PLANET TEXAS 2050

A UT Grand Challenge

Appendix: Project Publications

Performance Evaluation of Light Scattering $PM_{2.5}$ Sensors for Deployment in an Urban Sensing Network in Bangalore, India

Jonathan Gingrich[1], Mark Campmier[1], Advaitha Byereddy[1], Shayan Charolia[1], Heather Howton[1], Brian Mai[1], Meenakshi Kushwaha[2], Elbin Savio[2], Adithi Upadhya[2], Sreekanth Vakacherla[3], Julian Marshall[4], Joshua Apte[1]

[1]University of Texas at Austin, USA [2]ILK Consultancy Ltd., Bangalore, India [3]Centre for Science, Technology and Policy, Bangalore, India [4]University of Washington, USA

Motivation

Particulate matter (PM) causes a wide variety of adverse health effects and leads to decreased life expectancy[1]. While fine particle ($PM_{2.5}$) concentrations can vary sharply in space and time[2], the spatial density of most monitoring networks is low, especially in many developing world cities. Lower-cost (<$500) sensor systems could fill critical data gaps on $PM_{2.5}$. However, for many reasons, optical devices may have high associated uncertainty when estimating $PM_{2.5}$. Real-world performance evaluations of low-cost sensors for precision, accuracy, reproducibility, and stability can help quantify these uncertainties.

Aims

The aims of this project were to test 60 low-cost sensors in a polluted urban environment (Bangalore, India) for (i) the comparability of measurements between identical sensors, and (ii) for their response in the ambient aerosol relative to a federally equivalent reference monitor (BAM-1022). We also explored the effect of increasing averaging time on sensor agreement.

Results and Discussion

Comparison among 60 PurpleAirs

7 sensors failed at the beginning of the collocation period, resulting in 113 light scattering sensors used in the analysis. For the pairwise comparison between the sensor modules, the median R^2 was 0.99 for 1-hour averaging. The median NRMSE between all pairs was 0.02 and the average slope was 0.98. The data show that the PA sensors agree very well when compared to one another, except for a few sensors that read dramatically different from the others (~20%). The remainder are accurate when compared with one another.

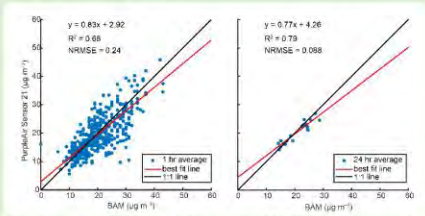

Figure 2. Comparisons between PurpleAir Sensor 21 and the BAM at 1-hour averaging (left) and 24-hour averaging (right). Sensor 21 represented one of the better performing low-cost sensors. The best fit line was determined using ordinary least squares (OLS) regression.

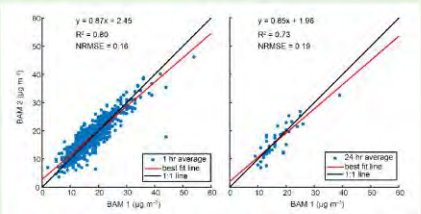

Figure 3. Comparisons between 2 BAM-1022s at 1-hour averaging (left) and 24-hour averaging (right). BAMs were collocated in ambient conditions for 41 days (941 total hours). The best fit line was determined using OLS. Increasing from 1 hr averaging to 24 hr averaging decreased the R^2 and increased the NRMSE.

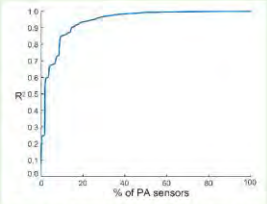

Figure 1. Distribution of 11025 pairwise R^2 values for the functional PurpleAir sensors. 80% of the R^2 values were above 0.9

Comparisons among BAMs

During the collocation between the two BAMs, the hour averaged R^2 was 0.79 and the daily averaged R^2 was 0.73. This indicates that even amongst federally equivalent method monitors, one sensor only accounts for 79% of the variance of the other, and this agreement does not improve with increasing time averaging.

Comparisons between PurpleAirs and BAM

Comparing the low-cost sensors to the BAM at an hourly average resulted in a median R^2 of 0.67, with a median NRMSE of 0.037 and an average slope of 0.81. Compared with the intra-comparisons between the low-cost sensors, these values are not as good, however, with increasing averaging time, these values also increase. At 24-hour averaging, the median R^2 between the BAM and the PurpleAirs is 0.79 with a median NRMSE of 0.034 and average slope of 0.72. This means that at 24-hour averaging, both a BAM and a PurpleAir can explain nearly the equivalent variance between one another.

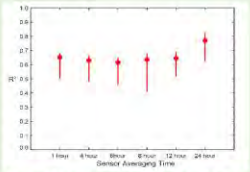

Figure 4. R^2 values between PurpleAirs and BAM as averaging time changes. The circle represents the mean R^2 and the line represents 5%-95% R^2 value.

Conclusions

A key motivation for this study was to assess whether PurpleAir sensors would be measurement approach for assessing spatial patterns of PM2.5 across Bangalore. The high precision of co-located sensors suggests that the PA sensor may be adequate for revealing these spatial patterns in a relative sense. Indeed, the PA-to-BAM relationship had approximately similar imprecision to the performance of two collocated BAM monitors. However, we detected a moderate bias when comparing the PA sensors to the BAM in this 1-month collocation, some of which could likely be corrected for with a calibration equation. The response of optical sensors depends on particle size and composition, which varies in space and time, such that some additional uncertainty will arise when sensors are distributed across the city. The next steps are to deploy the low-cost sensors (after removing the sensors that performed poorly during collocation) throughout the city. Meanwhile, we plan to test the long-term sensor performance of PurpleAirs using a series of sensors collocated with the BAM to determine the effect of sensor drift as well as mass loading on the sensors.

Methods

60 PurpleAir PAII light scattering particulate matter sensors were collocated on the roof of a 15-meter office building in the city of Bangalore, India. A Beta Attenuation Monitor (BAM 1022, MetOne Inc.) was collocated with the PurpleAir monitors from June 20 to July 9, 2019 for a total of 441 hours of collocation. All PurpleAirs have two light scattering sensors (Plantower PMS5003) within the unit, resulting in a total of 120 light scattering sensors tested. Ordinary least squares regression was applied to compare the sensor modules with one another and an R^2, slope, and normalized root mean square error were calculated between each sensor module. This analysis was repeated to compare the 120 sensor modules to the BAM. Finally, to see the effect of averaging time, the analysis was repeated for 4, 6, 8, 12, and 24-hour averaging time.

Finally, after the collocation was completed, a second BAM was installed next to the first BAM. From the period of July 9 to August 19, 2019 these two BAMs recorded hourly averaged $PM_{2.5}$ measurements (941 total hours). The R^2, Slope and NMRSE were compared between these two identical BAM monitors.

References

1. Apte et al., Env. Sci. Technol., 49 (2015) 8057-8066
2. Both et al., Env. Sci. Technol., 45 (2011) 5629-5626

Acknowledgements

The authors would like to thank CSTEP for hosting the sensors at the collocation site in Bangalore, to Jai Asundi CSTEP and David Eaton (University of Texas at Austin) for their assistance, and to the Presidents Award for Global Learning for funding this project.

PLANET TEXAS 2050

A UT Grand Challenge

Performance Evaluation and Distribution of Light Scattering $PM_{2.5}$ Sensors for Deployment in an Urban Sensing Network in Bangalore, India

Jonathan Gingrich[1*], Mark Campmier[1], Advaitha Byereddy[1], Shayan Charolia[1], Heather Howton[1], Brian Mai[1], Meenakshi Kushwaha[2], Elbin Savio[2], Adithi Upadhya[2], Sreekanth Vakacherla[3], Julian Marshall[4], Joshua Apte[1]

[1]University of Texas at Austin, USA [2]ILK Consultancy Ltd., Bangalore, India [3]Centre for Science, Technology and Policy, Bangalore, India [4]University of Washington, USA

*jon.gingric@utexas.edu

Motivation

Particulate matter (PM) causes a wide variety of adverse health effects and leads to decreased life expectancy[1]. While fine particle ($PM_{2.5}$) concentrations can vary sharply in space and time[2], the spatial density of most monitoring networks is low, especially in many developing world cities. Lower-cost (<$500) sensor systems could fill critical data gaps on $PM_{2.5}$. However, for many reasons, optical devices may have high associated uncertainty when estimating $PM_{2.5}$. After performance is compared with respect to a Federal Equivalent Method, the low-cost sensors can then be deployed in field to determine the spatial and temporal variability of $PM_{2.5}$.

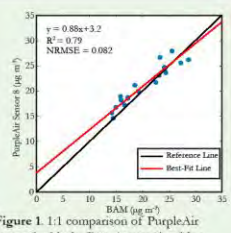

Figure 1. 1:1 comparison of PurpleAir sensor 8 with the Beta Attenuation Monitor (BAM) at daily averaging resolution. 19 days were used in this analysis.

Table 1. Sensor comparisons over 19 day collocation period. Averaging changed from hourly averaging to 24 hour averaging.

	Slope	Intercept	NRMSE	R^2
PA-PA hr	1.01	-0.19	0.02	0.99
PA-PA day	1.00	-0.06	0.02	0.99
PA-BAM hr	1.75	-14.74	0.04	0.01
PA-BAM day	0.88	3.25	0.03	0.74

Collocation Results

- 104 of the 120 sensor modules were used in the collocation period, as the remainder failed before collocation was completed. Pairwise correlation between the PurpleAir sensor modules was very good. Median R^2 values were 0.99 at hourly averaging and the hourly NRMSE was 0.02. About 20% of the sensors had R^2 values of less than 0.8.
- Correlations were poorer between the sensor modules and the BAM at hourly resolution. The median R^2 value was 0.01, and the median NRMSE was 0.04. Decreasing the time resolution to 24 hour averaging increases the R^2 to 0.74 with an NRMSE of 0.03.

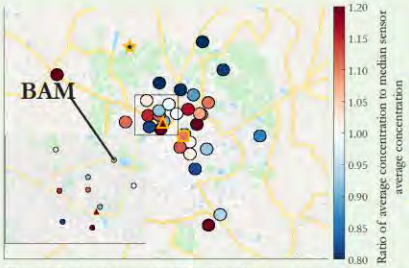

Figure 2. Ratios of average PM2.5 to the median sensor average from the period of August 19 to November 3, 2019. Inset is in the lower left corner. The location of the BAM is also identified.

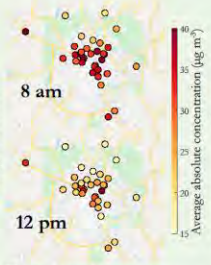

Figure 3. Absolute meanc concentrations at 8 am and 12 pm for all sensors.

Distributed Network Results

- 40 of the best performing sensors were distributed throughout the city. The data used in this poster is from August 18 to November 3, a total of 2.5 months. Figure 2 shows the map of the distributed network. For each sensor the mean concentration over the period was taken. The sensors were then normalized to the median sensor giving the color scale. The median sensor had an average concentration of 28 µg m^{-3}. The 10th and 90th percentile sensors had concentrations of 21 and 33 µg m^{-3}, respectively. This is a range of 43%.
- Figure 3 shows the average absolute concentrations at 8 am and 12 pm. The clean areas remain clean throughout the day compared to the more polluted areas.
- Figure 4 shows the diurnal averages of three sensors (locations indicated on figure 2). The background sensor shows a morning peak, but no afternoon peak. The urban center sensor has both peaks and a higher baseline than the background sensor. Finally, the near road sensor has the highest baseline, but shows a smaller afternoon peak.

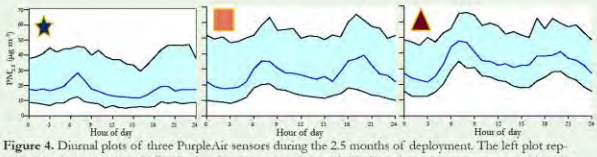

Figure 4. Diurnal plots of three PurpleAir sensors during the 2.5 months of deployment. The left plot represents background, the middle plot is a downtown location, and the right plot represents a near road location.

Conclusion

The high precision of co-located sensors suggests that the PA sensor may be adequate for revealing these spatial patterns in a relative sense. Indeed, the PA-to-BAM relationship had approximately similar imprecision to the performance of two collocated BAM monitors.
The response of optical sensors depends on particle size and composition, which varies in space and time, such that some additional uncertainty will arise when sensors are distributed across the city.
Sensors in the distributed network had a range of 43%. This shows that there is decent variation in $PM_{2.5}$ within the city. Furthermore the inset in Figure 2 shows that even within a relatively small area, there can be a wide amount of variation depending on the local sources.
All sensors see a similar morning peak, however, many sensors fail to see an evening peak.
This research shows that while unreliable at hourly averages, PurpleAir sensors can give some information as to the exposure to $PM_{2.5}$. In the future, the number of sensors will be increased to acheive greater spatial coverage of the city.

Methods

60 PurpleAir PAII light scattering particulate matter sensors were collocated on the roof of a 15-meter office building in the city of Bangalore, India. A Beta Attenuation Monitor (BAM 1022, MetOne Inc.) was collocated with the PurpleAir monitors from June 20 to July 9, 2019 for a total of 441 hours of collocation. Deming regression was applied to compare the 120 plantower sensor modules with one another and an R2, slope, and normalized root mean square error (NRMSE) were calculated between each sensor module. This analysis was repeated to compare the 120 sensor modules to the BAM. After collocation was completed, 40 sensors were distributed throughout the city. Overall average $PM_{2.5}$ concentrations were taken and hourly diurnal plots of $PM_{2.5}$ were created for each sensor in order to try and see if any sensor deviated from the rest.

References

1. Apte et al., Env. Sci. Technol., 49 (2015) 8057-8066
2. Both et al., Env. Sci. Technol., 45 (2011) 5629-5626

Acknowledgements

The authors would like to thank CSTEP for hosting the sensors at the collocation site in Bangalore, to Jai Asundi CSTEP and David Eaton (University of Texas at Austin) for their assistance, and to the Presidents Award for Global Learning for funding this project.

PLANET TEXAS 2050

A UT Grand Challenge

The University of Texas at Austin
Cockrell School of Engineering

Long-Term Performance Evaluation of the PurpleAir PA-II Sensor in New Delhi, India

Mark Campmier [1*], Shahzad Gani [1], Joshua Apte [1]

[1] Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Texas, USA

*mcampmier@utexas.edu

Motivation

– New Delhi, India is the most polluted megacity in the world, with annually-averaged $PM_{2.5}$ measurements of 140 µg/m³. Episodic events in the Autumn and Winter can exceed 1000 µg/m³.

– A dearth of continuous monitoring sites hinders understanding of intra-urban variability, and the relative contributions of local and regional sources.

– Advances in low-cost Optical Particle Counters (OPC), such as the popular PurpleAir PA-II offer a relatively cost-effective means of establishing a dense urban network. However, OPCs have well-documented sources of error, and newer OPCs such as the PA-II have not been extensively evaluated in extremely polluted environments.

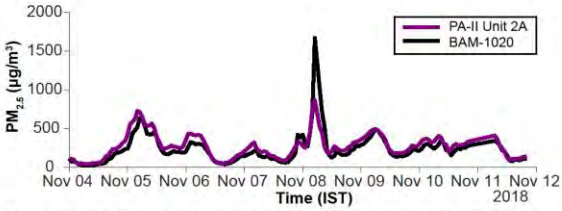

Figure 1: Time series of an extreme event in early November. Large regional sources coinciding with Diwali and unfavorable meteorology lead to days of concentrations exceeding 300 µg/m³ peaking at 1800 µg/m³.

Methodology

– From June 2018 - July 2019, we collocated a BAM-1020 at the US Embassy in New Delhi with 2 PA-II sensor package, each consisting of 2 PMS5003 Plantower units.

– We evaluated several univariate calibration algorithms using a 70:30 test to train split with cross-validation.

– Since the BAM-1020 mass concentration data is noisy at 1-hr averaging time, Orthogonal Distance Regression (ODR) was used instead of Ordinary Least Squares. ODR minimizes error for both the explantory and response variables, ensuring a robust regression relationship for variables with high uncertaintiy.

Results

– All four sensors showed strong 1:1 agreement ($R^2 > 0.95$). For calibration purposes, it is only one representative sensor is necessary, called "PA-II Unit 2A." Hourly BAM-PA performance is influenced by meteorology.

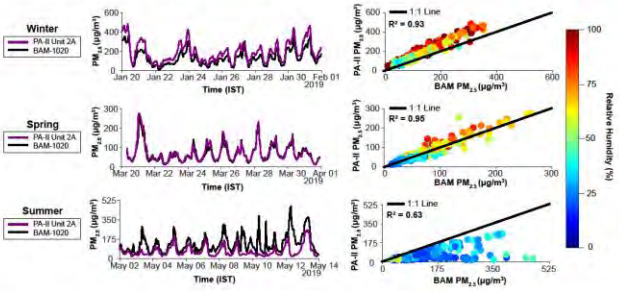

Figure 2: Time series of continuous monitoring data in three seasons, juxtaposed with 1-hr PA-II Unit 2A plotted against the collocated US Embassy BAM-1020 and colored for RH from the same time period. Pearson's R^2 shows high correlation during the high humidity Winter and Spring, but low correlation with systmemic underestimations by the PA-II in the Summer.

– Model performance in Fall, Winter and Spring is strong ($R^2 > 0.8$), but much weaker in the Summer and Monsoon seasons.

Table 1: Comprehensive summary of univariate model performance on the evaluation datasets. NRMSE refers to Normalized Root Mean Square Error, the square root of the mean square difference between the predicted and measured response variable divided by the mean of the measured response variable. "All" refers to models built with all seasons. Performance in the Fall, Winter, and Spring is much more robust than All, while Summer and Monsoon are weaker.

	1-hr			24-hr		
	Slope	R^2	NRMSE	Slope	R^2	NRMSE
All	1.34	0.8	0.43	1.32	0.78	0.45
Fall	1.12	0.84	0.28	1.12	0.84	0.28
Winter	1.27	0.89	0.15	1.23	0.87	0.18
Spring	1.09	0.91	0.2	1.08	0.86	0.27
Summer	0.54	0.48	0.48	0.53	0.56	0.51
Monsoon	0.36	0.15	0.31	0.41	0	0.29

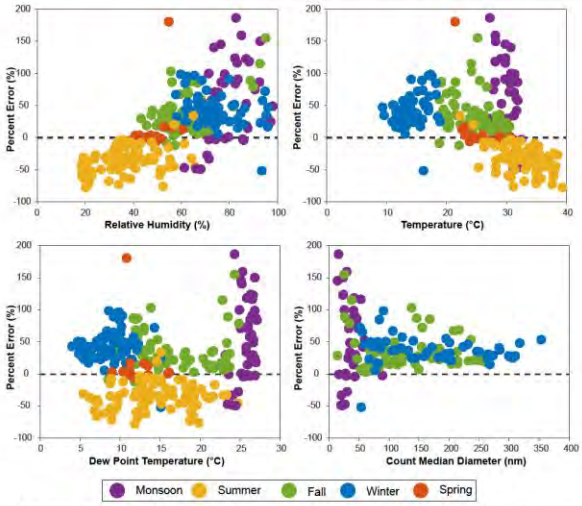

Figure 3: Scatter plots of daily (24-hour block average) percent error between the BAM and PA-II $PM_{2.5}$ with commonly used explanatory variables from literature. None show robust linear relationships ($R^2 < 0.2$). Intra-seasonal variabity for Temperature, RH, and DPT lack structure, pointing to poor explantory power of these variables in seasonal models.

– As expected, high RH results in high error (>50% error when RH >= 60%). However, lower RH also results in systemic underestimations. Higher temperatures are also associated with PA-II underestimations. There is no obvious relationship between the residuals and dew point, but the seasons do form rough clusters.

– The Count Median Diameter (CMD) data was collected at IIT Delhi (about 6 km away from the embassy collocation site). At low daily CMD, the error is unpredictable driven by the PA-II's limitations. As CMD increases, error converges at about +40%. The shift toward smaller CMD in the Monsoon season may help explain the PA-II's poor performance.

Conclusions

– The PA-II can measure $PM_{2.5}$ mass concentrations with strong precision, but variable accuracy. Meterology as well as the particle size distribution were both observed to have pronounced effects on the magnitude and consistency of error. Using season as an approximation for different meterological and aerosol regimes proved to increase model robustness during the Fall, Winter, and Spring at all block averages.

– Summer and Monsoon seasons did not yield significant improvements. Unfortunately, the monsoon season lacked enough data to build more complex models. Initial attempts to implement multivariate algorithms with the summer data have not resulted in significant improvements to NRMSE.

– The PA-II will be useful in estimating $PM_{2.5}$ extremes, given the relatively large expected CMD. PA-II corrections with RH, Temperature, and Dew Point should take account the strong non-linearity of these systems and inherant complexity of aerosol dynamics including composition, and particle size distribution.

References

Gani et al. "Submicron Aerosol Composition in the World's Most Polluted Megacity: the Delhi Aerosol Supersite Study." Atmospheric Chemistry and Physics, vol. 19, no. 10, 22 May 2019, pp. 6843–6859.

Jaiprakash et al. "Chemical Characterization of PM1.0 Aerosol in Delhi and Source Apportionment Using Positive Matrix Factorization." Environmental Science and Pollution Research, vol. 24, no. 1, Jan. 2017, pp. 445–462.

Malings et al. "Fine Particle Mass Monitoring with Low-Cost Sensors: Corrections and Long-Term Performance Evaluation." Aerosol Science and Technology, 10 June 2019.

Acknowledgments

We would like to thank the staff at the US Department of State at the Embassy in New Delhi, especially Ravindra Srivastava, for providing high quality reference grade data, and housing the PA-II units.

We would also like to our colleague from UT Austin, Sahil Bhandari, and our collegues from IIT Delhi, especially Purushottam Kumar and Nisar Baig for maintaing the site.

PLANET TEXAS 2050

A UT Grand Challenge

Long-Term Performance Evaluation of the PurpleAir PA-II Sensor in New Delhi, India

Mark Campmier [1*], Shahzad Gani [1], Joshua Apte [1]

[1] Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, Texas, USA

*mcampmier@utexas.edu

Motivation

– New Delhi, India is the most polluted megacity in the world, with annually averaged $PM_{2.5}$ measurements of 140 μg/m³. Episodic events in the Autumn and Winter can exceed 1000 μg/m³.

– A dearth of continuous monitoring sites hinders understanding of intra-urban varability, and the relative contributions of local and regional sources.

– Advances in low cost Optical Particle Counters (OPC), such as the popular PurpleAir PA-II offer a relatively cost effective means of establishing a dense urban network. However, OPCs have well-documented sources of error, and many newer OPCs such as the PA-II have not been extensively evaluated in extremely polluted environments.

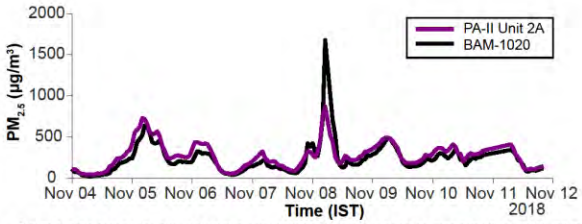

Figure 1: Time series of an extreme event in early Novemeber. Large regional sources coinciding with Diwali and unfavorable meterology lead to days of concentrations exceeding 300 μg/m³ peaking at 1800 μg/m³.

Methodology

– From June 2018 - July 2019, we collocated a BAM-1020 at the US Embassy in New Delhi with 2 PA-II sensor package, each consisting of 2 PM S5003 Plantower units.

– We evaluated several univariate and multivariate calibration algorithms using a 70:30 test to train split with cross-validation.

Univariate Models

– **Ordinary Least Squares (OLS):** Minimizes square of the vertical distance between the predicted y-axis variable and measured y-axis variable. Assumes the y-axis is the dependent variable with inherent measurement uncertainity, and the x-axis is the independent variable with no inherent measurement uncretainity.

– **Othogonal Distance Regression (ODR):** Minimizes square of the orthogonal distance between the predicted y-axis variable and true y-axis variable. The x-axis and y-axis variables are interchangeable and can be dependent variables with inherent measurement uncertainity, such as the BAM-1020 hourly measurements. Also called Deming or Total Least Squares Regression.

Multivariate Models

– **Multiple Linear Regression (MLR):** Extension of ODR with multiple explantory variables. While a relatively efficient and interpretable model, it can have difficulty accounting for inate piecewise behavior and can require extensive feature engineering.

– **Random Forest Regression (RFR):** A bagging ensemble of Regression Tree models - optimizes prediction based on naturally occuring splits in the response variable. More time intensive fitting than MLR, and contains more inherant model complexity, but is highly adaptive to non-linear trends and somewhat intrepretable.

Table 1: Comprehensive summary of univariate model performance on the evaluation datasets. NRMSE refers to Normalized Root Mean Square Error, the square root of the mean square difference between the predicted and measured response variable divided by the mean of the measured response variable. "All" refers to models built with all seasons. Performance in the Fall, Winter, and Spring is much more robust than All, while Summer and Monsoon are weaker.

		1-hr			24-hr		
		Slope	R^2	NRMSE	Slope	R^2	NRMSE
All	OLS	1.2	0.82	0.4	1.17	0.81	0.43
	ODR	1.34	0.80	0.43	1.32	0.78	0.45
Fall	OLS	1.04	0.85	0.26	0.95	0.86	0.24
	ODR	1.12	0.84	0.28	1.12	0.84	0.28
Winter	OLS	1.15	0.90	0.15	1.08	0.86	0.18
	ODR	1.27	0.89	0.15	1.23	0.87	0.18
Spring	OLS	1.03	0.91	0.21	1.00	0.85	0.27
	ODR	1.09	0.91	0.2	1.08	0.86	0.27
Summer	OLS	0.46	0.50	0.48	0.44	0.56	0.51
	ODR	0.54	0.48	0.48	0.53	0.56	0.51
Monsoon	OLS	0.21	0.13	0.31	0.21	0.06	0.27
	ODR	0.36	0.15	0.31	0.41	0.00	0.29

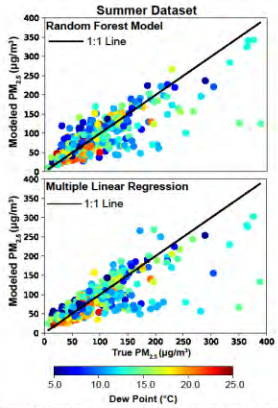

Figure 3: Scatter plots of modeled concentrations plotted against the BAM-1020 reference values, and colored for Dew Pont for RFR and MLR models.

– Univate models only use the block averaged PurpleAir PA-II PM2.5_CF_1 (μg/m³) channel to build the regression with the BAM-1022 1-hr concentration provided by AirNow.

– Model performance varies with season, with especially strong performance in the Spring. There is little difference between 1-hr and 24-hr model performance.

– Multivariate models used the PN0.3, PN0.5, PN1.0, PN2.5 particle number count bins, and Dew Point to achieve the strongest peformance

– The Summer dataset multivariate models had more robust performance than the univariate models - R^2 improved from 0.50 to 0.70 and 0.68, and NRMSE improved from 0.48 to 0.38 and 0.42 for RFR and MLR respectively.

Results

– All four sensors showed strong 1:1 agreement ($R^2 > 0.95$). For calibration purposes, it is only one representative sensor is necessary, called "PA-II Unit 2A." Hourly BAM-PA performance is strongly influenced by meterology, especially relative humidity (RH).

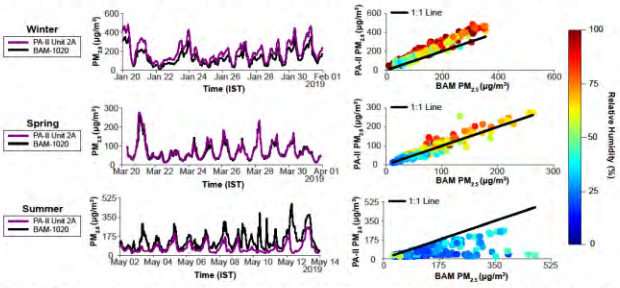

Figure 2: Time series of continuous monitoring data in three seasons, juxtaposed with 1-hr PA-II Unit 2A plotted against the collocated US Embassy BAM-1020 and colored for RH from the same time period.

Conclusions

– Univate models demonstrate low error and high correlation during the Autumn, Winter, and Spring. Autumn and Winter model slopes show the PA-II overestimate by more than 10% on 1-hr data. Both seasons are correlated with high RH. Daily data model perfromance is comparable to the 1-hr models.

– The multivariate models were useful in improving performance - both correlation and error - in the Summer, and could be deployed to adapt to evolving meterological regimes, aerosol composition, and sensor decay if univariate models are not robust.

References

Gani et al. "Submicron Aerosol Composition in the World's Most Polluted Megacity: the Delhi Aerosol Supersite Study." Atmospheric Chemistry and Physics, vol. 19, no. 10, 22 May 2019, pp. 6843–6859.

Jaiprakash et al. "Chemical Characterization of PM1.0 Aerosol in Delhi and Source Apportionment Using Positive Matrix Factorization." Environmental Science and Pollution Research, vol. 24, no. 1, Jan. 2017, pp. 445–462.

Mailings et al. "Fine Particle Mass Monitoring with Low-Cost Sensors: Corrections and Long-Term Performance Evaluation." Aerosol Science and Technology, 10 June 2019.

Acknowledgments

We would like to thank the AirNow staff at the US Department of State at the Embassy in New Delhi, especially Ravindra Srivastava, for providing high quality reference grade data, and housing the PA-II units .

We would also like to our collegue from UT Austin, Sahil Bhandari, and our collegues from IIT Delhi, especially Purushottam Kumar and Nisar Baig for maintaing the site.