An Updated Modeling Framework to Simulate Los Angeles Air Quality. Part 1: Model Development, Evaluation, and Source Apportionment.

Elyse A. Pennington[1], Yuan Wang[2,] Benjamin C. Schulze[3,4], Karl M. Seltzer[5], Jiani Yang[3,4], Bin Zhao[6,7], Zhe Jiang[8], Hongru Shi[9], Melissa Venecek[10], Daniel Chau[10], Benjamin N. Murphy[11], Christopher M. Kenseth[12], Ryan X. Ward[4], Havala O. T. Pye[11], and John H. Seinfeld[1,4]

[1]Department of Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

[2]Department of Earth System Science, Stanford University, Stanford, CA 94305

[3]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125

[4]Department of Environmental Science and Engineering, California Institute of Technology, Pasadena, CA 91125

[5]Office of Air and Radiation, US Environmental Protection Agency, Research Triangle Park, NC 27711

[6]State Key Joint Laboratory of Environmental Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China

[7]State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Beijing 100084, China

[8]Carbon Neutrality Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

[9]Key Laboratory of Middle Atmosphere and Global Environment Observation, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

[10]Modeling and Meteorology Branch, California Air Resources Board, Sacramento, CA 95814

[11]Office of Research and Development, US Environmental Protection Agency, Research Triangle Park, NC 27711

[12]Department of Chemistry, California Institute of Technology, Pasadena, CA 91125

Corresponding Authors: Yuan Wang (yzwang@stanford.edu) and John H. Seinfeld (seinfeld@caltech.edu)

Abstract

This study describes a modeling framework, model evaluation, and source apportionment to understand the causes of Los Angeles (LA) air pollution. A few major updates are applied to the Community Multiscale Air Quality (CMAQ) Model with high spatial resolution (1 km x 1 km). The updates include dynamic traffic emissions based on real-time on-road information and recent emission factors and secondary organic aerosol (SOA) schemes to represent volatile chemical

products (VCP). Meteorology is well-predicted compared to ground-based observations, and the emission rates from multiple sources (i.e., on-road, volatile chemical product, area, point, biogenic, and sea spray) are quantified. Evaluation of the CMAQ model shows that ozone is well-predicted despite inaccuracies in nitrogen oxide (NO_x) predictions. Particle matter (PM) is underpredicted compared to concurrent measurements made with an aerosol mass spectrometer (AMS) in Pasadena. Inorganic aerosol is well-predicted while SOA is underpredicted. Modeled SOA consists of mostly organic nitrates and products from oxidation of alkane-like intermediate volatility organic compounds (IVOCs) and has missing components that behave like less-oxidized oxygenated organic aerosol (LO-OOA). Source apportionment demonstrates that the urban areas of the LA Basin and vicinity are NO_x-saturated (VOC-sensitive) with the largest sensitivity of O_3 to changes in VOCs in the urban core. Differing oxidative capacities in different regions impact the nonlinear chemistry leading to PM and SOA formation, which is quantified in this study.

1. Introduction

Air quality is influenced by particle- and gas-phase species which can impact human and environmental health. Particulate matter (PM), or aerosols, affect human health (Lim et al., 2012), climate (Intergovernmental Panel on Climate Change, 2014), and visibility (Hyslop, 2009). A major fraction of PM in urban areas is organic (Q. Zhang et al., 2007), which itself is largely secondary in nature (Jimenez et al., 2009). Secondary organic aerosol (SOA) comprises thousands of species which are formed via complex chemistry that also produces ozone (O_3). O_3 is an oxidant which can damage human (Nuvolone et al., 2018) and plant (Sandermann Jr, 1996) health. Reactive organic gases (ROG) are necessary precursors to these pollutants and span a range of properties, including vapor pressure and oxygen-to-carbon ratio. Volatile organic compounds (VOCs) and nitrogen oxides (NO_x) control O_3 and SOA formation, and semivolatile organic compounds (SVOCs) and intermediate volatility organic compounds (IVOCs) have high potential to form SOA (Robinson et al., 2007).

The Los Angeles Basin has a long history of air pollution resulting from substantial anthropogenic emissions and unique meteorology. On-road mobile emissions have historically been the most important source of atmospheric pollution in the LA Basin, but emissions have decreased as emissions control technologies (i.e., catalytic converters) have improved, vehicle fuel efficiencies have increased, and electric vehicles have become more prevalent (Khare & Gentner, 2018). Other sources of emissions have become more important, particularly VOC and SVOC emissions from volatile chemical products (VCPs). VCPs are consumer and industrial products that utilize evaporative organics (Seltzer et al., 2021) and can form SOA (Qin et al., 2021). Asphalt emissions can also form SOA, and are likely important in LA where the urban land fraction and temperatures are both high (Khare et al., 2020). In addition to organic emission reductions, NO_x emissions from on-road vehicles have decreased. Moreover, NO_x emissions from off-road vehicles have become almost equally important to on-road NO_x emissions in LA (Khare & Gentner, 2018). As total emissions have decreased, ambient levels of most criteria pollutants have decreased, including NO_x, carbon monoxide (CO), and sulfur oxides (SO_x) (US EPA, 2013). However, O_3 in LA has increased in the past decade (US EPA, 2013) because of the nonlinear atmospheric chemistry leading to its formation (Seinfeld & Pandis, 2016; Le et al., 2020). The LA Basin also displays a temperature inversion layer which leads to strong atmospheric stability with a low flow rate out of the Basin. The complex interactions between emissions, meteorology, and chemistry will be investigated in this study.

Predicting air quality using chemical transport models (CTMs) is challenging. Developing a model that best represents the complexity of atmospheric chemistry—particularly SOA formation—in a reasonable computation time involves a tradeoff in chemical detail. Models exist which represent gas-phase and heterogeneous chemistry (e.g., Carter, 2010; Yarwood et al., 2010; Goliff et al., 2013, Keller & Evans, 2019), and researchers have traditionally modeled SOA formation from VOC oxidation (e.g., Odum et al., 1996; Carlton et al., 2010). An active area of research is the oxidation of SVOCs and IVOCs, which likely yield higher SOA than VOCs due to their lower volatility (e.g., Donahue et al., 2011; Murphy et al., 2017; Gentner et al., 2017). It is well-documented that SOA tends to be underpredicted in the Community Multiscale Air Quality (CMAQ) model (Appel et al., 2021) unless an empirical representation of anthropogenic SOA is introduced (Murphy et al., 2017), so a goal of model improvement is to increase SOA mass with improved understanding of sources and physiochemical processes. Representing the correct sources of SOA in a process-based approach is critical for model applications designed to inform control strategies. Recent works have developed new models to represent SOA formation from VCPs (Pennington et al., 2021) and mobile sector IVOCs (Lu et al., 2020), which reduced model SOA bias. The predicted chemistry leading to pollutant formation is highly nonlinear (Seinfeld & Pandis, 2016), and is additionally influenced by emission inventories that typically have high uncertainties (Qin et al., 2021); (Khare & Gentner, 2018). Recent work has improved the estimation of emission rates of VCP VOCs (Seltzer et al., 2021), on-road VOCs, NO_x, PM, and CO (California Air Resources Board, 2018), and on-road IVOCs (Zhao et al., 2016).

Detailed observational data that can be used to constrain model parameters governing chemical transformations is often lacking. While pollutants like O_3, $PM_{2.5}$, and NO_2 are regularly monitored throughout the United States (US EPA, 2013), these sites tend to be sparsely distributed. Components of $PM_{2.5}$ are generally only available on a daily-integrated basis, preventing diagnostic separation of daytime vs nighttime chemistry. Measurements of radical species and specific VOCs are only obtained during field campaigns, which are limited to a small region during a short time duration because they are very expensive to carry out. Even though the lack of in situ data makes it difficult to parameterize or evaluate models, it also underscores the importance of models. Models fill in the spatiotemporal gaps in our measurements and allow us to predict important air quality impacts.

The modeling period in this study covers April 2020, during the strict COVID-19 lockdown regulations in LA. On-road vehicle miles traveled (VMT) declined significantly during this month as many people remained at home (Caltrans, 2020), and this altered the composition of anthropogenic emissions and resulting pollutant levels (Parker et al., 2020). However, this period also experienced several weather patterns that are unusual to springtime months in LA, namely a rainy period and a very hot period. Untangling the relative impacts of decreased emissions versus meteorology is feasible using CTMs.

In the first part of this work, we use the CMAQ model to understand the current air quality of the Los Angeles Basin. Model inputs to CMAQ are developed to represent meteorology and emissions in 2020 and are evaluated against available data. CMAQ model predictions are presented throughout the Basin, while source apportionment studies describe the important sources of emissions. SOA formation in Pasadena is compared to detailed ground-based measurements. In Part 2 of this work, documented in a second article, the sensitivity of pollutants to reduced on-road and VCP emissions are further explored. The relative importance

of emissions and meteorology in dictating O_3 and PM concentrations during the COVID-19 pandemic are also investigated. The simulations investigated in part 2 can represent future emission scenarios and provide insight on helpful policies to mitigate air quality.

2. Methods

2.1 Model Development

The model framework is summarized in Figure 1 and detailed descriptions of each component are described below. CTM inputs include meteorology, emissions, chemical boundary conditions, and grid information. The CTM uses these inputs to predict concentrations which will be compared to hourly or daily observed data throughout the domain and specifically in Pasadena.

2.1.1 Chemical Transport Model

We use CMAQ version 5.3.2 (US EPA, 2020), which is documented and evaluated in Appel et al. (2021). The gas-phase chemical mechanism used here is SAPRC07TIC (Carter, 2010) (Xie et al., 2013), the organic aerosol-phase chemical mechanism is AERO7 (Pye et al., 2013; Pye et al., 2017; Murphy et al., 2017; Xu et al., 2018; Qin et al., 2021), the inorganic aerosol-phase chemical mechanism is ISORROPIA II (Fountoukis & Nenes, 2007), and the aqueous-phase chemical mechanism used is AQCHEM (Fahey et al., 2017). The M3Dry module is the air-surface exchange module used to represent the dry deposition of gas- and particle-phase species (Pleim & Ran, 2011; Appel et al., 2021) and uses the Noah land surface model (Alapaty et al., 2008). The Detailed Emissions Scaling, Isolation, and Diagnostic (DESID) module within CMAQ (Murphy et al., 2021) was used to modify emissions and in our source apportionment sensitivity simulations. The SAPRC07TIC_AE7 chemical mechanism used here was updated to include the emissions and chemistry of VCP species (Pennington et al. (2021) and IVOCs from on-road mobile sources (Lu et al. 2020). The organic aerosol (OA) chemical mechanism is summarized in Fig. S1.

2.1.2 Meteorology

Meteorological simulations are performed using the Weather Research and Forecasting (WRF) Model (Skamarock et al., 2008) version 4.2. Climatological input data are provided from the ERA5 Reanalysis Dataset (Hersbach et al., 2018, p. 5), which contains hourly data on a 0.25° x 0.25° grid at the surface and on 37 pressure levels from 100 to 1 hPa. The WRF configuration uses three nested domains to resample and simulate the meteorological variables from the input resolution to 16-km, 4-km, and then 1-km resolution (Figure 2A). The innermost 1 km x 1 km domain is the region of interest in this study and referred to as the LA domain (Figure 2A, C).

2.1.3 Emissions

On-road vehicles can be separated into two categories, light duty and heavy duty, based on the weight of the vehicle. Light duty vehicles are smaller, tend to be passenger cars, and tend to use gasoline fuel. On the other hand, heavy duty vehicles are larger, tend to be used for transport, and tend to use diesel fuel. These categories are represented separately in the model because there has been historical interest in understanding the class of vehicles and fuels to target for emissions regulations (e.g., Bahreini et al., 2012; Ensberg et al., 2014; Gentner et al., 2017; Lu et al., 2020). Additionally, because of the different uses of these types of vehicles, their driving and therefore emissions patterns differ spatially and temporally.

On-road mobile emissions are represented by the EMission FACtor (EMFAC2017) emissions inventory and model projected to year 2020 (California Air Resources Board, 2018). The projection to year 2020 includes 2020-specific meteorological effects on emission rates. The Emissions Spatial and Temporal Allocator (ESTA) model uses 1 km x 1 km spatial surrogates and California Vehicle Activity Database (CalVAD) temporal surrogates (Ritchie & Tok, 2016) to calculate hourly, gridded emissions on the LA domain. The speciation profiles used in ESTA include the surrogate NMOG (non-methane organic gases), which provides diagnostic information but is not used by the chemistry in CMAQ. To estimate emissions of alkane-like IVOC emissions, the unspeciated fraction of NMOG was used with information from Lu et al. (2020).

EMFAC and ESTA do not capture the effect of COVID-19 on vehicle use, so we modified the on-road emissions to include those changes. The California Performance Measurement System (PeMS) uses in-situ detectors distributed throughout California to measure vehicle usage metrics (Caltrans, 2020). One such metric is vehicle miles traveled (VMT), which measures the miles traveled by different vehicle types, e.g., light and heavy duty vehicles. VMT changed directly in response to COVID-19 policies and human behavior changes, so it can be used to reduce on-road emissions in response to the pandemic (Yang et al., 2021). VMT data were summed for all PeMS monitoring sites in the LA domain, separated into heavy duty and light duty vehicles (Figure 3a-b). VMT January through March (pre-pandemic) was relatively constant. These values were averaged and used as the baseline VMT, represented by the dashed black lines. VMT decreased in March as COVID-19 stay-at-home policies were implemented. VMT reached its lowest value in April and then slowly increased towards the baseline value. All weekly-averaged VMT values were divided by the baseline VMT value to obtain scaling factors which are a proxy for declining vehicle emissions resulting from the pandemic (Figure 3C). The VMT scaling factors are not identical for light duty and heavy duty vehicles, consistent with the rationale for separating these vehicle types. Light duty VMT decreased the most, since the pandemic primarily decreased the use of personal vehicles, with a lesser decrease of industrial transport vehicles' (i.e. heavy-duty vehicles) use.

VCP emissions are predicted using the VCPy model framework (Seltzer et al., 2021). VCPy version 1.1 (Seltzer et al., 2022) was used to calculate VOC emission rates for 2018 over the contiguous United States (CONUS) on a 4 km x 4 km grid, which were re-gridded to 1 km x 1 km to fit the LA domain grid. The year 2018 emissions are assumed to be representative of year 2020 emissions within the range of uncertainty present in VCPy.

Natural emissions are treated in-line in CMAQ using land surface descriptive files generated using the Spatial-Allocator tool (US EPA, 2017/2022). Gas-phase biogenic emissions and particle-phase sea spray emissions are modeled using the Biogenic Emission Inventory System (BEIS) version 3.6.1 (Bash et al., 2016). Particle-phase sea spray emissions are modeled according to the method of Gantt et al. (2015). Wildfire emissions were not included as this time period experienced limited wildfire activity. Lightning NO_x and windblown dust emissions are not turned on in the model. Dust makes up a small fraction of total PM loading. Hayes et al. (2013) showed that in Pasadena, dust makes up only 1.6% of total PM_1 by mass. Natural emissions are the lowest source of PM emissions (CARB, 2020), so windblown dust is a minor contributor to total PM. However, it is possible that muting the dust scheme could cause underestimations of $PM_{2.5}$ and PM_{10}. Previous work suggests that crustal elements, i.e. dust elements, do not have a large impact on modeled ammonium and nitrate concentrations, so

omitting these emissions should not have a large impact on other inorganic aerosol or gas-phase species. Previous work (e.g. Choi et al., 2009) has shown that lightning NO_x is nearly negligible over Southern California.

All other emissions are calculated using the California Air Resources Board (CARB) emissions inventory (CARB, 2020). The emissions inventory includes data from sources including off-road vehicles and equipment, agriculture, oil and gas production, industrial, and other sources. Annual emission rates were calculated for base year 2017 and scaled to year 2020 using the California Emissions Projection Analysis Model (CEPAM) growth and control data (CARB, 2020). The inventory is processed in the Sparse Matrix Operator Kernel Emissions (SMOKE) model version 4.8 (CMAS, 2020) using spatial and temporal surrogates from 2019. SMOKE calculates both gridded area source emissions as well as individual point source emissions, and their sum will be referred to as area+point emissions.

Emission rates and the importance of each emission source vary by pollutant and region. Domain-wide emission rates are given in Figure 4 and the spatial distribution of emissions is given in Fig. S2-7. All anthropogenic emissions peak during midday when people are most active. Biogenic VOC and NO emissions also peak midday corresponding to temperature. In contrast, sea spray emissions peak overnight as temperatures decrease and winds increase. Sea spray emissions are only located in the surf zone along the coastline (Fig. S5). Biogenic sources emit significant VOCs, comparable to those from VCPs. However, VCP emissions are largest over urban areas while biogenic VOC emissions are largest over remote regions (Fig. S7), and so will impact pollutant formation regionally. Area and point sources emit large amounts of all pollutants and comprise a variety of sources (Fig. S8-9). On-road vehicles emit large amounts of CO (Figure 4), but total CO emissions are dominated by off-road vehicles (Fig. S8). On-road vehicles also emit significant NO_x (Figure 4), similar in quantity to the individual area+point sources (i.e., boats, off-road, and trains) given in Fig. S8.

2.1.4 Initial & Boundary Conditions

A nested modeling setup was used to provide the boundary conditions for the Los Angeles Basin. The Los Angeles Basin is represented by the domain shown in Figure 2C, has a resolution of 1 km x 1 km, and is the domain of interest for this project. The initial and boundary conditions for the LA domain were provided by a coarse-resolution CMAQ simulation performed over a larger domain (Figure 2B). The outer domain covering southern and central California has a resolution of 4 km x 4 km and its air quality was simulated using the WRF and CMAQ scenarios described in Sections 2.1.1-2.1.2. The emissions for this domain match the emissions described in Jiang et al. (2021). Publicly-available seasonal average hemispheric CMAQ output was used as initial and boundary conditions for the California domain (Hogrefe et al., 2021). The CMAQ predictions from the coarse-resolution California domain were used as initial and boundary conditions for the inner, finer-resolution LA domain.

2.2 Observational Data

Observational data throughout the modeling domain are provided by the EPA AQS monitoring system (US EPA, 2013). These sites include measurements of O_3, CO, NO, NO_2, NO_y, SO_2, $PM_{2.5}$, PM_{10}, temperature, relative humidity, wind speed, and wind direction (not all sites contain all species at all times) and their locations are shown in Figure 2B-C. In addition, gas- and aerosol-phase measurements were collected concurrent to the modeling period in

Pasadena at Caltech. The Caltech air quality system (CITAQS) measures O_3, CO, NO, NO_2, NO_y, SO_2, and $PM_{2.5}$ (Parker et al., 2020).

Measurements of PM_1 (fine PM with diameters less than 1 μm) and its components (organic, NH_4, NO_3, SO_4, and Cl) were performed using an Aerodyne high resolution time-of-flight aerosol mass spectrometer (HR-ToF-AMS) as described in Schulze et al. (submitted, 2022). Briefly, the AMS measures submicron, non-refractory PM_1 (NR-PM_1) at high time resolution. During the 2020 measurement campaign, the AMS isokinetically sampled air from a stainless-steel line downstream of a 2.5 μm cut diameter Teflon-coated cyclone mounted on the roof of the Linde Laboratory at Caltech. Approximately 6 m of stainless steel tubing connected the cyclone to the inlet of the HR-ToF-AMS. Standard methods were used to correct the data for gas-phase interferences and composition-dependent collection efficiencies (Middlebrook et al., 2012). Daily detection limits for aerosol chemical classes were calculated as three times the standard deviation of 30-minute blank measurements made with a high-efficiency particle arrestance (HEPA) filter. Daily detection limits for OA ranged from ~0.1-0.3 μg m^{-3}. The ionization efficiency of nitrate and relative ionization efficiency of ammonium were calibrated weekly using 350 nm ammonium nitrate particles size selected with a differential mobility analyzer.

Positive matrix factorization (PMF) was applied to the OA mass spectral datasets to gain insight into OA sources. PMF results presented here were taken from a larger analysis of data collected in 2020 (April 8 – July 19, 2020). A detailed description of PMF solution selection is provided in Schulze et al. (2022). A total of five factors, corresponding to less-oxidized oxygenated OA (LO-OOA), more-oxidized oxygenated OA (MO-OOA), hydrocarbon-like OA (HOA), cooking-influenced OA (CIOA), and an organic-nitrate influenced LO-OOA (LO-OOA-ON), were extracted from the OA dataset. Factors were identified using correlations with known tracers and comparisons of mass spectral and diurnal profiles to those extracted previously in Los Angeles (Hayes et al., 2013) and other urban areas (Hu et al., 2016; J. Xu et al., 2016). For comparisons with model predictions, we combine the HOA and CIOA factors as primary OA (POA), though we note that SOA formed from low-volatility species may appear spectrally similar to HOA (Lambe et al., 2012), as discussed in Schulze et al. (2022).

Multiple statistics are used to compare modeled data to observed data. These are mean bias (MB), normalized mean bias (NMB), root mean square error (RMSE), and r^2 (the square of the Pearson correlation coefficient), defined below. In these equations, M is modeled data, O is observed data, $\bar{M}$ is the mean of the modeled data, $\bar{O}$ is the mean of the observed data, and N is the number of data points.

$$MB = \frac{1}{N}\sum_1^N (M - O) \quad (1)$$

$$Fractional\ NMB = \frac{\sum_1^N (M-O)}{\sum_1^N O} \quad (2)$$

$$NMB = \frac{\sum_1^N (M-O)}{\sum_1^N O} x100\% \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_1^N (M - O)^2} \quad (4)$$

$$r^2 = \frac{\left(\sum_1^N (M-\bar{M})(O-\bar{O})\right)^2}{\sum_1^N (M-\bar{M})^2 \sum_1^N (O-\bar{O})^2} \quad (5)$$

3. Results & Discussion

3.1 Evaluation of CTM Inputs

3.1.1 Meteorology

The WRF predictions are compared to the AQS observations and the model performs very well in predicting temperature. The NMB values of temperature, relative humidity, wind speed, and wind direction at all AQS sites are calculated in the LA domain (Figure 5), and statistics are averaged using all site data in Table S1. Temperature is predicted well, with very low bias (NMB = 3.8%) and low scatter (r^2 = 0.97). Relative humidity is moderately well-predicted, with low scatter (r^2 = 0.81) but nonnegligible bias (NMB = -21.3%). Errors in relative humidity will affect the water content of aerosols and the resulting partitioning of aqueous aerosol, and the concentrations of other inorganic aerosol components like ammonium, nitrate, and chloride.

Wind speed and direction tend not to be predicted well, with high bias and high scatter, but the error is highly variable between sites (Figure 5). Wind speed and direction error will potentially affect the transport between grid cells, and their impact on modeled pollutant concentrations is investigated in Section 3.2. To understand the source of wind speed error, the NMB was quantified in all 3 modeling domains (Fig. S10). Wind speed did not improve appreciably as the model resolution increased, and the spatial distribution of error remained consistent. This suggests that the model error lies with the input reanalysis data, and less with the model configuration. This further suggests that to improve model simulations, new reanalysis data should be used or observational nudging should be engaged when running WRF. However, using new reanalysis data may introduce error to other meteorological fields, whereas temperature is well-predicted by this model setup.

The domain-wide statistics (Table S1) capture data over a long time period and over sites with different meteorology, so the error at individual sites must be investigated when making site-specific comparisons. Despite the range of sites contained in these statistics, temperature is well-predicted. This is critical, as temperature has a substantial impact on atmospheric chemistry and reaction rates.

3.1.2 Coarse-Resolution Simulation Results

California coarse-resolution CMAQ simulation results provide the lateral chemical boundary conditions for the inner LA domain. Predicted pollutant concentrations from the coarse-resolution California simulation are compared to EPA AQS monitoring site data in Table 1. O_3 is well-predicted based on its low MB, NMB, and RMSE. CO, NO_x, and PM_{10} are all underpredicted (MB and NMB) with moderately high scatter (RMSE and r^2), while $PM_{2.5}$ is overpredicted. SO_2 is greatly overpredicted (MB and NMB). The accuracy of the region covering the Los Angeles Basin is of particular importance since that region will provide the initial and boundary conditions for the fine-resolution domain. Those results are compared to AQS measurements (Table 1) and demonstrate some different behaviors than the results of the full domain. NO_x is slightly better predicted, while still underestimated, but O_3 is now underpredicted and less accurate. Average $PM_{2.5}$ mass increases substantially, as expected due to the higher air pollution in LA compared to other regions in California. $PM_{2.5}$ also becomes greatly overpredicted in the model (MB and NMB) and will be considered when evaluating the results of

the fine-resolution simulation. The model bias remains approximately consistent for CO, SO_2, and PM_{10}.

3.2 Evaluation of Fine-Resolution Model Predictions

Model predictions are compared to EPA AQS measurements at 44 sites in the domain (Figure 5-6, Table S2). O_3 has low NMB at all sites (NMB = 10.2%) despite high scatter (r^2 = 0.30), and has the correct spatial distribution despite poorly predicted NO_x. NO, NO_2, and CO prediction error can be positive or negative depending on location. PM measurements are limited in the domain and will be investigated further in Sections 3.3-3.4. Domain-wide statistics are provided in Table S2. NO_x and VOC concentrations are highest in polluted and high-emitting regions, and O_3 titration by freshly emitted NO results in O_3 concentrations that are lower in the urban core than in surrounding areas. Fine PM (PM_1 and $PM_{2.5}$) are highest in the urban center, while PM_{10} concentrations increase over the ocean due to sea spray aerosol. Because of the potential overprediction of sea spray emissions, it is possible that PM_{10} is overpredicted. POA is highest over high-emission regions, while SOA is highest over downwind regions, displaying the importance of chemical aging during transport.

The impact of transport on modeled pollutant concentration was investigated by performing a sensitivity simulation with perturbed wind speed. The WRF wind speed was reduced by 25% (i.e., scaled by a factor of 0.75) in an effort to correct for some of the wind speed bias (Figure 5). A reduction of 25% was chosen to represent the correction required to bring modeled wind speed into the range of observed wind speed, as represented by the values in Table S1. The results are presented below in Figure 7 and Figure 8 and can be compared to the base case wind speed bias in Figure 5. Wind speed improved appreciably in response to the 25% reduction in their values throughout the domain. In spite of improved wind speed, modeled O_3 and $PM_{2.5}$ did not improve. This suggests that wind speed does not have a large effect on modeled pollutant concentrations, and bias in those concentrations is more likely caused by errors in modeled chemistry and/or emissions.

3.3 Evaluation of Aerosol Chemistry by Ground-Based Observations in Pasadena

Modeled PM_1 is underestimated due primarily to a large underestimation of OA. PM_1 mass and composition in Pasadena measured by AMS and predicted by CMAQ are compared in Figure 9. All predicted inorganic component (SO_4, NO_3, NH_4, Cl) concentrations are smaller by mass than observed values. Of note, PM_1 NO_3 is nearly well-predicted (Table S3) despite gaseous NO_x underpredictions (Table S4). The model additionally predicts “other” inorganic PM_1, which includes EC, soil, and crustal elements which is not measured at the Pasadena ground site. The overall PM_1 bias (NMB = -49.1%) is caused by the large underprediction of OA (NMB = -63.0%). POA is well-predicted (Figure 10A) and the diurnal trend matches predictions except during late night and early morning hours (Figure 10B). SOA is significantly underpredicted (Figure 10A) and has an accurate diurnal trend except during early morning (Figure 10B). During the day when emissions and photochemistry are at maximum, measured and observed SOA peaks. SOA decreases in the evening as emissions decrease. Despite lower photochemistry and emissions, SOA (and other pollutant levels) remain high at night due to low planetary boundary layer (PBL) height. The accurate representation of POA and poorer representation of SOA suggests that OA is better represented near source regions and diminishes in its effectiveness with distance from sources.

Detailed model speciation and source apportionment can be used to understand the major sources of OA precursors in Pasadena and the error in SOA predictions. Measured POA comprises cooking-influenced OA (CIOA) and hydrocarbon-like OA (HOA). CIOA peaks overnight due to the PBL height dilution effect during the day, while HOA remains high throughout the day due to high local primary emissions sources (Figure 10). Measured SOA comprises more-oxidized oxygenated OA (MO_OOA), less-oxidized oxygenated OA (LO_OOA), and LO_OOA associated with organic nitrates (LO_OOA_ON). MO_OOA is consistently one of the largest OA components, with little diurnal variation. LO_OOA is the largest SOA component and has a sharp peak midday, consistent with higher oxidation rates during midday. Modeled alkane-like IVOCs have a similar high peak around midday, although of a smaller magnitude (Figure 10D). LO_OOA_ON have a small midday peak suggesting some photochemical production, but the largest contribution from LO_OOA_ON is overnight. This could be due in part to the PBL effect, and may also be due to overnight NO_3 chemistry producing organic nitrates. This is consistent with the overnight peak of modeled organic nitrates (Figure 10D) and terpene- and glyoxal-derived SOA (Fig. S11), which are biogenic in nature. All other modeled SOA species except oligomers have low overnight mass and peak at midday, but their magnitudes are small which are likely a source of error in the CMAQ chemical mechanism. CMAQ lacks species which are behaving like LO-OOA, and the inclusion of additional SOA precursor species could improve SOA predictions (Pye et al., 2022). One potential source of error could be too-low yields of species that already exist in the model, such as aromatics, which have not been corrected for gas-phase wall losses (Zhang et al., 2014). Additional sources of error could include missing emissions, such as from asphalt which would peak during midday when temperatures are highest, consistent with LO-OOA.

3.4 LA Basin Source Apportionment

The impact of removing each emission source on O_3 is presented in Figure 11 and these changes can be understood by investigating the changes in NO_x, VOC, and OH (Fig. S13-15). The impact of sea spray is small because sea spray emits only particles, so those results are presented in Fig. S12. O_3 decreased everywhere in response to the removal of VCP and biogenic emissions. VCPs only emit VOCs, and so the elimination of VCP emissions leads to VOC decreases everywhere. In response, OH and NO_x concentrations increase, and the importance of transport and secondary aging processes is evident by the downwind location of most of the OH increase. The O_3 decrease resulting from VOC decreases is consistent with NO_x-saturated behavior, which has typically described highly-polluted urban areas. The removal of biogenic emissions has a similar response, as biogenic sources mainly emit VOCs. One exception lies in that biogenic sources also emit NO, so the VOC:NO_x ratio changes less and thus biogenics have a smaller impact on O_3 change than VCPs do. In both cases of VCP and biogenic emissions removal, the outer regions display less sensitivity as a reduction in VOCs results in a near-zero change in O_3.

On-road vehicles and area+point sources emit NO_x, VOC, particles, and other inorganic gas-phase species. When these emission sources are removed, VOC and NO_x concentrations decrease everywhere. In the urban core where VOC and NO_x concentrations are high, OH and O_3 increase in response to the combined on-road VOC and NO_x reductions. This is characteristic of the effect of large NO_x relative to VOC (Figure 4) reductions under NO_x-saturated conditions. In contrast, the outer regions display behavior closer to NO_x-limited behavior, where VOC and NO_x reductions result in OH and O_3 reductions. The reductions are small, suggesting that O_3 is not

sensitive to emission reductions in these regions. The elimination of area+point source emissions has a similar impact on O_3. OH and O_3 increase in the urban core, with a decrease of OH and O_3 in the outer regions. The importance of ships and the Long Beach Port is evident, but it is likely that shipping emissions of NO_x are overestimated relative to other area source emissions (Fig. S8) and so this impact may be overstated in these results.

$PM_{2.5}$ concentrations decrease everywhere in response to emission reductions (Figure 12). VCPs and biogenic sources emit only gas-phase species, so PM is formed exclusively via secondary processes. Biogenic PM is formed mostly over high emission areas like mountains, while VCP-derived PM is found in downwind regions, highlighting the importance of secondary formation during transport, similar to O_3 formation (Figure 11). PM from on-road and area+point sources is predominantly emitted directly because most of the impact to $PM_{2.5}$ is located in high emission regions. This is in spite of increased oxidation capacity in the high-emission regions (Fig. S13). So if the emissions are removed entirely, as in this study, $PM_{2.5}$ will decrease. However, if the emissions were not entirely removed, the increased OH and the nonlinearity of atmospheric chemistry could lead to increased PM. Sea spray particles are reduced along the coastline where waves break (Fig. S16).

Different species impact the $PM_{2.5}$ change from each emission source (Fig. S17). On-road sources primarily decrease the NO_3 and NH_4 components of $PM_{2.5}$, both by direct emission and emissions of gas-phase NO_x. The reduction of on-road VOCs has relatively little impact on the organic fraction of $PM_{2.5}$. Area+point emissions also reduce $PM_{2.5}$ NO_3 and NH_4, plus other direct emissions like POA and elemental carbon (EC). VCPs and biogenic sources emit only VOCs, so they impact mostly the SOA fraction of $PM_{2.5}$. The reduction of VOCs leads to increases in OH and NO_x and thus increases of $PM_{2.5}$ NO_3 and NH_4.

SOA decreases almost everywhere in response to the removal of emission sources but can increase in some high-emission regions (Figure 13). The SOA change from VCPs is downwind of the main emission regions. Biogenic SOA decrease is located mostly in remote, mountainous regions. Downwind SOA decreases when all on-road emissions are removed, but SOA in the downtown LA region increases. This occurs because it is NO_x-saturated and has increased OH concentrations (Fig. S13), which increases rates of VOC oxidation and therefore SOA formation. The SOA decrease from the removal of area+point emission sources is more widely distributed than the emissions themselves (Fig. S2-7), displaying the importance of SOA formation during transport.

SOA speciation varies throughout the domain and is dependent on location-specific emissions and meteorology (Fig. S18). The largest components of SOA are derived from alkane-like IVOCs, organic nitrates, and monoterpenes. Alkane-like IVOC concentrations are highest downwind of high-emissions regions, demonstrating the importance of secondary formation during transport. Organic nitrate concentrations are highest over high-emission areas where VOC and NO_x concentrations are largest. Monoterpene concentrations are more uniform and have both anthropogenic (i.e., VCP) and biogenic sources. Little SOA throughout the domain is formed from siloxanes, sesquiterpenes, or cloud processing. Biogenic SOA is primarily derived from sesquiterpenes, monoterpenes, and isoprene, and these aerosol species dominate over mountainous and remote areas in the outer regions of the domain.

SOA formation chemistry can be further understood by investigating the source apportionment of SOA components in Pasadena. The impact of removing each emission source

on each modeled SOA component is given in Table *2*. The main component of SOA—alkane-like IVOCs—originates particularly from VCPs and area+point emission sources. Alkane-like IVOCs are emitted from VCPs as low-volatility gases, while they are evaporated and oxidized POA from area+point emission sources. Organic nitrates have important contributions from VCPs and area+point emission sources, but are mostly formed from biogenic precursors. Despite VCP, biogenic, and area+point emission sources being highest during daytime, organic nitrates peak overnight due to nighttime NO_3 chemistry. In general, our modeling suggests SOA in LA is mostly driven by VCP, area, and point emission sources.

4. Conclusions

This study presents a new model framework to simulate air quality in Los Angeles. Past modeling studies of LA focus on 2010 to overlap with the CalNex campaign, and few exist which focus on SOA sources and speciation. We developed state-of-the-science inputs of meteorology, emissions, and boundary conditions, and show that these inputs are comparable to observations. Emissions are separated into 3 anthropogenic categories—VCP, on-road, and area+point—and 2 natural categories—gases and sea spray—allowing for source apportionment studies.

The model is set up for April 2020 and the results are compared to observations, aiming to better understand the chemistry leading to pollutant formation. Temperature and O_3 are very well-predicted, but NO_x and PM are underpredicted. In particular, OA is underpredicted in Pasadena when compared to AMS measurements. While POA is well-predicted, SOA is greatly under-predicted. The main components of modeled SOA are alkane-like IVOCs and organic nitrates, while other categories of SOA are likely underpredicted; for example, oxygenated IVOCs which have not been well-classified in laboratory settings.

This study stresses that improved model predictions will require updated chemistry and emissions. The chemistry of SVOCs is not well-understood, and better representations should be included in CMAQ as they are developed. SVOCs are also typically not represented in emission inventories, and while the VCP inventory used here utilizes new SVOC speciation profiles, the other categories of emissions did not specifically study SVOCs. The chemistry of oxygenated species has not been extensively studied, and should be focused on in future work due to the prevalence of oxygenated emissions and atmospheric constituents (Pennington et al., 2021). Some emissions from anthropogenic sources are likely underpredicted. For example, boats are estimated to emit more NO_x than off-road sources, but off-road sources should likely be the main area source of NO_x (Khare & Gentner, 2018). Also, many forms of asphalt emissions are not included in VCP or area sources, but likely will contribute significant SOA and therefore reduce modeled SOA bias (Khare & Gentner, 2018).

The source apportionment results convey important qualities about the VOC-NO_x regime of the LA atmosphere. The urban core of LA demonstrates NO_x-saturated behavior: NO_x reductions lead to O_3 increase, while VOC reductions lead to O_3 decrease. Outside of the urban core, O_3 decreases in response to any level of either NO_x or VOC removal, suggesting a regime that is less NO_x-saturated than the urban region, such as a regime lying close to the O_3-NO_x-VOC ridgeline in the VOC-sensitive regime (Seinfeld & Pandis, 2016). Reducing O_3 is a consistent goal for policymakers, and this work shows that O_3 in Los Angeles is reduced by the removal of VOCs. NO_x emission decreases remain important, as these decreases will move the Basin from a NO_x-saturated regime closer to a NO_x-sensitive regime. However, NO_x reductions

without concurrent or larger reductions in VOC concentrations will make O_3 pollution worse until the NO_x-sensitive regime is reached. VCPs emit the highest amount of VOCs from anthropogenic activities and thus may be particularly effective to target for reducing O_3. It is also important to consider the spatial distribution of emissions and reduction policies. Reducing NO_x and/or VOC emissions in the outer regions of the domain will have a lesser impact than reductions in the urban core, or may have an opposite effect, as demonstrated in this study. The increased oxidative capacity of the NO_x-saturated regions also has an impact on SOA formation and the formation of secondary inorganic components of PM. Focusing on emissions in the urban core is critical and will affect downwind regions. It should be noted that this study was performed in the spring season, which is not peak ozone season. Thus, results may differ in the summer months and further studies should investigate this period.

In Part 2 (Pennington et al., in prep), the new model framework is used to investigate future emission scenarios involving VCP and on-road vehicle emissions during the 2020 lockdown of the pandemic. VCP emissions have been quantified in multiple studies (i.e., Seltzer, Pennington, et al., 2021; McDonald et al., 2018), but none of these studies have investigated the implications of future VCP emissions. We reduce VCP emissions to investigate the impact on O_3, NO_x, PM, and SOA speciation. Additionally, we run the model in a "non-COVID" scenario, where on-road emissions are represented without COVID-induced VMT reductions. In this way, the impact of emissions versus meteorology on 2020 air quality can be distinguished. Understanding these possible outcomes can shape informed policy decisions.

Data Availability

These will be posted on Caltech's permanent data site.

- CMAQ source code
- WRF namelist files
- CA4km emission file
- All LA1km emission files (VCP, LDV, HDV, area, point)

Author Contributions

EAP, YW, and JHS designed and led the research project. EAP performed all model simulations and drafted the paper. EAP, YW, and JHS analyzed the data. BCS collected AMS data and performed PMF analysis. KMS provided VCP emissions. JY provided VMT data. ZJ and BZ provided emissions for the California 4 km x 4 km domain. MV provided the CARB emissions inventory and all SMOKE input files. DC provided the EMFAC emissions inventory. BNM and HOTP participated in useful research discussions and mentored EAP. CMK and RXW collected AMS data. All authors participated in useful research discussions and revised the paper.

Disclaimer

The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Acknowledgements

The authors would like to thank Leonardo Ramirez for his guidance on the CARB emission inventories, and Han Kim for introducing and explaining useful Python analysis tools. We're also grateful to John Crounse and Harrison Parker for managing the CITAQS station and collecting the CITAQS data used in this study. EAP and JHS acknowledge funding support from

Samsung Global Research Outreach Program. YW and JHS acknowledge funding support from the National Science Foundation (AGS-2103714). We also acknowledge high-performance computing support from NASA Pleiades.

Competing interests

YW is a member of the editorial board of Atmospheric Chemistry and Physics.

References

Alapaty, K., Niyogi, D., Chen, F., Pyle, P., Chandrasekar, A., & Seaman, N. (2008). Development of the Flux-Adjusting Surface Data Assimilation System for Mesoscale Models. *Journal of Applied Meteorology and Climatology*, *47*(9), 2331–2350. https://doi.org/10.1175/2008JAMC1831.1

Appel, K. W., Bash, J. O., Fahey, K. M., Foley, K. M., Gilliam, R. C., Hogrefe, C., Hutzell, W. T., Kang, D., Mathur, R., Murphy, B. N., Napelenok, S. L., Nolte, C. G., Pleim, J. E., Pouliot, G. A., Pye, H. O. T., Ran, L., Roselle, S. J., Sarwar, G., Schwede, D. B., … Wong, D. C. (2021). The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: System updates and evaluation. *Geoscientific Model Development*, *14*(5), 2867–2897. https://doi.org/10.5194/gmd-14-2867-2021

Bahreini, R., Middlebrook, A. M., Gouw, J. A. de, Warneke, C., Trainer, M., Brock, C. A., Stark, H., Brown, S. S., Dube, W. P., Gilman, J. B., Hall, K., Holloway, J. S., Kuster, W. C., Perring, A. E., Prevot, A. S. H., Schwarz, J. P., Spackman, J. R., Szidat, S., Wagner, N. L., … Parrish, D. D. (2012). Gasoline emissions dominate over diesel in formation of secondary organic aerosol mass. *Geophysical Research Letters*, *39*(6). https://doi.org/10.1029/2011GL050718

Bash, J. O., Baker, K. R., & Beaver, M. R. (2016). Evaluation of improved land use and canopy representation in BEIS v3.61 with biogenic VOC measurements in California. *Geoscientific Model Development*, *9*(6), 2191–2207. https://doi.org/10.5194/gmd-9-2191-2016

California Air Resources Board. (2018). *EMFAC2017 Volume III Technical Documentation: V1.0.2*. https://ww3.arb.ca.gov/msei/downloads/emfac2017-volume-iii-technical-documentation.pdf

Caltrans. (2020). *Caltrans PeMS*. https://pems.dot.ca.gov/

CARB. (2020). *Criteria Pollutant Emission Inventory Data | California Air Resources Board*. https://ww2.arb.ca.gov/criteria-pollutant-emission-inventory-data

Carlton, A. G., Bhave, P. V., Napelenok, S. L., Edney, E. O., Sarwar, G., Pinder, R. W., Pouliot, G. A., & Houyoux, M. (2010). Model Representation of Secondary Organic Aerosol in CMAQv4.7. *Environmental Science & Technology*, *44*(22), 8553–8560. https://doi.org/10.1021/es100636q

Carter, W. P. L. (2010). Development of the SAPRC-07 chemical mechanism. *Atmospheric Environment*, *44*(40), 5324–5335. https://doi.org/10.1016/j.atmosenv.2010.01.026

Choi, Y., Kim, J., Eldering, A., Osterman, G., Yung, Y. L., Gu, Y., & Liou, K. N. (2009). Lightning and anthropogenic NOx sources over the United States and the western North Atlantic Ocean: Impact on OLR and radiative effects. *Geophysical Research Letters*, *36*(17). https://doi.org/10.1029/2009GL039381

CMAS. (2020). *SMOKE (Sparse Matrix Operator Kerner Emissions) Modeling System*. CMAS: Community Modeling and Analysis System. https://www.cmascenter.org/smoke/index.cfm

Donahue, N. M., Epstein, S. A., Pandis, S. N., & Robinson, A. L. (2011). A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics. *Atmospheric Chemistry and Physics*, *11*(7), 3303–3318. https://doi.org/10.5194/acp-11-3303-2011

Ensberg, J. J., Craven, J. S., Metcalf, A. R., Allan, J. D., Angevine, W. M., Bahreini, R., Brioude, J., Cai, C., Coe, H., de Gouw, J. A., Ellis, R. A., Flynn, J. H., Haman, C. L., Hayes, P. L., Jimenez, J. L., Lefer, B. L., Middlebrook, A. M., Murphy, J. G., Neuman, J. A., … Seinfeld, J. H. (2013). Inorganic and black carbon aerosols in the Los Angeles Basin during CalNex. *Journal of Geophysical Research: Atmospheres*, *118*(4), 1777–1803. https://doi.org/10.1029/2012JD018136

Ensberg, J. J., Hayes, P. L., Jimenez, J. L., Gilman, J. B., Kuster, W. C., de Gouw, J. A., Holloway, J. S., Gordon, T. D., Jathar, S., Robinson, A. L., & Seinfeld, J. H. (2014). Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation. *Atmospheric Chemistry and Physics*, *14*(5), 2383–2397. https://doi.org/10.5194/acp-14-2383-2014

Fahey, K. M., Carlton, A. G., Pye, H. O. T., Baek, J., Hutzell, W. T., Stanier, C. O., Baker, K. R., Appel, K. W., Jaoui, M., & Offenberg, J. H. (2017). A framework for expanding aqueous chemistry in the Community Multiscale Air Quality (CMAQ) model version 5.1. *Geoscientific Model Development*, *10*(4), 1587–1605. https://doi.org/10.5194/gmd-10-1587-2017

Fountoukis, C., & Nenes, A. (2007). ISORROPIA II: A computationally efficient thermodynamic equilibrium model for K^+–Ca^{2+}–Mg^{2+}–NH_4^+–Na^+–SO_4^{2-}–NO_3^-–Cl^-–H_2O aerosols. *Atmospheric Chemistry and Physics*, *7*(17), 4639–4659. https://doi.org/10.5194/acp-7-4639-2007

Gantt, B., Kelly, J. T., & Bash, J. O. (2015). Updating sea spray aerosol emissions in the Community Multiscale Air Quality (CMAQ) model version 5.0.2. *Geoscientific Model Development*, *8*(11), 3733–3746. https://doi.org/10.5194/gmd-8-3733-2015

Gentner, D. R., Jathar, S. H., Gordon, T. D., Bahreini, R., Day, D. A., El Haddad, I., Hayes, P. L., Pieber, S. M., Platt, S. M., de Gouw, J., Goldstein, A. H., Harley, R. A., Jimenez, J. L., Prévôt, A. S. H., & Robinson, A. L. (2017). Review of Urban Secondary Organic Aerosol Formation from Gasoline and Diesel Motor Vehicle Emissions. *Environmental Science & Technology*, *51*(3), 1074–1093. https://doi.org/10.1021/acs.est.6b04509

Goliff, W. S., Stockwell, W. R., & Lawson, C. V. (2013). The regional atmospheric chemistry mechanism, version 2. *Atmospheric Environment*, *68*, 174–185. https://doi.org/10.1016/j.atmosenv.2012.11.038

Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W., Toohey, D. W., Flynn, J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglück, B., Taylor, J. W., Allan, J. D., Holloway, J. S., Gilman, J. B., Kuster, W. C., Gouw, J. A. de, … Jimenez, J. L. (2013). Organic aerosol composition and sources in Pasadena, California, during the 2010 CalNex campaign. *Journal of Geophysical Research: Atmospheres*, *118*(16), 9233–9257. https://doi.org/10.1002/jgrd.50530

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., & Thépaut, J.-N. (2018). *ERA5 hourly data on pressure levels from 1979 to present* [Computer software]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 10.24381/cds.bd0915c6

Hogrefe, C., Gilliam, R., Mathur, R., Henderson, B. H., Sarwar, G., Appel, K. W., Pouliot, G., Willison, J., Miller, R., Vukovich, J., Eyth, A., Talgo, K., Allen, C., & Foley, K. (2021). *CMAQv5.3.2 ozone simulations over the Northern Hemisphere: Model performance and sensitivity to model configuration*. 20th Annual CMAS Conference. https://drive.google.com/drive/folders/1A1ZzJE1t7OgwSezQNvy3rt9aATnXA0k2

Hu, W., Hu, M., Hu, W., Jimenez, J. L., Yuan, B., Chen, W., Wang, M., Wu, Y., Chen, C., Wang, Z., Peng, J., Zeng, L., & Shao, M. (2016). Chemical composition, sources, and aging process of submicron aerosols in Beijing: Contrast between summer and winter. *Journal of Geophysical Research: Atmospheres*, *121*(4), 1955–1977. https://doi.org/10.1002/2015JD024020

Hyslop, N. P. (2009). Impaired visibility: The air pollution people see. *Atmospheric Environment*, *43*(1), 182–195. https://doi.org/10.1016/j.atmosenv.2008.09.067

Intergovernmental Panel on Climate Change (Ed.). (2014). Anthropogenic and Natural Radiative Forcing. In *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 659–740). Cambridge University Press. https://doi.org/10.1017/CBO9781107415324.018

Jiang, Z., Shi, H., Zhao, B., Gu, Y., Zhu, Y., Miyazaki, K., Lu, X., Zhang, Y., Bowman, K. W., Sekiya, T., & Liou, K.-N. (2021). Modeling the impact of COVID-19 on air quality in southern California: Implications for future control policies. *Atmospheric Chemistry and Physics*, *21*(11), 8693–8708. https://doi.org/10.5194/acp-21-8693-2021

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., … Worsnop, D. R. (2009). Evolution of Organic Aerosols in the Atmosphere. *Science*, *326*(5959), 1525–1529. https://doi.org/10.1126/science.1180353

Keller, C. A., & Evans, M. J. (2019). Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10. *Geoscientific Model Development Discussions*, 1209–1225.

Khare, P., & Gentner, D. R. (2018). Considering the future of anthropogenic gas-phase organic compound emissions and the increasing influence of non-combustion sources on urban air quality. *Atmospheric Chemistry and Physics*, *18*(8), 5391–5413. https://doi.org/10.5194/acp-18-5391-2018

Khare, P., Machesky, J., Soto, R., He, M., Presto, A. A., & Gentner, D. R. (2020). Asphalt-related emissions are a major missing nontraditional source of secondary organic aerosol precursors. *Science Advances*, *6*(36), eabb9785. https://doi.org/10.1126/sciadv.abb9785

Lambe, A. T., Onasch, T. B., Croasdale, D. R., Wright, J. P., Martin, A. T., Franklin, J. P., Massoli, P., Kroll, J. H., Canagaratna, M. R., Brune, W. H., Worsnop, D. R., & Davidovits, P. (2012). Transitions from Functionalization to Fragmentation Reactions of Laboratory Secondary Organic Aerosol (SOA) Generated from the OH Oxidation of Alkane Precursors. *Environmental Science & Technology*, *46*(10), 5430–5437. https://doi.org/10.1021/es300274t

Le, T., Wang, Y., Liu, L., Yang, J., Yung, Y. L., Li, G., & Seinfeld, J. H. (2020). Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China. *Science*, *369*(6504), 702–706. https://doi.org/10.1126/science.abb7431

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M. A., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., … Ezzati, M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8

Lu, Q., Murphy, B. N., Qin, M., Adams, P. J., Zhao, Y., Pye, H. O. T., Efstathiou, C., Allen, C., & Robinson, A. L. (2020). Simulation of organic aerosol formation during the CalNex study: Updated mobile emissions and secondary organic aerosol parameterization for intermediate-volatility organic compounds. *Atmospheric Chemistry and Physics*, *20*(7), 4313–4332. https://doi.org/10.5194/acp-20-4313-2020

McDonald, B. C., Gouw, J. A. de, Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B., & Trainer, M. (2018). Volatile chemical products emerging as largest petrochemical source of urban organic emissions. *Science*, *359*(6377), 760–764. https://doi.org/10.1126/science.aaq0524

Middlebrook, A. M., Bahreini, R., Jimenez, J. L., & Canagaratna, M. R. (2012). Evaluation of Composition-Dependent Collection Efficiencies for the Aerodyne Aerosol Mass Spectrometer using Field Data. *Aerosol Science and Technology*, *46*(3), 258–271. https://doi.org/10.1080/02786826.2011.620041

Murphy, B. N., Nolte, C. G., Sidi, F., Bash, J. O., Appel, K. W., Jang, C., Kang, D., Kelly, J., Mathur, R., Napelenok, S., Pouliot, G., & Pye, H. O. T. (2021). The Detailed Emissions Scaling, Isolation, and Diagnostic (DESID) module in the Community Multiscale Air Quality (CMAQ) modeling system version 5.3.2. *Geoscientific Model Development*, *14*(6), 3407–3420. https://doi.org/10.5194/gmd-14-3407-2021

Murphy, B. N., Woody, M. C., Jimenez, J. L., Carlton, A. M. G., Hayes, P. L., Liu, S., Ng, N. L., Russell, L. M., Setyan, A., Xu, L., Young, J., Zaveri, R. A., Zhang, Q., & Pye, H. O. T. (2017). Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: Impacts on source strength and partitioning. *Atmospheric Chemistry and Physics*, *17*(18), 11107–11133. https://doi.org/10.5194/acp-17-11107-2017

Nuvolone, D., Petri, D., & Voller, F. (2018). The effects of ozone on human health. *Environmental Science and Pollution Research*, *25*(9), 8074–8088. https://doi.org/10.1007/s11356-017-9239-3

Odum, J. R., Hoffmann, T., Bowman, F., Collins, D., Flagan, R. C., & Seinfeld, J. H. (1996). Gas/Particle Partitioning and Secondary Organic Aerosol Yields. *Environmental Science & Technology*, *30*(8), 2580–2585. https://doi.org/10.1021/es950943+

Parker, H. A., Hasheminassab, S., Crounse, J. D., Roehl, C. M., & Wennberg, P. O. (2020). Impacts of Traffic Reductions Associated With COVID-19 on Southern California Air Quality. *Geophysical Research Letters*, *47*(23), e2020GL090164. https://doi.org/10.1029/2020GL090164

Pennington, E. A., Seltzer, K. M., Murphy, B. N., Qin, M., Seinfeld, J. H., & Pye, H. O. T. (2021). Modeling secondary organic aerosol formation from volatile chemical products. *Atmospheric Chemistry and Physics*, *21*(24), 18247–18261. https://doi.org/10.5194/acp-21-18247-2021

739 Pleim, J., & Ran, L. (2011). Surface Flux Modeling for Air Quality Applications. *Atmosphere*,
740 *2*(3), Article 3. https://doi.org/10.3390/atmos2030271
741 Pye, H. O. T., Murphy, B. N., Xu, L., Ng, N. L., Carlton, A. G., Guo, H., Weber, R., Vasilakos,
742 P., Appel, K. W., Budisulistiorini, S. H., Surratt, J. D., Nenes, A., Hu, W., Jimenez, J. L.,
743 Isaacman-VanWertz, G., Misztal, P. K., & Goldstein, A. H. (2017). On the implications
744 of aerosol liquid water and phase separation for organic aerosol mass. *Atmospheric*
745 *Chemistry and Physics*, *17*(1), 343–369. https://doi.org/10.5194/acp-17-343-2017
746 Pye, H. O. T., Pinder, R. W., Piletic, I. R., Xie, Y., Capps, S. L., Lin, Y.-H., Surratt, J. D., Zhang,
747 Z., Gold, A., Luecken, D. J., Hutzell, W. T., Jaoui, M., Offenberg, J. H., Kleindienst, T.
748 E., Lewandowski, M., & Edney, E. O. (2013). Epoxide Pathways Improve Model
749 Predictions of Isoprene Markers and Reveal Key Role of Acidity in Aerosol Formation.
750 *Environmental Science & Technology*, *47*(19), 11056–11064.
751 https://doi.org/10.1021/es402106h
752 Pye, H. O. T., Place, B. K., Murphy, B. N., Seltzer, K. M., D'Ambro, E. L., Allen, C., Piletic, I.
753 R., Farrell, S., Schwantes, R. H., Coggon, M. M., Saunders, E., Xu, L., Sarwar, G.,
754 Hutzell, W. T., Foley, K. M., Pouliot, G., Bash, J., & Stockwell, W. R. (2022). Linking
755 gas, particulate, and toxic endpoints to air emissions in the Community Regional
756 Atmospheric Chemistry Multiphase Mechanism (CRACMM) version 1.0. *Atmospheric*
757 *Chemistry and Physics Discussions*, 1–88. https://doi.org/10.5194/acp-2022-695
758 Qin, M., Murphy, B. N., Isaacs, K. K., McDonald, B. C., Lu, Q., McKeen, S. A., Koval, L.,
759 Robinson, A. L., Efstathiou, C., Allen, C., & Pye, H. O. T. (2021). Criteria pollutant
760 impacts of volatile chemical products informed by near-field modelling. *Nature*
761 *Sustainability*, *4*(2), Article 2. https://doi.org/10.1038/s41893-020-00614-1
762 Ritchie, S., & Tok, Y. C. (2016). *Development of a New Methodology to Characterize Truck*
763 *Body Types Along California Freeways* (11–316; p. 176). California Air Resources
764 Board. https://ww2.arb.ca.gov/sites/default/files/classic/research/apr/past/11-316.pdf
765 Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop,
766 A. P., Lane, T. E., Pierce, J. R., & Pandis, S. N. (2007). Rethinking Organic Aerosols:
767 Semivolatile Emissions and Photochemical Aging. *Science*, *315*(5816), 1259–1262.
768 https://doi.org/10.1126/science.1133061
769 Sandermann Jr, H. (1996). Ozone and Plant Health. *Annual Review of Phytopathology*, *34*(1),
770 347–366. https://doi.org/10.1146/annurev.phyto.34.1.347
771 Seinfeld, J. H., & Pandis, S. N. (2016). *Atmospheric Chemistry and Physics: From Air Pollution*
772 *to Climate Change* (3rd ed.). John Wiley & Sons, Inc.
773 Seltzer, K. M., Murphy, B. N., Pennington, E. A., Allen, C., Talgo, K., & Pye, H. O. T. (2022).
774 Volatile Chemical Product Enhancements to Criteria Pollutants in the United States.
775 *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.1c04298
776 Seltzer, K. M., Pennington, E., Rao, V., Murphy, B. N., Strum, M., Isaacs, K. K., & Pye, H. O.
777 T. (2021). Reactive organic carbon emissions from volatile chemical products.
778 *Atmospheric Chemistry and Physics*, *21*(6), 5079–5100. https://doi.org/10.5194/acp-21-
779 5079-2021
780 Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., & Barker, D. (2008). *A Description of*
781 *the Advanced Research WRF Version 3* (NCAR/TN-475+STR). University Corporation
782 for Atmospheric Research. http://dx.doi.org/10.5065/D68S4MVH
783 US EPA. (2020). *CMAQ* [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.4081737

US EPA. (2022). *Spatial Allocator v4.4 (June 2019 release)* [C++]. CMAS Center. https://github.com/CMASCenter/Spatial-Allocator/blob/14176784e03f7379d8c6a25f4ce7cfb2dd08128c/docs/User_Manual/README.md (Original work published 2017)

US EPA, O. (2013, August 1). *Air Quality System (AQS)* [Data and Tools]. US EPA. https://www.epa.gov/aqs

Xie, Y., Paulot, F., Carter, W. P. L., Nolte, C. G., Luecken, D. J., Hutzell, W. T., Wennberg, P. O., Cohen, R. C., & Pinder, R. W. (2013). Understanding the impact of recent advances in isoprene photooxidation on simulations of regional air quality. *Atmospheric Chemistry and Physics*, *13*(16), 8439–8455. https://doi.org/10.5194/acp-13-8439-2013

Xu, J., Shi, J., Zhang, Q., Ge, X., Canonaco, F., Prévôt, A. S. H., Vonwiller, M., Szidat, S., Ge, J., Ma, J., An, Y., Kang, S., & Qin, D. (2016). Wintertime organic and inorganic aerosols in Lanzhou, China: Sources, processes, and comparison with the results during summer. *Atmospheric Chemistry and Physics*, *16*(23), 14937–14957. https://doi.org/10.5194/acp-16-14937-2016

Xu, L., Pye, H. O. T., He, J., Chen, Y., Murphy, B. N., & Ng, N. L. (2018). Experimental and model estimates of the contributions from biogenic monoterpenes and sesquiterpenes to secondary organic aerosol in the southeastern United States. *Atmospheric Chemistry and Physics*, *18*(17), 12613–12637. https://doi.org/10.5194/acp-18-12613-2018

Yang, J., Wen, Y., Wang, Y., Zhang, S., Pinto, J. P., Pennington, E. A., Wang, Z., Wu, Y., Sander, S. P., Jiang, J. H., Hao, J., Yung, Y. L., & Seinfeld, J. H. (2021). From COVID-19 to future electrification: Assessing traffic impacts on air quality by a machine-learning model. *Proceedings of the National Academy of Sciences*, *118*(26). https://doi.org/10.1073/pnas.2102705118

Yarwood, G., Jung, J., Whitten, G. Z., Heo, G., Mellberg, J., & Estes, M. (2010). *Updates to the Carbon Bond Mechanism for Version 6 (CB6)*. 9th Annual CMAS Conference, Chapel Hill, NC.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T., Jayne, J. T., Miyoshi, T., Shimono, A., … Worsnop, D. R. (2007). Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes. *Geophysical Research Letters*, *34*(13). https://doi.org/10.1029/2007GL029979

Zhang, X., Cappa, C. D., Jathar, S. H., McVay, R. C., Ensberg, J. J., Kleeman, M. J., & Seinfeld, J. H. (2014). Influence of vapor wall loss in laboratory chambers on yields of secondary organic aerosol. *Proceedings of the National Academy of Sciences*, *111*(16), 5802–5807. https://doi.org/10.1073/pnas.1404727111

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A., & Robinson, A. L. (2016). Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines. *Environmental Science & Technology*, *50*(8), 4554–4563. https://doi.org/10.1021/acs.est.5b06247

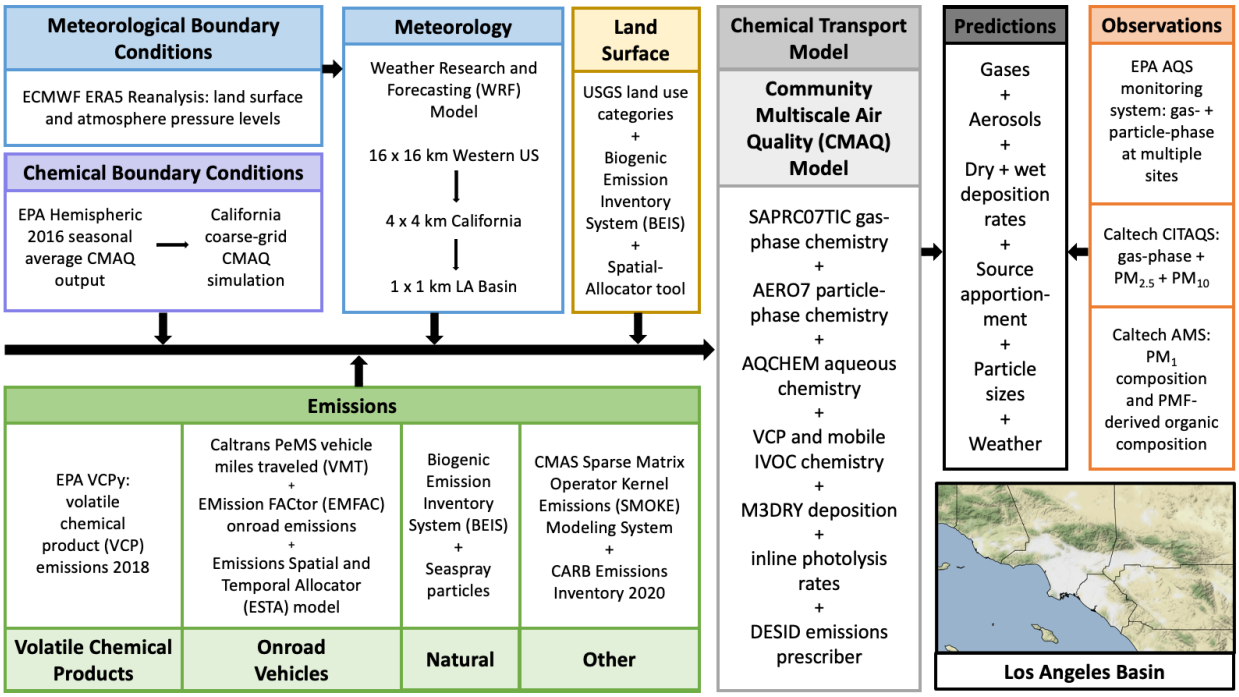

Figure 1: Model framework describing the inputs to CMAQ, CMAQ configuration, observational data, and modeling domain.

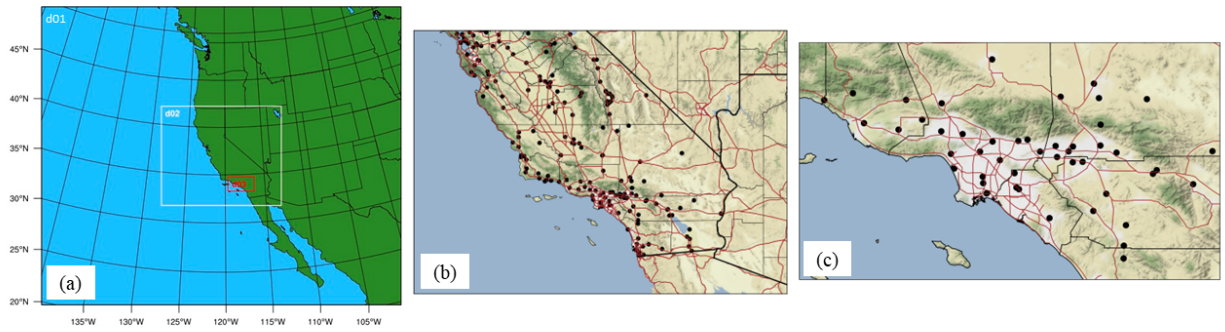

Figure 2: A) Three nested domains used in the WRF simulations. d01 has a horizontal resolution of 16 km, d02 has a resolution of 4 km, and d03 has a resolution of 1 km. B) California 4 x 4 km coarse-resolution domain. C) LA 1 x 1 km fine-resolution domain. Thick black lines are state borders and thin black lines are county borders. Black dots represent EPA AQS sites and red lines are freeways.

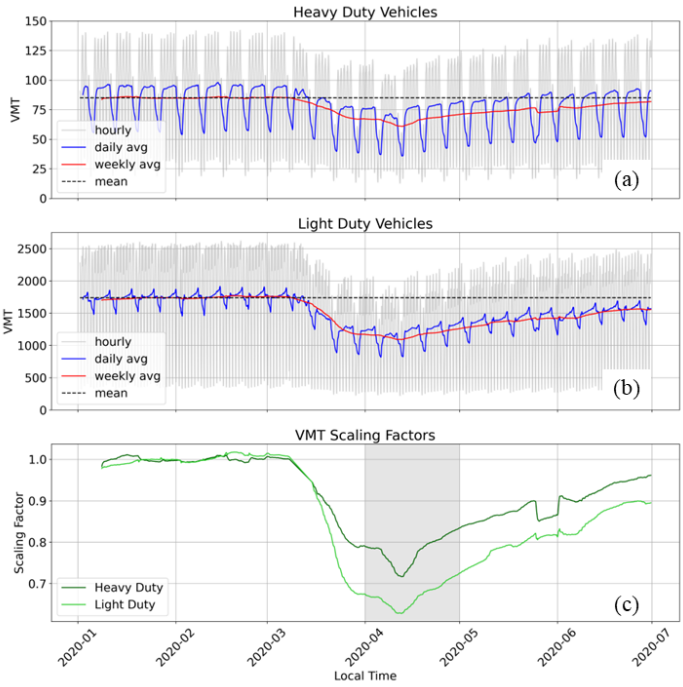

Figure 3: Hourly (gray), daily-averaged (blue), and weekly-averaged (red) VMT data (Caltrans, 2020) for A) heavy duty vehicles and B) light duty vehicles. VMT averaged January 1– March 1, 2020 is represented by the dashed black line. C) Weekly-averaged VMT divided by the January–March mean for heavy duty (dark green) and light duty (light green) vehicles. The gray shaded area covers the modeling period: April 1–30, 2020.

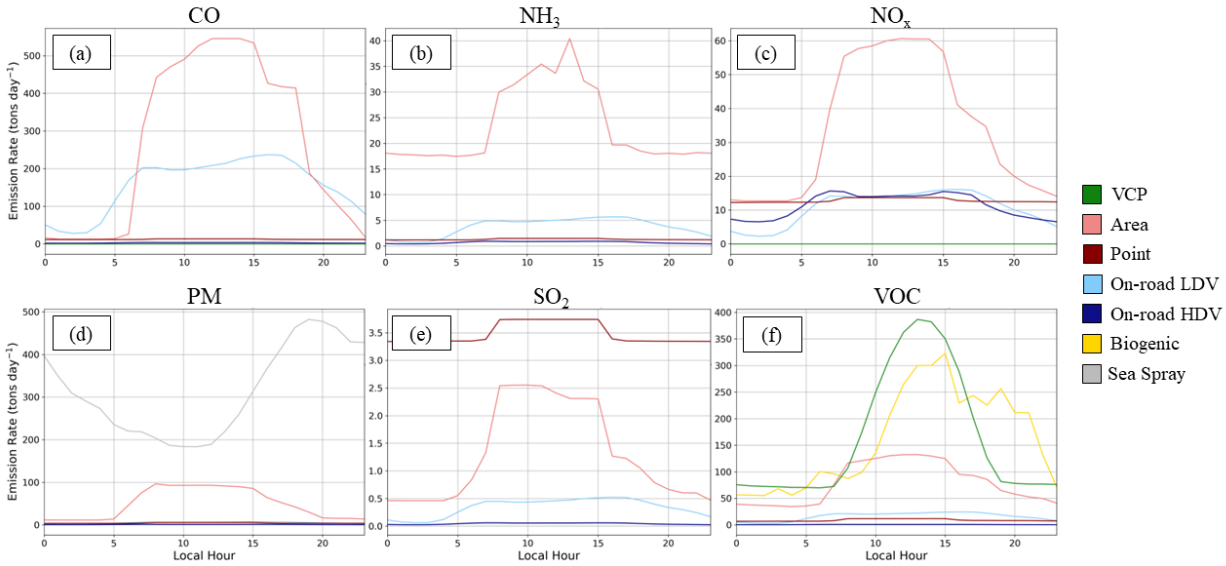

Figure 4: Diurnal variations of emission rates averaged April 1–30, 2020 and summed over the LA domain (with all ocean-covered cells removed) from all emission sources for A) CO, B) NH_3, C) NO_x, D) PM, E) SO_2, F) VOC.

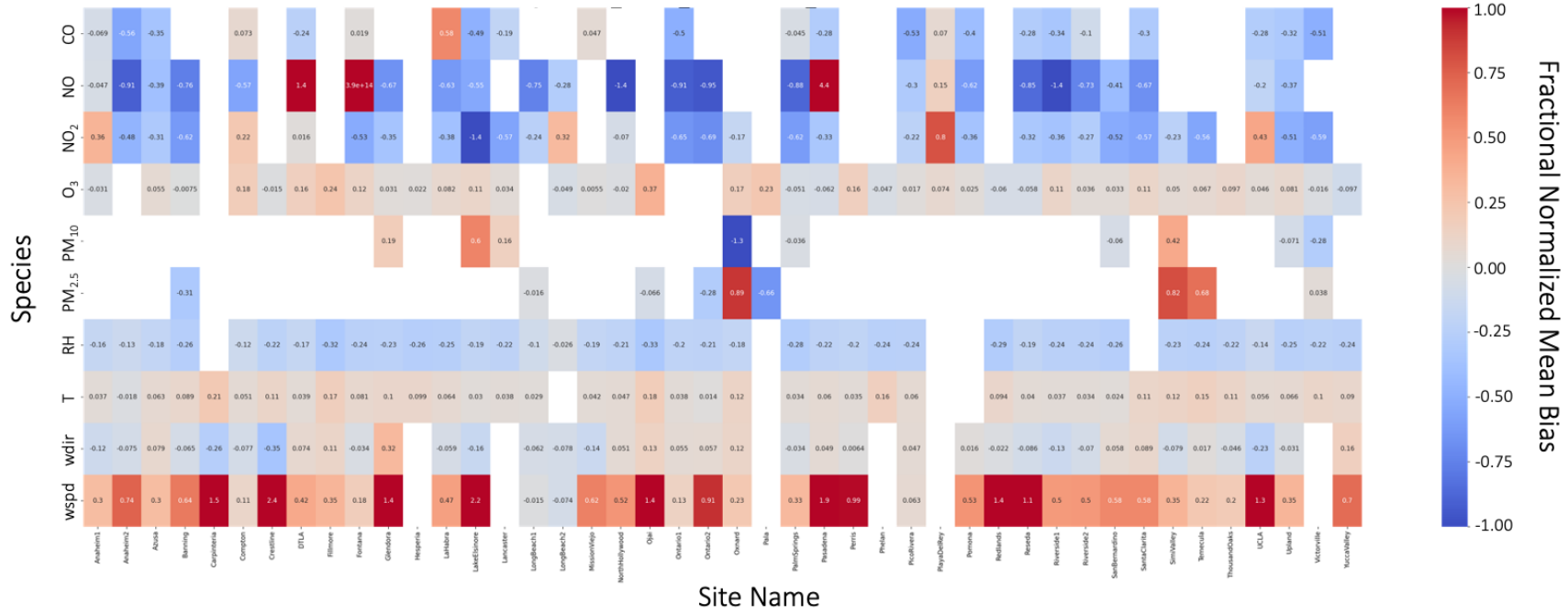

Figure 5: Fractional NMB of pollutants (rows) at all EPA AQS sites (columns) in the LA domain using daily-average values April 1–30, 2020. Empty boxes represent sites without measurements of the given pollutant.

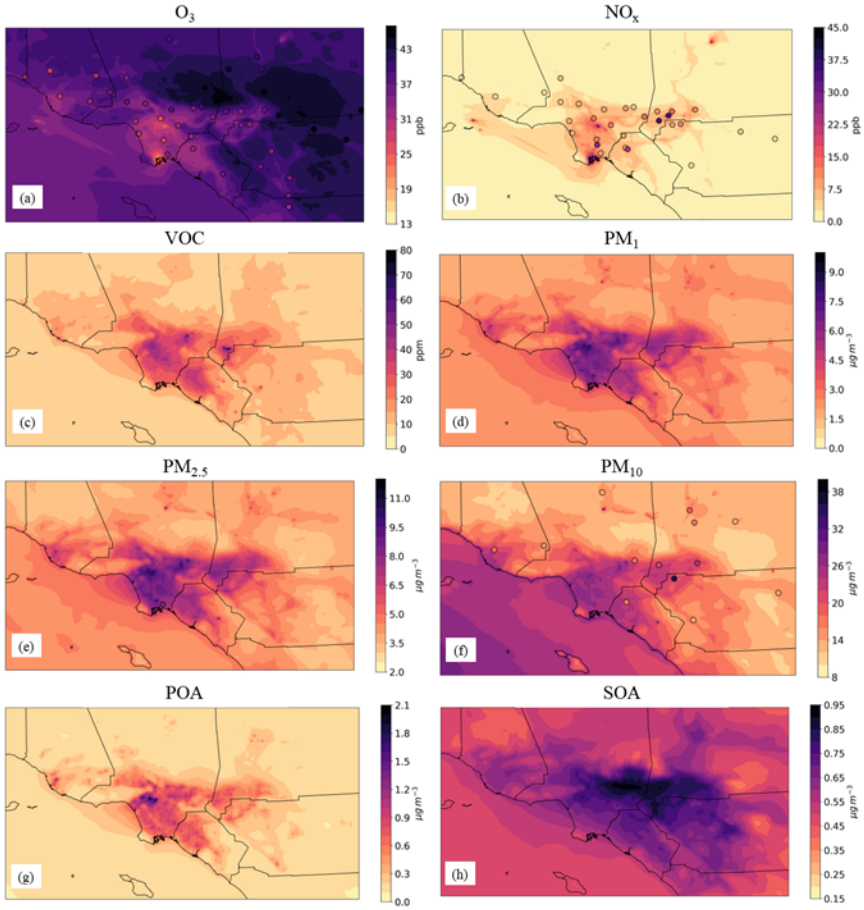

Figure 6: Time-averaged (April 1–30, 2020) CMAQ predicted concentration of A) O_3 (ppb), B) NO_x (ppb), C) total VOC (ppm), D) PM_1 (µg m^{-3}), E) $PM_{2.5}$ (µg m^{-3}), F) PM_{10} (µg m^{-3}), G) POA (µg m^{-3}), and H) SOA (µg m^{-3}). Circles depict the average concentration measured at the EPA AQS site at that location. There are no AQS measurements of VOCs, PM_1, POA, or SOA.

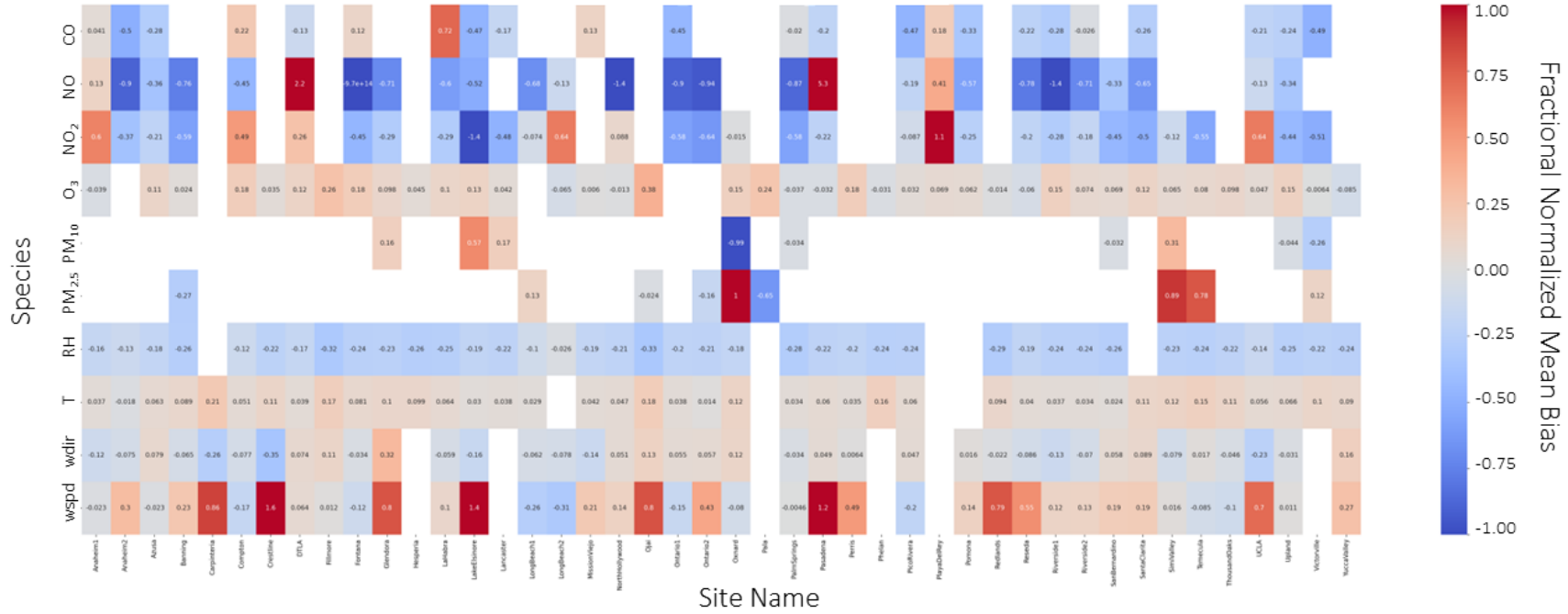

Figure 7: Fractional NMB of pollutants (rows) at all EPA AQS sites (columns) in the LA domain using daily-average values April 1–30, 2020. Empty boxes represent sites without measurements of the given pollutant. Results presented here use default wind speed scaled by a factor of 0.75.

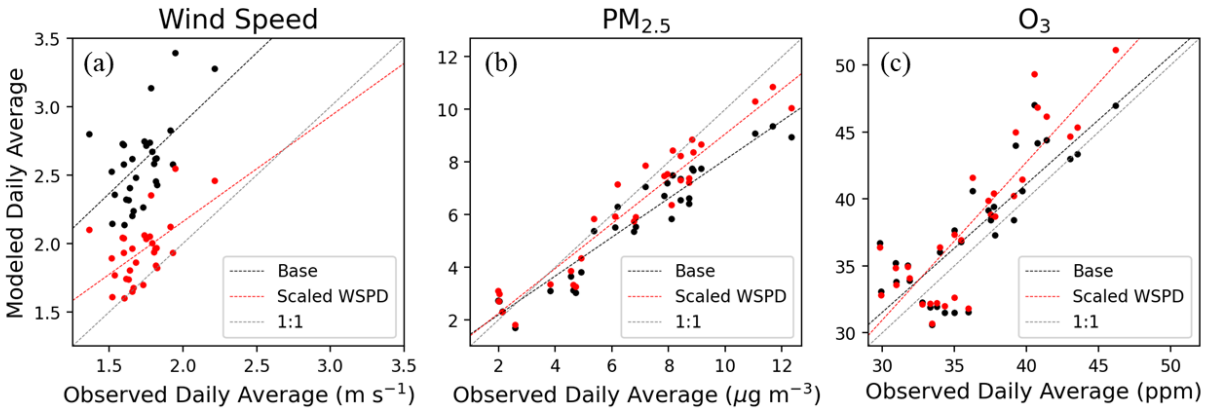

Figure 8: Daily-averaged modeled versus observed values of (a) wind speed, (b) $PM_{2.5}$, and (c) O_3. Black markers and lines represent data from the "base case" wind speed simulations. Red markers and lines represent data from the scaled (i.e. scaled by 0.75) wind speed simulations. Gray line represents the 1:1 modeled:observed line.

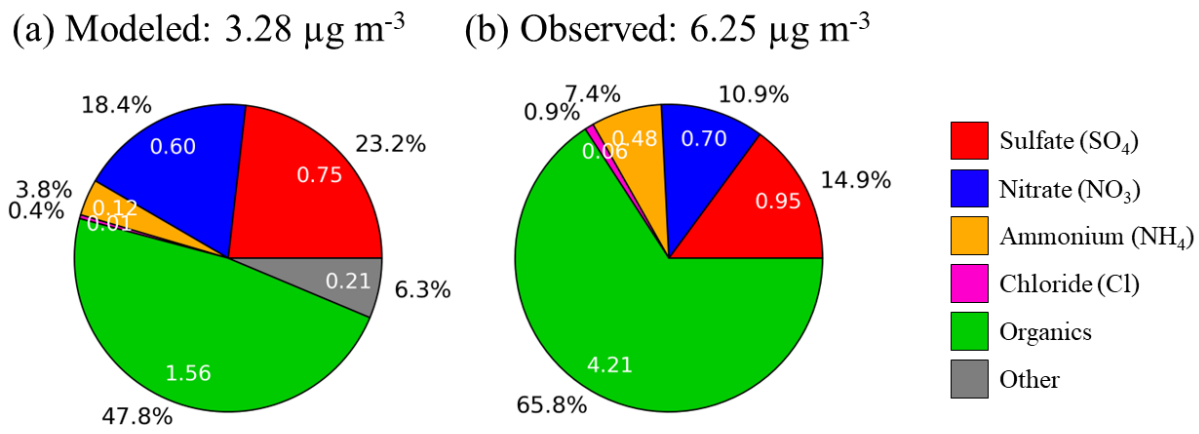

Figure 9: PM_1 composition averaged April 8–30, 2020 in Pasadena A) predicted by CMAQ and B) measured by AMS. Values inside the pie represent average mass values (μg m^{-3}) and values outside the pie represent the percentage of the total mass of each component.

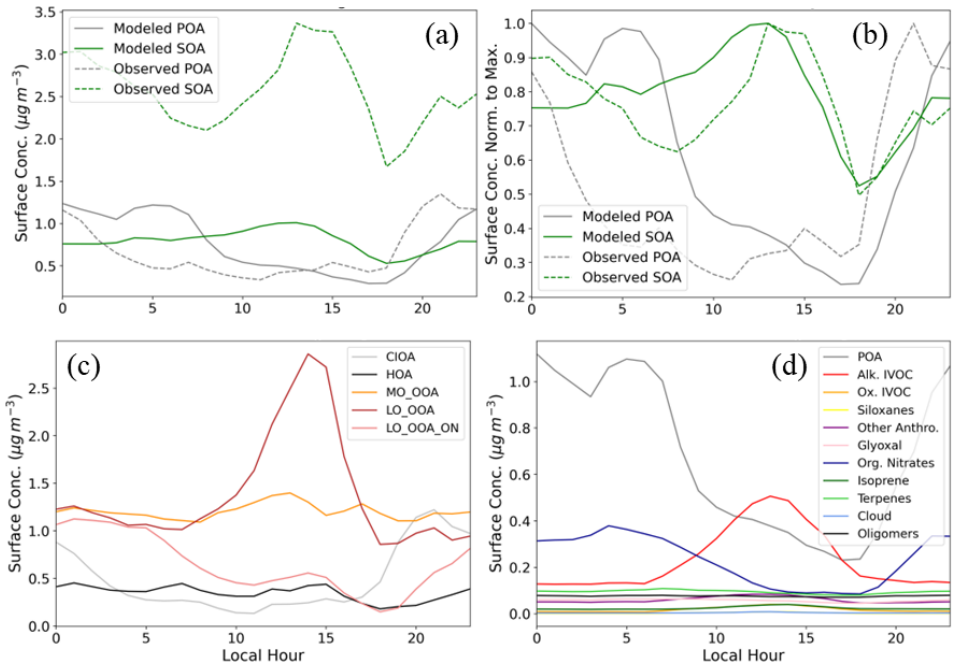

Figure 10: A) Modeled (solid) and measured (dashed) POA (gray) and SOA (green) diurnal variation in Pasadena. B) Modeled (solid) and measured (dashed) POA (gray) and SOA (green) diurnal variation in Pasadena. Surface concentration was normalized to the daily-maximum surface concentration. C) PMF-calculated POA and SOA speciation in Pasadena. D) Model-predicted POA and SOA speciation in Pasadena. All diurnal trends calculated April 8–30, 2020.

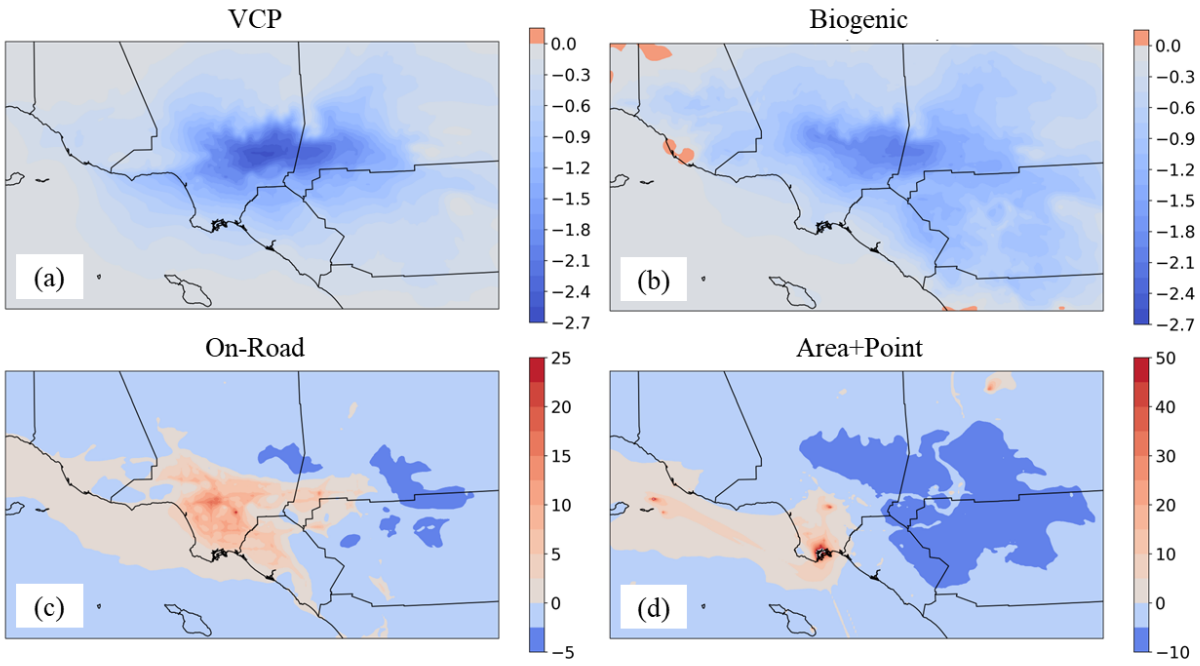

Figure 11: Percent change in average (April 1–30, 2020) predicted O_3 concentration averaged April 1–30, 2020 caused by removing each emission source: A) VCP, B) biogenic, C) on-road vehicles, and D) area+point.

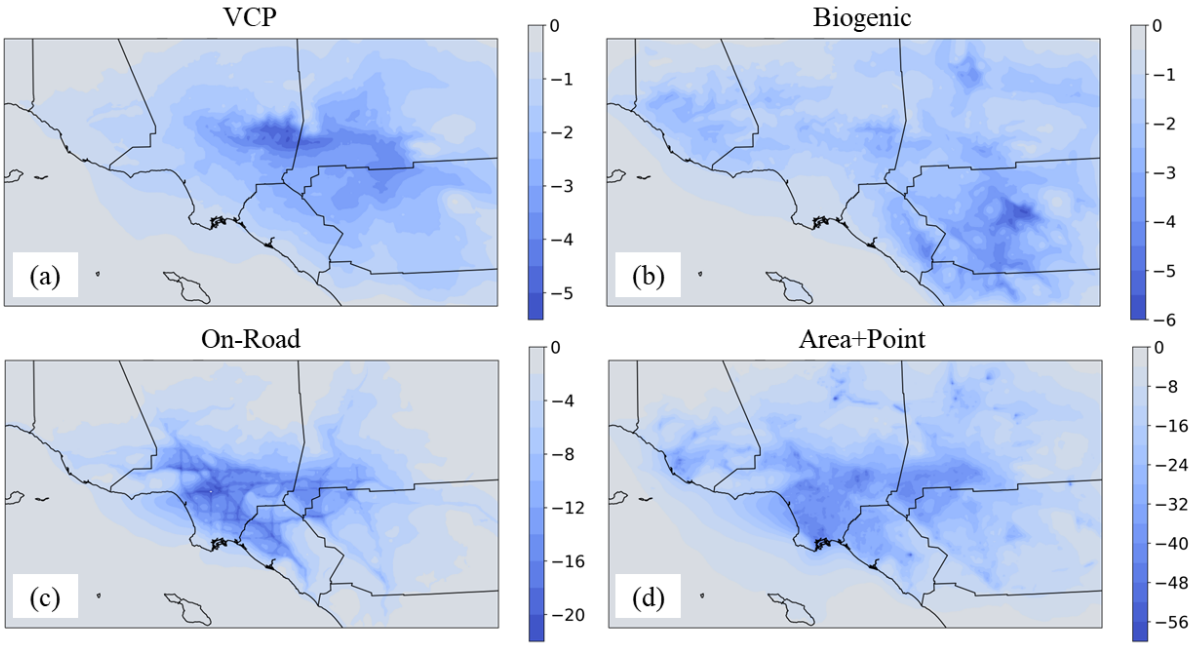

Figure 12: Percent change in average (April 1–30, 2020) predicted $PM_{2.5}$ concentration caused by removing each emission source: A) VCP, B) biogenic, C) on-road vehicles, and D) area+point.

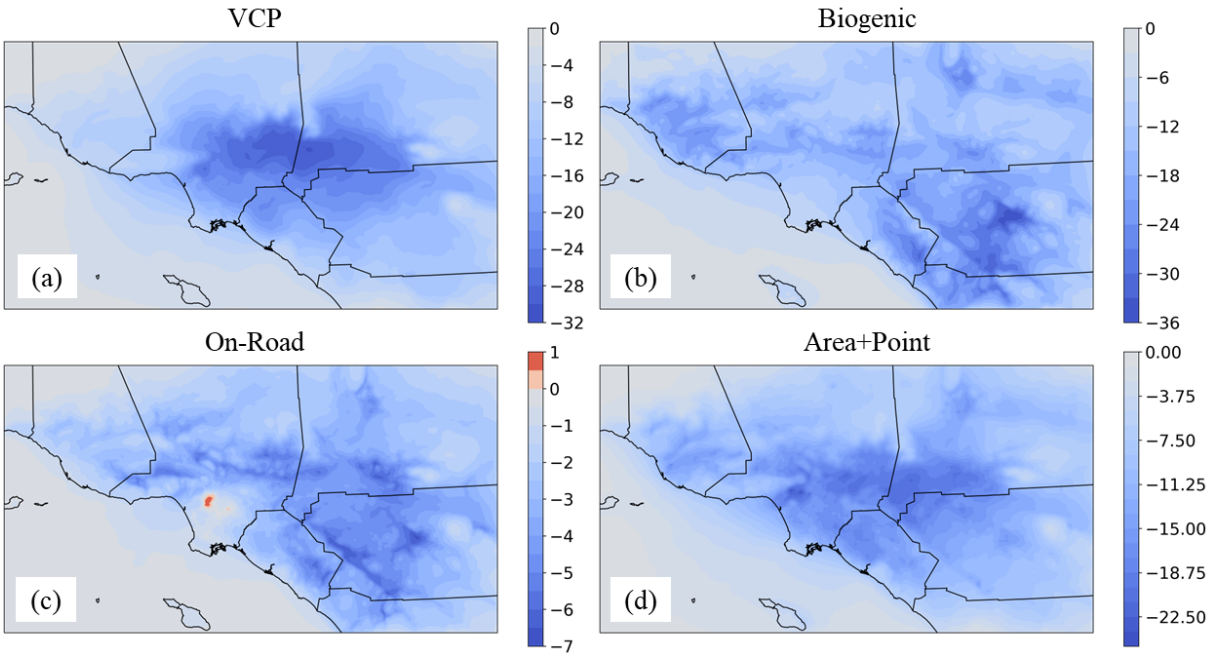

Figure 13: Percent change in average (April 1–30, 2020) predicted SOA concentration caused by removing each emission source: A) VCP, B) biogenic, C) on-road vehicles, and D) area+point.

Table 1: Statistical analysis of daily-averaged CMAQ predictions for the (top) CA coarse-resolution domain and (bottom) LA Basin subset of the California domain, compared to EPA AQS monitoring site data.

	O_3	CO	NO_x	SO_2	$PM_{2.5}$	PM_{10}
California Coarse-Resolution Simulation						
Number of Data Points	341	248	310	62	186	93
Observed Mean	32.6 ppb	221 ppb	9.09 ppb	0.095 ppb	5.29 μg m^{-3}	17.0 μg m^{-3}
Modeled Mean	33.1 ppb	140 ppb	7.88 ppb	0.217 ppb	7.21 μg m^{-3}	12.1 μg m^{-3}
MB	0.44 ppb	-81 ppb	-1.20 ppb	0.123 ppb	1.92 μg m^{-3}	-4.87 μg m^{-3}
NMB	1.36%	-36.5%	-13.2%	129%	36.3%	-28.7%
RMSE	6.37 ppb	99.2 ppb	8.07 ppb	0.160 ppb	5.41 μg m^{-3}	10.5 μg m^{-3}
r^2	0.23	0.40	0.37	0.15	0.51	0.28
Los Angeles Subset of California Coarse-Resolution Simulation						

Number of Data Points	126	134	155	31	36	33
Observed Mean	33.3 ppb	242 ppb	13.2 ppb	0.090 ppb	8.60 μg m^{-3}	21.2 μg m^{-3}
Modeled Mean	29.5 ppb	170. ppb	12.6 ppb	0.223 ppb	18.2 μg m^{-3}	15.5 μg m^{-3}
MB	-3.77 ppb	-72.2 ppb	-0.62 ppb	0.133 ppb	9.65 μg m^{-3}	-5.70 μg m^{-3}
NMB	-11.3%	-29.8%	-4.72%	147%	112%	-26.8%
RMSE	7.06 ppb	85.0 ppb	10.8 ppb	0.17 ppb	11.9 μg m^{-3}	8.36 μg m^{-3}
r^2	0.36	0.52	0.25	0.26	0.49	0.66

Table 2: Mass concentration change (ng m^{-3}) of SOA components averaged over the LA domain when each emission source is removed.

ng m^{-3}	VCP	Onroad	Biogenic	Sea Spray	Area+Point
Alkane-like IVOCs	-36.03	-4.89	1.29	-0.01	-23.76
Oxygenated IVOCs	-4.61	-0.17	0.03	0.002	-0.38
Siloxanes	-1.10	-0.09	0.006	-7.3×10^{-4}	-0.27
Glyoxal	-1.01	-1.05	-2.11	-0.10	-2.88
Other anthropogenic	-3.69	-0.71	-1.10	0.07	-2.63
Isoprene	-0.41	-0.29	-5.24	6.7×10^{-4}	-1.03
Monoterpenes	-2.41	0.56	-18.36	-0.01	-1.40
Sesquiterpenes	-0.13	-0.05	-0.15	-3.4×10^{-4}	-0.24
Organic nitrates	-10.52	-5.64	-42.53	0.14	-16.08
Oligomers	-0.83	-0.30	-1.35	7.9×10^{-4}	-0.90
Cloud-processed	-0.10	-0.10	-0.15	-1.8×10^{-4}	-0.26