Finding High-Pollution Hot-Spots in Areas of Generally low PM2.5 — The Bay Area Example

Robert B Chatfield, Robert F Esswein, Meytar Sorek-Hamer

NASA Ames Research Center, Atmospheric Science Branch

Frank Freedman, Minghui Diao, Ana Rivera

San Jose State University, Depts. of Meteorology and Geography

Saffet Tanrikulu

Bay Area Air Quality Management District

At the HAQAST2 Meeting in February 2017, co-author Saffet Tanrikulu challenged the NASA community to show how satellite data could aid him in assessing air pollution for the Bay Area Air Quality Management District (BAAQMD).

It's a tough challenge: the area is usually very clean, but with occasional violations of the 35 $\mu g\ m^{-3}$ standard for PM2.5 respirable aerosol, especially in winter:

Mixed layers vary on a few-km scale, responding to inflow from the ocean and the Central Valley, land-sea contrasts in stability to mixing, and multiple sub-basins defined by topography. We had considerable success in producing daily maps for the wintertime, but here unmodeled rms variability of ca. 5 µm proved not tolerable.

Instead, we looked to concentrate on highest concentrations in individual few-km areas. The PM2.5 monitoring-station data suggested separate regional and local influences. We expected that local influences might not be captured by the station.

Our approach is to provide informative maps based on satellite data. Also we aimed to separate out "transport error" and "source/composition error," using static views without transport or sources. This approach aims to test and improve, not replace 3-d source-driven models.

The aerosol optical thickness (AOT or   ) was obtained from the MODIS ftp site and we thank Alexi Lyapustin and Yujie Wang for their aid in producing and refining this product. The data is at 1 km resolution.

$$\tau_{Ext}$$

We aimed to quantify regional influences by simplifying the physics that describes the AOT-to-PM2.5 relationship in a near-surface mixed layer:

$$PM2.5 = \frac{\tau_{Ext}}{\Delta z_{ML} \cdot f\left(_{Composition,\ RH}\right)}$$

… where we divide the aerosol optical depth by the mixed-layer height (preferably also the mixing-layer height) and also by a function that describes the expansion of particles from PM monitor conditions to conditions appropriate for aerosol particles at various temperatures and relative humidities throughout the mixed layer.

Daily Relationships and a Mixed Effect Model

We use a daily calibration technique, where we seek the best-fitting function between AOT, extinction optical depth, (from MAIAC MODIS archive) and estimated for each day (i),

….however we assume uniform for that day at all locations. For the estimate we use all reporting PM2.5 stations (s).

$$\mathrm{PM2.5}_{i\,s} = f_i\left(\tau_{Ext\ i\,s}\right), \qquad \tau_{Ext\ i\,s} = \mathrm{AOT}_{i\,s}$$

instance (date)   *Station*

and we extend that daily f_i relation for all locations where there is MAIAC data. This creates maps (arrays) of estimated PM2.5 for the whole region.

$$PM2.5_{i\,s} = a \cdot \tau_{Ext_{is}} + c + \left(\alpha_i + \beta_i \cdot \tau_{Ext_{is}}\right) + \varepsilon_{is}$$

This expresses a uniform fixed effect, linear in AOT, and modifying linear relationships for each day, called "random effects" within a "mixed effects model" (MEM). Maps can still vary as MAIAC AOT varies. The robustness of the estimate are checked by leave-out-one cross validation.

The functions f_i vary significantly from day to day.

- *Varying slopes (α_i) allow for different mixing depths, particle composition, and relative humidity effects*
- *Varying intercepts (β_i) allow for detached aerosol layers higher in the troposphere, as long as they are approximately uniform, and some other variations.*
- *We're not sure why the variations are so high.*

Estimation of High Quantiles of Estimated PM2.5

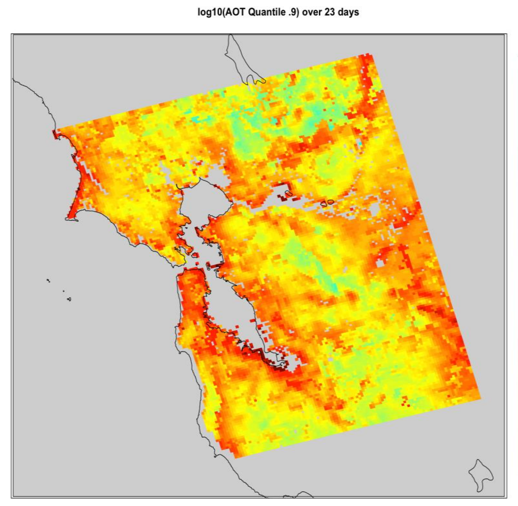
The linear relationships f_i perform well in providing regional patterns, but there is still 30–40% of rms residual error observed where we have checks, at the measurement stations.

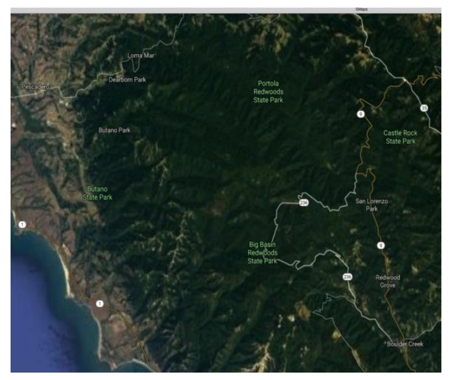
We had expected correlated variations between neighboring stations, but in fact they appeared more random. If there is additional AOT and so attributed PM2.5 in certain areas, they should appear in the maps. The magnitude of mapped variations will be underestimated, since the composition, relative humidity, and mixed-layer depth are not in the estimates.

Expecting very local variations, we took 23 days of MAIAC data, 23 arrays, calculated the associated PM2.5, and extracted the median, 75th percentile, and the 90th percentile at each point.

We then made maps of the 75th percentile and the 90th percentile, These are shown in a variety of ways in the images on the right. The values are shown in the maps as $\mu g\ m^{-3}$ but they should not be considered quantitative; the range should of values seen in the maps is no doubt diminished. Still, the pattern of variation appeared instructive.

Unweighted AOT values, 90th percentile values are less informative

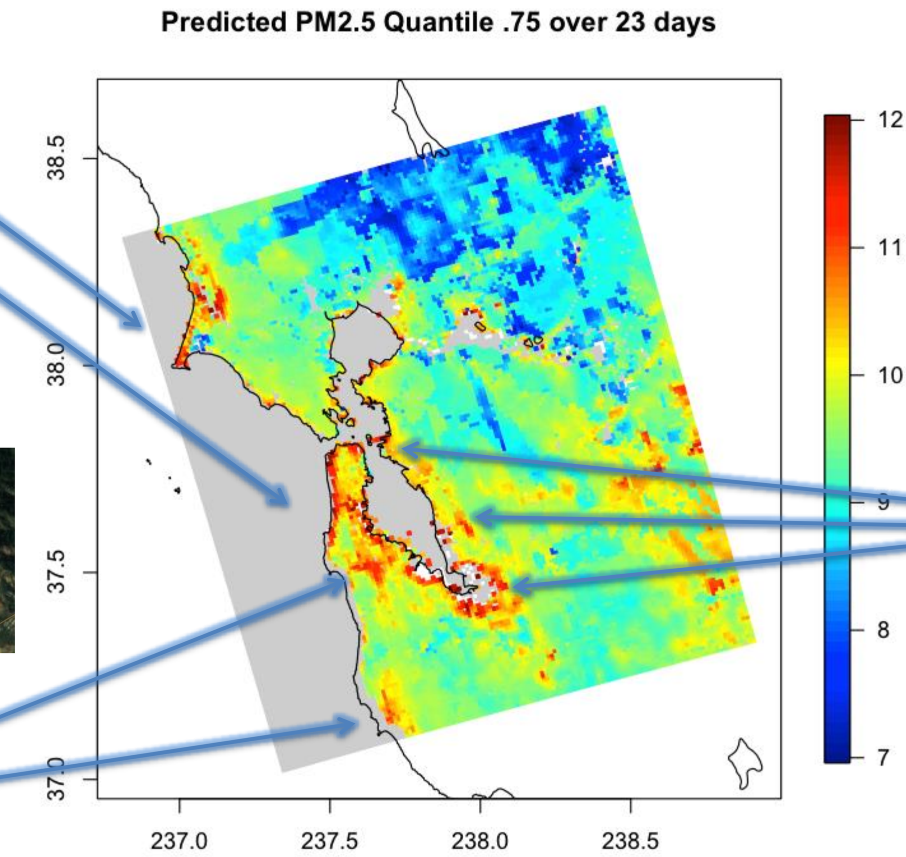




Please Examine the 70th and 90th percentile plots on the right.

Identified water areas (in the MAIAC data) have been removed. MAIAC treats water areas with a separate algorithm; water coloration in the Bay Area is complex. Some areas (salt flats in the South Bay, shoreline regions west of Vallejo, and show as very high PM2.5.

Beaches; sand? Surf?

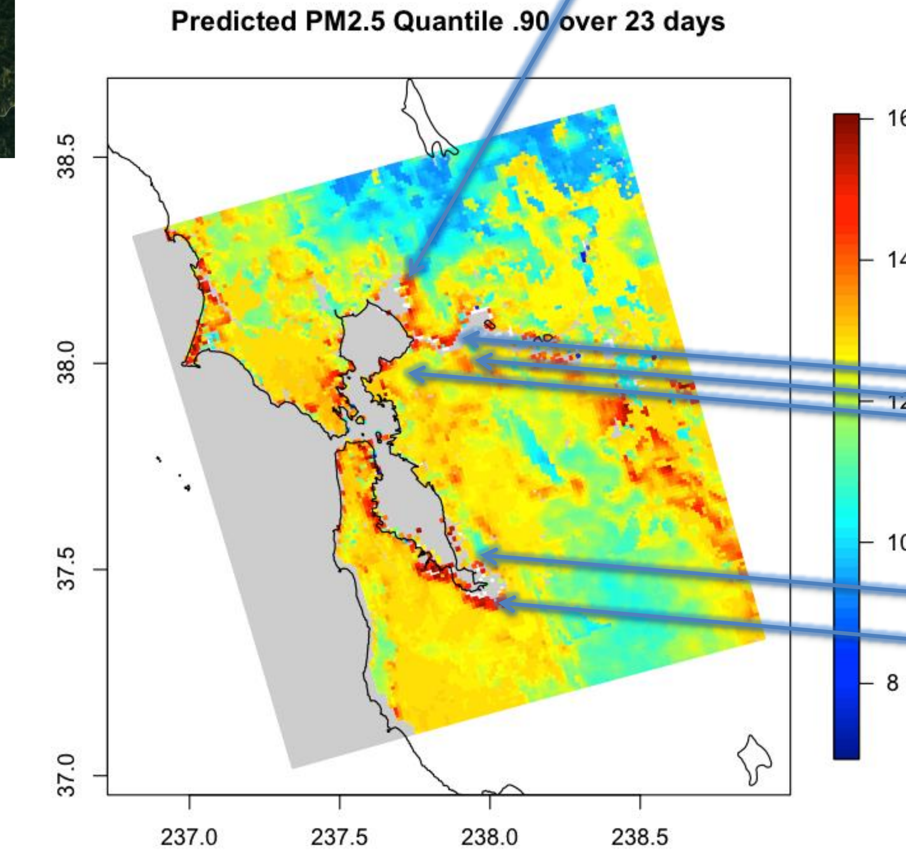


traffic jams??

Mystery! open land fill, quarry??

Open fires in Big Basin and Butano State Parks?

Napa Junction, Squab trucking transport hub



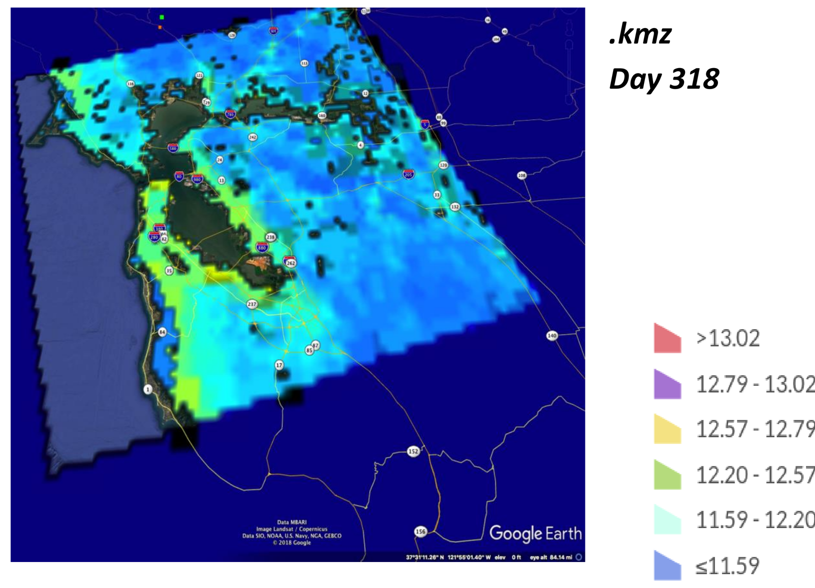
Valero, Tesoro, Phillips refineries

marshes salt evaporators,

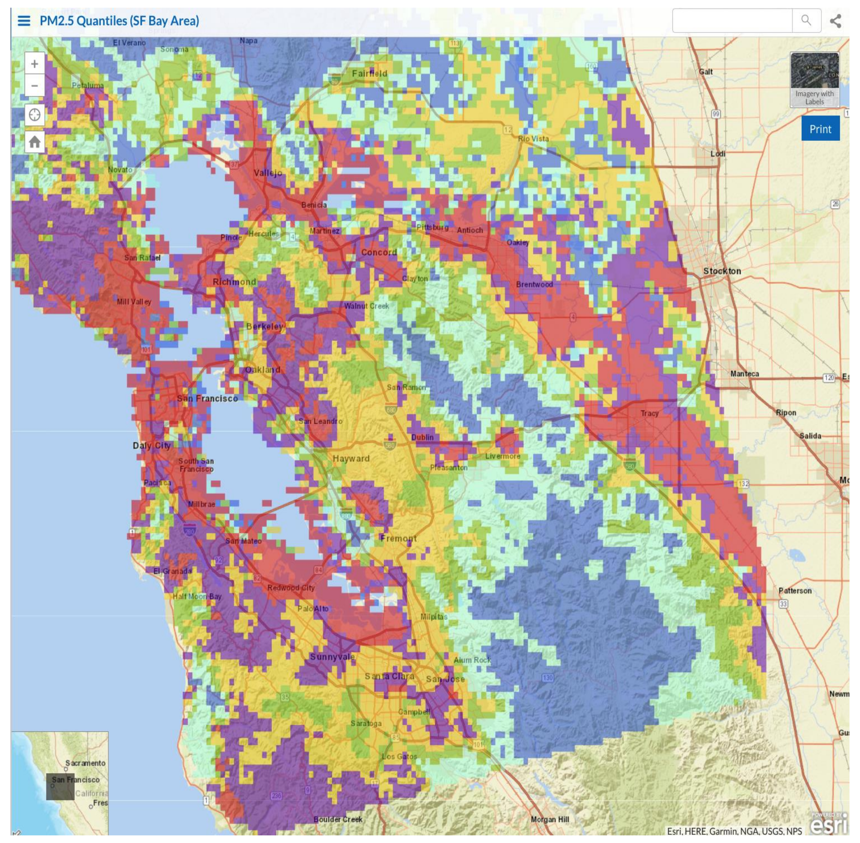
highway, urban

Data is available in .kmz Google Earth and ArcGIS formats. Top image: November 14, 2016. This shows native 1-km resolution

Lower image gives GIS information with a layer showing the 90th percentile.



.kmz Day 318



GIS on remapped grid

Conclusions: *Even in regions where PM2.5 estimated from MAIAC AOT which yield many results at noise level of 5 – 8 $\mu g\ m^{-3}$, the production of daily arrays and maps of PM2.5 can produce useful results. Median, 75th percentile and 90th percentile maps constructed from the daily maps were our tools. These suggest areas of poor air quality; many seem to correspond to likely high emissions sources. There is some rough comparability of the levels that they reach in these regions based on these maps, though they are not considered more than approximately quantitative. Coastal water bodies can give false highs.*

This algorithm uses the idea that regional levels of PM2.5 affecting most monitors can be estimated, but require a daily f_i relation relating AOT to PM2.5. High AOT regions then map with that f_i relation to estimate local PM2.5. High quantile maps drawn from daily maps will then be usefully estimated.

There are notable limits to the method we used. Sufficiency constancy of daily f_i, needs to be evaluated. Fine-scale estimation of mixing heights could help, but transport models are needed to establishment of equivalent mixed-layer depths when there are likely elevated layers. In turn, our maps can improve 3d transport models

This study used only a pilot set of 23 winter days, including enough high-particulate days to establish parameters well and also low and moderate days. High PM2.5 levels likely underestimated. Studies which study larger numbers of days, and also future deployment of relocatable monitors may help check these results.

Progress Report … please contact for finished products and conclusions:

Robert.B.Chatfield@nasa.gov

This work was encouraged and aided by Saffet Tanrikulu and his staff, but of course it cannot represent understandings of the BAAQMD.

We thank the Health and Air Quality Program and John Haynes for encouragement and support of this activity to this point.