Authors' reply.

We are extremely grateful for the helpful comments of both reviewers. We expect that they will find that we have addressed all comments, both general and detailed. In fact, the comments alerted us to examine and clarify several other details, for example the meaning of the intercepts and slopes as graphically analyzed, the alphas and the betas (Table 1 and Figure 8). Discussion of the comparison of the intensive Fresno measurements of Young et al. and our mapped valley-wide estimates now describe their results in some more detail.

We expect that the inclusion of the Supplementary Material will clarify how the P-3B NASA aircraft studies motivated our approach. Regretfully, the task of comparing our estimates to more detailed aircraft measurements would involve considerable additional explanation beyond the limits of a single paper.

We have attempted to improve all figures up to the standards requested. Figure 1 has a more recent and descriptive source; it now indicates the position of the San Joaquin Valley.

Satellite Mapping of PM2.5 Episodes in the Wintertime San Joaquin Valley: A "Static" Model Using Column Water Vapor

Robert B Chatfield[1], Meytar Sorek-Hamer[1,2], Robert F Esswein[1,3], and [4]Alexei Lyapustin[4]

[1]NASA Ames Research Center, Moffett Field, CA 94035, USA

[2] Universities Space Research Association, Moffett Field, CA, USA

[3]Bay Area Environmental Institute, Moffett Field, CA, USA

[4]NASA Goddard Space Flight Center, MD, USA

Abstract: The use of satellite Aerosol Optical Thickness (AOT) from imaging spectrometers has been successful in quantifying and mapping high PM2.5 (particulate matter mass < 2.5 μm diameter) episodes for pollution abatement and health studies. However, some regions have high PM2.5 but poor estimation success. The challenges in using Aerosol Optical Thickness (AOT) from imaging spectrometers to characterize PM2.5 worldwide was especially evident in the wintertime San Joaquin Valley (SJV). The SJV's attendant difficulties of high-albedo surfaces and very shallow, variable vertical mixing also occur in other significantly polluted regions around the world. We report on more accurate PM2.5 maps (where cloudiness permits) for the whole-winter period in the SJV, Nov 19, 2012–Feb 18, 2013. Intensive measurements by including NASA aircraft were made for several weeks in that winter, the DISCOVER-AQ California mission.

We found success with a relatively simple method based on calibration and checking with surface monitors and a characterization of vertical mixing, and incorporating specific understandings of the region's climatology.We estimate PM2.5 to within ~7 μg m^{-3} root-mean-square (rms) error and with *R* values of ~ 0.9, based on remotely sensed MAIAC (Multi-Angle Implementation of Atmospheric Correction) observations, and that certain further work will improve that accuracy. Mapping is at 1 km resolution. This allows a time sequence of mapped aerosols at 1 km for cloud-free days. We describe our technique as a "static estimation." Estimation procedures like this one, not dependent on well-mapped source strengths or on transport error, should help full source-driven simulations by deconstructing processes. They also provide a rapid method to create a long-term climatology.

Essential features of the technique are (a) daily calibration of the AOT to PM2.5 using available surface monitors, and (b) characterization of mixed-layer dilution using column water vapor (CWV, otherwise "precipitable water"). We noted that on multi-day timescales both water vapor and particles share near-surface sources and both fall to very low values with altitude; indeed, both are largely removed by precipitation. The existence of layers of H_2O or aerosol not within the mixed layer adds complexity, but mixed-effects statistical regression captures essential proportionality of PM2.5 and the ratio variable (AOT/CWV). Accuracy is much higher than previous statistical models, and can be extended to the whole Aqua-satellite data record. The maps and time-series we show suggest a repeated pattern for large valleys like the SJV — progressive stabilization of the mixing height after frontal passages: PM2.5 is somewhat more determined by day-by-day changes in mixing than it is by the progressive accumulation of pollutants (revealed as increasing AOT).

1. Introduction

The San Joaquin Valley (SJV) is an important agricultural area, characterized by poor air quality (Figure 1). The SJV gives an example of a region with frequent air pollution episodes, challenged by difficulties as varied particle characteristics with hard-to-quantify sources from domestic burning and spatially distinct ammonia and nitrate precursors. The 60,840 km^2 area (with approx. 4 million residents) is located southeast of San Francisco, between the Coastal Mountain Range to the west and the Sierra Nevada Range to the east (Sorek-Hamer et al., 2013). Figure 1 describes the particularly high particulate pollution characterizingthe San Joaquin Valley. Previous studies in this region reported a range of correlations between satellite-borne AOT and daily/ hourly collocated ground PM2.5 measurements in this region. Using linear tools resulted in little or no correlation (Engel-Cox et al., 2004; Ballard et al., 2008; Justice et al., 2009), while applying non-linear methods improved the correlation to R=0.71 (Sorek-Hamer et al., 2013).

More broadly, atmospheric particulate matter (PM) pollution in the respirable range, PM2.5, is recognized as a major threat to human health for some time (Brunekreef and Holgate, 2002, Dominici et al., 2006; Franklin, 2007 Kloog et al., 2013; Schwartz, et al., 1996; Zanobetti et al., 2009). Epidemiological studies have been hampered by the availability of relatively few PM2.5 measurement stations relative to the broad dispersal of populations affected. A variety of methods have been employed to estimate exposure, e.g., proximity-based using GIS, interpolation between sparse monitoring sites, land-use regression models, line- or area-dispersion plume models, 3-d atmospheric source-and-transport models, and models using information from imaging satellites, often including also land-use regression and proximity (Sorek Hamer et al., 2016). Sparse PM2.5 monitoring spatial networks may limit our ability to accurately assess human exposures to PM2.5, since concentrations measured at an outdoor site may be less representative of the subjects' exposures as the distance from the monitor increases (Bell et al., 2007; Lee et al., 2011).

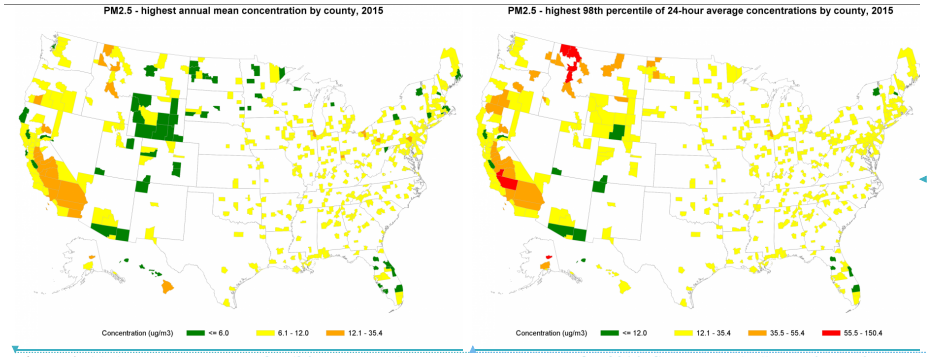
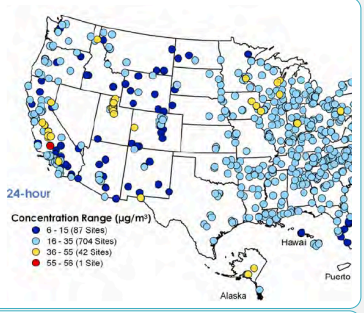


Figure 1. (a) Annual average PM2.5 (24-hr average) by county as observed for 2014: Source: EPA: "What is particle pollution and what types of particles are a health concern?" https://www.epa.gov/pmcourse/what-particle-pollution. Original description reads "U.S counties with high annual mean particle pollution concentrations in 2015. This map depicts fine particle pollution concentrations by U.S. county for 2015 based on long-term (annual) average concentrations. The map's color key is based on categories of the Air Quality Index (AQI) (see Patient Exposure and the Air Quality Index). All orange and red areas exceeded the annual ambient air quality standards for fine particle pollution during 2015 (b) 98th percentile concentrations by count for 2014 from the same source: Original description reads: "All orange and red areas exceeded the 24-hour ambient air quality standards for fine particle pollution during 2015. The map illustrates how likely it may be for a particular area to experience air quality advisories for particle



Deleted:

Formatted Table

Inserted Cells

Deleted: . Annual average PM2.5 (24-hr average) as estimated for 2010. Source: EPA. (https://www.epa.gov/outdoor-air-quality-data/interactive-map-air-quality-monitors, and RSIG, https://www.epa.gov/hesc/rsig-data-inventory)

pollution based on short-term averaging." The San Joaquin Valley comprises the area in red and the adjoining counties to the northeast and southwest; details are shown in later maps.

Formatted: Font: Times New Roman

For this reason, there has been extensive development of techniques to make best use of satellite-borne optical extinction as seen from moderate-resolution atmospheric imagers. Aerosol Optical Thickness (AOT) is typically reported as a vertical column integral of extinction above the ground footprint observed. Methods using AOT to assess exposure to PM showed early successes, but certain regions remained very poorly characterized (Engel-Cox et al., 2004, Liu et al., 2009, Gupta et al., 2006; Koelemeijer et al., 2006, Hoff and Christopher, 2009). Engel-Cox (2004) found correlations of AOT with PM2.5 for valleys along the US Pacific Coast ranged from –0.2 to +0.3, i.e., very little variance explained. MISR technology aided greatly (Liu et al., 2007), but yields mostly monthly averages over years (van Donkelaar et al., 2010), limiting event and epidemiological analysis.

AOT may be strongly affected by particles encountered well above the planetary boundary layer and different particulate composition. In addition, cloud cover severely limits the actual spatial coverage of AOT (Ford and Heald, 2016). Yet, in spite of these limitations (Jin et al, 2019), AOT has been employed extensively for assessing PM concentrations (e.g. Liu et al., 2018, Franklin et al., 2017; Van Donkelaar et al., 2015, 2016; Kloog et al., 2015, 2014; Hu et al., 2014; Sorek-Hamer et al., 2013; Hoff and Christopher 2009).

In regard to the SJV, considerable work has been published, since it was the site of two major intensive studies, CRPAQS (California Regional PM10/PM2.5 Air Quality Study, Chow et al,2006) and DISCOVER-AQ California (Deriving Information on Surface Conditions from COlumn and VERtically Resolved Observations Relevant to Air Quality, https://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html, more references below and at web site). There was a very useful analysis of particle composition for a well-instrumented Fresno surface site for this period (Young et al., 2016). This study added detail to the Watson and Chow (2002) analysis of an earlier intensive study of the area; in particular, the striking dominance of nitrate and organic aerosols in a regular diel pattern. Watson and Chow reference several publications describing that intensive. Johnson et al, (2014) made a three-dimensional modelling study of methane emissions that also helps describe the mixed-layer of the specific DISCOVER-AQ period). Lidar gives a very helpful view of complexities of submicron particle abundance and properties within the mixed layer and the uniformity of the mixed layer top (Sawamura et al., 2017).

Application of modelling with satellite AOT columns from different satellite platforms for the DISCOVER-AQ (included within our study period) was able to achieve R^2~0.8). These results were achieved for just the DISCOVER-AQ period of ~6 weeks and with separate sub-regions of the Central SJV. They highlight the complexity of composition and source-driven simulation (Friberg et al, 2018). The Friberg publication is highly recommended as a comparison to this effort, and has extensive references regarding the SJV and the details required for source-driven modelling.

There are several related goals in producing PM2.5 maps and assessing their accuracy. The Friberg et al., work primarily aimed to constrain CMAQ downwind of the surface air quality stations, and in particular, to constrain particle type as much as possible, along with concentration, using MISR constraints (Ralph Kahn, personal communication 2019). Our goal was to produce a large set of maps characterizing one winter in a particular setting,

inland Mediterranean valleys, with the aim of allowing air pollution professionals to understand particulate episodes and to improve sources and simulation details (e.g., transport error) for source-driven models. Goals of the Dalhousie group are to improve annual average exposure: they see that as the principal driver for health effect (Van Dankelaar et al., 2010, 2015, 2016). A main goal of NASA's MAIA (Multi-Angle Imager for Aerosols) mission is similarly deliver new data for a each-day mapping of PM2.5 exposure sufficient for full studies of health effects (Diner et al., 2018, https://maia.jpl.nasa.gov/). In pursuit of that goal for the MODIS Aqua dataset, we will indicate some preliminary, meteorology-based ideas for estimating high aerosol concentration when clouds prevent the use of remote sensing data.

Due to the complex meteorology of the San Joaquin flows and uncertainties surrounding the sources of ammonia, nitrogen oxides, and residential-burning smoke, we attempt to separate out some certain aspects of complex 3-d source-driven modeling (Bey et al., 2001, Nolte, 2015, Appel et al., 2017, Friberg et al. 2018) with a "static model" which does not attempt to simulate transport, but rather uses observational records related to vertical mixing and AOT. The spatial maps produced can give a more detailed check on the 3-d process modeling. They also allow the whole MODIS record of winters to be analyzed efficiently so as to reveal patterns and trends. We emphasize this and further extension the interpretation of satellite radiances, attempting to remain close to physical interpretations by using both MAIAC AOT and CWV retrievals. MAIAC Column Water Vapor (CWV) (Lyapustin et al., 2018) retrievals have been quite acceptability validated with the AERONET CWV measurements in higher CWV environments (Martins et al., 2017, 2018). It has not been previously recognized as a tool for improving ground PM estimation and in particular, in the SJV.

Deleted: , and references

1.1 Data

MAIAC AOT and CWV

The Multi-Angle Implementation of Atmospheric Correction (MAIAC) is an operational algorithm developed for MODIS Collection 6 (C6) data (Lyapustin et al., 2011a,b). This algorithm applies a dynamical time series technique to derive the MODIS surface bidirectional reflectance factor and atmospheric retrievals at a 1 km resolution, such as AOT, and CWV (Lyapustin et al., 2008; 2011b). MAIAC AOT retrievals present an expected error within 15% and relatively good correlation coefficient (R) with AERONET measurements in the study area (Lyapustin et al., 2018).

MAIAC data has been used from both Terra and Aqua satellite with a daily overpass at ~10:30 and ~13:30 local sun time (+ ca.1.5 hours), respectively. Data has been obtained for the period of Winter 2012-2013 (November 2012-April 2013). We surveyed this entire period and included, for estimation, all wintertime high-PM2.5 episodes for this specific winter, selecting November 19–February 18, as described in several later figures (discussed in context: Figures 3, 4, and 7).

Deleted: UTC,

AERONET AOT and CWV

AERONET (AErosol RObotic NETwork) is a global network of automatic sun-and-sky radiometers for aerosol monitoring (Holben et al., 1998). Direct sun measurements are used to compute the AOT values at seven wavelengths (340, 380, 440, 500, 675, 870, 1020 nm), while CWV retrievals are derived from the channel 940 nm (Schmid et al., 2006). The AERONET data were obtained for the study period with cloud screened and quality-assured at V3 Level 2 products. The AERONET AOT values were interpolated to a 550 nm using quadratic fits on a log–log scale. Details on instruments and monitoring sites of the DISCOVER-AQ campaign are available at: http://www.nasa.gov/mission_pages/discover-aq/instruments/index.html. Archived DISCOVER-AQ data are available at the NASA LaRC Science Data for Atmospheric Composition website: http://www-air.larc.nasa.gov/index.html.

Ground PM2.5 Concentrations

Hourly ground PM2.5 concentrations were obtained from the USA Environmental Protection Agency (EPA) at +-60 minutes from the satellite overpass. Data were obtained from stations that reported non-negative PM values over the whole study period. (https://aqs.epa.gov/aqsweb/airdata/download_files.html#Raw)

PBL

Momentum-based PBL depth, 10-m wind, and some CWV quantities for the model were taken from the archive of the NOAA Rapid Refresh (RAP) model available for this period. (Choice of MAIAC or RAP estimates is discussed later. The model archive had a nominal 13 km resolution resolved at a one-hour time interval, so that model quantities could be matched closely to the satellite overpass times. Unreported examination of the AERONET data for the period suggested that the temporal resolution of the MAIAC AOT was quite accurate. The HSRL2 aerosol data as described by Sawamura et al., (2017) suggested that depths of afternoon mixing tops were adequately described by a 13 km resolution model, as did adjacent spirals of the NASA P3-B aircraft as described by Michael Shook (Shook et al., 2013, See also Supplementary Material.). AOT could however vary on relatively short distance scales, e.g. within 0-2 km of roadways when winds were parallel to the road. We shall see the consequential variations in estimated PM2.5 later in the processed results.

Deleted: Update Cycle

Deleted: and

Deleted: temporal resolution of

Deleted:

Deleted:

Deleted:) .

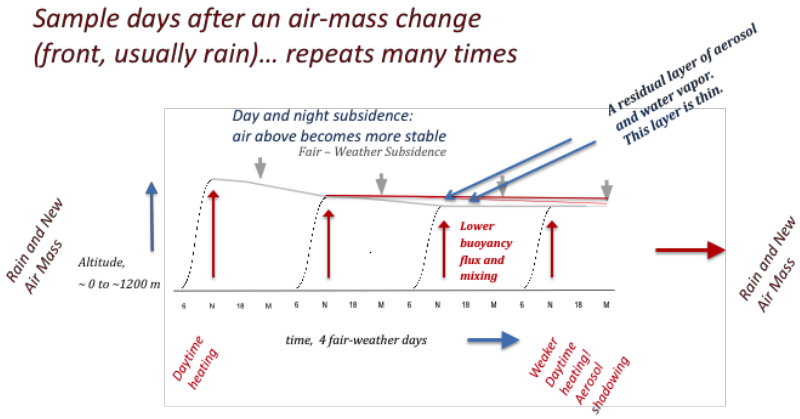
2. Motivating Meteorological Perspective

Koelenmeijer et al, 2006 give a succinct description of the relationship between AOT and dry particle mass. We adopt their simplification describe the relationship of AOT to PM2.5 using a simple equation where all particles are idealized as evenly mixed throughout a layer mixing to sensors near the ground, and the thickness of the mixed layer is Δz_{ML}

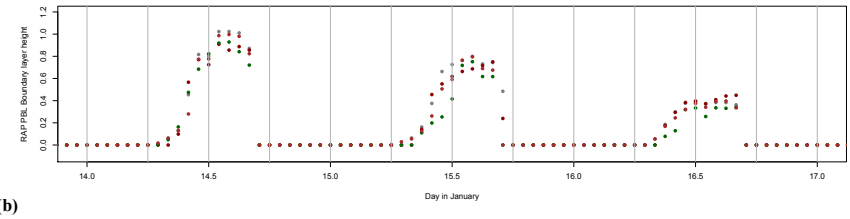
$$\mathrm{PM2.5} = f(\mathrm{AOT}) = \frac{\mathrm{AOT}}{\Delta z_{ML} \cdot M(Composition, RH)} \qquad \text{Eq. 1}$$

The factor in the denominator, M (for "magnification") describes the relationship of the optical extinction to "dry particle mass smaller than 2.5 µm aerodynamic diameter" which is the motivated definition of PM2.5. (PM2.5 also has a definition by a United States "Federal Reference Method" which is formulated to approximate the physical definition as closely as possible.) The factor *M* then is a function of particle composition and the extinction coefficients b_{Ext} associated with the components, one of which may be largely absorbed water. Particle composition and ambient relative humidity, RH, then interact with each other to determine the water content. It is significant that RH is a function of temperature and therefore altitude, with highest RH at the top of a well-mixed layer.

This work emphasizes and attempts to exploit features of regional aerosol haze palls that parallel features of aerosol mass and a different measure of water vapor.



(a)



(b)

Figure 2. (a) Conceptual figure describing the fair-weather PBL top for successive days in a clear-weather PM2.5 episode motivating this study. See text. (b) Simulation in the RAP model of planetary boundary layer height for momentum at three SJV PM2.5 stations (red: Tranquility, gray: Hanford, green: Bakersfield) Periods from 11 AM to 3:30 PM approximate the mixed layer for that period and time following, although advection may change the concentrations mixing to that height. Maximum PBL-top altitude may not be accurate for the station, but shape of diel profile is appropriate.

Deleted: .

Figure 2 illustrates a conceptual idea of the fair-weather simulation we focus on. Both regional particulate pollution and water vapor originate from the Earth's surface. Each tends to create relatively well mixed layers over several days, transported most significantly by a repeated daily cycle of mixing. The mixing of momentum is most active from just before noon through the mid-afternoon, creating an afternoon mixed layer, and water vapor and aerosol most typically mixes well up to this layer. Turbulent mixing depths vary from day to day, and these can create lofted layers of pollution cut off from the surface on the day of AOT and CWV observation. Flows in the San Joaquin can be greatly influenced by the nearby mountains, with flows day and night promoting some upslope transport of material which can recirculate, detached from mixing on following days. Consideration of subsidence of air into the San Joaquin mixed layer suggests a flow-through time for aerosol and water of 2–3 days for some situations (Caputi et al., 2018). Mixing of entrained and mixed layer air allows for continued accumulation of pollutant aerosol in the valley as Figure 2 shows.

Particles and water vapor are emitted and accumulate in the same region, and they are mixed similarly each midday and afternoon by convective stirring. The height of mixing can be determined by variations in the buoyancy flux from the surface and varying vertical subsidence velocities, responding to larger scale weather patterns, during successive days. Figure 2 does not show the effect of particle transport or water vapor transport for a specific location, but the PBL top, which is strongly controlled by local heat and water vapor fluxes at the surface.

If the mixing height is lower on succeeding days, then any water vapor and any particles at the top of the mixed layer are trapped in an “elevated layer” which does not mix to the surface. Other common ways in which elevated layers can be formed are mixing along the side of the valley (small -scale anabatic and katabatic winds) and by differential transport, i.e., wind shear. Fires, power plant plumes, and long-distance synoptic transport can form layers that are quite separated at higher altitudes in the troposphere. Eventually there is removal of both species. Wet removal of particles is particularly effective, and the specific humidity of the air is very effectively removed by the condensation accompanying cooling and rising, according to the Clausius Clapeyron equation. Similar processes then limit the vertical spread of particles and specific humidity.

3. Expected Variability of the AOT-PM2.5 relationship

Water vapor molecules also accumulate in the atmosphere over a period of several days (typically a somewhat longer period), and both aerosols and water vapor are cleared from a particular place by cloud removal processes (venting, rainfall) and by airmass replacement. In the case of high pressure systems in which air pollution episodes occur, such replacement is a common feature. If the other variables are available by measurement, e.g., airplane measurement such as in DISCOVER-AQ (https://discover-aq.larc.nasa.gov/data.html), Equation 1 can be solved for Δz_{ML}, defining an equivalent mixing height for particles. Similarly, we can write equivalent mixing depth of water vapor, $\Delta z_{e\,H2O}$:

$$\Delta z_{e\,H2O} = \int_0^{Top} \rho_{H2O}(z)\, dz \Big/ \bar{\rho}_{H2O}(ML, RTP) = \text{CWV} / \rho_{H2O}(ML, STP) \qquad \text{Eq. 2}$$

where CWV is in g / cm2 , $\rho_{H2O}(z)$ and $\bar{\rho}_{H2O}(ML, RTP)$ correspond to the vertically distributed water vapor and appropriately average water density of the mixed lay. CWV is available from the MAIAC analyses yielding AOT. Making the assumption that the heights are the nearly equivalent for water vapor and aerosol, we may write

$$\text{PM2.5} = f(AOT) = \frac{\text{AOT}}{\text{CWV}} \frac{\bar{\rho}_{H2O}(ML, RTP)}{M(Composition, RH)}. \quad \text{Eq. 3}$$

PM2.5 is calculated at EPA reference temperature (25 C) and pressure (1 atm), water vapor quantities in g cm–3 and $\Delta z_{e\ H2O}$ is in cm.

Work reported by Shook et al., 2018, described the vertical distribution of trace species with a vertical coordinate normalized to his estimated afternoon mixed layer top, This suggested to us that water vapor had vertical distributions that were usefully similar. The decline of water vapor was not as sharp, often showing a rapid decrease; the drop in scattering was dramatically rapid,

We found in ensuing work that approximating $\bar{\rho}_{H2O}(ML, RTP)$ by $\rho_{H2O}(z = 0, current\ conditions)$ added only a small amount to the variance explained by the regression given other limitations of the approximation. (Possibly relative humidity effects or the correlation of water density with temperature could be complicating correlated factors.)

We calibrate the relationship f(AOT) using data at official PM2.5 stations, and make the calibration daily. It is our observation that f varies only over a small range when there are several MODIS observations on the same day, and that it varies in a limited way between neighbouring stations in a local region. The definition of "region" is based on that similarity, and it suggests similarity ofΔz_{ML} and M, i.e., similar aerosol characteristics and boundary layer behavior. This similarity does not apply when the wind shifts greatly between times or between stations, e.g. when a front passes. Fortunately for our understanding of pollution episodes, frontal passage days tend not to have high PM2.5 .

We distil these understandings when we formulate a regression equation

$$\text{PM2.5}_{i\,s} = (a + \beta_i)\,(\text{AOT}_{is}/\text{CWV}_{is}) + \alpha_i + \varepsilon_{is} \quad \text{Eq. 4}$$

where the subscripts i describe "instance" or calendar date, and the subscripts s describe "station," so that AOT and PM2.5 form a two-dimensional table.

Given the independent nature of i and s, the regression must be solved by "mixed effects" methods described below. The subscript s need only be independent of i, so later we will use it to denote "situation" or the hour of the day when there are many observations made at one station on one day i. It is not assumed that the consecutive order of the day observations necessarily describe any continuity in i. Observations show that there is often continuity, but that the continuity is quickly broken when frontal passages or rain affect the region.

Writing Equation 3 in the form used for mixed-effects models, we separate a general term from the terms that depend on i or calendar date.

$$\text{PM2.5}_{\text{i s}} = a \cdot \text{AOT}_{is}/\text{CWV}_{is} + c + (\alpha_i + \beta_i \cdot \text{AOT}_{is}/\text{CWV}_{is}) + \varepsilon_{is} \quad \text{Eq. 5}$$

A commonly used shorthand is the Wilkinson and Rogers (1973) form, accepted by many software packages,

$$\text{PM2.5} \sim \text{AOT/CWV} + (\text{AOT/CWV} + 1 \mid \text{DOY}) \quad \text{Eq. 6}$$

Moved down [1]: This basic understanding does not fully explain the success of the mixed effects model that we observed for the San Joaquin Valley. Furthermore, analyses of the Baltimore-Washington area not described here suggest that it works more broadly. Both aerosol and especially water vapor often exhibit layers not in continual contact with surface monitors. These we will call "elevated layers." In-situ measurements on aircraft and also lidar measurements from ground lidars looking downward from aircraft (Sawamura et al., 2017), and satellite lidar (CALIPSO) reveal aerosol layers with significant optical thickness above the mixing layer. Similarly, airborne measurements in the DISCOVER-AQ intensive measurements of 2013 suggest a fraction of water vapor lies above the mixed layer for water. Allow these portions of total AOT and CWV layers to be quantified as AOTe and CWVe ('e' stands for elevated). There can be several individual layers. AOTe and CWVe refer to the total amounts of extinction and water vapor mass.

Deleted: Thus there is an approximate equation upon which to base regression estimation:¶
¶
PM2.5 = $\frac{(\text{AOT}-\text{AOTe})}{(\text{CWV}-\text{CWVe})} \frac{\bar{\rho}_{H2O}(ML,RTP)}{M(Composition,RH)}$ ... [1]

Deleted: neighboring

Deleted: distill

Deleted: 5

Formatted: Font: Italic

Deleted: "

Deleted: "

Formatted: Font: Italic

Deleted: "

Deleted: ,"

Moved down [2]: While elevated layers of water and aerosol are common, we will see that it appears that this regression equation allows rather good fits. This can happen when AOTe << AOT and CWVe << CWV to a sufficient degree, or else when there are approximate linear (slope + intercept) relationships obtaining between the numerator and denominator of Equation

Deleted: 4.

Moved down [3]: Essentially the terms are absorbed into constant parameters for the day, α_i and β_i, along with other parameters like M. AOTe and CWVe are considered to be essentially constant over the region. In fact, this degree of constancy can be taken to define the "region" of application.

Moved down [4]: We may these terms into constants α_i and β_i works under an implicit assumption of uniformity in

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Deleted: f

Deleted: 6

Deleted: 7

where DOY describes the calendar date subscript i. This formalism also describes the columns of the regression matrix to be solved.

It is tempting to generalize this relationship to recognize that there is often correlated behavior between stations, but with some constant offset

$$\mathrm{PM2.5}_{i\,s} = a \cdot \mathrm{AOT}_{is}/\mathrm{CWV}_{is} + c + (\alpha_i + \beta_i \cdot \mathrm{AOT}_{is}/\mathrm{CWV}_{is}) + \gamma_s + \varepsilon_{is} \qquad \text{Eq. 7}$$

However, if one allows such variations at monitoring stations, it can be difficult to decide what values of γ_s to use between stations. This is an attempt to describe "sub-regionality," that is, similar behaviour within a region modified by slight and geographically coherent variations which allow spatial interpolation.

For those not familiar with mixed-effects models, we mention that the procedure is similar to the use of dummy variables, where coefficients u_i multiply a set of discriminating variables, equal to 1 when i takes on the value of a particular instance/day, and 0 for all other instances. The mixed-effects techniques similarly solves a much larger regression equation, but has better theoretical development. Note that the number of observations is N_i times N_s, while the number of parameters is linear in N_i and N_s, where the N's signify the number of each. When N_i and $N_s > 5$, the problem becomes increasingly over-determined.

This basic understanding does not fully explain the success of the mixed effects model that we observed for the San Joaquin Valley. Furthermore, analyses of the Baltimore-Washington area not described here suggest that it works more broadly. Both aerosol and especially water vapor often exhibit layers not in continual contact with surface monitors. These we will call "elevated layers." In-situ measurements on aircraft and also lidar measurements from ground lidars looking downward from aircraft (Sawamura et al., 2017), and satellite lidar (CALIPSO) reveal aerosol layers with significant optical thickness above the mixing layer. Similarly, airborne measurements in the DISCOVER-AQ intensive measurements of 2013 suggest a fraction of water vapor lies above the mixed layer for water. Allow these portions of total AOT and CWV layers to be quantified as AOTe and CWVe ('e' stands for elevated). There can be several individual layers. AOTe and CWVe refer to the total amounts of extinction and water vapor mass. Thus, there is an approximate equation upon which to base regression estimation:

$$\mathrm{PM2.5} = \frac{(\mathrm{AOT} - \mathrm{AOTe})}{(\mathrm{CWV} - \mathrm{CWVe})} \frac{\bar{\rho}_{H2O}(ML, RTP)}{M(Composition, RH)} \qquad \text{Eq. 8}$$

While elevated layers of water and aerosol are common, we will see that it appears that this regression equation allows rather good fits. This can happen when AOTe << AOT and CWVe << CWV to a sufficient degree, or else when there are approximate linear (slope + intercept) relationships obtaining between the numerator and denominator of Equation 8. Essentially the terms are absorbed into constant parameters for the day, α_i and β_i, along with other parameters like M. AOTe and CWVe are considered to be essentially constant over the region. In fact, this degree of constancy can be taken to define the "region" of application. We may these terms into constants α_i and β_i works under an implicit assumption of uniformity in AOTe and CWVe throughout the region, or at least a uniform linear dependence with AOT and CWV.

Formatted: Font: Italic

Deleted: 8

Deleted: behavior

Deleted: Ni

Deleted: Ns,

Deleted: Ni

Deleted: Ns,

Formatted: Font: Italic

Deleted: Ni

Deleted: Ns >

Moved (insertion) [1]

Moved (insertion) [2]

Formatted: Indent: First line: 0"

Moved (insertion) [3]

Moved (insertion) [4]

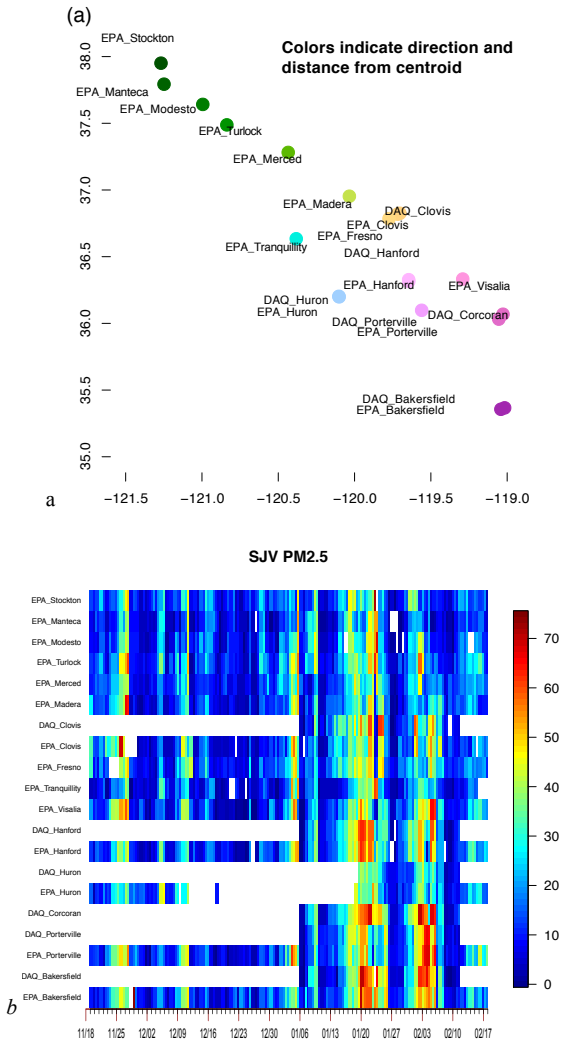


Figure 3. (a) Locations of stations in the SJV used; color coding allows nearby stations to be identified. (b) Matrix plot of PM2.5 at the sites for the period November 19, 2012 through February 18, 2013. (b) Another view that summarizes the variability of observed PM2.5 is shown in Figure 4(a).

4. Observations and an Overview of Pollution Episode Trends

In this section, we will show how components of a mixed effects model that utilize CWV contribute to its explanatory power. We examine the relationships and predictive ability for PM2.5 observed at SJV measurement stations for the winter season encompassing high-pollution periods, Nov 19, 2012 to February 17, 2013. Stations from Bakersfield in the South to Stockton in the North were included. Figure 3(b) shows several episodes affecting most of the Valley; one period with more stations reporting includes the DISCOVER-AQ period. This period has additional P3-B aircraft data which motivated this work, but are too lengthy to describe in this publication.

Figure 3(a) shows the locations of all stations used in this work; the stations include much of the Valley from Stockton to Bakersfield. Some of the stations labelled DAQ were in operation only during the DISCOVER-AQ California period. A color wheel was used to assign colors to the stations on the graph; this allows identification of stations' latitude, longitude, and proximity in later graphs comparing observations and our fitted values. Figure 3(b) describes the rise and fall of PM2.5 pollution using the station reports. The rows represent stations and are arranged north to south. Several major episodes are immediately seen, as well as differences in their intensity and timing of development. The DISCOVER-AQ observations were limited to the period shown, January 8th through February 10th, 2-13. Differences between the PM2.5 values observed at nearby stations, one DISCOVER-AQ, one California Air Resources Board (labelled "EPA" for the dataset origin) give an impression of local variability; differences between observations at Clovis are quite apparent.

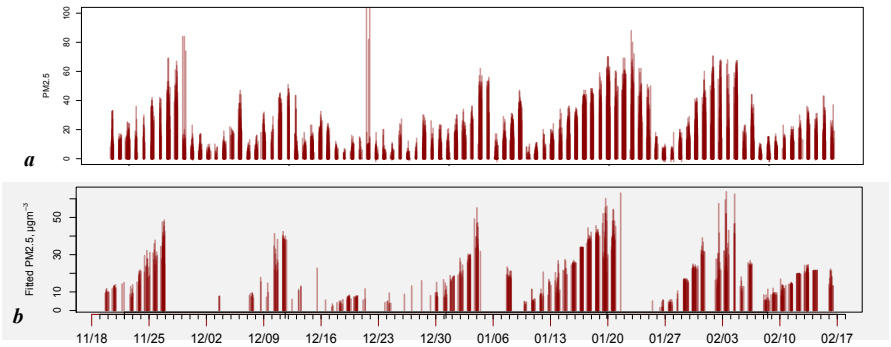


Figure 4. (a) PM2.5 as observed at all stations for the winter period extending from November 2012 to March 2013. The graph has vertical bars drawn with partial transparency, so that careful inspection of a single day describes all the observations in the Valley for the day. The observations contributing for each day may be seen inf Figure 3(b). (b) PM2.5 as fitted by the regression with slopes and intercepts, described further below.

Deleted: The figure

Deleted:

Table 1. Comparison of results using different terms in the mixed effects model

Wilkinson Shorthand	Terms	*R*	Remaining RMS Error	Panel Fig. 5
~ AOT	$PM2.5_{is} = a \cdot AOT_{is} + c + \varepsilon_{is}$	*0.40*	14.	(a)
~ AOT/CWV	$PM2.5_{is} = a \cdot AOT_{is}/CWV_{is} + c + \varepsilon_{is}$	*0.48*	13.	(b)
~ (1 | DOY)	$PM2.5_{is} = \alpha_i + \gamma_s + c + \varepsilon_{is}$	*0.78*	10.	(c)
~ AOT/CWV + (AOT/CWV – 1 | DOY)	$PM2.5_{is} = a \cdot AOT_{is}/CWV_{is} + c + \beta_i \cdot AOT_{is}/CWV_{is} + \varepsilon_{is}$	*0.88*	7.44	
~ AOT/CWV + (1 | DOY)	$PM2.5_{is} = a \cdot AOT_{is}/CWV_{is} + c + \alpha_i + \gamma_s + \varepsilon_{is}$	*0.85*	8.03	(d)
~ AOT/CWV + (AOT/CWV +1 | DOY)	$PM2.5_{is} = a \cdot AOT_{is}/CWV_{is} + c + (\alpha_i + \beta_i \cdot AOT_{is}/CWV_{is})$	*0.90*	6.72	(e)
~ AOT/CWV + (AOT/CWV +1 | DOY) + longitude	$PM2.5_{is} = a \cdot AOT_{is}/CWV_{is} + c + (\alpha_i + \beta_i \cdot AOT_{is}/CWV_{is}) + \gamma_s + \varepsilon_{is}$	*0.91*	6.44	(f)

Notes: (a) variables are described in the context of Equations 4–8 in the text.
(b) In all regressions with random effects (all but the first two regressions), the inclusion of the a and c variables suggests an over-fitting. Mixed effects convention emphasizes these "main effects" separately and therefore specifies there must be a single linear constraint on the terms such as a and α_i ; also, c and β_i. Importantly, in Section 6 and certain figures below, we describe the (more intuitive) combination of main and random effects, e.g. we graph $\alpha_i \leftarrow \alpha_i + a$ and $\beta_i \leftarrow \beta_i + c$.

Formatted: Centered
Formatted Table
Formatted: Font: Bold
Formatted: Font: Bold
Formatted: Font: Bold
Inserted Cells

5. Which Information Contributes to PM2.5 Maps

Using the MAIAC 1x1 km estimates for each day for the location of each aerosol monitoring station and the PM2.5 measured at overpass time for that day, we may solve the estimation equation, Eq. 6. The complete simulation of PM2.5 measurement at all stations where MAIAC data allowed is shown in Figure 4(b). The technique can be used for all years and the whole area of the SJV where MODIS data is available. We used the complete model as described in Equation 6, "slopes and "intercepts", but without any time-independent spatial variation allowed (γ_s). Three features deserve immediate comment. First, there are patterns of gradual increase of PM2.5 up to 45–80 µg m^{-3} followed by relatively sudden decrease to levels near 5 µg m^{-3}. Second, the regression technique using AOT / CWV, as estimated individually for each day, captures the variation rather well, for all days where estimates can be made. Individual, exotic high values are not captured. Third, there is a pattern where the end of an air pollution episode, showing very high values, is not captured by the technique. These are simply days where MODIS observations were not available, almost always due to cloud cover. We expect that these are readily explained in terms of weather phenomena especially typical of the American West during wintertime. Pollution episodes are ended with the approach of warm fronts with high clouds, followed in a few days by the cleansing effects of rain, air mass replacement, and higher wind. We will return to this topic later.

To understand what information is used by the technique and how important is that information, based on the series of regression estimates we present; we argue that there is a cumulative aspect to explanation. For example, when we include one statistical variable, e.g. α_i, describing variation by day but constant for all stations of the day, then the regression with AOT/CWV becomes much more informative. A general relation describing the slope of PM2.5 with AOT/CWV becomes more useful when an appropriate intercept (offset) is provided.

Consider both Table 1 and Figure 5 as they describe cumulative effects of adding information. First, we note that AOT alone is not very informative about PM2.5. This would seem to follow naturally from Equation 1, since variations in mixing depth and composition are not considered. Figure 5a show many station observations with high PM2.5 but low AOT, and vice versa. Slight but significant improvement is made when column water vapor, CWV, is introduced to provide some information on mixing depth and dilution. *R* improves to 0.48 but the remaining error is nearly as high. Still, some linear relationship begins to show for perhaps 60% of the data.

A side comment regarding significance: *R* and remaining RMS error (in µg m–3) are shown in Table 1. We also performed two other tests not tabulated. An analysis of the Kuhlbach-Liebler divergence (Hastie et al., 2009), where possible, suggested each successive test in the table clearly adds information regarding PM2.5. The number of observations justified the number of additional parameters. While the numerical values are difficult to compare to other examples of regression, they show similar trends as *R* and RMS error, i.e. accuracy becomes increasingly hard to improve as *R* increases. Another test was leave-out-one cross validation (Hastie et al., 2009). Each individual station was omitted, and the regression based on the remaining stations was tested against observations at that station. The cross-validated mean squared error was about 7.8 µg m^{-3} at most for the most informative regressions shown.

Deleted:

Deleted: γ_s

Now consider a popular alternative to the use of satellite data. The third regression shown estimates of satellite data to particulate estimation, this has been shown to surpass, or at least approximate the only α_i, i.e, assign a single PM2.5 estimate for each station based only on the individual day. Color-coded maps of PM2.5 drawn for a region have a single color which varies from day to day. In many applications of satellite data to particulate estimation, this has been shown to surpass, or at least approximate the results of use of AOT (Sorek-Hamer et al., 2017). $R \sim 0.78$, RMS error ~ 10 µg m–3. Its success emphasizes the regional similarity of conditions defining PM2.5 concentrations, and their extensive spatial correlation. An explanation is that respirable PM2.5 is defined by daily weather and orientation to major sources.ß

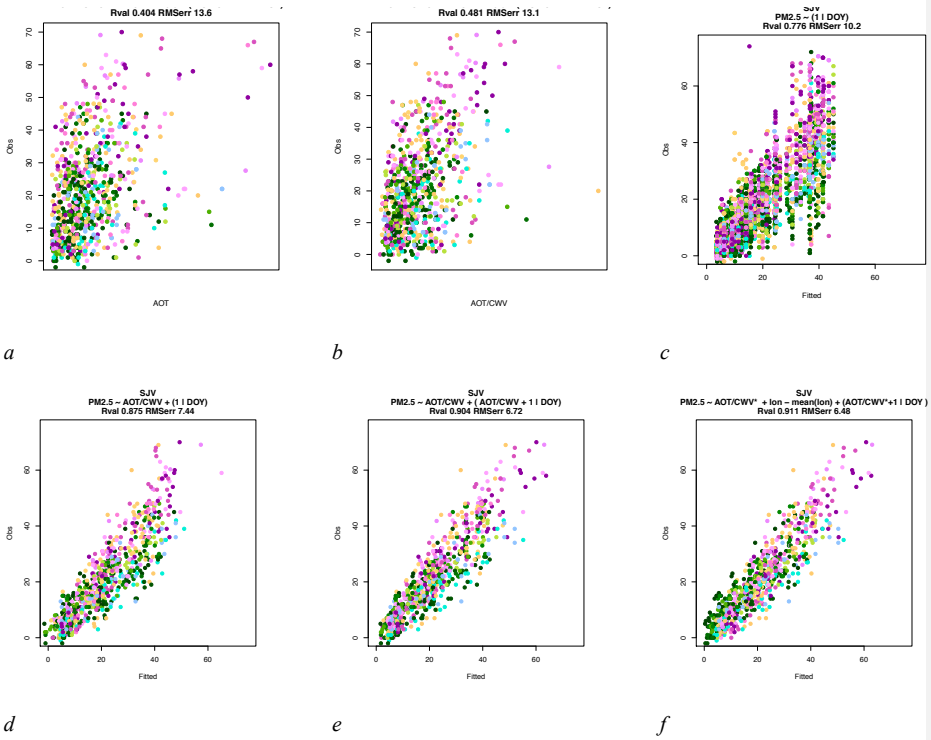


Figure 5. Progressive improvement of PM2.5 simulation showing the roles of daily calibration and AOT/CWV descriptions of aerosol vertical dispersion. Station observations (µg m^{-3}) are shown on the y axis, estimators on the horizontal. Note the progressive refinement of *R* and remaining rms error. See text. (a) Use of AOT only, an early methodology. (b) Some improvement using AOT/CWV but no daily calibration. (c) More improvement with daily calibration (mixed effects using intercepts α_i) (d) Clearly improved linearity when combing intercepts with AOT/CWV (e) Estimating daily "random" intercepts and slopes improves RMS error and R. (f) A simple description of variation within the region (longitude) aids the estimation slightly (RMS error ~ 6.48 µg m^{-3}, $R \sim 0.91$)

Once the regional similarity of pollutant conditions is recognized, it becomes appealing to combine information. The fourth estimate, Figure 5(d), does just this and shows a notable increase in R, 0.88, and decrease in

Deleted: fifth

Deleted: shown below

RMS error, 8.03. This is an approximately 50% decrease in error variance. In our situation, satellite data looks to be useful. The scatterplot of Figure 5(d) suggests distinctly more linear behavior.

An appealing alternative is to estimate only slope variations, β_i. This is nearly as useful as estimating just α_i, R ~ 0.85 RMS error ~ 10 μg m^{-3}. Each is useful. Do the two parameter estimations give distinct information?

Estimation of varying offsets β_i and sensitivities α_i does indeed help, reducing the variance by another 10%. Combining the use of AOT, CWV, and individual daily intercepts and slopes yields R ~ 0.90 and RMS error ~ 6.72 μg m^{-3}. Nevertheless, Figure 5(e) shows that certain stations have persistent deviations from the general swarm of points, Tranquility (pale green) is predicted high and Porterville and neighbors (red), are predicted low.

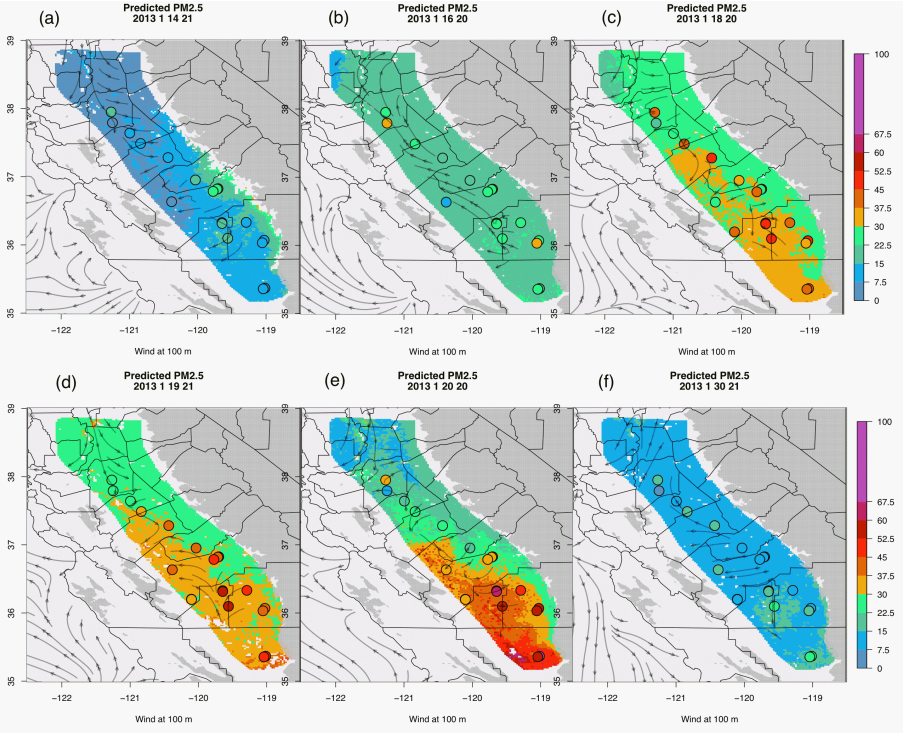


Figure 6. Estimated surface PM2.5 at 1 km indicated overpass times for the first wintertime episode in the San Joaquin Valley. Winds at 360 m agl are also shown. Estimated RMS error is 7 μg m^{-3} with a similar limit of detection. Filled circles show station PM2.5 . In this episode, the E-W correction based on the full dataset appears inappropriate, lowering mapped estimates in the east Valley. Error should decrease with improved understanding of geographic variability. Time stamp at top of image describes date and time in UTC format.

This analysis of residuals suggests that there may be spatial variations that can be specified for our stations, γ_s , but are general enough that they can be extended to maps. For this publication, we attempted a very simple variation, an east-west variation (longitude). This did improve the scatterplot for most stations, especially when considering values above ~10 μg m^{-3}. RMS error decreased slightly to 6.48, and the R estimate also rose slightly, to 0.911.

Deleted: 4c

Formatted: Superscript

Deleted: 4e

Formatted: Indent: First line: 0"

These changes are close to the range of sample variability. The maps shown in Figure 6 also show more convincing (subjective) agreement in magnitude and pattern. Nevertheless, many of the highest observations are underestimated by about 20%.

We used CWV rather than the RAP planetary boundary layer height for momentum, 11 to 15:30 local time. This was available in the 2012-2013 winter at times within a half hour of overpass time, However, this PBL height is not always recorded in the high-resolution RAP archive. We compared a regression very similar to the most detailed regression of Table 1, but using this PBL height. The formula used was

$$\mathrm{PM2.5}_{i\,s} = a \cdot \mathrm{AOT}_{is}/\mathrm{PBL}_{is} + c + (\alpha_i + \beta_i \cdot \mathrm{AOT}_{is}/\mathrm{PBL}_{is}) + \gamma_s + \varepsilon_{is} \qquad \text{Eq. 8}$$

With this, the R value was 0.917 and the RMS error was 6.25 μg m^{-3}; these are only insignificantly better than the CWV-based estimate R of 0.912; the RMS error was 6.43 μg m^{-3}. Mid-afternoon PBL depth is consequently useful. However, the CWV-based estimate may be used with all years of the MODIS data, while the best-available meteorology for PBL depth varies considerably, as high-resolution NOAA models advanced through the years.

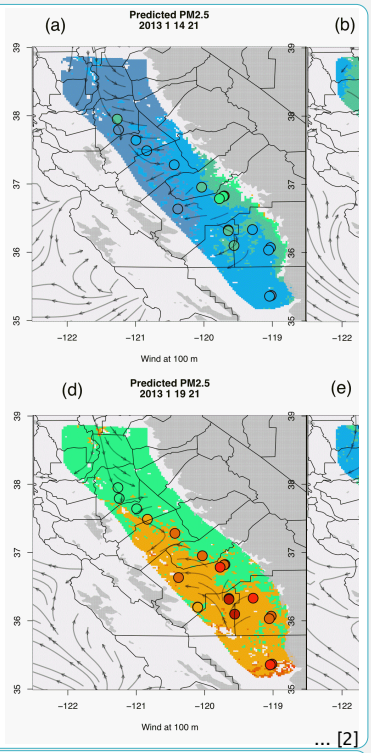
6. Results: Maps of Estimated PM2.5

The major purpose of this work, viz. to combine AOT, CWV, and daily calibration in order to allow maps of estimated PM2.5 for all regions where MODIS can provide optical thickness data. Results using the full model with α_i, β_i, and γ_s are shown (Figure 5f). Out of the 42 days in the calibration set, we consider 6 days of single major air pollution episode during middle of January, 2013, a period that was largely sampled by the DISCOVER-AQ ground and airplane samples. Detailed comparisons of the DISCOVER-AQ data would expand this work beyond a manageable size; such analysis is desirable. Winds are shown with streamlines and are obtained by interpolation from the RAP wind analyses.

We created 39 maps, six of which are shown in Figure 6. Accuracy is good. Residual Mean Squared Error, RMSE, ~7 μg m^{-3}. This dictated the 5 μg m^{-3} contour colors used: similar colors or neighboring colors show expectable agreement. Winds at 360 m for the hour of sampling have been superimposed on the maps.

There follows the description of just of one episode: On January 14, 2103, the Valley is clean (see also Figures 3 and 4). By January 16, 2013, light regional haze is accumulating, and the winds and mapped levels suggest some accumulation towards the south. On January 18, 2013, winds have veered: in the central Valley, pollution accumulates towards the east; in the south, transport is towards Bakersfield. On January 20, 2013, winds press the accumulating PM2.5 back towards the more populated east Valley. Several days following have increasing clouds (no maps). The first day, with advancing clouds overhead but no low clouds, no front, nor rain, retains high PM2.5 at the monitors. This pattern is seen for several wintertime pollution episodes in this region. When the clouds clear, the Valley is as clean as 14 Jan. In the maps for January 18, 19, and 20, the maps underestimate the highest values of PM2.5 by about 20%, as noted above.

Deleted:



Deleted: ... [2]

Deleted: ¶

Deleted: $(\alpha_i + \beta_i \cdot \mathrm{AOT}_{is}/\mathrm{CWV}/\mathrm{PBL}_{is})$

Deleted: ¶

Deleted: very slightly

Deleted: and

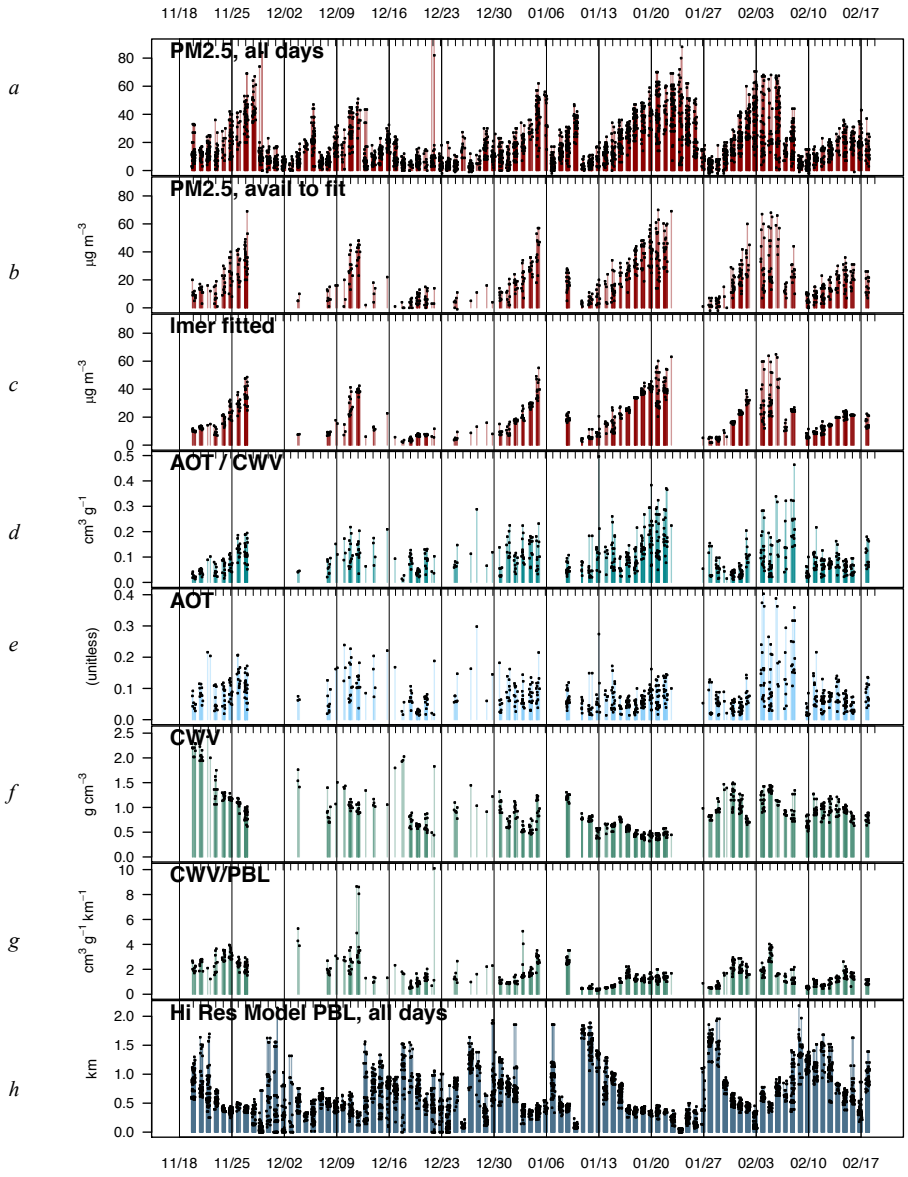


Figure 7. Time evolution of PM2.5 and related variables for ~8 intensifying particulate episodes during the winter of 2012–2013. Dots and vertical bars indicate variable values at individual stations when available. Blank regions reflect periods of cloud cover. (a) PM2.5, μg m^{-3}, as observed at stations (all dates) and (b) fitted PM2.5 on days and locations when MAIAC was available. (c) MAIAC AOT, 11:30 to 15:30 sun times. Note that increase is less pronounced than PM2.5, and varies between episodes. (d) CWV in g cm^{-2} or, colloquially, "cm of precipitable (liquid) water." (e) PBL

Formatted: Superscript

height for the noon-afternoon observations in this dataset. Morning PBL heights are much lower. (f) Ratio of CWV to PBL height (cm(H2O l) / km) , showing relative constancy over several days CWV time series resemble PBL height graphs.

7. Intensification of PM2.5 Episodes: Pollutant Accumulation vs Confinement

The well performing mixed effects models (equations 5 and 6) led us to examine the repeated development of air pollution episodes to a maximum, striking patterns seen in Figures 3, 4, and 6. How did the independent values for various models in Table 1 vary within episodes and between episodes? Our description of the development leads to some answers in Figure 7.

Deleted: Section

Figure 7a and 7b describe the development of the episodes. The time series of observed PM2.5 and fitted PM2.5 are repeated from Figure 4. The times with no data are essentially cloudy times. After periods of cloudiness, particulate values typically rise until the next period of clouds. There are 7–8 such periods of rising, or weather episodes. ("Episodes" can also refer to periods of highest particulate matter.) High values typically remain for 1–4 days after cloud obscuration. Figures 7c and 7b show the values fitted by our mixed effects regression and the values that are available for fitting. The time sequence as well as the magnitudes are in expectable agreement., but the variability between stations is smaller on some occasions (e.g., 1/17 and 02/15). Figures 7d shows that the time series of the ratio AOT/CWV develops from day to day as PM2.5 does, but suggests that these are modulated by differences in amplitude between weather episodes and sometime over several days within the weather episodes, e.g., 01/13 to 01/17 and 02/04 to 02/08. These explain the low overall correlation. In contrast, AOT shows little resemblance in the time series. Column water vapor, CWV, shows some tendency to decline during weather episodes (Figure 7e) notably the values at differing stations are more similar than those for AOT. Regionwide similarity in CWV within and above the afternoon mixed layer is an appealing explanation. Note the limited variability of the ratio CWV/PBL over 3–5 days and between stations. The afternoon PBL height itself is shown in Figure 7g. Note that it is often very low at the end of a cloudy period and then rises to high values ~1 km at the end of the cloudy period. We suggest that this reflects overcast skies and very limited convective mixing followed by rain and the introduction of new air masses with deeper mixing of water vapor in a less stable atmosphere.

Deleted: Cloud

Figure 7 describes differing causes of repeated PM2.5 buildup during cloud-free weather episodes. Progressive restriction of vertical mixing during clear-weather episodes acts to concentrate the effects of accumulated and recent pollution sources. The less stable air following a frontal passage feels increasing effects of strong subsidence, diminishing the mixing height. The threefold reduction in PBL height during major episodes, Figure 7d, nearly matches the 4-fold increases in PM2.5 during these periods (Figure 7a). MAIAC AOT shows variability between stations, and is reflected in local PM2.5. Winds redistribute particles and AOT. Figure 7 does not make clear the fate of aerosol, but it likely escapes with mountainside winds along the Valley. The entire set of maps suggests a flow to the south and stronger outflow near the Tejon pass east of Bakersfield. These mountainside winds likely may facilitate water vapor and aerosols escape the prevalent mixed layer.

This suggests a typical behavior for the San Joaquin and similar regions in winter. A cloudy disturbance (new air mass, rain, wind) stirs the lower troposphere. This initiates a high PBL mixing on the first clear days. Typical fair-weather subsidence begins. The surface buoyancy flux is too weak to maintain these relatively high mixed layer

tops; Afternoon PBL depths and mixed layer depths decrease day by day until a depth of 300-400 m is reached. (Figure 7d). Escape from the Valley may slow, allowing accumulation of pollution from within the region or from upwind. This further increases the surface PM2.5. Relatively local sources add to both AOT and PM2.5, and can transport them 50–100 km downwind, occasionally from east-valley sources to west-valley pollution hotspots (the map of Figure 6d). Both subsidence and surface buoyancy flux are broad-scale weather phenomena (~300 km) , and so AOT-to-PM2.5 relationships are similar on a given day with a given history of weather. Finally, warm-frontal rain approaches the region.

An examination of HSRL2 data for the DISCOVER-AQ period (Sawamura et al., 2017) suggests that there can be considerable vertical variability of aerosol extinction; the fact that AOT tends to average the whole afternoon mixed layer allows our generalized description to hold nevertheless.

Finally, we venture some ideas for filling in afternoon PM2.5 on days when MAIAC did not allow mapping due to cloud cover. Young et al. (2016) provided a thorough microphysical and chemical analysis for just the time period January 13 to February 11, 2013 — essentially the DISCOVER-AQ period — and just the fully instrumented UC Davis site deployed at Fresno. Their Figure 2, panels a, b, and e show time series of temperature (panel a), wind speed and direction (panel b), and particle mass for the period (panel e). Their measurements include periods of cloud cover and clearly show air mass transitions during rain and frontal passage (seen as wind shifts, commonly N to W to S) These time series suggest a meteorological plausible method to interpolate PM2.5 maps into cloud-covered days. These do compare to the panels a, b, and c in our Figure 7, describing observed and statistically estimated particulate mass at all stations including Fresno. PM2.5 drops to values below 10–15 µg m^{-3} whenever wind speeds rise to above ~2 m s^{-1} and the wind direction is from a quadrant (90 degree sector) centered on the north-northwest direction. Their Figure 2a also describes rainfall at the Fresno site. Particulate matter does drop by ~50% from the highest observed/estimated values at the end of the clear-sky period, and further when the winds rise to 2 m s^{-1} or higher. This behavior is most clearly observed in their graphs for the period January 23–January 27. Similar behavior is observed in the period February 6–February 11, although the episode has more complex increase than in the earlier, most intense episodes. The short spike up to 80 µg m^{-3} on the night of February 10 is not explained, and not reflected in the afternoon-only data of our Figure 7. Nevertheless, the averages shown by Young in Figure 2e do repeat the general observation that daily average PM2.5 and afternoon PM2.5 do tend to correlate well. For best-estimate maps of PM2.5, we suggest that the end-of-retrieval values of PM2.5 reduce gradually over a day or two. Maps of precipitation (e.g. from radar or other analyses) allow more detail. Estimates for a region should then fall to ~7 µg m^{-3} whenever sustained winds rise to > 2 m s^{-1} from the NNW or > 3 m s^{-1} from any direction. Such wind speeds are held to mark air mass replacement (e.g, frontal passage). These ideas remain suggestions since our analysis for a single winter may not provide enough instances. The whole Aqua MAIAC period is available, but currently beyond NASA's resources.

Deleted: DISCOVER-AQ

Deleted: (

Deleted: 14

Deleted:)

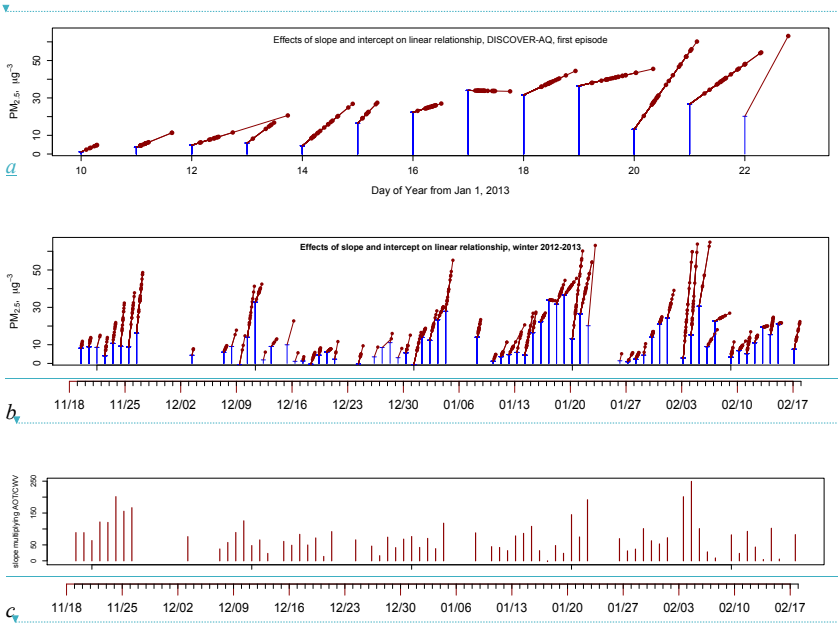
Deleted: should be compared

Deleted:

8. Variation of Random-effects Model Parameters

The preceding section gives some background so that we may understand the parameters for the random effect model. We will discuss the full Equation 4; results with mild spatial dependence (Equation 5) are very similar. The

intercept α_i and the slope for $\beta_i \cdot \mathrm{AOT}_{is}/\mathrm{CWV}_{is}$ are the same for each day and determine the fitted PM2.5 for the regression Equation 4. We exploit this to produce a "stork plot" like Figure 8. High α_i is shown by tall blue lines; high β_i is shown as a high slope. Variation in AOT/CWV contributes ~30–70% to the estimate on almost all days.



a

b

c

Figure 8. Roles of slopes and intercepts in a regression fit. (a) A "stork plot" for the clear-sky air-pollution episode mapped in Figure 6. Vertical blue lines indicate the contribution of the random intercept α_i to the total PM2.5 fitted in the model. These are the same for all geographical locations including the observation stations for any given day. The slope parameter β_i is the same for all geographical locations. (See note to Table 1.) Values of PM2.5 evaluated at the stations are shown by red dots along a line. Large values of AOT/CWV have wide vertical extent, and the corresponding high values of PM2.5 are shown as red dots at the upper right of each day's plot. Highly sloped lines indicate high β_i. (b) A stork plot for the whole wintertime interval evaluates, showing several clear-day episodes. (c) the values of β_i vary considerably. These slopes are shown as a time series.

The stork plot of Figure 8a illustrates a puzzling progression of parameter estimates day by day. For the first days, Jan 10–14, the slope parameter accounts for the largest contribution to PM2.5. For the second part of the period, Jan 15–19, the intercept term becomes progressively more important compared to the AOT/CWV dependence. The regression equation fit (Figure 7c) has difficulty in matching the observed PM2.5 (Figure 7b) variability between stations on these days although AOT (Figure 7e) shows moderate variability around low values, 0.03–0.05. (A side note: MAIAC AOT estimates should be particularly challenged at these low values.) Then, from the Jan 20–22, the intercept contribution diminishes and the AOT/CWV dependence becomes rather larger than typical. Referring back to Figure 7e, f, and g,, these variations seem explainable: the mixed layer decreases rapidly during the first period, then reaches a minimum at ~300 m. In the last three days, the AOT increases rapidly, though

Deleted:

Deleted: .

Deleted: geographcal

Deleted: days

Deleted:

Deleted:

Deleted: b

Deleted: shows

Deleted:

the mixing depth changes little. The following weather episode is notable for high and quite variable AOT (Figure 7e), and the fitting procedure does well.

9. Value of Improved CWV Data

At this point, concerns about the quality of the CWV estimate should be addressed. In our analysis of the difficult San Joaquin Valley, MAIAC CWV can be frequently low compared to AERONET CWV, some error can arise from the presence of clouds in neighboring footprints. In the figures and results shown CWV was based on the MAIAC data interpolated and extrapolated where cloud-contamination made the retrieval of lower accuracy (Lyapustin et al., 2018). Figure 6 shows some small-scale variability .RAP analyses of CWV could also be used at their 13-km model-imposed width with similar results, since CWV does not vary as rapidly spatially as AOT. A better direct use of the MAIAC CWV could uses spatial averaging with a width of 3 to 6 km. Random errors in the MAIAC CWV due to the low radiances used would be reduced; considerations of source patterns suggests that CWV might not truly vary at such small scales. Improved PM2.5 values could result. We are implementing this averaging.

As understanding of MAIAC CWV improves, its role in determining daily AOT-to-PM2.5 relationships should improve; calibration of MAIAC using sun-photometer measurements can be useful in the meantime (Just et al., 2019). Note also that assimilated CWV from the National Weather Service models is constrained empirically, and so not as reliant transport descriptions as is aerosol. Here are some constraints surface-station humidity measurements constraing CWV below 0.4–1 km, thermal-radiation sounders on the GOES (Geostationary Operational Environmental Satellite) satellites describe water vapor partial above that; radiosonde and GPS humidity sensors give further constraint. This allows GOES AOT estimates to be used with assimilated CWV, even though GOES lacks a reflective water vapor channel (S. Kondragupta, personal communication, 2018).

10. Conclusions

Goals: We sought broadly applicables methods to estimate PM2.5 maps from satellite AOT for very polluted regions poorly described by satellite data. Ths study focusrd on the whole polluted winter season of the San Joaquin Valley (SJQ), November 19, 2012 to February 18, 2013. We sought to fulfill the overarching goal of the whole DISCOVER-AQ mission— to find general relationships between extended satellite data observations and surface air pollutant concentrations and to evaluate their success. We found success with a simple methodology that follows the meteorology of regions like the SJQ. This success recommends an approach to the remote-sensing to PM2.5 analysis, investingating important pollution regions in terms of their meteorology and sources, but carrying over methods from similar regions. For example, the Po Valley of Itally and the Northern Gangetic Plane of India may respond similary to analyses based on detailed mixing height data and related distribution indicators.

Direct results: We found that a combination of information utilizing (1) optical depth, (2) measures of vertical dispersion, e.g. CWV, and (3) daily calibration of PM2.5 to predictors produced significantly better quantification of PM2.5 than a competitive no-satellite-use method which we named “regional correlation” since it produces un-featured maps of PM2.5 which vary only from day to day. Our maps of estimated PM2.5 extend for all cloud-free

periods November 19 2012 to February 18, 2013, essentially the whole pollution season for this winter. For that whole period, this first published attempt found good predictive value of $R \sim 0.9$ and rms error of 6.5 μg m^{-3}. Cross-validation suggested rms error of $\lesssim 7$ μg m^{-3}. Analysis of residuals suggested that better rms errors could be achieved if further workallowed for sub-regioality (use of smaller regions or a geographic characterization incorporating some spatial variation). Local variations in PM2.5 on the order of 1–3 km were noted using our method, but only when particulate accumulation could occur along-wind. Still, in order to estimate PM2.5 at $\lesssim$ 1-km scales, we expect that it will be necessary to use refined geographic information system methods (Kloog et al. 2014).

DISCOVER-AQ comparisons advisable: Our analyzed winter 2012–2103 period did include the more limited DISCOVER-AQ / California-2013 airborne-intensive study period, primarily focused on the area around Fresno. Analysis of that intensive suggested ideas (Shook et al., 2013) that motivated this work.The shorter DISCOVER-AQ period does deserve more detailed comparison to our results. Aircraft in-situ profiles of gas and particle composition, lidar profiles, very detailed surface measurements of particulate composition, and source-and-transport modeling all deserve comparison. The distribution of atmospheric particles and precursor gases is more complex than this work might suggest. Somehow averaging appears to allow our general methods. The development of concepts and the length of this workdo not allow for such comparison. We hope that research will be encouraged.

Usefulness of Column Water Vapor: A major finding was that the usefulness of CWV does not become apparent unless there is daily calibration of the AOT/CWV relationship to PM2.5. We attribute this primarily (a) details of CWV: e.g., CWV's dependence on mixed layer temperature on the timescale of days, (b) to CWV above the mixed layer for aerosols, presumably responding to other H_2O sources upwind, and (c) variations in composition: the relation of PM2.5 to light extinction. We believe that allowing for a full linear relationship each day for AOT/CWV to PM2.5, both slope and intercept effects, in a daily calibration allows regressioin to exploit portions of the PM2.5 vs. *f*(AOT/CWV) scatterplots that reveal proportionality. High-spatial-resolution estimates of the 11 AM–3:30 PM PBL heights for momentum may as helpful as CWV when available and when the PBL estimation has been examined for accuracy; this could be explored. Such PBL data is not available for the whole MODIS-Aqua period.(2004–present), while CWV is.

Accompanying insights on pollution episodes. We found that this approach allowed a broad description of the buildup of six air pollution episodes and the balance of the roles of accumulation of pollutants versus limited vertical mixing. Episodes were as in earlier descriptions (Watson and Chow, 2002). Each appears important in different phases of reptitive PM2.5-increase cycles. PM2.5 to AOT relationships suggest a few days residence time for particles (actually prticulate extinction) in the Valley. The first 1–3 days after MODIS described full cloud cover could still show high, slowly decreasing PM2.5. Unpublished analysis (see Young et al. 2018) suggests that this high PM2.5 dropped preciptously when surface winds rose to > 4 m s^{-1} from a quadrant centered on the NNW.

Best-estimate extensions to cloudy periods of the remote-sensing-based record can be made using the typical meteorology of the San Joaquin or presumably other areas, and verified by extensive checks. Widely available data mapping surface winds and precipitation suffice, and do not require that detailed meteorological modeling be available.

Formatted: Superscript
Deleted:
Deleted:
Deleted: 2012z
Deleted:
Deleted: =–
Deleted: isolate
Deleted: in some circumstances

Role of "Static" models: Our estimation approach aimed to avoid the use of modeling driven by source estimation and transport simulatoin. Principally we wished to provide dataseta that allowed independetn comparison to such three-d atmospheric chemistry models (e.g., Friberg, et al., 2018). When we used RAP-model CWV rather than spatially averaged or calibratee (Just et al., 2019, manuscript in progress) MAIAC CWV, that goal was not fully reached, although RAP CWV is strongly constrained by surface, satellite, snd other observations. An aspirational goal is to provide an economical, accurate, and calibrated estimation of PM2.5 for the whole MODIS Aqua period to date, and then beyond. The opportunities to use MISR,VIIRS, MAIA, and even geostationary imaging are appealing!

11. Acknowledgements;

We gratefully acknowledge the support from NASA's DISCOVER-AQ mission, followed by very encouraging continued interest and some partial support from the Health and Air Quality program management of NASA's Earth Science Applications Division. This allowed fulfilment of the prime DISCOVER-AQ objective to demonstrate the relevance of remote sensing to specific air pollution problems. We appreciate advice individuals in that program's Applied Science Team and from the GEO-CAPE mission formulation effort (aerosol focus). Aid from Yujie Wang (NASA GSFC) regarding the MAIAC processing of MODIS data were helpful. Michael Shook's analysis suggested the use of CWV. Recent comments on the draft paper by Qian Tan and Frank Freedman are also appreciated.

12. References

Appel, K. W., Napelenok, S. L., Foley, K. M., Pye, H. O. T., Hogrefe, C., Luecken, D. J., Bash, J. O., Roselle, S. J., Pleim, J. E., Foroutan, H., Hutzell, W. T., Pouliot, G. A., Sarwar, G., Fahey, K. M., Gantt, B., Gilliam, R. C., Heath, N. K., Kang, D., Mathur, R., Schwede, D. B., Spero, T. L., Wong, D. C., and Young, J. O.: Description and evaluation of the Community Multiscale Air Quality (CMAQ) model- ing system version 5.1, Geosci. Model Dev., 10, 1703–1732, https://doi.org/10.5194/gmd-10-1703-2017, 2017.

Ballard M., Newcomer M., Rudy J., Lake S., Sambasivam S., Strawa A.W., Schmidt C., Skiles J.W.: Understanding the correlation of San Joaquin air quality monitoring with aerosol optical thickness satellite measurements. ASPRS Annual Conference, Baltimore MD, 2008.

Bell M.L., Ebisu K., Belanger K.: Ambient air pollution and low birth weight in Connecticut and Massachusetts. Environ. Health Perspect., 115, 1118–1124, 2007.

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 assim- ilated meteorology: Model description and evaluation, J. Geophys. Res.-Atmos., 106, 23073–23095, 2001.

Brunekreef, B, Holgate, ST.: Air pollution and health. Lancet 360, 1233–1242. 2002.

Cappa C.D. and Zhang Q.: Characterization of PM2.5 Episodes in the San Joaquin Valley Based on Data Collected During the NASA DISCOVER-AQ Study in the Winter of 2013. Report to the California Air Resources Board Research Division Project # 14-307, 2018.

Caputi D. J., Faloona I., Trousdell J., Smoot J., Falk N., Conley S.: Residual Layer Ozone, Mixing, and the Nocturnal Jet in California's San Joaquin Valley. Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-854, under review, 2018.

Chow, J. C., Chen, L. W. A., Watson, J. G., Lowenthal, D. H., Magliano, K. A., Turkiewicz, K., Lehrman, D. E.: PM2.5 chemical composition and spatiotemporal variability during the California regional PM10/PM2.5 air quality study (CRPAQS), J. Geophys. Res.-Atmos., 111, D10S04, doi:10.1029/2005JD006457, 2006.

Chu D. A., Ferrare R., Szykman J., Lewis J., Scarino A., Hains J., Burton S., Chen G., Tsai T., Hostetler C., Hair J., Holben B., Crawford J.: Regional characteristics of the relationship between columnar AOD and surface $PM_{2.5}$ Application of lidar aerosol extinction profiles over Baltimore–Washington Corridor during DISCOVER-AQ. Atmospheric Environment, 101, 338-349, 2015.

Diner, D.J., S.W. Boland, M. Brauer, et al.:Advances in multiangle satellite remote sensing of speciated airborne particulate matter and association with adverse health effects: from MISR to MAIA. Journal of Applied Remote Sensing 12 (4), 2018

Dominici, F.; Peng, R.D.; Bell, M.L.; Pham, L;. McDermott, A.; Zeger, S.L.; Samet, J.M.: Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and Respiratory Diseases. JAMA, 295(10), 1127–1134, doi:10.1001/jama.295.10.1127, 2006.

Engel-Cox J.A., Holloman C. H., Coutant S.W., 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.

Ford B., and Heald C.L.: Exploring the uncertainty associated with satellite-based estimates of premature mortality due to exposure to fine particulate matter. Atmos. Chem. Phys., 16, 3499 –3523, doi:10.5194/acp-16-3499-2016, 2016.

Franklin M., Kalashnikova O., Garay M.: Size-resolved particulate matter concentrations derived from 4.4 km resolution size-fractionated Multi-angle Imaging SpectroRadiometer (MISR) aerosol optical depth over Southern California. Remote Sens. Environ., 196, 312–323, 2017.

Franklin M., Zeka, A., Schwartz. J.: Association between PM_{25} and all-cause and specific cause mortality in 27 US communities. Journal of Exposure Science & Environmental Epidemiology, 17,279-287, 2007.

Friberg M.D., Kahn R.A., Limbacher J. A., Appel K.W., Mulholland J. A.: Constraining chemical transport PM2.5 modeling outputs using surface monitor measurements and satellite retrievals: application over the San Joaquin Valley, Atmos. Chem. Phys., 18, 12891–12913, https://doi.org/10.5194/acp-18-12891-2018, 2018.

Gupta P., Christopher S.A., Wang J., Gehrig R., Lee Y.C., Kumar N.: Satellite remote sensing of particulate matter and air quality assessment over global cities. Atmos. Environ. 40, 5880–5892, 2006.

Hastie, Tibshirani, R. Tibshirani, and J. Friedman: The elements of statistical learning, New York. NY: Springer, 2009.

Hoff R.M. and Christopher A. S.: Remote Sensing of Particulate Pollution from Space: Have We Reached the Promised Land?, Journal of the Air & Waste Management Association, 59: 645-675, 2009.

Holben B.N., Eck T.F., Slutsker I., Tanré D., Buis J.P., Setzer A., Vermote E., Reagan J.A., Kaufman Y.J., Nakajima T, Lavenu F., Jankowiak I., Smirnov A., AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. Remote Sensing of Environment, Volume 66, Issue 1, 1-16, https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Hu X., Waller L.A., Lyapustin A., Wang Y., Al-Hamdan M.Z., Crosson W.L., Estes Jr. M.G., Estes S.M., Quattrochi D.A., Puttaswamy S.J., Liu Y.: Estimating ground-level PM2.5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. Remote Sens. Environ., 140, 220–232, doi:10.1016/j.rse.2013.08.032, 2014.

Jin X., Fiore A. M., Curci G., Lyapustin A., Civerolo K., Ku M., van Donkelaar A., Martin R. V.: Assessing uncertainties of a geophysical approach to estimate surface fine particulate matter distributions from satellite-observed aerosol optical depth. Atmos. Chem. Phys., 19, 295-313, https://doi.org/10.5194/acp-19-295-201, 2019.

Johnson M. S., Yates E.L., Iraci L.T., Loewenstein M., Tadić J.M., Wecht K.J., Jeong S., Fischer, M.L.: Analyzing source apportioned methane in northern California during Discover-AQ-CA using airborne measurements and model simulations. Atmos. Environ., 99, 248-256, doi:10.1016/j.atmosenv.2014.09.068, 2014.

Just A. C., De Carli M., Shtein A., Dorman M., Lyapustin A., Kloog I.: Correcting Measurement Error in Satellite Aerosol Optical Depth with Machine Learning for Modeling PM2.5 in the Northeastern USA. Remote Sensing, 10(5), 803, 2018.

Justice E., Huston L., Krauth D., Mack J., Oza S., Strawa A.W., Skiles J.W., Legg M., Schmidt C.: Investigating correlations between satellite-derived aerosol optical depth and ground PM2.5 measurements in California's San Joaquin Valley with MODIS Deep Blue. ASPRS Annual Conference. Baltimore, MD, 2009.

Kloog I., Chudnovsky A.A., Just A.C., Nordio F., Koutrakis P., Coull B.A., Lyapustin A., Wang Y., Schwartz J.: A new hybrid spatio-temporal model for estimating daily multi-year PM2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. Atmos. Environ., 95, 581–590, doi:10.1016/j.atmosenv.2014.07.014, 2014.

Kloog I., Sorek-Hamer M., Lyapustin A., Coull B., Wang Y., Just A.C., Schwartz J., Broday D.M.: Estimating daily PM2.5 and PM10 across the complex geo-climate region of Israel using MAIAC satellite-based AOD data. Atmos. Environ, 122, 409–416, doi:10.1016/j.atmosenv.2015.10.004, 2015.

Kloog I., Ridgway B., Koutrakis P., Coull B.A., Schwartz J.D. Long- and short-term exposure to PM25 and mortality. Epidemiology. 24, 555-561, 2013.

Koelemeijer R.B.A., Homan C.D., Matthijsen J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe. Atmospheric Environment. 40(27); 5304-5315, 2006.

Lee P., Pan L., Kim H., Tong D.: Intensive Campaigns Supported by Air Quality Forecasting Capability to Identify Chemical and Atmospheric Regimes Susceptible to Standard Violations. In: Steyn D., Mathur R. (eds) Air Pollution Modeling and its Application XXIII. Springer Proceedings in Complexity. Springer, Cham, DOI https://doi.org/10.1007/978-3-319-04379-1, 2014.

Lee, H.J., Liu, Y., Coull, B.A., Schwartz, J., Koutrakis, P. A novel calibration approach of MODIS AOD data to predict $PM_{2.5}$ concentrations. Atmos. Chem. Phys, 11, 7991–8002, 2011.

Liu B., Ma Y., Gong Y., Zhang M., Wang W., Shi Y.: Comparison of AOD from CALIPSO, MODIS, and Sun Photometer under Different Conditions over Central China. Scientific Reports. volume 8, 2045-2322, https://doi.org/10.1038/s41598-018-28417-7, 2018.

Liu Y . Franklin M . Kahn, R, Koutrakis, P.: Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: A comparison between MISR and MODIS. Remote Sensing Environ, 107. 33-44, 2007.

Liu Y., Paciorek C. J., Koutrakis P.: Estimating Regional Spatial and Temporal Variability of PM2.5 Concentrations Using Satellite Data, Meteorology, and Land Use Information. Environ. Health Perspect., 117(6), 886-892, doi:10.1289/ehp.0800123, 2009.

Lyapustin A, Wang Y, Korkin S, Huang D. MODIS Collection 6 MAIAC algorithm, Atmospheric Measurement Techniques, 11, October, 2018.

Lyapustin A., Martonchik J., Wang Y., Laszlo I., Korkin S.: Multi-Angle Implementation of Atmospheric Correction (MA- IAC): Part 1. Radiative Transfer Basis and Look-Up Tables. J. Geophys. Res., 116, D03210, doi:10.1029/2010JD014985, 2011a.

Lyapustin A., Wang Y., Laszlo I., Kahn R., Korkin S., Remer L., Levy R., Reid J. S.: Multi-Angle Implementation of Atmospheric Correction (MAIAC): Part 2. Aerosol Algorithm. J. Geo- phys. Res., 116, D03211, doi:10.1029/2010JD014986, 2011b.

Martins V.S., Lyapustin A., de Carvalho L.A., Barbosa C.C., Novo E.M.: Validation of high-resolution MAIAC aerosol product over South America. Journal of Geophysical Research: Atmospheres, 122, 7537-5, 2017.

Martins V.S., Novo EM, Lyapustin A, Aragão LE, Freitas SR, Barbosa CC.: Seasonal and interannual assessment of cloud cover and atmospheric constituents across the Amazon (2000–2015): Insights for remote sensing and climate analysis. ISPRS Journal of Photogrammetry and Remote Sensing, 2018.

Nolte, C. G., Appel, K. W., Kelly, J. T., Bhave, P. V., Fahey, K. M., Collett Jr., 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 com- position across sites in North America, Geosci. Model Dev., 8, 2877–2892, https://doi.org/10.5194/gmd-8-2877-2015, 2015

Sawamura P., Moore R. H., Burton S. P., Chemyakin E., Müller D., Kolgotin A., Ferrare R. A., Hostetler C. A., Ziemba L. D., Beyersdorf A. J., Anderson B. E.: HSRL-2 aerosol optical measurements and microphysical retrievals vs. airborne in situ measurements during DISCOVER-AQ 2013: an intercomparison study. Atmos. Chem. Phys., 17, 7229-7243, https://doi.org/10.5194/acp-17-7229-2017, 2017.

Schmid B., Ferrare R., Flynn C., Elleman R., Covert D., Strawa A., Welton E., Turner D., Jonsson H., Redemann J., Eilers J., Ricci K., Hallar A. G., Clayton M., Michalsky J., Smirnov A., Holben B., Barnard J.: How well can we measure the vertical profile of tropospheric aerosol extinction?. J. Geophys. Res, 111, D05S07 doi:10.1029/2005JD005837, 2006.

Schwartz J.: Air pollution and hospital admissions for respiratory disease. Epidemiology, 7, 20–28, 1996.

Shook M.: Daily evolution of boundary layer properties based on NASA DISCOVER-AQ airborne profiles over the California San Joaquin Valley. American Geophysical Union, Fall Meeting 2013, 2103, A43A-0228, 2013.

Sorek-Hamer M., Just A., Kloog I:, Satellite remote sensing in epidemiological studies, 1–6, Curr Opin Pediatr, 28(2), 228–234, doi:10.1097/MOP.0000000000000326, 2016.

Sorek-Hamer M., Broday D.M., Chatfield R., Esswein R.: Monthly analysis of PM ratio characteristics and its relation to AOD. Journal of the Air & Waste Management Association, Vol 67, 27-38, 2017.

Sorek-Hamer M., Strawa A.W., Chatfield R.B., Esswein R., Cohen A., Broday D.M.:. Improved retrieval of $PM_{2.5}$ from satellite data products using non-linear methods. Environmental Pollution. 182,417-423 2013.

Turco, R.P.: Earth Under Siege, From Air Pollution to Global Change, Oxford University Press, ISBN: 9780195142747, 2002.

Van Donkelaar A., Martin R.V., Brauer M., Kahn R., Levy R.C., Verduzco C., Villeneuve P.J.: Global estimates of ambient fine particulate matter concentrations from satellite based aerosol optical depth: Development and application. Environmental Health Perspectives, 118; 847-855, 2010.

Van Donkelaar A., Martin R. V., Spurr R.J.D., Burnett R.T.: High-Resolution Satellite-Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over North America. Environ. Sci. Technol, 49, 10482–91, doi:10.1021/acs.est.5b02076, 2015.

Van Donkelaar A., Martin R.V., Brauer M., Hsu N.C., Kahn R.A., Levy R.C., Lyapustin A., Sayer A.M., Winker D.M.: Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. Environmental Science & Technology, 50 (7), 3762-3772, DOI: 10.1021/acs.est.5b05833, 2016.

Watson J. G., and Chow, J. C.: A wintertime PM 2.5 episode at the Fresno, CA, supersite. Atmos. Environ., 36, 465-475, 10.1016/S1352-2310(01)00309-0, 2002.

Watson J.G.: Visibility: Science and Regulation. Journal of the Air & Waste Management Association, 52:6, 628-713, DOI: 10.1080/10473289.2002.10470813, 2002.

Wilkinson G. N., and C. E. Rogers: Symbolic Description of Factorial Models for Analysis of Variance. Journal of the Royal Statistical Society, Series C (Applied Statistics) 22, no. 3 392-99, doi:10.2307/2346786, 1973.

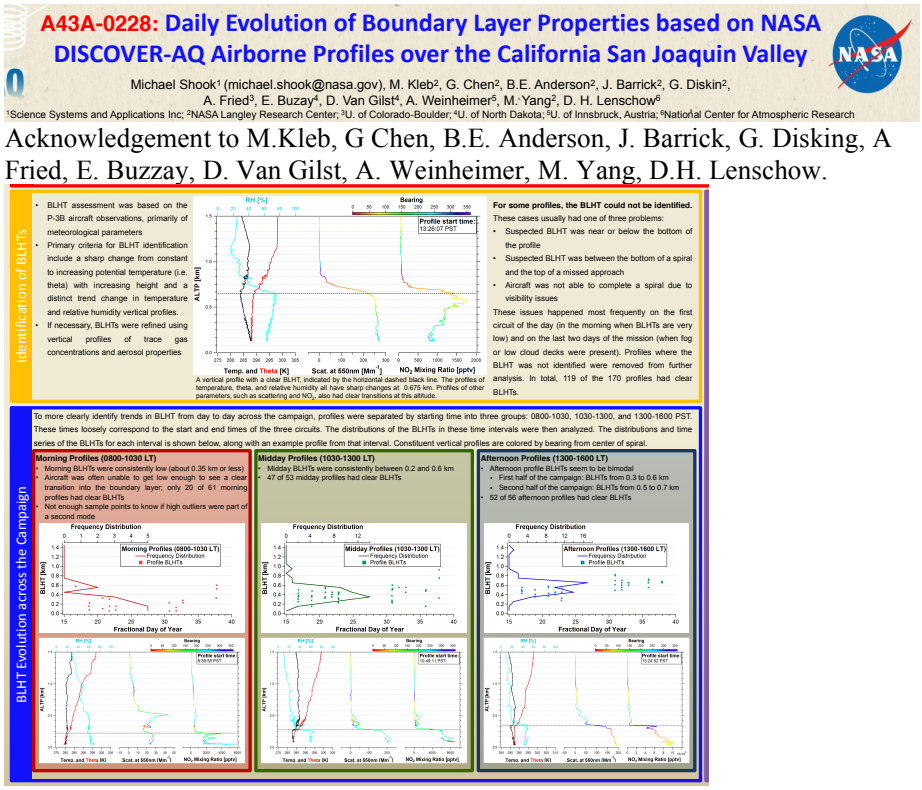
Young D. E., Kim H., Parworth C., Zhou S., Zhang X., Cappa C. D., Seco R., Kim S., Zhang Q. () Influences of emission sources and meteorology on aerosol chemistry in a polluted urban environment: results from DISCOVER-AQ California. Atmos. Chem. Phys., 16, 5427-5451, https://doi.org/10.5194/acp-16-5427-2016, 2016..

Zanobetti A., Franklin M., Koutrakis P., Schwartz J.: Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environmental Health, 8, 2009.

700

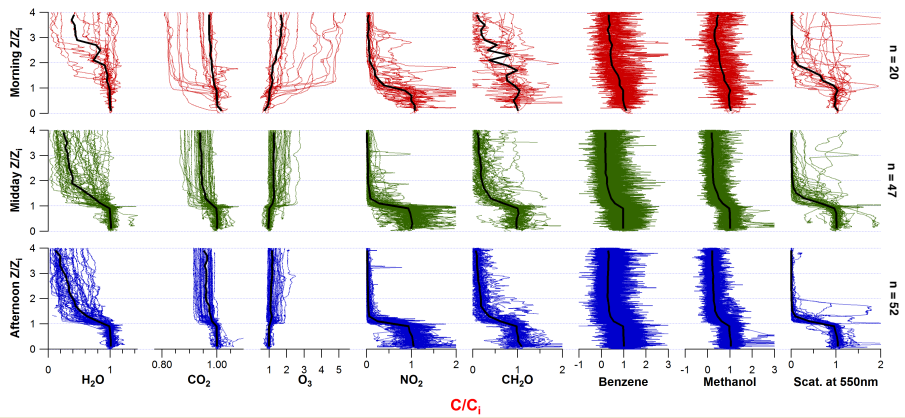
Supplementary Material

Included below are three portions of a poster presented by Michael Shook at the American Geophysical Union (Shook et al., 2013) soon after the DISCOVER-AQ measurements in California. That work gave a syncretic overview of many trace species measurement besides water vapor which motivated our investigation, as described in Section 1.1. . First, there is an acknowledgement of the many authors contributing measurements. Second, there is an explanation of methods of estimating a characteristic mixed layer height as observed of airplane measurements. Third, there are graphs of vertical profiles normalized by average tracer concentration and also by the depth of the mixed layer. The commonalities and some differences of the first tracer shown (water vapor) and the last (particulate scattering).



Acknowledgement to M.Kleb, G Chen, B.E. Anderson, J. Barrick, G. Disking, A Fried, E. Buzzay, D. Van Gilst, A. Weinheimer, M. Yang, D.H. Lenschow.

To visualize BL variability and vertical gradients, composite profiles for eight different constituents were created. Constituents were chosen to represent a variety of lifetimes and production/removal processes.



Composite profiles of different constituents scaled by the profile BLHT and the average concentration of the constituent within the boundary layer. The dotted, colored lines represent individual profiles, and the solid black lines represent the median profile for that constituent and time period.
Z = pressure altitude, Z_i = profile BLHT, C = constituent concentration, and C_i = average constituent concentration in the boundary layer

720

Page 8: [1] Deleted Rev 11/28/19 11:53:00 PM

Page 16: [2] Deleted Rev 11/28/19 11:53:00 PM