Presented at the 16th Annual CMAS Conference, Chapel Hill, NC, October 23-25, 2017

A SATELLITE-DISPERSION MODELING SYSTEM TO GENERATE HIGH-RESOLUTION DOWNSCALED PM2.5 FIELDS

Frank R. Freedman[1,2]*
[1]Department of Meteorology and Climate Sciences, San Jose State University, San Jose, CA, USA

Mohammad Al-Hamdan[3,2], Akula Venkatram[4], Seyedmorteza Amini[4], Sen Chiao[1,2]
[2]NASA MIRO Center for Applied Atmospheric Research and Education
[3]Universities Space Research Association, NASA Marshall Space Flight Center, Huntsville, AL, USA
[4]Department of Mechanical Engineering, University of California Riverside, Riverside, CA, USA

1. INTRODUCTION

There is growing interest in knowing air pollution levels at finer resolution. Enhanced methods to generate higher resolution and more spatially complete fields for PM2.5 and finer-sized particulate in particular would be beneficial since these are among the most harmful air pollutants for human health. Aerosol optical depth (AOD) data from the Moderate Resolution Imaging Spectrometer (MODIS) platform aboard the NASA Aqua and Terra polar orbiting satellites have been used to develop continuous PM2.5 fields at regional scale (e.g. van Donkelaar et al. 2015, Lee et al. 2016). Downscaled air pollution fields from regional fields using "hybrid" dispersion models, which combine outputs from regional and road-resolving models, have been generated over the last decade for certain urban metropolitan areas, and work is ongoing to enhance this capability (e.g. Batterman et al. 2015, Zhai et al. 2016, Bates et al. 2017, Hood et al. 2017).

As part of our NASA Health and Air Quality Applied Science Team membership (HAQAST, https://haqast.org/), our team is developing a hybrid dispersion modeling system that utilizes MODIS AOD to provide regional-scale PM2.5 fields and computationally efficient dispersion models to downscale a user-specified sub-region of this field to fine scale (~ 100 m resolution). To develop the system, daily satellite-informed regional PM2.5 fields are being developed over California utilizing MODIS Dark Target 3-km AOD retrievals (https://darktarget.gsfc.nasa.gov/products/land-3). The dispersion modeling to downscale a sub-region of this field combines Lagrangian backwards trajectory and line-source models, the latter currently configured to simulate local variations of fine particulate fields due to major California roadways.

Below we present a description of the system in its current state of development. We detail its design and demonstrate the procedure to downscale a sub-region within the satellite field by applying the system to a particular site and day. Currently, we have developed daily satellite-derived PM2.5 fields over California for 2016, with previous years and updates to 2017 planned. While demonstrated here for a particular site and day, the downscaling dispersion modeling system can be applied readily to multiple sites and over multi-day time periods (e.g. monthly).

2. SYSTEM DESIGN

The components of the system for combining satellite-derived regional PM2.5 fields with dispersion modeling to arrive at a downscaled representation within a certain sub-region of the satellite-derived field are shown in Figure 1. AOD from Dark Target 3-km are taken from the Aqua overpass (~ 1230-1430 LST daily), and are mapped to daily-average EPA FRM site PM2.5 measurements using a multi-regression model with meteorological variables from NLDAS-2 (https://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php) as co-variates. Daily PM2.5 fields at 3-km resolution are then constructed by applying spatial surfacing to the regressed PM2.5 field using the procedure of Al-Hamdan et al. (2009, 2014). Further details are given in Section 3.

**Corresponding author:* Frank R. Freedman, Department of Meteorology and Climate Sciences, San Jose State University, One Washington Square, San Jose, CA, 95192-0104; e-mail: frank.freedman@sjsu.edu

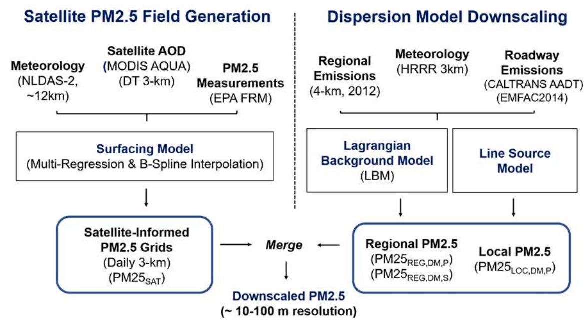


Figure 1: Components to construct regional PM2.5 fields from MODIS Dark-Target 3-km satellite retrievals and downscale these to fine-scale using Lagrangian backwards and line-source dispersion models.

The dispersion model fine-scale representation of a sub-region within the satellite-derived PM2.5 field is generated by combining outputs from two models: the Lagrangian Background Model (LBM, Pournazeri et al. 2014) for regional PM2.5, and the line-source model of Venkatram and Horst (2006) for local PM2.5 due to road traffic. Inputs to the dispersion modeling are a regional-scale gridded PM2.5 emission field, gridded meteorological fields, and traffic counts and vehicular emission factors. Year 2012 PM2.5 emission fields have been provided to us daily at 4-km resolution by South Coast Air Quality Management District (SCAQMD), and for summer and winter periods by the Bay Area Air Quality Management District (BAAQMD). Collectively, these cover the main population and PM2.5-affected areas of the state (Southern California including Imperial Valley, Bay Area, and Central Valley). Hourly gridded meteorological fields are from the initial forecast hour of the operational 3-km High-Resolution Rapid Refresh (HRRR, https://rapidrefresh.noaa.gov/hrrr/). Annual average daily traffic (AADT) for light-duty vehicles and trucks are from data provided by the California Department of Transportation (http://www.dot.ca.gov/hq/tsip/gis/datalibrary/#Highway) and emission factors from EMFAC2014 (https://www.arb.ca.gov/msei/categories.htm).

Collectively, the system therefore provides three independent measures of PM2.5 at a user-specified sub-region: a satellite-derived regional value ($PM25_{SAT}$), an LBM dispersion model regional value divided into primary and secondary portions ($PM25_{REG,DM,P}$ and $PM25_{REG,DM,S}$), and a line-source dispersion model local-scale field for primary roadway contributions ($PM25_{LOC,DM,P}$). To fuse the components, the dispersion modeling components are first summed:

$$PM25_{DM}(x,y) = PM25_{REG,DM,P} + PM25_{REG,DM,S} + PM25_{LOC,DM,P}(x,y).$$

The spatial average of this field over the sub-region, $\langle PM25_{DM} \rangle$, is then calculated. The final downscaled field is then constructed by scaling the dispersion model field so that the average matches the satellite-derived value:

$$PM25(x,y) = PM25_{DM}(x,y) \times [PM25_{SAT} / \langle PM25_{DM} \rangle]$$

The satellite-derived value, $PM25_{SAT}$, therefore provides the overall average value and the dispersion modeling, $PM25_{DM}$, the spatial distribution within the sub-region, distinguishing between PM2.5 entering the area from upwind ($PM25_{REG,DM,P}$ and $PM25_{REG,DM,S}$, assumed spatially uniform within the grid) and that produced locally ($PM25_{LOC,DM,P}$). Currently, we set the secondary regional PM2.5 to a multiple of the primary regional PM2.5 calculated by the LBM, i.e. $PM25_{REG,DM,S} = \alpha PM25_{REG,DM,P}$, with $\alpha = 2$ is currently set based on a rough ratio of inorganic secondary aerosol to vehicular and biomass organic components in analyses of speciated PM2.5 monitor data reported in previous studies (e.g. Hasheminassab et al. 2014). Work is ongoing to improve the secondary aerosol representation, considering both better empirical determinations and/or a physically-based parameterization built directly into LBM similar in design to what currently exists in the model for ozone chemistry (see Pournazeri et al. 2014).

3. SATELLITE-INFORMED PM2.5 FIELDS

We generated continuous spatial surfaces of daily PM2.5 on a 3-km grid for California for year 2016 by merging daily PM2.5 measurements from the U.S. EPA Air Quality System (AQS), AOD measurements from the MODIS instrument onboard the Aqua earth-orbiting satellite, and meteorological data from NLDAS-2. Leveraging MODIS-derived data to complement EPA ground observation data, we estimated the daily PM2.5 concentration fields by combining the data sources using multi-regression modeling, a B-spline smoothing model, a quality control procedure for the EPA AQS data and a bias adjustment procedure for MODIS/AOD-derived PM2.5 data. These algorithms are described in detail in Al-Hamdan et al. (2009, 2014). The output of the surfacing algorithm are continuous spatial surfaces of daily PM2.5 on a 3-km grid for California. Merging AOD with surface observations of PM2.5 not only provides a more complete daily representation of PM2.5 than either dataset alone would allow, but also reduces the errors in the PM2.5-estimated surfaces (Al-Hamdan et al., 2009, 2014). PM2.5 surfaces over California based on combined daily AOD retrievals from MODIS Aqua Dark Target (3-km) and AQS data for selected days as well as the annual mean surface of 2016 are shown in Figure 2.

The multi-regression model relating AOD to PM2.5 includes wind speed, wind direction, precipitation, temperature and relative humidity as co-variates. Fields for these variables are taken from the NLDAS-2 reanalysis. PM2.5 fields are then constructed by combining the multi-variate regression with the other parts of the procedure (B-spline smoothing, bias adjustment and QA/QC of EPA data) to produce the final satellite-derived PM2.5 fields as explained above and shown in Figures 2 and 3. Table 1 shows the Pearson correlation coefficient for single (AOD alone) versus multi-variable (AOD plus meteorological co-variates) regressions of Dark Target 3-km AOD to PM2.5. As seen, the addition of meteorological variables increases the coefficients in all cases except warm season (April through September) in Northern California. The values for single variate coefficients are similar to those found in other AOD to PM2.5 regressions for California (e.g. Lee et al. 2016).

Table 1: Correlation (R-values) between AOD and EPA AQS PM2.5 for single variable and multivariable regressions for three regions in California.

	NCAL		CCAL		SCAL	
Period/Season	Single Variable	Multi-variable	Single Variable	Multi -variable	Single Variable	Multi -variable
Full Year	0.21	0.27	0.18	0.40	0.26	0.37
Warm Season	0.34	0.23	0.30	0.35	0.25	0.44
Cold Season	0.16	0.39	0.35	0.52	0.35	0.36

*Development of single variable regression model utilizes data from 2003 – 2013, correlations in table are for data from 2014 – 2016. Development of multi-variable regression model utilizes data from 2010 – 2012, correlations in table are for data from 2013.

Validation analysis comparing final estimated satellite-informed PM2.5 grid cell values to the observed daily PM2.5 from non-FRM stations in the cell was performed. Non-FRM stations data are not used in constructing the PM2.5 surfaces, and therefore provide independent measurements for evaluation. Examples of such validation analysis plots and a map of the FRM and non-FRM validation site locations in the Southern California region are shown in Figure 3. This validation analysis demonstrated generally high correlations (r = 0.7 to 0.9). The improved correlations compared to values in Table 1

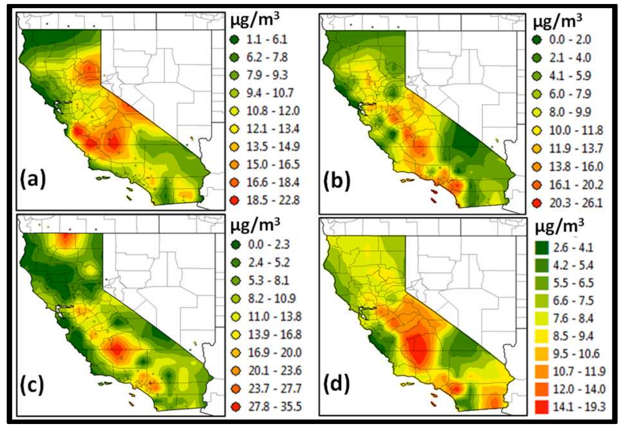


Figure 2. PM2.5 surfaces for selected days and annually for 2016 based on daily AOD retrievals from MODIS Aqua Dark Target (3-km). July 26, (a), September 18 (b), December 8 (c), and annual mean (d). Based on procedure of Al-Hamdan et al. (2009, 2014).

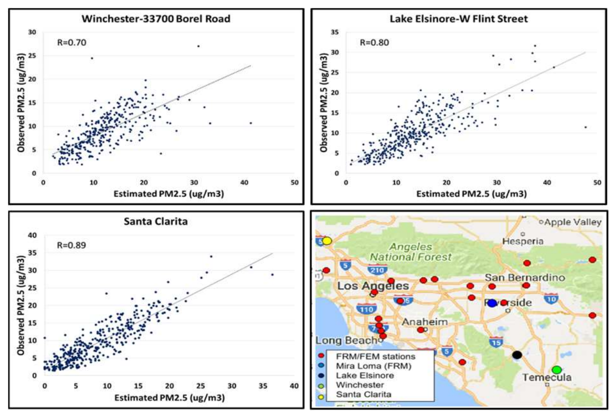


Figure 3. Validation analysis comparing satellite-informed PM2.5 grid cell values to observed daily PM2.5 from non-FRM stations in the cell. Map of site locations on lower right.

indicate the enhanced accuracy of the combined methodology (utilizing B-spline smoothing, bias adjustment and QA/QC of EPA data) over that from regression analysis alone.

4. DISPERSION MODEL DOWNSCALING

Downscaling a sub-region within the satellite-derived PM2.5 field is carried out by merging the PM2.5 values of the satellite field sub-region with output from two dispersion models: the LBM model and a fine-scale line-source model. LBM simulates the regional background PM2.5 entering the sub-region, whereas the fine-scale model simulates the local variations internal to the region. At this time, the fine-scale model is designed for road-traffic emissions, however other internal sources can be included by utilizing other available fine-scale dispersion models (e.g. AERMOD for point or area sources). We demonstrate the downscaling procedure for a single-day at a particular sub-region: December 8, 2016 over a 15 x 15-km area around the Mira Loma monitor in Riverside County (Southern CA).

a. LBM

LBM (Pournazeri et al. 2014) simulates regional primary pollution due to ground-level sources by advecting a column of air backwards in time, injecting pollutant into the column through surface emission flux along the backwards trajectory path. In its implementation here, LBM assumes pollution is mixed evenly through the depth of the column (assigned to the PBL depth) over all hours of backwards travel. In its full implementation, a shallower mixing depth is assumed for the first hour of backwards travel to account for incomplete vertical mixing of emissions emitted close to the receptor. LBM also enables chemical transformation of NOx and VOCs to ozone through a parameterized chemical reaction set. The extension of the parameterization to PM2.5 is ongoing work.

LBM simulated backwards trajectories of 0700 LST air arriving at Mira Loma on December 8, 2016 are shown in Figure 4. Shown are backwards trajectories for a central trajectory and two others either side of this from altering the input winds by plus/minus five degrees. The trajectories are over twelve backwards integration hours at an hourly time step using HRRR model input winds. These are overlaid on the annual average primary emission field for year 2012 supplied to us by SCAQMD so that the coupling of trajectory position with emissions is explicitly seen. While only the annual emission field is shown on the plot, the computation accounts for hourly emission variations by implementing a domain-averaged diurnal emission profile calculated from the supplied input gridded field.

LBM calculated PM2.5 concentrations at 0700 LST December 8, 2016 along with the daily average for the day computed from hourly calculations at three-hour increments (01, 04, 07, 10, 13, 16, 19 and 22 LST) are shown in Table 2. As seen, there is modest sensitivity of LBM results to the value of back-trajectory time step, and very small difference in results for central versus results calculating from the weighted-average of the central and two side trajectories. We take the "central trajectory, Δt = 60 min" daily-average concentration for further illustration of the

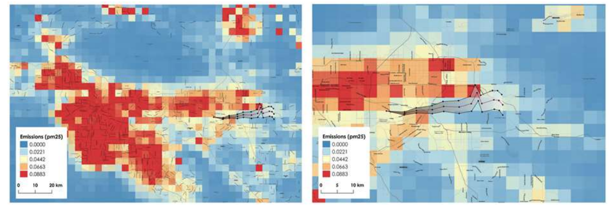


Figure 4: Lagrangian Background Model (LBM) back trajectories of air arriving at Mira Loma at 0700 LST on December 8, 2016 overlaid on the annual average primary emission field ($\mu g m^{-2} s^{-1}$). Simulations driven by HRRR model meteorology downloaded from http://hrrr.chpc.utah.edu/. Shown are central trajectory (dashed) and "side" trajectories (solid) based on altering input winds plus/minus five degrees.

Table 2: LBM Concentrations ($\mu g/m^3$)
December 8, 2016 at Mira Loma*

	Central Trajectory		Weighted-Average**	
Period	Δt = 60	Δt = 10	Δt = 60	Δt = 10
0700 LST	9.4	7.8	9.6	7.9
Daily-avg	5.6	5.6	5.7	5.7

* Δt is value of backwards integration time step used in simulation in minutes.
** Weights of 0.5, 0.25 and 0.25 assumed for central and two side trajectories, respectively. See Figure 5.

downscaling procedure: $PM25_{REG,DM,P}$ = 5.6 $\mu g/m^3$. Multiplying this by two to account for the secondary portion (see Section 2) gives $PM25_{REG,DM,S}$ = 11.2 $\mu g/m^3$. These two values are added to give 16.8 $\mu g/m^3$, which is in turn passed to the line-source model to complete the downscaling.

b. Line-Source Model

The primary local portion of PM2.5 ($PM25_{LOC,DM,P}$) is calculated using the line-source dispersion model of Venkatram and Horst (2006). The model is applied to downscale the satellite-derived PM2.5 concentration within a 15 by 15 km sub-region around Mira Loma on December 8, 2016. Downscaled fields to a grid of 93.75 m are generated (3 km divided into 32 sub-grids).

The major California freeways and traffic flow implemented into the fine-scale model are from CALTRANS traffic sensor data (http://www.dot.ca.gov/hq/tsip/gis/datalibrary/#Highway). We utilize the AADT reported for light-duty vehicles and trucks in separate line-source calculations for each vehicle type, which are then summed to give the total field. Freeway line-sources are defined as line segments connecting traffic sensors. Figure 6 illustrates the locations of

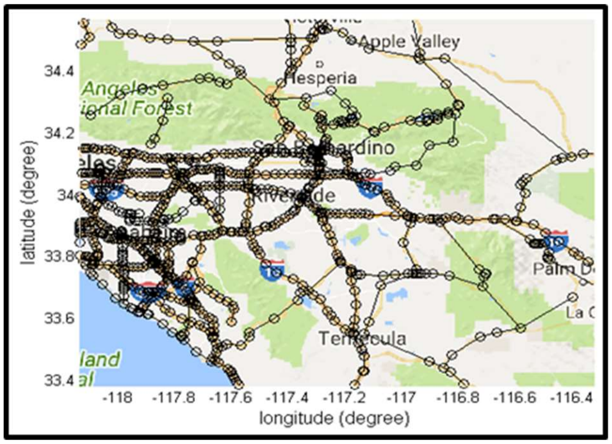


Figure 5: Major roadways implemented into the line-source dispersion model for the area around Riverside Country (Southern California). Circles indicate locations of mileposts where AADT is measured.

traffic sensors contained in the CALTRANS data shapefile in the Riverside County area.

The emission factors of light-duty vehicles and trucks are calculated using EMFAC2014 (https://www.arb.ca.gov/msei/categories.htm) as the average of emission factors for the various sub-categories of these vehicle types within the EMFAC2014 output, weighted by vehicles-mile-traveled for each vehicle sub-category. The calculated PM2.5 emission factors for light-duty vehicles and trucks are 0.0152 gr/km/vehicle and 0.1275 gr/km/vehicle, respectively.

Once line sources based on freeway segments are determined, the concentration contribution of each line source segment at the 93.75 m spaced receptors were simulated. Meteorological inputs were from the HRRR grid value around Mira Loma. For this illustration, we carried out simulations at the same three hourly increments as used for LBM results (Table 1), and averaged to arrive at the daily-average primary local-scale PM2.5 concentration field, $PM25_{LOC,DM,P}(x,y)$.

The satellite-derived daily PM2.5 concentration averaged to the sub-region on December 8, 2016 is 19.5 μg/m^3. This satellite-derived value, the LBM value of 16.8 μg/m^3, and the fine-scale concentration field are then merged using the equations in section 2. The resulting downscaled field is shown in Figure 6**.** As indicated, most of the local variation is confined to very close to the freeway.

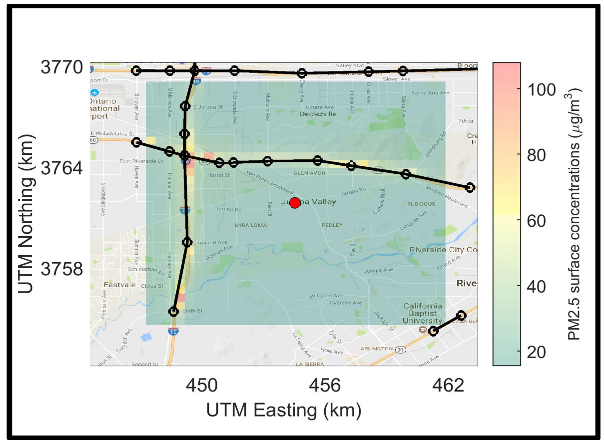
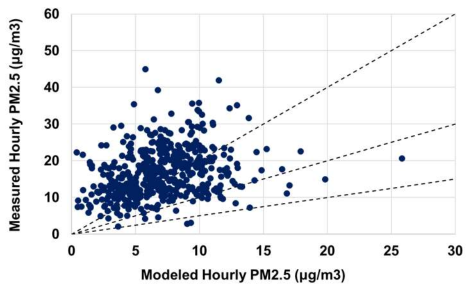


Figure 6: Downscaled PM2.5 concentrations within a 15 by 15 km sub-area of the satellite-derived field around Mira Loma CA (red dot) on December 8, 2016. Area average of field is 19.5 μg/m3, matching the satellite-derived value.

5. IMPLEMENTATION OF HRRR to LBM

In its original implementation, LBM is driven by observed meteorology from nearest monitor locations. Here, we extend its application to utilize gridded model winds from the initial hour of hourly real-time NOAA HRRR model output (see Section 2). Comparison of summer 2017 wind roses from the HRRR model with measured climatological summer-month period values (other years) at several Southern California sites were largely consistent (see https://www.cmascenter.org/conference//2017/slides/freedman_satellite-dispersion_2017.pdf).

The LBM with HRRR meteorology was evaluated over 60 simulation days from June 2 – September 21, 2017 for PM2.5 at Mira Loma. Eight hours per day at even three-hour increments were simulated for a total of 480 hours of simulations. Comparison of model concentrations with observations are shown in Figure 7. Hourly results show considerable scatter, while daily averages more coherent and generally increase as observations increase. Daily averages underestimate observations on average, as expected since the modeling does not account for secondary PM2.5.



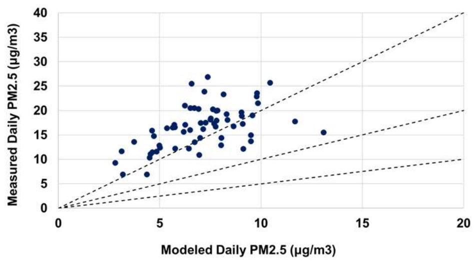


Figure 7: Hourly (top panel) and daily-average (bottom panel) LBM model primary PM2.5 concentrations driven by the NOAA HRRR model meteorological inputs vs. measurements at Mira Loma. Runs are for June 2 – Sept 20, 2017 at three-hourly intervals (1,4,7,10,13,16,19,22 local time, n = 480 hours, 60 days). For each run, back trajectories were calculated twelve hours at a one-hour time step.

REFERENCES

Al-Hamdan, M. Z. and co-authors (2014). Environmental Public Health Applications Using Remotely Sensed Data. *Geocarto International*, 29:85-98.

Al-Hamdan, M. Z. and co-authors (2009): Methods for Characterizing Fine Particulate Matter Using Ground Observations and Satellite Remote-Sensing Data: Potential Use for Environmental Public Health Surveillance. *Journal of the Air & Waste Management Association*, 59, 865–881.

Bates, J. and co-authors (2017): Development of Two Model Fusion Techniques Utilizing CMAQ and R-LINE to Obtain NOx, CO and PM2.5 Concentrations at 250m Resolution over Atlanta, GA. Presented at the *16th Annual CMAS Conference, Chapel Hill, NC, October 23-25, 2017*.

Batterman, S. and co-authors (2015): High Resolution Spatial and Temporal Mapping of Traffic-Related Air Pollutants. *International Journal of Environmental Research and Public Health*, 660 12, 3646–3666., doi:10.3390/ijerph120403646.

Hasheminassab, S. and co-authors (2014): Spatial and temporal variability of sources of ambient fine particulate matter (PM2.5) in California. *Atmos. Chem. Phys.*, 14, 12085–12097, doi:10.5194/acp-14-12085-2014

Hood, C. and co-authors (2017): Integrating regional and local modeling to create a high-resolution air quality forecasting system for Hong Kong. Presented at the *16th Annual CMAS Conference, Chapel Hill, NC, October 23-25, 2017*.

Lee, H. J., R. B. Chatfield, and A. W. Strawa, (2016): Enhancing the Applicability of Satellite Remote Sensing for PM2.5 Estimation Using MODIS Deep Blue AOD and Land Use Regression in California, United States. *Environmental Science & Technology*, 6546–6555, doi:10.1021/acs.est.6b01438.

Pournazeri, S., et al. (2014), A computationally efficient model for estimating background concentrations of NOx, NO2, and O3. *Environmental Modelling and Software*, 52, 19–37.

Venkatram, A. and W. Horst (2006): Approximating dispersion from a finite line source. *Atmospheric Environment,* 40. 2401-2408.

van Donkellar, A., and co-authors (2015): High-Resolution Satellite-Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over North America, *Environmental Science & Technology*, **49**, 10482 – 10491, doi: 10.1021/acs.est.5b02076

Zhai, X and co-authors (2016): Calibrating R-LINE model results with observational data to develop annual mobile source air pollutant fields at fine spatial resolution: Application in Atlanta. *Atmospheric Environment*, 147, 446 – 457.

ACKNOWLEDGEMENTS

The research was supported by the NASA Applied Science Program Health and Air Quality Applied Science Team (HAQAST, https://haqast.org, Award #NNX16AQ91G), and NASA-MIRO Center for Applied Atmospheric Research and Education (CAARE, Award #NNX15AQ02A). The authors acknowledge the provision of gridded emission fields from the South Coast Air Quality Management District and Bay Area Air Quality Management District (BAAQMD).