

Predicting Primary PM_{2.5} and PM_{0.1} Trace Composition for Epidemiological Studies in California

Jianlin Hu,[†] Hongliang Zhang,[†] Shu-Hua Chen,[‡] Christine Wiedinmyer,[§] Francois Vandenberghe,^{||} Qi Ying,[⊥] and Michael J. Kleeman^{*,†}

[†]Department of Civil and Environmental Engineering, University of California, Davis, One Shields Avenue, Davis California

[‡]Department of Land, Air, and Water Resources, University of California, Davis, One Shields Avenue, Davis, California

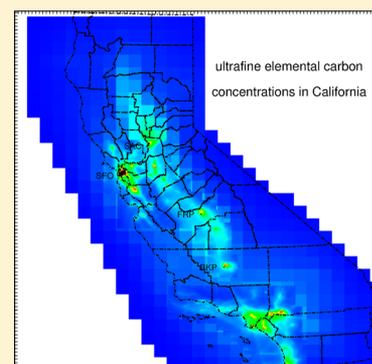
[§]Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, Colorado

^{||}Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado

[⊥]Zachry Department of Civil Engineering, Texas A&M University, College Station Texas

Supporting Information

ABSTRACT: The University of California—Davis_Primary (UCD_P) chemical transport model was developed and applied to compute the primary airborne particulate matter (PM) trace chemical concentrations from ~900 sources in California through a simulation of atmospheric emissions, transport, dry deposition and wet deposition for a 7-year period (2000–2006) with results saved at daily time resolution. A comprehensive comparison between monthly average model results and available measurements yielded Pearson correlation coefficients (R) ≥ 0.8 at ≥ 5 sites (out of a total of eight) for elemental carbon (EC) and nine trace elements: potassium, chromium, zinc, iron, titanium, arsenic, calcium, manganese, and strontium in the PM_{2.5} size fraction. Longer averaging time increased the overall R for PM_{2.5} EC from 0.89 (1 day) to 0.94 (1 month), and increased the number of species with strong correlations at individual sites. Predicted PM_{0.1} mass and PM_{0.1} EC exhibited excellent agreement with measurements ($R = 0.92$ and 0.94 , respectively). The additional temporal and spatial information in the UCD_P model predictions produced population exposure estimates for PM_{2.5} and PM_{0.1} that differed from traditional exposure estimates based on information at monitoring locations in California Metropolitan Statistical Areas, with a maximum divergence of 58% at Bakersfield. The UCD_P model has the potential to improve exposure estimates in epidemiology studies of PM trace chemical components and health.



INTRODUCTION

Epidemiological studies have identified positive correlations between exposure to ambient airborne particulate matter (PM) and increased health risk (see, for example, refs 1–5). Recent studies^{6–12} have attempted to link these health effects to particle size and/or composition using exposure estimates based on the measured ambient PM₁₀ or PM_{2.5} concentrations from central site monitors (the PM monitors at a site centrally located with respect to the area being studied), which are usually sparse in time (1 sample every 3 days), space (1 sample for a Metropolitan Statistical Area), chemical composition (no routine measurement of organic molecules), and source origin information (no routine estimation). In addition, important particle size distribution and chemical composition information is not routinely available, especially for the ultrafine particle size fraction ($D_p < 100$ nm; PM_{0.1}) that has been shown to have greater toxicity than larger particles.^{13–17} A more accurate estimate of exposure to detailed particle size fractions and chemical components would improve the power of future epidemiological studies.

A variety of statistical and mechanistic modeling techniques have been proposed to improve the accuracy of exposure estimates to air pollution. Land use regression models have

been developed to predict the spatial distribution of exposure to primary traffic PM on scales of hundreds of meters,^{18,19} but corresponding regression models for other important particle sources have not been widely demonstrated. Regression models also do not directly address the issues of data sparseness in time, particle size, and particle composition. Some land use regression models^{20,21} and dispersion models²² have been developed to estimate exposure to ultrafine particle number concentration, but the resulting epidemiological associations are not consistent with findings in toxicology studies. These results suggest that some feature other than particle number that is more directly correlated with ultrafine particles surface area may merit investigation. PM_{0.1} mass concentrations are more strongly correlated to surface area than number concentrations²³ but PM_{0.1} measurements have only been made in a few intensive study periods^{24–28} and one annual study²³ to date.

Received: February 13, 2013

Revised: February 11, 2014

Accepted: April 2, 2014

Published: April 2, 2014

Chemical transport models (CTMs) can predict the detailed size and chemical composition distribution of primary PM size fractions and composition with high temporal and spatial resolution. Recent studies have used CTMs to estimate air pollution exposure to PM_{2.5} mass and ozone (O₃) in the U.S.^{29,30} and Europe.³¹ No published studies to date have taken full advantage of the ability of CTMs to simultaneously estimate population exposure to multiple particle size fractions, chemical components, and source contributions, partly due to the uncertainties in the model inputs, simplified representation of the complex processes, and difficulty in evaluating the accuracy of such predictions.

The objective of this study is to develop the University of California —Davis_Primary (UCD_P) CTM to predict detailed particle size, composition, and source information over a 7-year time period that can be used in subsequent epidemiological studies for PM_{0.1} and PM_{2.5}. California is chosen as the focus area for the study because it has a large population exposed to high PM concentrations, accurate PM emissions inventories, and comprehensive ambient measurements for model evaluation. The ability of UCD_P predictions for mass and chemical component concentrations in the PM_{0.1} and PM_{2.5} size fractions is evaluated against ambient measurements; a companion study evaluates the model ability for PM source apportionment.³²

Model Description. The UCD_P model used in the current study was developed from the source-oriented UCD/CIT air quality model that has been successfully applied in several previous studies in the South Coast Air Basin (SoCAB) and the San Joaquin Valley (SJV) in California.^{33–38} In contrast to the conventional UCD/CIT model, the UCD_P model only tracks the primary PM (particles emitted directly into the atmosphere), and does not take into account the formation of secondary PM (produced by chemical reactions in the atmosphere). Therefore, the UCD_P model includes a complete description of emissions, advection, diffusion, dry and wet deposition, but excludes gas- and particle- phase chemistry, gas-to-particle conversion, nucleation, and coagulation. Sensitivity tests indicate that these omitted processes have little impact on predicted PM_{0.1} mass concentrations (Supporting Information (SI) Figure S8 and associated discussion). The UCD_P model is designed to track large numbers (+1000) of primary particle source contributions through the atmosphere while retaining size, composition, and source-origin information. The details of the standard algorithms used in the UCD/CIT family of models are provided in previous studies and therefore not repeated here. The formulation of advection and diffusion scheme is described by Kleeman and Cass,³³ the dry deposition scheme is described by Kleeman et al.,³⁹ the vertical advection scheme is described by Hu et al.,⁴⁰ and the wet deposition scheme is described by Mahmud et al.⁴¹ Verification of the individual processes through a comparison to theoretical calculations is described in the SI (Figure S1). Zhang and Ying⁴² developed a one-way nesting capability in the conventional UCD/CIT model and this feature was implemented into the UCD_P model. Concentration fields from model calculations in the coarse-resolution parent domain for every source type are saved in the boundary grid cells of the finer-resolution nested domains, and then are used as the boundary conditions for the model calculation in the nested domains.

Model Application. Model calculations were performed to predict primary PM concentrations for seven continuous years

(January first, 2000 to December 31st, 2006) in California. Figure 1a displays the air quality modeling domains used in the

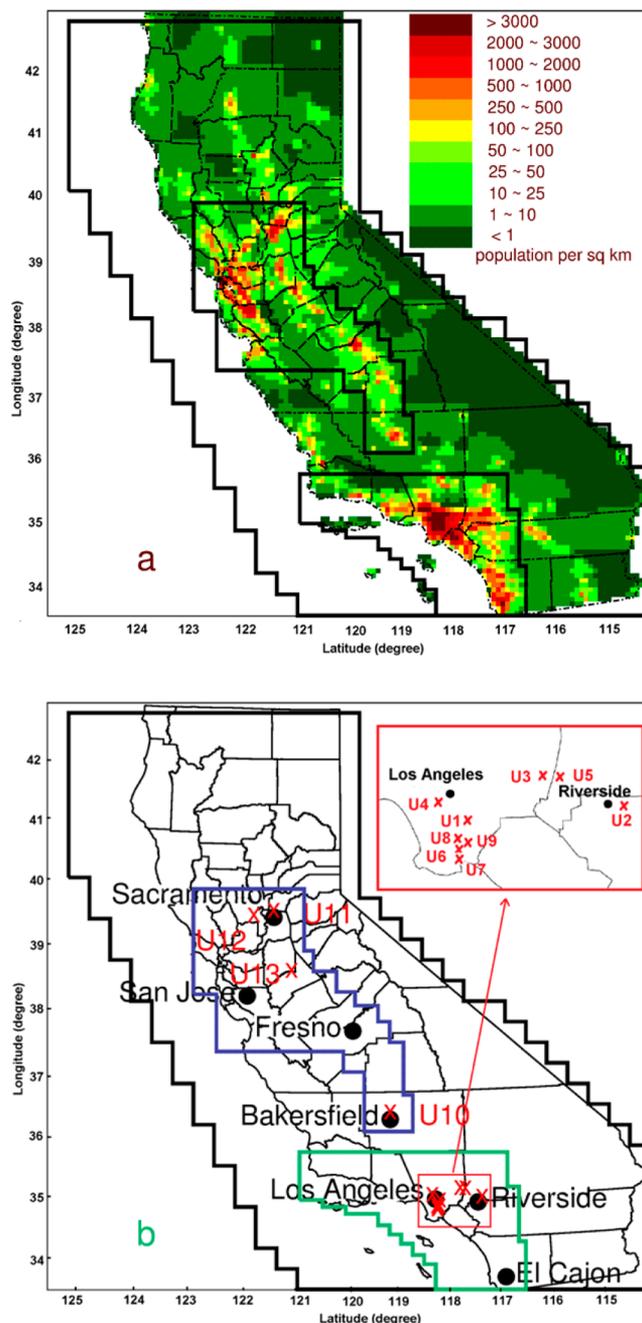


Figure 1. (a) Modeling domains (same as in panel b) and population density in California; (b) Modeling domains and measurement sites of fine PM (dark dots) and ultrafine PM (red crosses). In panel b, The CA_36km domain is outlined by black line, the SJV_4km domain is outlined by blue line, and the SoCAB_4km domain is outlined by green line. The names and locations (latitude/longitude) of the sites are listed in Table S2 in the SI.

present study as well as the population density in California. A one-way nesting technique was used with a parent domain of 36 km horizontal resolution that covered the entire state of California (referred to as CA_36km) and two nested domains of 4 km horizontal resolution that covered the SoCAB (referred to as SoCAB_4km) and San Francisco Bay Area + SJV + South Sacramento Valley air basins (referred to as SJV_4km). Over 92%

of California's population lives in the high-resolution (4 km) domains. The UCD_P model was configured with 10 vertical layers up to a height of 5 km above ground level in the CA_36km and the SoCAB_4km domains. The thickness of each vertical layer, from ground to top, was 35, 105, 140, 330, 390, 500, 500, 1000, 1000, and 1000 m. The UCD_P model was configured with nine vertical layers up to height of 4 km above ground level in the SJV_4km domain in order to reduce the memory requirements of the simulation. A sensitivity study determined that the configuration of the vertical domain at 4 km vs 5 km in the SoCAB caused <3% difference in predicted PM concentrations. Note that the use of relatively shallow vertical domains is only appropriate in regions with well-defined air basins and would not be appropriate for locations in the eastern U.S. or other regions with moderate topography.

Meteorological Fields. Hourly meteorological fields during the 7-year modeling period were simulated with the Weather Research and Forecasting (WRF) model version 3.1^{43,44} using two nested domains (2-way interaction) that had horizontal resolutions of 12 km and 4 km, respectively. North American Regional Reanalysis data with 32 km horizontal resolution and 3-h time resolution were used as initial and boundary conditions of the coarse 12 km domain. The WRF model was configured with 31 vertical layers up to 100 hpa (around 16 km above ground level). Previous studies^{45,46} showed that regional meteorological models such as MMS and WRF tend to overpredict the surface winds in regions with complex terrain such as the SJV, especially during stagnant events with low wind speed. Four-dimensional data assimilation⁴⁷ was utilized in the present simulations to anchor the model predictions to observed meteorological patterns. A previous study⁴⁰ demonstrated that four-dimensional data assimilation improved the meteorological predictions needed for air quality modeling in the SJV region, but four-dimensional data assimilation did not completely correct the bias during the events with low-speed surface winds that produce the highest pollution episodes. A recent study⁴⁸ found that increasing the surface friction velocity (u^*) by 50% improved the surface wind predictions in a complex-terrain domain that covers the state of Washington. Hu⁴⁹ adopted this method during simulations of the Central California Ozone Study 2000 that successfully reproduced ozone concentration trends in central California. In the present study, the ability to improve model performance for the major air-quality related meteorological parameters by increasing u^* was examined during a one-year sensitivity study for 2000. Figure S2 in the SI shows the comparison for the surface wind with and without modifications to u^* . The results confirm that increasing u^* lowers the mean wind bias from 1.15 m/s to -0.50 m/s, and lowers the root-mean-square error from 2.95 to 2.20 m/s. Therefore, the increased u^* method was also applied in 2001–2006 WRF simulations. Temperature, wind and humidity were exclusively evaluated against observations for all air basins in California for the seven years, and the results are summarized in Table S1 of SI. Mean fractional bias (MFB) of temperature and wind are generally within ± 0.15 , root-mean-square errors of temperature are around 4 °C, and root-mean-square errors of wind are generally lower than 2.0 m/s, especially in the air basins with high population density. This level of performance is consistent with a previous WRF study in California,⁵⁰ but is generally poorer than the performance of WRF in other regions with less extreme topography.^{51,52}

Emissions. Hourly particle-phase emissions from anthropogenic sources were generated using an updated version of the

emissions model described by Kleeman and Cass.⁵³ In the revised system, emission totals of individual emissions inventory code sources from the standard emissions inventories (including point, area, mobile, and dust emissions) provided by the California Air Resources Board were transformed into size-resolved emissions of total particle mass using measured source profiles (see, for example, refs 54–56). More detailed discussion of the emissions processing is presented in another study.⁵⁷ Particle size distributions emitted from each source were represented by eight discrete particle diameters centered within equally spaced logarithmic size intervals spanning the diameter range from 0.01 to 20 μm . PM emissions from each of the ~ 900 sources present in the California emissions inventory were tracked separately in the simulation as part of a comprehensive source apportionment study. The results of those source apportionment calculations are reported in a companion paper.³² The mass and density of size-resolved PM was tracked during model calculations, with composition profiles applied during postprocessing of results (see Section 4.1). High resolution emissions from wildfires and open burning (1×1 km)^{58,59} were included in this study, but emissions of biogenic particles and sea-salt particles were not included.

Ambient PM Measurements. Comprehensive model evaluations were performed for primary PM components, that is, elemental carbon and trace elements but not secondary components such as organic carbon, sulfate, nitrate, and ammonium because the model was not configured to predict these secondary species. $\text{PM}_{2.5}$ speciation measurements were obtained from the California Air Resources Board "2011 Air Quality Data DVD",⁶⁰ while ultrafine measurements were obtained from published literature. There were total of 13 $\text{PM}_{2.5}$ speciation sites in the 4 km domains during the modeling periods with a measurement frequency of 1 in 3 or 1 in 6 days; Measured $\text{PM}_{2.5}$ EC concentrations at five sites were found to be exactly 0.5 $\mu\text{g}/\text{m}^3$ on >80% of the measurement days suggesting corrupt or missing data at these locations. These 5 $\text{PM}_{2.5}$ sites were not included in the evaluation analysis. There were 13 sites in the 4 km domains with measurements available for $\text{PM}_{0.1}$ (ultrafine) and $\text{PM}_{0.18}/\text{PM}_{0.25}$ (quasi-ultrafine) during intensive operation periods typically shorter than 1 month. The ultrafine and quasi-ultrafine measurement sites are shown in Figure 1b, and a brief description of site locations and major nearby sources is included in SI Table S2. The sampling period and size cut information is shown in SI Table S8. MFB, mean fractional error (MFE), Pearson correlation coefficient (R), and coefficient of divergence (COD) were calculated as the statistical measures for model performance. R and MFB are the primary metrics used to evaluate the accuracy of model estimates in space and time.

RESULTS

$\text{PM}_{2.5}$ EC. The predicted and measured daily $\text{PM}_{2.5}$ EC concentrations exhibited seasonal variation at the seven available monitoring sites (Sacramento, San Jose, Fresno, Bakersfield, Los Angeles, Riverside, and El Cajon) with lower values during the summer and higher values during the winter (SI Figure S3). The predictions at the exact location of the monitors and the best fit predictions (closest to observations) within 12 km (i.e., three grid cells in the present study) of the observation site are shown in the figure. The difference between the predictions at the exact location of monitors and the best fit predictions conveys information about the spatial gradient of

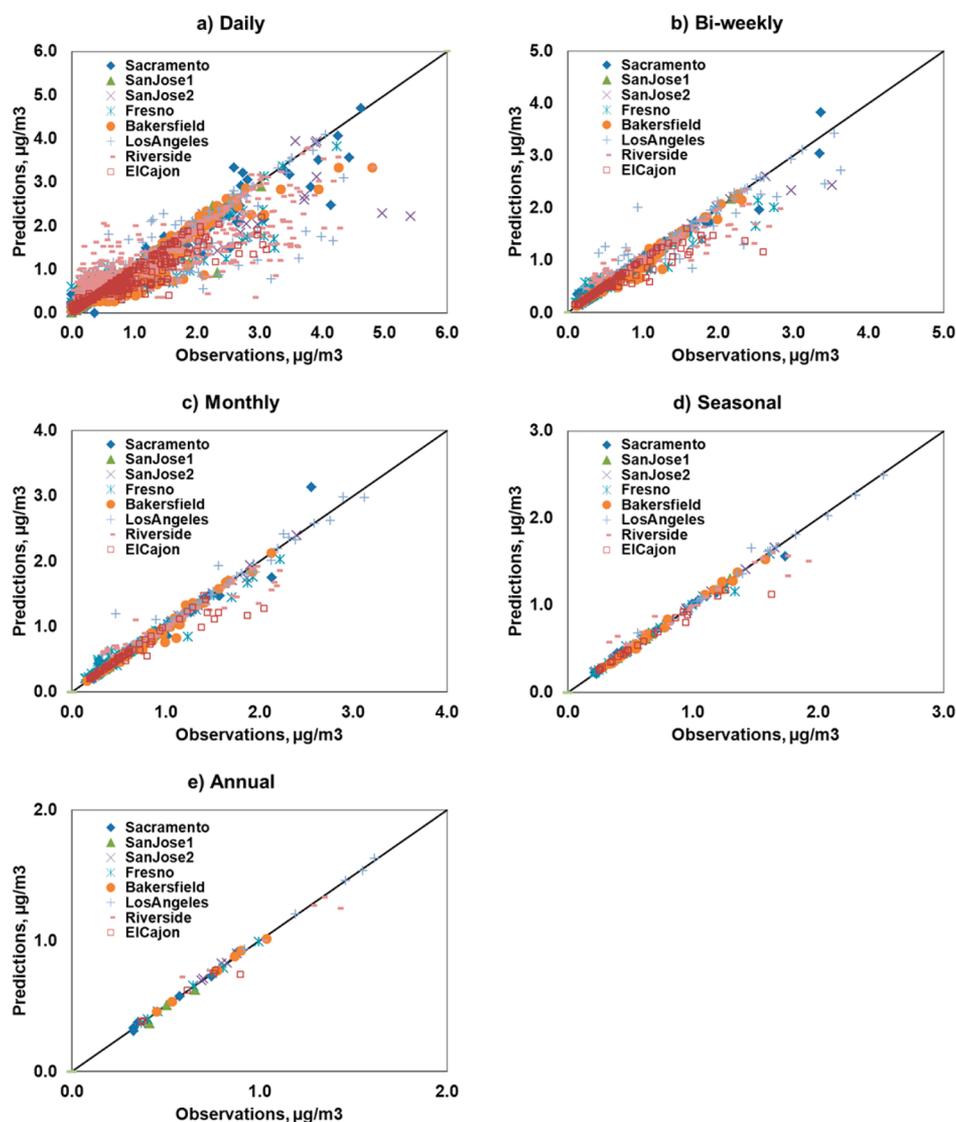


Figure 2. Observed and predicted $PM_{2.5}$ EC concentrations using averaging times of (a) 1 day; (b) 2 weeks; (c) 1 month; (d) 3 months; and (e) 1 year.

concentrations. Over longer time scales, EC concentrations were overpredicted in the winter of 2000–2002 at Sacramento, San Jose, Fresno and Los Angeles and under-predicted in the winter of 2005–2006 at Bakersfield, Riverside, and El Cajon. This issue is not related to nearby spatial gradients since the best fit predictions were used in this comparison. Long-term temporal trends likely reflect economic activity and goods movement patterns that are not currently represented in the emissions inventory.

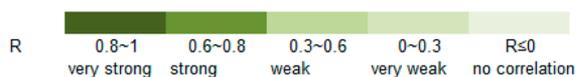
The choice of the time-averaging period influences the level of agreement between predictions vs measurements at the $PM_{2.5}$ speciation sites. Figure 2 compares observed EC concentrations and best fit predictions using averaging times of (a) 1 day, (b) 2 weeks, (c) 1 month, (d) 3 months, and (e) 1 year. Longer averaging times produce better agreement between model predictions and measurements because they remove the effects of random measurement errors at monitoring stations and variations in actual emissions rates that are not reflected in seasonally averaged emissions inventories. The overall correlation coefficient (R) for $PM_{2.5}$ EC at all sites increases from 0.89 for daily averages to 0.94 for

monthly averages, and 0.95 for annual averages (SI Table S3). The regression slope increases from 0.74 for daily averages to 0.92 for monthly averages, and 0.97 for annual averages. The regression slopes <1 indicate under-prediction of EC concentrations, especially at locations with the highest measured values. The Riverside site has the greatest EC under-predictions with a regression slope of 0.75. This site is within 0.5 km of a major freeway; model predictions based on 4 km grid resolution cannot represent the highest concentrations in the immediate vicinity of major highways. High emissions events and/or stagnant local meteorological conditions typically lead to high primary PM concentrations. These processes may not be represented accurately enough in the current study to capture the highest concentration events. MFB, MFE and COD decrease as the averaging period gets longer because the influence of extreme events is reduced (Table S3 in the SI).

$PM_{2.5}$ Trace Elements. Table 1 shows performance statistics (R and MFB) for EC and trace elements using monthly averaged best fit predictions at the eight monitoring sites (daily, biweekly, seasonal and annual statistics are included in Tables S4–S7 in the SI). At more than five individual sites

Table 1. Pearson Correlation Coefficients (R = Color) And Mean Fractional Bias (MFB = Numerical Value) Of Monthly Average $PM_{2.5}$ EC and Trace Elements at Individual $PM_{2.5}$ Speciation Measurement Sites

Species	Sacramento	SanJose1	SanJose2	Fresno	Bakersfield	LosAngeles	Riverside	ElCajon
EC	0.11	0.01	0.01	0.07	-0.02	0.00	0.13	-0.04
K	0.07	-0.09	-0.07	0.06	-0.10	-0.02	0.02	-0.16
CR	-0.05	0.12	0.06	0.19	0.03	0.12	0.20	0.02
ZN	0.25	0.05	0.09	-0.08	-0.04	-0.02	-0.15	0.08
FE	0.73	0.21	0.36	0.67	0.25	0.22	0.33	0.18
TI	0.43	0.05	0.19	0.39	0.10	0.11	0.17	0.09
AS	-0.13	-0.06	-0.02	-0.32	-0.65	0.02	0.01	0.07
CO	0.04	-0.02	0.02	-0.65	0.06	0.51	0.48	-0.91
SR	0.11	-0.06	-0.02	0.05	-0.10	-0.28	-0.12	-0.08
CA	0.42	-0.08	0.09	0.36	0.06	-0.08	-0.09	0.04
MN	0.60	0.17	0.45	0.63	0.38	0.16	0.25	0.11
AL	0.91	0.70	0.74	0.95	0.51	0.77	0.44	0.45
SI	0.86	0.30	0.60	0.80	0.42	0.32	0.18	0.29
CU	-0.33	-0.10	-0.21	-0.36	-0.63	-0.08	0.01	-0.43
NI	-0.18	-0.72	-0.28	-0.02	0.24	0.07	0.13	-0.49
PB	-0.34	-0.24	-0.09	-0.62	-0.66	0.03	0.06	-0.11
V	-0.45	-0.61	-0.82	0.02	-0.20	-0.30	-0.09	-0.44
MO	-0.60	-1.19	-1.14	-0.70	-0.76	-0.69	-0.08	-1.60
RB	-0.32	-0.76	-0.59	-0.29	-0.50	-0.26	-0.07	-0.70
BA	-1.09	-1.03	-0.38	-0.73	-1.26	-0.57	-0.94	-0.98
CD	-1.67	-0.76	-0.96	-1.76	-1.25	-0.79	-0.62	-1.37
MG	-1.23	-1.32	-1.28	-1.31	-1.56	-1.28	-1.55	-1.43
NA	-1.53	-1.86	-1.73	-1.40	-1.40	-1.57	-1.51	-1.78



(62.5% of the total eight sites), nine elements (potassium, chromium, zinc, iron, titanium, arsenic, calcium, manganese, and strontium) (39% of the total 23 elements analyzed in the study) have $R \geq 0.8$. Model performance varies strongly with location. Figure 3 shows that predictions at San Jose, Los

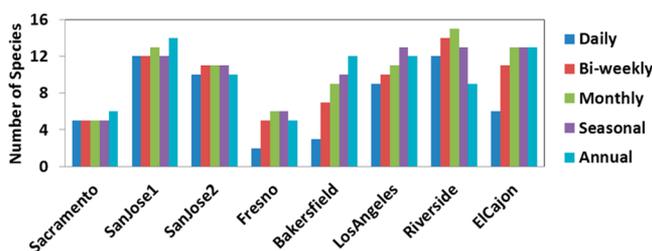


Figure 3. Number of species which have $R \geq 0.8$ and MFB within ± 0.3 at individual sites when using averaging times of (a) 1 day; (b) 2 weeks; (c) 1 month; (d) 3 months; and (e) 1 year.

Angeles, and Bakersfield have the best agreement with measurements, with $R \geq 0.8$ and MFB within ± 0.3 using daily averages for over ~ 10 species. Predictions at Sacramento, Fresno, and Bakersfield have the least agreement with measurement, with $R \geq 0.8$ and MFB within ± 0.3 for less than five species. Figure 3 also indicates that agreement between predictions and measurements for trace elements improves as averaging time increases. The number of species with $R \geq 0.8$ and MFB within ± 0.3 generally increases as the averaging time increases from 1 day to 3 months. Annual averages do not necessarily yield better results than monthly or seasonal averages, due to the fact that there are only (at most) seven data points in the annual average analysis and 1 or 2 outliers can significantly affect the results. Predicted and observed concentrations of the $PM_{2.5}$ compositions are compared in SI Figure S4. The spatial distribution of predicted $PM_{2.5}$ mass, EC, and element concentrations is shown in SI Figure S6, with significant differences observed for PM components emitted from different sources.

Ultrafine PM Mass and Element Carbon. A number of short-term ultrafine particle and quasi-ultrafine particle measurement campaigns were conducted in California during the modeling period^{24–28} (see Table S8 in SI for a summary). Model predictions were compared to the measured ultrafine PM mass concentrations pooled across all studies. Ultrafine PM number concentrations were not evaluated because nucleation processes were not enabled in model calculations. Figure 4 shows the measured and predicted ultrafine PM EC and mass concentrations. Best-fit model predictions are in excellent agreement with measured ultrafine EC concentrations, with $R = 0.94$. Predicted ultrafine PM mass also agrees well with measured values, with $R = 0.92$. The time series results of ultrafine EC and ultrafine mass (SI Figure S5) indicate excellent temporal agreement at most of the sites. The model predictions are not able to capture events when measured ultrafine or quasi-ultrafine PM mass is $>4 \mu\text{g}/\text{m}^3$ or $<1 \mu\text{g}/\text{m}^3$, and when measured ultrafine or quasi-ultrafine EC is $>1 \mu\text{g}/\text{m}^3$ or $<0.2 \mu\text{g}/\text{m}^3$. High ultrafine PM mass and EC concentrations were observed at urban sites that are located in central Los Angeles near industrial sources and busy highways. The current model calculations did not include plume in grid calculations and so point and line source emissions were instantaneously diluted into $4 \times 4 \text{ km}$ grid cells. Finer grid resolution combined with large eddy simulation would likely improve the model performance at sites influenced by neighborhood-scale emissions. The UCD_P model does not consider the processes of nucleation, coagulation, and gas-to-particle transfer in the simulations in order to make the detailed source apportionment calculations computationally feasible. Nucleation and condensation effects generally have minor impact on $PM_{0.1}$ because this particle size fraction is dominated by primary emissions from combustion sources. Nucleation is important for number concentrations which are dominated by particles with diameter $<0.05 \mu\text{m}$ but these particles contribute little to $PM_{0.1}$ mass concentrations. Condensation of semivolatile material generally takes place in the particle accumulation mode, not in the ultrafine size range due to the increase in the gas-phase

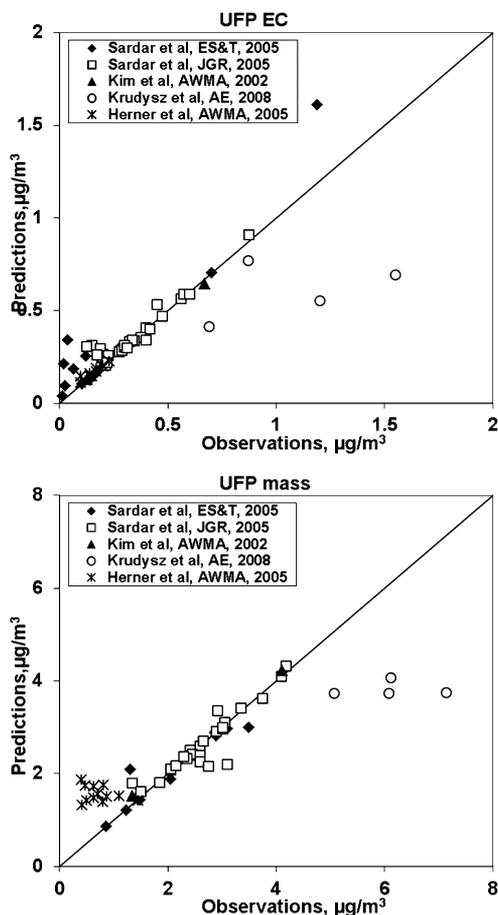


Figure 4. Observed and predicted ultrafine ($PM_{0.1}$) and quasi-ultrafine ($PM_{0.18}$, $PM_{0.25}$) EC and mass concentrations. $PM_{0.1}$ was reported by Sardar JGR, Kim, and Herner, $PM_{0.18}$ was reported by Sardar ES&T, and $PM_{0.25}$ was reported by Krudysz. Averaging times were 1 month, 2 weeks, 5 months, 2.5 months, and 3–4 days reported by Sardar JGR, Sardar ES&T, Kim, Krudysz, and Herner, respectively. The size cut and sampling period information is listed in SI Table S8.

concentration above small particles (Kelvin effect). Size-resolved source apportionment studies confirm that most secondary organic aerosols formation takes place at $D_p > 0.1 \mu m$.^{61–63} To examine the uncertainty in predicted $PM_{0.1}$ concentrations due to the omission of these processes, sensitivity tests were conducted that included gas-particle conversion and coagulation. Difference in the 7-year average concentrations of $PM_{0.1}$ EC and mass concentrations are generally less than 10% (SI Figure S8).

Sharper spatial gradients were found for $PM_{0.1}$ mass and all the components, compared to $PM_{2.5}$ mass and components (SI Figure S7). This reflects the larger deposition velocity for ultrafine particles due to Brownian diffusion.⁶⁴ $PM_{0.1}$ and $PM_{2.5}$ spatial patterns for certain elements such as Al and K were significantly different, reflecting different sources dominating each size fraction. Within the $PM_{0.1}$ size fraction, trace elements exhibited significantly different spatial patterns depending largely on their emissions sources.

DISCUSSION

Information from 8 available $PM_{2.5}$ speciation sites was used to verify the temporal and spatial accuracy of UCD_P predictions of $PM_{2.5}$ primary components (EC and trace elements). There were 66 $PM_{2.5}$ mass observation sites located inside the 4km domains that cannot be directly used to compare with UCD_P results because the secondary PM components (nitrate, sulfate, secondary organic aerosols, etc.) were not included in the UCD_P model. To help build confidence in the UCD_P results, we repeated the simulations based on the inputs for the 7 year study period using the full UCD/CIT model that includes the gas and PM chemistry, secondary PM components and gas-particle conversion. The full CTM predictions for total (=primary + secondary) $PM_{2.5}$ mass were compared to the measured concentrations at the 66 sites. Detailed results are discussed in a separate manuscript.⁵⁶ As a summary, 52 sites (79% of the total 66 sites) had MFB within ± 0.3 after a sensitivity analysis for dust emissions. Uncertainties in the emissions inputs and limitations in reproducing local

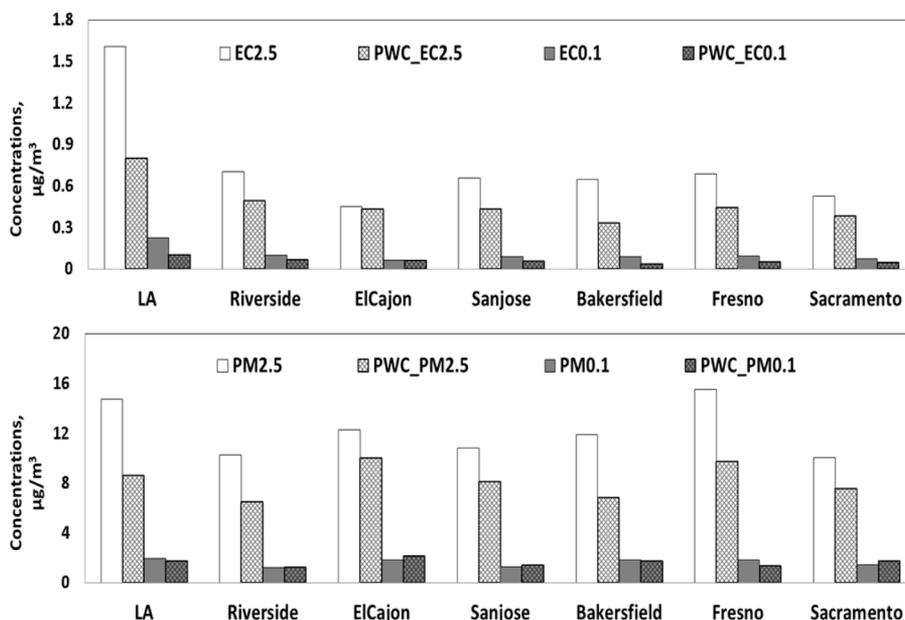


Figure 5. Central site concentrations vs population weighted concentrations in seven California metropolitan statistical areas.

meteorological conditions account for the poor performance at the remaining sites.

California's relatively accurate emissions inventory combined with a highly resolved source profile library and modified WRF predictions for meteorology yield model predictions for EC, potassium, chromium, zinc, iron, titanium, arsenic, calcium, manganese, and strontium that generally reproduce observed monthly/seasonal variations. The results of the current study demonstrate that not all CTM predictions for primary PM concentrations are suitable for inclusion in epidemiology studies. Sea-salt emissions need to be included for predictions of sodium, chloride, and other major components of sea-salt.⁶⁵ More accurate emissions inventories and source profiles would be needed to increase the accuracy of predictions for other species. One approach to account for poor emissions data is to combine model predictions with measurements at receptor sites in a statistical framework.⁶⁶ This new data "fusion" approach is undergoing initial evaluation in epidemiological studies, but was not employed here, because measurements for some particle size fractions/chemical components are not available even though they are of great interest in epidemiology applications.

In the current study, predicted PM trace components concentrations, which were in good agreement with measurements, had significantly different spatial patterns that would not be predicted by central site monitors or land use regression models optimized for traffic sources. These spatial patterns will lead to exposure misclassification in epidemiological studies. Annual-average population weighted concentrations (PWCs) were calculated to estimate the impact of the spatial heterogeneity. PWC is calculated as follows:

$$PWC = \frac{\sum_i C_i \times P_i}{\sum_i P_i} \quad (1)$$

where i is the grid point, C_i is the annual average concentration (as shown in SI Figure S4 and S5) and P_i the population density (as shown in Figure 1a) in grid cell i . PWC for EC and mass in $PM_{2.5}$ and $PM_{0.1}$ size fractions were calculated to test the effect of spatial heterogeneity. Figure 5 shows that PWCs of $PM_{2.5}$ EC in 7 California Metropolitan Statistical Areas were generally lower than central monitor predictions with a maximum bias of -50% in Los Angeles and -49% in Bakersfield, and an average bias of -33% . Model predictions for $PM_{2.5}$ EC are reasonably accurate at monitoring sites, with linear regression slopes of 1.02, 1.05, 0.99, 0.99, 0.98, 0.75, and 0.81 at Sacramento, San Jose, Fresno, Bakersfield, Los Angeles, Riverside, and El Cajon, respectively, and $R^2 > 0.92$ at all sites. These biases are caused by differences in predicted concentrations at the location of the central monitors vs other locations within the Metropolitan Statistical Area. PWCs of $PM_{0.1}$ mass revealed positive and negative biases associated with central site monitoring locations varying from -26% in Fresno to $+21\%$ in Sacramento. The difference between PWCs of $PM_{0.1}$ EC and central site concentrations were even greater, with an average difference of -39% , and specific values of -55% in Los Angeles, and -59% in Bakersfield. These biases have likely implication for epidemiological analysis of $PM_{0.1}$ health associations, suggesting that central site monitors may introduce uncertainty into this analysis.

Accurate CTM spatial representation is a key factor for exposure assessment. CTM grid resolutions of 4km provide more spatial information than central monitor measurements, but even 4 km resolution may not be sufficient for species and

sources with very sharp gradients. Land use regression models are widely used to construct spatial variations of air pollutants at ~ 100 m resolution for epidemiological studies.⁶⁷ Land use regression models have been mostly applied for studies of gas-phase pollutants emitted from traffic sources, but two recent studies used land use regression models in California for $PM_{2.5}$ mass.^{68,69} In the future, land use regression models could be combined with CTM predictions for detailed PM components and sources to predict concentration fields at spatial resolution finer than 4 km.

Model results agree better with measurements over longer averaging time and so they may be most useful in epidemiological studies with longer averaging times (for example ≥ 1 month). Day-to-day variations in concentrations driven by changes in weather patterns and/or unique source activity are difficult to predict at increasingly higher time resolution (for example ≤ 1 day). Model predictions have larger bias on days with the highest observed primary concentrations. Analysis of predicted meteorological fields reveals that the WRF model has difficulties predicting extreme stagnation events that produce high PM pollution (wind speed is overpredicted during these events). The CTM model temporal performance would be enhanced with improvements in meteorological modeling. In addition, the CTM model temporal performance would benefit from the development of high temporal resolution emissions inventories (for example, most of residential emissions are annual average in the current inventory).^{70,71}

Notwithstanding the areas for future improvement described above, the results of the current study are presently being evaluated in a number of epidemiology studies. The first such analysis to yield results has identified associations between primary PM exposure and term low birth weight in California's Los Angeles County.⁷² The results reveal a 2.5% increase in low birth weight risk associated with an interquartile range increase in primary $PM_{2.5}$ and $PM_{0.1}$ mass concentrations. Significant associations with low birth weight are also found for chemical compositions of EC, potassium, iron, and titanium in $PM_{2.5}$ and EC, potassium, iron, and chromium in $PM_{0.1}$. Future applications will evaluate the relationships between $PM_{2.5}$ and $PM_{0.1}$ composition and sources vs other health end points in California.

■ ASSOCIATED CONTENT

📄 Supporting Information

Meteorology evaluation metrics in California air basins (Table S1), Observation site information (Table S2), air quality evaluation metrics for EC and elements (Table S3), Pearson correlation coefficients and mean fraction bias of daily, biweekly, seasonal, and annual average $PM_{2.5}$ EC and trace elements at individual $PM_{2.5}$ speciation measurement sites (Tables S4–7), ultra- and quasi-ultra- fine PM measurements (Table S8), verification of the treatment in the UCD_P model for individual processes including dry deposition, wet deposition, advection, and diffusion (Figure S1), meteorology improvement by increasing friction velocity (u^*) in 2000 (Figure S2), time series of observed and predicted daily concentrations $PM_{2.5}$ EC at 7 sites (Figure S3), observed and predicted ultrafine EC and ultrafine mass concentrations (Figure S5), predicted spatial distributions of $PM_{2.5}$ mass, EC and trace elements (Figure S6), predicted spatial distributions of $PM_{0.1}$ mass, EC and trace elements (Figure S7). Impact of gas-particle transfer and coagulation on the 7 year average

concentrations of PM_{0.1} EC and PM_{0.1} mass (Figure S8). All model predictions described in the current study are available for download at faculty.engineering.ucdavis.edu/kleeman. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Phone: +1 530 752 8386; fax: +1 530 752 7872; e-mail: mjkleeman@ucdavis.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was funded by the United States Environmental Protection Agency under Grant No. 83386401. Although the research described in the article has been funded by the United States Environmental Protection Agency it has not been subject to the Agency's required peer and policy review and therefore does not necessarily reflect the reviews of the Agency and no official endorsement should be inferred.

REFERENCES

- (1) Adler, K. B.; et al. Interactions between respiratory epithelial-cells and cytokines—Relationships to lung inflammation. *Ann. N. Y. Acad. Sci.* **1994**, *725*, 128–145.
- (2) Dockery, D. W.; et al. An Association between Air-Pollution and Mortality in 6 United-States Cities. *New Engl. J. Med.* **1993**, *329* (24), 1753–1759.
- (3) Gordian, M. E.; et al. Particulate air pollution and respiratory disease in Anchorage, Alaska. *Environ. Health Perspect.* **1996**, *104* (3), 290–297.
- (4) Pope, C. A. Review: Epidemiological basis for particulate air pollution health standards. *Aerosol Sci. Technol.* **2000**, *32* (1), 4–14.
- (5) Pope, C. A.; Dockery, D. W. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manage. Assoc.* **2006**, *56* (6), 709–742.
- (6) Anderson, H. R. Differential epidemiology of ambient aerosols. *Philos. Trans. R. Soc., A* **2000**, *358* (1775), 2771–2785.
- (7) Brunekreef, B.; Forsberg, B. Epidemiological evidence of effects of coarse airborne particles on health. *Eur. Respir. J.* **2005**, *26* (2), 309–318.
- (8) de Hartog, J. J.; et al. Effects of fine and ultrafine particles on cardiorespiratory symptoms in elderly subjects with coronary heart disease—The ULTRA study. *Am. J. Epidemiol.* **2003**, *157* (7), 613–623.
- (9) Franklin, M.; Koutrakis, P.; Schwartz, J. The role of particle composition on the association between PM_{2.5} and mortality. *Epidemiology* **2008**, *19* (5), 680–689.
- (10) Ostro, B.; et al. Fine particulate air pollution and mortality in nine California counties: Results from CALFINE. *Environ. Health Perspect.* **2006**, *114* (1), 29–33.
- (11) Ostro, B.; et al. The effects of components of fine particulate air pollution on mortality in California: Results from CALFINE. *Environ. Health Perspect.* **2007**, *115* (1), 13–19.
- (12) Ostro, B. D.; Broadwin, R.; Lipsett, M. J. Coarse and fine particles and daily mortality in the Coachella Valley, California: A follow-up study. *J. Exposure Anal. Environ. Epidemiol.* **2000**, *10* (5), 412–419.
- (13) Dailey, L.; Devlin, R. B. Seasonal effects of ultrafine, fine, and coarse particulate matter (PM) on human primary airway epithelial cells. *Toxicol. Sci.* **2003**, *72*, 39–39.
- (14) Donaldson, K.; Li, X. Y.; Macnee, W. Ultrafine (nanometre) particle mediated lung injury. *J. Aerosol Sci.* **1998**, *29* (5–6), 553–560.
- (15) Ferin, J.; et al. Pulmonary tissue access of ultrafine particles. *J. Aerosol Med.* **1991**, *4* (1), 57–68.
- (16) Li, N.; et al. Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environ. Health Perspect.* **2003**, *111* (4), 455–460.
- (17) Oberdorster, G. Significance of particle parameters in the evaluation of exposure-dose-response relationships of inhaled particles. *Part. Sci. Technol.* **1996**, *14* (2), 135–151.
- (18) Briggs, D. J.; et al. Mapping urban air pollution using GIS: A regression-based approach. *Int. J. Geogr. Inf. Sci.* **1997**, *11* (7), 699–718.
- (19) Jerrett, M.; et al. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiology* **2005**, *16* (6), 727–736.
- (20) Hoek, G.; et al. Land use regression model for ultrafine particles in Amsterdam. *Environ. Sci. Technol.* **2011**, *45* (2), 622–628.
- (21) Rivera, M.; et al. Spatial distribution of ultrafine particles in urban settings: A land use regression model. *Atmos. Environ.* **2012**, *54*, 657–666.
- (22) Yuan, Y.; Zhu, Y.; Wu, J.; , Modeling traffic-emitted ultrafine particle concentration and intake fraction in Corpus Christi, Texas. *Chem. Prod. Process Model.* **2011**, *6*(1).
- (23) Kuwayama, T.; Ruehl, C. R.; Kleeman, M. J. Daily trends and source apportionment of ultrafine particulate mass (PM_{0.1}) over an annual cycle in a typical California City. *Environ. Sci. Technol.* **2013**, *47* (24), 13957–13966.
- (24) Kim, S.; et al. Size distribution and diurnal and seasonal trends of ultrafine particles in source and receptor sites of the Los Angeles basin. *J. Air Waste Manage. Assoc.* **2002**, *52* (3), 297–307.
- (25) Sardar, S. B.; et al. Size-fractionated measurements of ambient ultrafine particle chemical composition in Los Angeles using the NanoMOUDI. *Environ. Sci. Technol.* **2005**, *39* (4), 932–944.
- (26) Sardar, S. B.; Fine, P.M.; , Sioutas, C. Seasonal and spatial variability of the size-resolved chemical composition of particulate matter (PM₁₀) in the Los Angeles Basin. *J. Geophys. Res., Atmos.* **2005**, *110*(D7).
- (27) Herner, J. D.; et al. Size and composition distribution of airborne particulate matter in northern California: I—Particulate mass, carbon, and water-soluble ions. *J. Air Waste Manage. Assoc.* **2005**, *55* (1), 30–51.
- (28) Krudysz, M. A.; et al. Intra-community spatial variation of size-fractionated PM mass, OC, EC, and trace elements in the Long Beach, CA area. *Atmos. Environ.* **2008**, *42* (21), 5374–5389.
- (29) Sarnat, J. A.; et al. Associations between spatially resolved estimates of traffic-related pollution and acute morbidity: Assessing agreement of results among multiple exposure assignment approaches. *Epidemiology* **2011**, *22* (1), S31–S32.
- (30) Bravo, M. A.; et al. Comparison of exposure estimation methods for air pollutants: Ambient monitoring data and regional air quality simulation. *Environ. Res.* **2012**, *116*, 1–10.
- (31) Tainio, M.; et al. Future climate and adverse health effects caused by fine particulate matter air pollution: Case study for Poland. *Reg. Environ. Change* **2013**, 1–11.
- (32) Hu, J.; Zhang, H.; Chen, S. H.; Ying, Q.; Wiedinmyer, C.; Vandenberghe, F.; Kleeman, M. J. Identifying PM_{2.5} and PM_{0.1} sources for epidemiological studies in California. *Environ. Sci. Technol.* **2014**. DOI: 10.1021/es404810z.
- (33) Kleeman, M. J.; Cass, G. R. A 3D Eulerian source-oriented model for an externally mixed aerosol. *Environ. Sci. Technol.* **2001**, *35* (24), 4834–4848.
- (34) Held, T.; Ying, Q.; Kaduwela, A.; Kleeman, M. Modeling particulate matter in the San Joaquin Valley with a source-oriented externally mixed three-dimensional photochemical grid model. *Atmos. Environ.* **2004**, *38* (22), 3689–3711.
- (35) Ying, Q.; Kleeman, M. J. Source contributions to the regional distribution of secondary particulate matter in California. *Atmos. Environ.* **2006**, *40* (4), 736–752.
- (36) Kleeman, M. J.; et al. Source apportionment of secondary organic aerosol during a severe photochemical smog episode. *Atmos. Environ.* **2007**, *41* (3), 576–591.

- (37) Ying, Q.; et al. Verification of a source-oriented externally mixed air quality model during a severe photochemical smog episode. *Atmos. Environ.* **2007**, *41* (7), 1521–1538.
- (38) Ying, Q.; Lu, J.; Allen, P.; Livingstone, P.; Kaduwela, A.; Kleeman, M. Modeling air quality during the California Regional PM10/PM2.5 Air Quality Study (CRPAQS) using the UCD/CIT source-oriented air quality model—Part I. Base case model results. *Atmos. Environ.* **2008**, *42* (39), 8954–8966.
- (39) Kleeman, M. J.; Cass, G. R.; Eldering, A. Modeling the airborne particle complex as a source-oriented external mixture. *J. Geophys. Res., Atmos.* **1997**, *102* (D17), 21355–21372.
- (40) Hu, J.; Ying, Q.; Chen, J. J.; Mahmud, A.; Zhao, Z.; Chen, S. H.; Kleeman, M. J. Particulate air quality model predictions using prognostic vs. diagnostic meteorology in central California. *Atmos. Environ.* **2010**, *44* (2), 215–226.
- (41) Mahmud, A.; Hixson, M.; Hu, J.; Zhao, Z.; Chen, S. H.; Kleeman, M. J. Climate impact on airborne particulate matter concentrations in California using seven year analysis periods. *Atmos. Chem. Phys.* **2010**, *10* (22), 11097–11114.
- (42) Zhang, H. L.; Ying, Q. Source apportionment of airborne particulate matter in Southeast Texas using a source-oriented 3D air quality model. *Atmos. Environ.* **2010**, *44* (29), 3547–3557.
- (43) Skamarock, W. C.; Klemp, J. B.; Dudhia, J.; Gill, D. O.; Barker, D. M.; Duda, M. G.; Huang, X.-Yu; Wang, W.; Powers, J. G. A Description of the Advanced Research WRF Version 3, NCAR Technical Note NCAR/TN-475+STR, June 2008.
- (44) Wang, W.; Bruyère, C.; Duda, M.; Dudhia, J.; Gill, D.; Lin, H.-C.; Michalakes, J.; Rizvi, S.; Zhang, X. *The Advanced Research WRF (ARW) Version 3 Modeling System User's Guide*, January 2010.
- (45) Bao, J. W.; Michelson, S. A.; Persson, P. O. G.; Djalalova, I. V.; Wilczak, J. M. Observed and WRF-simulated low-level winds in a high-ozone episode during the Central California Ozone Study. *J. Appl. Meteorol. Climatol.* **2008**, *47* (9), 2372–2394.
- (46) Michelson, S. A.; Djalalova, I. V.; Bao, J. W. Evaluation of the summertime low-level winds simulated by MMS in the Central Valley of California. *J. Appl. Meteorol. Climatol.* **2010**, *49* (11), 2230–2245.
- (47) Liu, Y.; et al. *An Implementation of Obs-Nudging-Based FDDA into WRF for Supporting ATEC Test Operations*, 2005 WRF user workshop, Paper 10.7, 2005.
- (48) Mass, C. F. University of Washington: Seattle, WA, personal communication. 2010.
- (49) Hu, J.; Howard, C. J.; Mitloehner, F.; Green, P. G.; Kleeman, M. J. Mobile source and livestock feed contributions to regional ozone formation in Central California. *Environ. Sci. Technol.* **2012**, *46* (5), 2781–2789.
- (50) Zhao, Z.; et al. The impact of climate change on air quality-related meteorological conditions in California. Part I: Present time simulation analysis. *J. Clim.* **2011**, *24* (13), 3344–3361.
- (51) Molders, N. Suitability of the weather research and forecasting (WRF) model to predict the June 2005 fire weather for interior Alaska. *Weather Forecasting* **2008**, *23* (5), 953–973.
- (52) Coniglio, M. C.; et al. Evaluation of WRF model output for severe weather forecasting from the 2008 NOAA hazardous weather testbed spring experiment. *Weather Forecasting* **2010**, *25* (2), 408–427.
- (53) Kleeman, M. J.; Cass, G. R. Source contributions to the size and composition distribution of urban particulate air pollution. *Atmos. Environ.* **1998**, *32* (16), 2803–2816.
- (54) Kleeman, M. J. Size distribution of trace organic species emitted from biomass combustion and meat charbroiling. *Atmos. Environ.* **2008**, *42* (13), 3059–3075.
- (55) Robert, M. A.; Kleeman, M. J.; Jakober, C. A. Size and composition distributions of particulate matter emissions: Part 2—Heavy-duty diesel vehicles. *J. Air Waste Manage. Assoc.* **2007**, *57* (12), 1429–1438.
- (56) Robert, M. A. Size and composition distributions of particulate matter emissions: Part 1—Light-duty gasoline vehicles. *J. Air Waste Manage. Assoc.* **2007**, *57* (12), 1414–1428.
- (57) Hu, J.; Zhang, H.; Ying, Q.; Chen, S. H.; Wiedinmyer, C.; Vandenbergh, F.; Kleeman, M. J. Long-term particulate matter modeling for health effects studies in California—Part I: Model performance on temporal and spatial variations. manuscript in preparation, **2014**.
- (58) Wiedinmyer, C.; et al. The Fire INventory from NCAR (FINN): A high resolution global model to estimate the emissions from open burning. *Geosci. Model Dev.* **2011**, *4* (3), 625–641.
- (59) Hodzic, A.; et al. Wildfire particulate matter in Europe during summer 2003: Meso-scale modeling of smoke emissions, transport and radiative effects. *Atmos. Chem. Phys.* **2007**, *7* (15), 4043–4064.
- (60) Database: California Air Quality Data—Selected Data Available for Download. <http://www.arb.ca.gov/aqd/aqcd/aqcdcdld.htm> (accessed in 2011).
- (61) Ham, W. A.; Kleeman, M. J. Size-resolved source apportionment of carbonaceous particulate matter in urban and rural sites in central California. *Atmos. Environ.* **2011**, *45* (24), 3988–3995.
- (62) Kleeman, M. J.; et al. Source apportionment of Fine (PM_{1.8}) and ultrafine (PM_{0.1}) airborne particulate matter during a severe winter pollution episode. *Environ. Sci. Technol.* **2009**, *43* (2), 272–279.
- (63) Riddle, S. G.; et al. Size-resolved source apportionment of airborne particle mass in a roadside environment. *Environ. Sci. Technol.* **2008**, *42* (17), 6580–6586.
- (64) Slinn, W. G. N.; et al. Some aspects of transfer of atmospheric trace constituents past air-sea interface. *Atmos. Environ.* **1978**, *12* (11), 2055–2087.
- (65) Kelly, J. T.; et al. Simulating emission and chemical evolution of coarse sea-salt particles in the Community Multiscale Air Quality (CMAQ) model. *Geosci. Model Dev.* **2010**, *3* (1), 257–273.
- (66) Goldman, G. T.; et al. Characterization of ambient air pollution measurement error in a time-series health study using a geostatistical simulation approach. *Atmos. Environ.* **2012**, *57*, 101–108.
- (67) Hoek, G.; et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* **2008**, *42* (33), 7561–7578.
- (68) Moore, D. K.; et al. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *J. Environ. Monit.* **2007**, *9* (3), 246–252.
- (69) Beckerman, B.; et al. Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. *Atmos. Environ.* **2013**, *77*, 172–177.
- (70) Kuhlwein, J.; et al. Emission-modelling in high spatial and temporal resolution and calculation of pollutant concentrations for comparisons with measured concentrations. *Atmos. Environ.* **2002**, *36*, S7–S18.
- (71) Lenhart, L.; Friedrich, R. European emission data with high temporal and spatial resolution. *Water, Air, Soil Pollut.* **1995**, *85* (4), 1897–1902.
- (72) Laurent, O.; Hu, J.; Li, L.; Cockburn, M.; Escobedo, L.; Kleeman, M.; Wu, J. Sources and contents of air pollution affecting term low birth weight in Los Angeles County, California, 2001–2008. *Environ. Res.* **2014**, under review.