

1 **Gas and aerosol carbon in California: comparison of**
2 **measurements and model predictions in Pasadena and**
3 **Bakersfield**

4
5 **K.R. Baker¹, A.G. Carlton², T.E. Kleindienst³, J.H. Offenberg³, M.R. Beaver³, D. R.**
6 **Gentner⁴, A. H. Goldstein⁵, P. L. Hayes⁶, J. L. Jimenez⁷, J. B. Gilman⁸, J. A. de**
7 **Gouw⁸, M. C. Woody³, H.O.T. Pye³, J.T. Kelly¹, M. Lewandowski³, M. Jaoui⁹, P.S.**
8 **Stevens¹⁰, W.H. Brune¹¹, Y.-H. Lin¹², C.L. Rubitschun¹², J. D. Surratt¹²**

9 [1](Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency,
10 Research Triangle Park, NC)

11 [2](Dept. of Environmental Sciences, Rutgers University, New Brunswick, NJ)

12 [3](Office of Research and Development, U.S. Environmental Protection Agency, Research
13 Triangle Park, NC)

14 [4](Department of Chemical and Environmental Engineering, Yale University, New Haven, CT)

15 [5](Department of Civil and Environmental Engineering, University of California, Berkeley, CA)

16 [6](Département de Chimie, Université de Montréal, Montréal, Québec, Canada)

17 [7](Department of Chemistry & Biochemistry, and CIRES, University of Colorado, Boulder,
18 Colorado)

19 [8](Chemical Sciences Division, Earth System Research Laboratory, National Oceanic and
20 Atmospheric Administration, Boulder, CO)

21 [9](Alion Science and Technology, Inc., Research Triangle Park, NC)

1 [10](Center for Research in Environmental Science, School of Public and Environmental Affairs
2 and Department of Chemistry, Indiana University, Bloomington, IN)

3 [11](Department of Meteorology, Pennsylvania State University, University Park, PA)

4 [12](Department of Environmental Sciences and Engineering, Gillings School of Global Public
5 Health, University of North Carolina at Chapel Hill, Chapel Hill, NC)

6 Correspondence to: K.R. Baker (baker.kirk@epa.gov)

7

8 **ABSTRACT**

9 Co-located measurements of fine particulate matter (PM_{2.5}) organic carbon, elemental carbon,
10 radiocarbon (¹⁴C), speciated volatile organic compounds (VOCs), and OH radical during the
11 CalNex field campaign provide a unique opportunity to evaluate the Community Multiscale Air
12 Quality (CMAQ) model's representation of organic species from VOCs to particles. Episode
13 averaged daily 23-hr average ¹⁴C analysis indicate PM_{2.5} carbon at Pasadena and Bakersfield
14 during the CalNex field campaign was evenly split between contemporary and fossil origin.
15 CMAQ predicts a higher contemporary carbon fraction than indicated by the ¹⁴C analysis at both
16 locations. The model underestimates measured PM_{2.5} organic carbon at both sites with very little
17 (7% in Pasadena) of the modeled mass represented by secondary production, which contrasts with
18 the ambient based SOC/OC fraction of 63% at Pasadena.

19

20 Measurements and predictions of gas-phase anthropogenic species, such as toluene and xylenes,
21 are generally within a factor of 2, but the corresponding secondary organic carbon (SOC) tracer
22 (2,3-dihydroxy-4-oxo-pentanoic acid) is systematically underpredicted by more than a factor of 2.

1 Monoterpene VOCs and SOCs are underestimated at both sites. Isoprene is underestimated at
2 Pasadena and over predicted at Bakersfield and isoprene SOC mass is underestimated at both sites.
3 Systematic model underestimates in SOC mass coupled with reasonable skill (typically within a
4 factor of 2) in predicting hydroxyl radical and VOC gas phase precursors suggests error(s) in the
5 parameterization of semi-volatile gases to form SOC. Yield values (α) applied to semi-volatile
6 partitioning species were increased by a factor of 4 in CMAQ for a sensitivity simulation, taking
7 in account recent findings of underestimated yields in chamber experiments due to gas wall losses.
8 This sensitivity resulted in improved model performance for PM_{2.5} organic carbon at both field
9 study locations and at routine monitoring network sites in California. Modeled percent secondary
10 contribution (22% at Pasadena) becomes closer to ambient based estimates but still contains a
11 higher primary fraction than observed.

12

13 **1 INTRODUCTION**

14 Secondary organic aerosol (SOA) forms in the atmosphere during the gas-phase photooxidation of
15 volatile organic compounds (VOCs) that produce semi-volatile and water-soluble gases that
16 condense to form new particles or partition to pre-existing aerosol mass (Ervens et al., 2011). SOA
17 contributes to the atmospheric fine particulate matter (PM_{2.5}) burden, with subsequent effects on
18 air quality, visibility, and climate (Hallquist et al., 2009). Despite its importance and abundance,
19 ambient SOA mass is not well characterized by atmospheric models (Wagstrom et al., 2014). For
20 example, the Community Multiscale Air Quality (CMAQ) model consistently underpredicts
21 surface SOA mass concentrations for a variety of seasons and locations when compared to ambient

1 observational estimates (Carlton and Baker, 2011;Carlton et al., 2010;Hayes et al., 2014;Zhang et
2 al., 2014a).

3
4 SOA formation and the preceding gas-phase photooxidation chemistry are complex and often
5 involve multiple oxidation steps in the gas, aqueous, and particle phase as well as accretion
6 reactions in the particle phase that yield high molecular weight (MW) products. However, three-
7 dimensional photochemical models must represent the gas-phase chemistry and SOA formation in
8 a simplified fashion for computational efficiency (Barsanti et al., 2013). Gas-phase chemical
9 mechanisms employ “lumped” VOC species, categorized primarily according to reactivity (e.g.,
10 reaction rate constants with $\cdot\text{OH}$) (Carter, 2000;Yarwood et al., 2005), not product volatility or
11 solubility. Condensable SOA-forming oxidation products are typically represented with 2
12 products in the standard versions of publically available and routinely applied photochemical
13 modeling systems such as GEOS-CHEM (Chung and Seinfeld, 2002;Henze and Seinfeld, 2006)
14 and WRF-CHEM (Grell et al., 2005) and those employed in regulatory applications for rulemaking
15 such as CMAQ (Carlton et al., 2010) and the Comprehensive Air Quality Model with extensions
16 (CAMx) (ENVIRON, 2014). Given the relationships between precursor VOC, OH radical
17 abundance and SOA formation, it is important to simultaneously evaluate the model representation
18 of all three, in particular, within the context of how organic species evolve in the atmosphere to
19 diagnose persistent SOA model bias.

20
21 Recent studies have shown that warm season SOA mass concentrations are usually greater than
22 primary organic aerosol (POA) mass in the Los Angeles (Docherty et al., 2008;Hersey et al.,

1 2011;Hayes et al., 2013) and Bakersfield (Liu et al., 2012) areas. Gas-to-particle condensation of
2 VOC oxidation products dominate formation of summer SOA in Bakersfield (Liu et al., 2012;Zhao
3 et al., 2013) and up to a third of nighttime organic aerosols (OA) in Bakersfield are organic nitrates
4 (Rollins et al., 2012). Sources of warm season OA in Bakersfield include fossil fuel combustion,
5 vegetative detritus, petroleum operations, biogenic emissions, and cooking (Liu et al., 2012;Zhao
6 et al., 2013). Despite numerous studies based on observations and models, less consensus exists
7 regarding the largest sources of warm season SOA at Pasadena. Bahreini et al. (2012) concluded
8 that SOA at Pasadena is largely derived from gasoline engines with minimal biogenic and diesel
9 fuel contribution (Bahreini et al., 2012). Others concluded large contributions from gasoline fuel
10 combustion to SOA but also found notable contributions from diesel fuel combustion, cooking,
11 and other sources (Gentner et al., 2012;Hayes et al., 2013). Zotter et al. (2014) conclude that 70%
12 of the SOA in the urban plume in Pasadena is due to fossil sources, and that at least 25% of the
13 non-fossil carbon is due to cooking sources. Lower volatility VOC measurements made at
14 Pasadena have been estimated to produce approximately 30% of fresh SOA in the afternoon with
15 a large contribution from non-onroad diesel fueled sources (Zhao et al., 2014).

16

17 Chemical measurements of PM_{2.5} carbon, fossil and contemporary aerosol carbon fraction, OC
18 and its components, SOC tracers and speciated VOCs taken as part of the 2010 California Research
19 at the Nexus of Air Quality and Climate Change (CalNex) field study in central and southern
20 California (Ryerson et al., 2013) provide a unique opportunity to quantitatively evaluate modeled
21 organic predictions. These special study data combined with routine PM_{2.5} OC measurements in
22 California are compared with model estimates to gauge how well the modeling system captures
23 the gas and aerosol carbon burden using the standard CMAQ aerosol approach. The SOC

1 mechanism in the base version of CMAQ lends itself well to comparison with chemical tracers
2 because it retains chemical identity traceable to the precursor VOC (Carlton et al., 2010). Finally,
3 a CMAQ sensitivity simulation was performed where the yields of semi-volatile gases from VOC
4 oxidation were increased by a factor of 4 (Zhang et al., 2014b) to determine whether this may
5 ameliorate the model underprediction of secondary organic carbon (SOC) seen here and in other
6 studies (Ensberg et al., 2014).

7

8 **2 METHODS**

9 Predictions of speciated VOC, speciated COC, and aerosol-phase carbon are simultaneously
10 compared to co-located ambient measurements at two surface locations, one in Los Angeles
11 County (Pasadena) and one in the San Joaquin Valley (Bakersfield) air basin. The CMAQ
12 photochemical model is applied with a fine grid resolution (4 km sized grid cells) using emissions
13 from the 2011 National Emissions Inventory and 2010 specific point source information where
14 available.

15

16 **2.1 Model Background**

17 CMAQ version 5.0.2 (www.cmaq-model.org) was applied to estimate air quality in California
18 from May 5 to July 1, 2010, coincident with the CalNex Study. Gas-phase chemistry is simulated
19 with the SAPRC07TB condensed mechanism (Hutzell et al., 2012) and aqueous-phase chemistry
20 that oxidizes sulfur, methylglyoxal, and glyoxal (Carlton et al., 2008; Sarwar et al., 2013). The
21 AERO6 aerosol chemistry module includes ISORROPIAII (Fountoukis and Nenes, 2007)
22 inorganic chemistry and partitioning. The modeling system generally does well capturing ambient

1 inorganic gases and PM_{2.5} species during this time period at Pasadena and Bakersfield (Kelly et
2 al., 2014;Markovic et al., 2014).

3
4 Model predicted OC species are shown in Figure 1 by volatility bin (log of C*) and O:C ratio (see
5 Supporting Information for related details). Aqueous-phase species are shown with blue circles,
6 species largely fossil in origin are colored brown and those non-fossil in origin are green. A
7 general trend of increasing O:C ratio as volatility decreases is consistent with laboratory and field
8 measurements (Jimenez et al., 2009). The placement of the MGLY gem-diol vertically above gas-
9 phase MGLY in Figure 1 represents hydration processes. Aqueous-phase organic chemistry
10 represents multiple processes, including functionalization and oligomerization because some
11 photooxidation products are small carboxylic acids and others are high MW species (Tan et al.,
12 2010;Carlton et al., 2007).

13
14 VOC precursors for SOA include isoprene, monoterpenes, sesquiterpenes, xylenes, toluene,
15 benzene, alkanes, glyoxal, and methylglyoxal (Figure 1 right panel). Benzene, toluene, and xylene
16 form SOA precursors with high-NO_x (RO₂+NO) and low-NO_x (RO₂+HO₂) specific yields
17 (Carlton et al., 2010). CMAQ converts these precursors into multiple semi-volatile products
18 (Figure 1 middle panel) after a single oxidation step. These multiple products vary in terms of
19 assigned volatility and oxygen-to-carbon (O:C) ratio. All semi-volatile SOA mass oligomerizes in
20 CMAQ. After this process SOA identity is classified only as anthropogenic or biogenic, dependent
21 on the VOC precursor (see Figure S2). After oligomerization, the saturation vapor pressure (C*)
22 and OM:OC ratio associated with all of the 2-product semi-volatile SOA species change from the

1 individual values to the values assigned for non-volatile, non-partitioning oligomerized SOA mass
2 ($C^* \approx 0$; OM:OC = 2.1) (Carlton et al., 2010).

3
4 CMAQ VOCs and SOC species are paired in time and space with measurements (Table S2).
5 Modeled predictions are averaged temporally to match observations and extracted from the grid
6 cell where the monitor is located. Modeled toluene and xylene SOC are aggregated to match the
7 measured SOC tracer (2,3-dihydroxy-4-oxopentanoic acid) which is known to represent products
8 from both compounds and potentially other methylated aromatics (Kleindienst et al., 2004).
9 Because the original VOCs contributing to oligomerized species are not tracked by CMAQ,
10 biogenic oligomerized species mass is apportioned to parent VOC based on the fraction each semi-
11 volatile SOC species contributes to the total semi-volatile (non-oligomerized) biogenic SOC at
12 that time and location. The same technique is applied to anthropogenic SOC.

13

14 **2.2 Model Application**

15 The model domain covers the State of California and part of northwest Mexico using 4 km square
16 sized grid cells (Figure S1). The vertical domain extends to 50 mb using 34 layers (layer 1 top ~35
17 m) with most resolution in the boundary layer. Initial and boundary conditions are from a coarser
18 CMAQ simulation that used 3-hourly varying input from a global GEOS-Chem (v8-03-02) global
19 model (<http://acmg.seas.harvard.edu/geos/>) simulation for the same period (Henderson et al.,
20 2014). The coarser continental U.S. CMAQ simulation was run continuously from December 2009
21 and the first week of the finer 4 km CMAQ simulation was not used to minimize the influence of
22 initial chemical conditions. Gridded meteorological variables are generated using the Weather

1 Research and Forecasting model (WRF), Advanced Research WRF core (ARW) version 3.1
2 (Skamarock et al., 2008). Surface meteorology including temperature, wind speed, and wind
3 direction and daytime mixing layer height were well characterized by WRF in central and southern
4 California during this period (Baker et al., 2013).

5

6 Emissions are processed to hourly gridded input for CMAQ with the Sparse Matrix Operator
7 Kernel Emissions (SMOKE) modeling system (<http://www.cmascenter.org/smoke/>). Hourly solar
8 radiation and temperature estimated by the WRF model are used as input to the Biogenic Emission
9 Inventory System (BEIS) v3.14 to generate hourly emissions estimates of biogenic speciated VOC
10 and NO (Carlton and Baker, 2011). Electric generating point source emissions that report
11 continuous emissions monitor (CEM) data are used in the modeling to reflect 2010 emissions
12 information. Day specific fires are represented but minimally impacted air quality during this
13 period (Hayes et al., 2013). Mobile source emissions were generated using the SMOKE-MOVES
14 integration approach (United States Environmental Protection Agency, 2014) and then
15 interpolated between totals provided by the California Air Resources Board for 2007 and 2011.
16 Other anthropogenic emissions are based on the 2011 National Emissions Inventory (NEI) version
17 1 (United States Environmental Protection Agency, 2014). Primary mass associated with carbon
18 (non-carbon organic mass, NCOM) is estimated based on sector specific organic matter-to-organic
19 carbon (OM:OC) ratios (Simon and Bhave, 2012).

20

21 Emissions of primarily emitted PM_{2.5} OC and the sum of anthropogenic SOA precursors benzene,
22 toluene, and xylenes (BTX) are shown in Table 1 by source sector and area. Here, the southern

1 San Joaquin valley includes emissions from Kern, Tulare, Kings, and Fresno counties and the Los
2 Angeles area include emissions from Los Angeles and Orange counties. BTX emissions in both
3 areas are dominated by mobile sources (onroad and offroad) and area sources such as solvent
4 utilization and waste disposal (Table S1). Primary OC emissions are largely commercial cooking
5 (non-point area) in both locations with notable contribution from various types of stationary point
6 and mobile sources. BTX emissions are almost completely fossil in origin and primarily emitted
7 OC is split fairly evenly between contemporary and fossil origin in these areas based on the 2011
8 version 1 NEI (Table 1).

9

10 **2.3 Sampling and Analysis Methods**

11 CalNex ground-based measurements took place in Pasadena, CA, from 15 May – 15 June 2010
12 and in Bakersfield, CA, from 15 May – 30 June 2010. The Bakersfield sampling site was located
13 in a transition area of southeast Bakersfield between the city center and areas of agricultural
14 activity. The Pasadena sampling site was located on the California Institute of Technology campus
15 with the Los Angeles metropolitan area to the southwest and San Gabriel Mountains directly north
16 (see Figures S3).

17

18 An ambient-based approach is used here to estimate secondary OC from individual or groups of
19 similar hydrocarbons (Kleindienst et al., 2010). Concentrations of specific, tracer, compounds are
20 determined and used to estimate SOC contributions from the particular source groups based on
21 measured laboratory tracer-to-SOC mass fractions (Kleindienst et al., 2007). To conduct tracer-
22 based organic aerosol characterization, filter-based particulate matter sampling was conducted at
23 each site for 23-h periods starting at midnight (PDT) of the designated sampling day. In total, there

1 were 32 filter samples from Pasadena and 36 from the Bakersfield site (Lewandowski et al., 2013).
2 The filter sampling protocols have been described in detail elsewhere (Kleindienst et al., 2010).
3 For the analysis of the SOC tracer compounds, filters and field blanks were treated using the
4 derivatization method described by Kleindienst et al. (Kleindienst et al., 2007). The mass spectral
5 analysis for the organic compounds used as secondary molecular tracers has been described (Edney
6 et al., 2003). The method detection limit (MDL) for the SOC tracer species is 0.1 ng m^{-3} . Additional
7 details of this methodology are provided in the Supporting Information.

8

9 OC and elemental carbon (EC) concentrations were determined using the thermal-optical
10 transmittance (TOT) method (Birch and Cary, 1996) from 1.54 cm^2 punches of quartz filters
11 collected concurrent with the filters used for tracer analyses (hereafter referred to as UNC/EPA
12 OC). The outer non-loaded rings were removed from filter samples then sent to Woods Hole
13 Oceanographic Institute Accelerator Mass Spectrometry for ^{14}C analysis. The fraction of
14 contemporary carbon is provided for each daily total $\text{PM}_{2.5}$ carbon sample (Geron, 2009). The
15 contemporary carbon fraction is expressed as a percentage of an oxalic acid standard material that
16 represents the carbon isotopic ratio for wood growth during 1890 (Stuiver, 1983). To account for
17 the atmospheric ^{14}C enhancement due to nuclear bomb testing in the 1950s and 1960s, a factor of
18 1.044 was used to calculate the contemporary carbon fraction from the measured modern carbon
19 result (Lewis et al., 2004; Zotter et al., 2014). The fraction of contemporary $\text{PM}_{2.5}$ organic carbon
20 is estimated based on Zotter et al., (2014).

21

1 Two VOC datasets (one canister based, and one *in situ*) from each site were used in this analysis.
2 Three hour integrated (06:00 – 09:00 PDT) canister samples for VOC analysis were collected at
3 both sites. A total of 41 samples were collected at the Bakersfield site and 31 at Pasadena. The
4 offline VOC analysis details are given in the supplemental/supporting material. In Bakersfield,
5 online VOC mixing ratios were collected for 30 minutes on the hour, and analyzed via gas
6 chromatography-flame ionization detector (GC-FID) and gas chromatography-mass spectrometry
7 (GC-MS) (Gentner et al., 2012). In Pasadena, online VOC measurements were collected for 5
8 minutes every 30 minutes and analyzed via GC-MS (Borbon et al., 2013; Gilman et al., 2010).
9 Carbon monoxide measurements at Pasadena were determined using UV fluorescence (Gerbig et
10 al., 1999).

11
12 Hydroxyl (OH) and hydroperoxyl (HO₂) radical measurements were made at both locations using
13 Fluorescence Assay with Gas Expansion (FAGE). The Bakersfield OH measurements used in this
14 analysis were collected using the OH_{chem} method from the Penn State ground-based FAGE
15 instrument (Mao et al., 2012). The Pasadena HO_x observations were made using the Indiana
16 University FAGE instrument (Dusanter et al., 2009). HO₂ measurements from both instruments
17 could contain an interference from various RO₂, therefore when comparing the model output with
18 the observations, the sum of modeled HO₂ and RO₂ has been used (Griffith et al., 2013).

19
20 OC measurements from nearby Chemical Speciation Network (CSN) sites in Pasadena and
21 Bakersfield were also used for comparison purposes. The Los Angeles CSN site (60371103) was
22 approximately 9 miles from the CalNex site, and the Bakersfield CSN site (60290014) was

1 approximately 3 miles from the CalNex site (see Figures S3a and S3b in the supporting
2 information). The CSN network uses quartz-fiber filters and analyzes the carbon off-line using the
3 thermal-optical reflectance method. Aerodyne High-Resolution Time-of-Flight Aerosol Mass
4 Spectrometer (AMS) measurements of PM₁ OC made at Pasadena are described in Hayes et al.,
5 2013 and online Sunset PM_{2.5} OC measurements made at Bakersfield are described in Liu et al.,
6 2012.

7

8 **3 Results & Discussion**

9 The results and discussion are organized such that the contemporary and fossil components of
10 PM_{2.5} carbon at the Pasadena and Bakersfield sites are discussed, followed by model performance
11 for PM_{2.5} carbon, speciated VOC, and SOC tracer groups. Table 2 shows episode aggregated model
12 performance metrics for PM_{2.5} organic and elemental carbon, SOC tracers, total VOC, and select
13 VOC species. The results of a sensitivity increasing semi-volatile yields are presented throughout
14 and discussed in detail before finally providing an evaluation of PM_{2.5} carbon at all routine monitor
15 sites in California.

16

17 **3.1 Contemporary and Fossil Origins of PM_{2.5} Carbon**

18 Field campaign average total PM_{2.5} carbon measurements indicate nearly equal amounts of
19 contemporary and fossil contribution at Pasadena and Bakersfield. The field study average
20 contemporary fraction of 23-hr average PM_{2.5} total carbon samples is 0.51 at Bakersfield (N=35)
21 and 0.48 at Pasadena (N=25). The estimate for contemporary carbon fraction at Pasadena is

1 consistent with other ^{14}C measurements at this location for this period (Zotter et al., 2014) and
2 similar to measurements made at urban areas in the Southeast United States: Birmingham 52%
3 and Atlanta 63% contemporary carbon (Kleindienst et al., 2010).

4

5 Figure 2 shows observed daily 23-hr average elemental carbon and OC shaded by contemporary
6 and fossil component. The fractional contribution of contemporary carbon to total $\text{PM}_{2.5}$ carbon is
7 variable from day-to-day at the Pasadena site and steadily increases through the study period at
8 the Bakersfield location (first week average of 0.44 and final week average of 0.58). Some of the
9 contemporary carbon fraction measurements from Pasadena were above 1.0. These samples were
10 considered erroneous and not included in the analysis and suggest the possibility of positive biases
11 due to nearby sources (e.g. medical incinerator) in the area. It is possible some of the stronger day-
12 to-day variability in contemporary carbon fraction measurements at Pasadena may be related to
13 biases due to nearby “hot” sources. Higher time resolution ^{14}C measurements at Pasadena show an
14 increase in fossil fraction during the middle of the day related to increased emissions of fossil
15 $\text{PM}_{2.5}$ carbon precursors and SOA formation in the Los Angeles area (Zotter et al., 2014). $\text{PM}_{2.5}$
16 OC of fossil origin at Pasadena shows the strongest relationship to daily average temperature
17 (Figure S4a) compared with contemporary carbon, total carbon, and elemental carbon. At
18 Bakersfield the relationship between daily average temperature and fossil and contemporary
19 carbon is similar (Figure S4b) and not as strong as the relationship in Pasadena. Neither fossil nor
20 contemporary carbon concentrations show discernible patterns by day of the week at either
21 location (Figure S5).

22

1 Episode average ambient estimates of PM_{2.5} OC contemporary fraction (Bakersfield=0.54 and
2 Pasadena=0.51) are similar to the estimated contemporary fraction of the urban emissions of
3 primary PM_{2.5} OC (Bakersfield=.53 and Pasadena=.51), as noted in Table 1. Modeled
4 contemporary PM_{2.5} carbon is estimated by summing primarily emitted PM_{2.5} multiplied by the
5 contemporary fraction of urban area emissions (see Section 2.1 and Table 1) with model estimated
6 biogenic SOC species. The average baseline modeled contemporary fraction of PM_{2.5} OC in
7 Pasadena is 0.51 and Bakersfield 0.54, both of which are similar to average observation estimates.
8 However, the model shows little day to day variability in contemporary carbon fraction which does
9 not match observed trends (Figure S6).

10

11 **3.2 PM_{2.5} Carbon**

12 Figure 3 shows measured (UNC/EPA data) and modeled PM_{2.5} OC at Bakersfield and Pasadena.
13 Organic carbon measurements from co-located instruments (AMS at Pasadena measured PM₁ and
14 Sunset at Bakersfield measured PM_{2.5}) and a nearby CSN monitor are also shown in Figure 3. The
15 co-located AMS measurements compare well with the UNC/EPA PM_{2.5} organic carbon
16 measurements at Pasadena, while the concentrations measured at the nearby CSN site are
17 substantially lower. At Bakersfield UNC/EPA measurements compared with the nearby CSN
18 (episode average ~3 times lower) and co-located daily average Sunset (episode average 20%
19 lower) measured PM_{2.5} OC illustrate possible measurement artifacts in the CalNex measurements
20 at this location. These differences in measured concentration at Bakersfield may be related to filter
21 handling, variability in collected blanks, true differences in the OC concentrations since the CSN
22 site is spatially distinct, differences in the height of measurement (these CSN monitors are situated

1 on top of buildings), and differences in analytical methods since CSN sites use the thermal optical
2 reflectance (TOR) to operationally define OC and EC.

3

4 PM_{2.5} OC is underestimated at both CalNex locations (Figure 3), most notably at Bakersfield.
5 However, given the large differences in PM_{2.5} OC mass compared to co-located and nearby routine
6 measurements, it is not clear which measurement best represents ambient PM_{2.5} OC concentrations
7 and would be most appropriate for comparison with the model. The model generally compares
8 well to the CSN site nearest Pasadena and Bakersfield. PM_{2.5} elemental carbon is well
9 characterized by the model at Bakersfield (fractional bias = -19% and fractional error = 36%) and
10 over-estimated at Pasadena (fractional bias and error = 126%) (Figure S7). Since the emissions are
11 based on TOR and UNC/EPA measurements use the TOT operational definition of total carbon
12 some of the model overestimation may be related to the TOR method estimating higher elemental
13 carbon fraction of total carbon (Chow et al., 2001).

14

15 PM_{2.5} OC is mostly primary (Pasadena 93% and Bakersfield 88%) in the baseline model
16 simulation. AMS measurements at Pasadena suggest OC is mostly secondary in nature with an
17 average of 63% for the SVOOA and OOA components for this field study (Hayes et al., 2013).
18 Model estimated PM_{2.5} OC is largely from primarily emitted sources and contemporary in nature
19 based on the contemporary/fossil split of primary PM_{2.5} emissions near both sites (Figure S6).
20 Primarily emitted PM_{2.5} OC emissions sources near Pasadena and Bakersfield include mobile
21 sources, cooking, and dust based on emissions inventory information (Table 1). Some of these
22 sources of primarily emitted PM_{2.5} OC may be semi-volatile in nature. Model treatment of POA

1 as semi-volatile may improve the primary-secondary comparison with observations but would
2 likely exacerbate underpredictions of PM_{2.5} OC, unless oxidation and re-partitioning of the
3 products is considered (Robinson et al., 2007). The underestimation of SOC may result from
4 underestimated precursor VOC, poorly characterized oxidants, underestimated semi-volatile
5 yields, missing intermediate volatility VOC emissions (Stroud et al., 2014; Zhao et al., 2014), other
6 issues, or some combination of each.

7

8 **3.3 Gas-phase carbon**

9 Model estimates are paired with hourly VOC (Figure S8) and mid-morning 3-hr average VOC
10 (Figure S9) at both locations. Compounds considered largely fossil in origin including xylene,
11 toluene, and benzene are generally well predicted at both sites although these species tend to be
12 slightly overestimated at Pasadena and slightly underestimated at Bakersfield. Since emissions of
13 these compounds near these sites are mostly from mobile sources (Table 1), this suggests emissions
14 from this sector are fairly well characterized in this application.

15

16 Contemporary (biogenic) origin monoterpenes are underestimated at both sites while isoprene is
17 underestimated at Pasadena and has little bias at Bakersfield based on hourly measurements
18 (Figure S8; Table 2). Isoprene and monoterpene performance may be related to poorly
19 characterized meteorology that may not capture transport from nearby large emitting vegetation to
20 these monitor locations (Heo et al., 2015), deficiencies in emissions factors, or poorly
21 characterized vegetation. Speciated monoterpene measurements made at Bakersfield during this
22 field campaign suggest emissions of certain species were elevated at the start of this time period

1 due to flowering (Gentner et al., 2014b), which is a process not included in current biogenic
2 emissions models thus it may contribute to modeled monoterpene underestimates.

3

4 Other VOC species that are systematically underestimated include ethane, methanol, ethanol, and
5 acetaldehyde. Underprediction of methanol and ethanol in Bakersfield may be largely related to
6 missing VOC emissions for confined animal operations in the emission inventory (Gentner et al.,
7 2014a). Underestimates of oxygenated VOC compounds may indirectly impact SOC formation
8 through muted photochemistry (Steiner et al., 2008). Carbon monoxide tends to be underestimated
9 at both locations (Figure S8), possibly due to boundary inflow concentrations from the global
10 model simulation being too low.

11

12 **3.4 PM_{2.5} SOC tracers**

13 Figure 4 shows modeled and measured total PM_{2.5} OC mass. Measured mass explained by fossil
14 and contemporary SOC tracers are shown in the top row. The observed unexplained fraction
15 origins are a mixture of primary, secondary, fossil and contemporary origin. Modeled mass is
16 colored to differentiate primarily emitted OC and SOC. Estimates of SOC mass from a specific or
17 lumped VOC group (e.g. isoprene, monoterpenes, toluene) as estimated by specific tracer species,
18 hereafter called SOC tracer mass, comprise little of the measured or modeled PM_{2.5} OC at either
19 of these locations during this field study (Figure 4). Total SOC tracer estimates explain only 9%
20 of the total measured UNC/EPA PM_{2.5} OC at Pasadena and 5% at Bakersfield. The percentage of
21 mass explained by known secondary tracers is smaller than urban areas in the southeast United
22 States: Atlanta 27% and Birmingham 31% (Kleindienst et al., 2010). The portion of measured and

1 modeled PM_{2.5} carbon not identified with tracers may be from underestimated adjustment factors
2 related to previously uncharacterized SVOC wall loss in chamber studies (Zhang et al., 2014b),
3 unidentified SOC pathways, known pathways without an ambient tracer, and tracer degradation
4 between formation and measurement. Based on ¹⁴C measurements, this unidentified portion of the
5 measurements is likely comprised of both contemporary and fossil carbon in generally similar
6 amounts. Total modeled SOC explain only 12% of the PM_{2.5} carbon at Bakersfield and 7% at
7 Pasadena. As noted previously, AMS based observations suggest most OC is SOC (63%) at
8 Pasadena (Hayes et al., 2013) meaning both the SOC tracer measurements and model estimates
9 explain little of the SOC at this location.

10

11 Despite the relatively small component of total PM_{2.5} carbon explained by SOC tracers, a
12 comparison of measured and modeled SOC and precursor VOC provides additional opportunity
13 to better understand sources of PM_{2.5} carbon in these areas and begin to establish relationships
14 between precursors and resulting SOC formation. Ambient and model estimated SOC tracers and
15 daily average VOC precursors are shown in Figure 5 for Pasadena and Figure 6 for Bakersfield.
16 The model underestimates toluene and xylene SOC at both locations even though the VOC gas
17 precursors show an overprediction tendency at Pasadena and slight underestimation at Bakersfield.
18 Isoprene SOC is generally under predicted at both sites, in particular at Bakersfield. This is in
19 contrast to the slight overprediction of daily 24-hr average isoprene at Bakersfield. One
20 explanation may be that isoprene SOC is formed elsewhere in the region (e.g. the nearby foothills
21 of the Sierra Nevada where emissions are highest in the region), which would support the lack of
22 relationship between isoprene SOC and isoprene concentrations at Bakersfield (Shilling et al.,
23 2013). The lack of relationship could also be related to the reactive uptake kinetics of isoprene-

1 derived epoxydiols (IEPOX) (Gaston et al., 2014) and methacrylic acid epoxide (MAE). Since the
2 model does not include the reactive uptake of IEPOX and MAE and subsequent acid-catalyzed
3 aqueous phase chemistry it is likely isoprene SOC would be underestimated to some degree at both
4 sites (Karambelas et al., 2013;Pye et al., 2013). Of these channels the IEPOX channel is thought
5 to have the largest SOA production potential, but the chemistry in the LA basin proceeds almost
6 completely through the high-NO channel (Hayes et al., 2014), and thus IEPOX is not formed for
7 isoprene emitted within the LA basin. Consistent with that observation, the AMS tracer of IEPOX
8 SOA is only detected at background level in the LA basin.

9
10 Monoterpene VOC and monoterpene SOC are underestimated systematically at both locations
11 suggesting underpredictions of the VOC precursor translates to underestimates in SOC. As noted
12 previously, monoterpene measurements suggest an emissions enhancement related to flowering or
13 other emission events (e.g. harvest or pruning) (Gentner et al., 2014b) that is not included in current
14 biogenic emissions model formulations. The monoterpene measured tracer SOC group is based on
15 α -pinene products. Measured SOC at these sites could be from monoterpene species other than α -
16 pinene. A coincident study near Bakersfield indicates α - and β -pinene emissions represent a fairly
17 small fraction of total monoterpene emissions during this time period (Gentner et al., 2014b). SOA
18 yields in CMAQ for monoterpenes are heavily weighted toward α - and β - pinene, which may be
19 appropriate in most places, but not here where measurements show large contributions from
20 limonene, myrcene, and para-cymene. This is important because yields from different
21 monoterpenes vary and limonene has a much larger SOA yield than pinenes (Carlton et al., 2010).

1 Sesquiterpene VOC and SOC tracer (β -caryophyllenic acid) mass measurements were never above
2 the MDL at either site during CalNex, but the modeling system often predicts SOC from this VOC
3 group (Table 2, Figure S10b). The SOC tracer measurement methodology is more uncertain for
4 sesquiterpene products (Offenberg et al., 2009) and gas-phase sesquiterpenes would have oxidized
5 before reaching the measurement sites since sesquiterpene emitting vegetation exists in the San
6 Joaquin Valley (Ormeño et al., 2010). It is also possible that SOC is forming from sesquiterpenes
7 other than β -caryophyllene.

8

9 One potential explanation for an underestimation of SOC despite well characterized precursors
10 (e.g. toluene and xylenes) could be lack of available oxidants. Modeled anthropogenic SOC is
11 approximately 46% and 54% through the low-NO_x pathway at Pasadena and Bakersfield
12 indicating both hydroxyl and hydroperoxyl radical representation is important at these locations.
13 As shown in Figure 7, the model tends to overestimate the hydroxyl radical compared with
14 measurement estimates at Pasadena. Hydroperoxyl+peroxy radical measurements are
15 underestimated at Pasadena by a factor of 2 on average. The model overestimates preliminary
16 measurements of both hydroxyl (by nearly a factor of 2 on average) and hydroperoxyl+peroxy
17 radicals at Bakersfield. Model representation of hydroxyl radical at these locations during this time
18 period does not seem to be limiting VOC oxidation to semi-volatile products. Better agreement
19 between radical ambient and modeled estimates could result in less SOC produced by the model
20 and exacerbate model SOC underestimates. This suggests deficiencies other than radical
21 representation by the modeling system are more influential in SOC performance for these areas.
22 However, hydroperoxyl underestimates at Pasadena could lead to muted SOA formation through

1 low-NO_x pathways dependent on hydroperoxyl concentrations and contribute to model under-
2 estimates of SOC.

3

4 **3.5 Sensitivity Simulation**

5 OH is not underestimated in the model and biases in precursor VOC do not clearly translate into
6 similar biases in SOC (e.g. toluene and xylene VOC are overestimated at Pasadena but tracer SOC
7 for this group is underestimated) for these sites during this time period. Modeled SOC may partly
8 be underestimated due to the use of experimental SOC yields that may be biased low due to
9 chamber studies not fully accounting for SVOC wall loss (Zhang et al., 2014b). Even though Zhang
10 et al., 2014b showed results for one precursor to SOA pathway, as a sensitivity study here the yield
11 of all semivolatile gases are increased by a factor of 4. This was done by increasing in the mass-
12 based stoichiometric coefficients for each VOC to SOA pathway in the model to provide a
13 preliminary indication about how increased yields might impact model performance. A factor of
14 4 is chosen based on the upper limit related to SVOC wall loss in (Zhang et al., 2014b). Aside
15 from wall loss characterization, there are a variety of other aspects of chamber studies that could
16 result in underestimated yields including particle-phase accretion and aqueous phase chemistry
17 and differences in chamber and ambient humidity.

18

19 Model estimates of PM_{2.5} OC increase in urban areas and regionally when semivolatile yields are
20 increased. The sensitivity simulation results in episode average anthropogenic SOC increases by a
21 factor of 3 (benzene SOC at Pasadena) to 4.8 (toluene and xylene SOC at Pasadena) and biogenic
22 SOC increases between a factor of 5.1 (isoprene SOC at Pasadena) to 8.9 (monoterpene SOC at

1 Bakersfield). Model performance improves at the CalNex locations (Figures 3 and 4) and at routine
2 monitors throughout California (Figure 8). Average fractional bias improves from -63% to -25%
3 at routine monitor locations and fractional error is reduced from 75% to 51%.

4
5 The sensitivity simulation with increased semivolatile yields results in increased model estimated
6 secondary contribution as a percent of total PM_{2.5} carbon, but still does not conform to observation
7 based estimates that indicate PM_{2.5} carbon is largely secondary in nature at these sites (Liu et al.,
8 2012; Hayes et al., 2013). Modeled SOC in the sensitivity simulation explains 36% of the PM_{2.5}
9 OC at Bakersfield and 22% at Pasadena, which is larger than the baseline simulation by more than
10 a factor of 3. The model predicted percent contemporary fraction of PM_{2.5} carbon changed very
11 little due to this sensitivity. The model sensitivity results are not compared to SOC tracer group
12 estimates since the conversion of tracer concentrations to SOC concentrations would require a
13 similar adjustment and would result in similar relationships between model estimates and
14 observations.

16 **3.6 Aqueous and other SOC processes**

17 Measurements in Pasadena during the summer of 2009 suggest aqueous processes can be important
18 for SOC mass (Hersey et al., 2011). For the CalNex period at Pasadena, Washenfelder et al. (2012)
19 showed box model estimated 8-hr average SOC from aqueous-phase chemistry of glyoxal to be
20 between 0.0 and 0.2 $\mu\text{g m}^{-3}$ (Washenfelder et al., 2011) and Hayes et al. (2014) showed that the
21 observed SOA was not different between cloudy and clear morning days. CMAQ predicted 24-
22 hour average SOC from glyoxal and methyglyoxal through aqueous chemistry at Pasadena ranges

1 from 0.0 to 0.04 $\mu\text{g}/\text{m}^3$. CMAQ estimates of SOC from small carbonyl compounds via aqueous-
2 phase processes are within the range inferred from measurements.

3

4 Not all CMAQ SOC formation pathways can be included in this analysis. No observational
5 indicator exists for SOC derived from alkanes, benzene, glyoxal, and methylglyoxal since unique
6 tracer species have not been determined. Conversely, naphthalene/PAH SOC tracers were
7 measured, but not modeled in CMAQ. Measured naphthalene SOC at these sites is minor (Hayes
8 et al., 2014) which is consistent with other areas (Dzepina et al., 2009). Previous CMAQ
9 simulations predict that PAHs contribute less than 30 ng m^{-3} of SOA in Southern California in
10 summer (Pye and Pouliot, 2012), and thus including those pathways is unlikely to close the model-
11 measurement gap in total OC. 2-Methyl-3-Buten-2-ol (MBO) derived SOC concentrations (3-4
12 ngC m^{-3}) were low at both monitor locations throughout the campaign (Lewandowski et al., 2013).
13 MBO does not appear to notably contribute SOC at these locations during this time period, which
14 is consistent with low yields estimated in laboratory experiments (Chan et al., 2009). Organic
15 carbon emitted from marine biological activity is not included in this modeling assessment and
16 may contribute to some degree at Pasadena (Gantt et al., 2010) based on ship-based measurements
17 (Hayes et al., 2013).

18

19 **3.7 Regional PM_{2.5} Organic Carbon**

20 Including routine measurement data is important to characterize model performance beyond the
21 two CalNex monitor locations and provide broader context for PM_{2.5} carbon in California and
22 understand how the model performs and responds to perturbations at diverse locations. The highest

1 average modeled PM_{2.5} OC in California during this period is in the Los Angeles area (Figure 8).
2 The Sacramento and San Joaquin valleys also show higher concentrations of PM_{2.5} OC than more
3 rural parts of the State (Figure 8). Measurements made at routine monitor networks (Figure 8)
4 show similar elevated concentrations near Los Angeles, Sacramento valley, and San Joaquin
5 valley. These areas of elevated OC generally coincide with areas of the State that experience a
6 build-up of pollutants due to terrain features blocking air flow (Baker et al., 2013). The model does
7 not tend to capture the highest concentrations of measured PM_{2.5} OC in the central San Joaquin
8 valley, Imperial Valley, or at one CSN monitor in the northeast Sierra Nevada that is near large
9 residential wood combustion emissions (Figure S11). The model underestimates PM_{2.5} OC on
10 average across all CSN sites during this time period (fractional bias = -63% and fractional error =
11 75%). The modeling systems show a slight overprediction tendency (fractional bias = 27%) across
12 all CSN sites for PM_{2.5} elemental carbon in California during this period.

13

14 **4 Conclusions**

15 Total PM_{2.5} carbon at Pasadena and Bakersfield during the CalNex period in May and June of 2010
16 is fairly evenly split between contemporary and fossil origin. Total PM_{2.5} OC is generally
17 underestimated at both field study locations and at many routine measurement sites in California
18 and comparison with AMS observations suggest a large underestimate of SOC. Semivolatile yields
19 were increased by a factor of 4 based on recent research suggesting yields may be higher due to
20 updated accounting for SVOC wall loss. This sensitivity resulted in a better comparison to routine
21 and field study measurements. However, the model estimated OC is still largely primary in nature
22 and inconsistent with observation based approaches at these sites. A modeling study for the same

1 time period using different emissions, photochemical transport model, and SOA treatment also
2 show underestimated OA and SOA at Pasadena and underestimated SOA but comparable OA at
3 the Bakersfield location (Fast et al., 2014).

4

5 CMAQ predictions of individual VOCs are often not consistent with model performance for the
6 corresponding subsequent SOC species mass. Gas-phase mixing ratios of toluene and xylene are
7 well-predicted by CMAQ, typically within a factor of 2 of the observations at both sites. However,
8 measurement-based estimates of the corresponding SOC mass are consistently greater than model-
9 predicted mass. Mass concentrations of the isoprene SOC are systematically underpredicted, most
10 noticeably at Bakersfield, while model predictions of gas-phase isoprene are not biased in only
11 one direction to the same degree. Gas-phase monoterpenes and the related SOC species are
12 underpredicted at both CalNex monitoring sites. The hydroxyl radical is fairly well characterized
13 at Pasadena and systematically overestimated at Bakersfield suggesting oxidants are not limiting
14 SOC production in the model.

15

16 Episode average CMAQ model estimates of PM_{2.5} OC contemporary fraction at Pasadena and
17 Bakersfield are similar to radiocarbon measurements but lack day to day variability. CMAQ PM_{2.5}
18 OC is predominantly primary in origin which is contrary to findings from other studies that PM_{2.5}
19 OC in these areas are largely secondary in nature during this time period (Bahreini et al.,
20 2012; Hayes et al., 2013; Liu et al., 2012). Treatment of primarily emitted PM_{2.5} OC as semi-volatile
21 would likely result in total PM_{2.5} OC estimates that would be mostly secondary rather than primary.
22 Some model performance features including underestimated SOC may be related to less volatile

1 hydrocarbons emitted from some source categories missing from the emission inventory (Chan et
2 al., 2013;Gentner et al., 2012;Jathar et al., 2014;Zhao et al., 2014) or mischaracterized when
3 lumped into chemical mechanism VOC species (Jathar et al., 2014). A future intent is to simulate
4 this same period using a volatility basis set approach to treat primary OC emissions with some
5 degree of volatility and potential for SOC production and better account for sector specific
6 intermediate volatility emissions.

7

8 **Disclaimer**

9 Although this work was reviewed by EPA and approved for publication, it may not necessarily
10 reflect official Agency policy.

11

12 **Acknowledgements**

13 The authors would like to acknowledge measurements taken by Scott Scheller and the contribution
14 from Chris Misenis, Allan Beidler, Chris Allen, James Beidler, Heather Simon, and Rich Mason.
15 EPA, through its Office of Research and Development, funded and collaborated in the research
16 described here under Contract EP-D-10-070 to Alion Science and Technology. This work is
17 supported in part through EPA's STAR program, grant number RD83504101. PLH and JLJ were
18 supported by CARB 11-305.

19

20 **Supporting Information**

1 Additional model output, comparison with measurements and formulas used for data pairing are
2 provided in the Supporting Information.

3

4 **References**

- 5 Bahreini, R., Middlebrook, A. M., de Gouw, J. A., Warneke, C., Trainer, M., Brock, C. A., Stark,
6 H., Brown, S. S., Dube, W. P., Gilman, J. B., Hall, K., Holloway, J. S., Kuster, W. C., Perring, A.
7 E., Prevot, A. S. H., Schwarz, J. P., Spackman, J. R., Szidat, S., Wagner, N. L., Weber, R. J.,
8 Zotter, P., and Parrish, D. D.: Gasoline emissions dominate over diesel in formation of secondary
9 organic aerosol mass, *Geophysical Research Letters*, 39, 10.1029/2011gl050718, 2012.
- 10 Baker, K. R., Misenis, C., Obland, M. D., Ferrare, R. A., Scarino, A. J., and Kelly, J. T.: Evaluation
11 of surface and upper air fine scale WRF meteorological modeling of the May and June 2010
12 CalNex period in California, *Atmospheric Environment*, 80, 299-309, 2013.
- 13 Barsanti, K., Carlton, A., and Chung, S.: Analyzing experimental data and model parameters:
14 implications for predictions of SOA using chemical transport models, *Atmospheric Chemistry and
15 Physics*, 13, 12073-12088, 2013.
- 16 Birch, M., and Cary, R.: Elemental carbon-based method for monitoring occupational exposures
17 to particulate diesel exhaust, *Aerosol Science and Technology*, 25, 221-241, 1996.
- 18 Borbon, A., Gilman, J., Kuster, W., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros,
19 V., Lopez, M., and Sarda-Esteve, R.: Emission ratios of anthropogenic volatile organic compounds
20 in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and
21 Paris, *Journal of Geophysical Research: Atmospheres*, 118, 2041-2057, 2013.
- 22 Carlton, A. G., Turpin, B. J., Altieri, K. E., Reff, A., Seitzinger, S., Lim, H. J., and Ervens, B.:
23 Atmospheric Oxalic Acid and SOA Production from Glyoxal: Results of Aqueous Photooxidation
24 Experiments, *Atmospheric Environment*, 41, 7588-7602, 2007.
- 25 Carlton, A. G., Turpin, B. J., Altieri, K. E., Seitzinger, S. P., Mathur, R., Roselle, S. J., and Weber,
26 R. J.: CMAQ model performance enhanced when in-cloud SOA is included: comparisons of OC
27 predictions with measurements, *Environ. Sci. Technol.*, 42, 8798-8802, 2008.
- 28 Carlton, A. G., Bhave, P. V., Napelenok, S. L., Edney, E. O., Sarwar, G., Pinder, R. W., Pouliot,
29 G. A., and Houyoux, M.: Treatment of secondary organic aerosol in CMAQv4.7, *Environmental
30 Science and Technology*, 44, 8553-8560, 2010.
- 31 Carlton, A. G., and Baker, K. R.: Photochemical Modeling of the Ozark Isoprene Volcano:
32 MEGAN, BEIS, and Their Impacts on Air Quality Predictions, *Environmental Science &
33 Technology*, 45, 4438-4445, 10.1021/es200050x, 2011.

- 1 Carter, W. P. L.: Implementation of the SAPRC-99 chemical mechanism into the models-3
2 framework, U.S. Environmental Protection Agency, University of California, Riverside, CA,
3 USA, 2000.
- 4 Chan, A. W., Isaacman, G., Wilson, K. R., Worton, D. R., Ruehl, C. R., Nah, T., Gentner, D. R.,
5 Dallmann, T. R., Kirchstetter, T. W., and Harley, R. A.: Detailed chemical characterization of
6 unresolved complex mixtures in atmospheric organics: Insights into emission sources, atmospheric
7 processing, and secondary organic aerosol formation, *Journal of Geophysical Research:*
8 *Atmospheres*, 118, 6783-6796, 2013.
- 9 Chan, A. W. H., Galloway, M. M., Kwan, A. J., Chhabra, P. S., Keutsch, F. N., Wennberg, P. O.,
10 Flagan, R. C., and Seinfeld, J. H.: Photooxidation of 2-Methyl-3-Buten-2-ol (MBO) as a Potential
11 Source of Secondary Organic Aerosol, *Environmental Science & Technology*, 43, 4647-4652,
12 10.1021/es802560w, 2009.
- 13 Chow, J. C., Watson, J. G., Crow, D., Lowenthal, D. H., and Merrifield, T.: Comparison of
14 IMPROVE and NIOSH carbon measurements, *Aerosol Science & Technology*, 34, 23-34, 2001.
- 15 Chung, S. H., and Seinfeld, J. H.: Global distribution and climate forcing of carbonaceous aerosols,
16 *Journal of Geophysical Research-Atmospheres*, 107, AAC 14.11-AAC 14-33,
17 10.1029/2001jd001397, 2002.
- 18 Docherty, K. S., Stone, E. A., Ulbrich, I. M., DeCarlo, P. F., Snyder, D. C., Schauer, J. J., Peltier,
19 R. E., Weber, R. J., Murphy, S. N., Seinfeld, J. H., Grover, B. D., Eatough, D. J., and Jimenez, J.
20 L.: Apportionment of Primary and Secondary Organic Aerosols in Southern California during the
21 2005 Study of Organic Aerosols in Riverside (SOAR-1), *Environmental Science & Technology*,
22 42, 7655-7662, 10.1021/es8008166, 2008.
- 23 Dusanter, S., Vimal, D., Stevens, P., Volkamer, R., and Molina, L.: Measurements of OH and HO
24 2 concentrations during the MCMA-2006 field campaign—Part 1: Deployment of the Indiana
25 University laser-induced fluorescence instrument, *Atmospheric Chemistry and Physics*, 9, 1665-
26 1685, 2009.
- 27 Dzepina, K., Volkamer, R. M., Madronich, S., Tulet, P., Ulbrich, I. M., Zhang, Q., Cappa, C. D.,
28 Ziemann, P. J., and Jimenez, J. L.: Evaluation of recently-proposed secondary organic aerosol
29 models for a case study in Mexico City, *Atmospheric Chemistry and Physics*, 9, 5681-5709, 2009.
- 30 Edney, E. O., Kleindienst, T. E., Conner, T. S., McIver, C. D., Corse, E. W., and Weathers, W. S.:
31 Polar organic oxygenates in PM_{2.5} at a southeastern site in the United States, *Atmospheric*
32 *Environment*, 37, 3947-3965, 2003.
- 33 Ensberg, J., Hayes, P., Jimenez, J., Gilman, J., Kuster, W., de Gouw, J., Holloway, J., Gordon, T.,
34 Jathar, S., and Robinson, A.: Emission factor ratios, SOA mass yields, and the impact of vehicular
35 emissions on SOA formation, *Atmospheric Chemistry and Physics*, 14, 2383-2397, 2014.
- 36 Ervens, B., Turpin, B., and Weber, R.: Secondary organic aerosol formation in cloud droplets and
37 aqueous particles (aqSOA): a review of laboratory, field and model studies, *Atmospheric*
38 *Chemistry and Physics*, 11, 11069-11102, 2011.

1 Fast, J., Allan, J., Bahreini, R., Craven, J., Emmons, L., Ferrare, R., Hayes, P., Hodzic, A.,
2 Holloway, J., and Hostetler, C.: Modeling regional aerosol and aerosol precursor variability over
3 California and its sensitivity to emissions and long-range transport during the 2010 CalNex and
4 CARES campaigns, *Atmospheric Chemistry and Physics*, 14, 10013-10060, 2014.

5 Fountoukis, C., and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic
6 equilibrium model for $K^+-Ca^{2+}-Mg^{2+}-NH_4^+-Na^+-SO_4^{2-}-NO_3^- -Cl^- -H_2O$ aerosols,
7 *Atmospheric Chemistry and Physics*, 7, 4639-4659, 2007.

8 Gantt, B., Meskhidze, N., and Carlton, A.: The contribution of marine organics to the air quality
9 of the western United States, *Atmospheric Chemistry and Physics*, 10, 7415-7423, 2010.

10 Gaston, C. J., Riedel, T. P., Zhang, Z., Gold, A., Surratt, J. D., and Thornton, J. A.: Reactive Uptake
11 of an Isoprene-derived Epoxydiol to Submicron Aerosol Particles, *Environmental science &*
12 *technology*, 48, 11178-11186, 10.1021/es5034266, 2014.

13 Gentner, D., Ford, T., Guha, A., Boulanger, K., Brioude, J., Angevine, W., de Gouw, J., Warneke,
14 C., Gilman, J., and Ryerson, T.: Emissions of organic carbon and methane from petroleum and
15 dairy operations in California's San Joaquin Valley, *Atmospheric Chemistry and Physics*, 14,
16 4955-4978, doi:10.5194/acp-14-4955-2014, 2014a.

17 Gentner, D., Ormeño, E., Fares, S., Ford, T., Weber, R., Park, J.-H., Brioude, J., Angevine, W.,
18 Karlik, J., and Goldstein, A.: Emissions of terpenoids, benzenoids, and other biogenic gas-phase
19 organic compounds from agricultural crops and their potential implications for air quality,
20 *Atmospheric Chemistry and Physics*, 14, 5393-5413, 2014b.

21 Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S.,
22 Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A., and Goldstein, A.
23 H.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed
24 characterization of organic carbon emissions, *Proceedings of the National Academy of Sciences*
25 *of the United States of America*, 109, 18318-18323, 10.1073/pnas.1212272109, 2012.

26 Gerbig, C., Schmitgen, S., Kley, D., Volz-Thomas, A., Dewey, K., and Haaks, D.: An improved
27 fast-response vacuum-UV resonance fluorescence CO instrument, *Journal of Geophysical*
28 *Research: Atmospheres* (1984–2012), 104, 1699-1704, 1999.

29 Geron, C.: Carbonaceous aerosol over a *Pinus taeda* forest in Central North Carolina,
30 USA, *Atmospheric Environment*, 43, 959-969, 2009.

31 Gilman, J., Burkhardt, J., Lerner, B., Williams, E., Kuster, W., Goldan, P., Murphy, P., Warneke,
32 C., Fowler, C., and Montzka, S.: Ozone variability and halogen oxidation within the Arctic and
33 sub-Arctic springtime boundary layer, *Atmospheric Chemistry and Physics Discussions*, 10,
34 15885-15919, 2010.

35 Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder,
36 B.: Fully coupled "online" chemistry within the WRF model, *Atmospheric Environment*, 39, 6957-
37 6975, 10.1016/j.atmosenv.2005.04.027, 2005.

1 Griffith, S., Hansen, R., Dusanter, S., Stevens, P., Alaghmand, M., Bertman, S., Carroll, M.,
2 Erickson, M., Galloway, M., and Grossberg, N.: OH and HO₂ radical chemistry during PROPHET
3 2008 and CABINEX 2009—Part 1: Measurements and model comparison, *Atmospheric Chemistry
4 and Physics*, 13, 5403-5423, 2013.

5 Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen,
6 J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T.,
7 Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W.,
8 McFiggans, G., Mentel, T. F., Monod, A., Prevot, A. S. H., Seinfeld, J. H., Surratt, J. D.,
9 Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol:
10 current and emerging issues, *Atmospheric Chemistry and Physics*, 9, 5155-5236, 2009.

11 Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W.,
12 Toohey, D. W., Flynn, J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglueck, B., Taylor,
13 J. W., Allan, J. D., Holloway, J. S., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Massoli, P.,
14 Zhang, X., Liu, J., Weber, R. J., Corrigan, A. L., Russell, L. M., Isaacman, G., Worton, D. R.,
15 Kreisberg, N. M., Goldstein, A. H., Thalman, R., Waxman, E. M., Volkamer, R., Lin, Y. H.,
16 Surratt, J. D., Kleindienst, T. E., Offenberg, J. H., Dusanter, S., Griffith, S., Stevens, P. S., Brioude,
17 J., Angevine, W. M., and Jimenez, J. L.: Organic aerosol composition and sources in Pasadena,
18 California, during the 2010 CalNex campaign, *Journal of Geophysical Research-Atmospheres*,
19 118, 9233-9257, 10.1002/jgrd.50530, 2013.

20 Hayes, P. L., Carlton, A. G., Baker, K. R., Ahmadov, R., Washenfelder, R. A., Alvarez, S.,
21 Rappenglueck, B., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Zotter, P., Prevot, A. S. H.,
22 Szidat, S., Kleindienst, T. E., Offenberg, J. H., and Jimenez, J. L.: Modeling the formation and
23 aging of secondary organic aerosols in Los Angeles during CalNex 2010, *Atmospheric Chemistry
24 and Physics Discussions*, 2014.

25 Henderson, B., Akhtar, F., Pye, H., Napelenok, S., and Hutzell, W.: A database and tool for
26 boundary conditions for regional air quality modeling: description and evaluation, *Geoscientific
27 Model Development*, 7, 339-360, 2014.

28 Henze, D. K., and Seinfeld, J. H.: Global secondary organic aerosol from isoprene oxidation,
29 *Geophysical Research Letters*, 33, L09812, 10.1029/2006gl025976, 2006.

30 Heo, J., de Foy, B., Olson, M. R., Pakbin, P., Sioutas, C., and Schauer, J. J.: Impact of regional
31 transport on the anthropogenic and biogenic secondary organic aerosols in the Los Angeles Basin,
32 *Atmospheric Environment*, 103, 171-179, 2015.

33 Hersey, S. P., Craven, J. S., Schilling, K. A., Metcalf, A. R., Sorooshian, A., Chan, M. N., Flagan,
34 R. C., and Seinfeld, J. H.: The Pasadena Aerosol Characterization Observatory (PACO): chemical
35 and physical analysis of the Western Los Angeles basin aerosol, *Atmospheric Chemistry and
36 Physics*, 11, 7417-7443, 10.5194/acp-11-7417-2011, 2011.

37 Hutzell, W. T., Luecken, D. J., Appel, K. W., and Carter, W. P. L.: Interpreting predictions from
38 the SAPRC07 mechanism based on regional and continental simulations, *Atmospheric
39 Environment*, 46, 417-429, 10.1016/j.atmosenv.2011.09.030, 2012.

- 1 Jathar, S. H., Gordon, T. D., Hennigan, C. J., Pye, H. O. T., Pouliot, G., Adams, P. J., Donahue,
2 N. M., and Robinson, A. L.: Unspeciated organic emissions from combustion sources and their
3 influence on the secondary organic aerosol budget in the United States, *Proceedings of the National
4 Academy of Sciences*, 111, 10473-10478, 10.1073/pnas.1323740111, 2014.
- 5 Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H.,
6 DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M.,
7 Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin,
8 C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M.,
9 Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A.,
10 Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F.,
11 Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A.,
12 Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R.,
13 Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C.,
14 Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of Organic
15 Aerosols in the Atmosphere, *Science*, 326, 1525-1529, 10.1126/science.1180353, 2009.
- 16 Karambelas, A., Pye, H. O. T., Budisulistiorini, S. H., Surratt, J. D., and Pinder, R. W.: Isoprene
17 epoxydiol contribution to urban organic aerosol: evidence from modeling and measurements,
18 *Environmental Science & Technology Letters*, 1, 278-283, 10.1021/ez5001353, 2013.
- 19 Kelly, J. T., Baker, K. R., Nowak, J. B., Murphy, J. G., Markovic, M. Z., VandenBoer, T. C., Ellis,
20 R. A., Neuman, J. A., Weber, R. J., and Roberts, J. M.: Fine-scale simulation of ammonium and
21 nitrate over the South Coast Air Basin and San Joaquin Valley of California during CalNex-2010,
22 *Journal of Geophysical Research: Atmospheres*, 119, 3600-3614, 10.1002/2013JD021290, 2014.
- 23 Kleindienst, T. E., Conver, T. S., McIver, C. D., and Edney, E. O.: Determination of secondary
24 organic aerosol products from the photooxidation of toluene and their implications in ambient
25 PM_{2.5}, *Journal of Atmospheric Chemistry*, 47, 79-100, 2004.
- 26 Kleindienst, T. E., Jaoui, M., Lewandowski, M., Offenber, J. H., Lewis, C. W., Bhave, P. V., and
27 Edney, E. O.: Estimates of the contributions of biogenic and anthropogenic hydrocarbons to
28 secondary organic aerosol at a southeastern US location, *Atmospheric Environment*, 41, 8288-
29 8300, 10.1016/j.atmosenv.2007.06.045, 2007.
- 30 Kleindienst, T. E., Lewandowski, M., Offenber, J. H., Edney, E. O., Jaoui, M., Zheng, M., Ding,
31 X. A., and Edgerton, E. S.: Contribution of Primary and Secondary Sources to Organic Aerosol
32 and PM_{2.5} at SEARCH Network Sites, *Journal of the Air & Waste Management Association*, 60,
33 1388-1399, 10.3155/1047-3289.60.11.1388, 2010.
- 34 Lewandowski, M., Piletic, I. R., Kleindienst, T. E., Offenber, J. H., Beaver, M. R., Jaoui, M.,
35 Docherty, K. S., and Edney, E. O.: Secondary organic aerosol characterisation at field sites across
36 the United States during the spring–summer period, *International Journal of Environmental
37 Analytical Chemistry*, 93, 1084-1103, 2013.
- 38 Lewis, C. W., Klouda, G. A., and Ellenson, W. D.: Radiocarbon measurement of the biogenic
39 contribution to summertime PM-2.5 ambient aerosol in Nashville, TN, *Atmospheric Environment*,
40 38, 6053-6061, 10.1016/j.atmosenv.2004.06.011, 2004.

1 Liu, S., Ahlm, L., Day, D. A., Russell, L. M., Zhao, Y., Gentner, D. R., Weber, R. J., Goldstein,
2 A. H., Jaoui, M., Offenberg, J. H., Kleindienst, T. E., Rubitschun, C., Surratt, J. D., Sheesley, R.
3 J., and Scheller, S.: Secondary organic aerosol formation from fossil fuel sources contribute
4 majority of summertime organic mass at Bakersfield, *Journal of Geophysical Research-*
5 *Atmospheres*, 117, 10.1029/2012JD018170, 2012.

6 Mao, J., Ren, X., Zhang, L., Van Duin, D., Cohen, R., Park, J.-H., Goldstein, A., Paulot, F., Beaver,
7 M., and Crouse, J.: Insights into hydroxyl measurements and atmospheric oxidation in a
8 California forest, *Atmospheric Chemistry and Physics*, 12, 8009-8020, 2012.

9 Markovic, M., VandenBoer, T., Baker, K., Kelly, J., and Murphy, J.: Measurements and modeling
10 of the inorganic chemical composition of fine particulate matter and associated precursor gases in
11 California's San Joaquin Valley during CalNex 2010, *Journal of Geophysical Research:*
12 *Atmospheres*, 119, 6853-6866, 10.1002/2013JD021408, 2014.

13 Offenberg, J. H., Lewandowski, M., Edney, E. O., Kleindienst, T. E., and Jaoui, M.: Influence of
14 Aerosol Acidity on the Formation of Secondary Organic Aerosol from Biogenic Precursor
15 Hydrocarbons, *Environmental Science & Technology*, 43, 7742-7747, 10.1021/es901538e, 2009.

16 Ormeño, E., Gentner, D. R., Fares, S., Karlik, J., Park, J. H., and Goldstein, A. H.: Sesquiterpenoid
17 emissions from agricultural crops: correlations to monoterpenoid emissions and leaf terpene
18 content, *Environmental science & technology*, 44, 3758-3764, 2010.

19 Pye, H. O., Pinder, R. W., Piletic, I. R., Xie, Y., Capps, S. L., Lin, Y.-H., Surratt, J. D., Zhang, Z.,
20 Gold, A., and Luecken, D. J.: Epoxide pathways improve model predictions of isoprene markers
21 and reveal key role of acidity in aerosol formation, *Environmental science & technology*, 47,
22 11056-11064, 2013.

23 Pye, H. O. T., and Pouliot, G. A.: Modeling the Role of Alkanes, Polycyclic Aromatic
24 Hydrocarbons, and Their Oligomers in Secondary Organic Aerosol Formation, *Environmental*
25 *Science & Technology*, 46, 6041-6047, 10.1021/es300409w, 2012.

26 Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop,
27 A. P., Lane, T. E., Pierce, J. R., and Pandis, S. N.: Rethinking organic aerosols: Semivolatile
28 emissions and photochemical aging, *Science*, 315, 1259-1262, 2007.

29 Rollins, A., Browne, E., Min, K.-E., Pusede, S., Wooldridge, P., Gentner, D., Goldstein, A., Liu,
30 S., Day, D., and Russell, L.: Evidence for NO_x control over nighttime SOA formation, *Science*,
31 337, 1210-1212, 2012.

32 Ryerson, T. B., Andrews, A. E., Angevine, W. M., Bates, T. S., Brock, C. A., Cairns, B., Cohen,
33 R. C., Cooper, O. R., de Gouw, J. A., Fehsenfeld, F. C., Ferrare, R. A., Fischer, M. L., Flagan, R.
34 C., Goldstein, A. H., Hair, J. W., Hardesty, R. M., Hostetler, C. A., Jimenez, J. L., Langford, A.
35 O., McCauley, E., McKeen, S. A., Molina, L. T., Nenes, A., Oltmans, S. J., Parrish, D. D.,
36 Pederson, J. R., Pierce, R. B., Prather, K., Quinn, P. K., Seinfeld, J. H., Senff, C. J., Sorooshian,
37 A., Stutz, J., Surratt, J. D., Trainer, M., Volkamer, R., Williams, E. J., and Wofsy, S. C.: The 2010
38 California Research at the Nexus of Air Quality and Climate Change (CalNex) field study, *Journal*
39 *of Geophysical Research-Atmospheres*, 118, 5830-5866, 10.1002/jgrd.50331, 2013.

1 Sarwar, G., Fahey, K., Kwok, R., Gilliam, R. C., Roselle, S. J., Mathur, R., Xue, J., Yu, J., and
2 Carter, W. P. L.: Potential impacts of two SO₂ oxidation pathways on regional sulfate
3 concentrations: Aqueous-phase oxidation by NO₂ and gas-phase oxidation by Stabilized Criegee
4 Intermediates, *Atmospheric Environment*, 68, 186-197, 10.1016/j.atmosenv.2012.11.036, 2013.

5 Shilling, J. E., Zaveri, R. A., Fast, J. D., Kleinman, L., Alexander, M., Canagaratna, M. R., Fortner,
6 E., Hubbe, J. M., Jayne, J. T., and Sedlacek, A.: Enhanced SOA formation from mixed
7 anthropogenic and biogenic emissions during the CARES campaign, *Atmospheric Chemistry and
8 Physics*, 13, 2091-2113, 2013.

9 Simon, H., and Bhave, P. V.: Simulating the Degree of Oxidation in Atmospheric Organic
10 Particles, *Environmental Science & Technology*, 46, 331-339, 10.1021/es202361w, 2012.

11 Steiner, A. L., Cohen, R., Harley, R., Tonse, S., Millet, D., Schade, G., and Goldstein, A.: VOC
12 reactivity in central California: comparing an air quality model to ground-based measurements,
13 *Atmospheric Chemistry and Physics*, 8, 351-368, 2008.

14 Stroud, C. A., Liggio, J., Zhang, J., Gordon, M., Staebler, R. M., Makar, P. A., Zhang, J., Li, S.
15 M., Mihele, C., and Lu, G.: Rapid organic aerosol formation downwind of a highway: Measured
16 and model results from the FEVER study, *Journal of Geophysical Research: Atmospheres*, 119,
17 1663-1679, 2014.

18 Stuiver, M.: International agreements and the use of the new oxalic acid standard, *Radiocarbon*,
19 25, 793-795, 1983.

20 Tan, Y., Carlton, A. G., Seitzinger, S. P., and Turpin, B. J.: SOA from methylglyoxal in clouds
21 and wet aerosols: measurement and prediction of key products, *Atmospheric Environment*, 44,
22 5218-5226, 10.1016/j.atmosenv.2010.08.045, 2010.

23 http://www.epa.gov/ttn/chief/net/2011nei/2011_neiv1_tsd_draft.pdf. Accessed October 1, 2014.:
24 ftp://ftp.epa.gov/EmisInventory/2001v2CAP/2001emis/readme_2001emis.txt, access: October 2,
25 2014, 2014.

26 Wagstrom, K. M., Baker, K. R., Leinbach, A. E., and Hunt, S. W.: Synthesizing Scientific
27 Progress: Outcomes from US EPA's Carbonaceous Aerosols and Source Apportionment STAR
28 Grants, *Environmental science & technology*, 48, 10561-10570, 2014.

29 Washenfelder, R. A., Young, C. J., Brown, S. S., Angevine, W. M., Atlas, E. L., Blake, D. R., Bon,
30 D. M., Cubison, M. J., de Gouw, J. A., Dusanter, S., Flynn, J., Gilman, J. B., Graus, M., Griffith,
31 S., Grossberg, N., Hayes, P. L., Jimenez, J. L., Kuster, W. C., Lefer, B. L., Pollack, I. B., Ryerson,
32 T. B., Stark, H., Stevens, P. S., and Trainer, M. K.: The glyoxal budget and its contribution to
33 organic aerosol for Los Angeles, California, during CalNex 2010, *Journal of Geophysical
34 Research-Atmospheres*, 116, 10.1029/2011jd016314, 2011.

35 Yarwood, G., Rao, S., Yocke, M., and Whitten, G. Z.: Updates to the carbon bond chemical
36 mechanism: CB05, ENVIRON International Corporation, Novato, CA, 2005.

37 Zhang, H., Chen, G., Hu, J., Chen, S.-H., Wiedinmyer, C., Kleman, M., and Ying, Q.: Evaluation
38 of a seven-year air quality simulation using the Weather Research and Forecasting

- 1 (WRF)/Community Multiscale Air Quality (CMAQ) models in the eastern United States, *Science*
2 of The Total Environment, 473, 275-285, 2014a.
- 3 Zhang, X., Cappa, C. D., Jathar, S. H., McVay, R. C., Ensberg, J. J., Kleeman, M. J., and Seinfeld,
4 J. H.: Influence of vapor wall loss in laboratory chambers on yields of secondary organic aerosol,
5 *Proceedings of the National Academy of Sciences*, 111, 5802-5807, 2014b.
- 6 Zhao, Y., Kreisberg, N. M., Worton, D. R., Isaacman, G., Gentner, D. R., Chan, A. W., Weber, R.
7 J., Liu, S., Day, D. A., and Russell, L. M.: Sources of organic aerosol investigated using organic
8 compounds as tracers measured during CalNex in Bakersfield, *Journal of Geophysical Research:*
9 *Atmospheres*, 118, 11388-11398, 2013.
- 10 Zhao, Y., Hennigan, C. J., May, A. A., Tkacik, D. S., de Gouw, J. A., Gilman, J. B., Kuster, W.
11 C., Borbon, A., and Robinson, A. L.: Intermediate-Volatility Organic Compounds: A Large Source
12 of Secondary Organic Aerosol, *Environmental science & technology*, 48, 13743-13750, 2014.
- 13 Zotter, P., El-Haddad, I., Zhang, Y., Hayes, P. L., Zhang, X., Lin, Y. H., Wacker, L., Schnelle-
14 Kreis, J., Abbaszade, G., and Zimmermann, R.: Diurnal cycle of fossil and nonfossil carbon using
15 radiocarbon analyses during CalNex, *Journal of Geophysical Research: Atmospheres*, 119, 6818-
16 6835, 10.1002/2013JD021114, 2014.

17

18

1 Table 1. Episode total anthropogenic emissions of primarily emitted PM2.5 organic carbon and
 2 the sum of benzene, toluene, and xylenes by emissions sector group. The Los Angeles (LA) total
 3 includes Los Angeles and Orange counties. The southern San Joaquin Valley (SSJV) total
 4 includes Kern, Fresno, Kings, and Tulare counties. Residential wood combustion, fugitives, and
 5 non-point area PM2.5 emissions are largely contemporary in origin.

Sector	Primarily emitted PM2.5 organic carbon				Benzene + Toluene + Xylenes			
	SSJV (tons)	SSJV (%)	LA (tons)	LA (%)	SSJV (tons)	SSJV (%)	LA (tons)	LA (%)
Non-point area	139.9	33.8	410.1	40.8	326.7	37.2	1229.3	35.8
Onroad mobile	73.3	17.7	263.6	26.2	273.5	31.2	1190.9	34.6
Nonroad mobile	23.9	5.8	161.4	16.1	170.1	19.4	822.3	23.9
Point: non-electrical generating	61.3	14.8	56.3	5.6	68.3	7.8	177.7	5.2
Residential wood combustion	54.1	13.1	82.7	8.2	2.0	0.2	3.2	0.1
Oil & gas exploration and related	28.5	6.9	0.0	0.0	34.2	3.9	1.1	0.0
Fugitive dust	24.9	6.0	18.1	1.8	0.0	0.0	0.0	0.0
Commercial marine & rail	3.8	0.9	11.4	1.1	2.6	0.3	12.8	0.4
Point: electrical generating	4.3	1.0	1.7	0.2	0.1	0.0	1.0	0.0
Total Modern Carbon	218.9	52.9	510.9	50.8	2.0	0.2	3.2	0.1
Total Fossil Carbon	195.2	47.1	494.5	49.2	875.3	99.8	3435.1	99.9

6

7

1 Table 2. Episode average measured and modeled PM2.5 carbon, PM2.5 SOC groups, and VOC
 2 at the Pasadena and Bakersfield sites.

Specie	Model Run	Location	N	Observed (ugC/m3)	Predicted (ugC/m3)	Bias (ugC/m3)	Error (ugC/m3)	Fractional Bias (%)	Fractional Error (%)	r
Elemental Carbon	Baseline	Bakersfield	35	0.5	0.4	-0.1	0.1	-13	35	0.17
	Baseline	Pasadena	31	0.2	1.0	0.8	0.8	125	125	0.70
	Baseline	3STN/IMPROVE sites	220	0.2	0.6	0.6	0.6	77	87	0.47
Organic Carbon	Baseline	Bakersfield	35	5.4	0.8	-4.6	4.6	-144	144	0.11
	Baseline	Pasadena	31	3.6	2.0	-1.6	1.6	-53	53	0.73
	Baseline	3STN/IMPROVE sites	220	1.9	1.3	-0.6	0.9	-34	53	0.06
	Sensitivity	3STN/IMPROVE sites	220	1.9	1.7	-0.2	0.8	-11	42	0.32

Specie	Model Run	Location	N	Observed (ngC/m3)	Predicted (ngC/m3)	Bias (ngC/m3)	Error (ngC/m3)	Fractional Bias (%)	Fractional Error (%)	r
Isoprene SOC	Baseline	Bakersfield	36	96	21	-75	75	-126	128	0.36
		Pasadena	32	42	27	-15	25	-60	83	0.10
Monoterpene SOC	Baseline	Bakersfield	35	56	21	-35	37	-75	89	0.66
		Pasadena	32	82	21	-60	61	-89	93	0.55
Toluene+Xylene SOC	Baseline	Bakersfield	35	59	15	-44	44	-114	114	0.62
		Pasadena	32	125	36	-89	89	-100	100	0.82
Sesquiterpene SOC	Baseline	Bakersfield	41		17					
		Pasadena	41		7					
Benzene SOC	Baseline	Bakersfield	41		2					
		Pasadena	41		2					
Alkane SOC	Baseline	Bakersfield	41		12					
		Pasadena	41		22					
Cloud SOC	Baseline	Bakersfield	41		1					
		Pasadena	41		5					
Naphthalene SOC	Baseline	Bakersfield	36	43						
		Pasadena	32	114						

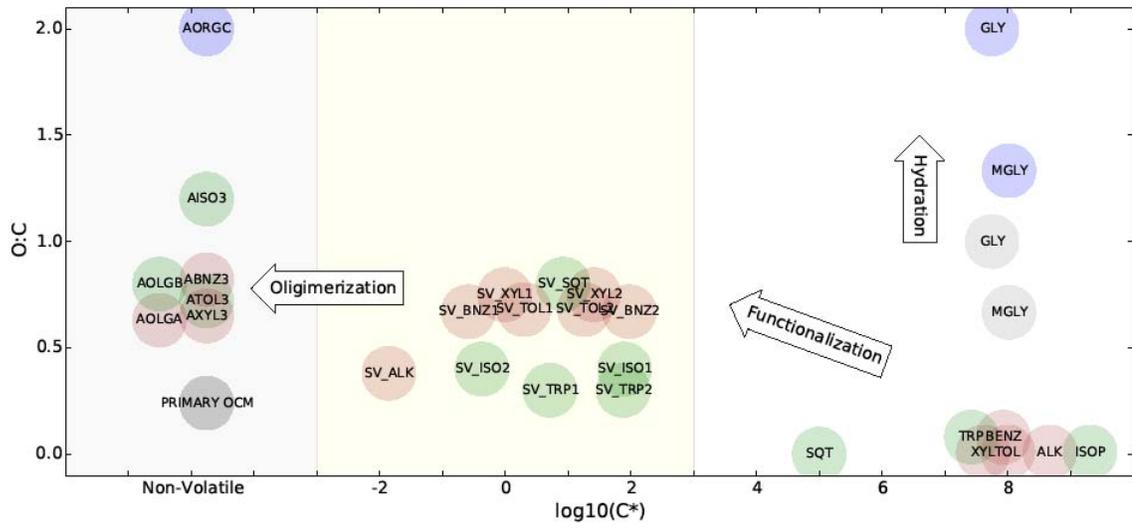
Specie	Model Run	Location	N	Observed (ppbC)	Predicted (ppbC)	Bias (ppbC)	Error (ppbC)	Fractional Bias (%)	Fractional Error (%)	r
Isoprene VOC 3-hr	Baseline	Bakersfield	5	0.1	0.3	0.2	0.2	79	79	0.79
		Pasadena	8	0.6	0.5	-0.2	0.5	0	84	-0.21
Monoterpene VOC 3-hr	Baseline	Bakersfield	37	1.4	0.5	-0.9	1.0	-72	89	0.25
		Pasadena	28	1.8	0.3	-1.5	1.6	-129	137	0.15
Toluene VOC 3-hr	Baseline	Bakersfield	41	4.3	2.7	-1.6	1.9	-48	55	0.44
		Pasadena	29	7.3	7.7	0.4	3.5	17	44	0.24
Xylene VOC 3-hr	Baseline	Bakersfield	41	4.3	1.8	-2.5	2.5	-82	83	0.34
		Pasadena	29	6.7	4.5	-2.1	2.6	-33	41	0.20
Benzene VOC 3-hr	Baseline	Bakersfield	41	1.2	1.3	0.2	0.5	6	38	0.14
		Pasadena	29	1.5	1.6	0.1	0.5	0	30	0.16
Total VOC 3-hr	Baseline	Bakersfield	41	186.9	63.7	-123.2	124.2	-95	97	0.37
		Pasadena	29	188.9	88.7	-100.1	100.1	-66	66	0.26
Isoprene VOC 1-hr	Baseline	Bakersfield	712	0.4	0.4	0.0	0.3	-21	83	0.15
		Pasadena	718	1.6	0.8	-0.9	1.7	-32	139	-0.10
Monoterpene VOC 1-hr	Baseline	Bakersfield	605	0.8	0.3	-0.6	0.7	-63	101	0.25
		Pasadena	707	0.7	0.2	-0.5	0.5	-105	111	0.05
Toluene VOC 1-hr	Baseline	Bakersfield	737	2.5	1.7	-0.8	1.5	-25	56	0.31
		Pasadena	717	4.0	6.1	2.0	2.8	36	54	0.23
Xylene VOC 1-hr	Baseline	Bakersfield	737	1.9	1.1	-0.7	1.2	-37	64	0.32
		Pasadena	718	3.2	3.4	0.2	1.7	2	51	0.15

3

4

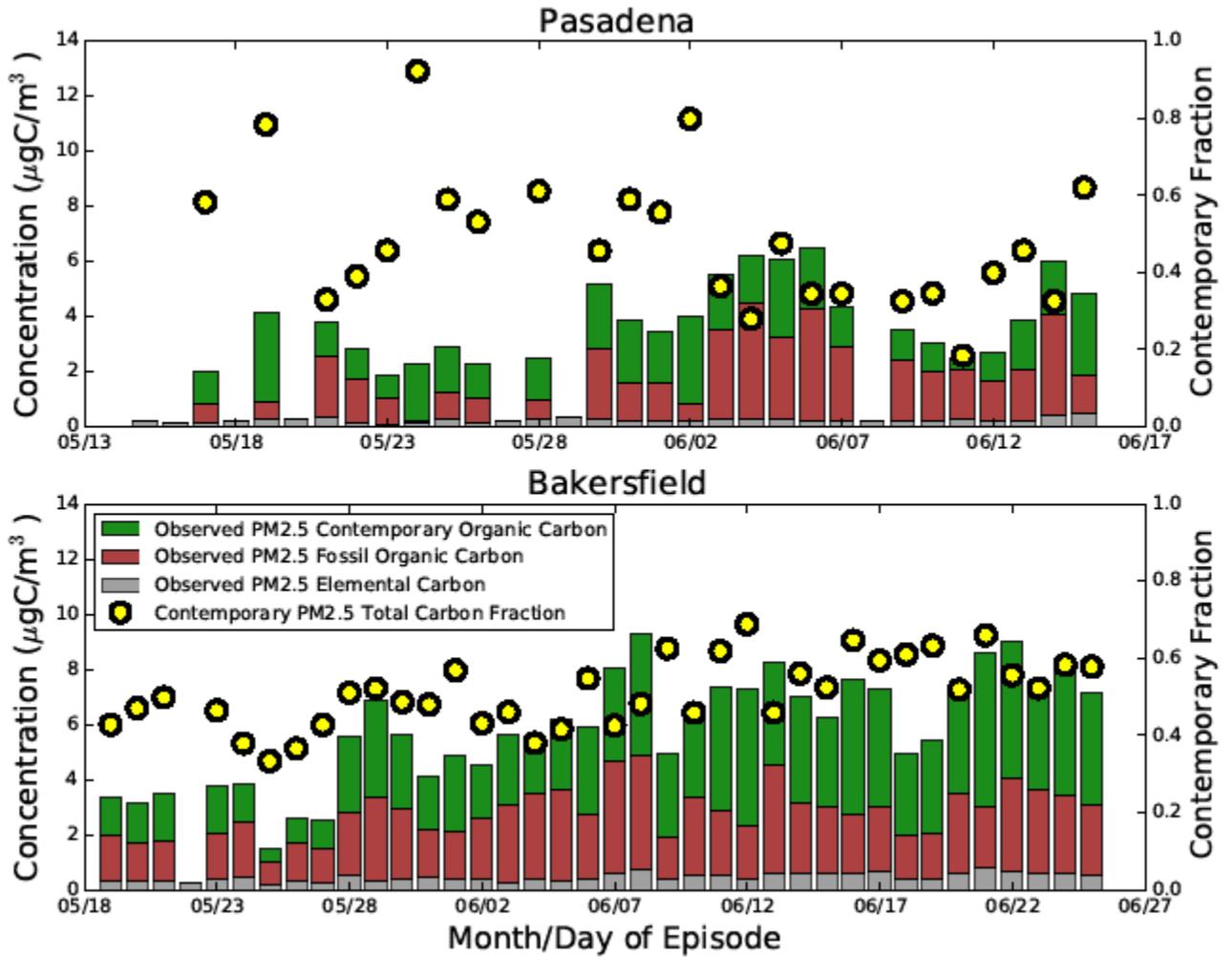
5

1 Figure 1. Gas (right panel), semi-volatile (middle panel), and particle phase (left panel) CMAQ
 2 organic carbon shown by saturation vapor pressure and O:C ratio. Compounds shown in blue
 3 exist in the aqueous phase, brown suggest generally fossil in origin, green generally
 4 contemporary in origin, and gray both contemporary and fossil in origin. Other known processes
 5 such as fragmentation are not shown as they are not currently represented in the modeling
 6 system.



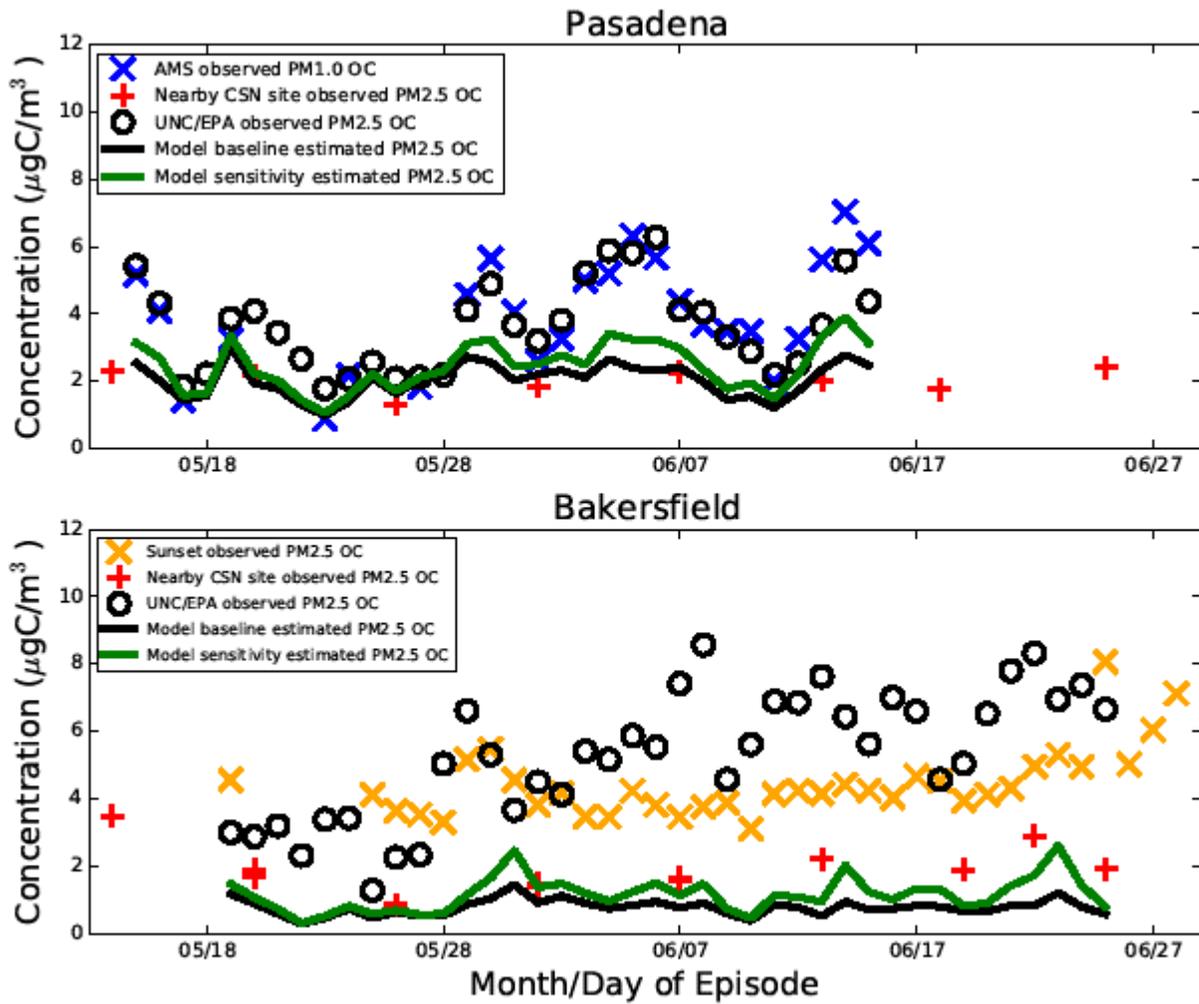
7

- 1 Figure 2. Observed daily 23-h average total carbon and contemporary total carbon fraction at
- 2 Pasadena and Bakersfield.



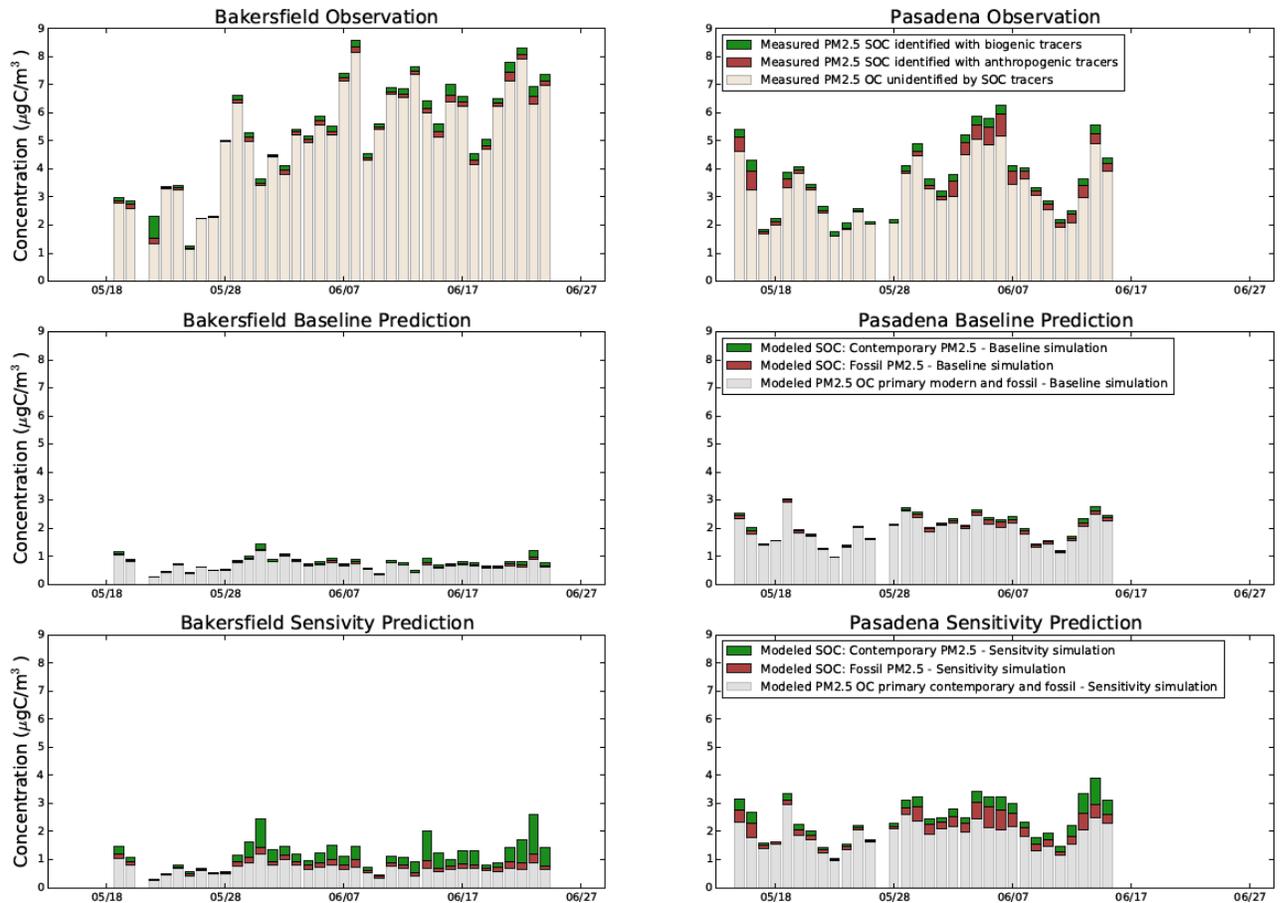
3

- 1 Figure 3. Model predicted and measured PM_{2.5} organic carbon at Pasadena and Bakersfield. The
- 2 nearby CSN measurements are intended to provide additional context and are not co-located with
- 3 CalNex measurements or model estimates.



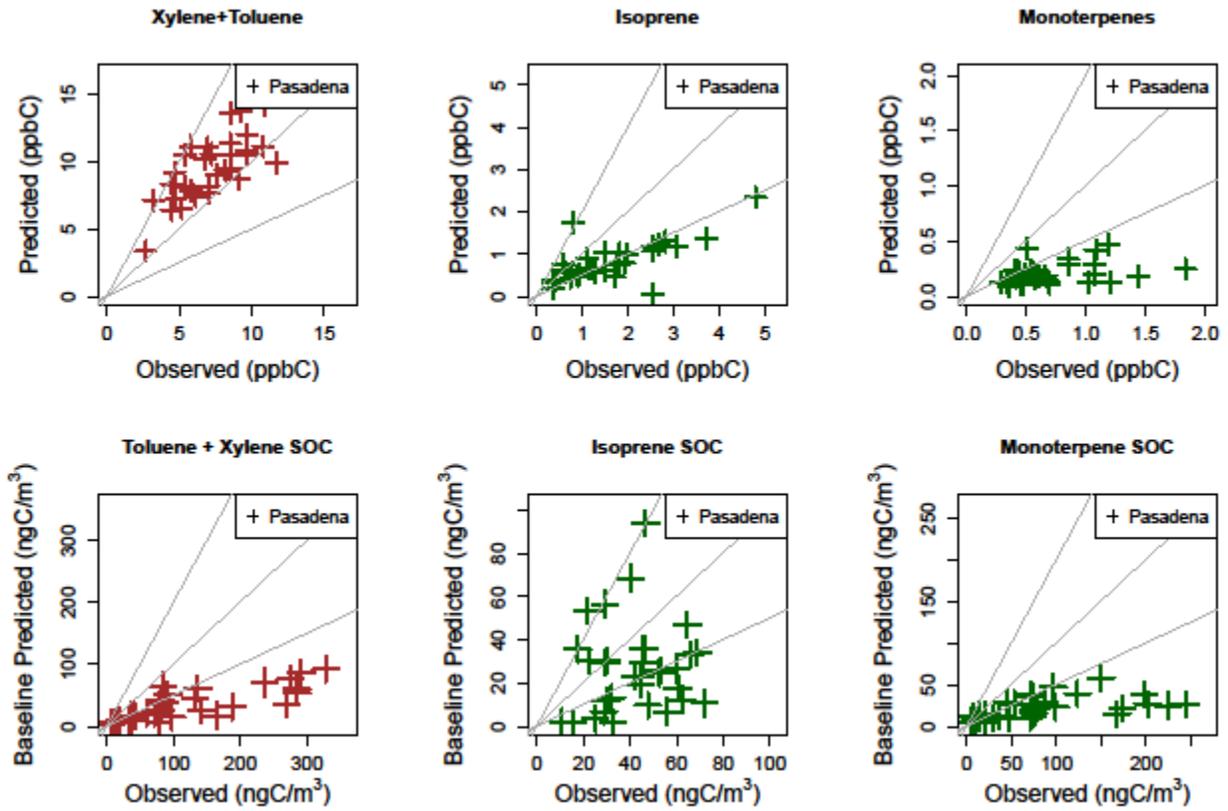
- 4
- 5
- 6

- 1 Figure 4. Observed (top row) and modeled (middle and bottom rows) PM_{2.5} organic carbon at
- 2 Pasadena and Bakersfield. Mass explained by SOA tracers shown in green (contemporary origin
- 3 tracers) and brown (fossil origin tracers). Gray shading indicates mass not explained by known
- 4 SOC tracers for observations and primarily emitted PM_{2.5} (both contemporary and fossil) for
- 5 modeled estimates. Middle row shows baseline model estimates and bottom row model
- 6 sensitivity results with increased SOA yields.



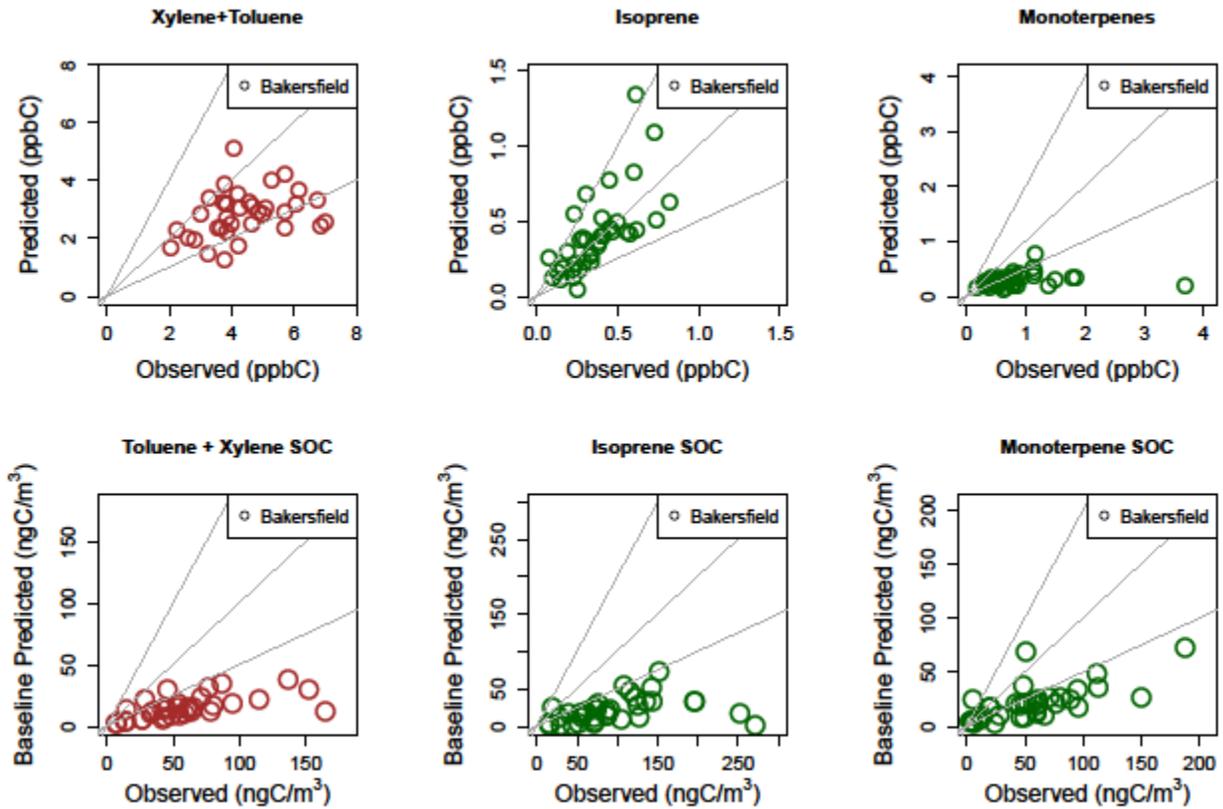
7

- 1 Figure 5. Comparison of CMAQ-predicted and measured VOC (daily average of hourly samples)
- 2 and corresponding SOC species (daily 23-hr average samples) for Pasadena. Comparison points
- 3 outside the gray lines indicate model predictions are greater than a factor of 2 different from the
- 4 measurements.



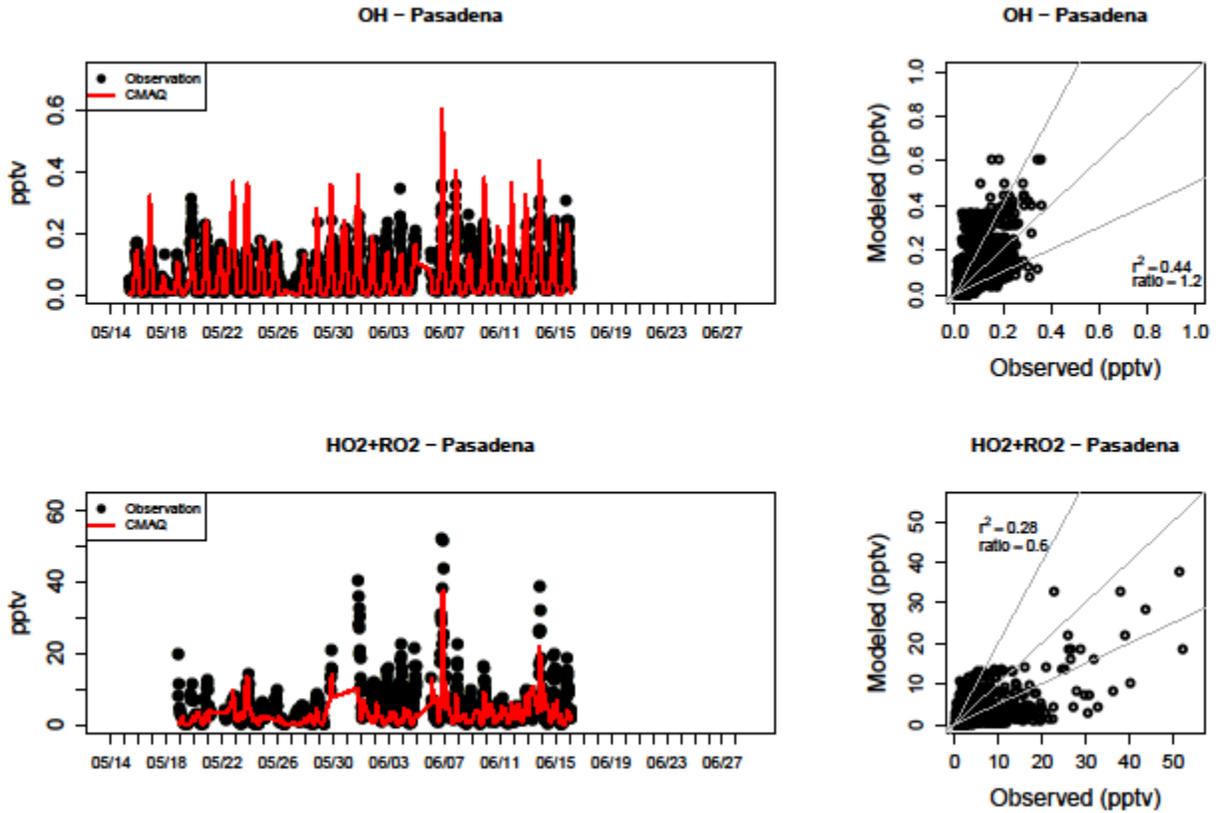
5

- 1 Figure 6. Comparison of CMAQ-predicted and measured VOC (daily average of hourly samples)
- 2 and corresponding SOC species (daily 23-hr average samples) for Bakersfield. Comparison
- 3 points outside the gray lines indicate model predictions are greater than a factor of 2 different
- 4 from the measurements.



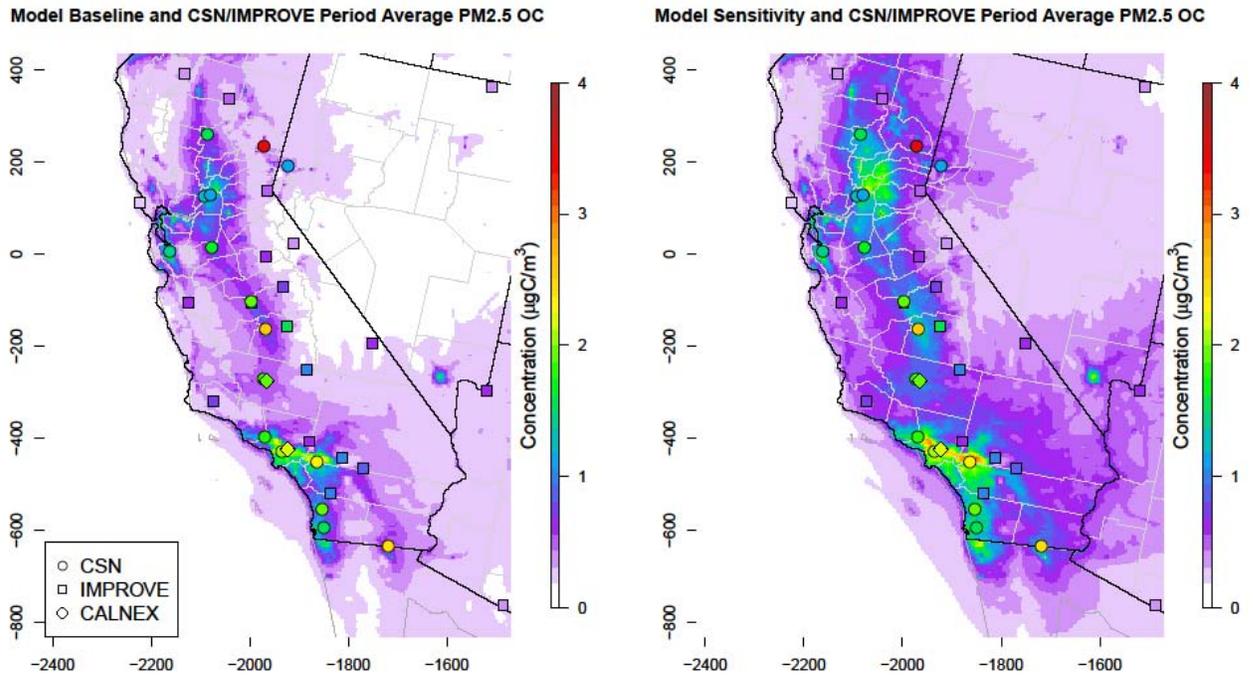
- 5
- 6
- 7

- 1 Figure 7. Measured and model estimated OH radical (top) and HO₂+RO₂ (bottom) at Pasadena.
- 2 The ratio shown on the scatterplots is the episode average model estimates divided by the episode average measured values.
- 3



4

- 1 Figure 8. May-June 2010 average observed and modeled PM_{2.5} organic carbon. Measurements
- 2 from both CalNex locations and routine networks including CSN (circles) and IMPROVE
- 3 (squares). Left panel shows baseline model predictions and right panel shows model estimates
- 4 with increased SOA yields.



- 5
- 6
- 7

Supporting Information

Measured and Model Estimated Gas and Aerosol Carbon in California: Pasadena and Bakersfield

K.R. Baker, A.G. Carlton, T.E. Kleindienst, J.H. Offenberg, M.R. Beaver, D. R. Gentner, A. H. Goldstein, P. L. Hayes, J. L. Jimenez, J. B. Gilman, J. A. de Gouw, M. C. Woody, H.O.T. Pye, J.T. Kelly, M. Lewandowski, M. Jaoui, P.S. Stevens, W.H. Brune, Y.-H. Lin, C.L. Rubitschun, J. D. Surratt

Organization of the Supporting Information

- Additional Tables (pages 2-5)
- Additional Figures (pages 6-20)
- Sampling Analysis and Methods (pages 21-22)
- CMAQ AE6 treatment for SOA presented by volatility and O:C ratio (pages 23-27)

Emission Inventory

Table S1a. Episode total emissions (tons) of primarily emitted PM2.5 organic carbon by area delineated by source classification code (SCC). The Los Angeles (LA) total includes Los Angeles and Orange counties. The southern San Joaquin Valley (SSJV) total includes Kern, Fresno, Kings, and Tulare counties. Only categories above 1% contribution shown.

Area	Tons	Percent	SCC7	SCC7 Description
SSJV	72.0	17.4	2302002	Industrial Processes;Food and Kindred Products: SIC 20;Commercial Cooking - Charbroiling
SSJV	54.1	13.1	2104008	Stationary Source Fuel Combustion;Residential;Wood
SSJV	49.3	11.9	2801500	Miscellaneous Area Sources;Agriculture Production - Crops;Agricultural Field Burning - whole field set on fire
SSJV	28.4	6.9	310004	Industrial Processes;Oil and Gas Production;Process Heaters
SSJV	25.4	6.1	2230074	Mobile Sources;Highway Vehicles - Diesel;Heavy Duty Diesel Vehicles (HDDV) Class 8A & 8B
SSJV	14.7	3.5	2201001	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Vehicles (LDGV)
SSJV	13.5	3.3	2294000	Mobile Sources;Paved Roads;All Paved Roads
SSJV	12.9	3.1	2282005	Mobile Sources;Pleasure Craft;Gasoline 2-Stroke
SSJV	12.9	3.1	101009	External Combustion Boilers;Electric Generation;Wood/Bark Waste
SSJV	11.3	2.7	202002	Internal Combustion Engines;Industrial;Natural Gas
SSJV	10.5	2.5	2201020	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 1 & 2 (M6) = LDGT1 (M5)
SSJV	8.9	2.1	305006	Industrial Processes;Mineral Products;Cement Manufacturing (Dry Process)
SSJV	8.8	2.1	2230073	Mobile Sources;Highway Vehicles - Diesel;Heavy Duty Diesel Vehicles (HDDV) Class 6 & 7
SSJV	6.2	1.5	2265003	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Industrial Equipment
SSJV	6.0	1.4	2610000	Waste Disposal, Treatment, and Recovery;Open Burning;All Categories
SSJV	5.4	1.3	2296000	Mobile Sources;Unpaved Roads;All Unpaved Roads
SSJV	5.4	1.3	2201040	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 3 & 4 (M6) = LDGT2 (M5)
LA	269.4	26.8	2302002	Industrial Processes;Food and Kindred Products: SIC 20;Commercial Cooking - Charbroiling
LA	95.9	9.5	2282005	Mobile Sources;Pleasure Craft;Gasoline 2-Stroke
LA	94.0	9.4	2201001	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Vehicles (LDGV)
LA	82.7	8.2	2104008	Stationary Source Fuel Combustion;Residential;Wood
LA	66.6	6.6	2307000	Industrial Processes;Wood Products: SIC 24;All Processes
LA	58.0	5.8	2201020	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 1 & 2 (M6) = LDGT1 (M5)
LA	38.4	3.8	2230074	Mobile Sources;Highway Vehicles - Diesel;Heavy Duty Diesel Vehicles (HDDV) Class 8A & 8B
LA	29.9	3.0	2201040	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 3 & 4 (M6) = LDGT2 (M5)
LA	23.8	2.4	2104006	Stationary Source Fuel Combustion;Residential;Natural Gas
LA	17.5	1.7	2103008	Stationary Source Fuel Combustion;Commercial/Institutional;Wood
LA	15.0	1.5	2230073	Mobile Sources;Highway Vehicles - Diesel;Heavy Duty Diesel Vehicles (HDDV) Class 6 & 7
LA	13.6	1.4	2294000	Mobile Sources;Paved Roads;All Paved Roads
LA	12.0	1.2	2302080	Industrial Processes;Food and Kindred Products: SIC 20;Miscellaneous Food and Kindred Products
LA	10.8	1.1	2265004	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Lawn and Garden Equipment
LA	10.7	1.1	2265006	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Commercial Equipment

Table S1b. Episode total emissions (tons) of the sum of benzene, toluene, and xylenes by area delineated by source classification code (SCC). The Los Angeles (LA) total includes Los Angeles and Orange counties. The southern San Joaquin Valley (SSJV) total includes Kern, Fresno, Kings, and Tulare counties. Only categories above 1% contribution shown.

Area	Tons	Percent	SCC7	SCC7 Description
SSJV	112.3	12.8	2201001	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Vehicles (LDGV)
SSJV	81.1	9.2	2201020	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 1 & 2 (M6) = LDGT1 (M5)
SSJV	71.3	8.1	2680001	Waste Disposal, Treatment, and Recovery;Composting;100% Biosolids (e.g., sewage sludge, manure, mixtures of these matls)
SSJV	44.2	5.0	2440000	Solvent Utilization;Miscellaneous Industrial;All Processes
SSJV	41.8	4.8	2201040	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 3 & 4 (M6) = LDGT2 (M5)
SSJV	35.7	4.1	2460000	Solvent Utilization;Miscellaneous Non-industrial: Consumer and Commercial;All Processes
SSJV	32.5	3.7	2282005	Mobile Sources;Pleasure Craft;Gasoline 2-Stroke
SSJV	31.9	3.6	2265004	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Lawn and Garden Equipment
SSJV	31.5	3.6	2310010	Industrial Processes;Oil and Gas Production: SIC 13;Crude Petroleum
SSJV	30.6	3.5	2620030	Waste Disposal, Treatment, and Recovery;Landfills;Municipal
SSJV	27.6	3.1	2265000	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Recreational Equipment
SSJV	25.6	2.9	309010	Industrial Processes;Fabricated Metal Products;Electroplating Operations
SSJV	20.6	2.4	2282010	Mobile Sources;Pleasure Craft;Gasoline 4-Stroke
SSJV	16.5	1.9	2401002	Solvent Utilization;Surface Coating;Architectural Coatings - Solvent-based
SSJV	15.1	1.7	2201070	Mobile Sources;Highway Vehicles - Gasoline;Heavy Duty Gasoline Vehicles 2B thru 8B & Buses (HDGV)
SSJV	14.8	1.7	2265005	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Agricultural Equipment
SSJV	14.7	1.7	2401005	Solvent Utilization;Surface Coating;Auto Refinishing: SIC 7532
SSJV	14.6	1.7	2401020	Solvent Utilization;Surface Coating;Wood Furniture: SIC 25
SSJV	12.0	1.4	2801500	Miscellaneous Area Sources;Agriculture Production - Crops;Agricultural Field Burning - whole field set on fire
SSJV	11.3	1.3	2265003	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Industrial Equipment
SSJV	10.9	1.2	2260004	Mobile Sources;Off-highway Vehicle Gasoline, 2-Stroke;Lawn and Garden Equipment
SSJV	10.6	1.2	2230074	Mobile Sources;Highway Vehicles - Diesel;Heavy Duty Diesel Vehicles (HDDV) Class 8A & 8B
SSJV	10.3	1.2	2415300	Solvent Utilization;Degreasing;All Industries: Cold Cleaning
SSJV	10.1	1.2	2265006	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Commercial Equipment
SSJV	10.1	1.2	2460500	Solvent Utilization;Miscellaneous Non-industrial: Consumer and Commercial;All Coatings and Related Products
SSJV	9.5	1.1	2501011	Storage and Transport;Petroleum and Petroleum Product Storage;Residential Portable Gas Cans
SSJV	9.2	1.0	2201080	Mobile Sources;Highway Vehicles - Gasoline;Motorcycles (MC)
LA	546.0	15.9	2201001	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Vehicles (LDGV)
LA	345.8	10.1	2201020	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 1 & 2 (M6) = LDGT1 (M5)
LA	244.8	7.1	2265004	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Lawn and Garden Equipment
LA	207.7	6.0	2282005	Mobile Sources;Pleasure Craft;Gasoline 2-Stroke
LA	201.2	5.9	2460000	Solvent Utilization;Miscellaneous Non-industrial: Consumer and Commercial;All Processes
LA	178.1	5.2	2201040	Mobile Sources;Highway Vehicles - Gasoline;Light Duty Gasoline Trucks 3 & 4 (M6) = LDGT2 (M5)
LA	114.0	3.3	2401050	Solvent Utilization;Surface Coating;Miscellaneous Finished Metals: SIC 34 - (341 + 3498)
LA	106.8	3.1	2282010	Mobile Sources;Pleasure Craft;Gasoline 4-Stroke
LA	104.9	3.0	2620030	Waste Disposal, Treatment, and Recovery;Landfills;Municipal
LA	88.2	2.6	2260004	Mobile Sources;Off-highway Vehicle Gasoline, 2-Stroke;Lawn and Garden Equipment
LA	84.1	2.4	2401005	Solvent Utilization;Surface Coating;Auto Refinishing: SIC 7532
LA	74.4	2.2	2265006	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Commercial Equipment
LA	70.6	2.1	2401090	Solvent Utilization;Surface Coating;Miscellaneous Manufacturing
LA	70.1	2.0	2460500	Solvent Utilization;Miscellaneous Non-industrial: Consumer and Commercial;All Coatings and Related Products
LA	69.1	2.0	2415300	Solvent Utilization;Degreasing;All Industries: Cold Cleaning
LA	63.3	1.8	2201070	Mobile Sources;Highway Vehicles - Gasoline;Heavy Duty Gasoline Vehicles 2B thru 8B & Buses (HDGV)
LA	58.7	1.7	306888	Industrial Processes;Petroleum Industry;Fugitive Emissions
LA	56.5	1.6	2401025	Solvent Utilization;Surface Coating;Metal Furniture: SIC 25
LA	55.9	1.6	2501011	Storage and Transport;Petroleum and Petroleum Product Storage;Residential Portable Gas Cans
LA	41.8	1.2	2440000	Solvent Utilization;Miscellaneous Industrial;All Processes
LA	40.4	1.2	2401040	Solvent Utilization;Surface Coating;Metal Cans: SIC 341
LA	39.8	1.2	2265003	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Industrial Equipment
LA	39.5	1.1	2401070	Solvent Utilization;Surface Coating;Motor Vehicles: SIC 371
LA	38.7	1.1	2201080	Mobile Sources;Highway Vehicles - Gasoline;Motorcycles (MC)
LA	36.9	1.1	2265001	Mobile Sources;Off-highway Vehicle Gasoline, 4-Stroke;Recreational Equipment

Matching Measured and Modeled Species

Table S2a. Matching grouped measured and modeled SOA species. Measured SOC contributions estimated by dividing the total tracer concentration by group with a photochemical reaction chamber derived SOC mass fraction for the same group.

Species	Measured Species	CMAQ Model Species
PM2.5 Species (ugC/m³)		
PM2.5 Elemental Carbon	EC	AECI + AECI
PM2.5 Organic Carbon	OC	(AXYL1J + AXYL2J + AXYL3J)/2.0 + (ATOL1J + ATOL2J + ATOL3J)/2.0 + (ABNZ1J + ABNZ2J + ABNZ3J)/2.0 + (AISO1J + AISO2J)/1.6 + AISO3J/2.7 + (ATRP1J + ATRP2J)/1.4 + ASQTJ/2.1 + 0.64*AALKJ + AORGCI/2.0 + (AOLGBJ + AOLGAJ)/2.1 + APOCI + APOCI
PM2.5 SOA species (ugC/m³)		
Isoprene	(2-methyl erythritol + 2-methyl trietol + 2-methylglyceric acid) * (1/0.063)	(AISO1J + AISO2J)/1.6 + AISO3J/2.7 + (AOLGBJ/2.1 * iso_fraction)
Monoterpenes	(3-acetyl hexanedioic acid + 3-hydroxyglutaric acid + 3-methyl-1,2,3-butanetricarboxylic acid + 3-acetyl pentanedioic acid + Pinic acid + 2-hydroxy-4,4-dimethylglutaric acid) * (1/0.231)	(ATRP1J + ATRP2J)/1.4 + (AOLGBJ/2.1 * trp_fraction)
Sesquiterpenes	(b-caryophellinic acid) * (1/0.023)	ASQT/2.1 + (AOLGB/2.1 * (1-trp_fraction-iso_fraction))
Toluenes + Xylenes	(2-3-diOH-4-oxo-pentanoic acid) * (1/0.0079)	(ATOL1J + ATOL2J + ATOL3J + AXYL1J + AXYL2J + AXYL3J)/2 + (AOLGAJ/2.1 * (tol+xyl_fraction))
Benzenes	n/a	(ABNZ1J + ABNZ2J + ABNZ3J)/2 + (AOLGAJ/2.1 * bnz_fraction)
Alkanes	n/a	AALKJ/1.56 + (AOLGAJ/2.1 * alk_fraction)
Methylglyoxal + glyoxal	n/a	AORGCI/2
Biomass related	levoglucosan * (1/0.126)	n/a
Naphthalene	(phthalic acid + isophthalic acid) * (1/0.0357)	n/a
$\text{iso_fraction} = (\text{AISO1J} + \text{AISO2J} + \text{AISO3J}) / (\text{AISO1J} + \text{AISO2J} + \text{AISO3J} + \text{ATRP1J} + \text{ATRP2J} + \text{ASQT})$ $\text{trp_fraction} = (\text{ATRP1J} + \text{ATRP2J}) / (\text{AISO1J} + \text{AISO2J} + \text{AISO3J} + \text{ATRP1J} + \text{ATRP2J} + \text{ASQT})$ $\text{bnz_fraction} = (\text{ABNZ1J} + \text{ABNZ2J} + \text{ABNZ3J}) / (\text{ATOL1J} + \text{ATOL2J} + \text{ATOL3J} + \text{AXYL1J} + \text{AXYL2J} + \text{AXYL3J} + \text{ABNZ1J} + \text{ABNZ2J} + \text{ABNZ3J} + \text{AALKJ})$ $\text{alk_fraction} = (\text{AALKJ}) / (\text{ATOL1J} + \text{ATOL2J} + \text{ATOL3J} + \text{AXYL1J} + \text{AXYL2J} + \text{AXYL3J} + \text{ABNZ1J} + \text{ABNZ2J} + \text{ABNZ3J} + \text{AALKJ})$ $\text{tol+xyl_fraction} = (\text{ATOL1J} + \text{ATOL2J} + \text{ATOL3J} + \text{AXYL1J} + \text{AXYL2J} + \text{AXYL3J}) / (\text{ATOL1J} + \text{ATOL2J} + \text{ATOL3J} + \text{AXYL1J} + \text{AXYL2J} + \text{AXYL3J} + \text{ABNZ1J} + \text{ABNZ2J} + \text{ABNZ3J} + \text{AALKJ})$		

Table S2b. Matching grouped measured and modeled VOC species.

Species	Measured Species	CMAQ Model Species
Gas Species (ppbC)		
Isoprene	Isoprene	5*ISOPRENE
Benzenes	Benzene	6*BENZENE
Monoterpenes	beta-pinene + alpha-pinene + 2-carene + limonene	10*TERP
Sesquiterpenes	n/a	15*SESEQ
Toluenes	Toluene + Ethylbenzene + Isopropylbenzene*0.78 + n-propylbenzene*0.78	7*TOLUENE
Xylenes	m-p xylene + o-xylene + m-ethyltoluene*0.89 + p-ethyltoluene*0.89 + 1,3,5-trimethylbenzene*0.89 + o-ethyltoluene*0.89 + 1,2,4-trimethylbenzene*0.89 + p-diethylbenzene*0.80	8*(MXYL + OXYL + PXYL)
Ethane	Ethane	2*ALK1
Ethene	Ethene	2*ETHENE
Formaldehyde	Formaldehyde	HCHO
Methanol	Methanol	MEOH
Ethanol	Ethanol	2*ETOH
Total VOC (non-methane)	Sum of all measured non-methane VOC species	(3.*ACETONE + 2.*ACETYLENE + 3.*ACROLEIN + 2.*ALK1 + 3.*ALK2 + 4.*ALK3 + 5.*ALK4 + 8.*ALK5 + 8.*ARO1 + 9.*ARO2 + 4.*BACL + 7.*BALD + 6.*BENZENE + 4.*BUTADIENE13 + 2.*CCHO + 2.*CCOOH + 7.*CRES + 2.*ETHENE + 2.*ETOH + 2.*GLY + HCHO + HCOOH + 5.*ISOPRENE + 5.*IPRD + 4.*MEK + MEOH + 4.*MACR + 3.*MGLY + 4.*MVK + 5.*OLE1 + 5.*OLE2 + 8.*(PXYL + MXYL + OXYL) + 6.*PRD2 + 3.*PROPENE + 3.*RCHO + 3.*RCOOH + 7.*TOLUENE + 9.*TRIMETH_BENZ124 + 6.0*RNO3)

Figure S1. Model domain and CALNEX monitor locations.

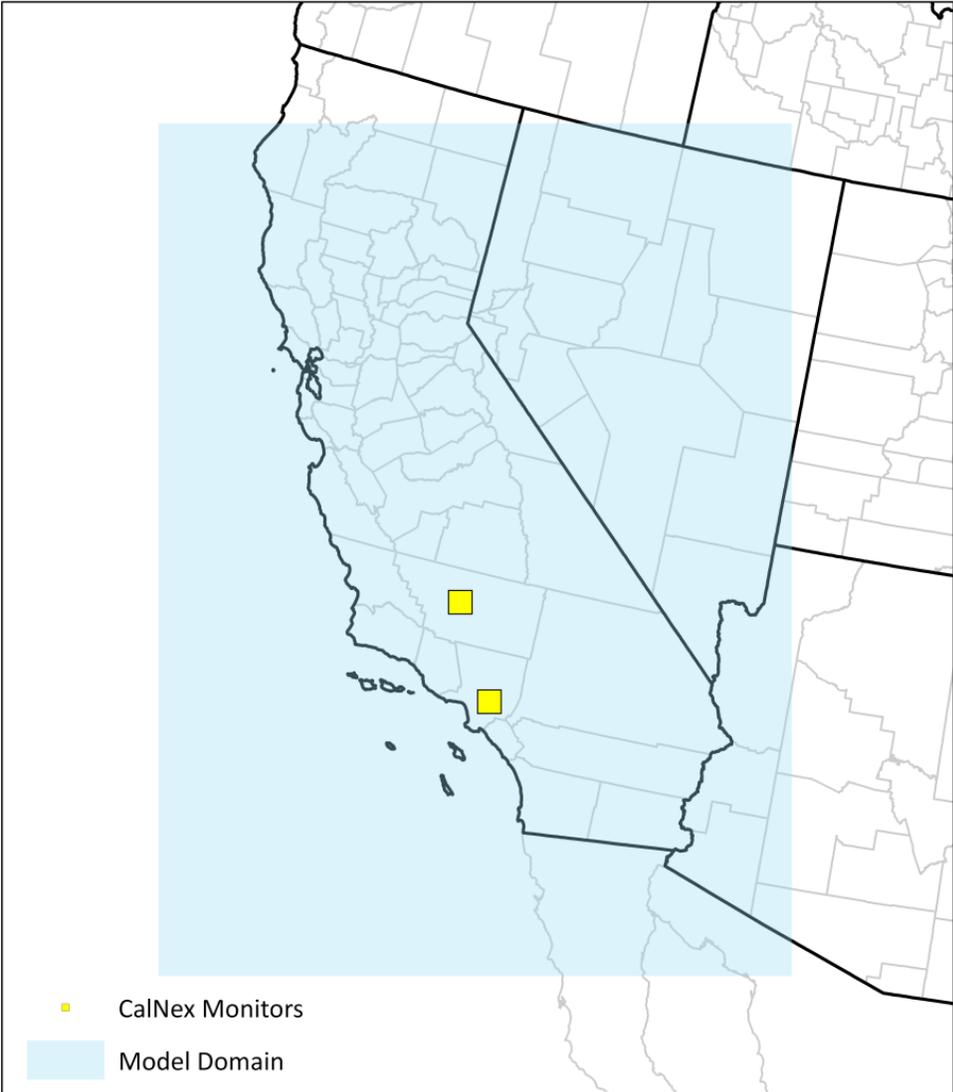


Figure S2. CMAQ SOA schematic diagram.

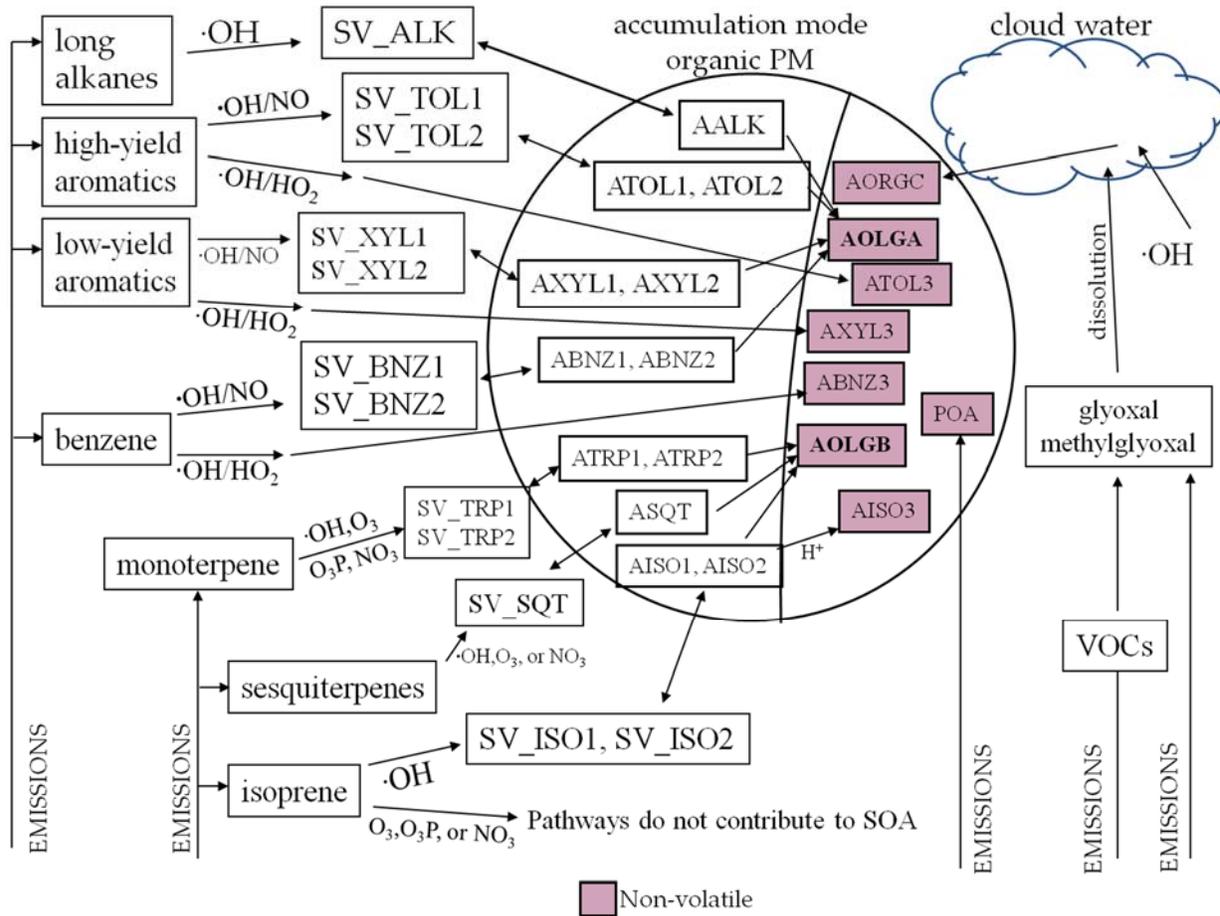


Figure S3a. Bakersfield CALNEX site and nearby CSN location.



Figure S4a. Observed daily average temperature compared with daily average PM2.5 fossil carbon, PM2.5 modern carbon, PM2.5 organic carbon, and PM2.5 elemental carbon at Pasadena.

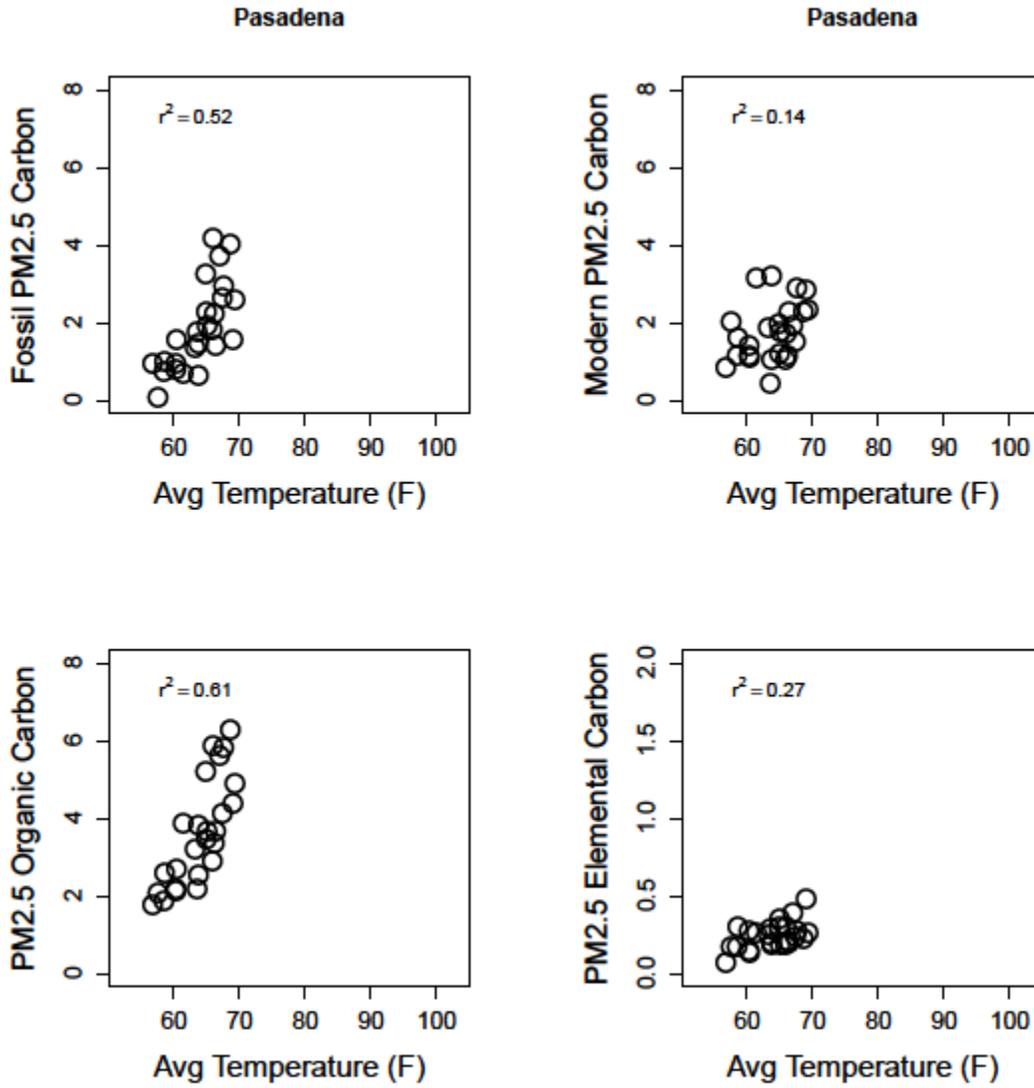


Figure S4b. Observed daily average temperature compared with daily average PM2.5 fossil carbon, PM2.5 modern carbon, PM2.5 organic carbon, and PM2.5 elemental carbon at Bakersfield.

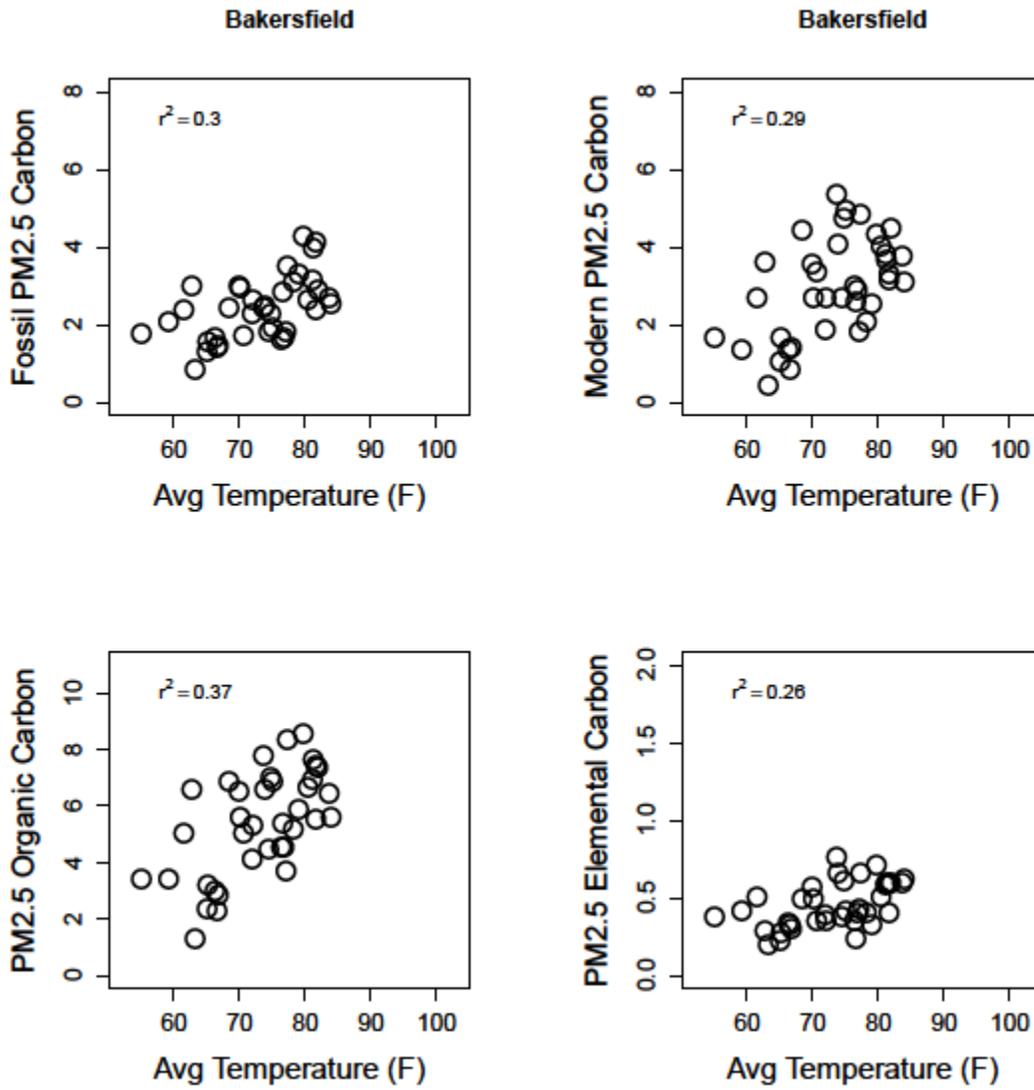


Figure S5. Observed modern carbon fraction, PM2.5 modern carbon, PM2.5 fossil carbon, and PM2.5 elemental carbon by day of the week for Pasadena and Bakersfield.

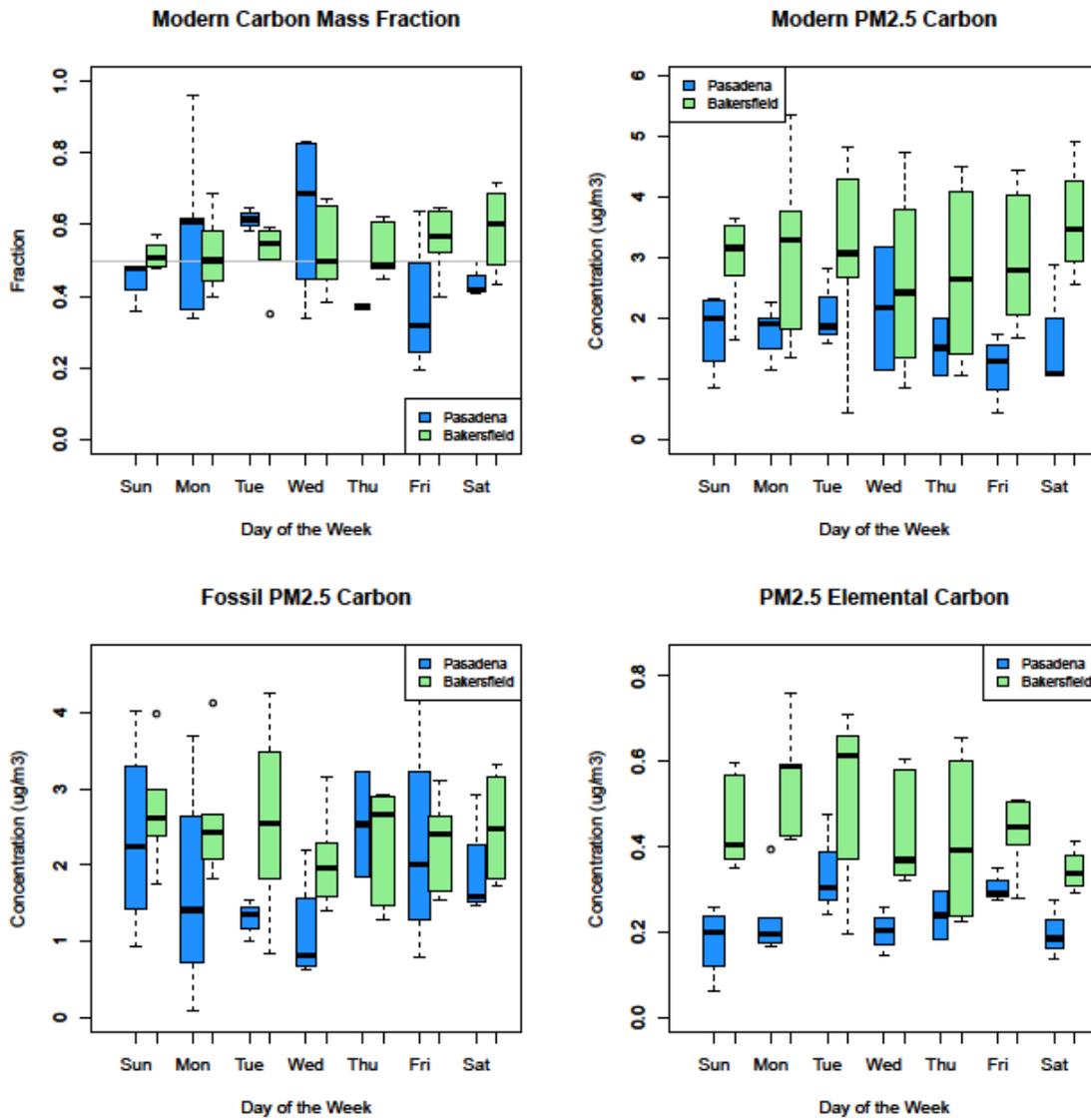


Figure S6. Model predicted and measured PM_{2.5} elemental carbon at Pasadena and Bakersfield.

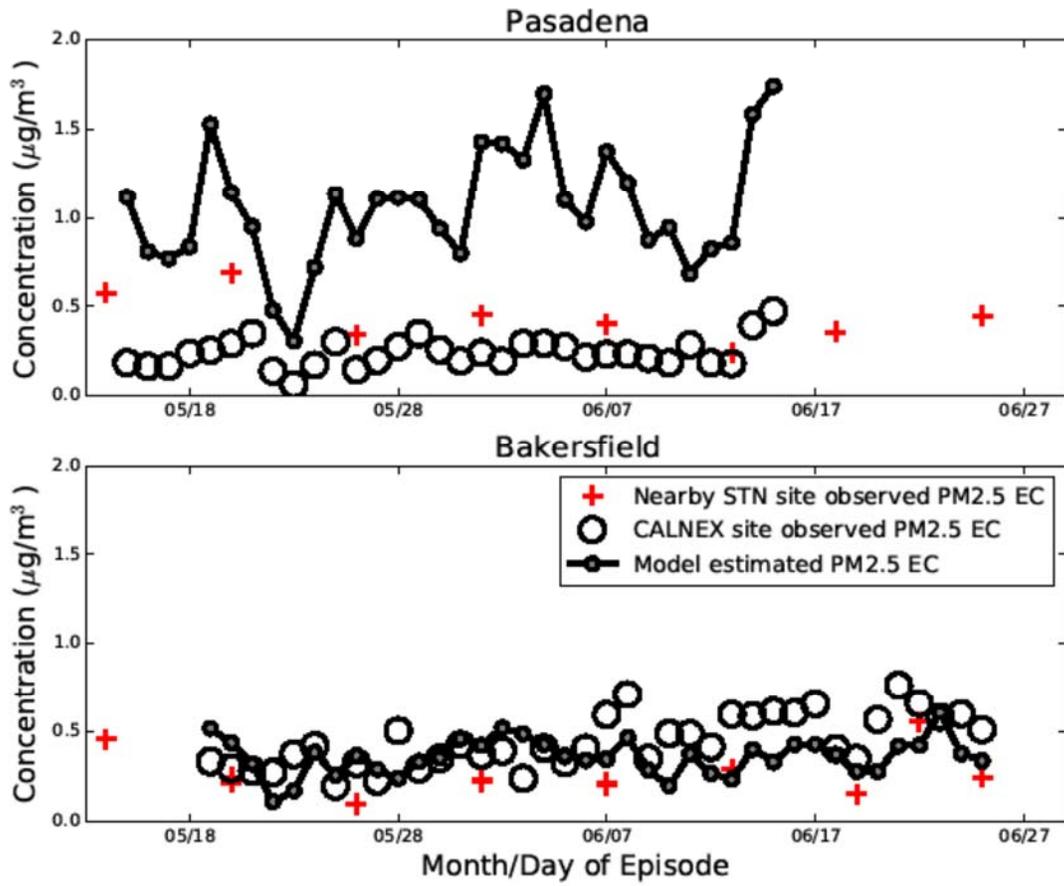


Figure S7. Model and observed hourly VOC paired in time and space.

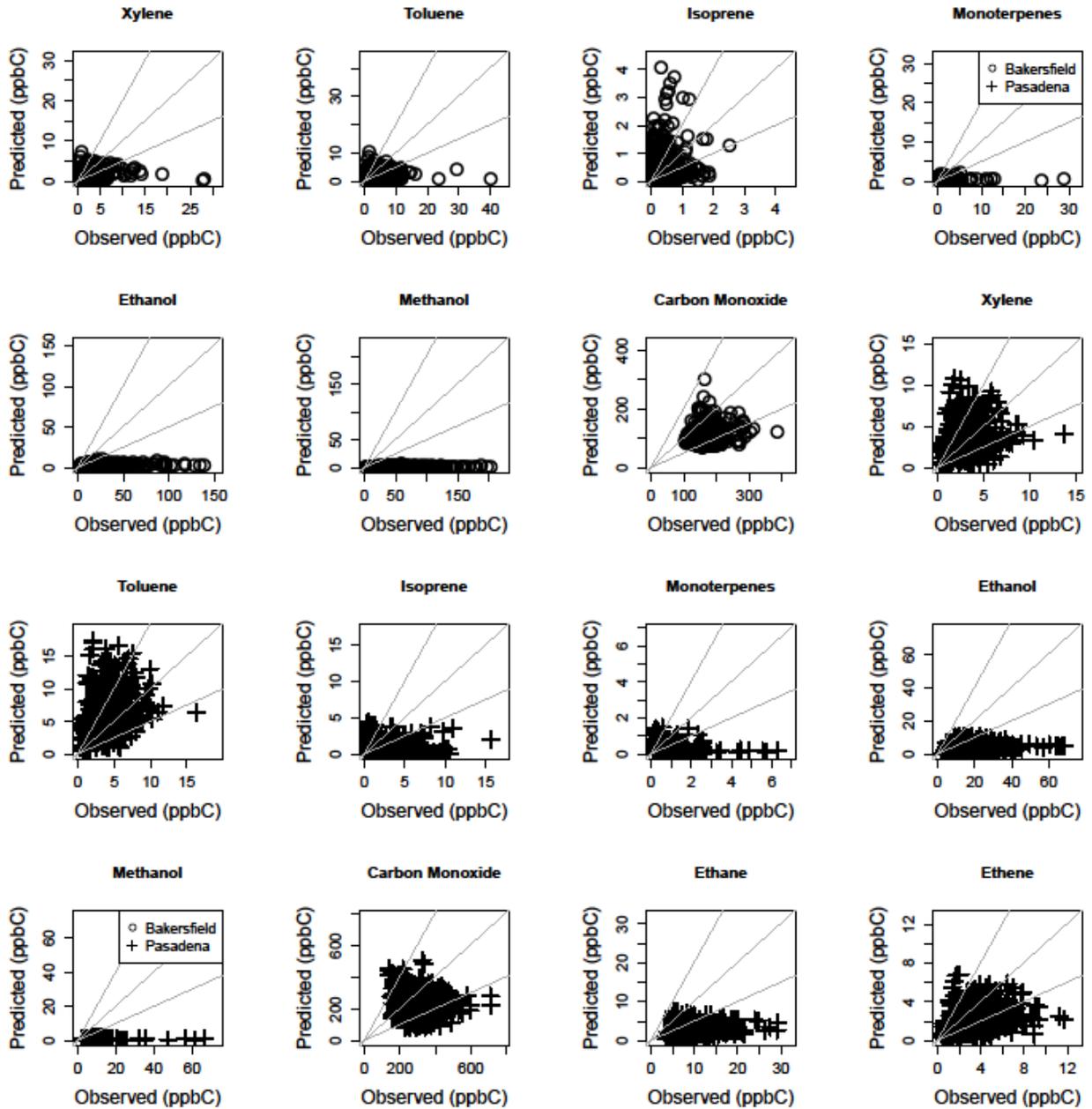


Figure S8. Model and observed mid-morning 3 hour average VOC paired in time and space.

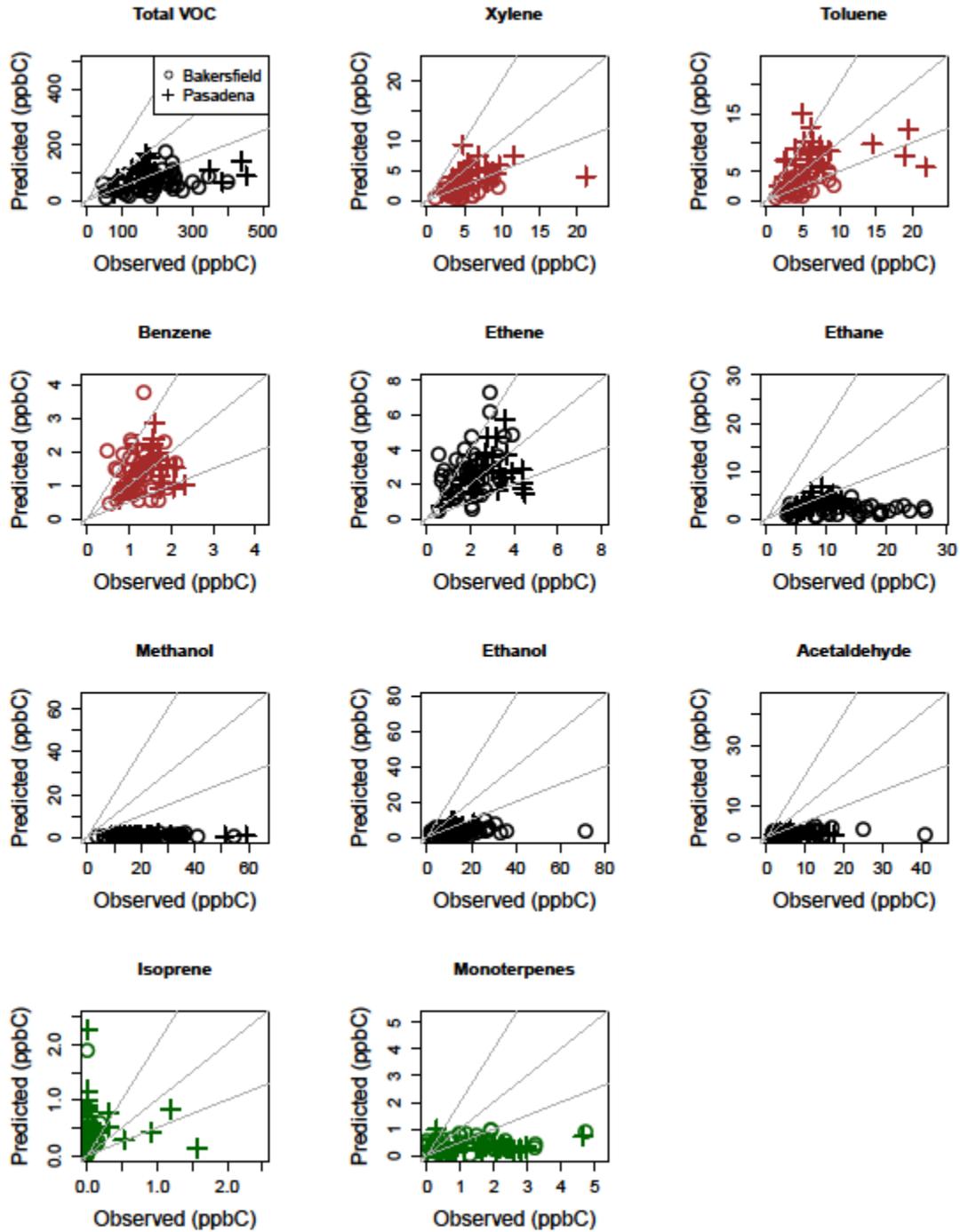


Figure S9a. June average model estimates.

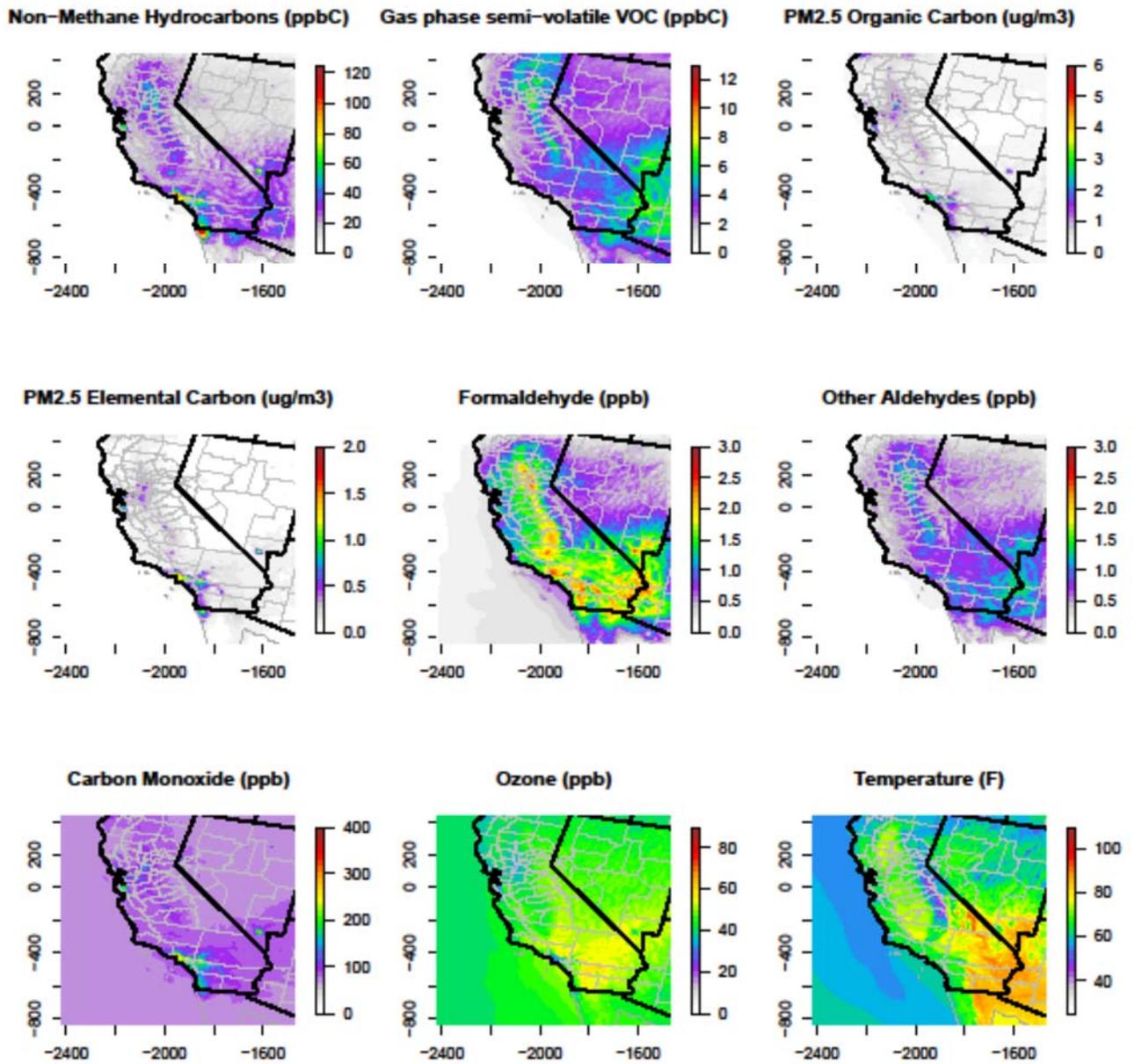


Figure S9b. June average model estimates. Biogenic VOC, semi-volatile products, and SOA.

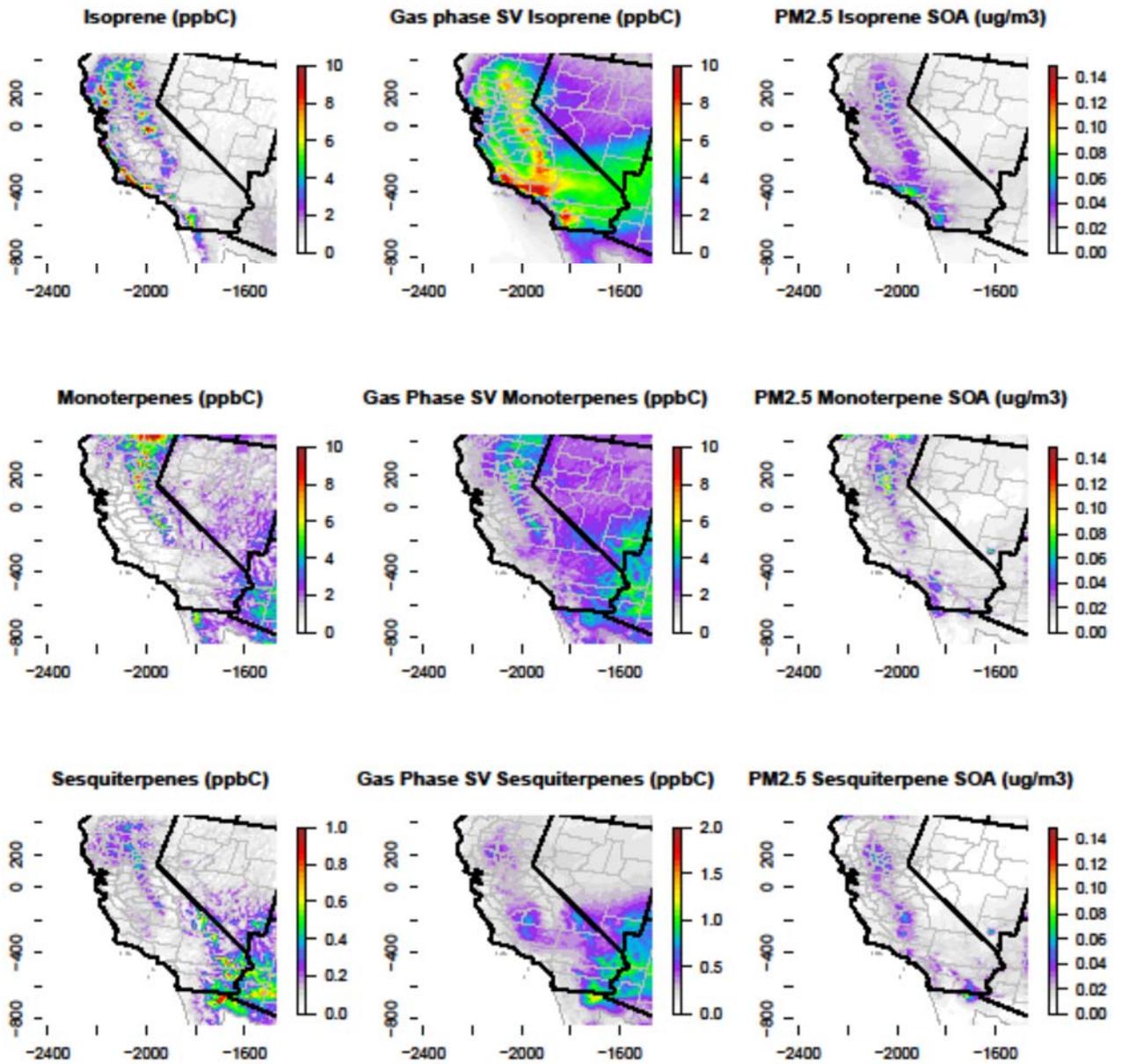


Figure S9c. June average model estimates. Anthropogenic VOC, semi-volatile products, and SOA.

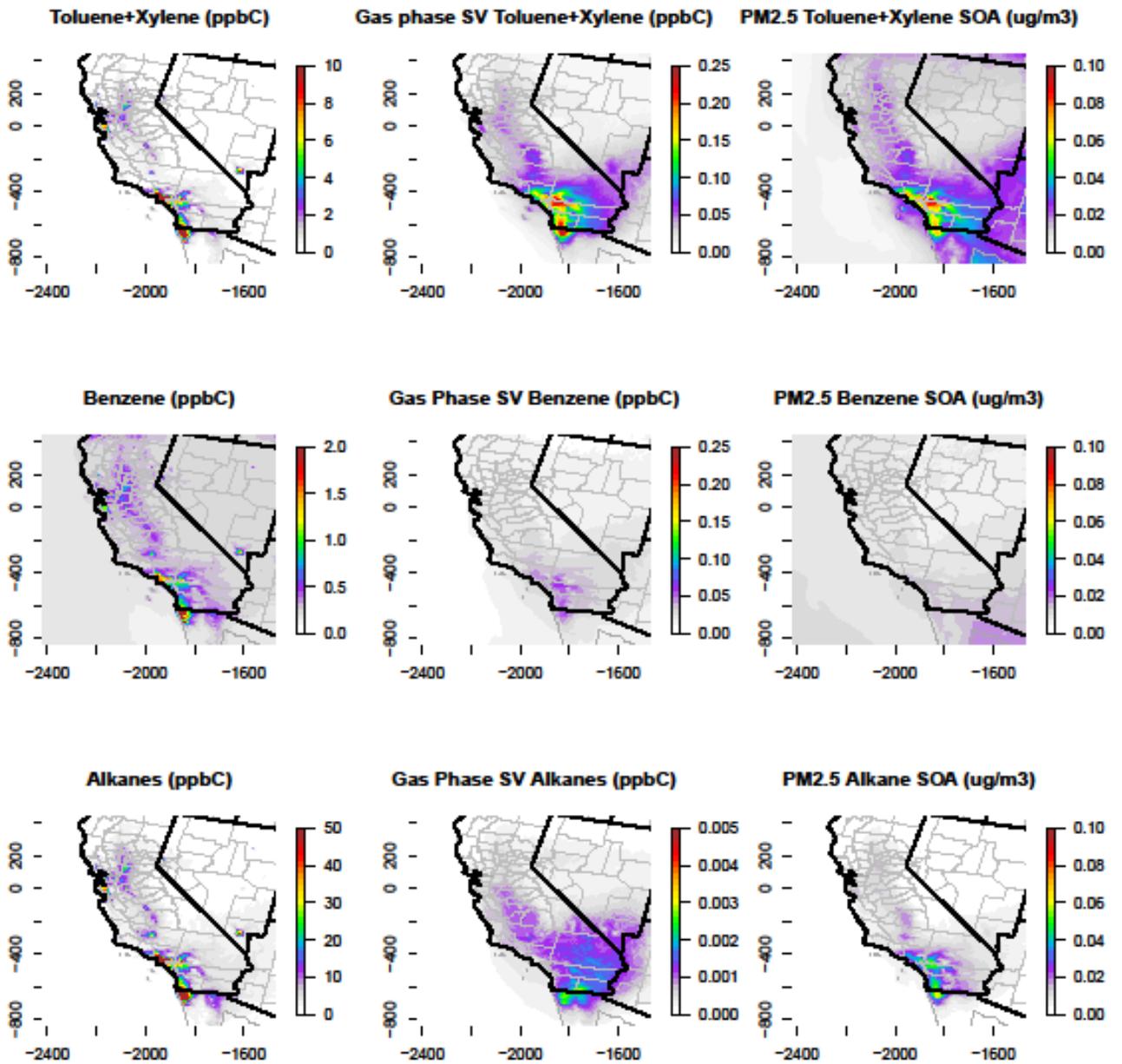
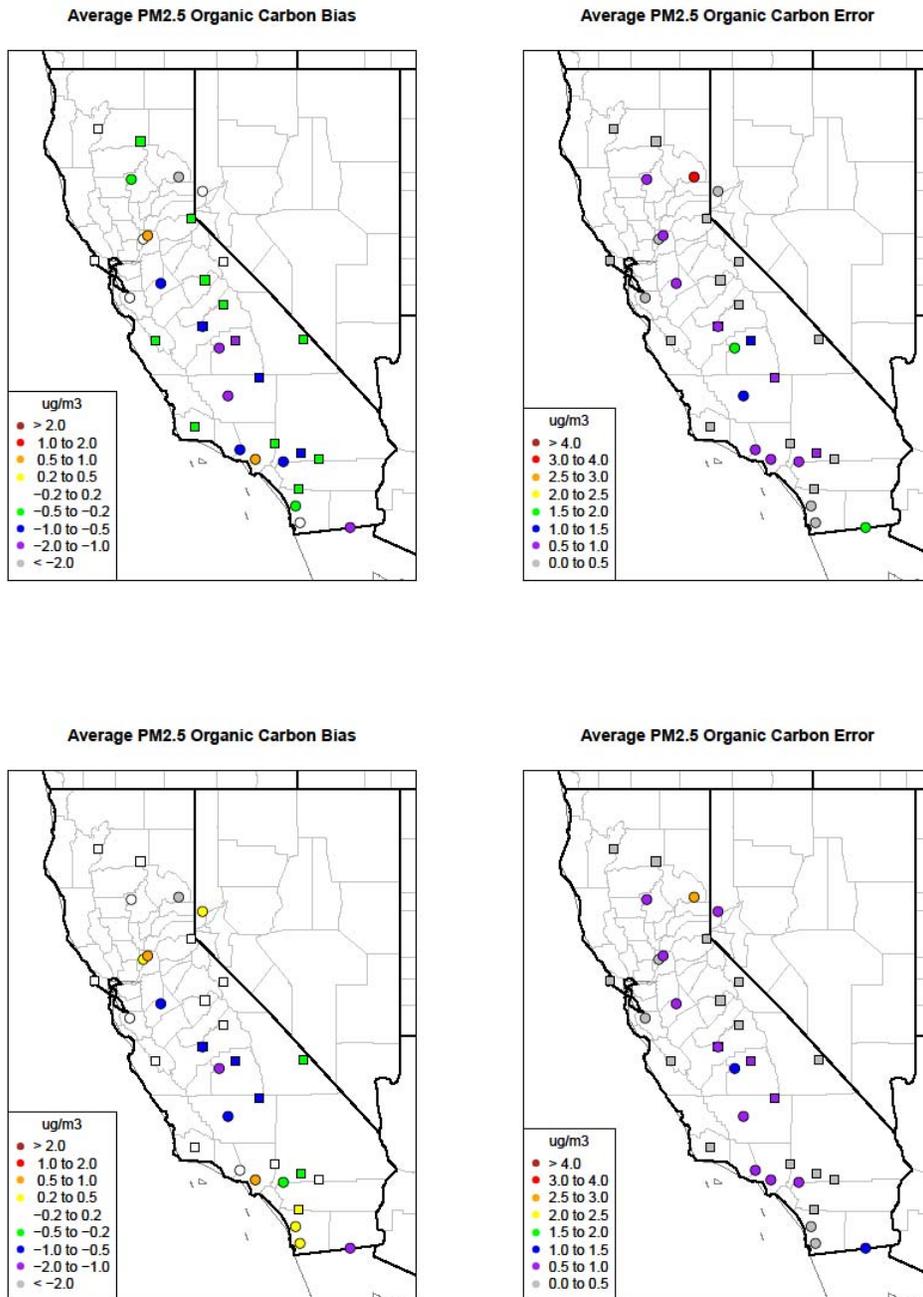


Figure S10. Episode average baseline (top row) and sensitivity (bottom row) model bias and error for PM_{2.5} organic carbon at CSN (circles) and IMPROVE (squares) sites. The large underestimate at the location in northeast California is related to model underestimates of elevated concentrations during a cold period in this area which may have resulted in increased residential fuel combustion. In general, the model underestimates organic carbon, most notably in urban areas. Rural areas dominated by biogenic sources show minimal bias and error.



Sampling and Analysis Methods

CalNex ground measurements took place in Pasadena, CA from 15 May – 15 June 2010 and in Bakersfield, CA from 15 May – 30 June 2010. Filter samples for SOA are integrated for 23-h periods starting at midnight of the designated sampling day. In total, there are 32 filter samples from Pasadena and 36 from the Bakersfield site. The sampling protocols have been described in detail elsewhere (Kleindienst, Lewandowski et al. 2010). PM_{2.5} was collected on quartz filters using high volume PM_{2.5} samplers operated at 0.226 m³ min⁻¹. Each sampler consists of a PM_{2.5} inlet (Tisch Environmental, Cleves, OH) followed by a 90-mm pre-combusted quartz filter (Pall-Life Sciences, East Hills, NY). Organic carbon denuders were not used in-line for these samples.

For the analysis of the SOA tracer compounds, filters and field blanks were treated using the method described by Kleindienst et al. (Kleindienst, Jaoui et al. 2007). Filters were Soxhlet extracted for 24 hours using 125 mL of a 1:1 (v/v) dichloromethane:methanol mixture. Prior to the extraction, cis-ketopinic acid and tetracosane-d₅₀ were added as internal standards. Filter extracts were rotary evaporated to a volume of 1 mL, then evaporated to dryness with ultrazero nitrogen. Extracts were derivatized with 250 µL BSTFA (1% TMCS catalyst) and 100 µL pyridine to give a final volume of 350 µL. The silylated extracts were analyzed by GC-ion trap mass spectroscopy (ITMS) in the methane-Cl mode. Further detailed aspects of the analytical methodology have been given by Jaoui et al. (Jaoui, Kleindienst et al. 2004).

The mass spectral analysis for the organic compounds used as secondary molecular tracers has been described (Edney, Kleindienst et al. 2003). The tracer compounds are grouped by major ion fragments (Kleindienst, Jaoui et al. 2007). Since standards do not exist for the majority of these compounds, the concentrations of all of the tracers were measured as ketopinic acid (KPA). GC-MS analysis for the tracer compounds was conducted using the total ion chromatogram (TIC) or, in cases where coelution or extremely weak signals occurred, by a selected ion technique. By this method, a factor is determined from laboratory samples of the non-co-eluted peaks that represent the fraction of the five ion intensity to the total ion intensity for each tracer compound. The use of the five-ion-to-TIC calibration factor was found to give more consistent concentration estimates than found using a single ion.

The actual or simulated TIC area was then used to calculate the tracer concentrations as KPA. An assessment of the accuracy of this technique has been estimated as 60% for pinic and norpinic acids (Jaoui, Kleindienst et al. 2005), two compounds for which independent standards were produced. The technique should not introduce major uncertainties into the contribution estimates from the field data, since the same analytical procedures were used to establish the laboratory-based mass fractions, thus, compensating for systematic errors found in laboratory and field samples. For the determination of the total carbon, a 1.45 cm² punch was taken from a representative quartz filter and analyzed using the National Institute of Occupational Safety and Health (NIOSH) or thermal-optical transmittance (TOT) method, as described by Birch and Cary (Birch and Cary 1996). Values for OC were corrected for filter background levels. The minimum detect limit for the SOA tracer species is 0.1 ng/m³.

Canisters for volatile organic compound (VOC) analysis were also collected at the two sites. A total of 41 samples were collected at the Bakersfield site and 31 at Pasadena. Samples were collected in evacuated (ca. 1 atm below ambient pressure) 6-liter canisters using Entech CS1200 samplers. Attached to the CS1200 samplers were Entech TM1100 systems that consisted of a battery operated off/on valve activated by a timer control element. The valve was open for a 3-h collection during the 6:00 – 9:00 a.m. (local daylight time) period. The samples were analyzed by GC-FID and reported compounds were

identified by a GC retention time index using a calibration table (CALTABLE) of 402 VOCs developed specifically for GC system. Compounds equal to or less than the detection limit of 0.05 ppbC were defined as zero for comparison presentation of paired experimental and modeled estimates.

References

Birch, M. E. and R. A. Cary (1996). "Elemental carbon-based method for monitoring occupational exposures to particulate diesel exhaust." Aerosol Sci. Technol. **25**: 221-241.

Edney, E. O., T. E. Kleindienst, T. S. Conner, C. D. McIver, E. W. Corse and W. S. Weathers (2003). "Polar organic oxygenates in PM_{2.5} at a southeastern site in the United States." Atmospheric Environment **37**(28): 3947-3965.

Jaoui, M., T. E. Kleindienst, M. Lewandowski and E. O. Edney (2004). "Identification and quantification of aerosol polar oxygenated compounds bearing carboxylic or hydroxyl groups. 1. Method development." Analytical Chemistry **76**(16): 4765-4778.

Jaoui, M., T. E. Kleindienst, M. Lewandowski, J. H. Offenberg and E. O. Edney (2005). "Identification and quantification of aerosol polar oxygenated compounds bearing carboxylic or hydroxyl groups. 2. Organic tracer compounds from monoterpenes." Environmental Science & Technology **39**(15): 5661-5673.

Kleindienst, T. E., M. Jaoui, M. Lewandowski, J. H. Offenberg, C. W. Lewis, P. V. Bhave and E. O. Edney (2007). "Estimates of the contributions of biogenic and anthropogenic hydrocarbons to secondary organic aerosol at a southeastern US location." Atmospheric Environment **41**(37): 8288-8300.

Kleindienst, T. E., M. Lewandowski, J. H. Offenberg, E. O. Edney, M. Jaoui, M. Zheng, X. A. Ding and E. S. Edgerton (2010). "Contribution of Primary and Secondary Sources to Organic Aerosol and PM_{2.5} at SEARCH Network Sites." Journal of the Air & Waste Management Association **60**(11): 1388-1399.

CMAQ AE6 treatment for SOA presented by volatility and O:C ratio

Modeled organic aerosol species, semi-volatile partitioning gas phase analogs, and VOC precursors are matched with a volatility bin assignment based on saturation vapor pressure (C^*) (Table S3). C^* values for semi-volatile and aerosol species are taken from (Carlton et al., 2010) and calculated for gas-phase VOCs based on (Pankow et al., 1994) (Table S3 and Equation S1). The saturation vapor pressure of sesquiterpenes has not been measured and bin assignment is somewhat arbitrary; sesquiterpenes are placed in the $\log(C^*)=5$ bin, two bins lower than monoterpenes and four bins lower than isoprene. Model reference C^* values are translated to local conditions at each sampling location by employing the species-specific enthalpy of vaporization (ΔH_{vap}). Mass concentrations of gas-phase species were calculated from the CMAQ reported mixing ratios (ppb_v):

$$C_i \left(\frac{\mu\text{g}}{\text{m}^3} \right) = X_i \frac{P}{RT} MW_i \quad [1]$$

Where X_i is the gas phase mixing ratio in ppb_v for species i , P and T are the location specific instantaneous pressure and temperature respectively, R is the ideal gas constant, and MW_i is the species molecular weight.

CMAQ predictions of organic material during CalNex are dominated by gas phase species. This is generally true and consistent with early measurements in California (Fraser et al., 1996). Episode average gas and particle organic carbon estimated by CMAQ for each ground site are shown in Figure S11 by saturation vapor pressure. Most of the carbon mass is either in the gas phase or non-volatile aerosol. Mass in the semi-volatile space is largely in the gas-phase. This suggests the potential exists for additional modeled SOC but this space is difficult to constrain due to limited available gas phase measurements. This Figure also shows gaps in the model representation of $\text{PM}_{2.5}$ organic carbon volatility distribution. Additional precursors with a saturation vapor pressure less than $10^8 \mu\text{g}/\text{m}^3$ (such as IVOCs) could ameliorate some of the SOC under prediction (Shrivastava et al., 2008).

Smog chamber experiments conducted at high mass loadings and the SOC parameterizations developed from those experiments, and used in CMAQ, may be a contributing factor to the absence of material in low volatility bins, possibly due to bin mis-assignment (Stanier et al., 2008). It would also seem that some material mapped to the “non-volatile” bin actually exhibit some partitioning and should be elsewhere in the distribution, i.e., $-4 < \log_{10} C^* < -1$. Missing SOC from PAHs (the naphthalene tracer) would contribute organic material to bins 0 and 2 with an O:C ratio ~ 0.5 (Chan et al., 2009b; Pye and Pouliot, 2012) and would not ameliorate the empty gaps. As noted previously, the modeling system does well at replicating total VOC but tends to underestimate aerosol carbon at these sites during this time period. For model applications requiring yields of species with saturation concentrations less than $0.1 \mu\text{g}/\text{m}^3$, experiments must be able to measure similarly low concentrations to provide robust parameters (Stanier et al., 2008).

References

- Carlton, A.G., Bhave, P.V., Napelenok, S.L., Edney, E.O., Sarwar, G., Pinder, R.W., Pouliot, G.A., Houyoux, M., 2010. Treatment of secondary organic aerosol in CMAQv4.7. *Environmental Science and Technology* 44, 8553-8560.
- Chan, A.W.H., Kautzman, K.E., Chhabra, P.S., Surratt, J.D., Chan, M.N., Crouse, J.D., Kuerten, A., Wennberg, P.O., Flagan, R.C., Seinfeld, J.H., 2009. Secondary organic aerosol formation from

photooxidation of naphthalene and alkylnaphthalenes: implications for oxidation of intermediate volatility organic compounds (IVOCs). *Atmospheric Chemistry and Physics* 9, 3049-3060.

Fraser, M.P., Grosjean, D., Grosjean, E., Rasmussen, R.A., Cass, G.R., 1996. Air quality model evaluation data for organics .1. Bulk chemical composition and gas/particle distribution factors. *Environmental Science & Technology* 30, 1731-1743.

Pankow, J.F., Isabelle, L.M., Buchholz, D.A., Luo, W.T., Reeves, B.D., 1994. Gas Particle Partitioning of Polycyclic Aromatic Hydrocarbons and Alkanes to Environmental Tobacco Smoke. *Environmental Science & Technology* 28, 363.

Pye, H.O.T., Pouliot, G.A., 2012. Modeling the Role of Alkanes, Polycyclic Aromatic Hydrocarbons, and Their Oligomers in Secondary Organic Aerosol Formation. *Environmental Science & Technology*.

Shrivastava, M.K., Lane, T.E., Donahue, N.M., Pandis, S.N., Robinson, A.L., 2008. Effects of gas particle partitioning and aging of primary emissions on urban and regional organic aerosol concentrations. *Journal of Geophysical Research: Atmospheres* (1984–2012) 113.

Stanier, C.O., Donahue, N.M., Pandis, S.N., 2008. Parameterization of secondary organic aerosol mass fractions from smog chamber data. *Atmospheric Environment* 42, 2276-2299.

Volatility Estimation

Table S3. Calculation of volatility bin for gas phase organic species.

Species Name	Molecular Weight (g/gmol)	Vapor Pressure (mmHg)	Saturation Vapor Pressure (c*)	log(c*)	Note
ISOP	68.12	550	2.05E+09	9.31	
ALK2,ALK3,ALK4	57.6	100	3.15E+08	8.50	MW=average of 36.7, 58.6, and 77.6
MGLY	108	27	1.60E+08	8.20	Hydrated methylglyoxal (higher MW)
MGLY	72.1	27	1.07E+08	8.03	
TOLUENE	92.1	20	1.01E+08	8.00	
GLY	92	18	9.06E+07	7.96	Hydrated glyoxal (higher MW)
BENZENE	78	20	8.54E+07	7.93	
GLY	58.1	18	5.72E+07	7.76	
MXYL,OXYL,PXYL	106.2	7	4.07E+07	7.61	
TERP	136	3.5	2.61E+07	7.42	
Napthalene	128.19	0.087	6.10E+05	5.79	Not included in CMAQv5.0.2
SESQ	204	0.01	1.12E+05	5.05	Assumed vapor pressure
ATRP2, SV_TRP2			1.34E+02	2.13	c* from Carlton et al, 2010
AISO1, SV_ISO1			1.16E+02	2.06	c* from Carlton et al, 2010
ABNZ2, SV_BNZ2			1.11E+02	2.05	c* from Carlton et al, 2010
AXYL2, SV_XYL2			3.45E+01	1.54	c* from Carlton et al, 2010
ASQT, SV_SQT			2.50E+01	1.40	c* from Carlton et al, 2010
ATOL2, SV_TOL2			2.13E+01	1.33	c* from Carlton et al, 2010
ATRP1, SV_TRP1			1.48E+01	1.17	c* from Carlton et al, 2010
ATOL1, SV_TOL1			2.33E+00	0.37	c* from Carlton et al, 2010
AXYL1, SV_XYL1			1.31E+00	0.12	c* from Carlton et al, 2010
AISO2, SV_ISO2			6.17E-01	-0.21	c* from Carlton et al, 2010
ABNZ1, SV_BNZ1			3.02E-01	-0.52	c* from Carlton et al, 2010
AALK, SV_ALK			2.00E-02	-1.70	c* from Carlton et al, 2010

Equation S1.

$$C_i^* = \frac{10^6 MW_i \xi_i p_{L,i}^o}{RT} = \frac{10^6 \left(\frac{\mu g}{g}\right) \times MW_i \left(\frac{g}{gmol}\right) \times \xi_i \times p_{L,i}^o (mmHG)}{62.36 \left(\frac{L mmHG}{K mol}\right) \times 293(K) \times \frac{m^3}{1000L}}$$

Note that ξ_i is assumed to be 1. This assumption is useful as an index for comparison among species but does not provide a complete description of the partitioning potential of water-soluble organic species.

Carbon and Oxygen Relationships

Table S4. Comparison of O:C ratio estimated from the assigned OM:OC ratios and estimated by counting the number of carbon, oxygen and hydrogen atoms.

Species	MW	OM/OC	# of C derived from parent compound	# of O integer dervied from MW less carbon mass	# of H integer dervied from MW less carbon & oxygen mass	O:C calculated from suspected chemical makeup	O:C (OM/OC-1)*(3/4)
AISO1	96	1.6	5	2	4	0.40	0.45
AISO2	96	1.6	5	2	4	0.40	0.45
ABNZ1	144	2	6	4	8	0.67	0.75
ABNZ2	144	2	6	4	8	0.67	0.75
ABNZ3	144	2	6	4	8	0.67	0.75
AALK	150	1.56	8	3	6	0.38	0.42
AISO3	162	2.7	5	6	6	1.20	1.28
ATOL1	168	2	7	5	4	0.71	0.75
ATOL2	168	2	7	5	4	0.71	0.75
ATOL3	168	2	7	5	4	0.71	0.75
ATRP1	168	1.4	10	3	0	0.30	0.30
ATRP2	168	1.4	10	3	0	0.30	0.30
AOLGA	176.4	2.1	8	5	0	0.63	0.83
AORGC	177	2	3	8	13	2.67	0.75
AXYL1	192	2	8	6	0	0.75	0.75
AXYL2	192	2	8	6	0	0.75	0.75
AXYL3	192	2	8	6	0	0.75	0.75
AOLGB	252	2.1	10	8	4	0.80	0.83
ASQT	378	2.1	15	12	6	0.80	0.83

NOTE: This analysis is somewhat limited in its application to AOLGA, AOLGB and AORGC. AOLGA and AOLGB represent oligomerization with a potentially changing number of carbon atoms. AORGC represents carboxylic acid and high molecular weight compounds (max O:C of 2). The assigned MW of ATRP1, ATRP2, AXYL1, AXYL2, AXYL3 and AOLGA is not able to accommodate H atoms consistent with the assigned OM:OC ratios.

Figure S11. Episode average CMAQ organic mass (baseline on left and sensitivity on right) plotted by saturation vapor pressure. Gaps in the distribution are associated with chemical properties that potentially provide insight regarding the identity missing organic aerosol mass noted in CMAQ predictions. Model baseline estimates shown at left and sensitivity simulation at right.

