

Performance Evaluations of PM_{2.5} Sensors in Research Triangle Park, NC

Samuel G. Frederick (ORAU/SSC) – frederick.samuel@epa.gov

Karoline K. (Johnson) Barkjohn (ORISE) – <u>johnson.karoline@epa.gov</u>

Cortina Johnson (Jacobs) – <u>johnson.cortina@epa.gov</u>

Robert W. Yaga (Jacobs) – yaga.robert@epa.gov

Andrea L. Clements (U.S. EPA) – <u>clements.andrea@epa.gov</u>



Overview

>Introduction

- Motivation for air sensor evaluation
- Particulate matter measurement and evaluation methods
- Evaluation site details
- Sensors chosen for this study

➤ Performance Evaluation of Sensor Types

- Data collection time period and uptime
- Precision within a group of identical sensors
- Performance against regulatory-grade instrument

≻Conclusions

- Summary of results
- Future work



Motivation: Air Quality Monitoring

- Rapidly increasing number of air quality monitors entering the market
 - Provide users with real-time air quality measurement
- Often used for non-regulatory applications which do not necessarily require the same rigorous standards of accuracy and reliability
 - Evaluation of air quality sensor measurements against regulatory methods is critical to ensuring measurement accuracy and precision











Sensor Evaluation: Contextualizing this Work

- Ongoing effort to evaluate sensor performance
- Evaluation covers a large number of sensor models and pollutants
 - These devices measure a varying number of pollutants including the following:*

```
• PM_{2.5}
• PM_{10}
• O_3
• NO_2, NO_x
• SO_2
• CO
```

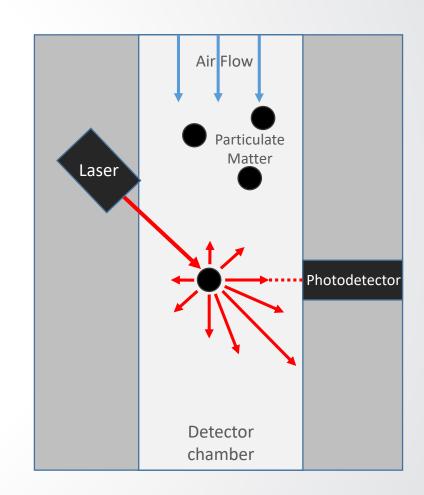
- Oftentimes devices also record meteorological data:*
 - Temperature
 - Relative Humidity

^{*}Not a comprehensive list of pollutants and meteorological values measured by air quality sensors



Particulate Matter Measurement

- Air quality sensors use optical scattering to measure PM. Despite cost effectiveness, this method presents limitations.
 - Optical scattered light proportional to the square of the size of particles for a fixed scattering angle
 - The size of volume of particles is converted to PM mass concentration often using a laboratory calibration
 - PM is often hygroscopic and can uptake water causing particles to grow in size
 - Increased particle size often results in an over-estimation of PM mass as compared to a dried sample resulting in a bias correlated to relative humidity (RH) and temperature
 - Regulatory-grade instruments are not as susceptible to this artifact because they are operated under controlled conditions: sample RH (30-40 %) and temperature (20 – 23 °C)





AIRS Evaluation Site

Air quality sensors collocated with regulatory-grade monitor

- Air sensors are tested in triplicate (or more)
 - Data recorded at ~1-min time resolution
 - Computed 1-hour averages for comparison
- Teledyne T640x PM Mass Monitor
 - Federal Equivalent Method (FEM) for PM_{2.5}
 - Scattered light spectrometry (heating element to reduce PM water content)
 - Default logging interval: 10-minute rolling averages (logged each minute)
 - Computed 1-hour averages for comparison





Goals of this Study

- 1-hour averaged PM_{2.5} performance for six air quality sensor types
 - Collocated sensors at AIRS evaluation site on RTP campus (Collocation period ≥ 30 days)
 - Determine precision within sensor types (≥ 3 sensors of the same type)
 - Evaluate performance against T640x
 FEM PM_{2.5} measurements
 - Inform users regarding future sensor implementation both internal and external to EPA

Device Manufacturer	Device Name	Internal PM Sensor	Time Resolution
Vaisala	AQT420	OEM	1 minute
Aeroqual	AQY	NOVA SDS011	1 minute
Clarity	Node-S	Plantower PMS6003	3 minutes
SENS-IT	Ramp	Plantower PMS5003	15 seconds
АРТ	Maxima	Plantower PMSA003	30 seconds
PurpleAir	PA-II-SD	Plantower PMS5003	2 minutes



Goals of this Study

- 1-hour averaged PM_{2.5} performance for six air quality sensor types
 - Collocated sensors at AIRS evaluation site on RTP campus (Collocation period ≥ 30 days)
 - Determine precision within sensor types (≥ 3 sensors of the same type)
 - Evaluate performance against T640x
 FEM PM_{2.5} measurements
 - Inform users regarding future sensor implementation both internal and external to EPA

Device Manufacturer	Device Name	Internal PM Sensor	Time Resolution
Vaisala	AQT420	OEM	1 minute
Aeroqual	AQY	NOVA SDS011	1 minute
Clarity	Node-S	Plantower PMS6003	3 minutes
SENS-IT	Ramp	Plantower PMS5003	15 seconds
АРТ	Maxima	Plantower PMSA003	30 seconds
PurpleAir	PA-II-SD	Plantower PMS5003	2 minutes



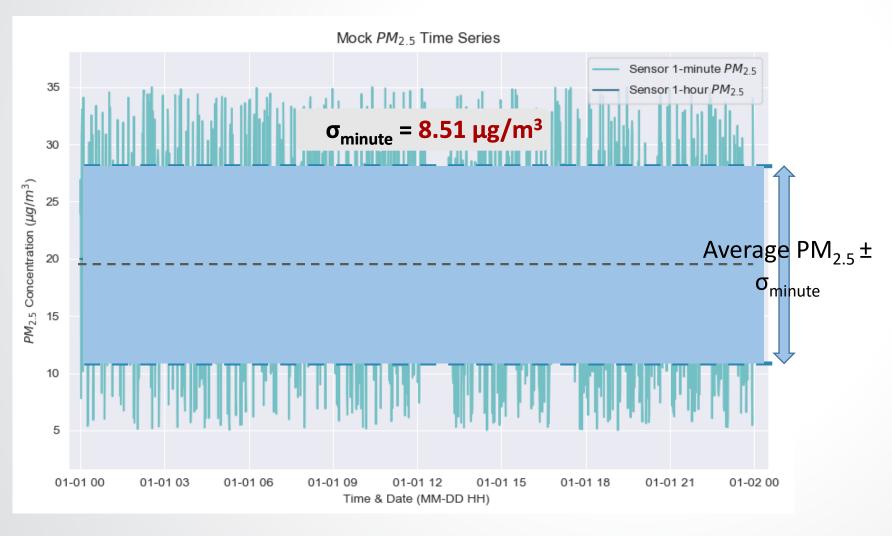
Variation in Minute-Resolution Data



 Data collected at minute intervals has a high degree of variation



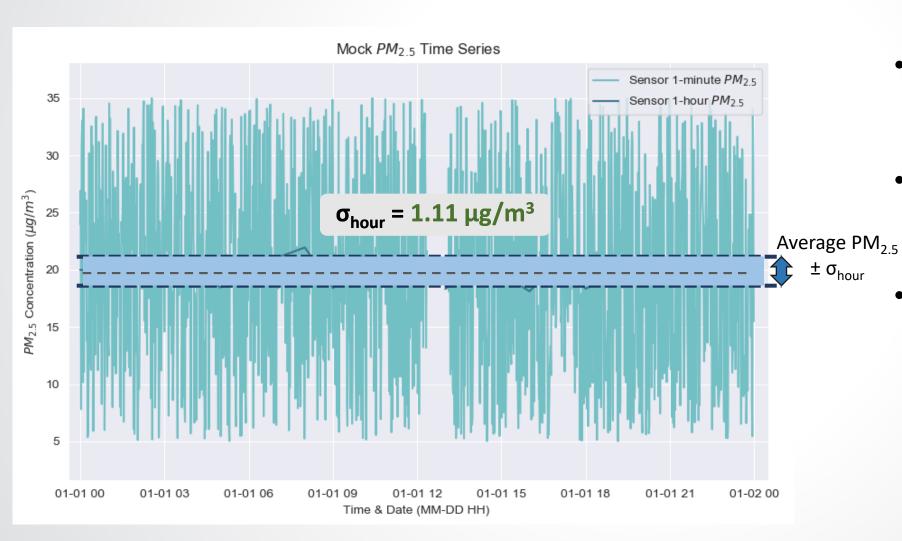
Variation in Minute-Resolution Data



- Data collected at minute intervals has a high degree of variation
- Standard deviation quantifies variation in the data

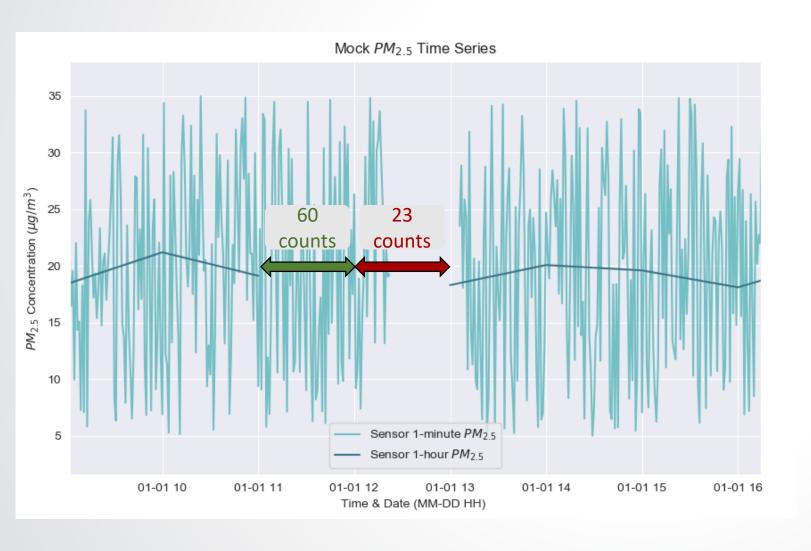


Variation in 1-Hour Averaged Data



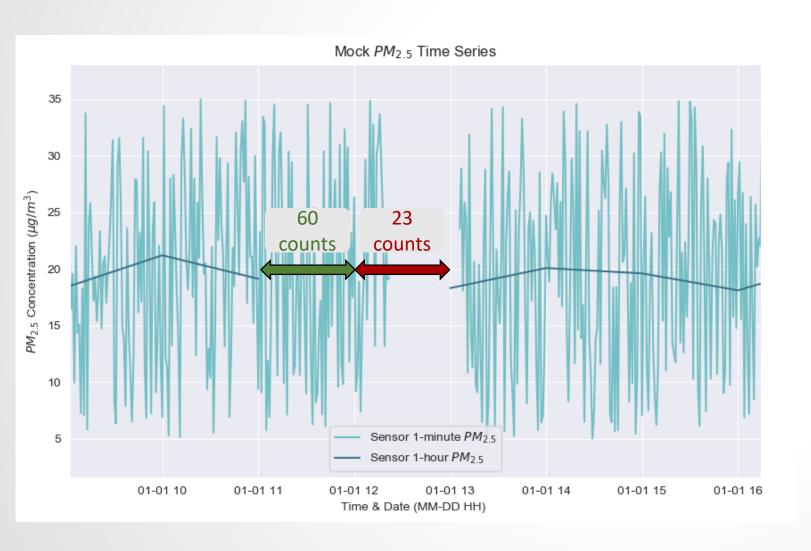
- Data collected at minute intervals has a high degree of variation
- Standard deviation quantifies variation in the data
- Computing 1-hour averages significantly reduces this variation.





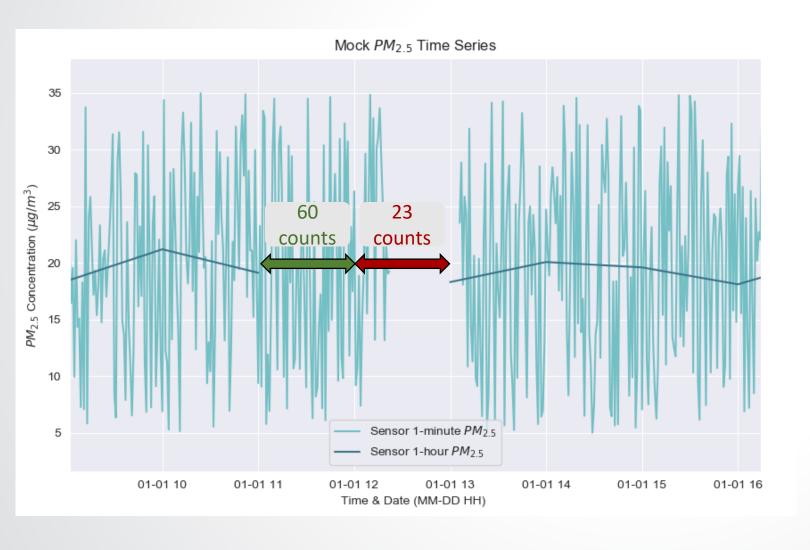
- Data may be incomplete for a given hour due to technical fault at the sensor level (e.g. loss of connection) or site level (e.g. loss of power, meteorological events)
- We want the 1-hour average to be representative of the 1-hour time period so we impose a completeness criteria requiring most of the data to be present before compute a 1-hour average





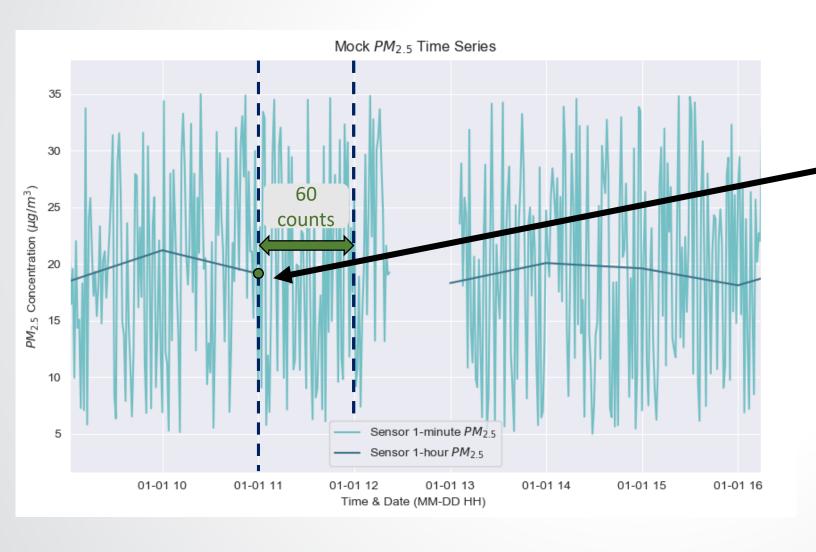
- Data may be incomplete for a given hour due to technical fault at the sensor level (e.g. loss of connection) or site level (e.g. loss of power, meteorological events)
- Requires enough data exist for a given hour to compute a 1-hour average
- For this work, we required data recorded within a given hour to be ≥90% complete in order to compute 1-hour averages





- Completeness Example
 - For 1-minute data, 60 data points are generated each hour.
 - 90% completeness requires
 ≥ 54 data points per hour
 to compute 1-hour average

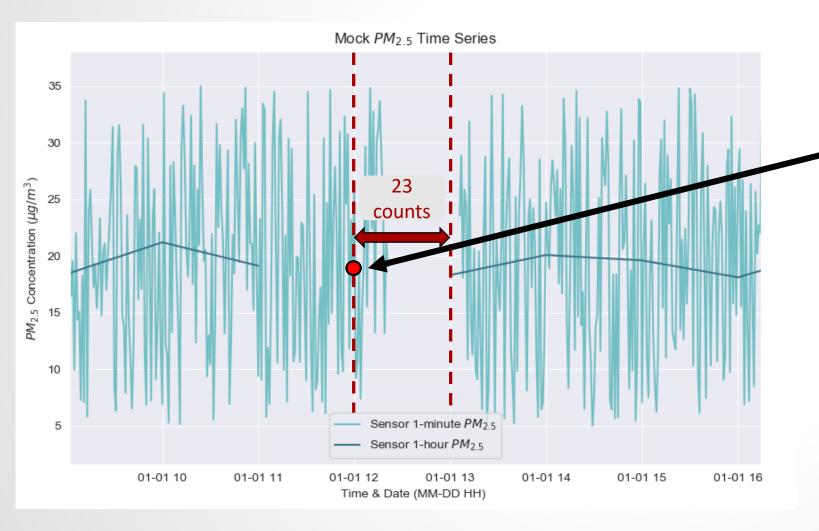




Completeness Example

For 1-minute data, this 1-hour average is computed from data collected between 11:00 – 11:59 and is assigned to 11:00 average.





- Completeness Example
 - Hourly periods with an insufficient number of data points are not averaged to 1-hour values



Sensor Performance Methods: Precision

Quantify agreement between sensors of the same type

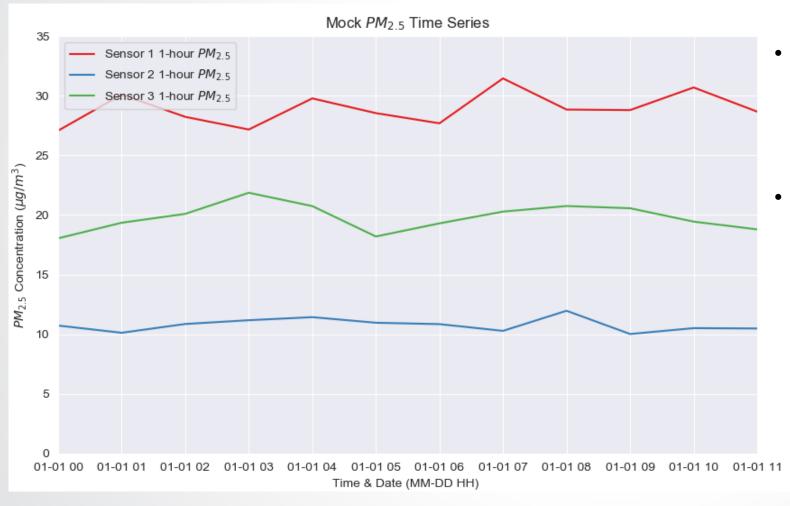
- Standard Deviation ($\sigma_{sensor\ type}$)

 Measure of the total variation in PM_{2.5} measurements calculated for all sensors of the same type
- Coefficient of Variation (CV_{sensor type})

Defined as $(\sigma_{sensor\ type})$ / (Mean PM_{2.5} concentration for sensor type), measure of sensor measurement variation relative to mean concentration value



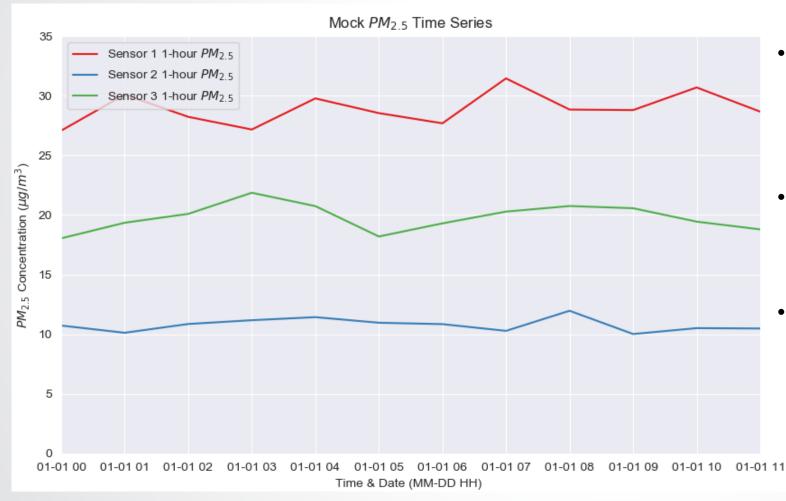
Measuring Sensor Precision



- An individual sensor's standard deviation doesn't inform how sensor measurements compare against other devices of the same type
- We compute the standard deviation of 1hour measurements across the entire set of data for a given sensor type
 - $\sigma_{sensor\ type}$



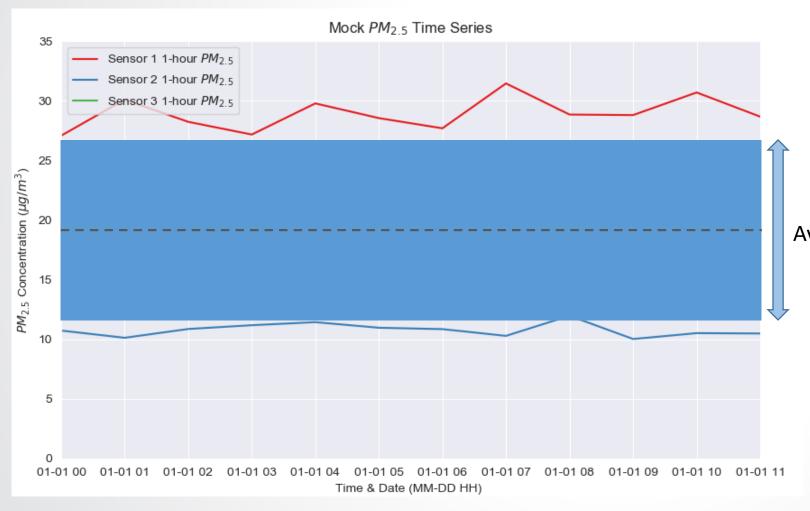
Measuring Sensor Precision



- An individual sensor's standard deviation doesn't inform how sensor measurements compare against other devices of the same type
- We compute the standard deviation of 1hour measurements across the entire set of data for a given sensor type
 - $\sigma_{sensor type}$
- Because the measurements made by the individual sensors range in offset bias and sensitivity, $\sigma_{sensor\ type}$ is normalized by the average concentration value for all sensors in a given type.
 - $CV_{sensor\ type} = \frac{G_{sensor\ type}}{Mean\ PM_{2.5} concentration}$



Measuring Sensor Precision: Low Precision



- $\sigma_{\text{sensor type}} = 7.35 \, \mu \text{g/m}^3$
- Average $PM_{2.5} = 19.79 \,\mu g/m^3$

•
$$CV = \frac{7.35 \ \mu g/m^3}{19.79 \ \mu g/m^3} \ x \ 100 = 37.14\%$$

Average $PM_{2.5} \pm \sigma_{sensor type}$



Measuring Sensor Precision: High Precision



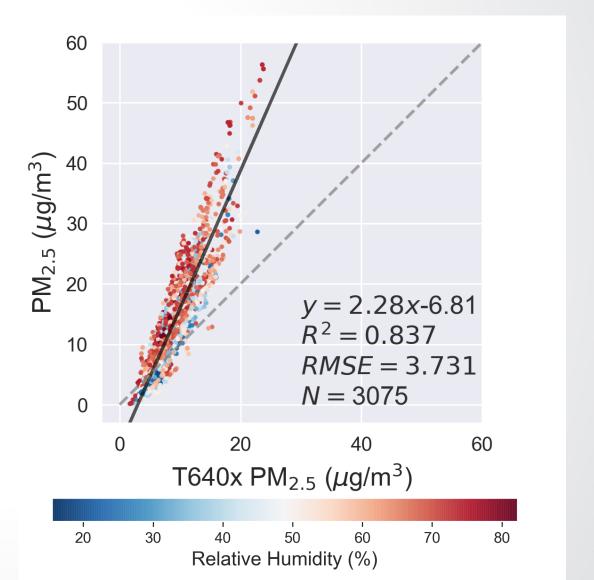
- $\sigma_{\text{sensor type}} = 1.95 \, \mu \text{g/m}^3$
- Average $PM_{2.5} = 13.04 \, \mu g/m^3$

•
$$CV = \frac{1.95 \,\mu g/m^3}{13.04 \,\mu g/m^3} \times 100 = 14.97\%$$

Average $PM_{2.5} \pm \sigma_{sensor type}$



Quantify how sensors perform against T640x regulatory monitor and monitor relative humidity influence

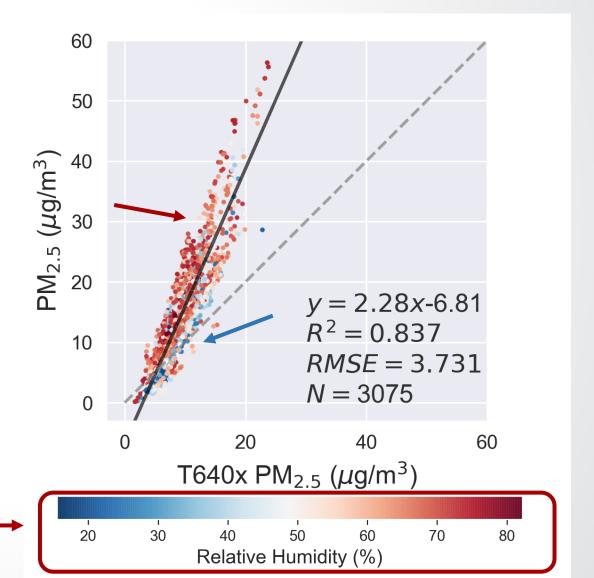




Quantify how sensors perform against T640x regulatory monitor and monitor relative humidity influence

Influence of relative humidity

1-hour averages are colored based on the average relative humidity experienced that hour. A strong separation of red and blue dots indicate a strong influence of humidity on the sensor measurement.

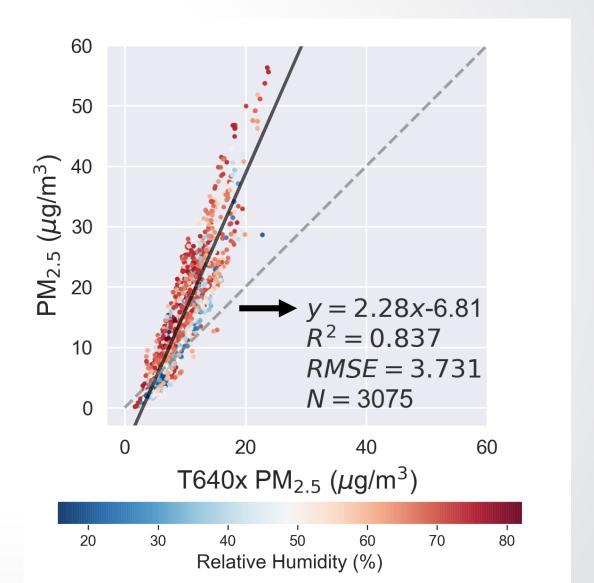




Quantify how sensors perform against T640x regulatory monitor and monitor relative humidity influence

Linear Regression (y = ax+ b)
 Slope tells us whether the sensor is under-reporting or over-reporting the monitor, intercept gives us a measure of the bias relative to the monitor

Slopes (a) near 1 and intercepts (b) near 0 are desired.

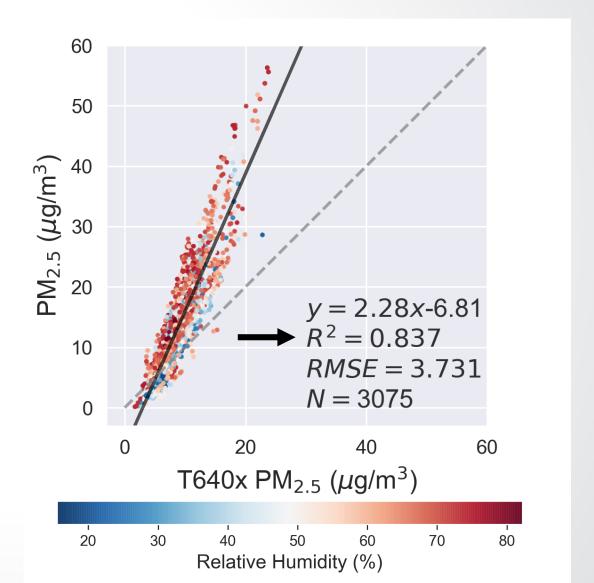




Quantify how sensors perform against T640x regulatory monitor and monitor relative humidity influence

- Linear Regression (y = ax+ b)
 Slope tells us whether the sensor is under-reporting or over-reporting the monitor, intercept gives us a measure of the bias relative to the monitor
- Coefficient of determination (R²)
 Measure of linear correlation between sensor PM values and regulatory monitor values

R² values range between 0 and 1. Values near 1 indicate strong linear relationships.

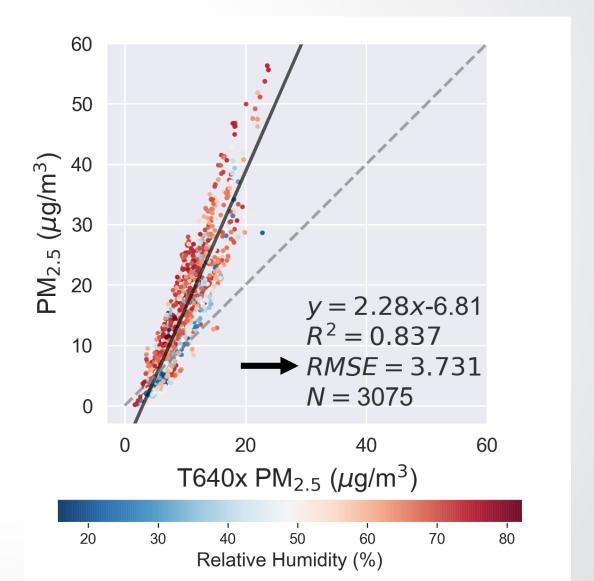




Quantify how sensors perform against T640x regulatory monitor and monitor relative humidity influence

- Linear Regression (y = ax+ b)
 Slope tells us whether the sensor is under-reporting or over-reporting the monitor, intercept gives us a measure of the bias relative to the monitor
- Coefficient of determination (R²)
 Measure of linear correlation between sensor PM values and regulatory monitor values
- Root Mean Square Error (RMSE)
 Measure of the spread of sensor PM data, squaring penalizes outliers

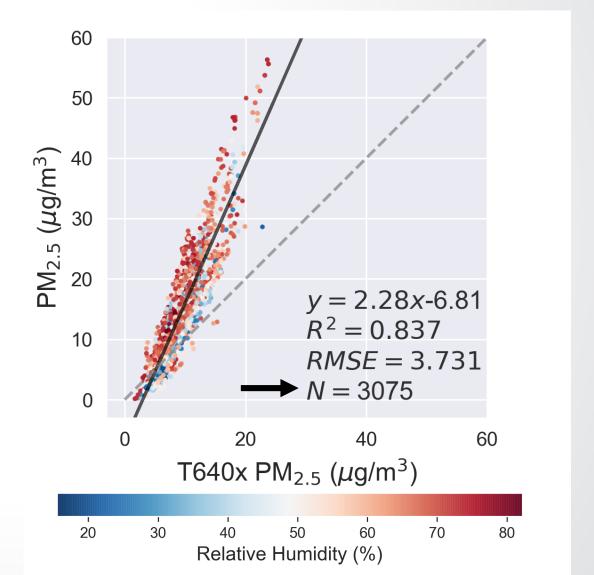
Lower RMSE is desirable.





Quantify how sensors perform against T640x regulatory monitor and monitor relative humidity influence

- Linear Regression (y = ax+ b)
 Slope tells us whether the sensor is under-reporting or over-reporting the monitor, intercept gives us a measure of the bias relative to the monitor
- Coefficient of determination (R²)
 Measure of linear correlation between sensor PM values and regulatory monitor values
- Root Mean Square Error (RMSE)
 Measure of the spread of sensor PM data, squaring penalizes outliers
- Total Number of Datapoints (N)
 The number of 1-hour averaged data points that satisfy the 90% completeness criteria for computing 1-hour averages





Evaluated Sensors: Vaisala AQT420

Sensor Number	Data Collection Period
1	2018/06/22 – 2018/07/31
2	2018/06/15 – 2018/07/31
3	2018/06/22 – 2018/07/31

Pollutants Measured:

• Cost:

~\$5200

• Time Resolution:

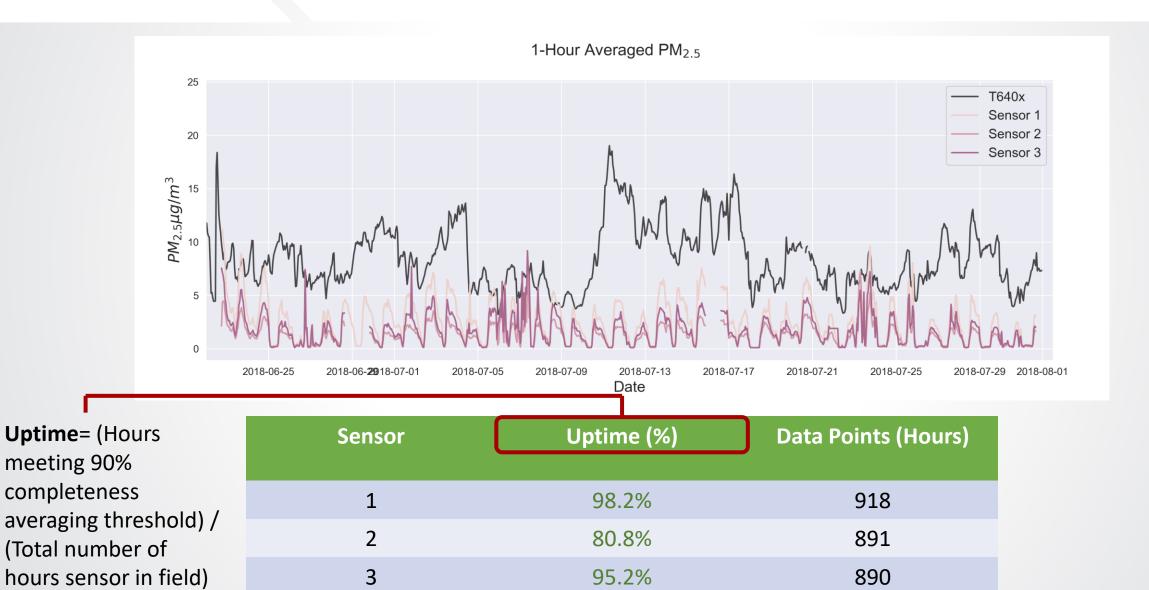
1-minute



AQT420 - Vaisala

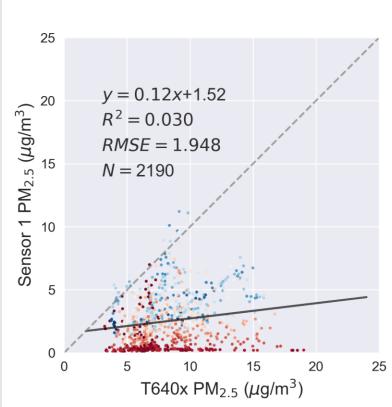


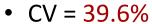
AQT420: Time Series and Uptime





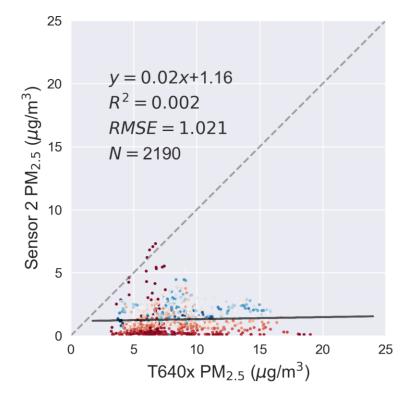
AQT420: Performance against Regulatory Monitor

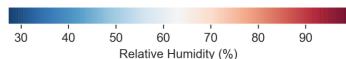


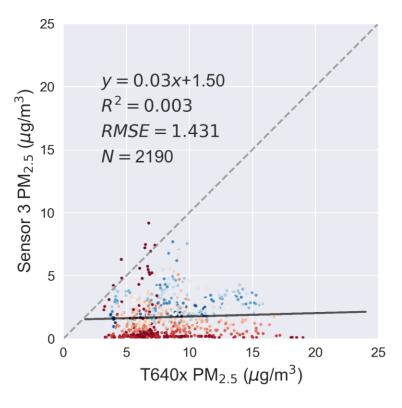


• $\sigma_{sensor\ type} = 0.73\ \mu g/m^3$

Vaisala Sensor vs. T640x Performance (Hourly PM_{2.5})









Vaisala AQT420: Takeaways

- Precision:
 - CV (39.6%) and $\sigma_{sensor\ type}$ (0.73 µg/m³) indicate poor agreement between individual AQT420 devices
- Accuracy:
 - Correlation between AQT420 and T640x PM_{2.5} measurements indistinguishable (R² ≤ 0.03)
- The AQT420 under-reports $PM_{2.5}$ concentrations relative to T640x by >90%
- Observations for users:
 - Strong correlation between AQT420 $PM_{2.5}$ values and temperature and relative humidity



Evaluated Sensors: Aeroqual AQY

Sensor Number	Data Collection Period
1	2019/05/02 – 2019/06/10
2	2019/05/02 – 2019/06/10
3	2019/05/02 – 2019/05/25

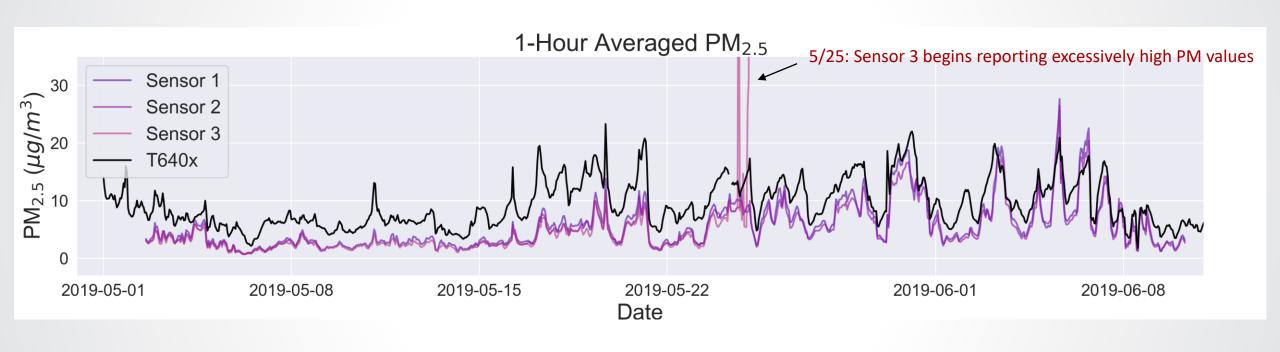
- Pollutants Measured
 PM_{2.5}, O₃, NO₂
- Cost ~\$3000
- Time Resolution: 1-minute



AQY - Aeroqual



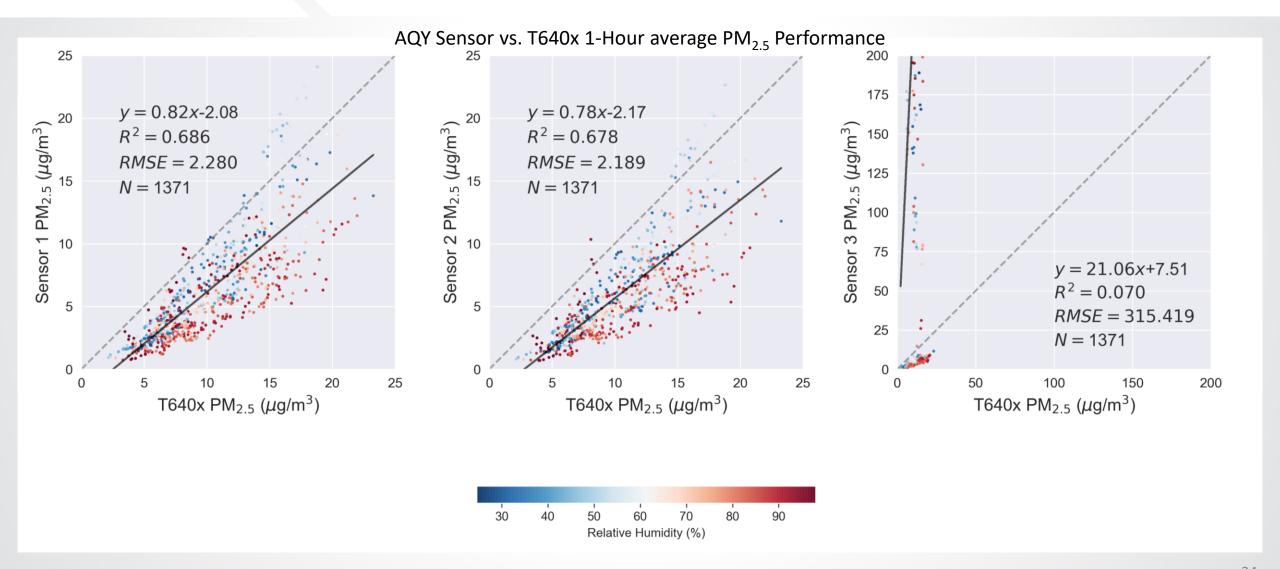
AQY: Time Series and Uptime



Sensor	Uptime (%)	Data Points (Hours)
1	99.9%	933
2	85.9%	802
3	58.1%	543

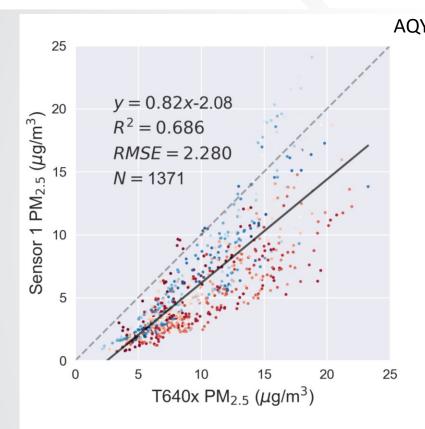


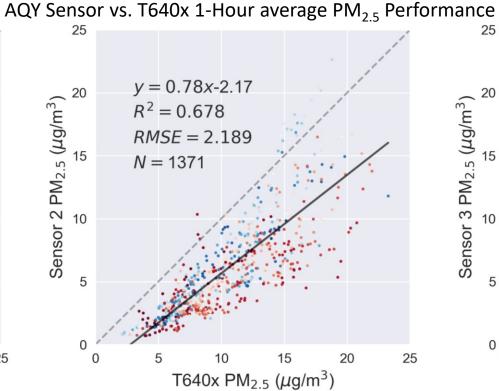
AQY: Performance against Regulatory Monitor

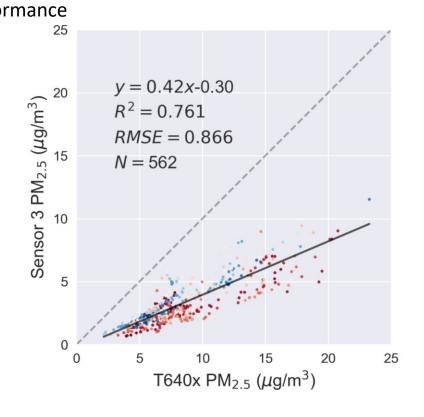




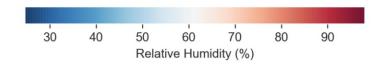
AQY: Performance against Regulatory Monitor







- Omitting Sensor 3 values past 5/25
 - CV = 28.0%
 - $\sigma_{sensor\ type} = 1.38\ \mu g/m^3$





Aeroqual AQY: Takeaways

Precision:

• CV (28.0%) and $\sigma_{sensor\ type}$ (1.38 µg/m³) indicate moderate agreement between individual AQY devices

Accuracy:

- Moderate Correlation (R²~0.68-0.76) between AQY and T640x PM_{2.5} measurements
- The AQY under-reports PM_{2.5} concentrations relative to T640x by \sim (20-60%)

Observations for users:

- High RH (> ~70%) causes increased underreporting of AQY
- Dramatic over-reporting of PM_{2.5} values observed for sensor #3 after 5/25, investigating potential causes



Evaluated Sensors: Clarity Node-S

Sensor Number	Data Collection Period
1-9	2019/06/10 – 2019/07/09

Pollutants Measured:

- Cost:
 - \$1000 annual subscription
- Time Resolution:

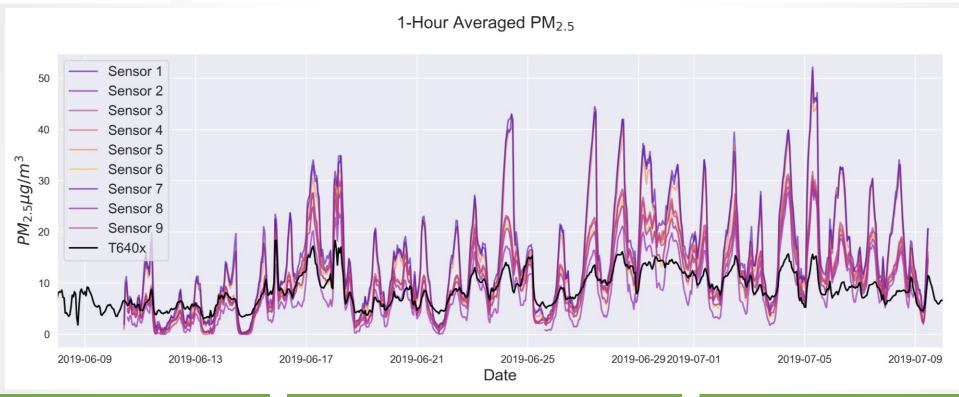
3 minutes



Clarity Node-S – Clarity Movement Co.



Clarity Node-S: Time Series and Uptime



Sensor	Uptime (%)	Data Points (Hours)
1	99.8%	696
2	99.7%	694
3	99.7%	695

Sensor	Uptime (%)	Data Points (Hours)
4	99.6%	693
5	99.8%	696
6	99.4%	693

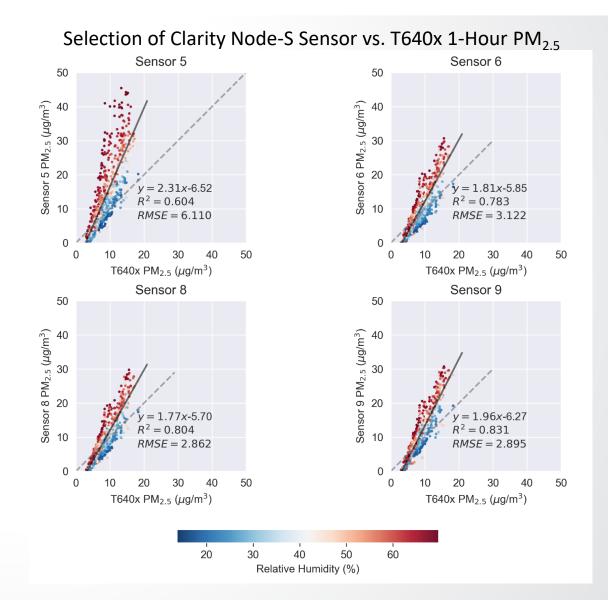
Sensor	Uptime (%)	Data Points (Hours)
7	99.7%	696
8	99.6%	694
9	99.6%	694



Clarity Node-S: Performance against Regulatory Monitor

Sensor Number	R² (1-hour PM _{2.5})	RMSE (1-hour PM _{2.5})
1	0.603	6.413
2	0.69	3.31
3	0.838	2.932
4	0.842	2.811
5	0.604	6.11
6	0.783	3.122
7	0.657	5.868
8	0.804	2.862
9	0.831	2.895

- Coefficient of Variation = 26.5%
- Standard Deviation = 3.15 μg/m³





Clarity Node-S: Takeaways

Precision:

• CV (26.5%) and $\sigma_{sensor\ type}$ (3.15 µg/m³) indicate moderate agreement between individual Node-S devices

Accuracy:

- Varying Correlation ($R^2 \sim 0.60$ -0.84) between Node-S and T640x PM_{2.5} measurements, some devices indicated moderate correlation ($R^2 \gtrsim 0.70$)
- The Node-S generally over-reports $PM_{2.5}$ concentrations relative to T640x but is highly variable (~80 150%)

Observations for users:

 Distinct relative humidity bias, increasing RH causes increased overreporting by Node-S



Evaluated Sensors: SENSIT RAMP

Sensor Number	Data Collection Period
1-6	2019/02/14 – 2019/04/10
7-9	2019/04/10 - 2019/05/30

Pollutants Measured:

• Cost:

~\$3000

Time Resolution:

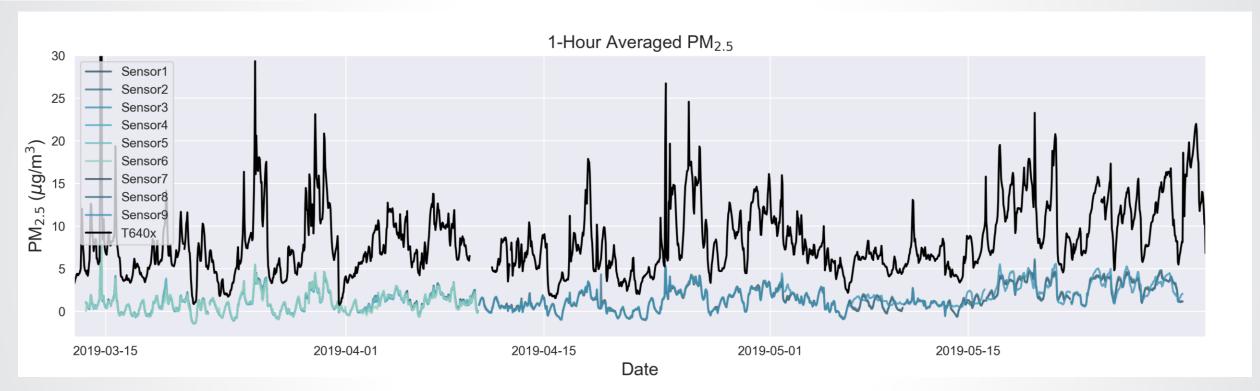
15 seconds



Ramp – SENSIT Above, sensors 1-6 deployed at AIRS



RAMP: Time Series and Uptime



Sensor	Uptime (%)	Data Points (Hours)
1	98.8%	660
2	98.8%	660
3	98.8%	660

Sensor	Uptime (%)	Data Points (Hours)
4	98.8%	660
5	98.8%	660
6	98.8%	659

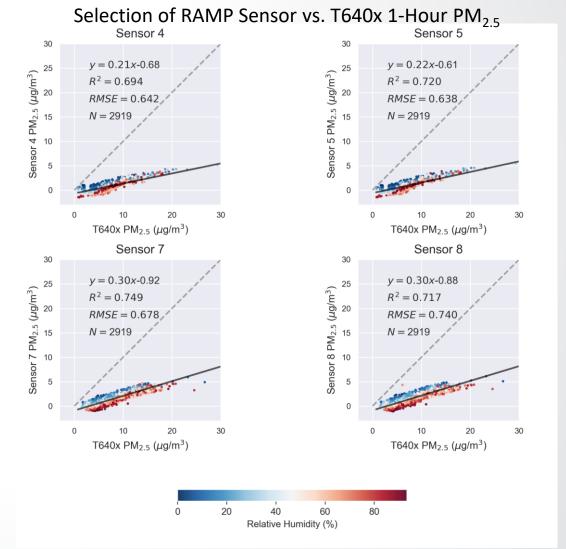
Sensor	Uptime (%)	Data Points (Hours)
7	99.2%	1188
8	99.2%	1183
9	99.2%	1188



RAMP: Performance against Regulatory Monitor

Sensor Number	R ₂ (1-hour PM _{2.5})	RMSE (1-hour PM _{2.5})
1	0.729	0.633
2	0.722	0.646
3	0.703	0.642
4	0.694	0.642
5	0.720	0.638
6	0.665	0.695
7	0.749	0.678
8	0.717	0.740
9	0.771	0.657

- Sensor Precision: Devices 1-6
 - Coefficient of Variation = 12.06%
 - Standard Deviation = 0.12 μg/m³
- Sensor Precision: Devices 7-9
 - Coefficient of Variation =15.3%
 - Standard Deviation = 0.25 μg/m³





SENSIT RAMP: Takeaways

Precision:

• CV (12.06%, 15.3%) and $\sigma_{sensor\ type}$ (0.12, 0.25 µg/m³) indicate good agreement between individual Ramp devices

Accuracy:

- Moderate Correlation (R²~0.66-0.77) between RAMP and T640x PM_{2.5} measurements
- The RAMP under-reports $PM_{2.5}$ concentrations relative to T640x by ~70-80%

Observations for users:

 Distinct relative humidity bias, increasing RH causes increased underreporting by RAMP



Evaluated Sensors: APT Maxima

Sensor Number	Data Collection Period
2-7, 9	2019/05/13 – 2019/06/10
8	2019/05/13 – 2019/05/30

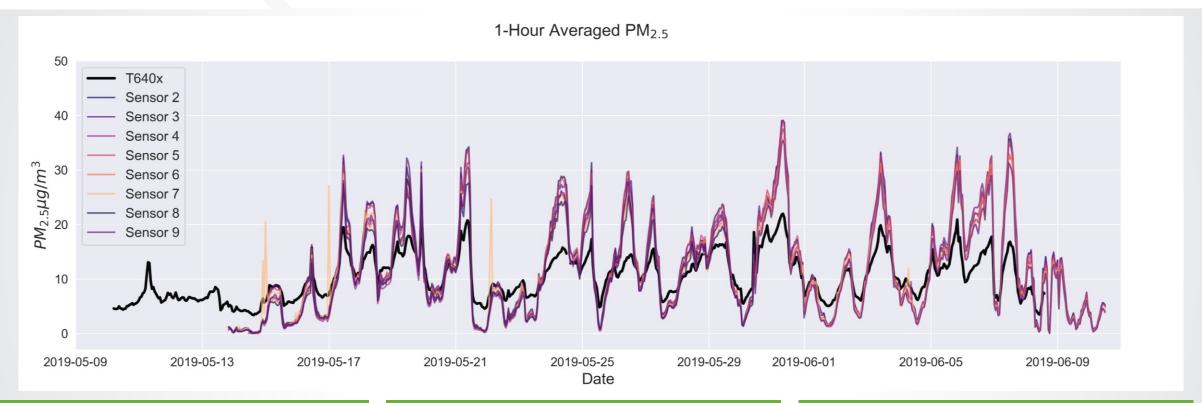
- Pollutants Measured PM₁, PM_{2.5}, PM₁₀
- Cost ~\$350
- Time Resolution
 30 seconds



Maxima – Applied Particle Technology Above, Maxima devices 4-7 deployed at AIRS



Maxima: Time Series and Uptime



Sensor	Uptime (%)	Data Points (Hours)
2	88.8%	665
3	88.8%	665
4	88.8%	665

Sensor	Uptime (%)	Data Points (Hours)
5	86.9%	651
6	88.8%	665
7	88.8%	665

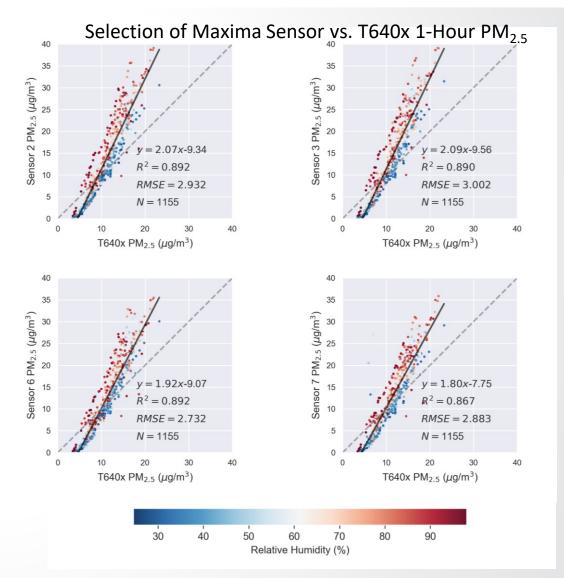
Sensor	Uptime (%)	Data Points (Hours)
7	88.8%	665
8	84.6%	400
9	88.8%	665



Maxima: Performance against Regulatory Monitor

Sensor Number	R ² (1-hour PM _{2.5})	RMSE (1-hour PM _{2.5})
2	0.892	2.932
3	0.890	3.002
4	0.886	2.835
5	0.888	2.933
6	0.892	2.732
7	0.867	2.883
8	0.913	2.112
9	0.894	2.645

- Coefficient of Variation = 9.22%
- Standard Deviation = $1.07 \,\mu\text{g/m}^3$





APT Maxima: Takeaways

Precision:

• CV (9.22%) and $\sigma_{sensor\ type}$ (1.07 µg/m³) indicate good agreement between individual Maxima devices

Accuracy:

- Good Correlation (R²~0.89) between Maxima and T640x PM_{2.5} measurements
- The Maxima has a large-magnitude intercept leading to an underestimation of $PM_{2.5}$ below ~10 µg/m³ and a large slope leading to over-reporting of $PM_{2.5}$ concentrations relative to T640x by ~90-110% above 10 µg/m³

Observations for users:

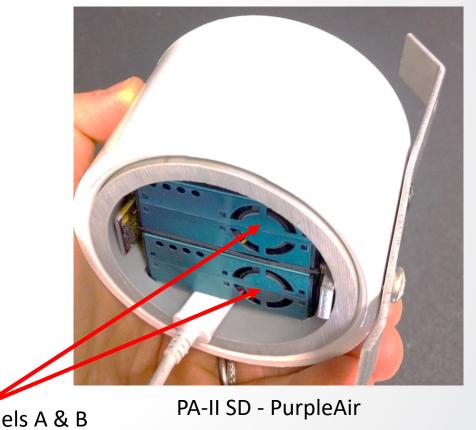
 Distinct relative humidity bias, increasing RH causes increased overreporting by Maxima



Evaluated Sensors: PurpleAir PA-II-SD

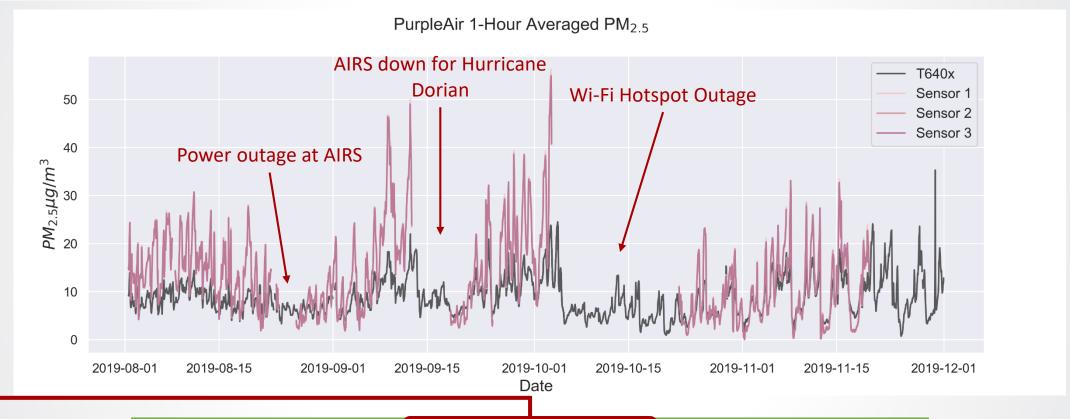
Sensor Number	Data Collection Period	
1	2019/08/01 – 2019/11/19	
2	2019/08/01 – 2019/11/19	
3	2019/08/01 – 2019/11/19	

- Pollutants Measured PM_{2.5}, PM₁₀
- Cost ~\$250
- Time Resolution2 minutes





PurpleAir: Time Series and Uptime



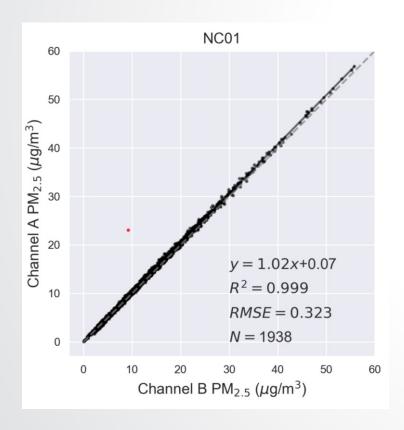
Uptime is artificially low based on site management decisions.

Sensor	Uptime (%)	Data Points (Hours)
1	73.16%	1938
2	73.17%	1939
3	73.22%	1941

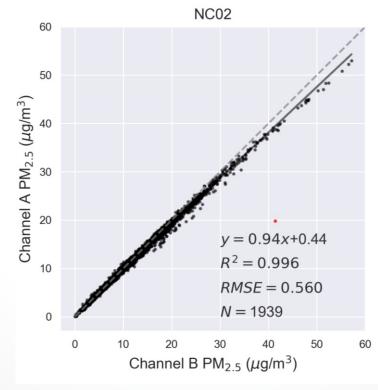


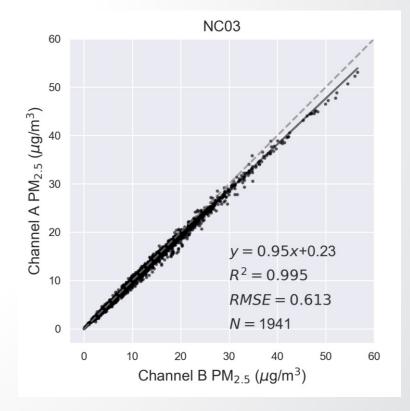
PurpleAir: Sensor-Channel Precision

- We determine the precision between sensor channels
 A and B to verify whether it is justified to average
 PM_{2.5} concentration values computed at two-minute intervals for both channels
- Ideal agreement between channels A and B would result in a linear one-to-one correlation. Plotting PM_{2.5} concentrations for channel A vs. channel B, we find good agreement between sensor channels.



Sensor Channel A & B Precision

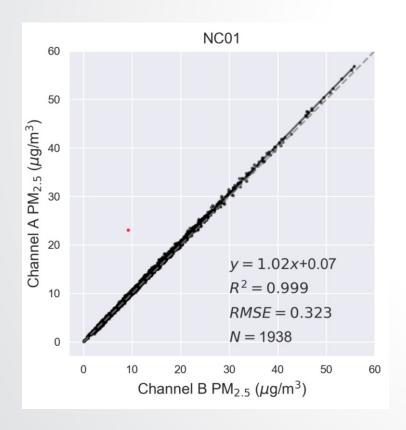




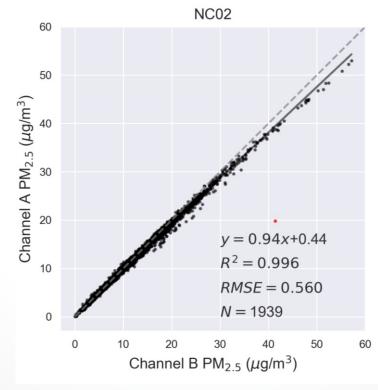


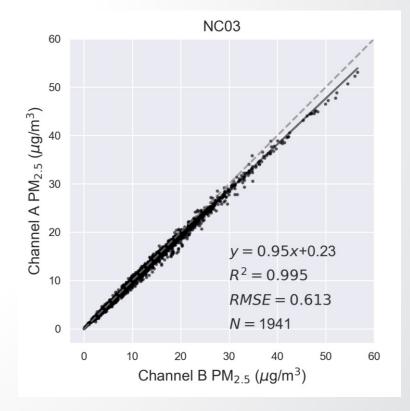
PurpleAir: Sensor-Channel Precision

- We determine the precision between sensor channels
 A and B to verify whether it is justified to average
 PM_{2.5} concentration values computed at two-minute intervals for both channels
- result in a linear one-to-one correlation. Plotting PM_{2.5} concentrations for channel A vs. channel B, we find good agreement between sensor channels.



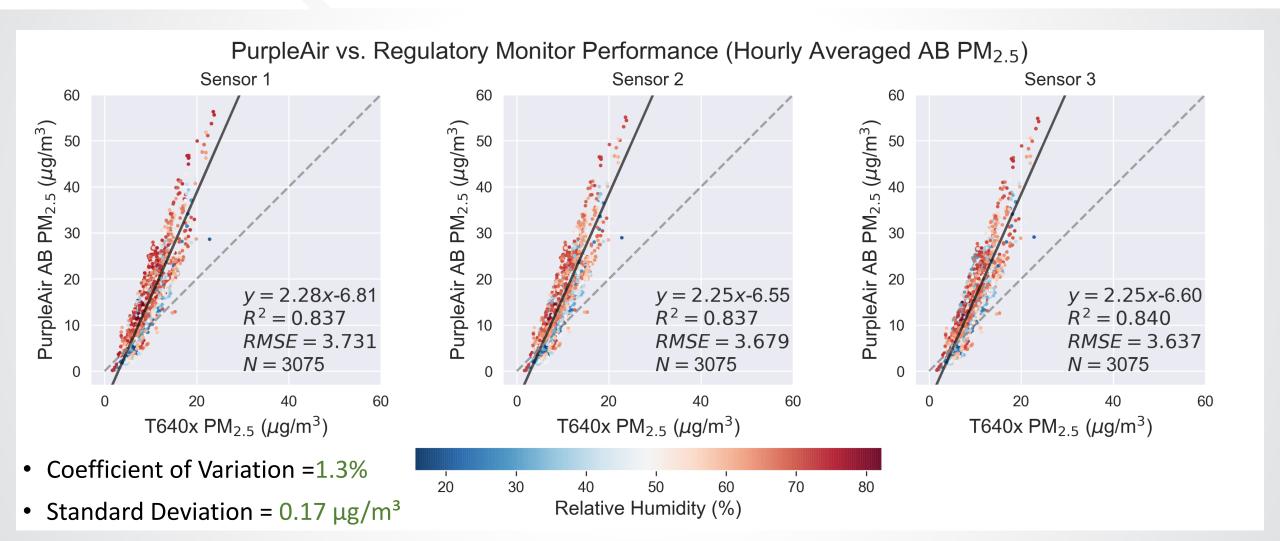
Sensor Channel A & B Precision







PurpleAir: A & B Averaged Performance against Regulatory Monitor





PurpleAir PA-II-SD: Takeaways

Precision:

• CV (1.3%) and $\sigma_{sensor\ type}$ (0.17 µg/m³) indicate good agreement between individual PA-II-SD devices

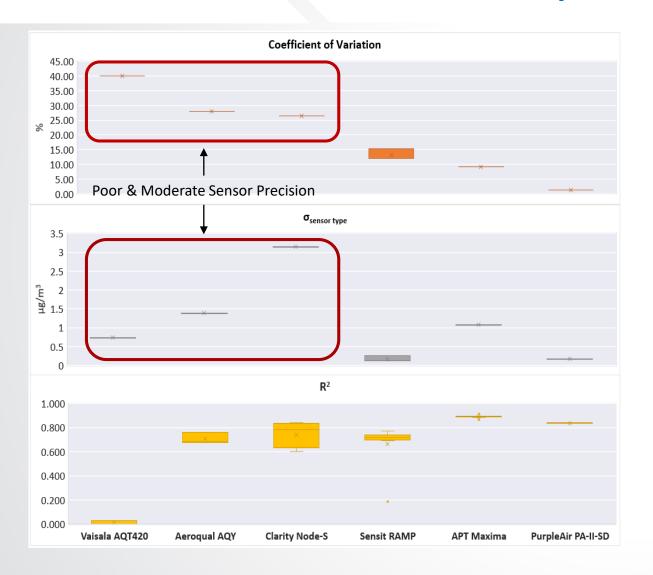
Accuracy:

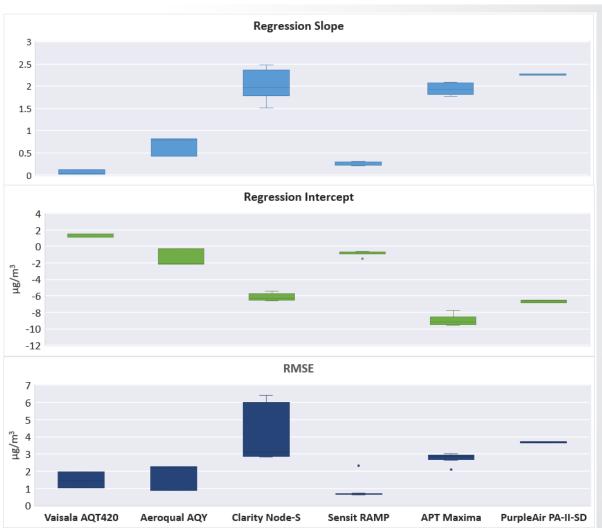
- Good Correlation (R²~0.84) between PA-II-SD and T640x PM_{2.5} measurements
- The PA-II-SD over-reports $PM_{2.5}$ concentrations relative to T640x by ~125%

Observations for users:

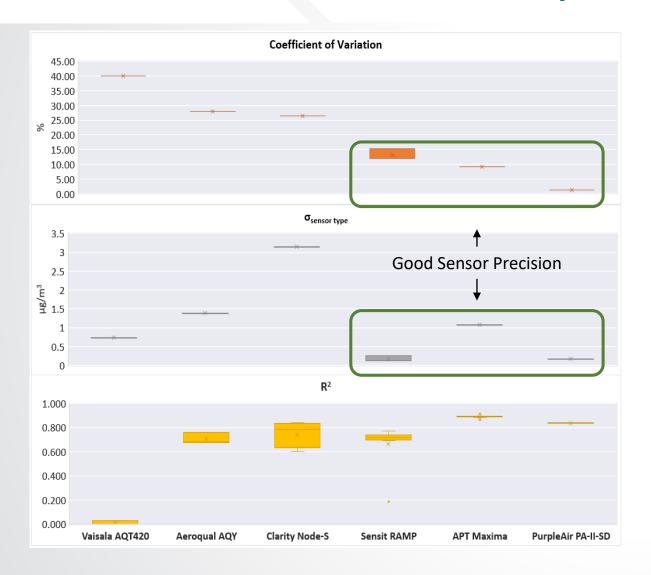
Relative humidity influence is less pronounced

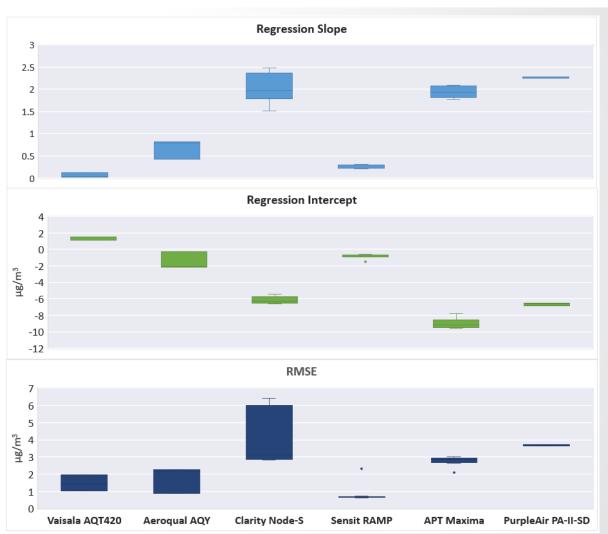




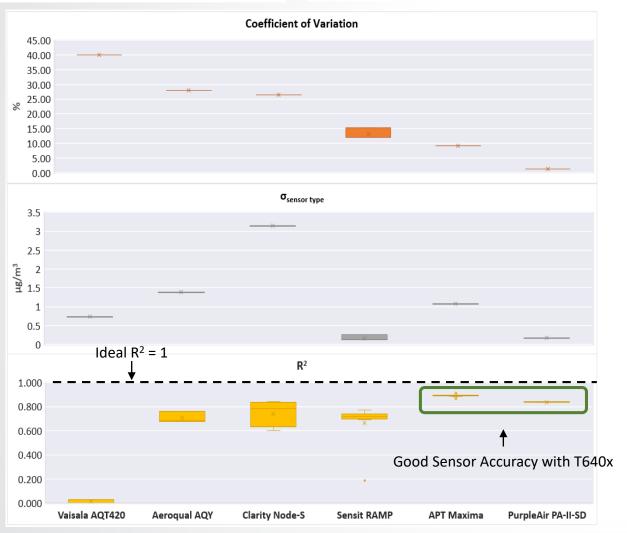


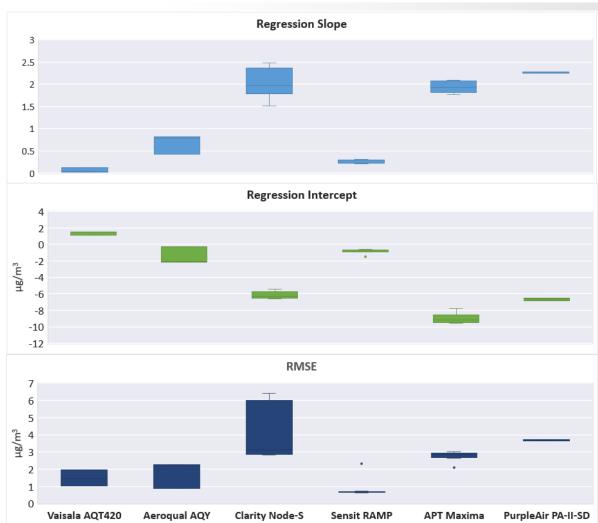




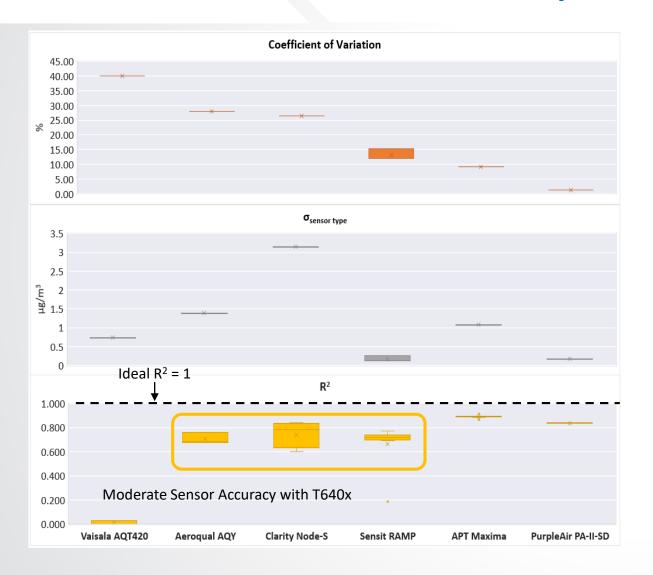


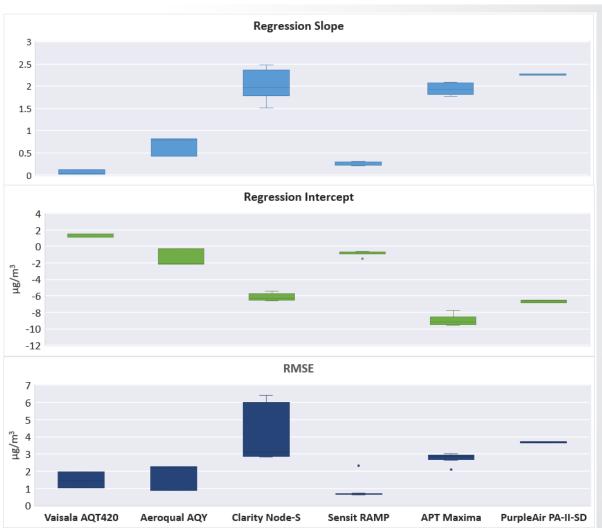




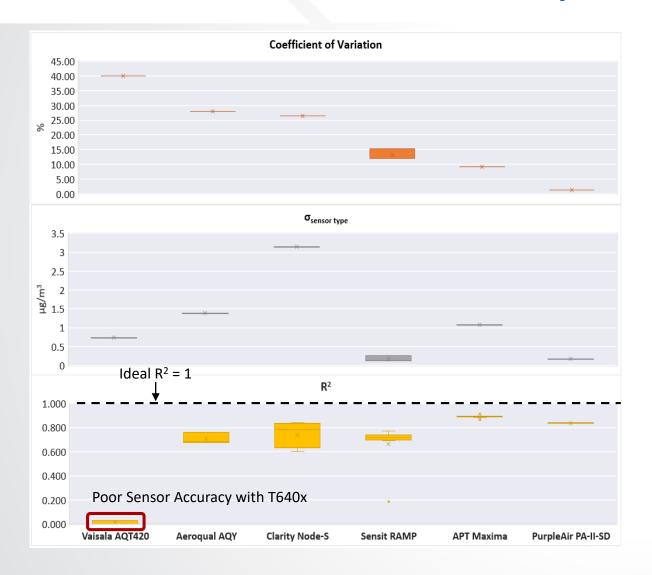


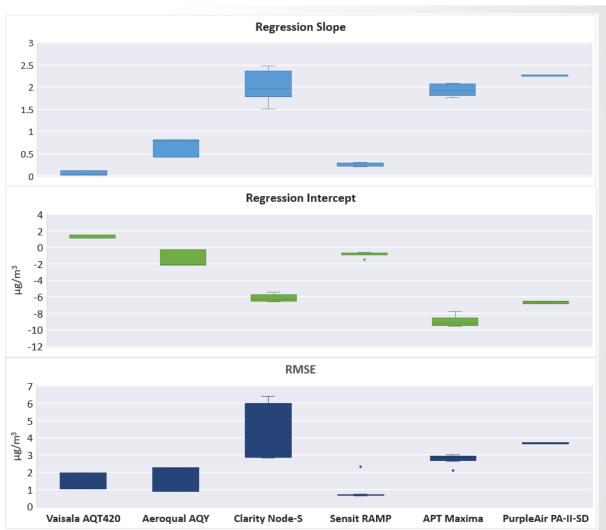




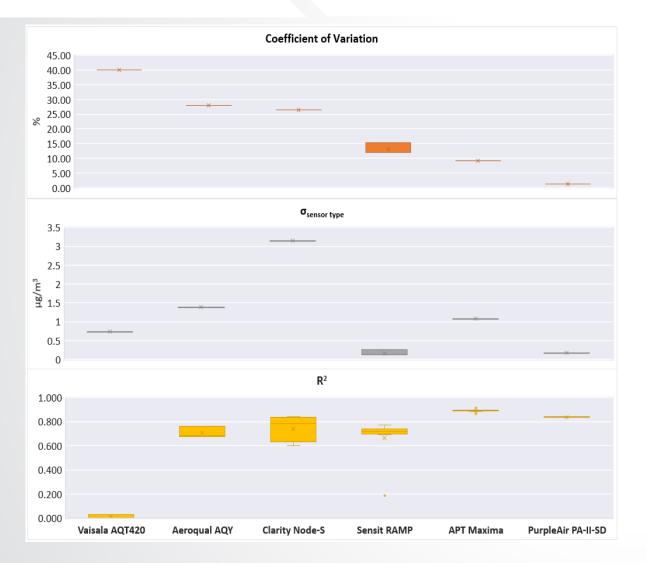


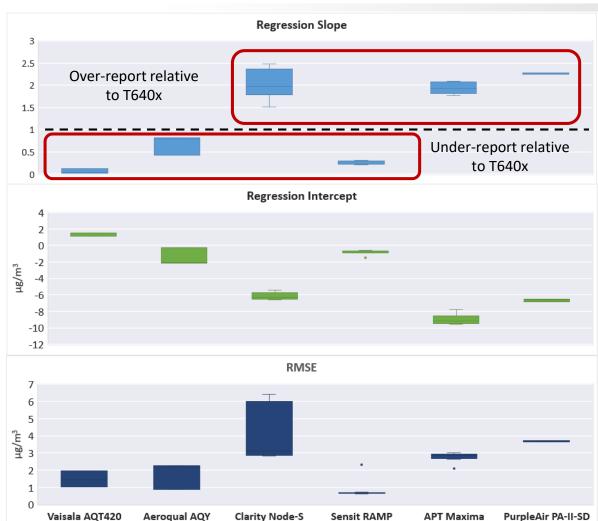




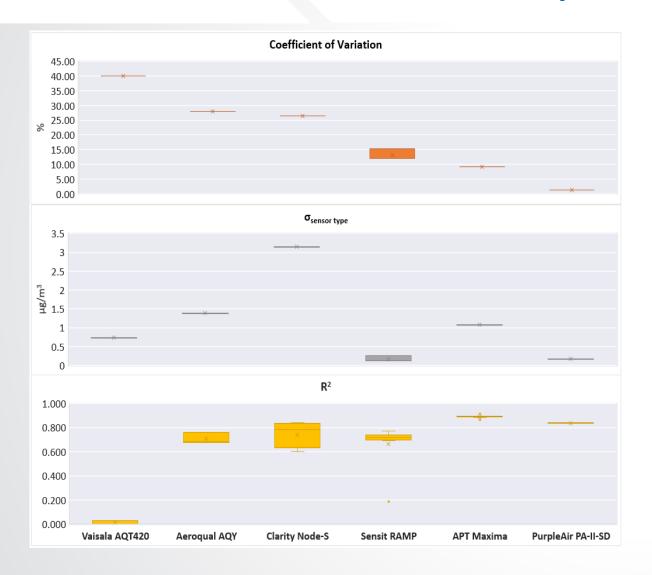


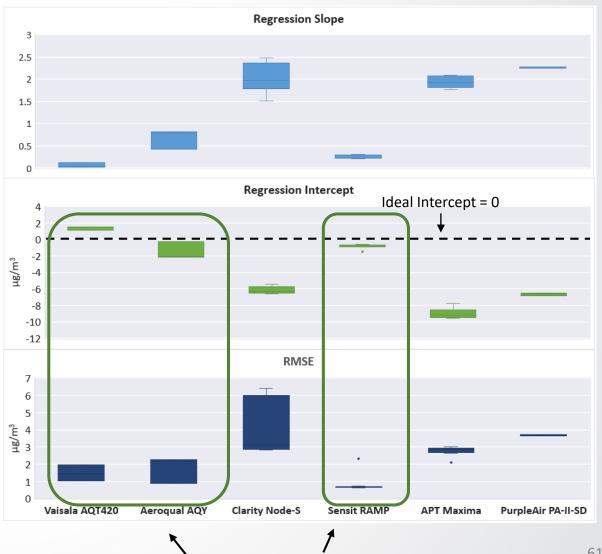






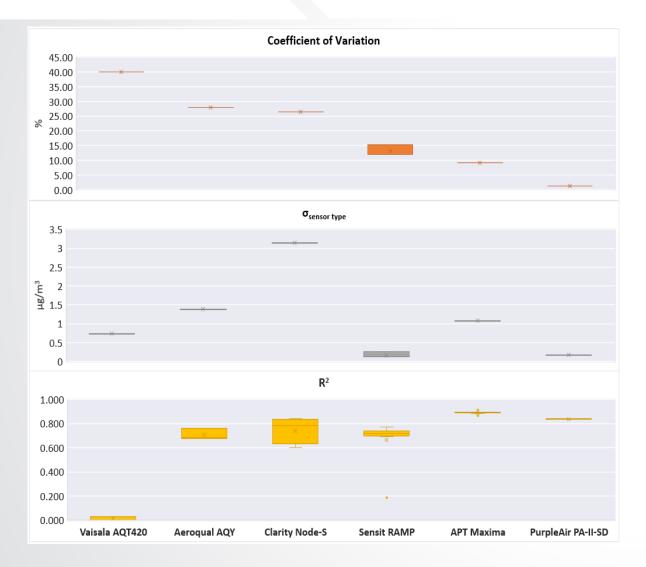


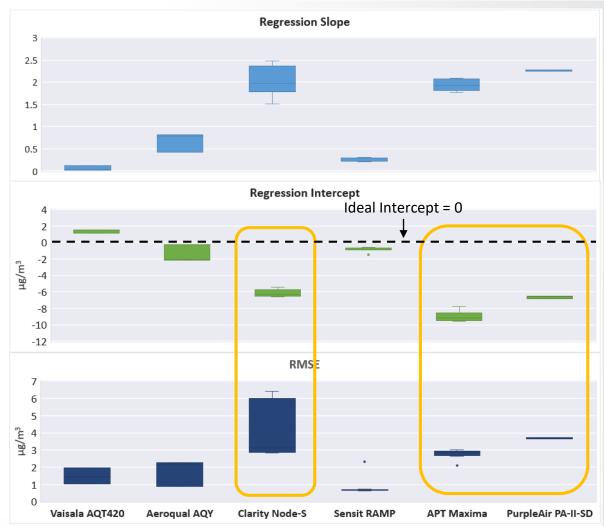




Good agreement relative to T640x









Future Work

- Evaluate additional pollutants measured by each sensor type (PM $_{10}$, O $_{3}$, NO $_{2}$ etc.)
- Further explore the influence of relative humidity and temperature on each sensor type
- Explore correction methods for PM_{2.5} sensor data appropriate for each sensor type
 - Is one correction equation appropriate for all sensors in a given type or does poor precision suggest a sensor-specific correction will be needed?
 - Is a linear correction method appropriate or is another regression method more suitable for the data?
- Ongoing sensor performance evaluations for additional sensors types



Application of this Work: Long-Term Performance

- Some of the sensors in this study are part of an ongoing project to determine the long-term performance of air quality sensors
 - Deployment at 7 air monitoring stations across the U.S. (North Carolina (NC), Georgia (GA), Delaware (DE), Arizona (AZ), Colorado (CO), Oklahoma (OK), Wisconsin (WI))
- The results of this study inform evaluation of sensor performance in the Long-Term Performance Project (LTPP)
 - > Project Lead: Andrea Clements

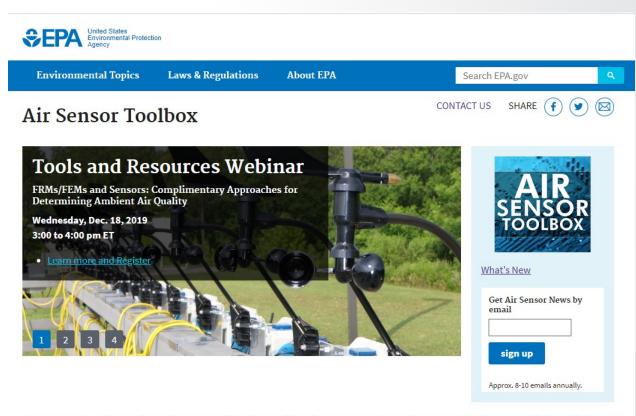




Resources

EPA's Air Sensor Toolbox

- Interpreting sensor readings and FAQs
- Documentation for choosing and operating air sensors
- Performance evaluations for air sensors
- Tools to analyze air sensor data
- Update coming this spring to improve navigation and will include new performance evaluations



Air sensor monitors that are lower in cost, portable and generally easier to operate than regulatory-grade monitors are widely used in the United States to understand air quality conditions. This website provides the latest science on the performance, operation and use of air sensor monitoring systems for technology developers, air quality managers, citizen scientists and the public. The EPA is involved in the advancement of air sensor technology, including performance evaluations of sensor devices and best practices for effectively using sensors. The information can help the public learn more about air quality in their communities.



Acknowledgements

- We extend sincere gratitude to the following individuals whose ongoing contributions make this work possible:
 - Joann Rice, EPA (OAQPS)
 - Solomon Ricks, EPA (OAQPS)
 - Ian VonWald, ORISE
 - Sensor Evaluation Group members

Disclaimer

The views expressed in this presentation are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. Any mention of trade names, products, or services does not imply an endorsement by the U.S. Government or the U.S. Environmental Protection Agency. The EPA does not endorse any commercial products, services, or enterprises.