APPENDIX G: HEALTH BENEFITS AND METHODOLOGY

APPENDIX G1: HEALTH BENEFITS ASSOCIATED WITH REDUCTIONS IN DIESEL PM EMISSIONS FROM OFF-ROAD DIESEL VEHICLES

APPENDIX G2: METHODOLOGY FOR ESTIMATING AMBIENT CONCENTRATIONS OF PARTICULATE MATTER FROM DIESEL-FUELED ENGINE EMISSIONS

APPENDIX G.1: HEALTH BENEFITS ASSOCIATED WITH REDUCTIONS IN DIESEL PM EMISSIONS FROM OFF-ROAD DIESEL VEHICLES

I. Overview

The estimation of premature death related to PM2.5 exposure presented below is based on a peer-reviewed methodology developed by the U.S. Environmental Protection Agency (U.S. EPA, 2010) and an incidents-per-ton (IPT) methodology (ARB, 2006). Staff estimates that approximately 1,900 tons of PM2.5 and 34,000 tons of NO_X emitted from off-road diesel vehicles will be reduced through implementation of this regulation in the years 2010-2029. The reduction of premature deaths associated with emission reductions of both primary PM and secondary PM (produced in the atmosphere from the precursor NO_X) are presented in Table 1.

Table 1: Estimate of Premature Deaths Avoided Associated with Emission Reductions from Implementation of the Off-Road Vehicle Regulation (2010-2029)*

Pollutant	# of Cases 95% C.I. (Lower Bound)	# of Cases (Mean)	# of Cases 95% C.I. (Upper Bound)
PM	190	250	310
NOx	170	220	270
Total	360	470	570

* Health effects from primary and secondary PM are labeled PM and NO_X, respectively.

Details of the methodology are described below. Sections II through VII describe how the health impact estimation methodology was developed, including development of the IPT factor and its application for the Off-Road Rule.

II. Incidence-per-ton Factors

ARB used a methodology known as the incidence-per-ton (IPT) to quantify the health benefits of regulatory controls to reduce PM2.5. This method, which was peer reviewed in 2006 (ARB, 2006), yields results similar to those of a more sophisticated modeling analysis, but can be used more efficiently. It is similar in concept to the methodology developed by the U.S. EPA for similar estimations (Fann et al., 2009). Details on the methodology used to calculate these estimates can be found in Appendix A of the Emission Reduction Plan for Ports and Goods Movement in California, as documented in the Emission Reduction Plan for Ports and Goods Movement (ARB, 2006).

The basis of the IPT methodology is the approximately linear relationship which holds between changes in emissions and estimated changes in health outcomes. This is a consequence of the following observations:

1. Across the range of ambient PM2.5 concentrations encountered in California, modeled changes in premature mortality are approximately proportional to changes in ambient pollutant concentrations.
2. For primary pollutants such as diesel particulate matter, changes in ambient concentrations are approximately proportional to changes in emissions.
3. For secondary pollutants such as ammonium nitrate aerosol, a linear relationship may be used as a first-order approximation to the relationship between ambient concentration and emissions of NO_X. There may be cases where the relationship between emission of NO_X and ammonium nitrate aerosol is greater than or less than linear.

Therefore, premature death is approximately proportional to emissions, and can be estimated by multiplying emissions by a scaling factor, the IPT factor. IPT factors are derived by calculating the premature death incidence associated with a PM2.5 source in an air basin, and dividing by the emissions of that PM2.5 source. The methodology for estimation of premature death incidence is described in Section III. The emission inventories for NO_X were developed for 2008, adjusted to account for changes due to the recession, updates to the ocean going vessels, trucks/buses, locomotives, and construction equipment models. Separate IPT factors are used for each air basin and for each pollutant. Since the total incidence of heath effects is proportional to population, the result is adjusted by the ratio of the population in the target year to the population in the base year for which the IPT factors were developed.

III. Mortality Incidence

Background

For estimating the health benefits of emission reductions brought about by implementation of this regulation, ARB applied the methodology used by U.S. EPA in the 2010 Quantitative Health Risk Assessment that was developed to estimate premature deaths associated with PM2.5 exposure. In this assessment of health benefits, ARB calculated estimates for premature cardiopulmonary death. ARB is emphasizing cardiopulmonary deaths because they are the most frequent causes of death, and category of deaths most strongly related to PM2.5 exposure (U.S. EPA, 2010).

Calculation of the premature death incidence associated with PM2.5 exposure requires baseline incidence rates, population data, ambient concentration of PM2.5, and a concentration-response function relating changes in PM2.5 exposure to changes in mortality incidence. These data are available in a spreadsheet as part of the rulemaking package. Calculations were made based on both primary and secondary PM2.5 exposure. The sources and derivation of these parameters are described in sections IV – VII.

Estimating population exposure to PM2.5

Estimation of the PM2.5 exposure is a several step process, involving estimation of the annual-average concentration at each PM2.5 monitor in the state, and interpolation of concentrations between monitors to estimate exposure for each census tract. Since diesel engines emit particles directly (primary particles), as well as gases that convert to PM2.5 through atmospheric chemical reactions (secondary particles), exposure estimates are made for both, in order to capture the full impact of diesel engines on premature death.

Population-weighted exposure to primary and secondary PM2.5 was estimated based on monitor-specific concentrations. Even with an extensive air quality monitoring network, the mortality quantification method requires estimation of exposure between monitors across a geographic area. ARB uses a standard spatial interpolation method known as inverse distance-squared weighting which was peer reviewed in 2007 (Shepard, 1968; Goodin and McRae, 1979). This method yields reasonable accuracy in estimating pollutant concentrations near monitoring stations, although when distance from the monitoring station increases, the uncertainty in the interpolated concentration also increases. This method gives more accurate estimates of concentration in areas with a large number of monitors with good spatial coverage and low variability in concentration. When data are abundant, most simple interpolation techniques give similar results (Jarvis et al., 2001). When data are sparse, however, the assumption made about the underlying variation in PM2.5 concentration, along with the choice of interpolation method and its parameters can be critical to avoid misleading results.

Aggregating results to county, air basin and state

To aggregate results from census tracts to larger geographical subdivisions such as counties or air basins, we used a GIS technique called areal interpolation. Areal interpolation is a procedure for translating spatial data from one set of geographical subdivisions to another when the boundaries do not exactly overlap. Numerous variants of the technique exist, but for the purpose of this analysis the simplest form, which uses area of polygon intersection, was employed (Goodchild and Lam, 1980; Fotheringham and Rogerson, 1994).

The precision of areal interpolation based on area of intersection depends on the relative size of the geographical subdivisions, and the homogeneity of the spatial distribution of the quantity being apportioned. In urban areas, where census tracts are small and population is distributed more evenly, areal interpolation to larger subdivisions such as air basins yields relatively precise estimates. In rural areas where the population is distributed unevenly over large census tracts, estimates are less precise.

IV. Population at the Census Tract Level

Age-resolved population data at the census tract level, for the 2000 Census, were obtained from the United States Census Bureau (U.S. Census Bureau). These were projected to 2006-2008 using age-resolved county population projections from the California Department of Finance (CDOF).

Age-specific growth factors for each county, for each year, were computed from the CDOF projections by dividing each county population for the target year by the county population for the year 2000. Since each census tract lies entirely in a county, these growth factors were applied to each census tract in the county, each age group separately. Population was projected for ten-year age groups 25-34 through 75-84, and for age 85 and older.

This method of projection reflects growth in overall county population, but does not model changes in population distribution within counties, such as expansion of urban areas into surrounding rural land.

V. Baseline Cardiopulmonary Mortality Incidence Rate

Baseline incidence rates vary by age bracket. Incidence was estimated separately for ten-year age groups 25-34 through 75-84, and age 85 and older.

Baseline cardiopulmonary mortality incidence rates were estimated at the county level from individual death records for the year 2005, obtained from the California Department of Public Health (CDPH). Cardiopulmonary mortality was defined as ICD9 codes 161-187 and 192-214.

The county of residence of the decedent was generally not recorded. However, the Federal Information Processing Standards (FIPS) city code and the ZIP code were usually recorded. The FIPS city code unambiguously identifies the county, but was sometimes invalid, unrecorded, or recorded as “unknown”. When the FIPS code was not available it was sometimes possible to identify the county from the ZIP code, but ZIP codes can overlap multiple counties. In cases where 90% or more of the area of the decedent’s zip code lay entirely within a county, the death was assigned to that county. A handful of records included invalid dates. The breakdown of records was as follows:

County identified by FIPS code	231,181	96.6%
County identified by ZIP code	4,196	1.8%
Unidentified or invalid	3,851	1.6%

Because the county could not be determined for 1.6% of the records, the incidence is slightly underestimated. No adjustment was made to compensate for excluded records.

In some cases the cardiopulmonary incidence was extremely low, because some counties only have a population of a few thousand, and the population is further subdivided into age groups. In such cases the variability of the incidence is high. However, since this represents a very small fraction of California's population the effect on statewide mortality estimates is negligible. Large counties show little year-to-year variability.

Baseline mortality rates are subject to other sources of uncertainty. For example, the baseline incidence level is treated as uniform throughout the county of interest. In addition, baseline incidence rates can change over time as lifestyles, health care, income, and other factors evolve.

VI. Annual Diesel PM Concentrations

Annual diesel PM concentrations were estimated indirectly from annual average NO_X concentrations by applying a conversion factor, α, of 0.022. Derivation of the conversion factor is described in Appendix G.2. Annual average NO_X concentrations for 2006-2008 were retrieved from ARB's air quality database (ARB ADAM).

Because there is statistical uncertainty in α, this uncertainty carries through in the mortality estimates from diesel PM. The overall uncertainty in the mortality incidence associated with diesel PM exposure was estimated by taking into account both the variability in α and the uncertainty in the β coefficient of the concentration-response function (described in Section III). Since this is very difficult to estimate analytically, parametric bootstrapping was used to obtain uncertainty estimates.

Parametric bootstrapping (Efron and Tibshirani, 1993) is a technique for estimating statistical properties of a complicated function of parameters with uncertain values. Probability distributions are assumed for the parameters, random samples are drawn from those distributions, and the function is calculated for each choice of parameters. The statistical properties of interest, in this case the confidence interval, can then be estimated from the values the function takes over the random sample.

In this case, α and β were assumed to have normal (Gaussian) distributions. The conversion factor α was assumed to have a mean of 0.022 and standard deviation of 0.05 as described in section VII. The mean of the β coefficient was taken to be 0.01293, and the standard deviation was back-calculated from the 0.95 confidence interval given in Appendix C of (U.S. EPA, 2009) as $(0.01587 - 0.01007) / (1.96 \times 2) = 0.001480$.

Using the R statistical programming environment (CRAN), 100,000 sets of values for α (one for each NO_X monitor) and values for β were randomly generated. The corresponding cardiopulmonary incidence for each choice of parameters was then computed. Since the sample exhibited a fairly symmetrical distribution, we computed the 0.025 and 0.975 quantiles of the sample to estimate the confidence interval.

VII. Annual PM2.5 ammonium nitrate concentrations

In addition to directly emitted PM, diesel exhaust contains NO_X, which is the precursor to nitrates, secondary diesel-related PM formed in the atmosphere. Secondary PM can lead to additional health impacts beyond those associated with directly emitted diesel PM. To quantify such impacts, staff developed annual average ammonium nitrate concentrations for 2006-2008. The concentrations were computed from ambient nitrate ion concentrations, using PM10 data combined from two sources: the regular air quality monitoring network and the IMPROVE visibility network (IMPROVE Visibility Network).

The ARB and air pollution control districts operate a network of PM10 monitors around the state, mostly in urban areas (ARB AQMN). PM10 samples are collected as 24-hour filter samples, once every six days, using size-selective inlet (SSI) sampler. Samples from some monitors are further analyzed to determine the concentration of nitrate and other constituents. During 2006-2008, nitrate data were available from 45 monitors. Since nitrate particles form a fine aerosol, essentially all of the nitrate mass falls into the PM2.5 fraction, so the PM10 nitrate concentration may be regarded as equivalent to PM2.5 nitrate concentration. SSI data were retrieved from ARB's ADAM air quality database (ARB ADAM).

In addition to urban PM10 nitrate monitoring, the national IMPROVE visibility network operated 18 PM2.5 nitrate monitors, mainly in national parks and other remote locations (IMPROVE Visibility Network). These instruments collect one sample every three days. The IMPROVE samplers are more efficient than the SSI samplers, and tend to recover a higher fraction of ambient ammonium nitrate than the SSI samplers. However, since the IMPROVE samplers are located at remote locations where PM2.5 concentrations are close to natural background levels, the effect of instrument bias is considered negligible, and the data were treated as equivalent to the SSI data. IMPROVE data were retrieved from the project web site (IMPROVE Visibility Network).

Daily samples were aggregated by monitor to obtain annual averages. In order to avoid potential seasonal bias due to missed samples, the samples were aggregated into quarterly means, and the four quarterly means were averaged to obtain annual means. For a quarterly average to be considered valid, the data were required to be at least 75% complete. For a year to be considered valid, all four valid quarters were required.

To convert from nitrate ion concentration to ammonium nitrate (NH_4NO_3) concentration, the annual averages were multiplied by the ratio of the molecular weight of ammonium nitrate to that of the nitrate ion, 1.29.

References

ARB ADAM. Air Quality Database web site. Available at: http://www.arb.ca.gov/adam/

ARB AQMN. Air Quality Monitoring Network web site. Available at: http://www.arb.ca.gov/aqd/aqmoninca.htm

CDOF. California Department of Finance population projection web site. Available at: http://www.dof.ca.gov/research/demographic/reports/view.php

CDPH. California Department of Public Health statistics web site. Available at: http://www.cdph.ca.gov/data/statistics/Pages/default.aspx

CRAN . The Comprehensive R Archive Network. Available at: http://cran.r-project.org/.

Efron E, Tibshirani R. 1993. An Introduction to the Bootstrap. Chapman and Hall. New York.

Fann N, Fulcher CM, Hubbell BJ. 2009. The influence of location, source, and emission type in estimates of the human health benefits of reducing a ton of air pollution Air Quality. Atmosphere & Health. 2:169-176.

Fotheringham S, Rogerson P, editors. 1994. Spatial Analysis and GIS. London. Taylor and Francis.

Goodchild M, Lam N. 1980. Areal interpolation: a variant of the traditional spatial problem. Geo-processing, 1: 297-312.

Goodin, WR, McRae, GJ, Seinfeld JH. 1979. A Comparison of Interpolation Methods for Sparse Data: Application to Wind and Concentration Fields. J Applied Meteor. 18:761-771.

IMPROVE Visibility Network web site. Available at: http://vista.cira.colostate.edu/improve/Overview/Overview.htm

Jarvis, CH Stuart, N. 2001. A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part !!: The interaction between number of guiding variables and the type of interpolation method. J. Appl. Metero. 40, 1075-1084.

Shepard D. 1968. A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 ACM National Conference. Pp. 517–524.

US Census Bureau. American Fact Finder web site. Available at: http://factfinder.census.gov/home/saff/main.html?_lang=en

U.S. EPA. 2009. Integrated Science Assessment for PM (Final Report). Available at: http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_2007_isa.html

U.S. EPA. 2010. Quantitative Health Risk Assessment for Particulate Matter Available at: http://www.epa.gov/ttn/naaqs/standards/pm/data/PM_RA_FINAL_June_2010.pdf

APPENDIX G.2: METHODOLOGY FOR ESTIMATING AMBIENT CONCENTRATIONS OF PARTICULATE MATTER FROM DIESEL-FUELED ENGINE EMISSIONS

Introduction

This appendix outlines a method to estimate annual average concentrations of diesel particulate matter (DPM) over large spatial scales. It consists of a simple variation of receptor modeling, which uses measurements of ambient chemical concentrations to infer source contributions, known as the tracer species method.[1] A basic assumption in this method is that the ambient concentration of a tracer species, *C*, may be used alone to infer the ambient concentration of a pollutant from a specific source, *S*:

$$S = \alpha C, \qquad (1)$$

where α is a scaling factor that is assumed to be independent of geographical location. In the estimation of DPM, we take *C* to be the annual-average ambient concentration of NO_x and *S* to be the annual-average ambient concentration of DPM less than 2.5 µm in diameter (DPM2.5). The factor α relates the concentration of PM2.5 produced by diesel-fueled engine emissions to the concentration of NO_x produced by all sources (NO_x is not a unique tracer for diesel emissions). In the following section, we demonstrate that estimates for α based on the emission inventory (EI) and on source apportionment (SA) studies agree within calculated uncertainties. We approximate the distribution of α values over counties by a Gaussian distribution with mean 0.023 and standard deviation 0.006 for the year 2000. Using data for 2008, we estimate a mean of 0.022 and standard deviation of 0.0005. The 2008 values are appropriate for estimating health impacts based on recent NO_x data. A single value for α and associated dispersion may be used to infer DPM concentrations from measurements of ambient NO_x concentrations in all air basins.

Background

The primary interest of the California Air Resources Board in the estimation of ambient DPM concentrations is for assessment of premature deaths. For this purpose, annual-average ambient concentrations of DPM are needed. These values are used to calculate lifetime average daily doses[2]; multiplication of the average daily inhalation dose over 70 years with a cancer potency factor gives inhalation cancer risk estimates. In previous estimates[3] of diesel particulate matter less than 10 µm in diameter (DPM10) concentrations, the Air Resources Board (1998) used a method based on ambient total PM10 concentrations. For this approach, one of two factors, rural or urban - which were determined from a combination of chemical mass balance source apportionment studies (CMB) and emission inventory estimates (EI) - was used to scale ambient PM10 measurement values to obtain estimates of DPM10 concentrations. Air basins that had DPM10 to total PM10 emissions (emission inventory estimates) greater or less than the base case had these DPM10 estimates scaled by another factor (that was determined from the EI): the ratio of air basin to base case value of the relative DPM10 to total PM10 emissions. Application of this method, therefore, depends on several elements, the most important of which are: 1) measurements of ambient PM10 concentrations, 2) previous source apportionment work in specific air basins (base cases), and 3) emission inventory estimates. These components are also the primary weaknesses of the method: 1) PM10 contains predominantly crustal material, and the fraction associated with diesel PM is very small - at most approximately 6.5%; 2) early CMB studies may not be as accurate as more recent organic marker species-based CMB methods; and 3) early emission inventory estimates may not be as accurate in accounting for all source emissions as more recent models.

We believe the use of scaled ambient NO_x concentrations is more direct than the PM10 method to estimate DPM concentrations for several reasons, foremost among these are the close linkage[4] of diesel-engine emitted NO_x to total emitted NO_x – about half of total NO_x emissions are from diesel sources – and the relatively good correlation of ambient with recent emission inventory estimates for α. In addition, NO_x emissions in California are predominantly from mobile sources, with diesel vehicle emissions accounting for approximately half the on-road mobile contribution and greater than 60% of the off-road contribution. The limited variation in different air basins of the diesel engine emissions contribution to total NO_x is a reflection of the similarity of on- and off-road fleet composition and activity. In this respect, California likely differs from other regions of the country in the scarcity of important point stationary sources (such as power plants and refineries). Contributions from such point sources would introduce proximity dependencies and preclude the use of a simple NO_x-scaling methodology to approximate DPM.

Methods

In this section, we develop estimates for the scaling factor α, the ratio of ambient DPM to total ambient NO_x concentrations. First, we compare the ratio of ambient concentrations $(DPM/NO_x)_{SA}$ from several source apportionment (SA) studies done in the late 1990's with the ratio of annual emissions $(DPM/NO_x)_{EI}$ from the calendar year 2000 emission inventory (EI). Currently, the source apportionment studies are considered the best available methods for determining ambient DPM concentrations (at selected monitoring sites); agreement between the SA and EI estimates of α is used to support the use of a single α value for the whole state of California. Second, based on this favorable comparison, we use the distribution of county emission inventory estimates for the $(DPM/NO_x)_{EI}$ to determine an average and standard deviation for α for a baseline year 2008.

In the following section, we estimate the ratio of DPM to NO_x concentrations in ambient air for two year-long and several short-term source apportionment modeling studies with co-located NO_x measurements. These studies utilize specific organic chemical species, or molecular markers, in chemical mass balance (CMB) apportionment of PM, which is considered to be essential for the accurate separation of gasoline from diesel-fueled engine emissions. Even with such molecular markers, however, CMB modeling relies on several important, though typically unverifiable, assumptions. First, all chemical species used in the CMB mass balance equations must be conserved during transport from source to receptor site. That is, these species in the particle phase must not be either depleted through volatilization or chemical reaction or increased through atmospheric reactions of precursor species. Second, all major sources that contribute chemical species used in CMB apportionment must be chemically characterized and included in the model, and the source profiles constructed must be sufficiently different from one another that problems of collinearity are not encountered. Recent work has investigated how well these assumptions are fulfilled in practice.[5] In these studies, several challenges in application of the CMB method have been determined: emissions from vehicles exhibit significant profile-to-profile variability, which causes difficulties in construction of a single source profile; molecular markers undergo significant oxidation across a range of experimental conditions, which indicates that similar aging effects may affect apportionment studies; CMB analysis is very sensitive to source profile selection; and well defined source profiles may not exist for some emission categories. Another source of uncertainty in the apportionment of PM to diesel emission sources in all these studies is in the off-road diesel source contribution. These sources are captured by CMB modeling only to the extent the emissions are similar in chemical composition to those of on-road diesel trucks, for which source profiles are available. In light of the emission inventory estimate that approximately half the diesel contribution to PM and NO_x is from off-road sources, this poorly understood aspect of SA modeling warrants further qualifications in all CMB estimates of DPM.

The first PM source apportionment work considered in this document was a year-long effort conducted as part of the Children's Health Study (CHS 1995), in which James Schauer carried out organic chemical PM CMB studies for eleven sites in the South Coast Air Basin.[6,7] Hence, eleven annual average values for DPM10 and NO_x concentrations are available from this work. Four of the sites are centrally located, while the rest are in more or less outlying areas. The second considered SA study was carried out as part of the California Regional Particulate Air Quality Study (CRPAQS 2000) by Desert Research Institute (DRI) in the San Joaquin Valley.[8] From this work, seven estimates of annual average DPM2.5 and associated NO_x are available. Most of these sites are in urban areas (with the exception of Bethel Island). Although J. Chow of DRI used a different methodology to measure elemental and organic carbon (IMPROVE method) than used by J. Schauer for CHS (NIOSH method), DRI utilized similar specific organic chemical markers for combustion sources. In addition to these long-term measurements, side-by-side, organic marker CMB modeling was done by E. Fujita and J. Schauer at two sites for one week each in Southern California in 1999 as part of the Diesel-Gasoline Particulate Matter Split Study (DGPM 2000).[9,10,11]

An unexpected result from the Diesel-Gasoline Particulate Split Study is that apportionment of PM2.5 depends on the specific carbon measurement method utilized (to determine relative organic/elemental carbon). Such differences in apportionment are currently not incorporated into uncertainty estimates. We also note that the DGPM study raised several important, but still unresolved, questions in the interpretation of CMB modeling results. Specifically, SA estimates may be very sensitive to the choice of source profiles used; e.g. the characteristics of the "average" driving cycle, categories of vehicles, composition of the fleet (e.g. inclusion of high emitter categories such as gasoline "smoker" vehicles) and, information about average high emitter organic species emissions. These aspects bear directly upon SA attribution estimates in a poorly understood manner. Results from several recent short-term apportionment studies that do not utilize CMB modeling are also included below; these studies provide further evidence for a wide range of DPM estimates. Based on a favorable comparison of SA and EI results, we develop an estimate of the DPM/ NO_x ratio based on the EI.

Results

Source apportionment of PM collected in the South Coast Air Basin was done by J. Schauer as part of the Children's Health Study (CHS) in 1995. The sampling sites are described in the CHS Final Report and represent eleven communities in the South Coast Air Basin; these include four urban sites (Mira Loma, North Long Beach, Riverside, Upland), two sites in a mountainous region (Alpine and Lake Arrowhead), one desert site (Lancaster), three rural coastal sites (Atascadero, Lompoc, and Santa Maria), and one rural inland site (Lake Elsinore). NO_x measurements and filter samples (organic chemical marker measurements) were taken at the same locations. Although each filter PM sample was collected over a two week interval, filters from each site were composited into three seasonal time periods. Each composited sample was analyzed for organic marker compounds and utilized in chemical mass balance source apportionment modeling. We concentrate on using annual average results from the apportionment study, and show seasonal trends below. Figure 1 shows site-to-site variation of source apportionment estimates of the ratio (annual average DPM10 concentration)/(annual average total NO_x concentration) from the CHS (1995). A straight average over all eleven sites of the ratio DPM10/NO_x, gives the mean value as 0.030 (0.010), where here and in the following text the value in parentheses denotes the standard deviation. An alternative estimate based on regression of DPM10 concentrations against ambient NO_x concentrations (over eleven sites) gives 0.027 (0.001); see Fig. 2. In this, and all following regressions, the intercept is set to zero, which makes the regression less sensitive to scatter and is physically meaningful, as one expects that diesel emissions tend to zero with total NO_x emissions.

As expected, the dispersion in α is much larger over individual measurements of DPM/NO_x than it is for the regression coefficient. It is unclear which choice of error is best for use in personal exposure estimates that use population weighting. The site-specific DPM/NO_x values, Fig. 1, are best estimates for local DPM/NO_x ratios, though specific meteorology and lack of population weighting may emphasize unrepresentative values. Similarly, DPM/NO_x ratios obtained from linear regression (with zero intercept) are highly influenced by data with large NO_x and/or DPM values. Because individual measurements for the ratio DPM/NO_x retain site-specific variability in concentrations, we believe the statistics from individual measurements are better estimates than regression coefficients for DPM exposure-related work. We take the standard deviation of the distribution of DPM/NO_x values as the measure of uncertainty in α for SA studies.

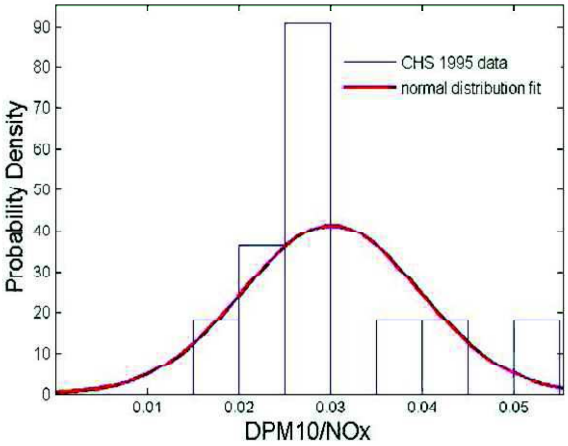

Figure 1

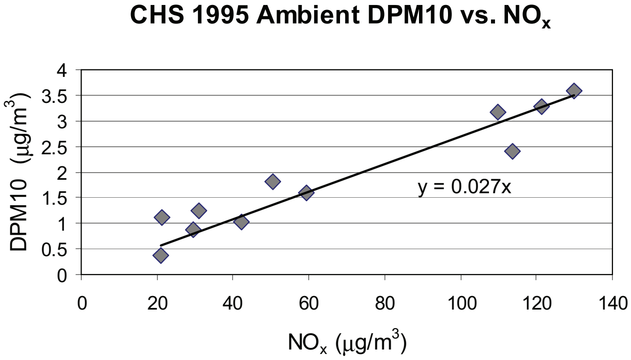

Figure 2

Figure 3 shows DPM10 vs. NO$_x$ plots of the CHS data for each of the three composited seasons (January – April; May – October; November – December). The slope exhibits a clear seasonal dependence with largest value in summer and smallest in winter. This variation can not be explained completely by EI estimates (summer, winter), which show much less seasonal variability, and indicates further sources of uncertainty in the use of short timescales for scaling NO$_x$.

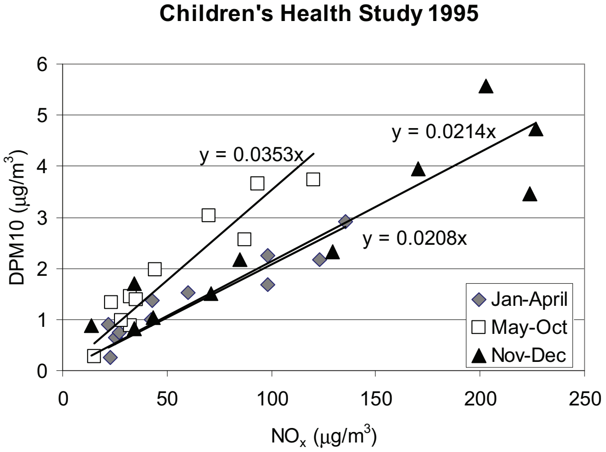

Figure 3

The other year-long SA estimate for α is from CRPAQS (DRI, 2000) for the San Joaquin Valley (SJV). As in the CHS, this study also utilizes organic chemical species CMB for apportionment, though it departs in many details of application, which affect apportionment results. Several aspects which bear directly on the results are usage of different carbon analysis methods (the National Institute for Occupational Safety and Health (NIOSH) method in CHS and the Interagency Monitoring of Protected Visual Environments (IMPROVE) protocol, also known as thermal/optical reflectance (TOR), in CRPAQS), different chemical markers and source profiles (see refs. 6-8). A straight average of the ratios of SA DPM2.5 to NO_x concentrations for seven sites in the SJV gives a value of 0.018 (0.008). Figure 4 shows a regression of SA ambient DPM against NO_x, which gives a slope of 0.016 (0.004). As for the previous SA work, we take the standard deviation (0.008) from the distribution of DPM/NO_x values as an indicator of the variability in ambient ratios.

We note that the relative variability of DPM/NO_x in both studies is very large: the coefficient of variation (standard deviation/average) ≈ 0.5. We believe this large uncertainty in SA estimates best captures local variation of source composition, mixing, chemical reactions and other factors. Hence, this order of uncertainty is expected in any estimate of DPM based on ambient NO_x concentrations.
A recent short-term SA modeling study investigated the sources of uncertainties in the relative contributions of diesel and gasoline vehicle emissions to PM2.5 in the South Coast (2001) – the Gasoline/Diesel PM Split Study.[9,10,11] In this work, James Schauer (University of Wisconsin, Madison) and Eric Fujita (Desert Research Institute) collected samples side-by-side for sources (57 light duty gasoline and 34 heavy duty diesel vehicles) and ambient air (two sites Los Angeles, N. Main, and Azusa), and carried out independent chemical and SA data analyses. The DRI study did not directly apportion PM2.5 to diesel sources; instead, apportionment of total carbon and elemental carbon were determined.

To estimate corresponding DPM2.5 concentrations, we scale reported diesel EC values by a factor determined from DRI's diesel source profile (DPM2.5 = 1.36 DEC2.5, where DEC2.5 denotes elemental carbon in PM2.5 apportioned to diesel engine emissions). The resulting SA results show a lack of agreement between DPM2.5 estimates: apportionment of PM to diesel emission sources by the two groups differ by approximately a factor of two; see Figures 5 and 6. Estimates for DPM2.5/NO_x are: 0.010 (0.003) for Schauer and 0.023 (0.004) for Fujita. Because ambient and vehicle emission samples were collected side-by-side, these results indicate that the disparity in DPM estimates is driven by differences in source apportionment methodology, which includes differences in carbon measurement methods (NIOSH vs. IMPROVE), organic marker chemical species, and source profiles for vehicles. Without *a priori* information about which method is more accurate, we believe both estimates should be weighted equally, giving DPM/NO_x = .017 (.009).

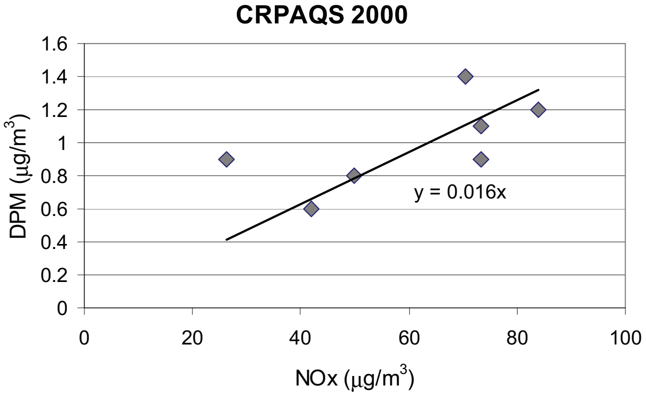

Figure 4

Two recent studies that used simpler methods to apportion PM2.5 to diesel sources were done by Livermore National Laboratory (LLNL) in 2007 and ARB's Monitoring and Laboratory Division (MLD) in 2003. LLNL utilized fossil carbon measurements, based on Carbon 14, and MLD utilized n-octadecane as a diesel tracer. In these studies, DPM concentrations were estimated to be on the order of 1 µg/m^3; these values would presumably support a lower DPM/NO_x ratio of .01 (with a likely relative uncertainty of 50%). More specifically, LLNL determined the average fossil elemental carbon (FEC) at Wilmington as approximately 1.05 µg/m^3 (based on the limited data), and the average FEC at Roseville as approximately 0.65 µg/m^3. Assuming that all FEC is from diesel emissions and that OC emissions from diesels are small in comparison, these values may be considered upper bound DPM concentrations. MLD's study yielded estimates of DPM for Wilmington as 1.2 µg/m^3 and Sacramento as 0.8 µg/m^3, and the statewide average as 1.0 µg/m^3. These values, however, differ by over a factor of two from the recent MATES III organic marker CMB estimate of >3 µg/m^3 in 2004-2005 (in Wilmington). Therefore, while these two independent estimates provide support for the lower end of DPM/NO_x ratio, considerable uncertainty remains in their interpretation.

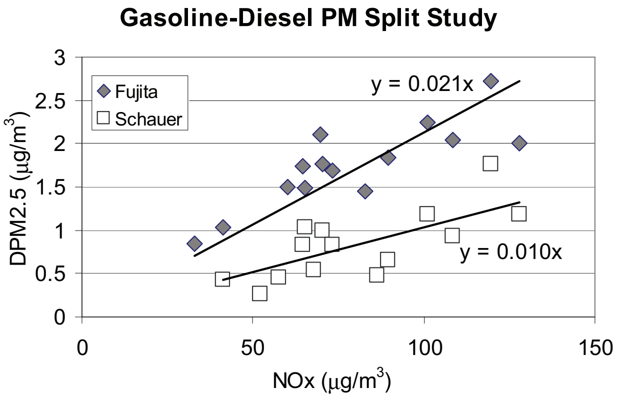

Figure 5

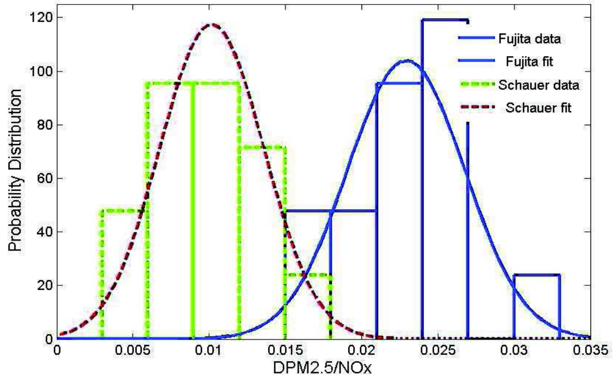

Figure 6

A comparison of the above SA estimates with the emission inventory can not be made directly, as emission inventory estimates are for whole counties while SA estimates are specific to monitoring sites and implicitly take into account meteorology, chemistry and deposition. Hence we compare average values for DPM/ NO_x from the previous SA studies with EI estimates of DPM to total NO_x emission ratios. For this purpose, the EI estimates for DPM and total NO_x emission rates for individual counties are utilized.[12] These estimates may be visualized as tons of pollutants emitted each day into a well-mixed box covering each county, with removal rates of DPM and NO_x proportionately the same. The assumption of approximately equal removal rates is difficult to verify, given that the rates are caused by a variety of processes: deposition, chemical reactions, and flow into and out of air basins. Further, while the atmospheric lifetimes for DPM and NO_x are typically very different (greater and less than a few days, respectively), which would bias the ratio of DPM/ NO_x toward higher values, the mean residence time of an air parcel in a coastal air basin is typically less than a day, which would dominate the reaction and deposition rates and effectively make the rate of removal for $\mathrm{NO_x}$ and PM the same. In the following, we assume this dominance of air parcel residence time on removal rates, and take the removal rates for $\mathrm{NO_x}$ and DPM as equal.

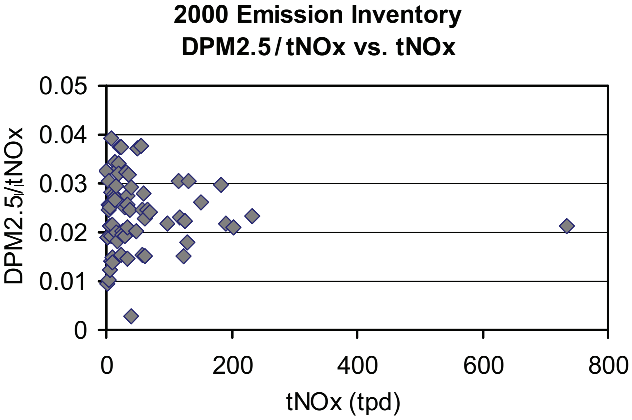

Figure 7

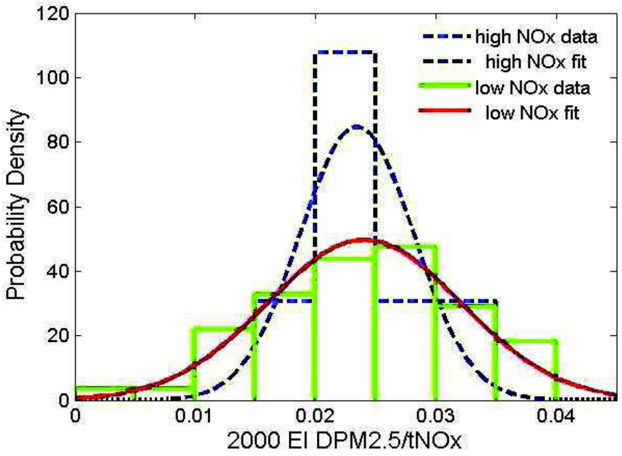

Figure 8

To compare the above source apportionment estimates of $\mathrm{DPM/NO_x}$ with emission inventory estimates, we utilize ARB emission inventory estimates for the year 2000 (the SA studies were conducted in 1995 CHS in SoCAB, 2000 CRPAQS in SJV, and 2001 GDPM in SoCAB). The emission inventory estimates incorporate spatial and temporal averaging over large scales and therefore may be used to estimate average ambient $\mathrm{DPM/NO_x}$ ratios directly (in the following figures for the emission inventory, we abbreviate total $\mathrm{NO_x}$ by $\mathrm{tNO_x}$). A plot of $(\mathrm{DPM/tNO_x})_e$ against $\mathrm{tNO}_{x\,e}$ for all counties in California is shown in Fig. 7. This scatter plot shows that the county-wide ratios $\mathrm{DPM/tNO_x}$ are clustered about an average and that the dispersion depends on the average annual $\mathrm{tNO_x}$ emission rate. The plot also shows that a separation of high-$\mathrm{NO_x}$ from lower-$\mathrm{NO_x}$ emission counties occurs with a division around an annual average of 80 tons per day.

(High-NO_x counties are listed in the next section of this document.) In California, NO_x inventories are dominated by on-road and off-road mobile sources; overall, diesel engine emissions contribute approximately half of the on-road NO_x emissions and greater than 60% of the off-road mobile emissions. To better capture exposure-related estimates of DPM/tNO_x, each county value is weighted by its population; weighted histograms are approximated by normal distributions. Figure 8 shows histograms for the ratio of DPM/tNO_x over high- and low-NO_x emission counties. The mean and standard deviation for these distributions are: 0.023 (0.003) for the high-NO_x counties and 0.023 (0.006) for the low-NO_x counties. Hence, population weighted distributions for α in high and low-NO_x counties may be described by normal distributions with same mean value and a dispersion that depends on NO_x emissions characteristics. To develop a single California-wide approximation, we take the (larger) dispersion for the ratio of DPM/tNO_x in low-NO_x counties as measure of the variability that is encountered locally within air basins.

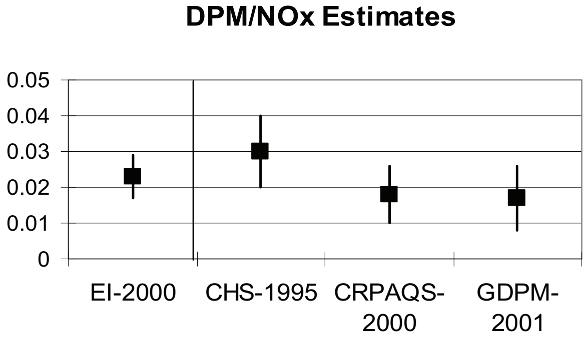

Figure 9

The above estimates of the ratio DPM/tNO_x from the 2000 EI population-weighted and SA studies compare well, given the relatively large uncertainties: 2000 EI county average 0.023 (0.006); and SA: 1995 CHS 0.030 (0.010), 2000 CRPAQS 0.018 (0.008), and 2001 GDPM 0.017 (.009). Figure 9 shows a plot of these estimates. This overall agreement in α values motivates adoption of a single scaling factor for the whole state of California for years close to 2000: the average from the 2000 EI estimates, α = 0.023 (0.006).

Conclusions

Based on the relatively good agreement between SA and EI estimates for the scaling factor α - the ratio of DPM to total NO_x - for years close to 2000, we propose the use of a single value of 0.023 for α in estimating the population-weighted annual average ambient DPM concentration in California. The resultant DPM estimates from this approach depend upon the network of ambient NO_x measurements from the ARB monitoring sites. In the following, we outline a method to calculate such averages. First, the annual average DPM concentration at each monitoring site is estimated as the product of annual average NO_x concentration value and α. The uncertainty associated with this DPM estimate is the product of the annual average NO_x measurement value and the low- NO_x county standard deviation, .006. The following twelve counties are considered high-NO_x emission counties (annual average NO_x > .80 tons per day): Los Angeles, San Bernardino, Kern, San Diego, Orange, Riverside, Alameda, Fresno, Santa Clara, Contra Costa, San Joaquin, and Sacramento. The remaining 46 counties are considered low-NO_x emission counties. From this set of spatially discrete DPM concentration estimates a smooth DPM concentration surface may be constructed using kriging or other interpolation methods. In remote areas without monitoring sites, the smoothing method may be modified to incorporate a minimum concentration, which would reflect a nonzero background value (or such areas may be removed, if the population is sufficiently small). Second, census data for California is used to approximate a population density surface (population fraction per unit area) and the (pointwise) product of the population density and DPM concentration surfaces is taken. This product may be integrated over any region and divided by the fraction of California population within that region to give a population-weighted average DPM concentration; in particular, integration of the product may be performed over the state to give an average population-weighted ambient DPM concentration. Once ambient diesel PM concentrations have been estimated for a baseline year (2000), linear rollback techniques may be used to project concentrations for future years.

A comparison of DPM concentration estimates for the year 2000 using the proposed NO_x-scaling method with the projections from the previous PM10-scaling method[3] is given in Table 1. The overall agreement between DPM concentration estimates for the six highest population air basins is very good. More specifically, the six highest population air basins contain over 90% of the population of California and contribute greater than 96% of the population weighted DPM concentration; in each of these air basins, the difference between the proposed and the previous DPM concentrations is less than 20% of the previous estimate. It should be noted that the previous estimates use a baseline year 1990 and are projected forward by a decade based on linear rollback, and so do not constitute the best approximation for year 2000. Greater variation of agreement between proposed and previous methods is found for lower population air basins. Many factors contribute to this variability, several of which are: the larger dispersion in the DPM to NO_x ratio (.006), uncertainty in application of PM10 scaling method to regions less similar to the SJV, and greater influence of localized emission sources. Altogether, the proposed, population-weighted DPM concentration for California is increased by 11% over the previous estimate. This high level of agreement between the population-weighted DPM estimates gives confidence that the proposed method is consistent with the previous technique and represents a viable approach to estimate DPM exposure.

A final application of the NO_x-scaling approach is to estimate α for a more recent year - 2008 (to be used as a baseline for health impact studies). An analysis of an updated ARB 2008 emissions inventory, similar to that carried out for 2000, yields the value $\alpha = 0.022$ (0.005). This result is based on an unweighted fitting of the distribution of county DPM2.5/NO_x ratios to a normal distribution; see Fig. 10. [Emissions from ocean going vessels were down-weighted by an adjustment factor, 0.10 in the South Coast and 0.25 in other areas, to reflect their decreased contribution to inland areas. [14]] As expected, because of the relatively slow turnover of diesel engines, the mean and standard deviations are similar to those for 2000. The low- and high-NO_x counties, exhibit the same α values: $\alpha = 0.022$ with (0.006) for low-NO_x counties and (0.003) for high-NO_x emission counties (the demarcation between low- and high-NOx counties was taken as 60 tpd NO_x). It is expected that α will slowly vary with year due to improvements in technology and turnover of emission sources, though large changes may be expected based on preferential DPM control measures. Such divergences will necessitate reexamination of estimates for α; although time intervals for such reviews can not be prescribed, revisions in α are necessary when the difference in estimates approaches the uncertainty (dispersion).

A rough comparison of this 2008 EI estimate may be made with measurements from recently completed field work – the Harbor Community Monitoring Study (HCMS).[13] This program was conducted in 2007 to characterize the spatial variations in concentrations of toxic air contaminants (TACs) and their co-pollutants within the communities of Wilmington, West Long Beach, and San Pedro in California's South Coast Air Basin. These communities were chosen because of the close proximity of residents to many emission sources, which include the Ports of Los Angeles and Long Beach, petroleum refineries, intermodal rail facilities and the greatest concentration of diesel traffic in the Los Angeles metropolitan area. To avoid sites heavily impacted by near sources, we concentrate on neighborhood sites from the saturation monitoring network operated by the Desert Research Institute (denoted LWIN, WCOL, WGUL, WLAK, WMAR, WMCD and WMCF in the report). This study consisted of 7-day

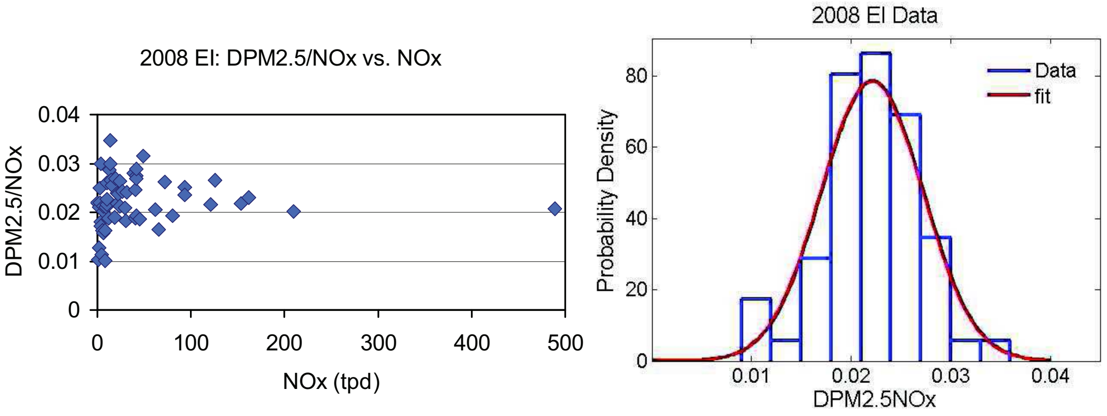

Figure 10

time-integrated sampling for four consecutive weeks in four seasons. NO_x was measured with Ogawa passive samplers and PM2.5 was collected on 7-day integrated Teflon and quartz filters with Airmetrics MiniVol samplers. Organic carbon (OC) and elemental carbon (EC) were analyzed by thermal optical reflectance (TOR) method using the IMPROVE (Interagency Monitoring of Protected Visual Environments) temperature/oxygen cycle (IMPROVE TOR). Several caveats should be noted in the interpretation of this data. First, the high density of emission sources in the HCMS area may produce ambient NO_x and DPM concentrations that are different from those in the greater region, and hence, less well described by EI estimates. Second, source apportionment was not carried out for this study; instead, the tracer used for diesel-engine PM emissions is elementary carbon (EC) from PM2.5, which is predominantly but not exclusively from diesel emissions. Under these limitations, DPM2.5 was estimated from EC concentrations as 65% of the total carbon from diesel engine emissions (DTC): EC=0.65 DTC. This conversion is adopted because the carbon fraction of EC in heavy-duty diesel truck emissions during city suburban driving has been measured as approximately 0.65 (City Suburban Heavy Vehicle Route dynamometer cycle, ref. 10) and the general agreement between TC collected on quartz filters and PM2.5 mass collected on Teflon filters for diesel emissions in dynamometer testing (Gasoline/Diesel PM Split Study).[9,10] A full consideration of various methods to estimate DPM using EC, and possibly other co-pollutants, was not attempted; such an investigation would yield a confidence interval for DPM estimates, and possibly provide a better case for comparison. A different scaling method was adopted by DRI in the Final Report for the HCMS saturation study.[11] In their work, ambient diesel particulate carbon (DPC) concentrations for each site (for a given season) were taken as directly proportional to measured EC concentrations, where the proportionality constant was determined from the slope of the correlation between total carbon and EC at the near road sampling locations. These slopes varied between 1.5 and 2.2. Diesel particulate matter was then estimated as the sum of the elemental carbon and organic matter portions of DPM:

$$\mathrm{DPM} = \mathrm{EC} + 1.46\,(\mathrm{DPC} - \mathrm{EC}), \qquad (2)$$

where 1.46 is the ratio of diesel particulate organic matter to DPC from the Gasoline/Diesel PM Split dynamometer testing of diesel trucks by DRI. A third estimate of DPM may be obtained from a method used in the Multiple Air Toxics Exposure Study III (MATES III) - a monitoring and evaluation study conducted in the South Coast Air Basin (Basin) from April 2004 to March 2005. In this study, a scaling factor for EC of 1.72 was developed based on the 2005 emission inventory ratio of diesel particulate to elemental carbon emissions. Application of these three DPM estimation methods to measurements from the neighborhood sites yields the following average values for the ratio of DPM concentration to NOx concentration: 0.038 (0.005), 0.025 (0.004), and 0.028 (0.004), for the DRI, EC/0.65, and MATES III methods, respectively. These results show that all sites exhibited similar DPM to NO_x ratios and that the largest source of uncertainty is in the choice of scaling factors for EC (and total carbon). The largest value of DPM/NO_x was obtained using DRI's methodology, which may be regarded as an estimate of the "upper-bound ambient concentration" of DPM.[12] In light of the caveats and simple EC-scaling, these estimates agree relatively well with the 2008 EI estimate of 0.022 (.005) and provide further support for use of a NO_x-scaling methodology (with associated uncertainty interval) to estimate DPM in California.

In summary, the proposed method to estimate ambient DPM concentrations has distinct advantages over the previous PM10 method as well as several important limitations. The primary strengths of the method include the significant connection between DPM and NO_x, simple application, estimates of uncertainty intervals, and ability to capture sub-county variations in DPM concentrations. In addition to these strengths, the approach is tied directly to the ARB emission inventory, and links bottom-up EI estimates with top-down SA estimates. Several limitations and caveats also bear on applications of the method. The limitations include all assumptions sufficient for application of EI estimates to ambient air, such as well-mixed air parcels on county-wide scales, proportional removal rates for NO_x and DPM (including air basin outflow), proportionally time-uniform emission rates for all NO_x and DPM sources, etc. Verification of these assumptions is in general not possible; instead, agreement between EI and SA estimates is taken as best available evidence for support. The uncertainty intervals produced by the estimation method are based on variations between low-NO_x counties and reflect differences in relative emission sources (primarily diesel vs. non-diesel mobile sources). As such, the uncertainty describes the confidence in α to accurately describe local NO_x emission sources. For areas outside California, in which the NO_x emission inventory has a significant contribution from non-mobile sources (e.g. power plants or refineries), the value of α is likely to be different from that for California and vary with source proximity. Further work is needed in strengthening the understanding of the contribution of various emission sources to ambient concentrations of both gases and particles. In this respect, source apportionment work that utilizes organic marker species is the probably best available approach; ideally, highly time-resolved studies would allow better characterization and support for single species scaling estimates, such as the NO_x-scaling method. Finally, off-road diesel sources, which are a large source of uncertainty in current CMB modeling, need to be included explicitly in future source apportionment studies.

References

1 Henry R., Lewis C., Hopke P., and Williamson H., *Review of Receptor Model Fundamentals* (1984) Atmospheric Environment 18, 1507-1515.

2 *Air Toxics Hot Spots Program Risk Assessment Guidelines, Appendix D Risk Assessment Procedures to Evaluate Particulate Emissions from Diesel-Fueled Engines, 2003,* Office of Environmental Health Hazard Assessment.

3 *Estimate of Ambient PM10 Concentrations from Directly Emitted Emissions from Diesel Engines, Appendix III, Part A, Exposure Assessment from Proposed Identification of Diesel Exhaust as a Toxic Air Contaminant* (1998). Report to the California Air Resources Board.

4 Marmur A., Unal A., Mulholland J.A., and Russell A., *Optimization-Based Source Apportionment of PM2.5 Incorporating Gas-to-Particle Ratios* (2005) Environmental Science and Technology 39, 3245-3254.

5 Robinson, A., Adams, P., and Donahue, N., Final Report: Source-Apportionment of Primary Organic Carbon in the Eastern United States Combining Receptor-Models, Chemical Transport Models, and Laboratory Oxidation Experiments (USEPA National Center for Environmental Research, 2007). The report and links to associated papers are available at: http://cfpub.epa.gov/ncer_abstracts/INDEX.cfm/fuseaction/display.abstractDetail/abstract/7503/report/F.

6 Manchester J., Schauer J., and Cass G. *Determination of the Elemental Carbon, Organic Compounds and Source Contributions to Atmospheric Particles During the Southern California Children's Health Study: Part B; The Distribution of Particle-Phase Organic Compounds in the Atmosphere and Source Contributions to Atmospheric Particulate Matter Concentrations During the Southern California Children's Health Study, 1995*. Final Report for California Air Resources Board.

7 Manchester-Neesvig J. B., Schauer J. J., and Cass G. R., *The Distribution of Particle-Phase Organic Compounds in the Atmosphere and their Use for Source Apportionment during the Southern California Children's Health Study* (2003) Journal of the Air and Waste Management Association. 53, 1065-1079.

8 Chow J., Chen L.W., Lowenthal D., Doraiswamy P., Park K., Kohl S., Trimble D., and Watson J. *California Regional PM10/PM2.5 Air Quality Study (CRPAQS): Initial Data Analysis of Field Program Measurements, Final Report, 2005* for the California Air Resources Board.

9 The DOE Gasoline/Diesel PM Split Study, Presentations by D. Lawson, E. Fujita and J. Schauer, California Air Resources Board Seminars webpage; and DOE/NREL Gasoline/Diesel PM Split Study webpage.

10 Fujita E., Campbell D., Arnott W.P., Chow J., and Zelinska B., *Evaluations of the Chemical Mass Balance Method for Determining Contributions of Gasoline and Diesel Exhaust to Ambient Carbonaceous Aerosols* (2007) Journal of the Air and Waste Management Association. 57, 721-740.

11 Lough G. and Schauer J., *Sensitivity of Source Apportionment of Urban Particulate Matter to Uncertainty in Motor Vehicle Emissions Profiles* (2007) Journal of the Air and Waste Management Association. 57, 1200-1213.

12 California Air Resources Board, 2007 Emission Inventory, Planning and Technical Support Division (extraction data 7/2008).

13 California Air Resources Board – Harbor Communities Monitoring Study website (http://www.arb.ca.gov/research/mobile/hcm/hcm.htm).

14 California Air Resources Board – Goods Movement Emission Reduction Plan, Appendix A (page A-52): (http://www.arb.ca.gov/planning/gmerp/gmerp.htm).