Responsive high-resolution air quality index mapping using model, regulatory monitor, and sensor data in real-time

Nico Schulte, Xiang Li, Jo Kay Ghosh, Philip M Fine and Scott A Epstein
South Coast Air Quality Management District, Diamond Bar, CA, United States of America
E-mail: nschulte@aqmd.gov

Keywords: air quality index, AQI, wildfire, purpleair, low-cost air quality sensor, air quality communication, Kriging

Abstract

Interpolated regulatory monitor measurements that are used to calculate the Air Quality Index (AQI) have low spatial resolution and are less accurate where regulatory monitors are sparsely distributed, especially during episodic events such as wildfires and dust storms. In this paper, an AQI map that is more responsive and accurate than current formulations with 5 km spatial and hourly temporal resolution is constructed for Southern California using interpolated concentration fields. The fine particle mass (PM$_{2.5}$) fields are calculated by combining regulatory-grade monitored concentrations, low-cost sensor measurements, and Community Multiscale Air Quality Model forecasts, using Residual Kriging (RK) interpolation to account for the uncertainty of each input component. The Ozone fields are calculated by combining regulatory monitor and model forecasted concentrations using RK. The interpolated concentration fields have root mean square interpolation errors (RMSE) of 5.59 $\mu$g m$^{-3}$ (PM$_{2.5}$) and 6.66 ppb (Ozone), smaller than corresponding RMSE of current inverse distance weighting methods used by the U.S. Environmental Protection Agency AirNOW system, and a previous interpolation method used by the South Coast Air Quality Management District. The inclusion of low-cost sensor measurements makes the AQI more responsive to wildfire events, with a PM$_{2.5}$ RMSE of 7.73 $\mu$g m$^{-3}$, a significant improvement from the other interpolation methods. The mass of particles with an aerodynamic diameter smaller than 10 micrometers (PM$_{10}$) calculated using natural neighbor interpolation of the difference, PM$_{10}$—PM$_{2.5}$, and carbon monoxide and nitrogen dioxide fields calculated using natural neighbor interpolation, have small but usually not statistically significant reductions of RMSE relative to the other interpolation methods. Quality control and correction procedures applied to the low-cost sensor measurements are described. Display of the blended AQI data on an interactive map in real time provides an important tool for the public to minimize their exposure to poor air quality in one of the most polluted regions in the United States.

1. Introduction

The Air Quality Index (AQI)\(^1\) (U.S. EPA 2018) remains one of the most important tools to communicate air pollution information to the public. School districts, public health officials, employers, and others use the AQI to make decisions about mitigating exposures when outdoor pollution levels are elevated. Air quality management agencies and other organizations use a wide variety of approaches to model and display the AQI based on data from regulatory air pollution monitors.\(^2\) One of the main
challenges is to display accurate information in areas where regulatory monitors are sparsely distributed. A second challenge is to ensure that the AQI displayed is responsive to short-term episodic events, such as large wildfires and dust storms, so that residents can use timely information to reduce their exposure. This paper presents a blended interpolation method to disseminate high-resolution and temporally responsive AQI maps for the South Coast Air Quality Management District (South Coast AQMD) jurisdiction in Southern California. This jurisdiction includes the 17 million people in the greater Los Angeles area, which can exceed federal standards for either ozone ($O_3$) or fine particle mass ($PM_{2.5}$) on over 150 d per year. Additionally, the region is prone to large wildfires and includes the Coachella Valley, where dust storms can cause very high particle concentrations. To our knowledge, this is the first AQI map to bring together information from regulatory monitors, low-cost air pollution sensors, and air quality models in a blended interpolation approach that maximizes the accuracy of the data and provides spatially detailed and temporally resolved AQI information.

In recent years, users’ demand for hyperlocal air pollution measurements have contributed to a proliferation of low-cost particulate matter (PM) sensors. In 2019 there were at least 653 outdoor PurpleAir (PurpleAir LLC 2020) sensors with publicly available PM mass concentration (i.e. $PM_{1.0}$, $PM_{2.5}$, $PM_{10}$) data in the South Coast AQMD jurisdiction and nearby areas. These sensors help to provide data in regions between regulatory monitors, which are sparsely distributed to cover a large geographical area. For example, the nearest regulatory monitor in sparsely populated areas of the South Coast AQMD jurisdiction can be up to 37 km and 48 km away for $O_3$ and $PM_{2.5}$, respectively. Although sensors are appealing because they provide very localized air pollution information and are inexpensive, uncalibrated individual sensor measurements should not be used to determine pollution concentration levels because these sensors are less accurate than regulatory methods and may not be sited to represent community-wide exposures. Websites report $PM_{2.5}$ sensor measurements using short averaging times, which should not be compared to the AQI for public health decision-making. In contrast, the AQI reported through agency websites such as airnow.gov (U.S. Environmental Protection Agency website) (U.S. EPA 2020a) and aqmd.gov (South Coast AQMD website) (South Coast AQMD 2020b) is derived using the NowCast method (U.S. EPA 2019, 2020c), a weighted moving average of the concentrations designed to approximate the averaging times of health-based ambient standards. While the accuracy of a single instantaneous low-cost sensor measurement may not be accurate enough for decision-making, calibration along with spatial and temporal averaging of multiple sensors provides useful information.

Interpolation is commonly used to estimate concentrations between monitors. Inverse distance weighting interpolation (IDW) is used to report AQI on airnow.gov, and Kriging, a common geostatistical interpolation method, is used to estimate concentrations for studies of health effects (Jerrett et al 2005). However, Kriging is not typically used to estimate concentrations for the AQI. The South Coast AQMD reports AQI using ‘proxy’ interpolation which is based on a lookup table that maps measurements at stations to defined geographic areas. Interpolation of monitor data provides the spatial resolution that users demand, but the accuracy depends on the density of the monitoring network and whether secondary pollutants are a key component of the species concentrations (Yu et al 2018). The proxy interpolation may not adequately represent the spatial variation of $PM_{2.5}$ in all parts of the South Coast AQMD jurisdiction because of the large distance to the nearest regulatory monitor relative to the spatial scale of $PM_{2.5}$ concentration variation. Hybrid methods that combine observed with model estimated concentrations can overcome this limitation (Yu et al 2018), and may be especially useful for primary pollutants such as $PM_{2.5}$ from wildfire smoke.

This paper uses a hybrid of regulatory monitor data, low-cost sensor data (subsequently referred to as ‘sensor data’), and model data to improve the spatial resolution of concentrations and AQI estimates. This interpolation is subsequently referred to as ‘blended’ interpolation. National Air Quality Forecast Capability Community Multiscale Air Quality model (CMAQ) forecasts issued by the National Oceanic and Atmospheric Administration (NOAA) (Lee et al 2017) are used for this paper. Our blended interpolation method is a hybrid approach that exploits the advantages of each data source, listed in table 1, and overcomes limitations in the accuracy of sensor and model data. This technique provides

---

3 Low-cost sensors usually have a linear response against reference instruments in laboratory tests but their response varies under ambient conditions (Morawska et al 2018). $PM_{2.5}$ concentrations measured by Original Manufacturer Equipment (OEM) sensors used in the PurpleAir had linear correlation coefficients ($R^2$) between 0.69–0.99 with the Federal Equivalent Method (FEM) GRIMM model 1.109 instrument in a laboratory study using 10-minute averages, $R^2$ between 0.83–0.93 with the GRIMM in an ambient study using 1-hour averages, and $R^2$ of 0.88 with Federal Reference Method (FRM) 24-hour integrated $PM_{2.5}$ measurements, although the sensor response varied with particle properties (Kelly et al 2017). The hourly PurpleAir $PM_{2.5}$ measurements had $R^2$ of 0.95 with those of a Met One Beta Attenuation Monitor (BAM) FEM instrument in an ambient study and was affected by systematic errors and impacts of relative humidity (Feenstra et al 2019). Systematic errors should be corrected to compare low-cost sensors measurements with FRM or FEM. Low-cost sensors should be tested and calibrated in the conditions in which they are used (Morawska et al 2018).
Table 1. Advantages and disadvantages of each data source used in the hybrid method for constructing AQI maps.

| Data source       | Accuracy               | Spatial coverage              | Pollutants measured or modeled | Usage in blended interpolation                                                                 |
|-------------------|------------------------|-------------------------------|--------------------------------|-------------------------------------------------------------------------------------------------|
| Regulatory monitors | High                   | Limited                       | All criteria pollutants       | Use monitoring data as gold standard where available                                             |
| Low-cost sensors  | Less accurate for absolute concentrations but reasonable accuracy for spatial trends | Good coverage for PM$_{2.5}$ only, for the South Coast AQMD jurisdiction | Only PM$_{2.5}$ readily available | Use calibrated sensor data in aggregate to improve spatial coverage                                |
| CMAQ model        | Reasonably skilled at predicting spatial trends, but may not accurately capture absolute concentrations | Uniform and highly resolved spatial coverage | Only PM$_{2.5}$ and O$_3$ readily available | Use modeled data to fill in areas without measurements, for PM$_{2.5}$ and O$_3$               |

a useful interpretation of sensor data that is consistent with measurement uncertainties leading to increased accuracy over other interpolation methods. This paper evaluates the accuracy of the resulting concentrations and compares the accuracy to proxy interpolation and IDW.

2. Methods

Regulatory monitor data, gridded model data, and PurpleAir sensor data are used in this study. Figure 1 shows the locations of all PurpleAir sensors, continuous regulatory monitoring sites for PM$_{2.5}$ and O$_3$ inside the South Coast AQMD jurisdiction, and the model grid cells. There is more dense coverage of sensors near the coast.

2.1. Interpolation

The core approach of the blended interpolation method is an interpolation of available regulatory monitor and sensor concentration data to a regular grid using model data to fill in the concentrations far from measurements. The model, measurements, quality control of the data, and corrections to the sensor data are described in section 2.2. The interpolation is conducted at the beginning of every hour and results in a gridded concentration field each hour for each of the five AQI pollutants: PM$_{2.5}$, O$_3$, PM$_{10}$, nitrogen dioxide (NO$_2$), and carbon monoxide (CO). After interpolation, an algorithm is applied to calculate the NowCast concentrations and corresponding AQI for PM$_{2.5}$, O$_3$, and PM$_{10}$ for each grid cell. For NO$_2$ and CO, the AQI for each grid cell is calculated based on the hourly concentrations. The reported AQI is represented by the maximum AQI of each of the individual pollutant AQIs.

PM$_{2.5}$ and O$_3$ concentrations are interpolated using Residual Kriging (RK) (Pournazeri et al 2014). RK first calculates the residuals, which are defined as the differences between the measured concentrations (regulatory monitors and sensors) and the corresponding modeled concentrations at the measurement locations. Then the residuals are interpolated to the grid cells without measurements using Simple Kriging (Cressie and Wikle 2015, p. 136). Finally, the interpolated residual field is added to the modeled concentration field, resulting in an estimate of the concentration in every grid cell. Other forms of Kriging are available (Li and Heap 2008, Singh et al 2011) but RK was chosen for simplicity.

Kriging requires a model of the covariance function of the process underlying the observed data. For the Simple Kriging used in this paper, the covariance function is assumed to depend only on the separation distance between two points. The experimental semivariogram is calculated from the residuals (both regulatory monitor and sensor) and fit to several semivariogram models, the model with the smallest mean square error is selected, and the covariance function is calculated using the mathematical relationship $\gamma(h) = C_Y(0) - C_Y(h)$, where $C_Y$ is the covariance and $\gamma$ the semivariogram of the process (Cressie and Wikle 2015, p. 130). The details of how the semivariogram is fit to the data are described in supplementary material section S.4. Kriging distinguishes between the covariance of the process underlying the measured data and the covariance of the measured data itself, because the latter includes a component of variance caused by measurement error. Measurement uncertainty is accounted for by subtracting the standard uncertainty of the difference between two residuals from the experimental semivariogram before fitting the semivariogram model and adding the standard uncertainty of the measurement to the covariance function during the Kriging interpolation.

RK is a convenient framework for blending regulatory monitor data, sensor data, and model data for two reasons. First, the Simple Kriging applied in this paper assumes that residuals follow a Gaussian distribution with a constant mean. Model residuals are likely consistent with this assumption, but the underlying concentration fields are not. Thus, a more complex interpolation scheme would need to be used to directly interpolate concentration data without using the residual
Figure 1. Locations of PM$_{2.5}$ and O$_3$ regulatory monitors, PurpleAir sensors, and model grid cells. For a map of gravimetric PM$_{2.5}$ sampler sites that are used for independent evaluation see supplementary figure S-15. Basemap from Esri (2020a). Sources: Esri, HERE, Garmin, USGS, Intermap, INCREMENT P, NRCan, Esri Japan, METI, Esri China (Hong Kong), Esri Korea, Esri (Thailand), NGCC, (c) OpenStreetMap contributors, and the GIS User Community.

method. Second, Simple Kriging allows the user to account for uncertainty of the concentration measurements.

In the absence of a readily available model, NO$_2$, and CO concentrations are interpolated using the ‘nn’ software package (Sakov 2020) that implements natural neighbor interpolation (Sibson 1981). Natural neighbor interpolation, based only on the geographic location of monitors, partitions the plane into the ‘closest’ regions around each monitor location and then calculates interpolation weights proportional to the area lost by each region when the interpolation point is added to the partitioned plane. For PM$_{10}$, PM$_{coarse} = PM_{10} - PM_{2.5}$ is interpolated using natural neighbors, where PM$_{2.5}$ is the blended field and PM$_{10}$ are the monitored values, and then the interpolated PM$_{coarse}$ is added to the PM$_{2.5}$ field. Since natural neighbor interpolation cannot extrapolate outside the domain where measurements are made, four virtual background data points are included. The four data points, shown in supplemental figure S-17 (available online at stacks.iop.org/ERL/15/1040a7/mmedia), are located at each corner of the domain in sparsely populated areas with concentrations set equal to the minimum of the observed concentrations. This method ensures that extrapolated concentrations approach a small ‘background’ concentration value at the edges of the domain. Concentrations and AQI values are only reported inside the South Coast AQMD jurisdiction, which is at least 250 km from the domain’s edge.

Interpolated concentrations more than 100 km from a valid regulatory monitor observation are removed because these concentrations are associated with higher uncertainty. The methods are designed to be automatically executed; during each hour, automatic scripts download data and calculate the interpolated concentrations for the previous hour. Interpolated concentrations for all hours between June 8, 2018 and March 31, 2020 are used for this study.

2.2. Input data
2.2.1. Regulatory monitor data
Measurements of PM$_{2.5}$, O$_3$, PM$_{10}$, NO$_2$, and CO are made at the locations shown in supplemental figures S-12, S-13, and S-14. Sites are located within the South Coast, Mojave Desert, Salton Sea, and Antelope Valley air basins. Measurements are conducted by the South Coast AQMD, the Imperial County Air Pollution Control District, the Mojave Desert AQMD, the Antelope Valley AQMD, the National Parks Service, and the California Air Resources Board (CARB) and are obtained from airnowtech.org to backfill data to the beginning of the study period and in near real-time, and from a CARB database in the real-time implementation. The dataloggers perform automatic tests on the measured values and may flag values as ‘suspect”—these values are automatically removed from the input data.

Most continuous regulatory monitors in the South Coast AQMD jurisdiction are sited to measure concentrations that are representative of regional concentrations, population exposures, or the highest concentrations within certain parts of the jurisdiction. Regulatory monitoring data collected at the four near-road stations in the South Coast Air Basin are excluded from the AQI calculation because these regulatory monitors are designed to measure local impacts of freeway traffic and are not representative of the grid cell that they reside in. When blending
the model data with the measurements, the standard uncertainty of the monitored concentrations, $\sigma_{\text{e}}$, is set equal to zero. This approach enforces the requirement that the interpolated concentration fields must be equal to the measured concentrations at the regulatory monitor locations. This is consistent with the fact that regulatory monitor data are the highest quality concentration measurements routinely available and represent the best estimate of the true concentration at that location. Setting $\sigma_{\text{e}} = 0$ is a method used for RK and does not imply that measurements have zero uncertainty for other purposes.

2.2.2. Model data
The $\text{PM}_{2.5}$ and $\text{O}_3$ model concentration fields are taken from the National Air Quality Forecast Capability CMAQ forecasts that are issued twice daily at 6 UTC and 12 UTC by NOAA at 12 km resolution (Lee et al. 2017). NOAA adjusts the CMAQ model fields to more closely match $\text{PM}_{2.5}$ and $\text{O}_3$ measurements at regulatory monitors using a method called bias correction (Huang et al. 2017). The NOAA/National Weather Service North American Mesoscale Forecast System (NAM), based on Nonhydrostatic Multiscale Model with Arakawa B-grid staggering (NMMB), provides the predicted meteorological fields. The data are automatically downloaded each day at 11 UTC and 17 UTC and are also backfilled to the beginning of the study period, and the latest available data for the hour of interest are used for the interpolation. The model data is reported on a uniform grid in a lambert conformal projection with cell size 5 km by 5 km and the same grid is used for the interpolation of the concentration data.

2.2.3. Sensor data
We only use $\text{PM}_{2.5}$ data from PurpleAir sensor devices due to their ubiquity in the region, although the interpolation methods can be applied to many types of sensor data. Sensor data is automatically downloaded from www.purpleair.com/json at the beginning of every hour and the data representing the 1-hour average $\text{PM}_{2.5}$ determined using the ‘ambient correction factor’ is extracted. Only sensors labeled as being outdoor are retained in the data set.

A minority of PurpleAir sensors produce poor quality data and are removed from the analysis. The criteria for removing data are described in section S.1 in the supplementary material. In short, sensors that have not passed the removal criteria are identified and removed by comparing simultaneous $\text{PM}_{2.5}$ measurements made by the two sensor channels of each PurpleAir.

Next the PurpleAir measurements are corrected for interference from relative humidity (RH) and measurement bias using calibration and correction curves determined from the collocation of several PurpleAir sensors with beta attenuation monitors (BAM), Met One BAM 1020 Federal Equivalent Method (FEM) at four sites and Thermo Scientific 5014i FEM at two sites. Relative humidity data is obtained from South Coast AQMD measurements, made by Met One 083D and Rotronic HC2-S3 sensors, at the sites shown in the supplemental material. The corrections are based on those developed by (Malings et al. 2020) with some modifications made for this work, and are described in the supplemental material section S.2. Collocated measurements of PurpleAir and BAM that are used to fit the model were made by the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) (South Coast AQMD 2020a) and other organizations.

In certain cases, it is useful to remove sensor data that is heavily impacted by local sources because the data is not representative of concentrations in a 5 km grid cell. However the PurpleAir data is not labeled with information that could be used to determine whether the concentrations are representative of the surrounding area. To reduce the effect of individual sensor measurements on the AQI, the sensor data within each grid cell is averaged. The average concentration in the grid cell is less sensitive to local sources that impact a few sensors, especially when there are many sensors in a grid cell. Sensor data from grid cells with less than three sensors that have passed the removal criteria are removed.

To account for uncertainty of the measurements, an estimate of the standard uncertainty of the grid cell averaged sensor data, $\sigma_{\text{e, sensor}}$, is used in the interpolation. To estimate $\sigma_{\text{e, sensor}}$, the standard uncertainty of the PurpleAir data, $\sigma_{\text{e,PA}}$, is combined with the standard error of the grid cell average:

\[
\sigma_{\text{e, sensor}}^2 = \sigma_{\text{e,PA}}^2 + \frac{\text{var}(\text{C})}{n} + \frac{1}{n^{0.63}} \left( \frac{0.2C^2}{n} - \frac{\text{var}(\text{C})}{n} \right) \tag{1}
\]

where $\sigma_{\text{e,PA}}$ is the component of the standard uncertainty of the PurpleAir caused by precision errors, $\text{var}(\text{C})$ is the sample variance of the PurpleAir measurements in a grid cell, $C$ is the average concentration, and $n$ is the number of sensors. The $\sigma_{\text{e,PA}}$ is estimated from the root mean square error of the correction/calibration curve fit derived from collocations of the PurpleAir with BAM and $\sigma_{\text{e,PA}}$ is estimated by comparing the two sensor channels of all PurpleAir sensors, as described in section S.3 in the supplementary material. The $\sigma_{\text{e,PA}}^2$ is subtracted to avoid double counting the component of uncertainty caused by precision errors in the third and fourth term in equation (1). The third and fourth terms in equation (1) interpolate between the sample variance, and an estimate of the sample variance, $0.2C^2$, where the calculated sample variance is the limit of the interpolation when $n$ is large and the estimated sample variance

...
is the limit of the interpolation when \( n \) is small. This is necessary because the sample variance is not a good estimate of the standard error of the grid cell average when \( n \) is small. The estimate, \( 0.2C^2 \), is determined by examining scatter plots of the sample variance against the squared grid cell average for all the PurpleAir data when \( n \) is greater than 10 (not shown). The power of 0.63 is chosen so that \( \text{var}(C) \) and \( 0.2C^2 \) are weighted equally when \( n = 3 \).

The methods used to quality control, calibrate, and correct PurpleAir sensor data are necessary to use the data, but since research on the subject is new and rapidly evolving, future improvements are likely. More advanced methods currently under development that use additional data to correct for variation of aerosol composition and meteorological variables other than relative humidity may improve the accuracy of the blended AQI.

2.3. Evaluation

To evaluate the interpolation accuracy, the root mean square interpolation errors (RMSE) of the concentrations are compared with the RMSE of two other interpolation methods: inverse distance weighting (IDW) and proxy interpolation.\(^4\) The RMSE is calculated for each pollutant using leave one out cross validation (LOOCV) (Efron and Tibshirani 1994, p. 224). Since the large spatial variability in \( \text{PM}_{2.5} \) makes it more difficult to predict and elevated concentrations can drive AQI values in the winter months, independent evaluation data sets are also used to evaluate \( \text{PM}_{2.5} \) performance.

2.3.1. Independent evaluation data for \( \text{PM}_{2.5} \)

Three independent evaluation data sets are used to calculate RMSE for \( \text{PM}_{2.5} \): (i) gravimetric 24-hour integrated \( \text{PM}_{2.5} \) measurements made at 11 regulatory monitor sites that do not have a collocated hourly BAM instrument, shown in figure 2, using R & P Model 2000 and 2025 and Thermo Electron Model RAAS2.5–300 Air Samplers, (ii) BAM measurements at the North Hollywood station and, (iii) BAM measurements at the Mission Viejo station. The gravimetric \( \text{PM}_{2.5} \) sampler data are independent of the hourly BAM data and sensor data that are used as input for the interpolation. South Coast AQMD gravimetric \( \text{PM}_{2.5} \) and BAM \( \text{PM}_{2.5} \) measurements are comparable for reporting AQI, as determined by linear correlation coefficients greater than 0.90 (South Coast AQMD 2019). Best fit lines to collocated data have intercepts between 1.09 and 2.71 \( \mu g \) m\(^{-3}\) and slopes between 0.95 and 1.07 (South Coast AQMD 2019). The measurement bias is expected to have only a small effect on the calculated RMSE, and it affects all interpolation methods in the same way, so the gravimetric data can be used to compare RMSE between methods. Measurements at North Hollywood and Mission Viejo are relatively new and are not used as input data for the interpolation or for the bias correction applied to the model data so the data at these stations is also independent. The gravimetric data was collected between June 8, 2018 and March 31, 2020, North Hollywood data was collected between October 11, 2019 and March 31, 2020, and Mission Viejo data was collected between October 29, 2019 and March 31, 2020.

2.3.2. Proxy interpolation

Proxy interpolation is a lookup table that maps measurements at regulatory ambient monitoring stations to defined geographic areas. The 37 geographic areas in the South Coast Air District that are assigned proxy stations are shown in supplementary figure S-18. Each area is assigned multiple stations based on proximity and representativeness of the regulatory monitor; higher priority stations are used first and, if data is unavailable, then lower priority stations are used.

2.3.3. Inverse distance weighting

Inverse Distance Weighting (IDW) is an interpolation method that calculates the interpolated concentration as

\[
C(s_0) = \frac{\sum_{i=1}^{N} w_i C(s_i)}{\sum_{i=1}^{N} w_i},
\]

where the \( C \) are the concentrations, \( s_0 \) is the location at which to interpolate, \( s_i \) are the locations where measurements are made, \( w_i \) are the interpolation weights, and \( N \) is the number of stations to use for the interpolation at \( s_0 \). The weights are calculated as \( w_i = d(s_0, s_i)^{-n} \), where \( d(s_0, s_i) \) is the distance between \( s_0 \) and \( s_i \), and \( n \) is a parameter. IDW is used by airnow.gov for their public-facing AQI map with \( n = 5 \) and \( N = 10 \) (U.S. EPA 2020b). IDW is calculated using the same parameter values in this work.

3. Results and discussion

There were between 103 and 688 outdoor PurpleAir sensors in the map domain on any given hour during the June 8 2018—March 31, 2020 time period, with the number of sensors increasing over the study period, where sensors are counted as the number of unique labels. Data completeness for regulatory monitors and PurpleAir are shown in table 2. The quality control procedures removed 12% of hourly PurpleAir data on average, at most 19% of PurpleAir data for one hour, and almost none of the regulatory monitor data. Regulatory monitor data is downloaded in near-real-time, and therefore is only screened by automated quality control procedures applied by the dataloggers, such as checks for out of range values.
and large changes in the values. Most of the regulatory monitor data that was missing was because the data was not available when it was downloaded from airnowtech.org, caused by routine zero and span precision checks conducted by some instruments, airnowtech.org service outages, and intermittent network connection errors in the real time data download scripts. Model data was always available for PM$_{2.5}$ and O$_3$.

The RMSE of each interpolation method and each pollutant calculated using LOOCV is shown in table 3. The RMSE is calculated at each monitor and then the RMSEs are averaged. Note that the LOOCV evaluated errors at the monitors and not the sensor locations. Bootstrapping is used to test the hypothesis that the average RMSE of blended and natural neighbor interpolation are significantly less than those of proxy and IDW, with p-values shown in table 3. The bootstrap is based on Algorithm 16.2 from (Efron and Tibshirani 1994, p. 240) with 50 000 resamples with replacement of the RMSE. Only monitors where data for all methods is available are used, which limits the number of monitors for PM$_{10}$, NO$_2$, and CO because proxy has no second priority monitors assigned at some locations. Blended interpolation has smaller LOOCV RMSE than other methods, 3.1 ppb, 1.63 µg m$^{-3}$, and 1.1 µg m$^{-3}$ smaller than proxy for O$_3$, PM$_{2.5}$, and PM$_{10}$, respectively, where the differences for O$_3$ and PM$_{2.5}$ are statistically significant, demonstrating that blended interpolation is more accurate than other interpolation methods. The RMSE of natural neighbor interpolation is not significantly different from proxy.
interpolation is slightly smaller than proxy and IDW interpolation for all pollutants, although some differences are not statistically significant. This demonstrates that the natural neighbor interpolated fields are as accurate as the proxy and IDW interpolation methods.

Independent evaluation data are used to provide another estimate of the interpolation error to verify the findings of the LOOCV analysis. The average RMSE of each method for PM2.5 calculated using the independent gravimetric, North Hollywood, and Mission Viejo data is shown in table 4. Bootstrapping is used to test the hypothesis that the average blended RMSE is smaller than that of Proxy and IDW using a similar method as for the cross validation. The blended RMSE is 1.05 µg m⁻³ smaller than proxy and 0.48 µg m⁻³ smaller than IDW when evaluated with gravimetric data. The blended RMSE is smaller than other methods at North Hollywood but it is 1.56 µg m⁻³ greater than IDW at Mission Viejo. The relatively high blended RMSE at Mission Viejo will be reduced in the future as Mission Viejo measurements are incorporated into the NOAA bias correction model.

The RMSE calculated using gravimetric data has some variation among monitor sites (figure 2). Confidence intervals for the RMSE at each site are calculated using bootstrapping. At 7 out of 11 sites, the blended interpolation has smaller RMSE than the other interpolation methods and the differences in RMSE are small (<0.12 µg m⁻³) at two other sites. The blended RMSE exceeds that of the other methods at Mission Viejo (0.29 µg m⁻³ larger than that of Proxy), which is expected based on the larger RMSE of blended interpolation than the other methods when evaluated with the Mission Viejo BAM measurements. IDW and Proxy RMSE are also smaller than blended at Azusa. The small differences between methods may be caused by sampling error as indicated by overlapping confidence intervals. Thus the blended interpolation is at least as accurate as Proxy and IDW interpolation at most locations in the South Coast AQMD jurisdiction.

| Pollutant | Proxy | IDW | NN | Blended | p-value Blended < Proxy | p-value Proxy < IDW | p-value Blended < IDW | Number of Monitors |
|-----------|-------|-----|----|---------|------------------------|---------------------|----------------------|-------------------|
| PM2.5 [µg m⁻³] | 7.60 | 7.73 | 6.46 | 5.94 | 0.017 | 0.028 | 0.0011 | 0.0047 | 14 |
| O₃ [ppb] | 10.2 | 8.75 | 8.23 | 7.12 | 0.0053 | 0.25 | <0.001 | 0.0014 | 27 |
| PM10 [µg m⁻³] | 34.7 | 36.1 | 33.0 | 33.6 | 0.42 | 0.36 | 0.45 | 0.40 | 9 |
| CO [ppm] | 0.171 | 0.191 | 0.163 | 0.40 | 0.22 | 0.30 | 0.14 | 9 |
| NO₂ [ppb] | 7.57 | 6.94 | 6.22 | 0.14 | 0.22 | 0.22 | 0.22 | 0.22 | 9 |

*p-value < 0.05. RMSE is significantly smaller than the reference interpolation method.*

Interpolated PMcoarse + PM2.5.*Configuration used in the South Coast AQMD AQI map.

Scatter plots comparing interpolated and hourly observed PM2.5 and O₃ at several monitor sites, shown in section S.5 in the supplementary material, are examined as an additional check that the blended interpolation performs adequately. These plots indicate that the blended interpolated concentrations of PM2.5 and O₃ are qualitatively similar to those of proxy interpolation.

One benefit of blending PM2.5 sensor data into the AQI is the improved responsiveness to wildfire smoke. The accuracy of the interpolated PM2.5 fields during wildfire conditions is estimated by calculating the RMSE using data from 22 d, listed in section S.6 in the supplemental material, during periods where wildfire smoke impacted the South Coast Air Basin. Days with active South Coast AQMD wildfire smoke advisories and visible smoke on the Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua or Terra satellite imagery between June 8, 2018 and March 31, 2020 are considered. The LOOCV RMSE on smoke days is 7.73 µg m⁻³ for blended, 10.4 µg m⁻³ for proxy, and 10.3 µg m⁻³ for IDW. The LOOCV RMSE is higher on smoke impact days but performance is significantly improved when using the blended interpolation. There is insufficient data to calculate the RMSE on smoke impact days with the other independent evaluation data.

The responsiveness of blended and IDW AQI maps to wildfire smoke is qualitatively compared in figure 3, which shows AQI for one hour when smoke from the Woolsey Fire impacted the South Coast Air Basin. The blended AQI is higher at central coastal locations because of the sensor data. The blended AQI is smaller at the northwest corner of the jurisdiction and in the Coachella Valley. Additional comparisons for the Woolsey Fire with IDW are shown in the supplemental material section S.7.

The blended interpolation method has a smaller RMSE than other methods because the low-cost sensor data and model data provide information about the local PM2.5 and ozone concentrations in between regulatory monitors and thus reduce errors caused by interpolation. In figure 3 the unhealthy
Table 4. Root mean square error of proxy, IDW, and blended interpolation method for PM$_{2.5}$.

| Method                      | Proxy | IDW | Blended | p-value Blended < Proxy | p-value Blended < IDW |
|-----------------------------|-------|-----|---------|-------------------------|----------------------|
| Gravimetric validation dataset | 4.64  | 4.07 | 3.59    | 0.0051$^b$              | 0.049$^b$            |
| North Hollywood validation dataset | 8.91  | 9.07 | 7.51    | 0.044$^a$               | 0.0029$^a$          |
| Mission Viejo validation dataset | 8.83  | 7.31 | 8.87    |                         |                      |

$p$-value $< 0.05$. RMSE is significantly smaller than the reference interpolation method.$^a$$^b$ 24-hour average.

AQI (red) along the coast southwest of Los Angeles in the blended map is driven by low-cost sensor data while the estimate at the same location on the proxy map is derived from the interpolation of nearby monitor data; the low-cost sensor measurements are likely more accurate than the simple interpolation of solely monitored values. Improvements in performance are also evident in the far eastern side of the jurisdiction (Eastern Riverside County). In the proxy map, concentrations recorded in the South Coast Air Basin, a far more urbanized region, are used to populate AQI values in the sparsely-populated Eastern Riverside County. Because of the absence of nearby monitors, the blended map relies on the modeled values to represent concentrations in Eastern Riverside County.

4. Conclusions

The weight of evidence presented in this study suggests that the blended interpolation is more accurate than two common methods (i.e. proxy interpolation and IDW) for interpolating air quality measurements, at most locations in the South Coast Air District for O$_3$, PM$_{2.5}$, and PM$_{10}$. Overall improvements in accuracy are approximately 30%, 22%, and 3% for O$_3$, PM$_{2.5}$, and PM$_{10}$, respectively, as RMSE relative to proxy interpolation, calculated using LOOCV. Accuracy improvements for PM$_{2.5}$ on wildfire smoke impact days are even larger, with approximately a 25% reduction in RMSE. Wildfire smoke often leads to widespread areas of PM$_{2.5}$ concentrations that exceed federal standards, and as a result, increased use of real-time AQI values by the public. Moreover, the blended interpolation provides more accurate AQI values at a higher spatial resolution than previous methods and will strengthen the scientific basis for decision making by schools, public health departments, outdoor workers, and residents to minimize personal exposure and protect public health.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Nico Schulte  https://orcid.org/0000-0001-6758-6825
Xiang Li  https://orcid.org/0000-0001-6797-7340
Scott A Epstein  https://orcid.org/0000-0002-1282-2015

References

Cressie N and Wikle C K 2015 *Statistics for Spatio-Temporal Data* (New York: Wiley)
Efron B and Tibshirani R J 1994 *An Introduction to the Bootstrap* (Boca Raton, FL: CRC Press)
Esri 2020a Terrain with Labels [basemap]. Scale Not Given (available at: www.arcgis.com/home/item.html?id=a52ab98763904006aa382d90c906fdd5)
Esri 2020b Streets [basemap]. Scale Not Given. (available at: www.arcgis.com/home/item.html?id=-3b93337983e9436f8db95038a8629af)
Feenstra B et al 2019 Performance evaluation of twelve low-cost PM$_{2.5}$ sensors at an ambient air monitoring site Atmos. Environ. 216 116946
Huang J et al 2017 Improving NOAA NAQFC PM$_{2.5}$ predictions with a bias correction approach Weather and Forecast. 32 407–21
Jerrett M et al 2005 A review and evaluation of intraurban air pollution exposure models J. Expo. Anal. Environ. Epidemiol. 15 185–204
Kelly K E et al 2017 Ambient and laboratory evaluation of a low-cost particulate matter sensor Environ. Pollut. 221 491–500
Lee P et al 2017 NAQFC developmental forecast guidance for fine particulate matter (PM2.5) Weather and Forecast. 32 343–60
Li J and Heap A D 2008 A Review of Spatial Interpolation Methods for Environmental Scientists (Canberra: Geoscience Australia)
Malings C, Tanzer R, Hauryliuk A, Saha P K, Robinson A L, Presto A A and Subramanian R 2020 Fine particle mass monitoring with low-cost sensors: corrections and long-term performance evaluation Aerosol. Sci. Technol. 54 160–74
Morawska L et al 2018 Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: how far have they gone? Environ. Int. 116 286–99
Pournazeri S et al 2014 A computationally efficient model for estimating background concentrations of NOX, NO2, and O3 Environ. Model. Softw. 52 19–37
PurpleAir LLC 2020 PurpleAir: real-time Air Quality Monitoring (available at: https://www2.purpleair.com/)
Sakov P 2020 Natural Neighbours interpolation (available at: https://github.com/sakov/nn-c) (Accessed: 31 March 2020)
Sibson R 1981 A Brief Description of Natural Neighbor Interpolation, Interpreting Multivariate Data (New York: Wiley)
Singh V et al 2011 A colkriging based approach to reconstruct air pollution maps, processing measurement station concentrations and deterministic model simulations Environ. Modell. Softw. 26 778–86
South Coast AQMD 2019 Air Quality Monitoring Network Plan Appendix C: PM2.5 Continuous Monitor Comparability Assessment and Request for Waiver (available at: www.aqmd.gov/home/air-quality/clean-air-plans/monitoring-network-plan)
South Coast AQMD 2020a AQ-SPEC Air Quality Sensor Performance Evaluation Center (available at: www.aqmd.gov/aq-spec) (Accessed: 17 April 2020)
South Coast AQMD 2020b South Coast AQMD Website (available at: www.aqmd.gov) (Accessed: 6 March 2020)
US EPA 2018 Technical Assistance Document for the Reporting of Daily Air Quality – The Air Quality Index (AQI) (Research Triangle Park, NC: US EPA)
US EPA 2019 O3-Nowcast (available at: https://github.com/USEPA/O3-Nowcast) (Accessed: 5 January 2020)
US EPA 2020a AirNow.gov. (available at: www.airnow.gov) (Accessed: 6 March 2020)
US EPA 2020b How are the map contours made? What interpolation method is used? (available at: www.airnow.gov/faqs/how-are-map-contours-made-what/) (Accessed: 6 March 2020)
US EPA 2020c How is the NowCast algorithm used to report current air quality? (available at: www.airnow.gov/faqs/how-nowcast-algorithm-used-report/) (Accessed: 5 January 2020)
Yu H, Russell A, Mulholland J, Odman T, Hu Y, Chang H H and Kumar N 2018 Cross-comparison and evaluation of air pollution field estimation methods Atmos. Environ. 179 49–60