wood burning appears to be the largest single source of wintertime PM_{2.5}, a conclusion supported both by ambient data and the emissions inventory.
- Ammonium nitrate is a major contributor to peak PM and a large contributor to annual PM.
- Ammonium sulfate is a substantial contributor to annual PM_{2.5}, but only a small contributor to peak PM.
- Wood burning and on- and off-road vehicles account for approximately 55% of annual PM_{2.5} and 75% of peak PM_{2.5}.

7.2 Ongoing work

The District has a number of projects underway to better understand the sources of PM_{2.5} in the Bay Area and to track progress in reducing PM_{2.5}.

- Performing annual telephone surveys to update and improve the estimates of emissions from wood burning. Future surveys can be used to help track progress in emissions reduction from wood burning.
- Implementing a grid-based particulate model. The model will provide information on the diffusion and dispersion of primary PM, formation of secondary PM and transport of PM from areas outside of the District.
- Developing an ammonia emissions inventory for the Bay Area that will be used for the PM model. The quality of these Bay Area-specific estimates of ammonia

emissions is expected to be better than those obtained from the use of the EPA or ARB methods.

- Performing a study of diesel emissions and concentrations through the District's Community Air Risk Evaluation (CARE) program will help quantify the relative contributions to PM_{2.5} from diesel and other fossil fuel sources.
- Analyzing winter and summer PM₁₀ filters for carbon-14. These measurements will add to our understanding of the relative PM contributions of wood smoke and cooking.
- Working with UC Davis to investigate the relationship between meteorology and PM transport, especially on wintertime high PM days.

8. REFERENCES

Chow, J.C., Watson, J.G., Crow, D., Lowenthal, D.H, Merrifield, T., 2001. Comparison of IMPROVE and NIOSH carbon measurements, *Aerosol Science and Technology*, 34: 23-34.

EPA/NARAL, 2002. Comparison of CARB and EPA speciated measurements: Performance Evaluation Sample audit of the CARB lab conducted by EPA/NAREL. Memo (9/12/02) from Michael S. Clark/NARAL to Jim Homolya / OAQPS. Reference found: <http://www.epa.gov/ttn/amtic/pmspec.html>.

EPA, 1999. Particulate Matter (PM_{2.5}) Speciation Guidance, Edition 1, Jim Homolya and Joann Rice, US EPA, October 7,1999. Final Draft.

Fairley, D., 1995. Chemical Mass Balance of Bay Area Particulates, draft Technical Report TR9501, April 1995.

Fairley, D., 2004. A look at C-14 comparisons, interoffice memo to Gary Kendall, 8/24/2004.

Fairley, D., 2005. Results from annual carbon-14 sampling, interoffice memo to Gary Kendall, et al., October 6, 2005.

Held, T., Ying, Q., Kaduwela, A., Kleeman, M.J., 2004. Modeling particulate matter in the San Joaquin Valley with a source-oriented externally mixed three-dimensional photochemical grid model. *Atmospheric Environment* (22): 3689-3711.

Watson, J.G. *et al.*, 1990. Receptor model technical series, volume III (1989 revision), CMB7 user's manual, EPA-450/40-90-004, January 1990.

Watson, J.G. 2008. Personal communication.

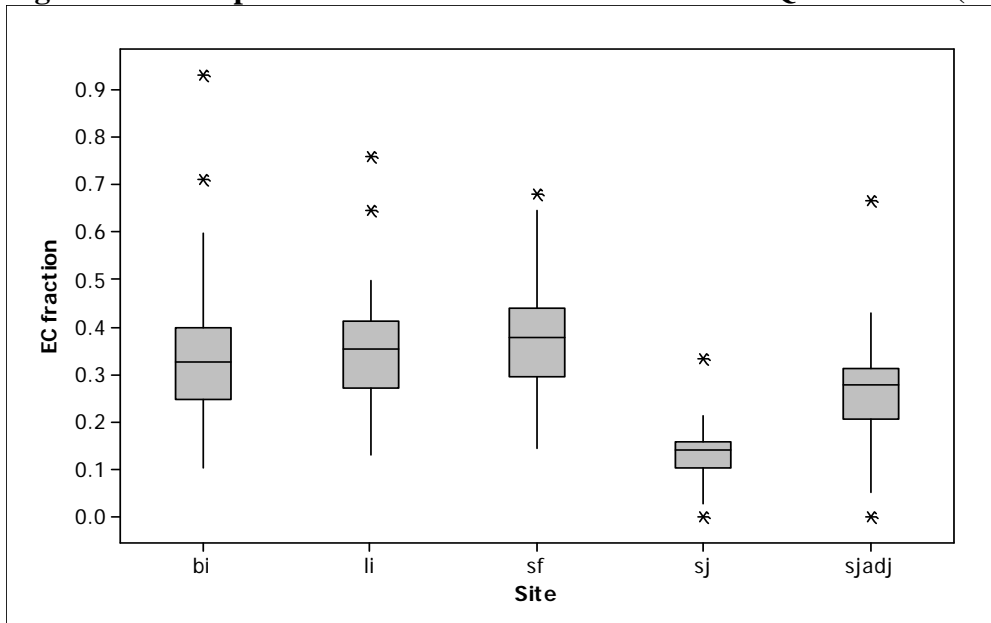
APPENDIX A. ADJUSTMENT OF OC AND EC DATA

Two different approaches have been developed to measure OC and EC on PM filters – the IMPROVE and NIOSH methods. Chow *et al.* (2001) showed that the NIOSH method attributes considerably more carbon to the OC fraction than the IMPROVE method does. CRPAQS and the Speciated Trends Network used the IMPROVE method for ambient samples, and CRPAQS and BAAQMD used the IMPROVE method for their source profiles. But the EPA has used the NIOSH method for its analysis of San Jose PM_{2.5} filters. CMB analysis requires consistent source and ambient measurements, so one set of measurements had to be adjusted. The choice was made to modify the NIOSH measurements, simply because there were fewer datasets to change.

Comparisons showed that the fraction of IMPROVE-measured EC out of total OC+EC averaged at least double that of the NIOSH method. Figure A1 shows boxplots of the ratios of EC to TC = OC+EC. As can be seen, the NIOSH method EC fraction is considerably less than that of the IMPROVE sites. The median and mean ratios for the IMPROVE sites are all somewhat more than double the NIOSH site (Table A1). Also shown are the EC/TC ratios from the District's Wintertime 1993-94 PM₁₀ Study, where all EC/OC measurements were done using the IMPROVE method. Although the ratios appear somewhat lower than they do at present, the ratios for SJ are certainly not lower than for the other sites. Thus, there is no reason to suspect that the EC fractions at San Jose are, in reality, systematically lower than for any of the other sites.

Because the PM_{2.5} ratios for the IMPROVE sites in Table A1 are at least twice what they are for SJ, we adjusted the SJ observations by doubling the EC numbers and subtracting off the corresponding amount from the OC, i.e., $EC^* = 2EC$ and $OC^* = OC - EC$. The boxplot of the ratios of the adjusted EC/OC values is shown in Figure A1.

Figure A1. Comparisons of EC/TC ratios for the CRPAQS and EPA (SJ) sites.



Note: SAdj represents the ratios for San Jose after adjustment.

Table A1. Mean and median EC/TC ratios

Site	n	mean ratio	median ratio	pm10 ratio*
SF	71	0.38	0.38	.28
LI	70	0.34	0.35	.24
BI	68	0.34	0.33	.22
SJ	188	0.13	0.14	.27

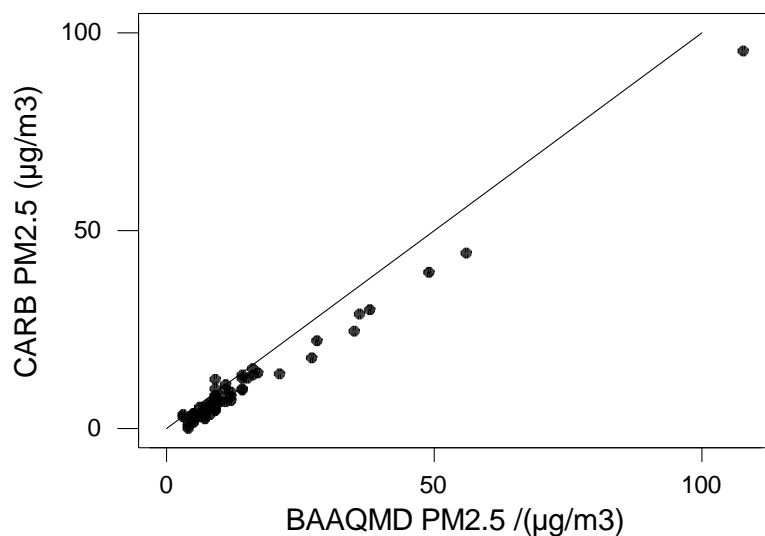
* From the Winter 1993-94 field study

APPENDIX B. COMPARISON OF MASS MEASUREMENTS

Various labs measured the data analyzed in this report. Parallel measurements were made at some sites. In particular, simultaneous PM_{2.5} mass measurements were made at District sites and two of the three CRPAQS sites, Livermore and San Francisco, and at the Speciation Trends Network data from the San Jose site.

We found high correlation between Livermore and San Francisco, and between Livermore and San Jose. Figure B1 shows a plot of Livermore PM_{2.5} mass measured by CARB versus mass measured by the District.

Figure B1. Comparison of PM_{2.5} mass measured by CARB and BAAQMD for Livermore, 2000



Note: 1-to-1 line also shown.

There are discrepancies, however, in the magnitudes of the measured mass for Livermore and San Francisco. Table B1 shows comparative statistics between BAAQMD measurements and those of CARB. Included are mass measurements and also sums of the major individual species, including nitrate, sulfate, EC, OC, aluminum, silicon, sodium and chlorine, more precisely, all of the compounds used in CMB fitting listed in Appendix C.

Table B1. Same-day comparisons of BAAQMD, CARB and EPA PM_{2.5} measurements

Site	N	Mean			Median			Ratios: Other/BAAQMD			Ratios: Sum of Spec./BAAQMD		
		BAA QMD mass	Other Mass ^a	Sum of Spec. ^b	BAA QMD mass	Other mass ^a	Sum of Spec. ^b	Med-ian	Low CI	Upp CI	Med-ian	Low CI	Upp CI
LI	64	13.6	10.3	11.3	9.0	6.2	9.1	0.70	0.65	0.76	0.93	0.84	1.04
SF	64	14.3	11.5	12.0	9.0	7.2	8.6	0.73	0.67	0.79	0.87	0.81	0.94
SJ	126	14.3	15.4	14.0	11.0	11.8	11.7	1.06	1.03	1.11	1.06	1.03	1.09

^a Other = CARB for LI and SF, and EPA (RTI International) for SJ.

^b Sum of Spec. = Sum of all major components, including nitrate, sulfate, EC, OC, aluminum, silicon, sodium and chlorine.

The table shows that CARB PM_{2.5} averages roughly 70% of BAAQMD. The ratio is somewhat higher for higher concentrations. For San Jose, the comparisons are much closer, with the EPA measurements averaging slightly higher than the District's. The sum of species should be less than the mass, because it omits components like oxygen in soil and hydrogen in OC. Yet for LI and SF, the mean and median sum of species is larger than the CARB total mass, and the medians are actually close to the BAAQMD total mass. Also note that the ratios of the sum of species to BAAQMD mass are closer to 1 than the total mass. Therefore, it was assumed that the CARB masses were underestimated, and the LI, SF and BI measurements were adjusted by $1.1x + 1.9$, where x is the CARB measurement.

APPENDIX C. CHEMICAL SPECIES USED IN THE CMB MODEL

Table C1 shows the number of times the measured concentrations of various species exceeded 1 and 2 standard deviations. Also shown, by asterisks, are those species used for fitting by the CMB model.

Table C1. Chemical species – frequency above the limits of detection.

species id	species	#>1sd	#>2sd	used for CMB*	Species id	species	#>1sd	#>2sd	used for CMB*
MS	mass	194	164		CU	copper	138	71	*
CL	chloride	139	67		ZN	zinc	134	108	*
NO3	nitrate	208	198	*	GA	gallium	1	0	
SO4	sulfate	208	202	*	AS	arsenic	0	0	
NH4	ammonium	187	170	*	SE	selenium	15	3	*
NA	sodium ion	158	140	*	BR	bromine	124	42	*
KP	potassium ion	107	73		RB	rubidium	1	0	
OC	organic carbon	178	139	*	SR	strontium	16	7	*
EC	elemental carbon	189	143	*	YT	yttrium	0	0	
NAX	sodium	79	43		ZR	zirconium	3	0	
MG	magnesium	104	33		MO	molybdenum	1	0	
AL	aluminum	34	6	*	PD	palladium	0	0	
SI	silicon	24	13	*	AG	silver	1	0	
PH	phosphorus	16	1		CD	cadmium	0	0	
SU	sulfur	208	208		IN	indium	0	0	
CL	chlorine	100	78	*	SN	tin	2	0	
KPX	potassium	162	82	*	SB	antimony	1	0	
CA	calcium	52	13	*	BA	barium	1	0	
TI	titanium	0	0		LA	lanthanum	0	0	
VA	vanadium	0	0		AU	gold	0	0	
CR	chromium	0	0		HG	mercury	0	0	
MN	manganese	13	3	*	TL	thallium	0	0	
FE	iron	47	24	*	PB	lead	55	13	*
CO	cobalt	0	0		UR	uranium	1	0	
NI	nickel	20	4	*					

* Asterisk indicates that the species was used in fitting the CMB model.

APPENDIX D. DETERMINING WEIGHTS APPLIED TO THE FITS

For each sample, a range of models were fit. As explained in section 4.1, the results were summarized as a weighted average, where the weights were based on the product of two probabilities – a chi-square for how closely individual calculated species matched the measured species, and a Gaussian for the difference between the total measured mass and the calculated total.

One limitation with this approach is that the true distribution of the chi-square statistic is difficult to estimate because it involves unknown statistical correlations among the chemical species. In other words, the chemical species are not statistically independent, and the lack of independence between chemical species suggests that the chi-square statistic might not have a chi-squared distribution. In fact, the CMB fits produced "chi-square" statistics that averaged about 0.5. If the model were correct, these statistics should be averaging about 1.0.⁹

A second issue is that a range of models is fit to the *same* sample. Some samples were inherently easy to fit, others difficult, so that, for some samples, a large number of chi-square statistics were very small, suggesting good fits, in others few if any chi-square statistics were small. Thus, the assumption that the chi-squared statistic has a chi-squared distribution yielded unrealistic probabilities.

We minimized this problem by indexing the chi-square fits to the best fit, which was arbitrarily given a value of 12 – the 25th percentile of a chi-square with 17 degrees of freedom. Each other chi-square statistic, S , was adjusted to 12 times its ratio to the minimum: $12 * S / S_{\min}$. This adjusted value was then assumed to have a chi-squared distribution.

For the second p-value, the difference between the measured and calculated masses was assumed to have a Gaussian distribution. However, occasionally there were large differences between the measured mass and any of the calculated masses – in a number of cases, the sum of the measured species was considerably larger than the measured mass, indicating that one or the other was substantially mismeasured resulting in a model that fit badly for individual components, nevertheless gave a better second p-value because the sum of the modeled species accidentally got closer to the mismeasured total PM.

To account for these anomalies, a "posterior" probability was calculated, for the likelihood that a problem had occurred. It was arbitrarily assumed that initially there is a 90% chance that the measurements are not grossly inaccurate. The validity of this assumption was checked by comparing the measured mass with the (adjusted) sum of the measured species. Specifically, the species were summed, except the OC and EC were

⁹ Actually, the statistic was had the form of a chi-square statistic divided by its degrees of freedom. Under the assumption that the terms in the chi-square come from a set of independent standard normal random variables, the expected value of a chi-square equals its degrees of freedom. So dividing by the degrees of freedom produces a statistic with mean 1.0.

multiplied by 4/3 to adjust for the fact that for two of the most common sources of ambient carbon – wood smoke and motor vehicle exhaust – the measured species account for about 75% of the measured mass (the remainder including oxygen and hydrogen), and Aluminum and Silicon were multiplied by 1/.6 to account for the fact that the measured geological species account for about 60% of the mass (much of the rest being oxygen).

A z-statistic was computed as the difference between the measured mass and sum of species (as described above) divided by their estimated standard error. A p-value was computed as $p = 2 * \Phi(|z|)$, and a posterior probability computed: $p_1 = .9p / (.9p + .1(1-p))$. If p is close to 1.0, the posterior is also close to 1.0. If p is equal to 0.1, indicating some chance that there are serious measurement discrepancies, then the posterior probability is 0.5. If p is small, 0.01 or less, then the posterior is also small.

To adjust the second term for this posterior probability of erroneous measurements, the Gaussian was raised to the power of the posterior: $[\Phi(w)]^{p_1}$, where $w = |\text{measured mass} - \text{calculated mass}| / (\text{estimated sd of difference})$. If p_1 is near 1, so there is good agreement between the measured mass and the sum of individual species, then $\Phi(w)$ is relatively unchanged. But if there is a large discrepancy, then p_1 is small, thereby shrinking $[\Phi(w)]^{p_1}$ toward 1, so that this term doesn't play much of a role in the overall fit.

The weight for an individual fit equals the product of its two p-values, divided by the sum of all the 1,100 or so products of p-values (for which all estimated coefficients are positive). The coefficients reported for a given sample is the weighted average of the coefficients. The coefficients' standard errors are estimated as the square root of the weighted average of the individual variances, i.e., the squares of the standard errors provided by the model. The definition of the weights, w_i , can be summarized as follows:

$$w_i = r_i / \sum r_j$$

where $r_i = P(X > 12S_i/S_{\min}) * [\Phi(w_i)]^{p_1}$, and X is a chi-squared random variable with degrees of freedom = # of fitted species - # of fitted source categories.

APPENDIX E. INITIAL CMB RESULTS AND SOURCE CATEGORY SELECTION

The initial CMB results showed sizeable uncertainties for some of the source categories. This indicated that the data should be reanalyzed with fewer categories (See Section 5). This appendix describes how this subset of source categories was chosen.

If several sources have similar relative quantities of chemical species, the CMB model will have difficulty distinguishing between them. This leads to an inflation in the associated standard errors, i.e. uncertainties. Wood smoke from different tree species, for example, may have very similar source profiles, at least on the species measured for this analysis.¹⁰

In the initial CMB runs, only one source from a class was included at a time, e.g., only one woodsmoke profile and only one auto exhaust profile. However, identification problems persisted. In particular, the profiles of auto exhaust, wood smoke and commercial cooking are similar enough that the CMB model may have difficulty distinguishing them. Of course, from the viewpoint of PM controls, the distinction is crucial. Thus, for the initial CMB runs, various combinations of these sources were tried.

In order to determine the magnitude of uncertainty, two quantities were estimated. One is the *total* uncertainty – the standard error of the quarterly averaged source coefficients. It was computed from the estimated coefficients themselves. If the coefficient average were a simple arithmetic mean, then this estimate would be the usual sample standard deviation. This standard error includes variation from 3 sources: model uncertainty, meteorological variation, and variation in underlying emissions.

The second quantity was estimated *model* uncertainty of the quarterly averaged source coefficients, based on combining estimated standard errors provided by the model for each coefficient for each fit. For a given coefficient and site, the combining formula was the square root of:

$$S_{\text{ave}}^2 = \frac{1}{16} \sum_{i=1}^4 \frac{1}{n_i^2} \sum_{j=1}^{n_i} s_{ij}^2$$

where n_i = number of fitted coefficients in quarter i , and s_{ij} = model estimated standard error for the j th sample in quarter i .

Figure E1 presents a comparison of total and model uncertainties for each source category for each site. For the top 3 source categories – ammonium nitrate, ammonium sulfate, and marine air – total uncertainty is much greater than modeled. This implies

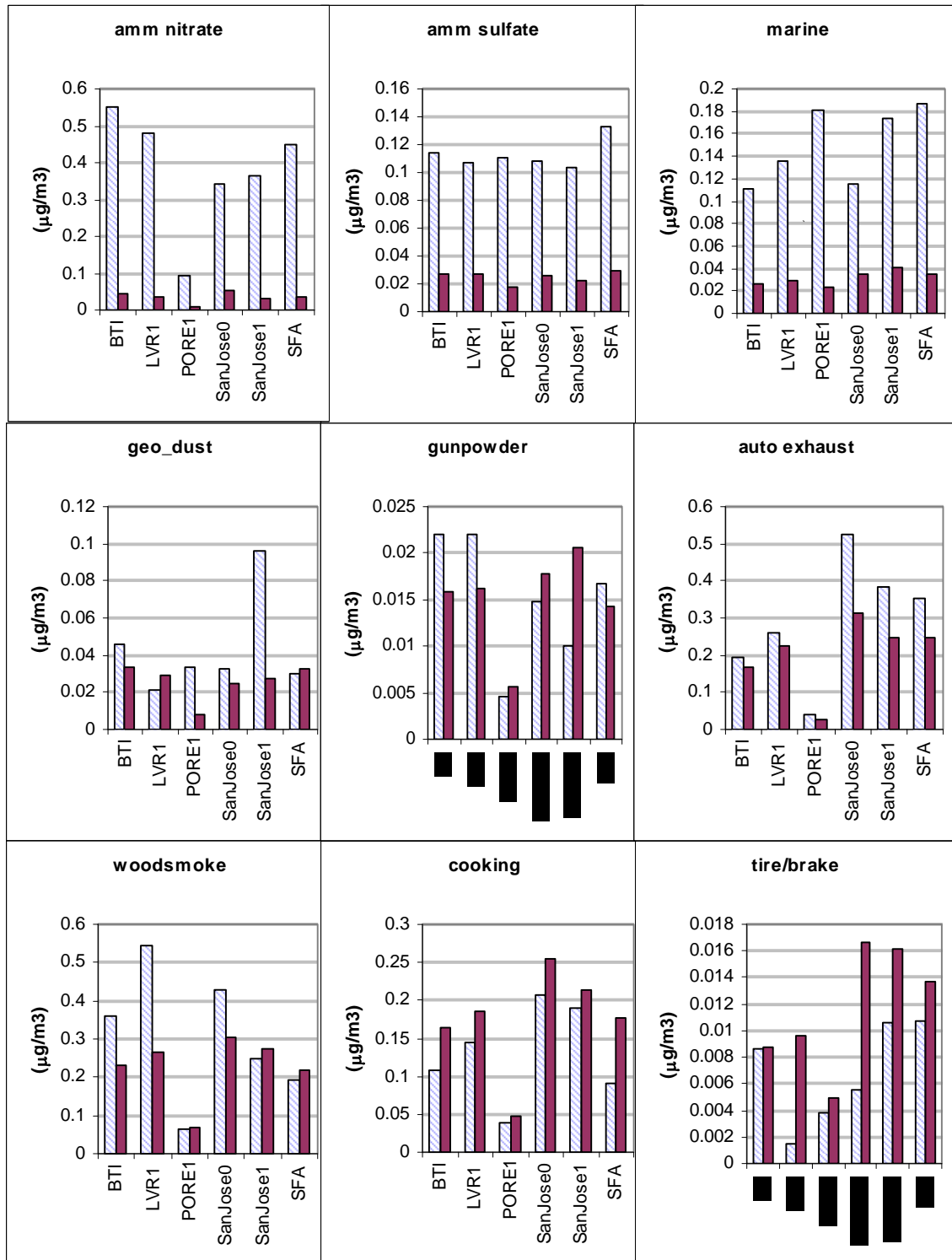
¹⁰ Some differences can be found by speciating the organic carbon. However, this was not done for this study. Wood smoke and cooking can be differentiated from fossil fuel PM with Carbon-14 analysis. Some of this has been done for Bay Area PM, but not for this study.

that there is little difficulty in distinguishing these sources and, moreover, there was substantial variability in the contributions from that source from sample to sample.

For the three correlated sources – auto exhaust, woodsmoke and commercial cooking – the modeled uncertainties are sometimes as large or larger than the total. (Of course, the *actual* modeled uncertainty must be less than or equal to the *actual* total.) This indicates that the model has difficulty in distinguishing among them. At some sites, however, it does appear the model can distinguish the first two sources – LI and BI for woodsmoke and SJ and SF for auto exhaust. But at no site could the commercial cooking source be distinguished above the model uncertainty. Therefore, the commercial cooking profile was dropped in the final CMB runs.

The final two sources – geological dust and tire/brake wear – also showed modeled uncertainties as large as total. This is not due to confounding with other sources but instead because they occur at such low concentrations that they are difficult to measure, that is, their distinguishing chemical species are not above the limits of detection. For example, for geological dust, silicon and aluminum are key species, consisting of about 25% and 10% of total mass, respectively. In the samples analyzed, these species occurred above the limits of detection in only 24 and 34 out of 208 cases, respectively. However, the model was able to distinguish geological dust at some sites. Geological dust has a unique signature, so it can be readily differentiated from the other sources, even at low concentrations. Because tire/brake wear did not occur above the detection limits of the model, it was eliminated from the analysis.

Figure E1. Model uncertainties vs. total uncertainty for fitted source categories



Note: Comparison of model uncertainties with total uncertainty for fitted source categories. Lined blue bars = total uncertainty. Solid red bars = modeled.

APPENDIX F. MODIFICATION OF CMB RESULTS BASED ON CARBON-14 ANALYSIS

Carbon-14 analysis combined with results from the previous source apportionment study and the fractions in the District's emissions inventory suggest that the CMB analysis of the current data over-estimated the fossil contribution to $PM_{2.5}$ vis-à-vis the vegetative burning component. The current CMB analysis used wood and motor vehicle source profiles from CRPAQS (a major PM study in northern California), which were developed for the Fresno area. This profile had a much higher soluble potassium content than the wood profile from our previous source apportionment study, which measured wood smoke from three Bay Area households. An in-house sensitivity analysis using wood smoke profiles from various studies showed a strong negative correlation between the soluble potassium content of the wood smoke and the estimated vegetative burning contribution. Thus, because the Fresno wood profile had a soluble potassium content that may be higher than that of Bay Area wood, there is a reason to believe that the current CMB analysis underestimates the wood smoke contribution vis-à-vis the fossil fuel contribution.

The Carbon-14 analysis provides measurements that can be used to adjust the CMB results. The new carbon content from PM_{10} filters from 2004 and 1998 was measured. The filters were composited from a set that included a sample representative of the 4 quarters in the year and the appropriate percentage of weekends and weekdays. The CMB analysis described here was on filters containing $PM_{2.5}$, not PM_{10} . However, the vast majority of ambient carbon (deriving from vegetative burning, fossil fuel burning, cooking or secondary organic) is largely in the $PM_{2.5}$ range. There is no reason to suspect that the $PM_{2.5}$ from the same periods would have a new to old carbon ratio different from PM_{10} .

Although the apportionment of carbonaceous $PM_{2.5}$ into fossil and vegetative sources contains large uncertainties, their *sum* is well-estimated by CMB. The annual mean CMB results were adjusted to have the same total carbon, but with the vegetative/fossil ratio equal to the new/old carbon ratio found in the University of Arizona analysis.

The San Francisco and Livermore sites have C-14 measurements for both 2004 and 1998. These were interpolated to provide annual vegetative burning fractions for 2000: $(2 \cdot .452 + .490)/3 = .465$ for San Francisco, and $(2 \cdot .521 + .572)/3 = .538$ for Livermore. For San Jose and Bethel Island, only the 2004 measurements were available. It was assumed that the annual percent change in C-14 fractions at these sites was the average for the Bay Area sites, namely an increase of 9.5%, yielding .513 for San Jose 2000, .521 for San Jose 2001, and .592 for Bethel Island.

The source apportionments for peak and seasonal $PM_{2.5}$ were also adjusted. The goal was to adjust the profiles so that new/old carbon ratio of the adjusted average equaled the ratio found in the Carbon-14 analysis, yet to remain as faithful as possible to the new/old ranking found in the CMB analysis. A method that fits these constraints is as follows.

For a given site, let x_i = the CMB vegetative burning estimate for day i , and y_i = the CMB fossil fuel estimate, and let x_i' and y_i' be the values modified to correspond to the Carbon-14 annual means. That is,

$$\frac{\bar{x}'}{\bar{x}' + \bar{y}'} = c, \quad (F1)$$

where c = the new/old carbon ratio, \bar{x}' = mean of the x_i' , and \bar{y}' = mean of the y_i' . We used the adjustment $x_i' = a_i * x_i$, where $a_i = (1 - k) * r_i + k$, where $r_i = x_i / (x_i + y_i)$. That is, the adjustment factor, a_i , is a linear function of the initial CMB vegetative fraction, r_i , with the property that if $r_i = 1$ (i.e., $y_i = 0$), then $a_i = 1$; i.e., there is no adjustment. As r_i shrinks to zero, a_i increases, provided $k > 1$. We still need $a_i * r_i \leq 1$, which is always satisfied provided $k \leq 2$.

Applying the formula F1 yields $k = \frac{c(\bar{x} + \bar{y}) - \bar{s}}{\bar{x} - \bar{s}}$, where \bar{s} = the mean of $r_i * x_i$.

Using this a_i formula for Bethel Island and Livermore produced values of k between 1 and 2. But applying the formula to San Francisco and San Jose produced values of $k > 2$. In these latter cases there were, in fact, a couple of instances where $a_i * r_i$ was > 1 . In these cases, a_i was set to $1/r_i$.

To summarize, for a given site, each pair of CMB fits x_i = vegetative burning concentration on day i and y_i = fossil fuel combustion concentration on day i were transformed to x_i' , y_i' as follows:

$$x_i' = a_i * x_i, \quad y_i' = x_i + y_i - x_i', \quad \text{where}$$

$$a_i = (1 - k) * r_i + k \quad (F2)$$

where $r_i = x_i / (x_i + y_i)$, $k = \frac{c(\bar{x} + \bar{y}) - \bar{s}}{\bar{x} - \bar{s}}$, with \bar{s} = the mean of $r_i * x_i$, and c = the vegetative fraction for 2000 for that site.