Supplemental Information For:

Fine Particle Mass Monitoring with Low-Cost Sensors: Corrections and Long-Term Performance Evaluation

Carl Malings[1], Rebecca Tanzer[1], Aliaksei Hauryliuk[1], Provat K. Saha[1], Allen L. Robinson[1], Albert A. Presto[1], R. Subramanian[1]

[1]Center for Atmospheric Particle Studies, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213. Email: subu@cmu.edu (Corresponding Author)

1. RAMP and PM Sensor Picture

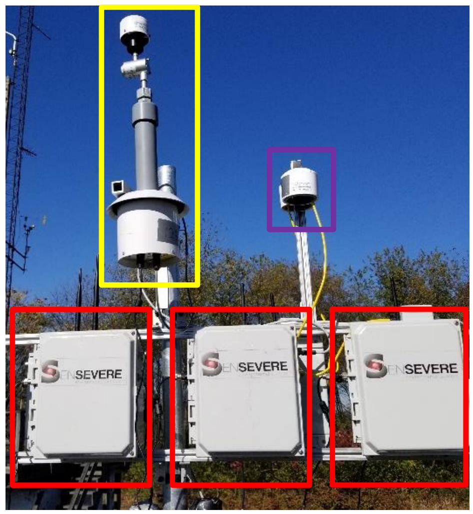

Figure S.1: Several RAMP monitors (red boxes) with connected Met-One NPM (yellow box) and PurpleAir (purple box) $PM_{2.5}$ sensors.

2. Correction Methods – Hygroscopic Growth Factor Computation

This hygroscopic growth factor is computed as:

$$\mathrm{fRH}(T, RH) = 1 + \kappa_{\mathrm{bulk}} \frac{a_w(T,RH)}{1-a_w(T,RH)} \tag{S.1}$$

where:

$$a_w(T, RH) = RH \exp\left(\frac{4\sigma_w M_w}{\rho_w RTD_{\mathrm{p}}}\right)^{-1} \tag{S.2}$$

κ_{bulk} is the hygroscopicity of bulk aerosol; $\kappa_{\mathrm{bulk}} = \sum_i \ x_i\kappa_i$ where. x_i and κ_i are the volume fraction hygroscopocity parameters of the i^{th} component comprising the particle. Organic, sulfate, nitrate and ammonium are assumed as the main components comprising the particle. The fractional contributions of these chemical components to $PM_{2.5}$ during summer, winter, and as an annual average (applied to other periods) are obtained from recent AMS measurements in Pittsburgh (Gu et al. 2018) and their hygroscopocity parameters are adopted from literature (Cerully et al. 2015; Petters and Kreidenweis 2007). a_w is the water activity parameter, estimated using Eq. (S.2), where σ_w, M_w, and ρ_w represent the surface tension, molecular weight and density of water, respectively; T is the absolute temperature, R is the ideal gas constant, RH is ambient relative humidity; D_p is the particle diameter, adopted as volume median diameter from long-term size distribution measurements using SMPS in Pittsburgh. Table S.1 lists different parameter values used in hygroscopic growth factor calculation.

Table S.1: Parameters used in hygroscopic growth factor calculation

Parameter	Value			Unit	Source
	Summer	Winter	Other		
κ_{OA}	0.15	0.15	0.15	-	(Cerully et al. 2015)
κ_{SO4}	0.5	0.5	0.5	-	(Petters and Kreidenweis 2007)
κ_{NO3}	0.6	0.6	0.6	-	(Petters and Kreidenweis 2007)
κ_{NH4}	0.5	0.5	0.5	-	(Petters and Kreidenweis 2007)
x_{OA}	0.64	0.41	0.53	-	(Gu et al. 2018)
x_{SO_4}	0.24	0.16	0.20	-	(Gu et al. 2018)
x_{NO_3}	0.04	0.29	0.165	-	(Gu et al. 2018)
x_{NH_4}	0.08	0.15	0.115	-	(Gu et al. 2018)
κ_{bulk}	0.26	0.34	0.30	-	
σ_w	0.072	0.072	0.072	N/m	
M_w	0.018	0.018	0.018	kg/mol	
ρ_w	1000	1000	1000	kg/m^3	
R	8.314	8.314	8.314	J/mol K	
D_p	200	200	200	nm	

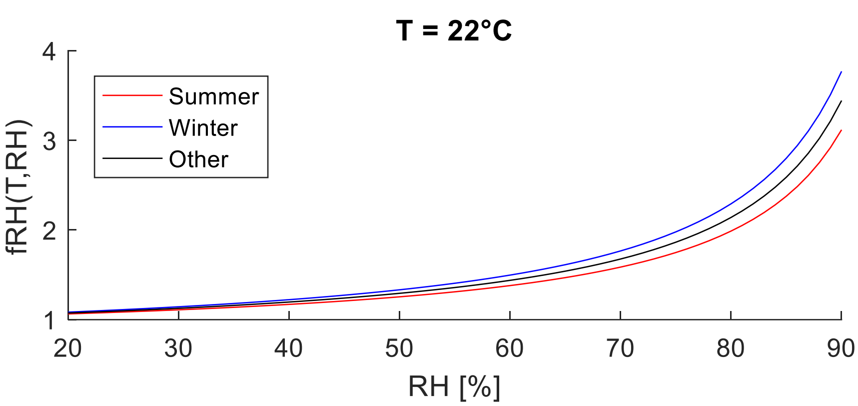

Figure S.2: Example of how the hygroscopic growth factor varies with humidity in summer, winter, and otherwise.

3. Correction Methods – Empirical Approach

Several explanatory factors were considered for the empirical correction method. Dewpoint DP was considered as a factor related to condensation that might serve as a proxy for the hygroscopic growth factor which is independent of aerosol composition. Furthermore, humidity is known to affect the performance of optical particle sensors directly (e.g. Jayaratne et al. 2018), and so relative humidity RH was included as a factor. Finally, temperature T was included as a factor since it has been observed to affect the performance of optical sensor components (Johnson et al. 2016; Jayaratne et al. 2018; Zheng et al. 2018).

Various combinations of the as-reported sensor readings and the above inputs into various functional forms and with different application thresholds were applied to generate correction equations. Two functional forms were considered: linear and quadratic regression models. Thresholds were considered to define different subsets of the domain over which different functional parameters could be applied, allowing for piecewise-linear or piecewise-quadratic functions. Models without thresholds were considered, as well as models with single or multiple threshold values chosen from among 5, 10, 15, 20, 30, 40, and 50 $\mu g/m^3$ (as determined from the raw sensor reading). For reference, ambient concentrations in Pittsburgh typically range from 3 to 20 $\mu g/m^3$.

Models were calibrated using a combination of data collected at both the Lawrenceville and Lincoln sites from half of the sensors deployed to each site (the “training” set); model performance was evaluated on the other half of sensors at these sites (the “testing” set). Performance metrics assessed for the various models are included as supplementary data. The performance of each correction model on the test sensor set was scored using a heuristic combining various performance metrics across both collocation sites and penalizing the complexity of the model (see the supplementary data for the resulting metrics). For selecting a final correction method for each type of sensor, performance across a range of concentrations experienced at both collocation sites was traded off against the complexity of the model (and therefore its propensity to overfit to training data).

4. Drift-Adjustment Methods

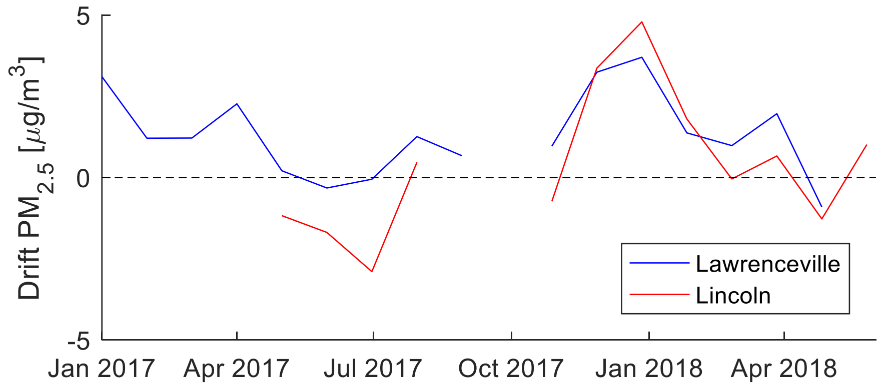

Figure S.3: Illustration of observed NPM sensor drift at the Lincoln and Lawrenceville sites. Drift is depicted as the difference in monthly average readings of the NPM sensor, corrected using Eq. (4), versus the collocated regulatory-grade instrument at each site.

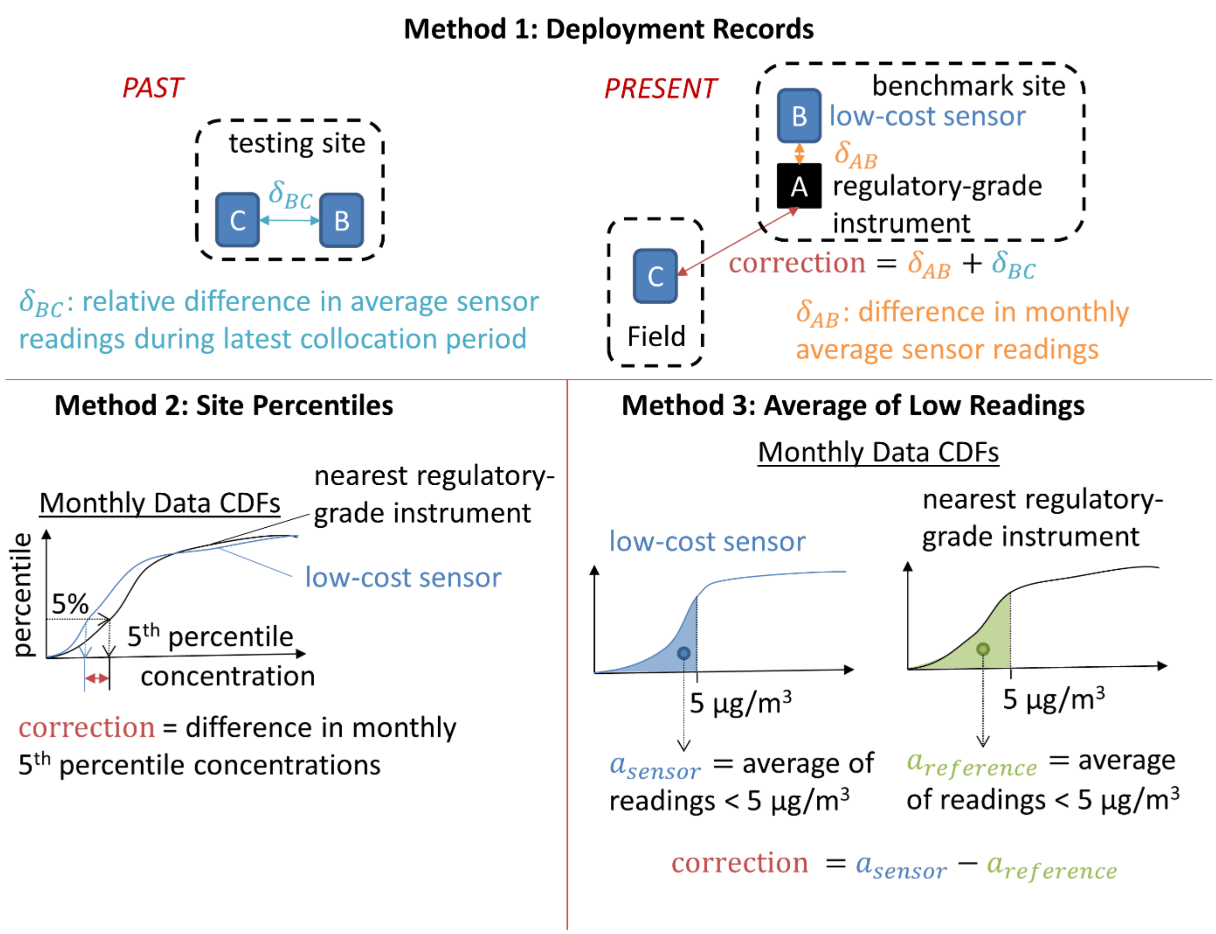

Figure S.4: Diagrams of the three proposed drift-adjustment methods.

5. Assessment metrics

For n measurements of concentration by the sensor (c) and reference ($\hat{c}$), bias is computed as:

$$\text{bias} = \frac{1}{n}\sum_{i=1}^{n}(c_i - \hat{c}_i) \tag{S.3}$$

mean absolute error (MAE) is evaluated as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|c_i - \hat{c}_i| \tag{S.4}$$

and the Pearson correlation coefficient (r) is evaluated as:

$$r = \frac{\sum_{i=1}^{n}\left(c_i - \frac{1}{n}\sum_{j=1}^{n} c_j\right)\left(\hat{c}_i - \frac{1}{n}\sum_{j=1}^{n}\hat{c}_j\right)}{\sqrt{\sum_{i=1}^{n}\left(c_i - \frac{1}{n}\sum_{j=1}^{n} c_j\right)^2}\sqrt{\sum_{i=1}^{n}\left(\hat{c}_i - \frac{1}{n}\sum_{j=1}^{n}\hat{c}_j\right)^2}}, \tag{S.5}$$

These statistics assess, respectively, the systematic differences between the sensor and reference measurements over time, the average absolute difference in measurements taken at the same time, and the degree of linearity between the measurements. Lower absolute values of bias and MAE denote better agreement, while a value of r close to 1 denotes stronger correlation.

Additionally, the following EPA bias and precision score metrics (Camalier et al., 2007) were used:

$$\text{Precision Score} = \sqrt{\frac{n\sum_{i=1}^{n}\delta_i^2 - \left(\sum_{i=1}^{n}\delta_i\right)^2}{n\chi^2_{0.1,n-1}}} \tag{S.6}$$

where $\chi^2_{0.1,n-1}$ denotes the 10$^{\text{th}}$ percentile of the chi-squared distribution with $n-1$ degrees of freedom, and:

$$\delta_i = 100\,\frac{c_i - \hat{c}_i}{\hat{c}_i} \tag{S.7}$$

The bias score is:

$$\text{Bias Score} = \frac{1}{n}\sum_{i=1}^{n}|\delta_i| + \frac{t_{0.95,n-1}}{n}\sqrt{\frac{n\sum_{i=1}^{n}\delta_i^2 - \left(\sum_{i=1}^{n}|\delta_i|\right)^2}{n-1}} \tag{S.8}$$

where $t_{0.95,n-1}$ is the 95$^{\text{th}}$ percentile of the t distribution with $n-1$ degrees of freedom. These precision and bias scores can be compared to performance guidelines for various sensing applications (Williams et al., 2014). For $PM_{2.5}$, requirements for educational monitoring (Tier I) are for precision and bias scores below 50%; for hotspot identification and characterization (Tier II) or personal exposure monitoring (Tier IV), these should be below 30%; for supplemental monitoring (Tier III), below 20%; and for regulatory monitoring (Tier V), below 10%.

6. Seasonal Changes in $PM_{2.5}$ fraction below 300 nm in Pittsburgh

Aerosol size distributions over the 10-300 nm mobility size range were measured with a TSI scanning mobility particle sizer (SMPS) at the CMU campus. $PM_{0.3}$ mass concentrations were estimated assuming a mobility density of 1 gm/cm^3 and spherical particles, and then corrected to the equivalent mass at 35% RH using the previously-discussed hygroscopic corrections. $PM_{2.5}$ mass concentrations were obtained from an NPM instrument attached to a RAMP co-located with the SMPS. These values were corrected using Eq. (1). For the winter months, the RAMP RH was assumed to be the same as the conditions inside the SMPS. For the summer months, we assumed that the SMPS RH was 15% higher (than the RAMP RH) inside the air-conditioned trailer where the SMPS operated. The SMPS/NPM comparison is further complicated by the fact that we are comparing an electrical mobility sizer to an optical sizer, but the overall result of higher sub-300 nm aerosol mass is consistent with previously reported results.

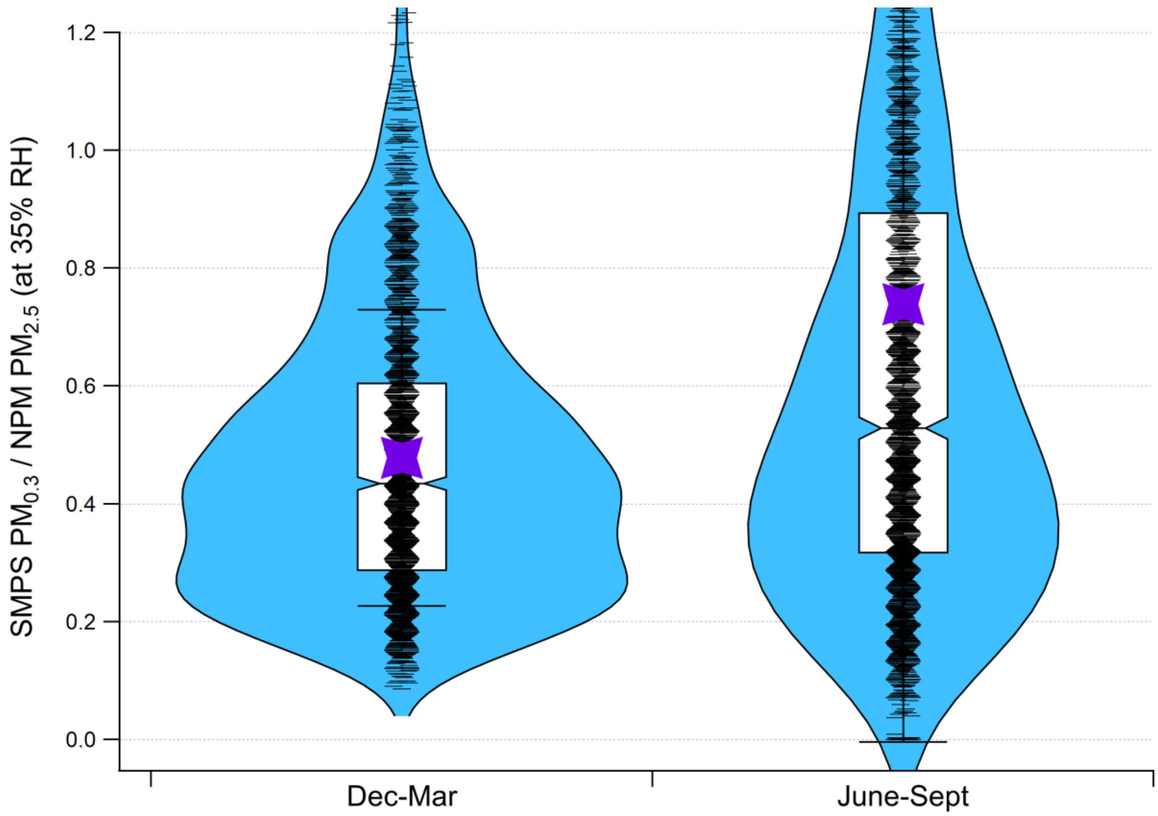

Figure S.5: Ratios of $PM_{0.3}$ to $PM_{2.5}$ based on summer and winter data collected in Pittsburgh. Individual data points are jittered; means are shown by the purple stars; whiskers represent one standard deviation of the data. Values greater than unity likely indicate data where our assumptions are no longer valid, but these are <25% of the data. The median $PM_{0.3}/PM_{2.5}$ is 0.43 in the winter and 0.53 in the summer. For an annual average concentration of ~10 µg/m^3, this represents a 1 µg/m^3 higher sub-300 nm fraction in the summer.

7. Results for Correction Methods

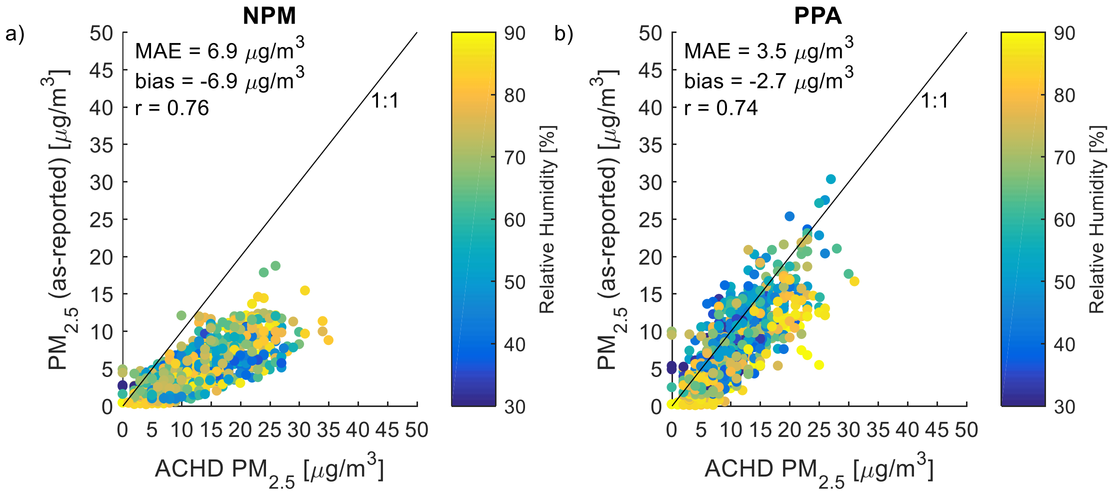

Figure S.6: Comparison of median one-hour-average NPM (a) and PPA (b) sensor readings to the BAM instrument during collocation at the Lawrenceville site after correction using a hygroscopic growth factor only (i.e. corrected measurement is raw measurement divide by fRH). Colors indicate relative humidity at the time of the measurements. Note that the NPM measurement corrected in this manner severely underestimates $PM_{2.5}$ concentration. For PPA sensors, while absolute errors are decreased relative to those of using the as-reported values directly, bias is also increased and correlation is reduced.

Table S.2: Coefficients for empirical correction equations

Coefficient	Value Estimate	Standard Deviation	Unit
α_0	0	2.9	$\mu g/m^3$
α_1	2.93	0.08	N/A
α_2	-0.11	0.08	$\mu g/{}^\circ C m^3$
α_3	0	0.08	$\mu g/\% m^3$
α_4	5.3×10^{-4}	1.5×10^{-4}	$m^3/\mu g$
α_5	-8.9×10^{-3}	1.2×10^{-3}	${}^\circ C^{-1}$
α_6	-2.7×10^{-2}	0.11×10^{-2}	$\%^{-1}$
α_7	2.9×10^{-3}	0.8×10^{-3}	$\mu g/{}^\circ C^2 m^3$
α_8	5.0×10^{-3}	1.0×10^{-3}	$\mu g/{}^\circ C \% m^3$
α_9	0	6.0×10^{-4}	$\mu g/\%^2 m^3$
β_0	75	11	$\mu g/m^3$
β_1	0.60	0.0090	N/A
β_2	-2.5	0.51	$\mu g/{}^\circ C m^3$
β_3	-0.82	0.11	$\mu g/\% m^3$
β_4	2.9	0.53	$\mu g/{}^\circ C m^3$
γ_0	21	2.1	$\mu g/m^3$
γ_1	0.43	0.013	N/A
γ_2	-0.58	0.090	$\mu g/{}^\circ C m^3$
γ_3	-0.22	0.023	$\mu g/\% m^3$
γ_4	0.73	0.098	$\mu g/{}^\circ C m^3$

The following figure summarizes the medians and ranges in performance of the corrected NPM and PPA hourly averaged data across both collocation sites, using all sensors deployed to both sites (as opposed to only the testing set), as well as specifying performance by different concentration ranges (0 to 10, 10 to 20, and higher than 20 μg/m^3). Correlation is typically better for NPM sensors (using either empirical correction equation), with r between 0.7 and 0.9, while for PPA sensors it ranges down to 0.5. Correlations also improve at higher concentrations. The MAE for both sensors are between 3 and 5 μg/m^3. MAE also tends to increase as concentrations increase, but the PPA sensors appear to be less affected than NPM at concentrations above 20 μg/m^3; however, considering there were only two PPA sensors at the Lincoln site (where these higher concentrations were more common) this may be a sample size artefact. Although unbiased over the full range, the corrected sensor readings tend to be positively biased at low concentrations and negatively biased at moderate concentrations. This is opposite to the trend seen before correction and may be due to overcorrections at the extremes.

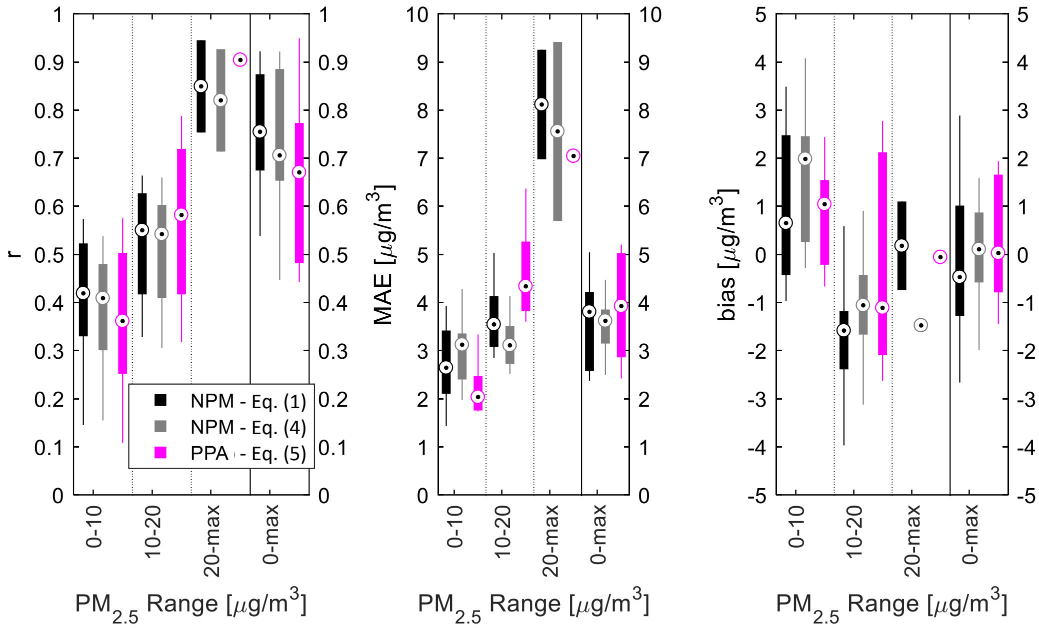

Figure S.7: Comparison of one-hour-average corrected sensor performance compared to BAM instruments during collocation at both the Lawrenceville and Lincoln sites. Performance metrics are plotted overall (0-max range) and by different $PM_{2.5}$ ranges (0-10, 10-20, 20-max). Results shown relate to a total of 32 NPM and 11 PPA sensors, and only consider sensors with at least five samples in the relevant range.

The following figures illustrate how the performance of the proposed correction approaches is affected if data from just one of the sites (Lincoln or Lawrenceville) is used to train the model, and it is then tested on data from the other site.

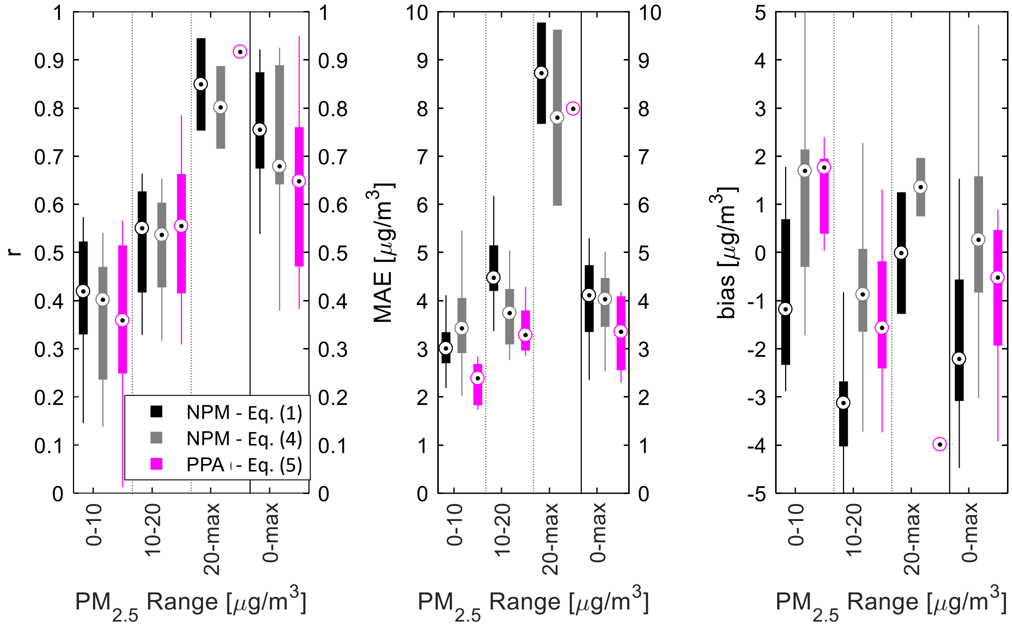

Figure S.8: Comparison of sensor performance compared to the BAM instrument during collocation at the Lawrenceville site, using correction models calibrated using only data collected at the Lincoln site. Performance is comparable in terms of correlation and MAE to models trained using data from both sites, although bias, especially using Eq. (1) for NPM sensors, is generally worse.

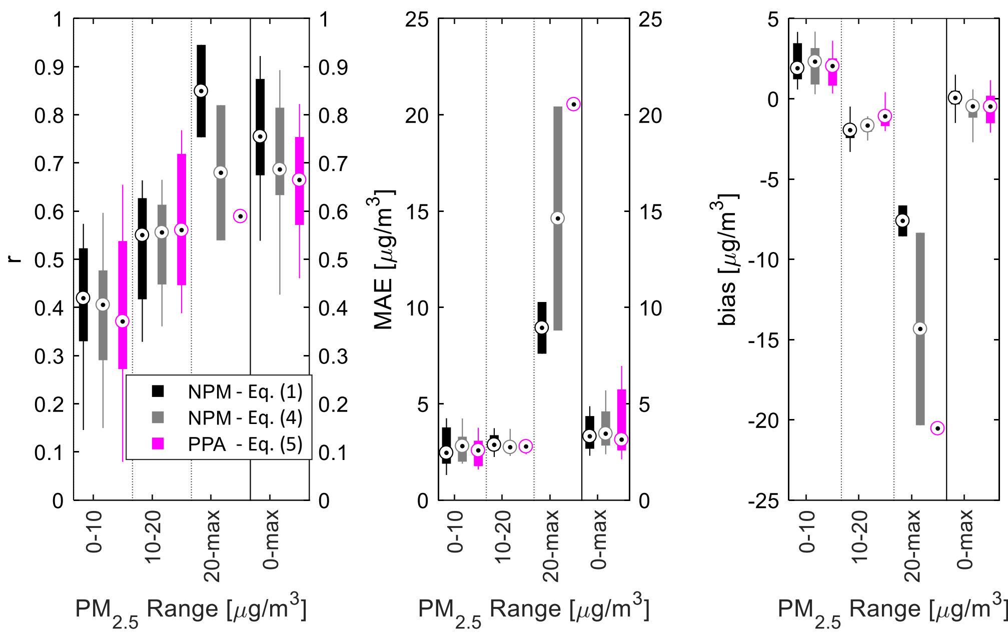

Figure S.9: Comparison of sensor performance compared to the BAM instrument during collocation at the Lincoln site, using correction models calibrated using only data collected at the Lawrenceville site. Performance is comparable except in the 20-max range, where performance is significantly worse than for models calibrated using data from both sites. This illustrates the importance of calibrating correction equations across the entire range of concentrations which might be expected during field deployments.

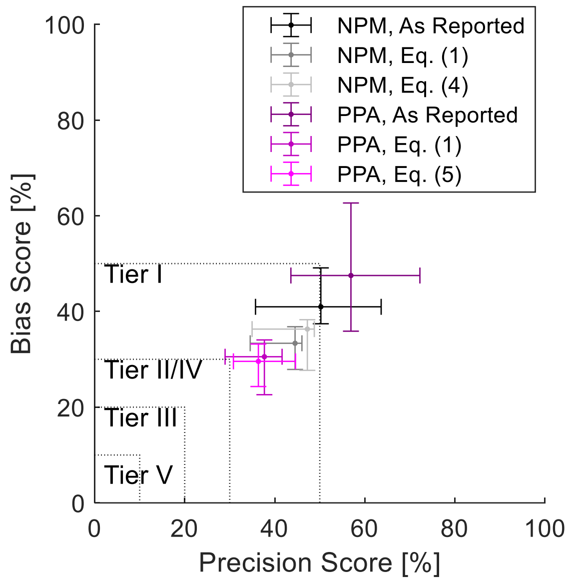

Figure S.10: Evaluation of EPA precision and bias score metrics for hourly-averaged data from NPM and PurpleAir sensors. Center-points of crosses indicate median performance, with arms indicating 25%-75% range. Following corrections, both instruments meet Tier I requirements for educational monitoring.

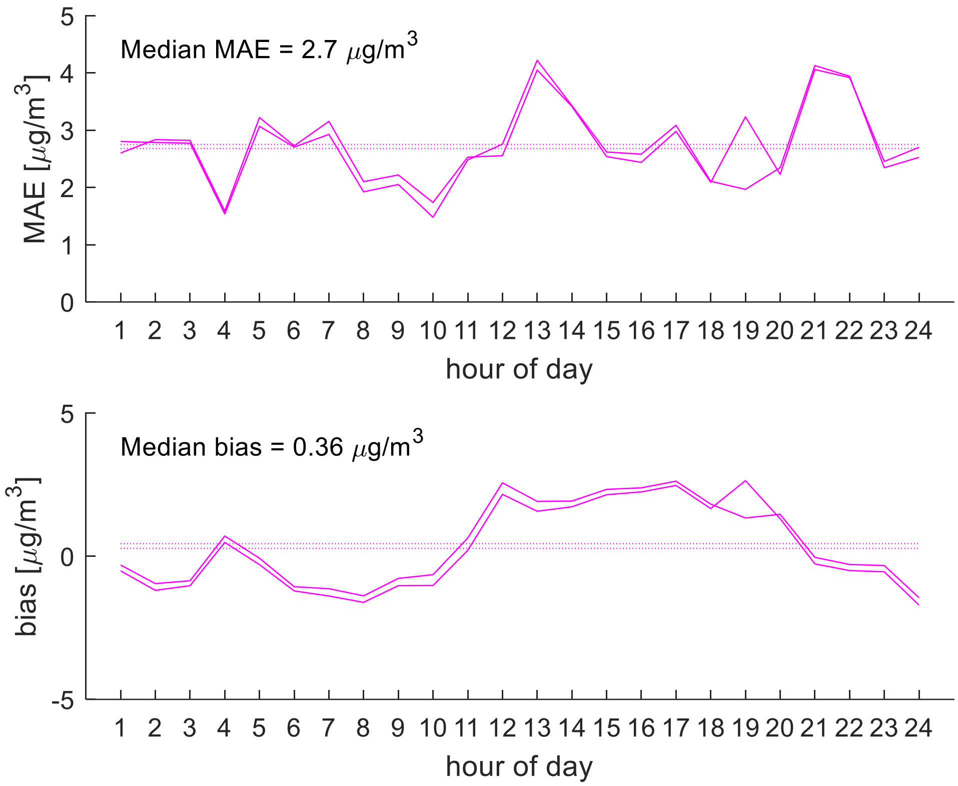

Figure S.11: Results of a performance evaluation of a pair of PurpleAir sensors at the Parkway East site. Results cover a data collection period of three weeks. Hourly-average bias and MAE are plotted as a function of time of day in the solid lines for the two sensors; dotted lines indicate the median performance throughout the day for each sensor. Median bias and MAE for both sensors are also listed in the figure. Corrections are performed using Eq. (1).

References

Cerully KM, Bougiatioti A, Hite JR, Guo H, Xu L, Ng NL, et al. 2015. On the link between hygroscopicity, volatility, and oxidation state of ambient and water-soluble aerosols in the southeastern United States. Atmospheric Chemistry and Physics 15:8679–8694; doi:10.5194/acp-15-8679-2015.

Gu P, Li HZ, Ye Q, Robinson ES, Apte JS, Robinson AL, et al. 2018. Intra-city variability of PM exposure is driven by carbonaceous sources and correlated with land use variables. Environmental Science & Technology; doi:10.1021/acs.est.8b03833.

Jayaratne R, Liu X, Thai P, Dunbabin M, Morawska L. 2018. The Influence of Humidity on the Performance of Low-Cost Air Particle Mass Sensors and the Effect of Atmospheric Fog. Atmospheric Measurement Techniques Discussions 1–15; doi:10.5194/amt-2018-100.

Johnson KK, Bergin MH, Russell AG, Hagler GSW. 2016. Using Low Cost Sensors to Measure Ambient Particulate Matter Concentrations and On-Road Emissions Factors. Atmospheric Measurement Techniques Discussions 1–22; doi:10.5194/amt-2015-331.

Petters MD, Kreidenweis SM. 2007. A single parameter representation of hygroscopic growth and cloud condensation nucleus activity. Atmospheric Chemistry and Physics 7:1961–1971; doi:10.5194/acp-7-1961-2007.

Zheng T, Bergin MH, Johnson KK, Tripathi SN, Shirodkar S, Landis MS, et al. 2018. Field evaluation of low-cost particulate matter sensors in high and low concentration environments. Atmospheric Measurement Techniques Discussions 1–40; doi:10.5194/amt-2018-111.