NASA Goddard Space Flight Center
Code 613, Greenbelt, MD 20771
mariel.d.friberg@nasa.gov

Dr. Qiang Zhang
Co-Editor
Atmospheric Chemistry and Physics

July 12, 2018

Dear Dr. Qiang Zhang:

We are pleased to submit our revised manuscript acp-2018-152, titled "Constraining Chemical Transport $PM_{2.5}$ Modeling Output Using Surface Monitor Measurements and Satellite Retrievals: Application over the San Joaquin Valley", for peer-review completion and potential final publication in *Atmospheric Chemistry and Physics* the revised version. We appreciate the referee comments and have updated the manuscript in response these comments.

Thank you for your consideration and assistance.

Sincerely,

Mariel D. Friberg, PhD
Georgia Institute of Technology / NASA Goddard Space Flight Center

Enclosed Files: Cover Letter+Response to Referees+Manuscript with tracked changes, Manuscript, and Supporting Information.

Comment on "Constraining Chemical Transport $PM_{2.5}$ Modeling Using Surface Monitor Measurements and Satellite Retrievals: Application over the San Joaquin Valley" *by* Mariel D. Friberg et al.

Author Comments (AC) to Referee Comments (RC) #1 – Anonymous Referee #2

Received and published: 1 June 2018

In this paper, the authors conducted a case study for six days over San Joaquin Valley to constrain model simulated PM2.5 using surface monitor measurements and satellite retrievals. They combined the aerosol products at 275 m spatial resolution from the MISR Research Aerosol retrieval algorithm, ground observations from EPA and the 2 km resolution simulations from WRF/CMAQ to improve the surface estimates of PM2.5, its major chemical component species estimates, and related estimates of uncertainty. The optimized results show good agreements with ground observations for both the total PM2.5 and the species. The method is sound and the results look reliable. I recommend considering this paper for publication upon response to the following comments:

AC1_0. We thank the reviewer for the valuable comments to improve our manuscript. Please see our itemized responses below.

Major comments:

RC1_1. This work is a case study and the authors selected several days with requirements for the MISR data: (1) relatively cloud-free conditions for more MISR coverage; (2) mid-visible AOD exceeds 0.15. They have mentioned in the manuscript that applying this method in other polluted regions are likely to meet common condition with AOD exceeding 0.15. However, what about the coverage issue? For days with limited MISR coverage, the MAIAC AOD used to fill the gap will also have a lot of missing. Then how will this method be applied? This should be discussed in the manuscript.

AC1_1. Where satellite data are missing or where the AOD is too low to provide reliable aerosol type from MISR, we must rely on the emissions-based CMAQ model, tuned, to the extent possible by satellite and surface measurement. Nevertheless, the satellite provides vastly more spatial coverage than the surface stations alone, and this is especially important downwind of major pollution sources. As such, our approach provides improvements where possible, but does not resolve all possible problems. This is now emphasized in the Conclusions section of the paper. The plotting coverage in Figures 6, 7, and S4 has been addressed. Following Figure 5, when FillSAT is not available, the optimized dataset reflects the fused (model + surface measurements) results.

RC1_2. What are the major advantages of this study compared to previous studies that combined information from the satellite retrieval, CTMs and ground observations together? The optimized results in this study seemed not to take advantage of the full coverage of the CTMs.

AC1_2. The physical approach introduced in this study complements the statistical approaches now widely used to take advantage of satellite coverage for air quality applications. Statistical approaches rely on surface-based data training sets to constrain parameters in statistical models, which are then applied elsewhere. Where training data are limited or entirely absent, there is great uncertainty with this approach. In other studies where satellite data are used to constrain a CTM, only the AOD or very limited aerosol-type constraints from the satellite is considered. The physical approach we present makes use of surface data where available, but unlike other approaches relies primarily on both AOD and particle property information contained in the satellite retrievals to constrain a complex, physically based atmospheric dispersion model. This is especially helpful over the vast areas where surface measurements of aerosol concentration and type are not available. We now emphasize this in the Introduction and Conclusions, and mention it in the Abstract.

Minor comments:

RC1_3. Page 1, line 30: Why is that EC have much worse performance compared to other species?

AC1_3. Largely emitted from incomplete combustion, EC is a spatially heterogeneous primary species whose particulate phase chemistry and physics is very complex and difficult to model. This is reflected in Table S8, which shows low spatial correlation values and high root mean square error comparison between ground monitors and CMAQ outputs for EC. Appel et al., (2008) discuss overprediction of EC in January and August over western US by CMAQ. EC also relies heavily on the emissions inventory, and although there have been great strides in the past five years or so to improve EC estimates in the emissions inventory (e.g., residential wood combustion), there are (or at least very likely) still large errors in the inventory relating to EC emissions.

RC1_4. Page4, line1: 1km or 275m?

AC1_4. Revised.

RC1_5. Page 6, line 19-20: Will this interpolation process introduce biases?

AC1_5. Yes, downscaling CMAQ outputs using any interpolation method inherently introduces biases. Three cross-validation techniques were employed to evaluate the biases of the optimized dataset with respect to ground observations.

RC1_6. Page 13, line 23: How is the MAIAC AOD scaled before gap-filling MISR AOD? This seems not to be mentioned in the manuscript.

AC1_6. We have revised Section 3.3 for clarity as follows:

"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 is missing retrievals due to the presence of clouds. The combined AOD product is more complete than the MISR or MAIAC product alone.

The Fig. S1 scatterplots show MISR-RA AOD retrievals are higher than those retrieved by MAIAC, and much closer to the AERONET ground-truth values, for the three case study days with highest AOD. These scatterplots reinforce the need to scale MAIAC-retrieved AOD before gap-filling MISR-retrieved AOD fields. Based on Fig. S1, a study-specific AOD adjustment was applied to the MAIAC data; in addition, a filter with an upper bound of 0.4 was used for MAIAC retrievals to reduce potential cloud contamination. On days when Aqua and Terra MAIAC C6v2 AOD retrievals on the 1 km fixed sampling grid were available, the MAIAC-Aqua AOD retrievals were used to fill in missing AOD in the MAIAC-Terra AOD maps (as MAIAC-Terra is closest in time to the MISR-RA retrieval) by linearly regressing values from a 15 x 15 MAIAC-Aqua grid cell region centered on the missing MAIAC-Terra cell value. The 1 km gap-filled MAIAC-Terra AOD maps were subsequently downscaled and spatially interpolated (via bilinear interpolation) to match the downscaled CMAQ 275 m × 275 m output grid, referred to herein as gap-filled MAIAC. Before combining retrieved AOD products, the 275 m × 275 m MISR-RA AOD at 558 nm was converted to 550 nm using the retrieved ANG product, and the dynamic sampling grid was re-gridded to match the downscaled CMAQ 275 m × 275 m grid. The gap-filled MAIAC product was then used to fill in gaps in the MISR-RA AOD product by linearly regressing values from a 15 x 15 gap-filled MAIAC grid cell region centered on the missing MISR-RA cell value. Larger gaps caused by cloud contamination in the satellite-retrieved AOD were filled using a 7 x 7 grid cell region of CMAQ-reconstructed AOD value, linearly regressed to the satellite-retrieved AOD. This procedure was repeated multiple times as needed until the satellite retrieval area within the SJV study region was filled, referred herein as $\tau_{FillSAT}$.

A unique component of this work involves the use of the MISR-RA aerosol species-specific groups. Consequently, 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 gap-filling procedure used for $\tau_{FillSAT}$."

RC1_7. In Section 3.4, there are a lot of sentences (e.g. line 25-27 on page 14) reported the evaluation results, which should not belong to the Method section.

AC1_7. The sections 3.4.1 and 3.4.2 comparisons are AERONET validation, critical to the choices made in subsequent steps and, thus, were kept in the Methods section.

RC1_8. Figure 6: Although the OPT results had better agreement with ground observations, it still lacks of spatial coverage, even on the selected days with more MISR coverage.

AC1_8. Please see the response to comment RC1_1 above.

Author Comments (AC) to Referee Comments (RC) #2 – Anonymous Referee #1

Received and published: 6 June 2018

The paper provided a rigorous and detailed analysis of using satellite data (MISR, MODIS), surface observations (AERONET, PM2.5 and aerosol speciation), and CMAQ to derive surface PM2.5 and surface PM speciation. The novelty of this paper, as pointed by the authors, is the use of aerosol type information retrieved from MISR research algorithm. This, however, is really not new, which is also acknowledged in the paper - past work by Liu et al. has used MSIR aerosol type already. The paper also developed several methods for data gap filling, data fusion, and reconstruction of surface PM2.5 and total AOD from CMAQ. To this reviewer, the most interesting part is indeed the latter, as it has been vague in past studies on how PM2.5 mass is indeed computed with CTM outputs.

AC2_0. We thank the reviewer for the encouragement and the valuable comments. All the comments have been addressed in the revised manuscript. Please see our itemized responses below.

The approach here is fundamentally different – Liu et al. used a statistical approach, whereas we present here a complementary, physical approach. The underlying model being refined here is the CTM (CMAQ) rather than a regression model. Furthermore, we use as model constraints the particle size and light-absorption information from MISR, in addition to the particle shape, in a novel manner consistent with the limitations of the data. This is now emphasized in the Introduction and Conclusions of the paper.

The paper has done an excellent job in organizing its structure and presenting the detailed analysis. The paper, however, can be further improved by acknowledging other work done in the past that used satellite observations and CTM together to improve estimate of surface PM2.5. In various places, simplification and summary of the results (from the supplements) can make the paper more easier to read, keep the text flow smoother, improve the clarity, and ultimately enable more readability.

The paper can be published after the following concerns/comments are fully addressed.

General concerns/comments:

RC2_1) The title of the paper. The work of this paper in essence is data fusion and statistical analysis by combining data from various sources. While CTM outputs are used, the satellite data here really didn't provide any constraint for improving CTM MODELING that entails emissions, meteorology, different atmospheric processes, and data assimilation. It is recommended to add 'outputs' after 'modeling' in the tittle to avoid confusion, or change the title to emphasize the data fusion part. This paper didn't improve any components in CTM modeling; instead, it belongs to research of "model output statistics" (MOS) to postprocess model outputs.

AC2_1. We have added the word "outputs" as recommended.

RC2_2. P2, L3. not sure what 'a systematic and practical approach' means here. As pointed by the first reviewer, there have been much work that combine satellite and ground-based observations already. Please see the summary paper by Hoff and Christopher (2010) prior to 2010 and many other works afterwards. Indeed, the study here is demonstrated for the days and locations that have field campaign data and fewer clouds (compared to many other regions that studied). So, further discussion of the application of the method here in other places is needed.

AC2_2. Please see the responses AC1_1, AC1_2, and AC2_0, above. There are fundamental differences with our approach that provide certain advantages. We have made a larger point of the differences and advantages in the revised text.

RC2_3. Overall, in what percentage spatially, the AOD values are filled based on MAIA AOD (and scaling factor based on MISR/MODIS AOD ratio)?

AC2_3. We gap-fill using results from the MODIS Multi-Angle Implementation of Atmospheric Correction advanced algorithm (**MAIAC;** Lyapustin et al., 2018). The MISR AOD case study retrievals had 70% or greater spatial coverage of the SJV boundary delineated in Figure 1 and AOD approximately above the 0.15 threshold. Thus, MAIAC (not MAIA) AOD was used to gap-fill between 0 to 30 % of the MISR AOD scene for each case. This information is on Page 8 Line 17) in the text.

Specific concerns/comments:

RC2_4. P1, L25. This is a bit confusing. AOD is at 2 km resolution, while aerosol mass type can be retrieved at 275 m resolution? why not AOD at 275 m?

AC2_4. We have removed the phrase "2 km resolution", which referred to the CTM output.

RC2_5. P1, L30. R2 is only one of the measures for agreement. How about mean bias and RMSE?

AC2_5. We have added RMSE to the abstract. The remaining statistics are reported in Table S8.

RC2_6. P2, L26. Also emissions and parametrization schemes, especially for CTM. See Ge et al., JGR, 2017.

AC2_6. We revised the sentence as "The accuracy of the simulated fields is also affected by the accuracy of the simulated meteorology, emissions, and of the physical and chemical parameterization schemes specified in the model (Cooke et al., 1999; Tong and Mauzerall, 2006; Monks et al., 2009)" and added citations.

RC2_7. P3, L4, it is worth mentioning that early studies, while neglecting these factors (speciation and vertical profile), indeed acknowledge the importance of these factors such as in Wang and Christopher (2003). The current writing gives readers an impression that these early studies didn't recognize the importance of these factors, which in not true. These factors have been recognized since the beginning (Wang and Christopher, 2003).

AC2_7. We revised the sentence as "Early space-based PM2.5 air quality studies directly correlated satellite-derived AOD from the MODerate resolution Imaging Spectroradiometer (MODIS) instruments and ground-level PM2.5 concentrations acknowledged, but did not account for, particle vertical distribution, day-to-day variations, and/or aerosol speciation (Chu et al., 2003;Wang and Christopher, 2003;Engel-Cox et al., 2004;Chu, 2006;Gupta and Christopher, 2009;Wallace and Kanaroglou, 2007;Schaap et al., 2009;Zhang et al., 2009;Hu and Rao, 2009;Tsai et al., 2011;Hu et al., 2014)" and added citations.

RC2_8. P3, L10. It is worth mentioning that all the work cited here has inconsistence of aerosol optical properties between models and satellite retrieval algorithms. Work has been done that uses CTMs to inform aerosol types for the retrieval from satellites, which in turn improve the estimate of surface PM2.5 from CTM. References include Drury et al. (2010, JGR), Wang et al. (2010, RSE), and van Donkelaar et al., 2013.

AC2_8. We added the sentence "Work has been done to improve estimates of surface $PM_{2.5}$ from CTM by improving the consistency of aerosol optical properties between models and satellite retrieval algorithms, as well as, using CTMs to inform satellite-retrieved aerosol types (Drury et al., 2010; Wang et al., 2010; Donkelaar et al., 2013). However, we map the MISR RA constraints on spherical light-absorbing, spherical non-absorbing, and non-spherical particles to the appropriate aerosol chemical species in the CTM, which is different from previous work." and added citations.

RC2_9. P5, L5. How long is the DISCOVR-AQ time period? In average, what are the percentage of days that MISR AOD has good spatial converge and AOD is higher than 0.15?

AC2_9. The DISCOVER-AQ SJV deployment ran from 16 January through 08 February 2013. Approximately half of the MISR retrievals met the case requirements of 70% or greater spatial coverage of the SJV boundary delineated in Figure 1 and AOD approximately above the 0.15 threshold. This information has been added to Page 6 Line 3 in the text.

RC2_10. P6, L10-20. How many layers in the vertical and in the boundary layer? What is fire emission inventory used? Is CTM outputs data saved at every hour?

AC2_10. The CMAQ domain consisted of 35 vertical layers with varying thickness extending from the surface to 50 hPa and an approximately 10 m midpoint for the lowest (surface) model layer. CMAQ outputs are saved hourly. This information has been added to the text. We used U.S. EPA 2011 NEI emissions data with 2013 updates to fire sources. The wildfire emissions used came from SMARTFIRE v2 (https://www.airfire.org/smartfire/).

RC2_11. P7, section 2.5.1. MISR-RA. How does MISR-RA AOD compare with MISR operational AOD? Does MSIR operational product offer the aerosol type retrievals? Using MISR operational product would seem more practical. It will be nice to have some justification here.

AC2_11. The current 4.4 km x 4.4 km MISR Standard Algorithm (SA) AOD product was not available at the time of the evaluation and is not available at higher resolution. The SA has greater inconsistencies in aerosol particle retrievals due to limitations in the aerosol climatology included in the algorithm (74 mixtures for the SA vs. over 700 for the RA), poorer surface-reflectance assumptions, issues with the radiometric calibration critical for aerosol-type retrievals that are corrected in the RA, details of the acceptance criteria, and the spatial resolution at which the algorithm is run. For more details, please see the series of papers by Limbacher and Kahn (2014; 2016; 2017; 2018). For particle-type retrievals, the RA performs considerably better than the SA. The information has been added to the text.

RC2_12. P11, equation 1. Does CMAQ offer concentration of Al, Si, Ca, Ti, etc? if not, please give the exact equation used in reconstructing CMAQ PM2.5 in this research, so readers don't have to refer to the supplement often.

AC2_12. Yes, CMAQ does offer these and other dust related concentrations. The aerosol group equations used in this study are included in the supplement for brevity and are specific to CMAQ v5.0.2.

RC2_13. P11, equation 2. How are the values for negative and positive terms in right-hand side of equation obtained in this study?

AC2_13. The equation 2 terms are calculated using CMAQ and WRF outputs. These equations are discussed in detail elsewhere (Frank, 2006;Chow et al., 2015) and are referenced appropriately in our paper.

RC2_14. P12, equation 2. fRH (upper case) in equation 3, but frh (lower case) in L15

AC2_14. We revised the hygroscopic growth factor parameter to f_{rh}.

RC2_15. P12, L24. This is not correct. The extinction per unit length is called extinction coefficient, and it is inversely proportional to visibility; see details in Kessner et al., Atmospheric Environment, 2013.

AC2_15. The sentences were revised as follows:
"The ambient particle extinction as a function of height is the sum of the ambient scattering and absorption with respect to altitude (z), which are the two terms in Eq. (3). When integrated over a horizontal path, the extinction per unit length is sometimes called the visibility, typically reported in Mm^{-1}. From Eq. 3, the dimensionless extinction AOD is obtained by multiplying the ambient particle extinction by the vertical atmospheric path height of each CMAQ layer."

RC2_16. P13, and P14; AOD gap filling using MODIS. How to scale MAIA AOD exactly? In cases where both Terra and Aqua MODIS have AOD, is it only Terra MODIS AOD used? Some details are needed here, including when the method works best and when may not work well (such as with large cloud cover).

AC2_16. We gap-fill using results from the MODIS Multi-Angle Implementation of Atmospheric Correction advanced algorithm (**MAIAC**; Lyapustin et al., 2018). Please see response AC_6 above, as Section 3.3 has been revised for clarity.

RC2_17. P14, L20-25. it will be good to show a scatter plot that summarizes the comparison for all days in one plot? Also, a plot showing the comparison for data filling only (e.g., in places/times that has no MISR AOD, but filled with MODIS AOD and through scaling/interpolation) can be good to show the improvement by combining both MODIS and MISR.

AC2_17. Please see Figures 3, S1, and S2. Aerosol airmass types and spatial distribution change over time. It is not clear to the authors what a scatterplot of all the days in one plot would contribute to the focus of this paper. The individual density scatter plots comparing AERONET, MISR, MAIAC, gap-filled MISR, and CMAQ are shown in Figures S1 and S2.

RC2_18. P15, L12. What happens in hours that have cloud? Daily AOD from AERONET has a clear-sky bias.

AC2_18. The AERONET clear-sky bias is a limitation of the satellite-based and AERONET AOD comparison. In the optimized dataset the pixels within the domain of study with no satellite-based retrievals rely on the fused CMAQ and ground observations. See also our response to RC1_1.

RC2_19. P15, L31 . not sure what 'sufficient' mean here?

AC2_19. We have removed the word "sufficient."

RC2_20. P15, L11. There are papers talking about diurnal variation of AOD. for example Kaufman et al. in GRL. Are the results here consistent with previous findings?

AC2_20. The context of diurnal variation of AOD in the Kaufman et al. (2000) paper is with respect to climate changes and thus compares annual aerosol measurements. The conclusion that a 10:30 AM AOD represents the diurnal average applies to specific location and/or scenarios. Fire events are some examples of situations where this conclusion does not apply. In our paper, we are interested in capturing diurnal variations with respect to short-term changes (please see Figures 4 and S2). Furthermore, it is difficult to compare diurnal variation conclusions from other study sites due to: (1) the unique weather pattern and pollution transport characteristic of the SJV (i.e., persistent inversion and very low PBL height), (2) differences in product version uncertainty (i.e., AERONET versions between this and earlier studies), and (3) disparity in satellite-retrieved spatial resolution (i.e., biases in earlier studies due to coarser spatial resolution).

RC2_21. equation 4. This equation is not correct. equation 3 won't give equation 4 as beta, fRH all depends on Z, and C(z) varies with Z.

AC2_21. It is correct that equation 4 depends on altitude. Therefore, we specifically use the height-stratified hygroscopic growth and specific dry extinction efficiency factors from Step 2 for equation 4 in Step 4. Equations 4 through 6 and their descriptions have been updated to make clear these are column-average dry particle concentrations.

RC2_22. for the results. It will be good to show the summary as several scatter plots respectively for PM2.5 and speciation in all days and sites in the main manuscript. Having summary statistics (such as R, RMSE, and mean bias) in figure. Are the results or improvement by MISR statistically significant?

AC_22. A limitation mentioned in the paper is that the study domain and timeframe did not offer a substantial quantity of suborbital observations for assessing the results statistically. Statistical power is a known issue in this study. Yet, we performed three separate statistical tests to establish as best we could the significance of the results. The comprehensive set of the various summary statistics can be found in the supplemental material.

RC2_23. P24, L5. Worthy mentioning recent studies that used VIIRS DNB to derive surface PM2.5 at night. see Wang et al., AE, 2016; Fu et al., 2018.

AC2_23. The following sentence has been added to the diurnal sampling segment of the conclusion: "Future research assessing diurnal sampling could benefit from the inclusion of Visible Infrared Imaging Radiometer Suite (VIIRS) instrument datasets, such as daylight-retrieved AOD (Jackson et al., 2013) and Day/Night Band as an estimate of $PM_{2.5}$ surface change (Wang et al., 2016)."

Constraining Chemical Transport $PM_{2.5}$ Modeling Outputs Using Surface Monitor Measurements and Satellite Retrievals: Application over the San Joaquin Valley

Mariel D. Friberg[1, 2], Ralph A. Kahn[1], James A. Limbacher[1, 3], K. Wyat Appel[4], James A. Mulholland[2]

[1]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2]School of Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[3]Science Systems and Applications Inc., Lanham, MD 20706, USA
[4]US EPA, Research Triangle Park, NC 27711, USA

Correspondence to: Mariel D. Friberg (mariel.d.friberg@nasa.gov)

Abstract. Advances in satellite retrieval of aerosol type can improve the accuracy of near-surface air quality characterization, by providing broad regional context and decreasing metric uncertainties and errors. The frequent, spatially extensive and radiometrically consistent instantaneous constraints can be especially useful in areas away from ground monitors and progressively downwind of emission sources. We present a physical approach to constraining regional-scale estimates of $PM_{2.5}$, its major chemical component species estimates, and related uncertainty estimates of chemical transport model (CTM, e.g. the Community Multi-scale Air Quality Model) outputs. This approach uses ground-based monitors where available, combined with aerosol optical depth and qualitative constraints on aerosol size, shape, and light-absorption properties from the Multi-angle Imaging SpectroRadiometer (MISR) on the NASA Earth Observing System's Terra satellite. The CTM complements these data by providing complete spatial and temporal coverage. Unlike widely used approaches that train statistical regression models, the technique developed here leverages CTM physical constraints such as conservation of aerosol mass and meteorological consistency, independent of observations. The CTM also aids in identifying relationships between observed species concentrations and emission sources.

Aerosol airmass types over populated regions of Central California are characterized using satellite data acquired during the 2013 San Joaquin field deployment of the NASA DISCOVER-AQ project. We investigate the optimal application of incorporating 275 m horizontal-resolution aerosol airmass-type maps and total-column aerosol optical depth from the MISR Research Aerosol retrieval algorithm (RA) into regional-scale CTM output. The impact on surface $PM_{2.5}$ fields progressively downwind of large single sources is evaluated using contemporaneous surface observations. Spatiotemporal R^2 and RMSE values for the model, constrained by both satellite and surface-monitor measurements based on 10-fold cross-validation, are 0.79 and 0.33 for $PM_{2.5}$, 0.88 and 0.65 for NO_3^-, 0.78 and 0.23 for SO_4^{2-}, 1.00 and 1.01 for NH_4^+, 0.73 and 0.23 for OC, and 0.31 and 0.65 for EC, respectively. 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 unconstrained CTM simulations, and provide finer spatial

Mariel Friberg 7/12/2018 5:13 PM
Deleted: . In addition

Mariel Friberg 7/12/2018 5:13 PM
Deleted: ,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: single-scattering albedo provided by multi-angle instruments, such as

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , can provide frequent, spatially extensive, instantaneous constraints on chemical transport models (CTMs), which can be 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

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , offering additional

Mariel Friberg 7/12/2018 5:13 PM
Deleted: (e.g.,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: ,

Mariel Friberg 7/12/2018 5:13 PM
Deleted:)

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , and aid

Mariel Friberg 7/12/2018 5:13 PM
Deleted: . Incorporating satellite aerosol information in the development of $PM_{2.5}$ concentration metrics can lead to a decrease in metric uncertainties and errors

Mariel Friberg 7/12/2018 5:13 PM
Deleted: This work focuses on the degree t ... [1]

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Southern

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Using the MISR Research Aeros ... [2]

Mariel Friberg 7/12/2018 5:13 PM
Deleted: and evaluate

Mariel Friberg 7/12/2018 5:13 PM
Deleted: a 2 km resolution,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , to obtain constrained fields of

Mariel Friberg 7/12/2018 5:13 PM
Deleted: . Contemporaneous surface obser ... [3]

Mariel Friberg 7/12/2018 5:13 PM
Deleted: discussed. The spatiotemporal R^2

Mariel Friberg 7/12/2018 5:13 PM
Deleted: % withholding

resolution. SO_4^{2-} cross-validation values showed the largest spatial and spatiotemporal R^2 improvement with a 43% increase. Assessing this physical technique in a well-instrumented region opens the possibility of applying it globally, especially over areas where surface air quality measurements are scarce or entirely absent.

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** using the satellite data to apply the technique

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 (Laden et al., 2006;Pope et al., 2009;Özkaynak et al., 2009). Even for the continental US, many ambient, ground-level fine particulate matter ($PM_{2.5}$) chemical datasets are acquired only once every three or six days, and data records at many sites are less than a decade or two long. In addition, the monitors tend to be concentrated in a small number of populated counties, with the exception of the Interagency Monitoring of Projected Visual Environment (IMPROVE) program sites located primarily in US national parks (Hand et al., 2011). Prior to 2009, instrument types and sensitivities varied from monitor to monitor and among monitoring networks (Chow et al., 2010), making comparisons and uncertainty assessment difficult.

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** for

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** considerably

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** s

Urban-level epidemiological time-series studies often span large geographic regions (Goldstein and Landovitz, 1977;Wade et al., 2006). Especially for long-term exposure analysis, broad regions within or downwind of urban and industrial centers are also of concern due to the presence of distributed populations, and natural and agricultural ecosystems. Characterizing spatial variability is fundamental to effectively conducting environmental epidemiologic studies and air quality assessments. 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 (Ito and Thurston, 1995;Pinto et al., 2004;Goldman et al., 2012). This is particularly relevant for regional-scale studies, where urban-to-rural gradients of ambient surface $PM_{2.5}$ and chemical species concentrations are often lacking.

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** the spatial variability (e.g.

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:**)

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** , which is fundamental to effectively conducting environmental epidemiologic studies and air quality assessments, can be

Although chemical transport model (CTM) simulations provide more complete spatial and temporal coverage than surface monitors, 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, emissions, and of the physical and chemical parameterization schemes specified in the model (Cooke et al., 1999;Monks et al., 2009;Tong and Mauzerall, 2006). 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 sparsely monitored areas.

> Mariel Friberg 7/12/2018 5:13 PM
> **Deleted:** parameters

Early space-based $PM_{2.5}$ air quality studies directly correlated satellite-derived AOD from the MODerate resolution Imaging Spectroradiometer (MODIS) instruments and ground-level $PM_{2.5}$ concentrations acknowledged, but did not account for, particle vertical distribution, day-to-day variations, and/or aerosol speciation (Chu et al., 2003;Wang and Christopher, 2003;Engel-Cox et al., 2004;Chu, 2006;Gupta and Christopher, 2009;Wallace and Kanaroglou, 2007;Schaap et al., 2009;Zhang et al., 2009;Hu and Rao, 2009;Tsai et al., 2011;Hu et al., 2014). This direct total-column AOD-to-surface $PM_{2.5}$ correlation approach works well when the aerosol is almost entirely concentrated in the near-surface boundary layer, but suffers when transported aerosol makes a significant contribution to the total-column AOD, or when the boundary layer is deep or variable on short timescales, as has been pointed out by Hidy et al. (2009). Other early studies used surface measurements (Al-Saadi et al., 2005) or CTMs (Liu et al., 2004;Koelemeijer et al., 2006;Mathur, 2008;Van Donkelaar et al., 2010;Drury et al., 2010;Wang et al., 2010; Van Donkelaar et al., 2013;Boys et al., 2014;Ma et al., 2014) to provide some constraint on aerosol vertical distribution, but did not account in detail for either spatial or temporal variations in the relationship between total-column AOD amounts and surface $PM_{2.5}$ concentrations, and provided very limited or no aerosol-type constraints. Work has been done to improve CTM estimates of surface $PM_{2.5}$ by improving the consistency of aerosol optical properties between models and satellite retrieval algorithms, as well as using CTMs to inform satellite-retrieved aerosol types (Van Donkelaar et al., 2013;Wang et al., 2010;Drury et al., 2010; Li et al., 2015). 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. Advanced statistical models that use land-use, meteorological, and relative humidity parameters have been applied to increase the accuracy of AOD-to-$PM_{2.5}$ estimates (Kumar et al., 2007;Di Nicolantonio et al., 2009;Lee et al., 2011;Kloog et al., 2012;Hu et al., 2014; Ma et al., 2014;Song et al., 2014;Lv et al., 2016;Ma et al., 2016). Several of these statistical models are location-specific, and most rely on surface-based data training sets to constrain parameters in statistical models, which are then applied elsewhere. Where training data are limited or entirely absent, there is significant uncertainty with this approach.

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 US, and constrained aerosol vertical distribution and speciated the spherical components with an aerosol transport model (Liu et al., 2007b;Liu et al., 2007a). Subsequent work applied MISR aerosol size and shape constraints over the Indian subcontinent and surrounding areas to map seasonal changes in aerosol type (Dey et al., 2012) and combined MISR particle shape and qualitative light-absorption information to make a first effort at mapping aerosol airmass types over an urban area, i.e., Mexico City (Patadia et al., 2013).

In the current study, we introduce and enhance a physical approach that takes advantage of satellite coverage over regional scales for estimating ambient $PM_{2.5}$ mass and associated chemically speciated concentrations, as needed in air quality

Mariel Friberg 7/12/2018 5:13 PM
Deleted: that were based upon space-based data simply…irectly correlated ground-level PM2 ... [4]

Mariel Friberg 7/12/2018 5:13 PM
Deleted: To estimate…n the current study, ... [5]

applications. The approach uses ground-based $PM_{2.5}$ measurements, where available, to anchor speciated, near-surface CTM aerosol concentrations. To help constrain the model outputs over extended regions, both MISR total-column AOD and qualitative, column-effective aerosol type observations are also applied when retrieval quality is adequate (generally, where mid-visible AOD values exceed 0.15). Specifically, we map the satellite-retrieved constraints on spherical light-absorbing, spherical non-absorbing, and non-spherical particles to the appropriate aerosol chemical species in the CTM, which is substantially different from previous work. Enhanced aerosol-type retrievals from the MISR Research Aerosol (MISR-RA) retrieval algorithm (Kahn et al., 2001;Limbacher and Kahn, 2014), at 275 m 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-RA coverage. The results account for spatiotemporal variability in $PM_{2.5}$ and 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 2 describes the datasets involved, Sect. 3 describes the method and technical approach, and Sect. 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 Sect. 5, and detailed data and ancillary documentation are provided in Supplemental Material.

2 Study Domain and Datasets

2.1 Study Domain

The San Joaquin Valley (SJV), which comprises the southern two-thirds of California's Central Valley (about 26,000 km^2), has long suffered from severe air pollution issues and is among the most studied air sheds in the US (Ngo et al., 2010;Chow et al., 2006). The SJV has complex topography and meteorology, particularly in winter, when low planetary boundary layer (PBL) heights and high pollutant mixing ratios create a challenging environment for chemical transport modeling (Hidy et al., 2009; Appel et al., 2017). 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 (Fig. 1). Although primarily a rural area, the eight counties that comprise the SJV are home to more than 4 million residents. Despite the semi-arid climate, the SJV is one of the world's most productive agricultural regions (Schoups et al., 2005). Its air shed frequently experiences high $PM_{2.5}$ concentrations during the winter due to the combination of relatively dry climate, shallow PBL heights, local source emissions, and the surrounding mountain ranges. The region 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 US (EPA, 2012).

Mariel Friberg 7/12/2018 5:13 PM
Deleted: is also applied, along with

Mariel Friberg 7/12/2018 5:13 PM
Deleted: (i.e., size, shape, and single scattering albedo)

Mariel Friberg 7/12/2018 5:13 PM
Deleted: .

Mariel Friberg 7/12/2018 5:13 PM
Deleted: 1 km

Mariel Friberg 7/12/2018 5:13 PM
Deleted: of

Mariel Friberg 7/12/2018 5:13 PM
Deleted: The SJV

Mariel Friberg 7/12/2018 5:13 PM
Deleted: SJV

The study period for this work was selected to coincide with the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ; http://discover-aq.larc.nasa.gov) field campaign, which ran from January 16th through February 8th 2013. This campaign was a joint collaboration between NASA, NOAA, US EPA, multiple universities, and several local organizations, with the goal of characterizing air quality in urban areas using satellite, aircraft, vertical profiler and ground-based measurements. Targeting the 2013 DISCOVER-AQ deployment period for this study provides considerable ground- and aircraft-based measurements for aerosols and fine particulate matter, which we apply as model constraints and for evaluation.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , January and February 2013,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: .

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 aircraft 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 20th, February 3rd, and February 5th. On lower-AOD days, MISR-RA aerosol type information has higher uncertainty for the current application and thus the analysis of speciated $PM_{2.5}$ focuses on the higher-AOD days. Of the three higher-AOD days, January 20th has the least cloud cover, followed by February 5th, so these days are 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, North Africa, and many major metropolitan areas.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: will be

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

2.2 Ground-based PM Mass and Speciated Measurements

This study focuses on $PM_{2.5}$ mass and the five components that dominate total $PM_{2.5}$ in the SJV: 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) and Federal Equivalence Method (FEM) mass from the Air Quality System (AQS; https://www.epa.gov/aqs), and (2) daily averaged total $PM_{2.5}$ and chemically speciated mass (measurements typically made every third or sixth day) from the Chemical Speciation Network (Solomon et al., 2014).

FRM compliant data from gravimetric filter-based samplers and FEM compliant data from continuous mass monitors provide spatial variability of $PM_{2.5}$ mass (EPA, 2004). 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°C). Monitor locations are shown in Fig. 1, and Table 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.

2.3 DISCOVER-AQ AERONET DRAGON

The AErosol RObotic NETwork (Holben et al., 1998) has 10 permanent sun photometer (SP) monitors operating in the study region. During the DISCOVER-AQ mid-January through mid-February 2013 deployment, these monitors were supplemented with an additional 14 temporary monitors 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 multiple spectral bands from the ultraviolet (~340 nm) to the near-infrared (~1640 nm), with an accuracy within ±0.015 (Eck et al., 1999).

We use version 2 level 2 (L2) AERONET/DRAGON AOD and Angstrom Exponent (ANG) data for the six study days. The L2 data were sun-calibrated after field deployment, cloud screened (Smirnov et al., 2000), and quality controlled. The AOD at 550 nm wavelength is calculated using a quadratic log-log fit to AERONET observations at shorter and longer wavelengths (Eck et al., 1999). Columnar AODs at 550 nm derived from AERONET are considered as AOD ground truth in this study.

2.4 Chemical Transport Model Simulations

Simulations of the coupled Weather Research and Forecasting model (WRF, Skamarock et al., 2008), version 3.4, and the Community Multiscale Air Quality model (CMAQ, Byun and Schere, 2006), version 5.0.2, were obtained from the US Environmental Protection Agency (EPA). These hourly atmospheric simulations, at 2 km × 2 km horizontal grid spacing cover the entire SJV and surrounding major cities during the months of January and February of 2013. The CMAQ domain consisted of 35 vertical layers with varying thickness extending from the surface to 50 hPa and an approximately 10 m midpoint for the lowest (surface) model layer. Concentration fields from the fixed 2 km × 2 km horizontal CMAQ grid were downscaled to a horizontal grid of 275 m x 275 m by linear interpolation and used as the reference grid for all subsequent analyses. Emission data were based on the 2011 EPA National Emissions Inventory (EPA, 2015) with 2013 updates to electric generating unit emissions, fire, and mobile sources. Biogenic emissions were generated in-line to CMAQ using the Biogenic Emissions Inventory System (BEIS; http://www.cmascenter.org) version 3.14, and the emissions were processed using the Sparse Matrix Operator Kernel Emissions (Houyoux et al., 2000) version 3.5. The carbon bond 2005 chemical mechanism used was CB05TULC (Sarwar et al., 2012;Whitten et al., 2010;Yarwood et al., 2005). 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. Boundary conditions for the 4 km

Mariel Friberg 7/12/2018 5:13 PM
Deleted: (UV, ~

Mariel Friberg 7/12/2018 5:13 PM
Deleted: (NIR, ~

Mariel Friberg 7/12/2018 5:13 PM
Deleted: (v2)

Mariel Friberg 7/12/2018 5:13 PM
Deleted: 550nm

Mariel Friberg 7/12/2018 5:13 PM
Deleted: 550nm

Mariel Friberg 7/12/2018 5:13 PM
Deleted: with 35 vertical layers

simulation were derived from a 36 km simulation covering the contiguous US, and BCs for the 36 km simulation were provided by a GEOS-Chem simulation (Bey et al., 2001) with the chemical species mapped to the corresponding CMAQ species (Appel et al., 2017).

The EPA conducted a model evaluation of CMAQ v5.0.2 with respect to the scientific updates to v5.1 (Appel et al., 2017). 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 (Model - Observations) in the SJV exceeded -10 µg m^{-3}. Errors in simulated PBL height and mixing were considered to be contributing factors to the January $PM_{2.5}$ underestimation in the SJV. Although CMAQ v5.0.2 is missing several secondary organic aerosol species of anthropogenic volatile organic carbon (i.e., long-chain alkanes and naphthalene) 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 µg m^{-3}) in the SJV (Appel et al., 2017). At the time this study was conducted, CMAQ v5.1 results were not yet available.

2.5 Satellite Observations

The primary satellite resource for this study is the MISR instrument. We supplement the MISR-RA aerosol data with results from the MODIS instruments. They offer more extensive spatial coverage and provide up to two observations per day (one in the morning, and one in the early afternoon), though with larger AOD uncertainty over land, and with no constraints on aerosol type over land (Levy et al., 2013). We describe these two data sources below.

2.5.1 MISR-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 (Diner et al., 1998). 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 the four spectral bands centered at 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 (Bruegge et al., 2007;Limbacher and Kahn, 2015;Limbacher and Kahn, 2017), the MISR-RA is uniquely capable of supporting air-quality applications by providing information about aerosol microphysical properties at regional scales. The Version 23 4.4 km x 4.4 km MISR Standard Algorithm (MISR-SA) AOD product was not available at the time of the evaluation and is not available at higher resolution. The MISR-SA has greater inconsistencies in aerosol particle retrievals due to limitations in the aerosol climatology included in the algorithm (74 mixtures for the MISR-SA vs. over 700 for the MISR-RA), poorer surface-reflectance assumptions, issues with the

Mariel Friberg 7/12/2018 5:13 PM
Deleted: United States

Mariel Friberg 7/12/2018 5:13 PM
Deleted: CMAQv5

Mariel Friberg 7/12/2018 5:13 PM
Deleted: – Research Aerosol Retrieval Algorithm (RA)

radiometric calibration critical for aerosol-type retrievals that are corrected in the MISR-RA, details of the acceptance criteria, and the spatial resolution at which the algorithm is run. More details are available in the series of papers by Limbacher and Kahn (2014; 2015; 2017). For particle-type retrievals, the MISR-RA performs considerably better than the MISR-SA.

High-resolution (275 m) results from the MISR-RA are used to constrain aerosol concentration 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 MISR-RA provides column-averaged information regarding aerosol properties under favorable retrieval conditions (i.e., cloud-free, low surface albedo, mid-visible AOD exceeding about 0.15) (Kahn and Gaitley, 2015;Kahn et al., 2010). This information amounts to constraints on particle shape (non-spherical (dust) vs. spherical AOD fraction), particle size (typically three-to-five bins, e.g., "small," "medium," and "large" AOD fraction, parameterized as the Angstrom Exponent, ANG), and particle light absorption (typically two-to-four bins, e.g., "dirty" and "clean," represented as single-scattering albedo, SSA = 1.0 – [absorbing AOD]/[total AOD]). 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). A brief summary of the MISR-RA retrieval process is provided in the supplemental Sect. S1.

Following the work of Patadia et al. (2013), we identify different aerosol airmasses by categorizing aerosol based on the qualitative particle size, shape, and light-absorption constraints described above. Specifically, for the purposes of this paper, the 14 aerosol components used by all 774 mixtures included in the refined MISR-RA aerosol climatology (Limbacher and Kahn, 2014) can be organized into three broad aerosol-type "groups": spherical light-absorbing, spherical non-absorbing, and non-spherical (cirrus is ignored in the current application). Especially at low-AOD, the MISR-RA-derived aerosol-type sensitivity amounts to no more than these three groupings (Kahn and Gaitley, 2015). However, 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 below in Sect. 3.2. From the point-of-view of retrieval sensitivity, these three categories map to common aerosol species as follows (Table S2): (1) Light-Absorbing Carbon (LAC), (2) Inorganic Ions (II) plus Organic Matter (OM) plus Sea-Salt (SS), and (3) dust. Section S2 in supplemental provides a description of how the aggregated AOD retrievals are computed for the spherical absorbing aerosol components, and separately for spherical non-absorbing aerosol components. It is well established that MISR AOD retrievals suffer biases for scenes with substantial cloud cover (Witek et al., 2013;Shi et al., 2014;Limbacher and Kahn, 2015). Consistent with both Witek et al. (2013) and Limbacher and Kahn (2015), we present results only for days where clouds cover less than 30% of the scene within the SJV as indicated by the MISR-RA cloud mask, excluding the rural areas that extend into the Sierra Nevada mountains.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: retrieval algorithm

Mariel Friberg 7/12/2018 5:13 PM
Deleted:]/

Mariel Friberg 7/12/2018 5:13 PM
Deleted:).

Mariel Friberg 7/12/2018 5:13 PM
Deleted:).

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Over the last several years, substantial research has been published indicating

Mariel Friberg 7/12/2018 5:13 PM
Deleted: in the presence of clouds

Mariel Friberg 7/12/2018 5:13 PM
Deleted: our

Mariel Friberg 7/12/2018 5:13 PM
Deleted: county

2.5.2 MODIS - MAIAC

To supplement the MISR-RA AOD values where MISR coverage is lacking, we adopt results from the MODIS Multi-Angle Implementation of Atmospheric Correction (MAIAC) advanced algorithm (Lyapustin et al., 2018;Lyapustin et al., 2012), which 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 version 2.0 (C6v2) 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. (2018; 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 Collection 6 (C6) Terra-to-Aqua cross-calibration (Lyapustin et al., 2018), 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 the surface Bi-directional Reflectance Distribution Function (BRDF; 0.466 µm), 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 MODIS cross-track scanned scenes (i.e., cloud-free conditions with relatively low AOD) over several days to acquire multi-angle sets of observations for each location. This allows MAIAC to retrieve the 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 (Kloog et al., 2014).

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

Mariel Friberg 7/12/2018 5:13 PM
Deleted: (C6)

Mariel Friberg 7/12/2018 5:13 PM
Deleted: 2011a,b

Mariel Friberg 7/12/2018 5:13 PM
Deleted:).

are unreliable for very low AOD conditions, over mountainous terrain, and over surfaces with very high albedo. The retrieval conditions that affect this study include low AOD and some cloud-contaminated scenes.

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-surface-measurement biases over large areas, especially regions far from concentrated surface monitors. 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-monitor 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 Fig. 2. Note that the left side of Fig. 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 stars in Fig. 2 indicate where uncertainties are estimated by comparison with AERONET and the EPA ground monitors, respectively. 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. In Step 2, spatially complete AOD and grouped AOD maps are produced for each of the six study days by combining MISR-RA 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 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-monitor data, satellite AOD snapshots and AOD grouped by aerosol type, and the CMAQ model simulations. The outputs are the fused ground-monitor, satellite, plus model $PM_{2.5}$ mass concentration field, and speciated versions of this field. A detailed description of the key steps follows.

Mariel Friberg 7/12/2018 5:13 PM
Moved (insertion) [1]

Mariel Friberg 7/12/2018 5:13 PM
Moved (insertion) [2]

Mariel Friberg 7/12/2018 5:13 PM
Moved up [1]: Note that the left side of Fig. 2 tracks the process for deriving total $PM_{2.5}$, whereas the right side presents the flow for speciated $PM_{2.5}$.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Blue and orange asterisks in Fig.

Mariel Friberg 7/12/2018 5:13 PM
Moved up [2]: 2 indicate where uncertainties are estimated by comparison with AERONET and the EPA ground monitors, respectively.

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 $PM_{2.5}$. This 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 (Frank, 2006;Hand et al., 2011;Hand et al., 2014;Malm et al., 2011). 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 US (Malm et al., 2011). Chow et al. (2015) provides a detailed literature review of the various mass reconstruction equations

Mariel Friberg 7/12/2018 5:13 PM
Deleted: .

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 (Chow et al., 1994;Chow and Egami, 1997;Andrews et al., 2000;Nolte et al., 2015); (2) organic matter (DeBell et al., 2006;Hand et al., 2011); (3) Elemental Carbon (EC), also referred to as light absorbing carbon (Bond and Bergstrom, 2006); (4) crustal material, which includes mineral and soil particles, referred to herein as dust (Malm et al., 1994;Malm et al., 2011); (5) sea salt (Hand et al., 2011); and (6) other elements (Simon et al., 2011), which, in the SJV during the study period, made a negligible contribution to $PM_{2.5}$. The respective references provide details as to how multipliers for each species were derived and summarize the evaluation performed for each major PM component.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: ,

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 S1 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 (Eq. (A) in Chow et al., 2015):

$$RM[\mu g\ m^{-3}] = \underbrace{[SO_4^{=}] + [NH_4^{+}] + [NO_3^{-}]}_{Inorganic\ Ions} + \underbrace{1.8[OC]}_{\substack{Organic\\ Matter}} + \underbrace{[EC]}_{\substack{Light\\ Absorbing\\ Carbon}} + \underbrace{1.8[Cl^{-}]}_{\substack{Sea\\ Salt}} + \underbrace{2.2[Al] + 2.49[Si] + 1.63[Ca] + 1.94[Ti] + 2.42[Fe]}_{Dust} \quad (1)$$

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 (Andrews et al., 2000;Rees et al., 2004;Tanner et al., 2004;El-Zanan et al., 2005). Using Eq. (1) to estimate OM from OC for CMAQ output allows for consistency with satellite-derived estimates; in the future we may expand the method to include various organic aerosol species explicitly in cases where we have more *in situ* data.

Following the framework of Eq. (1), the reconstructed $PM_{2.5}$ mass does not account for the positive and negative factors that affect gravimetric and speciated measurements (DeBell et al., 2006;Frank, 2006;Hand et al., 2011;Chow et al., 2015). To close the mass-balance difference between $PM_{2.5}$ FRM gravimetric mass and ambient mass (simulated and measured), the material balance Eq. (1) was adjusted to account for factors affecting gravimetric measurements (Eq. (10) in Frank, et al., 2006).

$$PM_{2.5FRM}\ [\mu g\ m^{-3}] = RM - ([\mathrm{NH_4^+}]_{loss} + [\mathrm{NO_3^-}]_{loss}) + [PBW] + [Blank_{FRM}] \quad (2)$$

where ammonium and nitrate volatilization are not captured by gravimetric measurements and thus, are accounted as negative artifacts. The particle bound water (PBW) 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 (Frank, 2006;Chow et al., 2015). The uncertainty estimated for the CMAQ- and satellite-based surface concentrations are discussed in Sect. 4.

3.2 Step 2 – CMAQ-based Columnar AOD and AOD Subcategorized into Species-related Groups Derived Using the Reconstructed Extinction Coefficient Method

Section 3.1 summarizes 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 AOD 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 US as part of IMPROVE. Since then this method has been continuously upgraded by several researchers (Malm et al., 1994;Malm et al., 2000;Malm et al., 2011;Song et al., 2008;Park et al., 2011). The process estimates extinction AOD using simulated concentrations of II, OM, SS, LAC, and Dust (Table S1) assuming externally mixed aerosols with respect to the modeled altitude (z), as follows:

$$\tau = \int \left\{ \underbrace{\sum_i \omega_{(z),i} \beta_{de,i} f_{rh(z),i} C_{(z)i}}_{particle\ scattering\ efficiency} + \underbrace{\sum_i (1-\omega_{(z),i}) \beta_{de,i} f_{rh(z),i} C_{(z)i}}_{particle\ absorption\ efficiency} \right\} dz \quad (3)$$

Mariel Friberg 7/12/2018 5:13 PM
Deleted: d

Mariel Friberg 7/12/2018 5:13 PM
Deleted: $\int \left\{ \underbrace{\sum_i \omega_i \beta_{de,i} f_{RH(z),i} C_{(z)i}}_{particle\ scattering\ efficiency} + \underbrace{\sum_i (1-\omega_i) \beta_{de,i} f_{RH(z),i} C_{(z)i}}_{particle\ absorption\ efficiency} \right\} dz$

where

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

i = chemical component

ω = single scattering albedo (SSA)

β_{de} = specific dry mass extinction efficiency [m^2 g^{-1}]

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

C = concentration of chemical component i as a function of height [g m^{-3}]

Equation 3 is further subdivided for dust by size in accordance with the CMAQ Aitken, accumulation, and coarse particles size categories (Park et al., 2011). The empirically based factors and their respective literature sources are summarized in Table S3. The WRF simulated relative humidity data, *rh(z)*, were used to evaluate the height-dependent hygroscopic growth factors. The ambient particle extinction as a function of height is the sum of the ambient scattering and absorption with respect to altitude (z), which are the two terms in Eq. (3). When integrated over a horizontal path, the extinction per unit length is sometimes called the visibility, typically reported in Mm^{-1}. From Eq. 3, the dimensionless extinction AOD is obtained by multiplying the ambient particle extinction by the vertical atmospheric path height of each CMAQ layer. These are added vertically to obtain columnar AOD values, which are compared to ground- and satellite-based AOD values in the following subsections to assess uncertainties.

The three CMAQ-based AOD AG (i.e., LAC, II+OM+SS, and Dust), indicated in Table S2, are calculated using the five major chemical components derived in Eq. (1). 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 Sect. 3.4 below.

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-RA-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 is missing retrievals due to the presence of clouds. The combined AOD product is more complete than the MISR-RA or MAIAC product alone.

The Fig. S1 scatterplots show MISR-RA AOD retrievals are higher than those retrieved by MAIAC, and much closer to the AERONET ground-truth values, for the three case study days with highest AOD. These scatterplots reinforce the need to scale MAIAC-retrieved AOD before gap-filling MISR-RA-retrieved AOD fields. Based on Fig. S1, a study-specific AOD

Mariel Friberg 7/12/2018 5:13 PM
Deleted: per mass

Mariel Friberg 7/12/2018 5:13 PM
Deleted: per unit length

Mariel Friberg 7/12/2018 5:13 PM
Deleted: per unit length,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: per unit length

Mariel Friberg 7/12/2018 5:13 PM
Deleted: length

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , referred herein as $\tau_{FillSAT}$,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: s

Mariel Friberg 7/12/2018 5:13 PM
Moved down [3]: A unique component of this work involves the use of the MISR-RA aerosol species-specific groups.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: The MISR-RA and MODIS-based MAIAC satellite-retrieved products were combined to improve spatial coverage. Before combining retrieved AOD products, the MISR-RA AOD at 558nm was converted to 550 nm using the retrieved ANG product. The MAIAC maps were downscaled and spatially interpolated (via bilinear interpolation) to match the downscaled model 275 m × 275 m grid. Because CMAQ and MAIAC have fixed sampling grids at each location, whereas the MISR grid varies when the same location of observed from different sub-spacecraft paths, 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, referred to herein as gap-filled MAIAC. The scatterplots in Fig. S1

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Fig. S1 reinforces

adjustment was applied to the MAIAC data; in addition, a filter with an upper bound of 0.4 was used for MAIAC retrievals to reduce potential cloud contamination. On days when Aqua and Terra MAIAC C6v2 AOD retrievals on the 1 km fixed sampling grid were available, the MAIAC-Aqua AOD retrievals were used to fill in missing AOD in the MAIAC-Terra AOD maps (as MAIAC-Terra is closest in time to the MISR-RA retrieval) by linearly regressing values from a 15 x 15 MAIAC-Aqua grid cell region centered on the missing MAIAC-Terra cell value. The 1 km gap-filled MAIAC-Terra AOD maps were subsequently downscaled and spatially interpolated (via bilinear interpolation) to match the downscaled CMAQ 275 m × 275 m output grid, referred to herein as gap-filled MAIAC. Before combining retrieved AOD products, the 275 m × 275 m MISR-RA AOD at 558 nm was converted to 550 nm using the retrieved ANG product and the dynamic sampling grid was re-gridded to match the downscaled CMAQ 275 m × 275 m grid. The gap-filled MAIAC product was then used to fill in gaps in the MISR-RA AOD product by linearly regressing values from a 15 x 15 gap-filled MAIAC grid cell region centered on the missing MISR-RA cell value. Larger gaps caused by cloud contamination in the satellite-retrieved AOD were filled using a 7 x 7 grid cell region of CMAQ-reconstructed AOD value, linearly regressed to the satellite-retrieved AOD. This procedure was repeated multiple times as needed until the satellite retrieval area within the SJV study region was filled, referred herein as $\tau_{FillSAT}$.

A unique component of this work involves the use of the MISR-RA aerosol species-specific groups. Consequently, we produce gap-filled, aerosol-type-grouped AODs from the original MISR-RA-based AG AODs using the model-based grouped AODs from Step 1 and following the same gap-filling procedure used for $\tau_{FillSAT}$.

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. This material is presented here rather than in Section 4 below because key decisions in the Method depend on the results of these comparisons.

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

Evaluation of MISR-RA (Limbacher and Kahn, 2014) and MAIAC (Lyapustin et al., 2011) 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 across the SJV (Fig. 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

Mariel Friberg 7/12/2018 5:13 PM
Deleted: .

Mariel Friberg 7/12/2018 5:13 PM
Moved (insertion) [3]

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Additionally

Mariel Friberg 7/12/2018 5:13 PM
Deleted: ,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: total AOD.

polar-orbiting coverage for a given day. Fig. 3 and Table S4 provide scatterplots and a statistical summary, respectively, of AERONET AOD collocated in time and space with the MISR-RA, MAIAC, gap-filled MISR-RA AOD (i.e., $\tau_{FillSAT}$), and CMAQ results. Although AERONET reports AOD at 550 nm, AOD values at 558 nm were calculated for comparison with the MISR-RA 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.

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 S4). 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-RA 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-RA and MAIAC satellite-retrieved AODs with AERONET also illustrates how gap-filling MISR-RA with scaled and gap-filled MAIAC retrievals produces a more consistent product. For example, the Fig. 3 subplot for February 5th shows that gap-filled MISR-RA (i.e., FillSAT) offers better agreement than gap-filled MAIAC at the averaged AERONET retrieved AOD value of 0.47. On this specific day and location there is no coincident MISR-RA retrieval, indicating that the gap-filled MISR-RA improvement is due to scaled and gap-filled MAIAC used to gap-fill the MISR-RA AOD snapshot. Further evident from Fig. 3, 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.

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 for the study region, we compare the satellite retrievals (MISR-RA, MAIAC, gap-filled MISR-RA) the with daylight-averaged AERONET-retrieved AOD results (Fig. S2). 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 (Fig. 4) 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.

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. For the MISR-RA comparison, we have only the same MISR-RA AOD retrieval snapshots as in Fig. 3. 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-RA overpasses are in some cases gap-filled with other scaled-MAIAC Terra/Aqua retrievals acquired during that day. A third satellite-retrieved AOD product is the gap-filled, primarily MISR-RA-derived AOD (*FillSAT*) described in Sect. 3.3. Also shown in Fig. S2 are the CMAQ reconstructed daylight-average AODs, described in Sect. 3.2.

Overall, the MISR-RA 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. Table S5 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-RA agree better statistically than the other datasets in terms of correlation and error relative to AERONET daylight-average values. Although the retrieved AODs for the MISR-RA and gap-filled MISR-RA 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-RA 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-RA-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 agreement with daylight-averaged ground truth to serve as proxies for the daylight-averaged values during the study period.

A procedure for fusing CMAQ model simulations with surface-based measurements is described briefly in Sect. S3 in the supplemental material, and in detail in Friberg et al. (2016). This procedure was applied to C_{SURF} and C_{CMAQ} (Fig. 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 (Fig. 4 and Table S6). 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

Mariel Friberg 7/12/2018 5:13 PM
Deleted: sufficient

Mariel Friberg 7/12/2018 5:13 PM
Deleted: C_{OBS}

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 is approximately less than 0.15. At these smaller AODs, a small absolute change in AOD will produce larger percent changes.

One possible reason for the scatter in Fig. 4 is the model representation of transported aerosol. Transported aerosol above the boundary layer is dependent on the lower BCs adopted in the model, and thus is not always well represented by CMAQ 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, warrants further investigation. 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) (Kahn et al., 2008).

3.5 Step 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 ($C^{PM_{2.5}FRM}_{FillSAT,z=0}$ and $C^{Speciated}_{FillSAT,z=0}$) are obtained from the total-column satellite AOD ($\tau_{FillSAT}$) and aerosol group AOD ($\tau^{AG}_{FillSAT}$) 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 Eq. (4), the column-average dry particle concentrations for the three aerosol groups ($\overline{C^{AG}_{FillSAT}}$) are calculated from the AODs, $\tau_{FillSAT}$ and $\tau^{AG}_{FillSAT}$, by reversing the reconstructed extinction process applied to model-only values in Step 2 (Eq. (3)). The same height-stratified hygroscopic growth and specific dry scattering or absorbing efficiency factors from Step 2 are used here for consistency. The column-average satellite-based AG concentrations ($\overline{C^{AG}_{FillSAT}}$) are further stratified into the five column-average representative PM chemical components ($Columnar\ C^{Speciated}_{FillSAT}$), defined in Step 1 according to Eq. (1), using the CMAQ-based species-to-aerosol group partition in Eq. (5). With $\overline{C^{Speciated}_{FillSAT}}$ defined, satellite-based total-column $PM_{2.5}$ ($\overline{C^{PM_{2.5}}_{FillSAT}}$) is obtained using Eq. (1). The satellite-derived column-average concentrations are then scaled to surface-level concentrations by relying on the vertical distribution of the CMAQ simulations of each species in Eq. (6). The

Mariel Friberg 7/12/2018 5:13 PM
Deleted: <~
Mariel Friberg 7/12/2018 5:13 PM
Deleted: or 0.2
Mariel Friberg 7/12/2018 5:13 PM
Deleted: needs to be investigated
Mariel Friberg 7/12/2018 5:13 PM
Deleted: $Surface\ C^{PM_{2.5}FRM}_{FillSAT}$ and $C^{Speciated}_{FillSAT}$
Mariel Friberg 7/12/2018 5:13 PM
Deleted: total-
Mariel Friberg 7/12/2018 5:13 PM
Deleted: $Columnar\ C^{AG}_{FillSAT}$)
Mariel Friberg 7/12/2018 5:13 PM
Deleted: total-
Mariel Friberg 7/12/2018 5:13 PM
Deleted: $Columnar\ C^{AG}_{FillSAT}$)
Mariel Friberg 7/12/2018 5:13 PM
Deleted: total-
Mariel Friberg 7/12/2018 5:13 PM
Deleted: $Columnar\ C^{Speciated}_{FillSAT}$
Mariel Friberg 7/12/2018 5:13 PM
Deleted: $Columnar\ C^{PM_{2.5}}_{FillSAT}$)
Mariel Friberg 7/12/2018 5:13 PM
Deleted: total-
Mariel Friberg 7/12/2018 5:13 PM
Deleted: apportioned

satellite-based surface-level $PM_{2.5}$ concentrations ($C^{PM_{2.5}}_{FillSAT,z=0}$) are adjusted to reflect $PM_{2.5\,FRM}$ concentrations using Eq. (2). These relationships were defined in terms of daily AOD and species concentrations.

$$\overline{C^{AG}_{FillSAT}} = \frac{\tau^{AG}_{FillSAT}}{\int(\beta_{de,i} f_{rh(z),i})dz} \quad (4)$$

$$\overline{C^{Speciated}_{FillSAT}} = \overline{C^{AG}_{FillSAT}}\left(\overline{C^{Speciated}_{CMAQ}}/\overline{C^{AG}_{CMAQ}}\right) \quad (5)$$

$$C^{Speciated}_{FillSAT,z=0} = \overline{C^{Speciated}_{FillSAT}}\left(C^{Speciated}_{CMAQ,z=0}/\overline{C^{Speciated}_{CMAQ}}\right) \quad (6)$$

3.6 Step 5 –Optimized $PM_{2.5}$ FRM and Speciated Concentrations, Derived by Fusing Satellite-constrained Values with Ground-monitor Data

The optimized concentration dataset (C_{OPT}) closely parallels the surface-measurement-constrained CMAQ simulation described in Eq. (S4). The C_{OPT} dataset is derived by constraining the results with the surface-monitor data near their locations, and weighting the satellite-constrained concentration values progressively more heavily away from available ground monitors (Fig. 5). Using Eq. (7), the six daily C_{CMAQ} fields coincident with the flight campaign span are replaced with the satellite-derived daily $C_{FillSAT}$ fields, as these were the days when retrieval conditions were adequate to use the data for the current application (See Sect. 2.1 above). With only 11.5% of the C_{CMAQ} fields changing due to contributions from the surface stations, the weighting factors (W; Eq. (S5)) and average temporal correlations between the simulations and surface observations (R_2; Eq. (S7)), across all monitors, did not need to be recalculated. Thus, for this study, C_{OPT} diverges from C_{FCMAQ} for 6 days out of the entire study time period.

$$C_{Opt_{s,t}} = \alpha\overline{C_{CMAQ_s}}^{\beta}\left[W_{s,t}\left\{\frac{C_{SURF_{s_m,t}}}{C_{SURF_{s_m}}}\right\}_{krig} + (1-W_{s,t})\left\{\frac{C_{FillSAT_{s,t}}}{C_{CMAQ_s}}\right\}\right] \quad (7)$$

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

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, and to identify biases that arise from different sampling frequencies and spatial distribution of monitors across the pollutants. First, a 10-fold 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 $PM_{2.5}$. Brief descriptions of these tests are given here; the results of the tests are discussed in Section 4.2 below.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Columnar $C^{AG}_{FillSAT} = \frac{\tau^{AG}_{FillSAT}}{\int \beta_{dsae,i} f_{RH,i} dz}$ (4) ... [6]

Mariel Friberg 7/12/2018 5:13 PM
Deleted: in

Mariel Friberg 7/12/2018 5:13 PM
Deleted: two-month

Mariel Friberg 7/12/2018 5:13 PM
Deleted: 2-month

Mariel Friberg 7/12/2018 5:13 PM
Deleted: $\left[W_{s,t}\left\{\frac{C_{OBS_{s_m,t}}}{C_{OBS_{s_m}}}\right\}_{krig} + (1-W_{s,t})\left\{\frac{C_{FillSAT_{s,t}}}{C_{CMAQ_s}}\right\}\right]$

Mariel Friberg 7/12/2018 5:13 PM
Deleted: tenfold

Mariel Friberg 7/12/2018 5:13 PM
Deleted: %

3.7.1 10-fold Cross-Validation

The dataset performance was evaluated using a 10-fold cross-validation analysis. For each of 10 independently run trials, a random 10% of the surface observations were held back per day and each method ("fused," i.e., surface measurements + model, and "optimized," i.e., surface + satellite measurements + model) 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 surface observations for each pollutant (Table 1), ranging from 44 for $PM_{2.5}$-OC to 779 for $PM_{2.5}$.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: 10% Holdout

Mariel Friberg 7/12/2018 5:13 PM
Deleted: tenfold 10% withholding

3.7.2 Leave-one-out Cross-Validation

As an alternative to the 10-WH method, the LOO withholding is applied to the five PM 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 Fig. 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.

3.7.3 Regional Holdout Cross-Validation

A regional withholding technique is used to evaluate fused and optimized $PM_{2.5}$ datasets, as monitor clustering affects the cross-validation results. For each of the cross-validation regions in Fig. 1, all but one of the monitors in a region is withheld, and this is repeated independently for each daily monitor and region. The approach approximates the evaluation of LOO when the distance between monitor locations is large (i.e., >50 km).

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-constrained results for $PM_{2.5}$ and $PM_{2.5}$ grouped by species are evaluated against EPA AQS and CSN ground observations, respectively. For the second set of analyses, cross-validation is used to evaluate satellite-constrained model performance. The main objectives of this section are (1) to evaluate the results of Steps 2-5 as much as possible, (for evaluation of 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 $PM_{2.5}$.

4.1 Comparison of Satellite-constrained 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}), model-fused-with-ground-monitor (C_{FCMAQ}), deconstructed satellite-constrained ($C_{FillSAT}$) and optimized (C_{OPT}; model + ground monitor + satellite) daily averaged $PM_{2.5}$ and speciated component concentrations with EPA AQS and CSN observations. Table S7 provides a statistical summary of the comparison between the ground truth and the modeled, fused, satellite-constrained, and optimized results, stratified by pollutant, day, and dataset. Fig. 6 presents concentration maps of the four aforementioned datasets with embedded ground-truth $PM_{2.5}$ values and their respective RGB images (depicting cloud coverage) for the three days with relatively high AOD in the study set (January 20th, February 3rd, and February 5th).

Focusing on the area within the SJV, the higher concentration gradients in $C_{FillSAT}$ are due to the application of satellite snapshots. The satellite-constrained concentration snapshots tend to provide more realistic spatial distributions of $PM_{2.5}$ compared to the unconstrained model values, C_{CMAQ}. 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 snapshot 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 primarily in low-AOD regions. This suggests that the technique will yield increasingly good results when applied in more heavily polluted areas around the globe. Fig. S3 presents scatterplots comparing the daily averaged models and the satellite-constrained 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 high-AOD days (Jan. 20th, Feb. 3rd, and Feb. 5th), Fig. S3 shows $C_{FillSAT}$ $PM_{2.5}$ is in general agreement with surface observations within the performance range of the model results, and the variability is minimal, especially compared to low AOD days. Of the three relatively high-AOD days, January 20th has the least amount of cloud contamination, whereas February 5th has the most. Following the Fig. 5 weighting between the datasets, the visible contributions of the C_{CMAQ} and $C_{FillSAT}$ datasets to the C_{FCMAQ} and C_{OPT} $PM_{2.5}$ fields in Fig. 6 occur at distances of a fifth to a half degree (20 to 50 km) beyond a monitor. At or near a ground observation, the C_{OPT} fields are weighted towards the interpolated surface-observation fields, whereas the influence of $C_{FillSAT}$ on C_{OPT} improves the regional behaviour and enhances the spatial gradient structure synoptically. For C_{CMAQ} and $C_{FillSAT}$, the estimated temporal variances are fairly constant and do not depend on distance to the surface observations. The surface observations, rather than model or satellite-based results, dominate the C_{FCMAQ} and C_{OPT} temporal correlations at and near monitor locations, whereas C_{CMAQ} and $C_{FillSAT}$ dominate at distances 20 to 50 km beyond a monitor. As such, the temporal correlations for C_{CMAQ}, $C_{FillSAT}$, C_{FCMAQ}, and C_{OPT} generally do not approach zero away from the surface stations. For example, on February 5th, the interpolated surface-observation field dominates both the satellite and CMAQ values in the C_{OPT} and C_{FCMAQ} $PM_{2.5}$ maps. The situation at Bakersfield on this day is a bit different. Here the assumed surface monitor uncertainty plays a role, as CMAQ reports a much lower value, the satellite contribution is weighted significantly some distance from the urban center, and the actual difference between the monitor and the C_{OPT} field is about 12.5%, though the contrast appears large due to the color scale. The satellite contribution is investigated further and quantified in the validation exercises of the next section, where we systematically decrease the dependence of C_{OPT} fields on surface observations.

Figures 7 and S4 provide speciated NO_3, NH_4, and SO_4 surface concentration maps for January 20th and February 3rd; respectively; ground-truth data, available only for February 3rd, are included in Fig. 7. For the evaluation of the modeled and 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-monitor sampling infrequency, as evident in the correlation ranges (Table S7). Fig. S4 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. Unlike for $PM_{2.5}$, there were no speciated monitor measurements available on January 20th, so the OPT results are equal to FillSAT (Fig. S4). Although the C_{CMAQ} and $C_{FillSAT}$ results show agreement around the locations of known emission sources, the satellite-derived aerosol concentrations at the surface show more realistic horizontal dispersion patterns, and the spatial distribution better reflects the likely influence of topographic features. Specifically, during SJV winters, wide horizontal uniformity of ammonium nitrate concentrations is characteristic of this air basin, due to the near-surface inversion (Watson and Chow, 2002). Particulate nitrate is known to form over non-urban areas when high ammonia emissions from the surface, and nitric acid, formed aloft during night-time decoupling, mix during the morning collapse of the inversion (Watson and Chow, 2002). Throughout the region, consecutive days with low PBL heights are known to produce increased and spatially more uniform concentrations of fine particulate matter, nitrate, and sulfate (Watson and Chow, 2002). The $C_{FillSAT}$ spatial structure and background concentration ranges of 10-15 $\mu g\ m^{-3}$ for nitrate and 4-5 $\mu g\ m^{-3}$ for ammonium (Fig. S4) reflect the aforementioned concentration dynamics. The differences between the model and satellite-constrained concentration gradients within the SJV are visible on January 20th and February 3rd, and the related surface mixing and plume dispersion are evident, especially in Fig. S4. Given the very limited speciated monitor measurements available, the Fig. S5 scatterplots show $C_{FillSAT}$ provides better agreement than the model and fused-model values.

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 the 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 the AERONET, MISR-RA, and MAIAC products used.

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 ground observations to satellite-derived surface concentrations for $PM_{2.5}$ mass and speciated particles over the western US. The statistical regression technique used 3-hour averaged CTM (GEOS-Chem) results coincident with Terra overpasses 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= 0.23, SO_4 R^2= 0.11, and OC R^2= 0.11. In our study, the spatial R^2 values for $PM_{2.5}$ averaged 0.53 across all days and 0.73 on Jan. 20th, the clearest day with high AOD. The spatial R^2 values for the $C_{FillSAT}$ speciated PM on February 12th, the only day for which we have more than one surface measurement, are 0.48 for NO_3, 0.10 for SO_4, 0.46 for OC, 0.63 for NH_4, and 0.41 for EC.

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

The model, fused, and optimized datasets are included in the 10-WH cross-validation comparison with the monitor data. The RMSE, MB, and the spatiotemporal, temporal, and spatial mean correlations for the five datasets are presented in Table S8. 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 component 10-WH cross-validation statistics are affected by low numbers of available observations, sampling frequency, and coincident satellite-retrieval data, particularly for NH_4 and EC. As a result, when compared to C_{CMAQ}, the $C_{OPT\ 10\text{-}WH}$ 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} show the largest improvement over the unconstrained model, with a 43% increase 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 error differences based on proximity to monitors. Overall, the statistical improvement between the CMAQ simulations and cross-validated datasets suggest the empirically based mass reconstruction factors, specific dry efficiencies, and SSA values adopted were adequate for the SJV domain. The 5-cities study 10-WH cross-validation spatiotemporal R^2 ranges were 0.81-0.89 for SO_4, 0.67-0.83 for $PM_{2.5}$, 0.52-0.72 for NO_3, 0.43-0.80 for NH_4, and 0.32-0.51 for OC (Friberg et al., 2017). In light of the 5-cities study, the results for relatively homogeneous pollutants of secondary origin of this study fall within these ranges.

Unlike 10-WH, LOO cross-validation results allow us to leverage the spatial distribution of monitor locations throughout the domain. Table 2 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 $PM_{2.5}$ component species and outperformed temporal R^2 for C_{FCMAQ} and $C_{FCMAQ\ LOO}$ values by 10 and 8%. NH_4 cross-validation performance is highest for monitor locations closest to the agricultural emission sources in the southern area of domain. This finding agrees with the general expectation that aerosol type uncertainties being lowest when the mid-visible AOD is higher than ~0.15. For SO_4, 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.

To explore the $PM_{2.5}$ $C_{FillSAT}$ impact on C_{OPT}, i.e., combining the surface monitor data with the 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 3). As expected, removing $PM_{2.5}$ clustered monitors increased the cross-validated dataset 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 Fig. 1. Proximity to emission sources, meteorology, and topography contribute to the performance differences between northern regions 1 and 2, and southern regions 3 and 4. The dominant primary $PM_{2.5}$ mass emission sources (i.e., residential wood combustion, and motor vehicles) as well as the major secondary aerosols in the SJV are associated with urban hotspots such as Fresno and Bakersfield (Chen et al., 2007). Winter wind speeds in the SJV are typically below 4 m s^{-1} (Watson and Chow, 2002). 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 (Cahill et al., 2011;Hayes et al., 1989). During the winter, regional transport occurs when the nocturnal boundary layer is decoupled from the air aloft; as a result, these higher wind speeds aloft tend not to ventilate the surface, intensifying pollutant surface concentrations throughout the SJV (Chow et al., 1999), whereas dust originating from desert sources to the east and southeast is likely transported aloft.

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, especially where the AOD is higher than in the SJV cases available for the current study. Based on these results, including the satellite data improves 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, especially in areas where surface measurements are lacking.

5 Conclusions

Even in the best-monitored urban areas, ground-based networks have limited spatial coverage. Building on earlier work that produced a method for fusing surface-based measurements with model simulations (Friberg et al., 2016; 2017), the current

Mariel Friberg 7/12/2018 5:13 PM
Deleted: the threshold of

Mariel Friberg 7/12/2018 5:13 PM
Deleted: of

Mariel Friberg 7/12/2018 5:13 PM
Deleted: to

Mariel Friberg 7/12/2018 5:13 PM
Deleted: s

Mariel Friberg 7/12/2018 5:13 PM
Deleted: the

Mariel Friberg 7/12/2018 5:13 PM
Deleted: ,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: ,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , especially over extended regions downwind of major pollution sources.

study relies on both satellite-derived AOD and particle property information contained in the satellite retrievals as additional constraints on the model outputs. The strength of the satellite data is broad spatial coverage, providing radiances that tend to have uniform quality over space and time compared to most suborbital observation datasets. The satellite provides vastly more spatial coverage than the surface stations alone, and this is especially important downwind of major pollution sources. The main limitations of the satellite data are lack of vertical discrimination in most situations, lack of diurnal coverage, and only crude aerosol-type sensitivity, especially at low AOD. The physical approach presented here uses CTM simulation along with surface-based measurements to address these limitations. Where satellite data are missing or where the AOD is too low to provide reliable aerosol type from the MISR-RA, the method relies on the model, tuned, to the extent possible, by satellite and surface measurements.

Satellite and ground-based aerosol measurements were combined with numerical model simulations to: (1) generate aerosol airmass type maps covering the central California test region for 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 monitors.

Satellite data helps capture $PM_{2.5}$ distributions over large, under-sampled or un-sampled regions, and its fusion with model results tends 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). We used retrievals from the MISR-RA, to take advantage of the higher spatial resolution and greater aerosol-type accuracy and precision compared to the standard products. However, to avoid over-interpreting the data, we classified the satellite aerosol-type results into three broad groups for application as a model constraint: spherical light-absorbing, spherical non-absorbing, and non-spherical. The satellite-constrained concentration maps are spatially consistent with topography, typifying localized hotspots over known urban areas, and exhibiting realistic dispersion patterns in the SJV. Comparison with daylight-averaged AERONET and diurnally averaged CMAQ modeling demonstrated that, for AOD >~ 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 was reduced on higher AOD days. These results suggest satellite retrievals in conjunction with model results can improve $PM_{2.5}$ spatial characterization 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-monitor data, for nitrate, ammonium, and possibly also sulfate.

For the current study, model-based aerosol vertical distributions were used to address the lack of profile measurements. However, model aerosol vertical distribution could be constrained on large scales with space-based stereo imaging (e.g.,

Mariel Friberg 7/12/2018 5:13 PM
Deleted: incorporates

Mariel Friberg 7/12/2018 5:13 PM
Deleted: species-related AOD

Mariel Friberg 7/12/2018 5:13 PM
Deleted: The main limitations

Mariel Friberg 7/12/2018 5:13 PM
Deleted: model

Mariel Friberg 7/12/2018 5:13 PM
Deleted: Satellites help

Mariel Friberg 7/12/2018 5:13 PM
Deleted: was

from 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. Comparison of diurnal variation results to other studies was hindered by the following factors: (1) the unique weather pattern and pollution transport characteristic of the SJV (i.e., persistent inversion and very low PBL height), (2) differences in product version uncertainty (i.e., AERONET versions between this and earlier studies), and (3) disparity in satellite-retrieved spatial resolution (i.e., biases in earlier studies due to coarser spatial resolution). Future research assessing diurnal sampling could benefit from the inclusion of Visible Infrared Imaging Radiometer Suite (VIIRS) instrument datasets, such as daylight-retrieved AOD (Jackson et al., 2013) and Day/Night Band as an estimate of $PM_{2.5}$ surface change (Wang et al., 2016). 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 of emission sources. It is physically based in that it leverages components of a CTM such as the meteorology, conservation of aerosol mass, and assumed emissions, and complements statistical approaches that rely on tuning parameters in a regression-type model. The new method 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) is a possible next step toward establishing the strengths and limitations of the method. 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.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: This

Mariel Friberg 7/12/2018 5:13 PM
Deleted: method

Mariel Friberg 7/12/2018 5:13 PM
Deleted: , and

Mariel Friberg 7/12/2018 5:13 PM
Deleted: are

Mariel Friberg 7/12/2018 5:13 PM
Deleted: s

Acknowledgements

We thank the CSRA for creating the WRF meteorological inputs and emissions data used in the various model simulations; the US EPA for creating the CMAQ simulations; the US EPA for establishing and maintaining the AQS and CSN sites used in this work and the AQS datasets; the DISCOVER-AQ project (doi:10.5067/Aircraft/DISCOVER- AQ/Aerosol-TraceGas) for providing some of the field data used in this work; and the NASA AERONET network and its principal investigators, as well as their staffs, for establishing and maintaining the AERONET sites used in this work. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. The work of M. Friberg is supported under Grant Number NNX13AR89H issued through the NASA Education Minority University Research Education Project (MUREP) as part of the NASA Harriett G. Jenkins Graduate Fellowship program. The work of R. Kahn is supported in part by NASA's Climate and Radiation Research and Analysis Program under H. Maring and NASA's Atmospheric Composition Program under R. Eckman. Last but not least, we thank our anonymous reviewers who provided comments that have helped us refine this paper.

References

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

Andrews, E., Saxena, P., Musarra, S., Hildemann, L. M., Koutrakis, P., McMurry, P. H., Olmez, I., and 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, 50, 648-664, 2000.

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., and Foroutan, H.: Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1, Geosci. Model Dev., 10, 1703, 2017.

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., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research: Atmospheres, 106, 23073-23095, 2001.

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

Boys, B., Martin, R., Van Donkelaar, A., MacDonell, R., Hsu, N., Cooper, M., Yantosca, R., Lu, Z., Streets, D., and Zhang, Q.: Fifteen-year global time series of satellite-derived fine particulate matter, Environmental science & technology, 48, 11109-11118, 2014.

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

Byun, D., and 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, 59, 51-77, 2006.

Cahill, T. A., Barnes, D. E., Spada, N. J., Lawton, J. A., and Cahill, T. M.: Very fine and ultrafine metals and ischemic heart disease in the California central valley 1: 2003–2007, Aerosol Science and Technology, 45, 1123-1134, 2011.

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

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

Chow, J. C., and 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.

Chow, J. C., Watson, J. G., Lowenthal, D. H., Hackney, R., Magliano, K., Lehrman, D., and 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, 49, 16-24, 1999.

Chow, J. C., Chen, L. W. A., Watson, J. G., Lowenthal, D. H., Magliano, K. A., Turkiewicz, K., and 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, 111, 2006.

Chow, J. C., Watson, J. G., Chen, L. W., Rice, J., and Frank, N. H.: Quantification of PM 2.5 organic carbon sampling artifacts in US networks, Atmos. Chem. Phys., 10, 5223-5239, 2010.

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

Chu, D. A., Kaufman, Y., Zibordi, G., Chern, J., Mao, J., Li, C., and Holben, B.: Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS), Journal of Geophysical Research: Atmospheres, 108, 2003.

Chu, D. A.: Analysis of the relationship between MODIS aerosol optical depth and PM2. 5 in the summertime US, Proceedings of the SPIE, 2006, 629903,

Cooke, W., Liousse, C., Cachier, H., and Feichter, J.: Construction of a 1× 1 fossil fuel emission data set for carbonaceous aerosol and implementation and radiative impact in the ECHAM4 model, Journal of Geophysical Research: Atmospheres, 104, 22137-22162, 1999.

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

DeBell, L. J., Gebhart, K. A., Hand, J. L., Malm, W. C., Pitchford, M. L., Schichtel, B. A., and 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, 217-218, 2006.

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

Di Nicolantonio, W., Cacciari, A., and Tomasi, C.: Particulate matter at surface: Northern Italy monitoring based on satellite remote sensing, meteorological fields, and in-situ samplings, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2, 284-292, 2009.

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., and Gerstl, S. A. W.: Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview, IEEE Transactions On Geoscience and Remote Sensing, 36, 1072-1087, 1998.

Donkelaar, A., Martin, R. V., Spurr, R. J., Drury, E., Remer, L. A., Levy, R. C., and Wang, J.: Optimal estimation for global ground-level fine particulate matter concentrations, Journal of Geophysical Research: Atmospheres, 118, 5621-5636, 2013.

Drury, E., Jacob, D. J., Spurr, R. J., Wang, J., Shinozuka, Y., Anderson, B. E., Clarke, A. D., Dibb, J., McNaughton, C., and Weber, R.: Synthesis of satellite (MODIS), aircraft (ICARTT), and surface (IMPROVE, EPA-AQS, AERONET) aerosol observations over eastern North America to improve MODIS aerosol retrievals and constrain surface aerosol concentrations and sources, Journal of Geophysical Research: Atmospheres, 115, 2010.

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

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

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

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

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

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

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, 56, 500-511, 2006.

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.: Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution, Environmental science & technology, 50, 3695-3705, 2016.

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.: Daily ambient air pollution metrics for five cities: Evaluation of data-fusion-based estimates and uncertainties, Atmospheric Environment, 158, 36-50, 2017.

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

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

Gupta, P., and Christopher, S. A.: Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach, Journal of Geophysical Research: Atmospheres, 114, 2009.

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., and 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, 0737-5352, 2011.

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

Hayes, T. P., Kinney, J. J., and Wheeler, N. J.: California surface wind climatology, California Air Resources Board, Aerometric Data Division, Aerometric Projects and Laboratory Branch, Meteorology Section, 1989.

Hidy, G. M., Brook, J. R., Chow, J. C., Green, M., Husar, R. B., Lee, C., Scheffe, R. D., Swanson, A., and Watson, J. G.: Remote sensing of particulate pollution from space: have we reached the promised land?, Journal of the Air & Waste Management Association, 59, 1130-1139, 2009.

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

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

Hu, X., Waller, L., Lyapustin, A., Wang, Y., and Liu, Y.: 10-year spatial and temporal trends of PM 2.5 concentrations in the southeastern US estimated using high-resolution satellite data, Atmos. Chem. Phys., 14, 6301-6314, 2014.

Hu, Z., and 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, 8, 26, 2009.

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

Jackson, J. M., Liu, H., Laszlo, I., Kondragunta, S., Remer, L. A., Huang, J., and Huang, H. C.: Suomi-NPP VIIRS aerosol algorithms and data products, Journal of Geophysical Research: Atmospheres, 118, 2013.

Kahn, R., Banerjee, P., and McDonald, D.: Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean, Journal of Geophysical Research, 106, 18219-18238, 2001.

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

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

Kahn, R. A., and Gaitley, B. J.: An analysis of global aerosol type as retrieved by MISR, J. Geophys. Res. Atmos., 120, 4248-4281, 2015.

Kloog, I., Melly, S. J., Ridgway, W. L., Coull, B. A., and 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, 11, 40, 2012.

Kloog, I., Chudnovsky, A. A., Just, A. C., Nordio, F., Koutrakis, P., Coull, B. A., Lyapustin, A., Wang, Y., and 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, 95, 581-590, 2014.

Koelemeijer, R., Homan, C., and Matthijsen, J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, Atmospheric Environment, 40, 5304-5315, 2006.

Kumar, N., Chu, A., and Foster, A.: An empirical relationship between PM2. 5 and aerosol optical depth in Delhi Metropolitan, Atmospheric Environment, 41, 4492-4503, 2007.

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

Lee, H., Liu, Y., Coull, B., Schwartz, J., and Koutrakis, P.: A novel calibration approach of MODIS AOD data to predict PM2. 5 concentrations, Atmos. Chem. Phys., 11, 7991, 2011.

Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Patadia, F., and Hsu, N.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989, 2013.

Li, S., R.A. Kahn, M. Chin, M.J. Garay, and Y. Liu, 2015. Improving MISR retrieved aerosol microphysical properties using GoCART data. Atm. Meas. Tech., 8, 1157–1171, doi:10.5194/amt-8-1157-2015.

Limbacher, J., and Kahn, R.: MISR empirical stray light corrections in high-contrast scenes, Atmos. Meas. Tech., 8, 2015.

Limbacher, J. A., and Kahn, R. A.: MISR research-aerosol-algorithm refinements for dark water retrievals, Atmos. Meas. Tech., 7, 3989-4007, 2014.

Limbacher, J. A., and 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., 10, 1539, 2017.

Liu, Y., Park, R. J., Jacob, D. J., Li, Q., Kilaru, V., and Sarnat, J. A.: Mapping annual mean ground-level PM2.5 concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States, J. Geophys. Res. Atmos., 109, 2004.

Liu, Y., Koutrakis, P., and 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, 57, 1351-1359, 2007a.

Liu, Y., Koutrakis, P., Kahn, R., Turquety, S., and 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, 57, 1360-1369, 2007b.

Lv, B., Hu, Y., Chang, H. H., Russell, A. G., and Bai, Y.: Improving the accuracy of daily PM2. 5 distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in North China, Environmental science & technology, 50, 4752-4759, 2016.

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

Lyapustin, A., Wang, Y., Laszlo, I., and Korkin, S.: Improved cloud and snow screening in MAIAC aerosol retrievals using spectral and spatial analysis, Atmos. Meas. Tech., 5, 843-850, 2012.

Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC Algorithm, Atmos. Meas. Tech. Discuss., 2018, 1-50, 2018.

Ma, Z., Hu, X., Huang, L., Bi, J., and Liu, Y.: Estimating ground-level PM2. 5 in China using satellite remote sensing, Environmental science & technology, 48, 7436-7444, 2014.

Ma, Z., Liu, Y., Zhao, Q., Liu, M., Zhou, Y., and Bi, J.: Satellite-derived high resolution PM2. 5 concentrations in Yangtze River Delta Region of China using improved linear mixed effects model, Atmospheric Environment, 133, 156-164, 2016.

Malm, W. C., Sisler, J. F., Huffman, D., Eldred, R. A., and Cahill, T. A.: Spatial and seasonal trends in particle concentration and optical extinction in the United States. Journal of Geophysical Research: Atmospheres, 99, 1347-1370, 1994.

Malm, W. C., Day, D. E., and 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, 50, 686-700, 2000.

Malm, W. C., Schichtel, B. A., and 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, 61, 1131-1149, 2011.

Mathur, R.: Estimating the impact of the 2004 Alaskan forest fires on episodic particulate matter pollution over the eastern United States through assimilation of satellite-derived aerosol optical depths in a regional air quality model, Journal of Geophysical Research: Atmospheres, 113, 2008.

Monks, P., Granier, C., Fuzzi, S., Stohl, A., Williams, M., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., and Bey, I.: Atmospheric composition change–global and regional air quality, Atmospheric Environment, 43, 5268-5350, 2009.

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

Nolte, C. G., Appel, K. W., Kelly, J. T., Bhave, P. V., Fahey, K. M., Collett, J. J. L., Zhang, L., and 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., 8, 2877-2892, 2015.

Özkaynak, H., Frey, H. C., Burke, J., and 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, 43, 1641-1649, 2009.

Park, R. S., Song, C. H., Han, K. M., Park, M. E., Lee, S. S., Kim, S. B., and 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., 11, 12275-12296, 2011.

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

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

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

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

Sarwar, G., Simon, H., Bhave, P., and Yarwood, G.: Examining the impact of heterogeneous nitryl chloride production on air quality across the United States, Atmos. Chem. Phys., 12, 6455-6473, 2012.

Schaap, M., Apituley, A., Timmermans, R., Koelemeijer, R., and Leeuw, G. d.: Exploring the relation between aerosol optical depth and PM 2.5 at Cabauw, the Netherlands, Atmos. Chem. Phys., 9, 909-925, 2009.

Schoups, G., Hopmans, J. W., Young, C. A., Vrugt, J. A., Wallender, W. W., Tanji, K. K., and Panday, S.: Sustainability of irrigated agriculture in the San Joaquin Valley, California, Proceedings of the National Academy of Sciences, 102, 15352-15356, 2005.

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

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

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and 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.

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

Solomon, P. A., Crumpler, D., Flanagan, J. B., Jayanty, R. K. M., Rickman, E. E., and 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, 64, 1410-1438, 2014.

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., and 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., 8, 6627-6654, 2008.

Song, W., Jia, H., Huang, J., and Zhang, Y.: A satellite-based geographically weighted regression model for regional PM2. 5 estimation over the Pearl River Delta region in China, Remote Sensing of Environment, 154, 1-7, 2014.

Tanner, R. L., Parkhurst, W. J., Valente, M. L., and Phillips, W. D.: Regional composition of PM 2.5 aerosols measured at urban, rural and "background" sites in the Tennessee valley, Atmospheric Environment, 38, 3143-3153, 2004.

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

Tsai, T.-C., Jeng, Y.-J., Chu, D. A., Chen, J.-P., and Chang, S.-C.: Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008, Atmospheric Environment, 45, 4777-4788, 2011.

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

Wade, K. S., Mulholland, J. A., Marmur, A., Russell, A. G., Hartsell, B., Edgerton, E., Klein, M., Waller, L., Peel, J. L., and 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, 56, 876-888, 2006.

Wallace, J., and Kanaroglou, P.: An investigation of air pollution in southern Ontario, Canada, with MODIS and MISR aerosol data, Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, 2007, 4311-4314,

Wang, J., and Christopher, S. A.: Intercomparison between satellite-derived aerosol optical thickness and PM2. 5 mass: implications for air quality studies, Geophys. Res. Lett., 30, 2003.

Wang, J., Xu, X., Spurr, R., Wang, Y., and Drury, E.: Improved algorithm for MODIS satellite retrievals of aerosol optical thickness over land in dusty atmosphere: Implications for air quality monitoring in China, Remote Sensing of Environment, 114, 2575-2583, 2010.

Wang, J., Aegerter, C., Xu, X., and Szykman, J. J.: Potential application of VIIRS Day/Night Band for monitoring nighttime surface PM2. 5 air quality from space, Atmospheric Environment, 124, 55-63, 2016.

Watson, J. G., and Chow, J. C.: A wintertime PM 2.5 episode at the Fresno, CA, supersite, Atmospheric Environment, 36, 465-475, 2002.

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

Witek, M. L., Garay, M. J., Diner, D. J., and 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, 118, 2013.

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

Zhang, H., Hoff, R. M., and Engel-Cox, J. A.: The relation between Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth and PM2. 5 over the United States: a geographical comparison by US Environmental Protection Agency regions, Journal of the Air & Waste Management Association, 59, 1358-1369, 2009.

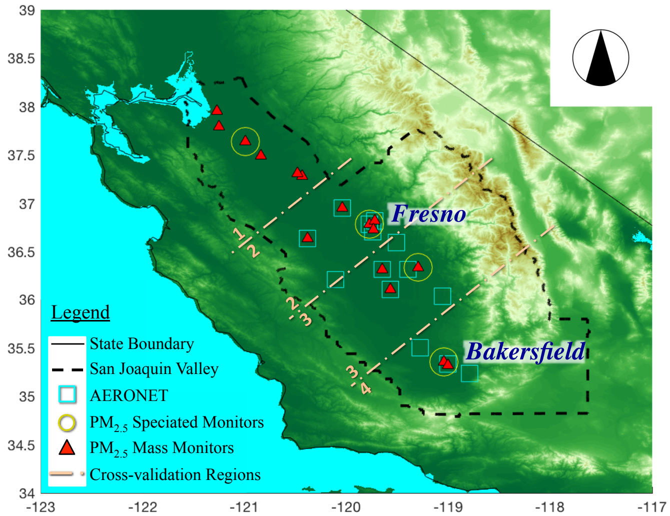


Figure 1: San Joaquin study area shows the ground elevation, EPA AQS and CSN monitors, and AERONET sites during the NASA DISCOVER-AQ flight campaign.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: two-month

Table 1: EPA AQS and CSN monitor summary statistics for 52 days (6 days).

Pollutant	No. of Monitors	Sampling Frequency	OBS	Mean	SD
$PM_{2.5}$, μg/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, μg/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, μg/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, μg/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, μg/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, μg/m^3	4 (4)	3 1-in-3; 1 1-in-6	44 (8)	5.25 (5.73)	3.09 (2.48)

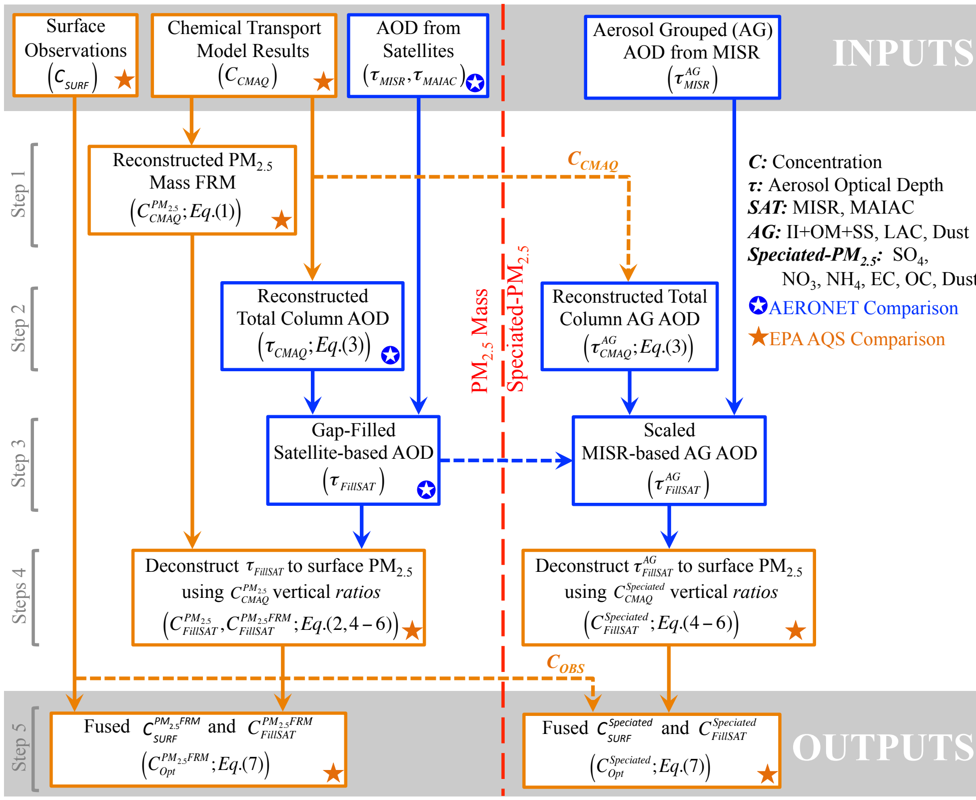
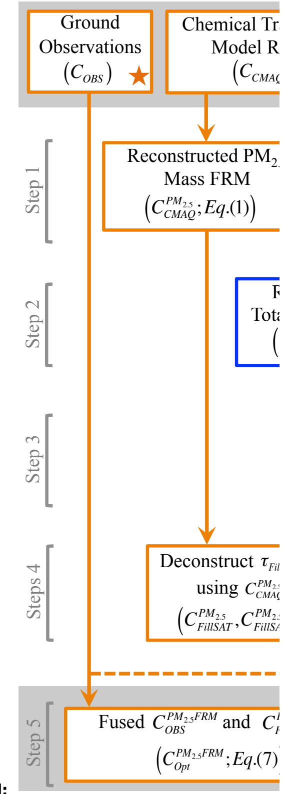


Figure 2: Methods flow chart connecting satellite retrieved AOD to Modeled AOD, PM2.5 Mass, and PM2.5 Speciated Mass. The parenthetical terms are defined in their respective steps in the methods section.

Mariel Friberg 7/12/2018 5:13 PM

Deleted:



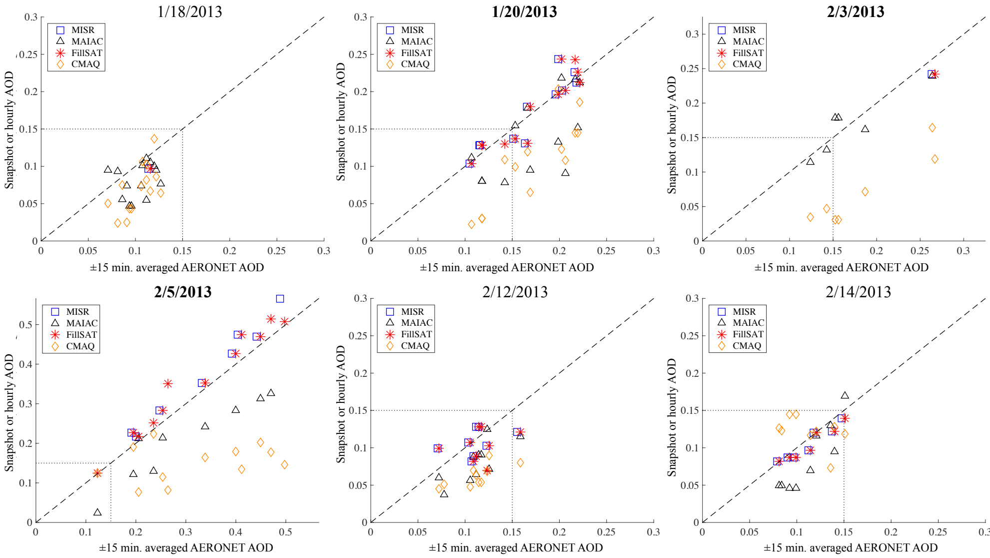


Figure 3: Scatterplot comparisons of AERONET coincidences with MISR-RA, MAIAC, gap-filled MISR-RA, and CMAQ results within ±15 minutes of Terra overpass time. The MAIAC and AERONET AOD comparisons are plotted at 550 nm, whereas the MISR-RA and AERONET AOD comparisons are at 558 nm; the dotted lines indicate the 0.15 AOD threshold; a 1:1 dashed lines are shown for reference.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: comparison is

Mariel Friberg 7/12/2018 5:13 PM
Deleted: while

Mariel Friberg 7/12/2018 5:13 PM
Deleted: comparison is

Mariel Friberg 7/12/2018 5:13 PM
Deleted: line indicates

Mariel Friberg 7/12/2018 5:13 PM
Deleted: line is

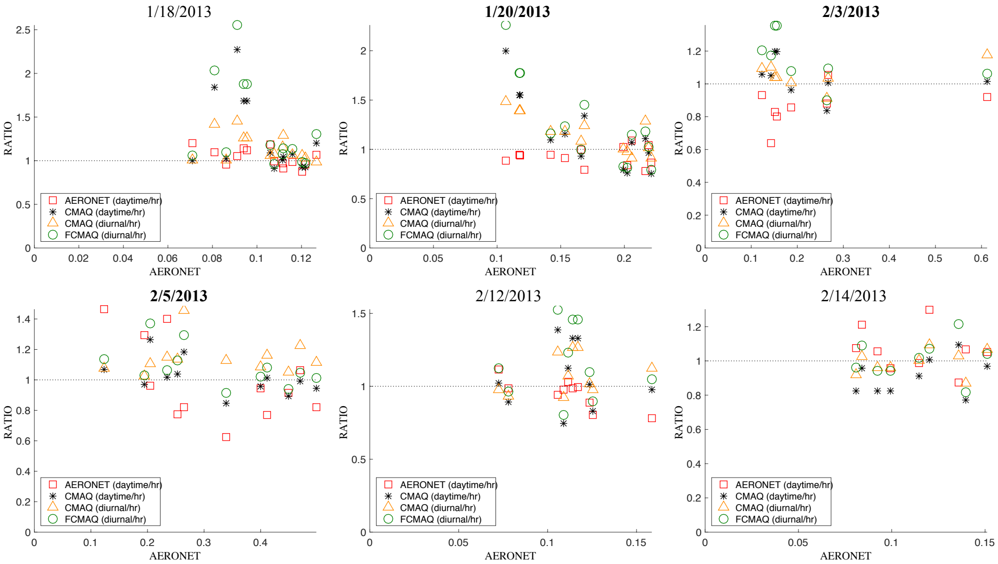


Figure 4: Scatterplots of daylight-averages to the Terra overpass time ratios versus AERONET AOD retrievals within ±15 minutes of Terra overpass time. Two ratios are shown for CMAQ: daytime average-to-hour ratio and diurnal average-to-hour ratio. The FCMAQ ratios shown are the FCMAQ diurnal to CMAQ hour values. The dashed unity lines are included for reference.

Mariel Friberg 7/12/2018 5:13 PM
Deleted: line is

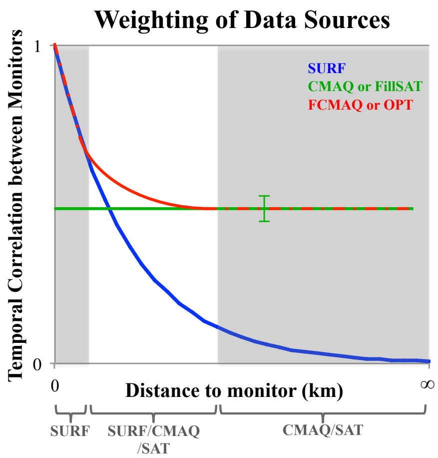
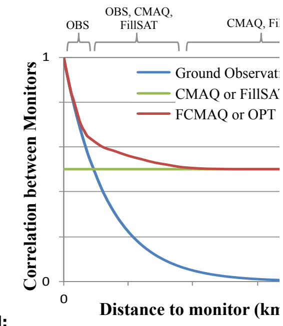


Figure 5: Theoretical plot of fused and optimized dataset weights as a function of distance from surface observations. Three regimes identify the contribution of each dataset towards improving concentration field estimates stratified by distance. As applied here, the exponential decay rate that reflects the temporal Pearson correlation between ground observations as a function of distance is species-specific. The temporal variations between ground observations and CMAQ or FillSAT are more consistent (shown as constant), independent of ground observations, and therefore, are not functions of distance to monitor. FCMAQ (surface measurements + model) and OPT (surface + satellite measurements + model) curves show how the strengths of the ground observations and other datasets are maximized using temporal correlations as each grid cell is a function of distance from a ground observation.

Mariel Friberg 7/12/2018 5:13 PM
Deleted:



Mariel Friberg 7/12/2018 5:13 PM
Deleted: The

Mariel Friberg 7/12/2018 5:13 PM
Deleted: a function

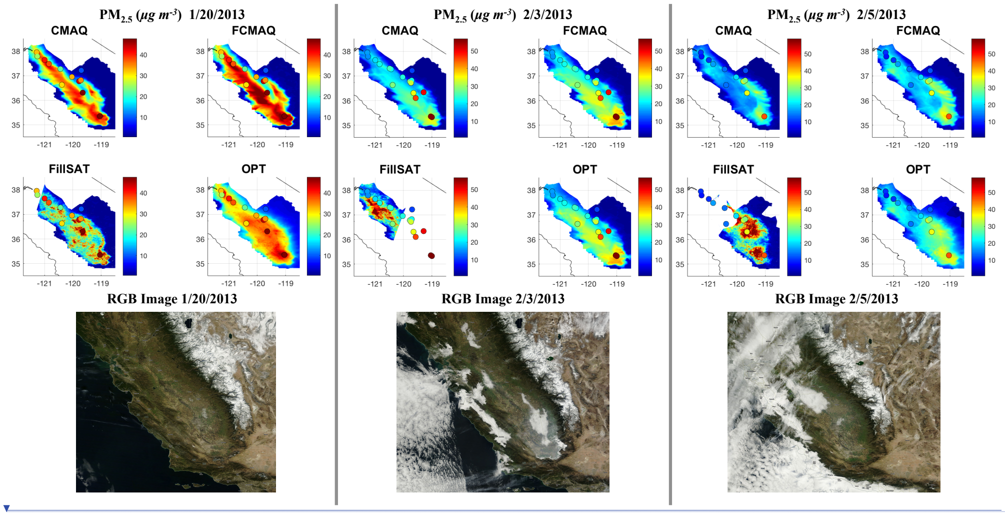
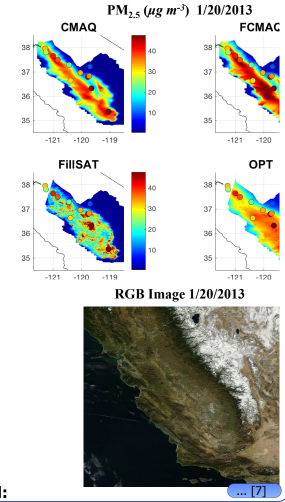


Figure 6: $PM_{2.5}$ FRM calculated concentration maps with monitor observations (filled circles) and RBG images for the three days during the study period with highest AOD. The resolution of the concentration maps is 275 m, whereas the size of the observation markers is ~ 0.1 degrees (~ 11.1 km).

Mariel Friberg 7/12/2018 5:13 PM

Deleted:



... [7]

Mariel Friberg 7/12/2018 5:13 PM

Deleted: are

Mariel Friberg 7/12/2018 5:13 PM

Deleted: are

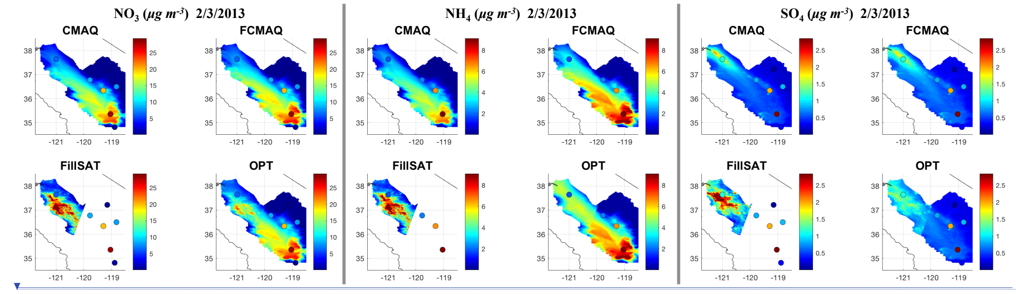


Figure 7: NH_4, SO_4, and NO_3 calculated concentration maps and monitor observations (filled circles) for February 3rd.

Mariel Friberg 7/12/2018 5:13 PM

Deleted:

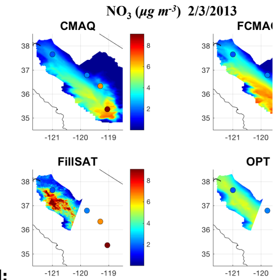


Table 2: Comparison of averaged temporal R^2, mean bias, and root means square error values between observations and leave-one-out cross-validation (LOO CV) for 52 days across all locations.

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

Table 3: Comparison of temporal R^2, mean bias, and root means square error $PM_{2.5}$ 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