USING GROUND-BASED OBSERVATIONS AND SATELLITE RETRIEVALS TO CONSTRAIN URBAN-TO-REGIONAL-SCALE AIR QUALITY CHEMICAL TRANSPORT MODELING

A Dissertation
Presented to
The Academic Faculty

by

Mariel D. Friberg

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
May 2018

USING GROUND-BASED OBSERVATIONS AND SATELLITE RETRIEVALS TO CONSTRAIN URBAN-TO-REGIONAL-SCALE AIR QUALITY CHEMICAL TRANSPORT MODELING

Approved by:

Dr. James A. Mulholland, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Yang Liu
Rollins School of Public Health
Emory University

Dr. Armistead G. Russell
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Ralph A. Kahn
Earth Sciences Climate and Radiation Laboratory
NASA Goddard Space Flight Center

Dr. Yuhang Wang
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Date Approved: [01 10, 2018]

To you.

ACKNOWLEDGEMENTS

This effort has been a long time in the making and an enterprise that would not have been possible without the assistance and cooperation of many people. First, I would like formally to thank the School of Civil and Environmental Engineering for the opportunity to pursue the Ph.D. at the Georgia Tech Institute of Technology and for their financial support in the form of research assistantships, a fellowship, and research facilities. Among the faculty, I am particularly grateful to James A. Mulholland, my academic advisor, who took on this challenge and continued to see it through. My gratitude is likewise extended to other committee members: to Ralph A. Kahn, my NASA advisor and mentor extraordinaire, for his patience and encouragement at several opportune times during the dissertation ordeal; to Armistead G. Russell for his advisement and group meetings; to Yang Liu for his mentorship; and Yuhang Wang for his supportive role. Additionally, I would like to recognize the support of the SCAPE research group at Georgia Tech and Emory University for their expertise and fellowship.

This work was supported under Grant Number R834799 issued through the US EPA and Grant Number NNX13AR89H issued through the NASA Office of Education Minority University Research Education Project (MUREP) as part of the NASA Harriett G. Jenkins Graduate Fellowship activity.

Two very special personal acknowledgements go to Phillip R. Stratton and Raquel Alcaraz-Figueroa for their unwavering support. And finally for my mother and father, Ruth I. Aponte-Diaz and Justin C. Friberg, a simple acknowledgement seems so totally inadequate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The overarching goal of this research is to improve urban- and regional-scale air quality modeling for health risk assessment using a combination of ground-station and satellite-based measurements. The integration of near-surface air pollution concentrations, emissions-based air quality model simulations, and satellite observations of column-integrated quantities will improve the accuracy of exposure metrics and promote the appropriate use of satellite data over extended areas for long time periods, especially where ground-based air quality measurement networks are limited or lacking. In the broader sense, this information will help public health scientists, policy makers, and monitoring agencies to research and implement better control strategies and regulations.

The first phase of this study (Friberg et al., 2016) demonstrated and assessed a systematic and practical approach to fusing surface-network measurements with chemical transport model (CTM) simulations to produce daily concentrations for five air pollutants of primary origin (NO_2, NO_x, CO, SO_2, and EC), and seven secondary pollutants (O_3, PM_{10} mass, $PM_{2.5}$ mass, SO_4, NH_4, NO_3, and OC) for use in cross-sectional epidemiological studies. A second study (Friberg et al., 2017) assessed the ability of the data fusion method to produce daily concentrations across five metropolitan areas from 2002 to 2008. In addition to the variety of pollutant types, the five cities represent a range of meteorological conditions, background aerosol conditions, population densities, and sampling-station distributions. Among the pollutant types, the primary pollutants tend to be more heterogeneously distributed over the urban regions than the secondary

ones. The resulting daily spatial field estimates of air pollutant concentrations and associated correlations were not only consistent with observations, emissions, and meteorology, but substantially improved CTM-derived results in areas without observations, for most pollutants and all cities. The data fusion improved daily metrics across all pollutants with the greatest improvements occurring for O_3 and $PM_{2.5}$.

The final study (Friberg et al., 2017, to be submitted) demonstrated and assessed an optimization technique, expanding upon the surface-station-model fusion technique, to estimate ambient $PM_{2.5}$ mass and associated chemically speciated concentrations for potential use in longitudinal epidemiological studies. The newest method constrains surface $PM_{2.5}$ and chemical-component CTM results, using both ground-station data to anchor speciated, near-surface aerosol concentrations, and total column aerosol optical depth (AOD). When the mid-visible AOD is high, the retrieved AOD from the Multi-angle Imaging SpectroRadiometer (MISR) Research Aerosol retrieval algorithm along with qualitative, column-effective aerosol type observations helped constrain the CTM over extended regions. The retrieved AOD had a horizontal resolution of 275m. The method was applied over a case study area in the San Joaquin Valley of California during NASA's DISCOVER-AQ field campaign in this region, on days when there was good satellite coverage and considerable suborbital data for validation of the approach. The accuracy of estimated concentrations and evaluation of the latest MISR aerosol retrieval algorithm ability to typify urban AOD, aerosol mixtures, and aerosol airmasses were examined by comparing the results with speciated ground observations and standard model fitting statistics. The results indicate that on days with high AOD and adequate

observing conditions, satellite retrievals improve simulated spatial distributions of $PM_{2.5}$ and chemical component concentrations.

CHAPTER 1. INTRODUCTION

Approximately 4.2 million deaths and 103.1 million years of shortened life expectancy were attributed to long-term exposure to ambient fine particle air pollution ($PM_{2.5}$; particle size less than 2.5 μm in the aerodynamic diameter) in 2015, making $PM_{2.5}$ the fifth-ranking mortality risk factor and sixth rank for lost years of healthy life in 2015.[1] As the top environmental risk factor from the Global Burden of Diseases Study, long-term poor air quality exposure epidemiological and toxicological studies play an important role in informing public policy for mitigation purposes and related health benefits.[1] Such health studies have shown that poor air quality poses a hazard to public health by increasing in cardiopulmonary morbidity and mortality,[2-19] contributing to the incidence and development of diabetes mellitus, adverse birth outcomes,[20-23] and posing a certain risk to public health even at levels below national standards).[23-27] In addition to knowledge gaps, the understanding of underlying mechanisms/pathways whereby gases and fine particulate matter cause adverse effects to public health (i.e., inducing intracellular oxidative stress, genotoxicity, and inflammatory responses;[28] there is also a critical need to assess the $PM_{2.5}$ chemically speciated health effects and the synergistic effects of air pollutants on multiple risk factors.

Large population epidemiologic studies of acute and chronic health effects related to ambient air pollution rely on accurate estimates of the spatial and temporal variation in a number of monitored air pollutants. Spatial coverage of air quality surface monitoring

networks is sparse, and some pollutants of interest are not measured daily, even in well-instrumented areas.

This sparseness in ground-level air quality monitoring data has led to efforts to improve spatiotemporal exposure estimates for epidemiological studies by incorporating more spatially robust datasets such as chemical transport models and satellite data. Emissions-based air quality modeling does not have the spatial and temporal discontinuities of the observations, but has well-known inaccuracies due to such limitations as insufficient characterization of meteorological and emission inputs as well as inexact representation of physical and chemical processes.

Improvements in surface air pollution spatial and temporal estimates, obtained by combining publicly available air monitoring data and chemical transport model (CTM) outputs, have the potential to provide added accuracy in assessing human health exposure risks and to aid regulatory efforts. High-resolution spatiotemporal air pollution datasets are in great demand by public health researchers due to the need to identify both specific drivers of adverse health effect associations, including multi-pollutant mixtures, and potential effect-modification factors. Accurate spatiotemporal characterization of ground-level ambient air pollution can help when comparing inter-city health risk estimates as well as intra-city effect-modification factors such as socio-economic status.[13, 29-31] Factors limiting studies of ambient air pollution and health effects relationships include the temporal and spatial completeness and accuracy of air quality estimates, and the availability of simulated concentrations for multiple pollutants. In general, narrower-

swath, polar-orbiting satellites will especially help more with long-term (i.e., chronic) exposure, whereas the CTMs will help with both short- and long-term exposure.

To investigate air pollution health effects on humans, population-based epidemiologic time-series studies often use exposure metrics developed from a single central monitoring site or from regulatory monitoring networks.[32-34] Although these data provide useful indicators for air pollution impacts at specific locations, the data are typically limited temporally, spatially, and chemically.[35] Given the sparse spatial coverage of air quality monitoring networks, urban-to-rural gradients are poorly characterized. In addition, the spatial network of monitors for each pollutant is different. The impact that measurement error has on health risk assessment depends on the amount and type of error, which in turn depend on the number and location of monitors.[36-37]

1.1 Chemical Transport Models

CTM simulations based on meteorological and emission data inputs and physical and chemical transport process modeling provide more complete spatial and temporal coverage than surface-network observations. However, such modeling is computationally intensive, often contains biases, and tends to minimize day-to-day variability.[38]

The U.S. Environmental Protection Agency (EPA) has joined with other agencies to provide hourly air pollutant concentration fields for the purpose of public health analysis over the contiguous U.S. and specific regions, at high-spatial resolutions ranging between 2 and 12 km, using the Community Multi-Scale Air Quality (CMAQ) model.[39-41]

The accuracy of these publicly available simulated fields, which varies by pollutant, in turn is affected by the accuracy of meteorological and emission inputs, and physical and chemical model specifications.[42]

1.2 Satellite Aerosol Data

The heterogeneity of aerosols places rigorous requirements on measurement systems to effectively characterize aerosols on regional and global scales.[43] Under favorable retrieval conditions, aerosol retrieval algorithms applied to satellite data can differentiate among dominant tropospheric aerosol types, at least crudely.[44] CTMs simulate aerosol type, aerosol optical depth (AOD) gradients, and near-surface aerosol concentration information, in part to help develop fine PM emission reduction strategies and better protect human health.[45-46] There is significant but as yet limited experience at applying space-based observations to constraining air quality, although such observations show great promise at reducing the uncertainties in regions-scale and long-term modeling efforts.[45, 47-48]

Advancements in the interpretation of satellite retrievals, such as aerosol types, can lead to further developments in near-surface air quality conditions diagnostics. Aerosol airmass types can be identified under some circumstances as: biomass burning, continental, dust, maritime, urban, and smoke-dust hybrid.[49] More generally, satellite-based particle-type classification amounts to optical constraints on particle size, shape, and light-absorption properties.[50] Improvements in characterizing near-surface air quality in turn can further our understanding of uncertainties related to ambient air pollution

exposure-response studies. Moderate Resolution Imaging Spectroradiometer (MODIS) performs global AOD retrievals. The standard Dark Target (DT) algorithm performs well over Case 1 waters and vegetated surfaces, but poorly over bright surfaces, and does not retrieve additional aerosol optical property information needed to identify aerosol air mass types.[51]

Although most satellite instruments provide only a single view, either toward nadir or the horizon, the NASA Earth Observing System's Multi-angle Imaging Spectroradiometer (MISR) employs nine push-broom cameras at nine different view angles directed toward Earth's surface. Each of MISR's cameras images in blue, green, red, and near-infrared spectral bands. Unlike MODIS, MISR is much less dependent on any assumed relationship of surface reflectance between bands because of its additional viewing angles. This is an especially important consideration over brighter land surfaces, such as most urban areas. MISR's unique configuration lends itself to studying Earth's aerosols, clouds, and surface. In particular, the instrument offers qualitative constraints on particle size, shape, and light-absorption under favorable retrieval conditions that make it possible to map out aerosol air masses, even over populated areas.[47]

This thesis focuses on the use of measurements to remove biases and generally improve the accuracy of air quality modeling by applying surface *in situ* and satellite measurements. Chapter 2 elaborates on the key objectives pursued for this thesis. Chapter 3 contains published work presenting a new method for fusing surface-based *in situ* measurements with an air quality model, and Chapter 4 contains published work applying this technique to multi-year air quality over five cities. In Chapter 5, the

approach is expanded to add spatially extensive MISR data are to the model constraints. The final chapter provides a summary of work to date, and prospects for future application of the methods developed here.

CHAPTER 2. OBJECTIVES

The overarching objective of this dissertation is to improve air quality modeling for health risk assessment, by developing techniques for applying surface-station data and satellite retrieval results to constraining near-surface conditions represented in chemical transport models (CTMs). By integrating near-surface air pollution concentrations, emissions-based air quality model simulations, and satellite observations of column-integrated quantities, we aim to improve the accuracy of exposure metrics on urban-to-regional scales. As part of this effort, we develop techniques that promote the use of satellite data to improve modeling in areas with limited ground-based air quality measurement networks. In the broader sense, this information is intended to help public health scientists, policy makers, and monitoring agencies to research and implement better control strategies and regulations. Figure 2.1 presents a flow chart of the four phases that compose the thesis research.

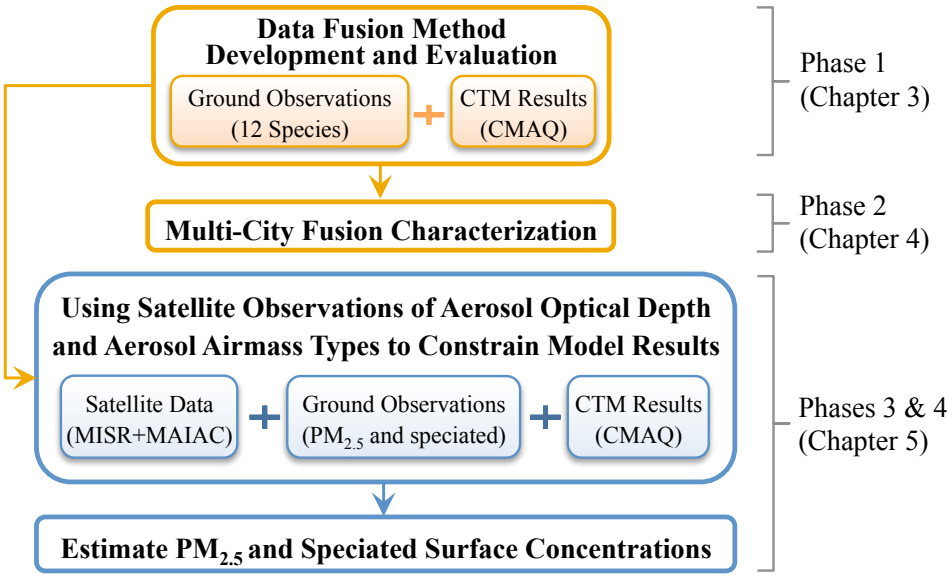

Figure 2.1 – Flow chart of thesis phases and associated chapters.

2.1 Phase 1: Data Fusion Method

Question. How can blending surface-station observations and CTMs improve the accuracy of multiple air pollution estimates away from monitor sites? ***Objective.*** Develop concentration estimates and evaluate error estimates over space and time for use in health studies by fusing observations and CTM simulations. Increase the temporal and spatial coverage of concentration metrics while still reflecting the observed concentrations. Case study was performed for Georgia. ***Rationale.*** Although observations provide low error at and in proximity to monitors, the dataset contains temporal sampling gaps and offers minimal spatial coverage. The regional air quality chemical transport model, Community Multiscale Air Quality (CMAQ), provides wider spatial coverage and relationships to emissions, but is largely unconstrained by observations.

2.2 Phase 2: Multi-city Fusion Characterization

Question. Can multi-city fused air quality characterization help explain between-city spatial heterogeneity? ***Objective.*** Evaluate applicability of the observation and CMAQ data fusion methodology to other locations. Characterize inter-city air quality heterogeneity and uncertainties for five cities using the data-fusion method. ***Rationale.*** Understand representativeness and variability among five fused multi-city concentrations. Elucidate the degree to which air quality exposure metric trends are generalizable among various locations that have different meteorology and emissions.

2.3 Phase 3: Aerosol Air Mass Type Mapping

Question. Can advancements in the interpretation of aerosol types from satellite retrievals improve the accuracy of near-surface-condition diagnostics by providing broad regional context? ***Objective.*** Following the work of *Patadia et al*. (2013) over Mexico City, characterize aerosol airmass types over populated regions of Southern California using satellite data, taking advantage of suborbital data acquired during the DISCOVER-AQ field campaign in 2013 to validate the results. Advance the application of aerosol types in satellite retrievals, which will lead to improved, regional-scale observational constraints on near-surface air quality. ***Rationale.*** Unlike most other *passive* remote sensing techniques currently available, MISR multi-angle and multi-spectral data can distinguish between surface and atmospheric signals and can map aerosol air mass types under favorable retrieval conditions. POLDER is another passive satellite imager that can help constrain aerosol type to some extent, but at much lower spatial resolution, making it

less useful for most air-quality studies. Having the satellite 2-D data can provide some spatial constraints to the model, especially useful in areas lacking ground monitors.

2.4 Phase 4: Relating Aerosol Mass Type Maps & AOD to Surface $PM_{2.5}$

Question. To what degree can regional-scale satellite and CTM data be combined to improve surface $PM_{2.5}$ estimates progressively downwind of sources? ***Objective.*** Develop, apply, and evaluate a method for incorporating aerosol air mass maps and total column aerosol optical depth into a regional-scale CTM, to obtain surface $PM_{2.5}$ constraints for potential use in regional-scale and long-term pollution-exposure health studies. ***Rationale.*** In addition to AOD that is commonly applied in such work, MISR retrievals contain information about aerosol optical properties, such as particle size and sphericity that help identify aerosol airmass types. CMAQ provides uniform spatial and temporal coverage, offers additional physical constraints independent of these observations, and identifies relationships between species concentrations and emissions. Incorporating satellite aerosol information in the development of $PM_{2.5}$ concentration metrics leads to a decrease in metric uncertainties and errors.

CHAPTER 3. METHOD FOR FUSING OBSERVATIONAL DATA AND CHEMICAL TRANSPORT MODEL SIMULATIONS TO ESTIMATE SPATIOTEMPORALLY-RESOLVED AMBIENT AIR POLLUTION

(Friberg, M.D., Zhai, X., Holmes, H.A., Chang, H.H., Strickland, M.J., Sarnat, S.E., Tolbert, P.E., Russell, A.G. and Mulholland, J.A. *Environmental Science & Technology, 50(7)*: 3695-3705. 20,6)

Abstract

Investigations of ambient air pollution health effects rely on complete and accurate spatiotemporal air pollutant estimates. Three methods are developed for fusing ambient monitor measurements and 12-km resolution chemical transport model (CMAQ) simulations to estimate daily air pollutant concentrations across Georgia. Temporal variance is determined by observations in one method, with the annual mean CMAQ field providing spatial structure. A second method involves scaling daily CMAQ simulated fields using mean observations to reduce bias. Finally, a weighted average of these results based on prediction of temporal variance provides optimized daily estimates for each 12-km x 12-km grid. These methods were applied to daily metrics of 12 pollutants (CO, NO_2, NO_x, O_3, SO_2, PM_{10}, $PM_{2.5}$, and five $PM_{2.5}$ components) over the state of Georgia for a seven-year period (2002-2008). Cross-validation demonstrates a wide range in optimized model performance across pollutants, with SO_2 predicted most poorly

due to limitations in coal combustion plume monitoring and modeling. For the other pollutants studied, 54-88% of the spatiotemporal variance (Pearson R^2 from cross-validation) was captured, with ozone and $PM_{2.5}$ predicted best. The optimized fusion approach developed provides daily spatial field estimates of air pollutant concentrations and uncertainties that are consistent with observations, emissions, and meteorology.

3.1 Introduction

High-resolution spatiotemporally-resolved air pollution datasets are needed by public health researchers for identification of specific drivers of adverse air pollution health effects, including multi-pollutant mixtures, and potential effect modifiers (e.g., synergistic effects of multiple risk factors). For example, accurate spatiotemporal characterization of ground-level ambient air pollution, when used in citywide health effect estimation, can facilitate inter-city comparison of health effect estimates. In addition, such data can be used for within-city effect estimation, and assessment of effect modification by spatial factors such as neighborhood socioeconomic status.[52-56] Major factors currently limiting such studies include lack of temporal and spatial completeness, and accuracy of air quality estimates. These limitations are addressed in this work.

To investigate air pollution health effects on humans, population-based epidemiologic time-series studies often use exposure measures derived from regulatory monitoring networks.[32-34] These data are limited temporally, spatially, and chemically.[35] Many pollutants of interest are measured infrequently, at very few locations, or not at all. With the typically sparse spatial coverage of air quality monitoring networks, spatial

variability, such as the urban-to-rural gradient, is poorly characterized. In addition, the spatial distribution of monitors for each pollutant is different. The impact of measurement error on health risk assessment is dependent on the amount and type of error, which depend in part on the number and location of monitors.[37, 57] Averaging of ambient monitor data can improve how well the exposure metric represents an area or a population.[58] However, such averaging requires estimating the spatial distribution of pollutants. Ambient monitor data can also be limited temporally. Such is the case for fine particulate matter ($PM_{2.5}$) and its components, which are often measured once every three or six days. Observation-based spatial and temporal interpolation approaches, such as spatial kriging and temporal linear interpolation, typically introduce excessive autocorrelation in the estimates.[11]

Chemical transport model (CTM) simulations provide more complete spatial and temporal coverage based on meteorological and emission data inputs and physical and chemical transport process modeling. However, such modeling is computationally intensive, has biases, and tends to under-estimate day-to-day variability.[38, 40] The U.S. Environmental Protection Agency (EPA) and Centers for Disease Control (CDC) have collaborated in the Public Health Air Surveillance Evaluation (PHASE) project to provide hourly air pollutant concentration fields for the eastern U.S. at a spatial resolution of 12-km for the period 2002-2008 using the Community Multi-Scale Air Quality (CMAQ) model for the purpose of public health analysis.[38, 40, 59] The accuracy of these publicly available simulated fields, which varies by pollutant, is affected by the accuracy of meteorological and emission inputs, and physical and chemical model specifications.[42]

Error in these fields can be characterized by comparison with observational data. Various statistical approaches have been applied to correct known biases in CMAQ estimates.[58, 60]

Statistical regression approaches combine ground-based measurements and model simulations to take advantage of the high degree of spatial and temporal completeness of the simulations.[59, 61-67] One example is the blending of satellite aerosol optical depth data, ground-based monitor data, and land use data [45, 58, 68] Geostatistical methods such as land use regression (LUR) have several limitations, including limited ability to separate impacts between pollutants, and potential introduction of spatially-varying confounders in epidemiologic studies.[69-72] A large number of measurements, often obtained from specific monitoring-intensive campaigns, are required in LUR applications to minimize predictor sensitivity and maximize model performance.[73-75] LUR applications are generally quite spatially refined and focused on estimating spatial variability in long-term air pollution exposures.

The limitations of spatially sparse and temporally incomplete ambient monitor data can be overcome by blending these data with estimates from a physically and chemically coherent model driven by emissions and meteorological inputs. Although the physical and chemical constraints imposed in a deterministic model are limited by the accuracy with which these processes can be specified, application across multiple pollutants over space and time is not as limited as empirical approaches that can lead to large errors if extrapolated beyond a calibrated domain. Chen et al. (2014) used an inverse distance error weighting method to adjust 36-km resolution CMAQ biases within 100-km of monitors for several pollutants and compared results with other regional

estimates.[76] In this paper, we evaluate two alternative methods of fusing observational data and CMAQ simulations to provide daily concentration fields for a variety of pollutants, including primary pollutants, and then combine the results to optimize the prediction of temporal variation over space. In addition, we quantify uncertainty in our estimates using a rigorous cross-validation analysis. Error can lead to underestimating variability in the exposure estimates and can result in substantial bias in health outcome analysis. Results from this work are being used in Georgia-wide studies of ambient air pollution and preterm births and emergency department visits.[77]

3.2 Methods

3.2.1 Ambient Monitor Measurements

Data for 12 pollutants of interest from several ground-based monitoring networks from 2002 to 2008 were obtained for the state of Georgia and for two sites near the Georgia border in Florida and Tennessee. The pollutants are five gases (CO, NO_2, NO_x, O_3, and SO_2), two particulate matter (PM) mass measurements (PM_{10}, $PM_{2.5}$), and five $PM_{2.5}$ components: sulfate (SO_4^{2-}), nitrate (NO_3^-), ammonium (NH_4^+), elemental carbon (EC), and organic carbon (OC). Data were obtained from the following networks: U.S. EPA's Air Quality System (AQS), which include the State and Local Air Monitoring System and Speciation Trends Network for $PM_{2.5}$ component measurements; the Southeastern Aerosol Research and Characterization (SEARCH) network;[78] the Interagency Monitoring of Protected Visual Environments (IMPROVE) network;[79] and the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) network.[80] For PM

measurements, Tapered Element Oscillating Microbalance (TEOM) mass data were calibrated to Federal Reference Method (FRM) data, and EC and OC measurements by the thermal optical transmittance (TOT) method were converted to be equivalent to thermal optical reflectance (TOR) measures.[81-82] Collocated measurements at several sites over various time periods demonstrate a high degree of precision across instruments and analytical procedures, and data from these networks have been used to assess spatial variability in previous work.[37, 57]

Figure 3.1 shows the location of monitors. The monitoring networks provided best spatial coverage for O_3 and $PM_{2.5}$, and very limited spatial coverage for CO, NO_2, and NO_x with data available only in the metropolitan Atlanta area.

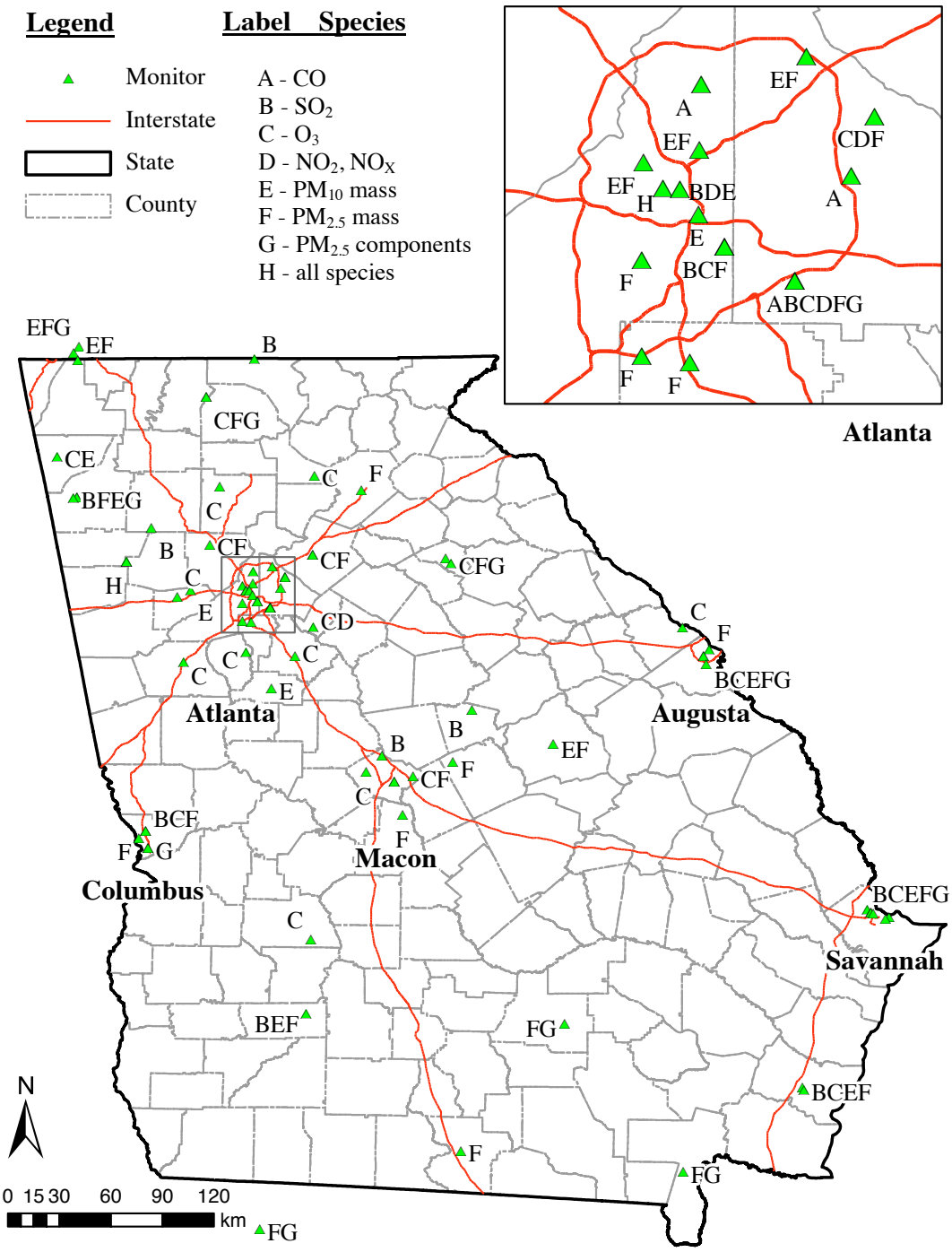

Figure 3.1 – Georgia study area and locations of ambient air quality monitors with observations during 2002-2008. SEARCH monitors at Jefferson St. in Atlanta and at Yorkville west of Atlanta are labeled H, indicating a comprehensive set of pollutant measurements are available at these sites.

On days with 16 or more hourly observations, missing hourly measurements of gases and PM mass were estimated by linear interpolation; on days with less than 16 hourly observations, data were treated as missing. Daily metrics, based on the National Ambient Air Quality Standards, were calculated for each pollutant based on those identified as primary exposure variables in the health analyses; these are one-hour maxima for CO, NO_2, NO_x, and SO_2, eight-hour daily maximum for O_3, and 24-hour averages for PM_{10} and $PM_{2.5}$ masses and $PM_{2.5}$ species. Table 3.1 provides the number of monitors, sampling frequency, summary statistics of all monitors used in this study, and the distribution of measurements at a central Atlanta monitoring site by pollutant.

Table 3.1 – Summary statistics of observed daily metrics (OBS) during the 2002-2008 time period for Georgia monitors used in this study.

Pollutant	No. of Monitors	Sampling Frequency	OBS	Comp. [a,c]	Mean [b,c]	IQR [b,c]
1-hr max NO_2, ppb	7	7 daily	16,045	90%	22.9	27
1-hr max NO_x, ppb	7	7 daily	16,031	90%	60	63.5
1-hr max CO, ppm	5	5 daily	9,582	75%	0.8	0.8
1-hr max SO_2, ppb	14	14 daily	24,018	67%	11	11.5
8-hr max O_3, ppb	27	2 daily Jan-Dec; 25 daily Mar-Oct	42,734	89%	47.6	22
PM_{10}, $\mu g/m^3$	20	2 daily; 1 1-in-3; 17 1-in-6	11,155	84%	23.4	14.8
$PM_{2.5}$, $\mu g/m^3$	42	15 daily; 20 1-in-3; 7 1-in-6	43,200	74%	14.8	9.4
$PM_{2.5}$-SO_4, $\mu g/m^3$	14	2 daily; 3 1-in-3; 9 1-in-6	9,306	81%	4.14	3.18
$PM_{2.5}$-NO_3, $\mu g/m^3$	14	2 daily; 3 1-in-3; 9 1-in-6	9,290	81%	0.7	0.56
$PM_{2.5}$-NH_4, $\mu g/m^3$	12	2 daily; 1 1-in-3; 9 1-in-6	7,716	79%	1.4	1.04
$PM_{2.5}$-EC, $\mu g/m^3$	14	2 daily; 3 1-in-3; 9 1-in-6	9,043	79%	0.89	0.7
$PM_{2.5}$-OC, $\mu g/m^3$	14	2 daily; 3 1-in-3; 9 1-in-6	9,017	78%	3.16	2.21

[a] Completeness indicates the fraction of available versus expected daily metrics for all monitors and days in the study.
[b] Mean and interquartile range (IQR) for the Atlanta central monitor at Jefferson St.
[c] Inclusive of days for which hourly linear interpolation was conducted to fill in missing hourly measurements (if <8 hrs were missing).

3.2.2 CMAQ Simulations

Hourly simulated concentrations for the pollutants of interest obtained from the PHASE project were developed using the CMAQ model at 12-km resolution for the 2002-2008 time period over the eastern U.S. CMAQ is a deterministic chemical transport model that provides spatially and temporally resolved simulated outputs. Daily metrics of interest for the 12 pollutants were calculated from the PHASE dataset. The U.S. EPA conducted a comprehensive model evaluation of CMAQ version 4.5 using several different key synoptic weather patterns. In that evaluation, ozone simulation results were consistent with previous evaluations, where the model captured daily trends but had difficulty capturing the upper and lower limits of the observations. CMAQ had a tendency to over-predict fine particulate matter mass during the winter and fall. EPA's evaluation of CMAQ model version 4.5 also showed particulate sulfate simulations improved when compared to results of the previous version due to changes in dry deposition velocity and cloud treatment. Particulate nitrate and ammonium were highly over-predicted by the CMAQ model version 4.5 during the fall. The evaluation also found unresolved secondary organic aerosol pathways in the model contributed to under-prediction of carbonaceous aerosols during the late spring and summer months.

3.2.3 *Data Fusion*

Ambient monitor observations provide very limited spatial information and information on temporal variation that decreases with increasing distance from monitors. CMAQ simulations, on the other hand, provide information that is independent of observations. Our approach to fusing observations and CMAQ simulations involves three steps (Figure 3.2). First, we develop fused concentration fields using daily

interpolated normalized observations to estimate temporal variation and the annual mean CMAQ field to estimate spatial variation (FC_1). Strengths of this method are that observations drive temporal variation and CMAQ simulations provide mean spatial variation, which are what these data do best. Weaknesses of this method are that temporal variation is predicted poorly far from observations and daily spatial variation is not captured well by a spatially sparse monitoring network. Second, we develop fused concentration fields using adjusted daily CMAQ simulations, correcting for annual and seasonal biases (FC_2). This method is less subject to the weaknesses of the first method in that prediction of temporal variation is largely independent of monitor location and daily spatial variation is predicted based on daily meteorology and emissions. However, inaccuracies in the CMAQ simulations limit the effectiveness of this method. Third, we combine these to produce optimized fused concentration fields (FC_{opt}) by computing a weighted average with the weight depending on the spatial autocorrelation of observations (which governs how well FC_1 predicts temporal variance) and the correlation between observations and CMAQ simulations (which governs how well FC_2 predicts temporal variance). By optimizing based on the prediction of temporal variation over space, these data fused estimates may be particularly useful in large time-series epidemiologic studies. Due to the different spatial scales of observations (point measurements) and CMAQ simulations (12-km resolution) and the limited number of observations available each day, prediction of daily spatial variance was not included in the optimization; corrections for annual bias and annual spatial variance in the CMAQ

simulations are included earlier in the data fusion approach. A detailed description of each step follows.

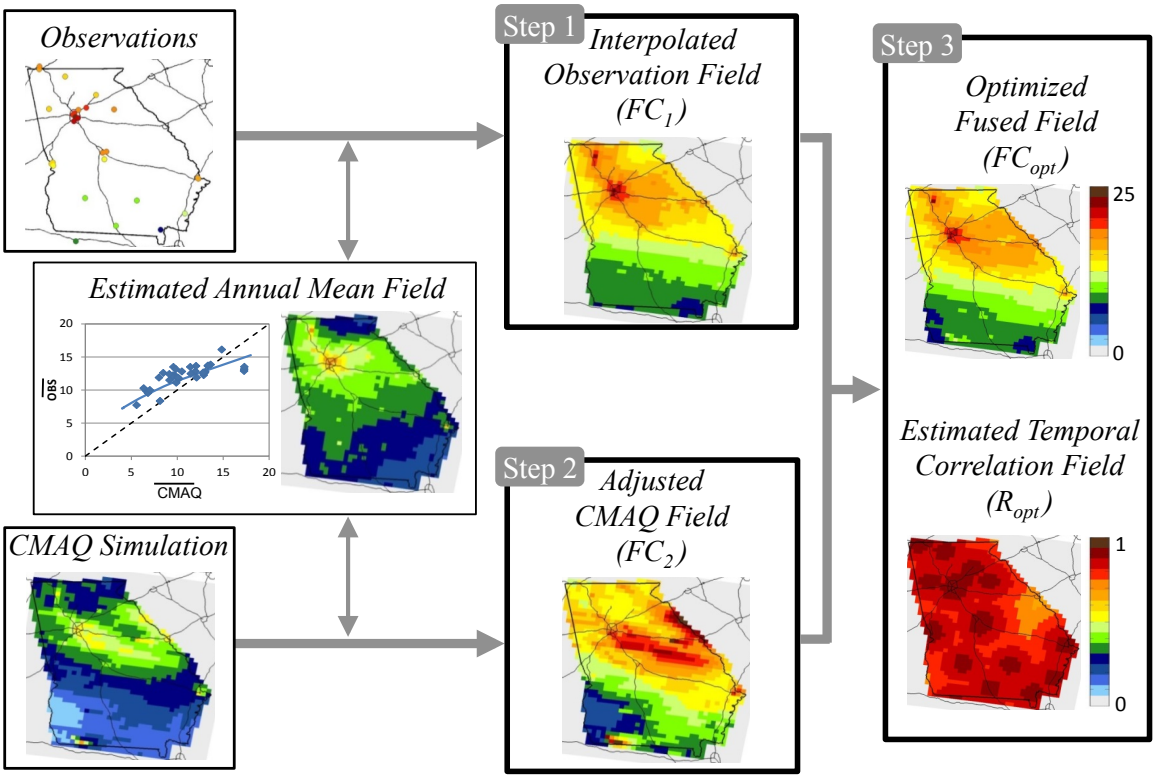

Figure 3.2 – Data fusion flow diagram including depiction of 24-hr $PM_{2.5}$ (µg/m3) fields for 7/23/08 at 12-km resolution.

3.2.3.1 Interpolated Observation Method (FC_1)

To obtain pollutant estimates over space (s) and time (t) with temporal variation driven by monitor data, we first normalize daily observations (OBS_m) at each monitor (m) to annual mean levels ($\overline{OBS_m}$). Next, normalized data are spatially interpolated by the *krig* method described below. Lastly, the interpolated field is denormalized using the CMAQ annual field adjusted to the annual mean observations ($\overline{FC}$), also described below.

Equation 3.1 describes this procedure for estimating daily concentration fields with temporal variation driven by observations and spatial structure governed by the adjusted annual mean CMAQ field.

$$FC_1(s,t) = \left(\frac{OBS_m(t)}{\overline{OBS_m}}\right)_{krig} \times \overline{FC(s)} \tag{3.1}$$

Daily spatial interpolation of the normalized observations was performed by ordinary kriging. Kriging is a geostatistical interpolation method based on the spatial autocorrelation of the data points to estimate trends. There are various types of kriging methods, each of which assumes a combination of known or unknown, constant or varying spatial mean function. In ordinary kriging, the observed data are assumed to have an unknown, but constant mean function (a zeroth-order polynomial trend that characterizes the observed data as having a constant spatial dependence), yielding a continuous surface of estimated points within a defined domain.[83] Similar to inverse distance weighting, ordinary kriging calculates the distance between the unknown point to be estimated and the observed data point. Ordinary kriging uses the value of the covariogram, a measure of deviation between pairs of observed data points as a function of distance, to calculate the weight of each of the surrounding measurements by minimizing standard error. MATLAB and mGstat, a geostatistical toolbox for MATLAB, were used to fit exponential semivariograms to the data by minimizing the error of the fitted variogram function. Creating a continuous surface of estimated points using ordinary kriging can lead to invalid spatial features. Such artifacts are a consequence of

the statistical smoothing technique used that does not account for the chemical and physical nature of the data. Normalization prior to kriging provides a smoother surface necessary with limited monitor coverage of the spatial domain. A more detailed spatial structure is obtained in denormalization using the mean CMAQ field.

Regression of annual mean measurements ($\overline{OBS_m}$) and CMAQ simulations at monitor locations ($\overline{CMAQ_m}$) provides parameters for an annual mean pollutant field model that captures the spatial pattern of emissions and annual effects of meteorological variables, correcting for CMAQ annual biases (eq. 3.2).

$$\overline{FC(s)} = \alpha_{year} \times \overline{CMAQ(s)}^{\beta} \tag{3.2}$$

Here, the overbar indicates temporal averaging (annual), β is a parameter derived for all years, and α_{year} is a regression parameter derived for each year (Table A.1). Spatial misalignment of measurements (point) and CMAQ simulations (12-km x 12-km) and a changing network of monitors over time are factors that contribute to model instability and led to the use of more constrained models (e.g. non-varying β) and zero intercepts. Linear models with slopes and intercepts varying each year provided similar results, although negative intercepts and greater variability across years were found. For pollutant gases and EC, linear models ($\beta = 1$) with zero intercepts were used; non-unity exponents were found to improve the mean spatial prediction of other PM pollutants (Figure 3.3). The annual mean regression models capture the spatial variance of CO, NO_2 and NO_x well, with a minimum goodness-of-fit R^2 of 0.80, 0.87, and 0.84, respectively. The

urban-to-rural concentration differences for these three primary pollutants are large. For SO_2 and PM_{10}, the simulated annual spatial models performed least well, with R^2 values of 0.2 and 0.1.

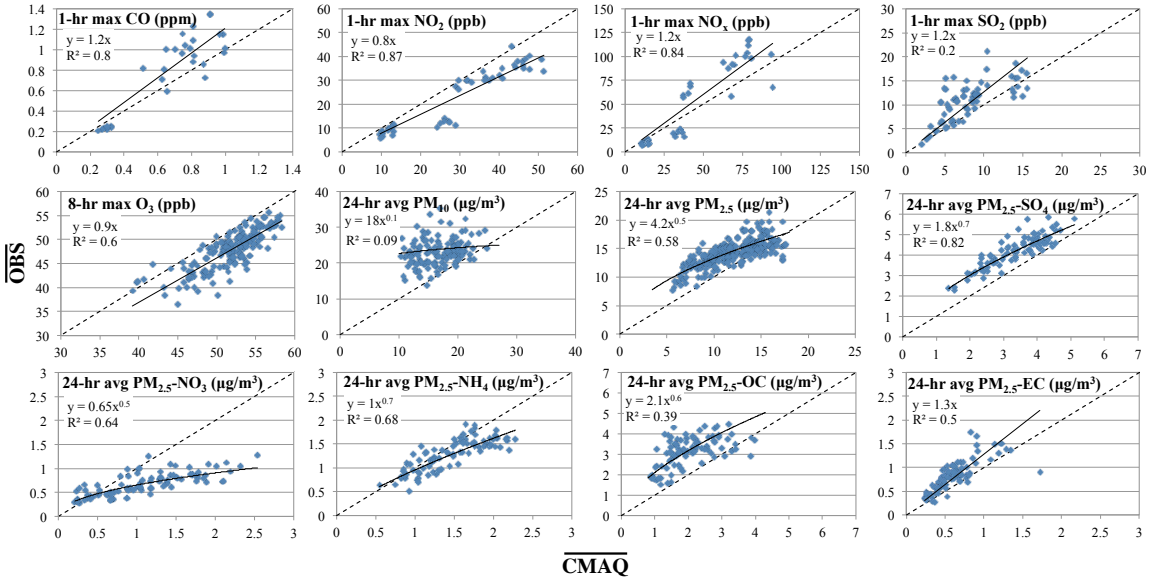

Figure 3.3 – Annual data used to estimate annual mean pollutant field parameters, 2002-2008. For each pollutant, an overall fitted trend for the seven-year time period is shown (solid curve); a 1:1 dashed line is shown for reference. Values of β and average values of α (eq. 3.2) are shown; for values of α_{year}, see Table A.1.

3.2.3.2 Scaled CMAQ Method (FC_2)

An alternative data fusion approach is to use the daily CMAQ fields and scale them to observations using the adjusted annual mean CMAQ fields corrected for seasonal bias (eq. 3.3).

$$FC_2(s,t) = CMAQ(s,t) \times \left(\frac{\overline{FC(s)}}{\overline{CMAQ(s)}} \right) \times \beta_{season}(t) \qquad (3.3)$$

Here, β_{season} is the seasonal correction function. Analysis of the differential pattern between the observations and CMAQ simulations across species showed the seasonal component could be described and minimized using a sinusoidal cycle. The seasonal bias, which follows a sinusoidal variation, was modelled as a smooth trigonometric function (eq. 3.4) with two fitted parameters: amplitude (A) and day of peak correction (t_{max}) (Table A.2).

$$\beta_{season}(t) = e^{A \times \cos\left[\frac{2\pi}{365.25}(t - t_{max})\right]} \tag{3.4}$$

Use of this correction factor removed seasonal trends in the residual errors of the FC_2; this is shown graphically in supplemental material (Figure A.1).

3.2.3.3 Optimized Fused Fields (FC_{opt})

Finally, to obtain the best predictions over space and time, FC_1 and FC_2 were averaged using a weighting function based on prediction of temporal variance. For the interpolated observation data fusion method, we use an exponential correlogram (eq. 3.5) with fitted parameters to model spatial autocorrelation isotropically. Correlogram plots for all pollutants are shown in the supporting material (Figure A.2).

$$R_{obs}(d) = R_{coll} e^{-d/r} \tag{3.5}$$

Here, R_{obs} is the temporal Pearson correlation between observations from monitors a distance d apart, R_{coll} is the intercept which results from instrument error (i.e., error as estimated by collocated instruments), and r is the range at which the correlation between

monitors has decreased to an *e-folding* of R_{coll} (Table A.3). We then estimate the prediction of temporal variance by the interpolated observation data fusion method by 3. 6.

$$R_1(s,t) \approx R_{coll} e^{-x(s,t)/r} \tag{3.6}$$

Here, R_1 is the estimated temporal correlation of FC_1 and ambient pollution and x is the distance to the nearest observation. R_1 varies over time and space because the x varies with grid cell location as well as with day since monitor observations may be missing on some days. The estimated prediction of temporal variance by the scaled CMAQ method, on the other hand, does not depend on distance to the observations since daily observations are not used. The average of the temporal correlations of CMAQ and *OBS* across all monitors (R_{cmaq}) is used to estimate R_2, which represents the estimated temporal correlation of FC_2 and ambient pollution (eq. 3.7).

$$R_{cmaq} = \frac{1}{N}\sum_{m=1}^{N} corr(OBS_m(t), CMAQ_m(t)) \approx R_2 \tag{3.7}$$

Here, N is the total number of monitors. The R_{cmaq} values derived in this study are provided in supporting material (Table A.3 and Figure A.3). We note that R_1 and R_2 are conservative estimates of temporal variance in that FC_1 uses information from multiple measurements and FC_2 incorporates annual and seasonal adjustments to CMAQ.

A weighting factor, W, defined by eq. 3.8 was used to estimate the degree to which the observation-based estimate (FC_1) predicts temporal variation relative to the CMAQ-based estimate (FC_2).

$$W(s,t) = \frac{R_1(s,t)\times(1-R_2)}{R_1(s,t)\times(1-R_2)+R_2\times(1-R_1(s,t))} \tag{3.8}$$

The weighting factor has a value of 1 when R_1 is 1, a value of 0.5 when $R_1 = R_2$, and a value of 0 when $R_1 = 0$. The factor varies over time and space because the distance to nearest observation varies over time (due to missing data and variation in measurement frequency) and space. The optimized fused field (FC_{opt}) is the weighted average of FC_1 and FC_2, given by eq. 3.9.

$$FC_{opt}(s,t) = W(s,t)\times FC_1(s,t) + \big(1-W(s,t)\big)\times FC_2(s,t) \tag{3.9}$$

Figure A.4 of the supporting material shows average W spatial fields for the study period. The optimized field resembles the FC_1 field near observations (where the weighting is large) and the FC_2 field far from observations (where the weighting is small). The temporal correlation of FC_{opt} and OBS, R_{opt}, is estimated by a similar weighting (eq. 3.10).

$$\begin{gathered}\text{if } R1 > R2, R_{opt}(s,t) = W(s,t)\times R_1(s,t) + (1-W(s,t))\times R_2; \\ else, R_{opt} = R_2\end{gathered} \tag{3.10}$$

Cross-validation shows this weighted average of R_1 and R_2 is a reasonable and conservative approximation of the temporal correlation of FC_{opt} with withheld *OBS* (Figure A.5), as will be discussed later. The lower threshold of R_2 avoids a minimum in R_{opt} as a function of distance to nearest observation, a minimum not supported in the cross-validation results.

3.2.4 Model Performance Characterization and Evaluation Statistics

Model performance was characterized and model evaluated for CMAQ predictions and our three fused concentration field outputs (FC_1, FC_2, FC_{opt}) using the following statistics.

$$percent\ mean\ bias = \frac{1}{N}\sum_{m=1}^{N}\left(\frac{\overline{modeled_m} - \overline{OBS_m}}{\overline{OBS_m}}\right) x100\% \tag{3.11}$$

$$percent\ RMSE = \frac{1}{N}\sum_{m=1}^{N}\left(\frac{\sqrt{\overline{(modeled_m - OBS_m)^2}}}{\overline{OBS_m}}\right) x100\% \tag{3.12}$$

$$RSQ = \big(corr(OBS, modeled)\big)^2 \tag{3.13}$$

$$monitor\ temporal\ R = corr(OBS_m(t), modeled_m(t)) \tag{3.14}$$

$$daily\ spatial\ R = corr(OBS_t, modeled_t) \tag{3.15}$$

$$mean\ spatial\ R\ = corr\big(\overline{OBS}, \overline{modeled}\big) \tag{3.16}$$

Here, the subscript *m* denotes the monitor, the subscript *t* denotes the day, and the overbar indicates the annual average. The mean bias (eq. 3.11) and root mean square error (*RMSE*) (eq. 3.12) statistics were normalized to the mean of the withheld observations at each monitor, and then averaged across monitors, to facilitate comparison across pollutants. To assess model prediction of variation over space and time, various Pearson correlation coefficients were calculated. *RSQ* represents a correlation over both space (monitor locations) and time (eq. 3.13). To evaluate model performance in predicting spatial and temporal variation separately, correlations of subsets of measured and modeled data were evaluated. Temporal *R* values at each monitor (eq. 3.14) were described by a mean and standard deviation across monitors. Spatial *R* values on each day with four or more monitors (eq. 3.15) were described by a mean and standard deviation across days. The spatial *R* of annual mean concentrations (eq. 3.16) was also calculated.

3.3 Results

Three sets of results are presented. First, CMAQ simulations are evaluated for the pollutant metrics and temporal and spatial domains of this study. Second, a comparison of monitor data and data fusion modeling outputs (FC_1, FC_2, FC_{opt}) that use the monitor data is presented to describe model performance characteristics. Third, results of a comprehensive cross-validation analysis are presented to evaluate data fusion model predictive capabilities both near to and far from observations.

3.3.1 CMAQ Simulation Performance

The CMAQ model evaluation conducted by the U.S. EPA was extended to the pollutants, metrics, and spatial and temporal domains of this study. Mean bias and Pearson correlation coefficients representing temporal variation, spatial variation of mean concentrations, and daily spatial variation provide a comparison of the 12-km CMAQ simulations and monitor data over the 2002-2008 study period (Figure 3.4); other evaluation metrics are shown in supporting material Table A.4.

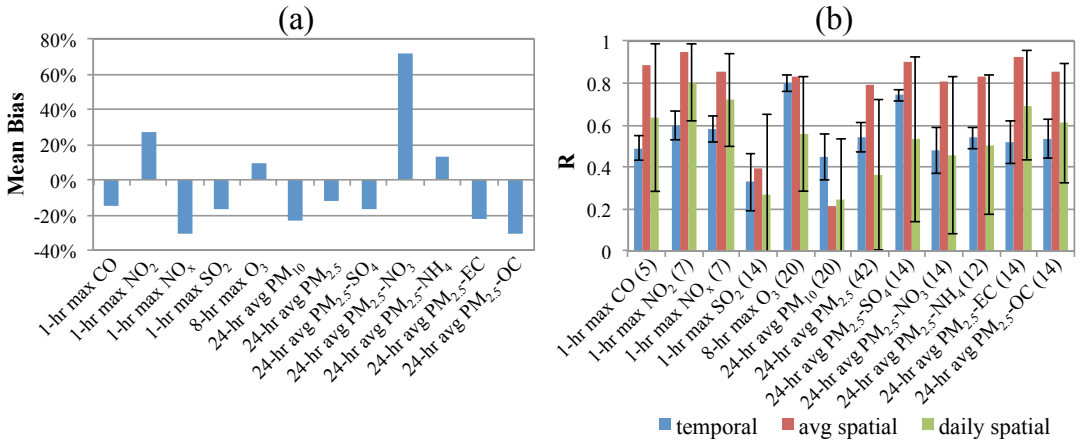

Figure 3.4 – Comparison of observations and CMAQ simulations at monitor locations for Georgia from 2002 to 2008. In the left panel (a), percent mean bias is shown. In the right panel (b), Pearson correlation coefficients (R) are shown for temporal variation, average spatial variation, and daily spatial variation when four or more observations are available. Average values and error bars representing two standard deviations are shown for temporal *R* values across monitors and for daily spatial *R* values across days. The numbers of monitors are indicated in parentheses.

CMAQ performance varied widely across the 12 pollutants studied. As expected, there was a large positive bias in the 24-hr nitrate simulations;[42] other mean biases were less than 30 percent. With the exception of PM_{10}, CMAQ simulated the spatial variation of mean concentrations better than daily spatial and temporal variations. CMAQ least

accurately simulated the spatial and temporal variations in 1-hr maximum SO_2 and 24-hr average PM_{10}, demonstrating known limitations in the CMAQ model to predict ground-level impacts of SO_2 plumes and biogenic PM_{10} source contributions.[84-85] Our spatial R^2 between annual mean $PM_{2.5}$ CMAQ simulations and measurements of 0.63 (R value shown in Figure 3.4b) is lower than the R^2 of 0.88 from a regionalized statistical approach developed for one year of annually averaged $PM_{2.5}$ data over the contiguous U.S. using an extensive database of geographic covariates.[86] This result is not surprising given the smaller geographic scale and absence of geographic covariates in our approach, which is more readily generalizable to other pollutants.

3.3.2 Data Fusion Model Performance Characterization

Before presenting model cross-validation results (shown subsequently in Figure 3.8), results are shown to characterize model performance. Outputs from the three data fusion models (FC_1, FC_2, and FC_{opt}) are compared with the monitor data in terms of the *RMSE* and the spatiotemporal, temporal, and spatial mean correlations in Figure 3.5; CMAQ simulation results are shown for comparison.

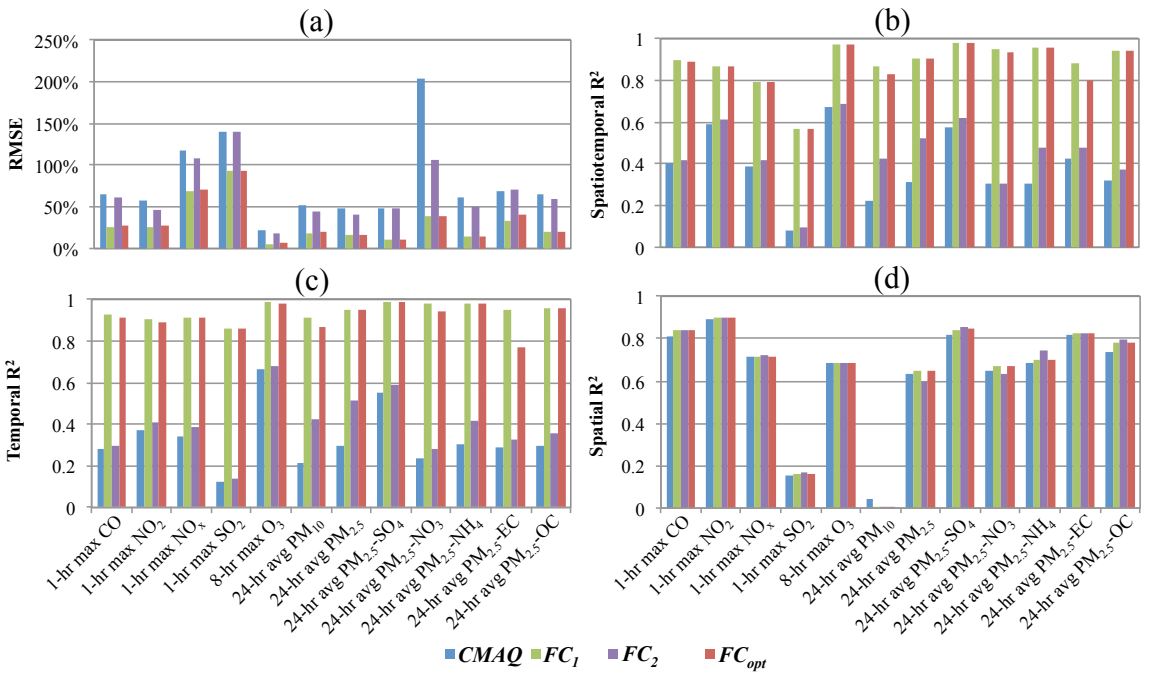

Figure 3.5 – Comparison between observations and simulations (CMAQ and three data fusion methods) for Georgia from 2002 to 2008 in terms of *RMSE* (panel a), spatiotemporal R^2 values (panel b), average temporal R^2 values (panel c), and spatial R^2 values of mean concentrations (panel d).

RMSE values are lower and correlations higher for the FC_1 and FC_{opt} methods, as expected, since these methods use the daily observations. Non-zero *RMSE* and correlation values less than one in these results demonstrate the impact of using the annual mean model, $\overline{FC}$, to provide spatial structure. The FC_2 method performs similarly to CMAQ, as expected, since the FC_2 method is a rescaling of the CMAQ simulations. Biggest improvements in spatiotemporal R^2 (Figure 3.5b) were with respect to PM_{10}, $PM_{2.5}$ and NH_4; for these, the improvements were largely due to seasonal corrections (Table A.4). The prediction of spatial variation in the annual mean fields is limited by

the ability of CMAQ to predict this variation; these limitations are greatest in the cases of 1-hr maximum SO_2 and 24-hr average PM_{10} (Figure 3.5d).

Average spatial FC_{opt} distributions of each of the 12 pollutant concentrations over the entire domain and time period are shown in Figure 3.6. Primary pollutants (NO_2, NO_x, SO_2, CO and EC) are spatially more heterogeneous than pollutants of secondary or mixed origin (PM_{10}, $PM_{2.5}$, SO_4, NO_3, NH_4 and OC), with greater urban-to-rural gradients. The impacts of major urban centers are evident in these 12-km resolved fields as indicated by the location of concentration peaks. On-road emission impacts are evident for the major interstates in the NO_2, NO_x, CO and EC fields, while impacts of coal combustion point sources are evident in the SO_2 field.

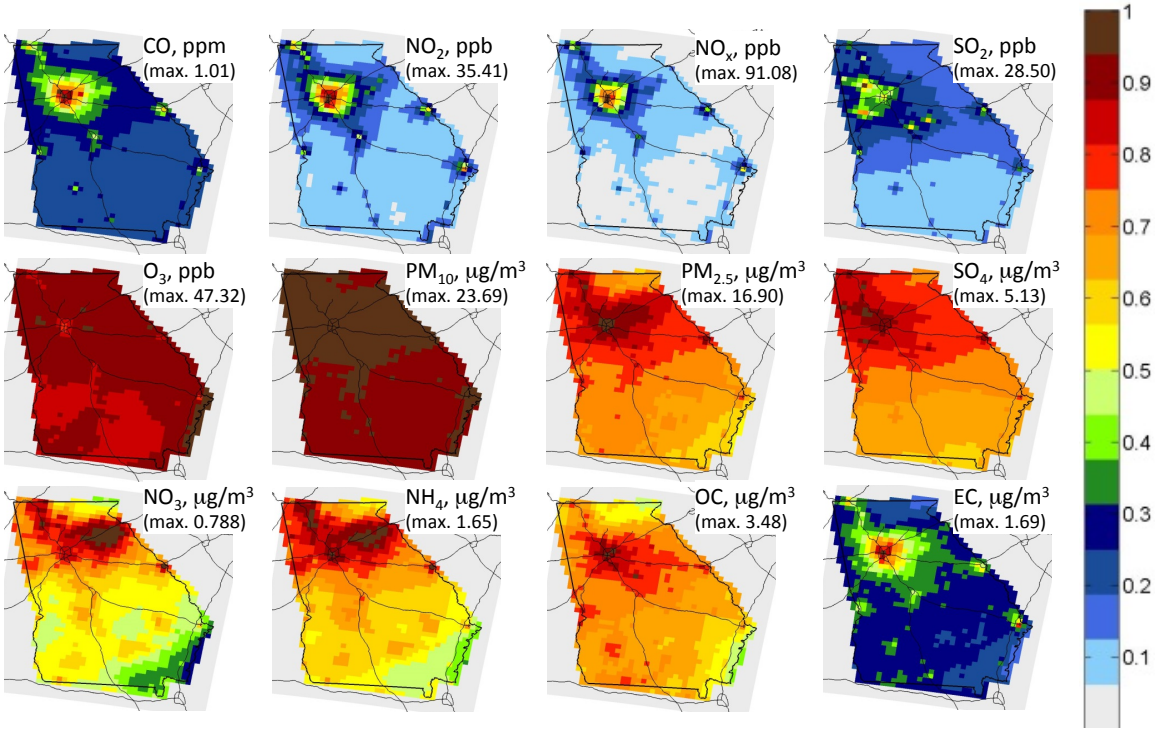

Figure 3.6 – Temporally averaged optimized concentration fields (FC_{opt}) normalized by their respective maximum, 2002-2008.

Estimated temporal correlation (R_{opt}) fields based on eq. 3.9 are shown in Figure 3.7. Two fields are shown for each of the pollutants that are not monitored with the same frequency across all sites; these pollutants are O_3 (all monitors were operational between March and November and only a few monitors were in December, January and February) and PM species (daily and 1-in-3 or 1-in-6 day monitors). The R_{opt} value is highest near a monitor and decreases as the distance to the nearest monitor increases. For pollutants with observations only in metro-Atlanta, the estimated temporal correlation for regions far from metro-Atlanta is that of the CMAQ simulations.

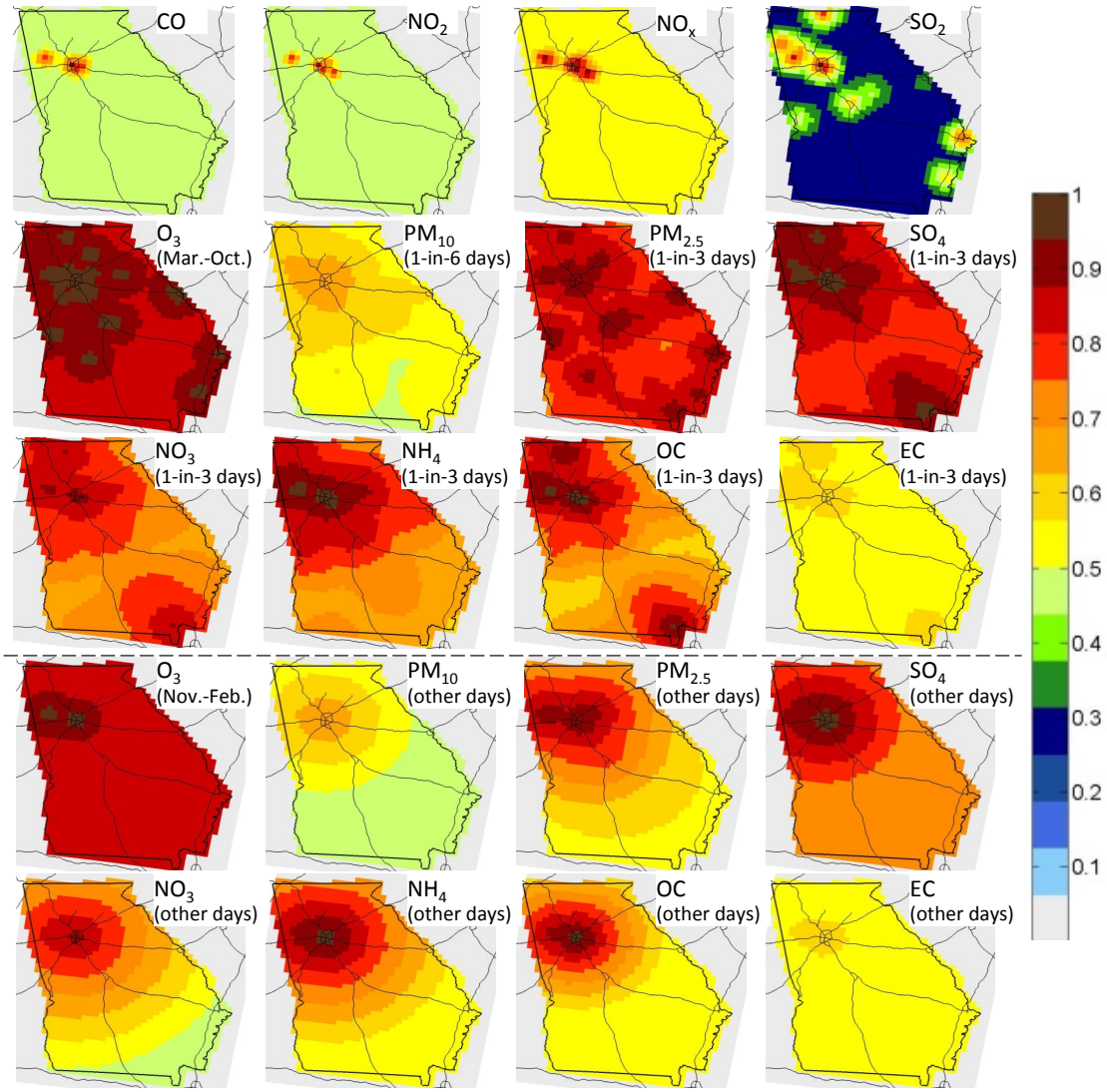

Figure 3.7 – Temporally averaged correlation fields (R_{opt}), 2002-2008.

3.3.3 Data Fusion Model Evaluation by Cross-Validation

Data fusion model performance was evaluated using a comprehensive tenfold 10% withholding cross-validation analysis. For each of 10 trials, 10% of the observations were randomly held back and each data fusion method was applied to simulate the withheld data. The results from the 10 trials were then combined to provide

robust cross-validation results that allow for the exploration of difference in errors based on proximity to monitors. Across monitors and days, the number of withheld data correspond to the number of observations for each pollutant (Table 3.1), ranging from 7,720 for $PM_{2.5}$-NH4 to 43,200 for $PM_{2.5}$.

The normalized mean bias and *RMSE,* and Pearson R^2 are shown for the three sets of data fusion predictions in Figure 3.8 for all monitors combined and for each of the SEARCH monitors (Jefferson St. in urban Atlanta and Yorkville in a rural area 60-km west of Atlanta). Analysis of the Jefferson St and Yorkville results separately allows for assessment of the effects of spatial clustering of monitors. All pollutants studied here were measured daily at these two sites, and other measurements of each pollutant were available nearer the Jefferson St. site than the Yorkville site.

The cross-validation results across all monitors indicate that average mean bias is less than 30 percent across pollutants, that 8-hr maximum ozone is predicted best (lowest *RMSE* and highest R^2), and that 1-hr maximum SO_2 is predicted worst (highest *RMSE* and lowest R^2). The optimized results from combining FC_1 and FC_2 tended to perform best in terms of lower RMSE and higher R^2, as expected. The temporal variation in pollutant concentrations were predicted better at Jefferson St. than at the Yorkville site, demonstrating the effect of having nearby measurements of a pollutant.

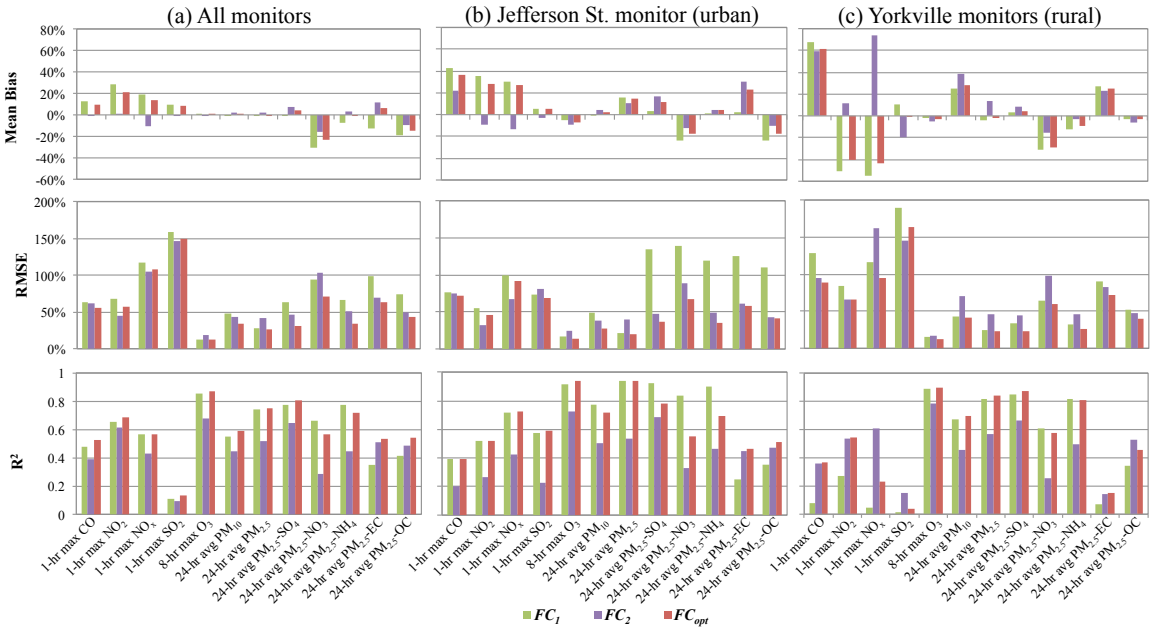

Figure 3.8 – Comparison of spatiotemporal Pearson R^2, RMSE, and mean fractional error values between observations and all simulated at (a) all withheld monitor locations, (b) withheld Jefferson St. monitor location, and (c) withheld Yorkville monitor location for 2002-2008.

3.4 Discussion

Our optimized data fusion approach combines two methods that have different strengths and limitations. The FC_1 fields have a structure similar to the annual mean CMAQ field and a temporal variation driven by that of the observations. These estimates agree very well with observations. The degree to which error is not zero and the correlations are not one at monitor locations is due to two factors. First, the annual mean model does not perfectly match observations. The CMAQ simulations represent 12-km x 12-km averages whereas the observations are point estimates. Hence, the annual mean field is smooth without capturing sub-grid impacts of sources. Second, the observations are assumed to have some error, based on the intercept of the correlogram, which

includes collocated instrument measurement data. Therefore, the weighting factor used for calculating FC_{opt} is not one. Error associated with the FC_1 estimates increases with increasing distance from observations. The FC_2 fields, on the other hand, have a spatial structure and a temporal variation similar to the daily CMAQ simulations. Error associated with these estimates is largely independent of distance from observations. The optimized prediction (FC_{opt}) is an average of the FC_1 and FC_2 predictions based on their estimated prediction of temporal variation (Eqs. 3.5 and 3.6, respectively).

The temporal correlation of optimized predictions and observations estimated by eq. 3.10 (R_{opt}) varies over time and space depending on the observations used in the estimates (average values are depicted in Figure 3.7). These R_{opt} estimates conservatively approximate those assessed by the cross-validation evaluation of our optimized data fusion method (supporting material Figure A.5).

The cross-validation results (Figure 3.8) suggest that optimized data fusion results overall have lower percent mean bias and percent *RMSE,* and higher Pearson R^2 values than the FC_1 and FC_2 results. The results also suggest that the clustering of monitors can lead to an inflated estimate of model performance, as can be seen by comparing results at Jefferson St. (Figure 3.8b) and Yorkville (Figure 3.8c) for CO, NO_2, and NO_x. With Jefferson St. data withheld, the Pearson R^2 values of the FC_1 estimates are higher than those of the FC_2 estimates due to nearby observations being available. With data from Yorkville withheld, however, the Pearson R^2 values of the FC_1 estimates are very low due to the lack of nearby observations. For the primary pollutants CO, NO_2, and NO_x,

therefore, the FC_2 estimates are better than the FC_1 estimates for most of the study area and weighted more in calculating the FC_{opt} estimates.

The major limitations in our data fusion approach are the spatial coverage of ambient measurements and the errors associated with CMAQ simulations. For example, CO and NO_x ambient monitors are located only in the metro-Atlanta region in Georgia. The cross-validation results indicate that the day-to-day variation of levels of pollutants of secondary origin (e.g., O_3, NO_3 and SO_4) can be better predicted than that of primary pollutants (e.g., NO_x, CO, SO_2 and EC) due to the spatial homogeneity of the secondary pollutants and high spatial autocorrelation. Limitations in CMAQ simulation of biogenic sources are consistent with the substantial improvements in spatiotemporal R^2 when PM_{10}, $PM_{2.5}$, and NH_4 CMAQ simulations are seasonally corrected (Table A.4). Variation in daily one-hour maximum SO_2 concentrations is predicted very poorly due to the inability of the monitor network and CMAQ simulations to capture the day-to-day impacts of coal combustion plumes at ground level. An implication for population-based acute health effect studies is that measurement error associated with spatial misalignment of SO_2 exposure is very high even when ambient monitor data and deterministic chemical transport model simulations are fused.

Another limitation in our data fusion approach is an over-prediction of spatial autocorrelation. In supporting material Figure A.6, spatial autocorrelation in observations, CMAQ simulations, and our data fusion results are compared. Results suggest the data fusion results have more spatial autocorrelation than observations or CMAQ simulations. This over-prediction of spatial autocorrelation is largely due to the

use of kriging in generating the FC_1 field estimates, and is more pronounced for the more spatially homogeneous pollutants such as PM_{10}, SO_4, and NH_4.

We use our cross-validation results to assess bivariate correlations as well, comparing the FC_{opt} bivariate correlations with those of observations and those of CMAQ simulations at the urban and rural SEARCH locations (Figure A.7). Overall, the CMAQ simulations tend to have higher bivariate correlations than observations, particularly in the cases in which one source category emits both pollutants. Data fusion tends to reduce the over-prediction of correlation between many pollutants inherent in CMAQ predictions.

The data fusion method presented here for blending monitor data and CMAQ simulations to develop daily spatial concentration fields of multiple pollutants is being used in ongoing epidemiologic investigations of acute health effects of ambient air pollution in Georgia and in four metropolitan areas.[35] Moreover, estimates of how well temporal variation is predicted, as a function of spatial location, allow for a measure of uncertainty to be used in the health analyses. The strengths of this data fusion approach are twofold: (1) ambient air exposure estimates are consistent with observations, emissions, and meteorology; (2) estimates are optimized based on simulation of temporal and spatial variation. Limitations of the approach include the number and location of ambient monitors, and the spatial resolution and accuracy of CTM simulations.

In this paper, we have described a new method to produce spatiotemporal air quality fields that fuse CMAQ model simulations with ambient monitoring data,

correcting model biases and optimizing prediction of temporal variation over space. We demonstrate that the method performs well estimating daily concentration fields of several pollutant gases and PM species, with the exception being SO_2, which is poorly characterized due to the limited ability of monitors and models to capture ground-level plume impacts. The approach is generalizable across pollutants with monitoring networks of limited and varied spatial coverage and sampling frequencies. Overall, the method is well suited for studies of the acute health effects of air pollution, providing spatiotemporal exposure estimates and uncertainties.

CHAPTER 4. DAILY AMBIENT AIR POLLUTION METRICS FOR FIVE CITIES: EVALUATION OF DATA-FUSION-BASED ESTIMATES AND UNCERTAINTIES

(Friberg, M.D., Kahn, R.A., Holmes, H.A., Chang, H.H., Sarnat, S.E., Tolbert, P.E., Russell, A.G. and Mulholland, J.A. *Atmospheric Environment, 158:* 36-50, 2017)

Abstract

Spatiotemporal characterization of ambient air pollutant concentrations is increasingly relying on the combination of observations and air quality models to provide well-constrained, spatially and temporally complete pollutant concentration fields. Air quality models, in particular, are attractive, as they characterize the emissions, meteorological, and physiochemical process linkages explicitly while providing continuous spatial structure. However, such modeling is computationally intensive and has biases. The limitations of spatially sparse and temporally incomplete observations can be overcome by blending the data with estimates from a physically and chemically coherent model, driven by emissions and meteorological inputs. We recently developed a data fusion method that blends ambient ground observations and chemical-transport-modeled (CTM) data to estimate daily, spatially resolved pollutant concentrations and associated correlations. In this study, we assess the ability of the data fusion method to produce daily metrics (i.e., 1-hr max, 8-hr max, and 24-hr average) of ambient air pollution that capture spatiotemporal air pollution trends for 12 pollutants (CO, NO_2,

NO_x, O_3, SO_2, PM_{10}, $PM_{2.5}$, and five $PM_{2.5}$ components) across five metropolitan areas (Atlanta, Birmingham, Dallas, Pittsburgh, and St. Louis), from 2002 to 2008.

Three sets of comparisons are performed: (1) the CTM concentrations are evaluated for each pollutant and metropolitan domain, (2) the data fusion concentrations are compared with the monitor data, (3) a comprehensive cross-validation analysis against observed data evaluates the quality of the data fusion model simulations across multiple metropolitan domains. The resulting daily spatial field estimates of air pollutant concentrations and uncertainties are not only consistent with observations, emissions, and meteorology, but substantially improve CTM-derived results for nearly all pollutants and all cities, with the exception of NO_2 for Birmingham. The greatest improvements occur for O_3 and $PM_{2.5}$. Squared spatiotemporal correlation coefficients range between simulations and observations determined using cross-validation across all cities for air pollutants of secondary and mixed origins are R^2 = 0.88-0.93 (O_3), 0.67-0.83 ($PM_{2.5}$), 0.81-0.89 (SO_4), 0.43-0.80 (NH_4), 0.52-0.72 (NO_3), 0.32-0.51 (OC), and 0.14-0.71 (PM_{10}).

Results for relatively homogeneous pollutants of secondary origin, tend to be better than those for more spatially heterogeneous (larger spatial gradients) pollutants of primary origin (NO_x, CO, SO_2 and EC). Generally, background concentrations and spatial concentration gradients reflect interurban airshed complexity and the effects of regional transport, whereas daily spatial pattern variability shows intra-urban consistency in the fused data. With sufficiently high CTM spatial resolution, traffic-related pollutants exhibit gradual concentration gradients that peak toward the urban centers. Ambient

pollutant concentration uncertainty estimates for the fused data are both more accurate and smaller than those for either the observations or the model simulations alone.

4.1 Introduction

Although associations between ambient air pollution and acute cardiorespiratory outcomes have been observed in numerous studies,[5-6, 9-11, 87-88] questions remain about the degree to which these findings can be generalized over multiple locations and whether the observed health effects are due to individual pollutants or to pollutants acting in combination. Accurately estimating air quality concentrations has direct implications for assessing the specific drivers of adverse, multi-pollutant health effects, as well as the synergistic effects of multiple risk factors. To estimate air quality concentrations and compare results across several metropolitan areas for use in epidemiological studies, a systematic and practical approach is needed for obtaining a quantitative understanding of the spatial and temporal correlation and variation of multiple ambient air pollutants.

Urban-level time-series epidemiological studies often span large geographic regions[36, 8936, 8136, 8935, 8834, 8733, 8632, 8531, 8431, 8430, 8329, 8228, 8127, 8026, 7925, 7824, 7723, 7622, 7521, 7420, 7319, 7317, 7315, 7112, 7010, 6810, 6810, 6810, 68 9, 69]. Assessing the health effects of ambient air pollution in these studies is frequently done using limited numbers of centrally sited monitors as surrogates for population exposure. Such networks might not adequately capture the spatiotemporal variability of pollutant concentrations and, therefore, might not accurately represent the outdoor air quality to

which the study subjects were exposed.[89] This may be especially true for pollutants of primary origin that vary considerably on short space and time scales.

Reliance on air pollution metrics developed from sparse monitoring data have limitations for application in health studies.[89-93] In general, the uncertainty of an air pollution metric depends on several factors, including instrument error, the spatiotemporal variability of the pollutant, and the method(s) used to develop such metrics. Error due to inadequately characterized spatial variability, which is often much larger than instrument error, can lead to substantial bias and reduced precision in health estimates.[94-95] Improving the characterization of pollutant concentration spatial variability can reduce the impact of the metric's uncertainty on epidemiologic results.[90, 95-97] In addition, for studies examining within-city heterogeneity of air pollution effects, such as effect modification by neighborhood socioeconomic status, spatially-resolved pollutant fields may reduce the potential for differential measurement error across neighborhoods compared to reliance on data from centrally-located monitors.[13, 29-30, 90, 96, 98-100]

High-resolution, spatiotemporally-resolved air pollution datasets can offer public health researchers representative concentration and uncertainty estimates that increase the accuracy of ambient pollution estimates used in retrospective epidemiological studies. By quantifying the heterogeneity of air pollutant concentrations and their uncertainties in multiple cities, the underlying common attributes and effects can be identified and characterized. We previously developed a data fusion approach that combines sparse observations with chemical transport model (CTM) results.[101] In this paper, we apply the

data fusion approach to a multicity time-series study to overcome some of the consequences of sparse air quality ground-based observations, measurement error (i.e., spatial misalignment), and other factors that affect exposure metric uncertainty and error. We obtain 12 km resolution, spatiotemporally-resolved air pollution datasets (i.e., concentration fields and their associated uncertainty) for several primary and secondary pollutants over five metropolitan areas, evaluate the results, and discuss the degree to which these findings are generalizable among locations.

4.2 Data

4.2.1 Study Design

Study areas encompass five metropolitan areas: Atlanta, GA; Birmingham, AL; Dallas, TX; Pittsburgh, PA; and St. Louis, MO-IL (Fig. 4.1). These non-coastal metropolitan areas were chosen because rich air quality datasets are available for each, and each offers distinct air quality characteristics that, taken together, represent many of the conditions found in larger U.S. metropolitan areas. Table 4.1 shows the population and land-area data, obtained from the 2010 US Census, for the high population density core study regions, which include counties with population densities greater than 50 km^{-2}. Core study areas include from 76% to 97% of the total population of each associated metropolitan statistical area (MSA) for each city, and from 36% to 87% of the total MSA land area (Table 4.1). Atlanta and Dallas have the largest study areas, and the Dallas study area has the highest population. Birmingham is smallest in population, and Birmingham and Pittsburgh have the smallest areas.

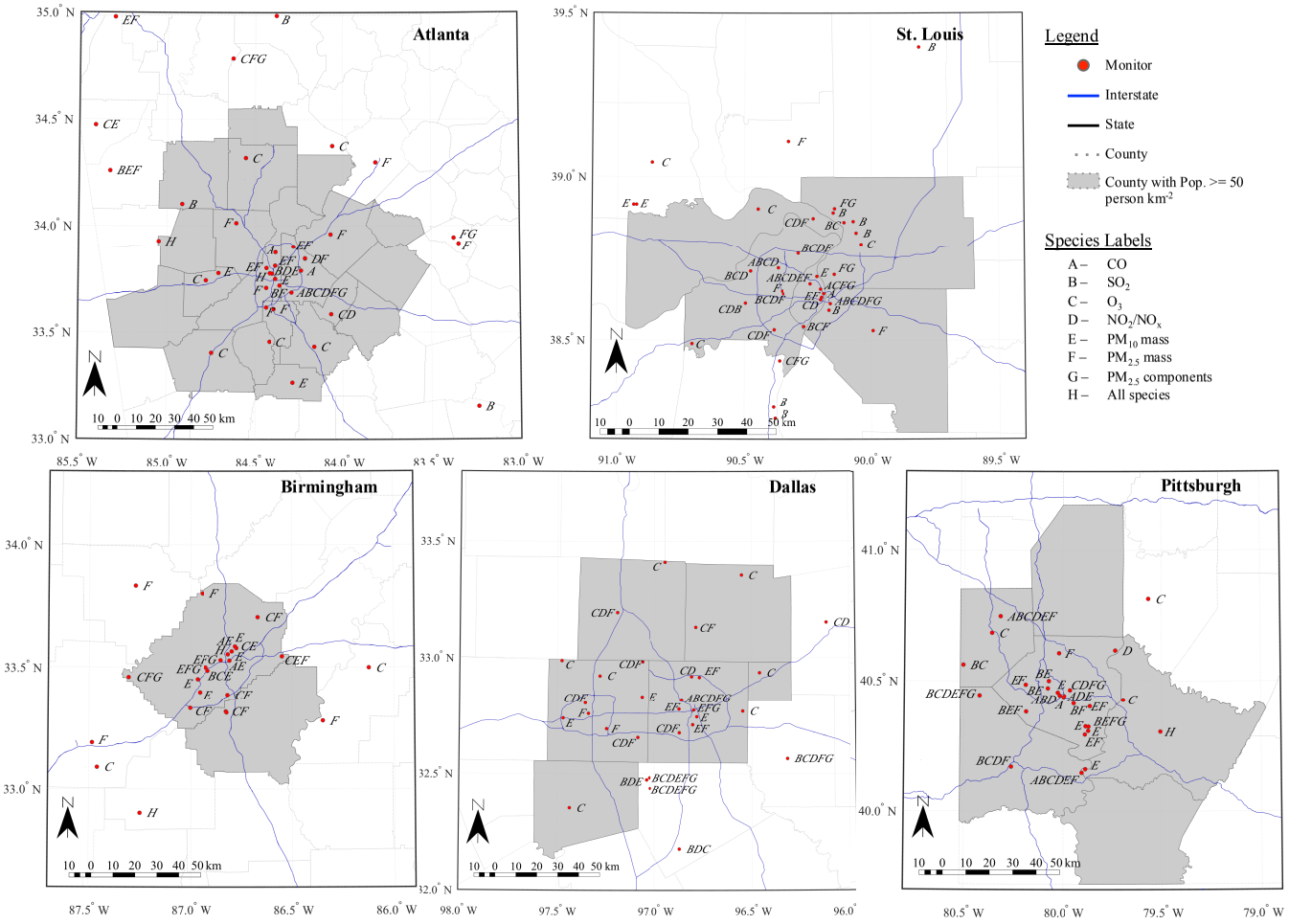

Figure 4.1 – Five city metropolitan statistical areas, counties, and monitor locations.

Table 4.1 – Population and size data for the five-cities study areas.

Cities	Counties No.	2010 Population	Area (km^2)
Atlanta, GA	28	5.6 M	26,873
Birmingham, AL	7	1.2 M	15,577
Dallas, TX	12	6.7 M	36,325
Pittsburgh, PA	7	2.45 M	14,607
St. Louis, MO-IL	15 + city	2.9 M	23,504

Average temperatures are similar across the cities, whereas seasonal mean precipitation and afternoon relative humidity show greater variation (Fig. 4.2).[102] Atlanta and Birmingham receive more precipitation than Dallas, Pittsburgh, and St. Louis. From 2002 through 2008, the average wind speed for Dallas was 9.5 m/s, compared to 8.7 m/s for St. Louis, 8.1 m/s for Atlanta, 7.3 m/s for Pittsburgh, and 6.1 m/s for Birmingham. Atlanta and Birmingham show similar mean temperature and relative humidity trends, but dissimilar mean wind speed and precipitation trends. Dallas has the highest temperatures and highest wind speeds, whereas Pittsburgh has the lowest temperatures and second-lowest wind speeds. St. Louis has the greatest summer-winter difference in temperature.

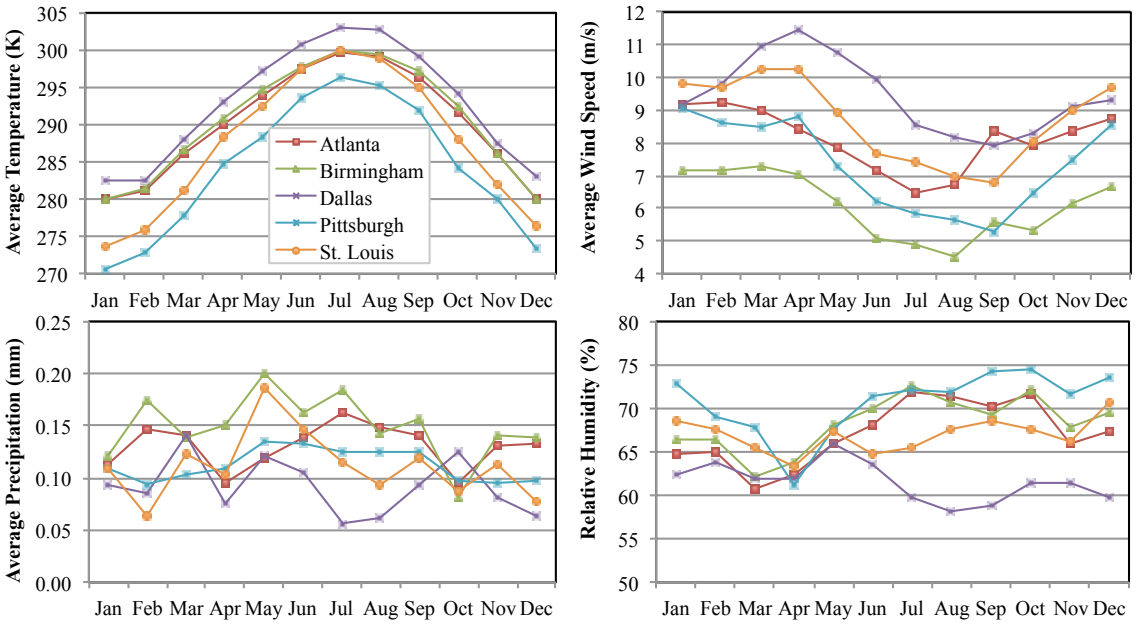

Figure 4.2 – Monthly mean climatology data for five cities for the 2002 to 2008 period from the NNDC CDO system database.

4.2.2 Air Pollution Observations

The current study focuses on 12 pollutants: five gases (CO, NO_2, NO_x, O_3, and SO_2), two particulate matter (PM) mass measurements (PM_{10}, $PM_{2.5}$), and five $PM_{2.5}$ components: sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), elemental carbon (EC), and organic carbon (OC). Ambient air concentrations for these 12 pollutants of interest were acquired for the 2002 to 2008 period from several ground-based monitoring networks within or near the MSA of each city. Data were obtained from the following networks: the State and Local Air Monitoring Stations (SLAMS) network; the Chemical Speciation Network (CSN);[103] the Southeastern Aerosol Research and Characterization (SEARCH) network;[104] the Interagency Monitoring of Protected Visual Environments (IMPROVE) network;[103] and the Assessment of Spatial Aerosol Composition in Atlanta (ASACA)

network.[105] For PM measurements, hourly Federal Equivalent Methods (FEMs) data were calibrated to Federal Reference Method (FRM) data,[106] and EC and OC measurements by the thermal optical transmittance (TOT) method were converted to be equivalent to thermal optical reflectance (TOR) measures.[107-108]

Data were treated as missing on days with fewer than 16 hourly observations for hourly sampled species. Otherwise the missing hourly measurements of gases and PM mass were estimated by linear interpolation within days with at least 16 hourly observations. Daily metrics, based on the National Ambient Air Quality Standards for criteria pollutants, were calculated for each pollutant of interest for the health analyses; these are one-hour maxima for CO, NO_2, NO_x, and SO_2, the eight-hour daily maximum for O_3, and 24-h averages for PM_{10} and $PM_{2.5}$ masses and $PM_{2.5}$ species. Details regarding the central site selected for each pollutant and imputations done to fill in for missing data can be found elsewhere.[100, 109-111] Table 4.2 provides the city-specific time range, number of observations based on daily data across all monitors, number of monitors, summary statistics and the distribution of measurements of all observations based on daily data across all monitors used in this study by city and pollutant. Monitor locations are also shown in Fig. 4.1.

Table 4.2 – Descriptive statistics for ground-based observations, categorized by metropolitan area and species for 2002-2008.

Pollutant (units)	City	Obs. No.	Monitor No.	Mean	Q_1	Q_2	Q_3	IQR	SD	Missing
CO 1-hr max. (ppm)	Atlanta, GA	9582	5	0.8	0.3	0.6	1.1	0.8	0.6	25%
	Birmingham, AL	11760	5	1.4	0.3	0.8	1.6	1.3	2.1	8%
	Dallas, TX	7087	4	0.6	0.3	0.5	0.8	0.6	0.5	19%
	Pittsburgh, PA	14258	6	0.6	0.3	0.5	0.8	0.6	0.5	7%
	St. Louis, IL-MO	6827	5	0.9	0.5	0.7	1.1	0.6	0.7	38%
NO_2 1-hr max. (ppb)	Atlanta, GA	16045	7	22.9	8	20	35	27	16.6	10%
	Birmingham, AL	3917	2	17.9	3.3	9	31.3	28	17.8	23%
	Dallas, TX	23693	13	21.7	12	20	30	18	12.5	17%
	Pittsburgh, PA	20208	9	25.6	17	25	33	16	12	12%
	St. Louis, IL-MO	14331	10	28.4	20	28	36	16	12.5	35%
NOx 1-hr max. (ppm)	Atlanta, GA	16032	7	0.06	0.01	0.03	0.07	0.06	0.08	10%
	Birmingham, AL	3914	2	0.04	0.00	0.01	0.05	0.04	0.05	23%
	Dallas, TX	23969	13	0.04	0.01	0.03	0.05	0.04	0.04	16%
	Pittsburgh, PA	20201	8	0.06	0.02	0.04	0.08	0.05	0.05	1%
	St. Louis, IL-MO	7777	7	0.05	0.02	0.04	0.06	0.04	0.05	49%
SO_2 1-hr max. (ppb)	Atlanta, GA	24018	16	11.00	2.00	5.00	13.50	11.50	15.10	41%
	Birmingham, AL	6804	3	12.20	3.50	7.60	15.80	12.30	14.50	11%
	Dallas, TX	8916	6	5.30	1.00	2.00	5.00	4.00	10.70	32%
	Pittsburgh, PA	28372	14	22.50	10.00	17.00	28.00	18.00	18.90	21%
	St. Louis, IL-MO	21182	16	16.10	4.00	9.00	17.00	13.00	26.50	40%
O_3 8-hr max. (ppb)	Atlanta, GA	42734	27	0.05	0.04	0.05	0.06	0.02	0.02	38%
	Birmingham, AL	24169	14	0.05	0.03	0.05	0.06	0.02	0.02	32%
	Dallas, TX	36274	20	0.04	0.03	0.04	0.05	0.02	0.02	17%
	Pittsburgh, PA	17404	11	0.05	0.03	0.05	0.06	0.02	0.02	38%
	St. Louis, IL-MO	17696	18	0.05	0.03	0.05	0.06	0.03	0.02	55%
PM_{10} 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	11155	20	23.40	15.00	22.00	29.80	14.80	11.50	78%
	Birmingham, AL	13375	13	33.00	18.00	27.00	41.00	23.00	22.70	60%
	Dallas, TX	1369	10	24.40	16.00	22.00	30.00	14.00	11.30	94%
	Pittsburgh, PA	26591	18	24.10	12.80	19.40	30.40	17.50	16.70	42%
	St. Louis, IL-MO	694	5	25.70	17.00	24.00	31.00	14.00	11.70	94%
$PM_{2.5}$ 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	43200	45	14.80	9.30	13.40	18.70	9.40	7.70	62%
	Birmingham, AL	20376	16	15.40	9.50	13.50	19.30	9.80	8.30	50%
	Dallas, TX	23533	22	10.80	6.90	9.70	13.60	6.70	5.50	51%
	Pittsburgh, PA	13744	15	15.30	8.60	12.80	19.00	10.50	10.00	64%
	St. Louis, IL-MO	16515	17	14.70	9.20	13.10	18.50	9.30	8.10	56%
$PM_{2.5}$-SO_4 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	9306	14	4.10	2.10	3.30	5.30	3.20	2.90	74%
	Birmingham, AL	5186	5	4.20	2.30	3.40	5.10	2.80	2.90	59%
	Dallas, TX	2216	5	3.20	1.60	2.70	4.20	2.60	2.10	80%
	Pittsburgh, PA	1810	4	5.20	2.50	3.90	6.30	3.80	4.10	82%
	St. Louis, IL-MO	2525	6	3.80	1.80	3.00	4.60	2.80	3.20	81%
$PM_{2.5}$-NO_3 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	9290	14	0.70	0.30	0.50	0.80	0.60	0.70	74%
	Birmingham, AL	5277	5	0.60	0.20	0.40	0.80	0.60	0.70	59%
	Dallas, TX	2216	5	0.70	0.20	0.40	0.80	0.60	1.10	80%
	Pittsburgh, PA	1732	4	1.40	0.50	0.90	1.80	1.30	1.30	83%
	St. Louis, IL-MO	2078	6	2.30	0.60	1.30	3.20	2.60	2.30	84%
$PM_{2.5}$-NH_4 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	7716	14	1.40	0.80	1.20	1.80	1.00	0.90	78%
	Birmingham, AL	5424	5	1.30	0.70	1.10	1.70	1.00	0.90	58%
	Dallas, TX	2215	5	1.20	0.60	1.00	1.60	1.00	0.90	80%
	Pittsburgh, PA	1732	4	2.00	1.10	1.70	2.50	1.50	1.40	83%
	St. Louis, IL-MO	2060	6	1.80	0.90	1.50	2.50	1.50	1.30	84%
$PM_{2.5}$-EC 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	9043	14	0.90	0.40	0.70	1.10	0.70	0.80	75%
	Birmingham, AL	5874	5	1.10	0.40	0.70	1.30	0.90	1.10	54%
	Dallas, TX	2252	5	0.60	0.30	0.50	0.80	0.50	0.40	79%
	Pittsburgh, PA	1501	4	0.90	0.40	0.60	0.90	0.50	1.20	85%
	St. Louis, IL-MO	2740	6	0.90	0.40	0.70	1.20	0.70	0.70	79%
$PM_{2.5}$-OC 24-hr avg. ($\mu g/m^3$)	Atlanta, GA	9017	14	3.20	1.80	2.80	4.00	2.20	2.00	75%
	Birmingham, AL	5615	5	4.00	2.40	3.30	4.80	2.40	2.50	56%
	Dallas, TX	2254	5	3.10	2.10	2.80	3.80	1.80	1.50	79%
	Pittsburgh, PA	1674	4	3.90	2.50	3.50	4.80	2.40	2.30	84%
	St. Louis, IL-MO	2392	6	3.70	2.30	3.40	4.60	2.20	1.80	82%

The number of observations and summary statistics are based on daily data across all monitors acquired for the 2002 to 2008 period with the exception of St. Louis (2002-2007) and Dallas (2003-2008). Q_1, Q_2, and Q_3 are the lower, median, and upper quartile of the observed values. The interquartile range (IQR) is the difference between the upper and lower quartiles. Missing refers to the percent of absent daily metrics over the study period. The larger PM mass and species fractions of missing data are a result of PM measurement sampling frequency of 3 or 6 days.

4.3 Methods

4.3.1 Chemical Transport Model Simulations

The U.S. Environmental Protection Agency (EPA) and Centers for Disease Control (CDC) collaborated in the Public Health Air Surveillance Evaluation (PHASE) project using the Community Multi-Scale Air Quality (CMAQ) model, to provide hourly air pollutant concentration fields over the eastern U.S., at a spatial resolution of 12 km and for the period that includes 2002-2008, for the purpose of public health analysis.[112-113] CMAQ is a deterministic chemical transport model that provides spatially- and temporally-resolved simulated outputs. The U.S. EPA conducted a comprehensive model evaluation of CMAQ version 4.5 for several different key synoptic weather patterns.[114-115] In that study, ozone simulation results were consistent with previous evaluations, where the model captured daily trends but had difficulty capturing the upper and lower limits of the observations. CMAQ had a tendency to over-estimate fine particulate matter mass during the autumn and winter. EPA's evaluation of CMAQ model version 4.5 also showed that particulate sulfate simulations were improved, compared to results of the previous version, due to changes in dry deposition velocity and cloud treatment. Particulate nitrate and ammonium were highly over-estimated by the CMAQ model version 4.5 during the fall. The evaluation also found that unresolved secondary organic aerosol pathways in the model contributed to under-estimation of carbonaceous aerosols during the late spring and summer months.

4.3.2 Data fusion exposure metrics

The applied data fusion approach blends ambient ground observations and modeled data from a CTM to estimate daily spatially-refined pollutant metrics and associated correlations. This hybrid method provides daily spatial field estimates of air pollutant concentrations and uncertainties that are consistent with observations, emissions, and meteorology; it is summarized briefly below, and is described in detail by *Friberg et al*.[24]

The optimized spatiotemporal concentration dataset (C_{opt}) is built using weighted fields of daily interpolated surface observation ratios and daily adjusted CMAQ result ratios that are rescaled to estimated annual mean fields (eq. 4.1):

$$\underbrace{C_{opt}{}_{s,t}}_{\substack{Daily\ Fused\\ Fields}} = \underbrace{\alpha_{year}\overline{C_{CMAQ_s}}^{\beta}}_{\substack{Estimated\ Annual\\ Mean\ Fields}} \left[\underbrace{W_{s,t}}_{\substack{Weighting\\ Factor}} \underbrace{\left\{ C_{OBS_{s_m},t} / \overline{C_{OBS_{s_m}}} \right\}_{krig}}_{\substack{Daily\ Interpolated\\ Obs.\ Ratio\ Fields}} + \underbrace{\left(1 - W_{s,t}\right)}_{\substack{Weighting\\ Factor}} \underbrace{\left\{ C_{CMAQ}{}_{s,t} / \overline{C_{CMAQ}{}_s} \right\} \gamma_{season}{}_t}_{\substack{Daily\ Adjusted\ CMAQ\\ Results\ Ratio\ Fields}} \right] \tag{4.1}$$

Here, α_{year} is a regression parameter that adjusts the amplitude to account for interannual differences, C_{CMAQ} represents CMAQ simulated concentrations, the overbar indicates annual temporal averaging, β is a regression parameter that accounts for interspecies nonlinearity differences, s indicates spatial location, t represents time, W is a weighting factor, C_{OBS} represents observed concentrations, s_m indicates monitor locations, and γ_{season} is a species-specific, seasonal correction function having amplitude, phase, and

sinusoidal form. Scaling the daily ratio fields by the spatially regressed annual mean observations reduces model biases. The estimated annual mean fields are developed from CMAQ-derived annual mean spatial fields adjusted to observed annual means using power regression models. These regression parameters are city-, pollutant-, and year-specific, because the spatial pattern of emissions, annual effects of meteorological variables, and CMAQ biases differ among the five cities (supplemental material). The exponent β is assumed to be constant across years, whereas α_{year} changes annually.

The daily-resolved observation-based ratio fields capture the robust temporal variance characterized by ground monitors. These concentration fields are calculated by spatially interpolating the normalized daily observed concentrations using kriging. The daily adjusted CMAQ result ratios capture the spatial variance while reducing bias. The optimization is based on a spatiotemporal weighting factor (W) that maximizes the degree to which the observation-based estimate captures temporal variation relative to the CMAQ-based estimate, as a function of distance from an observation (eq. 4.2). Due to missing data, the weighting factors vary over time as well as space. The temporal Pearson correlation fields of the daily observation-based fields, R_1, are derived using an exponential correlogram modeled to fit the isotropic spatial autocorrelation of the observations (eq. 4.3). The fitted parameters include the intercept which results from instrument error estimated by collocated instruments (R_{coll}), the distance between a grid centroid to the nearest observation on a given day (x), and the range (r) at which the correlation between monitors has decreased to an e-folding of R_{coll}. The term R_1 varies over space and time because the observation frequency varies among monitors. The

average of the temporal correlations of the CMAQ simulations and observations across all monitors (n_{s_m}) is used to estimate R_2, which represents the estimated temporal correlation of the daily adjusted CMAQ results ratio fields and ambient pollution (eq. 4.4). The spatiotemporal weighting factor is also applied to the observation- and CMAQ-based temporal correlation fields to quantify the uncertainties of the optimized spatiotemporal concentration dataset (R_{opt}; eq. 4.5).

$$W_{s,t} = \frac{R_{1_{s,t}}(1 - R_2)}{R_{1_{s,t}}(1 - R_2) + R_2\left(1 - R_{1_{s,t}}\right)} \tag{4.2}$$

$$R_{1_{s,t}} \approx R_{coll} e^{-x_{s,t}/r} \tag{4.3}$$

$$R_2 \approx \frac{1}{n_{s_m}} \sum_{s_m} corr\left(C_{OBS_{s_m}}(t), C_{CMAQ_{s_m}}(t)\right) \tag{4.4}$$

$$R_{opt_{s,t}} = \begin{cases} W_{s,t} R_{1_{s,t}} + \left(1 - W_{s,t}\right) R_2 & R_1 > R_2 \\ R_2 & R_1 =< R_2 \end{cases} \tag{4.5}$$

4.3.3 Taylor Diagram

For this study Taylor diagrams are used to visually summarize how well model results match observations and to evaluate performance changes among models.[116-117] Patterns between the modeled datasets relative to the observations can be quantified and graphically summarized in terms of four statistical measures of model performance that are defined in the supplemental material: spatiotemporal Pearson correlation (R) calculated using all spatial and temporal model values coincident with observations

(Supplemental eq. S3), centered pattern of the root-mean-square difference (eq. S7), standard deviation (eq. S6), and mean bias (eq. S8). Combined, these statistics provide key information about the differences in amplitude of their variance, degree of accuracy, correlation, and biases between the simulated and observed datasets into a single plot.

Measures of model performance are located in the first quadrant of the Taylor diagram when the simulated datasets are positively correlated with the observations (e.g., Fig. 4.3 and Supplemental Figs. B.2-B.5). The abscissa of the Taylor diagram represents a perfect correlation between the simulated data and the observations. The correlation between the modeled and observed datasets is given by the azimuthal position of the modeled dataset. The radial distance from the origin is proportional to the normalized standard deviation (NSD) of the modeled dataset. Normalization of the pollutant-city datasets by the corresponding standard deviation of the observations allows all the cities to be shown on the same plot. The distance between the modeled and observed data overall is proportional to the normalized centered (i.e., unbiased) root-mean square difference (NCRMSE), represented by a combination of R and NSD values as green concentric circles about the reference point. The reference point, plotted as a unit distance along the abscissa, represents ideal model-to-observation performance statistics (R=1, NSD=1, and NCRMSE=0). The data points nearest the reference point, the green bull's eye in Fig. 4.3, have the smallest NCRMSE values and thus provide the closest fit to the observed data. By adding the mean bias (MB) to the diagrams (encoded as color within the markers), the performance measures provide information about both

systematic and unsystematic errors. The MB color scale assignment changes per pollutant.

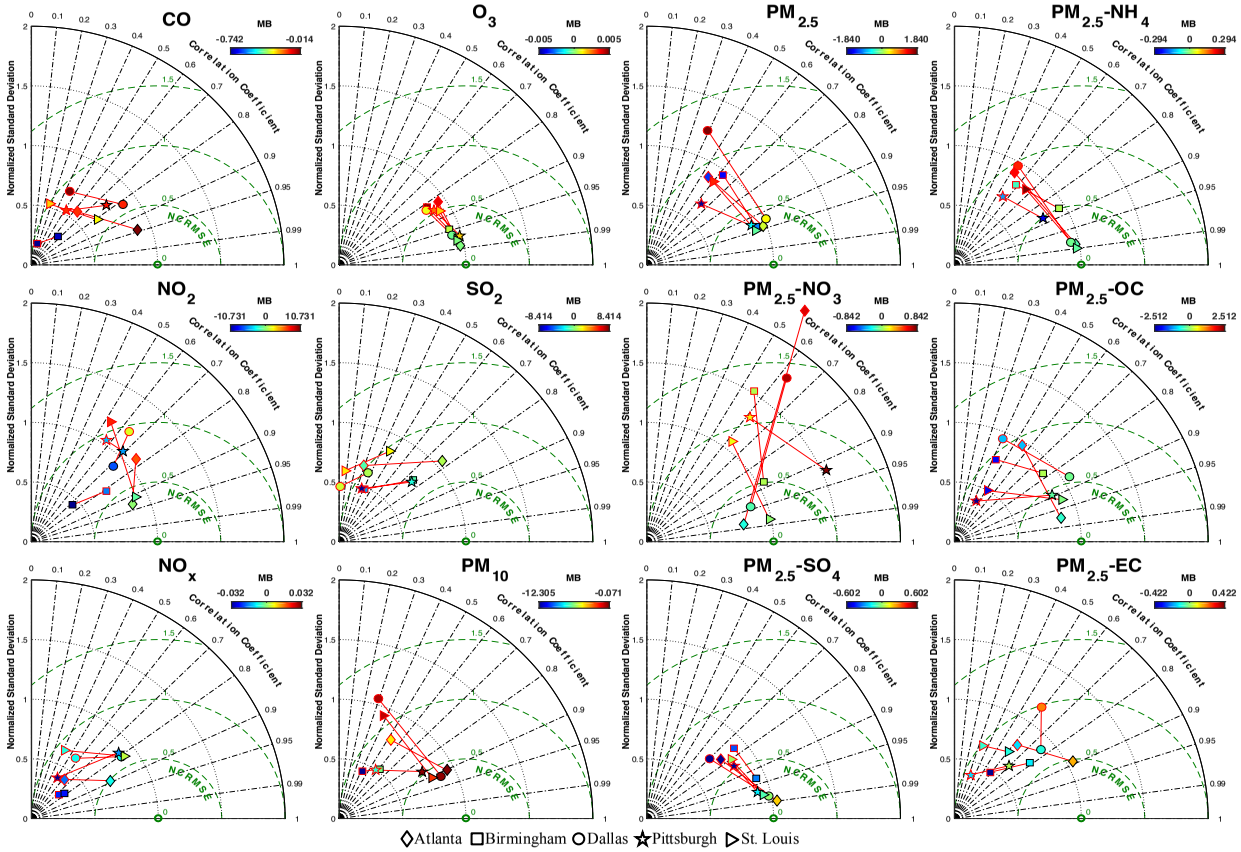

Figure 4.3– Spatiotemporal correlation coefficient (R) Taylor diagrams show normalized statistics for five cities, indicated by distinct symbol, for each pollutant. C_{CMAQ} symbols are outlined in red, whereas C_{opt} symbols are outlined in black. The symbol color indicates the MB. Correlation coefficient is represented by azimuth in these plots, NSD by the radial direction, and the superposed green dashed curves indicate NCRMSE. Perfect agreement is indicated by the bullseye along the horizontal axis.

4.4 Results

This study includes different numbers of monitors for each of the 12 pollutants across the 5 regional domains. Three sets of comparisons are performed. First, the model concentrations C_{CMAQ} are evaluated for each pollutant and metropolitan domain. Second, C_{opt} is compared with the monitor data, to characterize data fusion model performance. Third, results of a comprehensive cross-validation analysis are presented, to evaluate data fusion model simulative capabilities across multiple metropolitan domains. Although observations have inherent measurement errors, the model performance criteria are quantified in terms of how well the model results fit observed data.

4.4.1 CMAQ Simulation Performance

The CMAQ model evaluation conducted by the U.S. EPA was extended to include the multi-year spatial and temporal domains of the pollutant metrics for 12 pollutants: CO, NO_2, NO_x, O_3, SO_2, PM_{10}, $PM_{2.5}$, and fine particulate components NH_4, NO_3, SO_4, OC, and EC (Fig. 4.3, Supplemental Table B.4). The Taylor diagram panels (Fig. 4.3) graphically summarize the correlation, normalized centered root-mean-square difference, normalized standard deviation, and mean bias (eqs. B.1-B.7) for the C_{CMAQ} and C_{opt} results for each city, stratified by pollutant. This figure shows the spatiotemporal Pearson correlation, which assesses the model simulation of variation over space and time, coincident with observations. The simulations are distinguished by marker edge color, with C_{CMAQ} statistics outlined in red and C_{opt} statistics outlined in black. Each model-related result is further categorized into the five distinct city domains indicated by five different

marker shapes and orientations: Atlanta (diamond), Birmingham (square), Dallas (circle), Pittsburgh (star), and St. Louis (triangle).

The C_{CMAQ} statistical performance patterns across all cities vary widely across the 12 pollutants studied (Fig. 4.3). The C_{CMAQ} simulation statistics for ozone and sulfate show the strongest clustering pattern and best performance across all cities. Their spatiotemporal correlations to observations for all cities are greater than 0.81 and 0.71, and their RMSEs are lower than 0.011 ppb and 2.31 μg/m^3, respectively.

The squared temporal correlations (Table B.4) are also highest for ozone and sulfate. Other pollutants show lower correlations and larger spread of values across cities. With the exception of Dallas, the C_{CMAQ} simulation statistics for $PM_{2.5}$ also show a strong clustering pattern across all cities and relatively high spatiotemporal correlations to observations, with values above 0.55 (Dallas R=0.39). C_{CMAQ} statistical performance is poorest across all cities for SO_2 and PM_{10}. Their spatiotemporal correlations with observations for all cities have the lowest values, ranging between 0.1 to 0.41 and 0.29 to 0.57, whereas their RMSEs range between 11.76 and 29.88 ppb and 11.98-23.96 μg/m^3, respectively. C_{CMAQ} simulation RMSE and SD statistics were consistently high across all cities for NO_2, SO_2, PM_{10}, and $PM_{2.5}$ (Table B.4). C_{CMAQ} simulated the spatial, temporal, and spatiotemporal variations least accurately in 1-hr maximum SO_2 and 24-hr average PM_{10}, demonstrating known limitations in the CMAQ model to simulate ground-level impacts of SO_2 plumes and biogenic PM_{10} source contributions.[118-119]

4.4.2 Data Fusion Simulation Performance

As expected, C_{opt} improves the MB and NCRMSE compared to the C_{CMAQ} results across most cities and pollutants over the seven-year period (Fig. 4.3). The MB and NCRMSE underperformance of the C_{opt} dataset for NO_2 at Birmingham and NO_3 at Pittsburgh are a result of averaging over space (Fig. B.2). The C_{opt} improvements for NO_3 in Pittsburgh, where the NSD is poorer but R (Fig. 4.3) and R^2 (Table B.4) are closer to unity, is an example of the chosen tradeoffs between improving correlation and improving the standard deviation. Among the cities, the highest centered NCRMSE and NSD variability of C_{CMAQ} results is seen for NO_3. Spatiotemporally and temporally, Figs. 4.3 and B.2 show the C_{opt} dataset significantly improves the C_{CMAQ} results for nearly all pollutants and all cities (with the exception of NO_2 for Birmingham), with the greatest correlation improvements occurring for O_3 and $PM_{2.5}$ and the least for SO_2. Spatially, the variance in R^2 values (Table B.4) derives from the dependence of C_{opt} on C_{CMAQ}. For example, NO_2, NO_x, and SO_2 at Birmingham, and PM_{10} at St. Louis exhibit relatively high spatial R^2 values as compared to the other cities because they have the smallest number of monitoring sites for these particular pollutants, causing the optimization to rely heavily on C_{CMAQ} results rather than observations. Blanchard et al., (2014) found several Birmingham monitors for PM_{10}, $PM_{2.5}$, and NO_x to be affected by emissions from local sources.[43] The combination of CMAQ underestimating nitrogen oxides at the affected Birmingham monitor and the lack of monitors (only 2 monitors in Birmingham) leads to underperformance in the fusion dataset, with an increase in MB and RMSE. The C_{opt} statistics indicate a decrease in C_{opt} model performance relative to the C_{CMAQ} statistics for NO_x in Birmingham, where the NCRMSE increased and the correlation

decreased. Another pollutant that showed minimal C_{opt}-related performance improvements is PM_{10} for Birmingham, again due to sparse monitoring. The elevated PM_{10} concentrations in this study agree with the levels reported by Blanchard et al., (2014). Overall, the spatial R^2 values are fairly consistent across both modeled datasets, with SO_2 and PM_{10} being the lowest. Spatiotemporal correlation coefficient values for primary gases show wider variance (i.e., greater azimuthal coverage) across cities, whereas $PM_{2.5,}$ speciated $PM_{2.5}$, and O_3 show the least amount of variance and highest values (Fig. 4.3).

Density scatter plots of modeled results (C_{CMAQ}, C_{opt}) versus coincident observations for 2002-2008 per city and pollutant are shown in Fig. B.1. The regression correlation coefficients for C_{opt} are higher than the C_{CMAQ} results across all pollutants and cities (with the exception of NO_2 in Birmingham), indicating better fits at observation locations.

Average spatial C_{opt} distributions of each of the 12 pollutant concentrations and five cities during a seven year span are shown in Fig. 4.4. Primary pollutants (CO, NO_2, NO_x, SO_2, and EC) are spatially more heterogeneous than pollutants of secondary or mixed origin (O_3, PM_{10}, $PM_{2.5}$, SO_4, NO_3, NH_4 and OC), with greater urban-to-rural gradients, as might be expected. Coarse and nitrate particulates are examples of spatially homogenous pollutants showing minimal spatial variation across all cities. Spatially heterogeneous pollutants are characterized as those pollutants whose intraurban temporal correlations exhibit less uniformity over space (Fig. 4.5). The impact of major urban centers is evident in these 12-km resolved fields, as indicated by the locations of

concentration peaks. The R_{opt} temporal correlation fields based on eq. B.4 are shown in Fig. 4.5. The R_{opt} values are highest near monitors and decrease as the distance to the nearest monitor increases. For pollutants with observations only in metro-areas, the estimated temporal correlation for non-metro areas is that of the C_{CMAQ} in these plots.

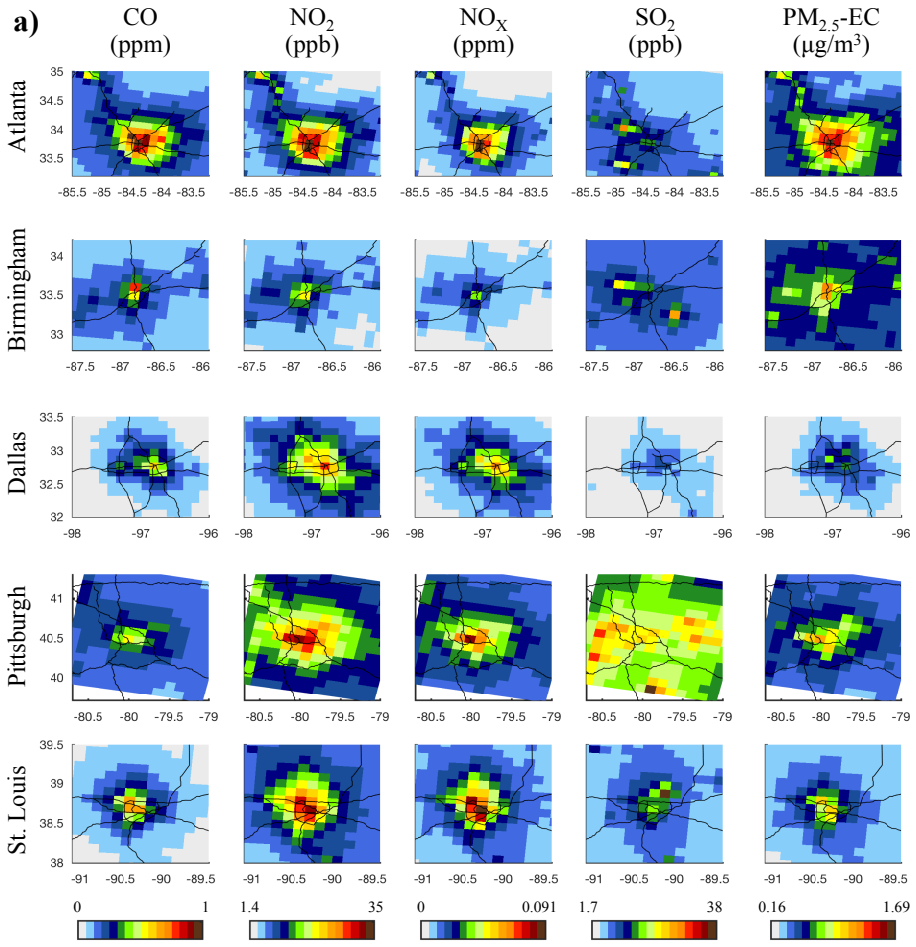

Figure 4.4 – Mean optimized concentration fields (C_{opt}) of (a) primary and (b) secondary or mixed origin pollutants, 2002-2008.

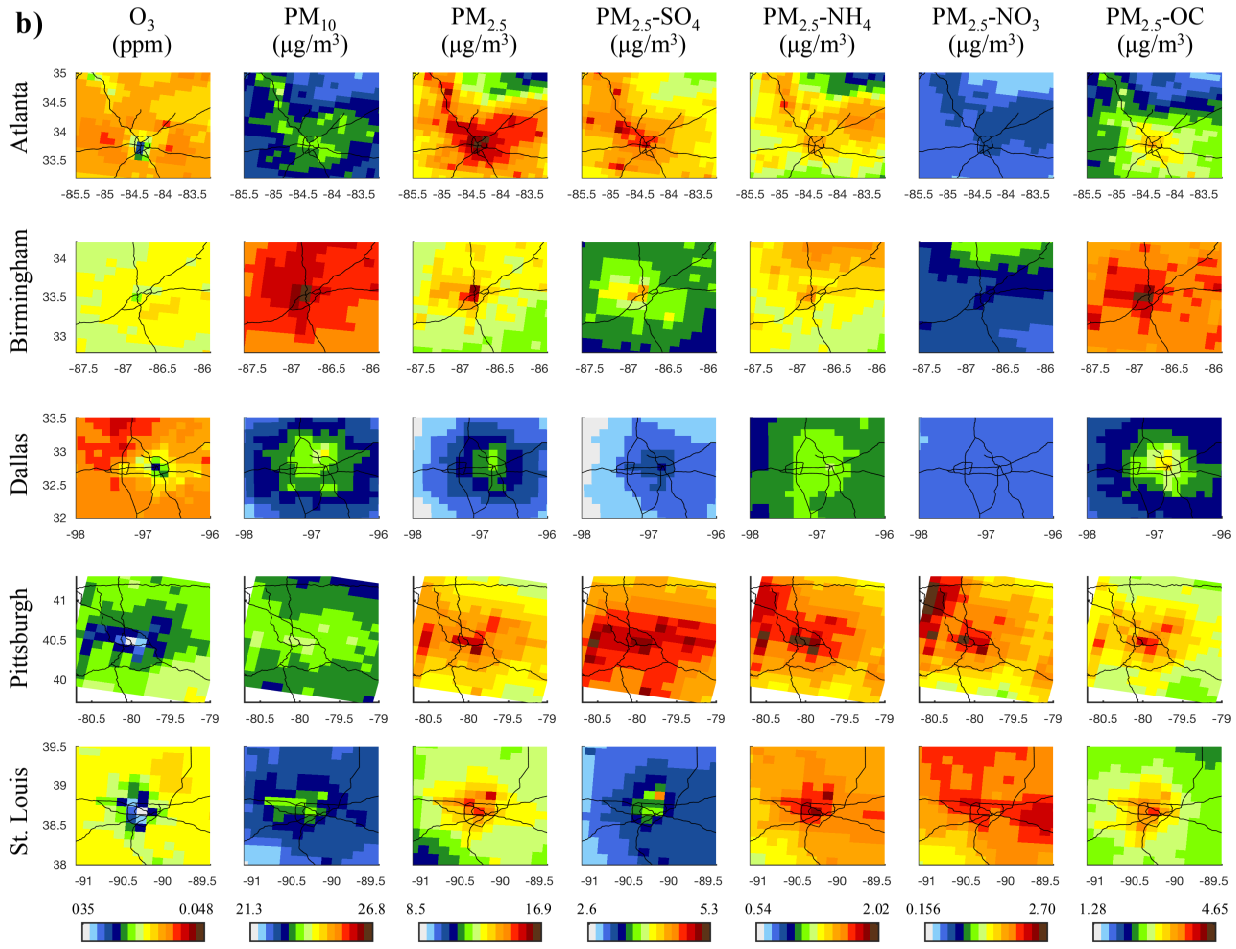

Figure 4.4 – (*continued*)

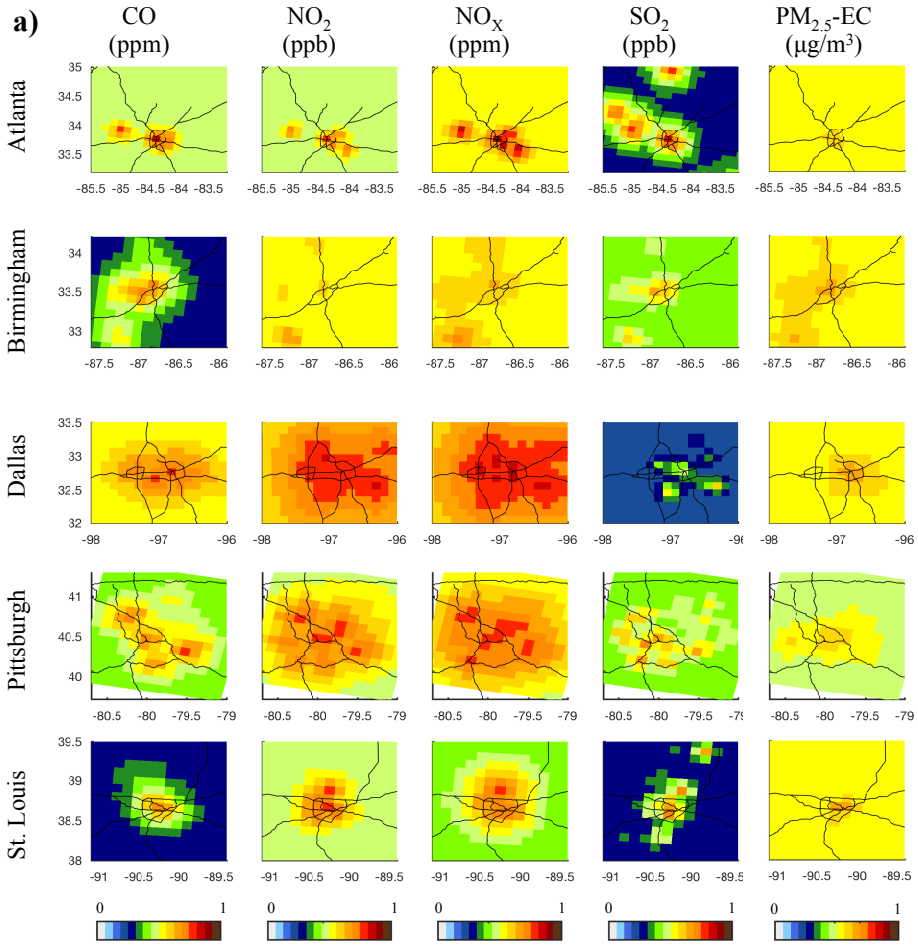

Figure 4.5 – Mean correlation coefficient fields (R_{opt}) of (a) primary and (b) secondary or mixed origin pollutants, 2002-2008.

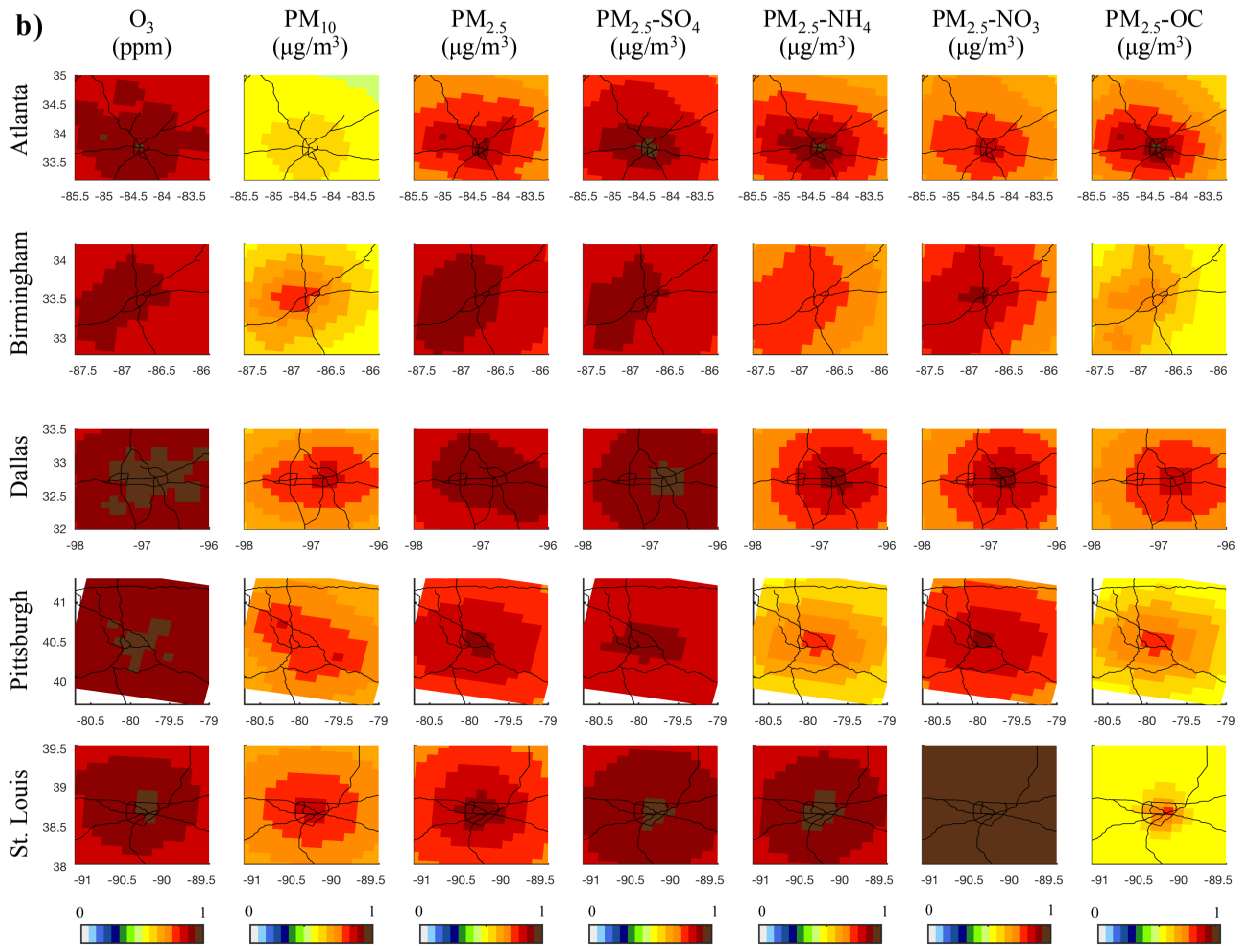

Figure 4.5 – (*continued*)

The interurban comparison of C_{opt} (Fig. 4.4) shows Atlanta and Dallas have the highest ozone levels and most notable urban area ozone titration effects by primary pollutants (e.g., NO_x) of the five cities. Atlanta followed by Pittsburgh and St. Louis have high NO_2/NO_x, and high $PM_{2.5}$ carbon levels, indicative of high motor vehicle pollution. On the other hand, Dallas has low levels of primary pollutants relative to the other cities in this study, likely due to its higher wind speeds. The wide range in wind speeds in Dallas, Atlanta and St. Louis results in different mixes of primary and secondary air pollutants from motor vehicles. Dallas is higher in ozone, derived in large part from motor vehicle emissions, and lower in primary motor vehicle emissions,[120-121] whereas St. Louis is higher in primary motor vehicle emissions (e.g. copper from brake linings).[122-125] Atlanta is high in both primary and secondary motor vehicle emission impacts.[36, 96, 106, 126-127] Pittsburgh is affected by coal-fired power plants, as evidenced by high SO_2 levels, and high sulfate and selenium (often used as a coal emission tracer) levels.[128-129] The high SO_2 and $PM_{2.5}$-SO_4 values for Pittsburgh are distributed around the region, reflecting the distribution of coal-fired power-plant sources. Birmingham is impacted by industrial sources, with high levels of a number of metals; its low wind speeds results in the slow primary $PM_{2.5}$ dispersion, contributing to high CO levels as well.[104, 130] Thus, the five cities included in this analysis have different ambient pollution mixtures that represent a broad range of large, non-coastal U.S. urban areas. Figs. 4.3 and B.2 show the statistical performance patterns across the five cities are particularly good for O_3, $PM_{2.5}$, NH_4, and SO_4, indicating CMAQ's ability to capture temporal trends. In general, the Copt statistics show improvements in correlation and NCRMSE across all

cities when compared to the C_{CMAQ} statistics, with the exception of the NO_x. The 12 km resolution CMAQ simulations used in this application are too coarse to capture the fine spatial gradients of NO_x and other primary pollutants emitted from mobile sources near roadways. Conversely, the 12 km resolution is adequate to capture the spatial uniformity of secondary aerosol. For O_3 and $PM_{2.5}$, the clustering of C_{opt}-related points around the reference point shows clear improvements across all cities, as compared to the C_{CMAQ} points.

The C_{opt} statistical performance patterns in Fig. B.2 are particularly high for O_3 and $PM_{2.5}$, and noticeable for SO_4, NH_4, NO_3, OC, and PM_{10} across all cities. Their range of spatiotemporal squared-correlations to observations across all cities are R^2 = 0.90-0.97 (O_3), 0.38-0.84 (PM_{10}), 0.86-0.90 ($PM_{2.5}$), 0.77-0.97 (NO_3), 0.87-0.98 (SO_4), 0.75-0.98 (NH_4), and 0.60-0.95 (OC), respectively. The C_{opt} for SO_2 has the poorest performance across all cities, with spatiotemporal R^2 ranging between 0.13 and 0.59, RMSE values ranging between 9.68 and 25.74 ppb, and SD values ranging between 6.62 and 22.59 ppb. Although the C_{opt} results for SO_2 show some improvements in correlations, the seemingly unchanged NCRMSE values cause a de-clustering effect in the Taylor plots. Dallas has the lowest background SO_2 concentrations, whereas Pittsburgh has the highest background SO_2 concentrations, which in turn lead to the highest sulfate concentrations. The localized SO_2 peaks coincide with stationary sources and exhibit steep spatial concentration gradients (Fig. 4.4).

The Taylor diagrams shown in Fig. B.2 depict the model results by monitor and temporal Pearson correlation for each city and pollutant. Normalization of the pollutant-

city datasets by the corresponding SD of the observations allows data for all the cities to be included in each panel. Comparison between the C_{CMAQ} and C_{opt} statistics indicates performance improvement and less heterogeneity across cities in the latter dataset, across all four statistics for the secondary pollutants. The primary pollutants, on the other hand, show more variability at the interurban level and across cities. Performance improvement trade-offs of the optimized dataset are more apparent in these panels, which graphically summarize the temporal correlation, NCRMSE, NSD, and MB (eqs. S1, S4, S6-S8) for the C_{CMAQ} and C_{opt} results for each city, classified by pollutant. For example, in Fig. B.2 the NO_x panel for St. Louis indicates that for one particular monitor, improvements in correlation come at the expense of worsening the NSD, RMSE, and MB values.

4.4.3 Data Fusion Cross-validation Performance

To evaluate the potential advantages of the data fusion model, the cross-validation spatiotemporal, temporal, and spatial correlations are compared for C_{opt} and C_{CMAQ}. The data fusion model performance is evaluated using a comprehensive tenfold 10% withholding cross-validation analysis per city and pollutant. That is, ten percent of the observations are randomly held back across all monitors and days per pollutant, city, and trial. The data fusion method is then applied to simulate the withheld data per pollutant, city, and trial. The results from the trials for each pollutant and city are then combined to provide robust cross-validation results that allow for the exploration of differences in errors based on proximity to monitors. Across monitors and days, the number of

withheld data points corresponds to the number of observations for each pollutant and city (Table 4.1), ranging from 695 for PM_{10} in St. Louis to 43,200 for $PM_{2.5}$ in Atlanta.

Fig. 4.6 summarizes the 10-fold cross-validation spatiotemporal correlation results for C_{opt} and coincident C_{CMAQ} for the twelve pollutant concentrations on days without model-observation pairs, or at locations without pollutant monitors. The cross-validation C_{opt} spatiotemporal variance (Pearson R^2) values across all cities are 0.88-0.93 (O_3), 0.14-0.71 (PM_{10}), 0.67-0.83 ($PM_{2.5}$), 0.52-0.72 (NO_3), 0.81-0.89 (SO_4), 0.43-0.80 (NH_4), and 0.32-0.51 (OC), respectively. Cross-validation C_{opt} metrics for SO_2 and CO are poorest across all cities with spatiotemporal R^2 from 0.0 to 0.19 and 0.0 to 0.53, respectively. Cross-validation demonstrates a wide range in optimized model performance across pollutants, with SO_2 simulated most poorly due to limitations in coal combustion plume monitoring and modeling. This performance analysis indicates that higher resolution CMAQ simulations are needed to resolve the spatial gradients of primary pollutant, especially those related to roadway emissions. The emission-source-related trends are discussed further in the following section. For the other cross-validated pollutants studied, 0.54-0.88 of the spatiotemporal R^2 was captured, with ozone and $PM_{2.5}$ simulated best.

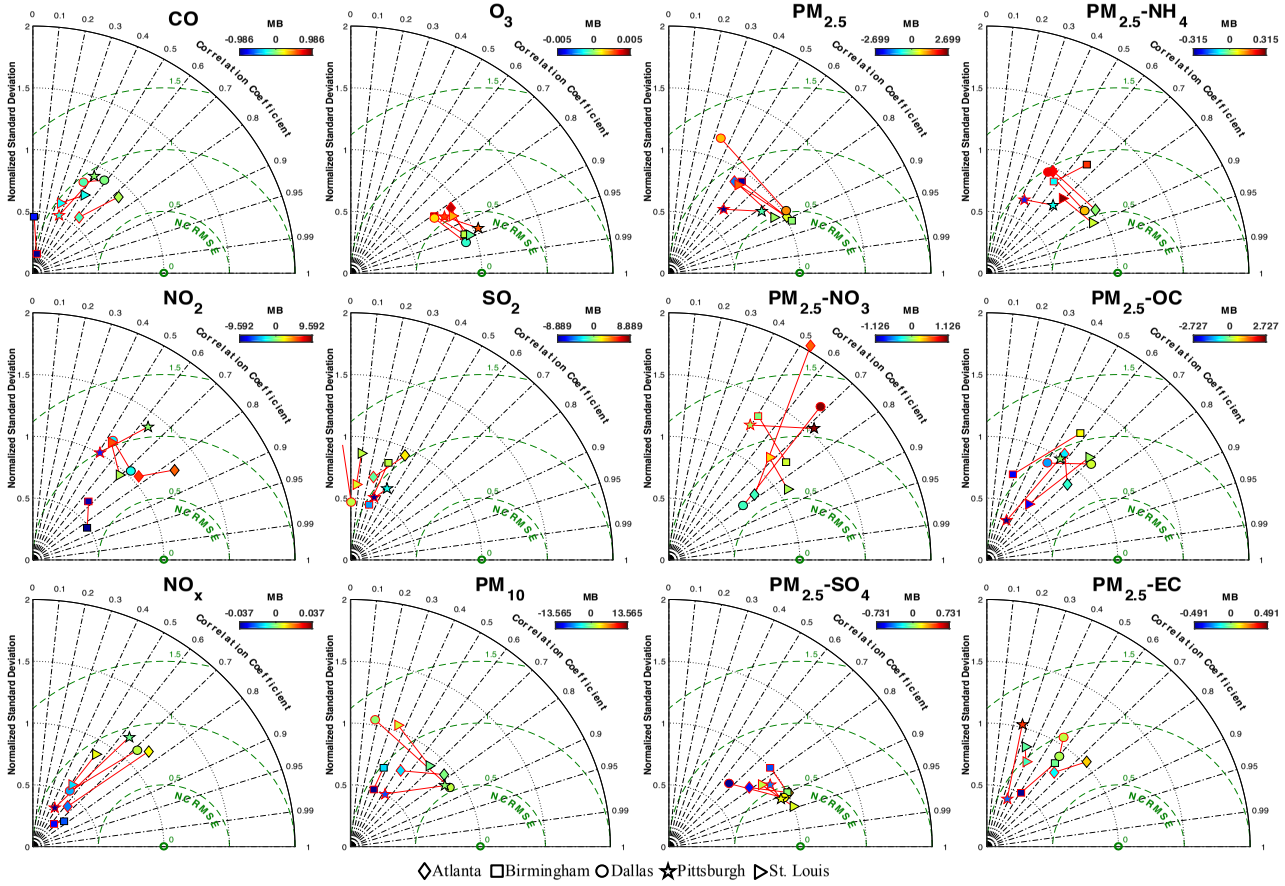

Figure 4.6 – Cross-validation spatiotemporal correlation coefficient (R)Taylor diagrams show normalized statistics for five cities, indicated by symbols, per pollutant. C_{CMAQ} symbols are outlined in red, whereas C_{opt} in black. The symbol color indicates the mean absolute gauge error.

Data fusion captures temporal variance overall better than CMAQ (Fig. B.3). Despite decreasing the number of available observations, data fusion increases the temporal correlations, as compared to unadjusted CMAQ values, due to the method's integration of observations. The C_{opt} SO_2 spatial correlations are the lowest of all the pollutants across all cities with the exception of Birmingham. The SO_2 spatial R^2 values are below 0.20 (eq. B.5, Table B.5). Similar to NO_x and NO_2, Birmingham's low SO_2 spatial correlation arises from having the fewest monitors (Table 4.2 and Fig. 4.4-4.5).

Based on Figs. B.4 and B.5, seasonally, the C_{opt} results are biased high for NO_3 across all locations in fall and biased low in the summer. C_{opt} results are biased low for EC and OC at the rural sites from May through August, although the EC biases are minimal. These are a result of the CMAQ systematic biases.[114-115] For Pittsburgh, seasonal C_{opt} trends for particulate mass and component concentration corroborate the results by Tang et al., (2004). Particulate nitrate dominated the $PM_{2.5}$ total mass concentrations in winter, whereas sulfate dominated in summer. Large regional $PM_{2.5}$ contributions to the Atlanta area are most notable during the summer months, supporting results from previous studies.[36] Several of the particulate sulfate concentration peaks coincide with the SO_2 peaks from power plants. Consistent with model results from Takahama et al. (2004), the conversion of SO_2 to sulfate is significantly faster in the summer across all cities (Fig. B.6 in supplemental material).

4.5 Discussion

Data fusion leads to improved spatiotemporal fields versus using unadjusted datasets, providing substantial improvements over CMAQ alone for nearly all pollutants and all cities, with the exception of NO_2 for Birmingham, which is affected by a paucity of measurement stations. The greatest spatiotemporal cross-validation correlation improvements occur for O_3 and $PM_{2.5}$, and the least for SO_2 (Table 4.3), probably because O_3 is a secondary and spatially more homogeneous pollutant with a high percent of available daily metrics, whereas SO_2 sources tend to be spatially dispersed and not very well characterized. In agreement with previous studies, the measurement error associated with spatial misalignment of SO_2 exposure is very high, even when ambient monitor data and deterministic chemical transport model simulations are fused. Some factors that limit the data fusion results are: (1) the spatial coverage of ambient measurements and the biases associated with CMAQ simulations (e.g., NO_x and NO_2 in Birmingham), (2) inflated estimates of model cross-validation performance linked to the clustering of monitors, such as $PM_{2.5}$ mass in Atlanta (Fig. 4.1), (3) over-estimation of spatial autocorrelation due to the use of kriging in the daily interpolated observation ratio field estimates, which tends to affect the more spatially homogeneous pollutants (e.g., PM_{10}, SO_4, and NH_4), and (4) situations where CMAQ has systematic biases and measurements that could correct the results are lacking (e.g., NO_3 and carbonaceous aerosols). Nevertheless, the fused spatiotemporal dataset substantially outperforms CMAQ results as an exposure metric (Table 4.3). Key themes identified with this multicity data fusion application that are related to physical aspects of the five urban

environments include: (1) airshed complexity, (2) spatial and spatiotemporal patterns, (3) emission-source-related trends, and (4) constrained uncertainty fields.

Table 4.3 – Range of spatiotemporal Pearson correlations coefficients (R^2) across all cities for simulation-based observations categorized by species, covering 2002-2008.

Pollutant	Spatiotemporal R^2		Cross-validation Spatiotemporal R^2	
	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}
CO	0.06-0.40	0.27-0.89	0.30-0.38	0.00-0.53
NO_2	0.28-0.66	0.48-0.87	0.26-0.59	0.40-0.72
NO_x	0.17-0.55	0.60-0.79	0.22-0.43	0.29-0.57
SO_2	0.00-0.17	0.13-0.59	0.00-0.11	0.00-0.19
O_3	0.67-0.75	0.90-0.97	0.66-0.74	0.88-0.93
PM_{10}	0.09-0.33	0.38-0.84	0.03-0.29	0.14-0.71
$PM_{2.5}$	0.15-0.42	0.86-0.90	0.12-0.39	0.67-0.83
$PM_{2.5}$-SO_4	0.50-0.71	0.87-0.98	0.44-0.71	0.81-0.89
$PM_{2.5}$-NO_3	0.31-0.39	0.77-0.97	0.24-0.47	0.52-0.72
$PM_{2.5}$-NH_4	0.27-0.44	0.75-0.98	0.19-0.48	0.43-0.80
$PM_{2.5}$-EC	0.11-0.42	0.37-0.79	0.14-0.42	0.07-0.55
$PM_{2.5}$-OC	0.16-0.32	0.60-0.95	0.08-0.34	0.32-0.51

4.5.1 Airshed Complexity

Unquestionably, differences in airshed complexity and meteorological conditions among the cities have a significant impact on the interurban spatial concentration gradient. The effect of high mean wind speed on Dallas is evident in the spatial fields of all species. With the exception of ozone, spatial concentration gradients in Dallas tend to be lower than the other cities. Synoptic-scale winds in Dallas are unimpeded by topographic obstacles, allowing cleaner air masses to flow more readily than in the other cities; this is most apparent in the spatial homogeneity across the PM constituents. Another aspect of airshed complexity affects the Pittsburgh results, where the influence of long-range transport from the Ohio River Valley is evident in the elevated background

concentrations of several particulate species. Ammonia sources in the North Central Plains affect background particulate ammonium concentrations in Pittsburgh and St. Louis, elevating the intraurban peaks. Findings by Tang et al. and Turner et al. corroborate the effect of regional transport on elevated and spatially homogenous summer $PM_{2.5}$ mass as well as several other major component (SO_4, NH_4, and OC) concentrations in Pittsburgh.[131, 133]

4.5.2 Spatial and Spatiotemporal Patterns

Assessment of the spatiotemporal exposure concentrations generated using the data fusion method provides largely consistent patterns of emissions that leads to consistency in daily spatial patterns of intraurban variability for several gaseous and particulate species across the five non-coastal metropolitan areas. In accordance with previous research, this study confirms that the secondary species (O_3, PM_{10} mass, $PM_{2.5}$ mass, sulfate, ammonium, and OC) are notably more spatially homogenous within each of the five cities than the primary species (Fig. 4.5).[36, 93, 133] Furthermore, the current study also shows that the spatial heterogeneity of primary species (CO, NO_2, NO_x, EC) is dominated by contributions from mobile sources, based on the high spatial gradients in species concentration from the city centers to the surroundings (Fig. 4.4). Nitrogen oxide episodes, and to a lesser extent carbon monoxide episodes, across all cities are of high intensity and short spatial scale. NO_x and NO_2, two traffic-related pollutants, exhibit gradual concentration gradients, with peaks coinciding with the urban centers. The fusion method's ability to capture trends is heavily reliant on the input datasets: observations and CMAQ results. It is no surprise that the temporal performance of the CMAQ results,

particularly for O_3, $PM_{2.5}$, NH_4, and SO_4, across the five cities is further improved by the additional, observational constraints applied by the fusion method (Table 4.3).

4.5.3 Emission-source-related Trends

Of particular interest is the comparison of traffic-related pollutants across cities. Although the Dallas metropolitan area has the largest population and total vehicle miles, Dallas does not show the highest traffic-related-pollutant spatial concentration gradients as a result of its airshed complexity. Atlanta, with its lower wind speeds and higher population density, produces consistently higher and non-uniformly distributed concentration gradients across traffic-related pollutants (e.g., CO, NO_x, NO_2, EC). However, the cities with the highest populations, Dallas and Atlanta, exhibit the highest spatial ozone concentration fields. High O_3 concentrations in Dallas persist into the fall months. Atlanta has the highest CO mean concentrations and larger spatial concentration gradients. Due to the spatial scale of the plume, the ability to correctly capture winds, as well as other plume and model characteristics, SO_2 is simulated most poorly among the pollutants studied, across all cities.

4.5.4 Constrained Uncertainty Fields

By capturing the spatial complexity of the more heterogeneous pollutants, the data fusion method combines the available observations with physically based simulations, thus increasing the accuracy of ambient pollution uncertainty estimates compared to use of observations or CMAQ results alone. The data fusion method constrains the uncertainty in the spatial-daily exposure concentration datasets. Reductions in

spatiotemporal uncertainty estimates are most notable for homogenous species with many monitors. The range of correlation field estimates for the more homogenous species across all cities are R = 0.81-0.96 (O_3), 0.44-0.84 (PM_{10}), 0.54-0.92 ($PM_{2.5}$), 0.49-0.97 (NO_3), 0.74-0.97 (SO_4), 0.55-0.97 (NH_4), and 0.50-0.96 (OC), respectively (Fig. 4.5).

4.6 Conclusion

In this study, the ability of the data fusion method to capture spatiotemporal air pollution trends for various pollutants across multiple metropolitan areas is assessed. The aim is to develop daily ambient air pollution metrics. Additionally, the spatiotemporal uncertainty of the results and the degree to which the results are generalizable between locations is evaluated. Using a cross-validation technique, the fused estimates and uncertainties are compared against observations, and the ability of the fused estimates to capture the multicity spatiotemporal air quality landscapes verified.

Adequate spatial and temporal characterization of local, regional, and global air quality source contributions bolster population-based acute health effect studies, which can be affected by exposure misclassification. Total reliance upon observations can be misleading during the development of spatial ambient air concentration metrics, especially due to spatial and/or temporal sampling limitations, but combined with emissions, meteorology, and chemical transport modeling, a clearer picture of the spatiotemporal air quality landscape is gained. The rationale for this evaluation requires that the fused observations and chemical transport modeling results adequately capture spatial and temporal variability of the air pollutants across five cities. Although this

method is applied using 12-km CMAQ fields, it can be readily adapted to finer-resolution modeling results.

CHAPTER 5. CONSTRAINING CHEMICAL TRANSPORT $PM_{2.5}$ MODELING USING SURFACE STATION MEASUREMENTS AND SATELLITE RETRIEVALS: APPLICATION OVER THE SAN JOAQUIN VALLEY

Abstract

Advances in satellite retrievals of aerosol type can improve the accuracy of near-surface air quality diagnostics, by providing a broad regional context. In addition to aerosol optical depth, qualitative constraints on aerosol size, shape, and single-scattering albedo provided by multi-angle instruments, such as the Multi-angle Imaging SpectroRadiometer (MISR) on the NASA Earth Observing System's Terra satellite, can provide frequent, spatially extensive, instantaneous constraints on chemical transport models (CTM), especially useful in areas away from ground monitors and progressively downwind of emission sources. CTMs (e.g. the Community Multi-scale Air Quality Modeling System) complement such data by providing complete spatial and temporal coverage, offering additional physical constraints (e.g., conservation of aerosol mass, meteorological consistency) independent of observations, and identifying relationships between observed species downwind and emission sources. Incorporating satellite aerosol information in the development of $PM_{2.5}$ concentration metrics can lead to a decrease in metric uncertainties and errors.

This work focuses on the potential spatial improvement to surface estimates of $PM_{2.5}$, its major chemical species estimates, and related estimates of uncertainty from combining regional-scale satellite and CTM data. Aerosol airmass types over populated regions of Southern California are characterized using satellite data acquired during the 2013 San Joaquin field deployment of the NASA DISCOVER-AQ project. Using the MISR Research Aerosol Retrieval algorithm (RA), we investigate and evaluate the optimal application of incorporating 275 m horizontal-resolution aerosol airmass-type maps and total column aerosol optical depth into a 2 km resolution, regional-scale CTM, to obtain constrained fields of surface $PM_{2.5}$. Contemporaneous suborbital observations and additional, high-resolution CTM simulations are used to evaluate the results. The impact of incorporating MISR aerosol data on the ability to characterize air quality progressively downwind of large single sources is discussed. The spatiotemporal R^2 values for the model, constrained by (1) both satellite and surface-station measurements and (2) only surface-station measurement (values in parentheses) based on 10% withholding, are 0.79 (0.80) for $PM_{2.5}$, 0.88 (0.88) for NO_3, 0.78 (0.79) for SO_4, 1.00 (1.00) for NH_4, 0.73 (0.75) for OC, and 0.31 (0.32) for EC. Regional cross-validation temporal and spatiotemporal R^2 results for the satellite-based $PM_{2.5}$ improve by 30% and 13%, respectively, in comparison to CTM simulations. SO_4 cross-validation values showed the largest spatial and spatiotemporal R^2 improvement with a 43% increase.

5.1 Introduction

To investigate air pollution health effects on humans, population-based epidemiologic time-series studies often use exposure measures derived from regulatory

monitoring networks.[134-136] Even for the continental U.S., many ambient, ground-level fine particulate matter ($PM_{2.5}$) chemical datasets are acquired only once every three or six days, and at many sites for less than a decade or two. In addition, they tend to be concentrated in a small number of populated counties, with the main exception of the isolated Interagency Monitoring of Projected Visual Environment (IMPROVE) program sites in national parks.[137] Prior to 2009, instrument types and sensitivities varied considerably from station to station and among monitoring networks,[138] making comparisons and uncertainty assessments difficult.

Urban-level epidemiological time-series studies often span large geographic regions.[36, 89] Especially for long-term exposure analysis, broad regions significantly downwind of urban and industrial centers are also of great importance due to the presence of distributed populations, and natural and agricultural ecosystems. Reducing exposure-metric error caused by inadequately characterized spatial variability, which is often much larger than instrument error, can substantially reduce bias and improve precision in epidemiologic results.[90, 95-96] This is particularly relevant for regional-scale studies, where the spatial variability of ambient surface $PM_{2.5}$ and chemical species concentrations, such as urban-to-rural gradients, fundamental to effectively conducting environmental epidemiologic studies and air quality assessments, can be lacking.

Although chemical transport model (CTM) simulations provide more complete spatial and temporal coverage than surface stations, they rely on uncertain inputs about pollution source characteristics that can contain significant biases. The accuracy of the simulated fields is also affected by the accuracy of the simulated meteorology, and of the

physical and chemical parameters specified in the model.[42] Errors in these fields can be identified and sometimes quantified by comparison with coincident ground- and aircraft-based observations. Under satisfactory retrieval conditions, satellite-derived aerosol optical depth (AOD), atmospheric scattering, light absorption, and extinction by suspended particles can be leveraged to constrain the columnar CTM simulations in less intensely monitored areas.

The earliest $PM_{2.5}$ air quality studies that were based upon spacecraft data simply correlated ground-level $PM_{2.5}$ concentrations and satellite-derived AOD from the MODIS instruments, without accounting for particle vertical distribution.[139-140] This worked well when the aerosol was almost entirely concentrated in the near-surface boundary layer, but not when transported aerosol made a significant contribution to the total-column AOD, or when the boundary layer was deep or variable on short timescales. Other early studies used surface measurements[141] or an aerosol transport model[48, 142] to provide some constraint on aerosol vertical distribution, but still did not account in detail for either spatial or temporal variations of the relationship, and provided no observational constraints on aerosol type. The Van Donkelaar et al. (2010) study used space-based CALIPSO lidar backscatter profiles to validate the GEOS-Chem model vertical distributions globally, aggregated over a four-year period.

The first papers to include some space-based aerosol type information along with AOD from satellites for air quality applications used the Multi-angle Imaging SpectroRadiometer (MISR) spherical vs. non-spherical distinctions to separate airborne dust from spherical particles over the continental U.S., and constrained aerosol vertical

distribution and speciated the spherical components with an aerosol transport model.[45-46] Subsequent work applied MISR aerosol size and shape constraints over the Indian subcontinent and surrounding areas to map seasonal changes in aerosol type[143] and combined MISR particle shape and qualitative light absorption information to make a first effort at mapping aerosol airmass types over an urban area, Mexico City.[47]

To estimate ambient $PM_{2.5}$ mass and associated chemically speciated concentrations on regional scales, a systematic and practical approach is developed and evaluated in this paper. It uses ground-station $PM_{2.5}$ measurements, where available, to anchor speciated, near-surface CTM aerosol concentrations. To help constrain the model over extended regions, MISR total column AOD is also applied, along with qualitative, column-effective aerosol type observations where mid-visible AOD values exceed 0.15. Enhanced aerosol-type retrievals from the MISR Research Aerosol retrieval algorithm (RA), at 1 km horizontal resolution, are at the heart of this new approach.

To demonstrate the method, we apply it over a case study area in the San Joaquin Valley of California during the DISCOVER-AQ field campaign in this region, on six days when there is good MISR coverage. The results account for spatiotemporal variability in $PM_{2.5}$ and in the chemical component concentrations. The accuracy of estimated concentrations and evaluation of the latest MISR-RA ability to typify urban AOD, aerosol mixtures, and aerosol airmasses, are examined by comparing the results with speciated ground observations and standard model fitting statistics. Section 5.2 describes the datasets involved, Section 5.3 describes the method and technical approach, and Section 5.4 presents results and validation for our test cases. Conclusions, along with

a brief discussion of prospects for wider application of this approach, are given in Section 5.5.

5.2 Study Domain and Datasets

5.2.1 Study Domain

The San Joaquin Valley (SJV), which comprises the southern two-thirds (about 10,000 sq. miles) of California's Central Valley, has long suffered from severe air pollution issues and is among the most studied air sheds in the U.S.[144-145] It has complex topography, and especially in winter, low PBL heights and high pollutant mixing ratios, creating challenges for chemical transport modeling.[146] This region is surrounded by the Sierra Nevada to the east, the Diablo and Temblor Ranges to the west, the Tehachapi Mountains to the south, and the Sacramento Valley to the north (Figure 5.1).[147] Although primarily a rural area, the eight counties that comprise the SJV are home to more than 4 million residents. This arid to semiarid valley is one of the world's most productive agricultural regions. The SJV air shed frequently experiences high $PM_{2.5}$ concentrations during the winter, due to the combination of the surrounding mountain ranges, relatively dry climate, shallow PBL heights, and local source emissions. The SJV has been in violation of the $PM_{2.5}$ National Ambient Air Quality Standards for $PM_{2.5}$ annual standard since their inception in 1997, and is the largest $PM_{2.5}$ nonattainment area in the continental U.S.[148]

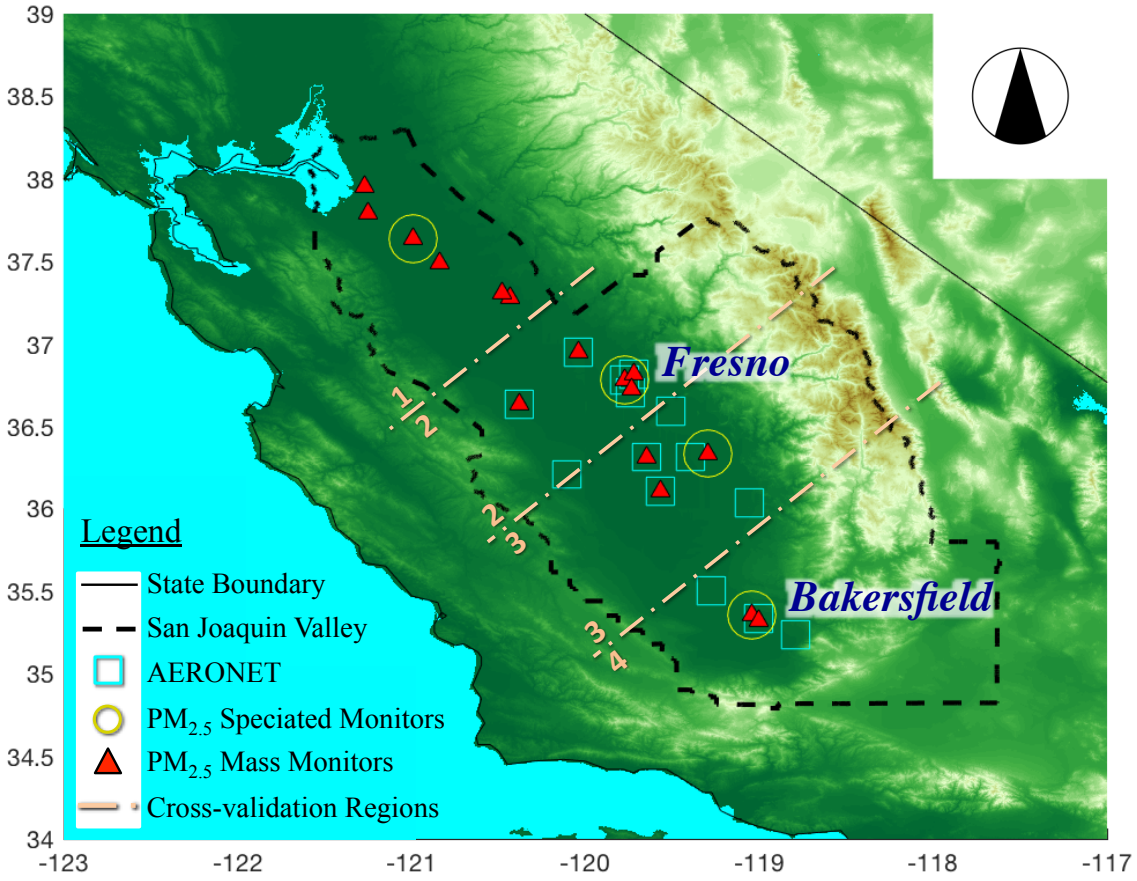

Figure 5.1 – San Joaquin study area shows the ground elevation, EPA AQS and CSN monitors, and AERONET sites during the two-month NASA DISCOVER-AQ flight campaign

The study period, January and February 2013, was selected to coincide with the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ; http://www.nasa.gov/) campaign. This campaign was a joint collaboration between NASA, NOAA, U.S. EPA, and several local organizations, with the goal of characterizing air quality in urban areas using satellite, aircraft, vertical profiler and ground-based measurements (http://discover-aq.larc.nasa.gov). Targeting the 2013 DISCOVER-AQ period for this study provides

additional ground- and aircraft-based measurements for aerosols and fine particulate matter, which we apply as model constraints and validation.

We analyze data for six days during the DISCOVER-AQ period for which (1) MISR observations were made over the study region, (2) coincident ground and air observations were acquired, including extensive field-campaign data, and (3) the key observational requirements of relatively cloud-free conditions and the presence of aerosols from different sources are met. Of the six days for which we have MISR coverage, the mid-visible AOD exceeds 0.15 on three days: January 20, February 03, and February 05. On lower-AOD days, MISR aerosol type information is too uncertain for the current application. Of the three higher-AOD days, January 20 has the least cloud cover, followed by February 05, so these will be the main focus of detailed analysis. The method developed here can in the future be applied to many other polluted regions of the world where AOD exceeding 0.15 is common, such as south and East Asia and North Africa.

The ground-based, aircraft, and simulation data used in this study are described briefly in the rest of this section, along with the MISR research aerosol retrieval (RA) product.

5.2.2 Ground-based PM Mass and Speciated Measurements

This study focuses on $PM_{2.5}$ measurements and its five major components: sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), elemental carbon (EC), and organic carbon (OC). Data files of ambient aerosol particulate matter species concentrations for sites

within the SJV for January and February 2013 were obtained from two EPA sources: (1) daily averaged $PM_{2.5}$ Federal Reference Method (FRM) mass from the Air Quality System (AQS; https://www.epa.gov/aqs), and (2) daily averaged total PM and chemically speciated mass (measurements typically made every third or sixth day) from the Chemical Speciation Network (CSN).[103]

FRM compliant data from gravimetric filter-based samplers and Federal Equivalence Method (FEM) compliant data from continuous mass monitors provide spatial variability of $PM_{2.5}$ mass.[149] The $PM_{2.5}$ FRM mass is determined gravimetrically by weighing particles on filters pre- and post-deployment. They are equilibrated at a constant relative humidity (30-40%) and temperature (20-23 degrees C). Monitor locations are shown in Figure 5.1, and Table 5.1 lists monitor summary statistics. Daily $PM_{2.5}$ concentrations measured by the FRM method are considered $PM_{2.5}$ ground truth, i.e., their uncertainties are small compared to those of the other $PM_{2.5}$ values used in this study.

Table 5.1 – EPA AQS and CSN monitor summary statistics for 52 days (6 days).

Pollutant	No. of Monitors	Sampling Frequency	OBS	Mean	SD
$PM_{2.5}$, $\mu g/m^3$	22 (21)	13 daily; 6 1-in-3; 3 1-in-6	779 (95)	21.20 (28.31)	13.33 (13.51)
$PM_{2.5}$-SO_4, $\mu g/m^3$	7 (6)	6 1-in-3; 1 1-in-6	86 (11)	0.77 (1.13)	0.46 (0.69)
$PM_{2.5}$-NO_3, $\mu g/m^3$	7 (6)	6 1-in-3; 1 1-in-6	86 (11)	7.27 (9.81)	6.11 (7.38)
$PM_{2.5}$-NH_4, $\mu g/m^3$	5 (4)	4 1-in-3; 1 1-in-6	54 (7)	2.07 (3.65)	2.25 (3.32)
$PM_{2.5}$-EC, $\mu g/m^3$	4 (4)	3 1-in-3; 1 1-in-6	44 (8)	1.28 (1.14)	0.77 (0.34)
$PM_{2.5}$-OC, $\mu g/m^3$	4 (4)	3 1-in-3; 1 1-in-6	44 (8)	5.25 (5.73)	3.09 (2.48)

5.2.3 DISCOVER-AQ AERONET DRAGON

The AErosol RObotic NETwork (AERONET)[150] has 10 permanent sun photometer (SP) stations operating in the study region. During the DISCOVER-AQ 2013 deployment, these stations were supplemented with an additional 14 temporary stations termed the Distributed Regional Aerosol Gridded Observation Network (DRAGON) to provide a more regionally dense dataset for satellite validation and *in situ* comparisons (Fig. 1). AERONET/DRAGON SPs measure AOD in eight spectral bands (340, 380, 440, 500, 670, 870, 1020 and 1640 nm), with an accuracy within ±0.015.[151]

We use version 2 (v2) level 2 AERONET/DRAGON AOD and Angstrom Exponent (AE) data for the six study days. The level 2 data were sun-calibrated after field deployment, cloud screened,[152] and quality controlled. The AOD at 550nm wavelength is calculated using a quadratic log-log fit.[151] Columnar AODs at 550nm derived from AERONET are considered as AOD ground truth in this study.

5.2.4 Chemical Transport Model Simulations

Simulations of the coupled Weather Research and Forecasting model (WRF),[153] version 3.4, and the Community Multiscale Air Quality (CMAQ) model,[154] version 5.0.2, were obtained from the U.S. Environmental Protection Agency (EPA). These atmospheric simulations, at a 2 km × 2 km horizontal grid spacing with 35 vertical layers, cover the entire SJV and surrounding major cities during the months of January and February of 2013. Concentration fields from the fixed 2 km × 2 km horizontal CMAQ grid were downscaled to a horizontal grid of 275m x 275m by linear interpolation and used as the reference grid for all subsequent analyses. Emission data was based on the

2011 EPA National Emissions Inventory [155] with 2013 updates to electric generating unit emissions, fire, and mobile sources. Biogenic emissions were generated using the Biogenic Emissions Inventory System (BEIS) version 3.14 (http://www.cmascenter.org), and the emissions were processed using the Sparse Matrix Operator Kernel Emissions (SMOKE; https://www.cmascenter.org/smoke/) version 3.5. [156] The carbon bond 2005 chemical mechanism used was CB05TULC.[157-159] The lateral Boundary Conditions (BCs) for the 2-km simulation were derived from a coupled WRF-CMAQ simulation with 4-km × 4-km horizontal grid spacing, covering the entire state of California and the surrounding areas. BCs for the 4-km simulation were derived from a 36-km simulation covering the contiguous United States, and BCs for the 36-km simulation were provided by a hemispheric GEOS-Chem[160] simulation with the chemical species mapped to the corresponding CMAQ species.[146]

The U.S. EPA conducted a model evaluation of CMAQ v5.0.2 with respect to the scientific updates to v5.1.[146] In that study, fine particulate matter simulations were biased low compared to observed concentrations over the SJV during the winter months. Winter $PM_{2.5}$ average mean bias (CMAQv5.0.2-Observations) in the SJV exceeded -10 μg/m3. PBL height and mixing were considered to be contributing factors to the January $PM_{2.5}$ underestimation in the SJV. Although CMAQv5.0.2 is missing several secondary organic aerosol species of anthropogenic volatile organic carbon (i.e., AALK1, AALK2, APAH1, APAH2 and APAH3) in its aerosol module (AERO6 v5.0.2), the mass contribution of these species to $PM_{2.5}$ during the winter was minimal (less than ±0.5 $\mu g/m^3$) in the SJV.[146] At the time this study was conducted, CMAQ v5.1 results were not yet available.

5.2.4.1 Constrained CMAQ simulation using ground-based observations

Ambient ground observations are used to constrain CTM model simulations (i.e., CMAQ), to estimate daily, spatially refined pollutant metrics and associated correlations. The method provides daily spatial field estimates of air pollutant concentrations and uncertainties that are consistent with observations at the ground stations, as well as with assumed emissions and modeled meteorology; it is summarized briefly below, and is described in detail by Friberg et al. (2016).

The constrained spatiotemporal concentration "fused" dataset (C_{FCMAQ}) is built using weighted fields of daily interpolated surface observation ratios and daily adjusted CMAQ result ratios that are rescaled to the estimated two-month study period mean fields (eq. 5.1):

$$C_{FCMAQ_{s,t}} = \alpha \overline{C_{CMAQ_s}}^{\beta} \left[W_{s,t} \left\{ \frac{C_{OBS_{s_m,t}}}{\overline{C_{OBS_{s_m}}}} \right\}_{krig} + (1 - W_{s,t}) \left\{ \frac{C_{CMAQ_{s,t}}}{\overline{C_{CMAQ_s}}} \right\} \right] \qquad (5.1)$$

Here, α is a regression parameter that adjusts the amplitude to account for inter-monthly differences, C_{CMAQ} represents CMAQ simulated concentrations, the overbar indicates two-month temporal averaging, β is assumed to be a constant, species-specific regression parameter that accounts for interspecies nonlinearity differences, s indicates spatial location, t represents time, W is a weighting factor, C_{OBS} represents observed concentrations, and s_m indicates monitor locations. Neither inter-monthly nor seasonal corrections were applied. Scaling the daily ratio fields by the spatially regressed two-

month mean observations reduces model biases. The estimated mean fields are developed from CMAQ-derived mean spatial fields adjusted to observed means using power regression models for the two-month time period of the current study. These regression parameters are species-specific, because CMAQ biases differ among the PM species.

The daily-resolved, observation-based ratio fields capture the robust temporal variance characterized by ground monitors. These concentration fields are calculated by spatially interpolating the normalized, daily-observed concentrations using kriging. As shown by Friberg et al. (2017), the daily-adjusted CMAQ result ratios capture the spatial variance while reducing bias. The optimization is based on a spatiotemporal weighting factor (W) that maximizes the degree to which the observation-based estimate captures temporal variation relative to the CMAQ-based estimate, as a function of distance from the observation (eq. 5.2). Due to missing data, the weighting factors vary over time as well as space. The temporal Pearson correlation fields of the daily observation-based fields, R_1, are derived using an exponential correlogram modeled to fit the isotropic spatial autocorrelation of the observations (eq. 5.3). The fitted parameters include the intercept that results from instrument error, estimated by collocated instruments (R_{coll}), the distance from a grid centroid to the nearest observation on a given day (x), and the range (r) at which the correlation between monitors has decreased to an e-folding value of R_{coll}. The term R_1 varies over space and time because the observation frequency varies among monitors. The average of the temporal correlations between the CMAQ simulations and observations across all monitors (n_{s_m}) is used to estimate R_2, which

represents the estimated temporal correlation of the daily adjusted CMAQ results ratio fields and ambient pollution (eq. 5.4). The spatiotemporal weighting factor is also applied to the observation- and CMAQ-based temporal correlation fields to quantify the uncertainties of the optimized spatiotemporal concentration dataset (R_{FCMAQ}; eq. 5.5).

$$W_{s,t} = \frac{R_{1_{s,t}}(1 - R_2)}{R_{1_{s,t}}(1 - R_2) + R_2\left(1 - R_{1_{s,t}}\right)} \tag{5.2}$$

$$R_{1_{s,t}} \approx R_{coll} e^{-x_{s,t}/r} \tag{5.3}$$

$$R_2 \approx \frac{1}{n_{s_m}} \sum_{s_m} corr\left(C_{OBS_{s_m}}(t), C_{CMAQ_{s_m}}(t)\right) \tag{5.4}$$

$$R_{FCMAQ_{s,t}} = \begin{cases} W_{s,t} R_{1_{s,t}} + (1 - W_{s,t}) R_2 & R_1 > R_2 \\ R_2 & R_1 =< R_2 \end{cases} \tag{5.5}$$

5.2.5 Satellite Observations

The primary satellite resource for this study is the Multi-angle Imaging SpectroRadiometer (MISR) instrument. We supplement the MISR aerosol data with results from the MODerate resolution Imaging Spectroradiometer (MODIS) instruments. They offer more extensive spatial coverage and provide somewhat greater diurnal sampling, though with larger AOD uncertainty over land, and with no constraints on aerosol type over land. We describe these two data sources below.

5.2.5.1 MISR – Research Aerosol Retrieval Algorithm (RA)

MISR was launched along with the first MODIS instrument aboard Terra, the flagship satellite of NASA's Earth Observing System (EOS), in December 1999.[161] Since then, Terra has maintained a sun-synchronous orbit, descending from North-to-South over the equator at a local time of ~10:30 AM. MISR measures upwelling short-wave radiance from Earth at nine distinct view angles along the line-of-flight (±70.5°, ±60.0°, ±45.6°, ±26.1°, and nadir), in each of in four spectral bands (446, 558, 672, and 866 nm). The one nadir, four forward, and four aft-viewing push-broom cameras take approximately 7 minutes to image a given 380 km wide swath of Earth. Due to swath size, it takes MISR about a week to obtain global coverage. Owing to its multi-spectral, multi-angular capabilities, high spatial resolution (up to 275 m), and highly accurate radiometric calibration,[162-164] MISR is uniquely capable of supporting air-quality applications, by providing information about aerosol microphysical properties at regional scales.

High-resolution (275 meter) results from the MISR Research Aerosol retrieval algorithm (RA) are used to constrain aerosol amount and type for the CMAQ model. Because of MISR's ability to sample over a large range of scattering angles (i.e., between about 60° and 160° at midlatitudes), the RA provides column-averaged information regarding aerosol absorption, particle size, and shape (non-spherical dust vs. spherical aerosol) under favorable retrieval conditions.[44, 50] Although passive satellite remote sensing can only provide information about aerosol type in two dimensions (column-averaged), a chemical transport model can be used to apportion the amount of aerosol near the surface (e.g., Liu et al., 2007a; van Donkelaar et al., 2010; this study).

5.2.5.1.1 MISR RA Retrieval Process

This subsection provides a brief summary of Limbacher and Kahn (2014; 2017). The basic principle of the MISR RA involves comparing the observed MISR top-of-atmosphere (TOA) reflectances with a pre-built look-up-table (LUT) of simulated TOA reflectances (analogous to a scaled, unitless radiance), and selecting only the aerosol optical depths and mixtures that meet certain goodness-of-fit criteria. To match the MISR-observed TOA reflectances ($\rho_{\lambda,c}^{TOA}$) in four spectral bands (λ) and each of up to nine cameras (c), simulated atmospheric and surface contributions to the TOA reflectance are calculated for a range of possible conditions and tested against the observations. The modeled portion of reflected light that reaches the instrument without interacting with the surface is the path reflectance ($\rho_{\lambda,c}^{path}$), and the modeled portion of reflected light that interacts with the surface is designated $\rho_{\lambda,c}^{surf}$. The *RA* uses a single minimization parameter (*M*) to self-consistently retrieve aerosol amount and type, as well as surface reflectance. For any given aerosol optical depth (AOD, or τ) and aerosol mixture combination, the minimization parameter *M* can be represented as:

$$M(\tau) = \sum_{\lambda} \sum_{c} \frac{w_{\lambda,c} * \left[\rho_{\lambda,c}^{TOA} - (\rho_{\lambda,c}^{path} + \rho_{\lambda,c}^{surf})\right]^2}{Unc_{\lambda,c}^2 * \left[\sum_{\lambda} \sum_{c} w_{\lambda,c}\right]} \tag{5.6}$$

The channel-specific weights are $w_{\lambda,c}$, and the assumed uncertainty of the entire model/measurement system is $Unc_{\lambda,c}$.

Because $\rho_{\lambda,c}^{surf}$ is not known a priori, this term must be determined before *M* can be computed. This two-step process involves first invoking the principle of angular shape similarity to compute a representation of the surface, by assuming that the angular shape of the Bi-directional Reflectance Distribution Function (BRDF) is fixed, but allowing it to vary spectrally.[165-166] Applying this additional constraint, the $\rho_{\lambda,c}^{surf}$ term of equation 5.6 is expanded into $\rho_{\lambda,c}^{surf} = E_{\lambda}^{BOA} * A_{\lambda} * T_{\lambda,c}^{up}$, where E_{λ}^{BOA} is the bottom-of-atmosphere (BOA) downward-directed irradiant reflectance, A_{λ} is the spectral albedo, and $T_{\lambda,c}^{up}$ is the azimuthally integrated transmittance from BOA to the MISR camera of interest. Because A_{λ} is not a function of view angle, we compute the optimal A_{λ} analytically for each band/AOD/mixture combination by taking the derivative of equation 5.6 with respect to A_{λ} and setting this equal to zero, yielding:

$$A_{\lambda} = \frac{\sum_c \left[\frac{w_{\lambda,c}}{Unc_{\lambda,c}^2} * T_{\lambda,c}^{up} * \left(\rho_{\lambda,c}^{TOA} - \rho_{\lambda,c}^{path} \right) \right]}{E_{\lambda}^{BOA} * \sum_c \left[\frac{w_{\lambda,c}}{Unc_{\lambda,c}^2} * \left(T_{\lambda,c}^{up} \right)^2 \right]} \tag{5.7}$$

The second step of this process requires that we modify $\rho_{\lambda,c}^{surf}$ such that $\rho_{\lambda,c}^{surf} = E_{\lambda}^{BOA} * A_{\lambda} * L_c * T_{\lambda,c}^{up}$, where L_c is a normalized, spectrally invariant but angularly varying, modulation of the surface albedo. This approximation simply implies that although the brightness of the surface can change with view angle, its color does not. Because A_{λ} and L_c cannot be calculated simultaneously, we instead use equation 5.7 to calculate a first

guess for A_λ, and then take the first derivative of equation 5.6 with respect to L_c, setting it equal to 0, and calculating L_c as follows:

$$L_c = \frac{\sum_\lambda \left[\frac{w_{\lambda,c}}{Unc_{\lambda,c}^2} * E_\lambda^{BOA} * A_\lambda * T_{\lambda,c}^{up} * \left(\rho_{\lambda,c}^{TOA} - \rho_{\lambda,c}^{path} \right) \right]}{\sum_\lambda \left[\frac{w_{\lambda,c}}{Unc_{\lambda,c}^2} * \left(E_\lambda^{BOA} * A_\lambda * T_{\lambda,c}^{up} \right)^2 \right]} \tag{5.8}$$

Substituting A_λ, L_c, and $\rho_{\lambda,c}^{surf} = E_\lambda^{BOA} * A_\lambda * L_c * T_{\lambda,c}^{up}$ into (6) yields the minimization parameter *M* for a particular AOD and aerosol mixture over land. (Note that L_c represents the angular dependence of the surface BRDF at the specific MISR view-angles.) The algorithm then selects the best fitting AOD for each of the 774 aerosol mixtures described in Limbacher and Kahn (2014), and saves the AOD, surface albedo, and associated goodness-of-fit parameter (M) for each mixture.

5.2.5.1.2 Applying the MISR RA Retrieval Results to Constraining the Air Quality Model

For comparison with the CTM, we compute aggregate AOD, Angstrom Exponent (ANG), absorption aerosol optical depth (AAOD), and non-spherical aerosol optical depth values from the RA results. As described below, we also compute aggregate AOD retrieved for the spherical absorbing aerosol components, and separately for spherical non-absorbing aerosol components. These aggregated parameters are calculated by weighting the respective parameters for each passing mixture by 1/M, such that better fitting mixtures are weighted more heavily than poorer fitting ones. The threshold value

of M used to determine passing mixtures is set to 1.25 * M_{min} + 0.25.[167] Because aerosol retrievals are affected by a range of conditions such as solar and viewing geometry, surface brightness, AOD, and aerosol type, we highlight below some of the key factors that help determine aerosol retrieval sensitivity.[50]

- *Surface brightness* – As the surface becomes brighter, the algorithm loses some sensitivity to all retrieved aerosol properties (including AOD). This occurs because, other things being equal, the contribution of $\rho_{\lambda,c}^{path}$ relative to $\rho_{\lambda,c}^{surf}$ decreases as surface brightness increases.
- *Number of cameras used* – To separate the two terms, the RA uses the property that $\rho_{\lambda,c}^{path}$ generally increases at steeper view angles, whereas $\rho_{\lambda,c}^{surf}$ generally decreases at steeper view angles. This also means that the number of cameras used can influence retrieval sensitivity to AOD and aerosol type.
- *Scattering Angle Range* – Other things being equal, a greater range of angles sampled by the MISR cameras relative to the solar direction offers higher confidence in the retrieved particle properties. As the aerosol scattering phase function peaks in the forward direction, retrieval sensitivity also tends to increase as the minimum scattering angle sampled decreases.
- *Retrieved Aerosol Optical Depth* – Mixture sensitivity diminishes when AOD is below about 0.15 or 0.2, although this also depends on other retrieval conditions.

Following the work of Patadia et al. (2013), we identify different aerosol airmasses by categorizing aerosol according to size (Angstrom Exponent), shape (non-

spherical AOD fraction), and single-scattering albedo (SSA, or 1.0 – AAOD/AOD). For the purposes of this paper, the 14 aerosol components used by all 774 mixtures[167] can be organized into three broad aerosol-type "groups": spherical light-absorbing, spherical non-absorbing, and non-spherical (cirrus is ignored). Especially at low-AOD, the MISR-derived aerosol-type constraints amount to broad classification (i.e., lower sensitivity) into light-absorbing spherical, non-absorbing spherical, and non-spherical particle components.[50] The general microphysical properties of the three broad aerosol groups (AG) can be associated with specific chemical species identified in the chemical transport model results, as described in Section 5.3.2 below. From the point-of-view of retrieval sensitivity, these three categories are as follows (Table 5.2): (1) Light-Absorbing Carbon (LAC), (2) Inorganic Ions (II) plus Organic Matter (OM) plus Sea-Salt (SS), and (3) dust. Over the last several years, substantial research has been published indicating that MISR AOD retrievals suffer biases in the presence of clouds.[163, 168-169] Consistent with both Witek et al. (2013) and Limbacher and Kahn (2015), we present results only for days where clouds (as indicated by our cloud mask) cover less than 30% of the scene within the SJV excluding the county areas that extend into the Sierra Nevada.

Table 5.2 – Classification of aerosol groups according to MISR RA microphysical property components.

Aerosol Group	**Light-Absorbing Carbon**		**Inorganic Ions + Organic**					**Dust**	
Component Shape	Spherical							Non-Spherical	
Component Absorption	Strongly Absorbing	Moderately Absorbing	Non-Absorbing					Varying Absorption Dust	
Component Description	Smoke Pollution	Smoke Pollution						Transported	Coarse
Component Effective Radius (μm)	0.12	0.12	0.06	0.12	0.26	0.57	1.28	0.75	2.40
Single-Scattering Albedo (558 nm)	0.82	0.91	1.00					0.98	0.90

*See Limbacher and Kahn (2014) for detailed component definitions.

5.2.5.2 MODIS - MAIAC

To supplement the MISR AOD values where MISR coverage is lacking, we adopt results from the MODIS Multi-Angle Implementation of Atmospheric Correction (MAIAC) advanced algorithm.[170] It uses time-series analysis and a combination of pixel- and image-based processing to improve the accuracy of cloud detection, aerosol retrievals and atmospheric correction for surface retrievals. The following is a brief overview of the MAIAC Collection 6 (C6) version 2.0 (v2) June 2017 North America release aerosol product. The current study uses the MAIAC Atmospheric Properties Products (MCD19A2), which provide AOD at 0.55 μm. A more detailed description of the MAIAC theoretical background and processing steps can be found in Lyapustin et al. (2011a,b; 2012).

After extensive characterization of the MODIS-observed surface background, the MODIS Level 1B data are gridded to a fixed sinusoidal projection at 1 km horizontal resolution in order to observe the same grid cell over time. Working with a fixed grid not only facilitates the use of polar-orbiting observations as if they were "geostationary", it also simplifies comparison of these datasets to fixed-grid model results and other measurements. In addition to the MODIS instrument on the Terra satellite, a second MODIS flies aboard NASA's Aqua satellite, which crosses the equator on the dayside at 1:30 PM local time. As a consequence of residual de-trending and MODIS C6 Terra-to-Aqua cross-calibration,[172] MAIAC currently processes MODIS C6 Terra and Aqua jointly as a single sensor. In addition to considerably greater spatial coverage than MISR, this joint product offers some diurnal spread in sampling relative to the MISR snapshots.

For the time series analysis, MAIAC utilizes a 4–16 day sliding window technique of scenes from multiple MODIS overpasses to retrieve surface BRDF at 0.466 μm wavelength, and spectral regression coefficients (SRCs; 0.466 μm and 2.13 μm), allowing MAIAC to retrieve AOD at 1 km spatial resolution. Unlike instruments that collect nearly simultaneous observations using push broom scanning, the MAIAC algorithm uses the sliding window technique of consecutive clear (i.e., cloud-free conditions with relatively low AOD) MODIS cross-track scanned scenes over several days to acquire multi-angle sets of observations for each location. This allows MAIAC to retrieve BRDF from an accumulated, multi-angle set of observations. Working under the assumption that surface reflectance changes rapidly over space and slowly over time (e.g., seasonal changes) helps the MAIAC internal dynamic land-water-snow classification. The algorithm produces well-characterized surface reflectance that improves cloud masking, and outperforms traditional pixel-level cloud detection techniques that rely on spatiotemporal analysis.[173]

Although AOD is originally retrieved in the MODIS Blue band B3 at 0.47μm, MAIAC offers a standardized and validated AOD product at 0.55μm. With the exception of smoke and dust aerosol detection, the current algorithm does not retrieve AOD over surfaces occurring at altitudes higher than 3.5 km. Like many satellite-based aerosol retrievals, MAIAC retrievals are unreliable for very low AOD conditions, over mountainous terrain, and over surfaces with high albedo. The retrieval conditions that affect this study include low AOD and cloud-contaminated scenes.

5.3 Methods

Air quality ground observations are spatially sparse, and are often temporally incomplete. CTM simulations provide information that is independent of these observations, and are consistent with meteorology and assumed emissions. But they can contain biases, and can have difficulty capturing the spatial structure of aerosol dispersion downwind of sources. Satellites offer spatially extensive, mainly column-effective aerosol amount and type, that, if included appropriately, can reduce or eliminate fused-model biases over large areas, especially regions far from concentrated surface stations. As there are gaps in the satellite products due to clouds and other retrieval-related issues, we use the model to help complete variable fields at several stages of the process. We also use the model to estimate the near-surface components of column-effective satellite values, and use ground-station data to constrain and to evaluate the results.

Our approach to fusing surface and satellite-based observations with CMAQ simulations involves five steps, illustrated in Figure 5.2. First, total column AOD and groupings of model aerosol species that match the spherical light absorbing, spherical non-absorbing, and non-spherical satellite aerosol-type AG are reconstructed from the simulated datasets. (Note that the left side of Figure 5.2 tracks the process for deriving total $PM_{2.5}$, whereas the right side presents the flow for speciated $PM_{2.5}$. Blue and orange asterisks in Figure 5.2 indicate where uncertainties are estimated by comparison with AERONET and the EPA ground stations, respectively.) In Step 2, spatially complete AOD and grouped AOD maps are produced for each of the six study days by combining MISR and MAIAC satellite retrievals with scaled values of the modeled AOD and AG

AOD products from Step 1, respectively, to fill any remaining gaps. In Step 3, we reconstruct $PM_{2.5}$ Mass FRM from the simulated concentration dataset. Step 4 deconstructs the satellite-based total column AOD and grouped AOD to surface $PM_{2.5}$ and grouped $PM_{2.5}$ mass concentrations using the CTM vertical and speciated vertical distributions, respectively. The fifth and final step involves blending daily averaged ambient ground observations and satellite-based total and grouped $PM_{2.5}$ mass concentrations to estimate daily, spatially refined $PM_{2.5}$ mass and speciated pollutant concentrations.

Overall, the inputs are the speciated ground-station data, satellite AOD snapshots and AOD grouped by aerosol type, and the CMAQ model simulations. The outputs are the fused ground-station, satellite, plus model $PM_{2.5}$ mass concentration field, and speciated versions of this field. A detailed description of the key steps follows.

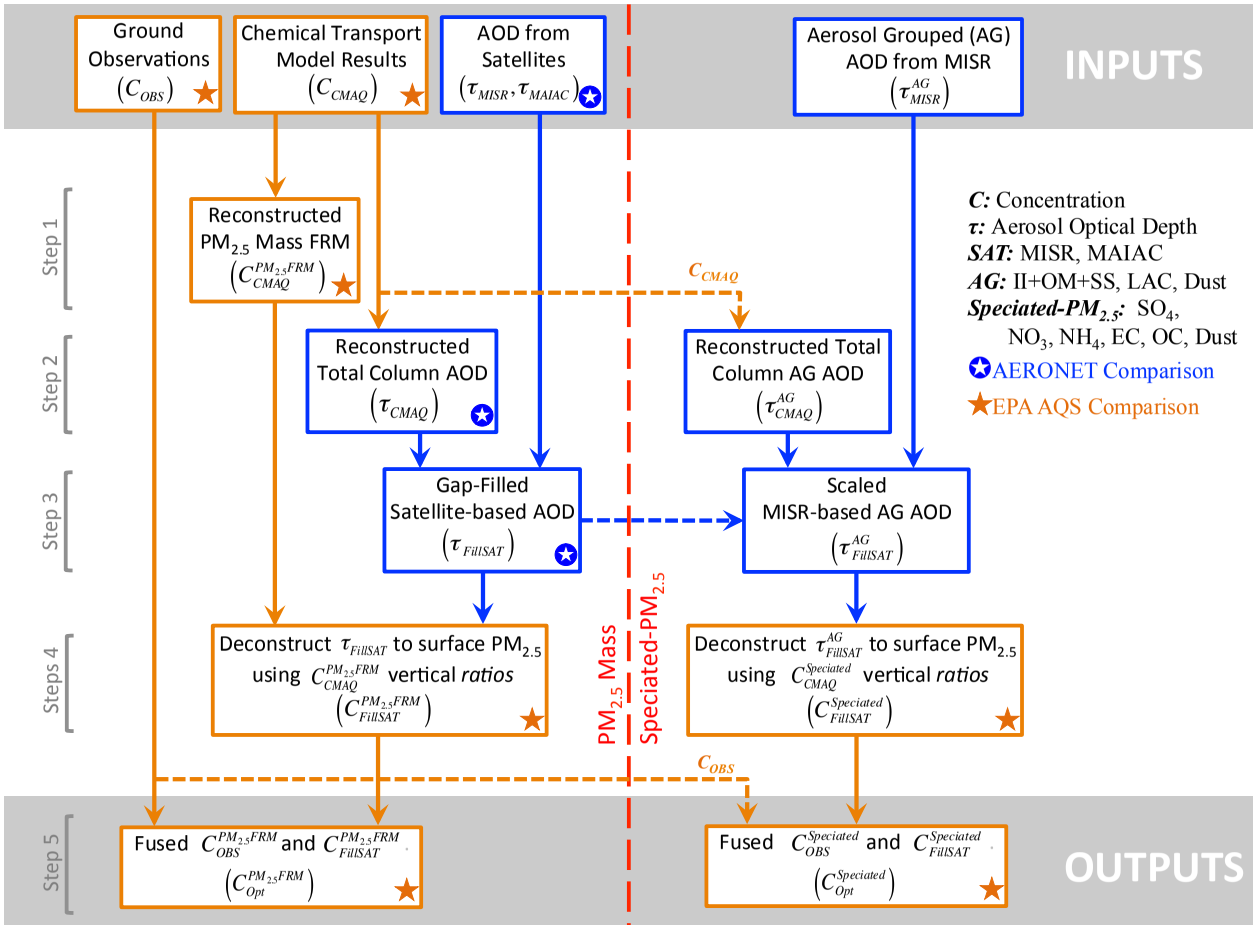

Figure 5.2 – Methods flow chart connecting satellite retrieved AOD to Modeled AOD, PM2.5 Mass, and PM2.5 Speciated Mass. The parenthetical terms are defined in their respective step in the methods section.

5.3.1 Step 1 – CMAQ- and Surface-derived $PM_{2.5}$ using reconstruction method

A commonly applied $PM_{2.5}$ mass reconstruction (RM) method, also termed mass closure or material balance, is used to compare the sum of major aerosol components to gravimetrically measured fine particulate matter. The approach also accounts for unmeasured or non-simulated species to avoid double counting. Beginning with Countess et al. (1980), the RM method is used to evaluate measurements, characterize spatiotemporal chemical gradients, estimate source contributions to PM, and calculate visibility impairment due to near-surface aerosol. Additionally, the reconstructed $PM_{2.5}$ mass provides insight into the spatial variations among the speciated data.[108, 137, 175-176] The development of this method, along with the differences between reconstructed and gravimetric mass in the CSN and IMPROVE data sets, have been extensively studied in the U.S.[108] Chow et al. (2015) provides a detailed literature review of the various mass reconstruction equations.

For the purposes of this study, the RM equation focuses on the following five representative chemical components, with the relevant references cited: (1) inorganic ions;[177-180] (2) organic matter;[137, 181] (3) Elemental Carbon (EC), also referred to as light absorbing carbon;[182] (4) crustal material, which includes mineral and soil particles, referred to herein as dust;[108, 183] (5) sea salt;[137] and (6) other elements ,[184] which, in the SJV during the study period, made a negligible contribution to PM. The respective references provide details as to how multipliers for each species were derived, and summarize the evaluation performed for each major PM component.

In addition to the measured aerosol species of interest, WRF-CMAQ model outputs for relative humidity, temperature, and speciated aerosol vertical distribution were used in the $PM_{2.5}$ mass reconstruction and as needed in the other analysis steps described hereafter. The RM method, excluding negligible "other" elements, was used to compare ground observations, CMAQ results, and satellite derived concentrations. Table C.1 in supplemental material provides a summary of the aerosol equations used for the ground monitor data and CMAQv5.0.2 simulations. The RM equation used is as follows (Chow et al., 2015, equation A):

$$RM[\mu g/m^3] = \underbrace{[SO_4^-] + [NH_4^+] + [NO_3^=]}_{Inorganic\ Ions} + \underbrace{1.8[OC]}_{\substack{Organic \\ Matter}} + \underbrace{[EC]}_{\substack{Light \\ Absorbing \\ Carbon}} + \underbrace{1.8[Cl^-]}_{\substack{Sea \\ Salt}} + \underbrace{2.2[Al] + 2.49[Si] + 1.63[Ca] + 1.94[Ti] + 2.42[Fe]}_{Dust} \quad (5.9)$$

For each of the major chemical components involved, Chow et al. 2015 covers in detail the factors and assumptions required for the RM calculation, and those contributing to the comparison with gravimetric mass measurements. These factors include the OM/OC ratio assumptions, carbon sampling and analysis artifacts, ammonium and nitrate volatilization, limitations of using chloride to estimate sea salt content, and water retention by hygroscopic species on filters.[179, 185-187]

Following the framework of equation 5.9, the reconstructed $PM_{2.5}$ mass does not account for the positive and negative factors that affect gravimetric and speciated measurements.[137, 175, 181, 188] To close the mass-balance difference between $PM_{2.5}$ FRM gravimetric mass and ambient mass (simulated and measured), the material balance (eq. 5.9) was adjusted to account for factors affecting gravimetric measurements (Frank, et al., 2006, eq. 10).

$$PM_{2.5FRM}\ [\mu g/m^3] = RM - ([\mathrm{NH_4^+}]_{loss} + [\mathrm{NO_3^-}]_{loss}) + [PBW] + [Blank_{FRM}] \quad (5.10)$$

where ammonium and nitrate volatilization are not captured by gravimetric measurements and thus, are accounted as negative artifacts. The particle bound water (PWB) is the water retained on the filter when particles are sampled and weighed for mass concentration. This concentration is dependent on ionic composition and relative humidity dependent species equilibrium prior to laboratory weighing. $Blank_{FRM}$ accounts for the passively collected mass value on "blank" filters. The limitations and uncertainties of the reconstruction method broken down by major chemical components are discussed in detail elsewhere.[175, 188] The uncertainty estimated for the CMAQ- and satellite-based surface concentrations are discussed in Section 5.4.

5.3.2 Step 2 – CMAQ-based columnar AOD and AOD subcategorized into species-related groups derived using the reconstructed extinction coefficient method

Section 5.3.1 summarized the method applied to calculate the five representative component surface mass concentrations from the surface observations; these components are also used to derive total-column aerosol optical depth from CMAQ (τ_{CMAQ}). First proposed by Malm et al. (1994), the reconstructed extinction coefficient method was designed to investigate the spatial and temporal variability of haze and visibility impairment in the U.S. as part of IMPROVE. Since then this method has been continuously upgraded by several researchers.[108, 183, 189-191] The process estimates extinction AOD using simulated concentrations of II, OM, SS, LAC, and Dust (Table C.1) assuming externally mixed aerosols with respect to the modeled altitudes (z), as follows:

$$\tau = \int \left\{ \underbrace{\sum_i \omega_i \beta_{de,i} f_{RH(z),i} C_{(z)i}}_{particle\ scattering\ efficiency} + \underbrace{\sum_i (1-\omega_i) \beta_{de,i} f_{RH(z),i} C_{(z)i})}_{particle\ absorption\ efficiency} \right\} dz \quad (5.11)$$

where

τ = aerosol extinction optical depth (AOD) at 550 nm

i = chemical component

ω = single scattering albedo (SSA)

β_{de} = specific dry extinction efficiency per mass [m^2/g]

f_{rh} = hygroscopic growth factors as a function of height

C = concentration of chemical component i as a function of height [$\mu g/m^3$]

Equation 5.11 is further subdivided for dust by size in accordance with the CMAQ Aitken, accumulation, and course particles size categories.[191] The empirically

based factors and their respective literature sources are summarized in Table 5.3. The WRF relative humidity, *rh(z)*, simulations were used to evaluate the height-dependent hygroscopic growth factors. The ambient particle extinction per unit length is the sum of the ambient scattering and absorption per unit length, which are the two terms in equation 5.11. When integrated over a horizontal path, the extinction per unit length is sometimes called the visibility, typically reported in [Mm^{-1}]. From equation 5.11, the dimensionless extinction AOD is obtained by multiplying the ambient particle extinction per unit length by the vertical atmospheric path length of each CMAQ layer. These are added vertically to obtain columnar AOD, which are compared to ground- and satellite-based AOD values in the following subsections to assess uncertainties.

Table 5.3 – The empirically based factors and their respective literature sources (eq. 11).

Species	Specific dry efficiency factor	Hygroscopic growth factor f (RH)=bscat(wet)/bscat(dry)	Single scattering albedo (ω)
Inorganic Ions	3 Malm et al. 2007; Chow et al. 2015	Malm et al., 2011; Song et al., 2008	1
Organic Matter	4 Malm et al. 2007; Chow et al. 2015	Zamora and Jacobson, 2013	ω_{OM}=.99 Sun et al.,2007; Bond et al., 2006
Sea Salt	1.37 Malm et al. 2007;Chow et al. 2015	Park et al., 2014	ω_{SS}=1
Light Absorbing Carbon	10 Malm et al. 2007; Chow et al. 2015	0	ω_{LAC}=Temp*0.928 (Conant et al., 2003); αabs=7.5±1.2 m^2/g^{-1} (Bond and Bergstrom, 2006)
Dust	stratifying by size (Park et al., 2011); βi ~regressed value (Tegen et al., 1996)	0	regressed value (Tegen et al., 1996)

The three CMAQ-based AOD AG (i.e., LAC, II+OM+SS, and Dust), indicated in Table 5.2, are calculated using the five major chemical components derived in the RM equation (eq. 5.9). The CMAQ-based total-column AOD AG aggregate is equivalent to the total-column CMAQ-based AOD. Assessment of the uncertainties in these

quantities, using a combination of ground-based and satellite total-column measurements, is given in Section 5.3.4 below.

5.3.3 Step 3 – Gap-Filled Satellite-derived AOD and Grouped AOD, using Scaled CMAQ-based AOD

To obtain a spatially complete AOD map for each case-study day, we combine the MISR-retrieved, MAIAC-retrieved, and CMAQ-based reconstructed AOD products, as CMAQ can simulate values in all grid boxes, regardless of cloud cover, surface brightness, terrain, and aerosol optical thickness. The most relevant factor affecting spatially complete satellite-retrieved AOD in this study are missing retrievals due to the presence of clouds. The combined AOD product, referred herein as $\tau_{FillSAT}$, is more complete than the MISR or MAIAC products alone.

A unique component of this work involves the use of the MISR RA aerosol species-specific groups. The MISR-RA as well as the MODIS-based MAIAC satellite-retrieved products were combined to improve spatial coverage. Before combining retrieved AOD products, the MAIAC maps were downscaled and spatially interpolated (via bilinear interpolation) to match the downscaled model 275 m × 275 m grid, and the 275 m × 275 m MISR maps were also re-gridded to match the downscaled model grid. On days when multiple Aqua and Terra MAIAC C6v2 1-km AOD retrievals were available, the MAIAC-Aqua AOD retrievals were used to fill in missing AOD in the MAIAC-Terra AOD maps closest in time to the MISR RA retrieval by linearly regressing values from a 15 x 15 grid cell region centered on the missing cell value. The scatterplots

in Figure 5.3 show MISR RA AOD retrievals are higher than those retrieved by MAIAC, and much closer to the AERONET values, for the three case study days with highest AOD. Figure 5.3 reinforces the need to scale MAIAC-retrieved AOD before gap-filling MISR-retrieved AOD fields. Based on Figure 5.3, a study-specific AOD filter with an upper bound of 0.4 was used for MAIAC retrievals to reduce potential cloud contamination. Within the SJV, larger gaps caused by cloud contamination in the satellite-retrieved AOD were filled with the model-reconstructed AOD using linearly regressed values from a 7 x 7 grid cell region, repeated multiple times as needed until the SJV study region was filled.

Additionally, we produce gap-filled, aerosol-type-grouped AODs from the original MISR-based AG AODs using the model-based grouped AODs from Step 1, and following the same procedure used for total AOD.

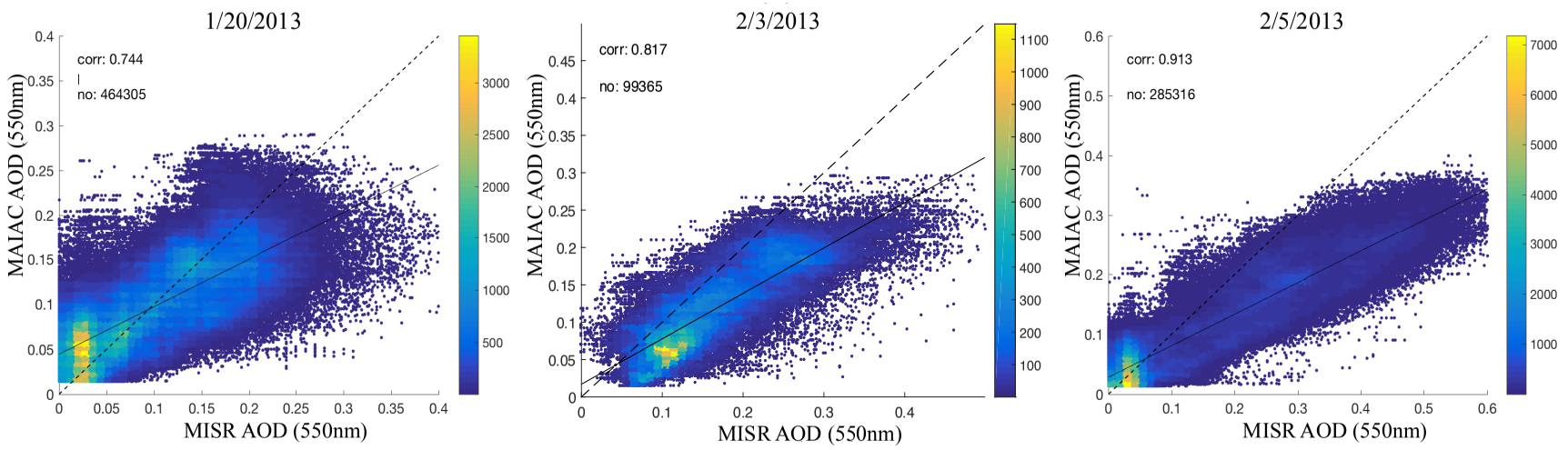

Figure 5.3 – Density scatterplots comparing MISR and MAIAC retrieved AOD for January 20th, February 3rd, and February 5th. A solid linear regression line and a 1:1 dashed line are shown for reference.

5.3.4 Uncertainty Estimates for Model-Reconstructed and Satellite Total-Column Quantities

Two sets of intermediate analyses are presented where surface-based *in situ* as well as column-integrated observations are provided as ground truth (i.e., their uncertainties are small compared to those of the other values used in this study). First, satellite-retrieved AOD snapshots are evaluated against coincident AERONET observations. Second, a comparison between daylight-averaged AERONET AOD data, satellite-retrieved AOD snapshots, and model-reconstructed diurnal AOD is presented to determine how well the snapshots represent diurnal values in the study region.

5.3.4.1 Comparison between Satellite-, CMAQ-reconstructed, and Ground-based Total-Column AOD Snapshots at Coincident Times

Evaluation of MISR RA[167] and MAIAC[171] AOD has been performed extensively before, but not specifically for the study region, where we have considerable ground-truth data. Overall, there were 14 AERONET sites over the SJV (Figure 5.1) during the six case study days. The number of coincident satellite- and ground-AOD observations is dependent on the swath width of each satellite instrument, the retrieval algorithm used, and the polar-orbiting coverage for a given day. Figure 5.4 and Table 5.4 provide scatterplots and a statistical summary, respectively, of AERONET AOD collocated in time and space with the MISR RA, MAIAC, gap-filled MISR AOD, and CMAQ results. Although AERONET reports AOD at 550 nm, AOD values at 558 nm were calculated for comparison with the MISR AOD retrievals. Only those Terra MAIAC AOD retrievals

that were temporally coincident with MISR RA retrievals were used in this comparison. A window of ± fifteen minutes was applied to select AERONET measurements as spatiotemporally coincident with the satellite overpass, and corresponding CMAQ hourly, reconstructed AOD values were used.

Table 5.4 – Summary Statistics for AERONET Coincidences with MISR and MAIAC within 15 minutes of Terra overpass time, for AOD≥0.15 and (all AOD).

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
1/18/13	MISR	- (1)	- (0.097)	- (-)	- (-)	- (0.017)	- (0.153)	- (-0.017)	- (0.017)	- (-0.153)	- (0.153)
	MAIAC	- (14)	- (0.08)	- (0.022)	- (0.261)	- (0.032)	- (0.314)	- (-0.022)	- (0.027)	- (-0.218)	- (0.267)
	FillSAT	- (1)	- (0.097)	- (-)	- (-)	- (0.019)	- (0.168)	- (-0.019)	- (0.019)	- (-0.168)	- (0.168)
	CMAQ	- (14)	- (0.07)	- (0.032)	- (0.63)	- (0.041)	- (0.398)	- (-0.033)	- (0.035)	- (-0.321)	- (0.344)
1/20/13	MISR	8 (11)	0.191 (0.172)	0.04 (0.048)	0.854 (0.913)	0.021 (0.019)	0.113 (0.113)	0.002 (0.004)	0.015 (0.014)	0.01 (0.021)	0.082 (0.082)
	MAIAC	9 (13)	0.161 (0.138)	0.049 (0.054)	0.334 (0.658)	0.056 (0.052)	0.288 (0.302)	-0.034 (-0.034)	0.04 (0.039)	-0.173 (-0.196)	0.206 (0.226)
	FillSAT	9 (13)	0.197 (0.174)	0.041 (0.05)	0.872 (0.923)	0.022 (0.019)	0.111 (0.109)	0.002 (0.002)	0.017 (0.015)	0.01 (0.01)	0.088 (0.085)
	CMAQ	9 (13)	0.133 (0.107)	0.043 (0.057)	0.625 (0.852)	0.07 (0.072)	0.358 (0.418)	-0.062 (-0.066)	0.063 (0.066)	-0.319 (-0.381)	0.324 (0.385)
2/3/13	MISR	1 (1)	0.242 (0.242)	- (-)	- (-)	0.02 (0.02)	0.078 (0.078)	-0.02 (-0.02)	0.02 (0.02)	-0.078 (-0.078)	0.078 (0.078)
	MAIAC	4 (6)	0.189 (0.167)	0.034 (0.044)	0.854 (0.89)	0.025 (0.021)	0.131 (0.123)	0 (-0.004)	0.025 (0.02)	-0.002 (-0.021)	0.13 (0.116)
	FillSAT	1 (1)	0.242 (0.242)	- (-)	- (-)	0.025 (0.025)	0.093 (0.093)	-0.025 (-0.025)	0.025 (0.025)	-0.093 (-0.093)	0.093 (0.093)
	CMAQ	6 (8)	0.121 (0.101)	0.107 (0.098)	0.971 (0.975)	0.166 (0.151)	0.608 (0.635)	-0.152 (-0.137)	0.152 (0.137)	-0.555 (-0.575)	0.555 (0.575)
2/5/13	MISR	8 (8)	0.377 (0.377)	0.128 (0.128)	0.99 (0.99)	0.045 (0.045)	0.134 (0.134)	0.04 (0.04)	0.04 (0.04)	0.119 (0.119)	0.119 (0.119)
	MAIAC	8 (9)	0.23 (0.207)	0.077 (0.1)	0.914 (0.927)	0.1 (0.1)	0.316 (0.338)	-0.088 (-0.089)	0.09 (0.091)	-0.277 (-0.302)	0.283 (0.307)
	FillSAT	11 (12)	0.37 (0.35)	0.114 (0.13)	0.977 (0.982)	0.04 (0.038)	0.117 (0.119)	0.032 (0.03)	0.032 (0.03)	0.096 (0.093)	0.096 (0.093)
	CMAQ	11 (12)	0.154 (0.151)	0.048 (0.047)	0.257 (0.31)	0.212 (0.203)	0.628 (0.635)	-0.185 (-0.169)	0.185 (0.169)	-0.546 (-0.528)	0.546 (0.529)
2/12/13	MISR	1 (8)	0.121 (0.107)	- (0.017)	- (0.392)	0.034 (0.022)	0.221 (0.196)	-0.034 (-0.005)	0.034 (0.02)	-0.221 (-0.042)	0.221 (0.18)
	MAIAC	1 (10)	0.115 (0.08)	- (0.027)	- (0.764)	0.043 (0.036)	0.272 (0.323)	-0.043 (-0.032)	0.043 (0.032)	-0.272 (-0.285)	0.272 (0.286)
	FillSAT	1 (9)	0.121 (0.103)	- (0.021)	- (0.223)	0.037 (0.028)	0.236 (0.244)	-0.037 (-0.012)	0.037 (0.024)	-0.236 (-0.107)	0.236 (0.209)
	CMAQ	1 (10)	0.08 (0.062)	- (0.015)	- (0.731)	0.078 (0.052)	0.494 (0.464)	-0.078 (-0.049)	0.078 (0.049)	-0.494 (-0.441)	0.494 (0.441)
2/14/13	MISR	- (7)	- (0.105)	- (0.022)	- (0.953)	- (0.01)	- (0.089)	- (-0.007)	- (0.008)	- (-0.062)	- (0.074)
	MAIAC	1 (9)	0.169 (0.086)	- (0.045)	- (0.898)	0.018 (0.036)	0.122 (0.317)	0.018 (-0.027)	0.018 (0.032)	0.122 (-0.243)	0.122 (0.279)
	FillSAT	1 (7)	0.14 (0.105)	- (0.022)	- (0.953)	0.011 (0.012)	0.075 (0.104)	-0.011 (-0.009)	0.011 (0.01)	-0.075 (-0.081)	0.075 (0.084)
	CMAQ	1 (9)	0.118 (0.122)	- (0.021)	- (-0.478)	0.033 (0.039)	0.217 (0.345)	-0.033 (0.009)	0.033 (0.033)	-0.217 (0.078)	0.217 (0.289)

Pearson Correlation (R); Modeled results (M); Observations (O); Root Mean Square Error (RMSE)= $\sqrt{\frac{1}{N}\sum_N (M-O)^2}$; Mean Bias (MB) = $\frac{1}{N}\sum_N (M-O)$; Mean Absolute Error (MAE) = $\frac{1}{N}\sum_N |M-O|$; Normalized RMSE (NRMSE) = $\frac{\sqrt{\frac{1}{N}\sum_N (M-O)^2}}{\frac{1}{N}\sum_N (O)}$; Normalized Mean Bias (NMB) = $\frac{\sum_N (M-O)}{\sum_N (O)}$; Normalized Mean Absolute Error by obs. (NME) = $\frac{\sum_N |M-O|}{\sum_N (O)}$.

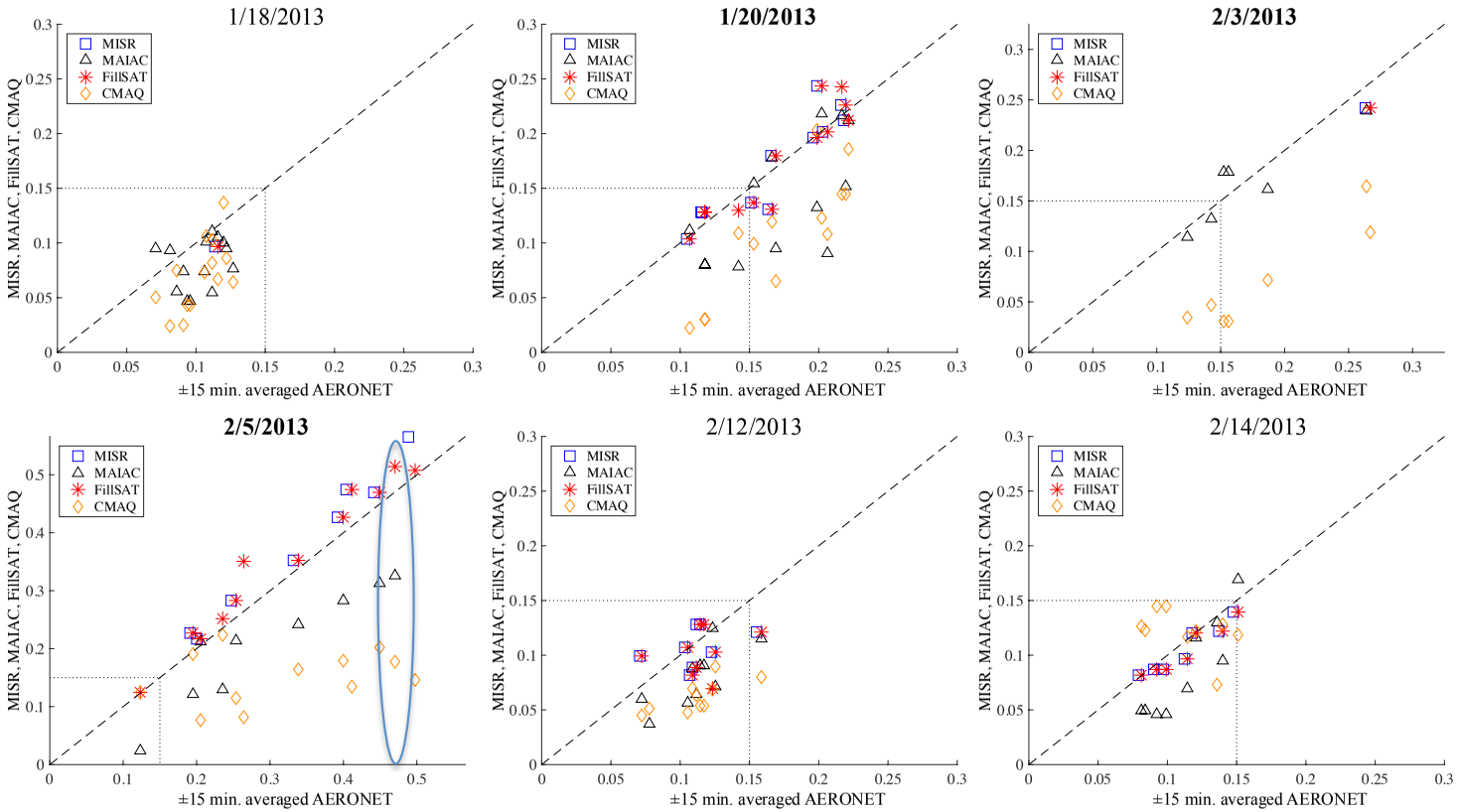

Figure 5.4 – Scatterplot comparison of AERONET coincidences with MISR, MAIAC, gap-filled MISR, and CMAQ results within ±15 minutes of Terra overpass time. The subplot for 2/5/2013 shows that gap-filled MISR offers better agreement than gap-filled MAIAC at the AERONET retrieved AOD value of 0.5 (circled in the lower left plot in the figure). The MAIAC and AERONET AOD comparison is plotted at 550 nm, while the MISR and AERONET AOD comparison is at 558 nm; the dotted line indicates the 0.15 AOD threshold; a 1:1 dashed line is shown for reference.

Overall, the MISR RA AOD compares well with coincident AERONET AOD, and tends to outperform MAIAC statistically over the SJV across all our case-study days (Table 5.4). The two best-case days for this analysis are January 20th and February 5th, where AERONET AOD values were relatively high (AOD≥0.15) and there were multiple coincident MISR retrievals across the region. On these days, MAIAC underestimates AOD compared to AERONET, whereas MISR RA slightly overestimates AOD. Specifically, for January 20th and February 5th, the MISR RA-to-AERONET AODs had an overall R of 0.91 and 0.99, and a NME of 0.08 and 0.12, respectively. For MAIAC, the corresponding values are an overall R of 0.66 and 0.93, and a NME of 0.23 and 0.31, respectively.

The comparison of MISR and MAIAC satellite-retrieved AODs with AERONET also illustrates how gap-filling MISR with scaled and gap-filled MAIAC retrievals produces a more consistent product. For example, the Figure 5.4 subplot for February 5th shows that gap-filled MISR offers better agreement than gap-filled MAIAC at the AERONET retrieved AOD value of 0.5 (circled in the lower left plot in the figure). On this specific day and location there is no coincident MISR retrieval, indicating that the gap-filled MISR improvement is due to scaled and gap-filled MAIAC used to gap-fill the MISR AOD snapshot. Further evident from Figure 5.4, the CMAQ reconstructed values systematically underestimate AOD relative to AERONET in nearly all cases and exhibit greater scatter, hinting at the possible value of applying the measurements as constraints on the model simulations.

5.3.4.2 Comparison of Satellite-based AOD Snapshots with Daylight-average Ground-based AOD and with Daylight- and Diurnal-average Model-based AOD

Unlike aerosol radiative forcing, which depends on daytime solar heating, conditions during the full diurnal cycle are relevant for many air quality applications. However, AERONET as well as the satellites acquire AOD data only during daylight hours, when the sun is well above the horizon. To test the feasibility of using satellite-based AOD snapshot retrievals as proxies for AOD averaged over daylight hours, we compare the satellite retrievals (MISR, MAIAC, gap-filled MISR) the with daylight-averaged AERONET-retrieved AOD results (Figure 5.5). We subsequently compare the model daylight- and diurnal-average AODs, as well as the AERONET daylight-average AODs, with the respective short-term values from these data sources (Figure 5.6), to assess how well snapshot values represent AOD for entire days in the study region. In places where the snapshots are substantially different from the daylight-average or diurnal-average AOD values, scaled model results would be required to complete the diurnal air quality picture.

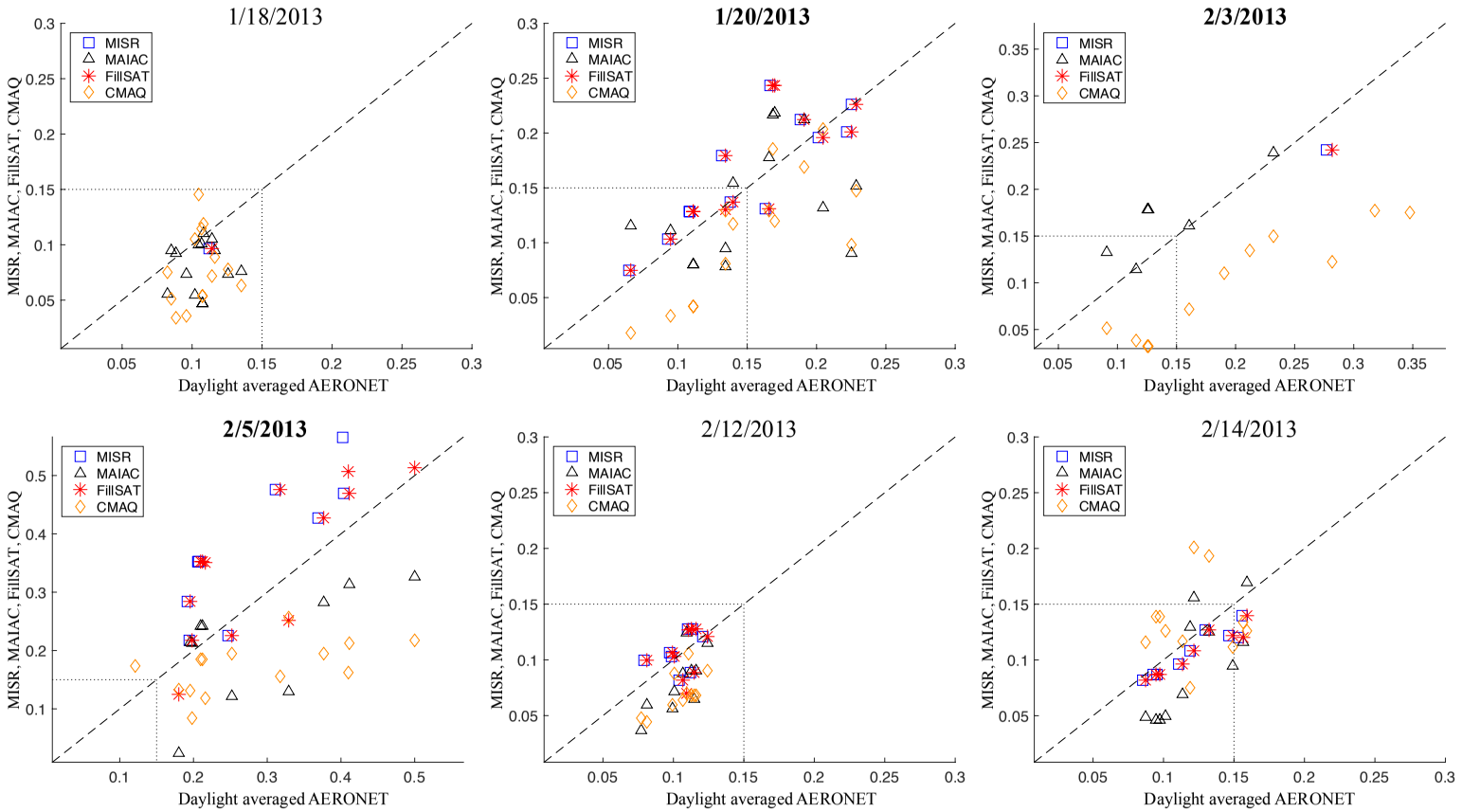

Figure 5.5 – Snapshot vs. daylight-average AOD. Scatter plot comparison of MISR, MAIAC, gap-filled MISR snapshots with AERONET and CMAQ daylight-averaged results. The MAIAC and AERONET AOD comparison is plotted at 550 nm, while the MISR and AERONET AOD comparison is at 558 nm; the dotted line indicates the 0.15 AOD threshold; a 1:1 dashed line is shown for reference.

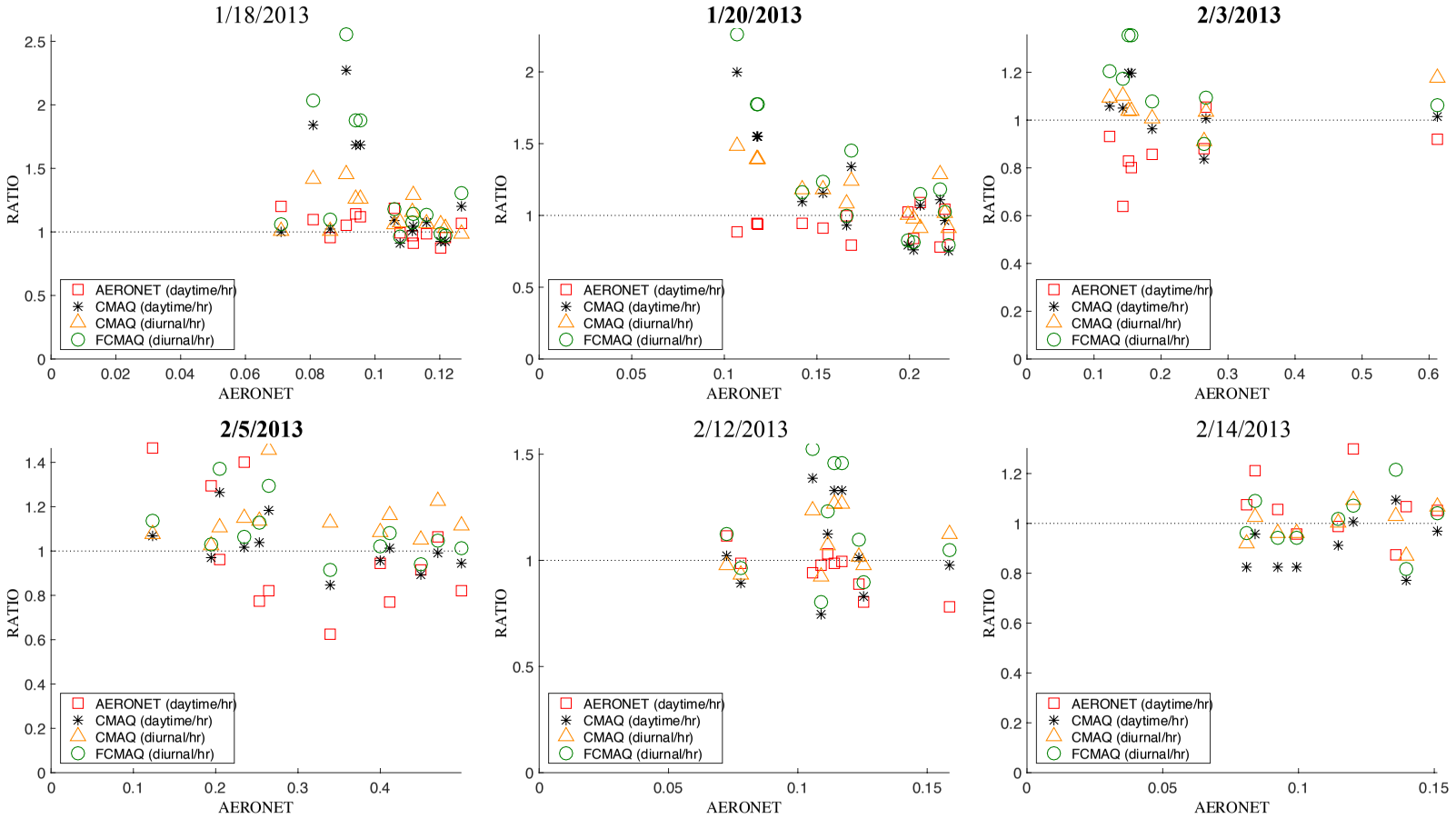

Figure 5.6 – Scatterplot of daylight-averages to the Terra overpass time ratios versus AERONET AOD retrievals within ±15 minutes of Terra overpass time. Two ratios are shown for CMAQ: daytime average-to-hour ratio and diurnal average-to-hour ratio. The FCMAQ ratio show is the FCMAQ diurnal to CMAQ hour values. The dashed unity line is shown for reference.

For the initial comparison, all retrieved AERONET values per each of the six case-study days were averaged to obtain a daylight average at each of the 14 sites (Figure 5.5). For the MISR comparison, we have only the same MISR RA AOD retrieval snapshots as in Figure 5.4. For the study cases, MAIAC can have multiple Terra and Aqua retrievals over the region during one day, occurring at different times, due to the wide MODIS swath. As such, MAIAC Terra-retrieved AOD "coincident" with MISR overpasses are in some cases gap-filled with other scaled-MAIAC Terra/Aqua retrievals acquired during that day (see section 5.3.3 above). The third satellite-retrieved AOD product shown in Figure 5.5 is the gap-filled, primarily MISR-derived AOD (*FillSAT*) described in section 5.3.3. Also shown in Figure 5.5 are the CMAQ reconstructed daylight-average AODs, described in section 5.3.2.

Overall, the MISR and FillSAT values are very nearly identical, and they tend to serve as better proxies for the daylight-average AERONET values than CMAQ for the study cases shown in Figure 5.5. Table 5.5 contains a statistical summary of the scatterplot data. For the two best days of January 20th and February 5th, the retrieved AODs for MISR RA and gap-filled MISR agree better statistically than the other datasets in terms of correlation and errors relative to AERONET daylight-average values. Although the retrieved AODs for the MISR RA and gap-filled MISR slightly outperform MAIAC for the specific case study days, this relationship is likely to change for different domains and time periods. As such, the technique for gap-filling MISR AOD might need to be dynamic in weighting the MAIAC AOD retrievals when applied to other regions. For January 20th and February 5th, the gap-filled MISR-to-daylight-average-AERONET

AODs had overall R values of 0.81 and 0.78, and NME of 0.16 and 0.28, respectively. This comparison indicates the satellite-retrieved AOD quantities are in sufficient agreement with daylight-averaged ground truth to serve as proxies for the daylight-averaged values during the study period.

Table 5.5 – Summary Statistics for AERONET Coincidences with all datasets using daylight average AOD≥0.15 (All AOD).

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
1/18/13	MISR	- (1)	- (0.097)	- (-)	- (-)	- (0.016)	- (0.141)	- (-0.016)	- (0.016)	- (-0.141)	- (0.141)
	MAIAC	- (14)	- (0.08)	- (0.022)	- (0.076)	- (0.036)	- (0.336)	- (-0.025)	- (0.028)	- (-0.24)	- (0.261)
	FillSAT	- (1)	- (0.097)	- (-)	- (-)	- (0.018)	- (0.156)	- (-0.018)	- (0.018)	- (-0.156)	- (0.156)
	CMAQ	- (14)	- (0.078)	- (0.033)	- (0.206)	- (0.043)	- (0.403)	- (-0.028)	- (0.037)	- (-0.265)	- (0.348)
1/20/13	MISR	6 (12)	0.202 (0.163)	0.039 (0.053)	0.326 (0.847)	0.036 (0.031)	0.188 (0.203)	0.007 (0.013)	0.027 (0.022)	0.038 (0.084)	0.137 (0.148)
	MAIAC	7 (14)	0.171 (0.137)	0.049 (0.052)	-0.785 (0.4)	0.07 (0.056)	0.363 (0.366)	-0.022 (-0.016)	0.059 (0.046)	-0.114 (-0.107)	0.305 (0.304)
	FillSAT	7 (14)	0.208 (0.167)	0.039 (0.054)	0.114 (0.81)	0.044 (0.034)	0.225 (0.222)	0.014 (0.014)	0.034 (0.025)	0.074 (0.088)	0.176 (0.16)
	CMAQ	7 (14)	0.151 (0.108)	0.038 (0.059)	-0.127 (0.776)	0.062 (0.057)	0.323 (0.375)	-0.043 (-0.045)	0.048 (0.047)	-0.221 (-0.293)	0.247 (0.309)
2/3/13	MISR	1 (1)	0.242 (0.242)	- (-)	- (-)	0.034 (0.034)	0.123 (0.123)	-0.034 (-0.034)	0.034 (0.034)	-0.123 (-0.123)	0.123 (0.123)
	MAIAC	2 (6)	0.2 (0.167)	0.055 (0.044)	- (0.85)	0.005 (0.035)	0.026 (0.247)	0.004 (0.026)	0.004 (0.026)	0.02 (0.181)	0.02 (0.184)
	FillSAT	1 (1)	0.242 (0.242)	- (-)	- (-)	0.039 (0.039)	0.139 (0.139)	-0.039 (-0.039)	0.039 (0.039)	-0.139 (-0.139)	0.139 (0.139)
	CMAQ	10 (14)	0.197 (0.152)	0.108 (0.117)	0.954 (0.965)	0.175 (0.154)	0.501 (0.545)	-0.152 (-0.13)	0.152 (0.13)	-0.435 (-0.462)	0.435 (0.462)
2/5/13	MISR	9 (9)	0.374 (0.374)	0.12 (0.12)	0.832 (0.832)	0.113 (0.113)	0.401 (0.401)	0.093 (0.093)	0.098 (0.098)	0.332 (0.332)	0.348 (0.348)
	MAIAC	10 (10)	0.211 (0.211)	0.095 (0.095)	0.617 (0.617)	0.115 (0.115)	0.402 (0.402)	-0.076 (-0.076)	0.095 (0.095)	-0.265 (-0.265)	0.33 (0.33)
	FillSAT	13 (13)	0.35 (0.35)	0.124 (0.124)	0.777 (0.777)	0.095 (0.095)	0.323 (0.323)	0.057 (0.057)	0.082 (0.082)	0.195 (0.195)	0.278 (0.278)
	CMAQ	13 (14)	0.172 (0.172)	0.047 (0.045)	0.603 (0.544)	0.146 (0.142)	0.5 (0.505)	-0.121 (-0.109)	0.121 (0.116)	-0.414 (-0.388)	0.414 (0.415)
2/12/13	MISR	- (8)	- (0.107)	- (0.017)	- (0.391)	- (0.016)	- (0.156)	- (0.003)	- (0.014)	- (0.025)	- (0.135)
	MAIAC	- (10)	- (0.08)	- (0.027)	- (0.752)	- (0.03)	- (0.289)	- (-0.024)	- (0.027)	- (-0.234)	- (0.263)
	FillSAT	- (10)	- (0.105)	- (0.021)	- (0.286)	- (0.02)	- (0.184)	- (-0.002)	- (0.017)	- (-0.021)	- (0.154)
	CMAQ	- (11)	- (0.07)	- (0.018)	- (0.668)	- (0.037)	- (0.352)	- (-0.034)	- (0.034)	- (-0.329)	- (0.329)
2/14/13	MISR	2 (9)	0.13 (0.108)	0.014 (0.021)	- (0.941)	0.026 (0.017)	0.17 (0.137)	-0.025 (-0.013)	0.025 (0.013)	-0.161 (-0.11)	0.161 (0.11)
	MAIAC	2 (11)	0.143 (0.096)	0.038 (0.046)	- (0.759)	0.029 (0.039)	0.186 (0.326)	-0.015 (-0.025)	0.025 (0.036)	-0.095 (-0.21)	0.161 (0.294)
	FillSAT	2 (9)	0.13 (0.108)	0.014 (0.021)	- (0.942)	0.029 (0.019)	0.185 (0.152)	-0.028 (-0.016)	0.028 (0.016)	-0.177 (-0.128)	0.177 (0.128)
	CMAQ	2 (11)	0.13 (0.134)	0.005 (0.036)	- (0.058)	0.028 (0.043)	0.178 (0.354)	-0.028 (0.013)	0.028 (0.038)	-0.176 (0.11)	0.176 (0.316)

*MISR, MAIAC, and SAT datasets are not daytime averages, but Terra overpass backfilled values.

A procedure for fusing CMAQ model simulations with surface-based measurements is described briefly in Section 5.2.4.1 above, and in detail in Friberg et al. (2016). This procedure was applied to C_{OBS} and C_{CMAQ} (Figure 5.2) to produce C_{FCMAQ}, also referred to as FCMAQ. The additional step allows us to assess how the spatially extensive satellite data affects the results compared to the model constrained only by local surface observations.

To estimate how well the AOD snapshots might characterize the diurnal-average AOD, diurnal-to-hourly ratios for CMAQ and FCMAQ are plotted against AERONET retrieved AODs acquired within 15 minutes of the satellite overpasses for each case (Figure 5.6 and Table 5.6). AERONET ratios are plotted as well. The diurnal model and daylight AERONET AOD values are divided by AODs at Terra overpass time within the hour and within 15 minutes for the model and AERONET ratios, respectively. On January 18th and 20th, FCMAQ and daytime CMAQ ratios exhibit the high variability at locations where AERONET ratios were near unity, suggesting that CMAQ diurnal-to-hour ratio are at times spatially biased. But generally, based on model performance, snapshots acquired at Terra overpass time tend to fall within 10% - 20% of the diurnal-average value, except in some cases when the AOD at overpass time <~0.15 or 0.2. At these smaller AODs, a small absolute change in AOD will produce larger percent changes.

Table 5.6 – Summary of daytime or diurnal to Terra overpass hour ratios at AERONET locations for AOD≥0.15 (All AOD).

Date	Dataset	Ratio
1/18/13	AERONET (daytime-to-hr)	- (1.038)
	CMAQ (daytime-to-hr)	- (1.264)
	CMAQ (diurnal-to-hr)	- (1.152)
	FCMAQ (diurnal-to-hr)	- (1.376)
1/20/13	AERONET (daytime-to-hr)	0.928 (0.928)
	CMAQ (daytime-to-hr)	0.985 (1.159)
	CMAQ (diurnal-to-hr)	1.068 (1.158)
	FCMAQ (diurnal-to-hr)	1.052 (1.264)
2/3/13	AERONET (daytime-to-hr)	0.89 (0.864)
	CMAQ (daytime-to-hr)	1.037 (1.041)
	CMAQ (diurnal-to-hr)	1.034 (1.05)
	FCMAQ (diurnal-to-hr)	1.141 (1.153)
2/5/13	AERONET (daytime-to-hr)	0.945 (0.988)
	CMAQ (daytime-to-hr)	1.011 (1.015)
	CMAQ (diurnal-to-hr)	1.149 (1.143)
	FCMAQ (diurnal-to-hr)	1.082 (1.087)
2/12/13	AERONET (daytime-to-hr)	0.783 (0.95)
	CMAQ (daytime-to-hr)	0.978 (1.065)
	CMAQ (diurnal-to-hr)	1.123 (1.08)
	FCMAQ (diurnal-to-hr)	1.049 (1.161)
2/14/13	AERONET (daytime-to-hr)	1.053 (1.064)
	CMAQ (daytime-to-hr)	0.968 (0.91)
	CMAQ (diurnal-to-hr)	1.069 (0.992)
	FCMAQ (diurnal-to-hr)	1.04 (1.011)

One possible reason for the scatter in Figure 5.6 is the model representation of transported aerosol. Transported aerosol above the boundary layer is not always well represented by the CMAQ model in this region. For example, the model results indicate minimal vertical distribution of dust aerosol, concentrating all the dust within the planetary boundary layer on the study days, whereas transported dust above the boundary layer is likely to be the major non-spherical aerosol species in this region and season

(e.g., Liu et al., 2007b). Any biases in dust AOD retrievals are compounded by inaccuracies in the model-based vertical distributions that are applied during the total-column-to-surface decomposition step. The impact of errors in the adopted vertical distribution of aerosols on these results, beyond the scope of the current paper, needs to be investigated further. Model aerosol vertical distribution can be further constrained by taking advantage of upwind aerosol elevation retrievals from space-based stereo imaging (MISR), in places where the aerosol sources produce visible plumes, and downwind aerosol layer heights from space-based lidar (e.g., CALIPSO).[192]

5.3.5 Steps 4 – Deconstructed Total-column Satellite-measured AOD to Surface $PM_{2.5}$ Mass and Speciated Concentrations

Using CMAQ-based aerosol vertical profiles, near-surface concentrations ($Surface\ C_{FillSAT}^{PM_{2.5}FRM}$ and $C_{FillSAT}^{Speciated}$) are obtained from the total column satellite AOD ($\tau_{FillSAT}$) and aerosol group AOD ($\tau_{FillSAT}^{AG}$) by the following three intermediate steps. As in previous work, the key step amounts to using model-derived ratios of total-column to near-surface aerosol distributions to obtain near-surface values constrained by total-column measurements (e.g., Liu et al., 2004; Van Donkelaar et al., 2010).

In equation 5.12, the total column dry particle concentrations for the three aerosol groups ($Columnar\ C_{FillSAT}^{AG}$) are calculated from the AODs, $\tau_{FillSAT}$ and $\tau_{FillSAT}^{AG}$, by reversing the reconstructed extinction process in Step 2 (eq. 11). The same hygroscopic growth and specific dry scattering or absorbing efficiency factors are used here for consistency. The total column satellite-based AG concentrations ($Columnar\ C_{FillSAT}^{AG}$)

are further stratified into the five total column representative PM chemical components ($Columnar\ C_{FillSAT}^{Speciated}$), defined in Step 1 according to equation 5.9, using the CMAQ-based species-to-aerosol group partition (eq. 5.13). With $Columnar\ C_{FillSAT}^{Speciated}$ defined, satellite-based total column $PM_{2.5}$ ($Columnar\ C_{FillSAT}^{PM_{2.5}FRM}$) is obtained using equation 5.10. The satellite-derived total column concentrations are then apportioned to surface-level concentrations by relying on the vertical distribution of the CMAQ simulations of each species (eq. 5.14). These relationships were defined in terms of daily AOD and species concentrations.

$$Columnar\ C_{FillSAT}^{AG} = \frac{\tau_{FillSAT}^{AG}}{\beta_{de,i} f_{RH,i} dz} \tag{5.12}$$

$$Columnar\ C_{FillSAT}^{Speciated} = Columnar\ C_{FillSAT}^{AG} \left(\frac{Columnar\ C_{CMAQ}^{Speciated}}{Columnar\ C_{CMAQ}^{AG}} \right) \tag{5.13}$$

$$Surface\ C_{FillSAT}^{Speciated} = Columnar\ C_{FillSAT}^{Speciated} \left(\frac{Surface\ C_{CMAQ}^{Speciated}}{Columnar\ C_{CMAQ}^{Speciated}} \right) \tag{5.14}$$

5.3.6 Step 5 – Optimized $PM_{2.5}$ FRM and speciated concentrations by fusing satellite-constrained values with ground-station data

The optimized concentration dataset (C_{Opt}) closely parallels the constrained CMAQ simulation using ground-based observations described in Section 5.2.4.1 (eq. 5.1). The C_{Opt} dataset is derived by constraining the results with the surface-station data

near their locations, and weighting the satellite-constrained concentration values progressively more heavily away from available ground stations.

Using equation 5.15, the six daily C_{CMAQ} fields in the two-month span are replaced with the satellite-derived daily $C_{FillSAT}$ fields. With only 11.5% of the C_{CMAQ} fields changing, the average temporal correlations between the simulations and observations across all monitors are not recalculated (R_2; eq. 5.4), nor are the weighting factors (W; eq. 5.2). Thus, for this study, C_{Opt} diverges from C_{FCMAQ} for 6 days out of the entire 2-month time period.

$$C_{Opt_{s,t}} = \alpha \overline{C_{CMAQ_s}}^{\beta} \left[W_{s,t} \left\{ \frac{C_{OBS_{s_m,t}}}{\overline{C_{OBS_{s_m}}}} \right\}_{krig} + \left(1 - W_{s,t}\right) \left\{ \frac{C_{FillSAT_{s,t}}}{\overline{C_{CMAQ_s}}} \right\} \right] \qquad (5.15)$$

Using the techniques described in the next section, we assess the performance of the optimized surface concentrations in the results section.

5.3.7 Evaluation of Optimized Datasets by Cross-Validation

Three cross-validation techniques are used to evaluate how well the optimized datasets represent diurnal values to elucidate biases that arise from different sampling frequencies and spatial distribution of monitors across the pollutants. First, a tenfold 10% withholding (10-WH) technique is applied to all species. Then a Leave-One-Out (LOO) cross-validation method is used for all the species with the exception of $PM_{2.5}$. Finally, a Regional Holdout (RH) is used only for $PM_{2.5}$.

5.3.7.1 10-fold 10% Holdout Cross-Validation

The dataset performance was evaluated using a tenfold 10% withholding cross-validation analysis. For each of 10 independently run trials, a random 10% of the observations were held back per day and each method (i.e., fused and optimized) was applied to simulate the withheld data. The results from the 10 trials were then combined to provide cross-validation results that allow for the exploration of difference in errors based on proximity to monitors. Across monitors and days, the holdout number corresponds to the number of observations for each pollutant (Table 5.1), ranging from 44 for $PM_{2.5}$-OC to 779 for $PM_{2.5}$.

5.3.7.2 Leave-one-out Cross-Validation

As an alternative to the 10-WH method, the LOO withholding is applied to the five particulate matter components to overcome the sampling and spatial scarcity. By withholding one location at a time, this location-based cross-validation technique can provide information on how well the CMAQ simulations and satellite-derived concentrations of the fused and optimized datasets, respectively, represent diurnal values at locations further than 50 km from other monitors (see speciated monitor locations in Figure 5.1). With some sites containing more than one monitor, collocated monitors were considered one location and, thus, all monitors at a location were withheld for LOO. This cross-validation technique does not provide much insight when the nearest monitor is in close proximity, as is the case with the $PM_{2.5}$ mass monitors.

5.3.7.3 Regional Holdout Cross-Validation

A regional withholding technique is used to evaluate fused and optimized $PM_{2.5}$ datasets affected by monitor clustering. For each of the cross-validation regions in Figure 5.1, all but one of the monitors in a region is withheld and is repeated independently for each daily monitor and region. This approach approximates the evaluation of LOO when the distance between monitor locations is large (i.e., >50 km).

5.4 Results

Two sets of analyses are presented where surface-based *in situ* observations are provided as ground truth (i.e., their uncertainties are small compared to those of the other values used in this study). First, modeled and deconstructed satellite-derived results for $PM_{2.5}$ and $PM_{2.5}$ grouped by species are evaluated against EPA AQS and CSN ground observations. The main objectives of this section are (1) to evaluate the results of Steps 2-5 as much as possible, (for Step 1, see Friberg et al. 2017), (2) to assess where, and to what degree, the satellite data help constrain the model $PM_{2.5}$ over an extended region, and (3) where mid-visible AOD values exceed 0.15, to also evaluate the satellite-constrained, speciated $PM_{2.5}$. For the second set of analyses, cross-validation is used to evaluate satellite-based model capabilities.

5.4.1 Comparison of Satellite-derived and Model-based Daily $PM_{2.5}$ and Speciated Component Surface Concentrations to Average Daily Ground Truth

We compare now the model-based (C_{CMAQ}, C_{FCMAQ}) and the intermediate satellite-derived ($C_{FillSAT}$) daily averaged $PM_{2.5}$ and speciated component concentrations (Figure 5.2) with EPA AQS and CSN observations. Table 5.7 provides a statistical summary of

the comparison between the ground truth and the modeled, fused, and satellite-constrained results, stratified by pollutant, day, and dataset. Figure 5.7 presents concentration maps with embedded ground truth $PM_{2.5}$ values and their respective RGB images for the three days with relatively high AOD in the study set (January 20th, February 3rd, and February 5th).

Table 5.7 – Statistical summary of comparison between AQS or CSN daily concentrations coincidences and each modeled or satellite derived dataset stratified by pollutant and day.

PM2.5_FRM

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
1/18/13	CMAQ	8	36.33	10.84	0.19	12.27	0.34	0.23	9.01	0.01	0.25
	FCMAQ	8	38.35	9.65	0.11	12.30	0.34	2.25	8.60	0.06	0.24
	FillSAT	8	29.33	5.25	0.89	8.56	0.24	-6.77	7.18	-0.19	0.20
	Opt	8	38.96	5.91	0.85	6.03	0.17	2.87	5.27	0.08	0.15
1/20/13	CMAQ	12	25.09	15.07	0.82	13.42	0.39	-9.36	10.76	-0.27	0.31
	FCMAQ	12	27.18	15.41	0.80	12.50	0.36	-7.27	10.59	-0.21	0.31
	FillSAT	12	23.19	14.23	0.86	14.28	0.41	-11.27	12.21	-0.33	0.35
	Opt	12	32.40	14.49	0.95	6.02	0.17	-2.06	4.57	-0.06	0.13
2/3/13	CMAQ	8	21.97	2.35	0.30	4.40	0.21	1.34	4.04	0.06	0.20
	FCMAQ	8	23.63	3.50	0.74	4.16	0.20	2.99	3.83	0.15	0.19
	FillSAT	8	23.09	9.32	0.72	6.85	0.33	2.46	5.26	0.12	0.26
	Opt	8	22.71	5.55	0.74	4.08	0.20	2.08	3.19	0.10	0.15
2/5/13	CMAQ	8	22.20	13.56	0.94	6.55	0.24	-4.74	4.76	-0.18	0.18
	FCMAQ	8	24.49	13.90	0.92	5.62	0.21	-2.44	4.35	-0.09	0.16
	FillSAT	8	38.55	27.23	0.59	23.75	0.88	11.62	16.70	0.43	0.62
	Opt	8	26.50	8.16	0.73	8.80	0.33	-0.43	6.10	-0.02	0.23
2/12/13	CMAQ	20	14.86	4.42	0.58	4.70	0.28	-1.89	3.73	-0.11	0.22
	FCMAQ	20	16.04	4.19	0.53	4.51	0.27	-0.71	3.46	-0.04	0.21
	FillSAT	20	11.82	3.72	0.19	7.45	0.44	-4.93	6.41	-0.29	0.38
	Opt	20	16.42	5.76	0.84	3.10	0.19	-0.33	2.38	-0.02	0.14
2/14/13	CMAQ	7	24.25	9.94	0.92	4.92	0.22	2.19	4.76	0.10	0.22
	FCMAQ	7	26.29	10.16	0.90	6.27	0.28	4.23	5.76	0.19	0.26
	FillSAT	7	11.50	5.34	0.88	12.58	0.57	-10.56	11.96	-0.48	0.54
	Opt	7	17.42	11.59	0.98	5.10	0.23	-4.64	4.64	-0.21	0.21

NH4

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
2/3/13	CMAQ	1	2.80	-	-	1.03	0.58	1.03	1.03	0.58	0.58
	FCMAQ	1	3.60	-	-	1.83	1.03	1.83	1.83	1.03	1.03
	FillSAT	1	2.58	-	-	0.81	0.46	0.81	0.81	0.46	0.46
	Opt	1	4.84	-	-	3.07	1.74	3.07	3.07	1.74	1.74
2/12/13	CMAQ	3	1.96	0.29	0.93	0.36	0.21	0.29	0.29	0.18	0.18
	FCMAQ	3	2.48	0.40	0.99	0.82	0.49	0.81	0.81	0.49	0.49
	FillSAT	3	1.48	0.50	-0.79	0.79	0.47	-0.19	0.66	-0.11	0.40
	Opt	3	3.61	2.15	1.00	2.37	1.42	1.94	1.94	1.16	1.16

Table 5.7 – (*continued*)

NO3

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
2/3/13	CMAQ	1	8.05	-	-	1.88	0.30	1.88	1.88	0.30	0.30
	FCMAQ	1	8.63	-	-	2.46	0.40	2.46	2.46	0.40	0.40
	FillSAT	1	7.41	-	-	1.24	0.20	1.24	1.24	0.20	0.20
	Opt	1	7.96	-	-	1.79	0.29	1.79	1.79	0.29	0.29
2/12/13	CMAQ	7	4.58	2.62	0.97	0.76	0.18	0.39	0.63	0.09	0.15
	FCMAQ	7	5.05	2.57	0.94	1.27	0.30	0.87	0.94	0.21	0.22
	FillSAT	7	3.32	2.11	0.69	2.11	0.50	-0.86	1.57	-0.21	0.37
	Opt	7	5.30	3.61	1.00	1.31	0.31	1.11	1.11	0.27	0.27

SO4

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
2/3/13	CMAQ	1	1.55	-	-	0.41	0.36	0.41	0.41	0.36	0.36
	FCMAQ	1	1.39	-	-	0.25	0.22	0.25	0.25	0.22	0.22
	FillSAT	1	1.43	-	-	0.29	0.26	0.29	0.29	0.26	0.26
	Opt	1	1.36	-	-	0.22	0.19	0.22	0.22	0.19	0.19
2/12/13	CMAQ	7	0.62	0.18	0.97	0.08	0.11	-0.05	0.05	-0.07	0.07
	FCMAQ	7	0.63	0.10	0.96	0.12	0.18	-0.04	0.09	-0.05	0.13
	FillSAT	7	0.56	0.25	-0.31	0.37	0.55	-0.11	0.29	-0.16	0.43
	Opt	7	0.72	0.24	0.75	0.16	0.24	0.05	0.10	0.08	0.15

EC

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
2/3/13	CMAQ	1	0.00	-	-	1.10	1.00	-1.10	1.10	-1.00	1.00
	FCMAQ	1	0.00	-	-	1.10	1.00	-1.10	1.10	-1.00	1.00
	FillSAT	1	0.13	-	-	0.97	0.88	-0.97	0.97	-0.88	0.88
	Opt	1	1.91	-	-	0.81	0.74	0.81	0.81	0.74	0.74
2/12/13	CMAQ	4	0.00	0.00	0.70	0.89	1.01	-0.88	0.88	-1.00	1.00
	FCMAQ	4	0.00	0.00	0.69	0.89	1.01	-0.88	0.88	-1.00	1.00
	FillSAT	4	0.00	0.00	0.64	0.88	1.00	-0.88	0.88	-1.00	1.00
	Opt	4	1.53	0.30	0.73	0.67	0.77	0.65	0.65	0.74	0.74

OC

Date	Dataset	No. Coincident Observations	Mean	SD	Spatial R	RMSE	NRMSE	MB	MAE	NMB	NME
2/3/13	CMAQ	1	6.00	-	-	0.30	0.05	0.30	0.30	0.05	0.05
	FCMAQ	1	5.25	-	-	0.45	0.08	-0.45	0.45	-0.08	0.08
	FillSAT	1	5.52	-	-	0.18	0.03	-0.18	0.18	-0.03	0.03
	Opt	1	4.77	-	-	0.93	0.16	-0.93	0.93	-0.16	0.16
2/12/13	CMAQ	4	4.58	0.97	0.77	0.99	0.26	0.80	0.82	0.21	0.22
	FCMAQ	4	4.30	0.63	0.76	0.78	0.21	0.52	0.65	0.14	0.17
	FillSAT	4	3.84	0.91	-0.68	1.54	0.41	0.06	1.46	0.02	0.39
	Opt	4	3.34	0.59	0.96	0.61	0.16	-0.44	0.51	-0.12	0.13

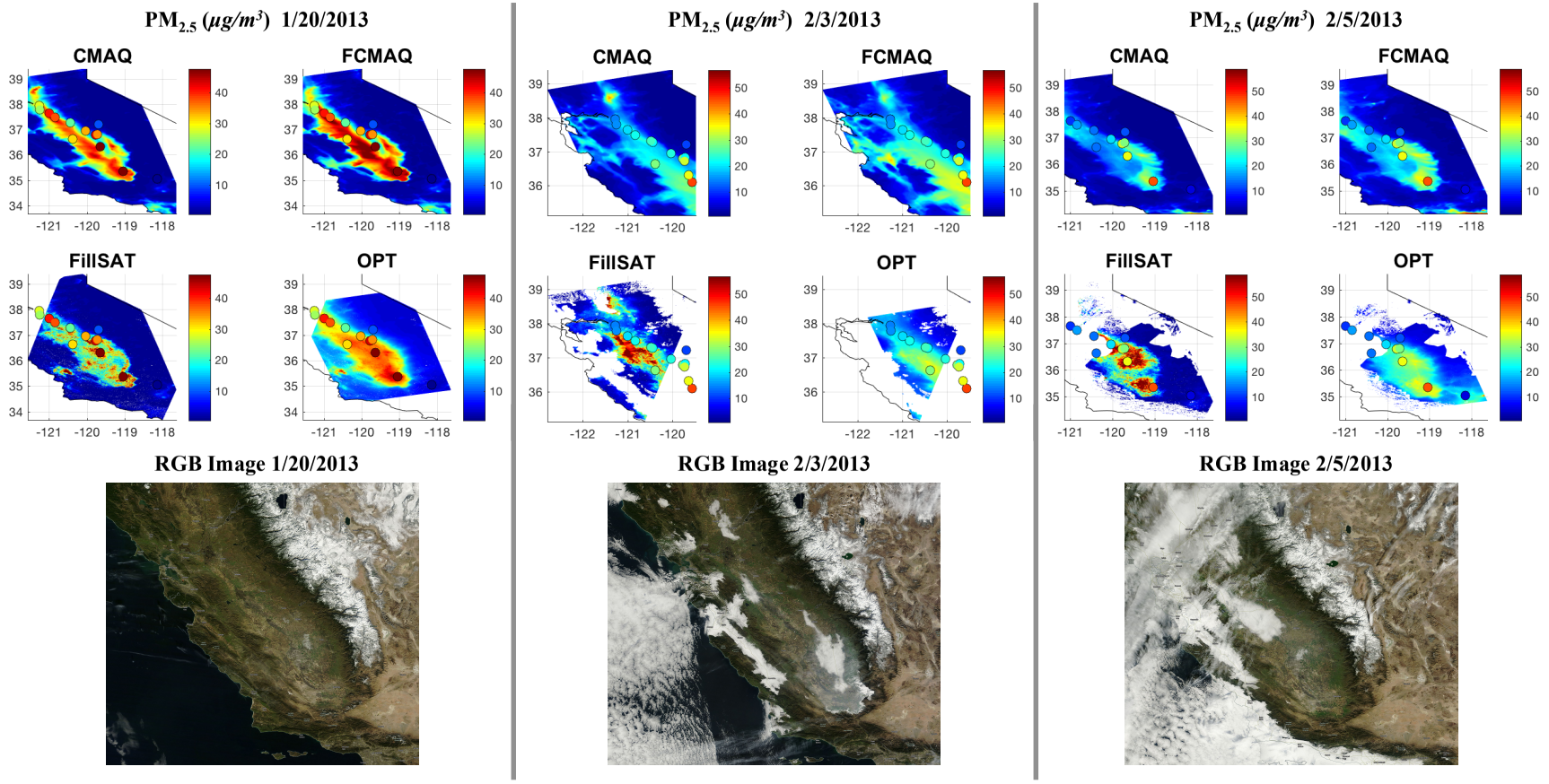

Figure 5.7 – PM2.5 FRM calculated concentration maps with monitor observations (filled circles) and RBG images for the three days with highest AOD.

Focusing on the area within the SJV, the higher concentration gradients in FillSAT are due to the application of satellite snapshots. Even so, the satellite-constrained concentration snapshots tend to provide more realistic spatial distributions of $PM_{2.5}$ compared to the unconstrained model values. Specifically, the $C_{FillSAT}$ maps show greater dynamic ranges of values, with localized hotspots over known urban areas, such as Bakersfield (35.4˚ N lat., 119.0˚ W long.) on January 20th and February 5th, and Fresno (36.7˚ N lat., 119.8˚ W long.) on February 3rd. The satellite-constrained results also tend agree better with available surface measurements in other high-AOD areas, but cloud contamination and the lack of satellite diurnal sampling affect the $C_{FillSAT}$ values in low-AOD regions. This suggests that the technique will yield increasingly good results when applied in more heavily polluted populated areas around the globe. Figure 5.8 presents scatterplots comparing the daily averaged models and the satellite-derived snapshots of near-surface $PM_{2.5}$ to ground monitor values. They indicate than diurnal variability is significant in some places and times, but not in others. For days with high AOD (Jan. 20th, Feb. 3rd, and Feb. 5th), Figure 5.8 shows that, with the exception of a few outliers, the satellite-derived PM2.5 is in agreement with observations within the performance range of the model results, and the variability is minimal compared to low AOD days.

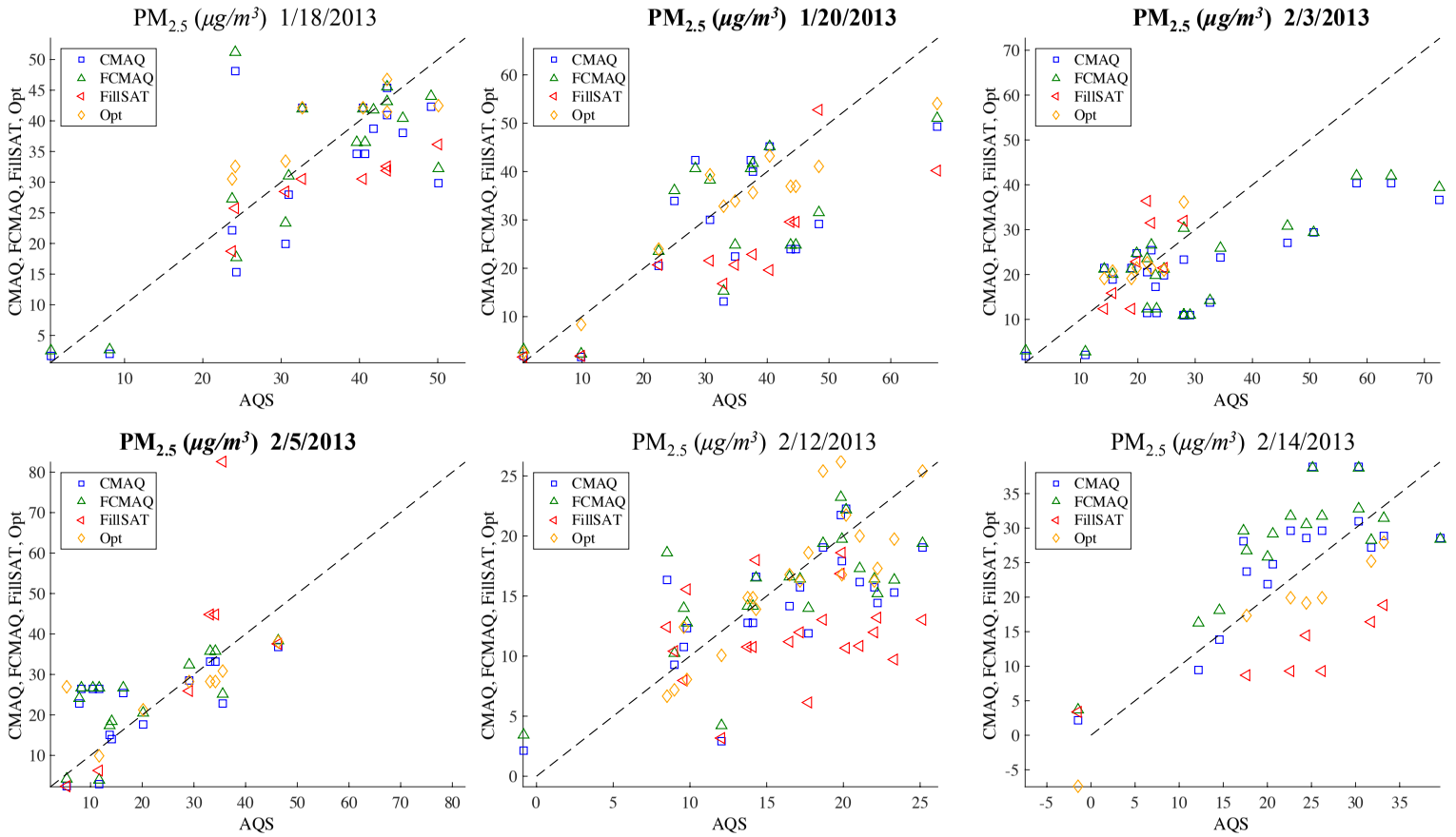

Figure 5.8 – Scatterplot comparison of diurnally-averaged AQS $PM_{2.5}$ coincidences with CMAQ, FCMAQ, CMAQSAT, and FCMAQSAT are show for all days. The satellite datasets are derived from Terra overpass snapshots. A 1:1 dashed line is shown for reference.

Overall, the mean concentrations and normalized mean absolute errors for each pollutant are comparable across the dataset per day, with the exception of EC. Of the three days with relatively high AOD levels, January 20th had the least amount of cloud contamination, whereas February 5th had the most. The satellite-derived fused dataset, $C_{FillSAT}$, outperformed the other datasets for $PM_{2.5}$ on January 20th and February 3rd, with R values of 0.87 and 0.70, and NME of 0.25 and 0.38, respectively. Figures 5.9 and 5.10 provide speciated NO_3, NH_4, and SO_4 maps for January 20th and February 3rd; respectively; ground-truth data available only for February 3rd, are included in Figure 5.10.

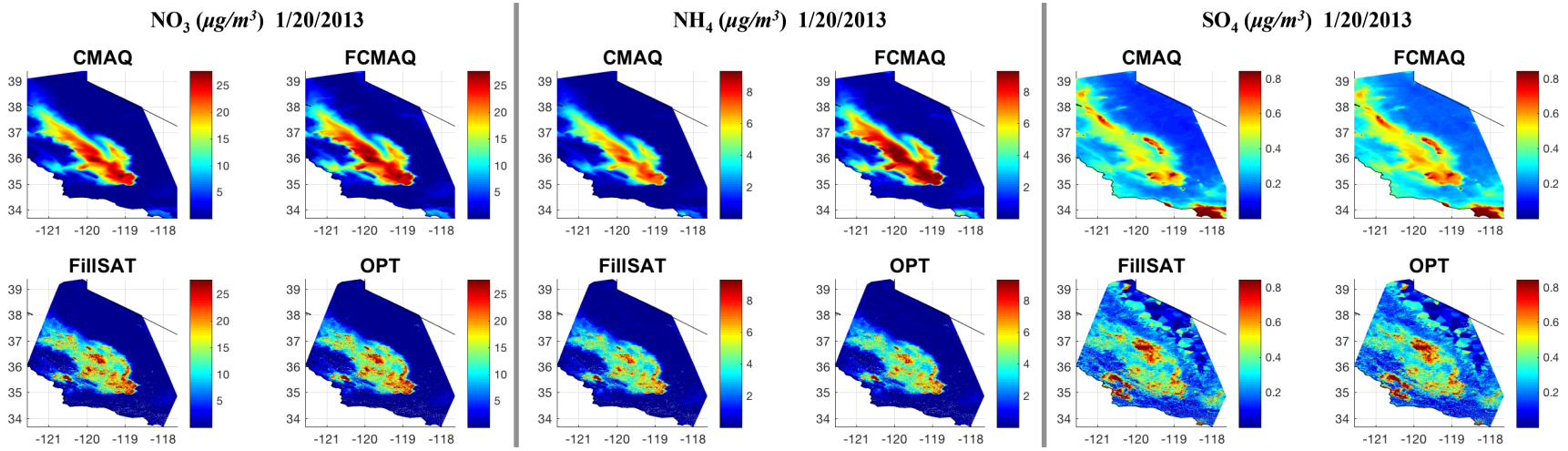

Figure 5.9 –NH_4, SO_4, and NO_3 calculated concentration maps for January 20th, a day void of speciated ground-observations.

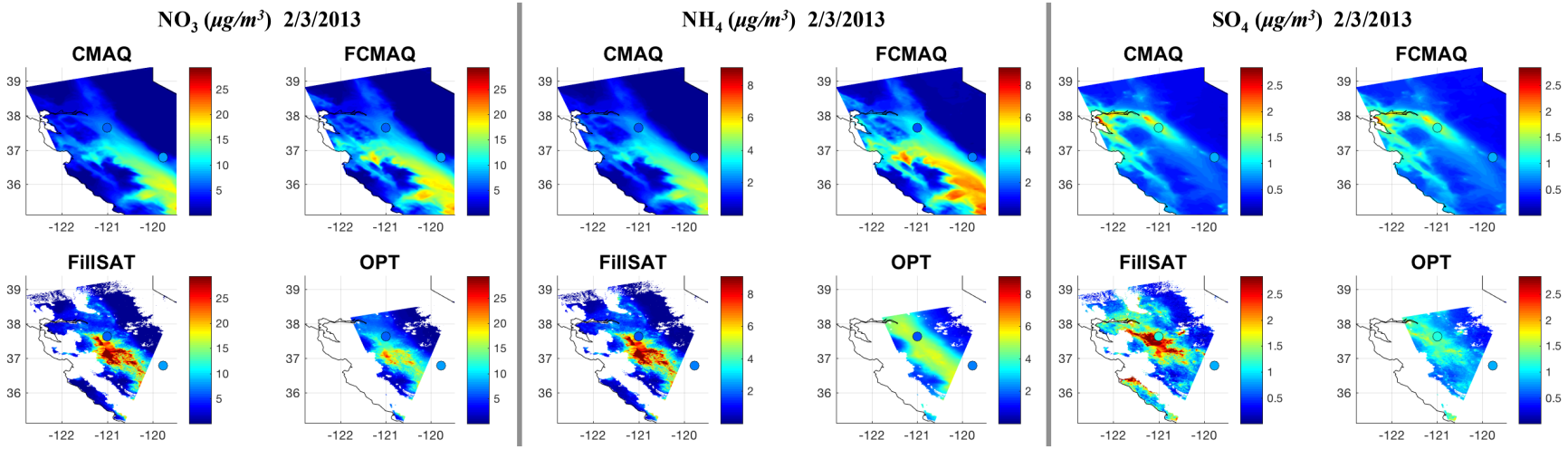

Figure 5.10 – NH_4, SO_4, and NO_3 calculated concentration maps and monitor observations (filled circles) for February 3rd.

For the evaluation of the satellite-constrained surface concentrations, sparse ground observations of speciated PM have a large impact, especially on the high-AOD days. This is compounded by ground-station sampling infrequency, as evident in the correlation ranges (Table 5.7). Figure 5.9 demonstrates the ability of satellite aerosol retrievals to characterize the spatial distributions of speciated aerosol airmass types more realistically and consistently than the models across all three species. Although the model and satellite results show agreement around the location of known emission sources, the satellite-derived aerosol concentrations at the surface show more realistic horizontal mixing patterns, and the spatial distribution better reflects the likely influence of topographic features. For example, the difference between the model and satellite-constrained concentration gradients within the SJV are visible on February 3rd. For this day (Figure 5.10), there is some disagreement between the satellite-derived and modeled concentration gradients, surface mixing, and plume dispersion. The scatterplots in Figure 5.11 show that, at least for the very limited speciated monitor measurements available, the satellite-constrained concentrations provide better agreement than the model and fused-model values.

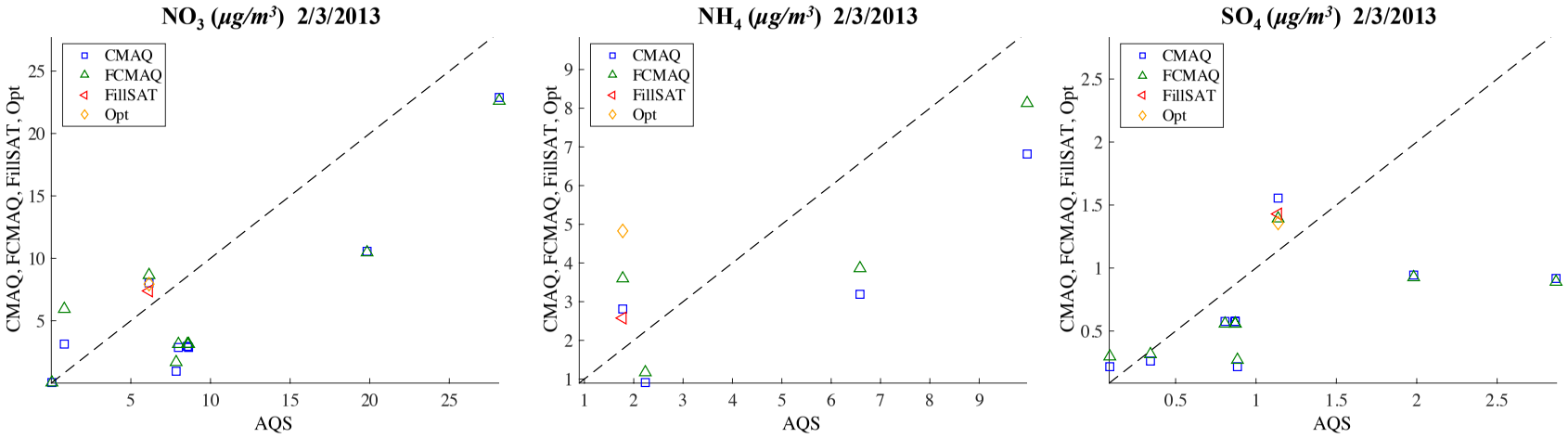

Figure 5.11 – Scatterplot comparison of diurnally-averaged AQS NO_3, NH_4, and SO_4 coincidences with CMAQ, FCMAQ, CMAQSAT, and FCMAQSAT are show for January 3rd. The satellite datasets are derived from Terra overpass snapshots. A 1:1 dashed line is shown for reference.

Comparing the results of the current analysis with previous studies that attempt to apply satellite data to surface air quality assessment is a challenge for the following reasons: (1) limited, non-overlapping case study domains; (2) disparity in spatial resolution at which the analyses are performed, which can bias pixel-to-point comparisons; (3) limited number of ground-truth observations; (4) prevalence of statistics that were averaged over entire seasons or years; (5) lack of actual surface-concentration statistics reported for the satellite-derived values (i.e., many studies report correlations just between satellite-derived, total-column AOD and surface-based $PM_{2.5}$) and (6) where AOD is the satellite-reported quantity used, algorithm version differences between AERONET, MISR, and MAIAC.

With regard to performance comparisons, the statistical-regression-technique study by Liu et al. (2007b; herein referred to as Liu2007b) is the most similar to the current analysis. Liu2007b compares 54 satellite-derived and ground observations for $PM_{2.5}$ mass and speciated particles for the western US. The statistical regression technique used 3-hour averaged CTM (GEOS-Chem) results coincident with Terra overpass for 2005 at 2° by 2.5° spatial resolution. The Liu2007b regression results with removed outliers were as follows: $PM_{2.5}$ R^2=0.21, NO_3 R^2= .23, SO_4 R^2= 0.11, and OC R^2= .11. In our study, the R^2 values for $PM_{2.5}$ are at least 0.41 for all days, and 0.61 for the three days with highest AOD. The R^2 values for the speciated PM on February 12th, the only day for which we have more than one surface measurement, are 0.41 for NO_3, 0.25 for SO_4, 0.61 for OC, 0.69 for NH_4, and 0.67 for EC.

5.4.2 Comparison of CMAQ, Fused, and Optimized Datasets to Observed Concentrations

Outputs from CMAQ simulations and the two measurement-constrained models (C_{FCMAQ}, C_{Opt}) are included in the 10-WH cross-validation comparison, with the monitor data for reference. The RMSE, MB, and the spatiotemporal, temporal, and spatial mean correlations for the five datasets are presented in Table 5.8. The spatiotemporal R^2 $C_{Opt\ 10\text{-}WH}$ values are 0.79 for $PM_{2.5}$, 0.88 for NO_3, 0.78 for SO_4, 1.0 for NH_4, 0.73 for OC, and 0.31 for EC. The similarities among the $PM_{2.5}$ speciated components 10-WH cross-validation statistics are affected by low numbers of observations available, sampling frequency, and coincident satellite-retrieval data, particularly for NH_4 and EC. As a result, the cross-validated EC results show a 40% increase in spatial R^2 and 10% decrease in spatiotemporal R^2, whereas the cross-validation spatiotemporal R^2 values for NH_4 are biased high. The SO_4 spatial and spatiotemporal R^2 cross-validation results for both C_{FCMAQ} and C_{Opt} have the largest improvement over the unconstrained model, with a 43% increase as compared to the CMAQ simulation performance. The $PM_{2.5}$ temporal and spatiotemporal R^2 cross-validation results are 30% and 13% higher than the CMAQ simulations. The C_{Opt} results from the 10-WH cross-validation would normally provide robust cross-validation results that allow for the exploration of differences in errors based on proximity to monitors. Overall, the statistical improvement between the raw CMAQ simulations and cross-validated datasets suggest the empirically based mass reconstruction factors, specific dry efficiencies, and SSA values were adequate for the SJV domain, as the resulting improvements demonstrate.

Table 5.8 – Comparison of temporal R2, spatial R2, spatiotemporal R2, mean bias, and root means square error values between observations and all simulation, including 10-fold 10% holdout cross-validation (10-WH CV), at all monitor locations and for 52 days.

Species	Dataset	Temporal R^2	Spatial R^2	Spatiotemporal R^2	Mean Bias	RMSE
$PM_{2.5}$	cmaq	0.65	0.87	0.67	0.00	0.45
	fcmaq	1.00	0.99	0.88	0.00	0.27
	opt	1.00	0.96	0.87	-0.02	0.28
	fcmaq 10-WH CV	0.95	0.68	0.80	0.24	0.32
	opt 10-WH CV	0.95	0.69	0.79	0.24	0.33
NH4	cmaq	0.70	0.67	0.61	0.29	1.69
	fcmaq	1.00	0.99	0.95	0.71	0.81
	opt	1.00	0.88	0.93	0.53	0.81
	fcmaq 10-WH CV	-	0.98	1.00	1.88	1.02
	opt 10-WH CV	-	0.98	1.00	1.86	1.01
SO4	cmaq	0.29	0.22	0.36	0.00	0.51
	fcmaq	0.99	0.98	0.94	0.05	0.19
	opt	0.99	0.96	0.92	0.01	0.19
	fcmaq 10-WH CV	0.98	0.74	0.79	0.11	0.23
	opt 10-WH CV	0.98	0.73	0.78	0.11	0.23
NO3	cmaq	0.69	0.77	0.79	0.13	0.67
	fcmaq	1.00	0.99	0.99	0.28	0.40
	opt	1.00	0.93	0.98	0.21	0.40
	fcmaq 10-WH CV	-	0.79	0.88	0.40	0.63
	opt 10-WH CV	-	0.79	0.88	0.42	0.65
OC	cmaq	0.62	0.78	0.63	-0.10	0.31
	fcmaq	1.00	0.99	0.98	-0.12	0.14
	opt	1.00	0.95	0.97	-0.16	0.15
	fcmaq 10-WH CV	-	0.98	0.75	-0.14	0.22
	opt 10-WH CV	-	0.97	0.73	-0.13	0.23
EC	cmaq	0.62	0.44	0.43	0.25	0.68
	fcmaq	1.00	0.99	0.97	0.68	0.67
	opt	1.00	0.98	0.97	0.65	0.65
	fcmaq 10-WH CV	-	0.94	0.32	0.49	0.62
	opt 10-WH CV	-	0.93	0.31	0.50	0.65

Pearson Squared Correlation (R^2); Modeled results (M); Observations (O); Covariance (cov); Standard Deviation (SD); monitor location(m_i); day observed (t); Number of monitors (N)

Temporal $R = \frac{1}{N}\sum_{i=1}^{N} \frac{cov\left(M_{m_i,t}, O_{m_i,t}\right)}{SD_{M_{m_i,t}} SD_{O_{m_i,t}}}$; Spatial $R = \frac{cov\left(\left(\frac{1}{N}\sum_N M_m\right)_{t'}, \left(\frac{1}{N}\sum_N O_m\right)_t\right)}{SD_{\left(\frac{1}{N}\sum_N M_m\right)_t} SD_{\left(\frac{1}{N}\sum_N O_m\right)_t}}$; Spatiotemporal $R = \frac{cov\left(M_{m,t}, O_{m,t}\right)}{SD_{M_{m,t}} SD_{O_{m,t}}}$;

Unlike 10-WH, LOO cross-validation results allow us to leverage the spatial distribution of monitor locations throughout the domain. Table 5.9 shows the LOO temporal R^2, MB, and RMSE values averaged across monitor locations. The NH_4 $C_{Opt\ LOO}$ results improved the most across the PM2.5 component species and outperformed temporal R^2 C_{FCMAQ} and $C_{FCMAQ\ LOO}$ values by 10 and 8%. NH_4 cross-validation performance is highest for monitor locations closest to the emission sources (i.e., agricultural), in the southern area of domain. This finding is consistent with aerosol type uncertainties being lowest when AOD is higher than the threshold of 0.15. CMAQ has a difficult time simulating SO_4 and the cross-validation for both $C_{FCMAQ\ LOO}$ and $C_{Opt\ LOO}$ datasets show significant improvements in temporal R^2 and RMSE. For NO_3, temporal R^2 of $C_{FCMAQ\ LOO}$ is slightly higher than that of $C_{Opt\ LOO}$, whereas the opposite is true for MB. The OC $C_{Opt\ LOO}$ results are mixed between locations, whereas the EC $C_{Opt\ LOO}$ shows improvements across all locations.

Table 5.9 – Comparison of averaged temporal R^2, mean bias, and root means square error values between observations and leave-one-out cross-validation (LOO CV) for 52 days across all locations.

Species	Dataset	Temporal R^2	Mean Bias	RMSE
NH4	cmaq	0.52	0.43	0.94
	fcmaq	1.00	0.91	1.24
	opt	1.00	0.70	1.13
	fcmaq LOO CV	0.56	0.90	1.44
	opt LOO CV	0.62	0.71	1.39
SO4	cmaq	0.28	0.02	0.57
	fcmaq	1.00	0.00	0.12
	opt	0.99	-0.09	0.11
	fcmaq LOO CV	0.75	-0.06	0.41
	opt LOO CV	0.63	-0.13	0.36
NO3	cmaq	0.73	0.16	0.49
	fcmaq	1.00	0.26	0.35
	opt	1.00	0.12	0.31
	fcmaq LOO CV	0.89	0.14	0.39
	opt LOO CV	0.85	0.02	0.38
OC	cmaq	0.68	-0.08	0.36
	fcmaq	1.00	-0.11	0.14
	opt	1.00	-0.15	0.13
	fcmaq LOO CV	0.68	-0.12	0.34
	opt LOO CV	0.70	-0.14	0.30
EC	cmaq	0.52	0.31	0.53
	fcmaq	1.00	0.74	0.85
	opt	1.00	0.69	0.83
	fcmaq LOO CV	0.74	0.84	0.87
	opt LOO CV	0.76	0.80	0.88

To explore the $PM_{2.5}$ $C_{FillSAT}$ impact of C_{Opt}, i.e., combining the surface station data with CMAQ simulation plus satellite results, the spatial cross-validation performance assessment of $PM_{2.5}$ C_{Opt} was expanded to include Regional Holdout (RH), which minimizes the effect of clustered monitors on statistics (Table 5.10). As expected, removing $PM_{2.5}$ clustered monitors increased the cross-validated datasets reliance of

C_{FCMAQ} and C_{Opt} on C_{CMAQ} and $C_{FillSAT}$, thus decreasing temporal R^2 values. $PM_{2.5}$ $C_{Opt\ RH}$ results are similar for the C_{CMAQ} and $C_{FCMAQ\ RH}$ datasets, with temporal R^2 values of 0.71-0.84 for $C_{FCMAQ\ RH}$ and 0.72-0.83 for $C_{Opt\ RH}$. Improvements in the cross-validation results with respect to CMAQ simulations are observed for the northern half of the SJV domain, regions 1 and 2 in Figure 5.1. Proximity to emission sources, meteorology, and topography contribute to the performance differences between the northern regions 1 and 2, and southern regions 3 and 4. The main $PM_{2.5}$ mass emission sources (i.e., secondary aerosols, residential wood combustion, and motor vehicles) in the SJV are associated with urban hotspots, such as Fresno and Bakersfield.[193] Winter wind speeds in the SJV are typically below 4 m/s. As compared to the southern portion of the SJV, the wind speed is slightly higher and is more consistently southeasterly in the northern part of the domain. During the winter, regional transport occurs when the nocturnal boundary layer is decoupled from the air aloft as a result of higher wind speeds aloft than at the surface, intensifying pollutant mixing throughout the SJV.[194] These vertical and spatial wintertime wind patterns imply dust is likely transported aloft.

Table 5.10 – Comparison of temporal R2, mean bias, and root means square error PM2.5 values between observations and all simulation, including regional holdout cross-validation (RH CV) for 52 days.

$PM_{2.5}$	Dataset	Temporal R^2	Mean Bias	RMSE
Region 1	cmaq	0.68	0.17	0.40
	fcmaq	1.00	0.10	0.15
	opt	1.00	0.09	0.15
	fcmaq RH CV	0.71	-0.10	0.46
	opt RH CV	0.73	-0.12	0.46
Region 2	cmaq	0.63	-0.04	0.33
	fcmaq	0.99	0.05	0.18
	opt	0.99	0.03	0.16
	fcmaq RH CV	0.75	0.05	0.33
	opt RH CV	0.72	0.03	0.36
Region 3	cmaq	0.77	-0.11	0.30
	fcmaq	1.00	-0.15	0.17
	opt	1.00	-0.17	0.17
	fcmaq RH CV	0.76	0.06	0.31
	opt RH CV	0.76	0.02	0.32
Region 4	cmaq	0.82	-0.11	0.34
	fcmaq	1.00	-0.19	0.24
	opt	1.00	-0.23	0.23
	fcmaq RH CV	0.84	-0.07	0.41
	opt RH CV	0.83	-0.11	0.39

In summary, these results suggest the optimization method is a viable way of constraining CTM simulations using satellite-retrieved information where ground observations are not available. Based on these results, including the satellite data improve short- and long-term spatiotemporal air quality metrics for $PM_{2.5}$ mass, and long-term air quality metrics for $PM_{2.5}$ speciated components.

5.5 Conclusions

Even in the best-monitored urban areas, ground-based networks have limited spatial coverage, especially over extended regions downwind of major pollution sources. Building on earlier work that produced a method for fusing surface-based measurements with model simulations (Friberg et al., 2016; 2017), the current study adds satellite-derived AOD and species-related AOD information as additional constraints on the model. The strength of the satellite data is broad spatial coverage that tends to have uniform quality, at least for the measured radiances, over space and time. The main limitations are lack of vertical discrimination in most situations, lack of diurnal coverage, and only crude aerosol-type sensitivity, especially at low AOD. The approach presented here uses model simulation along with the measurements to address these limitations.

Satellite and ground-retrieved aerosol measurements were combined with the numerical model simulations to: (1) generate aerosol airmass type maps covering the central California test region for the DISCOVER-AQ campaign time period in 2013, (2) explore the viability of using satellite data to improve aerosol airmass type mapping over extended regions, and (3) contribute regional context to what is known about air pollution sources and trends from point sampling stations.

Satellites help capturing $PM_{2.5}$ over large, under-sampled or un-sampled regions, and the combined results tend to represent spatial gradients better than the unconstrained model. Applied appropriately, satellite data can also improve speciated $PM_{2.5}$ where AOD is sufficiently high (generally mid-visible AOD >~0.15 in the study region). The satellite-constrained concentration maps are spatially consistent with topography, typifying localized hotspots over known urban areas, and exhibit realistic mixing patterns

in the SJV. Comparison with daylight-averaged AERONET and diurnally averaged CMAQ modeling demonstrated that, for AOD $\gtrsim$ 0.15 and with outliers removed, the satellite-derived snapshots represent the diurnal values within 10 – 20 % for the study cases. Furthermore, satellite-derived $PM_{2.5}$ is in agreement with surface observations, to within the scatter of unconstrained model results, and variability is reduced on higher AOD days. These results suggest satellite retrievals can improve model performance for $PM_{2.5}$ and speciated components in situations where the AOD is sufficiently high. The satellite aerosol retrievals also represent the spatial distributions of speciated aerosol airmass types more realistically and consistently than the unconstrained model and the model constrained only by surface-station data, for nitrate, ammonium, and possibly sulfate, at least qualitatively.

For the current study, model-based aerosol vertical distributions were used to address the lack of measured values. However, model aerosol vertical distribution could be constrained on large scales with space-based stereo imaging (e.g., MISR) near emission sources, at least where plumes are visible in the imagery, and with space-based lidar (e.g., CALIPSO) downwind of sources. Diurnal sampling, the second major limitation in the current satellite application, can be assessed and corrected where needed with a model that has been scaled to available satellite snapshots. Eventually, AOD and possibly speciated AOD from geostationary platforms will provide at least daylight if not fully diurnal values.

Under adequate observing conditions, the technique presented here improves the representation of pollutant spatial distributions in air quality models downwind from

emission sources. This physically based method also offers the ability to compare satellite-derived $PM_{2.5}$ and speciated concentrations directly to surface measurements. Although the study domain and timeframe did not offer the high AOD levels where this method would work best, the SJV offered a substantial quantity of suborbital observations for assessing the results, due to the DISCOVER-AQ campaign.

Expanding this work by applying the technique to the other areas with key ground measurements (i.e., Baltimore DISCOVER-AQ campaign) are possible next steps. The technique takes advantage of the stable (i.e., consistent) long-term satellite observations that offer global coverage, and provides speciated constraints based on retrieved microphysical properties for AOD retrievals above about 0.15. Once the aforementioned analyses are completed, the method will likely be applied to a selection of globally distributed urban regions that are downwind of sources, in locations where particulate pollution levels tend to be high.

CHAPTER 6. CONCLUSION AND FUTURE DIRECTIONS

By constraining CTMs using a combination of ground-station and satellite-based measurements, we can improve urban- and regional-scale ambient air quality metrics for health risk assessment. Epidemiologists use these metrics as surrogates for population exposure, to investigate the associations between air pollution and health outcomes. Thus, addressing exposure misclassification and quantifying their uncertainty can improve relative risk analyses of health effects associated with air pollution. This dissertation explores the impact of constraining CTM results using ground-based observations and satellite retrievals to improve spatial and temporal air quality metrics.

In Chapter 3, a method for blending surface-station observations and CTMs was developed, and the accuracy of daily air quality estimates for twelve pollutants away from monitor sites was evaluated for the state of Georgia over a period of seven years. The strength of this data fusion approach is that the predicted temporal variance is optimized over space. The limitations are the number and location of ambient monitors and the accuracy and resolution of the CTM used. In addition to the robust tenfold 10% withholding validation, future cross-validation analyses of the data fusion approach should include techniques that investigate the bias associated with spatial clustering of monitors and ensure the distance between withheld and non-withheld monitors is varied to sample the fusion technique temporal correlations across the various distance regimes.

Chapter 4 expanded the application of the data fusion method developed in Chapter 3 to five metropolitan. We applied the data fusion method to daily spatial fields

for 12 pollutants, quantifying spatial and temporal ambient air concentration metrics and uncertainty across five cities. Using the data fusion results, this chapter characterized inter- and intra-city air quality heterogeneity and identifying key physical aspects contributing the underlying relationships of the five urban environments. Reported cross-validation results showed the estimated datasets were consistent with temporal strengths of the ground monitor data and the increase in spatial coverage was in agreement with the CTM inputs, emissions and meteorology. The evaluation and understanding of the data fusion method was enhanced further through the use of statistical tools that compared performance changes between the CTM results and fusion results relative to observations.

With thorough data fusion method applications and evaluations established in Chapters 3 and 4, expansion of the method to include satellite-based measurements was investigated. Chapter 5 assessed improvements in near-surface PM2.5 estimates progressively downwind of sources at the regional-scale provided when constraining CTM data using aerosol type information from satellite retrievals.

Overall, this dissertation makes it possible to (1) identify seasonal and meteorological differences in observed associations, (2) explore various explanations for the multi-city differences, and (3) identify common underlying relationships among air quality characteristics across the cities. Satellite-retrieved aerosol data allowed us to (1) produce aerosol airmass type maps for the southern California test region and time period, (2) demonstrate the viability of using satellite data for mapping out aerosol air mass types in urban regions, (3) constrain a regional aerosol transport model using the

satellite-derived maps, and (4) contribute regional context to what is known about air pollution sources and trends from point sampling stations.

By integrating satellite-, aircraft-, and surface-based AOD measurements with models in several ways, we helped further our understanding of effective and practical approaches to using satellite-retrieved aerosol data products for air quality and public health applications. Advancing our understanding of how to optimize the microphysical aerosol column-integrated constraints derived from satellite observations and their relationship with near-surface air pollution concentrations promotes the appropriate use of satellite data for the vast areas with limited ground-based air quality measurement networks.

Satellites offer stable platforms capable of producing global, long-term data records. As such, space-based aerosol air mass mapping will be used on an extended basis to help reduce the unconstrained assumptions and biases of regional air-quality models, especially downwind of major pollution sources, where surface sampling tends to be sparse or non-existent. In future work, the satellite-fusion method will be applied to other areas with key ground measurements and higher-AOD situations, using several NASA aerosol data products that offer global coverage. The discrepancy between snapshots and daily averaged concentrations is a limitation in the current satellite application. This drawback will be further assessed and corrected where needed with a scaling technique. In the future geostationary platforms will provide AOD and possibly speciated AOD for least daylight if not fully diurnal values.

The feasibility of meeting the method requirements will be evaluated for the Baltimore DISCOVER-AQ campaign, for which 1km CTM simulations are available. In addition to Baltimore, the method will be applied to other urban locations that offer robust air quality data, CTM simulations, and high AOD levels. From there, the application areas will extend to a selection of globally distributed urban and surrounding regions that are affected by major pollution sources.

The satellite-constrained model approach lends itself to long time-series exposure studies. Thus, the viability of connecting the application locations and period with coexistent chronic or long-term health studies will be pursued. The ability of geostationary platform AOD and possibly aerosol grouped AOD to provide daylight if not fully diurnal values will also be investigated. By advancing our understanding of regional air quality and pollution transport through the use of aerosol mapping, we also inherently advance our understanding of changes in Earth's radiation balance along with air quality, thus addressing one of the overarching science goals of current Earth Science.

APPENDIX A. SUPPLEMENTAL MATERIAL FOR METHOD FOR FUSING OF OBSERVATIONAL DATA AND CHEMICAL TRANSPORT MODEL SIMULATIONS TO ESTIMATE SPATIOTEMPORALLY-RESOLVED AMBIENT AIR POLLUTION

Table A.1 – Estimated annual mean model fitted regression parameters for 2002 through 2008 for 12 km resolution.

Species	O_3	NO_2	NO_x	CO	SO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	EC	OC
β	1.0	1.0	1.0	1.0	1.0	0.1	0.5	0.7	0.5	0.7	1.0	0.6
α_{2002}	0.92	0.82	1.4	1.3	1.2	15	4.2	1.9	0.6	1.0	1.4	2.1
α_{2003}	0.90	0.78	1.3	1.3	1.1	17	4.2	1.8	0.6	1.0	1.4	2
α_{2004}	0.92	0.79	1.3	1.1	1.1	17	4.3	1.8	0.5	1.0	1.4	2
α_{2005}	0.91	0.81	1.3	1.2	1.2	17	4.3	1.8	0.5	1.0	1.4	1.9
α_{2006}	0.95	0.88	1.3	1.3	1.2	18	4.2	1.7	0.5	1.0	1.4	1.7
α_{2007}	0.93	0.69	1.0	1.0	1.1	20	4.2	1.8	0.5	1.0	1.4	1.6
α_{2008}	0.92	0.79	1.0	1.1	1.1	17	3.6	1.8	0.4	1.0	1.0	1.5

Table A.2 – β_{season} fitted parameters for 2002 through 2008 for 12 km resolution.

Species	O_3	NO_2	NO_x	CO	SO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	EC	OC
A	0.05	0.1	0.2	0.1	0.3	0.5	0.4	0.2	0.3	0.4	0.2	0.3
t_{max}	70	80	10	350	50	190	190	190	30	200	280	190

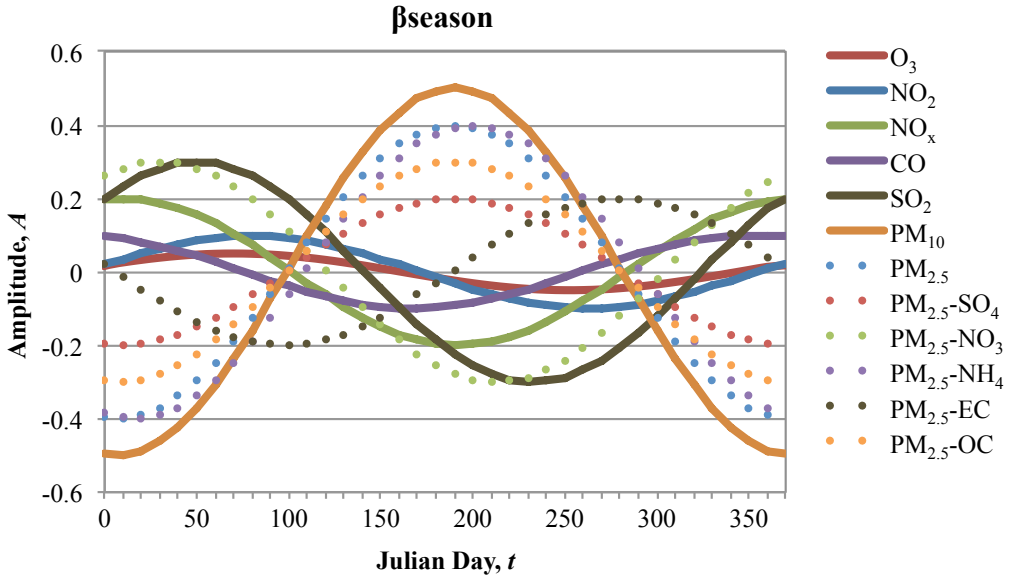

Figure A.1 – Seasonal corrections in CMAQ simulations, 2002-2008.

Table A.3 – Fitted parameters for 2002 through 2008 for 12 km resolution.

Species	O_3	NO_2	NO_x	CO	SO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	EC	OC
R_{coll}	0.97	0.89	0.98	0.91	0.9	0.69	0.93	0.98	0.90	0.98	0.61	0.98
r (km)	590	36	56	48	50	560	450	590	560	480	590	330
R_{cmaq}	0.81	0.49	0.52	0.47	0.30	0.43	0.54	0.74	0.48	0.54	0.52	0.51

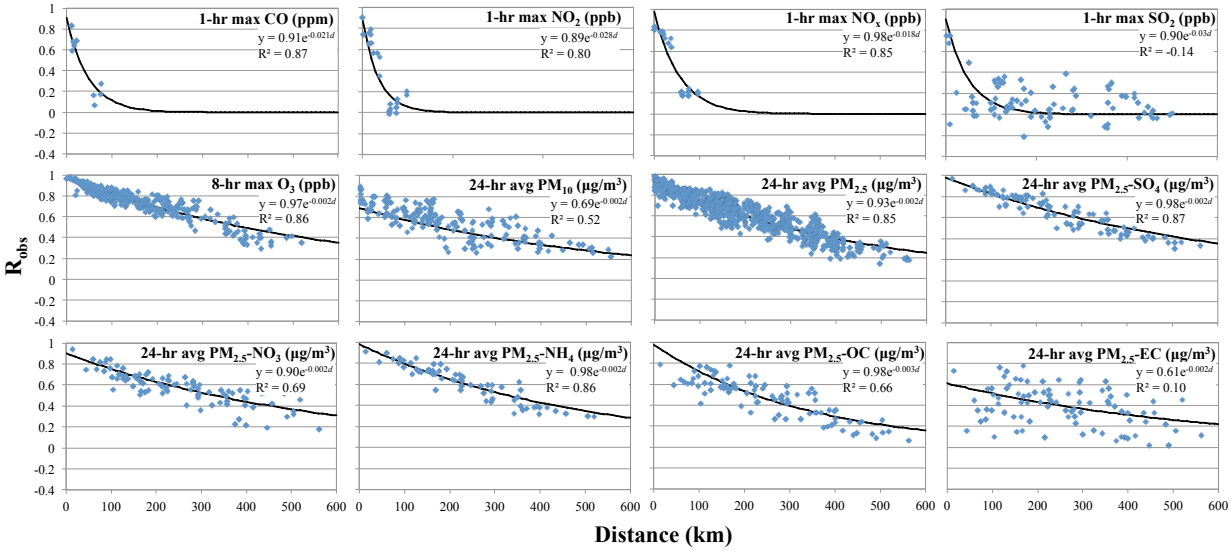

Figure A.2 – Correlograms (R_{obs}), 2002-2008. Curves represent fitted exponential correlograms.

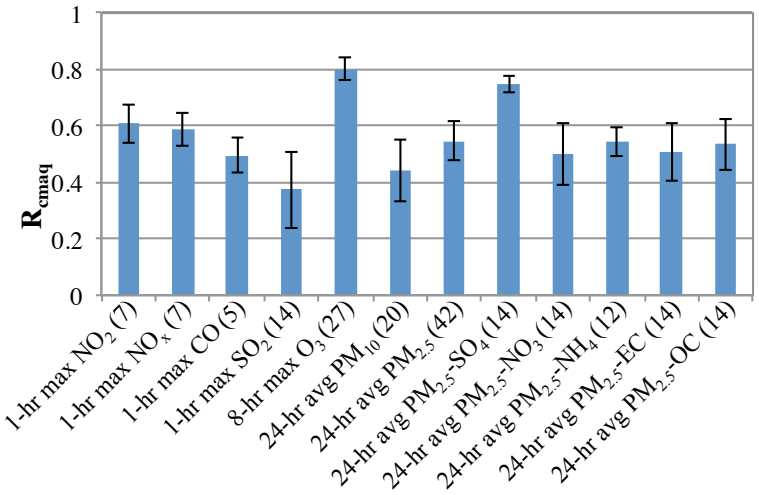

Figure A.3 – Temporal CMAQ-OBS Pearson Rcmaq, 2002-2008. Error bars represent two standard deviations. Number of monitors indicated in parentheses.

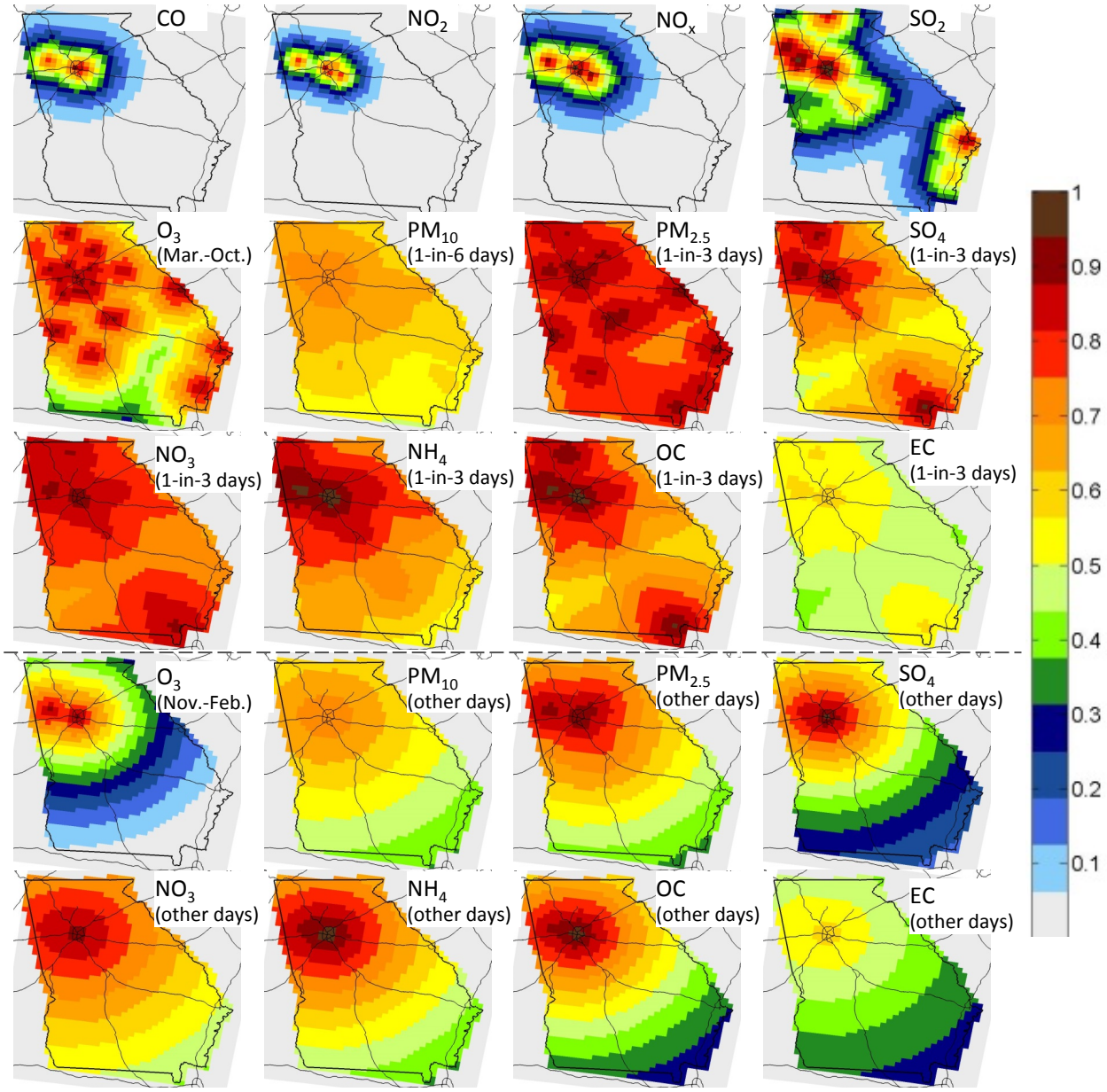

Figure A.4 – Average weighting factor (W) fields, 2002-2008.

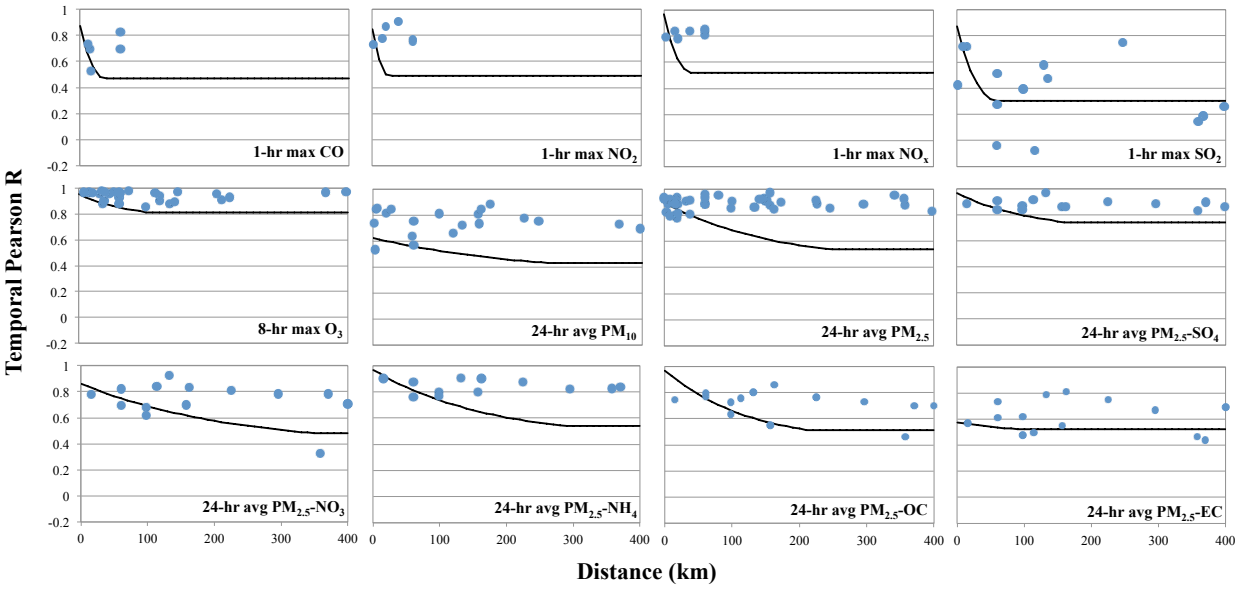

Figure A.5 –Temporal correlation of optimized data fusion results and withheld observations plotted as a function of the distance to the nearest monitor (blue points). The black curves represent the estimated temporal correlation using eq. 10.

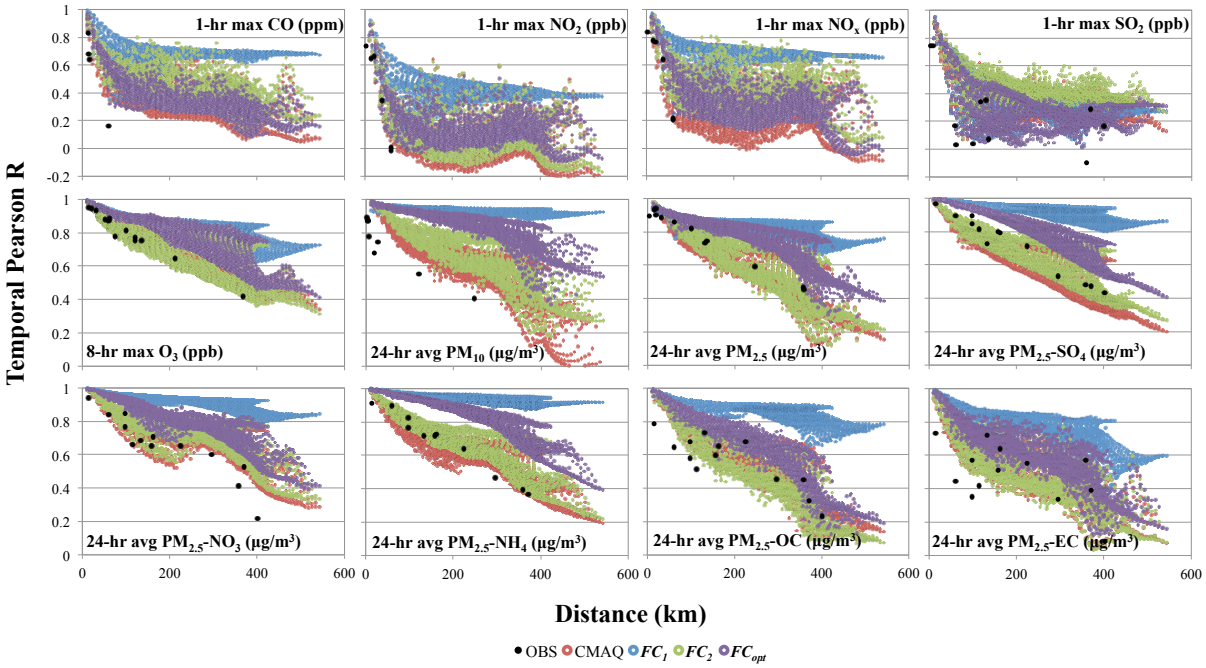

Figure A.6 – Spatial autocorrelation in observations, CMAQ simulations, interpolated observation results, adjusted CMAQ results, and optimized data fusion results. The temporal Pearson R for daily concentrations at the Jefferson St. monitor or grid cell and all other monitors or grid cells is plotted versus distance to Jefferson St.

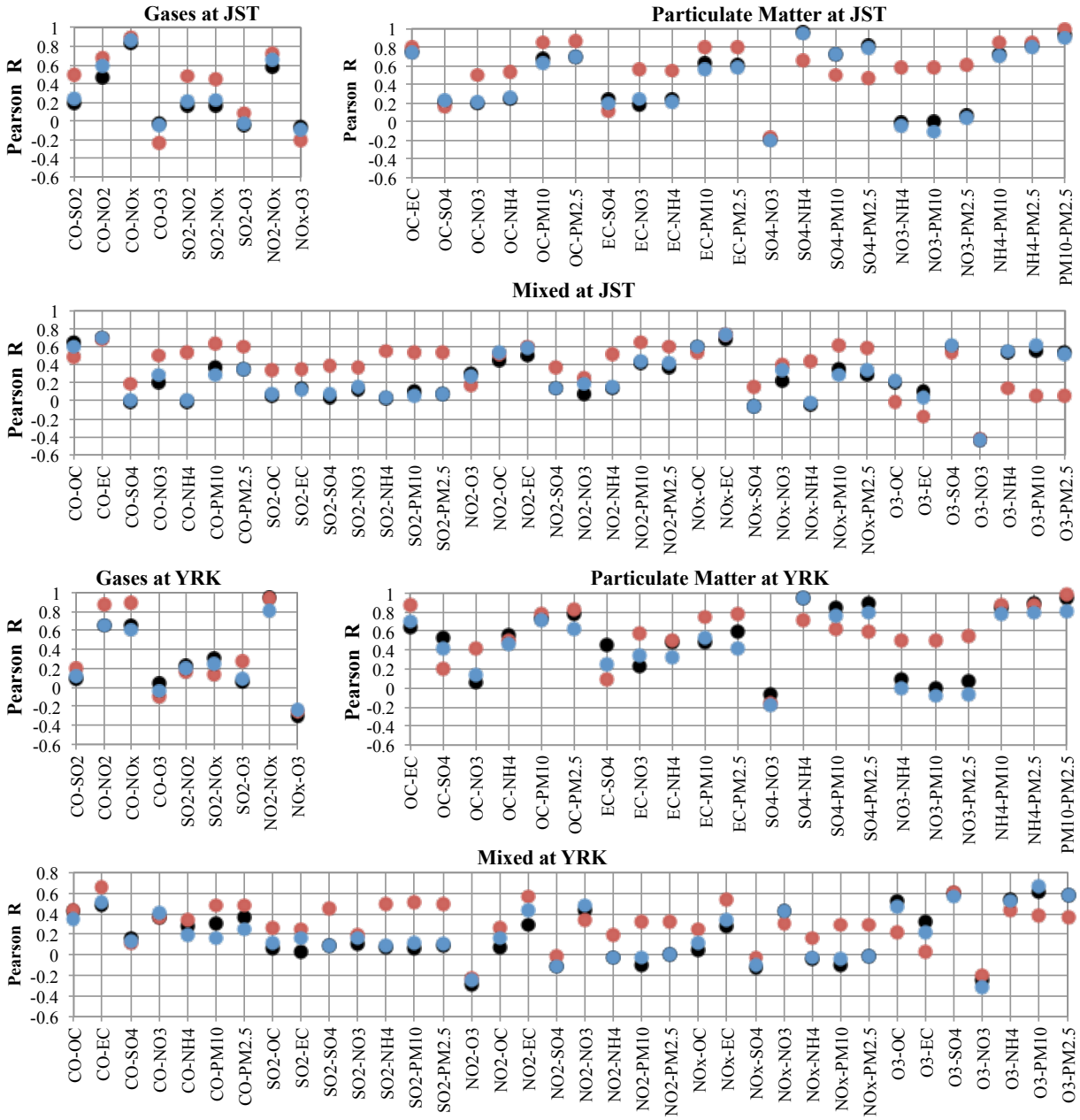

Figure A.7 – Bivariate correlations at Jefferson St. (JST) and Yorkville (YRK), 2002-2008. Temporal Pearson correlation in observations (black points), CMAQ simulations (red points), and optimized data fusion results (blue points).

Table A.4 – Evaluation metrics, 2002-2008. Metrics are shown for the population of monitors and coincident cell for each pollutant (numbers shown in parentheses).

Metric	Modeled dataset/ Species	1-hr max. NO_2, ppb (7)	1-hr max. NO_x, ppb (7)	1-hr max. CO, ppm (5)	1-hr max. SO_2, ppb (14)	8-hr max. O_3, ppb (27)	24-hr avg. PM_{10}, µg/m^3 (20)	24-hr avg. $PM_{2.5}$, µg/m^3 (42)	24-hr avg. $PM_{2.5}$-SO_4, µg/m^3 (14)	24-hr avg. $PM_{2.5}$-NO_3, µg/m^3 (14)	24-hr avg. $PM_{2.5}$-NH_4, µg/m^3 (12)	24-hr avg. $PM_{2.5}$-EC, µg/m^3 (14)	24-hr avg. $PM_{2.5}$-OC, µg/m^3 (14)
Mean Fractional Error, MFE $\frac{1}{N}\sum_1^N \frac{\lvert modeled - OBS\rvert}{\left(\frac{modeled+OBS}{2}\right)}$	*CMAQ*	0.45	0.57	0.40	0.78	0.18	0.42	0.39	0.35	0.87	0.43	0.48	0.61
	CMAQ w/ annual adj.	0.39	0.60	0.40	0.79	0.17	0.36	0.35	0.32	0.90	0.42	0.43	0.46
	FC_2	0.38	0.60	0.41	0.79	0.16	0.32	0.30	0.31	0.93	0.39	0.44	0.43
	FC_{opt}	0.24	0.44	0.24	0.40	0.05	0.15	0.11	0.07	0.26	0.10	0.25	0.14
	WH FC_{opt}	0.32	0.51	0.38	0.73	0.10	0.23	0.17	0.21	0.48	0.26	0.37	0.30
Mean Fractional Bias, MFB $\frac{1}{N}\sum_1^N \frac{(modeled - OBS)}{\left(\frac{modeled+OBS}{2}\right)}$	*CMAQ*	0.29	0.04	-0.07	0.13	0.11	-0.24	-0.14	-0.15	-0.05	0.15	-0.23	-0.48
	CMAQ w/ annual adj.	0.08	0.22	0.09	0.23	0.03	0.01	0.00	0.07	-0.60	0.04	0.03	-0.21
	FC_2	0.08	0.22	0.09	0.23	0.02	0.02	0.00	0.07	-0.58	0.05	0.03	-0.21
	FC_{opt}	0.09	0.16	0.07	0.05	0.01	0.02	0.02	0.03	-0.23	-0.01	0.08	-0.09
	WH FC_{opt}	0.09	0.17	0.09	0.15	0.02	0.03	0.02	0.07	-0.25	0.04	0.09	-0.09
Normalized Mean Error, NME $\frac{\sum_1^N \lvert modeled - OBS\rvert}{\sum_1^N OBS}\,100\%$	*CMAQ*	43.6%	58.9%	41.4%	77.5%	16.5%	37.3%	35.2%	32.1%	117%	44.8%	42.8%	47.4%
	CMAQ w/ annual adj.	33.8%	59.7%	41.3%	81.3%	15.0%	35.3%	35.0%	31.1%	64.3%	40.3%	43.9%	41.6%
	FC_2	33.1%	59.9%	41.8%	81.0%	14.7%	31.4%	28.4%	31.9%	68.9%	36.4%	45.3%	38.7%
	FC_{opt}	19.2%	37.2%	20.2%	42.2%	4.49%	14.3%	10.9%	6.84%	23.4%	9.23%	26.0%	13.9%
	WH FC_{opt}	26.1%	45.1%	36.1%	75.6%	8.57%	21.7%	15.6%	20.8%	40.6%	23.8%	39.3%	28.6%
Normalized Mean Bias, NMB $\frac{\sum_1^N (modeled - OBS)}{\sum_1^N OBS}\,100\%$	*CMAQ*	27.2%	-29.8%	-17.9%	-16.7%	8.58%	-23.2%	-11.6%	-16.8%	72.1%	13.8%	-17.5%	-30.3%
	CMAQ w/ annual adj.	1.03%	-13.4%	-2.24%	-4.65%	0.14%	-0.14%	0.96%	3.09%	-22.7%	-1.39%	9.11%	-9.51%
	FC_2	1.21%	-10.7%	-1.21%	-3.27%	-0.57%	4.57%	2.44%	7.98%	-11.2%	3.17%	10.8%	-11.0%
	FC_{opt}	1.10%	-12.7%	-2.14%	-4.48%	0.03%	-0.10%	0.96%	2.94%	-21.3%	-1.58%	9.80%	-10.6%
	WH FC_{opt}	1.15%	-12.9%	-1.81%	-4.27%	0.03%	0.60%	1.03%	4.85%	-18.4%	-0.03%	10.1%	-11.0%
Mean Bias, MB $\frac{1}{N}\sum_1^N (modeled - OBS)$	*CMAQ*	6.23	-0.02	-0.14	-1.84	0.004	-5.42	-1.72	-0.70	0.50	0.19	-0.16	-0.96
	CMAQ w/ annual adj.	0.24	-0.01	-0.02	-0.51	6.9E-05	-0.03	0.14	0.13	-0.16	-0.02	0.08	-0.30
	FC_2	0.28	-0.01	-0.01	-0.36	-2.7E-04	1.07	0.36	0.33	-0.08	0.04	0.10	-0.35
	FC_{opt}	0.25	-0.01	-0.02	-0.49	1.5E-5	-0.02	0.14	0.12	-0.15	-0.02	0.09	-0.33
	WH FC_{opt}	0.26	-0.01	-0.01	-0.47	1.3E-5	0.14	0.15	0.20	-0.13	0.00	0.09	-0.35
Root Mean Square Error, RMSE $\sqrt{\frac{1}{N}\sum_1^N (modeled - OBS)^2}$	*CMAQ*	13.41	0.07	0.52	15.30	0.01	11.98	7.23	2.01	1.42	0.86	0.61	2.08
	CMAQ w/ annual adj.	10.63	0.07	0.50	15.63	0.01	11.50	7.38	1.90	0.69	0.78	0.61	2.06
	FC_2	10.41	0.06	0.50	15.41	0.01	10.48	6.06	2.00	0.74	0.70	0.63	1.89
	FC_{opt}	6.23	0.04	0.22	10.29	0.003	4.80	2.42	0.44	0.27	0.19	0.37	0.62
	WH FC_{opt}	8.40	0.05	0.43	15.38	0.006	7.70	3.59	1.42	0.48	0.49	0.56	1.40
Spatiotemporal R^2 $corr(OBS, modeled)^2$	*CMAQ*	58.8%	38.7%	40.2%	7.8%	67.1%	22.3%	30.9%	57.4%	30.5%	30.3%	42.4%	31.9%
	CMAQ w/ annual adj.	59.9%	40.5%	41.8%	7.9%	67.2%	22.8%	30.7%	58.8%	27.9%	30.4%	48.2%	33.9%
	FC_2	61.5%	42.0%	41.7%	9.2%	68.7%	42.7%	52.2%	61.7%	30.9%	47.8%	48.1%	37.2%
	FC_{opt}	86.7%	79.2%	89.1%	56.7%	97.0%	82.7%	90.2%	98.0%	93.3%	95.5%	79.9%	94.0%
	WH FC_{opt}	68.8%	56.7%	52.9%	13.7%	87.1%	58.6%	75.1%	80.5%	56.8%	71.6%	53.0%	53.9%

APPENDIX B. SUPPLEMENTAL MATERIAL FOR DAILY AMBIENT AIR POLLUTION METRICS FOR FIVE CITIES: EVALUATION OF DATA-FUSION-BASED ESTIMATES AND UNCERTAINTIES

B.1 Definitions of performance measures used in Taylor diagrams

Model pollutant concentration values within grid cells were compared to coincident observations using the following evaluation metrics:

$$Mean\ Bias\ (MB) = \overline{C_{SIM_{m,t}} - C_{OBS_{m,t}}} \tag{B.17}$$

$$Root\ Mean\ Square\ Error\ (RMSE) = \sqrt{\overline{\left(C_{SIM_{m,t}} - C_{OBS_{m,t}}\right)^2}} \tag{B.2}$$

$$Spatiotemporal\ R = \frac{cov(C_{SIM_{m,t}}, C_{OBS_{m,t}})}{SD_{C_{SIM_{m,t}}} SD_{C_{OBS_{m,t}}}} \tag{B.3}$$

$$Temporal\ R = \frac{1}{N}\sum_{i=1}^{N} \frac{cov\left(C_{SIM_{m_i,t}}, C_{SIM_{m_i,t}}\right)}{SD_{C_{SIM_{m_i,t}}} SD_{C_{SIM_{m_i,t}}}} \tag{B.4}$$

$$Spatial\ R = \frac{cov\left(\left(\overline{C_{SIM_m}}\right)_t, \left(\overline{C_{OBS_m}}\right)_t\right)}{SD_{\left(\overline{C_{SIM_m}}\right)_t} SD_{\left(\overline{C_{OBS_m}}\right)_t}} \tag{B.5}$$

Here, C_{SIM} indicates the collocated simulated grid cell concentration (either C_{CMAQ} or C_{opt}), t indicates the day observed, SD is the standard deviation, N is the number of

monitors, the overbar indicates the mean, and m_i indicates a specific monitor (m). The mean bias (eq. B.1) and root mean square error (RMSE) (eq. B.7) statistics are averaged across monitors and/or days depending on the analysis type (i.e., temporal, spatial, or spatiotemporal). To assess model simulation of variation over space and time, various Pearson correlation coefficients (R) were calculated (eqs. B.3 and B.5). Correlations of subsets of measured and modeled data were evaluated to assess model performance in simulating spatial and temporal variation separately. Temporal R values at each monitor (eq. B.4) were described by a mean across monitors. Spatial R values were described by a mean across days (eq. B.5).

The linear correlation coefficient, R, is bound by the range $-1 \leq R \leq 1$ and is a common measure of the degree to which the simulated concentration relates to the observed values (eq. B.3). In general, as the simulated and observed temporal concentration datasets approach agreement, R converges towards 1 on the Taylor diagram. The second summary statistic used is the normalized standard deviation (NSD) of the modeled dataset:

$$NSD = \frac{SD_{C_{SIM_{m,t}}}}{SD_{C_{OBS_{m,t}}}} \tag{B.6}$$

The centered pattern of the root-mean-square difference (CRMSE) is the root-mean-square difference with the overall bias removed. The CRMSE can be used to detect changes in general model performance when comparing results across various models to observations. The normalized CRMSE (NCRMSE) complements the model bias (eq.

B.1), making it possible to plot statistics for different domains on the same plot and providing a gauge for error, though not providing information about the bias itself. The NCRMSE term is calculated as follows:

$$NCRMSD = \frac{\sqrt{\overline{\left[\left(C_{SIM_{m,t}} - \overline{C_{SIM_{m,t}}}\right) - \left(C_{OBS_{m,t}} - \overline{C_{OBS_{m,t}}}\right)\right]^2}}}{SD_{C_{OBS_{m,t}}}} \quad \text{(B.7)}$$

According to eq. B.7 and the inherent bias in the C_{CMAQ} results, the NCRMSE between the modeled and observed datasets may approach, but never reach a value of 0. With the addition of the mean bias (eq. B.1) in the Taylor diagram, increases in bias between model performances can be tracked, and general error patterns among models or domains can be detected.

Table B.1 – Estimated annual mean model-fitted regression parameters by city and species, for 12 km resolution.

City	Parameter	O_3	NO_2	NO_x	CO	SO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	EC	OC
Atlanta	β	1.0	1.0	1.0	1.0	1.0	0.1	0.5	0.7	0.5	0.7	1.0	0.6
	α_{2002}	0.92	0.82	1.4	1.3	1.2	15	4.2	1.9	0.6	1.0	1.4	2.1
	α_{2003}	0.90	0.78	1.3	1.3	1.1	17	4.2	1.8	0.6	1.0	1.4	2
	α_{2004}	0.92	0.79	1.3	1.1	1.1	17	4.3	1.8	0.5	1.0	1.4	2
	α_{2005}	0.91	0.81	1.3	1.2	1.2	17	4.3	1.8	0.5	1.0	1.4	1.9
	α_{2006}	0.95	0.88	1.3	1.3	1.2	18	4.2	1.7	0.5	1.0	1.4	1.7
	α_{2007}	0.93	0.69	1.0	1.0	1.1	20	4.2	1.8	0.5	1.0	1.4	1.6
	α_{2008}	0.92	0.79	1.0	1.1	1.1	17	3.6	1.8	0.4	1.0	1.0	1.5
Birmingham	β	1.0	1.0	1.0	1.0	1.0	0.1	0.5	0.7	0.7	0.7	1.0	0.6
	α_{2002}	0.90	0.80	1.3	1.2	1.3	20	4.0	1.6	1.2	1.3	2	3.3
	α_{2003}	0.90	0.80	1.3	1.2	1.3	20	4.0	1.6	1.2	1.1	1.5	3.1
	α_{2004}	0.90	0.80	1.3	1.2	1.3	20	4.0	1.6	1	1.1	1.6	3
	α_{2005}	0.90	0.80	1.3	1.2	1.3	20	4.2	1.6	0.9	1.2	1.3	2.7
	α_{2006}	0.92	0.80	1.3	1.2	1.4	20	4.2	1.6	0.8	1.2	1.5	2.7
	α_{2007}	0.92	0.80	1.3	1.2	1.6	20	4.0	1.6	0.8	1.2	1.7	2.5
	α_{2008}	0.92	0.80	1.3	1.2	1.6	20	4.0	1.6	0.8	1.3	1.6	2.7
Dallas	β	1.0	1.0	1.0	1.0	1.0	0.1	0.5	0.7	0.5	0.7	1.0	0.6
	α_{2003}	0.96	0.80	1.3	1.0	1.0	18	3.3	1.7	0.6	1.0	0.8	2.5
	α_{2004}	0.96	0.80	1.3	1.0	1.0	18	3.3	1.7	0.6	0.9	0.8	2.4
	α_{2005}	0.96	0.80	1.3	1.0	1.0	18	3.4	1.8	0.5	1.1	0.9	2.4
	α_{2006}	0.96	0.80	1.3	1.0	1.0	18	3.2	1.6	0.5	0.9	0.8	2.2
	α_{2007}	0.95	0.80	1.3	1.0	1.0	16	3.1	1.6	0.5	0.9	0.8	1.7
	α_{2008}	0.95	0.80	1.3	1.0	1.0	16	3.6	2.1	0.6	1.2	0.6	2.0
Pittsburgh	β	1.0	1.0	1.0	1.0	1.0	0.1	0.5	0.7	0.7	0.7	1.0	0.6
	α_{2002}	1.00	1.10	2.0	1.2	1.5	18	4.2	1.7	1.7	1.2	1.4	3.6
	α_{2003}	1.00	1.10	2.0	1.2	1.5	18	4.2	1.8	1.8	1.2	1.4	3.4
	α_{2004}	0.96	1.10	2.0	1.2	1.5	18	4.2	1.7	1.7	1.2	1.4	3.3
	α_{2005}	0.96	1.10	2.0	1.2	1.5	18	4.3	1.8	1.8	1.2	1.4	3.4
	α_{2006}	0.96	1.10	1.8	1.2	1.5	18	4.2	1.6	1.6	1.2	1.4	3.3
	α_{2007}	0.96	1.10	1.5	1.2	1.5	18	4.2	1.9	1.9	1.2	1.4	2.9
St. Louis	β	1.0	1.0	1.0	1.0	1.0	0.1	0.5	0.7	0.7	0.7	1.0	0.6

Table B.2 – Fitted γ_{season} parameters for 2002 through 2008 for 12 km resolution for Atlanta only (eq. 4.1).

Species	O_3	NO_2	NO_x	CO	SO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	EC	OC
A	0.05	0.1	0.2	0.1	0.3	0.5	0.4	0.2	0.3	0.4	0.2	0.3
t_{max}	70	80	10	350	50	190	190	190	30	200	280	190

Table B.3 – Fitted parameters for 2002 through 2008 for 42 km resolution (eq. 4.1).

City	Parameter	O_3	NO_2	NO_x	CO	SO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	EC	OC
Atlanta	R_{coll}	0.97	0.89	0.98	0.91	0.90	0.69	0.93	0.98	0.90	0.98	0.61	0.98
	x (km)	590	36	56	48	50	560	450	590	560	480	590	330
	R_2	0.81	0.49	0.52	0.47	0.30	0.43	0.54	0.74	0.48	0.54	0.52	0.51
Birmingham	R_{coll}	0.98	0.90	0.90	0.90	0.90	0.90	0.95	0.95	0.95	0.90	0.80	0.90
	x (km)	340	50	80	60	40	220	710	910	500	500	120	170
	R_2	0.82	0.52	0.55	0.25	0.39	0.53	0.57	0.77	0.30	0.52	0.51	0.52
Dallas	R_{coll}	0.98	0.90	0.90	0.90	0.90	0.90	0.95	0.98	0.95	0.95	0.80	0.90
	x (km)	770	200	260	140	20	430	1000	1400	530	560	110	630
	R_2	0.84	0.55	0.49	0.51	0.20	0.27	0.42	0.73	0.58	0.54	0.53	0.48
Pittsburgh	R_{coll}	0.98	0.90	0.90	0.90	0.90	0.90	0.95	0.95	0.95	0.90	0.80	0.90
	x (km)	710	90	120	60	30	260	630	480	560	210	80	220
	R_2	0.85	0.47	0.49	0.41	0.40	0.65	0.68	0.84	0.61	0.55	0.48	0.48
St. Louis	R_{coll}	0.98	0.90	0.90	0.90	0.90	0.90	0.95	0.98	0.98	0.98	0.80	0.90
	x (km)	480	60	80	40	20	420	590	2000	2500	910	60	80
	R_2	0.86	0.45	0.41	0.30	0.26	0.50	0.62	0.82	0.62	0.52	0.51	0.52

Table B.4 – Performance statistics for simulation-based observations categorized by metropolitan area and species, covering 2002-2008.

Pollutant	City	MB		RMSE		SD		Temporal R^2		Spatial R^2		Spatiotemporal R^2	
		C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}
CO	Atlanta	-0.14	-0.02	0.52	0.22	0.37	0.58	0.28	0.91	0.81	0.84	0.40	0.89
	Birmingham	-0.74	-0.83	1.96	2.02	0.34	0.59	0.11	0.47	0.66	0.53	0.06	0.27
	Dallas	-0.01	-0.13	0.43	0.30	0.32	0.41	0.26	0.73	0.03	0.35	0.20	0.67
	Pittsburgh	-0.15	-0.12	0.42	0.32	0.24	0.36	0.18	0.63	0.77	0.82	0.27	0.59
	St. Louis	-0.26	-0.29	0.71	0.50	0.35	0.42	0.10	0.80	0.02	0.02	0.08	0.65
NO_2	Atlanta	6.23	0.25	13.41	6.18	17.96	14.30	0.37	0.89	0.89	0.90	0.59	0.87
	Birmingham	-5.38	-10.73	11.79	17.01	12.99	8.06	0.29	0.50	1.00	1.00	0.66	0.53
	Dallas	1.79	-6.56	11.99	11.17	15.04	11.35	0.31	0.51	0.74	0.71	0.41	0.51
	Pittsburgh	-4.08	-4.78	11.99	10.78	12.42	12.57	0.23	0.58	0.58	0.48	0.33	0.48
	St. Louis	7.13	-0.92	15.16	5.26	14.80	11.29	0.22	0.87	0.83	0.81	0.28	0.83
NO_X	Atlanta	-0.02	-0.01	0.07	0.04	0.04	0.06	0.34	0.92	0.72	0.72	0.39	0.79
	Birmingham	-0.02	-0.02	0.05	0.05	0.02	0.02	0.32	0.60	1.00	1.00	0.55	0.60
	Dallas	-0.01	-0.01	0.03	0.03	0.03	0.04	0.25	0.62	0.72	0.70	0.33	0.65
	Pittsburgh	-0.03	-0.02	0.06	0.04	0.02	0.05	0.25	0.76	0.47	0.42	0.26	0.61
	St. Louis	-0.01	0.00	0.04	0.03	0.03	0.04	0.18	0.85	0.56	0.72	0.17	0.66
SO_2	Atlanta	-1.84	0.73	15.30	10.71	8.84	16.07	0.13	0.88	0.16	0.01	0.08	0.59
	Birmingham	-3.73	-0.88	13.79	9.68	6.99	11.28	0.17	0.57	0.97	0.96	0.17	0.56
	Dallas	1.11	0.60	11.76	10.34	4.94	6.62	0.05	0.43	0.00	0.00	0.00	0.13
	Pittsburgh	-8.41	-1.67	19.56	12.56	9.07	14.35	0.17	0.64	0.00	0.02	0.14	0.57
	St. Louis	2.30	2.00	29.88	25.74	15.79	22.59	0.08	0.75	0.00	0.01	0.01	0.21
O_3	Atlanta	0.00	0.00	0.01	0.00	0.01	0.02	0.67	0.99	0.68	0.72	0.67	0.97
	Birmingham	0.01	0.00	0.01	0.01	0.01	0.01	0.67	0.91	0.45	0.46	0.67	0.90
	Dallas	0.00	0.00	0.01	0.00	0.01	0.02	0.70	0.93	0.28	0.18	0.69	0.93
	Pittsburgh	0.00	0.00	0.01	0.00	0.02	0.02	0.72	0.95	0.65	0.67	0.73	0.94
	St. Louis	0.00	0.00	0.01	0.00	0.02	0.02	0.74	0.96	0.84	0.87	0.75	0.95
PM_{10}	Atlanta	-5.42	-0.07	11.98	4.85	8.69	10.54	0.21	0.85	0.05	0.11	0.22	0.81
	Birmingham	-12.31	-6.37	23.96	19.12	9.87	11.90	0.30	0.71	0.34	0.23	0.18	0.38
	Dallas	-0.80	-0.43	13.87	4.63	11.93	9.97	0.14	0.92	0.02	0.01	0.09	0.84
	Pittsburgh	-7.15	-0.45	15.44	8.72	8.27	12.74	0.42	0.87	0.06	0.00	0.33	0.74
	St. Louis	-0.99	-2.36	12.69	5.64	10.92	9.45	0.26	0.93	1.00	1.00	0.14	0.82
$PM_{2.5}$	Atlanta	-1.72	0.24	7.23	2.65	7.20	7.65	0.30	0.93	0.63	0.29	0.31	0.89
	Birmingham	-1.43	-0.95	7.19	3.06	7.97	7.59	0.33	0.90	0.61	0.64	0.39	0.88
	Dallas	1.84	0.42	7.12	2.21	6.79	5.64	0.19	0.90	0.22	0.38	0.15	0.86
	Pittsburgh	-1.81	-0.47	7.86	3.76	6.67	8.87	0.46	0.92	0.13	0.02	0.42	0.86
	St. Louis	1.25	-0.17	6.93	2.62	7.02	7.29	0.40	0.95	0.21	0.16	0.36	0.90
$PM_{2.5}$-SO_4	Atlanta	-0.70	0.20	2.01	0.50	2.20	3.13	0.55	0.98	0.82	0.59	0.57	0.98
	Birmingham	-0.38	-0.30	1.94	1.08	2.58	2.64	0.59	0.90	0.85	0.91	0.58	0.87
	Dallas	-0.60	0.00	1.62	0.41	1.50	2.08	0.53	0.97	0.02	0.36	0.50	0.96
	Pittsburgh	-0.60	-0.18	2.31	1.07	3.36	3.70	0.70	0.94	0.00	0.13	0.71	0.94
	St. Louis	0.04	-0.07	1.92	0.70	2.66	3.02	0.67	0.98	0.09	0.14	0.64	0.95
$PM_{2.5}$-NO_3	Atlanta	0.50	-0.17	1.42	0.26	1.58	0.57	0.24	0.96	0.65	0.80	0.31	0.97
	Birmingham	0.06	0.06	0.90	0.36	1.07	0.74	0.30	0.93	0.71	0.93	0.31	0.77
	Dallas	0.76	-0.09	1.73	0.40	1.99	0.98	0.35	0.95	0.28	0.09	0.39	0.89
	Pittsburgh	0.17	0.84	1.35	1.25	1.66	1.94	0.38	0.86	0.50	0.35	0.38	0.85
	St. Louis	0.16	0.03	2.04	0.43	2.42	2.22	0.38	0.98	0.28	0.41	0.39	0.96
$PM_{2.5}$-NH_4	Atlanta	0.19	-0.03	0.86	0.17	0.86	0.90	0.31	0.98	0.68	0.41	0.30	0.97
	Birmingham	-0.05	0.02	0.78	0.46	0.76	0.88	0.37	0.81	0.51	0.82	0.35	0.75
	Dallas	0.20	-0.01	0.86	0.18	0.84	0.81	0.30	0.97	0.02	0.35	0.27	0.96
	Pittsburgh	-0.13	-0.24	1.18	0.72	0.95	1.11	0.31	0.83	0.53	0.16	0.31	0.76
	St. Louis	0.29	-0.03	1.04	0.19	1.10	1.26	0.43	0.98	0.72	0.78	0.44	0.98
$PM_{2.5}$-EC	Atlanta	-0.16	0.16	0.61	0.43	0.62	0.86	0.29	0.76	0.82	0.77	0.42	0.79
	Birmingham	-0.42	-0.13	0.98	0.68	0.52	0.82	0.26	0.66	0.82	0.83	0.34	0.62
	Dallas	0.20	-0.09	0.45	0.28	0.47	0.36	0.29	0.70	0.87	0.87	0.35	0.58
	Pittsburgh	-0.13	0.03	1.13	0.86	0.46	0.73	0.24	0.76	0.30	0.38	0.11	0.48
	St. Louis	-0.07	-0.09	0.68	0.55	0.44	0.48	0.24	0.72	0.09	0.17	0.12	0.37
$PM_{2.5}$-OC	Atlanta	-0.96	-0.39	2.08	0.64	1.98	1.70	0.29	0.97	0.74	0.48	0.32	0.95
	Birmingham	-1.95	0.21	3.12	1.66	1.93	2.29	0.23	0.78	0.32	0.38	0.18	0.60
	Dallas	-0.96	-0.09	1.82	0.81	1.37	1.54	0.23	0.85	0.01	0.21	0.16	0.74
	Pittsburgh	-2.51	-0.21	3.23	1.05	0.87	1.96	0.23	0.87	0.34	0.31	0.21	0.79
	St. Louis	-2.04	-0.02	2.57	0.71	0.92	1.69	0.25	0.86	0.55	0.53	0.27	0.85

Table B.5 – Descriptive statistics for cross-validated simulation-based observations categorized by metropolitan area and species.

Pollutant	City	MB		RMSE		SD		Temporal R^2		Spatial R^2		Spatiotemporal R	
		C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}	C_{CMAQ}	C_{opt}
CO	Atlanta	-0.15	0.08	0.53	0.46	0.37	0.58	0.30	0.48	0.78	0.82	0.38	0.53
	Birmingham	-0.99	-0.71	2.40	2.52	0.36	1.01	0.12	0.06	0.44	0.01	0.03	0.00
	Dallas	-0.10	0.00	0.40	0.35	0.33	0.37	0.24	0.45	0.09	0.03	0.21	0.34
	Pittsburgh	-0.19	0.02	0.49	0.46	0.25	0.45	0.13	0.26	0.74	0.51	0.15	0.26
	St. Louis	-0.26	-0.23	0.64	0.58	0.37	0.45	0.12	0.33	0.07	0.01	0.12	0.28
NO_2	Atlanta	6.18	5.10	13.58	13.60	18.12	22.43	0.37	0.54	0.92	0.85	0.59	0.69
	Birmingham	-8.93	-9.59	16.30	15.14	11.65	8.97	0.21	0.33	1.00	1.00	0.45	0.72
	Dallas	-2.75	-2.13	13.35	9.78	14.36	13.04	0.23	0.56	0.62	0.60	0.29	0.52
	Pittsburgh	-6.70	0.12	14.04	13.39	12.49	17.17	0.20	0.46	0.51	0.49	0.26	0.40
	St. Louis	5.40	0.89	14.35	9.94	14.47	12.32	0.24	0.47	0.84	0.85	0.28	0.48
NO_X	Atlanta	-0.02	0.01	0.07	0.07	0.04	0.10	0.37	0.51	0.73	0.57	0.40	0.57
	Birmingham	-0.03	-0.02	0.06	0.05	0.01	0.02	0.27	0.35	1.00	1.00	0.43	0.57
	Dallas	-0.02	0.00	0.04	0.03	0.02	0.05	0.23	0.56	0.68	0.55	0.29	0.51
	Pittsburgh	-0.04	0.00	0.06	0.05	0.02	0.07	0.25	0.59	0.42	0.31	0.22	0.41
	St. Louis	-0.01	0.01	0.05	0.05	0.03	0.04	0.21	0.53	0.86	0.43	0.26	0.29
SO_2	Atlanta	-0.76	2.44	16.68	16.26	10.80	14.70	0.17	0.33	0.01	0.00	0.06	0.19
	Birmingham	-4.06	0.95	15.13	16.05	7.03	12.64	0.11	0.16	0.98	1.00	0.09	0.12
	Dallas	0.84	2.19	11.86	15.32	5.04	10.01	0.07	0.07	0.00	0.00	0.00	0.00
	Pittsburgh	-8.89	-2.19	20.96	18.37	10.56	12.63	0.15	0.29	0.02	0.14	0.11	0.19
	St. Louis	1.61	0.91	29.91	33.16	16.05	22.77	0.10	0.16	0.01	0.00	0.00	0.01
O_3	Atlanta	0.00	0.00	0.01	0.01	0.02	0.02	0.67	0.91	0.69	0.54	0.68	0.89
	Birmingham	0.01	0.00	0.01	0.01	0.01	0.01	0.64	0.89	0.45	0.46	0.66	0.88
	Dallas	0.00	0.00	0.01	0.00	0.01	0.02	0.67	0.94	0.28	0.52	0.67	0.93
	Pittsburgh	0.00	0.00	0.01	0.01	0.01	0.02	0.71	0.92	0.54	0.46	0.71	0.88
	St. Louis	0.00	0.00	0.01	0.01	0.02	0.02	0.73	0.93	0.89	0.59	0.74	0.89
PM_{10}	Atlanta	-4.56	0.00	11.10	7.56	8.49	10.83	0.26	0.63	0.22	0.27	0.29	0.60
	Birmingham	-13.57	-4.25	24.49	21.62	10.64	14.80	0.23	0.49	0.46	0.33	0.13	0.14
	Dallas	0.11	1.48	15.67	6.61	12.51	10.82	0.34	0.85	0.04	0.76	0.03	0.71
	Pittsburgh	-8.38	0.20	16.07	9.14	8.14	14.17	0.37	0.79	0.04	0.21	0.28	0.68
	St. Louis	1.73	-0.75	13.65	8.88	12.15	10.35	0.27	0.58	1.00	1.00	0.12	0.46
$PM_{2.5}$	Atlanta	-1.40	0.49	7.17	3.71	7.03	7.94	0.34	0.85	0.27	0.15	0.31	0.80
	Birmingham	-2.61	0.07	7.69	3.57	7.74	8.61	0.33	0.88	0.55	0.53	0.36	0.83
	Dallas	0.94	1.19	6.96	3.11	6.43	5.69	0.19	0.85	0.31	0.34	0.12	0.76
	Pittsburgh	-2.70	-0.32	8.25	5.80	6.65	8.70	0.41	0.81	0.12	0.22	0.39	0.67
	St. Louis	1.48	0.14	6.75	3.80	6.85	7.09	0.33	0.82	0.25	0.17	0.35	0.76
$PM_{2.5}$-SO_4	Atlanta	-0.55	0.28	1.91	1.35	2.30	3.00	0.59	0.82	0.68	0.88	0.62	0.81
	Birmingham	-0.45	0.02	1.96	1.27	2.81	2.82	0.66	0.86	0.49	0.49	0.59	0.81
	Dallas	-0.73	0.16	1.84	0.95	1.55	2.18	0.54	0.81	0.09	0.42	0.44	0.83
	Pittsburgh	-0.38	0.15	1.77	1.32	2.90	2.95	0.72	0.85	0.84	0.90	0.71	0.83
	St. Louis	0.08	0.13	1.73	1.00	2.56	2.96	0.71	0.92	0.54	0.63	0.66	0.89
$PM_{2.5}$-NO_3	Atlanta	0.61	-0.16	1.48	0.51	1.59	0.65	0.29	0.66	0.91	0.85	0.28	0.61
	Birmingham	-0.01	0.16	0.90	0.62	1.01	0.90	0.25	0.73	0.84	0.72	0.26	0.56
	Dallas	1.13	-0.16	2.11	0.89	2.42	1.03	0.51	0.79	0.49	0.28	0.47	0.63
	Pittsburgh	0.05	1.06	1.62	1.83	1.76	2.15	0.36	0.69	0.95	0.34	0.24	0.52
	St. Louis	0.33	0.07	1.96	1.30	2.54	2.40	0.45	0.78	0.91	0.86	0.46	0.72
$PM_{2.5}$-NH_4	Atlanta	0.23	0.00	0.88	0.48	0.86	0.87	0.31	0.75	0.74	0.78	0.28	0.72
	Birmingham	-0.08	0.20	0.81	0.86	0.82	1.06	0.34	0.61	0.74	0.37	0.32	0.43
	Dallas	0.25	0.11	1.01	0.58	0.94	0.90	0.30	0.76	0.36	0.65	0.24	0.68
	Pittsburgh	-0.22	-0.08	1.35	1.07	0.95	1.08	0.26	0.51	0.72	0.71	0.19	0.46
	St. Louis	0.31	0.04	1.03	0.60	1.12	1.20	0.48	0.80	0.33	1.00	0.48	0.80
$PM_{2.5}$-EC	Atlanta	-0.13	0.15	0.66	0.63	0.66	0.85	0.32	0.43	0.76	0.84	0.42	0.55
	Birmingham	-0.49	-0.02	1.03	0.87	0.53	0.90	0.22	0.34	0.79	0.53	0.26	0.37
	Dallas	0.05	0.04	0.46	0.40	0.49	0.43	0.30	0.49	0.75	0.71	0.30	0.36
	Pittsburgh	-0.27	0.32	1.13	1.48	0.49	1.20	0.15	0.25	0.64	0.12	0.14	0.07
	St. Louis	-0.06	-0.04	0.64	0.70	0.49	0.56	0.29	0.34	0.10	0.01	0.16	0.11
$PM_{2.5}$-OC	Atlanta	-0.84	-0.44	1.98	1.44	1.97	1.64	0.34	0.56	0.35	0.00	0.33	0.50
	Birmingham	-2.07	0.63	3.43	2.82	1.85	3.23	0.16	0.44	0.05	0.02	0.08	0.33
	Dallas	-1.23	0.40	1.88	1.27	1.36	1.66	0.32	0.71	0.12	0.18	0.26	0.51
	Pittsburgh	-2.73	0.06	3.51	2.27	0.87	2.42	0.26	0.53	0.26	0.22	0.19	0.32
	St. Louis	-2.08	0.12	2.56	1.57	1.01	2.08	0.38	0.57	0.50	0.27	0.34	0.47

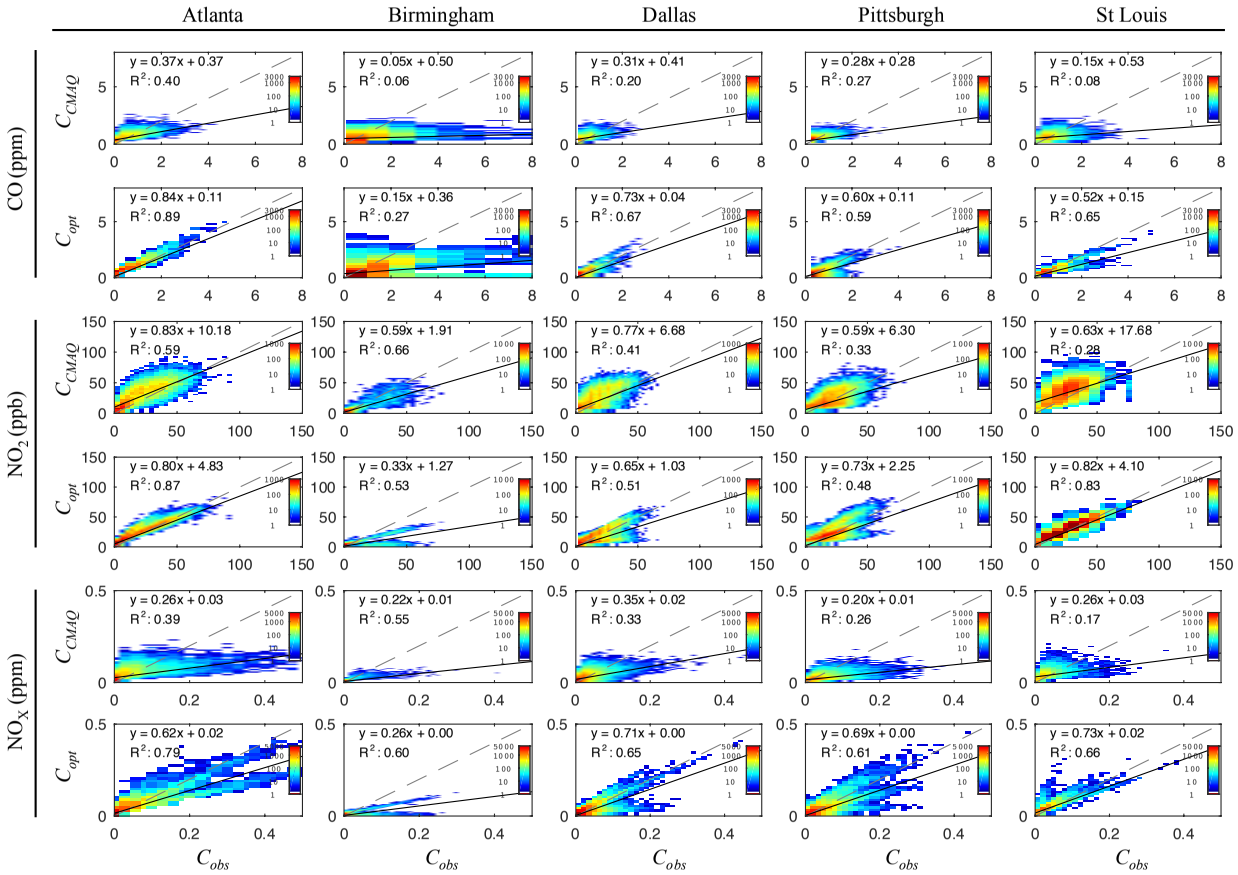

Figure B.1 – Density scatter plots of modeled results (C_{CMAQ}, C_{opt}) versus coincident observations for 2002-2008 per city and pollutant contoured using a color scale to show the range of point densities. The regression-line fits and correlation coefficients are given in the upper left of each plot. The solid line indicates the regression line, whereas the dashed line is the 1:1 line.

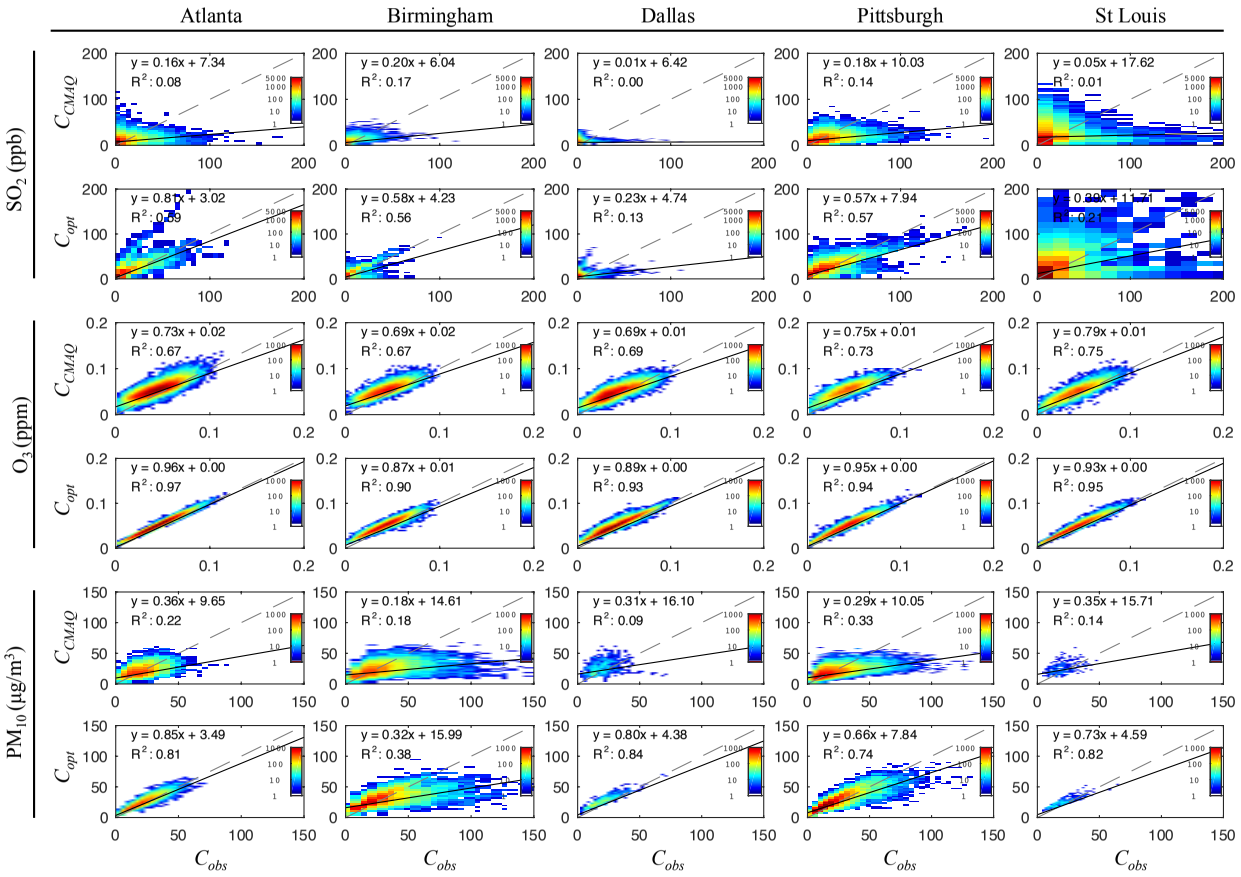

Figure B.1 – (*continued*)

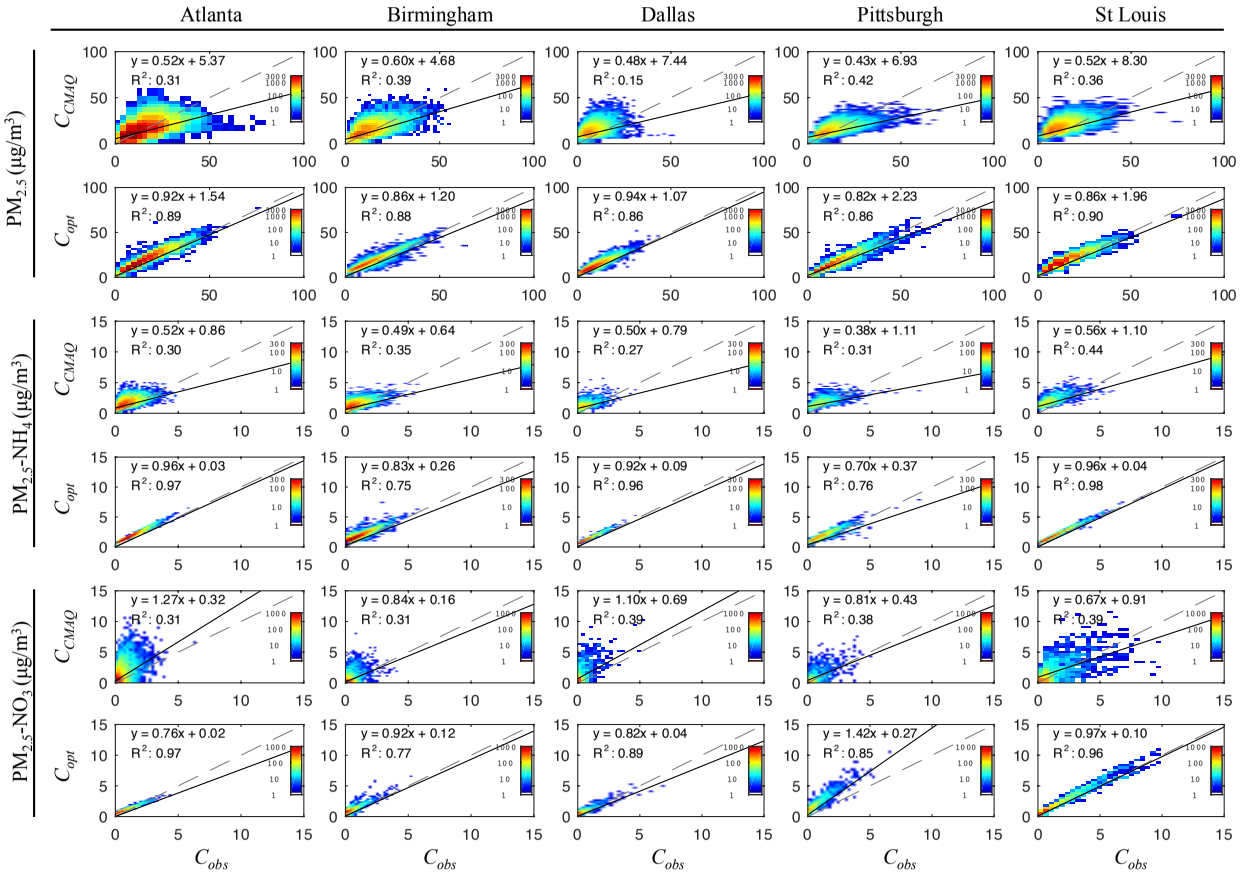

Figure B.1 – (*continued*)

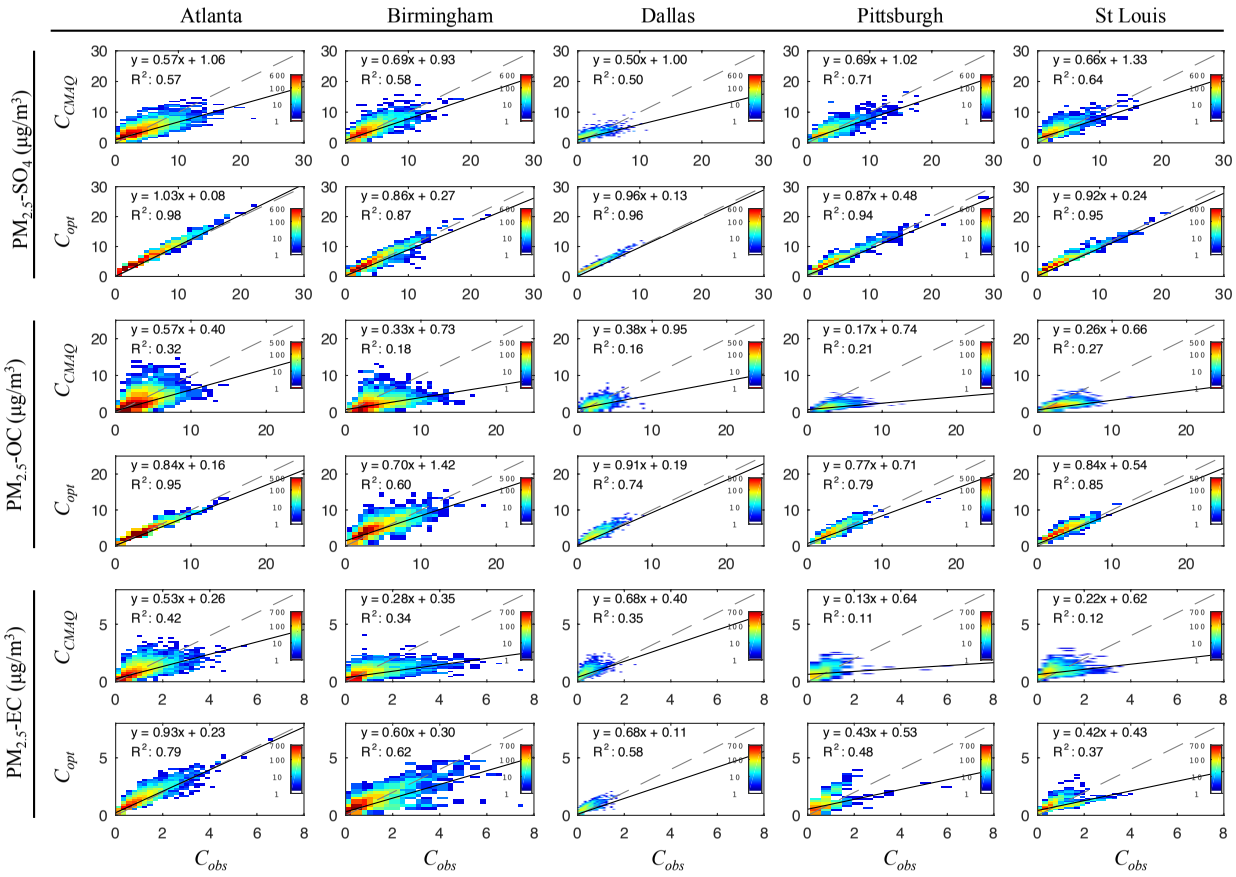

Figure B.1 – (*continued*)

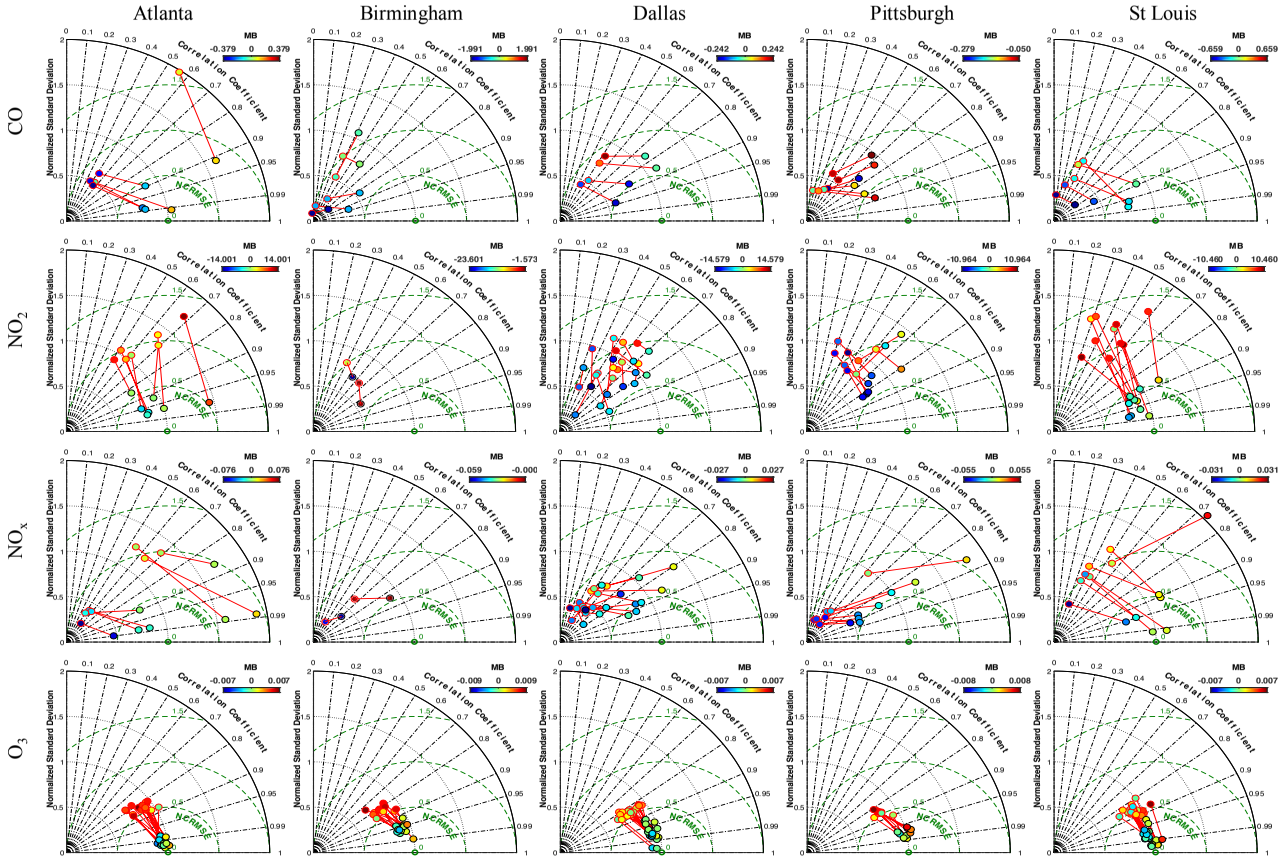

Figure B.2 – Temporal correlation coefficient (R) Taylor diagrams show normalized modeling statistics for each monitor separated by city and pollutant. C_{CMAQ} symbols are outlined in red, whereas C_{opt} symbols are outlined in black. The symbol color indicates the mean bias.

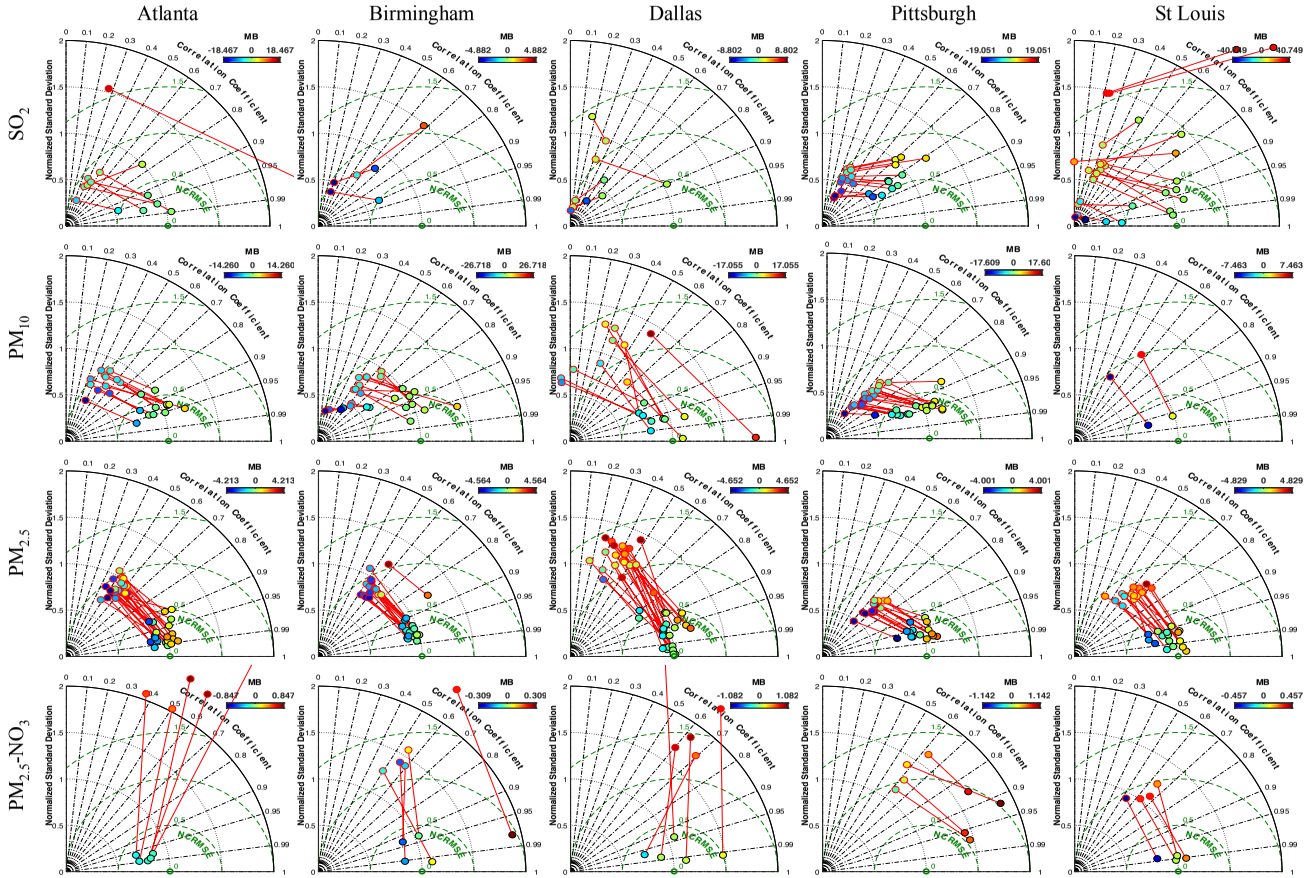

Figure B.2 – (*continued*)

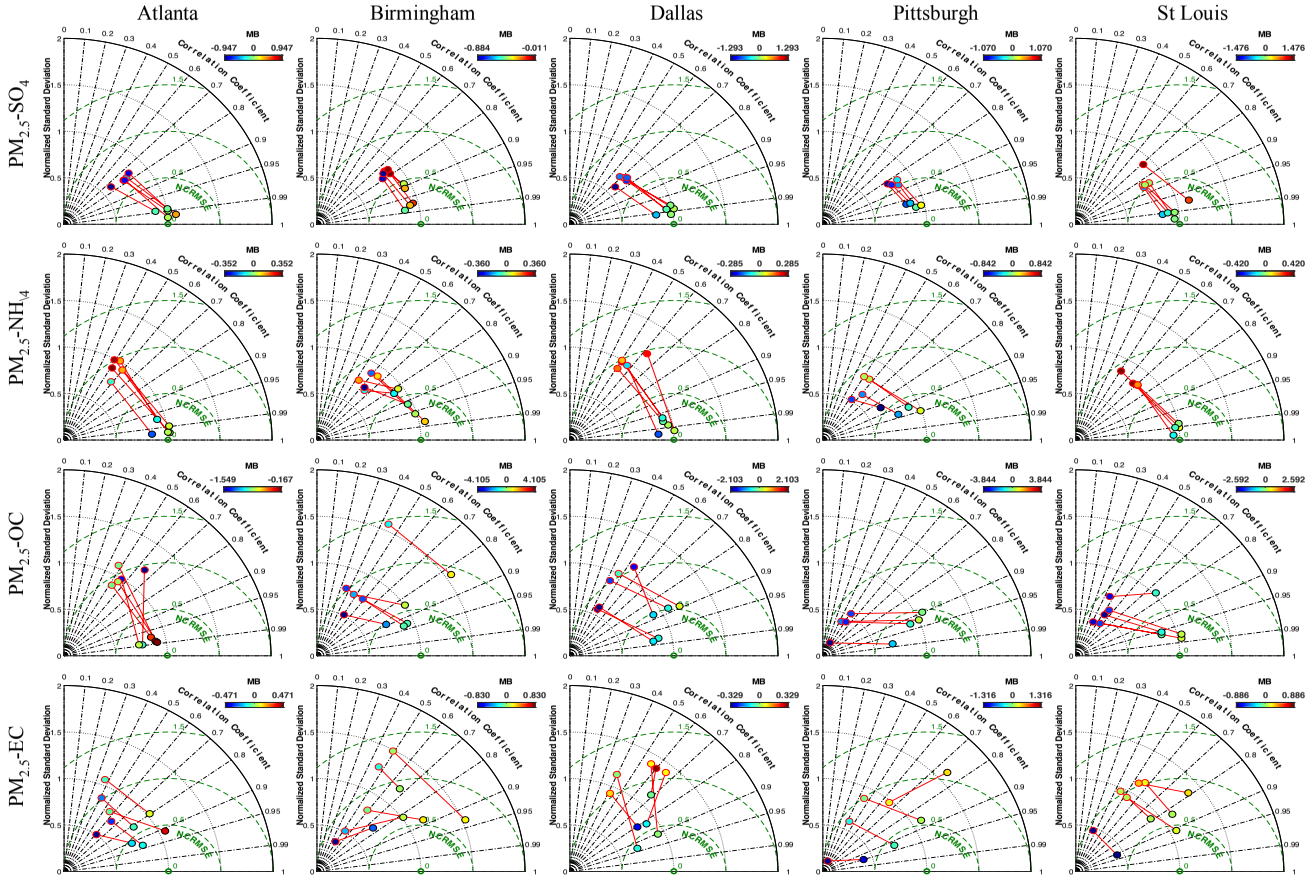

Figure B.2 – (*continued*)

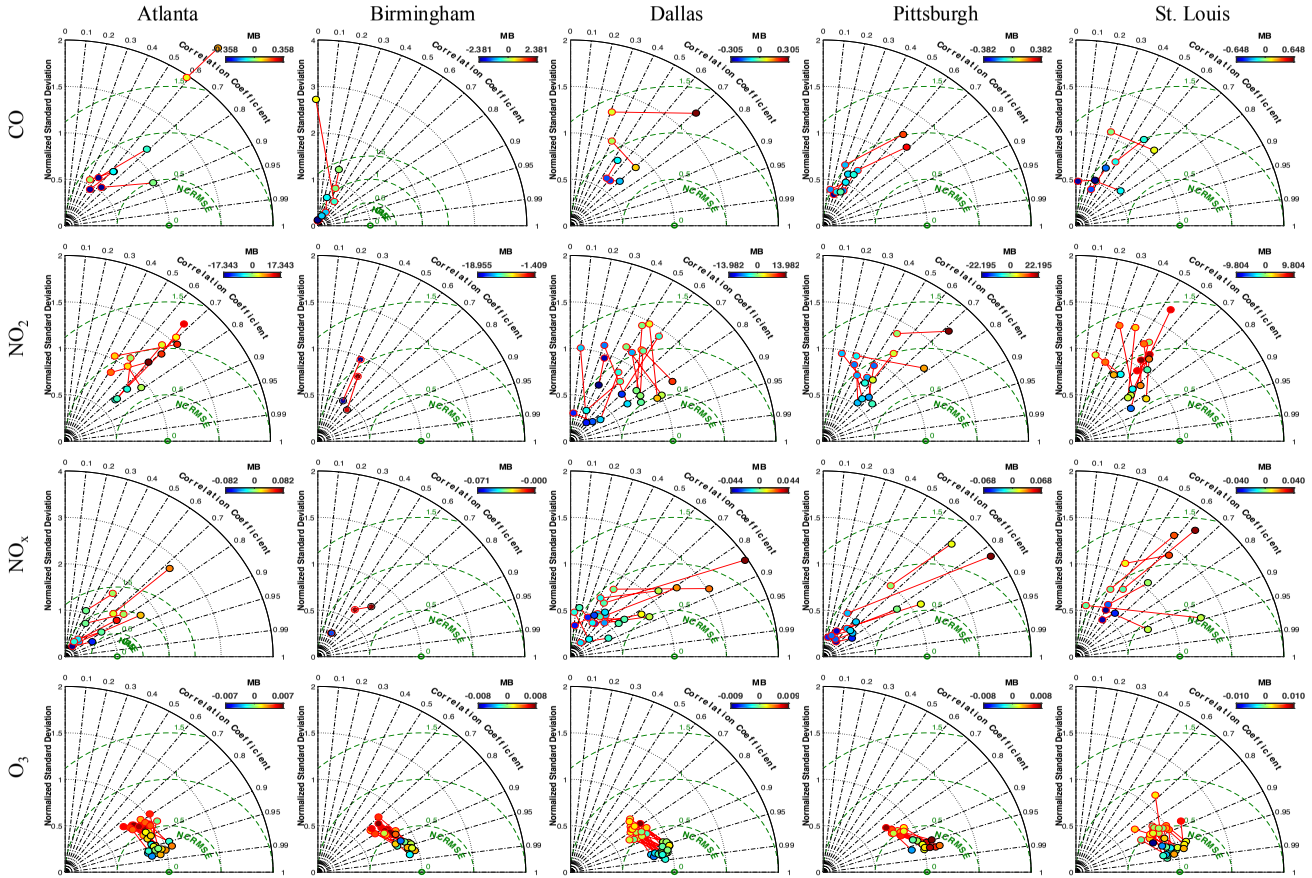

Figure B.3 – Cross-validation temporal correlation coefficient (R) Taylor diagrams show normalized modeling statistics for each monitor separated by city and pollutant. C_{CMAQ} symbols are outlined in red, whereas C_{opt} symbols are outlined in black. The symbol color indicates the mean bias.

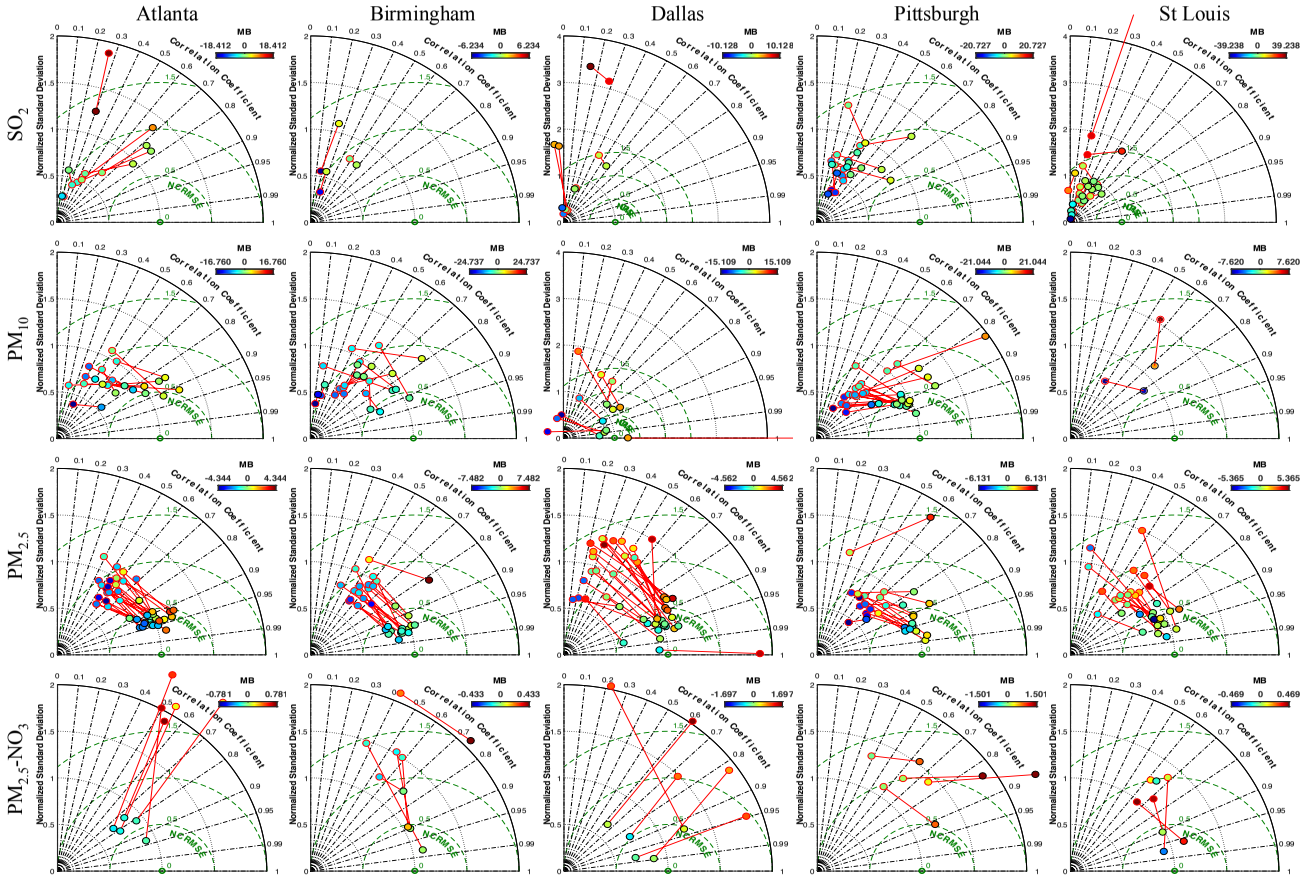

Figure B.3 – (*continued*)

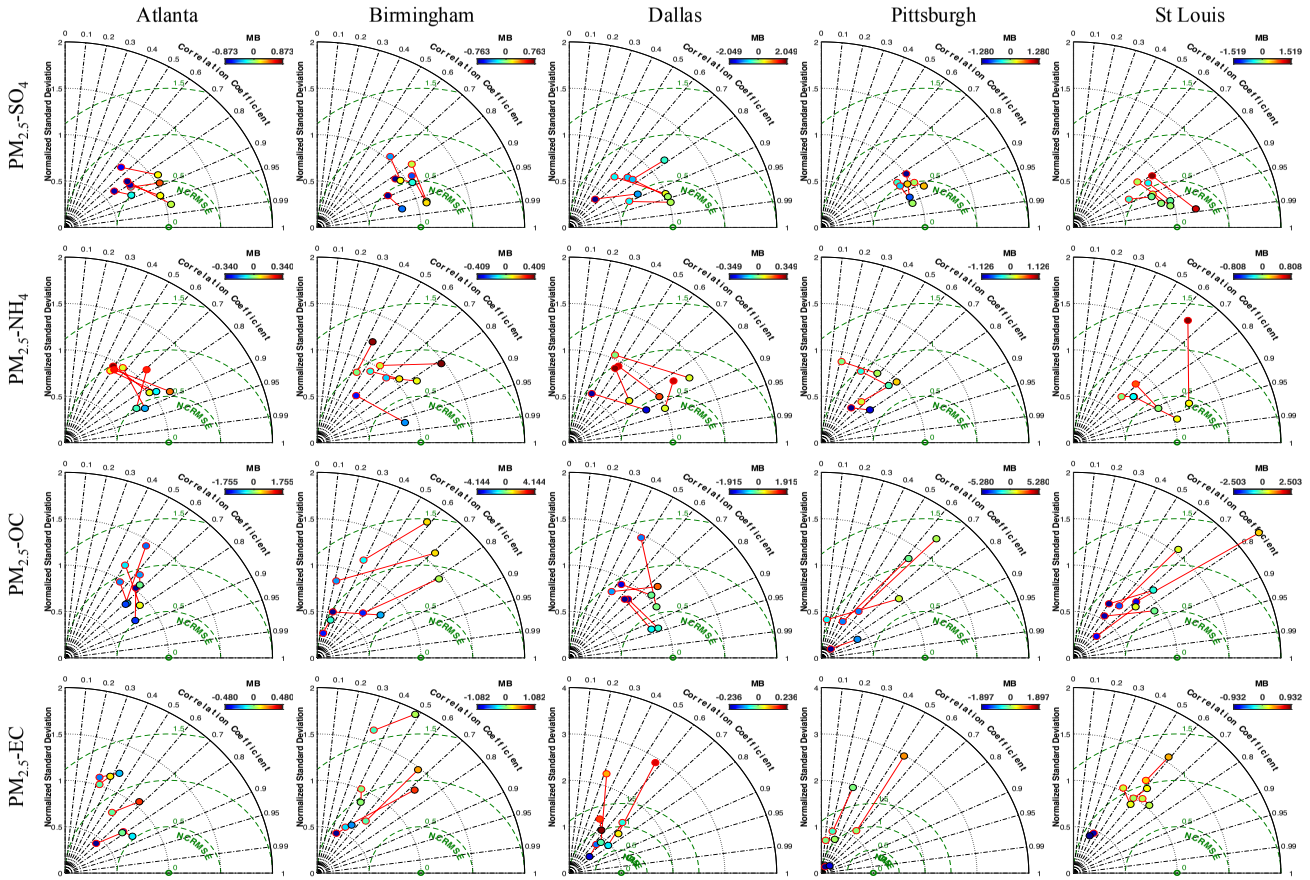

Figure B.3 – (*continued*)

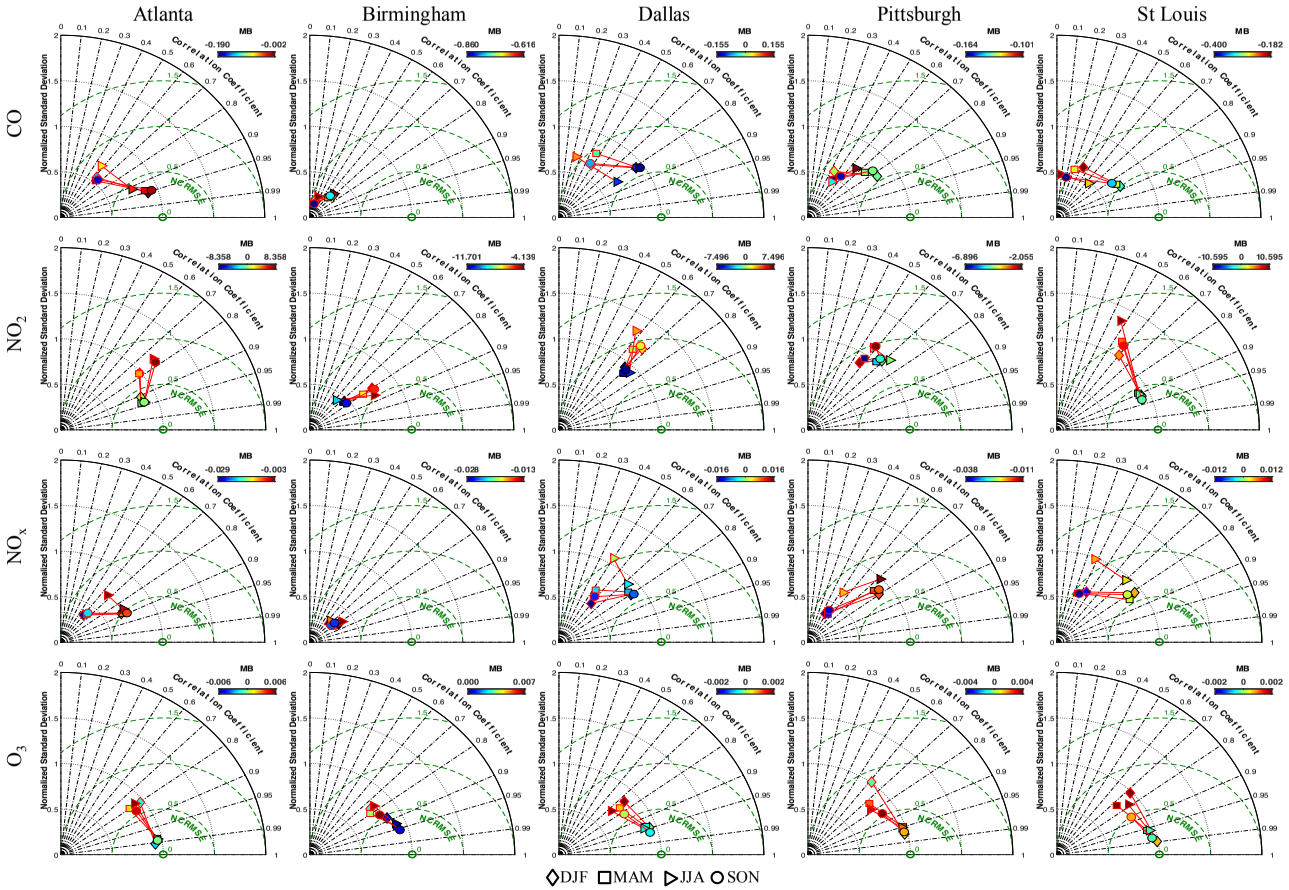

Figure B.4 – Seasonal spatiotemporal correlation coefficient (R) Taylor diagrams show normalized modeling statistics for each season separated by city and pollutant. C_{CMAQ} symbols are outlined in red, whereas C_{opt} symbols are outlined in black. The symbol color indicates the mean bias.

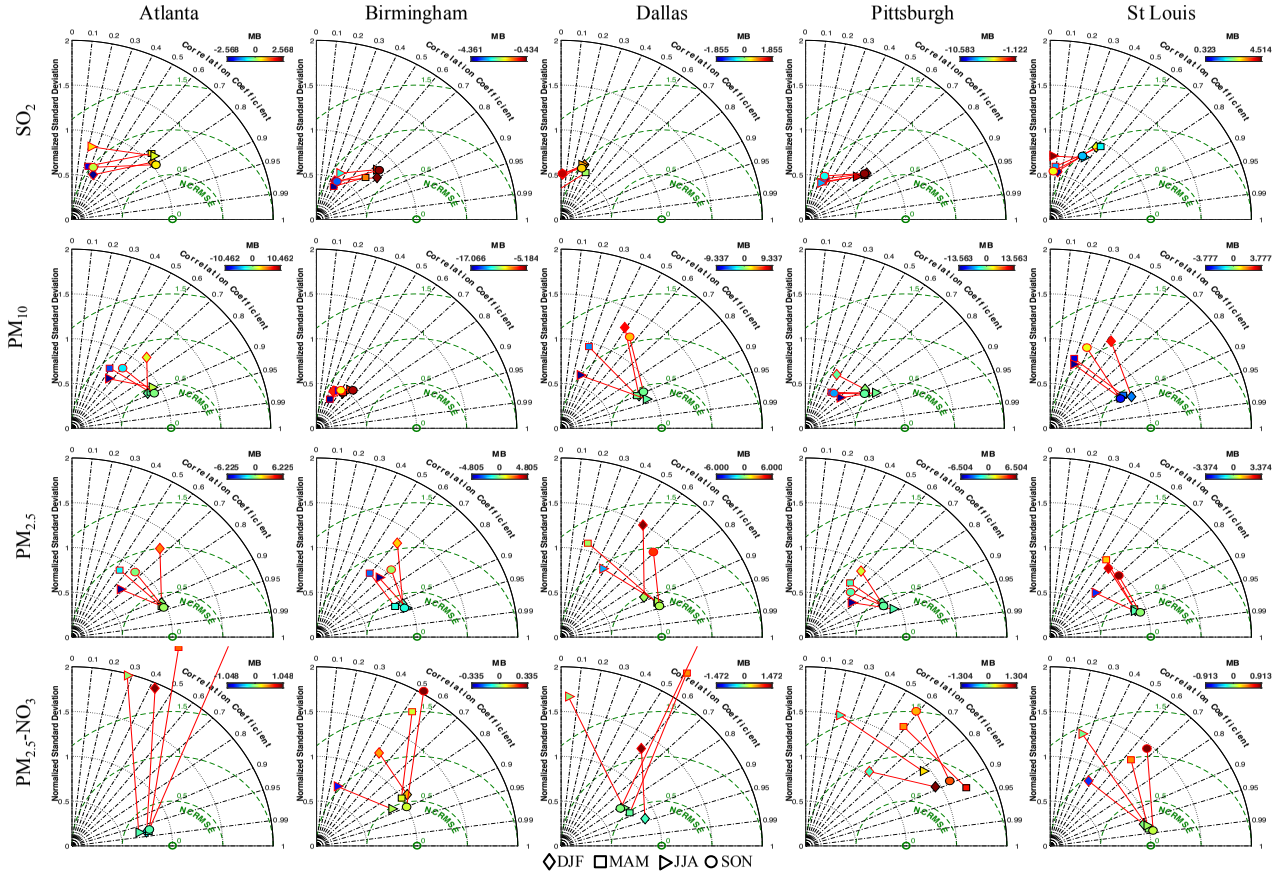

Figure B.4 – (*continued*)

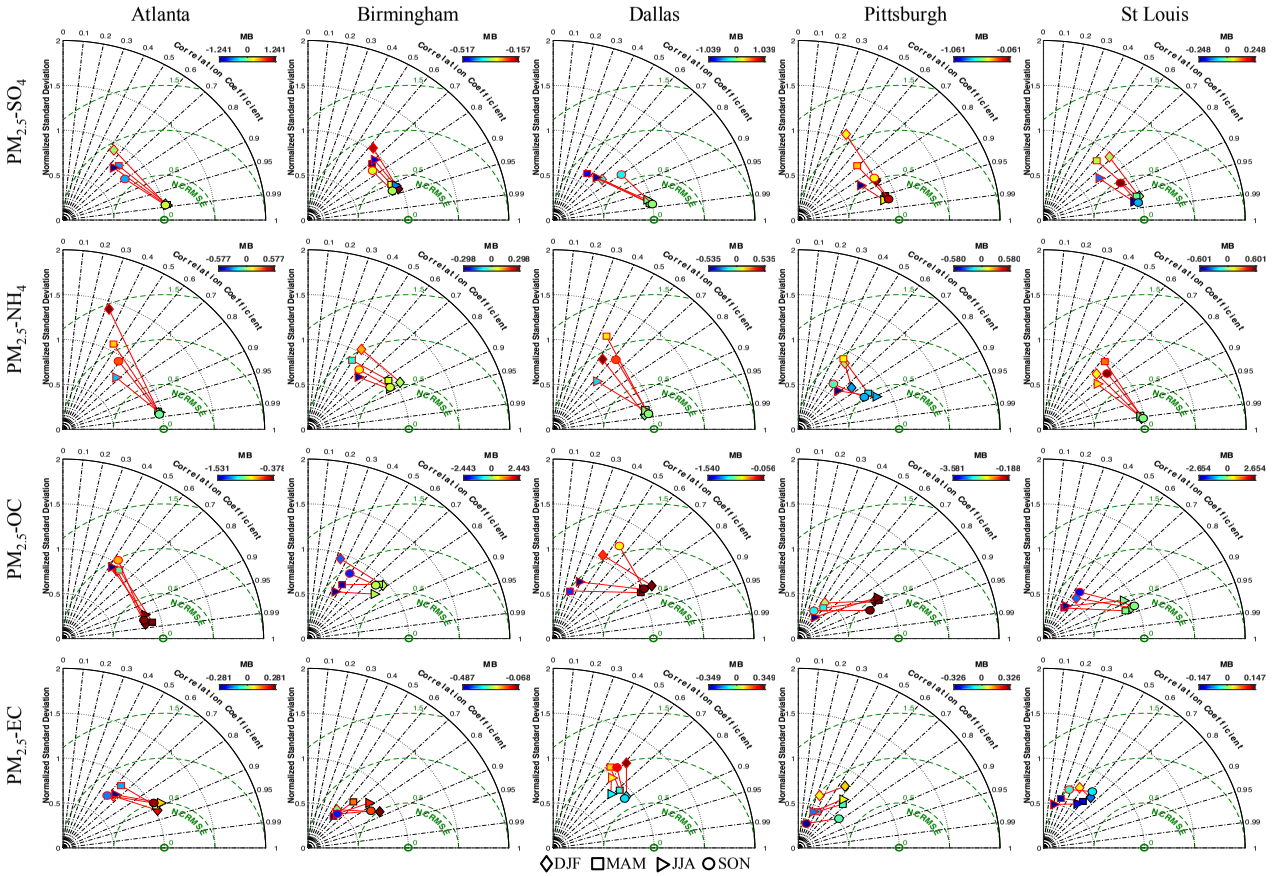

Figure B.4 – (*continued*)

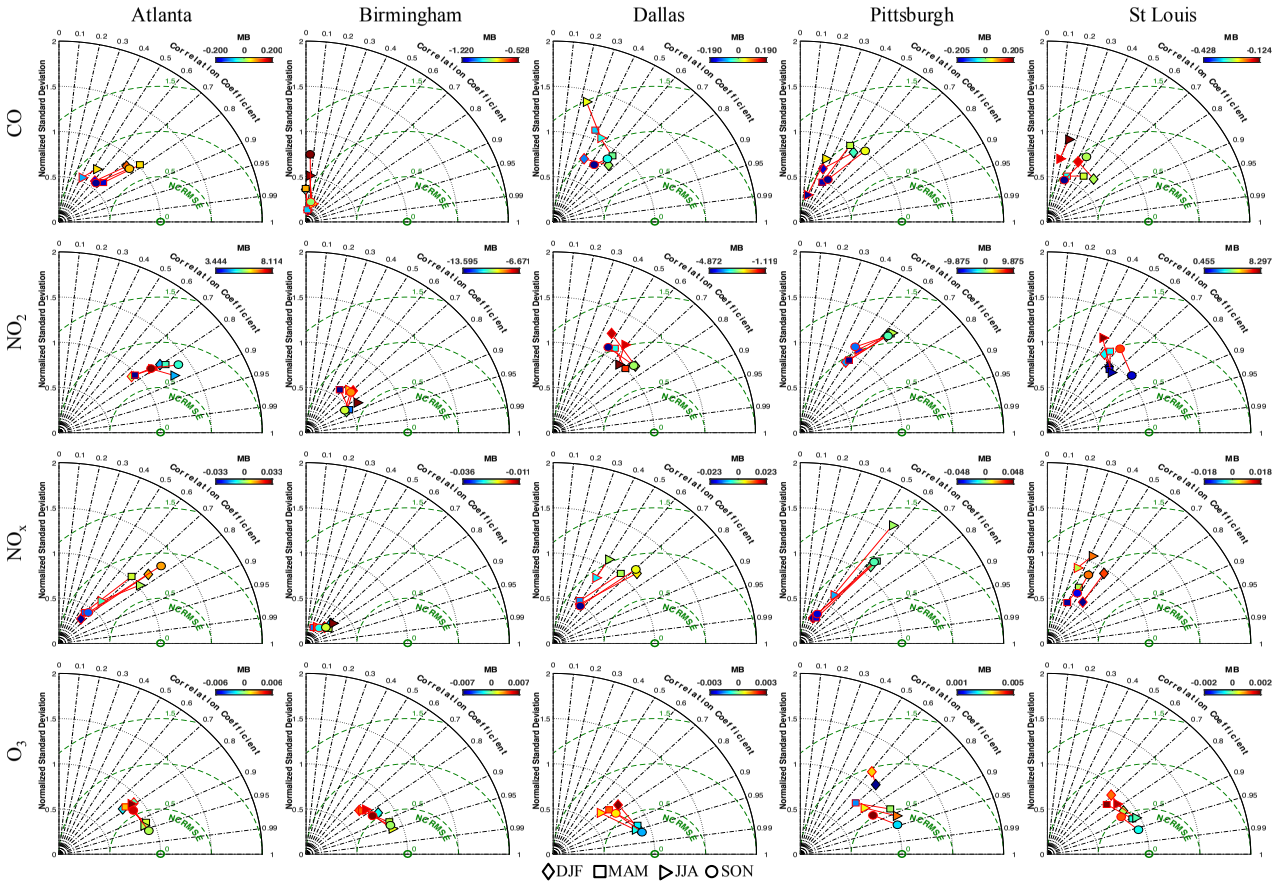

Figure B.5 – Cross-validation seasonal spatiotemporal correlation coefficient (R) Taylor diagrams show normalized statistics for each season separated by city and pollutant. C_{CMAQ} symbols are outlined in red, whereas C_{opt} symbols are in black. The symbol color indicates the mean absolute error.

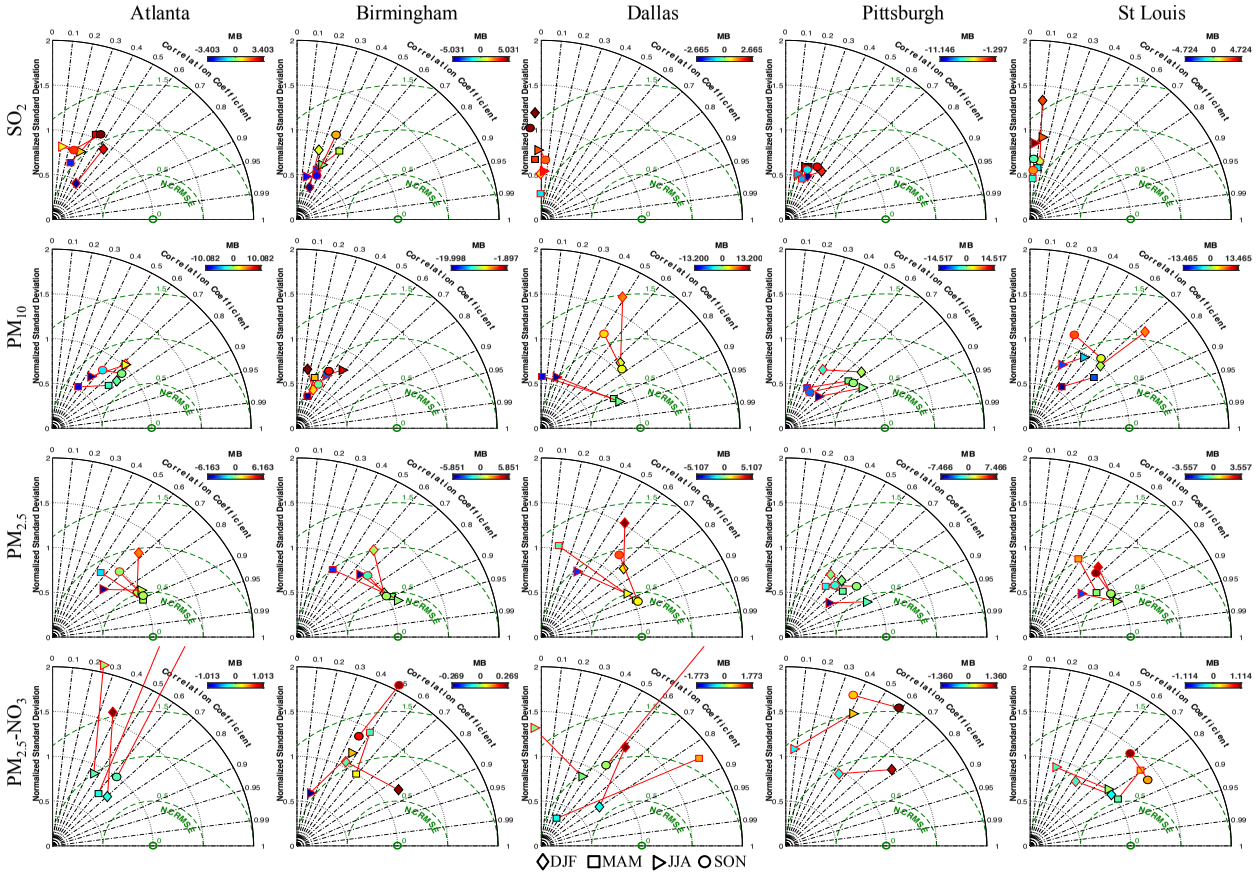

Figure B.5 – (*continued*)

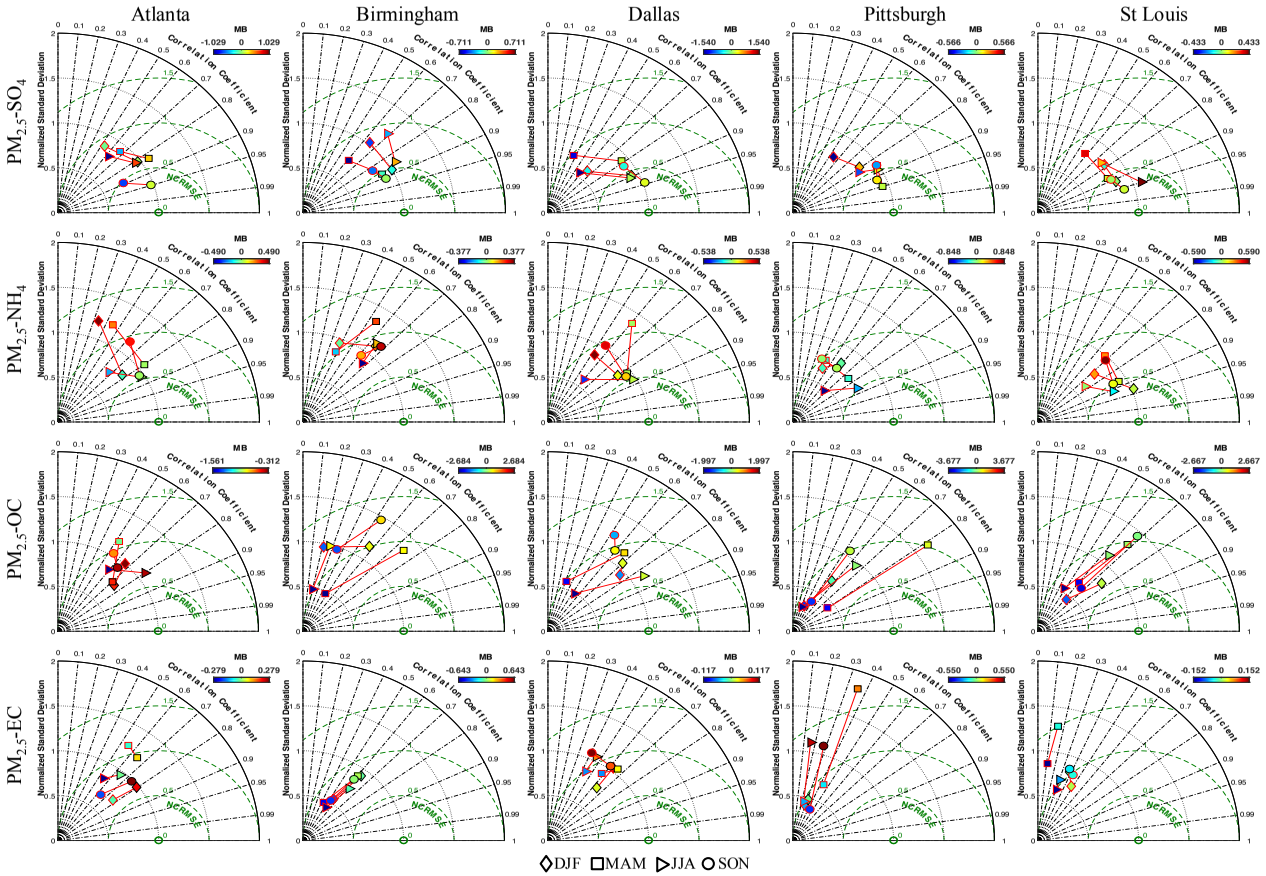

Figure B.5 – (*continued*)

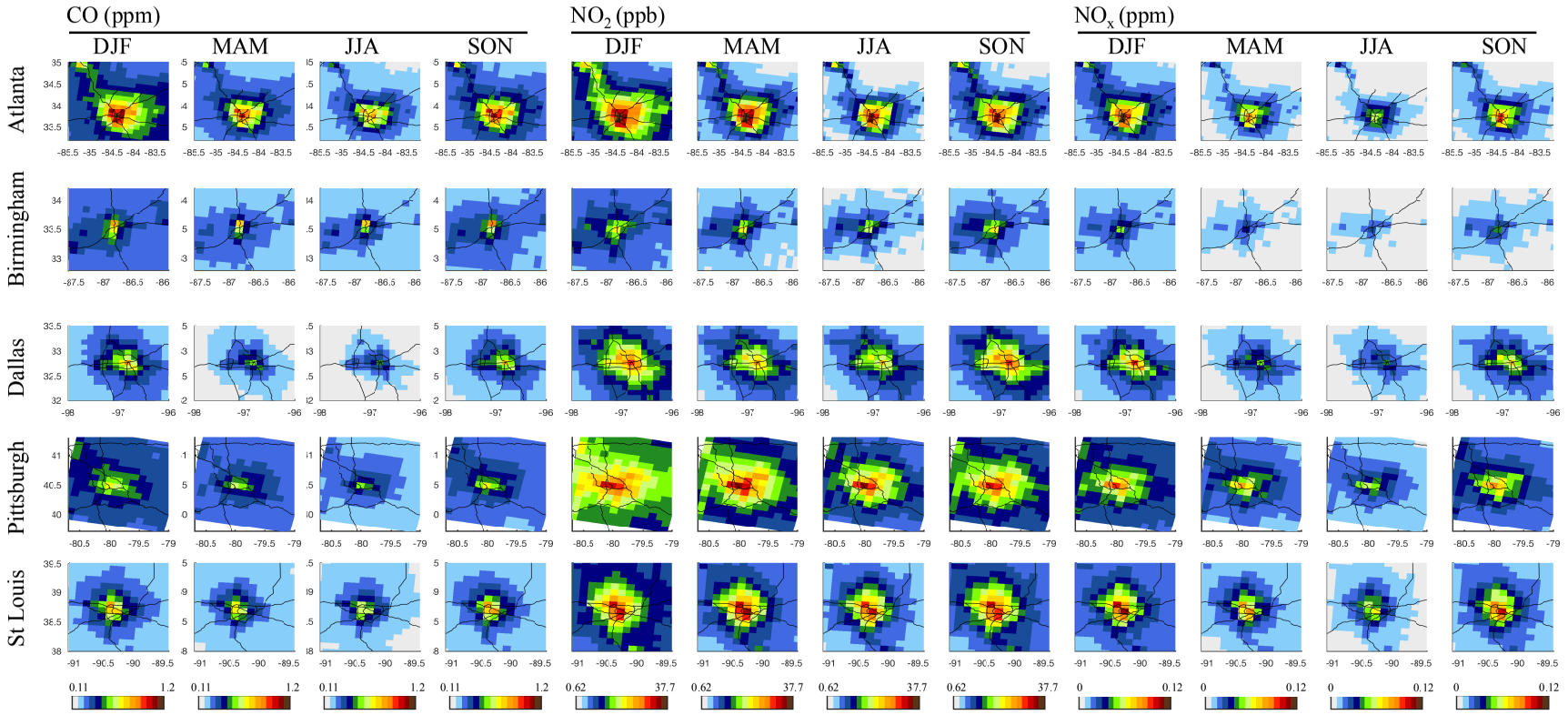

Figure B.6 – Mean seasonal optimized concentration fields (C_{opt}), 2002-2008. Seasons were calculated by averaging the results over three-month periods (DJF-Winter, MAM-Spring, JJA-Summer, and SON-Fall).

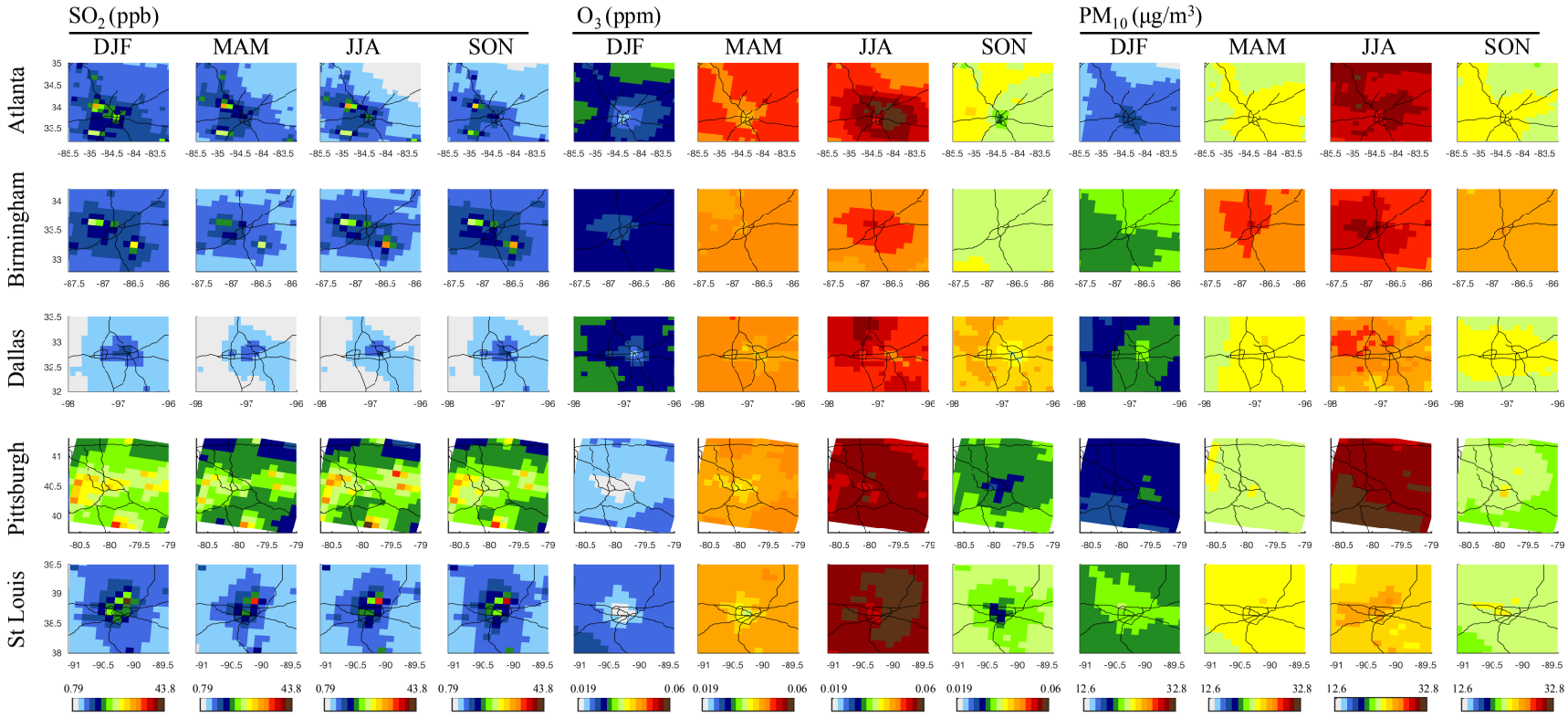

Figure B.6 – (*continued*)

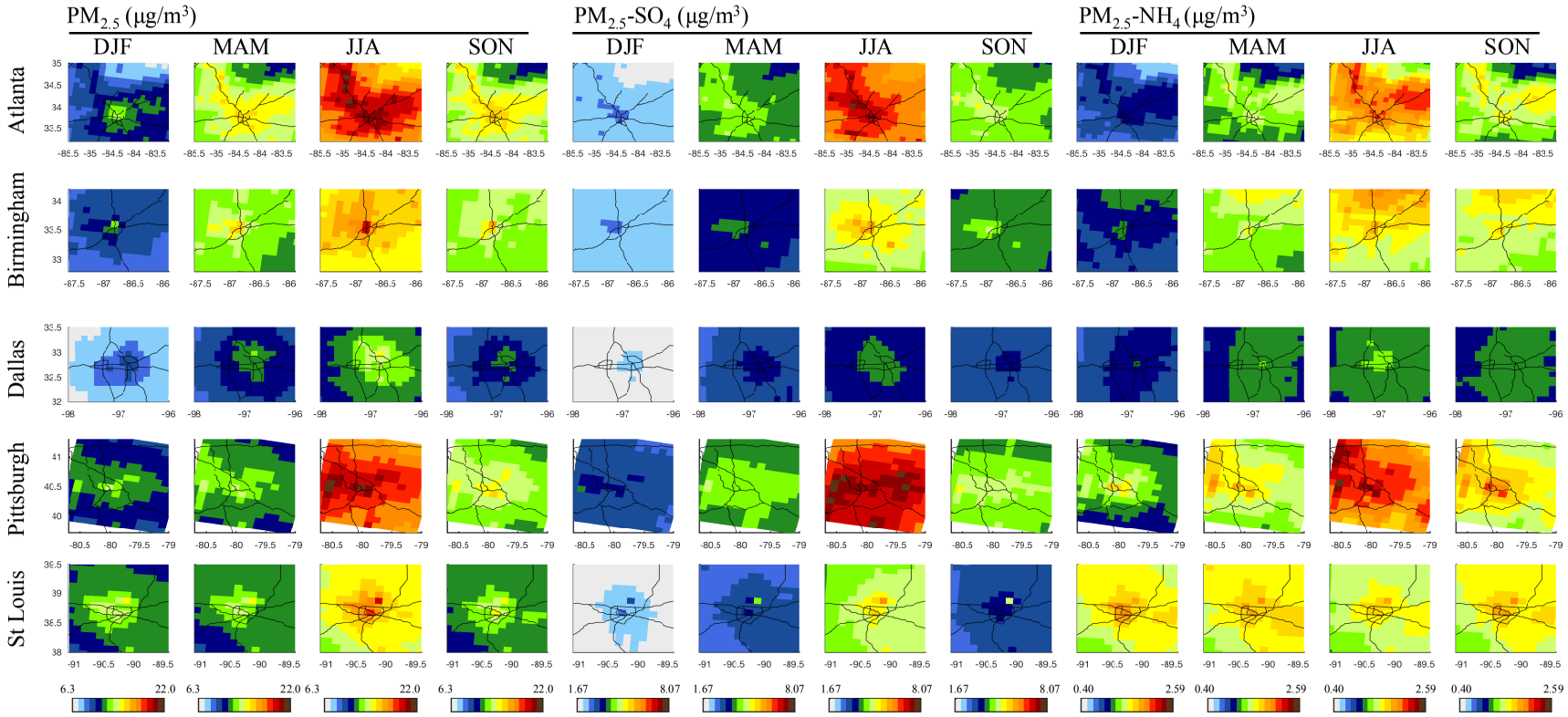

Figure B.6 – (*continued*)

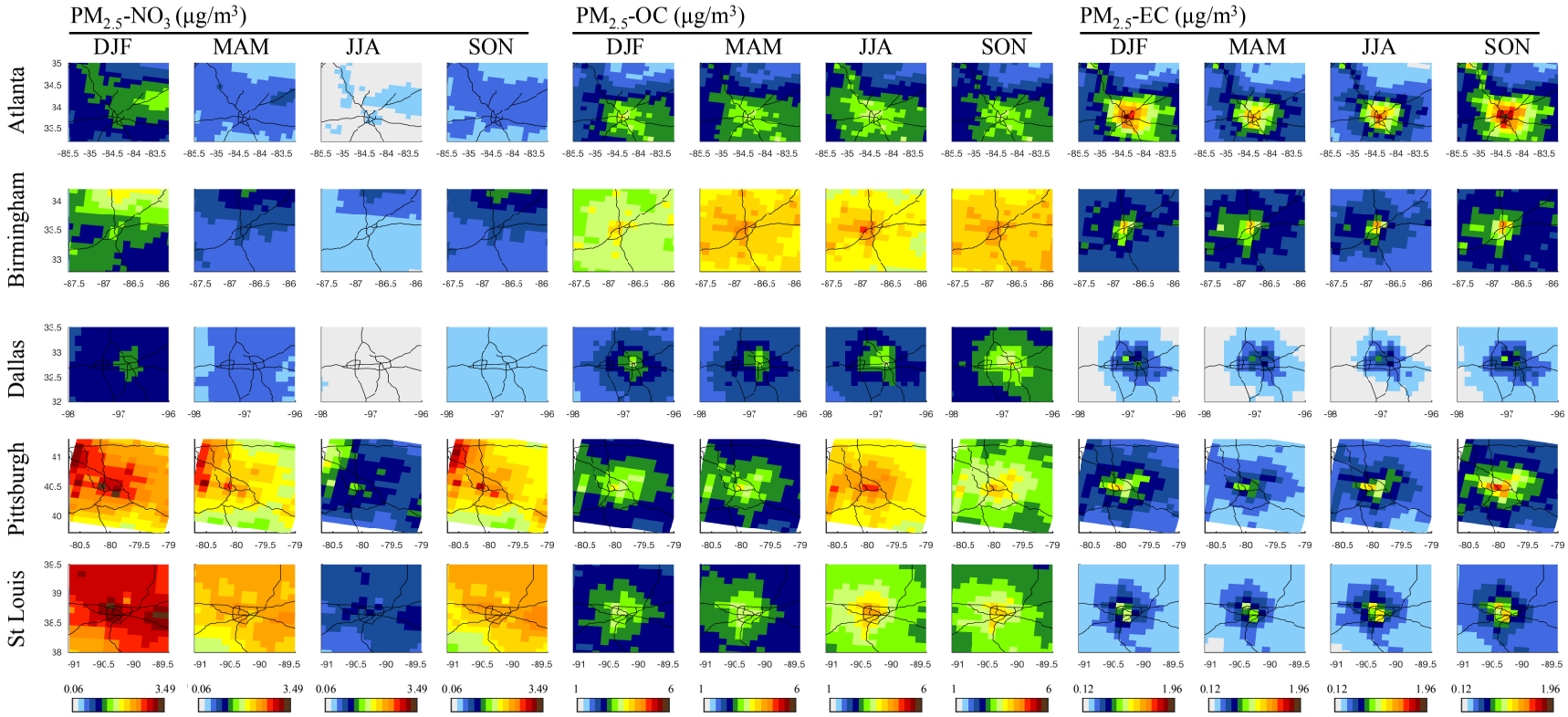

Figure B.6 – (*continued*)

APPENDIX C. SUPPLEMENTAL MATERIAL FOR CONSTRAINING CHEMICAL TRANSPORT $PM_{2.5}$ MODELING USING SURFACE STATION MEASUREMENTS AND SATELLITE RETRIEVALS: APPLICATION OVER THE SAN JOAQUIN VALLEY

Table C.1 – Aerosol groupings and definitions used in the $PM_{2.5}$ mass reconstruction for EPA CSN and CMAQ output species.

Aerosol Groups	Aerosols	CSN Monitor Representation	CMAQ Species Representation [1]
Inorganic Ions (II)	Sulfate	1.375 ´ SO4	1.375 ´ (ASO4I+ASO4J+ASO4K)
	Nitrate	1.29 x NO3	1.29 ´ (ANO3I+ANO3J+ANO3K)
	Ammonium	NH4	ANH4I+ANH4J+ANH4K
Light absorbing Carbon (LAC)	Elemental Carbon (EC)	EC	AECI+AECJ
Organic Matter (OM)	Organic Carbon (OC)	1.8 x OC	AALKJ+ABNZ1J+ABNZ2J+ABNZ3J+AISO1J+AISO2J+AISO3J+AOLGAJ+AOLGBJ+AORGCJ+ASQTJ+ATOL1J+ATOL2J+ATOL3J+ATRP1J+ATRP2J+AXYL1J+AXYL2J+AXYL3J+(1.4-0.2)*(APOCI+APOCJ)
Sea Salt (SS)	Sea Salt	1.8Cl + 1.4486Na+ 1.63Ca+ 1.2(K-0.6Fe)	ANAI+ANAJ+ACLI+ACLJ+ASEACAT+ACAJ+AKJ+AMGJ
Dust	Dust	2.49Si+1.94Ti +2.42Fe	AOTHRI+AOTHRJ+ACORS+ASOIL+AFEJ+ASIJ+ATIJ+AMNJ+APNCOMI+ APNCOMJ

[1] Model species in the Aitken, accumulation, and coarse size modes are indicated by the letters I, J, and K, respectively.

REFERENCES

1. Cohen, A. J.; Brauer, M.; Burnett, R.; Anderson, H. R.; Frostad, J.; Estep, K.; Balakrishnan, K.; Brunekreef, B.; Dandona, L.; Dandona, R., Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* **2017,** *389* (10082), 1907-1918.

2. Kloog, I.; Ridgway, B.; Koutrakis, P.; Coull, B. A.; Schwartz, J. D., Long-and short-term exposure to PM2. 5 and mortality: using novel exposure models. *Epidemiology (Cambridge, Mass.)* **2013,** *24* (4), 555.

3. Madrigano, J.; Kloog, I.; Goldberg, R.; Coull, B. A.; Mittleman, M. A.; Schwartz, J., Long-term exposure to PM2. 5 and incidence of acute myocardial infarction. *Environmental health perspectives* **2013,** *121* (2), 192.

4. Dominici, F.; Peng, R. D.; Bell, M. L.; Pham, L.; McDermott, A.; Zeger, S. L.; Samet, J. M., Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Jama* **2006,** *295* (10), 1127-1134.

5. Bell, M. L.; Zanobetti, A.; Dominici, F., Evidence on Vulnerability and Susceptibility to Health Risks Associated With Short-Term Exposure to Particulate Matter: A Systematic Review and Meta-Analysis. *American Journal of Epidemiology* **2013,** *178* (6), 865-876.

6. Krall, J. R.; Anderson, G. B.; Dominici, F.; Bell, M. L.; Peng, R. D., Short-term Exposure to Particulate Matter Constituents and Mortality in a National Study of U.S. Urban Communities. *Environmental Health Perspectives* **2013**.

7. Gold, D. R.; Litonjua, A.; Schwartz, J.; Lovett, E.; Larson, A.; Nearing, B.; Allen, G.; Verrier, M.; Cherry, R.; Verrier, R., Ambient pollution and heart rate variability. *Circulation* **2000,** *101* (11), 1267-1273.

8. Crouse, D. L.; Peters, P. A.; van Donkelaar, A.; Goldberg, M. S.; Villeneuve, P. J.; Brion, O.; Khan, S.; Atari, D. O.; Jerrett, M.; Pope III, C. A., Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. *Environmental health perspectives* **2012,** *120* (5), 708.

9. Zanobetti, A.; Schwartz, J., The Effect of Fine and Coarse Particulate Air Pollution on Mortality: A National Analysis. *Environmental Health Perspectives* **2009,** *117* (6), 898-903.

10. Schwartz, J.; Austin, E.; Bind, M.-A.; Zanobetti, A.; Koutrakis, P., Estimating causal associations of fine particles with daily deaths in Boston. *American journal of epidemiology* **2015**, kwv101.

11. Sarnat, J. A.; Koutrakis, P.; Suh, H. H., Assessing the relationship between personal particulate and gaseous exposures of senior citizens living in Baltimore, MD. *Journal of the Air & Waste Management Association* **2000,** *50* (7), 1184-1198.

12. Peters, A.; Dockery, D. W.; Muller, J. E.; Mittleman, M. A., Increased particulate air pollution and the triggering of myocardial infarction. *Circulation* **2001,** *103* (23), 2810-2815.

13. Brook, R. D.; Rajagopalan, S.; Pope, C. A.; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; Holguin, F.; Hong, Y.; Luepker, R. V.; Mittleman, M. A.; Peters, A.; Siscovick, D.; Smith, S. C.; Whitsel, L.; Kaufman, J. D.; on behalf of the American Heart Association Council on, E.; Prevention, C. o. t. K. i. C. D.; Council on Nutrition, P. A.; Metabolism, Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement From the American Heart Association. *Circulation* **2010,** *121* (21), 2331-2378.

14. Burnett, R. T.; Pope III, C. A.; Ezzati, M.; Olives, C.; Lim, S. S.; Mehta, S.; Shin, H. H.; Singh, G.; Hubbell, B.; Brauer, M., An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environmental health perspectives* **2014,** *122* (4), 397.

15. Laden, F.; Neas, L. M.; Dockery, D. W.; Schwartz, J., Association of fine particulate matter from different sources with daily mortality in six US cities. *Environmental health perspectives* **2000,** *108* (10), 941.

16. Evans, J.; van Donkelaar, A.; Martin, R. V.; Burnett, R.; Rainham, D. G.; Birkett, N. J.; Krewski, D., Estimates of global mortality attributable to particulate air pollution using satellite imagery. *Environmental research* **2013,** *120*, 33-42.

17. Janssen, B. G.; Godderis, L.; Pieters, N.; Poels, K.; Kiciński, M.; Cuypers, A.; Fierens, F.; Penders, J.; Plusquin, M.; Gyselaers, W., Placental DNA hypomethylation in association with particulate air pollution in early life. *Particle and fibre toxicology* **2013,** *10* (1), 22.

18. Hu, Z.; Rao, K. R., Particulate air pollution and chronic ischemic heart disease in the eastern United States: a county level ecological study using satellite aerosol data. *Environmental Health* **2009,** *8* (1), 26.

19. Adar, S. D.; Sheppard, L.; Vedal, S.; Polak, J. F.; Sampson, P. D.; Roux, A. V. D.; Budoff, M.; Jacobs Jr, D. R.; Barr, R. G.; Watson, K., Fine particulate air pollution and the progression of carotid intima-medial thickness: a prospective cohort study from

the multi-ethnic study of atherosclerosis and air pollution. *PLoS Med* **2013,** *10* (4), e1001430.

20. Kloog, I.; Melly, S. J.; Ridgway, W. L.; Coull, B. A.; Schwartz, J., Using new satellite based exposure methods to study the association between pregnancy PM 2.5 exposure, premature birth and birth weight in Massachusetts. *Environmental Health* **2012,** *11* (1), 40.

21. Watterson, T. L.; Sorensen, J.; Martin, R.; Coulombe Jr, R. A., Effects of PM2. 5 collected from Cache Valley Utah on genes associated with the inflammatory response in human lung cells. *Journal of Toxicology and Environmental Health, Part A* **2007,** *70* (20), 1731-1744.

22. Chen, Y.; Ebenstein, A.; Greenstone, M.; Li, H., Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences* **2013,** *110* (32), 12936-12941.

23. Zanobetti, A.; Dominici, F.; Wang, Y.; Schwartz, J. D., A national case-crossover analysis of the short-term effect of PM 2.5 on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environmental Health* **2014,** *13* (1), 38.

24. Franklin, P. J., Indoor air quality and respiratory health of children. *Paediatric respiratory reviews* **2007,** *8* (4), 281-286.

25. Fann, N.; Baker, K. R.; Fulcher, C. M., Characterizing the PM 2.5-related health benefits of emission reductions for 17 industrial, area and mobile emission sectors across the US. *Environment international* **2012,** *49*, 141-151.

26. Brook, R. D.; Cakmak, S.; Turner, M. C.; Brook, J. R.; Crouse, D. L.; Peters, P. A.; Van Donkelaar, A.; Villeneuve, P. J.; Brion, O.; Jerrett, M., Long-term fine particulate matter exposure and mortality from diabetes in Canada. *Diabetes care* **2013,** *36* (10), 3313-3320.

27. Elliott, C. T.; Copes, R., Burden of mortality due to ambient fine particulate air pollution (PM 2.5) in interior and Northern BC. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique* **2011**, 390-393.

28. Feng, C.; Li, J.; Sun, W.; Zhang, Y.; Wang, Q., Impact of ambient fine particulate matter (PM 2.5) exposure on the risk of influenza-like-illness: a time-series analysis in Beijing, China. *Environmental Health* **2016,** *15* (1), 17.

29. Bell, M. L.; Samet, J. M.; Dominici, F., Time-series studies of particulate matter. *Annu Rev Public Health* **2004,** *25*, 247-280.

30. Brunekreef, B.; Holgate, S. T., Air pollution and health. *Lancet* **2002,** *360* (9341), 1233-1242.

31. Pope, I. I. I. C. A.; Dockery, D. W., Health effects of fine particulate air pollution: lines that connect. *Journal of the Air & Waste Management Association* **2006,** *56* (6), 709-742.

32. Pope, C. A., 3rd; Ezzati, M.; Dockery, D. W., Fine-particulate air pollution and life expectancy in the United States. *N Engl J Med* **2009,** *360* (4), 376-86.

33. Laden, F.; Schwartz, J.; Speizer, F. E.; Dockery, D. W., Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *Am J Respir Crit Care Med* **2006,** *173* (6), 667-72.

34. Ozkaynak, H.; Glenn, B.; Qualters, J. R.; Strosnider, H.; McGeehin, M. A.; Zenick, H., Summary and findings of the EPA and CDC symposium on air pollution exposure and health. *J Expo Sci Environ Epidemiol* **2009,** *19* (1), 19-29.

35. Ozkaynak, H.; Isakov, V.; Baxter, L.; Graham, S. E.; Sarnat, S. E.; Sarnat, J. A.; Mulholland, J.; Turpin, B.; Rich, D. Q.; Lunden, M., Evaluating Alternative Exposure Metrics Used for Multipollutant Air Quality and Human Health Studies. *Air Pollution Modeling and Its Application Xxii* **2014**, 65-71.

36. Wade, K. S.; Mulholland, J. A.; Marmur, A.; Russell, A. G.; Hartsell, B.; Edgerton, E.; Klein, M.; Waller, L.; Peel, J. L.; Tolbert, P. E., Effects of instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, Georgia. *Journal of the Air & Waste Management Association* **2006,** *56* (6), 876-888.

37. Goldman, G. T.; Mulholland, J. A.; Russell, A. G.; Srivastava, A.; Strickland, M. J.; Klein, M.; Waller, L. A.; Tolbert, P. E.; Edgerton, E. S., Ambient air pollutant measurement error: characterization and impacts in a time-series epidemiologic study in Atlanta. *Environ Sci Technol* **2010,** *44* (19), 7692-8.

38. Marmur, A.; Unal, A.; Mulholland, J. A.; Russell, A. G., Optimization-Based Source Apportionment of PM2.5Incorporating Gas-to-Particle Ratios. *Environmental Science & Technology* **2005,** *39* (9), 3245-3254.

39. McMillan, N. J.; Holland, D. M.; Morara, M.; Feng, J., Combining numerical model output and particulate data using Bayesian space–time modeling. *Environmetrics* **2010,** *21* (1), 48-65.

40. Appel, K. W.; Foley, K. M.; Bash, J. O.; Pinder, R. W.; Dennis, R. L.; Allen, D. J.; Pickering, K., A multi-resolution assessment of the Community Multiscale Air Quality (CMAQ) model v4.7 wet deposition estimates for 2002-2006. *Geoscientific Model Development* **2011,** *4* (2), 357-371.

41. Appel, K. W.; Napelenok, S. L.; Foley, K. M.; Pye, H. O. T.; Hogrefe, C.; Luecken, D. J.; Bash, J. O.; Roselle, S. J.; Pleim, J. E.; Foroutan, H.; Hutzell, W. T.;

Pouliot, G. A.; Sarwar, G.; Fahey, K. M.; Gantt, B.; Gilliam, R. C.; Kang, D.; Mathur, R.; Schwede, D. B.; Spero, T. L.; Wong, D. C.; Young, J. O., Overview and evaluation of the Community Multiscale Air Quality (CMAQ) model version 5.1. *Geosci. Model Dev. Discuss.* **2016**, 1-41.

42. Tong, D. Q.; Mauzerall, D. L., Spatial variability of summertime tropospheric ozone over the continental United States: Implications of an evaluation of the CMAQ model. *Atmospheric Environment* **2006,** *40* (17), 3041-3056.

43. Kahn, R.; Banerjee, P.; McDonald, D., Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean. *Journal of Geophysical Research* **2001,** *106* (D16), 18219-18238.

44. Kahn, R. A.; Gaitley, B. J.; Garay, M. J.; Diner, D. J.; Eck, T. F.; Smirnov, A.; Holben, B. N., Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *Journal of Geophysical Research: Atmospheres (1984–2012)* **2010,** *115* (D23).

45. Liu, Y.; Koutrakis, P.; Kahn, R.; Turquety, S.; Yantosca, R. M., Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 2 - A case study. *Journal of the Air & Waste Management Association* **2007,** *57* (11), 1360-1369.

46. Liu, Y.; Koutrakis, P.; Kahn, R., Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 1—Method development. *Journal of the Air & Waste Management Association* **2007,** *57* (11), 1351-1359.

47. Patadia, F.; Kahn, R. A.; Limbacher, J. A.; Burton, S. P.; Ferrare, R. A.; Hostetler, C. A.; Hair, J. W., Aerosol airmass type mapping over the Urban Mexico City region from space-based multi-angle imaging. *Atmos. Chem. Phys.* **2013,** *13* (18), 9525-9541.

48. Van Donkelaar, A.; Martin, R. V.; Brauer, M.; Kahn, R.; Levy, R.; Verduzco, C.; Villeneuve, P. J., Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environmental health perspectives* **2010,** *118* (6), 847.

49. Omar, A. H.; Winker, D. M.; Kittaka, C.; Vaughan, M. A.; Liu, Z.; Hu, Y.; Trepte, C. R.; Rogers, R. R.; Ferrare, R. A.; Lee, K.-P.; Kuehn, R. E.; Hostetler, C. A., The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm. *J. Atmos. Oceanic Technol.* **2009,** *26* (10), 1994-2014.

50. Kahn, R. A.; Gaitley, B. J., An analysis of global aerosol type as retrieved by MISR. *J. Geophys. Res. Atmos.* **2015,** *120* (9), 4248-4281.

51. Li, J.; Carlson, B. E.; Lacis, A. A., Application of spectral analysis techniques in the intercomparison of aerosol data: Part III. Using combined PCA to compare spatiotemporal variability of MODIS, MISR, and OMI aerosol optical depth. *Journal of Geophysical Research: Atmospheres* **2014,** *119* (7), 4017-4042.

52. Bell, M. L.; Samet, J. M.; Dominici, F., Time-series studies of particulate matter. *Annu Rev Public Health* **2004,** *25*, 247-80.

53. Brook, R. D.; Rajagopalan, S.; Pope, C. A., 3rd; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; Holguin, F.; Hong, Y.; Luepker, R. V.; Mittleman, M. A.; Peters, A.; Siscovick, D.; Smith, S. C., Jr.; Whitsel, L.; Kaufman, J. D.; American Heart Association Council on, E.; Prevention, C. o. t. K. i. C. D.; Council on Nutrition, P. A.; Metabolism, Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **2010,** *121* (21), 2331-78.

54. Brunekreef, B.; Holgate, S. T., Air pollution and health. *Lancet* **2002,** *360* (9341), 1233-42.

55. Pope, C. A.; Dockery, D. W., Health Effects of Fine Particulate Air Pollution: Lines that Connect. *Journal of the Air & Waste Management Association* **2006,** *56* (6), 709-742.

56. Sarnat, S. E.; Sarnat, J. A.; Mulholland, J.; Isakov, V.; Ozkaynak, H.; Chang, H. H.; Klein, M.; Tolbert, P. E., Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. *J Expo Sci Environ Epidemiol* **2013,** *23* (6), 593-605.

57. Wade, K. S.; Mulholland, J. A.; Marmur, A.; Russell, A. G.; Hartsell, B.; Edgerton, E.; Klein, M.; Waller, L.; Peel, J. L.; Tolbert, P. E., Effects of instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, Georgia. *J Air Waste Manag Assoc* **2006,** *56* (6), 876-88.

58. Ivy, D.; Mulholland, J. A.; Russell, A. G., Development of ambient air quality population-weighted metrics for use in time-series health studies. *Journal of the Air & Waste Management Association* **2008,** *58* (5), 711-720.

59. McMillan, N. J.; Holland, D. M.; Morara, M.; Feng, J. Y., Combining numerical model output and particulate data using Bayesian space-time modeling. *Environmetrics* **2010,** *21* (1), 48-65.

60. Marshall, J. D.; Nethery, E.; Brauer, M., Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmospheric Environment* **2008,** *42* (6), 1359-1369.

61. Ozkaynak, H.; Baxter, L. K.; Dionisio, K. L.; Burke, J., Air pollution exposure prediction approaches used in air pollution epidemiology studies. *J Expo Sci Environ Epidemiol* **2013,** *23* (6), 566-72.

62. Isakov, V.; Crooks, J.; Touma, J.; Valari, M.; Ozkaynak, H.; Sarnat, S. E.; Sarnat, J.; Kewada, P.; Mulholland, J., Development and evaluation of alternative metrics of ambient air pollution exposure for use in epidemiologic studies. *Air Pollution Modeling and its Application XXI* **2012**, 681-686.

63. Crooks, J. L.; Ozkaynak, H., Simultaneous statistical bias correction of multiple PM2.5 species from a regional photochemical grid model. *Atmospheric Environment* **2014,** *95*, 126-141.

64. Fuentes, M.; Raftery, A. E., Model evaluation and spatial interpolation by Bayesian combination of observations with outputs from numerical models. *Biometrics* **2005,** *61* (1), 36-45.

65. Lindstrom, J.; Szpiro, A. A.; Sampson, P. D.; Oron, A. P.; Richards, M.; Larson, T. V.; Sheppard, L., A Flexible Spatio-Temporal Model for Air Pollution with Spatial and Spatio-Temporal Covariates. *Environ Ecol Stat* **2014,** *21* (3), 411-433.

66. Sampson, P. D.; Szpiro, A. A.; Sheppard, L.; Lindstrom, J.; Kaufman, J. D., Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmospheric Environment* **2011,** *45* (36), 6593-6606.

67. Wilton, D.; Szpiro, A.; Gould, T.; Larson, T., Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. *Science of the Total Environment* **2010,** *408* (5), 1120-1130.

68. Hoek, G.; Fischer, P.; Van Den Brandt, P.; Goldbohm, S.; Brunekreef, B., Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J Expo Anal Environ Epidemiol* **2001,** *11* (6), 459-69.

69. Solomon, P. A.; Chameides, W.; Weber, R.; Middlebrook, A.; Kiang, C. S.; Russell, A. G.; Butler, A.; Turpin, B.; Mikel, D.; Scheffe, R.; Cowling, E.; Edgerton, E.; St John, J.; Jansen, J.; McMurry, P.; Hering, S.; Bahadori, T., Overview of the 1999 Atlanta Supersite Project. *Journal of Geophysical Research-Atmospheres* **2003,** *108* (D7).

70. Hoek, G.; Beelen, R.; de Hoogh, K.; Vienneau, D.; Gulliver, J.; Fischer, P.; Briggs, D., A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* **2008,** *42* (33), 7561-7578.

71. Moore, D. K.; Jerrett, M.; Mack, W. J.; Kunzli, N., A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. *J Environ Monit* **2007,** *9* (3), 246-52.

72. Johnson, M.; Isakov, V.; Touma, J. S.; Mukerjee, S.; Ozkaynak, H., Evaluation of land-use regression models used to predict air quality concentrations in an urban area. *Atmospheric Environment* **2010,** *44* (30), 3660-3668.

73. Byun, D.; Schere, K. L., Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Reviews* **2006,** *59* (1-6), 51-77.

74. Basagana, X.; Rivera, M.; Aguilera, I.; Agis, D.; Bouso, L.; Elosua, R.; Foraster, M.; de Nazelle, A.; Nieuwenhuijsen, M.; Vila, J.; Kunzli, N., Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment* **2012,** *54*, 634-642.

75. Wang, M.; Beelen, R.; Eeftens, M.; Meliefste, K.; Hoek, G.; Brunekreef, B., Systematic evaluation of land use regression models for NO(2). *Environ Sci Technol* **2012,** *46* (8), 4481-9.

76. Chen, G.; Li, J.; Ying, Q.; Sherman, S.; Perkins, N.; Rajeshwari, S.; Mendola, P., Evaluation of observation-fused regional air quality model results for population air pollution exposure estimation. *Sci Total Environ* **2014,** *485-486* (C), 563-74.

77. Hao, H.; Chang, H. H.; Holmes, H. A.; Mulholland, J. A.; Klein, M.; Darrow, L. A.; Strickland, M. J., Air Pollution and Preterm Birth in the U.S. State of Georgia (2002-2006): Associations with Concentrations of 11 Ambient Air Pollutants Estimated by Combining Community Multiscale Air Quality Model (CMAQ) Simulations with Stationary Monitor Measurements. *Environmental Health Perspectives* **2016,** *124* (6), 875-880.

78. Hansen, D. A.; Edgerton, E. S.; Hartsell, B. E.; Jansen, J. J.; Kandasamy, N.; Hidy, G. M.; Blanchard, C. L., The Southeastern Aerosol Research and Characterization Study: part 1--Overview. *J Air Waste Manag Assoc* **2003,** *53* (12), 1460-71.

79. Solomon, P. A.; Crumpler, D.; Flanagan, J. B.; Jayanty, R. K. M.; Rickman, E. E.; McDade, C. E., U.S. National PM2.5Chemical Speciation Monitoring Networks—CSN and IMPROVE: Description of networks. *Journal of the Air & Waste Management Association* **2014,** *64* (12), 1410-1438.

80. Butler, A. J.; Andrew, M. S.; Russell, A. G., Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal of Geophysical Research-Atmospheres* **2003,** *108* (D1), SOS 3-1-SOS 3-11.

81. Chow, J. C.; Watson, J. G.; Chen, L. W.; Chang, M. C.; Robinson, N. F.; Trimble, D.; Kohl, S., The IMPROVE_A temperature protocol for thermal/optical carbon analysis: maintaining consistency with a long-term database. *J Air Waste Manag Assoc* **2007,** *57* (9), 1014-23.

82. Malm, W. C.; Schichtel, B. A.; Pitchford, M. L., Uncertainties in PM2.5 Gravimetric and Speciation Measurements and What We Can Learn from Them. *Journal of the Air & Waste Management Association* **2011,** *61* (11), 1131-1149.

83. Cressie, N., Spatial Prediction and Ordinary Kriging. *Mathematical Geology* **1988,** *20* (4), 405-421.

84. Foley, K. M.; Roselle, S. J.; Appel, K. W.; Bhave, P. V.; Pleim, J. E.; Otte, T. L.; Mathur, R.; Sarwar, G.; Young, J. O.; Gilliam, R. C.; Nolte, C. G.; Kelly, J. T.; Gilliland, A. B.; Bash, J. O., Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geoscientific Model Development* **2010,** *3* (1), 205-226.

85. Carlton, A. G.; Bhave, P. V.; Napelenok, S. L.; Edney, E. O.; Sarwar, G.; Pinder, R. W.; Pouliot, G. A.; Houyoux, M., Model representation of secondary organic aerosol in CMAQv4.7. *Environ Sci Technol* **2010,** *44* (22), 8553-60.

86. Sampson, P. D.; Richards, M.; Szpiro, A. A.; Bergen, S.; Sheppard, L.; Larson, T. V.; Kaufman, J. D., A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM2.5 concentrations in epidemiology. *Atmospheric Environment* **2013,** *75* (Supplement C), 383-392.

87. Carbajal-Arroyo, L.; Miranda-Soberanis, V.; Medina-Ramon, M.; Rojas-Bracho, L.; Tzintzun, G.; Solis-Gutierrez, P.; Mendez-Ramirez, I.; Hurtado-Diaz, M.; Schwartz, J.; Romieu, I., Effect of PM10 and O3 on infant mortality among residents in the Mexico City Metropolitan Area: a case-crossover analysis, 1997-2005. *Journal of Epidemiology & Community Health* **2011,** *65* (8), jech. 2009.101212.

88. Peng, R. D.; Samoli, E.; Pham, L.; Dominici, F.; Touloumi, G.; Ramsay, T.; Burnett, R. T.; Krewski, D.; Le Tertre, A.; Cohen, A., Acute effects of ambient ozone on mortality in Europe and North America: results from the APHENA study. *Air Quality, Atmosphere & Health* **2013,** *6* (2), 445-453.

89. Goldstein, I. F.; Landovitz, L., Analysis of air pollution patterns in New York City—I. Can one station represent the large metropolitan area? *Atmospheric Environment (1967)* **1977,** *11* (1), 47-52.

90. Ito, K.; Thurston, G., Daily PM10/mortality associations: an investigations of at-risk subpopulations. *J Expo Anal Environ Epidemiol* **1995,** *6* (1), 79-95.

91. Lipfert, F. W.; Wyzga, R. E., Air Pollution and Mortality: The Implications of Uncertainties in Regression Modeling and Exposure Measurement. *Journal of the Air & Waste Management Association* **1997,** *47* (4), 517-523.

92. Carrothers, T. J.; Evans, J. S., Assessing the Impact of Differential Measurement Error on Estimates of Fine Particle Mortality. *Journal of the Air & Waste Management Association* **2000,** *50* (1), 65-74.

93. Zeger, S. L.; Thomas, D.; Dominici, F.; Samet, J. M.; Schwartz, J.; Dockery, D.; Cohen, A., Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environmental Health* **2000,** *108* (5), 419-426.

94. Goldman, G. T.; Mulholland, J. A.; Russell, A. G.; Srivastava, A.; Strickland, M. J.; Klein, M.; Waller, L. A.; Tolbert, P. E.; Edgerton, E. S., Ambient Air Pollutant Measurement Error: Characterization and Impacts in a Time-Series Epidemiologic Study in Atlanta. *Environmental Science & Technology* **2010,** *44* (19), 7692-7698.

95. Goldman, G. T.; Mulholland, J. A.; Russell, A. G.; Gass, K.; Strickland, M. J.; Tolbert, P. E., Characterization of ambient air pollution measurement error in a time-series health study using a geostatistical simulation approach. *Atmospheric environment* **2012,** *57*, 101-108.

96. Pinto, J. P.; Lefohn, A. S.; Shadwick, D. S., Spatial variability of PM2. 5 in urban areas in the United States. *Journal of the Air & Waste Management Association* **2004,** *54* (4), 440-449.

97. Strickland, M. J.; Gass, K. M.; Goldman, G. T.; Mulholland, J. A., Effects of ambient air pollution measurement error on health effect estimates in time-series studies: a simulation-based analysis. *Journal of Exposure Science and Environmental Epidemiology* **2015,** *25* (2), 160-166.

98. O'Lenick, C. R.; Winquist, A.; Mulholland, J. A.; Friberg, M. D.; Chang, H. H.; Kramer, M. R.; Darrow, L. A.; Sarnat, S. E., Assessment of neighbourhood-level socioeconomic status as a modifier of air pollution–asthma associations among children in Atlanta. *... of Epidemiology and ...* **2016**, jech-2015-206530.

99. Pope, I. I. I. C. A.; Dockery, D. W., Healthe Effects of Fine Particulate Air Pollution: Lines that Connect. *Journal of the Air & Waste Management Association* **2006,** *56* (6), 709-742.

100. Sarnat, S. E.; Sarnat, J. A.; Mulholland, J.; Isakov, V.; Ozkaynak, H.; Chang, H. H.; Klein, M.; Tolbert, P. E., Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. *Journal of Exposure Science and Environmental Epidemiology* **2013,** *23* (6), 593-605.

101. Friberg, M. D.; Zhai, X.; Holmes, H. A.; Chang, H. H.; Strickland, M. J.; Sarnat, S. E.; Tolbert, P. E.; Russell, A. G.; Mulholland, J. A., Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution. *Environmental science & technology* **2016,** *50* (7), 3695-3705.

102. Online., N. C. D. National Climatic Data Center, NESDIS, NOAA, U.S. Department of Commerce. . http://www7.ncdc.noaa.gov/CDO/

dataproduct

103. Solomon, P. A.; Crumpler, D.; Flanagan, J. B.; Jayanty, R. K. M.; Rickman, E. E.; McDade, C. E., U.S. National PM 2.5Chemical Speciation Monitoring Networks—CSN and IMPROVE: Description of networks. *Journal of the Air & Waste Management Association* **2014,** *64* (12), 1410-1438.

104. Hansen, D. A.; Edgerton, E. S.; Hartsell, B. E.; Jansen, J. J.; Kandasamy, N.; Hidy, G. M.; Blanchard, C. L., The Southeastern Aerosol Research and Characterization Study: Part 1—Overview. *Journal of the Air & Waste Management Association* **2003,** *53* (12), 1460-1471.

105. Butler, A. J.; Andrew, M. S.; Russell, A. G., Daily sampling of PM2. 5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal of Geophysical Research: Atmospheres (1984–2012)* **2003,** *108* (D7), SOS 3-1-SOS 3-11.

106. Hauck, H.; Berner, A.; Gomiscek, B.; Stopper, S.; Puxbaum, H.; Kundi, M.; Preining, O., On the equivalence of gravimetric PM data with TEOM and beta-attenuation measurements. *Journal of Aerosol Science* **2004,** *35* (9), 1135-1149.

107. Chow, J. C.; Watson, J. G.; Chen, L. W. A.; Chang, M. C. O.; Robinson, N. F.; Trimble, D.; Kohl, S., The IMPROVE_A Temperature Protocol for Thermal/Optical Carbon Analysis: Maintaining Consistency with a Long-Term Database. *Journal of the Air & Waste Management Association* **2007,** *57* (9), 1014-1023.

108. Malm, W. C.; Schichtel, B. A.; Pitchford, M. L., Uncertainties in PM2. 5 gravimetric and speciation measurements and what we can learn from them. *Journal of the Air & Waste Management Association* **2011,** *61* (11), 1131-1149.

109. Metzger, K. B.; Tolbert, P. E.; Klein, M.; Peel, J. L.; Flanders, W. D.; Todd, K.; Mulholland, J. A.; Ryan, P. B.; Frumkin, H., Ambient air pollution and cardiovascular emergency department visits. *Epidemiology* **2004,** *15* (1), 46-56.

110. Peel, J. L.; Tolbert, P. E.; Klein, M.; Metzger, K. B.; Flanders, W. D.; Todd, K.; Mulholland, J. A.; Ryan, P. B.; Frumkin, H., Ambient Air Pollution and Respiratory Emergency Department Visits. *Epidemiology* **2005,** *16* (2), 164-174.

111. Tolbert, P. E.; Mulholland, J. A.; Macintosh, D. L.; Xu, F.; Daniels, D.; Devine, O. J.; Carlin, B. P.; Klein, M.; Butler, A. J.; Nordenberg, D. F., Air quality and pediatric emergency room visits for asthma and Atlanta, Georgia. *American Journal of Epidemiology* **2000,** *151* (8), 798-810.

112. McGeehin, M. A.; Qualters, J. R.; Niskar, A. S., National environmental public health tracking program: bridging the information gap. *Environmental Health Perspectives* **2004,** *112* (14), 1409.

113. Talbot, T. O.; Haley, V. B.; Dimmick, W. F.; Paulu, C.; Talbott, E. O.; Rager, J., Developing consistent data and methods to measure the public health impacts of ambient air quality for Environmental Public Health Tracking: progress to date and future directions. *Air Quality, Atmosphere & Health* **2009,** *2* (4), 199-206.

114. Appel, K. W.; Gilliland, A. B.; Sarwar, G.; Gilliam, R. C., Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance Part I - Ozone. *Atmospheric Environment* **2007,** *41* (40), 9603-9615.

115. Appel, K. W.; Bhave, P. V.; Gilliland, A. B.; Sarwar, G.; Roselle, S. J., Evaluation of the community multiscale air quality (CMAQ) model version 4.5: Sensitivities impacting model performance; Part II - particulate matter. *Atmospheric Environment* **2008,** *42* (24), 6057-6066.

116. Chang, J. C.; Hanna, S. R., Air quality model performance evaluation. *Meteorol Atmos Phys* **2004,** *87* (1), 167-196.

117. Taylor, K. E., Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres (1984–2012)* **2001,** *106* (D7), 7183-7192.

118. Carlton, A. G.; Bhave, P. V.; Napelenok, S. L.; Edney, E. O.; Sarwar, G.; Pinder, R. W.; Pouliot, G. A.; Houyoux, M., Model representation of secondary organic aerosol in CMAQv4.7. *Environmental Science & Technology* **2010,** *44* (22), 8553-8560.

119. Foley, K. M.; Roselle, S. J.; Appel, K. W.; Bhave, P. V.; Pleim, J. E.; Otte, T. L.; Mathur, R.; Sarwar, G.; Young, J. O.; Gilliam, R. C., Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geosci. Model Dev.* **2010,** *3* (1), 205-226.

120. Qin, Y.; Walk, T.; Gary, R.; Yao, X.; Elles, S., C 2–C 10 nonmethane hydrocarbons measured in Dallas, USA—seasonal trends and diurnal characteristics. *Atmospheric Environment* **2007,** *41* (28), 6018-6032.

121. Kemball-Cook, S.; Parrish, D.; Ryerson, T.; Nopmongcol, U.; Johnson, J.; Tai, E.; Yarwood, G., Contributions of regional transport and local sources to ozone exceedances in Houston and Dallas: Comparison of results from a photochemical grid model to aircraft and surface measurements. *Journal of Geophysical Research: Atmospheres* **2009,** *114* (D7).

122. Adachi, K.; Tainosho, Y., Characterization of heavy metal particles embedded in tire dust. *Environment international* **2004,** *30* (8), 1009-1017.

123. Garg, B. D.; Cadle, S. H.; Mulawa, P. A.; Groblicki, P. J.; Laroo, C.; Parr, G. A., Brake wear particulate matter emissions. *Environmental Science & Technology* **2000,** *34* (21), 4463-4469.

124. Sternbeck, J.; Sjödin, Å.; Andréasson, K., Metal emissions from road traffic and the influence of resuspension—results from two tunnel studies. *Atmospheric Environment* **2002,** *36* (30), 4735-4744.

125. Lee, J. H.; Hopke, P. K., Apportioning sources of PM 2.5 in St. Louis, MO using speciation trends network data. *Atmospheric Environment* **2006,** *40*, 360-377.

126. Duncan, B. N.; Stelson, A. W.; Kiang, C. S., Estimated contribution of power plants to ambient nitrogen oxides measured in Atlanta, Georgia in August 1992. *Atmospheric Environment* **1995,** *29* (21), 3043-3054.

127. Kirby, C.; Greig, A.; Drye, T., Temporal and spatial variations in nitrogen dioxide concentrations across an urban landscape: Cambridge, UK. *Environmental Monitoring and Assessment* **1998,** *52* (1-2), 65-82.

128. Pekney, N. J.; Davidson, C. I.; Robinson, A.; Zhou, L.; Hopke, P.; Eatough, D.; Rogge, W. F., Major source categories for PM2. 5 in Pittsburgh using PMF and UNMIX. *Aerosol Science and Technology* **2006,** *40* (10), 910-924.

129. Pekney, N. J.; Davidson, C. I.; Zhou, L.; Hopke, P. K., Application of PSCF and CPF to PMF-modeled sources of PM2. 5 in Pittsburgh. *Aerosol Science and Technology* **2006,** *40* (10), 952-961.

130. Blanchard, C. L.; Tanenbaum, S.; Hidy, G. M., Spatial and temporal variability of air pollution in Birmingham, Alabama. *Atmospheric Environment* **2014,** *89*, 382-391.

131. Tang, W.; Raymond, T.; Wittig, B.; Davidson, C.; Pandis, S.; Robinson, A.; Crist, K., Spatial variations of PM2. 5 during the Pittsburgh air quality study. *Aerosol Science and Technology* **2004,** *38* (S2), 80-90.

132. Takahama, S.; Wittig, A. E.; Vayenas, D. V.; Davidson, C. I.; Pandis, S. N., Modeling the diurnal variation of nitrate during the Pittsburgh Air Quality Study. *Journal of Geophysical Research: Atmospheres* **2004,** *109* (D16).

133. Turner, J. R.; Allen, D. T., Transport of Atmospheric Fine Particulate Matter: Part 2—Findings from Recent Field Programs on the Intraurban Variability in Fine Particulate Matter. *Journal of the Air & Waste Management Association* **2008,** *58* (2), 196-215.

134. Laden, F.; Schwartz, J.; Speizer, F. E.; Dockery, D. W., Reduction in Fine Particulate Air Pollution and Mortality. *American Journal of Respiratory and Critical Care Medicine* **2006,** *173* (6), 667-672.

135. Pope, C. A.; Ezzati, M.; Dockery, D. W., Fine-Particulate Air Pollution and Life Expectancy in the United States. *N Engl J Med* **2009,** *360* (4), 376-386.

136. Özkaynak, H.; Frey, H. C.; Burke, J.; Pinder, R. W., Analysis of coupled model uncertainties in source-to-dose modeling of human exposures to ambient air pollution: A PM 2.5 case study. *Atmospheric environment* **2009,** *43* (9), 1641-1649.

137. Hand, J. L.; Copeland, S. A.; Day, D. E.; Dillner, A. M.; Indresand, H.; Malm, W. C.; McDade, C. E.; Moore Jr, C. T.; Pitchford, M. L.; Schichtel, B. A., IMPROVE (Interagency Monitoring of Protected Visual Environments): Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States. *Report V, CIRA Report ISSN* **2011**, 0737-5352.

138. Chow, J. C.; Watson, J. G.; Chen, L. W. A., Quantification of PM 2.5 organic carbon sampling artifacts in US networks. *Atmospheric ...* **2010**.

139. Engel-Cox, J. A.; Holloman, C. H.; Coutant, B. W.; Hoff, R. M., Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmospheric Environment* **2004,** *38* (16), 2495-2509.

140. Chu, D. A. In *Analysis of the relationship between MODIS aerosol optical depth and PM2. 5 in the summertime US*, 2006; p 629903.

141. Al-Saadi, J.; Szykman, J.; Pierce, R. B.; Kittaka, C.; Neil, D.; Chu, D. A.; Remer, L.; Gumley, L.; Prins, E.; Weinstock, L., Improving national air quality forecasts with satellite aerosol observations. *Bull. Amer. Meteor. Soc.* **2005,** *86* (9), 1249-1261.

142. Liu, Y.; Park, R. J.; Jacob, D. J.; Li, Q.; Kilaru, V.; Sarnat, J. A., Mapping annual mean ground-level PM 2.5concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *J. Geophys. Res. Atmos.* **2004,** *109* (D22), n/a-n/a.

143. Dey, S.; Di Girolamo, L.; van Donkelaar, A.; Tripathi, S. N.; Gupta, T.; Mohan, M., Variability of outdoor fine particulate (PM 2.5) concentration in the Indian subcontinent: a remote sensing approach. *Remote Sensing of Environment* **2012,** *127*, 153-161.

144. Ngo, M.; Pinkerton, K.; Freeland, S.; Geller, M.; Ham, W.; Cliff, S.; Hopkins, L.; Kleeman, M.; Kodavanti, U.; Meharg, E., Airborne particles in the San Joaquin Valley may affect human health. *California Agriculture* **2010,** *64* (1), 12-16.

145. Chow, J. C.; Chen, L. W. A.; Watson, J. G.; Lowenthal, D. H.; Magliano, K. A.; Turkiewicz, K.; Lehrman, D. E., PM2. 5 chemical composition and spatiotemporal variability during the California Regional PM10/PM2. 5 Air Quality Study (CRPAQS). *Journal of Geophysical Research: Atmospheres* **2006,** *111* (D10).

146. Appel, K. W.; Napelenok, S. L.; Foley, K. M.; Pye, H. O. T.; Hogrefe, C.; Luecken, D. J.; Bash, J. O.; Roselle, S. J.; Pleim, J. E.; Foroutan, H., Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1. *Geosci. Model Dev.* **2017,** *10* (4), 1703.

147. Galloway, D.; Riley, F. S., San Joaquin Valley, California. *Land subsidence in the United States: US Geological Survey Circular* **1999,** *1182*, 23-34.

148. EPA, U. S. Green Book PM-2.5 (2012) Area Information https://www3.epa.gov/airquality/greenbook/kbtc.html.

149. EPA, U. S., Air Quality Criteria for Particulate Matter (Final Report). **2004**.

150. Holben, B. N.; Eck, T. F.; Slutsker, I.; Tanre, D.; Buis, J. P.; Setzer, A.; Vermote, E.; Reagan, J. A.; Kaufman, Y. J.; Nakajima, T., AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment* **1998,** *66* (1), 1-16.

151. Eck, T. F.; Holben, B. N.; Reid, J. S.; Dubovik, O.; Smirnov, A.; O'Neill, N. T.; Slutsker, I.; Kinne, S., Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. *Journal of Geophysical Research: Atmospheres* **1999,** *104* (D24), 31333-31349.

152. Smirnov, A.; Holben, B. N.; Eck, T. F.; Dubovik, O.; Slutsker, I., Cloud-screening and quality control algorithms for the AERONET database. *Remote Sensing of Environment* **2000,** *73* (3), 337-349.

153. Skamarock, W. C.; Klemp, J. B.; Dudhia, J.; Gill, D. O.; Barker, D. M.; Duda, M. G.; Huang, X. Y.; Wang, W.; Powers, J. G., A description of the Advanced Research WRF Version 3, NCAR technical note, Mesoscale and Microscale Meteorology Division. *National Center for Atmospheric Research, Boulder, Colorado, USA* **2008**.

154. Byun, D.; Schere, K. L., Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Reviews* **2006,** *59* (2), 51-77.

155. EPA, U. S. 2011 National Emissions Inventory, Version 2 Technical Support Document. https://www.epa.gov/sites/production/files/2015-2010/documents/nei2011v2012_tsd_2014aug2015.pdf.

156. Houyoux, M. R.; Vukovich, J. M.; Coats, C. J.; Wheeler, N. J.; Kasibhatla, P. S., Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. *Journal of Geophysical Research: Atmospheres* **2000,** *105* (D7), 9079-9090.

157. Sarwar, G.; Simon, H.; Bhave, P.; Yarwood, G., Examining the impact of heterogeneous nitryl chloride production on air quality across the United States. *Atmos. Chem. Phys.* **2012,** *12* (14), 6455-6473.

158. Whitten, G. Z.; Heo, G.; Kimura, Y.; McDonald-Buller, E.; Allen, D. T.; Carter, W. P.; Yarwood, G., A new condensed toluene mechanism for Carbon Bond: CB05-TU. *Atmospheric Environment* **2010,** *44* (40), 5346-5355.

159. Yarwood, G.; Rao, S.; Yocke, M.; Whitten, G., Updates to the carbon bond chemical mechanism: CB05. *Final report to the US EPA, RT-0400675* **2005,** *8*.

160. Bey, I.; Jacob, D. J.; Yantosca, R. M.; Logan, J. A.; Field, B. D.; Fiore, A. M.; Li, Q.; Liu, H. Y.; Mickley, L. J.; Schultz, M. G., Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research: Atmospheres* **2001,** *106* (D19), 23073-23095.

161. Diner, D. J.; Beckert, J. C.; Reilly, T. H.; Bruegge, C. J.; Conel, J. E.; Kahn, R. A.; Martonchik, J. V.; Ackerman, T. P.; Davies, R.; Gerstl, S. A. W., Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. *IEEE Transactions On Geoscience and Remote Sensing* **1998,** *36* (4), 1072-1087.

162. Limbacher, J. A.; Kahn, R. A., Updated MISR dark water research aerosol retrieval algorithm–Part 1: Coupled 1.1 km ocean surface chlorophyll a retrievals with empirical calibration corrections. *Atmos. Meas. Tech.* **2017,** *10* (4), 1539.

163. Limbacher, J. A.; Kahn, R. A., MISR calibration issues in high-contrast scenes, and empirical corrections. *Atmospheric Measurement ...* **2015**.

164. Bruegge, C. J.; Diner, D. J.; Kahn, R. A.; Chrien, N.; Helmlinger, M. C.; Gaitley, B. J.; Abdou, W. A., The MISR radiometric calibration process. *Remote Sensing of Environment* **2007,** *107* (1), 2-11.

165. Diner, D. J.; Martonchik, J. V.; Kahn, R. A.; Pinty, B.; Gobron, N.; Nelson, D. L.; Holben, B. N., Using angular and spectral shape similarity constraints to improve MISR aerosol and surface retrievals over land. *Remote Sensing of Environment* **2005,** *94* (2), 155-171.

166. North, P. R. J.; Briggs, S. A.; Plummer, S. E.; Settle, J. J., Retrieval of land surface bidirectional reflectance and aerosol opacity from ATSR-2 multiangle imagery. *IEEE Transactions On Geoscience and Remote Sensing* **1999,** *37* (1), 526-537.

167. Limbacher, J. A.; Kahn, R. A., MISR research-aerosol-algorithm refinements for dark water retrievals. *Atmos. Meas. Tech.* **2014,** *7* (11), 3989-4007.

168. Witek, M. L.; Garay, M. J.; Diner, D. J.; Smirnov, A., Aerosol optical depths over oceans: A view from MISR retrievals and collocated MAN and AERONET in situ observations. *Journal of Geophysical Research: Atmospheres* **2013,** *118* (22).

169. Shi, Y.; Zhang, J.; Reid, J.; Liu, B.; Hyer, E., Critical evaluation of cloud contamination in the MISR aerosol products using MODIS cloud mask products. *Atmos. Meas. Tech.* **2014,** *7* (6), 1791-1801.

170. Lyapustin, A. I.; Wang, Y.; Laszlo, I.; Hilker, T.; Hall, F. G.; Sellers, P. J.; Tucker, C. J.; Korkin, S. V., Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction. *Remote Sensing of Environment* **2012,** *127*, 385-393.

171. Lyapustin, A.; Wang, Y.; Laszlo, I.; Kahn, R.; Korkin, S.; Remer, L.; Levy, R.; Reid, J. S., Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. *Journal of Geophysical Research* **2011,** *116* (D3), D03211.

172. Lyapustin, A.; Wang, Y.; Xiong, X.; Meister, G.; Platnick, S.; Levy, R.; Franz, B.; Korkin, S.; Hilker, T.; Tucker, J., Scientific impact of MODIS C5 calibration degradation and C6+ improvements. **2014**.

173. Kloog, I.; Chudnovsky, A. A.; Just, A. C.; Nordio, F.; Koutrakis, P.; Coull, B. A.; Lyapustin, A.; Wang, Y.; Schwartz, J., A new hybrid spatio-temporal model for estimating daily multi-year PM 2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. *Atmospheric Environment* **2014,** *95*, 581-590.

174. Countess, R. J.; Wolff, G. T.; Cadle, S. H., The Denver winter aerosol: a comprehensive chemical characterization. *Journal of the Air Pollution Control Association* **1980,** *30* (11), 1194-1200.

175. Frank, N. H., Retained Nitrate, Hydrated Sulfates, and Carbonaceous Mass in Federal Reference Method Fine Particulate Matter for Six Eastern U.S. Cities. *Journal of the Air & Waste Management Association* **2006,** *56* (4), 500-511.

176. Hand, J. L.; Schichtel, B. A.; Malm, W. C.; Copeland, S.; Molenar, J. V.; Frank, N.; Pitchford, M., Widespread reductions in haze across the United States from the early 1990s through 2011. *Atmospheric environment* **2014,** *94*, 671-679.

177. Chow, J. C.; Watson, J. G.; Fujita, E. M.; Lu, Z.; Lawson, D. R.; Ashbaugh, L. L., Temporal and spatial variations of PM2. 5 and PM10 aerosol in the Southern California air quality study. *Atmospheric Environment* **1994,** *28* (12), 2061-2080.

178. Chow, J. C.; Egami, R. T., San Joaquin Valley 1995 integrated monitoring study: Documentation, evaluation, and descriptive data analysis of PM10, PM2. 5, and precursor gas measurements. *Prepared for the California Regional Particulate Air Quality Study,*

California Air Resources Board, Sacramento, CA, by the Desert Research Institute, Reno, NV **1997**.

179. Andrews, E.; Saxena, P.; Musarra, S.; Hildemann, L. M.; Koutrakis, P.; McMurry, P. H.; Olmez, I.; White, W. H., Concentration and composition of atmospheric aerosols from the 1995 SEAVS experiment and a review of the closure between chemical and gravimetric measurements. *J Air Waste Manag Assoc* **2000,** *50* (5), 648-664.

180. Nolte, C. G.; Appel, K. W.; Kelly, J. T.; Bhave, P. V.; Fahey, K. M.; Collett, J. J. L.; Zhang, L.; Young, J. O., Evaluation of the Community Multiscale Air Quality (CMAQ) model v5.0 against size-resolved measurements of inorganic particle composition across sites in North America. *Geosci. Model Dev.* **2015,** *8* (9), 2877-2892.

181. DeBell, L. J.; Gebhart, K. A.; Hand, J. L.; Malm, W. C.; Pitchford, M. L.; Schichtel, B. A.; White, W. H., Spatial and seasonal patterns and temporal variability of haze and its constituents in the United States: Report IV. *Cooperative Institute for Research in the Atmosphere* **2006**, 217-218.

182. Bond, T. C.; Bergstrom, R. W., Light absorption by carbonaceous particles: An investigative review. *Aerosol science and technology* **2006,** *40* (1), 27-67.

183. Malm, W. C.; Sisler, J. F.; Huffman, D., Spatial and seasonal trends in particle concentration and optical extinction in the United States. *Journal of ...* **1994**.

184. Simon, H.; Bhave, P. V.; Swall, J. L.; Frank, N. H.; Malm, W. C., Determining the spatial and seasonal variability in OM/OC ratios across the US using multiple regression. *Atmos. Chem. Phys.* **2011,** *11* (6), 2933-2949.

185. Rees, S. L.; Robinson, A. L.; Khlystov, A.; Stanier, C. O.; Pandis, S. N., Mass balance closure and the Federal Reference Method for PM 2.5 in Pittsburgh, Pennsylvania. *Atmospheric Environment* **2004,** *38* (20), 3305-3318.

186. Tanner, R. L.; Parkhurst, W. J.; Valente, M. L.; Phillips, W. D., Regional composition of PM 2.5 aerosols measured at urban, rural and “background” sites in the Tennessee valley. *Atmospheric Environment* **2004,** *38* (20), 3143-3153.

187. El-Zanan, H. S.; Lowenthal, D. H.; Zielinska, B.; Chow, J. C.; Kumar, N., Determination of the organic aerosol mass to organic carbon ratio in IMPROVE samples. *Chemosphere* **2005,** *60* (4), 485-496.

188. Chow, J. C.; Lowenthal, D. H.; Chen, L. W. A.; Wang, X.; Watson, J. G., Mass reconstruction methods for PM2.5: a review. *Air Quality Atmosphere and Health* **2015,** *8* (3), 243-263.

189. Malm, W. C.; Day, D. E.; Kreidenweis, S. M., Light scattering characteristics of aerosols as a function of relative humidity: Part I—A comparison of measured scattering

and aerosol concentrations using the theoretical models. *Journal of the Air & Waste Management Association* **2000,** *50* (5), 686-700.

190. Song, C. H.; Park, M. E.; Lee, K. H.; Ahn, H. J.; Lee, Y.; Kim, J. Y.; Han, K. M.; Kim, J.; Ghim, Y. S.; Kim, Y. J., An investigation into seasonal and regional aerosol characteristics in East Asia using model-predicted and remotely-sensed aerosol properties. *Atmos. Chem. Phys.* **2008,** *8* (22), 6627-6654.

191. Park, R. S.; Song, C. H.; Han, K. M.; Park, M. E.; Lee, S. S.; Kim, S. B.; Shimizu, A., A study on the aerosol optical properties over East Asia using a combination of CMAQ-simulated aerosol optical properties and remote-sensing data via a data assimilation technique. *Atmos. Chem. Phys.* **2011,** *11* (23), 12275-12296.

192. Kahn, R. A.; Chen, Y.; Nelson, D. L.; Leung, F. Y.; Li, Q.; Diner, D. J.; Logan, J. A., Wildfire smoke injection heights: Two perspectives from space. *Geophys. Res. Lett.* **2008,** *35* (4).

193. Chen, L.-W. A.; Watson, J. G.; Chow, J. C.; Magliano, K. L., Quantifying PM2. 5 source contributions for the San Joaquin Valley with multivariate receptor models. *Environmental science & technology* **2007,** *41* (8), 2818-2826.

194. Chow, J. C.; Watson, J. G.; Lowenthal, D. H.; Hackney, R.; Magliano, K.; Lehrman, D.; Smith, T., Temporal Variations of PM 2.5, PM 10, and Gaseous Precursors during the 1995 Integrated Monitoring Study in Central California. *Journal of the Air & Waste Management Association* **1999,** *49* (9), 16-24.