

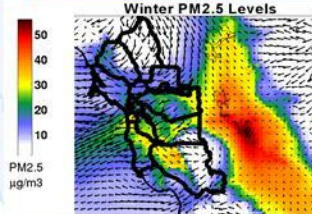
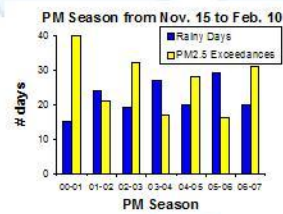
BAY AREA
AIR QUALITY
MANAGEMENT
DISTRICT

939 Ellis Street, San Francisco, CA 94109

Research and Modeling Section Publication No. 201009-010-PM

Trends in Bay Area Ambient Particulates

November, 2011



Prepared by:

David Fairley, Statistician

Reviewers:

Scott Beaver, BAAQMD
Philip Martien, BAAQMD

Reviewed and approved by:

Saffet Tanrikulu, Manager, Research and Modeling Section



Table of Contents

Summary	1
1. Introduction.....	8
2. PM Measurements in the Bay Area	9
3. PM Standards	14
4. Statistical Methods.....	16
5. PM _{2.5} Trends	20
6. PM ₁₀ Trends	28
7. Trends in PM components	31
8. PM _{2.5} Trends in Central California Air Basins.....	40
9. Comparisons with the Emissions Inventory.....	44
10. Trends in PM quantiles	48
11. Trends in PM Health Effects.....	52
12. Recommendations.....	53
References.....	54
Appendix A. Uncertainty of 24-hour design value trend estimates	55
Appendix B Methodology and Meteorological Variables Considered in Wintertime Regression	58
Appendix C – Numbers of days with PM ₁₀ Measurements	61
Appendix D. Use of Dichotomous PM _{2.5} Data for Trend Analysis	62
Appendix E. Adjustment of nitrate measurements for the impact of refrigeration.....	64

Summary

Airborne particles present the greatest estimated health risk of any air pollutant in the San Francisco Bay Area. Understanding the trends in ambient concentrations of this particulate matter (PM) is a key to judge the effectiveness of past and future air quality regulations. This report analyzes the trends in Bay Area PM from 1990-2010.

Background

PM is a complex pollutant, not only deriving from many sources, but also appearing in a wide variety of sizes and compositions, and forming via multiple pathways. PM was one of the six pollutants chosen in the original 1970 Clean Air Act for regulation and adoption of standards.

The original PM standard was set for total suspended particles (TSP). But over the next 40 years a mountain of scientific evidence has shown that the tiniest PM particles are the most dangerous¹. Thus, the TSP standard was changed to a PM₁₀ standard (particles less than 10 microns in diameter) in 1987, and a PM_{2.5} standard (particles less than 2.5 microns in diameter, aka fine PM) was promulgated in 1997. As a result, the District began routine, multi-site PM₁₀ monitoring in 1989 and PM_{2.5} monitoring in 1999.

Historically, PM has been measured on filters. A filter is pre-weighed then placed in a sampler that draws air through the filter, typically for 24 hours. Then the filter is post-weighed and the difference in weights converted to concentration by dividing by the volume of air drawn. More recently, instruments have been developed to measure PM continuously. However, PM standards are still based on 24-hour concentrations.

There are two forms of PM standards – annual and 24-hour. The annual standard is designed to cap individuals' total PM exposure; the 24-hour standard, individuals' peak PM exposure. Corresponding to each standard is a *design value* – a statistic computed from the set of PM measurements that is compared with the standard to determine compliance. For the annual standard, the design value is the quarterly averaged annual mean averaged over the three most recent years. For the 24-hour PM_{2.5} standard, the design value is the 98th percentile of the PM measurements averaged over the three most recent years. Thus, we focus on the annual means and 98th percentiles in the trend analysis below.

Ambient PM is a mixture of many components that transform, coalesce and agglomerate in the atmosphere. In contrast to the scientific evidence for an increased risk as particle size decreases there is little information for determining which component of PM poses the highest risk; current PM standards implicitly assume that all fine PM poses a risk irrespective of composition.

Ambient PM is comprised of two formation categories: *primary* PM – PM that is directly emitted, and *secondary* PM – PM that forms through atmospheric transformations of gases and liquids. It is also comprised of two source categories: *anthropogenic* PM – PM from human activity; and biogenic, or

¹ The USEPA has compiled the relevant studies into a 2000 page science assessment for the latest update to the National Ambient Air Quality Standards (EPA 2009).

“natural”, PM – PM from natural sources. Table S1 shows the main components in these categories in the Bay Area.

Table S1. Main components of Bay Area PM_{2.5} by category

	Anthropogenic	Biogenic/“natural”
Primary	carbonaceous PM from combustion of fossil fuels and wood, some geological dust, tire and brake wear	marine air (sea salt and sulfate), some geological dust, wildfires
Secondary	ammonium nitrate and ammonium sulfate	secondary carbonaceous (from vegetative emissions)

The District has been making measurements of a number of PM components, including nitrate, sulfate, ammonium, chloride (a marker for sea salt) since 1990 and, more recently, elemental and organic carbon. This allowed the tracking of the contributions of these individual components to the total PM.

Results

PM_{2.5} trends: Figure S1 shows the trend in the District’s design value relative to the national annual PM_{2.5} standard. The District’s design value dropped from 14 µg/m³ for 1999-2001 to 10 µg/m³ for 2008-2010, a 28% reduction, with a 90% confidence interval of 21% to 33%. As can be seen, the District met the national annual PM_{2.5} standard for the whole period.

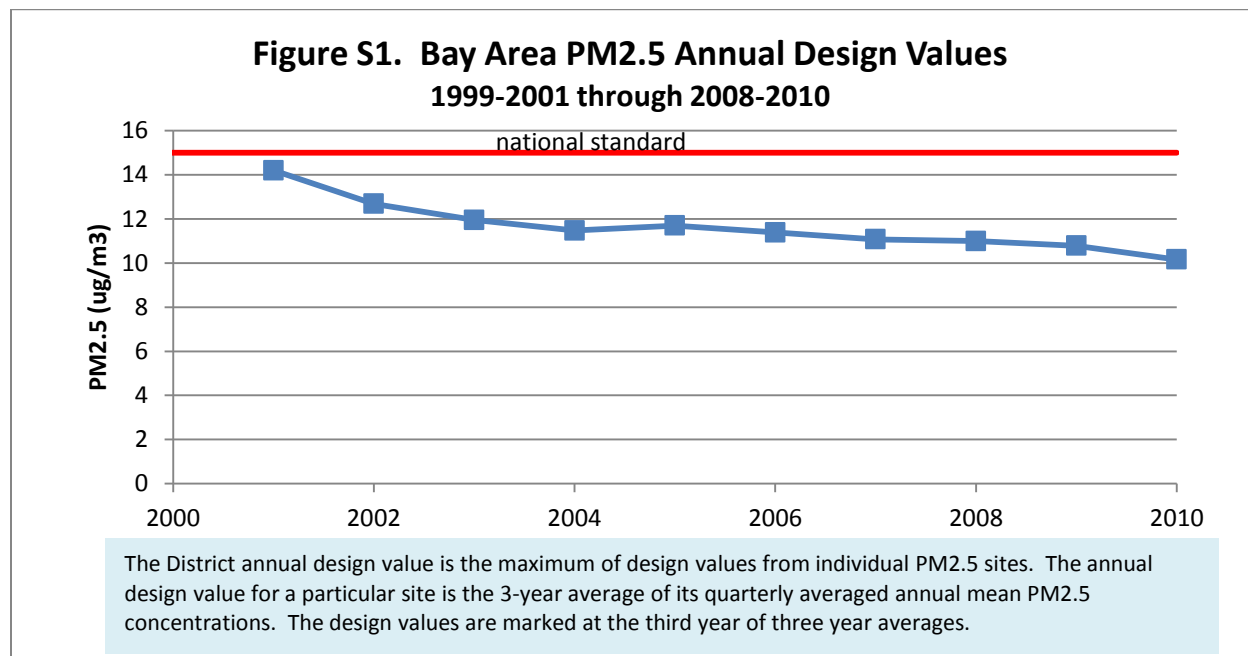
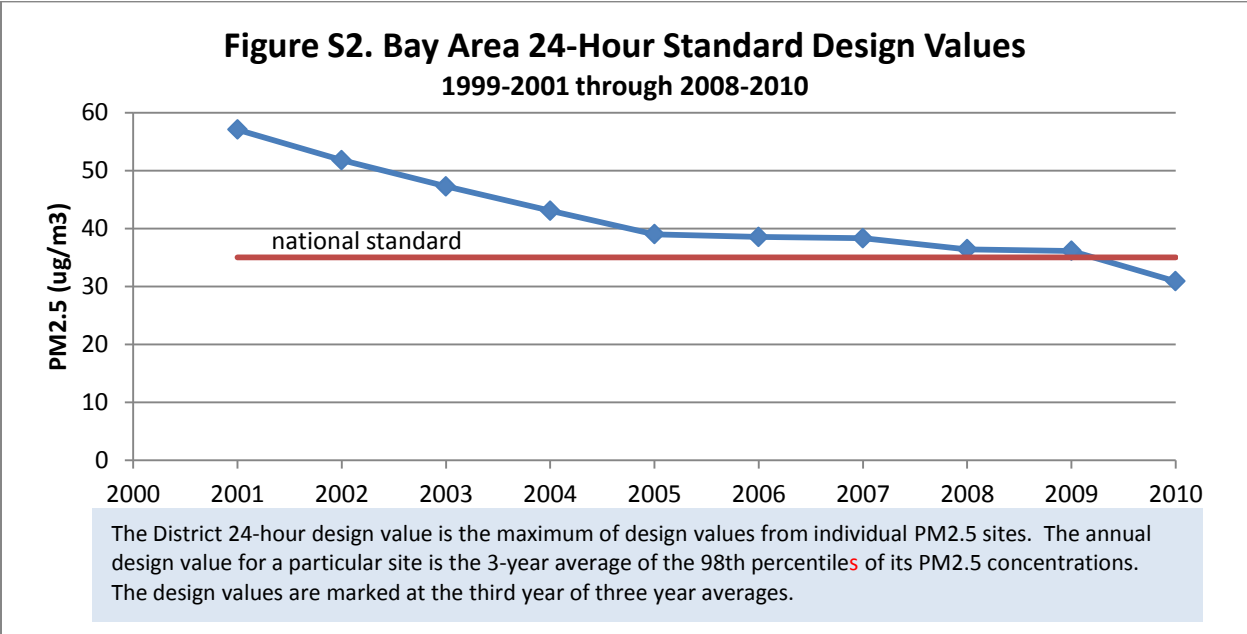


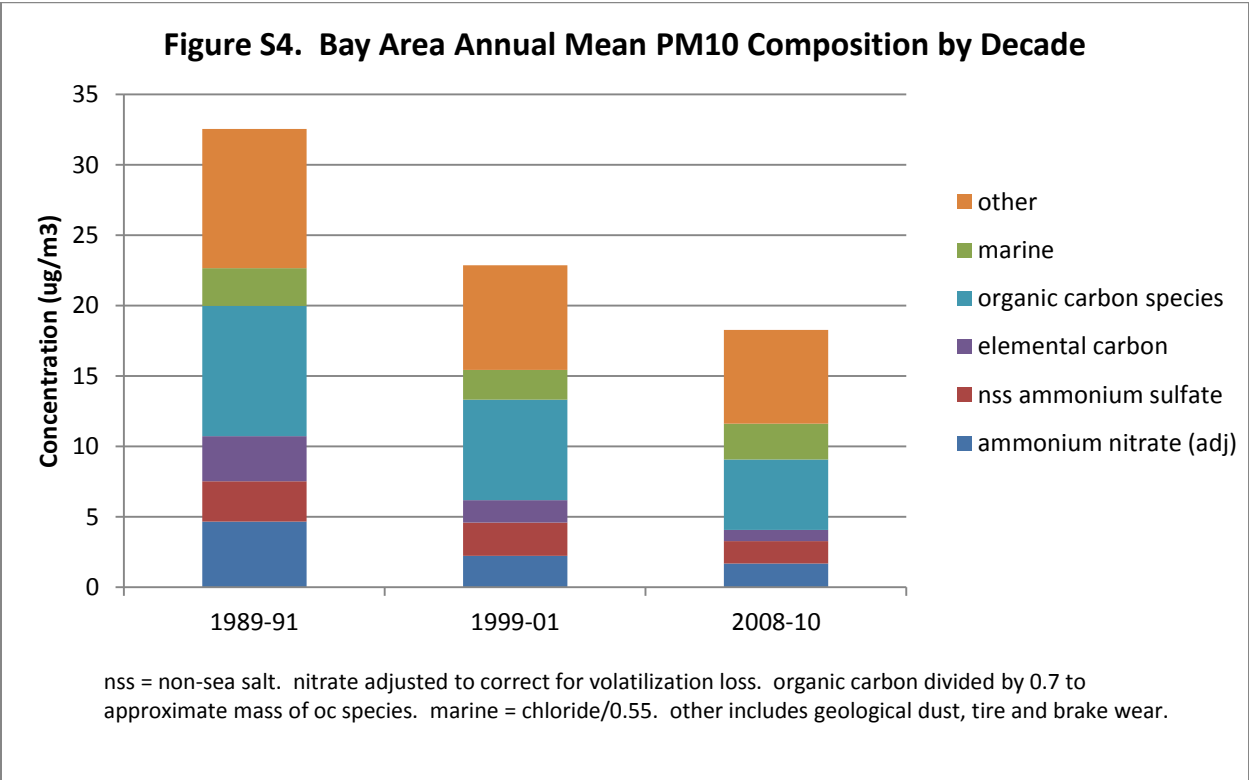
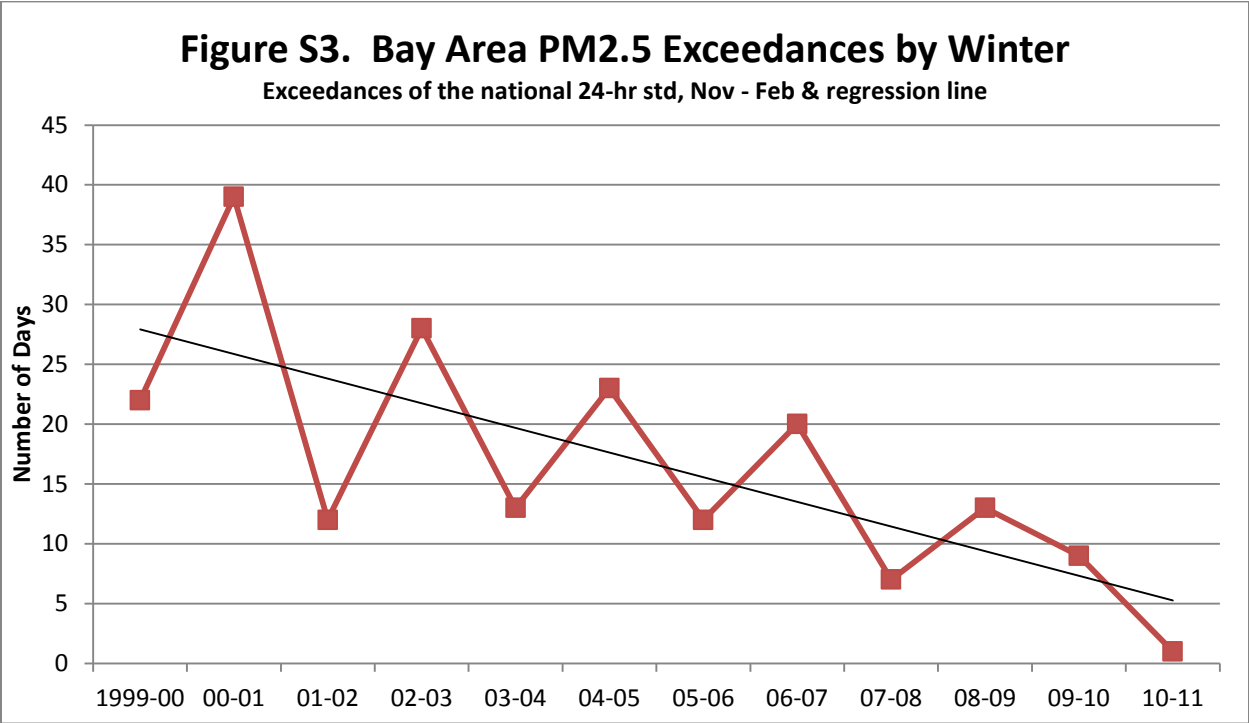
Figure S2 shows the District’s design values relative to the national 24-hour PM_{2.5} standard. For the earliest period available, 1999-2001, the design value was 57 µg/m³. By 2008-2010, it was 31 µg/m³, which meets the standard. Although there is considerable year-to-year variation in PM_{2.5} concentrations, the District is likely to continue to meet this standard. The reduction from 1999-2001 to 2008-2010 was 46%, with a 90% confidence interval of 29% to 63%.



We also investigated trends in number of days exceeding the 24-hour standard. Figure S3 shows the trend in exceedances of the standard across 11 winters from 1999-2000 through 2010-2011. There is considerable winter-to-winter weather-driven variation, giving the graph a saw-tooth pattern. But overlain is a clear downtrend, from an average of over 24 exceedances for the first 3 winters to under 8 exceedances in the most recent 3 winters, a 68% decrease. This decrease is statistically significant, and becomes more obvious with adjustment for meteorology.

PM₁₀ Trends: Bay Area annual mean PM₁₀ was reduced from 33 µg/m³ in 1989-1991 to 18 µg/m³ in 2007-2009, a total reduction of 44% representing an annual reduction of 3% per year. These reductions are a result of reductions in each of the anthropogenic components tracked by the District: ammonium nitrate, ammonium sulfate, and elemental and organic carbon.

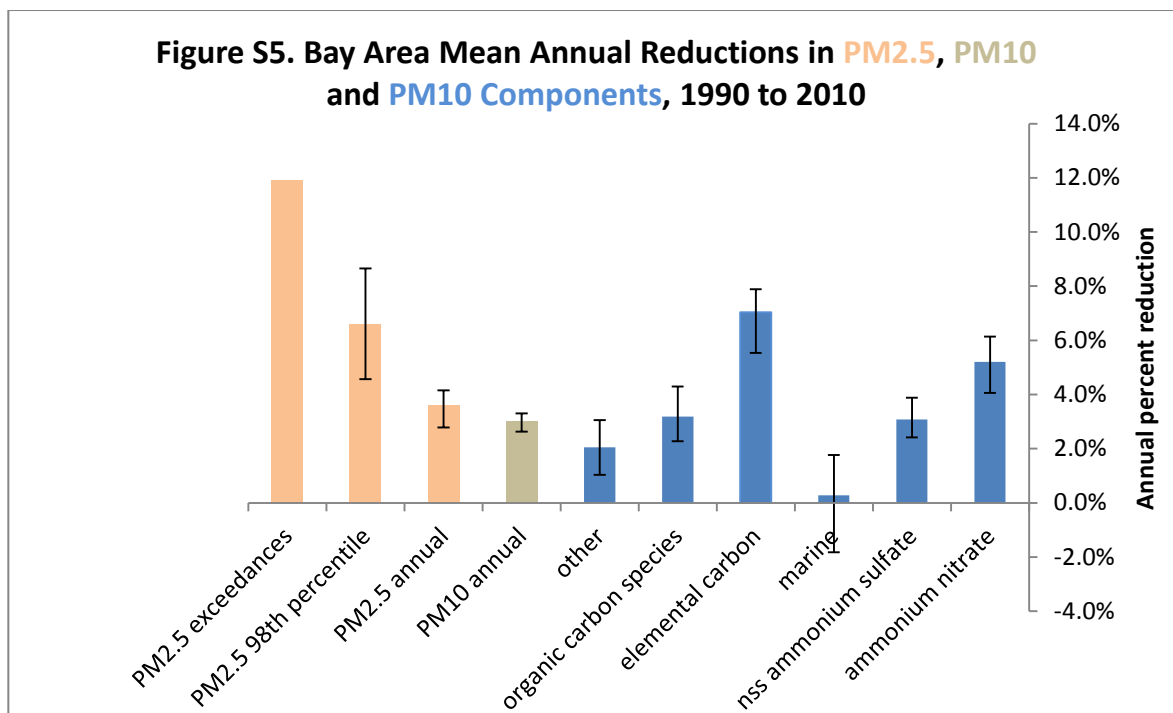
Figure S4 shows PM₁₀ composition by decade. There have been substantial reductions in each of the anthropogenic components.



PM_{2.5} – PM₁₀ comparison: Figure S5 shows the annual percent reductions in PM_{2.5} summary statistics from 2000 to 2010 and PM₁₀ components from 1990 to 2010 along with 90% confidence intervals. Reductions are shown on an annual basis to make the PM_{2.5} and PM₁₀ reductions comparable. Reductions

in annual mean PM_{2.5} and PM₁₀ are not statistically different, although PM_{2.5} may have a larger rate of reduction, because it has a lower percentage of components not expected to decrease – marine air and geological dust. Note that the rate of decrease in 98th percentile is larger than for the annual average, so that peak PM values appear to be particularly affected by anthropogenic emissions reductions.

Among the PM components that the District tracks, the greatest decrease is in elemental carbon (EC), which dropped about 7% per year from 1990 through 2010. The largest single source of Bay Area EC is diesel exhaust. It appears that diesel exhaust PM may have been reduced more than predicted in the emissions inventory, which estimates a decrease of only a couple percent per year. This conclusion is strengthened by the observation that organic carbon (OC) has been reduced less than EC. Unlike EC, Bay Area OC derives largely from other sources, including wood burning, cooking, and gasoline exhaust. The fact that EC was reduced at roughly double the rate of OC suggests that sources with high EC/OC ratios, specifically diesel, were reduced more than sources with low EC/OC ratios. Another bit of evidence is that the ratio of carbon from fossil fuels to carbon from non-fossil sources (including woodsmoke and meat cooking) has decreased, based on Carbon-14 analysis.



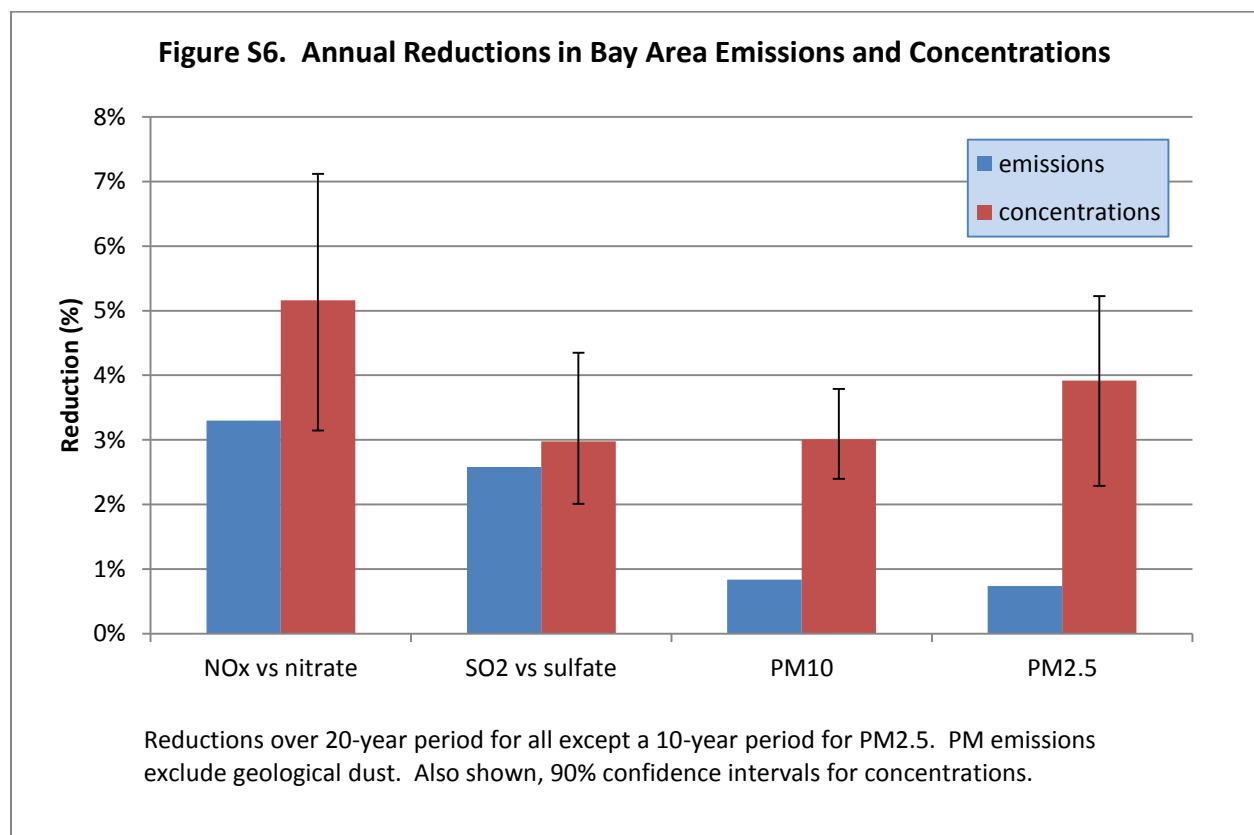
Ammonium nitrate, one of the largest components of Bay Area PM_{2.5}, has also been reduced substantially, with a rate of about 5% per year. Organic carbon and non-sea salt sulfate have been reduced about 3% per year, and other PM₁₀ species (including geological dust, tire and brake wear) have been reduced about 2% per year. As expected, the marine air component shows no change.

Comparison with other air basins: The Bay Area has shown more consistent progress in reducing PM_{2.5} than its two neighboring air basins to the east – the Sacramento and San Joaquin valleys. Like the District, these areas have PM_{2.5} measurements dating back to 2000. Over the past decade, Sacramento’s annual PM_{2.5} has been reduced by an amount similar to that of the Bay Area, and it meets the national

annual PM_{2.5} standard. But its 24-hour design value hasn't dropped significantly and it still violates the national 24-hour standard. The San Joaquin Valley has seen little progress in reducing its annual mean PM_{2.5} concentrations, and it is still well above the national annual standard. San Joaquin has reduced its 24-hour design value significantly, but it is still well above the standard.

Comparison with emissions inventory trends: Figure S6 shows a comparison of annual reductions in ambient nitrate, sulfate, PM₁₀, and PM_{2.5} compared with reductions in emissions of NO_x, SO₂, PM₁₀ and PM_{2.5}. The percent decreases in sulfate match those of SO₂. Although the point estimate of the decrease in nitrate is greater than that of NO_x, the two are not statistically different, owing to nitrate's large year-to-year variability.

Reductions in ambient PM₁₀ and PM_{2.5} concentrations are much greater than the reductions registered in the emissions inventory, even excluding road dust (which the inventory estimates to have been increasing). Although ambient PM contains secondary compounds, which have been reduced, there has been a substantial reduction in concentrations of PM directly emitted from anthropogenic sources. So there is a substantial difference between ambient and emissions inventory PM reductions for some source categories.



Health effects: Ambient PM has a major impact on the health of individuals in the Bay Area. It is implicated in a variety of serious health effects, including triggering chronic bronchitis, asthma, heart attacks, and premature death. The number of deaths affected is on the same order as that of homicide or traffic accidents.

We estimate that reductions in PM_{2.5} from 1990 to 2010 have reduced the number of premature deaths in the Bay Area by more than 3,000 per year². Life expectancy in the Bay Area has grown by 5 years in that time, from 76 years in 1990 to 81 years in 2010. We estimate that reductions in PM resulted in a 6-month increase in life expectancy, so that about a tenth of the total improvement is due to reductions in PM_{2.5} concentrations.

² Specifically, the number of deaths that occurred in 2010 with 2010 PM_{2.5} concentrations is estimated to be 3,000 fewer than would have occurred if PM_{2.5} concentrations had remained at 1990 levels.

1. Introduction

Airborne particulate matter (PM) constitutes the class of ambient air pollutants with the most serious impact on the health of Bay Area residents. PM health effects range from lower and upper respiratory symptoms including asthma and chronic bronchitis to heart attacks and death. These illnesses result in thousands of absences from school, days of lost work, and hospital visits annually. We estimate that the number of deaths in the Bay Area from elevated PM levels rivals that of motor vehicle fatalities or homicides.

The Bay Area attains the national annual $PM_{2.5}$ standard, but has continued to violate California's annual standard. The Bay Area is in non-attainment of the national 24-hour $PM_{2.5}$ standard. It meets the national PM_{10} standard, but violates California's PM_{10} standards. Efforts at the regional, state and national levels have reduced PM concentrations. This analysis estimates the amount of reductions in ambient PM as well as in its different chemical and size components.

1.1 History of PM standards and PM measurements

PM was one of the six "criteria" pollutants regulated in the original Clean Air Act of 1970. Standards were set for each of the pollutants and a procedure established for periodic review and update based on new scientific evidence, a process that continues 40 years later.

The original PM standards were set in terms of total suspended particulates (TSP). Evidence gradually accumulated that smaller particles appeared to account for most of the health impacts. Thus, in 1987, the EPA replaced the TSP standard with a standard for PM_{10} (particles less than 10 microns in diameter), and in 1997, it added a standard for even smaller particles, $PM_{2.5}$ (particles less than 2.5 microns in diameter). (See Lippmann 2011.) Today, there is mounting evidence that ultrafine particles (particles less than about 0.1 micron in diameter) may even have greater health impacts.

1.2 Time intervals for PM standards and measurements

Until recently, PM was measured by collecting particles on filters over 24 hours, and PM concentrations were estimated by weighting the filters before and after collecting the particles. Because this process is labor-intensive, measurements have not been made every day at every station. At most stations, they were made either on a 1-in-3 or 1-in-6 day schedule.

Air pollution trends are typically hard to discern over a period of a few years. Day-to-day and year-to-year variation in meteorology can lead to large swings in pollutant concentrations. Therefore, we conducted analyses of trends over the entire period of $PM_{2.5}$ measurements: 1999-2010. For PM_{10} , data are sufficient to estimate trends only by decade: 1990s and 2000s.

PM also varies in composition including: carbonaceous particles from combustion of fossil fuels and wood; geological dust particles; sea salt particles; tire and brake wear particles. *Secondary* particles are those formed by chemical reactions or physical interactions in the atmosphere. Ammonium nitrate and

ammonium sulfate account for the bulk of secondary PM in the Bay Area. Evaluating component trends is useful for understanding the overall PM trend, which is the composite of these component trends. The District has measured nitrate and sulfate on a 1-in-6 day schedule since 1989. It has also measured organic and elemental carbon, two other key PM components, since 2004.

1.3 Seasonality

There are large swings in PM concentrations by season. Almost all the Bay Area's highest PM concentrations occur in the wintertime. The reasons are several: 1. Wood burning is a significant source, and occurs mainly in winter. 2. Winter includes periods of atmospheric stagnation and stability, with particles trapped close to the surface that build up over period of days. 3. Ammonium nitrate particles form in colder weather. 4. Significant transport occurs from the Central Valley, as stagnant periods coincide with drainage of air through the Carquinez Strait and various passes out to the Pacific Ocean.

The District defines the *Winter Spare the Air* season as the months of November through February. Analysis of exceedances of the 24-hour PM_{2.5} standard is restricted to these months.

2. PM measurements in the Bay Area

The set of PM measurements made in the Bay Area has changed over time. This section discusses what measurements have been made and where and when.

2.1 PM measurements

The District began measuring PM₁₀ concentrations at a number of sites in 1989 on a 1-in-6 day schedule. In addition to total PM₁₀ concentrations, a set of ions has been measured: nitrate, sulfate, ammonium, and chloride. Starting in 1995, potassium was added, and elemental and organic carbon (EC and OC) in 2004. In recent years, the number of sites measuring PM₁₀ has been substantially reduced.

The District began measuring PM_{2.5} in 1999. Until recently, the primary measurements were filter-based. A filter is put in a sampler that draws a steady and monitored airflow through it for 24 hours. The PM_{2.5} concentration is estimated from the post-pre weight difference of the filter divided by the total air flow, yielding a measurement of ambient PM in micrograms per cubic meter ($\mu\text{g}/\text{m}^3$). These measurements are called Federal Reference Method (FRM) and have been the ones used for determining compliance with the national air quality standards.

The PM_{2.5} measurement schedule was complex, with some sites having daily wintertime measurements and others once in 3 days. All sites had a 1-in-6 day measurement schedule from April through September.

Starting in 2001, the District began operation of BAM units – continuous PM_{2.5} analyzers at Livermore, San Jose and San Francisco. These units collect PM_{2.5} data on an hourly basis. Initially, the measurements were not accepted by EPA for reference to the national PM_{2.5} standards because of large uncertainty in this type of measurements. Recently, the accuracy has been significantly improved and the

District upgraded most of its BAM instruments. Since 2009, the PM_{2.5} measurements from the upgraded BAM instruments have been accepted as a federal equivalent measurement method (FEM) by EPA.

2.1.1 Carbon measurements

The District commenced measurement of elemental and organic carbon (EC and OC, respectively) in 2004 for PM₁₀ filters. Previously, only sporadic measurements were made.

TSP measurements were discontinued in 1998, but another measurement, Coefficient of Haze (COH), continued well after the turn of the century. COH measurements had been made by the District dating back to the 1950s. COH is a measurement based on the difference in light transmitted through paper before and after PM has been deposited on it; it is highly correlated with the elemental carbon in the particles. Thus, it can serve as a surrogate measure for EC.

2.1.2 Other measurements

Speciated PM_{2.5} measurements have been made at San Jose since February, 2000 as part of the Speciation Trends Network. Measurements include all higher elements, a set of ions, and organic and elemental carbon. A similar set of measurements has been made at Point Reyes, part of the IMPROVE network, dating back to before 1989.

In late 2008, the District started a program of speciated PM_{2.5} measurements at Livermore and Vallejo. West Oakland was added in February, 2009.

Extensive measurements were made as part of the California Regional Particulate Air Quality Study (CRPAQS), late 1999 through early 2001. These measurements included speciated PM_{2.5} measurements at several Bay Area sites that can serve as a baseline for comparison.

2.1.3 Measurements by site and date

Table 2 presents information about the PM measurements made in the Bay Area. The measurements have been made for different lengths of time with changes in the number of sites. Thus, continuous records have not always been available. Figure 2.1 shows a map with the sites.

Table 2. Measurements and measurement history at Bay Area PM monitoring stations between 1989 and 2010.

Site	PM _{2.5} Filter (frequency) ^a	PM _{2.5} FEM	PM _{2.5} Species ^b	PM ₁₀ Mass+Ions	PM ₁₀ Carbon	COH
Berkeley				2008-2010		
Bethel Island			12/1999-1/2001	1989-2010	7/2004-2010	1989 – 3/1992
Concord	3/1999-2010 (1)			1989-2010	10/2005-2010	1989 – 1/2004
Fremont	3/1999-7/2009 (3)	8/2009-2010		3/1989 – 6/2008		1989 – 1/2004
Gilroy	3/2007-10/2009 (3)	11/2009-2010				
Livermore ^c	12/1999-9/2008 (3) 10/2008-6/2010 (1)	7/2010-2010	12/1999-1/2001 9/2008-2010	1989 – 6/2008	10/2005-2/2008	1989 – 8/1995
Napa				1989-2010	10/2005-2010	1989 – 1/2004
Oakland (East)	11/2007-9/2009	10/2009-2010				
Oakland (West)			2/2009-2010			
Point Reyes	1989-2010 (3)		1989-2010			
Redwood City	3/1999-7/2008	10/2009–2010		1989 – 6/2008		1989 – 2006
Richmond/San Pablo ^d				3/1989 – 2010		1989 – 8/1995
San Francisco	3/1999-9/2009 (1)	10/2009-2010	12/1999-1/2001	1989-2010	7/2004-2010	1989 – 5/2006
San Jose ^e	1999-2010 (1) 1990-1997 ^f		2/2000-2010	1989-2010	7/2004-2010	1989 – 4/2002
SJ – Tully Rd	3/1999-9/2006			1989-2006	7/2004-2005	
San Rafael		11/2009-2010		1989-2010	11/2008-2010	1989 – 8/1995
Santa Rosa	2/1999-10/2009 (3)	11/2009-2010		7/1994-6/2008		1989 – 1/2004
Vallejo			9/2008-2010	12/1994 – 6/2008	10/2005-2/2008	1989 – 8/1995

^a (frequency): October-March frequency 1 = every day, 3 = every 3rd day; April-September frequency every 6th day at all sites.

^b Species include all elements with atomic weight of sodium or higher; ions including nitrate, sulfate, ammonium; and elemental and organic carbon.

^c Site moved in 2000 from Old Railroad Road to Rincon.

^d Richmond site closed in 5/1997, replaced by San Pablo site at El Portal Shopping Ctr through 8/2002. Replaced by San Pablo site at Rumrill Blvd 9/2002.

^e San Jose site at 4th St through 4/2002, replaced by site at Jackson St in 10/2002 (no San Jose data from 5/2002 through 9/2002).

^f A dichot sampler operated at San Jose from 1990 through 1997.

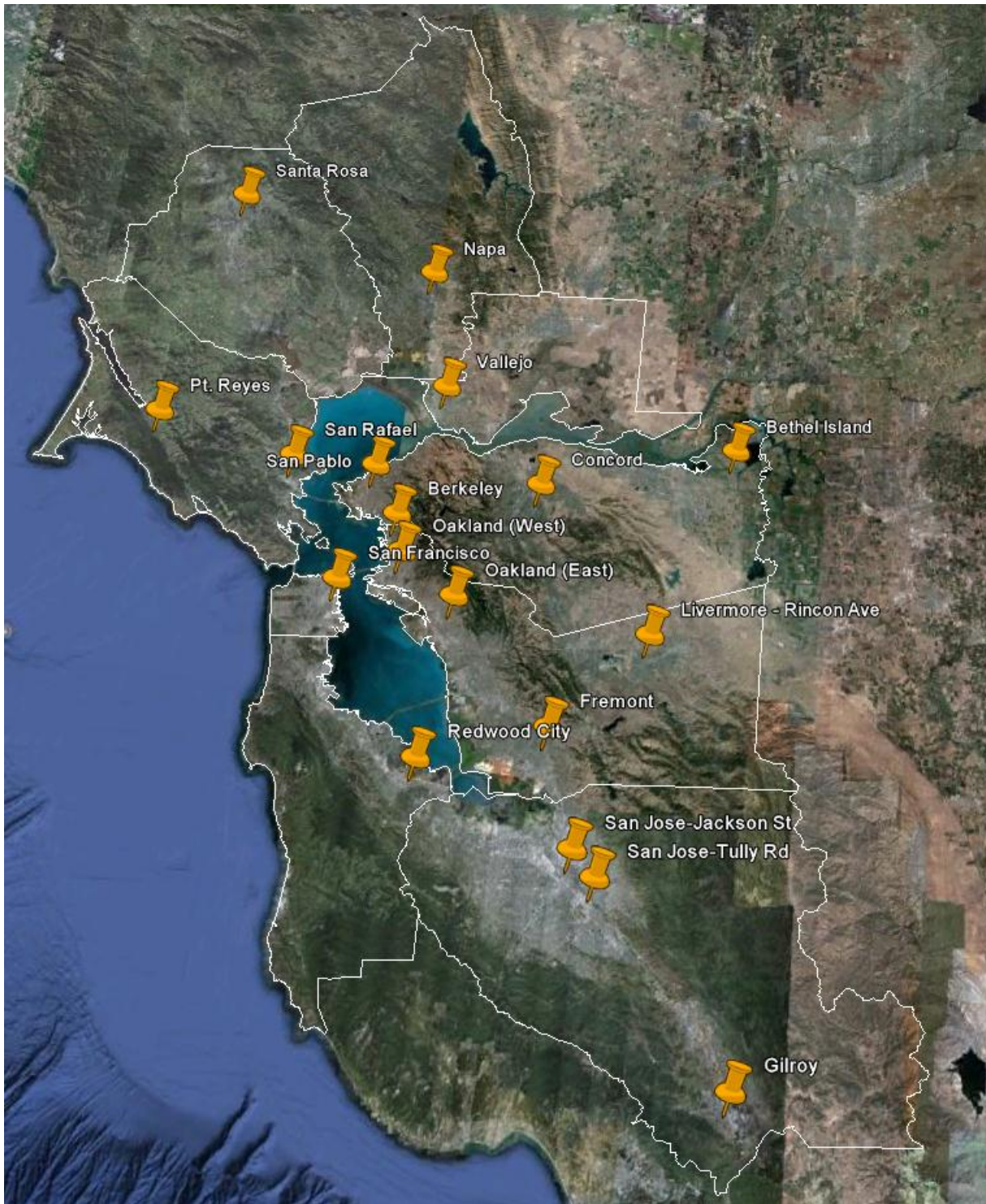


Figure 2.1. Bay Area PM monitoring sites and District boundaries.

2.2 Methodology for combining data for trend analysis

As Table 2 shows, the amount of data collected on PM and PM species has varied over time. Several strategies were used to yield as long a record as possible and to ensure that the data were comparable over time. Here is a summary of the issues and the methodology employed to deal with them.

In general, a subset of sites was chosen that had long data records. A 2-way ANOVA was used to fill in missing values, then the average taken across sites for each year. See section 4 for a discussion of the 2-way ANOVA technique.

2.2.1 PM₁₀ and PM₁₀ components

PM₁₀ filters were collected from 1989 through 2007 at many sites and analyzed for PM₁₀ mass, nitrate, sulfate, and chloride. Measurements were discontinued at about half the sites in 2008. A 2-way ANOVA was used to fill in the missing values to extend the record through 2010.

2.2.2 PM₁₀ nitrate

Since 2009, the District has refrigerated PM₁₀ filters after the sample had been collected. Before 2009, PM₁₀ filters were unrefrigerated. The impact of the change could be estimated by comparison with PM_{2.5} measurements made at San Jose for the STN network where the filters have always been refrigerated. There appears to have been little impact for other PM₁₀ components, but for nitrate there was a clear difference, with substantial volatilization loss from the unrefrigerated filters. A method was developed to adjust the pre-2009 measurements to make them comparable to the current ones. (See Appendix E.)

2.2.3 PM_{2.5} measurements

The District began its program of collecting PM_{2.5} measurements in 1999. Before that there had been some limited PM_{2.5} measurements made in a special study in San Jose from 1990 to 1997. Comparisons with concurrently measured PM₁₀ showed that the special study PM_{2.5} registered concentrations about 20% lower than the PM_{2.5} measured since 1999. Appendix D presents an analysis and develops an adjustment method.

2.2.4 Carbon measurements

In 2004, the District began measuring EC and OC on PM₁₀ filters. Before 2004, these were only measured in special studies, but the District had a very long record of another measurement, Coefficient of Haze (COH). COH is an optical measurement that is highly correlated with carbon, EC in particular. A period of overlap in EC and COH measurements facilitated the development of a method to estimate EC concentrations from COH measurements. This allowed the extension of EC measurements back to 1990.

3. PM standards

The set of PM standards is complex. There are both national and California standards for both PM_{2.5} and PM₁₀ for both the annual average and for 24 hours (with a couple of exceptions). Here is a table with the level of the various standards.

Table 3. Levels of national and State PM standards

	PM _{2.5}		PM ₁₀	
	annual	24-hour	annual	24-hour
National	15.0 µg/m ³	35 µg/m ³	N/A	150 µg/m ³
California	12.0 µg/m ³	N/A	20 µg/m ³	50 µg/m ³

Differences among the standards and how they are met are explained below.

3.1 National PM standards

The annual PM_{2.5} national ambient air quality standard (NAAQS) is 15.0 µg/m³. Compliance with the standard is determined as follows. The standard is met by the Bay Area if and only if it is met at every Bay Area site collecting EPA-approved PM_{2.5} measurements. For a given site, the standard is met if its annual design value does not exceed the standard. The *annual design value* is the 3-year average of annual averages of the site’s PM_{2.5} measurements, where each annual average is the mean of the four quarterly averages³. For example, the 2010 annual design value would be the average of the 2008, 2009 and 2010 annual averages.

The 24-hour PM_{2.5} NAAQS is 35 µg/m³. Compliance with the standard is determined as follows. The standard is met by the Bay Area if and only if it is met at every site collecting EPA-approved PM_{2.5} measurements. For a given site, the standard is met if its 24-hour design value does not exceed the standard. The *24-hour design value* is the 3-year average of annual 98th percentiles of the site’s PM_{2.5} measurements⁴.

There is also a 24-hour PM₁₀ NAAQS of 150 µg/m³. This standard is met if the 3-year average of the 99th percentile values at each site is less than or equal to 150 µg/m³.

Additional details of the standards are given in the Code of Federal Regulations, Title 40, Part 50, Appendix N. (CFR 40, 2011)

3.2 California PM standards

California standards are also based on 24-hour PM_{2.5} and PM₁₀ measurements, but differ from the NAAQS in both level and form. Like the PM NAAQS, the California annual standards are based on 3 years of data. But the

³ Each quarter must have measurements for 75% or more of the scheduled samples.

⁴ The 98th percentile is simply the sample 98th percentile if the sampling schedule has the same frequency year round (for example the 8th highest value if sampling is every-day). If the sampling frequency differs by season (as it has with many Bay Area sites), the 98th percentile is the determined from the empirical distribution. (See US EPA 1999, page 27).

California standards are violated if the annual average exceeds the standard in any of the 3 years. The California annual $PM_{2.5}$ standard is $12 \mu\text{g}/\text{m}^3$. Currently, California has no 24-hour $PM_{2.5}$ standard.

The California annual PM_{10} standard is $20 \mu\text{g}/\text{m}^3$. Its 24-hour PM_{10} standard is $50 \mu\text{g}/\text{m}^3$. Unlike the 24-hour PM_{10} NAAQS, the 24-hour California PM_{10} standard is violated if any PM_{10} measurement exceeds it, with certain exceptions. Among the exceptions are values that are expected to occur less than once per year on average. These are determined by a statistic termed the *expected peak day concentration* (EPDC), computed from a year's worth of data. Any measured value above a site's EPDC is excluded. The highest non-excluded value is the site's design value for that year. A site meets the standard if none of its past three year design values exceed the standard.

4. Statistical methods used for trend analysis

For both the national annual and 24-hour PM_{2.5} standards, the design value is a 3-year average – the average of annual means for the annual standard, the average of 98th percentiles for the 24-hour standard. Thus, the focus of this analysis is on methods to estimate trends in these averages, and also to estimate the uncertainty in these trend estimates.

Let t represent years from the initial year, so that, for example, for PM_{2.5} as the years range from 1999, 2000, ..., 2010, t ranges from 0, 1, ..., 11. Let Y_t be a summary statistic for year t for a given site. Here Y_t is either the annual mean or the 98th percentile for year t .

Some assumption about the form of the trend is necessary for statistical inference. For this analysis, we assume

$$Y_t = m_0(1-r)^t e_t \quad (4.1)$$

where m_0 is the initial mean [symbolically, $E(Y_0)$], r is the annual rate of reduction and e_t is a random error term with mean 1.

There is a plethora of ways to estimate trend. These include:

1. Regression vs. year: Y_0, Y_1, \dots, Y_k against $t = 0, 1, \dots, k$
2. Regression on the log scale: $\ln(Y_0), \ln(Y_1), \dots, \ln(Y_k)$ against $t = 0, 1, \dots, k$.
3. Comparing the 3 most recent years with the earliest 3 years: the absolute reduction, $\bar{Y}_R - \bar{Y}_E$, or the relative reduction, $1 - \bar{Y}_R/\bar{Y}_E$, where $\bar{Y}_R = \frac{Y_k + Y_{k-1} + Y_{k-2}}{3}$ and $\bar{Y}_E = \frac{Y_0 + Y_1 + Y_2}{3}$
4. Using a permutation approach: Let $W_t(r) = Y_t/(1-r)^t$. Find the value of r so that $\bar{W}_R(r) - \bar{W}_E(r) = 0$, or $1 - \bar{W}_R/\bar{W}_E = 0$.
5. A nonparametric approach, e.g., using Kendall's tau: $\tau(Y_1, Y_2, \dots, Y_k) = \sum_{s=1}^{k-1} \sum_{t=s+1}^k \text{Sgn}(Y_s < Y_t)$, where $\text{sgn}(x)$ is the sign function, $\text{sgn}(x) = -1$ if $x < 0$, $=0$ if $x = 0$ and $=1$ if $x > 0$. Find the value of r so that $\tau(W_1(r), W_2(r), \dots, W_k(r)) = 0$.

All these were used, in different parts of the analysis.

4.1 Focus of trend analysis

This analysis focuses on progress toward meeting the federal PM standards. For PM_{2.5} and PM₁₀ the statistical tests applied used design values from the most recent and earliest years available. For PM components, the tests used means for the most recent and earliest years.

4.2 Inference for ratios -- permutations

In light of this discussion, we consider inferences about the design values, \bar{Y}_R and \bar{Y}_E . Statistical inference for the difference in design value for a given site is straightforward – using a two-sample t-test. But to report progress as

a difference of so many micrograms/cubic meter is difficult for a non-expert to understand and, further, the degree of progress will depend on the typical levels, which vary from compound to compound.

The alternative it to use the relative progress, that is, the relative reduction, \bar{Y}_R/\bar{Y}_E . The interpretation of the relative reduction is straightforward and it makes it possible to compare the reductions of various PM components. The statistical inference is not as straightforward, however. A parametric approach is possible, one that assumes the sample means are normal. But unlike the inference for the difference in means, inference on the ratio involves a distribution with unknown parameters that, because of the small samples, can only be roughly approximated.

Here we use a non-parametric approach based on the permutation distribution. Assume that the data fit model (4.1) above. Then, if the true reduction rate is r , the set of W_t , where $W_t = Y_t/(1-r)^t$, will be identically distributed. Therefore, if we compute $D(r) = \bar{W}_R/\bar{W}_E$, with high probability it will be within the range of the ratios $D^*(r) = \bar{W}_R^*/\bar{W}_E^*$ where the W_t^* are random permutations of the W_t , that is, permuting the values among different years.

A confidence interval for r can thus be found as the set of r such that $D(r)$ is not unusual among the permuted values $D^*(r)$. More precisely, the set $\{r: r_l \leq r \leq r_u\}$ such that $P(D^*(r_l) \leq D(r_l)) = \alpha/2$, and $P(D^*(r_u) \geq D(r_u)) = \alpha/2$ is a $100(1-\alpha)\%$ confidence interval for r .⁵

4.3 Inference for District design values

The design values consist of 3-year averages, so the methods discussed above are appropriate for the design values for individual sites. But the design value for the District presents a more difficult problem. As the maximum of 3-year averages, it does not have the distribution of a simple average.⁶ One inference method, using regression against year, is described in Appendix A.

Here we use a generalization of the ratio-permutation method. Effectively, what that method does is to consider permutations of years. This same approach can be applied to the District design value. See Appendix A.

4.4 Methods for trends in PM_{2.5} exceedances, including meteorological adjustment

We also consider trends in the number of exceedances of the 24-hour PM_{2.5} standard. Exceedances occur almost exclusively in the winter months. Therefore, we limited the analysis to November through February. Because the high PM season overlaps years, we considered the number of exceedances by winter, rather than by year.

⁵ For the PM_{2.5} case, we are using six W_t 's to compute the two means. With 12 years, there are a total of $\binom{12}{3}\binom{9}{3}=18,480$ distinct ways to pick the six, so there will be at most 18,480 distinct values of $D^*(r)$, one of which is, of course, $D(r)$. If we chose a 90% confidence interval, then $\alpha/2 = 0.05$. Since $0.05 \times 18480 = 924$, we want to find r_l so that $D(r_l)$ is the 924th smallest of the $D^*(r_l)$ values, and r_u so that $D(r_u)$ is the 924th largest of the $D(r_u)$ values. We wrote a program that found the confidence interval iteratively – guessing values of r and, for each guess, cycling through all the possible permutations and computing $D^*(r)$ for each – until the value of $D(r)$ was close to the value of the 924th smallest $D^*(r)$; then repeating the process until the value of $D(r)$ was close to the value of the 924th largest $D^*(r)$.

⁶ In contrast, the 3-year average of yearly maxima would have the distribution of averages.

Exceedances are relatively infrequent events, so their distribution can be reasonably modeled by the Poisson distribution. Poisson regressions against year were performed.

Exceedances are highly dependent on meteorology, leading to large year-to-year fluctuations. Thus a key question is how much of any trend is due to changes in emissions and how much due to a trend in meteorology. A measure was developed that provides a rough first estimate of the meteorological potential, namely the number of light wind, no rain days. Specifically, we averaged the 24-hour average wind speeds at three sites: Bethel Island, Pleasanton and San Carlos. A light wind day was defined as one with an average wind speed less than 5 mph. We used rainfall from San Jose, and defined a “no rain” day as one with less than 0.02 inches. Because the highest $PM_{2.5}$ tends to occur only after several days of buildup, we defined a *PM_{2.5} conducive day* as a day where there had been three or more days in a row with both light winds and no rain up to and including that day.

We also performed a more comprehensive analysis using multiple regression to predict daily $PM_{2.5}$ concentrations from meteorological and temporal variables. The goal was to estimate how many exceedances would have occurred each year if the meteorology was the same every year. The methodology is explained in Appendix B.

4.5 Annual decreases for comparability

Because the number of years available for trend analysis varied for the different PM species, the *annual* reduction rates are also provided, where the annual rate is defined as $r = 1 - (\bar{Y}_R/\bar{Y}_E)^{1/(k-1)}$.

4.6 Estimation of missing values

In looking at trends in annual statistics across sites, ignoring these missing data could result in biases. To reduce the potential bias, an ANOVA model was fit to the data, namely

$$Y_{ij} = \mu + \alpha_t + \beta_j + e_{ij}$$

where Y_{ij} is an annual statistic for site j in year t . The missing Y_{ij} , were estimated iteratively 1. computing $\bar{Y}_{t.}, \bar{Y}_{.j}$ and $\bar{Y}_{..}$, where the dot indicates which index over which the average is taken, then 2. estimating the missing values by $\bar{Y}_{t.} + \bar{Y}_{.j} - \bar{Y}_{..}$. Steps 1 and 2 were repeated until the estimates for missing values converged.

4.7 Statistical inference issues

For all of the above techniques, statistical independence of the Y_t 's is required to estimate uncertainty. Although $PM_{2.5}$ values can be highly correlated one day to the next, this correlation decays within a few days. Thus, annual means from different years, which have at most a very few measured values within a few days, should be close to statistically independent of each other. Successive-year 98th percentiles may be correlated if a stagnant period occurs right around January 1, but it is unlikely to affect many years. Thus, day-to-day autocorrelation should not invalidate the inferences made.

Some factors do have a potential impact, however. One was a change of instrumentation. In 2009, continuous FEM measurements replaced filter measurements. Analysis where parallel measurements were made showed that

the FEM measurements were higher by a random amount up to $3 \mu\text{g}/\text{m}^3$ (Fairley 2011). Because of this, FEM measurements were not included in the $\text{PM}_{2.5}$ trend analysis.

Another potential factor is the moving of a station. A couple of key stations were moved – the downtown San Jose site was moved from 4th Street to Jackson Street in 2002. The Livermore site was moved in 2000. Based on comparisons using other sites as controls, it appears these changes didn't have much impact on measured PM concentrations.

The same concerns applied to $\text{PM}_{2.5}$ measurements made by other agencies – the $\text{PM}_{2.5}$ filter measurements made for CRPAQS, for the STN instrument at San Jose, and for the DRI measurements at Livermore, Oakland and Vallejo. These measurements were also excluded.

The degree of uncertainty will be expressed as 90% confidence intervals.

5. PM_{2.5} trends

This section presents trends in annual mean and 98th percentile PM_{2.5} design values, and also in numbers of exceedances of the 24-hour PM_{2.5} standard.

5.1 Annual trends

Figure 5.1 shows the quarterly averaged annual means for District sites and Point Reyes. Also shown is a 3-year average of BA urban levels. The average PM_{2.5} across the District sites shows a steady reduction from 12.2 µg/m³ for 1999-2001 to 8.6 µg/m³ for 2008-2010, a reduction of 30%, or a 3.9% per year compounded rate with a 90% confidence interval of 2.1% to 5.6%.

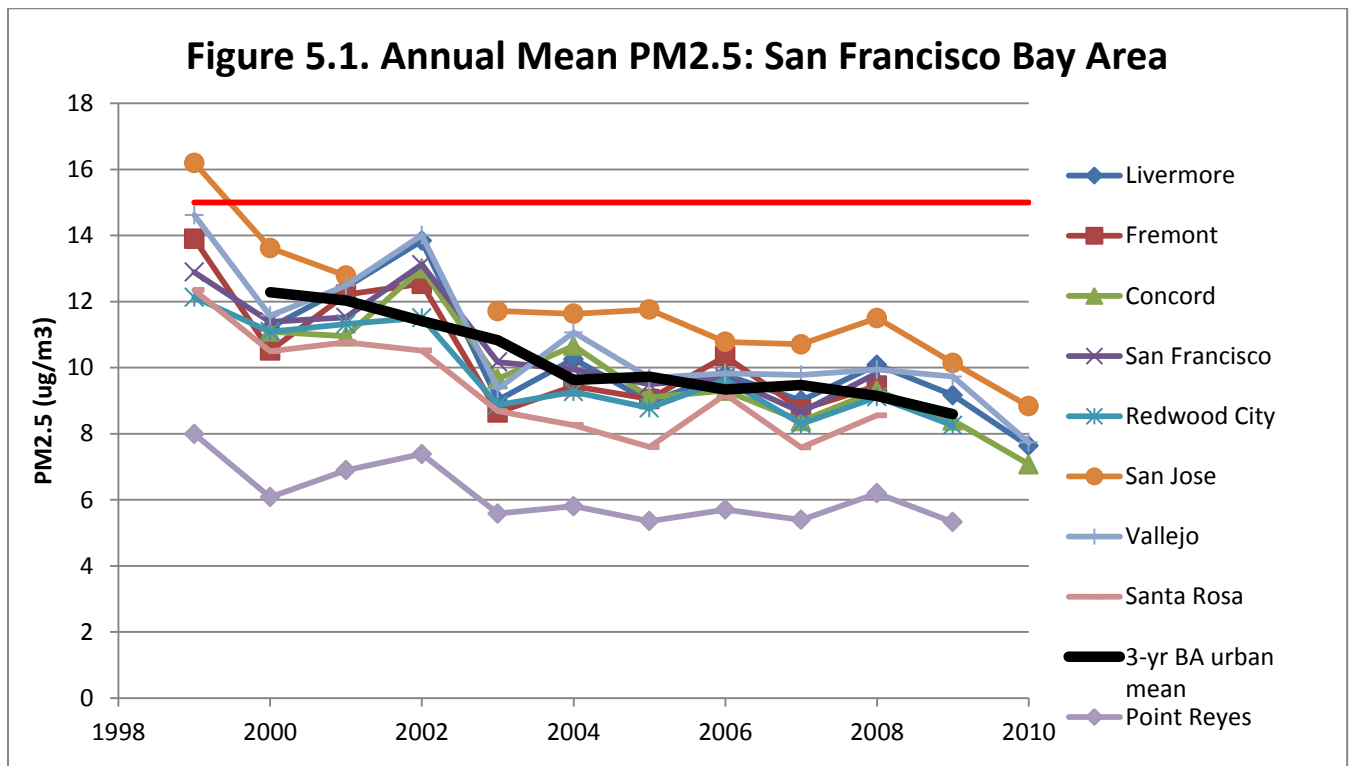
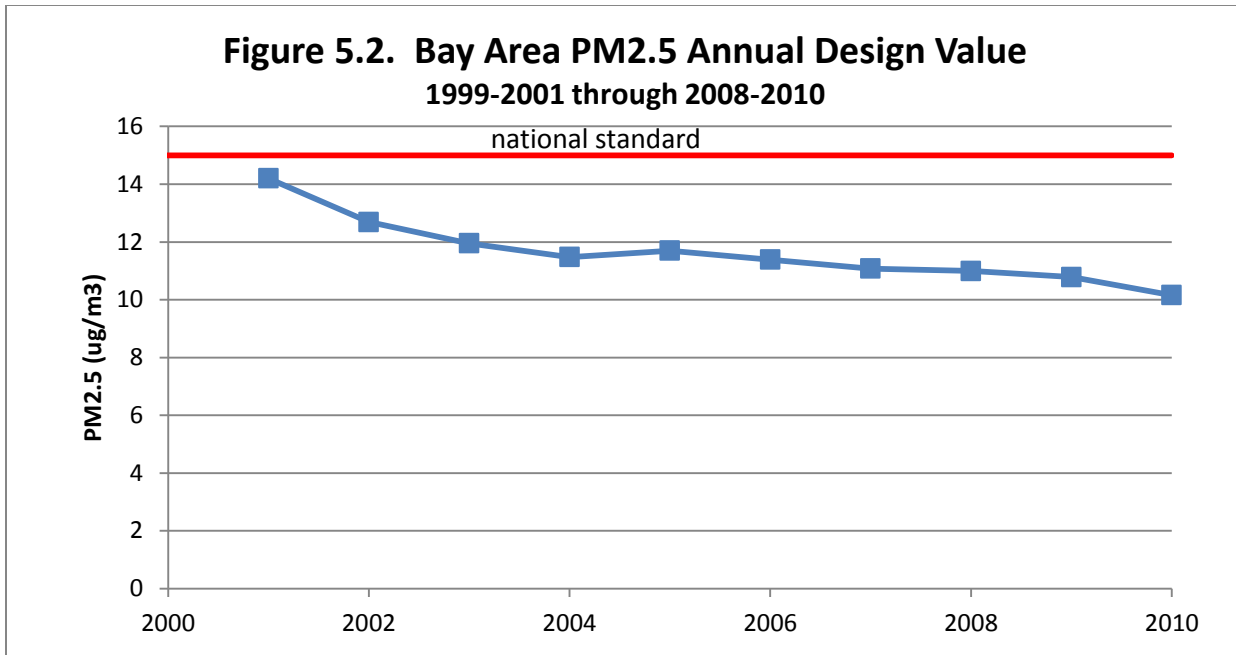


Figure 5.2 shows the District design values relative to the annual standard from 1999-2001 through 2008-2010. There has been an almost monotonic reduction, from 14.2 µg/m³ for 1999-2001 to 10.2 µg/m³ for 2008-2010, a reduction of 28% or 3.7% per year, with a 90% permutation confidence interval 1.7% to 5.1%.

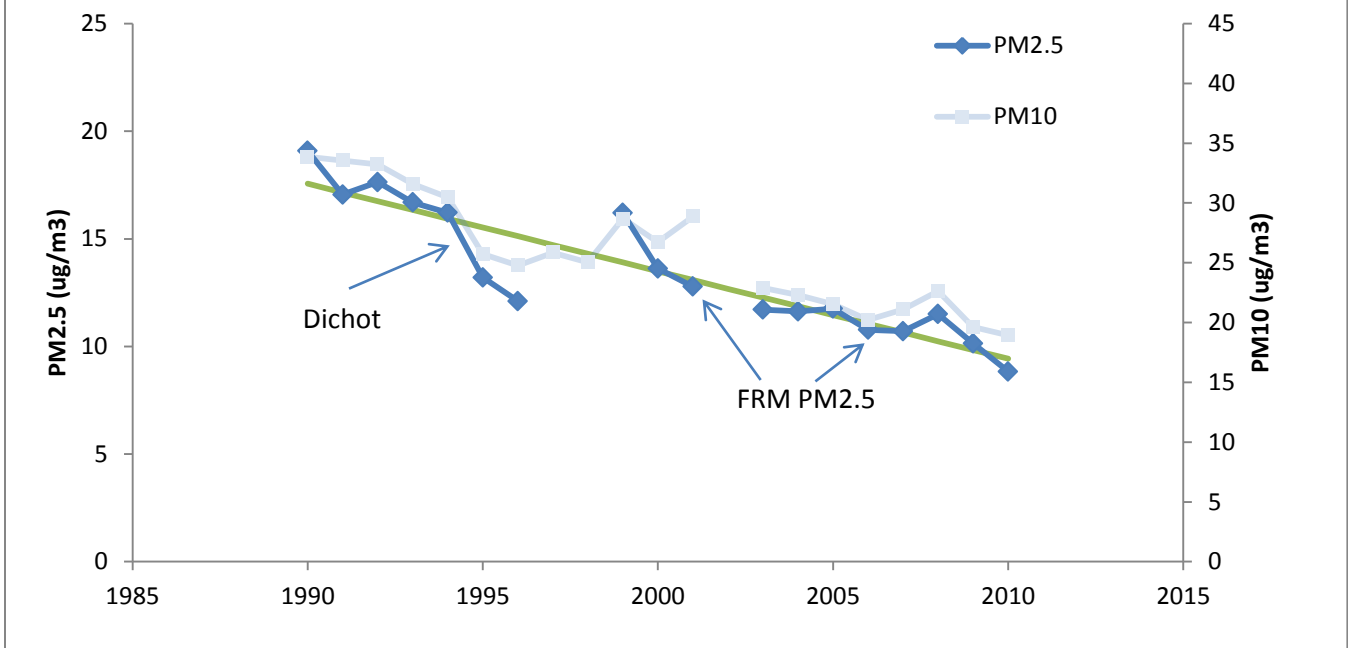


5.1.1 Longer-term trends

There were only limited PM_{2.5} data collected in the Bay Area before 1999. There is a continuous record from Point Reyes extending back into 1988, but it is of limited value being a background site. PM_{2.5} data were collected at Fremont from 1988 through 1994 as part of the National Acid Deposition Network, but Fremont is not one of the sites where routine PM_{2.5} measurements have been made. At San Jose, however, PM_{2.5} measurements were collected with a dichotomous sampler from 1990 through 1997, and also since 1999 as part of the District's routine monitoring network.

Figure 5.3 shows the annual concentrations for San Jose, where the dichot measurements for 1992-97 have been divided by 0.78 according to the analysis in Appendix D. Also shown are the PM₁₀ annual averages from the District FRM sampler. The relationship between annual mean PM_{2.5} and PM₁₀ has held fairly constant over time. A regression line shows that there has been a consistent trend in PM_{2.5} over the 20-year period from 1990 through 2010, a reduction from 17.9 µg/m³ in 1990-1992 to 10.1 µg/m³ in 2008-2010, representing a 43% decrease.

**Figure 5.3. Annual PM2.5 Concentrations
San Jose 1990-2010**



5.2 Trends in 24-hour design value

Figure 5.4 shows the annual 98th percentile concentrations at District sites. It also shows a 3-year moving average of the average of all sites plotted vs. middle year. The plot shows a sudden dip in 2003, but some of this is because the meteorology was much less conducive to high PM in 2003 than 2002. The 3-year average dropped from 47 $\mu\text{g}/\text{m}^3$ for 1999-2001 to 27 $\mu\text{g}/\text{m}^3$ for 2008-2010, a 41% reduction, or 5.7% per year.

Figure 5.4. PM2.5 98th Percentiles: San Francisco Bay Area

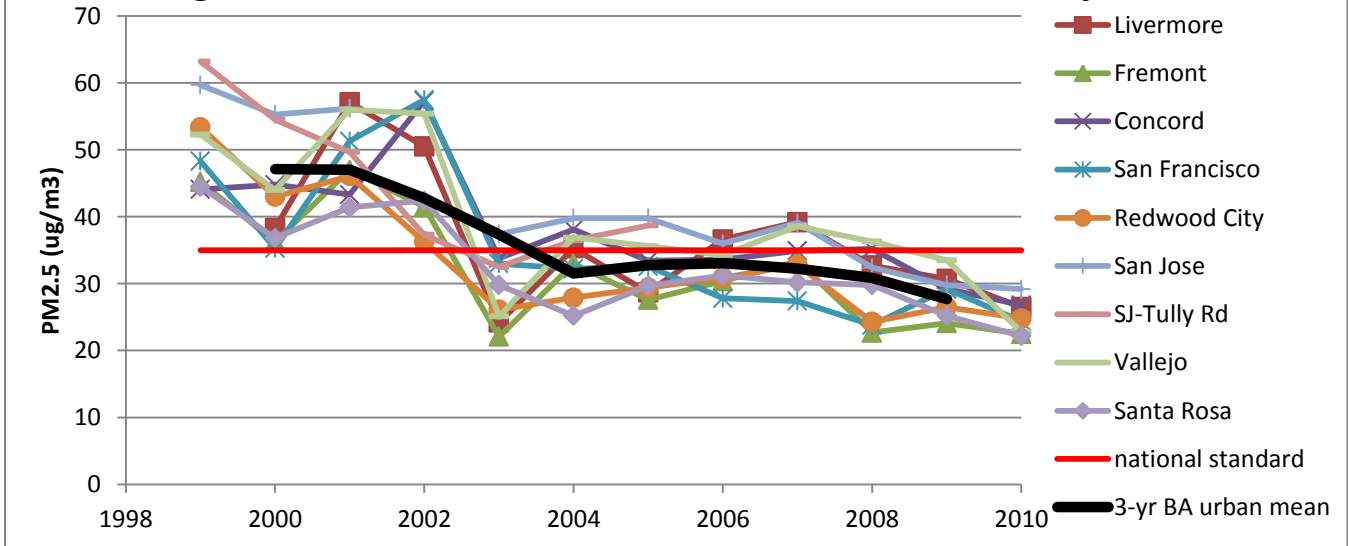
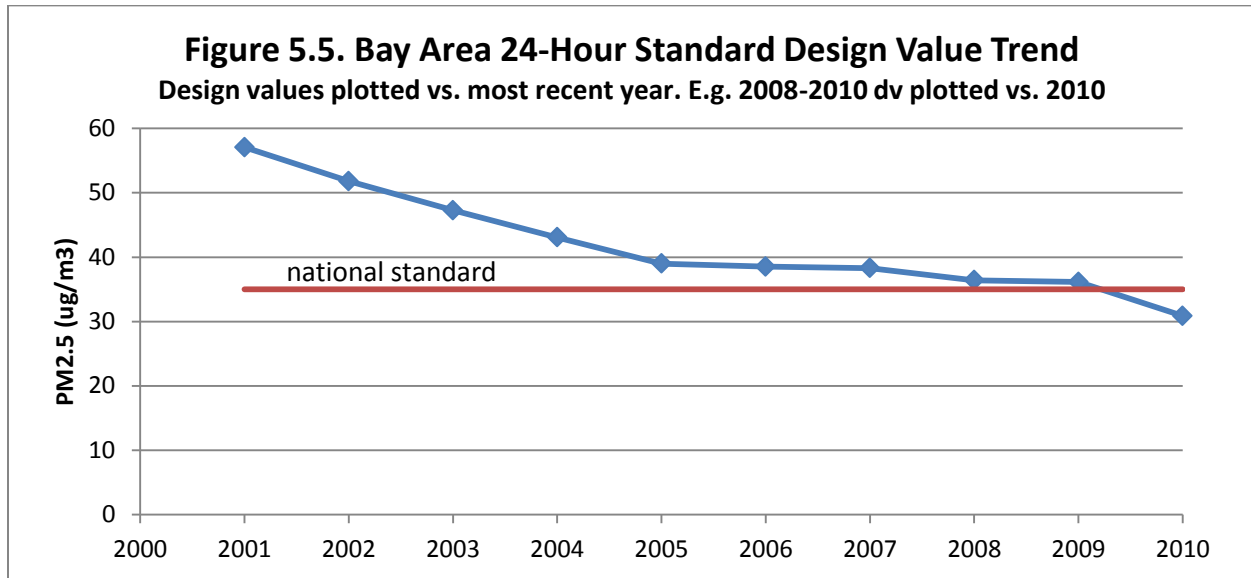


Figure 5.5 shows the trend in the Bay Area 24-hour design values. The values fell almost monotonically, from 57 $\mu\text{g}/\text{m}^3$ for 1999-2001 to 30 $\mu\text{g}/\text{m}^3$ for 2008-2010, a reduction of almost 50%. The annual reduction was 6.7% per year with a 90% confidence interval of 5.2% to 8.4%.



5.3 Trends in winter PM_{2.5}

As discussed earlier, virtually all of the Bay Area’s highest PM_{2.5} concentrations occur on winter days. These days are distinctive not only in their meteorology, but also their mix of PM_{2.5} – woodsmoke and ammonium nitrate predominating. Thus, we focus on wintertime trends, more precisely the trends from November through February. For this purpose, we look at trends in counts of exceedances by winter.

The number of exceedances per winter is highly variable because of meteorological fluctuations. A simple method to track the meteorological effect is to compare the number of exceedances with the number of PM conducive days. A more sophisticated analysis was done to account for meteorology, where formulas were established for each winter that related daily maximum PM_{2.5} to various weather and time-of-year factors. Such formulas represent estimates of the potential for high PM_{2.5} under given meteorological conditions, so applying each of the formulas across the same set of winter conditions allows us to look at the trend in the average potential, which we will term *Average Exceedance Potential*.

5.3.1 Trends in PM_{2.5} exceedances

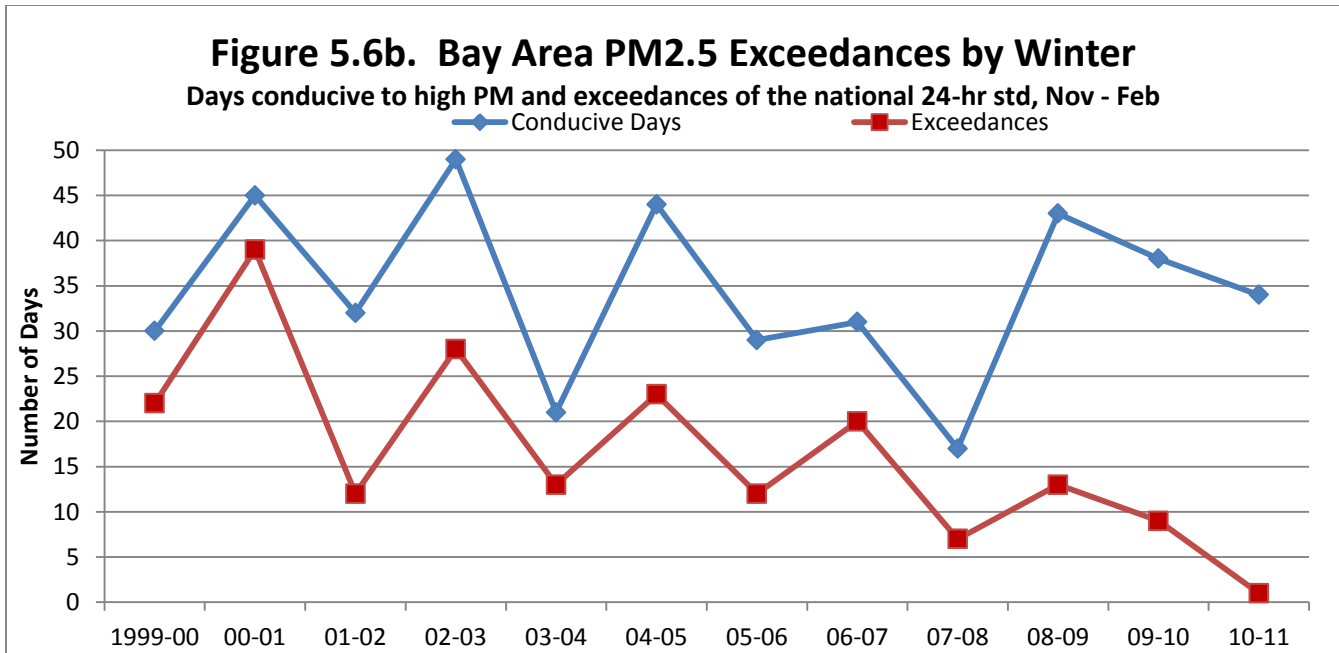
Figure 5.6a shows the number of days PM_{2.5} exceeded the 35 $\mu\text{g}/\text{m}^3$ standard for each winter 1999-2000 through 2010-2011. The sawtooth pattern in the numbers of exceedances is mostly due to meteorology rather than changes in emissions. A reduction in the numbers of exceedances is clear, and the Poisson regression line shows a reduction from a predicted 30.6 exceedances in 1999-2000 to 7.4 exceedances in 2010-2011.

Figure 5.6a. Bay Area PM_{2.5} Exceedances by Winter
 Exceedances of the 24-hour NAAQS and Poisson regression fit



5.3.2 PM_{2.5} conducive days

Figure 5.6b shows the number of PM_{2.5} conducive days along with the number of exceedances. Although not perfectly correlated, the PM_{2.5} conducive days generally follow the same sawtooth pattern as the PM_{2.5} exceedances. However, unlike the decline in exceedances, there is no long-term trend in conducive days; there are actually more conducive days in the winters of 2008-2009 through 2010-2011 than in 1999-2000 through 2001-2002.



5.3.3 Evaluating exceedance potential

To estimate potential, we began with the District-wide daily maximum PM_{2.5} concentrations.

The independent variables include a range of meteorological measurements made at various District sites. These were reduced to 24-hour summaries like 24-hour averaged wind speed and 24-hour resultant wind direction. We also considered the 24-hour rainfall at the City of San Jose meteorological station and measurements from the Oakland RAOB sounding⁷. In addition we considered the time lags of some of these variables. We also included temporal variables: day of week; holidays; and time of year, specifically, the number of days from January 1st. See Appendix B for methodology.

Results

A set of trial variables was selected. (See Appendix B.) Stepwise regressions were performed and the strong predictors – those with small p-values – were selected. The resulting regression formula was

$$f = \ln(\text{dmax}) = 1.79 - 0.00819 \text{ jd} - 0.0263 \text{ aveys} + 1.41 \text{ 1/aveys} + 0.202 \text{ sjrn}<.02 + 3.77 \text{ 1/3dyws} + 0.250 \text{ pbwd}<90 - 0.00118 \text{ fuws}\&\text{pb}<90$$

The coefficient values, definitions, uncertainties and t-values are shown in Table 5.1.

⁷ The RAOB soundings measurements made twice a day (4am and 4pm) at Oakland Airport using instruments on a balloon. The measurements are recorded as the balloon reaches certain air pressure levels.

Table 5.1. Predictors for District maximum PM_{2.5}.

	Predictor	Description	Estimated Coefficient (a _i)	SE Coef	t	p-value
x ₀	Constant		1.790	0.110	16.33	0
x ₁	jd	# of days from Jan 1	-0.00819	0.00063	-12.96	0
x ₂	avews	3-site average of 24-hour ave ws	-0.0263	0.0087	-3.04	0.002
x ₃	1/avews	reciprocal of x ₂	1.409	0.300	4.70	0
x ₄	sj rain<.02	indicator of san jose rain < 0.02 inches	0.2017	0.0283	7.12	0
x ₅	1/3dayavews	reciprocal of 3-day average wind speed	3.768	0.221	17.07	0
x ₆	san pablo wd<90°	indicator of san pablo wind direction between 0° and 90°	0.2504	0.0391	6.41	0
x ₇	fuws&pb<90	standard deviation at fort funston times x ₆	-0.00118	0.00037	-3.16	0.002

The adjusted R² was 65.9%.

A plot of the residuals vs. predicted shows a relatively even scatter, indicating no serious problems with heteroskedasticity (non-constant variance). The first order autocorrelation of the residuals was 0.39. This is statistically significant, but small enough that the variables in Table 5.1 should all still be statistically non-zero.

Regressions were run for each individual winter using these same 7 variables. The adjusted R²s ranged from 57% to 82%. The variables jd and 1/3dyws are consistently significant from winter to winter, the others intermittently so. The predictions based on all variables were used. So, for each winter,

$$f_w = a_{0w} + a_{1w} * x_1 + a_{2w} * x_2 + \dots + a_{7w} * x_7$$

These formulas were then applied to each of the winters 1999-2000 through 2010-2011 (Steps 3 and 4 above). The results are shown in Table B2 in Appendix B. Figure 5.7 compares the predicted number of exceedances in winter w with the actual number in winter w. There is a high correlation between the predicted number of exceedances and the actual number. The closeness suggests that the Gaussian error assumption works well.

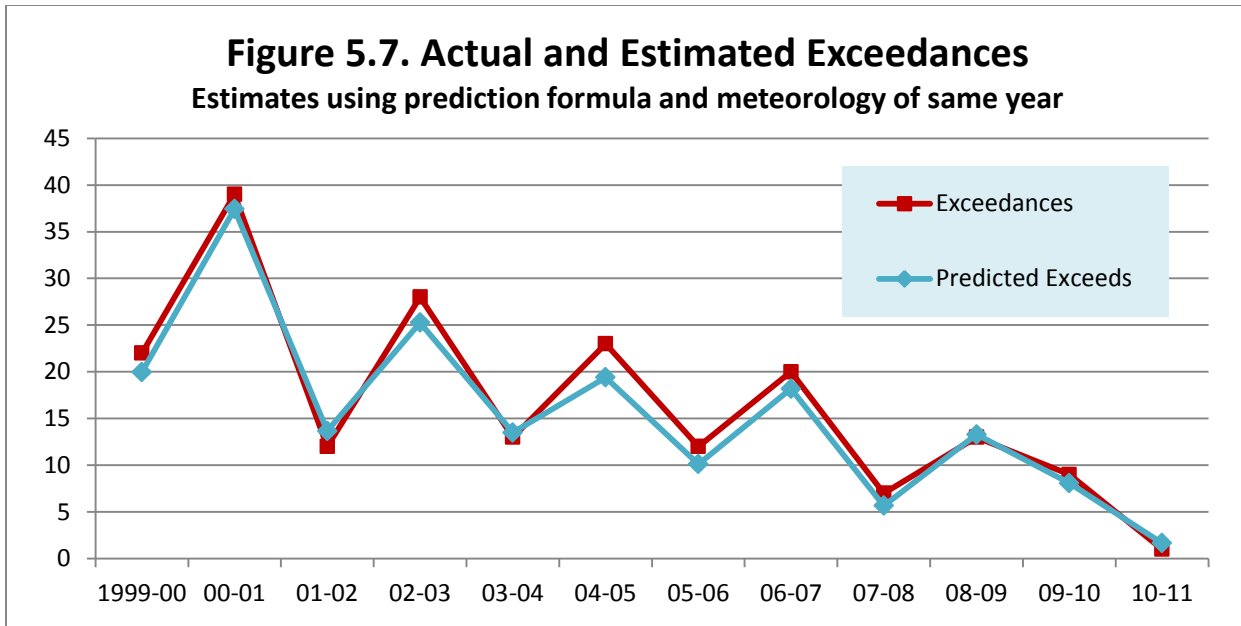
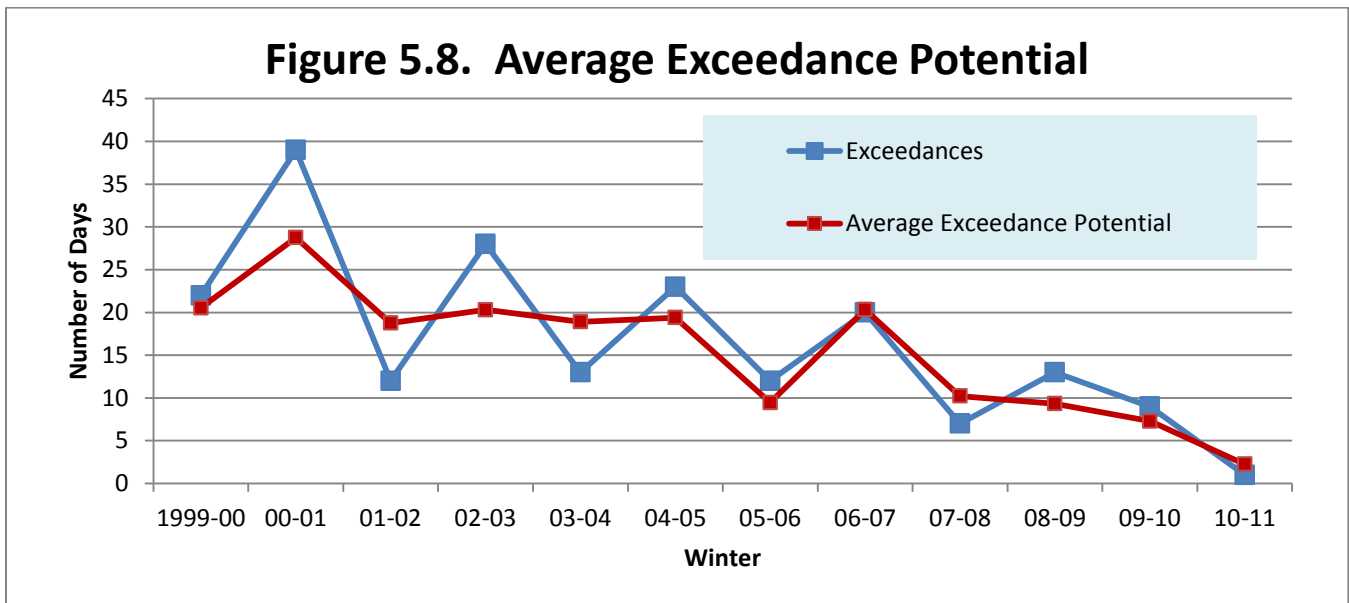


Figure 5.8 shows the estimated average exceedance potential for each winter. In most winters, this potential is closer to a linear trend line than the observed number of exceedances. Whereas the correlation of actual exceedances with year is 0.72, the correlation of exceedance potential with year is 0.86, reinforcing the conclusion that the observed trend is real. A regression of log of exceedance potential vs. year has a slope of -0.16, indicating an annual rate of decrease in the number of exceedances of $1 - e^{-0.16} = 0.15$, i.e. 15% per year.



6. PM₁₀ trends

For this trend analysis, we consider the period from 1989 through 2007 for peak PM₁₀ and 1989-2010 for mean PM₁₀. In order to establish consistency across years, for the peak PM₁₀ analysis, we only considered days where measurements were made at least 8 sites. As the table in Appendix C shows, the number of such days is relatively consistent across years.

6.1. Trends in mean PM₁₀

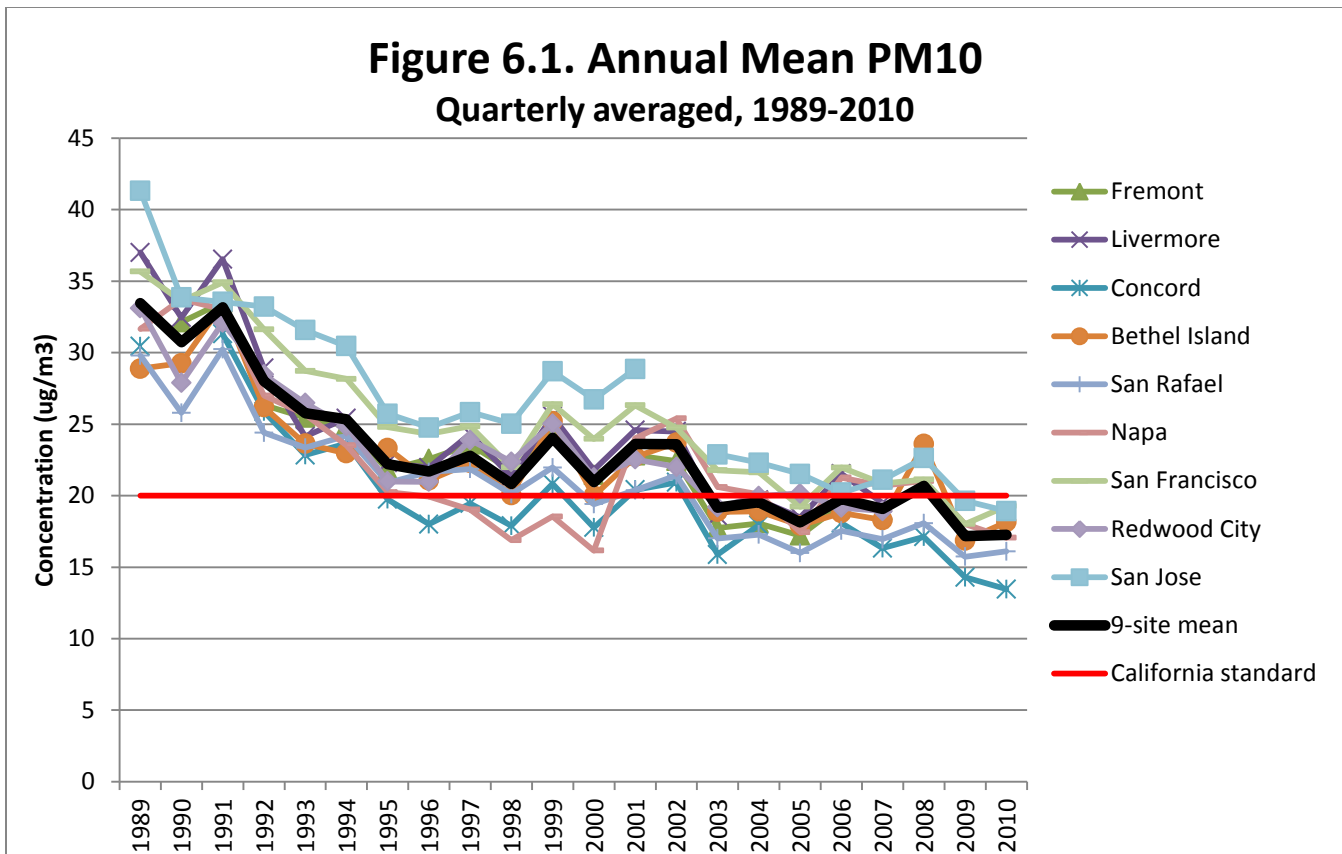
Figure 6.1 shows quarterly averaged annual PM₁₀ concentrations from 1989 through 2010. The solid red line shows the average of 9 sites, where missing values were filled in by repeatedly fitting a 2-way ANOVA and replacing the missing values with the 2-way ANOVA estimates.

The reductions were approximately 3% per year in the 1990s and 2% per year from 2000 through 2010, both statistically significant. The 9-site average dropped from 33 µg/m³ in 1989-1991 to 18 µg/m³ in 2008-2010, a reduction of 44%.

The District is close, but does not meet, the California annual PM₁₀ standard. Its 18 µg/m³ design value seems to indicate it met the standard, but it hasn't yet, because first the standard is based on the maximum average over the most recent 3 years, and second the California standard is based on PM₁₀ measured locally, which is generally somewhat higher. Bethel Island registered the highest annual average of 24 µg/m³ in 2008. However, the maximum average since then was 20.4 µg/m³ for San Jose in 2009.⁸

The 2008 maxima were impacted by wildfires. Assuming that these don't recur or that wildfire-influenced concentrations are excluded as exceptional events, then the District should reliably meet the California annual standard in a very few years.

⁸ California PM₁₀ standards are based on local PM₁₀ concentrations, using local temperature and pressure, whereas national PM₁₀ uses a standardized temperature and pressure. California values were not available for the whole period but were available for the most recent data. The local and measurements are very highly correlated and are 1% to 3% higher than the standardized. Both the California PM_{2.5} standard and the PM_{2.5} NAAQS are based on local measurements.



6.2. Trends in peak PM₁₀

Figure 6.2 shows the number of exceedances of the California 24-hour PM₁₀ standard. It also shows the predictions from a Poisson regression. The reduction in exceedances has been considerable. The Poisson regression shows a statistically significant decrease from almost 15 days per year in 1989 to 3 days per year in 2007, a reduction of approximately 8% per year.

Figure 6.3 shows the trend in District design values relative to the California standard. There has been a significant reduction over the entire 1989-2007 period. The trend in the maximum design value from 2000 to 2007 is not statistically significant and the downtrend in the 6-site average is borderline significant.

The overall drop, from an average maximum design value of 157 $\mu\text{g}/\text{m}^3$ for 1989-1991 to 78 $\mu\text{g}/\text{m}^3$ for 2005-2007 was 50% for an annual rate of 4.3% per year (90% confidence interval 2.4% to 6.5%). Nevertheless, the District still violates the standard by a considerable margin.

Figure 6.2 Exceedances of the California PM10 Standard, 1989-2007
 Days exceeding the 24-hour standard with measurements at at least 8 sites

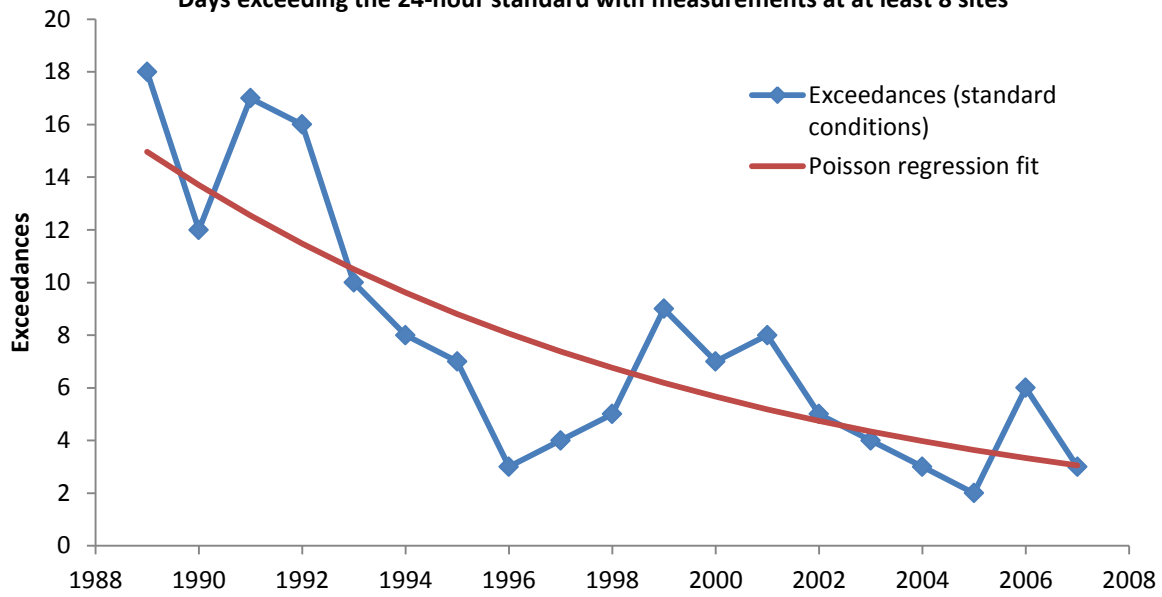
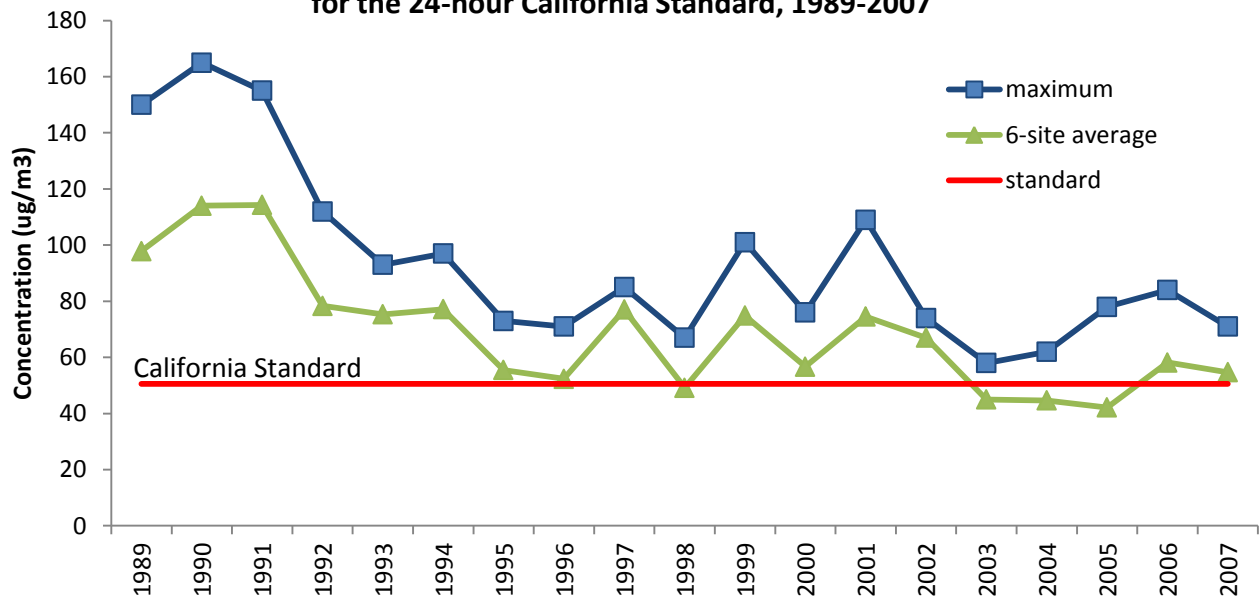


Figure 6.3. Bay Area PM10 Design Values for the 24-hour California Standard, 1989-2007

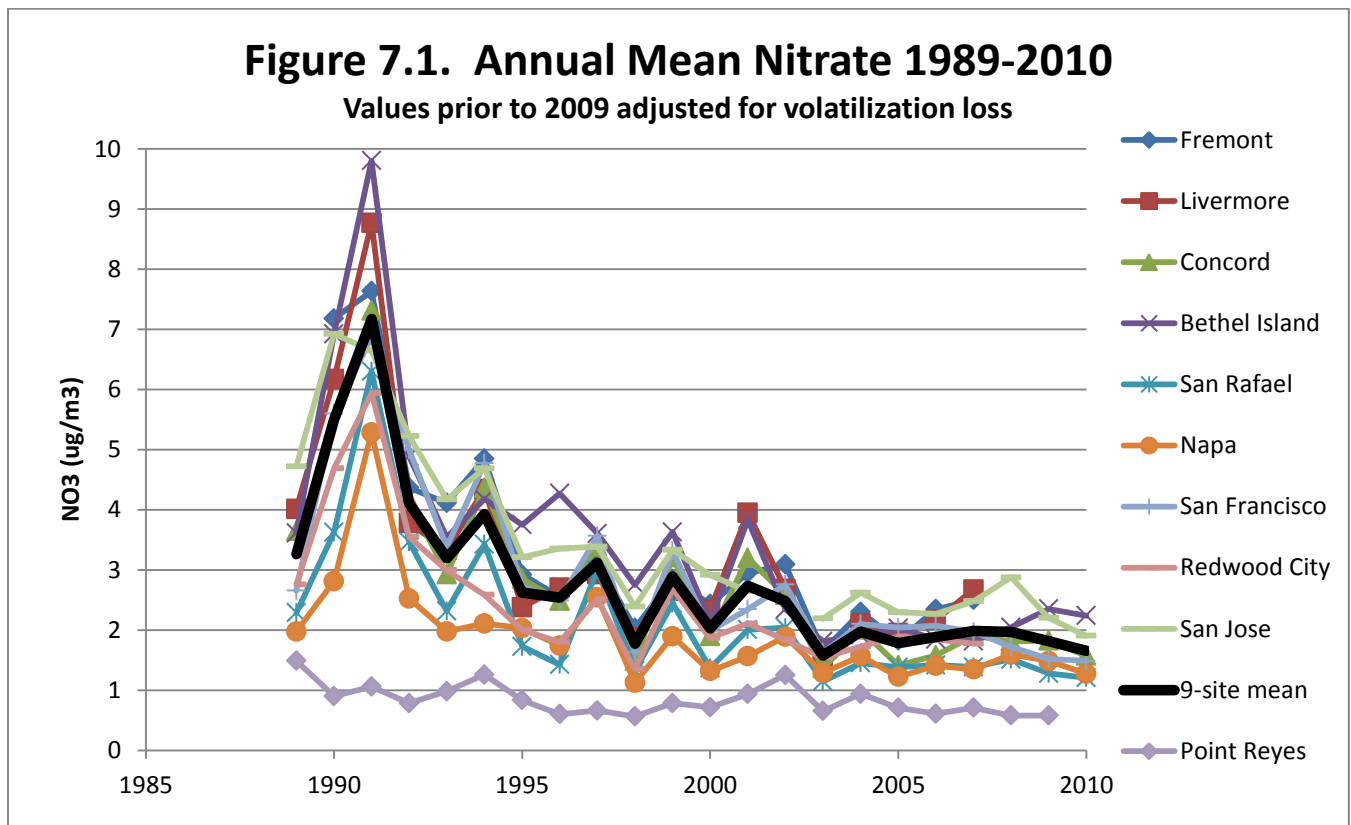


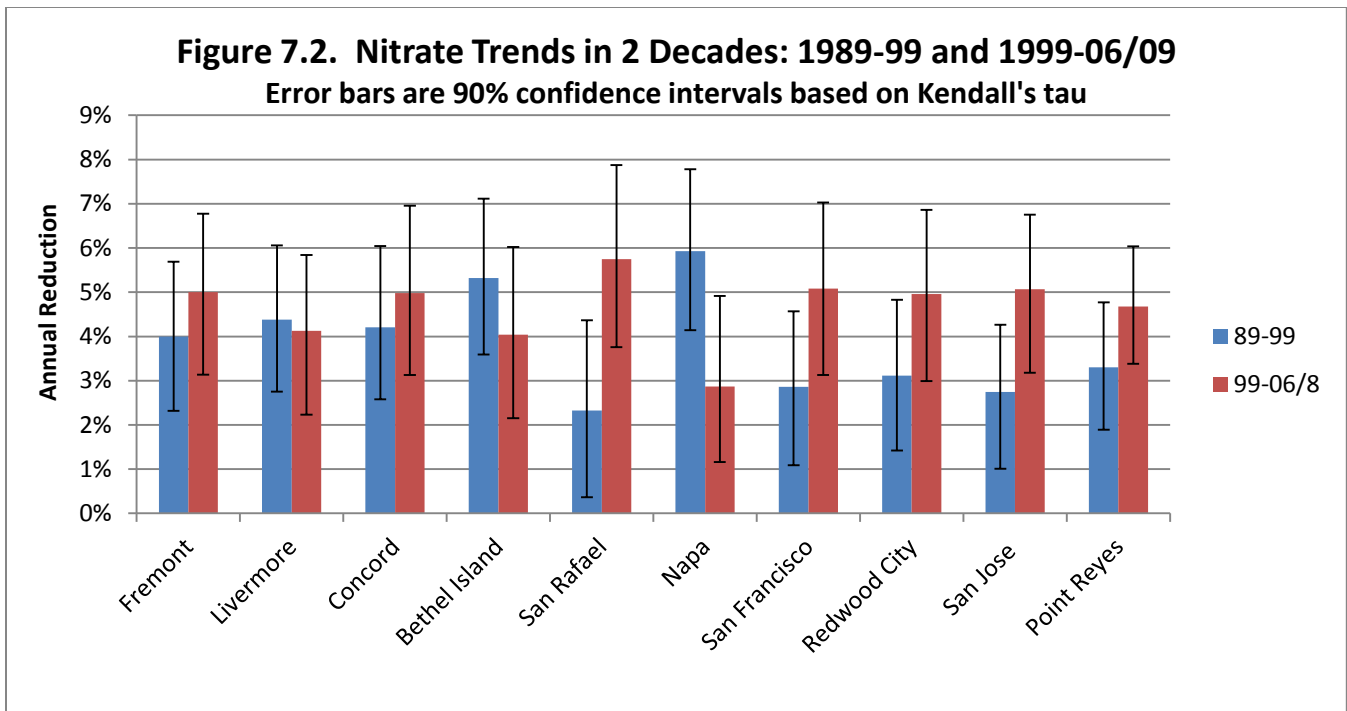
7. Trends in PM components

7.1. Nitrate trends

Figure 7.1 shows the trends in annual mean nitrate concentrations. All sites show a downtrend although there is considerable year-to-year variation. The 9-site mean nitrate was reduced from 5.3 $\mu\text{g}/\text{m}^3$ in 1989-1991 to 2.6 $\mu\text{g}/\text{m}^3$ in 1999-2010 to 1.8 $\mu\text{g}/\text{m}^3$ in 2008-2010. These constitute reductions of 52% from 1990 to 2000, 29% from 2000 to 2009, and 66% from 1990 to 2009, an overall reduction of nearly two-thirds (a 90% confidence interval of 48% to 83%). On an annual basis, nitrate was reduced at a compound average rate of 5.2% per year from 1990 through 2010 with a confidence interval of 2.4% to 7.0%.

A comparison of trends by decade (Figure 7.2) shows no significant difference between the rates in the 1990s and the 2000s.



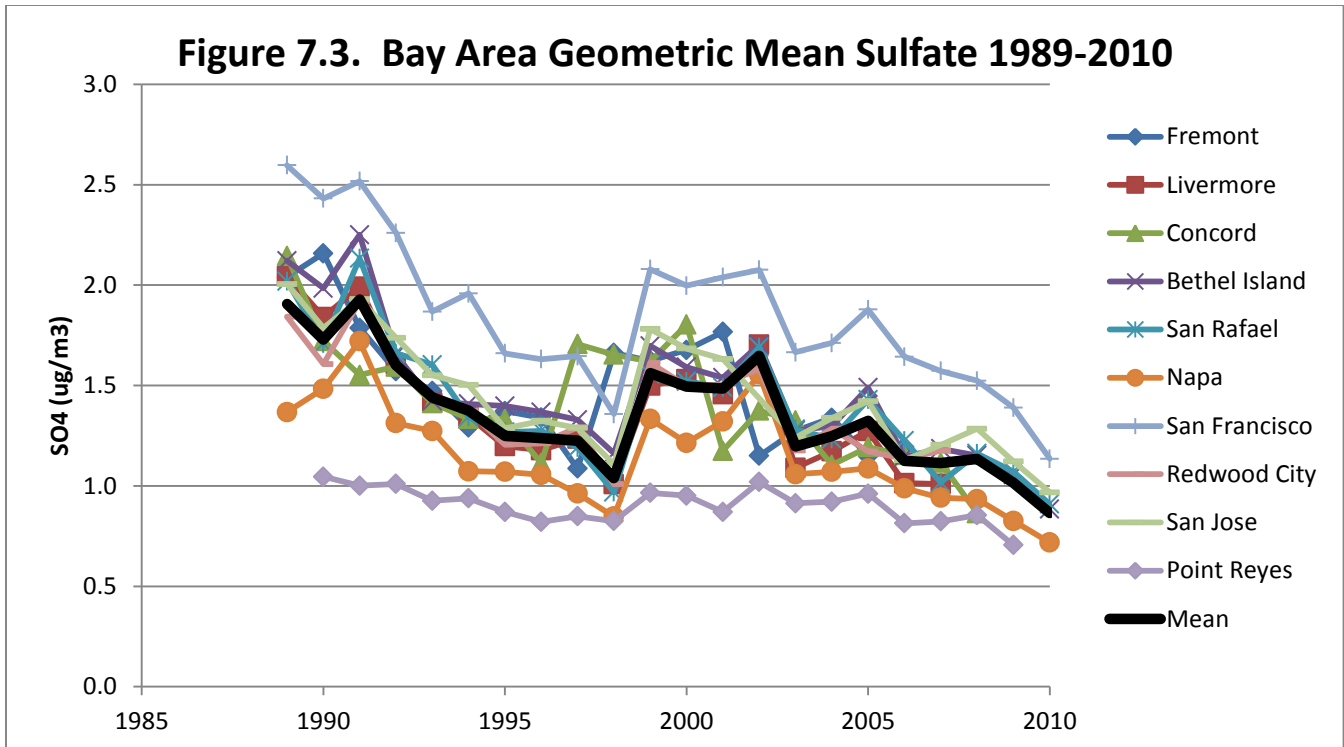


7.2. Sulfate trends

Sulfate is another large component of Bay Area PM. As with nitrate, sulfate is formed from reactions with ammonia, where SO_2 converts to sulfuric acid, combining with ammonia to produce ammonium sulfate $(\text{NH}_4)_2\text{SO}_4$.

There are a couple of differences from nitrate to keep in mind: 1. A sizeable amount of Bay Area sulfate derives from direct emissions from ships, and also as a naturally-occurring components of the salts in marine air. 2. Sulfate is less volatile, so that its concentration on filters is not affected by whether or not the filter was refrigerated.

Figure 7.3 shows the trend in mean sulfate. Although there is considerable variation, largely from meteorology, there is a substantial 20-year decline in mean sulfate concentrations. The 9-site mean sulfate dropped from $2.3 \mu\text{g}/\text{m}^3$ in 1990 to $1.9 \mu\text{g}/\text{m}^3$ in 2000 to $1.3 \mu\text{g}/\text{m}^3$ in 2009. The reduction from 1990 to 2000 was 17%; the reduction from 2000 to 2009 was 29%, and the overall reduction was 41% (with a 90% confidence interval of 27% to 51%). On an annual basis, the average reduction rate between 1990 and 2010 was 3.0% per year with a 90% confidence interval of 2.1% to 4.1%.



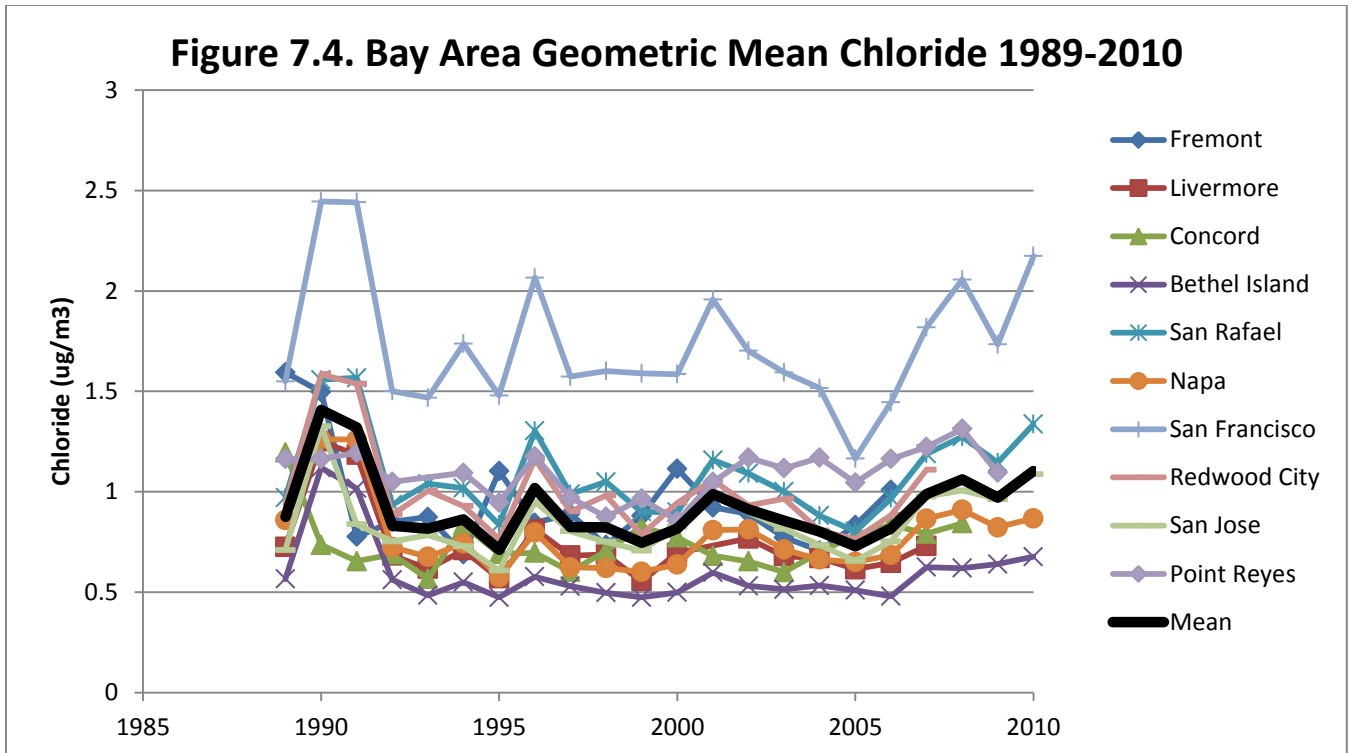
7.3. Chloride trends

Chloride is a major constituent of the Bay Area’s marine air. In fresh marine air, it constitutes about 55% of the non-aqueous mass; other key components are sodium 31%, sulfate (8%), magnesium (4%), calcium and potassium (1% each).

Chloride has no other significant sources, so its presence is a marker for marine air. But as marine air passes over the Bay Area’s urban centers, much of the sodium chloride is replaced by sodium nitrate. Thus, in areas some distance from the ocean or Bay, or at times when the air has been stagnant, the chloride concentration tends to under-represent the marine component.

As Bay Area chloride concentrations derive from marine air, we don’t expect a secular trend. As Figure 7.4 shows, this is indeed the case. The mean of Bay Area geometric mean chloride was 1.2 $\mu\text{g}/\text{m}^3$ for 1989-1991, 0.9 $\mu\text{g}/\text{m}^3$ for 1999-2001 and 1.1 $\mu\text{g}/\text{m}^3$ for 2008-2010. The annual rate of reduction for the whole period, from 1990 through 2010, was 0.5% per year. This reduction is not statistically significant. However, with the reductions in nitrate and sulfate, chloride is now as large a contributor to Bay Area PM_{10} as these.

Among the sites, San Francisco stands out, being surrounded by salt water on three sides. Point Reyes also has high levels given that it’s $\text{PM}_{2.5}$ not PM_{10} , and the chloride in marine air is largely in the coarse (> 2.5 micron diameter) fraction.

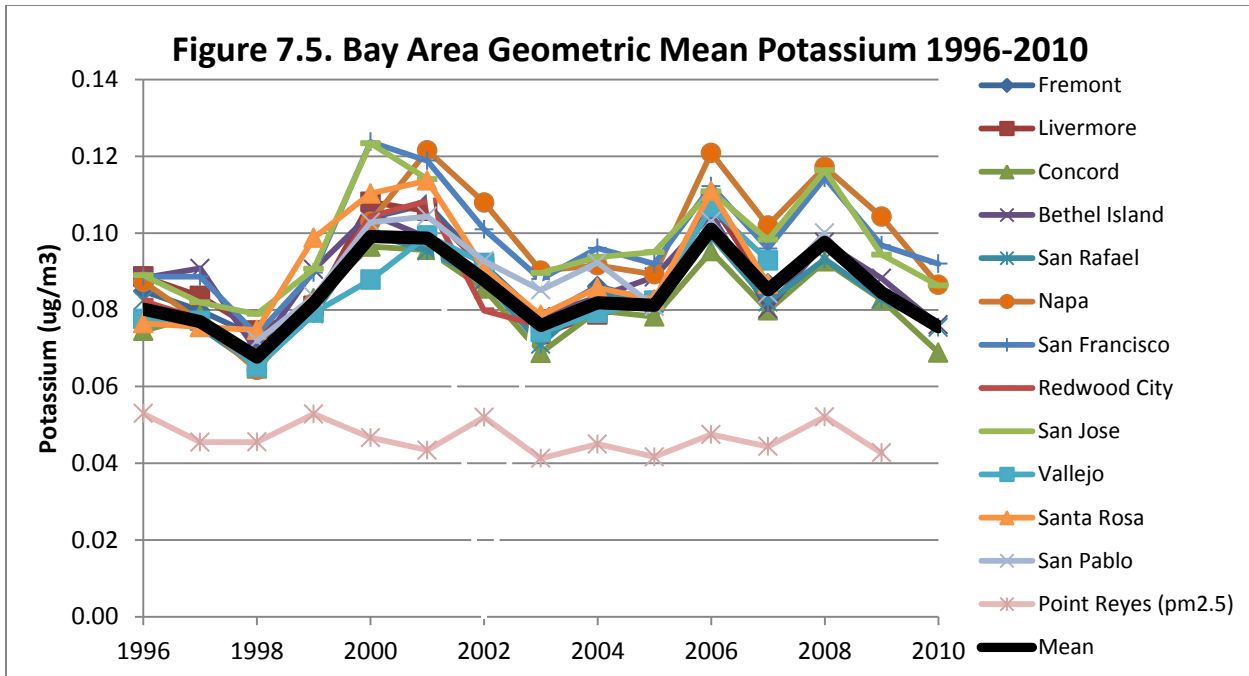


7.5 Potassium trends

The District began regular potassium ion measurements on PM_{10} filters mid-1995 with the idea that it could serve as a marker for wood burning in the Bay Area. Although this might have been effective in some areas, for the Bay Area, the presence of substantial amounts of soluble potassium in marine air obscures the woodsmoke signal.

Figure 7.5 shows mean soluble potassium concentrations from 1996 through 2010 at Bay Area sites. The 1996-1998 mean was $7.5 \mu\text{g}/\text{m}^3$ compared with $8.7 \mu\text{g}/\text{m}^3$ for 2008-2010. The 1.1% per year increase is not statistically significant. Unlike with chloride in Figure 7.4, San Francisco doesn't stand out above and Point Reyes does stand out below the concentrations at other sites. This indicates that there are substantial non-marine sources. Whereas concentrations of nitrate, sulfate and chloride at Napa are all at or below average, it often registers the highest potassium concentrations – with woodsmoke being the likely source; the fact that San Francisco potassium levels aren't higher, despite the fact potassium is a component of marine air, suggests that woodsmoke is not a big source there.

The District has instituted winter Spare the Air alerts that ban wood burning on certain days. It also has had an extensive educational campaign on the negative health effects of woodsmoke. Responses to the District's wintertime surveys have indicated that fewer residents burned wood, with the extent of the reduction being 25% to 50% since the winter of 2006-2007. The absence of a downtrend in soluble potassium suggests that reductions in woodsmoke emissions may have been more modest than the survey indicates.



7.6 Trends in elemental and organic carbon

The record of elemental and organic carbon (EC and OC) in the Bay Area is spotty and not totally consistent. What follows is an analysis utilizing and sometimes piecing together what is available. It should be considered preliminary.

7.6.1. Data

The sources of EC and OC measurements are the following:

1. Starting in 2004, the District began a program of measuring EC and OC on its PM₁₀ filters at some sites: Bethel Island, Livermore, Napa, San Francisco, San Jose, San Pablo, and Vallejo. These measurements continue to the present except for Livermore and Vallejo, where PM₁₀ sampling was discontinued in 2008.
2. At San Jose, EC and OC measurements have been made from PM_{2.5} filters since February 2000. But: a) the site moved in 2002, so the pre- and post-move measurements are not strictly comparable, and b) the EC and OC measurement method was changed⁹ in April 2009, a significant enough change so that those after April 2009 are not really comparable.
3. At Point Reyes, consistent EC and OC measurements have been made for the IMPROVE network on PM_{2.5} filters since 1989.
4. EC and OC measurements were made on PM_{2.5} filters at Bethel Island, Livermore and San Francisco from late 1999 to early 2001 for the CRPAQS study.

⁹ From the NIOSH method to the IMPROVE method. All other EC/OC measurements use the IMPROVE method.

5. EC and OC measurements have been made for the District by DRI on PM_{2.5} filters at Livermore and Vallejo since September 2008 and at West Oakland since February 2009.

6. In addition to filter measurements, the District operated instruments whose measurements are well-correlated with carbon particles, especially EC. There has been an aethalometer operating at San Jose-Jackson Street since 2004. Up until the mid-2000s there had been a set of coefficient of haze (COH) analyzers operating at many sites, including six that operated through 2004.

7.6.2 Analysis

Data were averaged by quarter for all quarters with sufficient measurements. Then the quarterly averages were averaged over 3-year periods. Table 7.1 shows the earliest and latest averages and the annual rates of decrease.

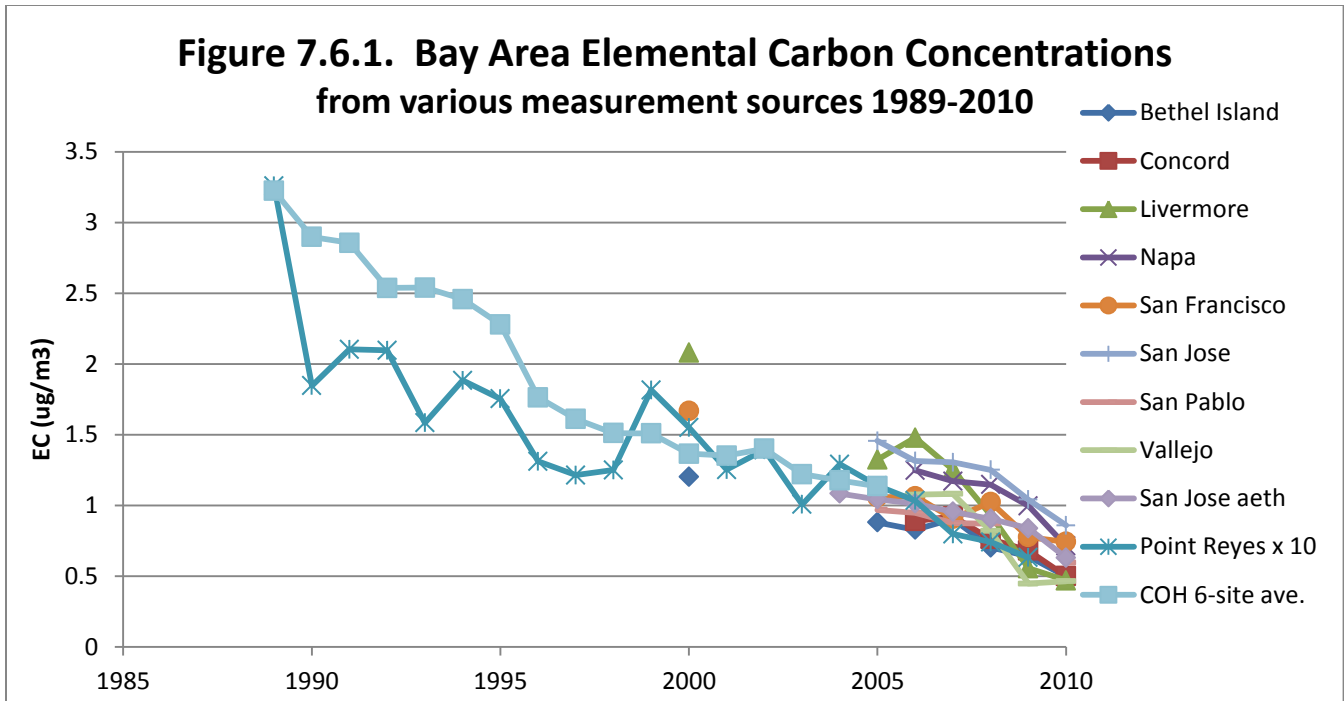
There is considerable site-to-site variation in trends, especially in OC. Some of it probably comes from piecing together data from more than one lab, but there seem to be other factors as well.

Among filter measurements, the median change in OC is a decrease of about 4% per year. The change in EC is a decrease of about 8% per year. The San Jose aethalometer shows a decrease of 5% per year compared with filter-based measurement reduction of 9% per year. The COH instruments show the likely reduction in EC during the 1990s as 7% to 8% per year, about the same as the filter measurements indicate for the 1st decade of the 21st century.

Table 7.1. Trends in OC and EC at Bay Area sites.

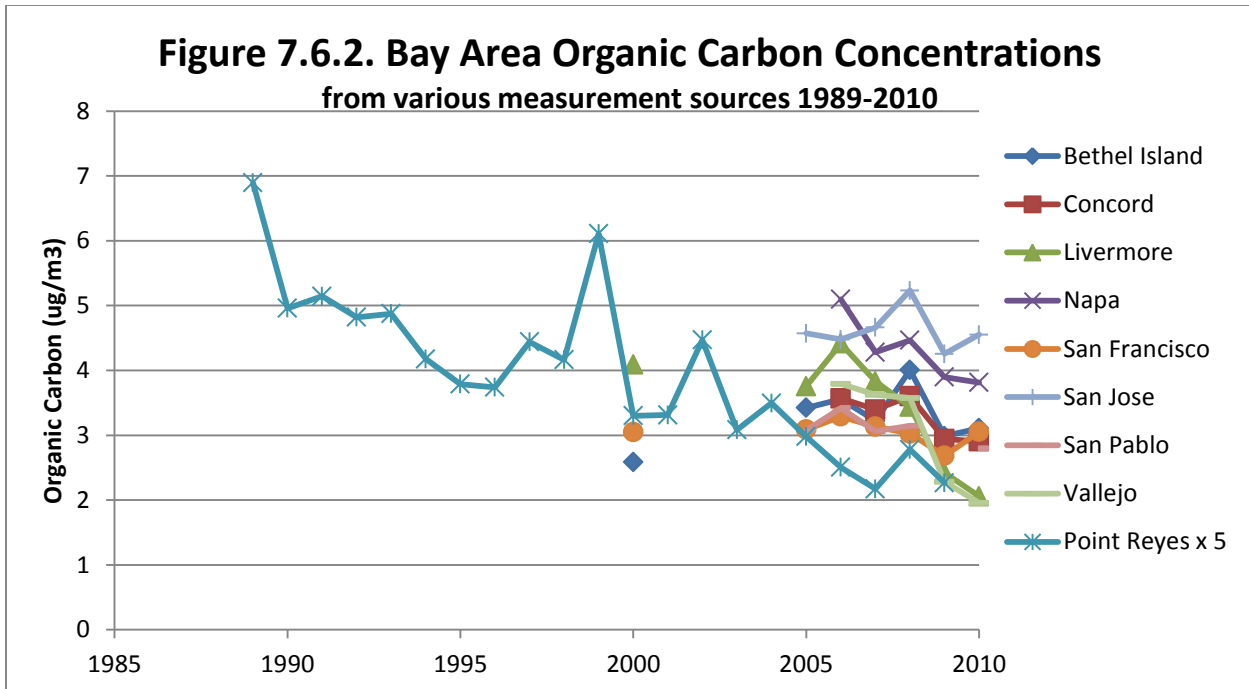
Site	Data Source	1 st OC	Last OC	Annual Reduction	1 st EC	Last EC	Annual Reduction
Bethel Island	CRPAQS PM _{2.5} 99-01/ District PM ₁₀ 08-10	2.91	3.36	-1.6%	1.34	0.61	8.4%
Concord	District PM ₁₀ 05-07 / 08-10	3.50	3.35	1.4%	0.86	0.77	3.7%
Livermore	CRPAQS PM _{2.5} 99-01/DRI PM _{2.5} 08-10	4.54	2.41	6.8%	2.39	0.61	14.1%
Napa	District PM ₁₀ 05-07/08-10	4.57	4.08	3.7%	1.18	0.95	7.0%
Point Reyes	IMPROVE 97-99/07-09	0.98	0.48	6.9%	0.14	0.07	7.3%
San Francisco	CRPAQS PM _{2.5} 99-01/District PM ₁₀ 08-10	3.41	2.91	1.7%	1.85	0.76	9.5%
San Jose	District 04-06/08-10	4.54	4.52	0.1%	1.48	1.02	8.9%
San Pablo	District 04-06/07-09	3.38	3.01	3.8%	0.96	0.84	4.3%
Vallejo	District PM ₁₀ 04-06/DRI PM _{2.5} 08-10	3.75	2.32	14.8%	1.08	0.58	18.8%
Median				3.7%			7.8%
SJ aethalometer	District 04-06/08-10				1.08	0.84	4.9%
COH	7-site mean 89-91/99-01				3.16	1.43	7.6%

Figure 7.6.1 shows a plot of EC values from the various sources from 1990 through 2010. Note that Point Reyes values have been multiplied by 10 to put them on a comparable scale. The graph shows what appears to be a steady reduction in EC. The larger than average reductions from 2008 to 2009 at Livermore and Vallejo may represent the change from District PM₁₀ measurements to DRI PM_{2.5} measurements.



The COH mean for 1989-1991 was 2.99 $\mu\text{g}/\text{m}^3$. A mean of Concord, Napa and San Francisco EC, the 3 of the 6 COH sites with EC measurements from PM₁₀ filters had a mean of 0.82 $\mu\text{g}/\text{m}^3$ for 2008-2010. This represents a reduction of 73%. For Point Reyes, EC dropped from 2.4 $\mu\text{g}/\text{m}^3$ in 1989-1991 to 0.69 $\mu\text{g}/\text{m}^3$ in 2007-2009, a reduction of 70%.

Figure 7.6.2 shows trends in OC concentrations. For most sites, the data record begins in 2004. As discussed above, we do see a median downtrend of about 4% a year among these sites. The Point Reyes site has a long and consistent record. The figure shows its OC values times 5 to make it comparable with the concentrations at other sites. On its original scale, Point Reyes OC averaged 0.99 $\mu\text{g}/\text{m}^3$ for 1990-1992 and 0.48 $\mu\text{g}/\text{m}^3$ for 2007-2009, representing a reduction of about 50%. This represents an annual rate of decrease of 4.1% per year, in line with the reduction rate at District urban sites.



7.6.3 Carbon-14 trends

In addition to EC/OC measurements, there has been a limited amount of measurement of Carbon-14. C-14 is a radioactive isotope of carbon with a half-life of 5,730 years. It is constantly renewed in the atmosphere by cosmic radiation. C-14 is taken up by plants, and through them animals including humans. Therefore, wood and meat have the modern fraction of C-14, whereas fossil fuels, whose carbon was absorbed millions of years ago, have virtually no C-14. Thus, the amount of C-14 in the carbon of filter samples provides an excellent estimate of the fraction of carbon from “new” vs. fossil sources.

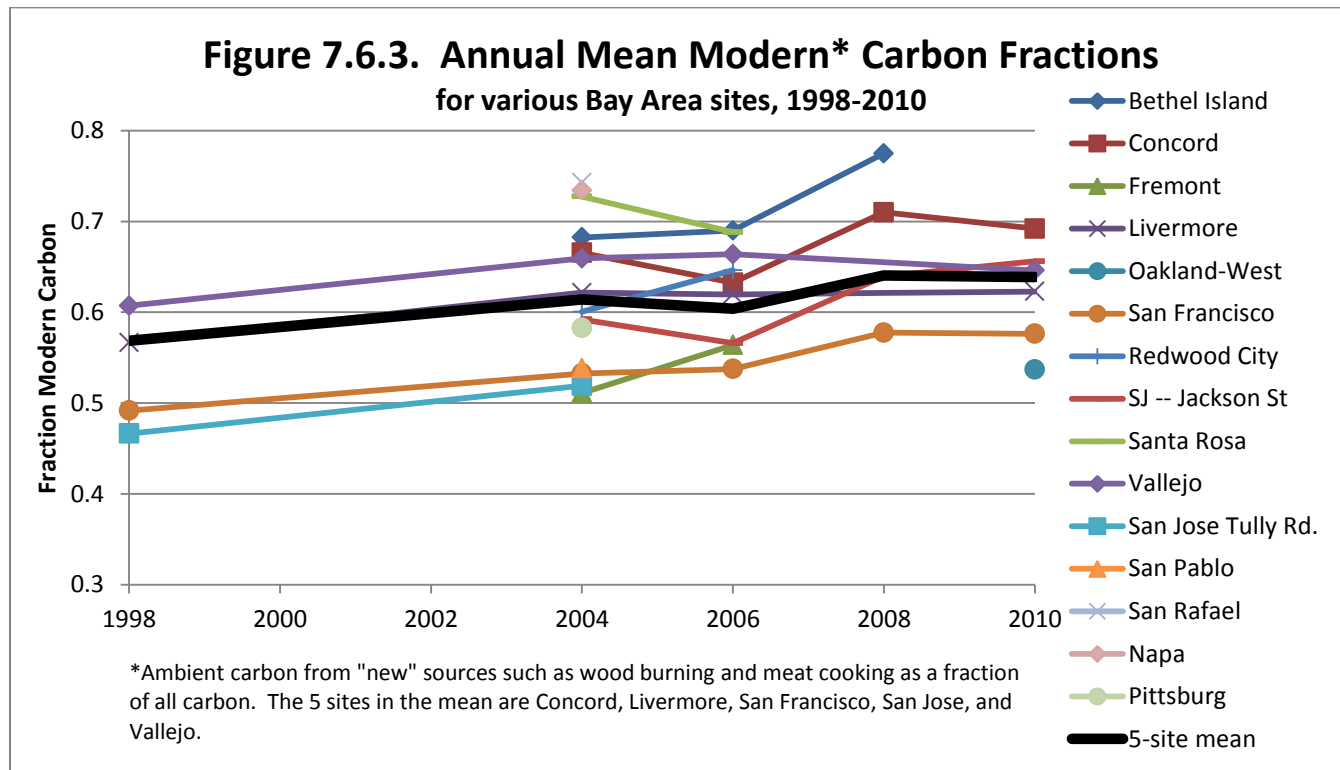
The District has had measurements made by the University of Arizona’s Accelerator Mass Spectrometry facility. A significant portion of these measurements have been made, not on a single filter, but on a composite of a random sample of filters taken around the year. This not only helps gain an understanding of annual average fractions, it also saved money in the high per-sample cost. Such samples have been made on a sporadic basis on filters dating back to 1998.

Figure 7.6.3 shows the fraction of modern carbon in samples from various Bay Area sites from 1998-2010. It is likely that the sites with the highest modern carbon fractions – Bethel Island, Napa, Santa Rosa and San Rafael – are impacted by significant woodsmoke, and conversely for the sites with low carbon fractions, including San Francisco and Oakland.¹⁰

Also shown is the mean of the 5 sites that had measurements made in at least 4 years: Concord, Livermore, San Francisco, San Jose, and Vallejo. Missing site-years were imputed iterating a 2-way ANOVA. The mean value

¹⁰ Air quality sites can’t represent a whole city or metropolitan area, even one as small as San Francisco (10 km x 10 km). Where a monitoring site is placed can make a substantial difference. The SF site, for example, is located in an industrial area with few homes nearby.

increases almost monotonically from 0.57 in 1998 to 0.63 in 2010, a 12% increase. This suggests that PM from fossil combustion has been reduced faster than PM from wood burning or cooking.



8. PM_{2.5} trends in Central California air basins

The Sacramento and San Joaquin valleys border the Bay Area and there is considerable air exchange between them. In summer, the typical pattern is westerly winds blowing from the Bay Area into these valleys. In the winter, during periods of stagnation when PM concentrations are high, the pattern tends to reverse with easterly flow draining from the valleys out into the Pacific Ocean.

Thus, there is interest in how the PM concentrations in these Central Valley air basins compare with those of the Bay Area. Figures 8.1a and 8.1b show the annual means for the Sacramento and San Joaquin valleys. Comparing with the Bay Area means in Figure 5.1, we see that Sacramento means are similar, and San Joaquin urban means are higher.

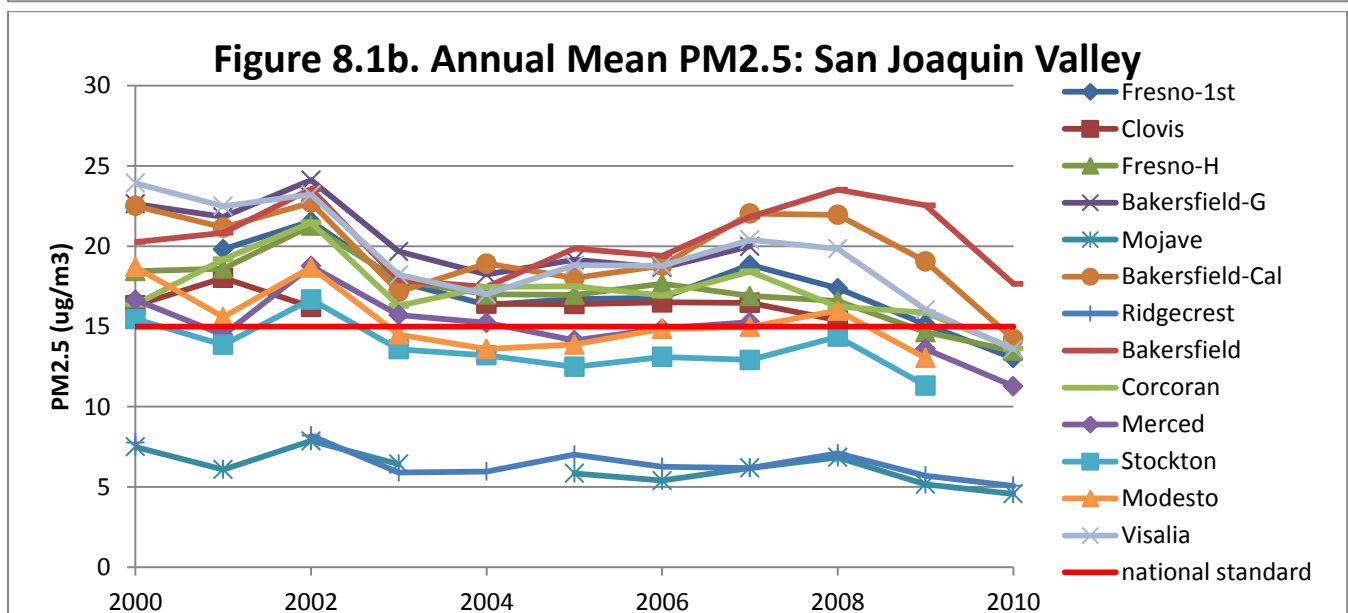
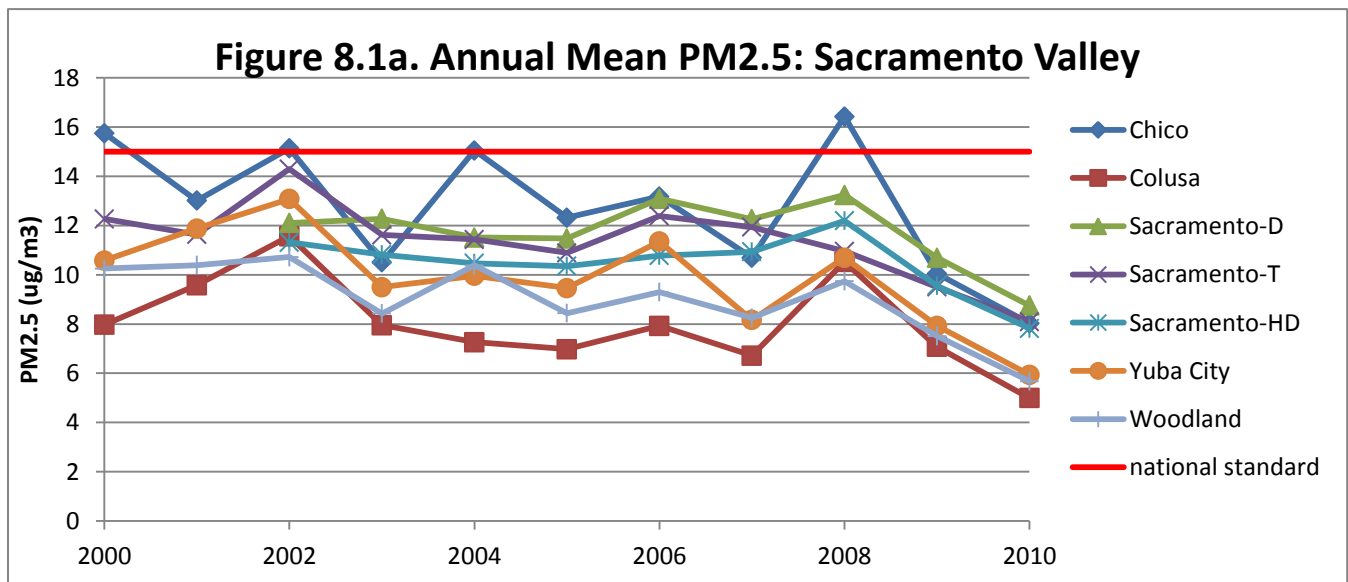
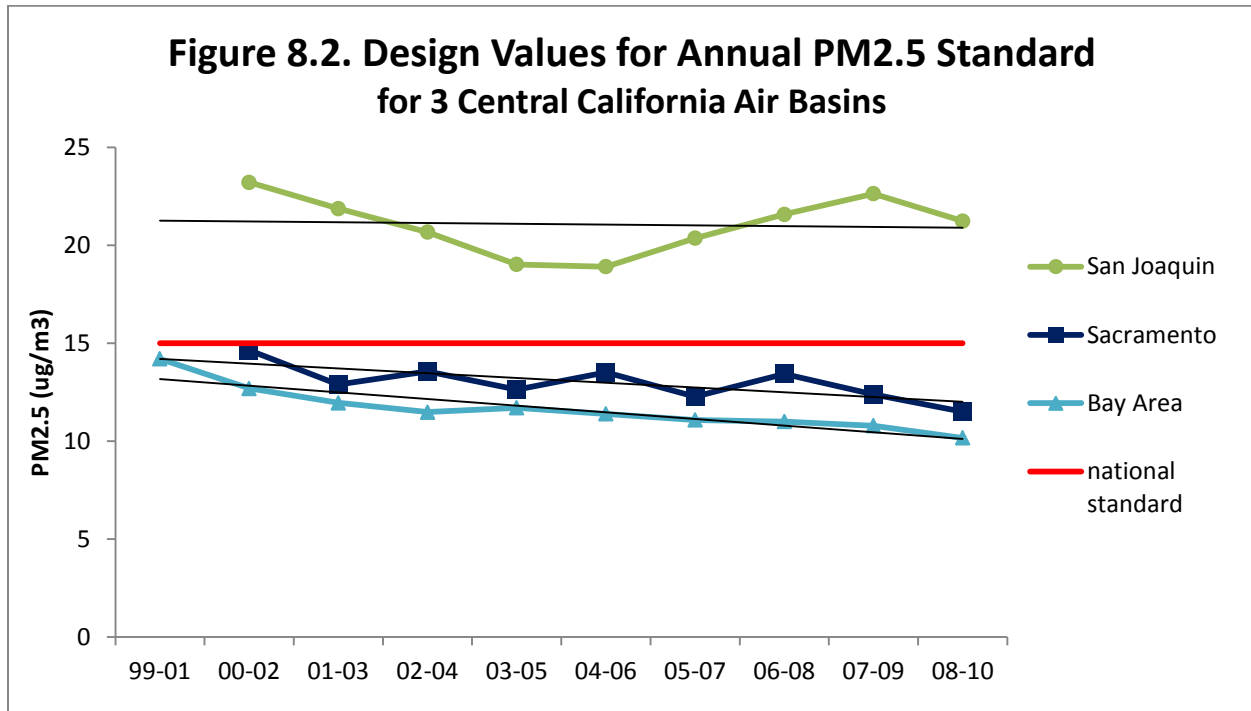


Figure 8.2 shows the trend in Bay Area mean $PM_{2.5}$ compared with trends in the Central Valley. The trends in the Bay Area and Sacramento are significant and similar: 2.8% per year (90% confidence interval 1.5% to 4.1%) for the Bay Area and 1.9% per year (90% confidence interval 0.5% to 3.2%) for Sacramento. The trend for San Joaquin is virtually 0 and not statistically significant.



8.1 98th Percentile comparison

Figures 8.3a and 8.3b show $PM_{2.5}$ 98th percentiles for Sacramento and San Joaquin valleys. These are comparable to Figure 5.3 for the Bay Area. Sacramento and the Bay Area are similar – with values above the standard in 1999 falling below the standard by 2010. The San Joaquin Valley also shows a decrease, but its sites all continue to violate the standard.

Figure 8.3a. PM2.5 98th Percentiles: Sacramento Valley

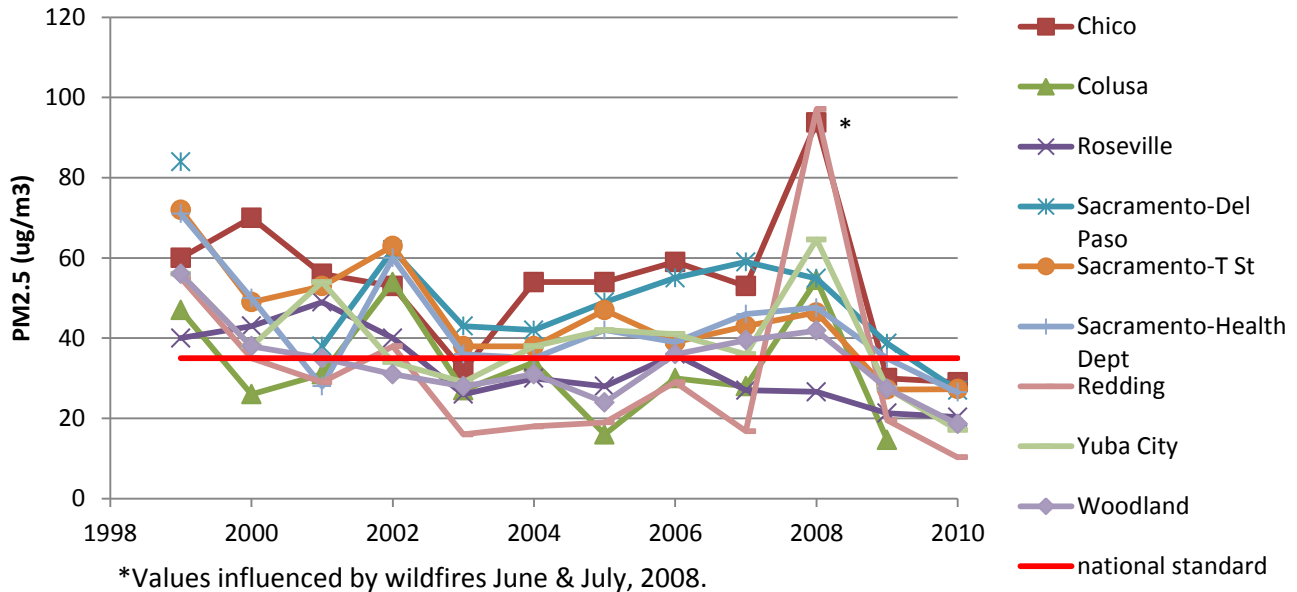


Figure 8.3b. PM2.5 98th Percentiles: San Joaquin Valley

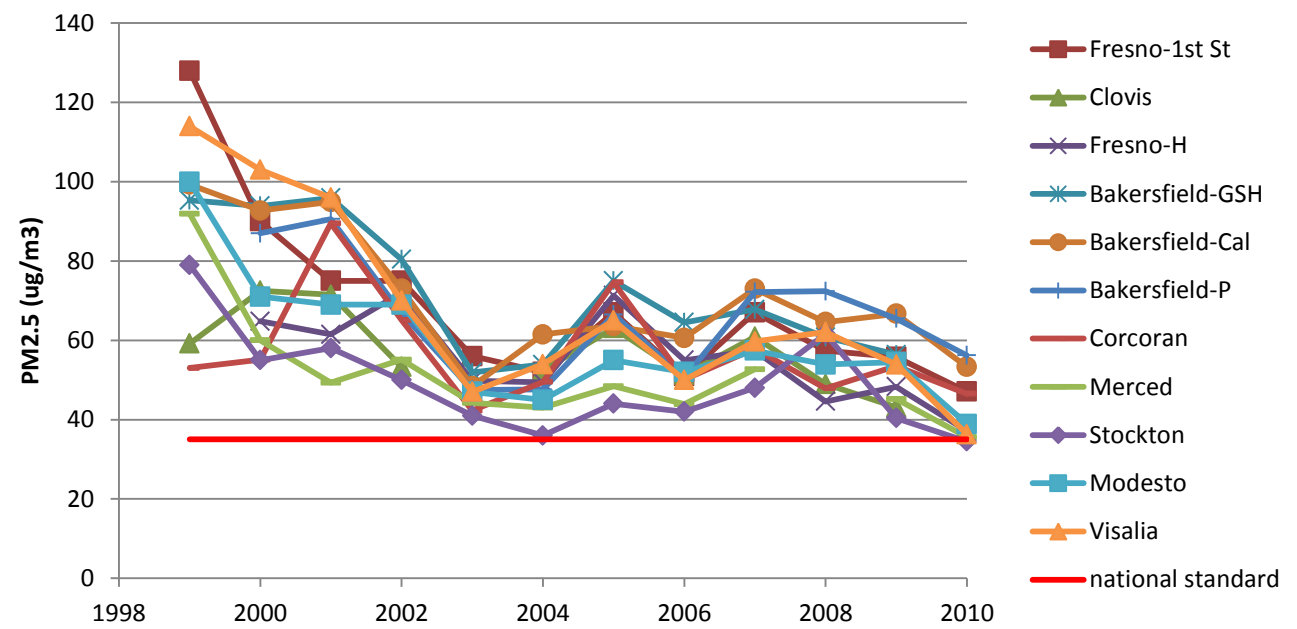
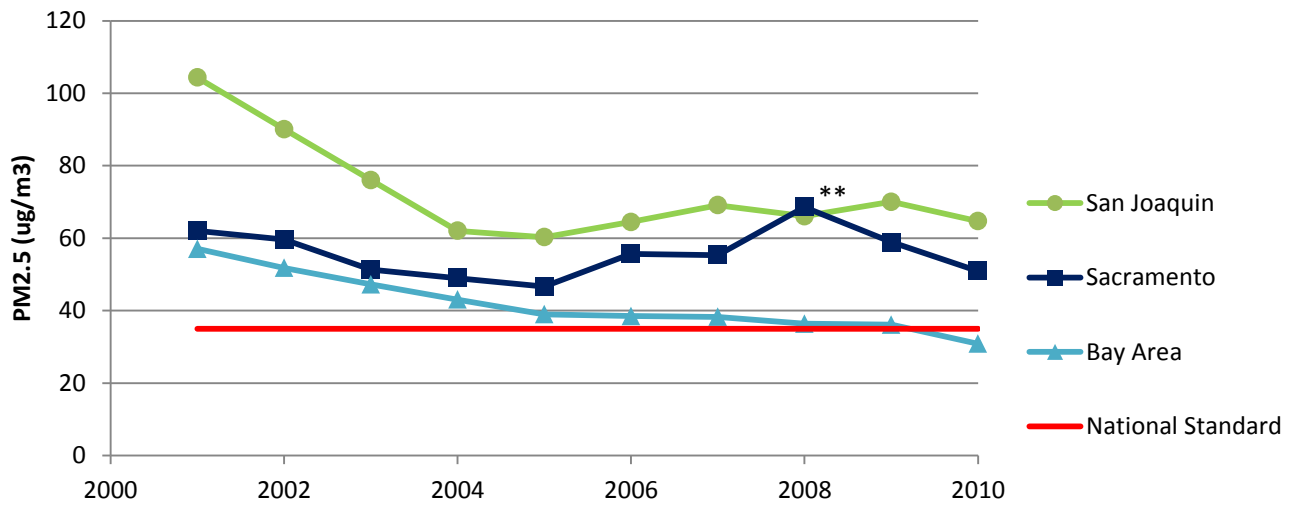


Figure 8.4 shows the 24-hour design values for the three air basins. The Bay Area’s design value has decreased 5.8% per year from 2000 to 2010 (90% confidence interval 3.6% to 7.8%) and met the standard for 2008-2010. The design value in the San Joaquin Valley shows a decrease of 3.8% per year that is borderline statistically significant. Sacramento’s design value shows no decrease, partly because of the 2008 wildfires. Excluding the wildfire months of June and July, the decrease is 3.3% per year, still not statistically significant.

Figure 8.4 Design Values* for 24-Hour PM2.5 Standard for 3 Central California Air Basins.



* Design values are 3-year averages of 98th percentiles, plotted vs. most recent year.

**Values for 2008-2010 influenced by wildfires in summer 2008.

9. Comparisons with the emissions inventory

Trends estimated from ambient measurements and the emissions inventory have somewhat complementary strengths and limitations. Comparing these trends can serve as a reality check on both.

Ambient trends have the strength that they are based on directly measured pollutant concentrations. But they are limited in the number of locations, and subject to considerable variation from meteorology. Emissions inventory trends, on the other hand, are based on a comprehensive set of sources covering the entire Bay Area and are unaffected by meteorology. But they are based on engineering calculations that, inevitably, don't perfectly match actual emissions, and don't account for atmospheric reactions and transport.

Comparing total PM concentrations with emissions is complex, both because PM is a mixture of many compounds, and because some PM is not directly emitted. The analysis of ammonium nitrate and sulfate is also complex, since although each involves only one compound, each has two precursors and several chemical pathways for their formation.

Figure 9.1 compares trends in nitrate (a 10-site average) and NOx emissions from 1988 to 2010. NOx emission estimates are taken from a report, BAAQMD 2008. Reductions in NOx don't automatically imply reductions in nitrate, because nitrate concentrations are a function of concentrations of other compounds especially ammonia. But the similar trends (including a correlation of 0.74) suggests that there is a causal link.

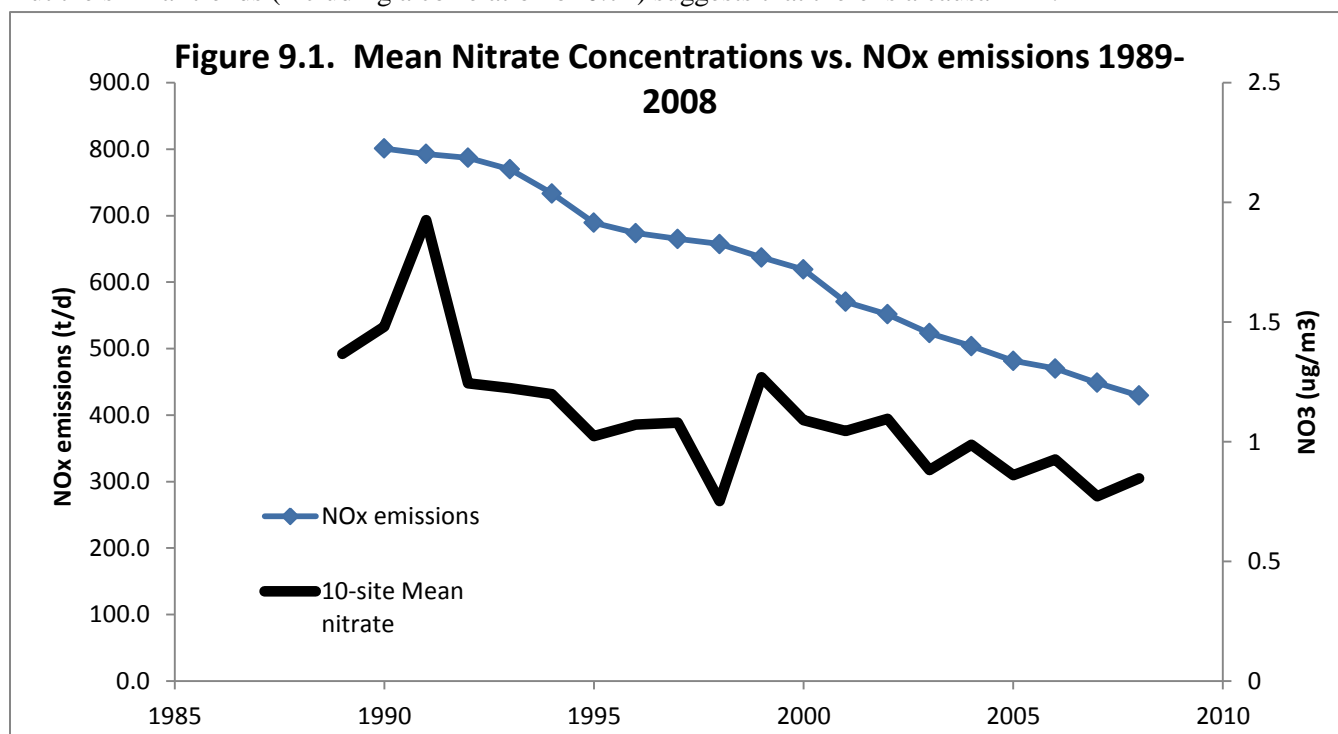
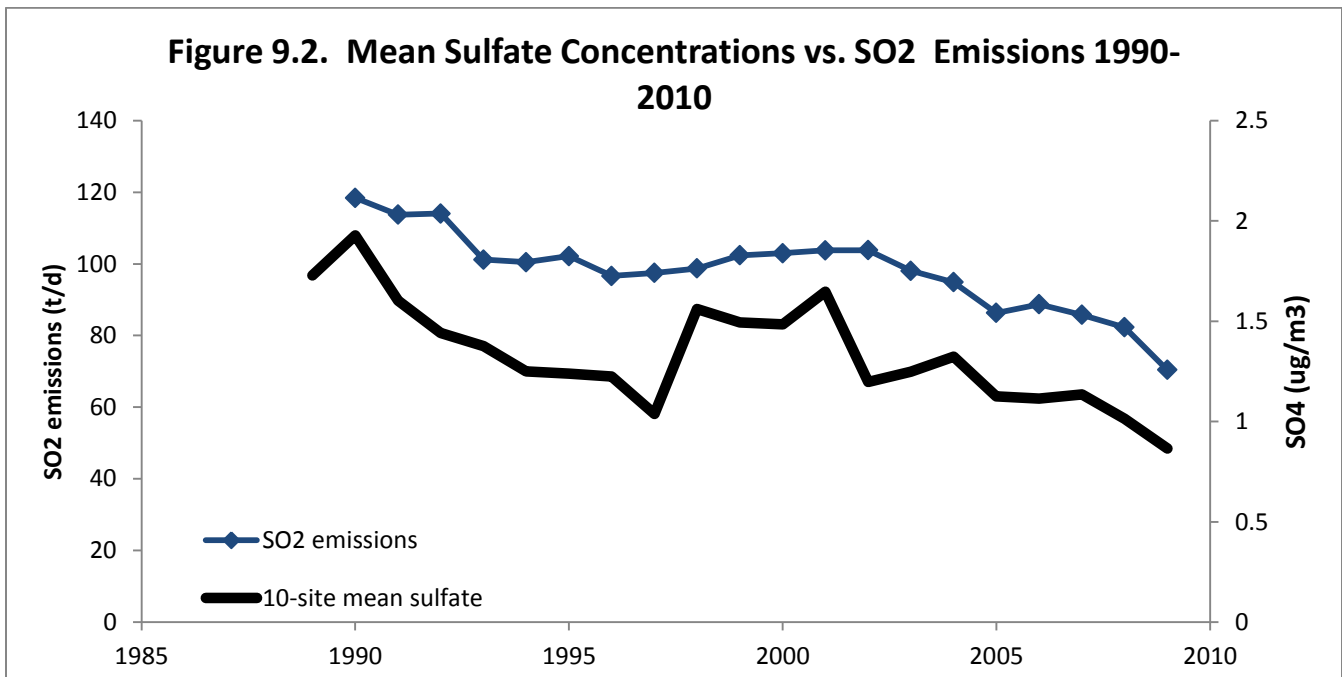
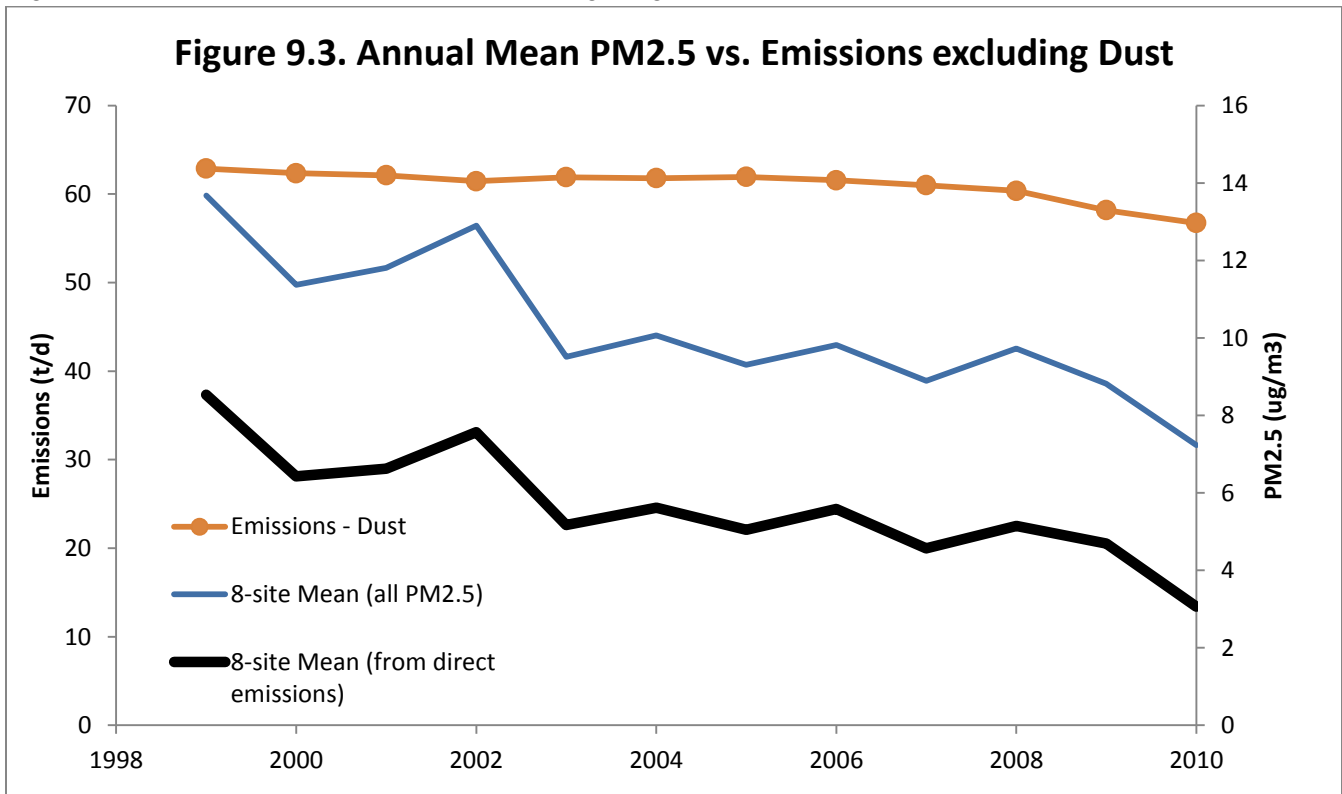


Figure 9.2 shows trends in sulfate concentrations and SO2 emissions. As with nitrate/NOx we don't expect a match, necessarily. But again we see similar trends (correlation 0.83).



Figures 9.3 and 9.4 show PM_{2.5} and PM₁₀, where geological dust is excluded from emissions.



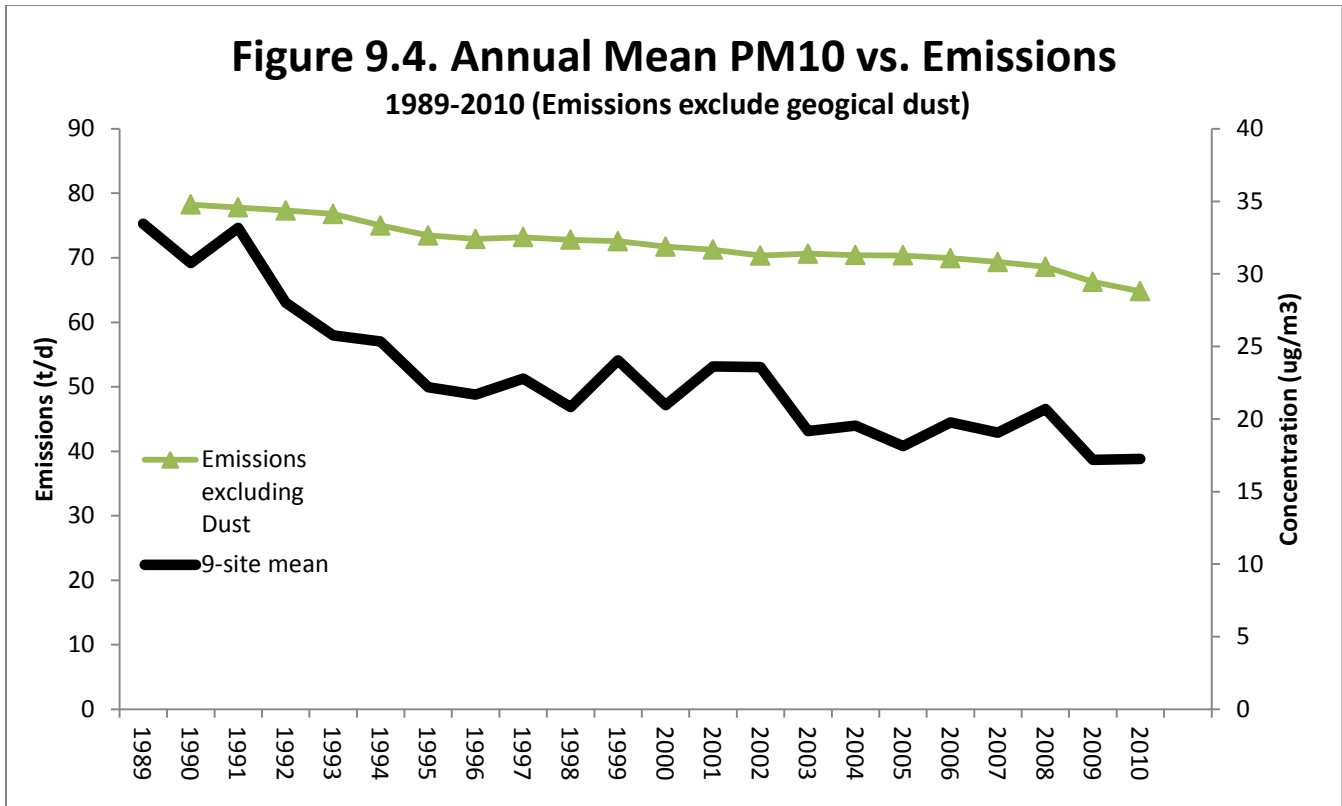


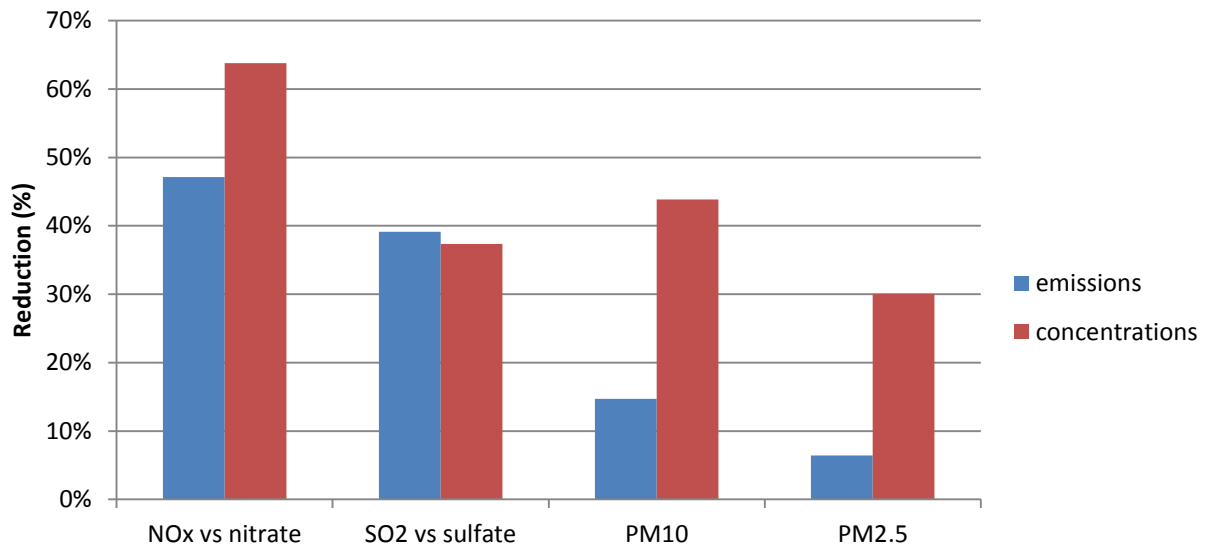
Figure 9.5 shows a comparison of percent reductions in emissions and concentrations. The reductions for SO₂ vs. sulfate are quite similar. The trends for NO_x vs. nitrate appear different, but are within statistical uncertainty.¹¹ Some of the difference derives from the choice of summary statistic.¹²

The PM trends are quite different, however. Even excluding geological dust emissions (which is estimated to increase), the trend in PM emissions is relatively slight – a reduction of about 15% for a 20-year period for PM₁₀ and 6% over a 10-year period for PM_{2.5}. In contrast, PM₁₀ concentrations have fallen over 40% in 20 years, and PM_{2.5} has fallen 30% in 10 years. Moreover, excluding nitrate and sulfate, which aren't emitted directly, and chloride, which is from marine air not counted in the inventory, the PM_{2.5} reductions are *greater*.

¹¹ The rate of NO_x emissions was within a 95% confidence interval for the rate of nitrate emissions.

¹² The reductions in *geometric* mean nitrate was 45%, almost identical with the NO_x decrease. The difference with the arithmetic mean derives from the difference in rates that different parts of the nitrate distribution have been reduced (see section 10).

9.5. Comparison of reductions in emissions and concentrations



Reductions over 20-year period for all except a 10-year period for PM2.5. PM emissions exclude geological dust.

10. Trends in PM quantiles

The above analysis shows downtrends in both coarse and fine PM and in key anthropogenic components. But downtrends in peak PM generally appeared greater than trends in the mean. This section explores how that might occur.

Figure 10.1 shows P-P plots¹³ from the six Bay Area sites with PM₁₀ measurements stretching from 1989 through 2010. The plot compares the PM₁₀ distributions from 2000-2010 with 1989-1999. The lines $y = x$ are drawn for comparison. In every case, the data fall below the line $y = x$, indicating that PM₁₀ has been reduced across the board. But in every case, the lines of data bend away from $y = x$ for the larger percentiles (starting at about the 75th percentile), indicating that there has been a greater percentage reduction in the top quarter of the distribution than in the rest.

Figure 10.2 shows percentile-percentile plots comparing 2000-2010 nitrate and sulfate percentiles with 1989-1999 for two Bay Area sites. The sites were chosen to represent the eastern and western parts of the Bay Area. The eastern part is affected more by Central Valley PM, which has high nitrate levels but less sulfate. San Francisco sulfate is elevated by shipping in the Bay. The plots for the two sites are remarkably similar in shape, suggesting that causal factors resulting in lower nitrate are widespread. Both nitrate and sulfate follow the pattern of PM₁₀, with across-the-board reductions in the distribution, but greater reductions for the high percentiles. The bend for nitrate is more extreme, indicating large reductions in peak nitrate, smaller relative reductions for peak sulfate.

¹³ A Q-Q plot is a method for comparing two distributions. It is a scatterplot showing the quantiles of one distribution against the quantiles of another. It pairs the first percentile of one distribution with the first percentile of the other, the 2nd percentile with the second percentile, and so on up to the 99th vs. the 99th. If the two distributions are equal, the data will lie on the line $y = x$. If one distribution is a simple shift of the other, then the Q-Q plot will lie on a straight line, offset from the line $y = x$ and parallel to it. If the difference is a simple proportion (e.g., all values have been reduced by 10% across the board) then again the Q-Q values will lie on a straight line, but this time through the origin. Q-Q plots that are not straight lines indicate a more complex difference that affects different parts of the distribution differently.

Figure 10.1. Percentile-Percentile Plots Comparing PM10 in 2 Decades
2000-2010 vs. 1989-1999 PM10 percentiles from 6 Bay Area sites

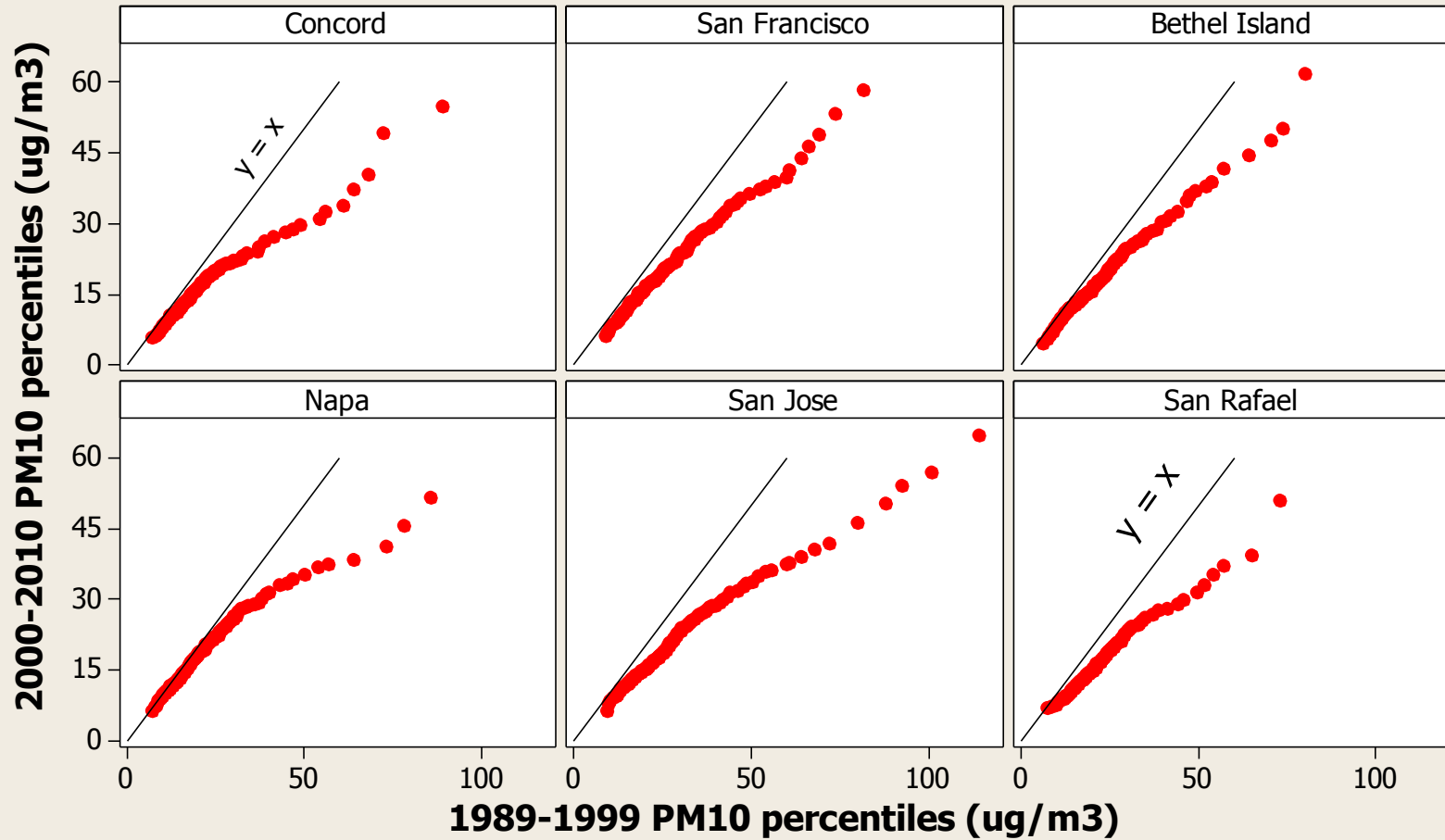
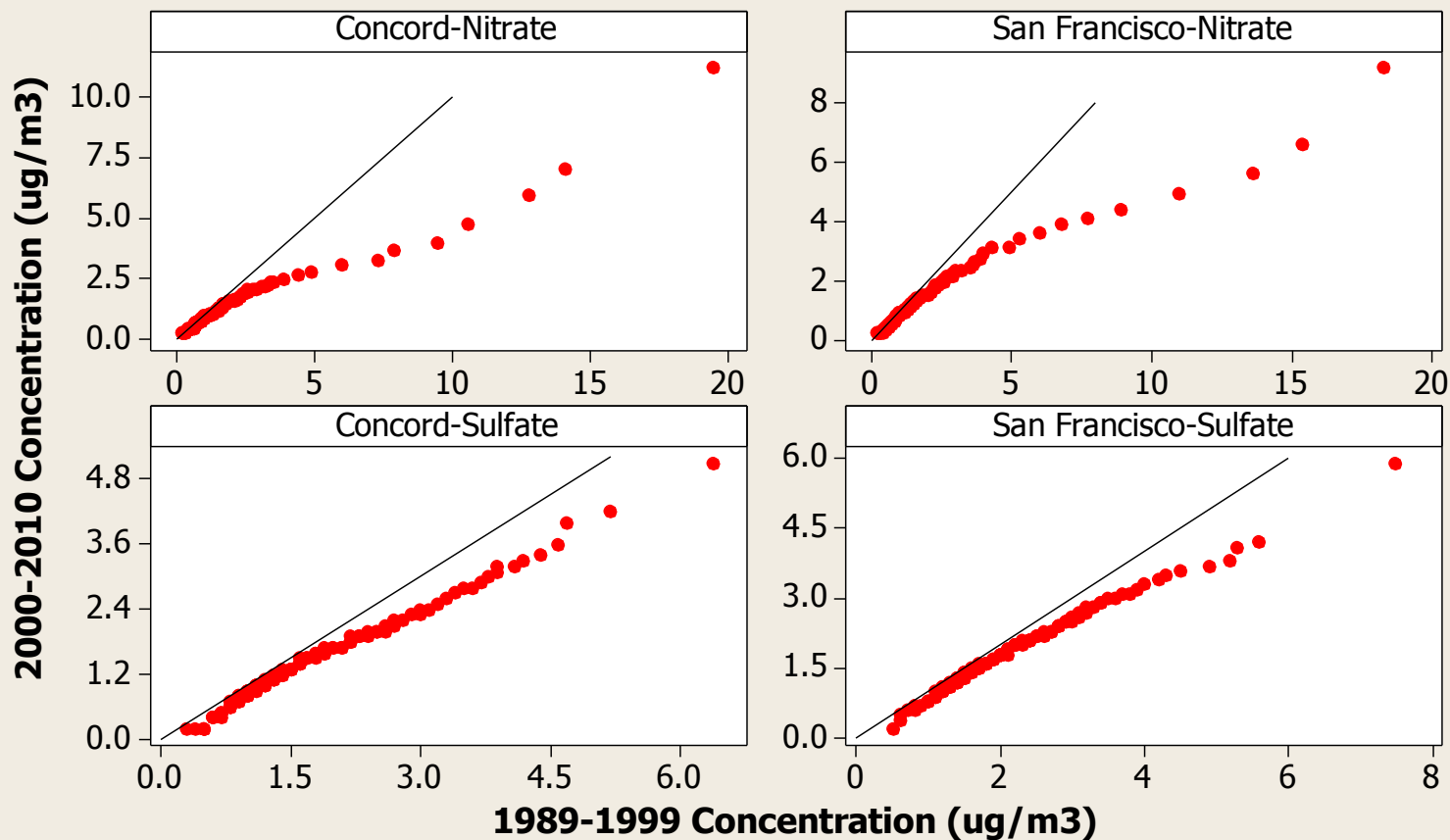


Figure 10.2. Percentile-Percentile Plots of Comparing NO₃ and SO₄ for 2 Decades
2000-2010 vs. 1989-1999 Nitrate and Sulfate Percentiles at 2 Bay Area Sites

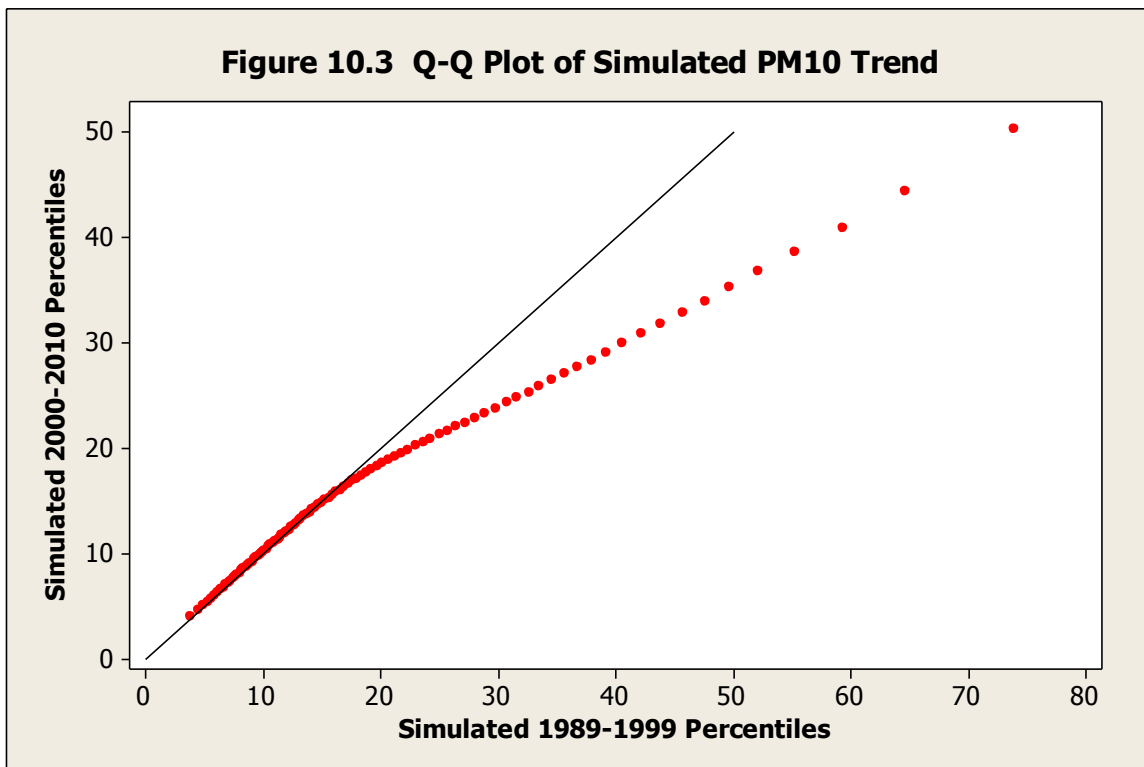


One consequence of this analysis is to show that peak PM has declined faster than lower percentiles. Thus, the percentage reduction in the 98th percentile is greater than the reduction in the mean (or median).

What do these plots imply about the dynamics of PM trends? An approach to answering this question is to find a simulated process with a similar P-P plot. One hypothesis is that the hockey-stick pattern derives from reductions in the distribution of anthropogenic PM against a constant distribution of background PM. Various simulations were performed attempting to mimic the appearance of the Q-Q plots in Figures 10.1 and 10.2. Notable was the failure of the rollback model $Y_j = X_{back} + c_j * X_{anthro}$, in which the trend was modeled as a reduction in c_j where $j = 1989-1999$ and $2000-2010$. This led not to a hockey stick but, rather, roughly to the line $y = dx$.

It was noted that PM concentrations are much higher in winter than the rest of the year. Our next simulation assumed that there was no change in $\frac{3}{4}$ of the year and a large change in the other $\frac{1}{4}$. Specifically we assumed that X had a lognormal(0, .5) distribution. $Y_{1989-1999} = 12 * (X_{back} + 3 * X_{anthro})$, $Y_{2000-2010} = 12 * (X_{back} + X_{anthro})$ for the 3 winter months, and $Y_{1989-1999} = X_{back}$ and $Y_{2000-2010} = X_{back}$ for the other 9 months. The factor 12 was selected to scale the Y to the PM10 range seen in Figure 10.1. Figure 10.3 is based on a simulation of 10,000 winter values and 30,000 non-winter values for $Y_{1989-1999}$ and $Y_{2000-2010}$.

Although the simulation is not intended to be a comprehensive model for PM_{10} , its similarity with the patterns in Figure 10.1 suggests this combination as an explanation of the behavior of the Q-Q plot.



11. Trends in PM health effects

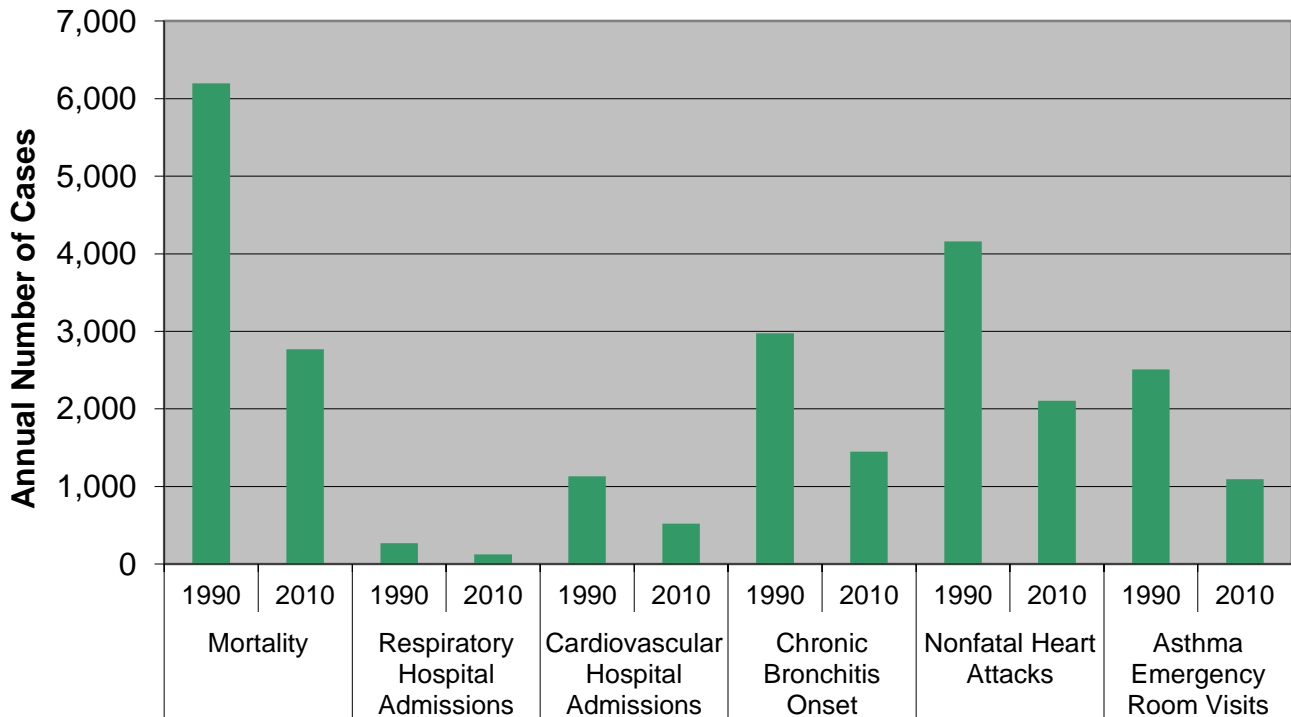
The health effects of breathing ambient PM have been extensively researched both in large-scale epidemiological studies and animal experiments. The US EPA has concluded that: "...the epidemiological evidence continues to support likely causal associations between PM_{2.5} and PM₁₀ and both mortality and morbidity from cardiovascular and respiratory diseases..." (US EPA 2004, Page 9-48, Paragraph 1).

Dose-response relations have been established in various studies (e.g., Hall *et al.* 2008, Stratus 2008) for a range of health endpoints. We used the values developed for the US EPA in their BenMAP program (US EPA 2008).

Although a no-effects threshold has not been found, we assumed that effects were limited to PM_{2.5} concentrations above the Bay Area background, which we estimated to be the 5.5 µg/m³ mean measured at the Point Reyes site. We estimated the 1990 PM_{2.5} values by scaling the 2008-2010 values by 17.9/10.2, the ratio estimated for San Jose in section 5.1.1. Because we look at the excess above 5.5 µg/m³ (i.e., 17.9-5.5 vs. 10.2-5.5) the reduction from 1990 to 2010 in concentration above the background is about 2/3.

Figure 11 shows estimates of the number of annual cases of various health effects caused by elevated PM_{2.5} concentrations in the Bay Area. Note that the reduction in the number of cases estimated is substantial, changing from over 6,000 in 1990 to under 3,000 in 2010.

Figure 11. Trends in Health Effects from Ambient PM
Annual number of cases of health effects 1990 and 2010 in the Bay Area



Over the past 20 years, Bay Area life expectancy has increased by almost 5 years, from 75.7 in 1990 to 80.5 in 2011, due to a variety of factors. Of the overall increase in life expectancy during this period, we estimate that the improvements in air quality can be credited with extending average life expectancy in the Bay Area by 6 months.

12. Recommendations

The ability to analyze ambient air quality trends depends crucially on the existence of the appropriate air quality data. The ideal is to have a long-term set of measurements made with a constant methodology at a large number of sites that represent an area. But the primary driving force behind the District's monitoring network are national and State requirements. New instruments and sites must be added periodically, and this requires tradeoffs to stay within a reasonable monitoring budget. Nevertheless, there are some principles that could be followed even in an era of tight budgets:

- When one measurement method is replaced by another, as with COH by EC, or filter PM_{2.5} by continuous PM_{2.5}, it is valuable to *provide a period of overlap* at several sites where both measurements are made will help to establish how to correlate the two so that their trends can be melded together.¹⁴
- From an analytical point of view, a way to economize would be to *make some measurements periodically rather than continuously*. An example is our speciated PM_{2.5} measurements at Livermore, Vallejo and Oakland. Valuable as these are, they could be discontinued for a couple of years and then resumed.
- *Sites with high PM should have monitors that yield official measurements*. The District generally complies with this principle, but West Oakland is an exception, having no FEM or FRM monitor even though it may have (or have had) the highest annual average PM_{2.5} in the District. A dramatic reduction in PM is likely to have occurred at this site, but the lack of official measurements makes it impossible to document with certainty.
- There is real value in *continued monitoring at fixed monitoring locations*. They provide the backbone of any trend analysis; without a set of these fixed points, trend analysis becomes difficult or impossible.

¹⁴ The EPA has a program to compare FEM measurements with FRM, but the relationship varies by region, so additional analysis for the Bay Area can be useful. The District made COH and EC measurements in addition to EPA-required sampling; the COH-EC relationship has been studied, but also varies by region.

References

BAAQMD (2008). *Base Year 2008 Emissions Inventory Summary Report*, May 2011. [Available from BAAQMD, 939 Ellis St., San Francisco, CA 94109].

CFR 40 (2011). *Code of Federal Regulations*, Title 40 Protection of Environment Part 50 National Primary and Secondary Air Quality Standards: 50.6, 50.7 and Appendix N. Available at http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&tpl=/ecfrbrowse/Title40/40cfr50_main_02.tpl [accessed 1/3/12]

Fairley D (2011). E-mail to Eric Stevenson & Glen Colwell, 2/7/2011.

Hall JV, Brajer V, Lurmann, FW (2008). *The benefits of meeting clean air standards in the South Coast and San Joaquin Valley air basins*. Institute for Economic and Environmental Studies, Cal State Fullerton, November, 2008.

Lippmann M (2011). "Total Suspended Particles (TSP)." In *Encyclopedia of Public Health*. Ed. Lester Breslow. Gale Cengage, 2002. Available at: <http://www.enotes.com/public-health-encyclopedia/total-suspended-particles-tsp> [accessed: 9/14/2011]

Stratus (2008). *Recommended health benefit assessment methods for the 2007 AQMP socioeconomic assessment – Final report*, prepared by Stratus Consulting, Leland Deck, for the South Coast AQMD, Diamond Bar, CA, July, 2008.

US EPA (1999). *Guideline on data handling conventions for the PM NAAQS*. [Available from Office of Air Quality Planning and standards, EPA-454/R-99-008, April 1999].

US EPA (2004). *Air Quality Criteria for Particulate Matter*. National Center for Environmental Assessment-RTP Office, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711; report no. EPA/600/P-99/002bF, October 2004. Available at: http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_cr_cd.html.

US EPA (2008). *BENMAP Users Manual Appendices*, September, 2008. Prepared for the Office of Air Quality Planning and Standards, USEPA by Abt Associates, Inc.

US EPA (2009). *Integrated Science Assessment: Final Report..* National Center for Environmental Assessment-RTP Division, Office of Research and Development, U.S. EPA, Research Triangle Park, NC 27711. report no. EPA/600/R-08/139F. December 2009. Available at: http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_2007_isa.html.

Westerling, AL; Hildago HG, Cayan DR, Swetnam TW (2006). "Warming and earlier spring increase U.S. forest wildfire activity," *Science* 8/18/2006, Vol. 313 no. 5789 pp. 940-943.

Appendix A. Uncertainty of 24-hour design value trend estimates

The District design value for either the annual or 24-hour PM_{2.5} NAAQS is the maximum of 3-year averages. Its distribution does not lend itself to off-the-shelf statistical tests. We've developed two test methods that should be appropriate – one a test based on permuting years, the other a generalized regression.

Permutation test for District design values

Here we use a generalization of the ratio-permutation method described in Section 4.2. Effectively, what that method does is to consider permutations of years. This same approach can be applied to the District design value.

Let $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{kt})$ be a vector of statistics from k sites for year t (the annual means or 98th percentiles). We compute $w_{it} = x_{it}/(1-r)^t$, creating a vector $\mathbf{w}_t = (w_{1t}, w_{2t}, \dots, w_{kt})$. For each site, we compute its design values for the 1st 3 years and last 3 years and, from these, compute the District design value for the 1st 3 years, W_E , and the last 3 years, W_R , and then compute the ratio W_E/W_R . We then take permutations of the vectors \mathbf{w}_t , that is, permuting years, so that, for example, the vector for 2007 is chosen as the permutation representing 1999 ($\mathbf{w}_{1999}^* = \mathbf{w}_{2007}$), the vector for 2002 is chosen as the permutation representing 2000 ($\mathbf{w}_{2000}^* = \mathbf{w}_{2002}$), and so on. For each of these permutations, we compute W_E^* and W_R^* , and W_R^*/W_E^* . We find a 90% confidence interval (r_l, r_u) by with the 5th and 95th percentiles. So, for PM_{2.5}, where there are 18,480 permutations, we find r_l so that W_R/W_E equals the 924th smallest of the W_R^*/W_E^* , and r_u so that W_R/W_E equals the 924th largest of the W_R^*/W_E^* .

Regression method for District design values

The trends in 24-hour 98th percentiles and design values were made by simple linear regression against year, that is, fitting a straight line through the data. Using simple linear regression to fit the line presents no problem provided the trends are roughly linear, as they appear to be.

But difficulty arises evaluating the uncertainty in the estimates of the design value. Simple linear regression makes several assumptions about the data, one being that the values are statistically independent, or at least uncorrelated. But design values are based on running 3-year averages, so that design values less than 3 years apart use overlapping measurements and hence are not independent. For example, the 1999-01 design value uses the 1999, 2000 and 2001 98th percentiles, the 2000-2002 design value also uses the 2000 and 2001 98th percentiles along with the 2002. Design values three or more years apart can be assumed to be statistically independent. (Those three years apart share one winter, so their 98th percentiles may be slightly correlated but likely not much.)

Now, if this were the only issue, it could be taken care of relatively simply, since a running 3-year mean is a linear transformation that meshes smoothly with the linear regression model. But the design values are means on the original scale, not the log scale, and the log of the means doesn't equal the mean of the logs, in general. Furthermore, the District design value is the maximum of the sites' design values, and the site representing the maximum can switch from one 3-year period to the next.

Here is one approach to estimating design value uncertainty. This is intended to be applied to the District's design value, but could also be used for individual sites. Let the design value for the years t-2, t-1, and t be D_t , and let $V_t = \ln(D_t)$. Assume a linear trend in the mean of V_t :

$$E(V_t) = \beta_0 + \beta_1(t - t_0)$$

where t_0 is chosen as the average value of the years. Assume that the variance of V_t is constant: $\text{Var}(V_t) = \sigma_0^2$. Moreover, assume that the 1st and 2nd order autocovariances are $\text{Cov}(V_t, V_{t+1}) = \sigma_1$ and $\text{Cov}(V_t, V_{t+2}) = \sigma_2$, and that $\text{Cov}(V_t, V_{t+k}) = 0$ for $k > 2$. Let Σ be the variance covariance matrix of V_1, \dots, V_n , so that Σ has entries $\Sigma_{tt} = \sigma_0^2$ on the diagonal, $\Sigma_{t+1} = \sigma_1 = \Sigma_{t+1t}$ one off the diagonal, $\Sigma_{t+2} = \sigma_2 = \Sigma_{t+2t}$ two off the diagonal, and $\Sigma_{st}=0$ elsewhere. The regression coefficients are estimated as

$$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)' = (X'X)^{-1}X'V$$

where X is a $n \times 2$ matrix with a first column of 1's, and a second column with $t - t_0$. The variance-covariance matrix of $\hat{\beta}$ is

$$(X'X)^{-1}X'\Sigma X(X'X)^{-1}. \quad (\text{A1})$$

So, we need to estimate Σ , i.e., σ_0^2 , σ_1 , and σ_2 . Here we use the regression residuals $r_t = v_t - (\hat{\beta}_0 + \hat{\beta}_1(t - t_0))$. The idea is to use

$$SSE_0 = \sum_{t=1}^n r_t^2, \quad SSE_1 = \sum_{t=1}^{n-1} r_t r_{t+1}, \text{ and } SSE_2 = \sum_{t=1}^{n-2} r_t r_{t+2},$$

find the expected values of each of these sums in terms of σ_0^2 , σ_1 , and σ_2 , then equate the expected values with the sample values and solve for the unknown σ_0^2 , σ_1 , and σ_2 .

$$\begin{aligned} E(SSE_0) &= E \mathbf{r}'\mathbf{r} = E\mathbf{V}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{V} = \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\Sigma] \\ &= \sigma_0^2 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')] + \sigma_1 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{K}] + \sigma_2 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{L}] \end{aligned} \quad (\text{A2})$$

where $\text{tr}()$ is the trace, $\mathbf{K} = (k_{ij})$, with $k_{ij} = 1$ if $|i-j|=1$ and 0 otherwise, $\mathbf{L} = (l_{ij})$ with $l_{ij} = 1$ if $|i-j|=2$, 0 otherwise.

$$SSE_1 = \mathbf{r}'\mathbf{M}'_{-1}\mathbf{M}_n\mathbf{r} = \mathbf{V}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-1}\mathbf{M}_n[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{V}$$

where \mathbf{r} is the vector of residuals, \mathbf{M}_{-1} is an $(n-1) \times n$ matrix created by removing the 1st row from the $n \times n$ identity matrix \mathbf{I} , and \mathbf{M}_n is the matrix created by removing the last row of \mathbf{I} . Then

$$\begin{aligned} E(SSE_1) &= E\mathbf{r}'\mathbf{M}'_{-1}\mathbf{M}_n\mathbf{r} = E\mathbf{V}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-1}\mathbf{M}_n[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{V} \\ &= \text{tr}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-1}\mathbf{M}_n[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\Sigma \\ &= \sigma_0^2 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-1}\mathbf{M}_n[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']] + \sigma_1 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-1}\mathbf{M}_n[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{K}] + \sigma_2 \\ &\text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-1}\mathbf{M}_n[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{L}] \end{aligned} \quad (\text{A3})$$

$$SSE_2 = \mathbf{r}'\mathbf{M}'_{-2}\mathbf{M}_{(n-1)}\mathbf{r} = \mathbf{V}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-2}\mathbf{M}_{(n-1)}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{V}$$

where M_{-2} is an $(n-2) \times n$ matrix created by removing the 1st 2 rows of I and $M_{-(n-1)}$ is the matrix created by removing the last 2 rows of I . Then

$$\begin{aligned}
 E(SSE_2) &= E\mathbf{r}'\mathbf{M}'_{-2}\mathbf{M}_{-(n-1)}\mathbf{r} = E\mathbf{V}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-2}\mathbf{M}_{-(n-1)}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{V} \\
 &= \text{tr}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-2}\mathbf{M}_{-(n-2)}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\Sigma \\
 &= \sigma_0^2 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-2}\mathbf{M}_{-(n-1)}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \\
 &+ \sigma_1 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-2}\mathbf{M}_{-(n-1)}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{K} \\
 &+ \sigma_1 \text{tr}[(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{M}'_{-2}\mathbf{M}_{-(n-1)}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{L}
 \end{aligned} \tag{A4}$$

Although messy, this is straightforward to compute, leaving us with 3 linear equations in 3 unknowns A_2 , A_3 , and A_4 , to estimate σ_0^2 , σ_1 , and σ_2 :

$$\begin{aligned}
 a_{00}\hat{\sigma}_0^2 + a_{01}\hat{\sigma}_1 + a_{02}\hat{\sigma}_2 &= SSE_0 \\
 a_{10}\hat{\sigma}_0^2 + a_{11}\hat{\sigma}_1 + a_{12}\hat{\sigma}_2 &= SSE_1 \\
 a_{20}\hat{\sigma}_0^2 + a_{21}\hat{\sigma}_1 + a_{22}\hat{\sigma}_2 &= SSE_2
 \end{aligned}$$

For $n = 10$, the table of a_{ij} is:

Table of a_{ij} for $n = 10$.

	j=0	j=1	j=2
i=0	8	-1.6	-1.2121
i=1	-3.2	5.8666	-2.29818
i=2	-2.4242	-2.29818	5.76514

Note that $a_{00} = 8$, i.e., the well-know result that $E(SSE_0) = (n-2) \sigma_0^2$.

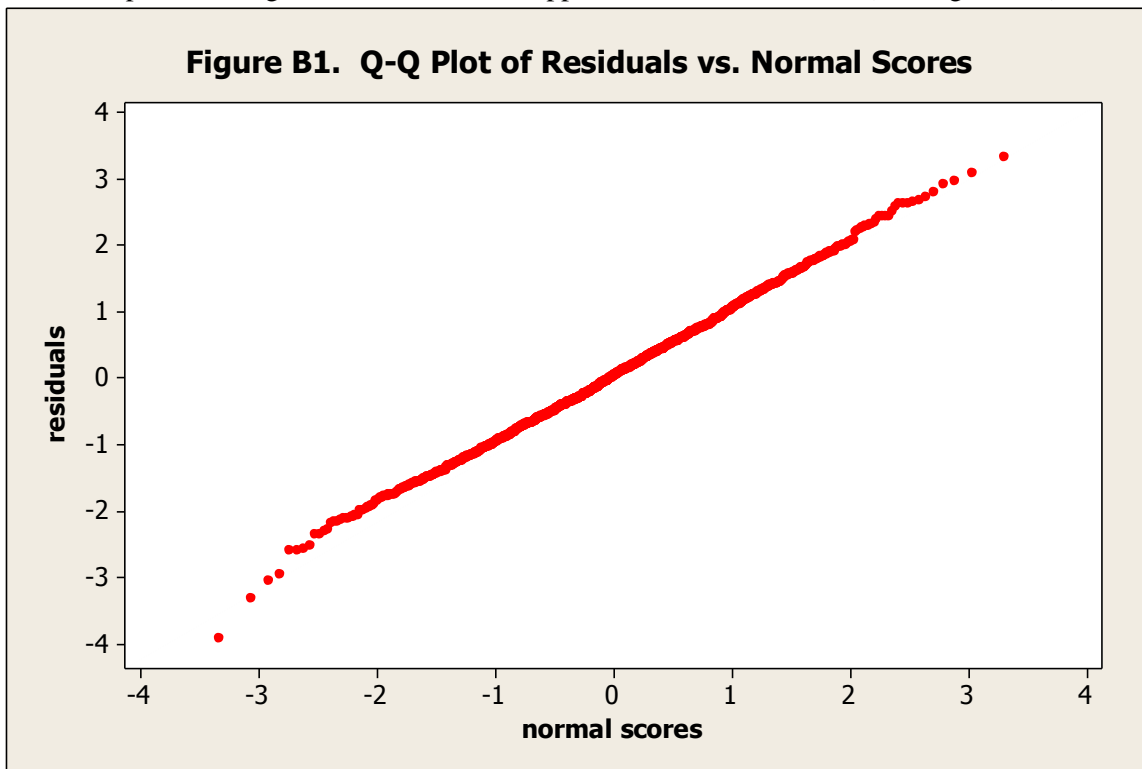
The observed values for the SSE's were: $SSE_0 = 0.021718$, $SSE_1 = 0.008228$, and $SSE_2 = -0.0007205$. Solving yields: $\hat{\sigma}_0^2 = 0.0048977$, $\hat{\sigma}_1 = 0.0036650$, and $\hat{\sigma}_2 = 0.0023658$. This allows us to estimate the standard deviations of $\hat{\beta}_0$ and $\hat{\beta}_1$ using term A1: 0.0391 and 0.01205. The fitted value of the slope was $\hat{\beta}_1 = -0.0589$, so a 90% confidence interval is $-0.0589 \pm t_{8,95} (0.01205) = -0.0803$ to -0.0365 or, roughly, reductions between 4% and 8%.

Appendix B. Methodology for meteorological variables considered in wintertime regression

Methodology for estimating meteorological potential

The key idea of the method is for each winter, w , to find a formula, f_w , that predicts what $PM_{2.5}$ concentrations would be expected from a given set of meteorological variables (e.g. wind speeds, wind directions, rainfall). Then we cycle through the meteorology of each of the winters, computing what f_w predicts for each. This provides an estimate of the number of exceedances that would be expected in winter w , averaged across the different winter seasons. That is, we are trying to estimate what would have happened in winter w under a variety of weather conditions. We can then look at these expected exceedances across winters $w = 1999-00, 2000-01, \dots, 2010-11$, and again look at the trend. Here are the steps in the process:

Step 1. Choose basic model and set of meteorological and other predictors using all winters pooled. For the dependent variable we started with the daily maximum $PM_{2.5}$ concentration. We found that the natural log transform produced regression residuals that appear close to Gaussian¹⁵. (See Figure B1.)



Step 2. Fit the model for each winter, w , yielding prediction functions, f_w , and regression standard deviations s_w .

¹⁵ We first tried Weibull regression, assuming that the regression errors were better modeled with a Weibull distribution, as demonstrated by Cox and Chu (1993) for ozone data. But Q-Q plots of the sorted residuals r_k vs. $\ln(-\ln(k/(n+1)))$ was not straight. Regressions on the original scale have the usual problems of heteroskedasticity and skewness. We found that a log transform resulted in a remarkably strong correlation with normal scores (see Figure B in Appendix B) – indicating that the residual distribution of these $PM_{2.5}$ data are well-approximated by a log-normal distribution.

Step 3. Apply f_w to each winter v 's meteorology, and get predictions $p_{1wv}, p_{2wv}, \dots, p_{n_wv}$ for the $\ln(\text{PM}_{2.5})$ where $n = 120$ or 121 depending on whether winter v contained a February 29th.

Step 4. Estimate the expected number of exceedances from each formula f_w . That is, estimate the average across different winters' meteorologies of the exceedance potential of winter w : Estimate the probability of exceedance for each day j : $\Phi[(p_{jwv} - \ln(35.5))/s_w]$, $j = 1, \dots, n_v$, where $\Phi(x)$ is the Gaussian cumulative distribution function. The value 35.5 is chosen based on the $35 \mu\text{g}/\text{m}^3$ standard; any $\text{PM}_{2.5}$ measurement greater than this exceeds the standard. The expected number of exceedances for a given winter v is just the summation of these probabilities:

$$\sum_{j=1}^{n_v} \Phi\left(\frac{p_{jwv} - \ln(35.5)}{s_w}\right).$$

Step 5. We define the *exceedance potential* of a winter, w , as the average across all winters, v , of the expected number of exceedances:

$$\text{Exceedance potential for winter } w = \frac{1}{12} \sum_{v=1}^{12} \sum_{j=1}^{n_v} \Phi\left(\frac{p_{jwv} - \ln(35.5)}{s_w}\right)$$

Table B1. **Meteorological Variables Considered**

Variable	Description
days from Jan 1	number of days between the date and January 1 (e.g., Jan 5 = 4, Dec 29 = 3)
avews	mean of 24-hour average wind speeds from San Carlos, Pleasanton & Bethel Island
1/avews	the reciprocal of avews
3dyavews	the 3-day average of avews
1/3dyws	the reciprocal of 3dyavews
nd<5	the number of consecutive days (up to 3) where the wind speed was < 5 mph
sjrain	san jose rainfall in inches
sjrn<.02	=1 if sj rainfall was < 0.02 inches, =0 if sj rainfall \geq 0.02 inches
pb-result	san pablo 24-hour resultant wind direction
pbres<90	=1 if pb-result was < 90°, =0 if pb-result was \geq 90°.
funwdsd	standard deviation of Fort Funston wind direction
fuws&pb<90	= funwdsd if pb-result < 90, =0 otherwise
we+hol	= 1 if a weekend or holiday, =0 if weekday
r1day	Scott's r1 day
r2day	Scott's r2 day
r3day	Scott's r3 day
vday	Scott's v day
zday	Scott's z day
950mb height	Height in meters of the 950 mb pressure measured during the 4am RAOB sounding
1000mb height	Height in meters of the 1000 mb pressure measured during the 4am RAOB sounding

Table B2. Exceedance potential by winter formula and meteorology

		Formula for winter:												Condu- -cive- -ness:
Winter	1999- 2000	2000- 2001	2001- 2002	2002- 2003	2003- 2004	2004- 2005	2005- 2006	2006- 2007	2007- 2008	2008- 2009	2009- 2010	2010- 2011		
Meteorology of winter:	1999-00	20	25	18	17	17	18	9	19	11	10	7	3	14.6
	2000-01	28	37	26	29	27	27	15	28	17	16	14	5	22.5
	2001-02	16	22	14	13	13	14	6	15	6	6	3	1	10.7
	2002-03	24	34	22	25	23	23	12	24	13	11	10	3	18.8
	2003-04	15	23	14	16	14	14	6	15	6	6	4	1	11.0
	2004-05	21	32	19	22	20	19	9	20	9	8	6	2	15.5
	2005-06	19	26	17	20	18	18	10	19	11	9	8	3	14.8
	2006-07	18	25	17	17	16	17	8	18	9	9	7	2	13.6
	2007-08	13	18	12	10	11	12	5	12	6	5	3	1	9.0
	2008-09	27	38	25	28	26	26	13	28	14	13	11	3	21.2
	2009-10	23	33	21	24	22	22	10	23	11	10	8	2	17.4
2010-11	22	32	21	22	21	21	9	22	10	9	7	2	16.5	
Exceedance Potential:	20.5	28.8	18.8	20.3	18.9	19.4	9.5	20.3	10.2	9.3	7.3	2.3		

Appendix C. Numbers of days with PM₁₀ Measurements

Table C. Numbers of days with PM₁₀ measurements at District monitoring sites

	fr	li	cc	bi	sr	np	sf	rc	sj	va	st	8+days*
1989	47	60	60	60	60	60	61	60	178	0	0	56
1990	61	61	60	61	61	61	61	61	174	0	0	56
1991	59	60	59	60	61	60	60	60	168	0	0	57
1992	61	61	61	61	61	61	61	61	61	0	0	60
1993	61	61	59	61	61	61	61	61	61	0	0	56
1994	60	61	60	60	61	59	61	61	61	9	31	59
1995	61	61	61	61	61	61	61	61	61	61	61	61
1996	61	61	61	61	61	61	61	61	61	61	61	61
1997	61	61	61	61	61	60	61	61	61	61	61	61
1998	61	61	61	61	61	60	61	61	60	61	61	61
1999	60	60	60	60	60	60	60	60	60	60	59	59
2000	61	62	61	61	61	61	61	61	61	61	60	61
2001	60	61	61	61	61	61	61	61	61	61	61	61
2002	61	61	61	61	61	61	61	61	14	61	61	61
2003	61	61	61	61	61	61	61	61	61	61	62	61
2004	61	61	58	61	61	61	61	63	61	60	61	61
2005	61	61	21	61	61	64	61	63	61	59	61	61
2006	66	61	61	60	61	64	61	63	61	60	61	61
2007	60	60	58	60	60	61	59	61	60	58	60	60

* Number of days with at least 8 sites with PM₁₀ measurements.

Appendix D. Use of dichotomous PM_{2.5} data for trend analysis

Analysis of long-term PM_{2.5} trends is hampered by the lack of routinely collected PM_{2.5} before 1999. However, some PM_{2.5} measurements were made at one site, San Jose, from 1990 through 1997. The sampling equipment was different, however, and this complicated the analysis.

Although there are no dichot PM_{2.5} overlapping with routine filter measurements, dichot PM₁₀ measurements overlap with FRM PM₁₀. Figure D1 shows a comparison by year for 1990 through 1997. First note that the correlation is quite high, over 90%. Next notice that the points lie near the line $y = x$ for 1990 and 1991 but fall below it for 1992-1997. A regression through the origin produced a slope of 0.98 for 1990-1991 and 0.78 for 1992-1997. Table D1 shows a comparison of mean values on days when both instruments made measurements. The ratios of the means average 0.99 for 1990-1991 and 0.78 for 1992-1997.

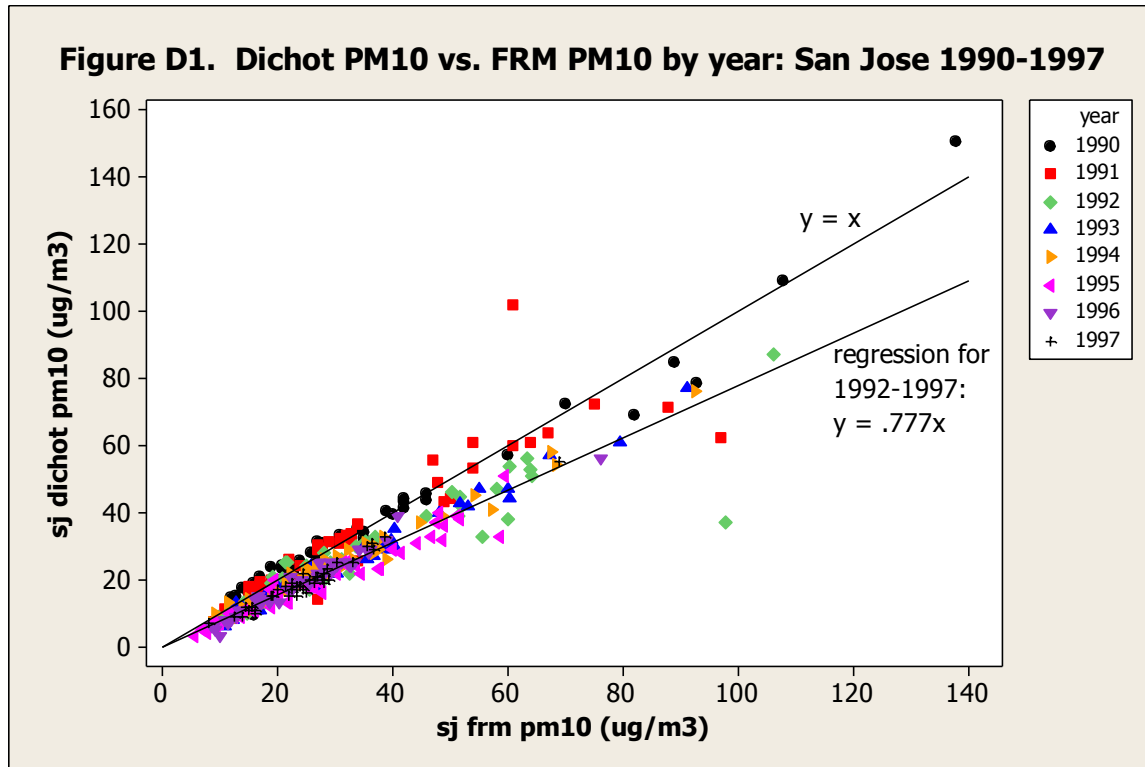
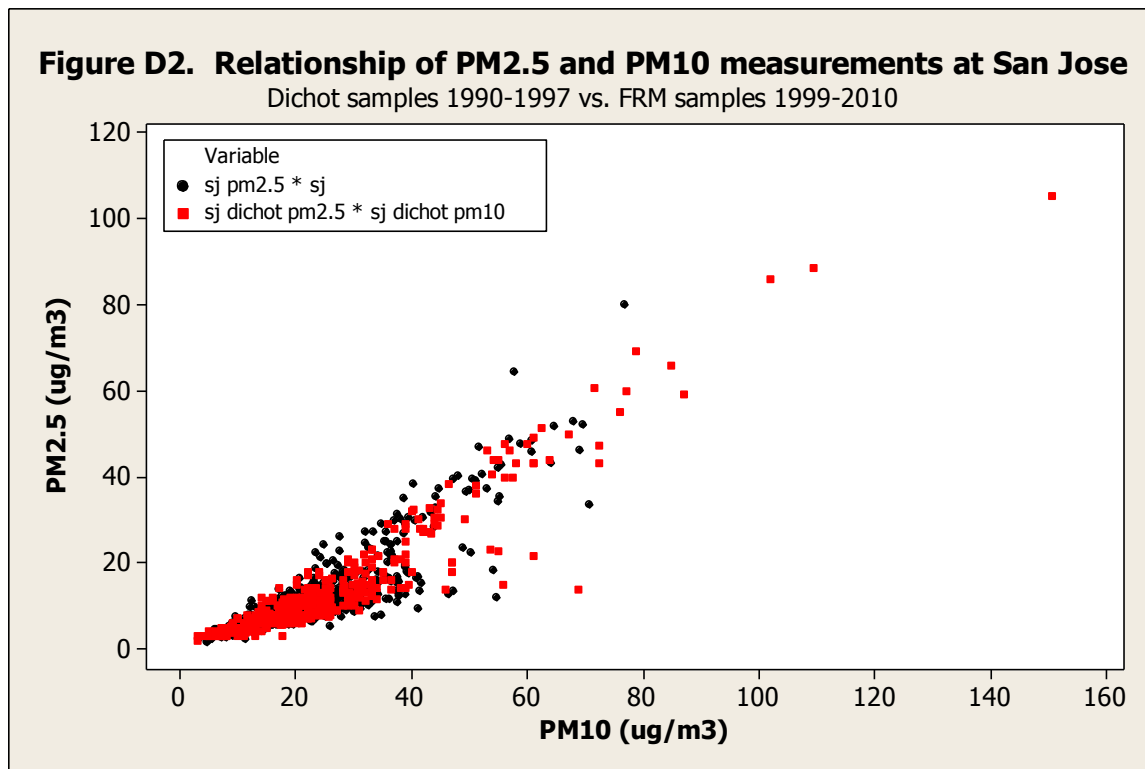


Table D1. FRM and Dichot PM₁₀ annual means on days when both had measurements

	sj pm10	dichot pm10	dichot/frm ratio
1990	33.54	33.29	0.99
1991	31.49	30.82	0.98
1992	32.94	26.39	0.80
1993	30.51	24.29	0.80
1994	28.46	23.33	0.82
1995	26.02	18.71	0.72
1996	24.53	19.15	0.78

1997	25.21	19.32	0.77
------	-------	-------	------

Figure D2 shows the relationship between $PM_{2.5}$ and PM_{10} for both the dichot and the FRM samplers. The patterns are very similar, with a lower $PM_{2.5}/PM_{10}$ slope for small PM values and a larger slope for high PM values, with a few exceptions. Thus, the $PM_{2.5}/PM_{10}$ relationship is roughly the same for the two measurement methods, so that the ratio of mean $PM_{2.5}$ to mean PM_{10} appears very similar (dichot $PM_{2.5}/PM_{10} \cong$ frm $PM_{2.5}/PM_{10}$). Thus, for 1990 and 1991 we assume that the dichot $PM_{2.5}$ represents what the FRM $PM_{2.5}$ would have measured, and for 1992-1997, it appears reasonable to use the approximation $\text{mean frm } PM_{2.5} \cong \text{mean dichot } PM_{2.5} (\text{frm } PM_{10}/\text{dichot } PM_{10}) \cong \text{mean dichot } PM_{2.5}/0.78$.



Appendix E. Adjustment of nitrate measurements for the impact of refrigeration

In 2009, the District started refrigerating PM₁₀ filters after sample collection. The goal was to prevent nitrate from volatilizing before the filters could be weighed. To assess the impact of the change, we compared San Jose PM₁₀ nitrate measurements against PM_{2.5} nitrate measurements, where the filters have always been refrigerated after sample collection, as a control. The change caused by refrigeration is substantial and appears to have fixed the volatilization problem.

Figure E1 shows PM₁₀ nitrate values relative to the corresponding PM_{2.5} values, the black circles for measurements before 2009, the red squares for 2009-2010. Before 2009, many PM₁₀ nitrate values were well below the corresponding PM_{2.5} nitrate value; in 2009-2010, they lie close to the line $y = x$, i.e., close to the corresponding PM_{2.5} nitrate value.

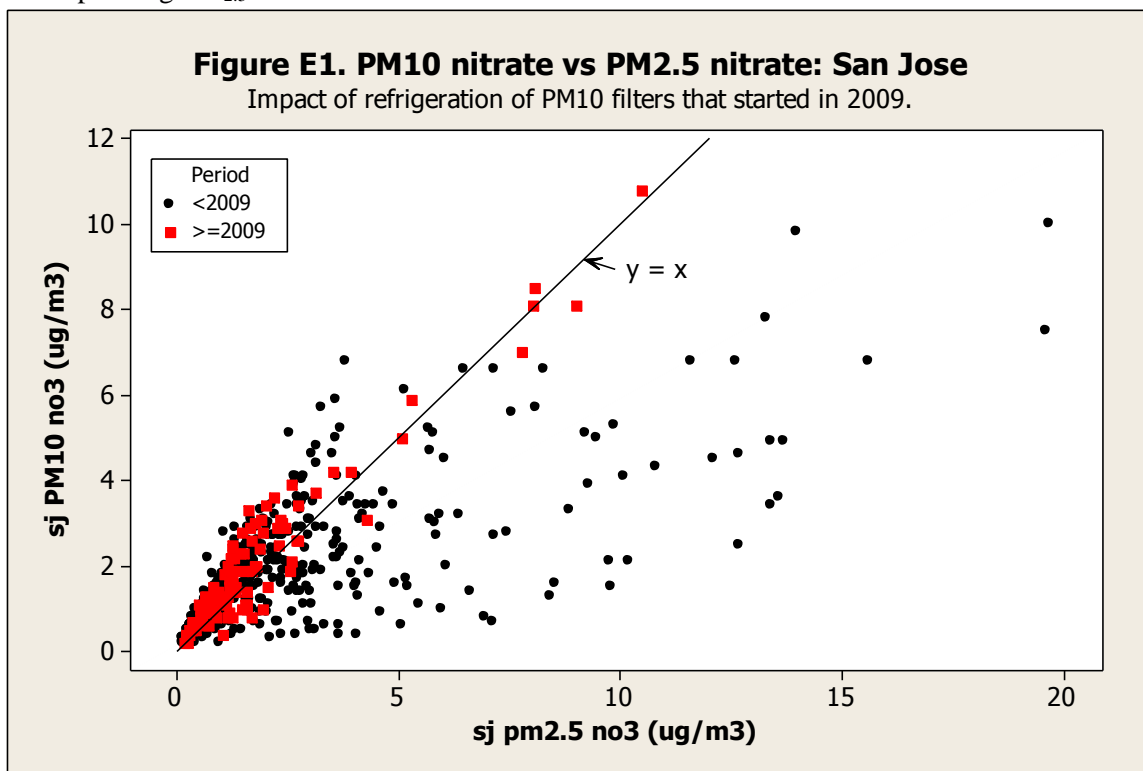


Figure E2 shows the relationship before 2009 split into winter-season (Nov-Feb) and non-winter. Simple linear regression lines for each season are also shown. The effect of non-refrigeration is apparent for both seasons, but more pronounced in the winter months (significantly more gradual slope). Also note that the effect appears non-linear, with PM₁₀ nitrate matching PM_{2.5} nitrate for small concentrations ($\leq 1 \mu\text{g}/\text{m}^3$).

Figure E3 shows the same plot but for 2009-2010. For both seasons, PM₁₀ nitrate matches PM_{2.5} nitrate, and there is no statistical difference between the seasons. The figure also shows a non-parametric fit – that is, a curve representing a best guess as to the true, underlying relation between PM₁₀ nitrate and PM_{2.5} nitrate. The curve starts at (0,0) but goes above the line $y = x$ for smaller concentrations, then descends back to the $y = x$ line. That is, PM₁₀ nitrate measurements are somewhat higher than the corresponding PM_{2.5} nitrate unless the concentrations

are high. Of course, since $PM_{2.5}$ is a subset of PM_{10} , there's no reason why they should be equal. It does raise the question of the extent that nitrate lies in the coarse fraction of PM_{10} .

Fig. E2. PM_{10} NO₃ vs. $PM_{2.5}$ NO₃: Winter and Non-Winter pre-2009

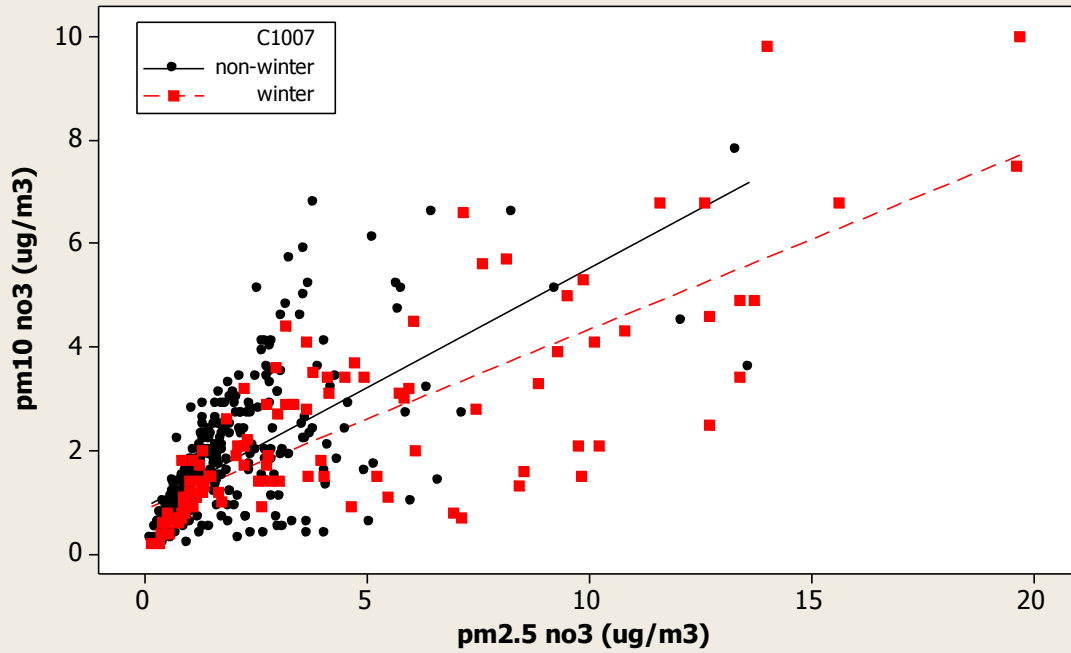
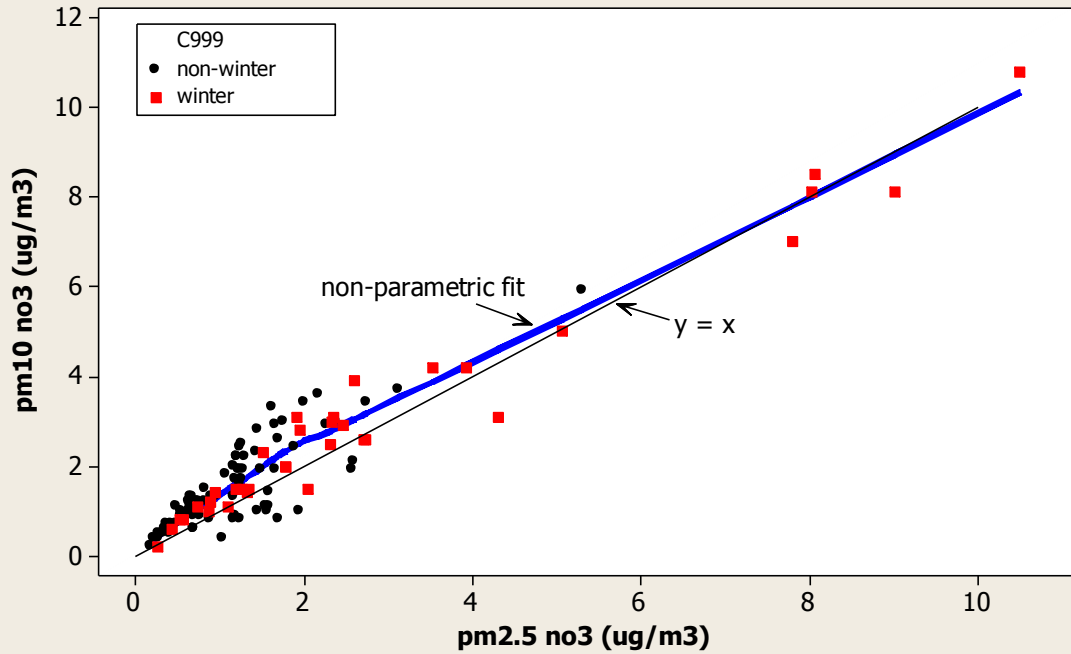


Fig. E3. PM_{10} NO₃ vs. $PM_{2.5}$ NO₃: Winter and Non-Winter 2009-10

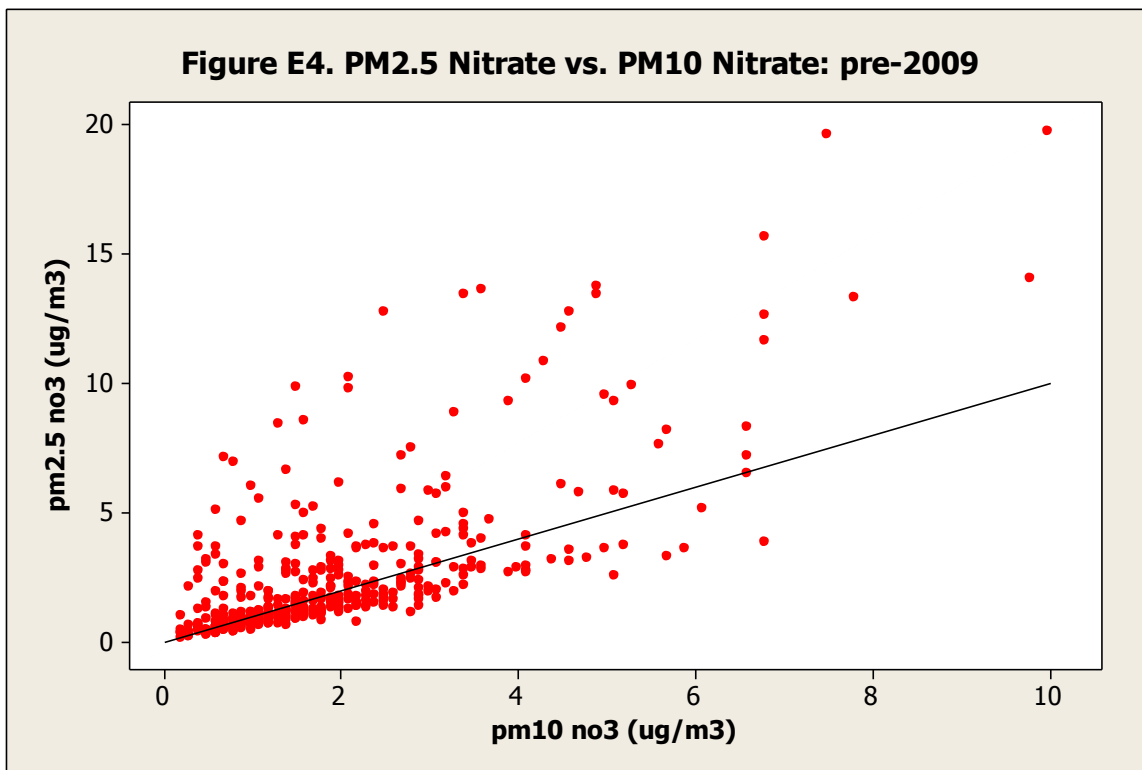


Adjusting pre-2009 nitrate measurements

Although the refrigeration of PM₁₀ nitrate filters appear to have solved one problem, it created another, namely a discontinuity in measured nitrate trends. To make past measurements closer to being comparable, we can attempt to adjust them based on the relationship with PM_{2.5} nitrate. This would also have the advantage of allowing us to get a better estimate of historical nitrate concentrations and what fraction these are of total PM₁₀.

It should be noted that we are making several assumptions – one being that the relationship between PM₁₀ nitrate and PM_{2.5} nitrate at San Jose would carry over to other sites. Also, the finding that PM₁₀ nitrate is scarcely larger than PM_{2.5} nitrate is somewhat questionable; we have been told that a larger percentage was in the coarse fraction. Also, if we adjust PM₁₀ nitrate, we presumably want to adjust PM₁₀ concentrations too, since it also is impacted by the loss of nitrate.

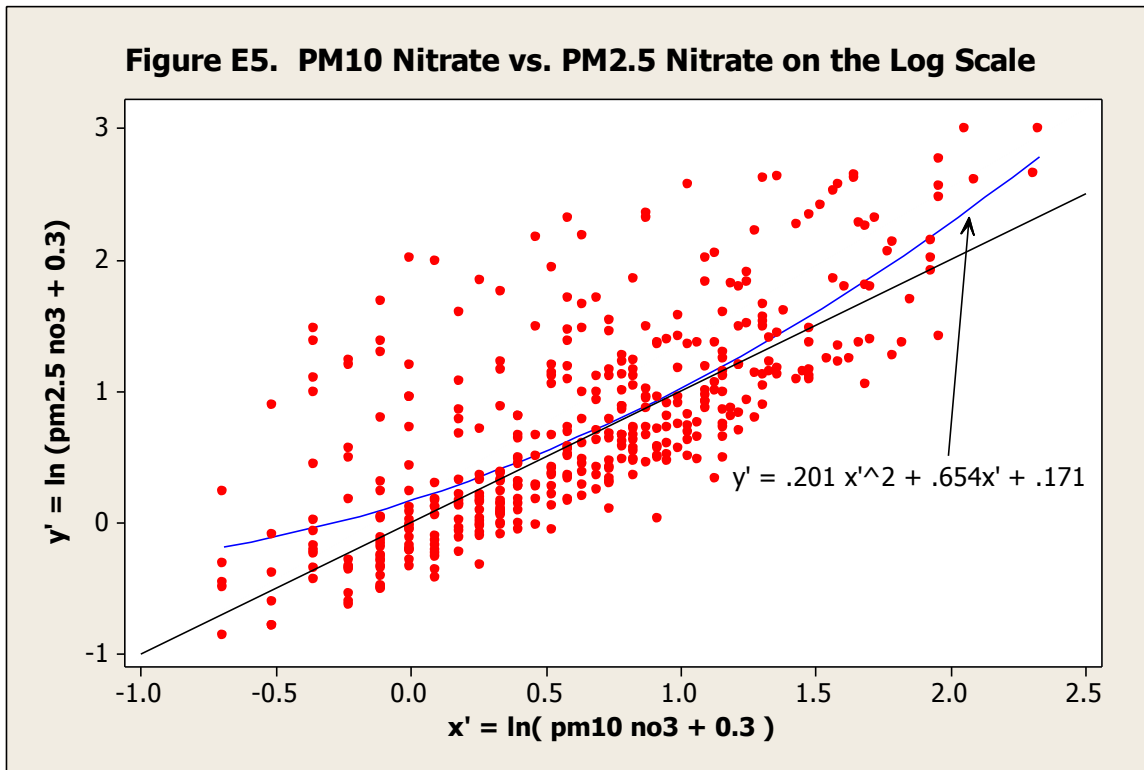
With these caveats, we plunge ahead. Figure E4 shows the pre-2009 relationship of PM_{2.5} nitrate to PM₁₀ nitrate (just flipping the axes from Figure E2). Notice that the deviations from the line $y=x$ are highly asymmetric, with many values below, but close to the line, and a few values far above the line. Also, the variability increases with increasing concentration.



Standard statistical techniques like regression are suited to Gaussian variability or, at least, symmetric and short-tailed distributions.

A log transform comes part way to solving this problem. Figure E5 shows the same data but transformed so that $x' = \ln(\text{pm10 no3} + 0.3)$ and $y' = \ln(\text{pm2.5 no3} + 0.3)$. [Adding 0.3 moderates the influence of the smallest

observations.] As previously discussed the winter and non-winter relationships are different, so separate regressions were fit.



For the winter, a simple linear regression was fit: $y' = 1.21 x' + .177$. For non-winter, a quadratic regression was fit: $y' = .202 * x'^2 + .504 * x' + .177$. In both cases, the standard error was .4945. Transforming back to the original scale:

$$\text{adjusted winter PM10 NO3} = e^{1.21 x' + .177} - 0.3,$$

and

$$\text{adjusted non-winter PM10 NO3} = e^{.202 * x'^2 + .504 * x' + .177} - 0.3.$$