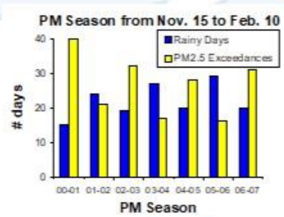
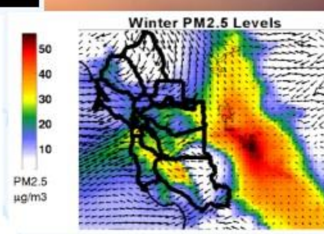
BAY AREA AIR QUALITY MANAGEMENT DISTRICT
375 Beale Street, San Francisco, CA 94105

Air Quality Modeling and Analysis Section Publication No. 201901-017-PM

Fine Particulate Matter Data Analysis and Regional Modeling In the San Francisco Bay Area to Support AB617

January, 2019





Prepared by the Air Quality Modeling and Analysis Section:
Saffet Tanrikulu, Manager
Stephen Reid, Senior Atmospheric Modeler
Bonyoung Koo, Senior Atmospheric Modeler
Yiqin Jia, Atmospheric Modeler
James Cordova, Air Quality Meteorologist
Jeff Matsuoka, Research Analyst
Yuanyuan Fang, Statistician

Executive Summary

E1. Background

The adoption of Assembly Bill 617 (AB617) established collaborative programs to reduce community exposure to air pollutants in neighborhoods most impacted by air pollution. Air District staff have been working closely with the California Air Resources Board (ARB), other local air districts, community groups, community members, environmental organizations, regulated industries, and other key stakeholders to reduce harmful air pollutants in Bay Area communities.

The purpose of this data analysis and regional modeling effort is to support the District's AB617 activities by assessing pollutant formation, quantifying the relative contribution of emission sources to ambient pollution levels, and assessing population exposures and the benefits of emission controls in impacted communities around the Bay Area. Our initial assessments focus on fine particulate matter ($PM_{2.5}$) concentrations in West Oakland, and follow-up analyses will include air toxics evaluations in West Oakland and expansion of our technical assessments to other communities.

For the $PM_{2.5}$ analyses, we evaluated ambient meteorological and air quality data, and applied the U.S. EPA's Community Multi-Scale Air Quality (CMAQ) model to simulate pollutant concentrations at a 1-km horizontal resolution over the entire Bay Area for 2016. Then we repeated the simulation with West Oakland's anthropogenic emissions removed from the modeling inventory, leaving all other model input parameters unchanged. We calculated annual average $PM_{2.5}$ concentrations using the output of each simulation. The first simulation provided the annual average $PM_{2.5}$ concentrations for 2016 over the entire Bay Area, which will be used for $PM_{2.5}$ exposure analyses and health impacts assessments. The second simulation provided an estimate of background $PM_{2.5}$ levels in West Oakland (i.e., the $PM_{2.5}$ concentrations that would exist in the absence of local West Oakland sources). Background $PM_{2.5}$ concentrations will then be combined with local-scale modeling of West Oakland sources using the AERMOD dispersion model to provide a complete picture of $PM_{2.5}$ levels in the community and the relative contribution of different emission sources to those levels.

E2. Major Findings

E2.1 Regional $PM_{2.5}$ Concentrations

The CMAQ model generally captured the observed $PM_{2.5}$ pattern within the 1-km domain (Figure E1). High concentrations in both simulations and observations are evident in the northern San Joaquin Valley, along the I-580 and I-880 corridors from Richmond to the Oakland Airport, along the I-101 corridor near Redwood City, and in the San Jose metropolitan area. In the Sacramento area, the model shows overestimation biases and $PM_{2.5}$ concentrations do not compare as well to observations as in the Bay Area. For Sacramento and other counties outside the Bay Area, we relied on the ARB's emission inventories, and further evaluation of these data may be warranted. The model also shows high concentrations along the I-880 corridor from Oakland Airport to San Jose and along the Delta from Antioch to Brentwood, although observations are unavailable in these areas.

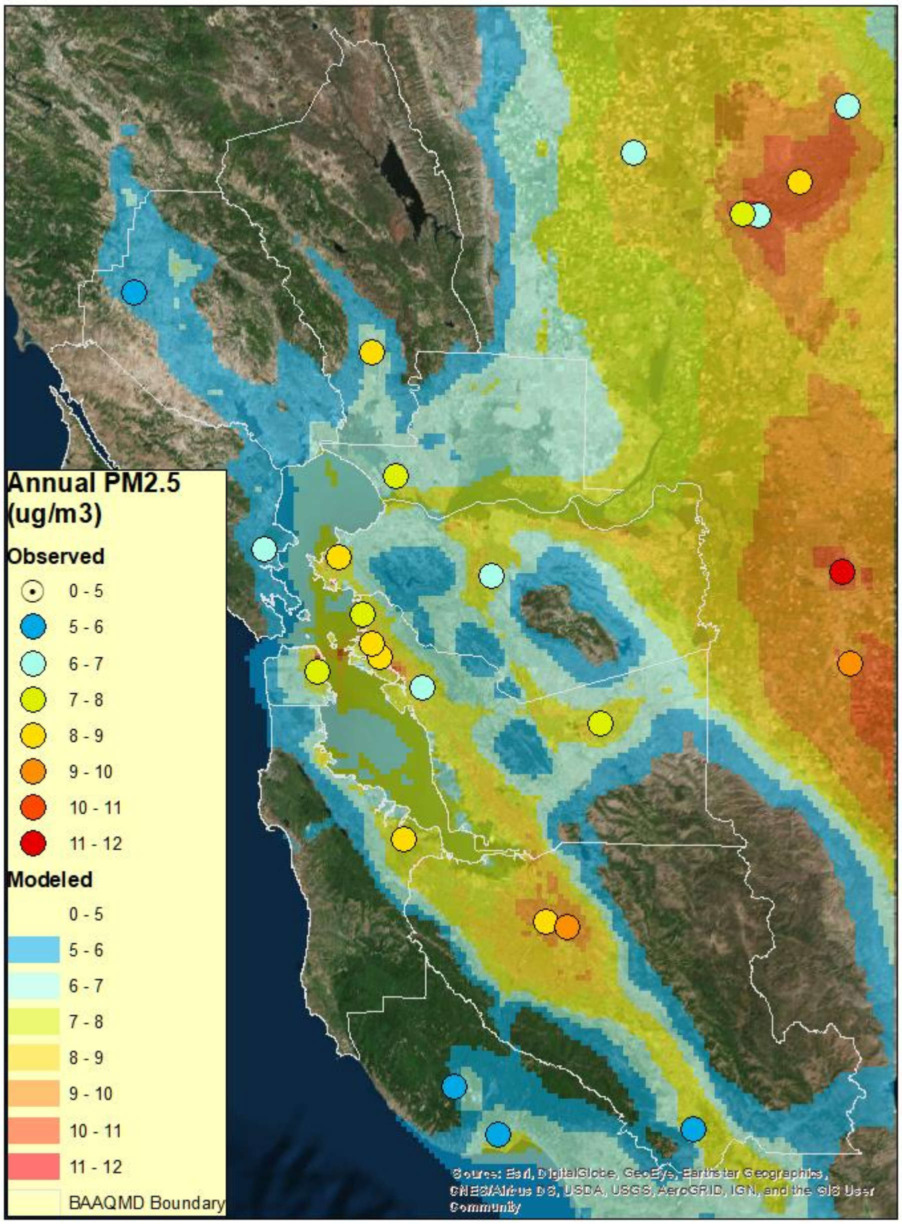


Figure E1: Spatial distribution of simulated and observed annual average $PM_{2.5}$ concentrations within the 1-km modeling domain.

Site by site comparisons between the simulations and observations (Figure E2) show that at most Bay Area sites (including the West Oakland Air Monitoring Station), the simulated annual average $PM_{2.5}$ concentrations are within ±1.0 μg/m^3 of observations. At a few sites (Concord, Oakland and Gilroy), the annual average $PM_{2.5}$ concentrations were overestimated, and at one site (Napa), the annual average PM2.5 concentration was underestimated by as much as 2.1 µg/m^3. Causes of these over and underestimations are under investigation.

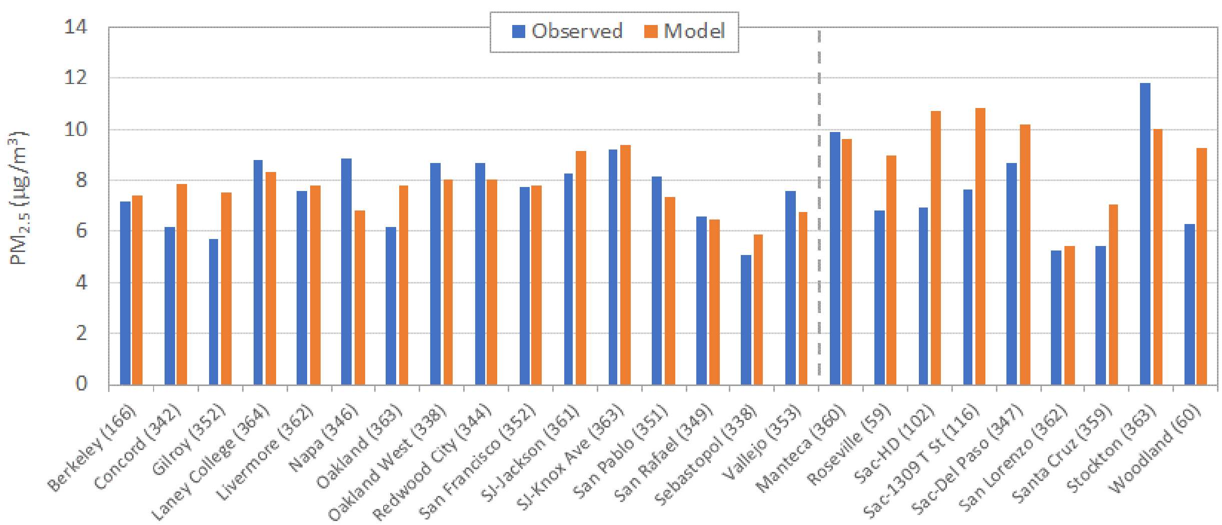


Figure E2: Annual mean observed vs. modeled $PM_{2.5}$ concentrations at monitoring sites within the 1-km modeling domain. The number of valid observations is shown in parentheses for each site.

E2.2 Estimating Background $PM_{2.5}$ in West Oakland

Figure E3 shows the annual average $PM_{2.5}$ concentrations for the base case within the West Oakland local-scale modeling domain that will be used for AERMOD. The highest and lowest annual average $PM_{2.5}$ concentrations are 9.3 µg/m^3 and 7.1 µg/m^3, respectively. A concentration gradient is evident within the domain. Cells with relatively higher concentrations extend along the eastern boundary and northwestern corner of the domain. A concentration gradient is also evident in the West Oakland community, an area within the red border in the figure. The eastern half of the community has slightly higher concentrations than the western half.

The spatial distribution of the annual average $PM_{2.5}$ concentrations is similar to the spatial distribution of West Oakland's emissions (Figure E4). The Chinatown area in the southeastern corner of the West Oakland local-scale domain has the highest emissions and concentrations. The cell along the southern boundary with the area's lowest concentration (7.1 µg/m^3) also has the lowest emissions (1.4 lbs/day).

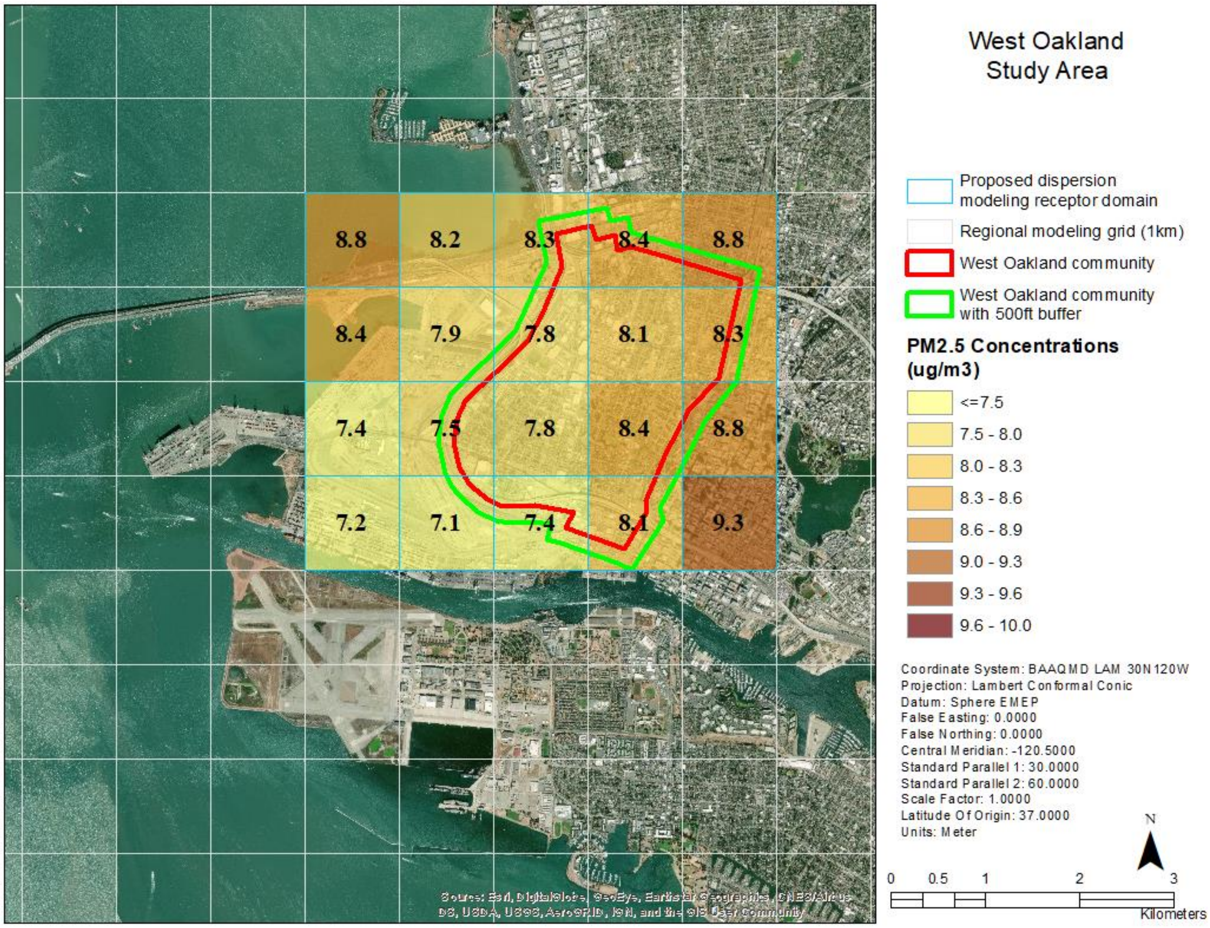


Figure E3: Spatial distribution of the simulated annual average $PM_{2.5}$ concentrations in the West Oakland modeling domain.

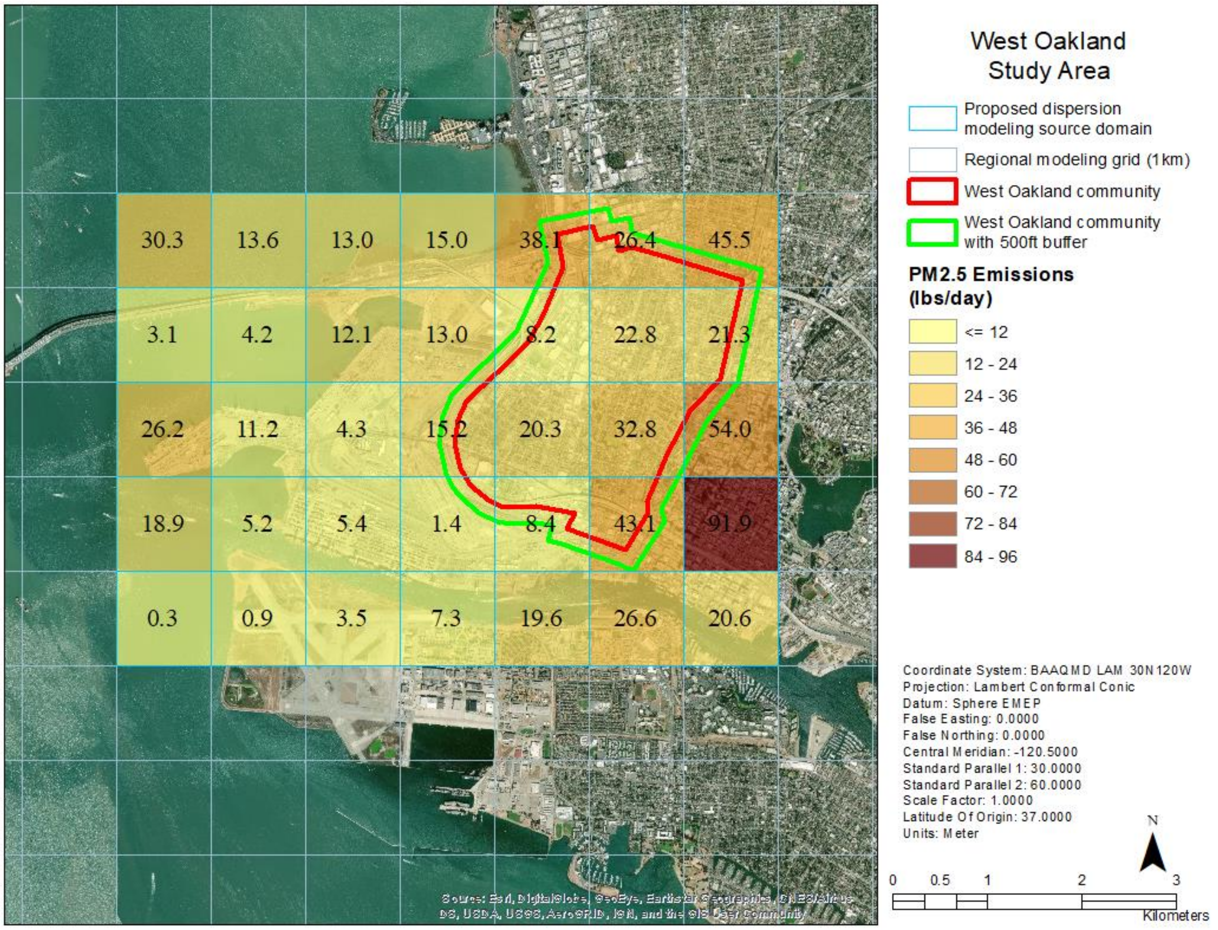


Figure E4: Spatial distribution of annual average $PM_{2.5}$ emissions in West Oakland.

Figure E5 shows the annual average $PM_{2.5}$ concentrations for the control case, i.e., a simulation without West Oakland's anthropogenic emissions. Compared to Figure E3, the spatial gradient in the annual average concentrations decreased significantly in the absence of West Oakland emissions across the local-scale domain. The location of the maximum annual average $PM_{2.5}$ concentrations has shifted from Chinatown to near the Bay Bridge, suggesting the influence of transport from the northwest corner of the domain.

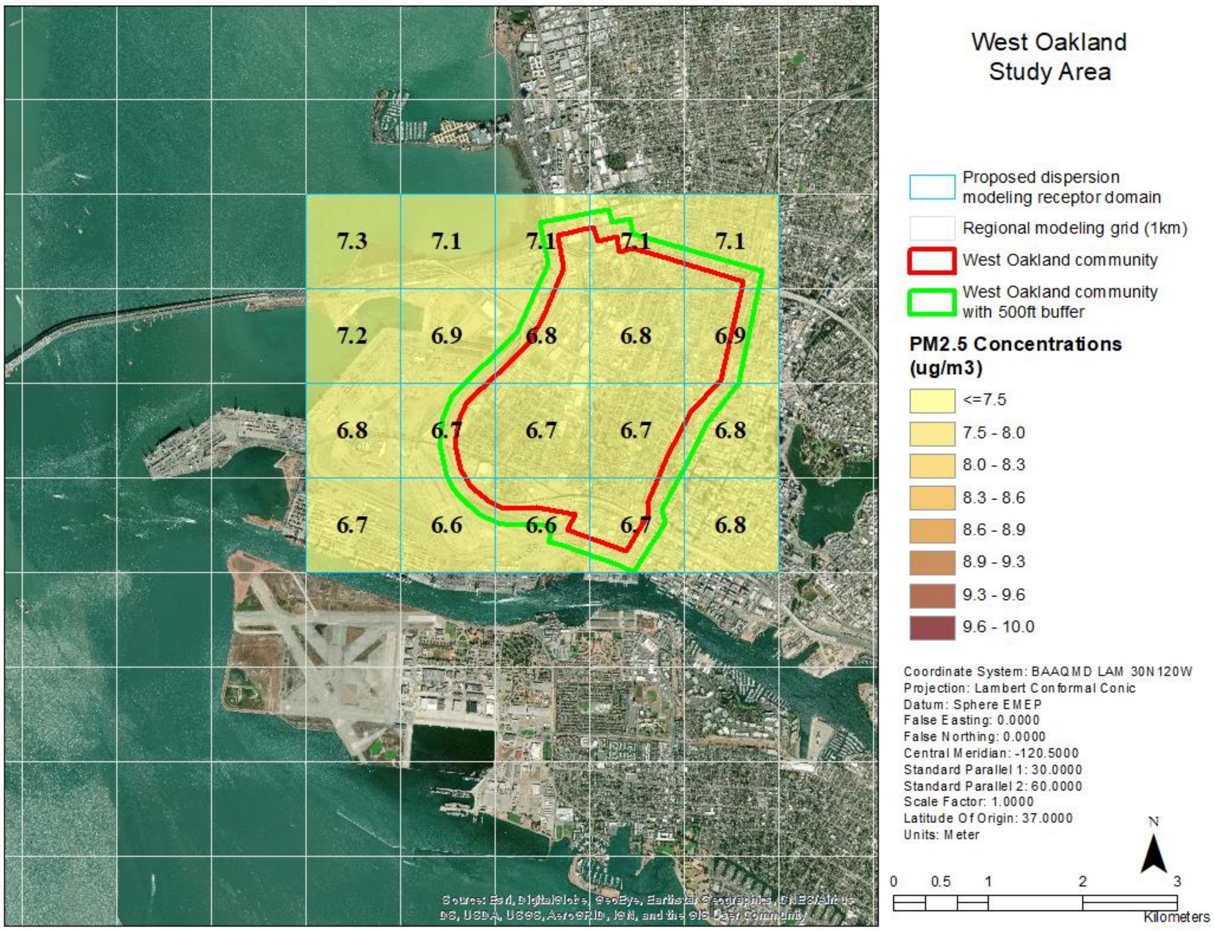


Figure E5: Spatial distribution of the simulated $PM_{2.5}$ concentrations without West Oakland's anthropogenic emissions.

Figure E6 shows the difference between the base and control cases. Based on the figure, the Chinatown area would benefit the most (2.5 µg/m^3) from zeroing out all anthropogenic emissions in the West Oakland local-scale domain. The West Oakland community (within the red border) would benefit by $PM_{2.5}$ reductions ranging from 0.8 µg/m^3 to 1.7 µg/m^3. The southwest corner of the modeling domain would be the least benefitted area, with a reduction of about 0.5 µg/m^3.

Note that these $PM_{2.5}$ concentrations and reductions represent the average value across a 1x1 km grid cell. Higher concentrations and reductions are possible at the sub-grid cell level, and these finer-scale gradients will be investigated with the local-scale AERMOD modeling.

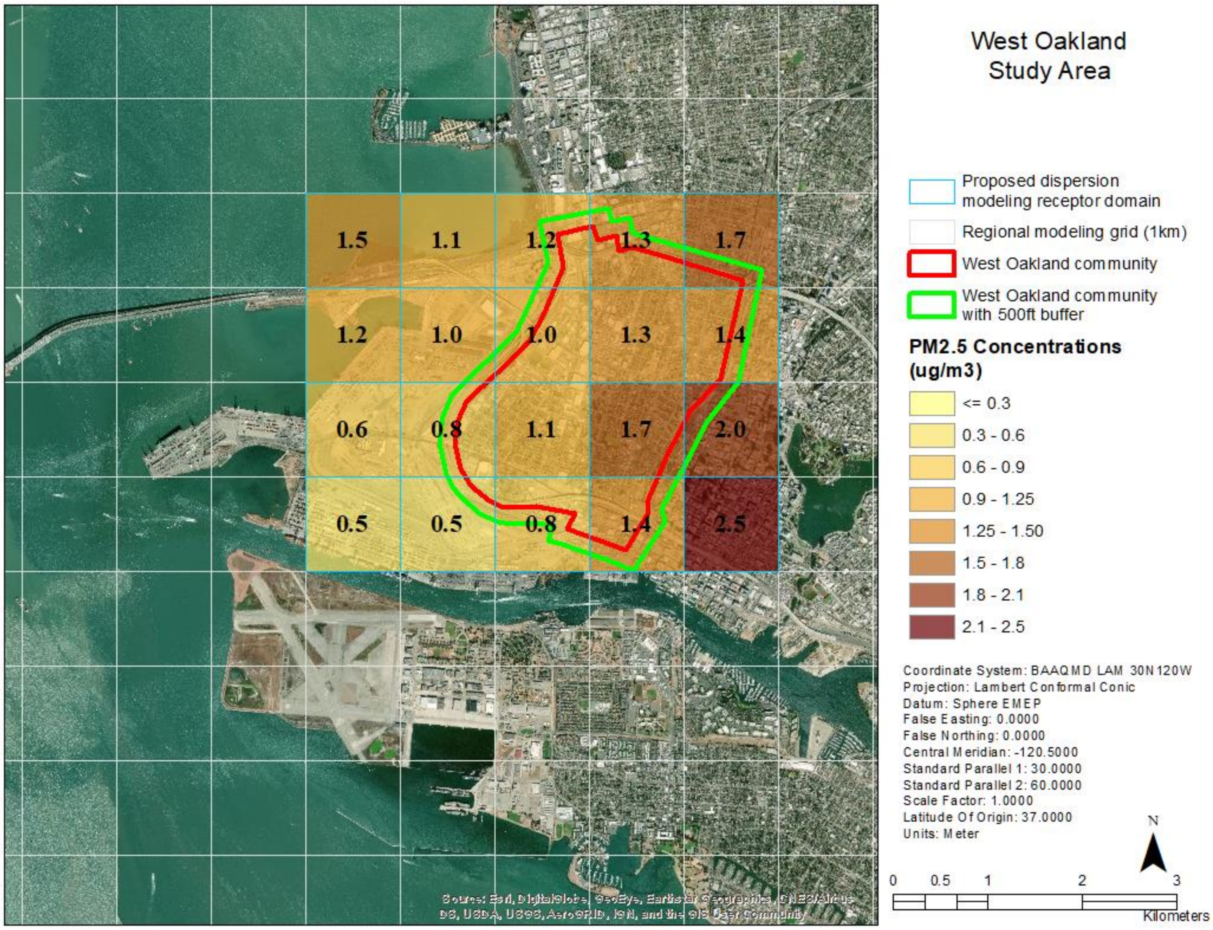
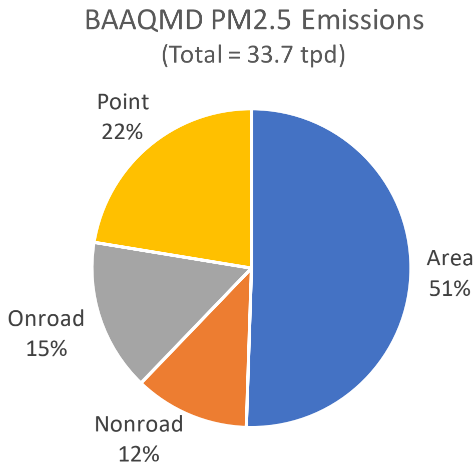


Figure E6: Difference between the simulated annual average base and control case $PM_{2.5}$ concentrations.

E3. Discussion

West Oakland is a unique area in terms of its geographic location, emissions, meteorology and air quality. In the West Oakland local-scale domain, annual average $PM_{2.5}$ emissions are 0.35 tons per day (tpd), about 1% of the Bay Area total. Onroad and nonroad mobile sources account for 66% of total $PM_{2.5}$ emissions. Area sources account for 24% of total $PM_{2.5}$ emissions, a significantly smaller percentage compared to the Bay Area total $PM_{2.5}$ emissions (Figure E7).

West Oakland is also impacted by pollutant transport from outside sources for all seasons. During spring, summer and fall, prevailing winds from the west, northwest and, to a lesser degree, from the southwest transport pollutants from downtown San Francisco, the San Francisco Peninsula, and shipping emissions from the Pacific Ocean and the Bay. During winter, occasional easterly airflow transports polluted air from the Central Valley through the Delta. West Oakland is also open for sea salt intrusion, which mostly occurs during spring, and the transport of wildfire emissions from the Sierras, other northern California locations and state of Oregon during the wildfire season. Transport to West Oakland from southern California, neighboring counties and intercontinental transport are also possible.[1]



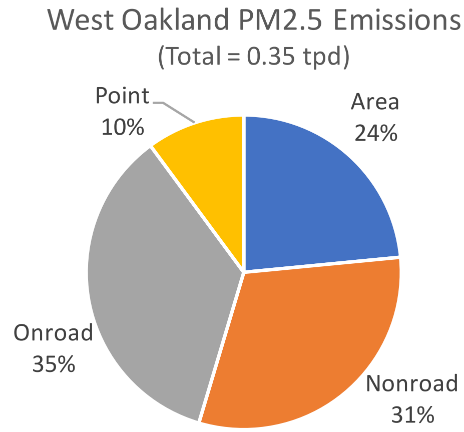


Figure E7: $PM_{2.5}$ emissions by source sector for the District (left) and West Oakland (right).

February, September and December usually exhibit the highest $PM_{2.5}$ concentrations in West Oakland (Figure E8). PM is elevated in February because of the contribution of wood burning emissions, secondary PM formation and near stagnant atmospheric conditions. Elevated PM in September is mainly influenced by wildfire emissions. In December, PM levels are significantly influenced by wood burning and cooking, which generally increases during the holidays, and relatively calm and foggy atmospheric conditions.

The remaining months exhibit PM levels around 8 $\mu g/m^3$, except July, August and October. The strong afternoon seabreeze in July and August lowers concentrations through atmospheric mixing, while October is a month with relatively low wind speeds and highly variable wind directions. The usual transport from nearby sources are not dominant during this month.

The CMAQ model is generally able to replicate the month-to-month variation in observed $PM_{2.5}$ concentrations in West Oakland (Figure E8). The model slightly overestimates PM during winter months and underestimates PM during summer months, a pattern that is typical of the CMAQ modeling system. The somewhat significant underestimation in September is likely due to lack of wildfire emissions in the CMAQ simulations.

[1] Note that this analysis did not seek to quantify the impact of various sources of transported pollution on West Oakland. Rather, to be consistent with AB617 goals, the focus was on the impact of local emissions.

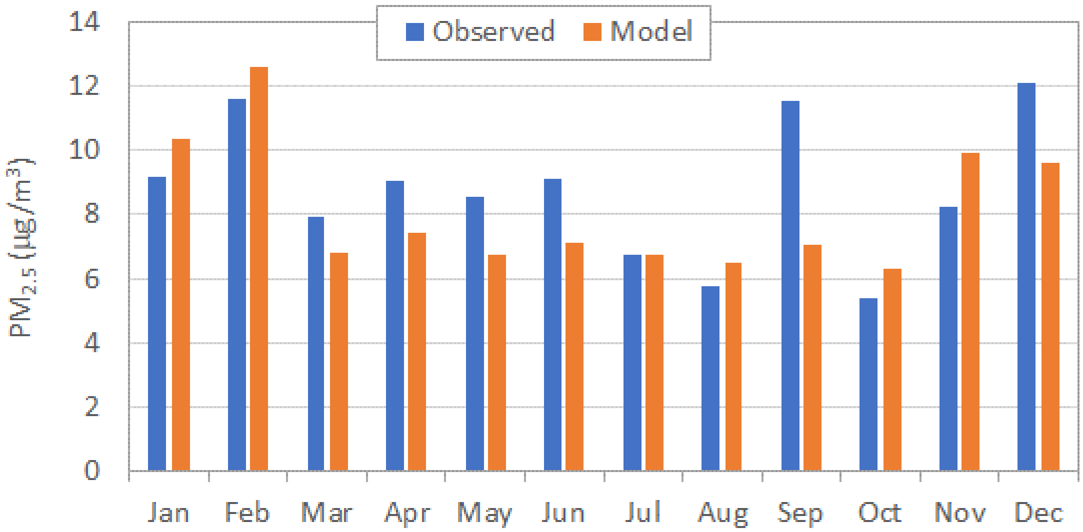


Figure E8: Monthly average simulated and observed $PM_{2.5}$ concentrations in West Oakland.

Fine Particulate Matter Data Analysis and Regional Modeling in the San Francisco Bay Area to Support AB617

1. Introduction

The adoption of Assembly Bill 617 (AB617) established collaborative programs to reduce community exposure to air pollutants in neighborhoods most impacted by air pollution. Air District staff have been working closely with the California Air Resources Board (ARB), other local air districts, community groups, community members, environmental organizations, regulated industries, and other key stakeholders to reduce harmful air pollutants in Bay Area communities.

The purpose of this data analysis and regional modeling effort is to support the District's AB617 activities by assessing pollutant formation, quantifying the relative contribution of emission sources to ambient pollution levels, and assessing population exposures and the benefits of emission controls in impacted communities around the Bay Area. Our initial assessments focus on fine particulate matter ($PM_{2.5}$) concentrations in West Oakland, and follow-up analyses will include air toxics evaluations in West Oakland and expansion of our technical assessments to other communities.

For the $PM_{2.5}$ analyses, we evaluated ambient meteorological and air quality data, and applied the U.S. EPA's Community Multi-Scale Air Quality (CMAQ) model to simulate pollutant concentrations at a 1-km horizontal resolution over the entire Bay Area for 2016 (Figure 1.1). Then we repeated the simulation with West Oakland's anthropogenic emissions removed from the modeling inventory, leaving all other model input parameters unchanged. We calculated annual average $PM_{2.5}$ concentrations using the output of each simulation. The first simulation provided the annual average $PM_{2.5}$ concentrations for 2016 over the entire Bay Area, which will be used for $PM_{2.5}$ exposure analyses and health impacts assessments. The second simulation provided an estimate of background $PM_{2.5}$ levels in West Oakland (i.e., the $PM_{2.5}$ concentrations that would exist in the absence of local West Oakland sources).

Background $PM_{2.5}$ concentrations will be combined with local-scale modeling of West Oakland sources using the AERMOD dispersion model to provide a complete picture of $PM_{2.5}$ levels in the community and the relative contribution of different emission sources to those levels. Figure 1.2 shows the AERMOD modeling domain for West Oakland. The area outlined in blue represents the "source domain," and all significant emissions sources in that area will be modeled in the AERMOD simulations. The red hatched area represents the "receptor domain," or the area for which pollutant concentrations will be calculated by AERMOD.

The application of the CMAQ model involves the preparation of meteorological and emissions inputs, model runs, analysis of simulated pollutant concentrations, and the evaluation of model performance via comparison between simulated and observed pollutant concentrations. A simulation year of 2016 was selected because (1) this is a recent year that is likely to be representative of current conditions in West Oakland and other communities; and (2) special measurement studies that took place in 2016 provide additional ambient data to support evaluations of model performance.

District staff have been applying and evaluating the CMAQ model in the Bay Area over the last several years, along with the Weather Research and Forecasting (WRF) model, which provides meteorological inputs for CMAQ. Findings from previous modeling work are documented in a District report on $PM_{2.5}$ data analysis and modeling (Tanrikulu et al., 2009) and in the District's 2017 Clean Air Plan (BAAQMD, 2017). Both the CMAQ and WRF models were tested and evaluated for many cases in the Bay Area and their performance has been iteratively improved. The 2016 simulations used the best-performing configuration of the model. The 2016 emissions inputs have been updated to reflect ARB's most recent estimates and have been evaluated to the extent possible.

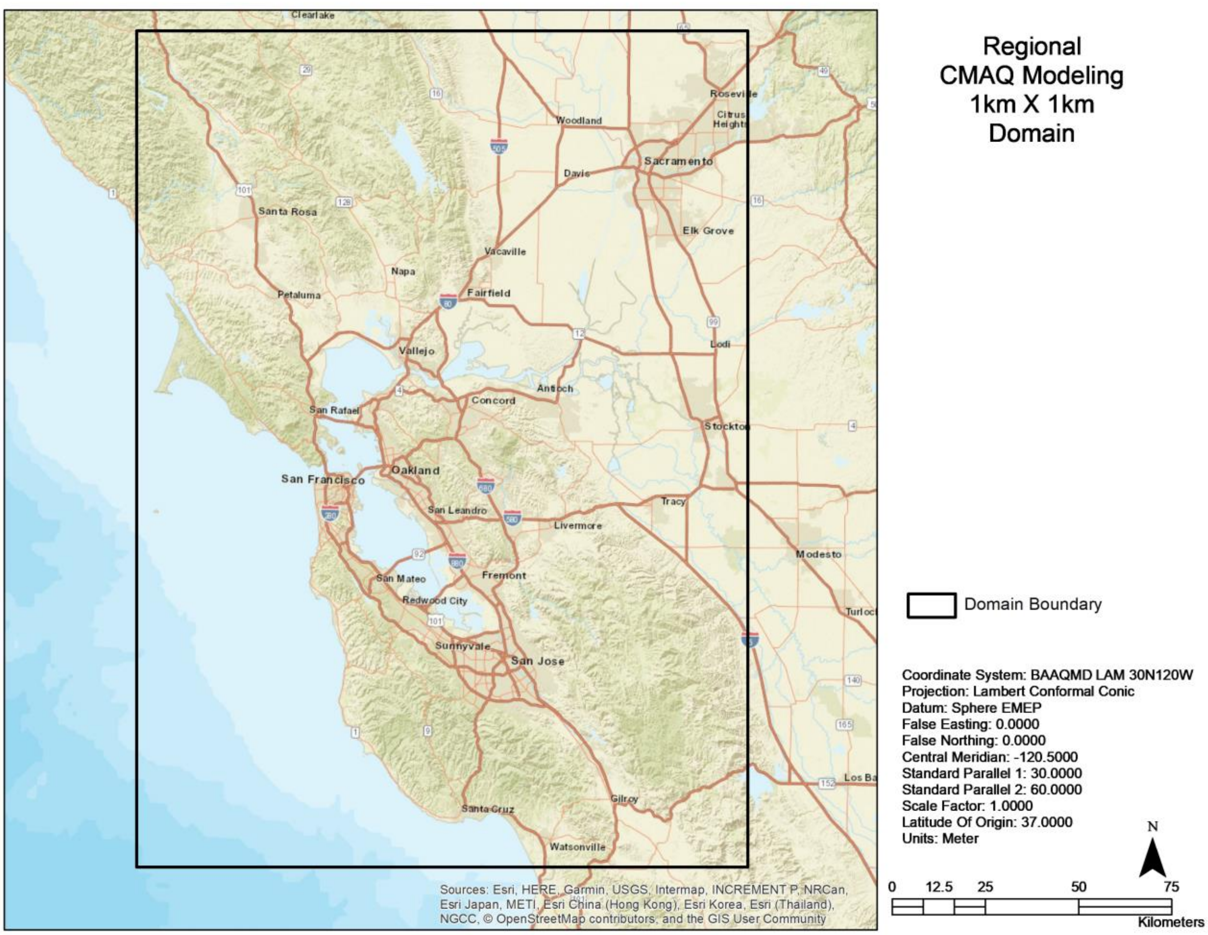


Figure 1.1: The regional 1-km modeling domain used for CMAQ simulations.

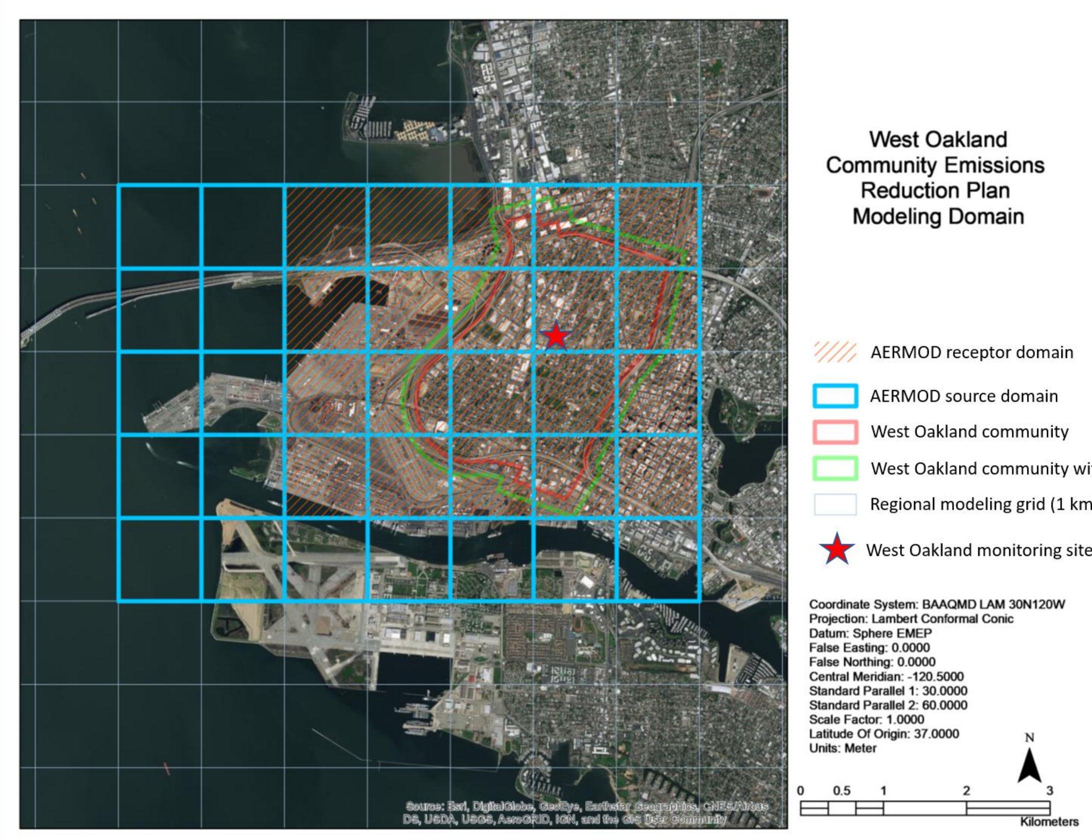


Figure 1.2: The West Oakland AERMOD modeling domain. The area outlined in blue represents the AERMOD source domain, and the red hatched area represents the AERMOD receptor domain.

1.1 $PM_{2.5}$ and Its Health Impacts

$PM_{2.5}$ is a complex mixture of suspended particles and liquid droplets in the atmosphere that have an aerodynamic diameter of 2.5 microns (µm) or less. An individual particle typically begins as a core or nucleus of carbonaceous material, often containing trace metals. These *primary* (directly emitted) particles usually originate from the incomplete combustion of fossil fuels or biomass. Layers of organic and inorganic compounds then deposit onto a particle, causing it to grow in size. These layers are largely comprised of *secondary* material that is not emitted directly. Secondary PM instead forms from chemical reactions of precursor gases released from combustion, agricultural activities, household activities, industrial sources, vegetation, and other sources. As a particle grows larger, gravity eventually causes it to be deposited onto a surface. Naturally emitted dust particles generally have diameters too large to be classified as $PM_{2.5}$.

Major human health outcomes resulting from $PM_{2.5}$ exposure include: aggravation of asthma, bronchitis, and other respiratory problems, leading to increased hospital admissions; cardiovascular symptoms, including chronic hardening of arteries and acute triggering of heart attacks; and decreased life expectancy, potentially on the order of years. Smaller particles have increasingly more severe impacts on human health as compared to larger particles. This occurs in part because smaller particles can penetrate more deeply into the human body. For the Bay Area, public health impacts from $PM_{2.5}$ may well exceed the combined impacts of all other currently regulated air pollutants.

District staff have previously evaluated the health and monetary impacts of $PM_{2.5}$ concentrations in the Bay Area for 2010. Findings of this evaluation are documented in a report by Tanrikulu, et al. (2011).

1.2 Formation of $PM_{2.5}$ in the Bay Area

In the Bay Area, $PM_{2.5}$ concentrations can build up during winter months (December, January and February) under stable atmospheric conditions that trap pollutants near the ground. Winters with frequent stagnant periods tend to have a higher number of days with elevated $PM_{2.5}$ than winters with more periods of windy and stormy conditions. Consecutive stagnant, clear winter days are typically required for $PM_{2.5}$ episodes to develop. $PM_{2.5}$ episodes are regional in nature and impact most Bay Area locations.

The Chemical Mass Balance (CMB) model was previously applied for $PM_{2.5}$ source apportionment using specialized measurements mostly obtained during the years 1999-2014. CMB is a statistical receptor model that uses speciated $PM_{2.5}$ measurements to estimate the contribution of individual source categories to observed $PM_{2.5}$ levels. CMB analyses for the Bay Area showed that primary combustion sources (both fossil fuels and biomass) were the largest $PM_{2.5}$ contributors in all seasons. The biomass combustion contribution to peak $PM_{2.5}$ levels was about 2-4 times higher during winter than for other seasons. Secondary $PM_{2.5}$ levels were mostly elevated during the winter months, with ammonium nitrate being the key component of wintertime secondary $PM_{2.5}$. This semi-volatile $PM_{2.5}$ component is stable in its solid form during the cooler winter months. Secondary ammonium sulfate $PM_{2.5}$ levels were generally low ($< 1\text{-}2\ \mu g/m^3$) but non-negligible. Sea salt, geological dust, and tire and brake wear contributed minimally to $PM_{2.5}$ concentrations (Tanrikulu et al., 2009).

Meteorological cluster analysis, a data mining technique, was implemented to determine how weather patterns impact $PM_{2.5}$ levels. Clustering was applied to measurements from every winter day across more than 10 years. This method provided a robust representation of how prevailing weather conditions affected the development of $PM_{2.5}$ episodes. Such episodes generally developed under: stable atmospheric conditions inhibiting vertical dispersion; clear and sunny skies favoring enhanced secondary $PM_{2.5}$ formation; and pronounced overnight drainage (downslope) flows off the Central Valley rims, causing low-level air in the Central Valley to empty through the Delta and into the Bay Area along its eastern boundary. Atmospheric transitions of aloft weather systems profoundly influenced the surface winds that determine $PM_{2.5}$ levels. Surface conditions stagnated whenever an upper-level high pressure system moved over Central California. Persisting high pressure conditions allowed $PM_{2.5}$ buildup, and Bay Area 24-hour elevated $PM_{2.5}$ generally occurred after 2-4 days.

A refined cluster analysis further characterized the upwind Central Valley conditions during Bay Area episodes. Two distinct inter-regional air flow patterns were associated with different types of Bay Area episodes. Most elevated PM days were associated with winds from the Sacramento Valley to the northeast entering the Bay Area through the Delta. Peak $PM_{2.5}$ levels typically occurred along the Delta and at San Jose for this type of episode. A minority of elevated PM days were associated with winds from the San Joaquin Valley from the southeast entering the Bay Area through the Delta. Peak $PM_{2.5}$ levels typically occurred along the Delta and in the East Bay (at Livermore, Concord, Vallejo or San Rafael, and to a lesser degree at Oakland and San Francisco) for this type of episode. The remaining relatively moderate episodes could not be associated with any distinct inter-regional transport pattern linking the Bay Area and surrounding air basins.

2. Observations and Data Analysis

2.1 Ambient Measurements

Both meteorological and air quality data have been continuously collected in the Bay Area and surrounding regions for many years. In 2016, there were twenty-six PM monitoring stations within the 1-km modeling domain - sixteen in the Bay Area and ten outside the region. Table 2.1 lists PM monitoring stations used in this study with their annual and quarterly average $PM_{2.5}$ values. Figure 2.1 shows the spatial distribution of monitored annual average $PM_{2.5}$ concentrations for 2016. A complete list of monitoring stations, types of measurements, and the purpose of their use in this study is provided in Appendix A. The air quality monitoring network plan published by BAAQMD (Knoderer et al., 2017) provides additional details on the District's monitoring network.

All ambient data used in this study were subjected to quality assurance checks and validated prior to being used. These data were used for the development of a conceptual model of PM formation in the region, establishment of relationships among emissions, meteorology and air quality, evaluation of models, and four-dimensional data assimilation (FDDA), in which meteorological observations are used by the meteorological model to "nudge" simulations toward observations.

Hourly average data are used for most analyses and model evaluation, but monthly, quarterly or annual averages are presented here for brevity.

2.2 Data Analysis

In 2016, the annual average $PM_{2.5}$ concentrations (Table 2.1) at two Bay Area air monitoring stations (Sebastopol and Gilroy) were between 5 $\mu g/m^3$ and 6 $\mu g/m^3$. These two sites captured the lowest $PM_{2.5}$ levels in the Bay Area. At three other air monitoring stations (Concord, Oakland and San Rafael), $PM_{2.5}$ concentrations were between 6 $\mu g/m^3$ and 7 $\mu g/m^3$, and at four other stations (Berkeley Aquatic Park, Livermore, San Francisco and Vallejo), they were between 7 $\mu g/m^3$ and 8 $\mu g/m^3$. At the remaining seven stations (Napa, San Pablo, Laney College, Oakland West, Redwood City, San Jose - Jackson and San Jose - Knox Avenue), $PM_{2.5}$ levels were above 8 $\mu g/m^3$. San Jose - Knox Avenue had the highest Bay Area annual average $PM_{2.5}$ concentration (9.2 $\mu g/m^3$).

Outside of Napa, the stations with annual average $PM_{2.5}$ concentrations above 8 $\mu g/m^3$ extend from the north Bay to the south Bay. Previous analyses showed that $PM_{2.5}$ levels at these locations were influenced by local sources and the transport of pollutants from the Central Valley. Elevated concentrations at Napa are mostly due to local residential wood burning and the transport of PM from both residential wood burning and wildfire emissions.

While $PM_{2.5}$ levels at several Bay Area stations, such as Laney College, West Oakland and Livermore, showed little change from one quarter to another, another set of stations (including Napa, Vallejo and San Francisco) had significant differences between quarters (Table 2.1). These stations are impacted by transport and seasonal changes in meteorology and/or emissions, such as wood burning.

Table 2.1: PM stations in the 1-km modeling domain with their annual and quarterly average $PM_{2.5}$ values.

Station Name	**$PM_{2.5}$ Averages ($\mu g/m^3$) for 2016**				
Stations in the Bay Area	**ANNUAL**	**QTR_01**	**QTR_02**	**QTR_03**	**QTR_04**
Berkeley Aquatic Park	7.2	--[a]	--[a]	7.7	6.6
Concord	6.2	6.0	4.3	4.6	9.4
Gilroy	5.7	5.9	6.1	6.8	4.1
Laney College	8.8	8.9	9.4	8.7	8.1
Livermore	7.6	7.4	7.2	8.4	7.3
Napa	8.9	6.5	7.2	10.4	11.1
Oakland	6.2	5.2	5.9	6.4	7.2
Oakland West	8.7	9.6	8.9	7.6	8.6
Redwood City	8.7	6.8	10.3	10.6	6.7
San Francisco	7.8	8.5	8.1	5.9	8.4
San Jose - Jackson	8.3	8.0	8.0	8.8	8.4
San Jose - Knox Avenue	9.2	9.0	8.6	9.9	9.2
San Pablo	8.1	7.6	8.9	7.8	8.2
San Rafael	6.6	7.0	6.1	5.9	7.1
Sebastopol	5.1	4.9	4.6	4.0	6.5
Vallejo	7.6	8.4	5.6	6.0	10.2
Stations outside the Bay Area					
Manteca	9.9	10.8	7.5	8.8	12.3
San Lorenzo Valley Middle School	5.3	5.4	5.2	4.7	5.8
Roseville - N Sunrise Ave	6.8	6.7	5.7	6.7	8.3
Sacramento Health Department - Stockton Blvd.	6.9	7.8	5.7	6.6	8.3
Sacramento - 1309 T Street	7.6	7.2	5.6	7.1	10.9
Sacramento - Bercut Drive	--[a]	--[a]	--[a]	--[a]	14.6
Sacramento - Del Paso Manor	8.7	8.6	6.1	7.2	13.2
Santa Cruz	5.4	5.8	5.9	5.3	4.5
Stockton - Hazelton	11.8	13.9	8.2	10.0	15.2
Woodland - Gibson Road	6.3	5.2	5.4	8.1	6.9

[a]Data missing or invalidated.

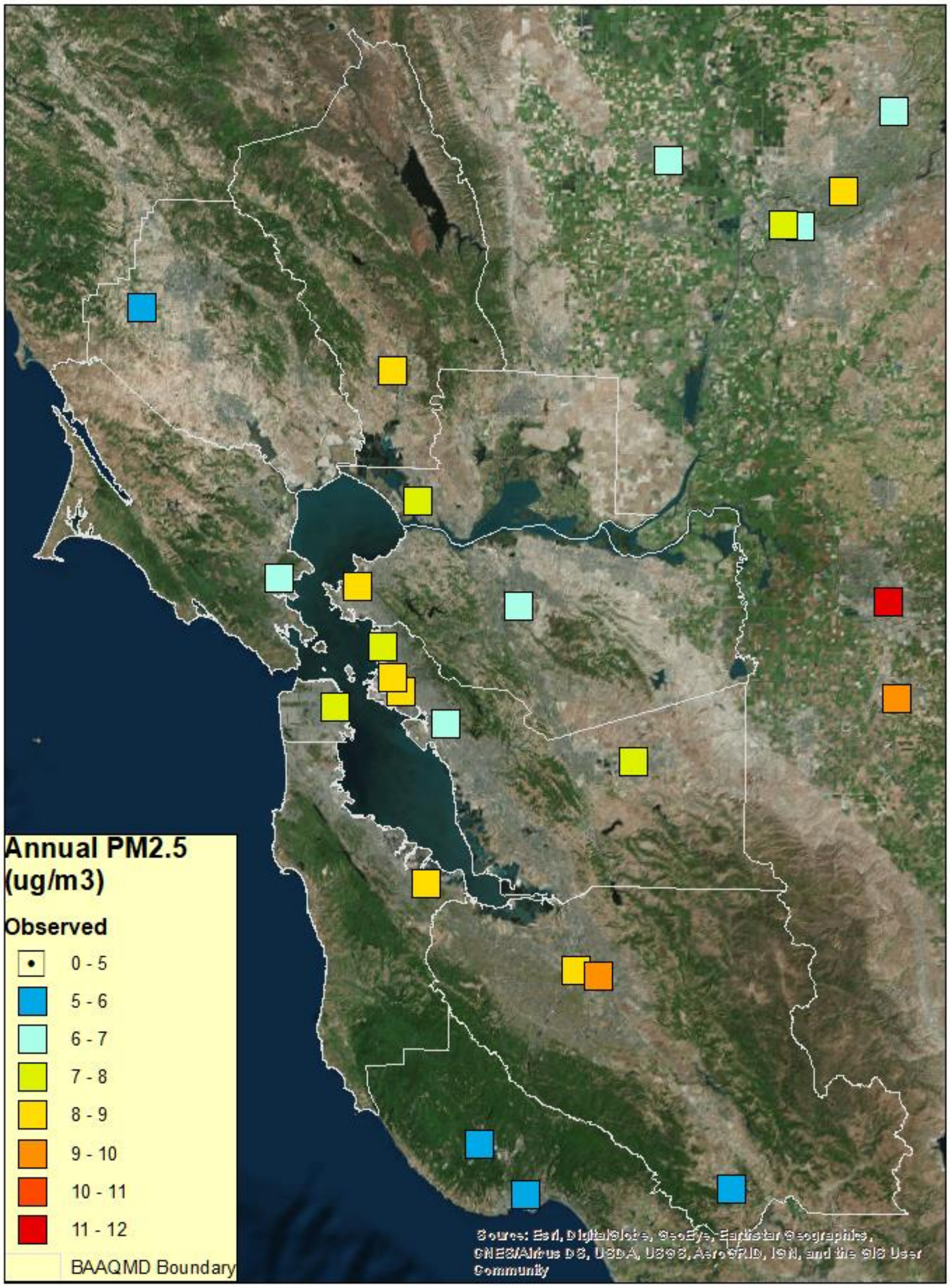


Figure 2.1: Spatial distribution of observed annual average $PM_{2.5}$ concentrations for 2016 within the 1-km modeling domain.

3. Modeling

3.1 Emissions Inventory Preparation

The 2016 modeling emissions inventory includes estimates for area sources,[2] point sources, onroad mobile sources, nonroad mobile sources, and biogenic sources. The inventory was assembled from a variety of data sources, including the District's in-house emissions estimates, emissions data from ARB, and outputs from ARB's EMFAC2017 model (see Table 3.1). ARB emissions data were used for all anthropogenic sources in non-BAAQMD counties with the exception of onroad mobile sources. County- and facility-level ARB emissions data for the entire state of California were downloaded from ARB's FTP site in June 2018. These data were in SMOKE-ready format and were consistent with the current version of ARB's online repository of emissions inventory data that was developed to support the preparation of ozone State Implementation Plans (SIPs) for non-attainment areas in California. At the time of this work, the latest version of ARB's SIP inventory was version 1.05, which was prepared for a base year of 2012 and projected to various future years, including 2016.[3]

For area sources, ARB's county-level emissions estimates for residential wood combustion in BAAQMD counties were adjusted to account for the impact of the District's winter Spare the Air program, which prohibits wood burning when air quality is forecast to be unhealthy. This adjustment was based on survey-based wood combustion emissions estimates developed by District staff, comparisons with wood combustion emissions estimates for other air districts, and discussions with ARB staff. Additional details in residential wood combustion emissions are provided in Appendix B.

For onroad mobile sources, ARB's EMFAC2017 model was run for the entire state of California for each month in 2016 to produce county-level, month-specific emissions estimates. EMFAC2017 reports emissions by vehicle type and emission mode (e.g., idling, running exhaust, brake wear, tire wear). EMFAC2017 outputs were converted to SMOKE-ready format using a Perl script developed by BAAQMD staff.

For point sources, ARB emissions estimates for BAAQMD counties were replaced by detailed, in-house data prepared in California Emission Inventory Development and Reporting System (CEIDARS) format. These point source emissions data were representative of calendar year 2012 and were projected to 2016 using industry-specific growth factors from ARB's California Emissions Projection Analysis Model (CEPAM). The CEIDARS data were converted to SMOKE-ready format using a Perl script developed by BAAQMD staff.

Biogenic emissions estimates were prepared using EPA's Biogenic Emission Inventory System (BEIS), version 3.61, which estimates emissions from vegetation and soil using land use data;

---

[2] Area sources are stationary sources such as dry cleaners that are too small or numerous to treat as individual point sources.
[3] See http://www.arb.ca.gov/app/emsinv/2016ozsip/2016ozsip/.

vegetation-specific emission rates for isoprene and other species; and gridded, hourly meteorological data from the WRF model.

3.1.1 SMOKE Processing

Emissions inventory data assembled from the sources described above were processed through version 4.5 of the SMOKE emissions processor to develop CMAQ-ready emissions inputs for each day of 2016. SMOKE uses the processing steps described below to convert "raw" emissions inputs to the spatial, temporal, and chemical resolution required by CMAQ or an equivalent air quality model.

Table 3.1: Summary of data sources used to develop the 2016 modeling inventories.

Region	Source Sector	Data Source
BAAQMD Counties[a]	Area	ARB county-level emissions estimates, with adjustments made to residential wood combustion emissions
	Nonroad	ARB county-level emissions estimates
	Onroad	County-level, month-specific EMFAC2017 outputs
	Point	In-house CEIDARS data
	Biogenic	Hourly outputs from EPA's BEISv3.61 model
Non-BAAQMD Counties	Area	ARB county-level emissions estimates
	Nonroad	ARB county-level emissions estimates
	Onroad	County-level, month-specific EMFAC2017 outputs
	Point	ARB facility-level emissions estimates
	Biogenic	Hourly outputs from EPA's BEISv3.61 model

[a]Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, and Santa Clara counties, plus the southern portion of Sonoma County and the western portion of Solano County.

<u>Spatial allocation</u>

SMOKE assigns county- or facility-level emissions to individual grid cells in the modeling domain. For point sources, emissions are assigned to grid cells based on the location coordinates (i.e., latitude and longitude) of emission release points. For county-level area, nonroad, or onroad emissions estimates, SMOKE allocates emissions to grid cells using spatial allocation factors developed from "surrogate" geospatial data sets such as land use or socioeconomic data. Geospatial data sets used to develop the surrogates used in SMOKE include land use data from the Association of Bay Area Governments (ABAG) (Reid, 2008). For counties in the District's jurisdiction, gridded surrogate data were available at 1-km grid resolution. However, for counties outside the District, only 4-km surrogate data were available, so these data were parsed to create a set of pseudo 1-km surrogates.

<u>Temporal allocation</u>

SMOKE assigns annualized or average day emissions to the specific dates and hours being modeled using temporal profiles that reflect source-specific activity patterns by month, day of week, and hour of day. Temporal profiles from ARB's CEIDARS database were used in SMOKE to temporally allocate area, point, and nonroad mobile source emissions. For onroad mobile sources, temporal profiles that ARB developed from its California Air Resources Board Vehicle Activity Database (CalVAD) were used.[4]

Chemical speciation

SMOKE disaggregates total organic gas (TOG) and $PM_{2.5}$ emissions into a series of model species that CMAQ uses to represent atmospheric chemistry. For the 2016 CMAQ modeling, speciation profiles developed for the SAPRC07 chemical mechanism were applied to TOG emissions from all sources, and profiles developed for the AERO6 aerosol module were applied to $PM_{2.5}$ emissions from all sources.

The SMOKE system includes an implementation of EPA's Biogenic Emission Inventory System (BEIS), version 3.61, which estimates biogenic emissions using land use data; vegetation specific emission rates for isoprene and other species; and gridded, hourly meteorological data from the WRF model. BEISv3.61 was run within SMOKE to prepare 2016 biogenic emissions estimates for the 1-km modeling domain.

Once SMOKE runs were completed, a number of quality assurance checks were performed on the resulting emissions data. First, plots of gridded emissions were generated to examine the spatial distribution of emissions. Similarly, diurnal plots were generated to examine hourly variations in emissions by source sector and to ensure that the patterns make sense. In addition, SMOKE's SMKREPORT utility was used to generate tabular summaries of emissions by pollutant, county, grid cell, hour, and source category. This information was used in a variety of ways, including:

- Comparing emissions before and after key processing steps to ensure that any changes in the mass of emissions make sense. For example, for counties that only partially lie within the modeling domain, total emissions should decrease after the gridding step.
- Sorting emissions by source category code (SCC) or facility ID to identify key contributors to total emissions for each pollutant and to identify potential outliers.
- Summarizing emissions by pollutant and county to ensure that geographic distributions make sense (e.g., SO_2 emissions are highest in Contra Costa County where refineries are concentrated).
- Extracting emissions for grid cells in the West Oakland AERMOD modeling domain to identify key sources and compare emissions by source sector with the District as a whole.

These checks identified several issues, including $PM_{2.5}$ hotspots at two landfills in eastern Alameda County. Our modelers worked with staff from the District's Emissions & Community Exposure Assessment section to correct the emissions from these and other point sources.

[4] The CalVAD database fuses available data sources such as Caltrans Weigh-in-Motion (WIM) data and Highway Performance Monitoring System (HPMS) data to produce a best estimate of vehicle activity by class.

3.1.2 Emissions Summaries

This subsection provides emissions density plots and summary tables for $PM_{2.5}$, and similar information for additional pollutants can be found in Appendix B. Figure 3.1 shows annual average $PM_{2.5}$ emissions for the 1-km modeling domain. Table 3.2 summarizes the annual average $PM_{2.5}$ emissions by county and source sector, as reported by the SMOKE emissions model. Within the District's jurisdiction, annual average $PM_{2.5}$ emissions total 33.7 tons per day (tpd). The area source sector accounts for about half of this total (17 tpd), and individual source categories that are key contributors to total $PM_{2.5}$ emissions include residential wood combustion, fugitive dust from roadways and construction sites, and commercial cooking.

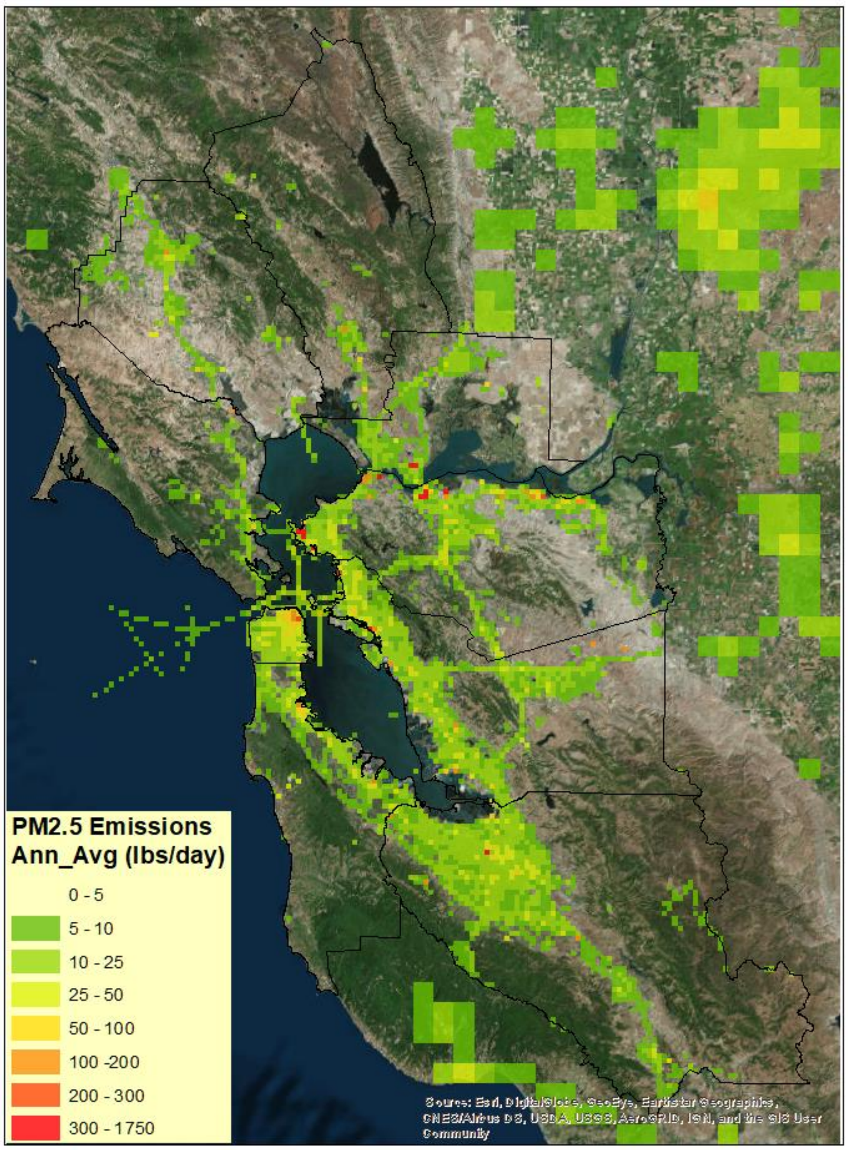


Figure 3.1: Spatial distribution of annual average $PM_{2.5}$ emissions for the 1-km modeling domain.

Table 3.2: Summary of 2016 $PM_{2.5}$ anthropogenic emissions (tpd) by geographic area and source sector.

Geographic Area	**Area**	**Nonroad**	**Onroad**	**Point**	**Total**
Alameda	3.0	0.5	1.4	1.3	6.2
Contra Costa	3.1	0.5	0.8	4.2	8.7
Marin	0.8	0.2	0.2	0.1	1.3
Napa	0.8	0.2	0.1	0.1	1.2
San Francisco	1.2	1.0	0.3	0.1	2.7
San Mateo	1.4	0.5	0.5	0.4	2.7
Santa Clara	3.9	0.6	1.3	0.7	6.5
Solano[a]	1.3	0.1	0.3	0.5	2.1
Sonoma[a]	1.4	0.3	0.3	0.2	2.2
BAAQMD Subtotal	*17.0*	*3.9*	*5.2*	*7.5*	*33.7*
Non-BAAQMD Counties	23.7	2.2	2.9	2.4	31.2
Domain Total	**40.7**	**6.1**	**8.0**	**9.9**	**64.9**

[a]Emissions totals for Solano and Sonoma counties only include the portion of those counties in BAAQMD's jurisdiction.

For the West Oakland AERMOD modeling domain, annual average $PM_{2.5}$ emissions total 0.35 tpd, or about 1% of the BAAQMD total. Figure 3.2 shows that the distribution of emissions by source sector in West Oakland differs from the District as a whole. In West Oakland, onroad and nonroad mobile sources account for 66% of total $PM_{2.5}$ emissions, while the same sources only account for 27% of total $PM_{2.5}$ emissions districtwide. Figure 3.3 shows the spatial distribution of $PM_{2.5}$ emissions across the 1-km grid cells that coincide with the local-scale AERMOD modeling domain.

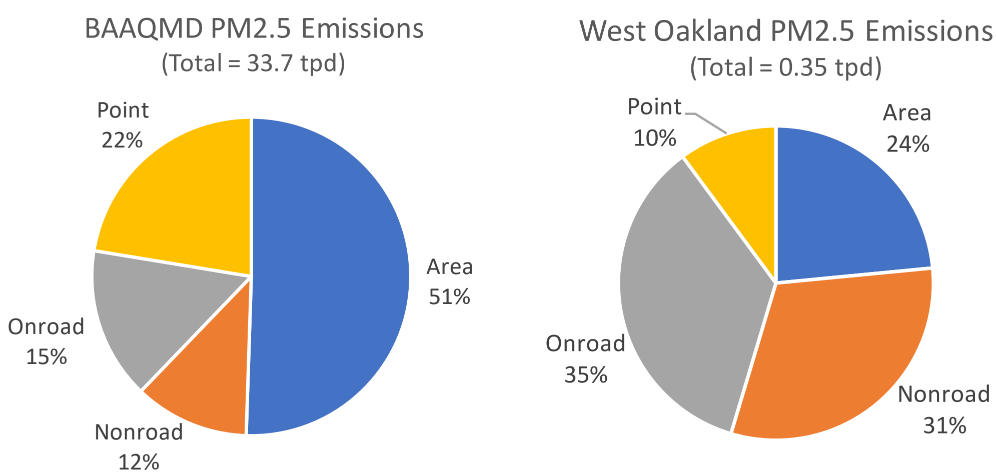


Figure 3.2: $PM_{2.5}$ emissions by source sector for the District (left) and West Oakland (right).

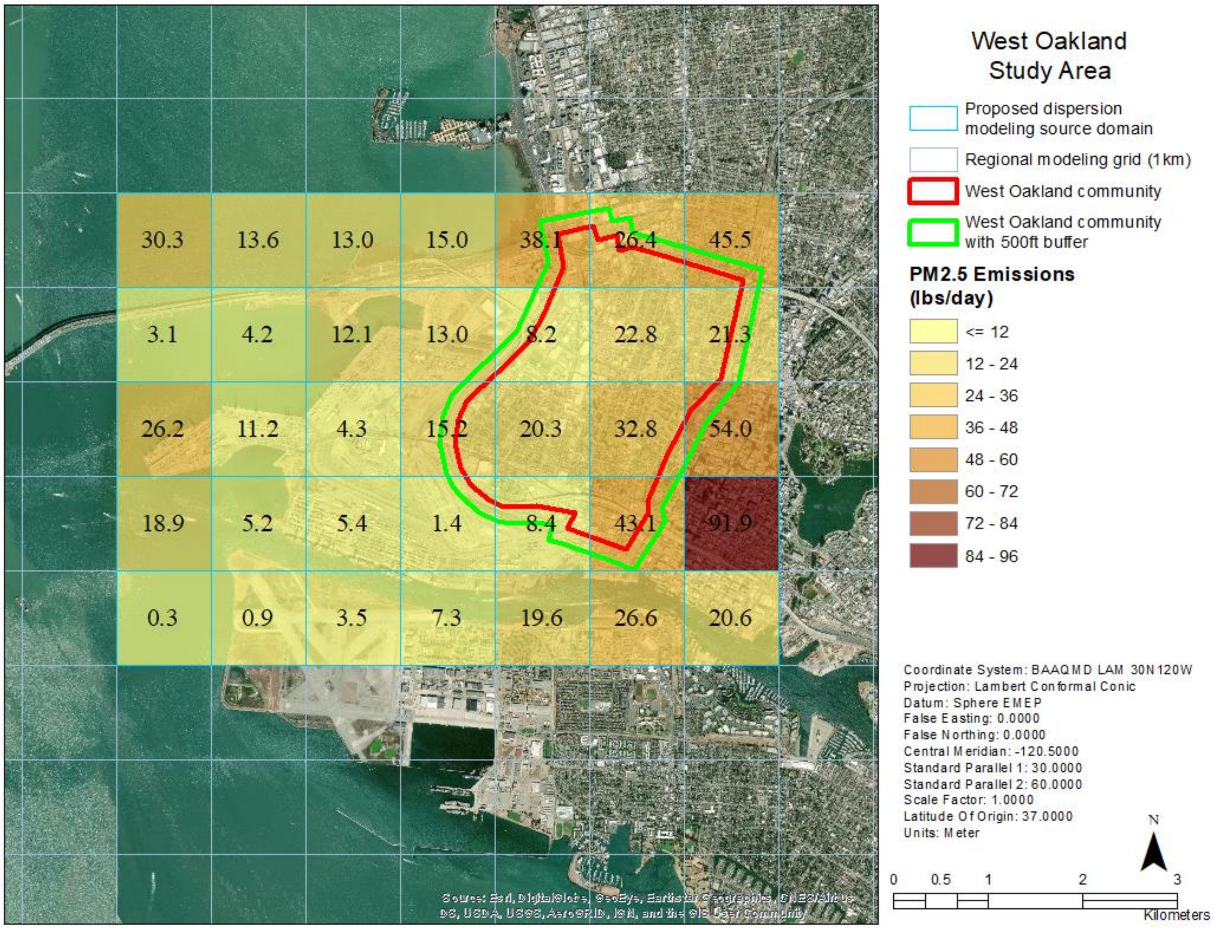


Figure 3.3: Spatial distribution of annual average $PM_{2.5}$ emissions in West Oakland.

3.2 Meteorological Modeling

The Weather Research and Forecasting (WRF) Model version 3.8 was used to prepare meteorological inputs to CMAQ. Four nested modeling domains were used (Figure 3.4). The outer domain covered the entire western United States at 36-km horizontal grid resolution to capture synoptic (large-scale) flow features and the impact of these features on local meteorology. The second domain covered California and portions of Nevada at 12-km horizontal resolution to capture mesoscale (sub-regional) flow features and their impacts on local meteorology. The third domain covered Central California at 4-km resolution to capture localized air flow features. The 4-km domain included the Bay Area, San Joaquin Valley, and Sacramento Valley, as well as portions of the Pacific Ocean and the Sierra Nevada mountains. The fourth domain covered the Bay Area and surrounding regions at 1-km resolution. All four domains employed 50 vertical layers with thickness increasing with height from the surface to the top of the modeling domain (about 18 km).

Meteorological variables are estimated at the layer midpoints in WRF. The thickness of the lowest layer nearest the surface was about 25 m. Thus, meteorological variables near the surface were estimated for a height of about 12.5 m above ground level. The model configuration was tested using available physics options, including: (1) planetary boundary layer processes and time-based evolution of mixing heights; (2) choice of input database for WRF; (3) four-dimensional data assimilation (FDDA) strategy; (4) horizontal and vertical diffusion; (5) advection scheme; and (6) initial and boundary conditions. The final choice of options was the one proved to best characterize meteorology in the domain.

WRF was applied for 2016 to estimate parameters required by the air quality model, including hourly wind speed and direction, temperature, humidity, cloud cover, rain and solar radiation levels. Observations are assimilated into the model during the simulations to minimize the difference between simulations and real-world measurements. Two types of nudging methods were employed (analysis and observation). The NCEP North America Mesoscale (NAM) 12-km analyzed meteorological fields were used for analysis nudging as well as for initializing the model. The NCEP ADP Global Surface and Upper Air Observational Weather Data were used for observational nudging. A list of these stations for the 1-km domain is given in Appendix A.

The analysis nudging was applied to the 36-km and 12-km domains. Frequency of surface analysis nudging was every three hours, while the frequency of 3D analysis nudging was every six hours. The 3D analysis nudging of winds was performed over all model layers, but the 3D analysis nudging of temperature and humidity was limited to layers above the planetary boundary layer. The observation nudging of wind was applied to all four domains every three hours.

The WRF model was rigorously evaluated for accuracy. Observations used to evaluate WRF were taken from the EPA's Air Quality System, the BAAQMD meteorological network, and the National Climate Data Center. A list of these stations for the 1-km domain is given in Appendix A. Hourly and daily time series plots of observed and simulated wind, temperature and humidity were generated at each observation station and compared to each other hour by hour and day by day. Simulated hourly areal plots of wind, temperature, humidity, planetary boundary layer height, pressure and other fields were generated and quantitatively compared against observations where observations were available.

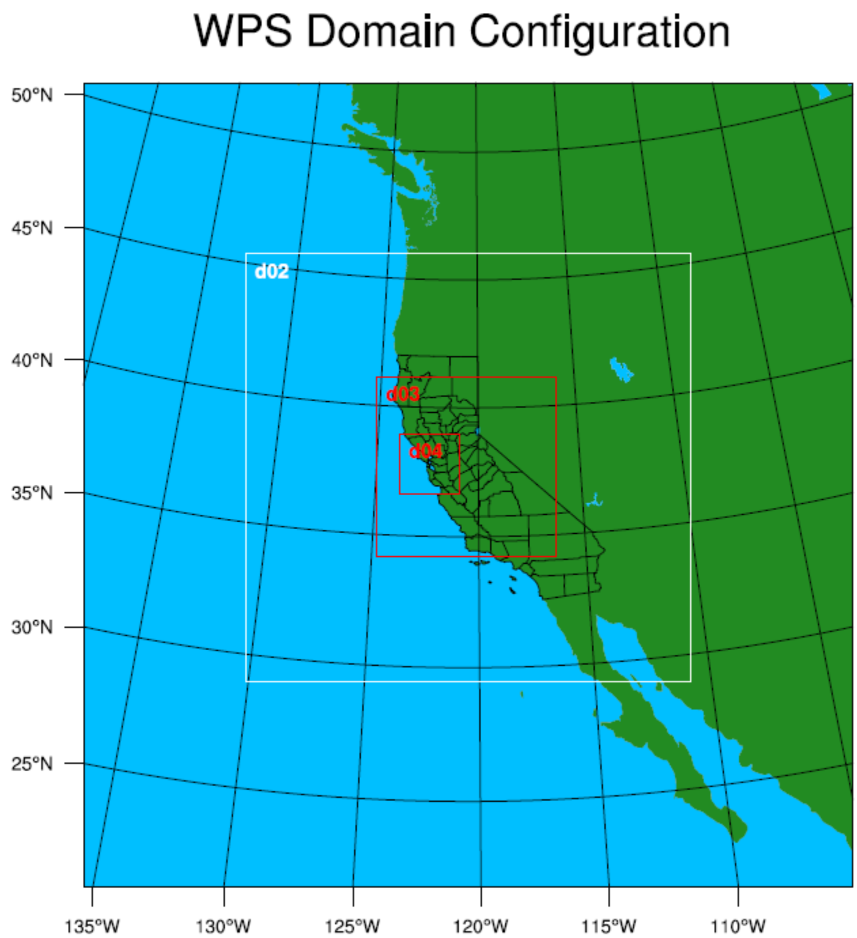


Figure 3.4: Nested WRF modeling domains.

These plots were also qualitatively evaluated for known meteorological features of the modeling domain, especially at 4-km and 1-km resolutions. These features include slope flows, channeled flows, sea breeze and low-level jet. The vertical profile of observed and simulated meteorological fields was compared at several upper air meteorological stations, including Oakland, Medford, Reno and Las Vegas, and at a temporary station established at Bodega Bay. RAMBOLL's METSTAT program (Emery et al., 2001) was used to statistically evaluate the performance of WRF. The statistical metrics used in this evaluation are defined in Appendix C.

The WRF model performed reasonably well in every evaluation category. The estimated bias, gross error, root mean square error (RMSE), and index of agreement (IOA) are within established criteria for acceptable model performance for every day of 2016. In other words, performance obtained from the Bay Area applications of WRF is similar or slightly better than performance obtained from applications elsewhere, available from literature.

Performance statistics were generated for all days of 2016. Samples from ten winter (December 1-10) and ten summer (July 1-10) days are shown in Table 3.3, while Table 3.4 shows maximum and minimum skill scores of WRF for all of 2016.

Simulated and observed wind speed, wind direction, temperature and humidity, for both hourly and annual time periods, were compared at each station. Time series comparisons for West Oakland, Vallejo and San Jose are shown in Appendix C. These three stations were selected because West Oakland is the area of interest for this study. Vallejo is strategically located in the Delta to capture air flow between the Bay Area and the Central Valley. San Jose is an important sub-region of the Bay Area representing outflows. Good model performance at Vallejo and San Jose is critical to simulate representative meteorological features in West Oakland.

The WRF model was also compared against upper air measurements at Oakland, a site operated by the National Weather Service with twice daily upper air measurements and at Bodega Bay, a temporary site established by the California Baseline Ozone Transport Study with daily ozonesonde and meteorological measurements. Simulated winds, temperatures and humidity matched these upper air measurements very well. Details can be found in Appendix C.

Table 3.3: Sample (December 1-10 and July 1-10) statistical scores of WRF.

Parameter	**Metric**	**Units**	**12/01**	**12/02**	**12/03**	**12/04**	**12/05**	**12/06**	**12/07**	**12/08**	**12/09**	**12/10**
Wind Speed	Bias	(m/s)	-0.21	-0.43	-0.21	-0.68	-0.35	-0.38	-0.67	-0.32	-0.35	-0.77
Wind Speed	Gross Error	(m/s)	1.28	1.61	1.05	1.08	0.91	1.13	1.16	1.49	0.93	1.33
Wind Speed	RMSE	(m/s)	1.68	2.07	1.4	1.5	1.22	1.49	1.48	1.88	1.19	1.7
Wind Speed	IOA	--[a]	0.78	0.75	0.64	0.67	0.66	0.73	0.71	0.7	0.66	0.79
Wind Direction	Bias	(deg)	2.21	-1.32	-6.78	3.74	4.73	13.09	3.22	5.08	-6.37	6.45
Wind Direction	Gross Error	(deg)	32.17	24.42	57.64	45.58	42.43	43.08	32.72	50.97	38.39	22.33
Temperature	Bias	(K)	1.26	1.25	2.6	0.85	0.35	1.76	2.08	0.75	-0.31	-0.68
Temperature	Gross Error	(K)	32.17	1.72	2.66	1.47	1.18	1.85	2.26	1.64	1.43	0.92
Temperature	RMSE	(K)	2.09	2.25	3.2	1.87	1.53	2.41	2.71	2.09	1.82	1.15
Temperature	IOA	--[a]	0.93	0.91	0.87	0.94	0.94	0.88	0.82	0.89	0.88	0.87
Parameter	**Metric**	**Units**	**7/01**	**7/02**	**7/03**	**7/04**	**7/05**	**7/06**	**7/07**	**7/08**	**7/09**	**7/10**
Wind Speed	Bias	(m/s)	-0.96	-1.12	-0.96	-1.15	-1.04	-0.95	-1.08	-0.91	-1.27	-0.91
Wind Speed	Gross Error	(m/s)	1.35	1.48	1.37	1.53	1.44	1.43	1.51	1.41	1.7	1.44
Wind Speed	RMSE	(m/s)	1.81	1.91	1.77	2.02	1.99	1.86	2.01	1.8	2.26	1.88
Wind Speed	IOA	--[a]	0.71	0.76	0.75	0.71	0.71	0.75	0.71	0.7	0.71	0.74

Wind Dir	Bias	(deg)	-0.16	0.36	-3.34	-1.97	-3.95	-1.52	0.97	3.73	3.51	4.82
Wind Dir	Gross Error	(deg)	28.1	27.05	24.63	27.85	24.76	23.07	20.44	21.91	21.83	28.44
Temperature	Bias	(K)	1.3	1.47	0.24	0.39	1.2	1.09	0.87	0.34	0.06	0.59
Temperature	Gross Error	(K)	1.89	1.9	1.21	1.14	1.58	1.36	1.12	1.14	0.87	1.17
Temperature	RMSE	(K)	2.43	2.34	1.56	1.48	1.99	1.69	1.39	1.43	1.09	1.54
Temperature	IOA	-- [a]	0.97	0.96	0.98	0.98	0.96	0.97	0.98	0.98	0.99	0.98

[a]The Index of Agreement (IOA) is a dimensionless quantity.

Table 3.4: Maximum and minimum statistical scores of WRF for 2016.

Parameter	Metric	Units	Max	Min
Wind Speed	Bias	(m/s)	1.02	-0.99
Wind Speed	Gross Error	(m/s)	2.03	0.57
Wind Speed	RMSE	(m/s)	2.5	0.76
Wind Speed	IOA	--	0.92	0.4
Wind Direction	Bias	(deg)	14.84	-12.56
Wind Direction	Gross Error	(deg)	81.37	13.5
Temperature	Bias	(K)	2.99	-1.62
Temperature	Gross Error	(K)	3.24	0.72
Temperature	RMSE	(K)	5.22	0.9
Temperature	IOA	--	0.98	0.47

3.3 Air Quality Modeling

Air quality modeling was conducted using the U.S. EPA's Community Multiscale Air Quality (CMAQ) modeling system version 5.2. Two nested domains were used. The outer domain coincides with the third domain of the meteorological model and covers the Bay Area, San Joaquin Valley, and Sacramento Valley, as well as portions of the Pacific Ocean and the Sierra Nevada mountains at 4-km horizontal resolution. The inner domain covers the Bay Area and surrounding regions at 1-km horizontal resolution.

Both CMAQ modeling domains had 28 vertical layers. Below 1,500 m, the CMAQ layers match the WRF layers, while some upper-level meteorological model layers above 1,500 m were collapsed while preparing meteorological inputs for CMAQ to reduce computational time. This is a common practice in air quality modeling, as pollutant levels in layers aloft are relatively low and do not significantly impact concentrations at the surface. The thickness of CMAQ model layers was also increasing with height from the surface to the top of the modeling domain (about 18 km). The thickness of the first layer of CMAQ was kept the same as in WRF (about

25 m), meaning that pollutant concentrations are estimated at around 12.5 m above the surface (the midpoint of the first layer).

The outer domain provides initial conditions and hourly boundary conditions to the 1-km domain. As a result, the inner domain accounts for the contribution of emissions sources outside of the Bay Area to Bay Area PM levels. The outer domain was initialized and its boundary conditions were updated every six hours using outputs from a global air quality model (MOZART), available from the National Center for Atmospheric Research (NCAR).

CMAQ simulates both primary and secondary $PM_{2.5}$, with secondary $PM_{2.5}$ formation being dependent upon photochemistry. The chemical mechanism used in this simulation was the Statewide Air Pollution Research Center version 2007 (SAPRC-07) mechanism. Secondary PM formulation was simulated using the Models-3 AE6 aerosol module.

Each month was simulated separately to distribute simulations over 12 computer nodes for computational efficiency. Each monthly simulation includes the last three days of the previous month as a spin-up period except for January, which includes the last five days of December 2015. Model outputs from the spin-up periods were not used in analyses.

3.3.1 CMAQ Evaluation

The CMAQ model was rigorously evaluated for accuracy. Observations used to evaluate CMAQ were taken from the District's Data Management System and the EPA's Air Quality System. Hourly and daily time series plots of observed and simulated $PM_{2.5}$ concentrations were generated at each observation station and compared to each other hour by hour and day by day. This evaluation also provided an opportunity to identify gaps in measurements and outliers. Hourly, daily, monthly, quarterly and annual average spatial plots of PM and precursor concentrations were generated for observed and simulated values, and simulated values were quantitatively compared against observations where observations were available.

These plots were also qualitatively evaluated for known air quality features that may be impacted by meteorology, emissions, chemistry and other environmental parameters. Examples include local and regional transport of pollutants, proximity of polluted areas to emission sources such as freeways, and the behavior of atmospheric chemistry.

Various statistical metrics were used to evaluate the performance of CMAQ. Standard statistical measures used for CMAQ evaluation are described in EPA's latest modeling guidance (EPA, 2018) and in Appendix D. These metrics were applied for daily average simulated $PM_{2.5}$ concentrations over quarterly and annual periods. The CMAQ model performed reasonably well, meeting the performance goals proposed by Boylan and Russell (2006) and criteria by Emery et al. (2017), two well-known references for PM model evaluation. Figure 3.5 shows fractional bias, fractional error, normalized mean bias and normalized mean error for quarterly and annual periods. The performance goals (dashed lines) and criteria (solid lines) are also shown in the figure as references.

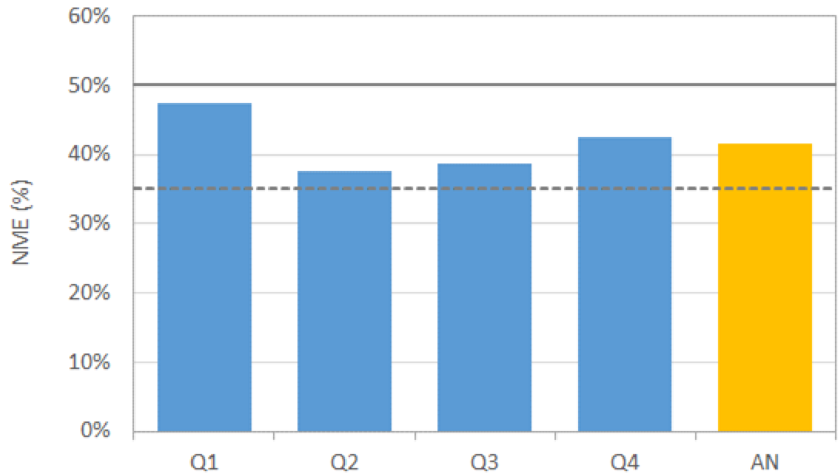


Figure 3.5: Quarterly and annual performance statistics for simulated $PM_{2.5}$ over the 26 air monitoring sites within the 1-km modeling domain with performance goals (dashed lines) and criteria (solid lines) proposed by Boylan and Russell (2006) and Emery et al. (2017). FB stands for fractional bias, FE fractional error, NMB normalized mean bias and NME normalized mean error.

Additional comparisons between simulated and observed $PM_{2.5}$ are discussed in Section 4 (Results) and in Appendix D. These comparisons largely focus on three selected Bay Area sites (West Oakland, Vallejo and San Jose) that are particularly relevant to West Oakland study.

4. Results

Comparison between the annual average simulated and observed $PM_{2.5}$ concentrations (Figure 4.1) shows that the CMAQ model generally captured the observed $PM_{2.5}$ pattern within the 1-km domain. High concentrations in both simulations and observations are evident in the northern San Joaquin Valley, along the I-580 and I-880 corridors from Richmond to the Oakland Airport, along the I-101 corridor near Redwood City, and in the San Jose metropolitan area. In the Sacramento area, the model shows overestimation biases and $PM_{2.5}$ concentrations do not compare as well to observations as in the Bay Area. For Sacramento and other counties outside the Bay Area, we relied on the ARB's emission inventories, and further evaluation of these data may be warranted. The model also shows high concentrations along the I-880 corridor from Oakland Airport to San Jose and along the Delta from Antioch to Brentwood, although observations are unavailable in these areas.

Site by site comparisons between the model predictions and observations (Figure 4.2) show that at most Bay Area sites, the simulated annual average $PM_{2.5}$ concentrations are within ±1.0 $\mu g/m^3$ of observations. At a few sites (Concord, Oakland and Gilroy), the annual average $PM_{2.5}$ concentrations were overestimated, and at one site (Napa), the annual average $PM_{2.5}$ concentration was underestimated by as much as 2.1 $\mu g/m^3$.

Further analyses of model output showed that at sites with overestimated $PM_{2.5}$, both primary and secondary $PM_{2.5}$ concentrations appear to be overestimated. This suggests that there may be multiple causes of overestimation. Primary $PM_{2.5}$ concentrations can be overestimated due to overestimation of emissions, transport, and stability of the atmosphere. Secondary $PM_{2.5}$ can be overestimated due to overestimation of precursor emissions, chemical conversion of the precursors to PM, transport of secondary $PM_{2.5}$ or its precursors, and stability of the atmosphere.

Underestimation of $PM_{2.5}$ at Napa is likely due to an underestimation of wood burning emissions or the transport of wildfire emissions to the North Bay. Wildfire emissions are not included in the modeling emissions inventory.

Figure 4.3 shows time-series plots of observed and simulated daily $PM_{2.5}$ concentrations at three key Bay Area sites relevant to the West Oakland study: West Oakland, Vallejo and San Jose. It is evident from this figure that $PM_{2.5}$ is generally overestimated during winter months and underestimated during summer months. However, on a monthly average basis, the CMAQ model is generally able to replicate the month-to-month variation in observed $PM_{2.5}$ concentrations in West Oakland (Figure 4.4). The somewhat significant underestimation in September is likely due to lack of wildfire emissions in the CMAQ simulations.

During winter months, especially in February, the atmosphere is relatively sunny, calm and cool in the Bay Area, ideal conditions for the formation of secondary PM and for allowing ammonium nitrate to remain in particle form. All of these suspected causes of overestimation are under further investigation and evaluation.

Wintertime overestimation and summertime underestimation of $PM_{2.5}$ by the WRF-CMAQ couple have also been reported elsewhere, and developers of the modeling system are aware of this problem (Appel et al., 2017; Simon et al., 2012). Efforts are underway by the model developers and District staff to minimize errors and to improve model performance for both winter and summer.

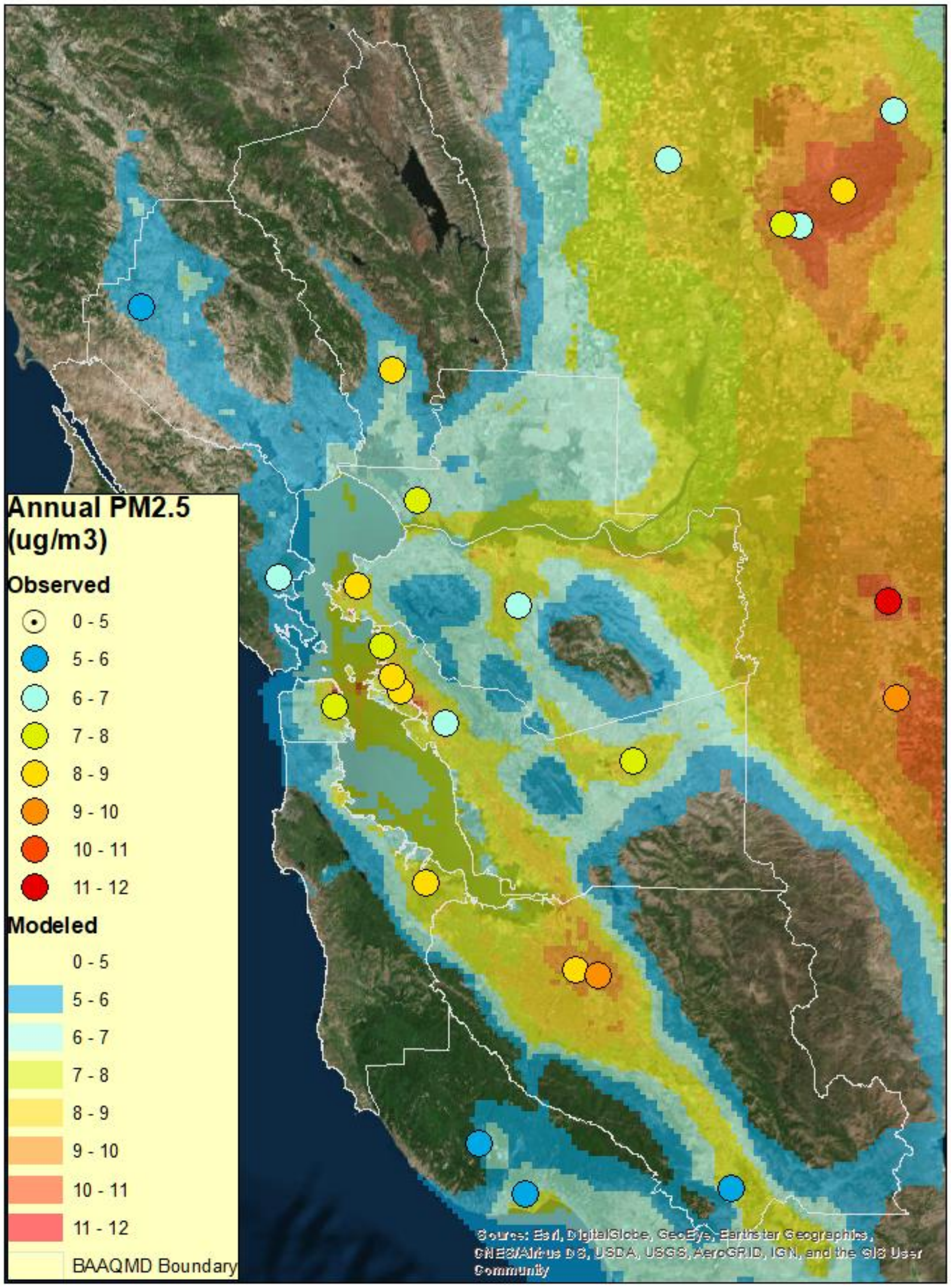


Figure 4.1: Spatial distribution of simulated and observed annual average $PM_{2.5}$ concentrations within the 1-km modeling domain.

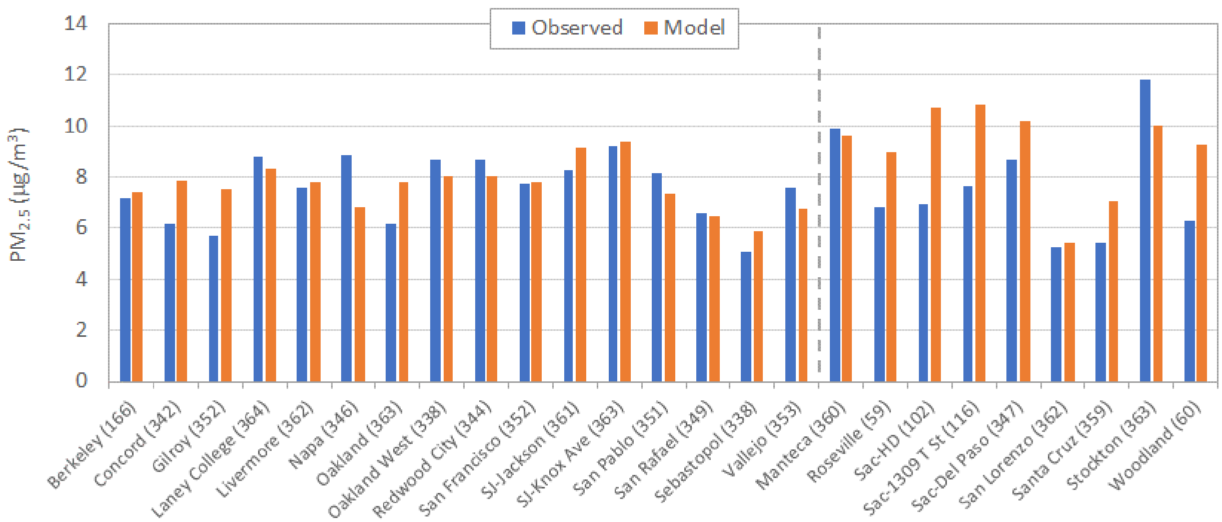
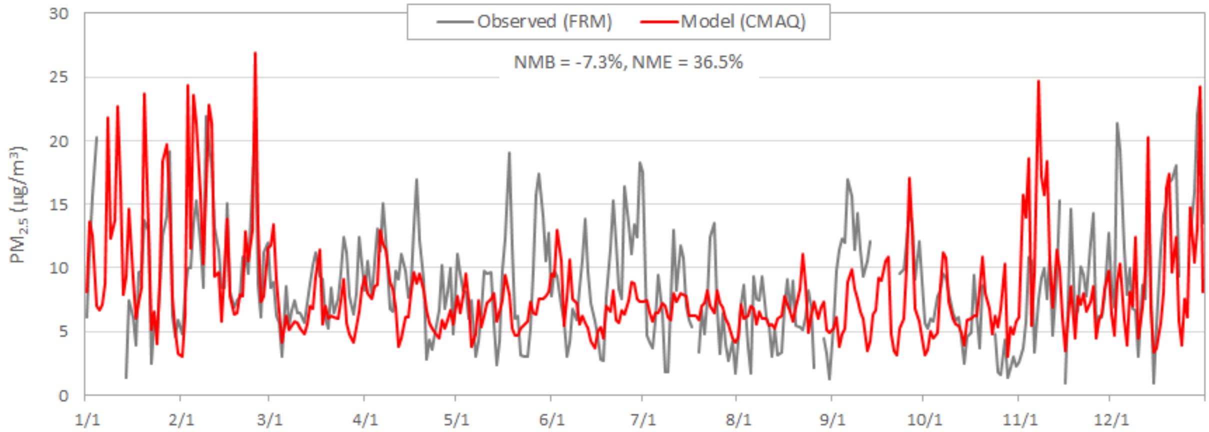
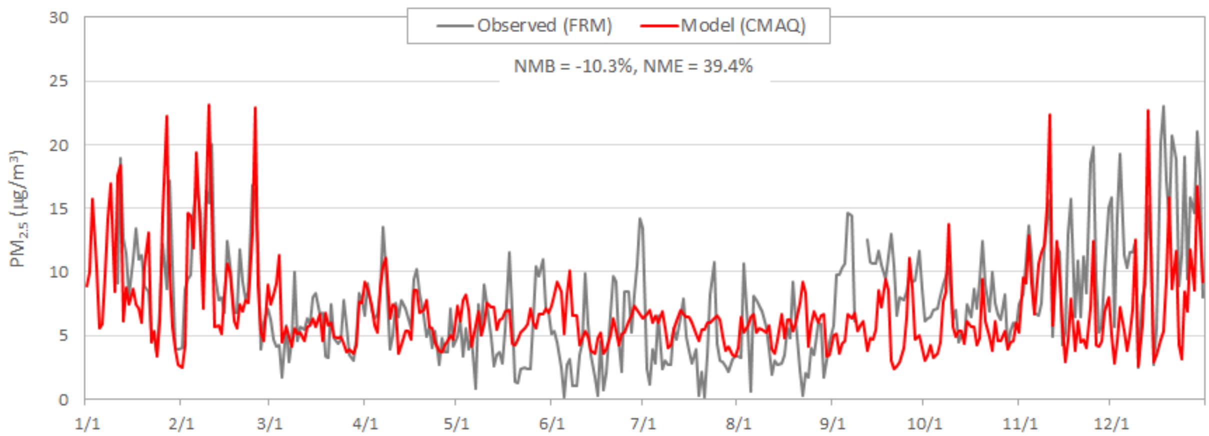


Figure 4.2: Annual mean observed vs. modeled $PM_{2.5}$ concentrations at monitoring sites within the 1-km modeling domain. The annual means are calculated over the days with valid observations. The number of valid observations is shown in parentheses for each site. The Berkeley site is missing observations for January through June. Two Sacramento sites (Health Department - Stockton Blvd. and 1309 T Street) have observations every 3rd day. The Roseville and Woodland sites have observations every 6th day. The Sacramento - Bercut Drive site has data only for December, so its annual mean is not shown.

(a) Oakland West



(b) Vallejo



(c) San Jose – Knox Avenue

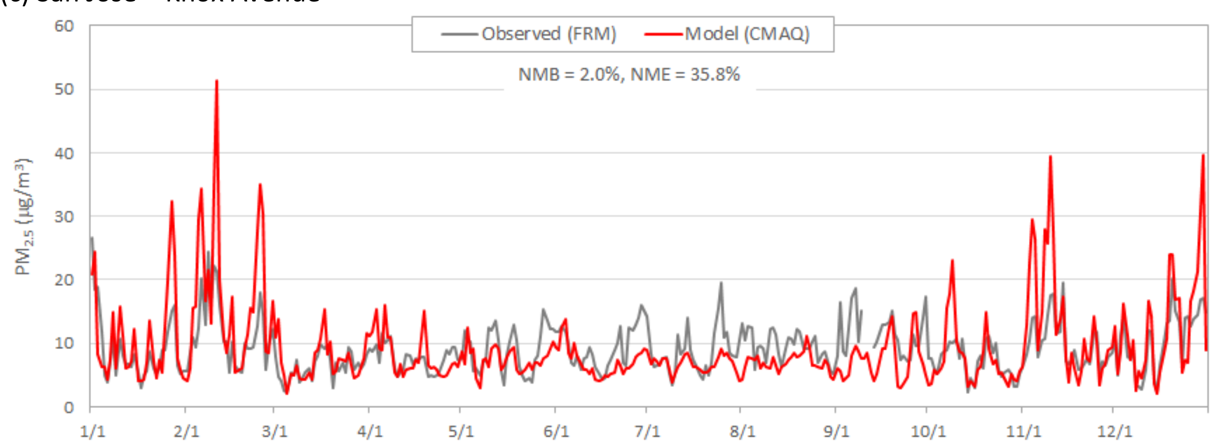


Figure 4.3: Time-series plots of observed vs. modeled daily $PM_{2.5}$ concentrations at (a) Oakland West, (b) Vallejo, and (c) San Jose monitoring sites.

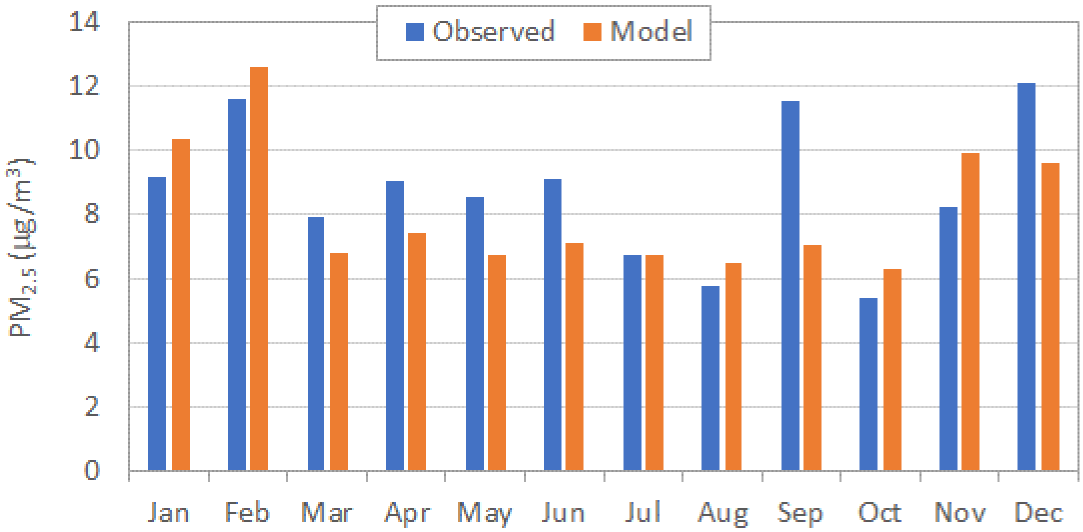


Figure 4.4: Monthly average simulated and observed $PM_{2.5}$ concentrations in West Oakland.

4.1 Estimating Background $PM_{2.5}$ in West Oakland

As mentioned in Section 1, we have simulated pollutant concentrations at a 1-km horizontal resolution over the entire Bay Area for 2016 (base case). Then we repeated the simulation with all anthropogenic emissions removed from the modeling inventory in the West Oakland source domain (control case), leaving all other model input parameters unchanged.

Figure 4.5 shows the annual average $PM_{2.5}$ concentrations for the base case within the West Oakland receptor domain. The highest and lowest annual average $PM_{2.5}$ concentrations are 9.3 µg/m^3 and 7.1 µg/m^3, respectively. A concentration gradient is evident within the domain. Cells with relatively higher concentrations extend along the eastern boundary and northwestern corner of the domain. A concentration gradient is also evident in the West Oakland community, an area within the red border in the figure. The eastern half of the community has slightly higher concentrations than the western half.

The spatial distribution of the annual average $PM_{2.5}$ concentrations is similar to the spatial distribution of West Oakland's emissions (Figure 3.3). The Chinatown area in the southeastern corner of the West Oakland domain has the highest emissions and concentrations. The cell along the southern boundary with the area's lowest concentration (7.1 µg/m^3) also has the lowest emissions (1.4 lbs/day).

Figure 4.6 shows the annual average $PM_{2.5}$ concentrations for the control case, i.e., a simulation without West Oakland's anthropogenic emissions. Compared to Figure 4.5, the spatial gradient in the annual average concentrations decreased significantly in the absence of West Oakland emissions across the receptor domain. The location of the maximum annual average $PM_{2.5}$ concentrations has shifted from Chinatown to near the Bay Bridge, suggesting the influence of transport from the northwest corner of the domain.

Figure 4.7 shows the difference between the base and control cases. Based on the figure, the Chinatown area would benefit the most (2.5 µg/m^3) from zeroing out all anthropogenic emissions in the West Oakland source domain. The West Oakland community (within the red border) would benefit by $PM_{2.5}$ reductions ranging from 0.8 µg/m^3 to 1.7 µg/m^3. The southwest corner of the receptor domain would be the least benefitted area, with a reduction of about 0.5 µg/m^3.

Note that these $PM_{2.5}$ concentrations and reductions represent the average value across a 1x1 km grid cell. Higher concentrations and reductions are possible at the sub-grid cell level, and these finer-scale gradients will be investigated with local-scale AERMOD modeling.

Bias in the simulated annual average $PM_{2.5}$ concentrations for both base and control cases are expected to be similar. Since reductions are estimated from the difference between the two simulations, the impact of model bias on estimated reductions is expected to be insignificant.

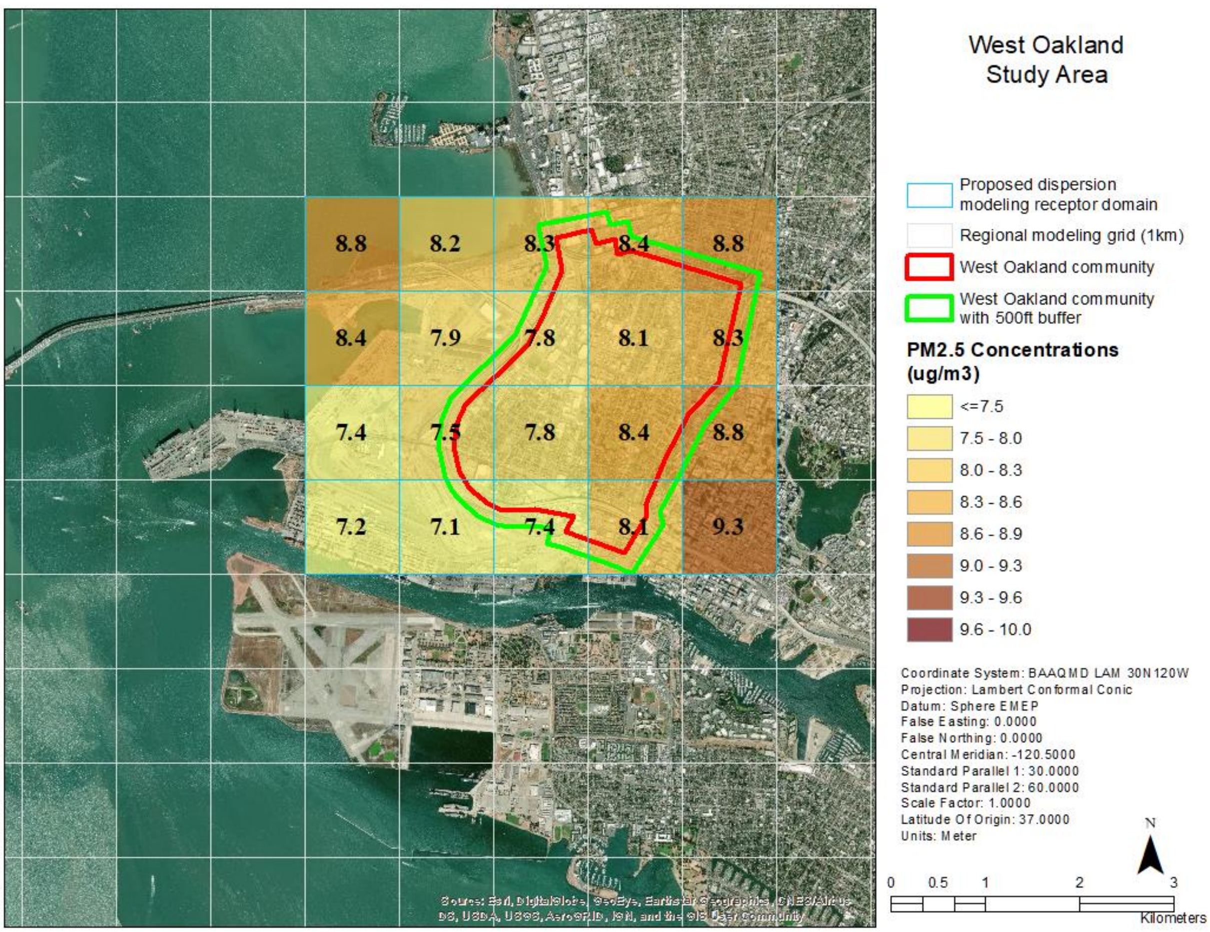


Figure 4.5: Spatial distribution of the simulated annual average $PM_{2.5}$ concentrations in the West Oakland receptor domain (base case).

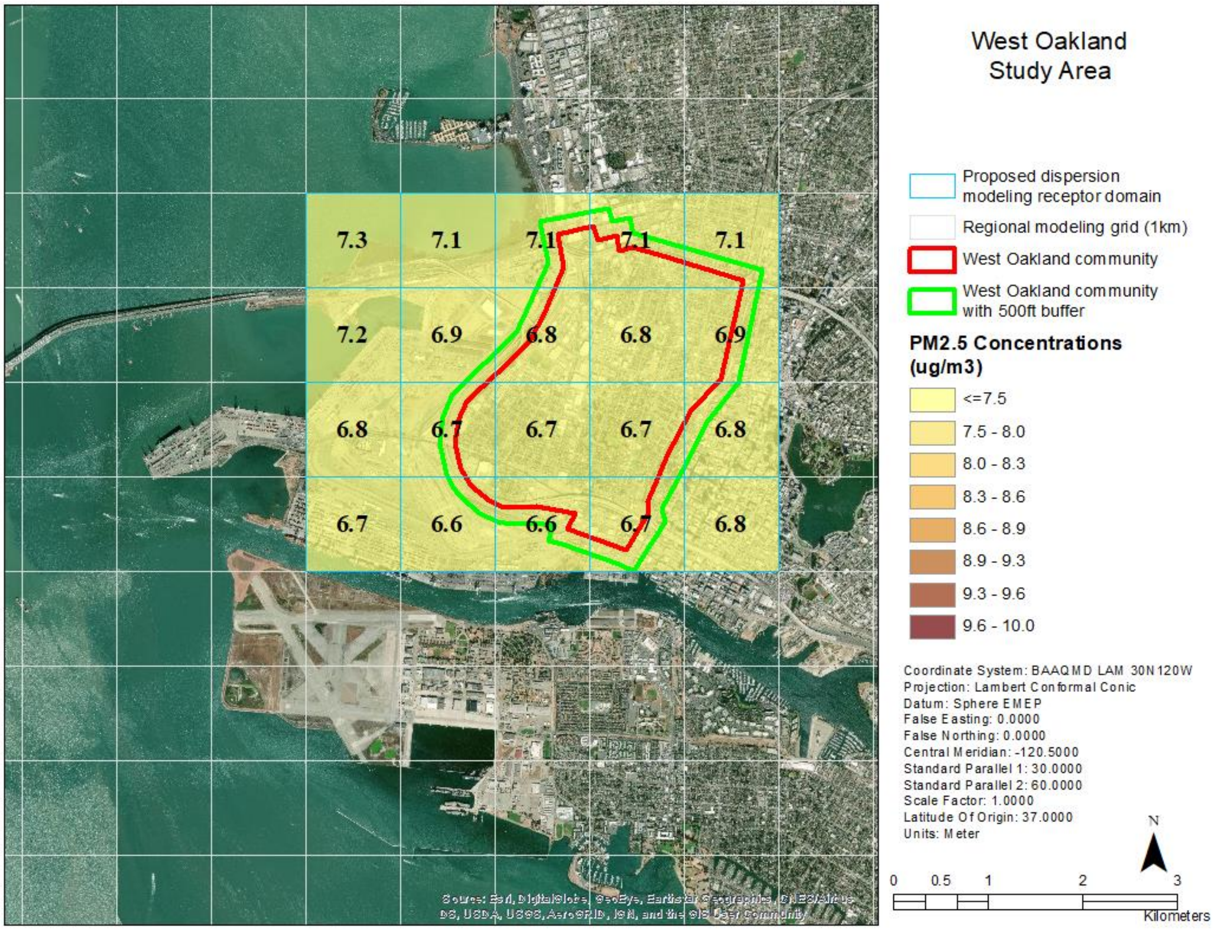


Figure 4.6: Spatial distribution of the simulated $PM_{2.5}$ concentrations without West Oakland's anthropogenic emissions (control case).

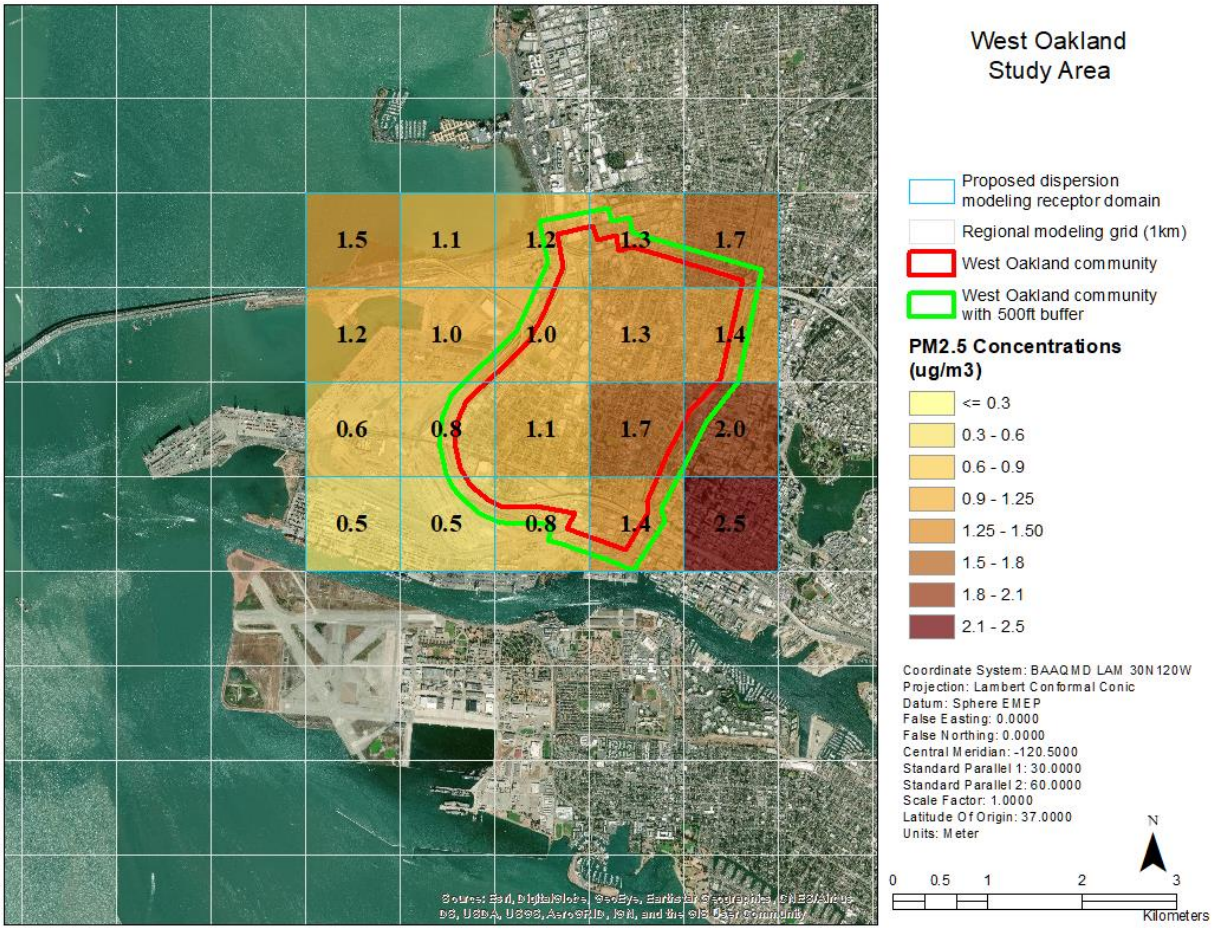


Figure 4.7: Difference between the simulated annual average base and control case $PM_{2.5}$ concentrations.

Appendix A – Observational Data

A1. Description of Observations

Table A1 lists all aerometric stations within the 1-km modeling domain from which data were used in this study. It also shows data sources, data types, and purpose of the data. Under the monitoring location column, the first two subsections list $PM_{2.5}$ stations within and outside of the Bay Area. The subsequent subsections list meteorological measurement stations within and outside of the Bay Area, followed by a list of upper air measurement stations. Meteorological measurement stations within the Bay Area are further separated based on whether or not they are operated by the District. Some stations measure both $PM_{2.5}$ and meteorology. These stations are listed under both the $PM_{2.5}$ and meteorology measurement sections and identified through checkmarks under columns titled “PM” and “Met.”

Hourly $PM_{2.5}$ and meteorological data were obtained from the District’s Data Management System (DMS) in October 2018, and hourly $PM_{2.5}$ and meteorological data were obtained from the U.S. EPA’s Air Quality System (AQS) at around the same time. Hourly meteorological data were also obtained from the NCAR/UCAR ADP data archive and twice daily upper air data were obtained from NOAA’s National Climatic Data Center.

The ADP and National Climatic Data Center data were used for the four-dimensional data assimilation (FDDA) in the WRF model. FDDA is a method to nudge the WRF model results towards observations. The WRF model includes a post-processing utility computer program that prepares the ADP and NCDC data for FDDA. The utility program also quality assures and quality checks both the ADP and National Climatic Data Center data. Meteorological data from the other sources listed in Table A1 were not used in FDDA because they do not include pressure, which is required for the nudging process. All the observed meteorological data listed in the column titled “Met” were used for WRF model validation using a software tool called METSTAT. The METSTAT program has a module for applying consistency checks to the data being used.

$PM_{2.5}$ data obtained from DMS and AQS were compared against each other and no differences were found. $PM_{2.5}$ data were used for both data analysis and CMAQ model validation. For consistency in data format, only data downloaded from AQS were used for data analysis and model validation.

As explained in the main text of this document, time series and spatial plots of simulated and observed hourly $PM_{2.5}$ concentrations were generated and compared against each other. This process allowed identification of gaps and outliers in the $PM_{2.5}$ data. Statistical formulas were developed in an Excel spreadsheet to evaluate the observations, such as assessing gaps in measurements, calculating daily, monthly, seasonal and annual averages, and assessing high and low values. Statistical formulas were also developed to assess bias, normalized bias and root mean square error in simulations by comparing simulated values to observations.

Table A1: Description of observations used in this study.

Monitoring Location	Source	PM	Met (ws, wd, t, rh)*	FDDA	Model Validation	
					WRF	CMAQ
San Francisco Bay Area PM Stations						
Berkeley Aquatic Park	DMS	x				x
Concord	DMS	x	x			x
Gilroy	DMS	x				x
Laney College	DMS	x				x
Livermore	DMS	x	x			x
Napa	DMS	x	x			x
Oakland	DMS	x				x
Oakland West	DMS	x				x
Redwood City	DMS	x				x
San Francisco	DMS	x				x
San Jose - Jackson	DMS	x				x
San Jose - Knox Avenue	DMS	x				x
San Pablo	DMS	x				x
San Rafael	DMS	x				x
Sebastopol	DMS	x				x
Vallejo	DMS	x	x			x
PM Stations Outside the San Francisco Bay Area						
Manteca	AQS	x				x
Roseville - N Sunrise Ave	AQS	x	x			x
Sacramento Health Department - Stockton Blvd.	AQS	x				x
Sacramento - 1309 T Street	AQS	x	x			x
Sacramento - Bercut Drive	AQS	x				x
Sacramento - Del Paso Manor	AQS	x	x			x
San Lorenzo Valley Middle School	AQS	x				x
Santa Cruz	AQS	x				x
Stockton - Hazelton	AQS	x	x			x
Woodland - Gibson Road	AQS	x				x
BAAQMD Met Stations						
Bethel Island	DMS		x		x	
Chabot	DMS		x		x	
Concord	DMS	x	x		x	
Fairfield	DMS		x		x	
Ft. Funston	DMS		x		x	
Livermore	DMS	x	x		x	
Napa	DMS	x	x		x	
Oakland STP	DMS		x		x	
Patterson Pass	DMS		x		x	
Pleasanton	DMS		x		x	
Pt. San Pablo	DMS		x		x	

Monitoring Location	Source	PM	Met (ws, wd, t, rh)*	FDDA	Model Validation	
					WRF	CMAQ
Rio Vista	DMS		x		x	
San Carlos	DMS		x		x	
San Martin	DMS		x		x	
San Ramon	DMS		x		x	
Sonoma Baylands	DMS		x		x	
Vallejo	DMS	x	x		x	
Valley Ford	DMS		x		x	
Non-BAAQMD Met Stations in the Bay Area						
Berkeley Lab	DMS		x		x	
Concord KCCR	ADP		x	x	x	
Hayward KHWD	ADP		x	x	x	
Livermore KLVK	ADP		x	x	x	
Moffett NASA/Mountain View KNUQ	ADP		x	x	x	
Napa KAPC	ADP		x	x	x	
Novato KDVO	ADP		x	x	x	
Oakland KOAK	ADP		x	x	x	
Palo Alto KPAO	ADP		x	x	x	
Petaluma KO69	ADP		x	x	x	
San Carlos KSQL	ADP		x	x	x	
San Francisco KSFO	ADP		x	x	x	
San Francisco STP	DMS		x		x	
San Jose KSJC	ADP		x	x	x	
San Jose/Reid KRHV	ADP		x	x	x	
San Martin KE16	ADP		x	x	x	
Santa Rosa KSTS	ADP		x	x	x	
Travis AFB KSUU	ADP		x	x	x	
Met Stations Outside the Bay Area						
Davis - UCD Campus	AQS		x		x	
Davis KEDU	ADP		x	x	x	
Elk Grove - Bruceville Road	AQS		x		x	
Half Moon Bay KHAF	ADP		x	x	x	
Hollister KCVH	ADP		x	x	x	
Lincoln KLHM	ADP		x	x	x	
Mather Field KMHR	ADP		x	x	x	
McClellan AFB KMCC	ADP		x	x	x	
Modesto KMOD	ADP		x	x	x	
Roseville - N Sunrise Ave	AQS	x	x		x	
Sacramento - 1309 T Street	AQS	x	x		x	
Sacramento - Del Paso Manor	AQS	x	x		x	
Sacramento KSAC	ADP		x	x	x	
Salinas KSNS	ADP		X	X	x	
Sanford MUNI KSMF	ADP		x	x	x	

Monitoring Location	Source	PM	Met (ws, wd, t, rh)*	FDDA	Model Validation	
					WRF	CMAQ
Stockton KSCK	ADP		x	x	x	
Stockton - Hazelton	AQS	x	x		x	
Tracy - Airport	AQS		x		x	
Vacaville KVCB	ADP		x	x	x	
Watsonville KWVI	ADP		x	x	x	
Upper Air Stations						
Oakland Sounding	NCDC		x	x	x	

*ws=wind speed; wd=wind direction; t=temperature; rh=relative humidity.

A2. Spatial Distribution of Observation Stations

Figure A1 shows spatial distribution of meteorological observation stations in the 1-km modeling domain. They are grouped based on whether they are operated by BAAQMD or other agencies (non-BAAQMD) and whether they are inside or outside of the District boundaries. BAAQMD sites collocated with $PM_{2.5}$ measurements are also marked.

The spatial distribution of $PM_{2.5}$ monitoring stations in the 1-km modeling domain is shown in Figure A2. These stations are grouped based on whether they are inside or outside of the District boundaries. All stations within the District are operated by the District.

Note that both the meteorological and air quality models have nested domains. Meteorological and air quality measurements outside of the 1-km domain were obtained from various databases and used for FDDA and model evaluation along with data for the 1-km domain. These data are stored on modeling computers and are available on request.

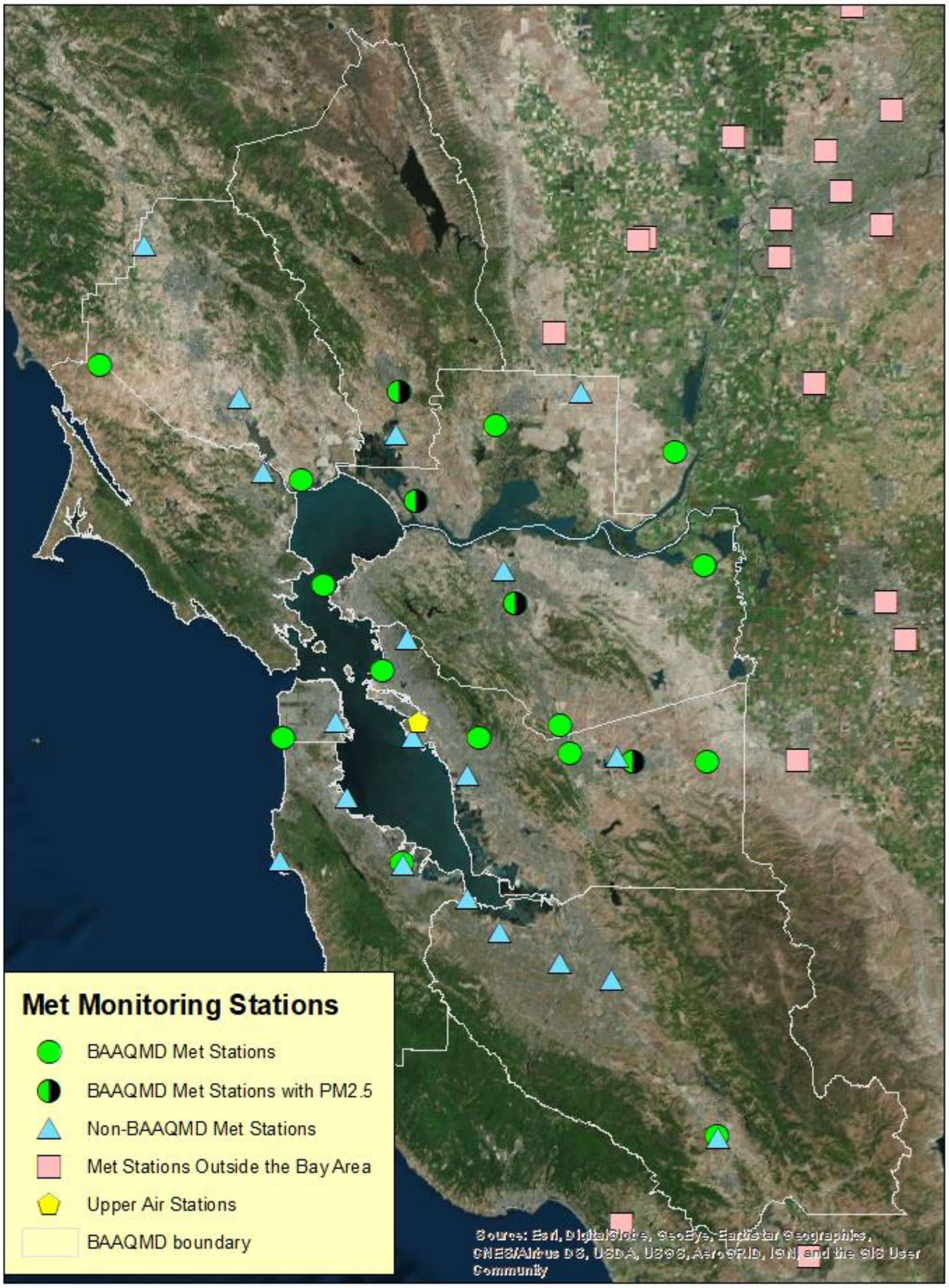


Figure A1: Spatial distribution of meteorological monitoring sites in the 1-km modeling domain.

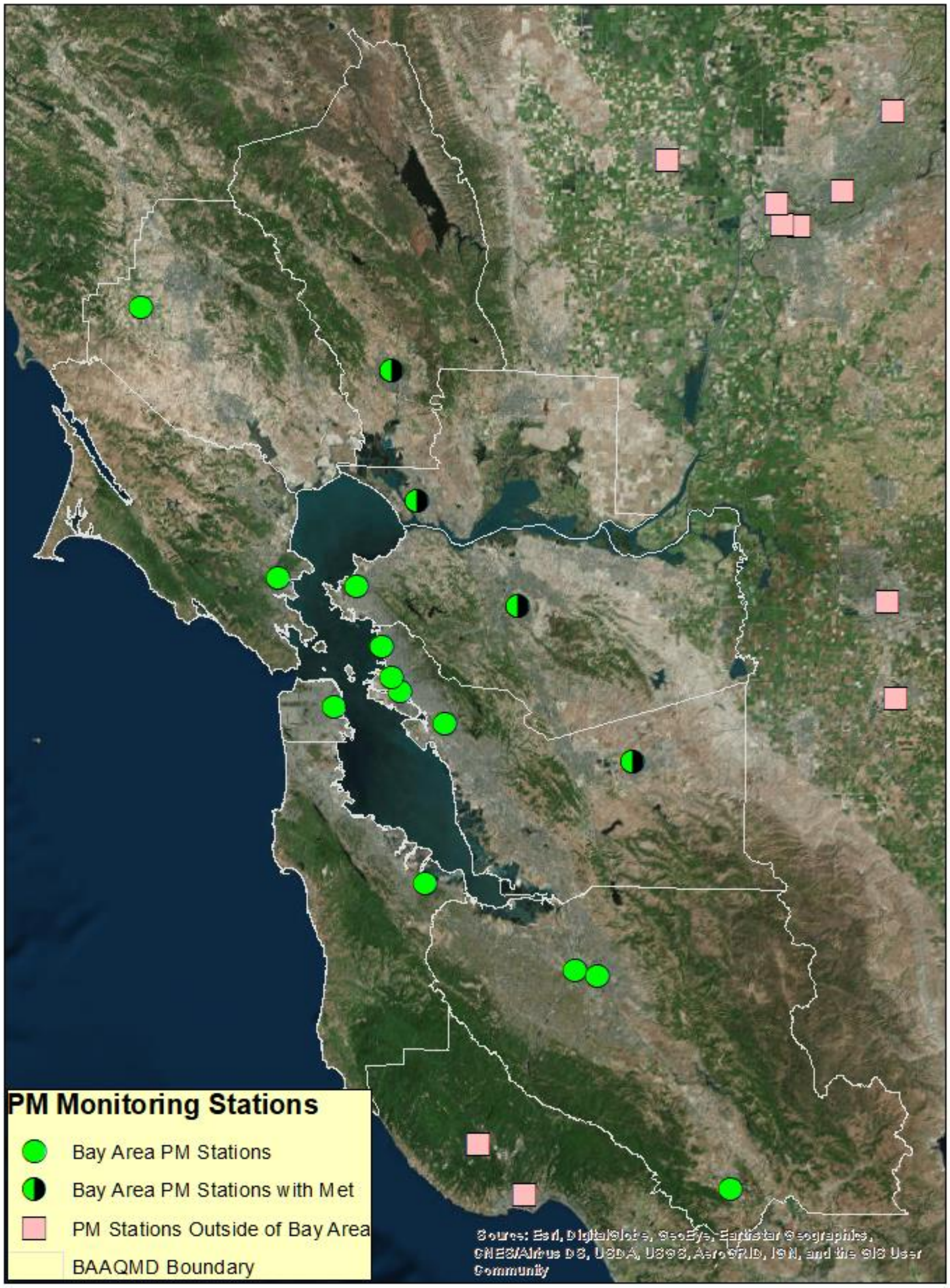


Figure A2: Spatial distribution of $PM_{2.5}$ monitoring stations in the 1-km modeling domain.

Appendix B – Emissions Inventory

This appendix provides additional details on the development of emissions estimates for residential wood combustion, an important PM2.5 source during winter pollution episodes. This appendix also contains additional summary tables and emissions density plots that characterize the emissions inventory used for the 2016 CMAQ modeling.

B1. Residential Wood Combustion

ARB emissions estimates for residential wood combustion are based on county-level populations of fireplaces and woodstoves, wood consumption rates by device type, and emission factors that represent the quantity of emissions per ton of fuel burned. Where possible, ARB estimates device populations and wood consumption rates using local survey data, such as data collected as part of the District's Spare the Air Tonight Study (BAAQMD, 2007). ARB residential wood combustion emissions estimates for the District were compared to internal estimates derived from survey data, as well as estimates from neighboring air districts.

Figure B1 shows annual average $PM_{2.5}$ emissions from residential wood combustion for the two BAAQMD inventories, as well as inventories for the Sacramento Metropolitan Air Quality Management District (SMAQMD) and the San Joaquin Valley Unified Air Pollution Control District (SJVUAPCD). The internal BAAQMD inventory is somewhat higher than the ARB inventory for BAAQMD, and both BAAQMD inventories are significantly higher than emissions estimates for SMAQMD and SJVUAPCD. After additional investigations and discussions with ARB, it was determined that:

- ARB's $PM_{2.5}$ emissions estimates for winter compared well with the District's internal estimates; however, the District's estimates for summer (which were extrapolated from winter survey results based on temperature data) are significantly higher than ARB's estimates. These summer estimates do not appear to be realistic and result in the higher annual average $PM_{2.5}$ emissions in the District's internal inventory.
- Higher residential wood combustion emissions estimates for BAAQMD relative to SMAQMD and SJVUAPCD likely result from a failure to account for the impact of the District's Spare the Air Program.

Based on these findings, it was decided that ARB's residential wood combustion estimates for the District would be reduced by 50% as an initial estimate of the impact of the District's Spare the Air program. Figure B2 shows the final monthly average $PM_{2.5}$ emissions from residential wood combustion that were included in the CMAQ modeling inventories. Monthly average emissions range from 0.20 tpd in August to 8.39 tpd in January, with an annual average of 3.93 tpd.

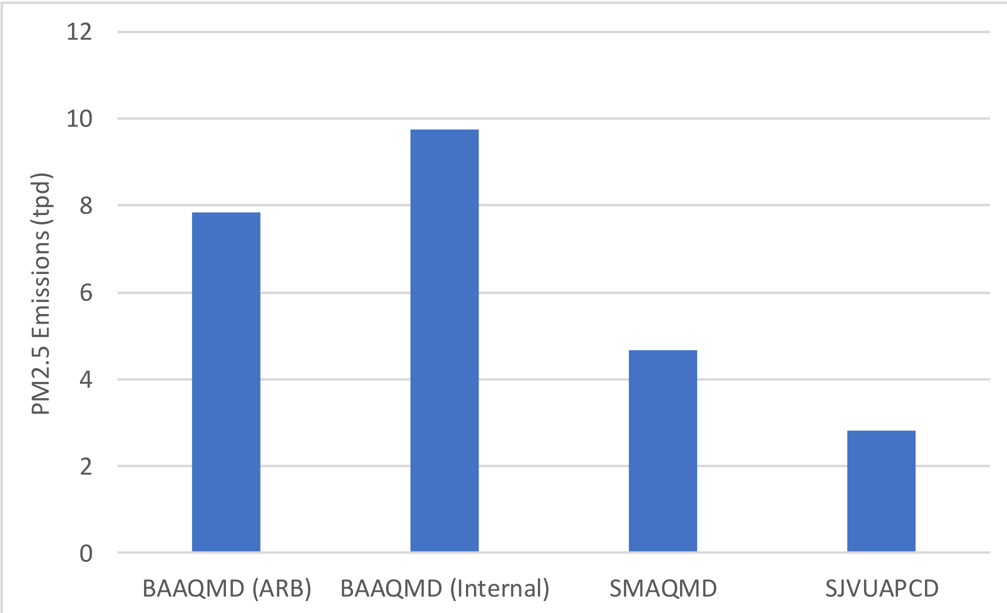


Figure B1: Comparison of 2016 annual average $PM_{2.5}$ emissions inventories for residential wood combustion.

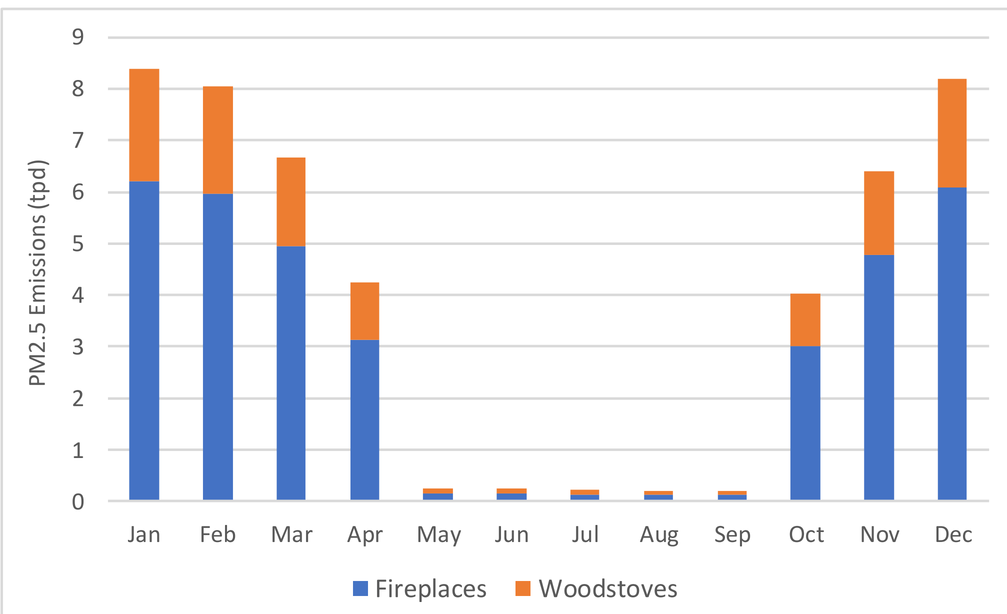


Figure B2: 2016 monthly average $PM_{2.5}$ emissions for BAAQMD from residential wood combustion.

B2. Emissions Inventory Summaries

This section provides additional summary information on the emissions inventory used for the 2016 CMAQ modeling. Tables B1 through B4 show emissions of $PM_{2.5}$ precursors (TOG, NO_x, SO_2, and NH_3) by geographic area and source sector. Key sources of TOG emissions include landfills, natural gas transmission losses, petroleum refining, and solvent usage. Key sources of NO_x emissions include onroad and nonroad mobile sources, especially diesel-powered vehicles.

Key sources of SO_2 emissions include petroleum refining and ocean-going vessels. Key sources of NH_3 emissions include farming operations such as livestock waste and fertilizer application.

Table B1: Summary of 2016 TOG emissions (tons/day) by geographic area and source sector.

Geographic Area	Area	Nonroad	Onroad	Point	Total
Alameda	42.6	8.0	12.4	64.1	127.1
Contra Costa	51.0	6.3	7.7	43.9	109.0
Marin	16.2	2.9	2.5	15.2	36.8
Napa	6.8	1.9	1.4	3.6	13.6
San Francisco	18.1	6.2	2.9	3.0	30.2
San Mateo	18.9	7.1	4.4	25.1	55.5
Santa Clara	57.6	8.1	12.2	50.8	128.7
Solano	14.3	2.1	2.6	8.7	27.7
Sonoma	25.4	3.0	3.5	9.1	41.1
BAAQMD Subtotal	*251.0*	*45.6*	*49.6*	*223.4*	*569.6*
Non-BAAQMD Counties	502.5	22.7	29.0	80.9	635.1
Domain Total	**753.5**	**68.3**	**78.6**	**304.3**	**1,204.7**

Table B2: Summary of 2016 NO_x emissions (tons/day) by geographic area and source sector.

Geographic Area	Area	Nonroad	Onroad	Point	Total
Alameda	3.4	12.0	27.2	3.2	45.8
Contra Costa	4.3	11.8	13.8	15.5	45.4
Marin	0.9	3.8	3.6	0.3	8.5
Napa	0.3	2.2	3.0	0.2	5.7
San Francisco	2.1	34.4	4.5	1.4	42.4
San Mateo	2.1	18.9	6.7	0.7	28.5
Santa Clara	4.2	10.0	22.2	8.5	45.0
Solano	1.0	3.9	5.4	3.8	14.1
Sonoma	0.9	7.9	6.7	0.4	15.9
BAAQMD Subtotal	*19.3*	*104.8*	*93.0*	*34.0*	*251.2*
Non-BAAQMD Counties	19.0	37.3	60.5	4.2	121.1
Domain Total	**38.4**	**142.1**	**153.5**	**38.3**	**372.3**

Table B3: Summary of 2016 SO_2 emissions (tons/day) by geographic area and source sector.

Geographic Area	Area	Nonroad	Onroad	Point	Total
Alameda	0.1	0.4	0.2	1.4	2.0
Contra Costa	0.1	0.9	0.1	16.6	17.7
Marin	0.0	0.0	0.0	0.1	0.2
Napa	0.0	0.0	0.0	0.0	0.0
San Francisco	0.1	0.4	0.0	0.1	0.6
San Mateo	0.1	0.9	0.1	0.1	1.1
Santa Clara	0.1	0.1	0.2	2.9	3.3
Solano	0.0	0.2	0.0	0.4	0.6

Sonoma	0.0	0.1	0.0	0.0	0.2
BAAQMD Subtotal	*0.5*	*3.1*	*0.7*	*21.6*	*25.9*
Non-BAAQMD Counties	1.6	0.3	0.4	1.9	4.2
Domain Total	**2.1**	**3.4**	**1.1**	**23.4**	**30.1**

Table B4: Summary of 2016 NH_3 emissions (tons/day) by geographic area and source sector.

Geographic Area	Area	Nonroad	Onroad	Point	Total
Alameda	3.0	0.0	1.5	0.4	4.9
Contra Costa	3.1	0.0	0.9	2.1	6.1
Marin	2.5	0.0	0.3	0.3	3.0
Napa	0.5	0.0	0.2	0.1	0.8
San Francisco	1.4	0.0	0.3	0.0	1.7
San Mateo	1.3	0.0	0.5	0.2	2.1
Santa Clara	3.6	0.0	1.6	1.5	6.7
Solano	1.9	0.0	0.3	0.1	2.3
Sonoma	3.3	0.0	0.4	0.3	4.0
BAAQMD Subtotal	*20.7*	*0.1*	*6.0*	*5.0*	*31.7*
Non-BAAQMD Counties	46.9	0.0	3.3	7.0	57.3
Domain Total	**67.6**	**0.1**	**9.3**	**12.0**	**89.0**

B3. Emissions Density Plots

This section provides emissions density plots that show the spatial distribution of key $PM_{2.5}$ precursors (an emissions density plot for primary $PM_{2.5}$ is provided in the main body of this report in Figure 3.1). The emissions density plot for NO_x (Figure B3) shows elevated emissions along major freeways and shipping lanes and in urban cores. Note that high emissions along shipping lanes in San Pablo Bay to the south of Marin County may be overestimated due to the spatial surrogate used for commercial marine vessel emissions, which does not include offshore shipping lanes for Sonoma County.

The emissions density plot for SO_2 (Figure B4) shows the presence of point sources in the 1-km domain that emit this pollutant, as well as emissions from commercial marine vessels along shipping lanes. The emissions density plot for NH_3 (Figure B5) shows elevated emissions in San Joaquin County in the eastern part of the modeling domain, an area with significant agricultural activity.

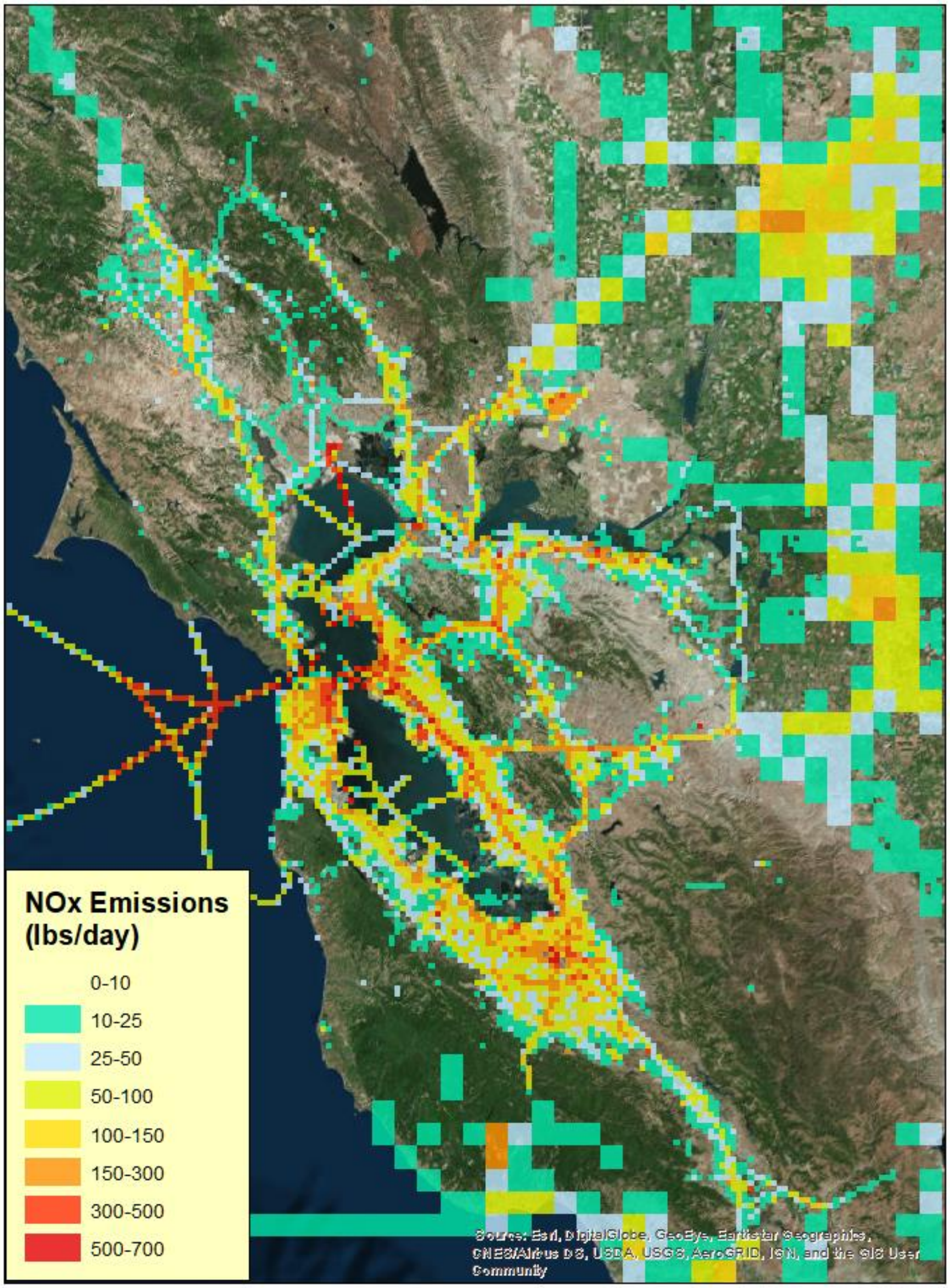


Figure B3: Spatial distribution of annual average NO_x emissions for the 1-km modeling domain.

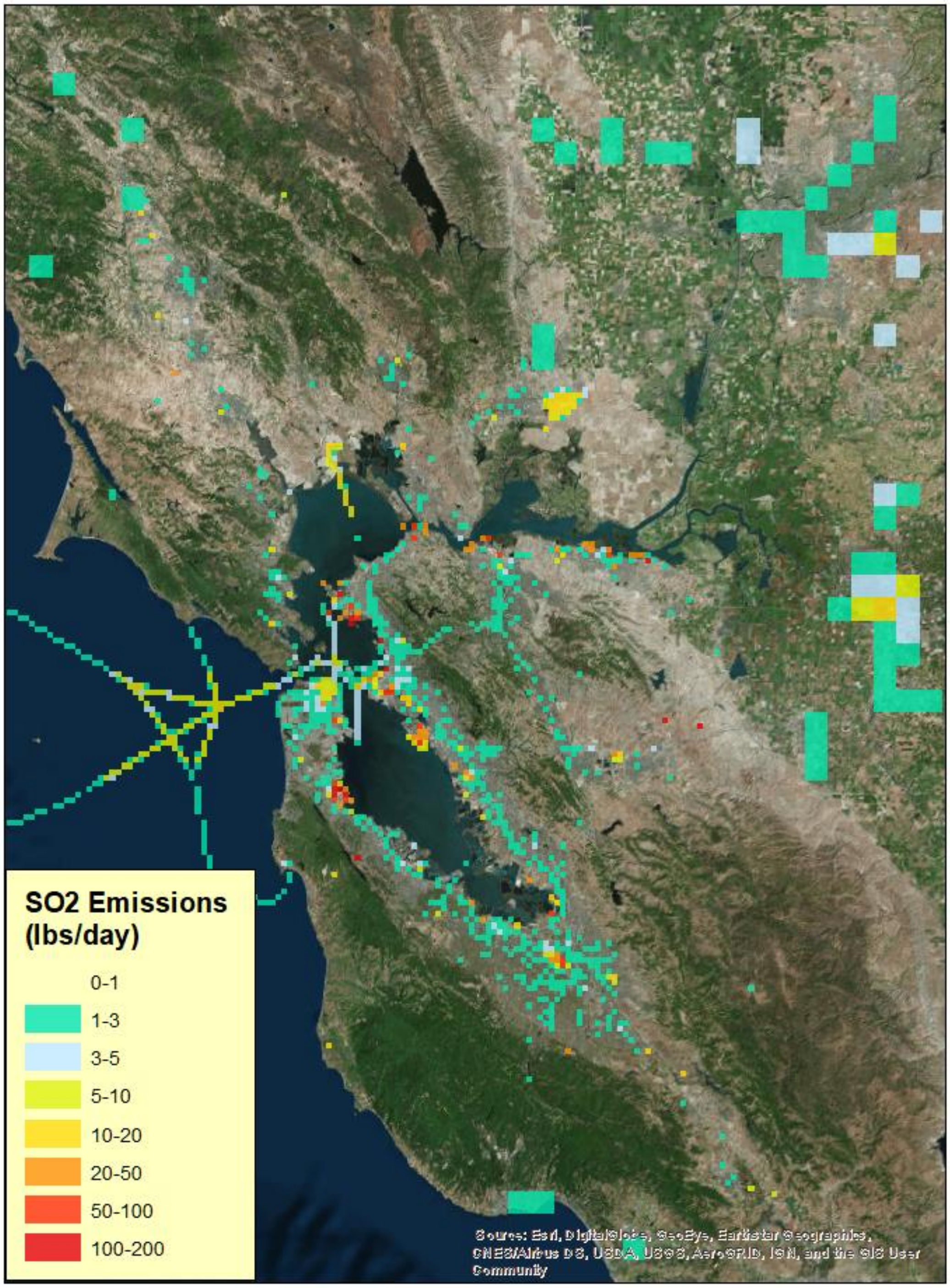


Figure B4: Spatial distribution of annual average SO_2 emissions for the 1-km modeling domain.

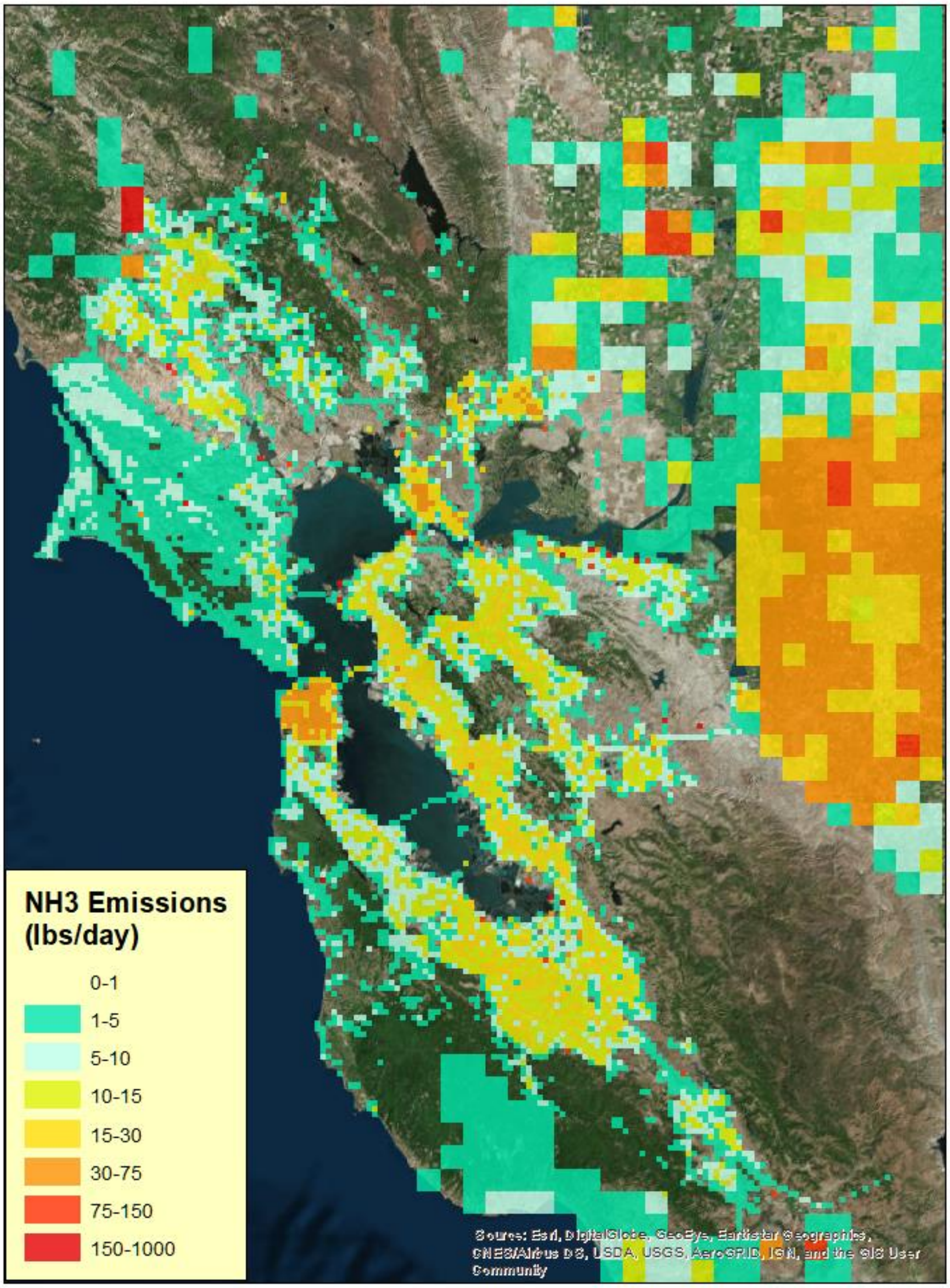


Figure B5: Spatial distribution of annual average NH_3 emissions for the 1-km modeling domain.

APPENDIX C – Meteorological Model Evaluation

C1. Statistical Evaluation

The ENVIRON METSTAT program (Emery et al., 2001) was used to compare the WRF-generated meteorological fields against hourly surface observations archived at NCAR. METSTAT is a statistical analysis software package that calculates and graphically presents statistics such as mean observation, mean simulation, bias error, gross error, and index of agreement.

Hourly time series of observed and simulated surface-layer wind and temperature are presented to evaluate the model performance. Statistics are defined as follows:

<u>Mean observation (M_o)</u>: calculated from all sites with valid data within a given analysis region and for a given time period (hourly or daily):

$$M_o = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I} O_j^i$$

where O_j^i is the individual observed quantity at site *i* and time *j*, and the summations are over all sites (*I*) and time periods (*J*).

<u>Mean prediction (M_p)</u>: calculated from simulation results that are interpolated to each observation used to calculate the mean observation (hourly or daily):

$$M_p = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I} P_j^i$$

where P_j^i is the individual simulated quantity at site *i* and time *j*. Note that mean observed and simulated winds are vector-averaged (for east-west component *u* and north-south component *v*), from which the mean wind speed and mean resultant direction are derived.

<u>Bias error (B)</u>: calculated as the mean difference in prediction-observation pairings with valid data within a given analysis region and for a given time period (hourly or daily):

$$B = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}\left(P_j^i - O_j^i\right)$$

<u>Gross Error (E)</u>: calculated as the mean *absolute* difference in prediction-observation pairings with valid data within a given analysis region and for a given time period (hourly or daily):

$$E = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}\left|P_j^i - O_j^i\right|$$

Note that the bias and gross error for winds are calculated from the predicted-observed residuals in speed and direction (not from vector components *u* and *v*). The direction error for a given prediction-observation pairing is limited to range from 0 to ±180°.

Root Mean Square Error (RMSE): calculated as the square root of the mean squared difference in prediction-observation pairings with valid data within a given analysis region and for a given time period (hourly or daily):

$$RMSE = \left[\frac{1}{IJ} \sum_{j=1}^{J} \sum_{i=1}^{I} \left(P_j^i - O_j^i \right)^2 \right]^{1/2}$$

The RMSE, as with the gross error, is a good overall measure of model performance.

Index of Agreement (IOA): calculated following the approach of Willmont (1981). This metric condenses all the differences between model estimates and observations within a given analysis region and for a given time period (hourly and daily) into one statistical quantity. It is the ratio of the total RMSE to the sum of two differences – between each prediction and the observed mean, and each observation and the observed mean:

$$IOA = 1 - \left[\frac{IJ \cdot RMSE^2}{\sum_{j=1}^{J} \sum_{i=1}^{I} \left| P_j^i - M_o \right| + \left| O_j^i - M_o \right|} \right]$$

Viewed from another perspective, the index of agreement is a measure of the match between the departure of each prediction from the observed mean and the departure of each observation from the observed mean. Thus, the correspondence between predicted and observed values across the domain at a given time may be quantified in a single metric and displayed as a time series. The index of agreement has a theoretical range of 0 to 1, the latter score suggesting perfect agreement.

C2. Time Series Comparisons

To further evaluate model performance and to understand the meteorology at specific regions of concern, the simulated results were compared to wind and temperature measurements from monitoring sites in West Oakland, San Jose and Vallejo. Figures C2 through C9 show time series comparing daily average WRF-simulated surface wind speed and temperature to observations at Oakland, San Jose and Vallejo for each quarter of 2016.

The WRF-simulated wind and temperature matched the observed trends very well for the whole year of 2016. There were no significant differences between the predicted and the observed values. The best performance was observed at Vallejo site, especially for wind speed performance. Underestimations of wind speed were noticeable at Oakland and San Jose throughout 2016. Investigations into this wind speed performance problem are on-going.

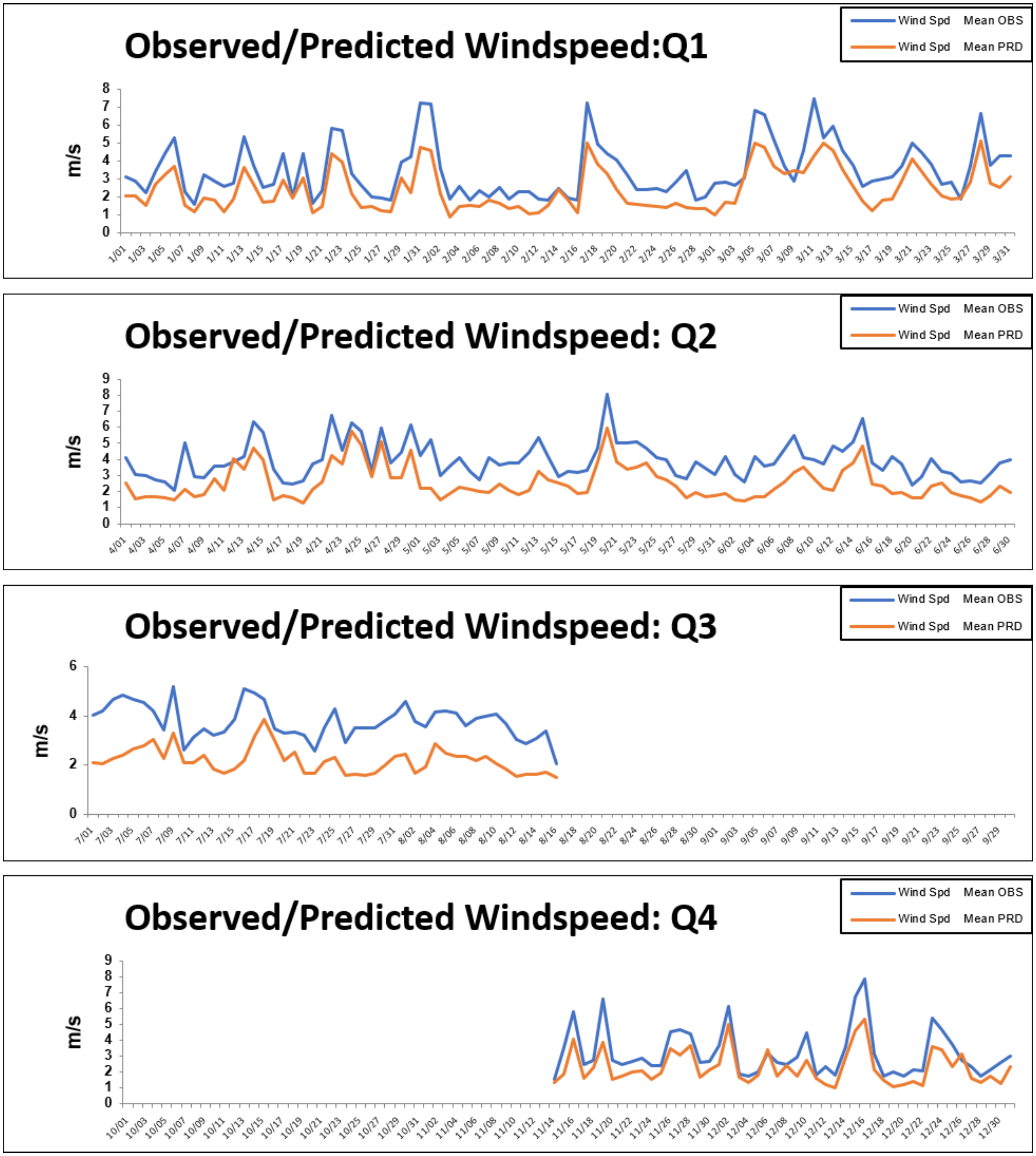
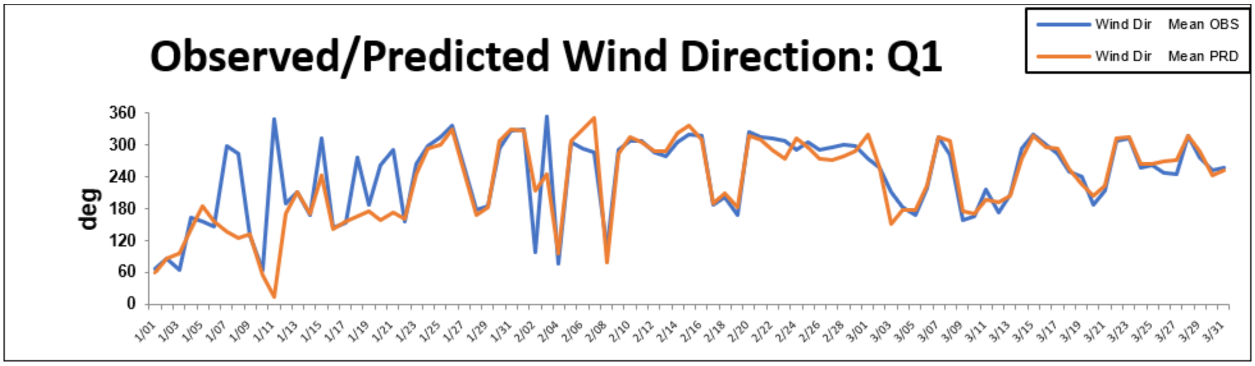
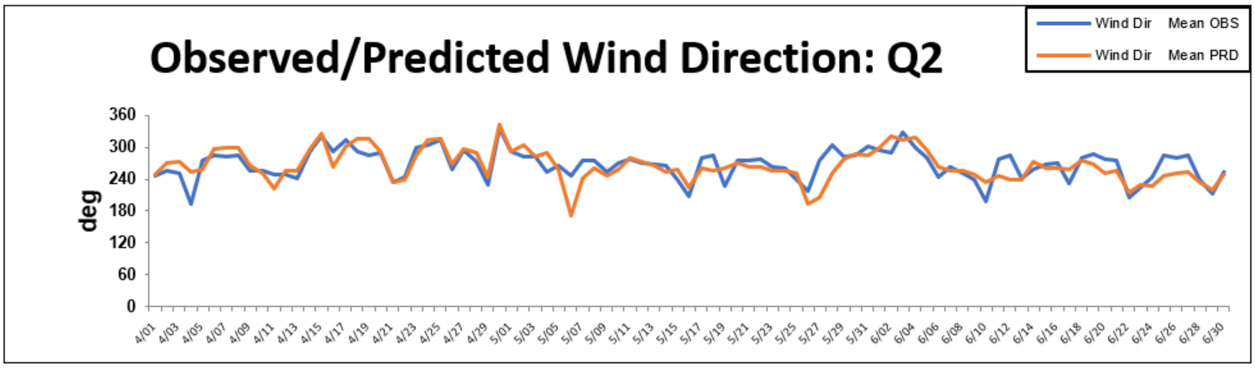
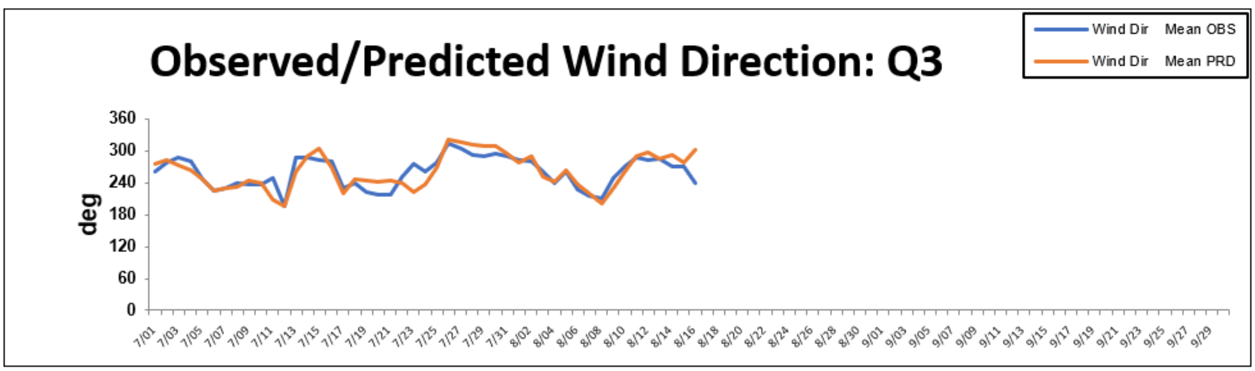


Figure C1: Daily time series of observed and simulated wind speed at West Oakland for 2016 are displayed quarterly. Observation data from mid-August through mid-November were not available. “Mean OBS” is for all observations averaged over the 1-km domain. “Mean PRD” is for all prediction fields at the observation sites averaged over the 1-km domain.







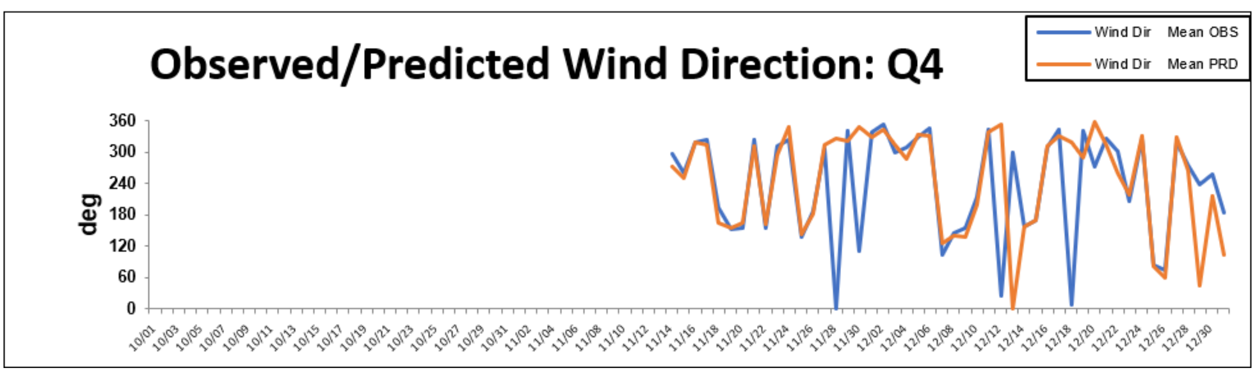


Figure C2: Daily time series of observed and simulated wind direction at West Oakland for 2016.

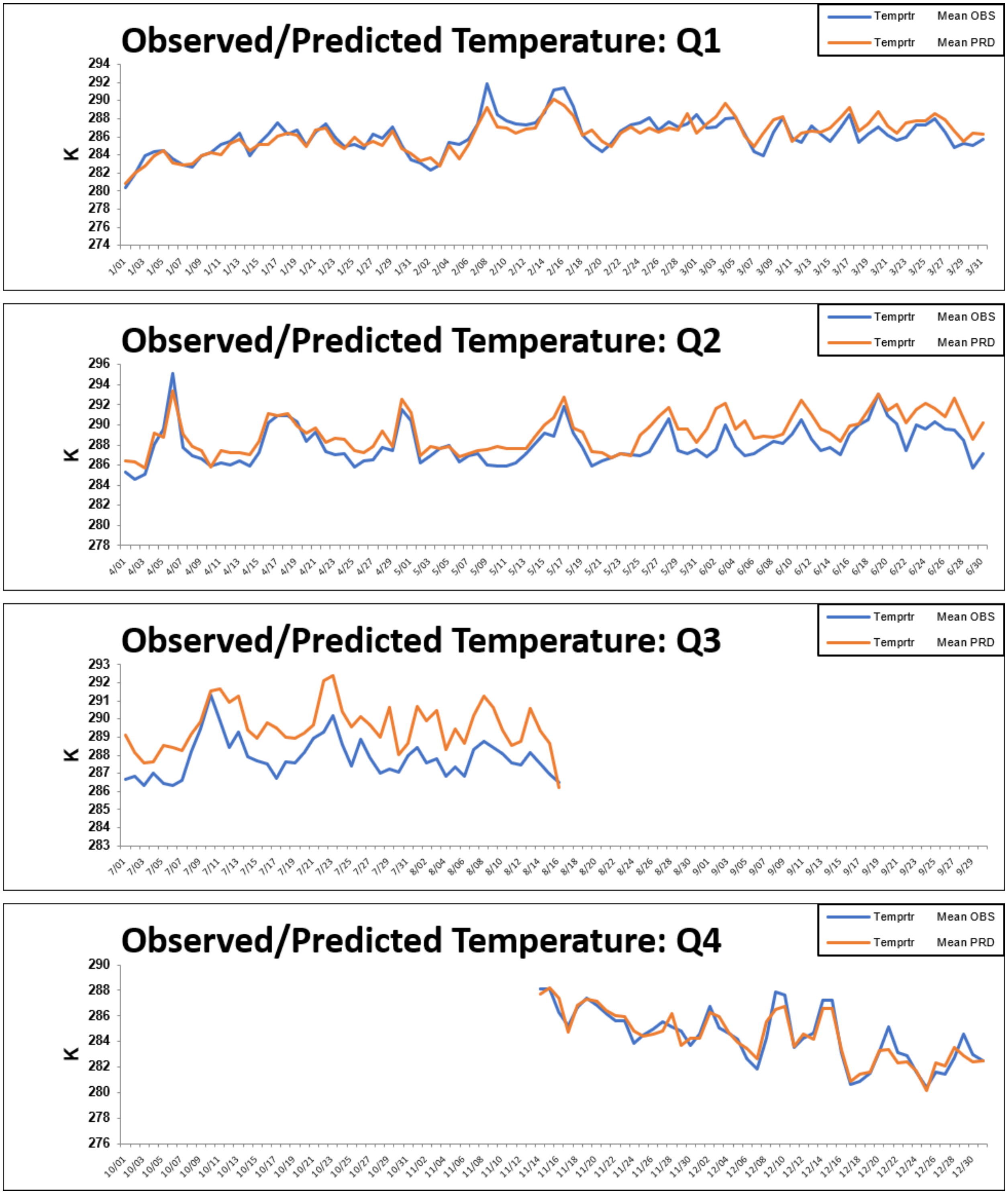


Figure C3: Daily time series of observed and simulated temperatures at West Oakland for 2016.

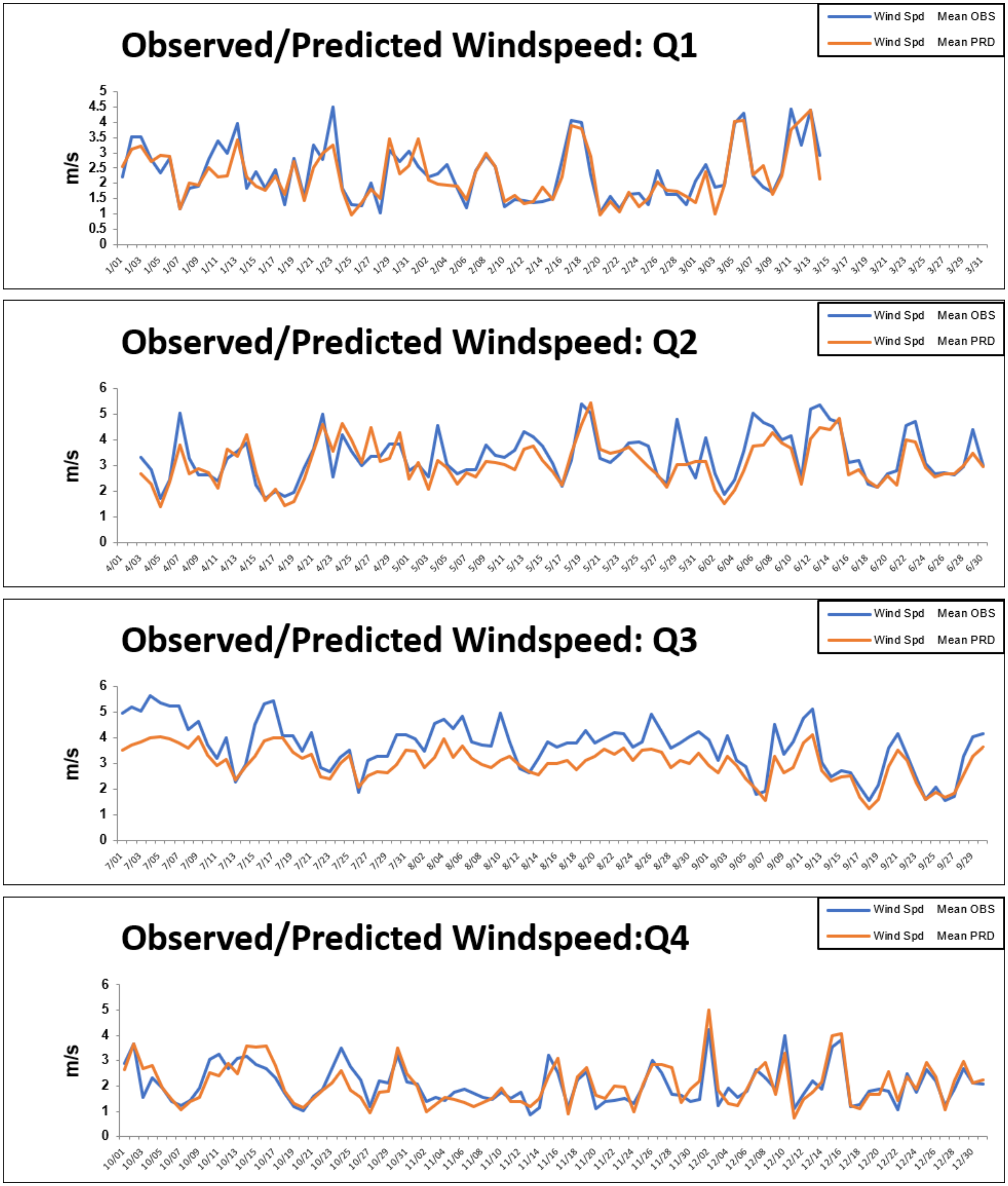


Figure C4: Daily time series of observed and simulated wind speed at Vallejo for 2016.

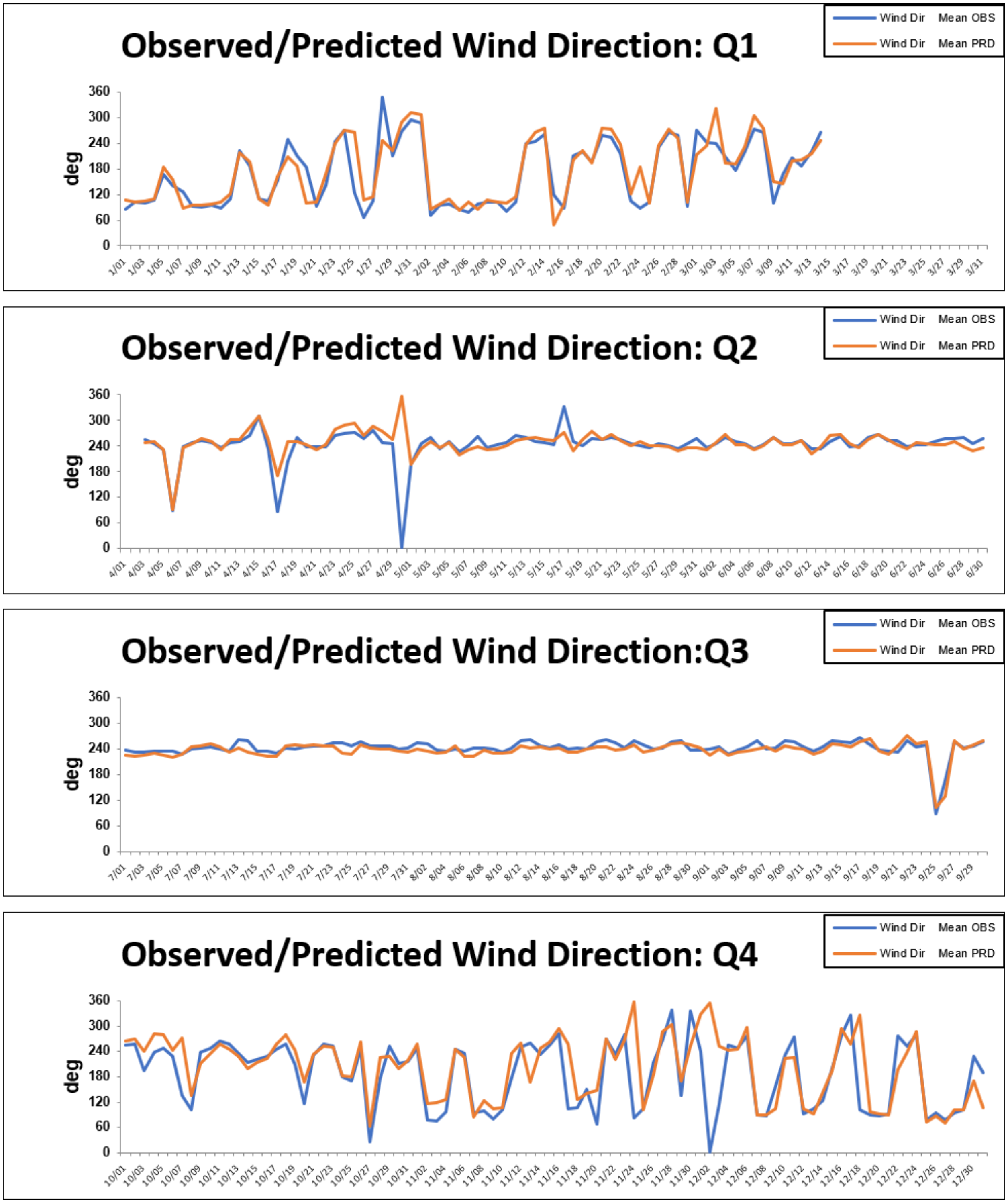


Figure C5: Daily time series of observed and simulated wind direction at Vallejo for 2016.

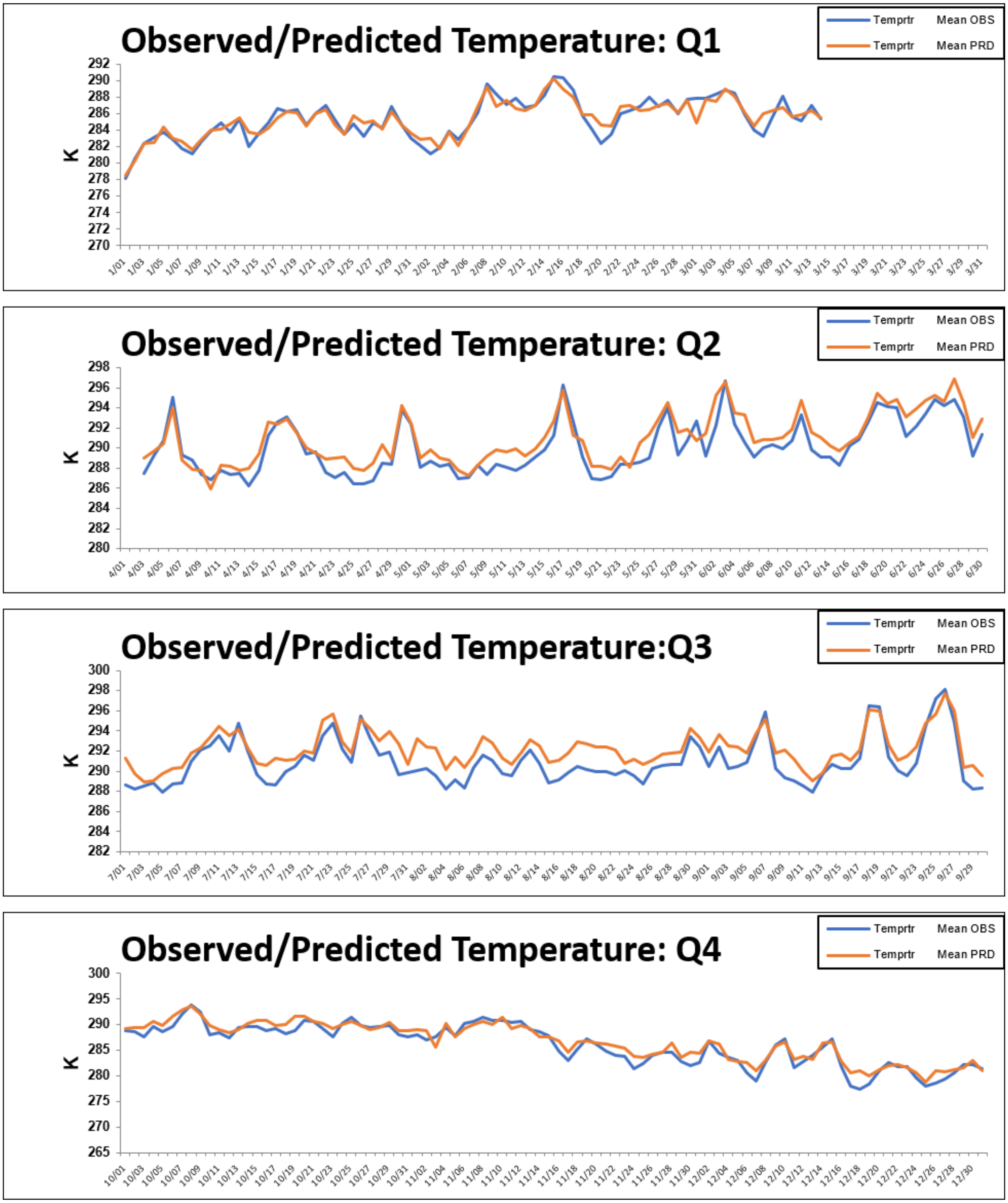


Figure C6: Daily time series of observed and simulated temperatures at Vallejo for 2016.

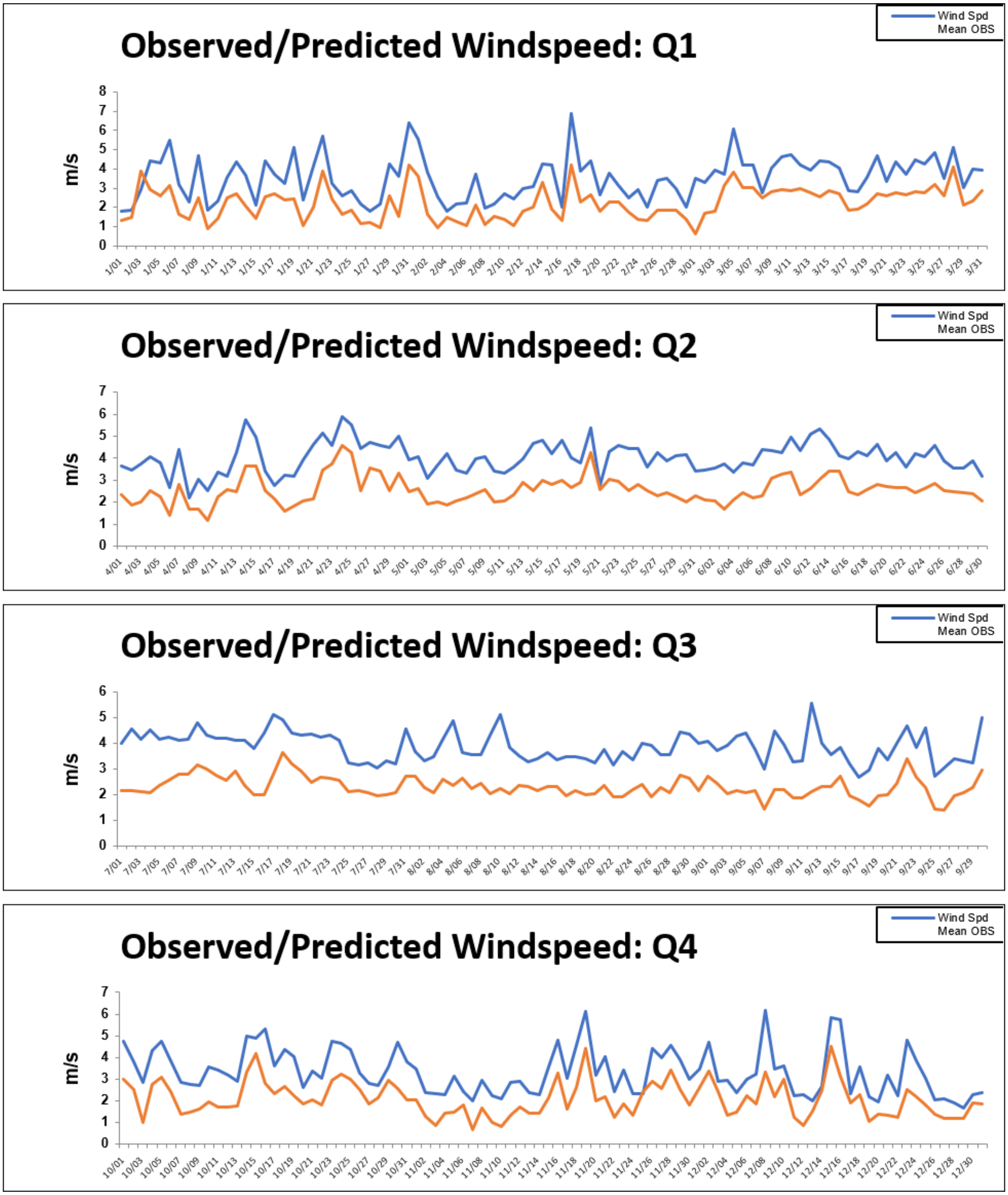


Figure C7: Daily time series of observed and simulated wind speed at San Jose for 2016.

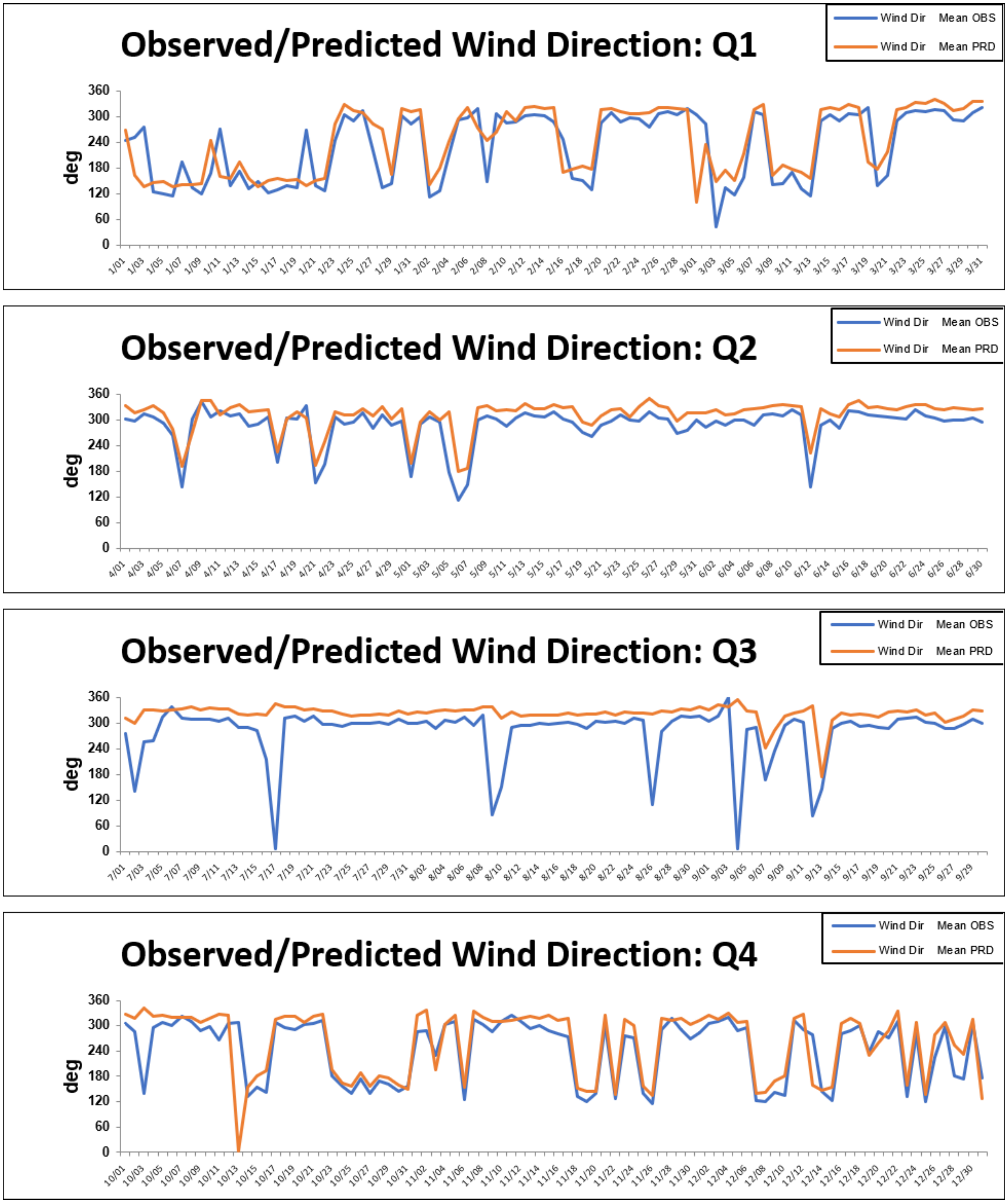


Figure C8: Daily time series of observed and simulated wind direction at San Jose for 2016.

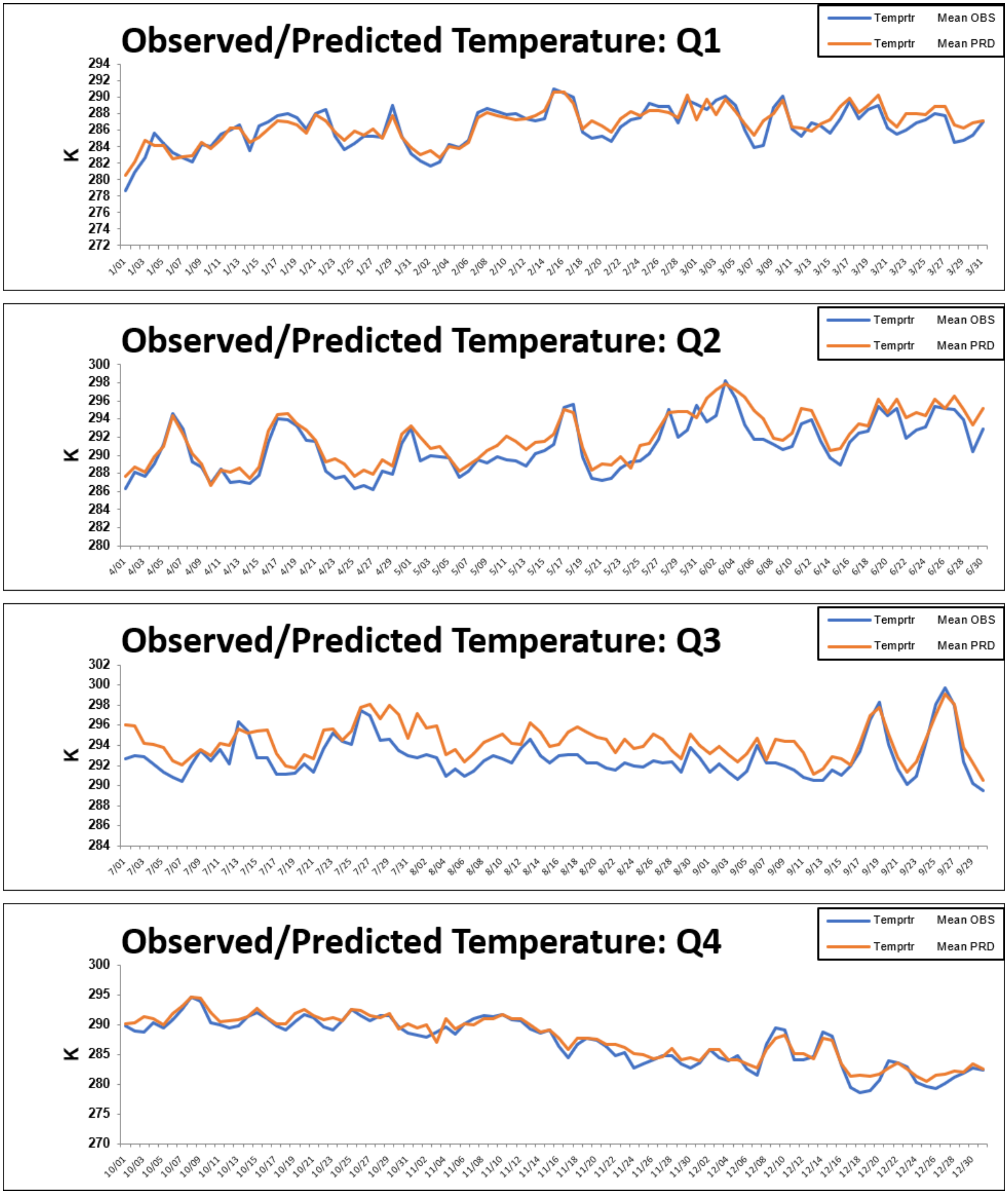


Figure C9: Daily time series of observed and simulated temperatures at San Jose for 2016.

C3. Evaluating WRF Against Upper Air Measurements

There were two upper air stations within the 1-km WRF modeling domain that were operating in 2016. One of them was in Oakland, where the National Weather Service made twice daily measurements at 00 GMT and 12 GMT (4:00 pm and 4:00 am PST, respectively) throughout the year. The other station was at Bodega Bay, where midday measurements were made from May through August, 2016. This was a temporary station established in support of the California Baseline Ozone Transport Study.

Outputs from the 1-km WRF model were compared against measurements at both stations. Day by day, simulations matched observations exceptionally well. Figures C10 and C11 show simulated and observed upper air meteorological data from one winter day (January 10, 2016 at 12 GMT) and from one summer day (June 4, 2016 at 12 GMT) at Oakland. Simulated temperature and dew point (dashed lines) follow observations (solid lines) very well.

Figure C12 shows observed and simulated temperatures at 1:00 pm at Bodega Bay. The simulated temperature matches observations very well.

These are randomly selected plots for the purpose of displaying observed vs. simulated meteorological parameters. They do not necessarily show the best or worst match between the simulation and observations.

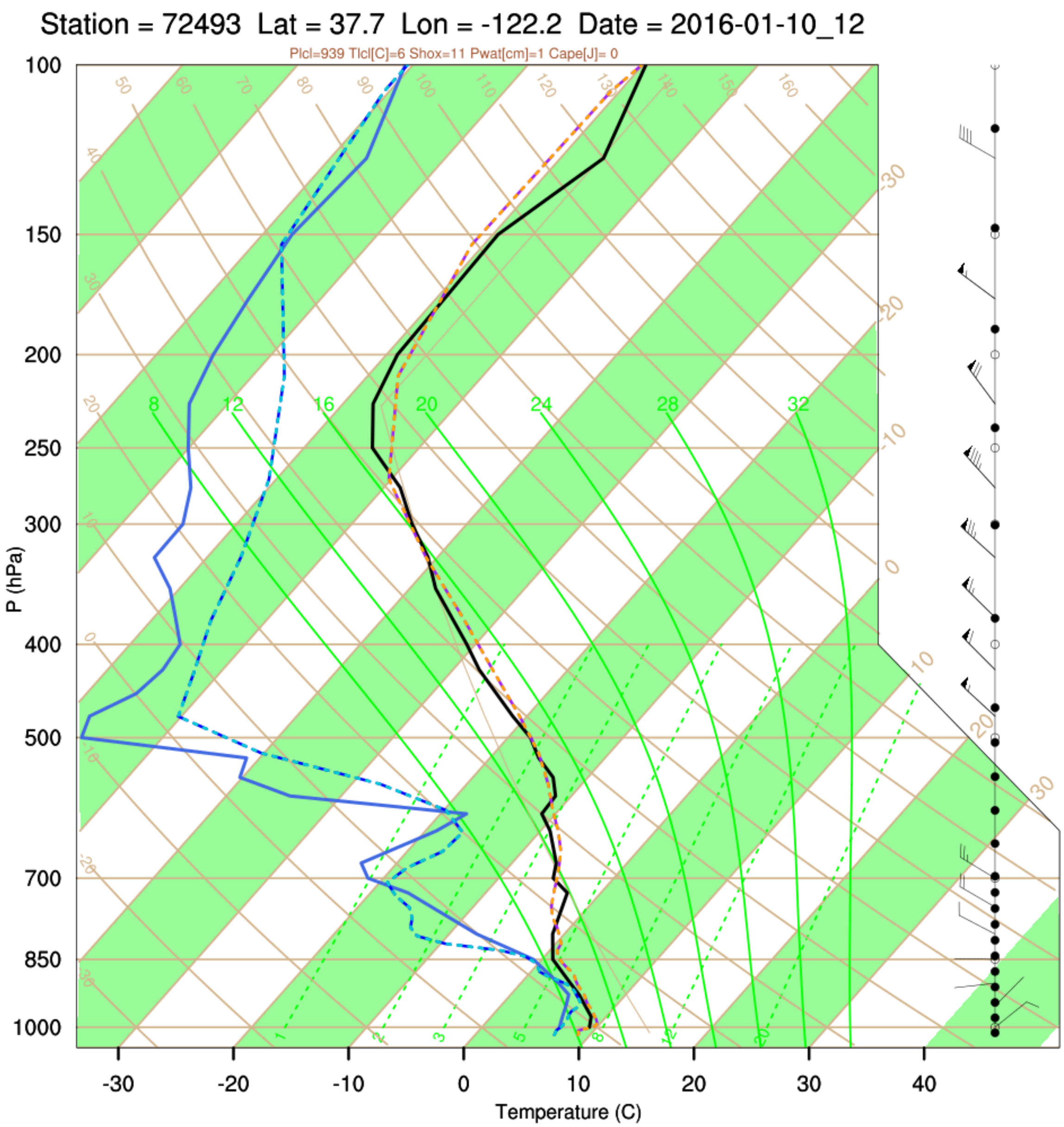


Figure C10: A skew-T plot showing simulated (dashed lines) and observed (solid lines) temperatures (orange and black) and humidity (blue) at Oakland on January 10, 2016 at 12 GMT. Observed wind barbs at pressure levels are shown on the right y-axis.

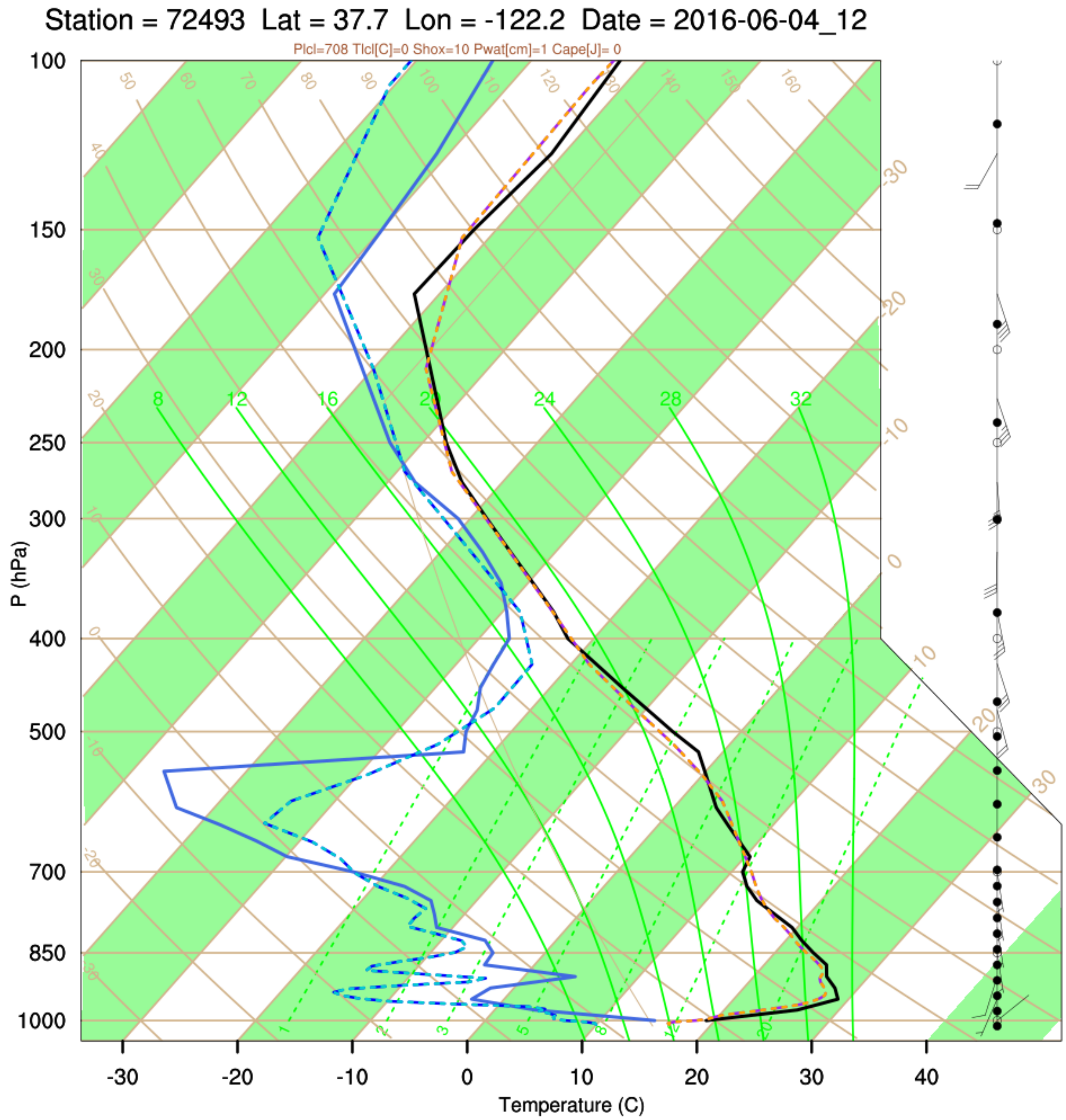


Figure C11: A skew-T plot showing simulated (dashed lines) and observed (solid lines) temperatures (orange and black) and humidity (blue) at Oakland on June 4, 2016 at 12 GMT. Observed wind barbs at pressure levels are shown on the right y-axis.

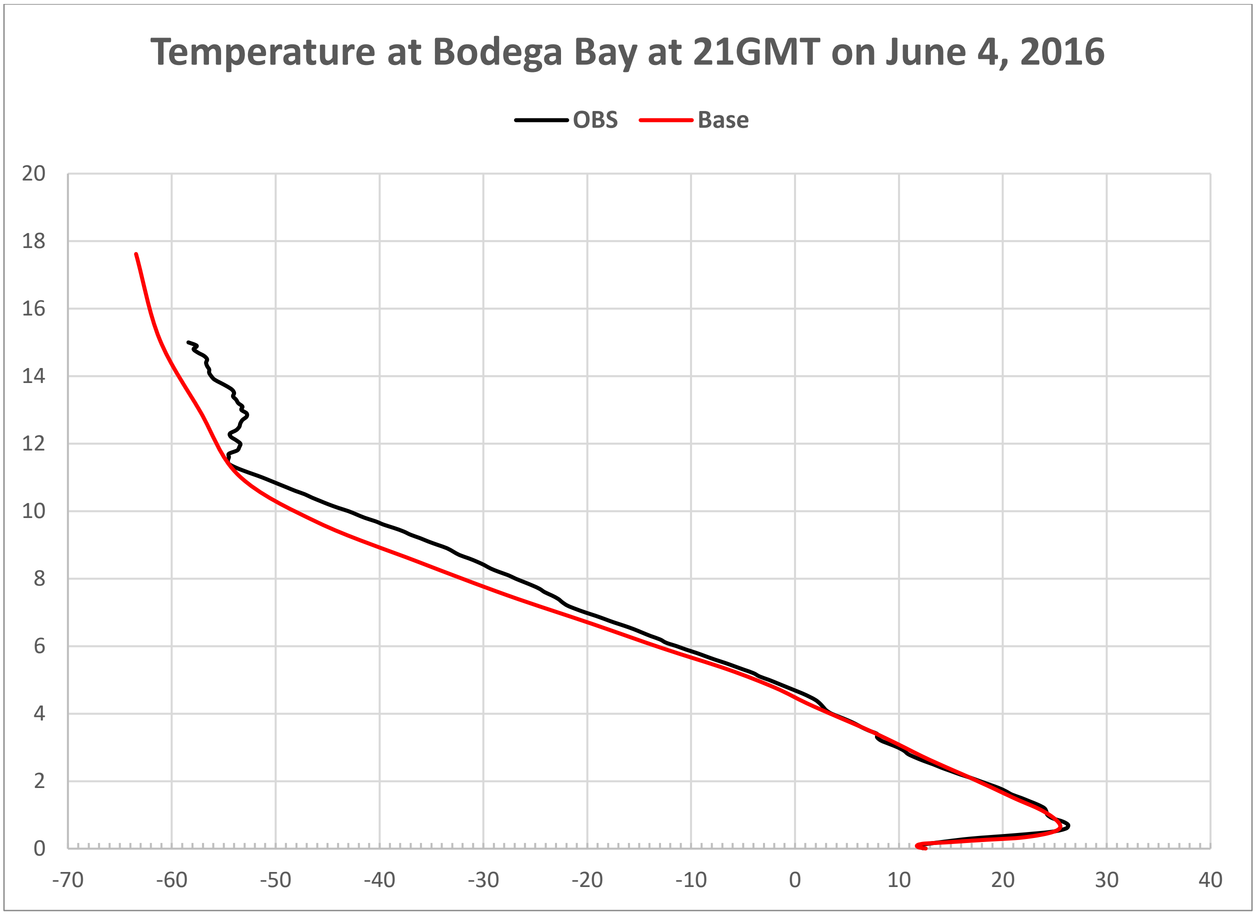


Figure C12: A plot showing simulated (red) and observed (black) temperatures at Bodega Bay at 1:00 pm PST.

Appendix D – Evaluation of the CMAQ Model

D1. Statistical Metrics

Table D1 shows statistical metrics used for CMAQ evaluation. Statistical metrics were calculated from paired daily observed and simulated $PM_{2.5}$ concentrations over quarterly and annual periods. Q1, Q2, Q3 and Q4 represent the 1st, 2nd, 3rd and 4th quarters, respectively. They are defined as January-March, April-June, July-September and October-December.

Table D1. Quarterly and annual statistical model performance metrics.

Metric	Definition[1]	Q1	Q2	Q3	Q4	Annual
Mean bias (MB, μg/m³)	$\frac{1}{N}\sum(P_i - O_i)$	2.3	-0.3	-1.2	0.4	0.3
Mean error (ME, μg/m³)	$\frac{1}{N}\sum\lvert P_i - O_i\rvert$	3.6	2.7	2.9	3.6	3.2
Root mean square error (RMSE, μg/m³)	$\sqrt{\frac{1}{N}\sum(P_i - O_i)^2}$	5.5	3.4	3.7	5.3	4.6
Fractional bias (FB, %)	$100 \times \frac{2}{N}\sum\frac{P_i - O_i}{P_i + O_i}$	23%	5%	-8%	4%	6%
Fractional error (FE, %)	$100 \times \frac{2}{N}\sum\frac{\lvert P_i - O_i\rvert}{P_i + O_i}$	40%	40%	43%	41%	41%
Normalized mean bias (NMB, %)	$100 \times \frac{\sum(P_i - O_i)}{\sum O_i}$	30%	-4%	-16%	4%	4%
Normalized mean error (NME, %)	$100 \times \frac{\sum\lvert P_i - O_i\rvert}{\sum O_i}$	47%	38%	39%	42%	42%
Correlation coefficient (r)	$\frac{\sum\left[\left(P_i - \overline{P}\right)\left(O_i - \overline{O}\right)\right]}{\sqrt{\sum\left(P_i - \overline{P}\right)^2 \sum\left(O_i - \overline{O}\right)^2}}$	0.69	0.35	0.32	0.61	0.56

[1] The summations are taken over all pairs of predictions (P_i) and valid observations (O_i) by site and day, and N is the total number of data pairs. Overbars represent means over the N data.

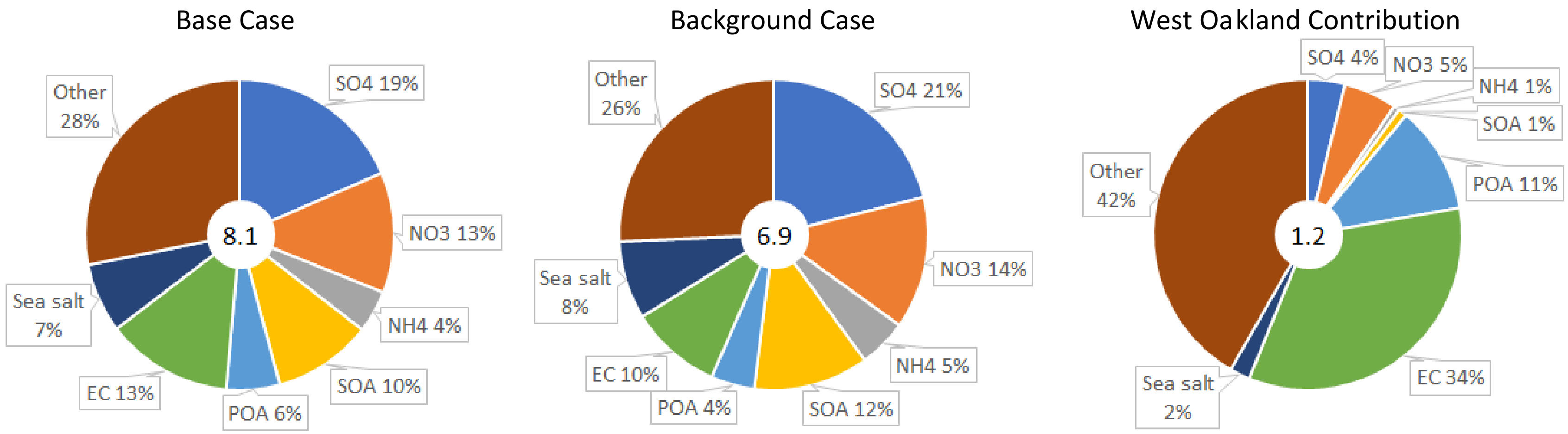
The annual mean bias in simulated $PM_{2.5}$ concentrations is 0.3 μg/m³. On a quarter by quarter basis, the mean bias ranges from -0.3 to 2.3 μg/m³. Among the quarters, Q1 has the highest bias. As explained in the main text, the model is significantly overestimating $PM_{2.5}$ during winter months, especially in February. Possible reasons for the overestimation are under investigation.

Overall, the model shows acceptable $PM_{2.5}$ performance, meeting the goals by Boylan and Russell (2006) and criteria by Emery et al. (2017) for the whole year as well as all 4 quarters.

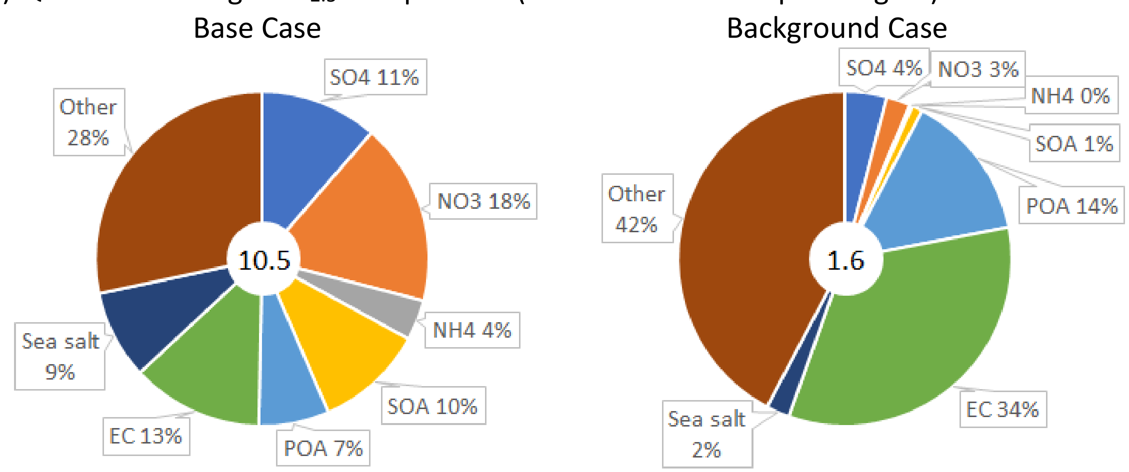
D2. West Oakland $PM_{2.5}$ Composition

Figure D1 shows annual and quarterly average $PM_{2.5}$ compositions over the West Oakland receptor domain for the base and control (i.e., a simulation without West Oakland's anthropogenic emissions) cases as well as the West Oakland contributions (i.e., the difference between the base and control cases). The "Other $PM_{2.5}$" fractions (primary $PM_{2.5}$ mass other than carbonaceous material and sea salt; mostly fugitive dust in this region) are generally the largest component except for the 3rd quarter, where sulfate is the dominant $PM_{2.5}$ component. Secondary $PM_{2.5}$ fractions (ammonium sulfate, ammonium nitrate, and secondary organic aerosol) account for approximately half of total $PM_{2.5}$ mass (ranging from 41% to 63%). The base and control cases exhibit similar $PM_{2.5}$ compositions, indicating that the regional background influence is dominating. The West Oakland contributions are heavily weighted by primary fractions (84% to 93%) from the local sources.

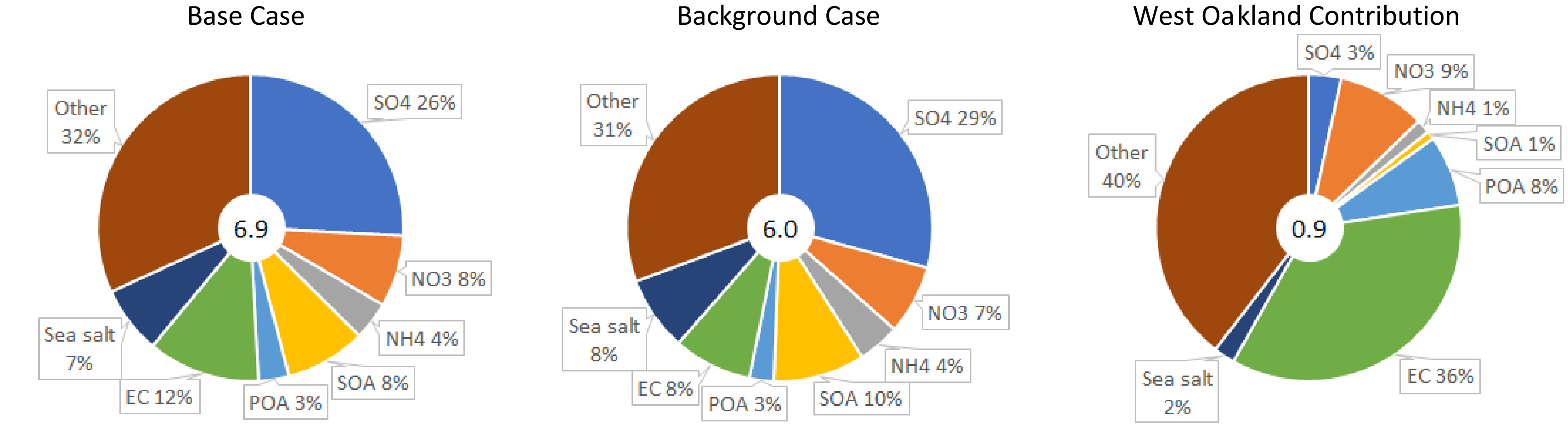
(a) Annual Average $PM_{2.5}$ Composition (West Oakland Receptor Region)



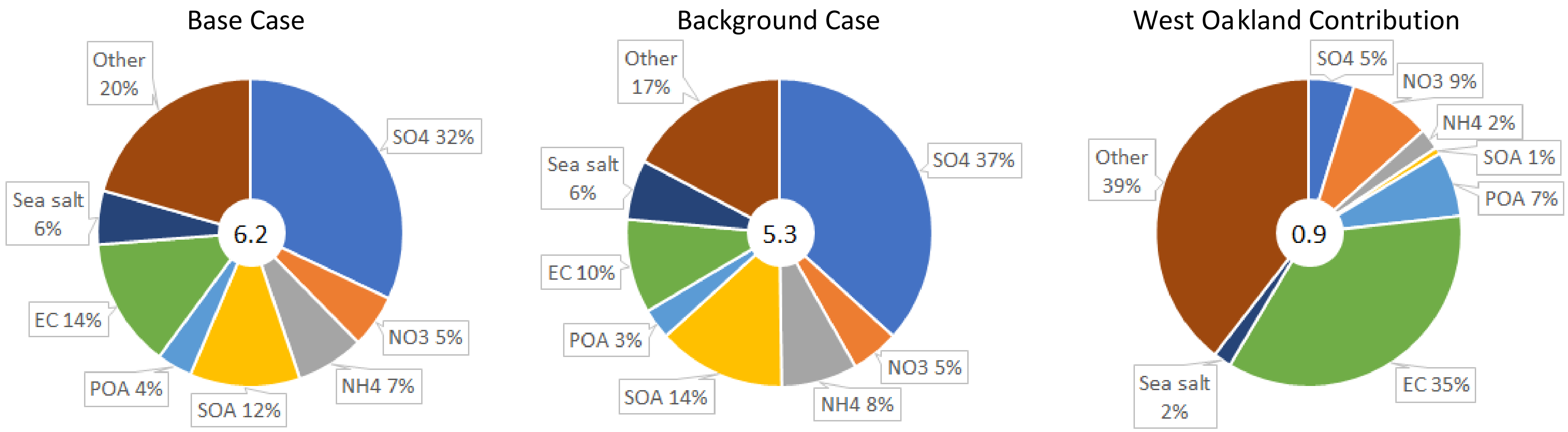
(b) Quarter 1 Average $PM_{2.5}$ Composition (West Oakland Receptor Region)



(c) Quarter 2 Average $PM_{2.5}$ Composition (West Oakland Receptor Region)



(d) Quarter 3 Average $PM_{2.5}$ Composition (West Oakland Receptor Region)



(e) Quarter 4 Average $PM_{2.5}$ Composition (West Oakland Receptor Region)

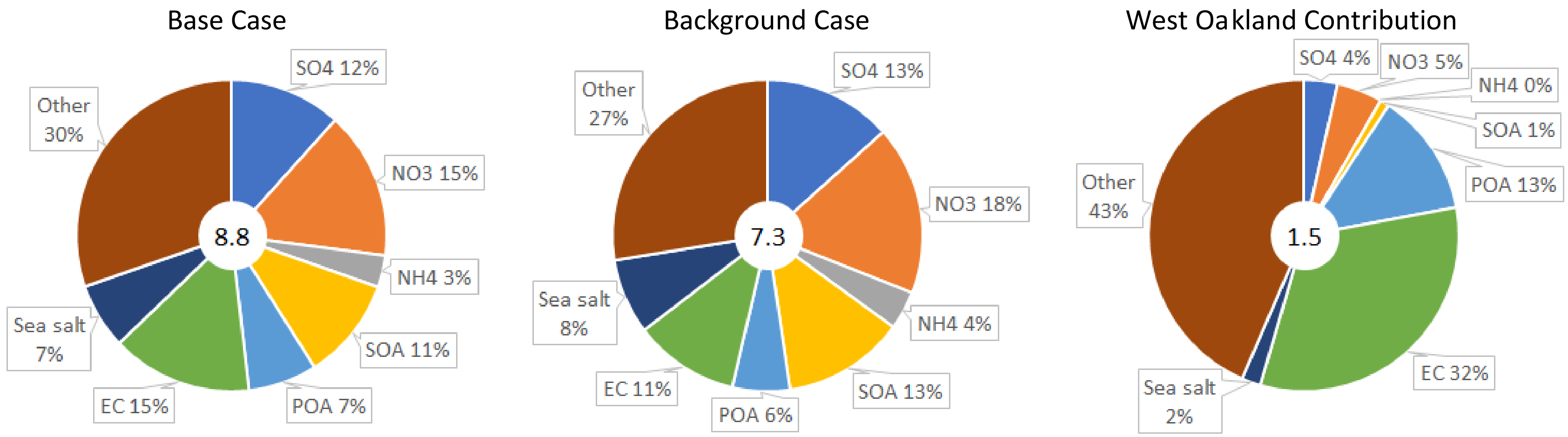


Figure D1: Annual and quarterly average $PM_{2.5}$ compositions over the West Oakland Receptor Region for the base and control cases and their differences (i.e., contributions from the West Oakland anthropogenic emissions). Numbers in the center are total $PM_{2.5}$ concentrations in $\mu g/m^3$.

References

Appel, W., Napelenok, S., Foley, K., Pye, H., Hogrefe, C., Luecken, D., Bash, J., Roselle, S., Pleim, J., Foroutan, H., Hutzell, B., Pouliot, G., Sarwar, G., Fahey, K., Gantt, B., Gilliam, R.C., Heath, N., Kang, D., Mathur, R., Schwede, D., Spero, T., Wong, D., and Young, J., 2017. Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1. Geoscientific Model Development. Copernicus Publications, Katlenburg-Lindau, Germany, 10:1703-1732.

BAAQMD, 2007. Spare the air tonight study, 2007-2008 winter wood smoke season. Bay Area Air Quality Management District, San Francisco, CA (March).

BAAQMD, 2017. Spare the air, cool the climate. Bay Area Air Quality Management District, San Francisco, CA (April).

Boylan, J.W. and Russell, A.G., 2006. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmos. Environ., 40, 4946-4959, doi:10.1016/j.atmosenv.2005.09.087.

Emery, C., Tai, E., and Yarwood, G., 2001. Enhanced meteorological modeling and performance evaluation for two Texas ozone episodes. Prepared for the Texas Natural Resource Conservation Commission by ENVIRON International Corp., Novato, CA (August).

Emery, C., Liu, Z., Russell, A.G., Odman, M.T., Yarwood, G. and Kumar, N., 2017. Recommendations on statistics and benchmarks to assess photochemical model performance. J. Air & Waste Manage. Assoc., 67:5, 582-598, doi:10.1080/10962247.2016.1265027.

EPA, 2018. Modeling guidance for demonstrating air quality goals for ozone, $PM_{2.5}$, and regional haze. U.S. Environmental Protection Agency, Research Triangle Park, NC (EPA-454/R-18-009).

Knoderer, C., Nguyen, D., Alrick, D., and Hoag, K., 2017. 2016 air monitoring network plan. Bay Area Air Quality Management District, San Francisco, CA (July).

Reid, S.B., 2008. Documentation of the preparation of year-2005 emission inventories of toxic air contaminants for the San Francisco Bay Area. Technical memorandum prepared for the Bay Area Air Quality Management District by Sonoma Technology, Inc., Petaluma, CA (January).

Simon, H., Baker, K., and Phillips, S., 2012. Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. Atmos. Environ., 61, 124-139, doi:10.1016/j.atmosenv.2012.07.012.

Tanrikulu, S., Soong, S., Tran, C., and Beaver, S., 2009. Fine particulate matter data analysis and modeling in the Bay Area. Bay Area Air Quality Management District, San Francisco, CA (October).

Tanrikulu, S., Tran, C., and Beaver, S., 2011. Health impact analysis of fine particulate matter in the San Francisco Bay Area. Bay Area Air Quality Management District (September).