Field Evaluation
Airly

South Coast AQMD

AQ-SPEC
Air Quality Sensor Performance Evaluation Center

Background

- From 11/14/2020 to 01/09/2021, three **Airly** multi-pollutant sensor units were deployed at the South Coast AQMD stationary ambient monitoring site in Rubidoux and were run side-by-side with Federal Equivalent Method (FEM) and Federal Reference Method (FRM) instruments measuring the same pollutants
- Airly (*3 units tested*):
  - Gas Sensors: (Electrochemical; non-FEM)
  - Particle Sensor: Light Scattering (non-FEM; PMS5003 by Plantower)
  - Each unit measures: O_3 (ppb), NO_2 (ppb), $PM_{1.0}$, $PM_{2.5}$ and PM_{10} (µg/m^3), T (°C), RH (%)
  - Units also measure pressure
  - Unit cost: $1000
  - Time resolution: 5-min
  - Units IDs: 1107, 1108, 1109

- South Coast AQMD Reference instruments:
  - O_3 instrument (FEM); cost: ~$7,000
    - Time resolution; 1-min
  - NO_X instrument (FRM); cost: ~$11,000
    - Time resolution: 1-min
  - MetOne BAM (FEM $PM_{2.5}$ & FEM PM_{10}); cost: ~$20,000
    - Time resolution: 1-hr
  - Teledyne API T640 (FEM $PM_{2.5}$); cost: $21,000
    - Time resolution: 1-min
  - Met station (T, RH, P, WS, WD); cost: ~$5,000
    - Time resolution: 1-min

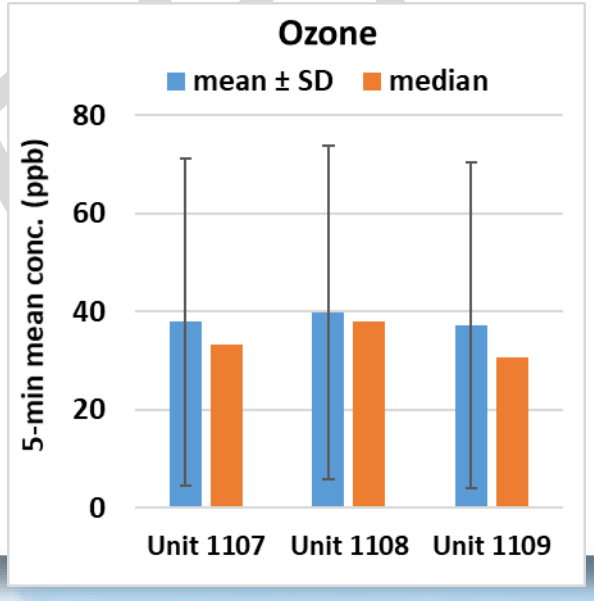
Ozone (O_3) in Airly

Data validation & recovery

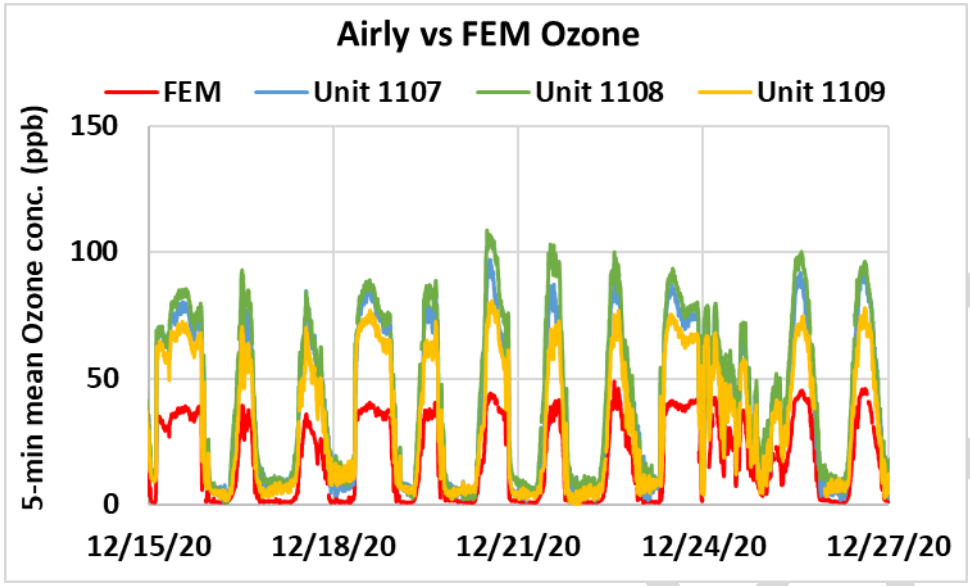
- Basic QA/QC procedures were used to validate the collected data (i.e., obvious outliers, negative values, and invalid data-points were eliminated from the data-set)
- Data recovery for ozone from Unit 1107, Unit 1108 and Unit 1109 was ~ 98%, 82% and 98%, respectively.

Airly; Intra-model variability

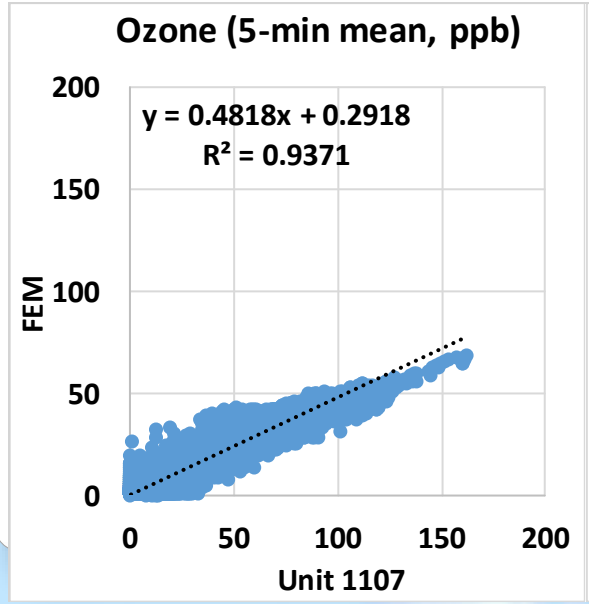
- Absolute intra-model variability was ~ 1.3 ppb for the ozone measurements (calculated as the standard deviation of the three sensor means)
- Relative intra-model variability was ~ 3.4% for the ozone measurements (calculated as the absolute intra-model variability relative to the mean of the three sensor means)

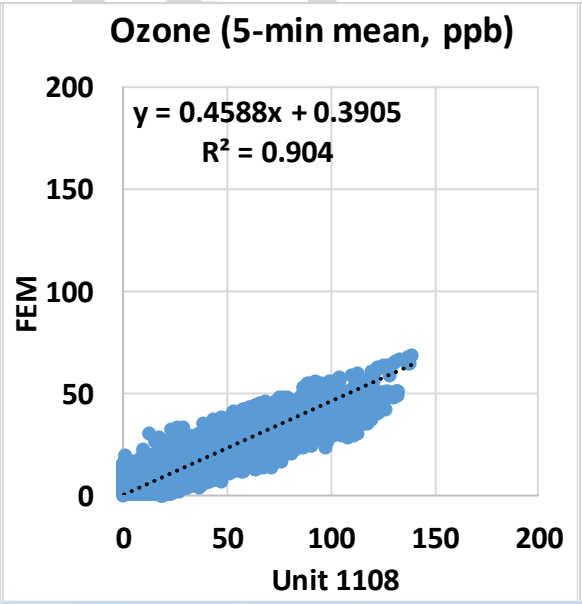


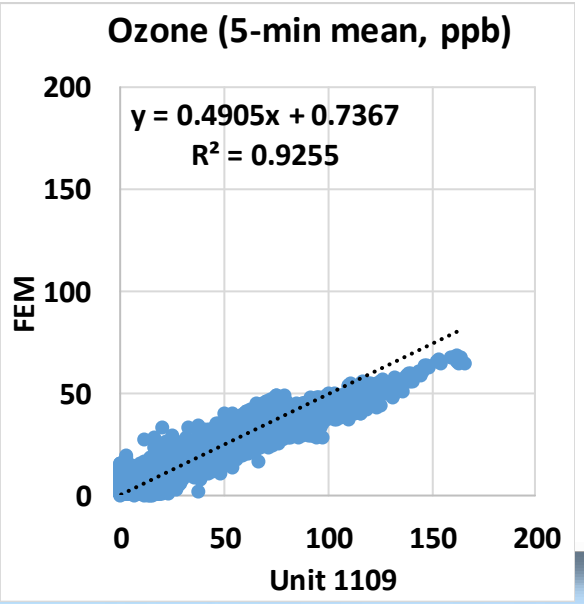
Airly vs FEM (Ozone; 5-min mean)



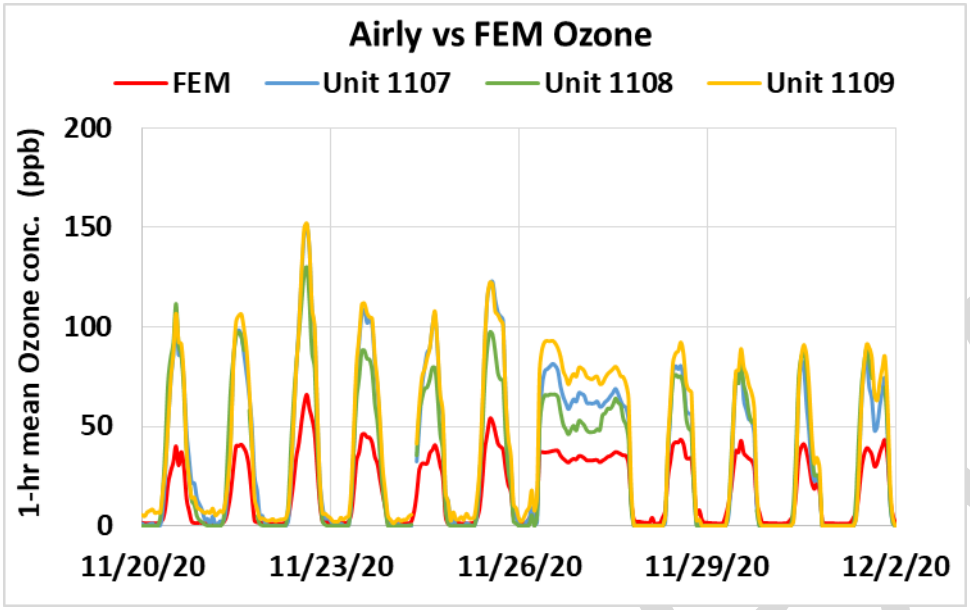
- Airly sensors showed very strong correlations with the corresponding FEM ozone data ($0.90 < R^2 < 0.94$)
- Overall, the Airly sensors overestimated the ozone concentration as measured by the FEM ozone instrument
- The Airly sensors seemed to track the diurnal ozone variations as recorded by the FEM instrument



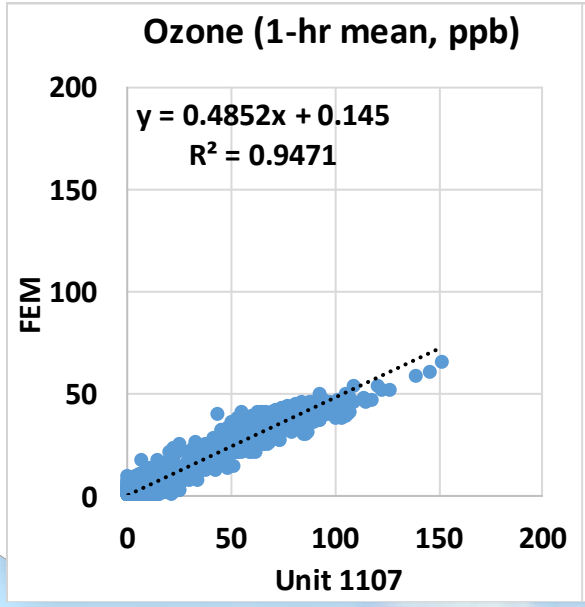


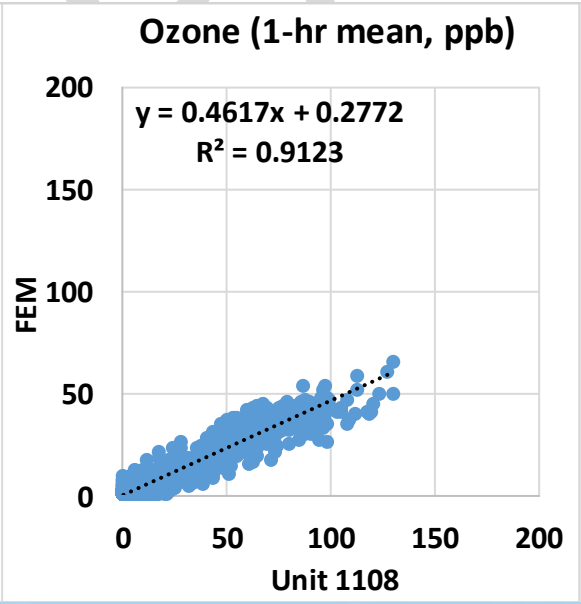


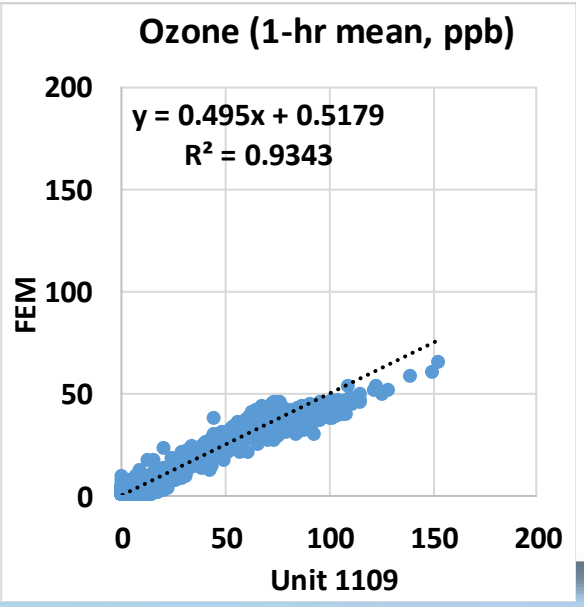
Airly vs FEM (Ozone; 1-hr mean)



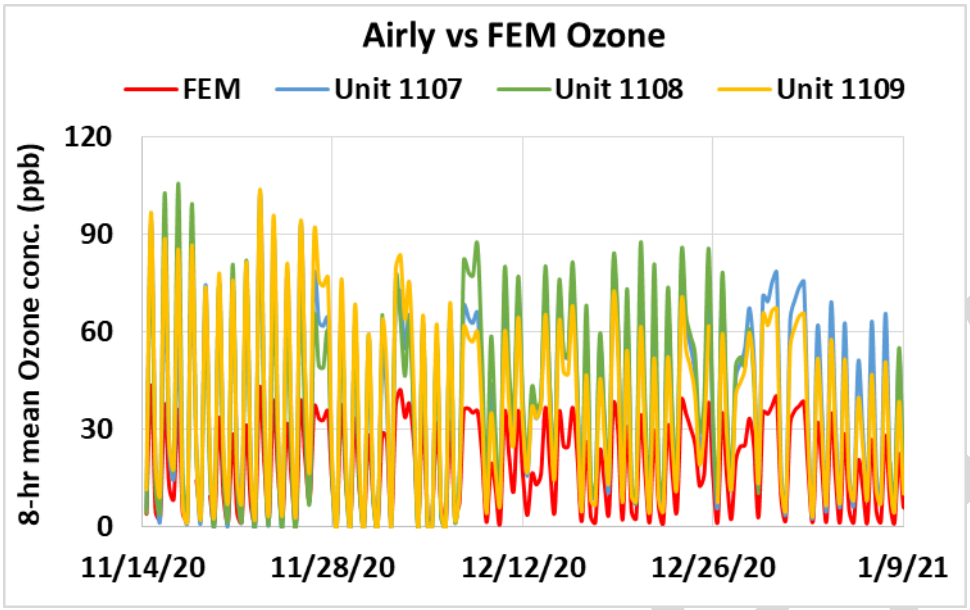
- Airly sensors showed very strong correlations with the corresponding FEM ozone data ($0.91 < R^2 < 0.95$)
- Overall, the Airly sensors overestimated the ozone concentration as measured by the FEM instrument
- The Airly sensors seemed to track the diurnal ozone variations as recorded by the FEM instrument



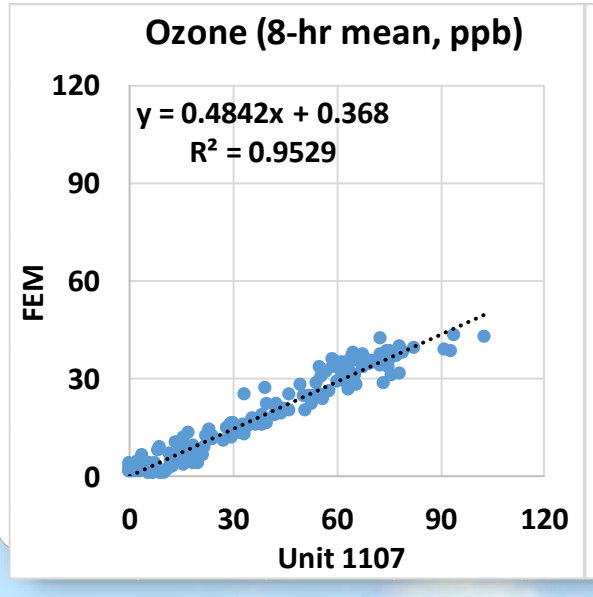


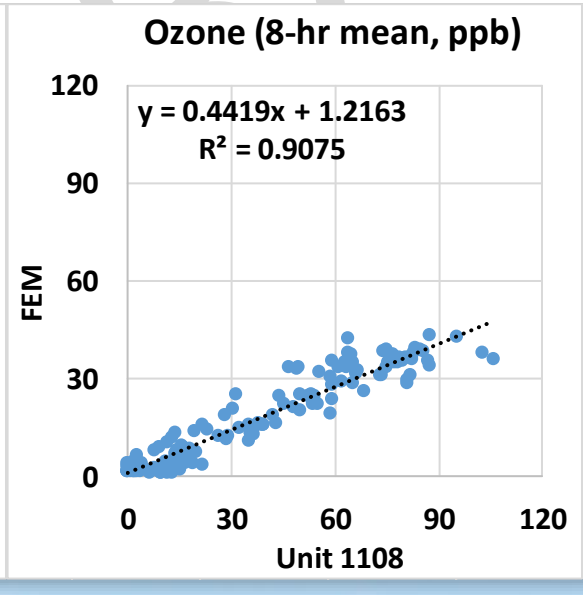


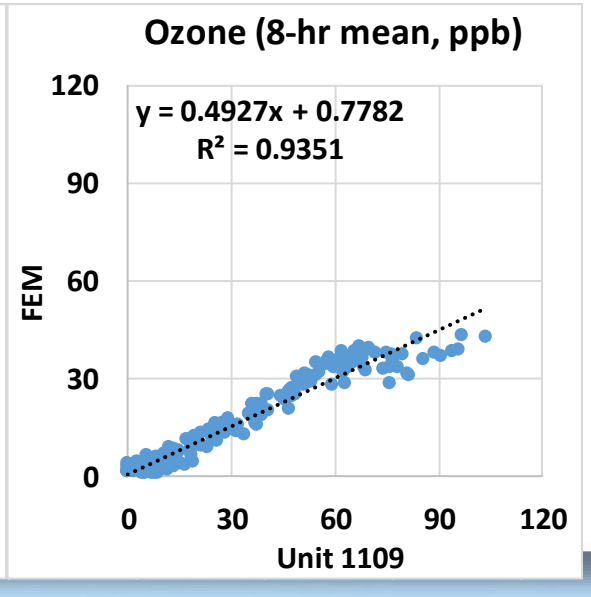
Airly vs FEM (Ozone; 8-hr mean)



- Airly sensors showed very strong correlations with the corresponding FEM ozone data ($0.90 < R^2 < 0.96$)
- Overall, the Airly sensors overestimated the ozone concentration as measured by the FEM instrument
- The Airly sensors seemed to track the diurnal ozone variations as recorded by the FEM instrument







Summary: Ozone

	Average of 3 Sensors, Ozone		Airly vs FEM, Ozone						FEM Ozone (ppb)		
	Average (ppb)	SD (ppb)	R^2	Slope	Intercept	MBE[1] (ppb)	MAE[2] (ppb)	RMSE[3] (ppb)	FEM Average	FEM SD	Range during the field evaluation
5-min	38.3	33.0	0.90 to 0.94	0.46 to 0.49	0.29 to 0.74	18.8 to 22.0	19.3 to 22.9	40.5 to 43.0	19.5	16.3	0.4 to 68.9
1-hr	38.3	32.6	0.91 to 0.95	0.46 to 0.50	0.15 to 0.52	18.0 to 21.0	18.4 to 21.8	24.4 to 28.1	18.6	16.1	0.9 to 65.9
8-hr	38.1	28.6	0.91 to 0.95	0.44 to 0.49	0.37 to 1.22	17.6 to 20.5	17.9 to 20.8	22.7 to 24.2	18.6	13.4	1.3 to 43.6

[1] Mean Bias Error (MBE): the difference between the sensors and the reference instruments. MBE indicates the tendency of the sensors to underestimate (negative MBE values) or overestimate (positive MBE values).
[2] Mean Absolute Error (MAE): the absolute difference between the sensors and the reference instruments. The larger MAE values, the higher measurement errors as compared to the reference instruments.
[3] Root Mean Square Error (RMSE): another metric to calculate measurement errors.

Nitrogen Dioxide (NO_2) in Airly

Data validation & recovery

- Basic QA/QC procedures were used to validate the collected data (i.e., obvious outliers, negative values, and invalid data-points were eliminated from the data-set)
- Data recovery for NO_2 from Unit 1107, Unit 1108 and Unit 1109 was ~ 98%, 82% and 98% respectively.

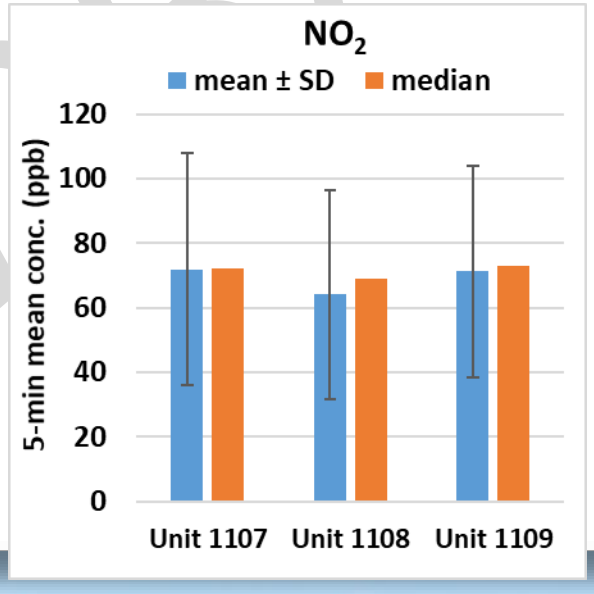
Airly; Intra-model variability

- Absolute intra-model variability was ~ 4.3 ppb for the NO_2 measurements

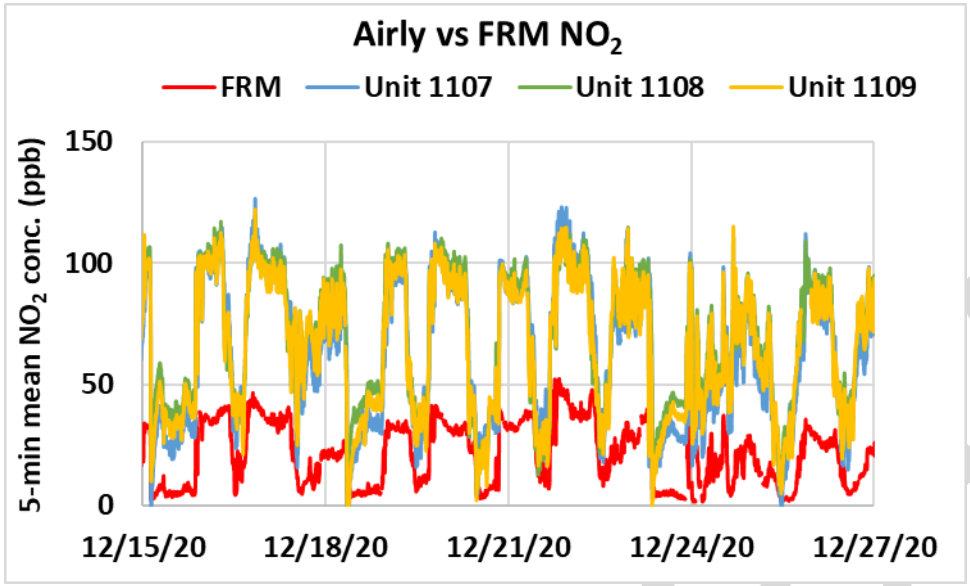
(calculated as the standard deviation of the three sensor means)

- Relative intra-model variability was ~ 6.3% for the NO_2 measurements

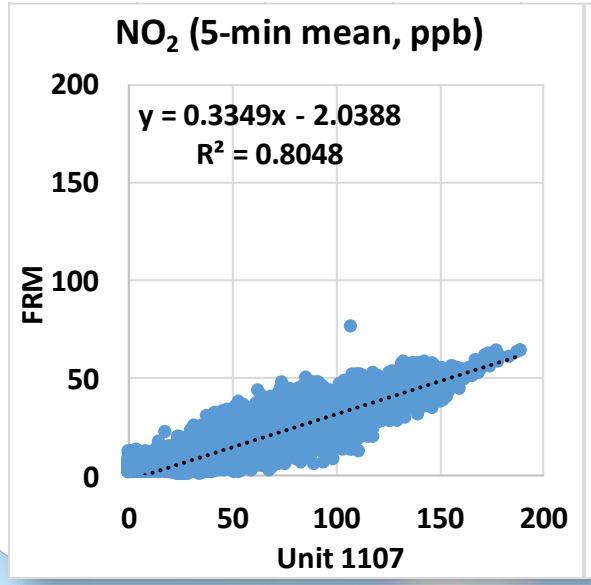
(calculated as the absolute intra-model variability relative to the mean of the three sensor means)

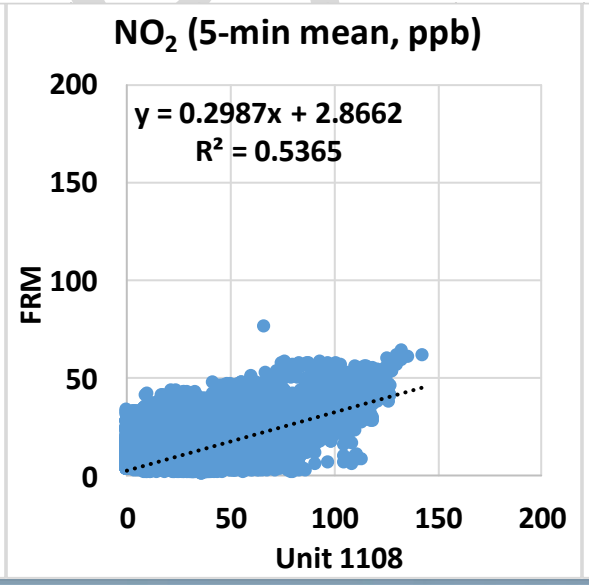


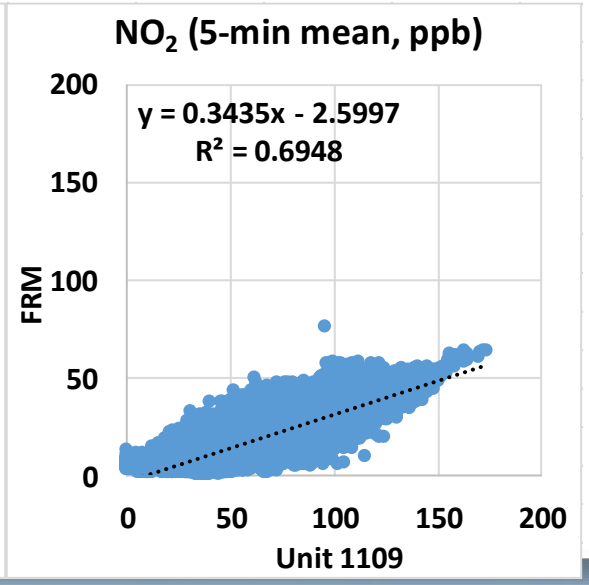
Airly vs FRM (NO_2; 5-min mean)



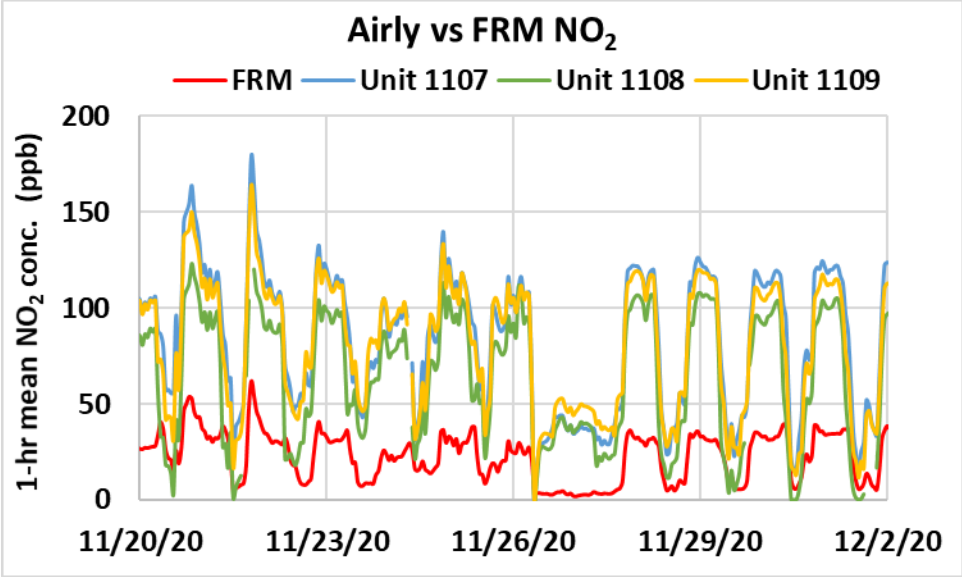
- Airly sensors showed moderate to strong correlations with the corresponding FRM NO_2 data ($0.53 < R^2 < 0.81$)
- Overall, the Airly sensors overestimated the NO_2 concentration as measured by the FRM instrument
- The Airly sensors seemed to track the diurnal NO_2 variations as recorded by the FRM instrument



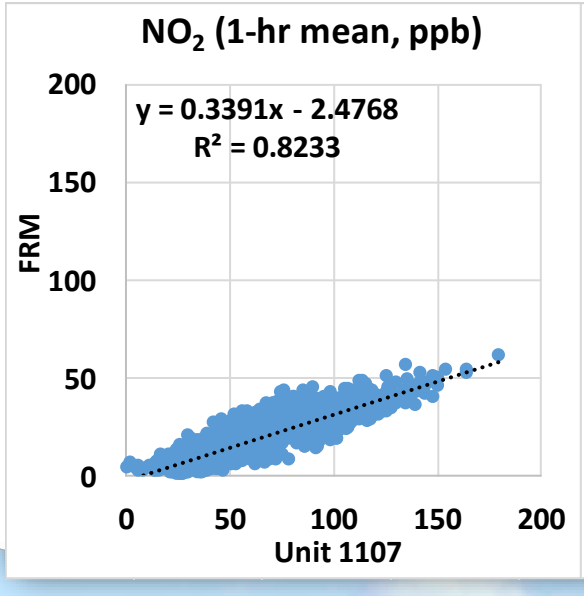


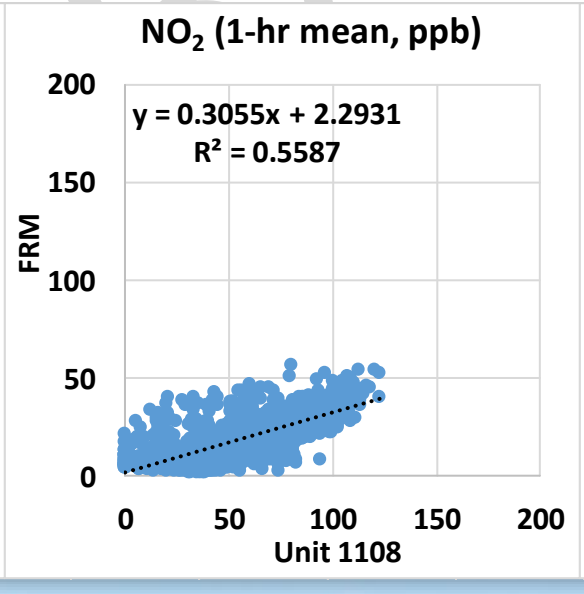


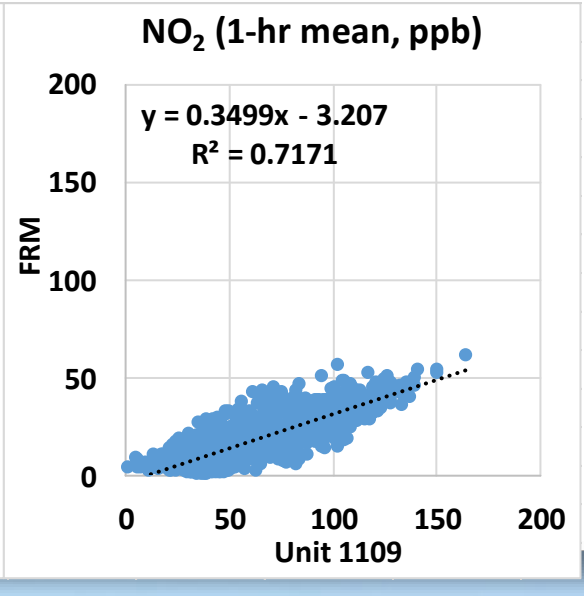
Airly vs FRM (NO_2; 1-hr mean)



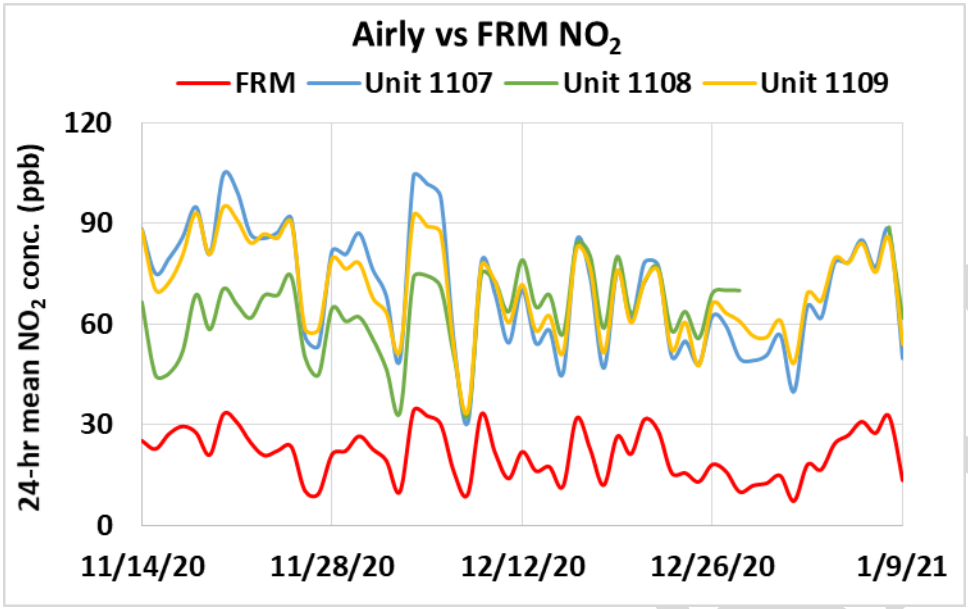
- Airly sensors showed moderate to strong correlations with the corresponding FRM data ($0.55 < R^2 < 0.83$)
- Overall, the Airly sensors overestimated the NO_2 concentration as measured by the FRM instrument
- The Airly sensors seemed to track the diurnal NO_2 variations as recorded by the FRM instrument



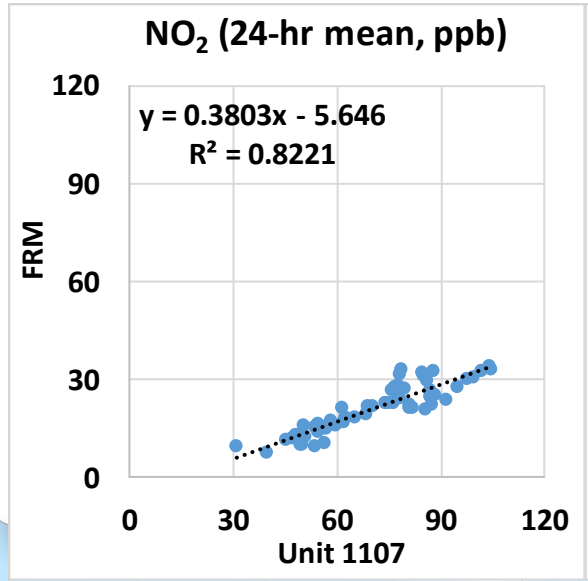


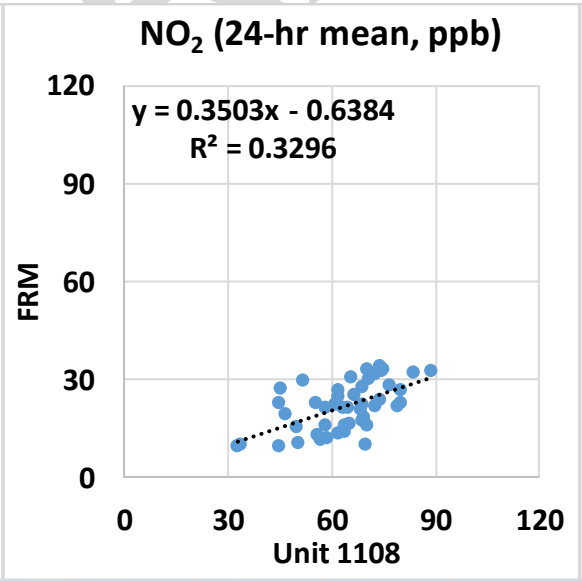


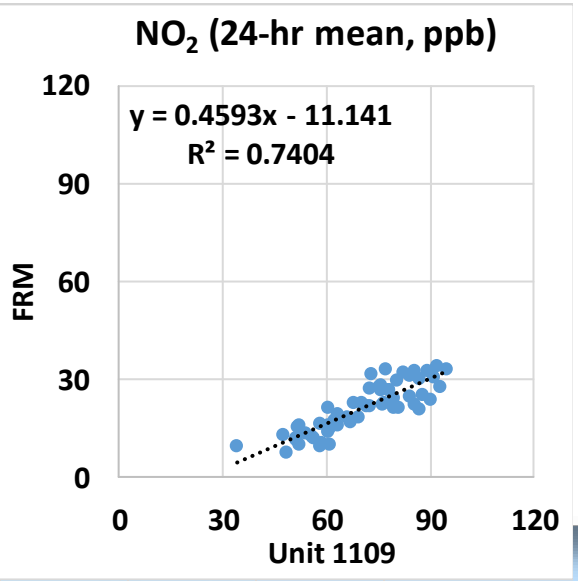
Airly vs FRM (NO_2; 24-hr mean)



- Airly sensors showed weak to strong correlations with the corresponding FRM data ($0.32 < R^2 < 0.83$)
- Overall, the Airly sensors overestimated the NO_2 concentration as measured by the FRM instrument
- The Airly sensors seemed to track the diurnal NO_2 variations as recorded by the FRM instrument







Summary: NO_2

	Average of 3 Sensors, NO_2		Airly vs FRM, NO_2						FRM NO_2 (ppb)		
	Average (ppb)	SD (ppb)	R^2	Slope	Intercept	MBE[1] (ppb)	MAE[2] (ppb)	RMSE[3] (ppb)	FRM Average	FRM SD	Range during the field evaluation
5-min	68.6	33.2	0.54 to 0.80	0.30 to 0.34	-2.6 to 2.9	40.9 to 48.1	42.4 to 48.1	70.4 to 86.3	21.2	13.1	1.0 to 76.3
1-hr	68.6	32.3	0.56 to 0.82	0.31 to 0.35	-3.2 to 2.3	42.5 to 49.3	43.6 to 49.3	48.5 to 54.5	21.6	12.8	1.3 to 62.1
24-hr	68.4	14.6	0.33 to 0.82	0.35 to 0.46	-11.1 to -0.64	42.1 to 49.5	42.1 to 49.5	43.3 to 50.8	21.3	7.5	7.4 to 34.3

[1] Mean Bias Error (MBE): the difference between the sensors and the reference instruments. MBE indicates the tendency of the sensors to underestimate (negative MBE values) or overestimate (positive MBE values).
[2] Mean Absolute Error (MAE): the absolute difference between the sensors and the reference instruments. The larger MAE values, the higher measurement errors as compared to the reference instruments.
[3] Root Mean Square Error (RMSE): another metric to calculate measurement errors.

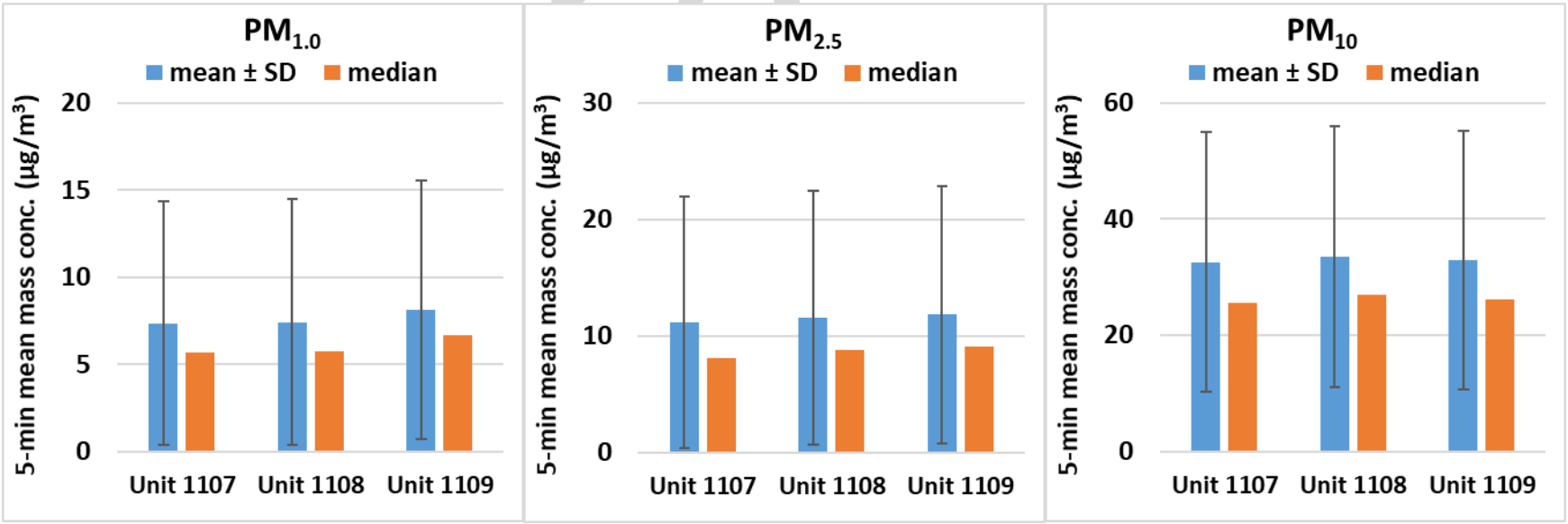
PM in Airly

Data validation & recovery

- Basic QA/QC procedures were used to validate the collected data (i.e., obvious outliers, negative values, and invalid data-points were eliminated from the data-set)
- Data recovery from Unit 1107, Unit 1108 and Unit 1109 was ~ 100% for all PM fractions.

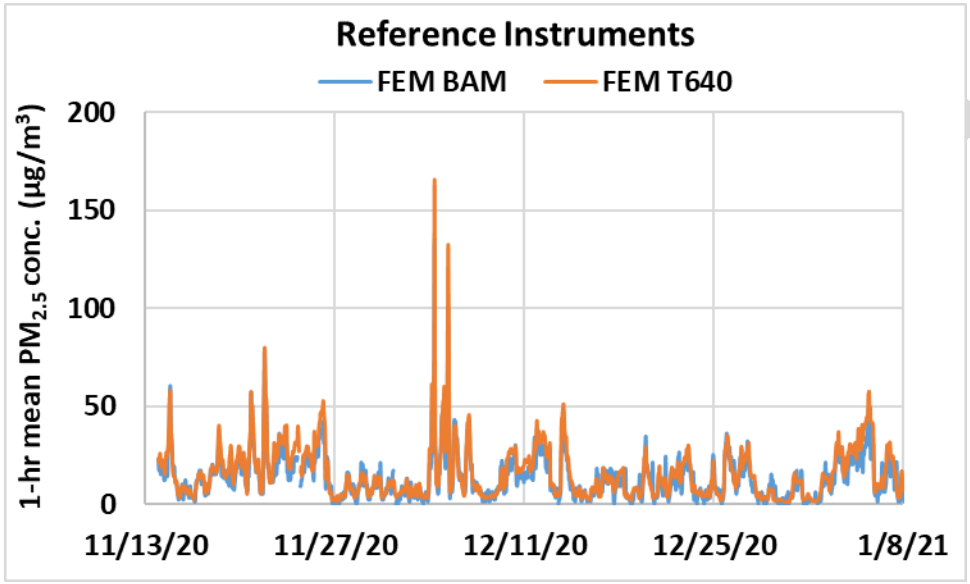
Airly; Intra-model variability

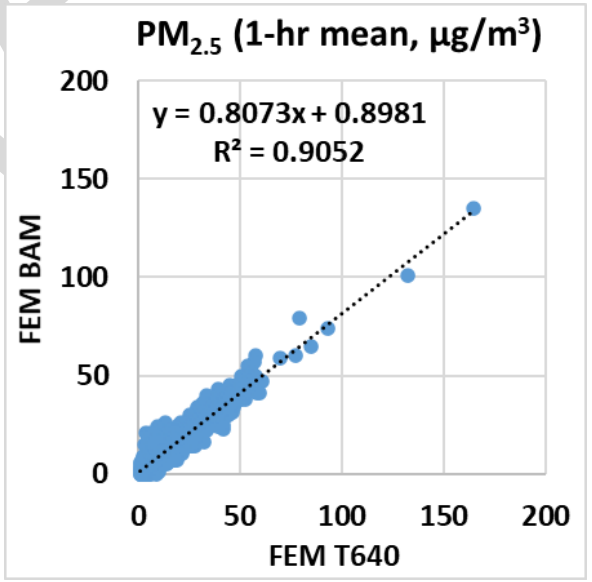
- Absolute intra-model variability was ~ 0.36, 0.29 and 0.41 µg/m^3 for the $PM_{1.0,}$ $PM_{2.5}$ and PM_{10}, respectively. (calculated as the standard deviation of the three sensor means)
- Relative intra-model variability was ~ 4.7%, 2.5% and 1.3% for the $PM_{1.0,}$ $PM_{2.5}$ and $PM_{10,}$ respectively. (calculated as the absolute intra-model variability relative to the mean of the three sensor means)



Reference Instruments: $PM_{2.5}$ FEM BAM & FEM T640

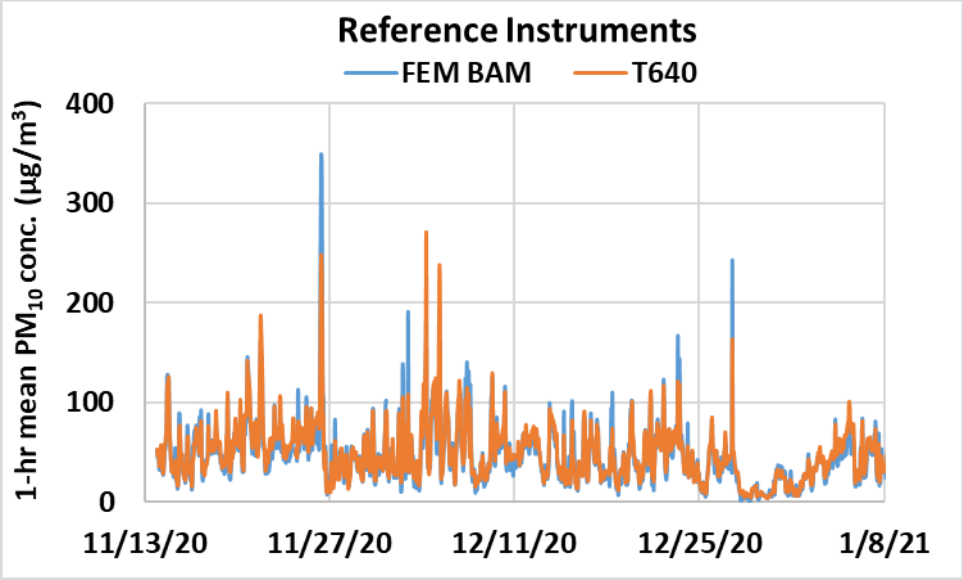
- Basic QA/QC procedures were used to validate the collected data (i.e. obvious outliers, negative values and invalid data-points were eliminated from the data-set)
- Data recovery for $PM_{2.5}$ from FEM BAM and FEM T640 is ~97% and 100%, respectively
- Very strong correlations between FEM BAM and FEM T640 for $PM_{2.5}$ measurements (R^2 ~ 0.91)

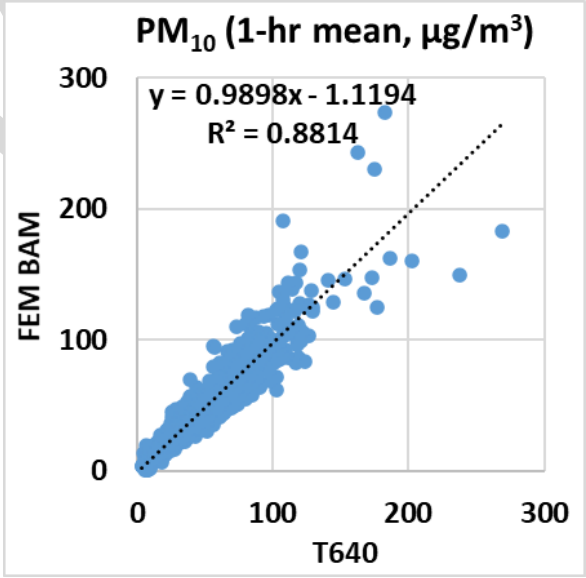




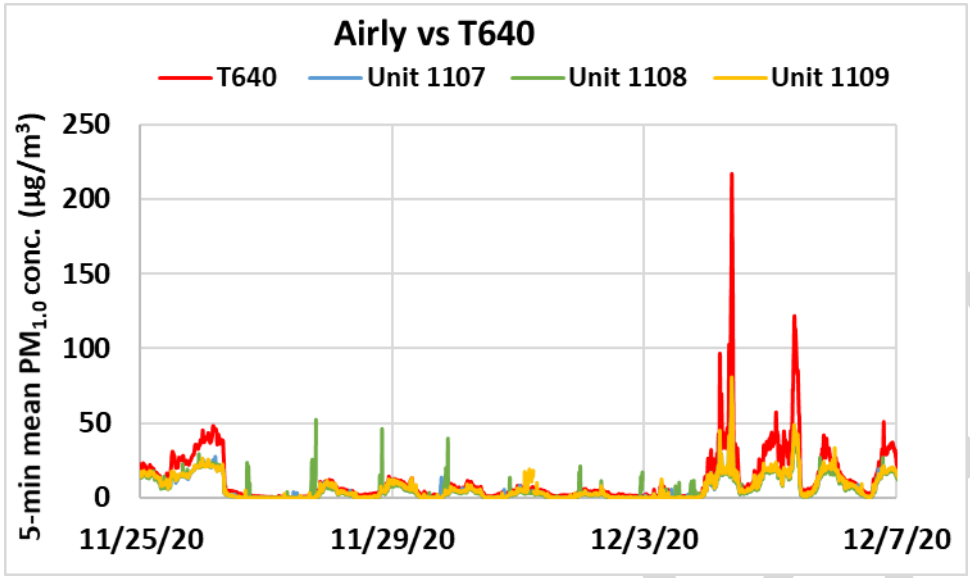
Reference Instruments: PM_{10} FEM BAM & T640

- Basic QA/QC procedures were used to validate the collected data (i.e. obvious outliers, negative values and invalid data-points were eliminated from the data-set)
- Data recovery for PM_{10} from FEM BAM and T640 is ~99% and 100%, respectively
- Strong correlations between FEM BAM and T640 for PM_{10} measurements (R^2 ~ 0.88)

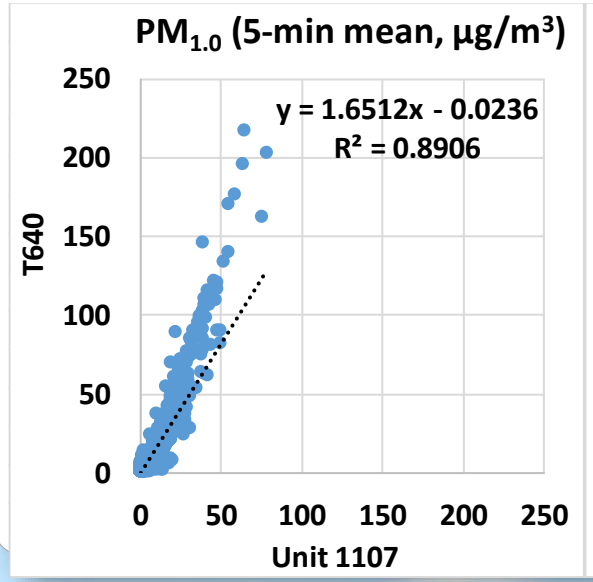


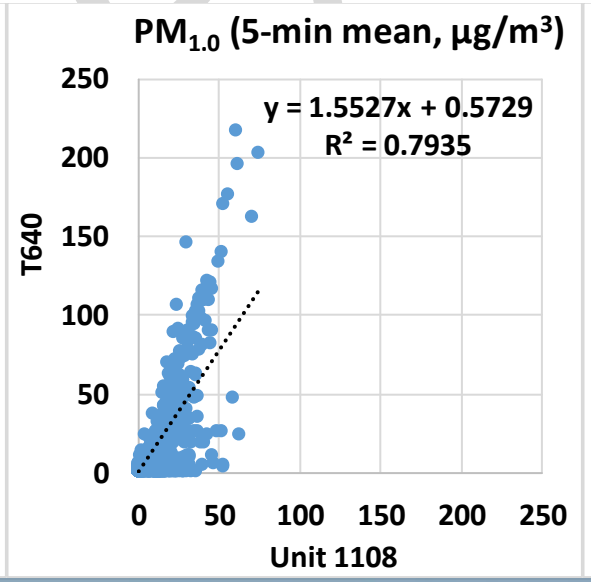


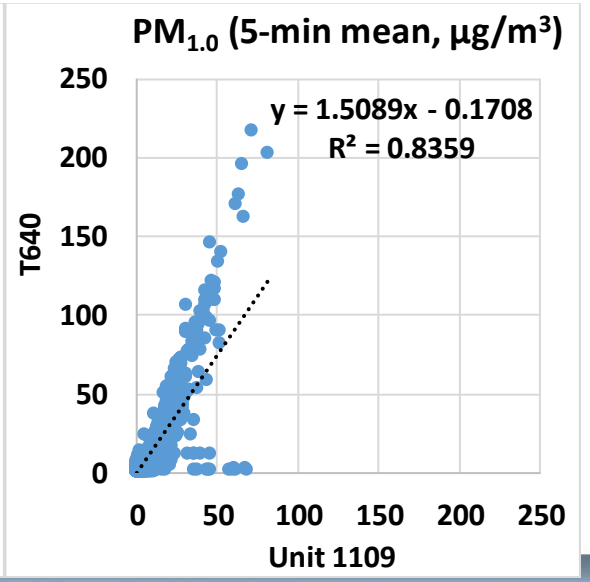
Airly vs T640 ($PM_{1.0}$; 5-min mean)



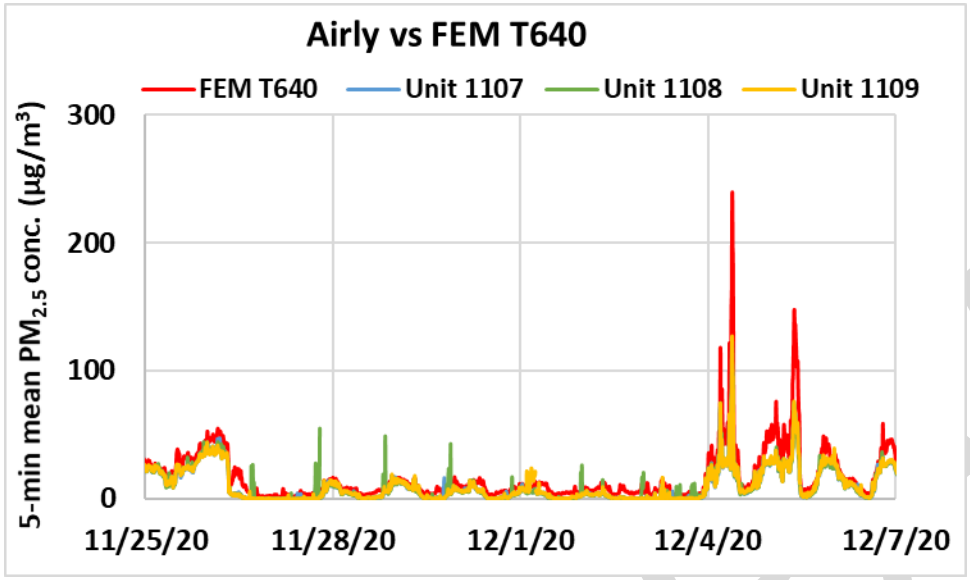
- Airly sensors showed strong correlations with the corresponding T640 data ($0.79 < R^2 < 0.90$)
- Overall, the Airly sensors underestimated the $PM_{1.0}$ mass concentration as measured by the T640
- The Airly sensors seemed to track the diurnal $PM_{1.0}$ variations as recorded by the T640



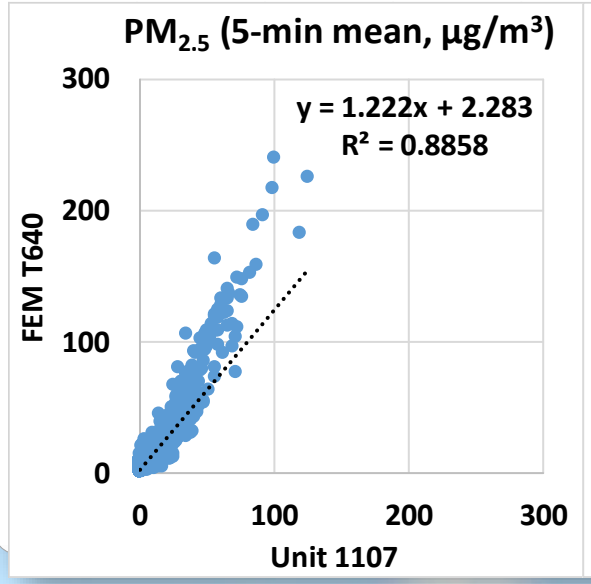


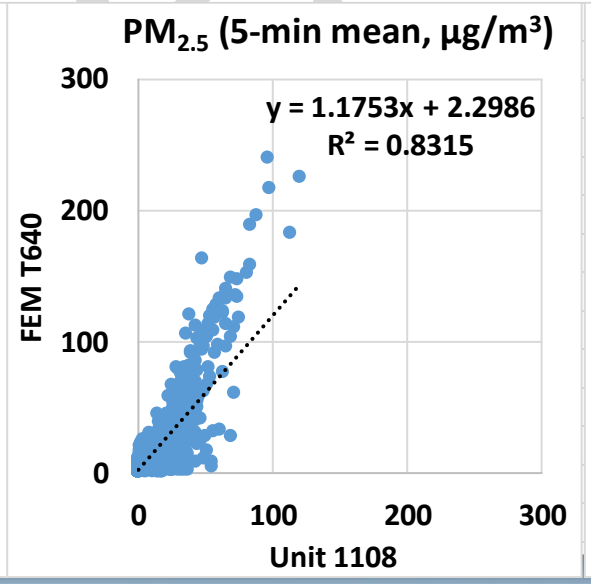


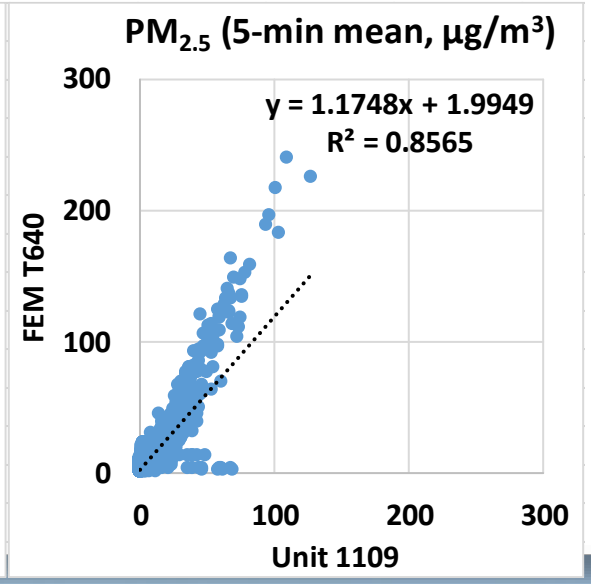
Airly vs FEM T640 ($PM_{2.5}$; 5-min mean)



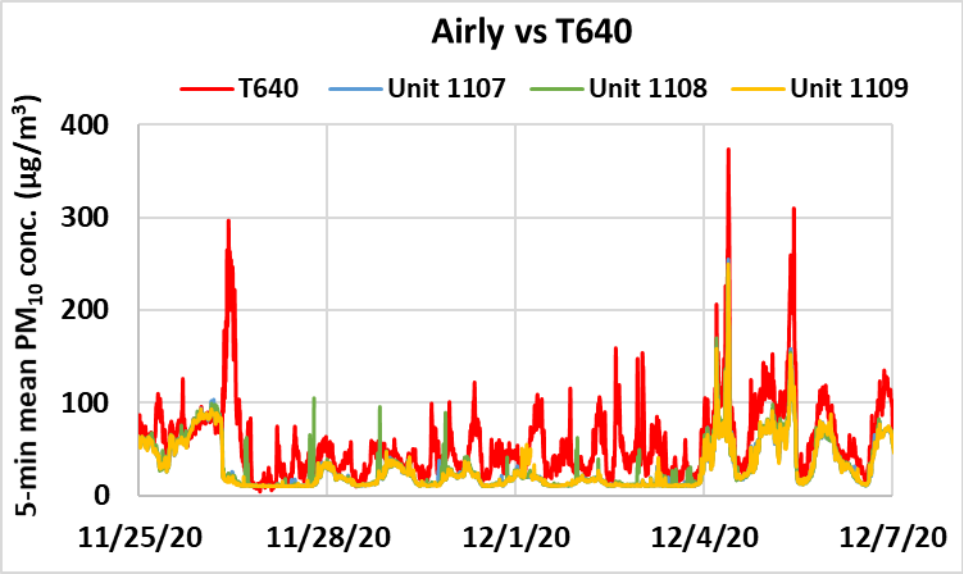
- Airly sensors showed strong correlations with the corresponding FEM T640 data ($0.83 < R^2 < 0.89$)
- Overall, the Airly sensors underestimated the $PM_{2.5}$ mass concentration as measured by the FEM T640
- The Airly sensors seemed to track the diurnal $PM_{2.5}$ variations as recorded by the FEM T640



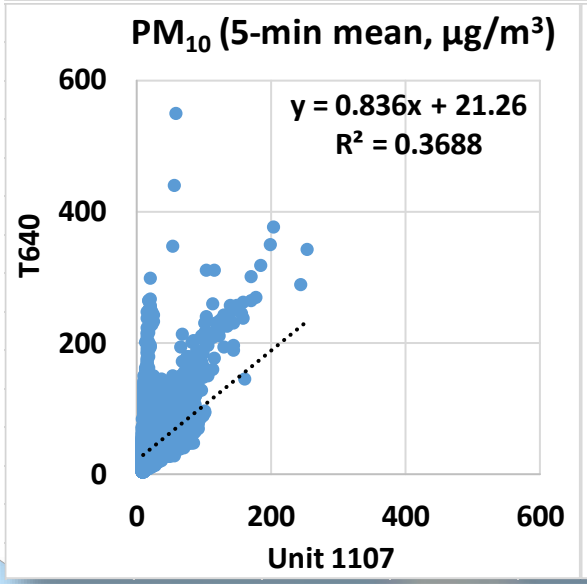


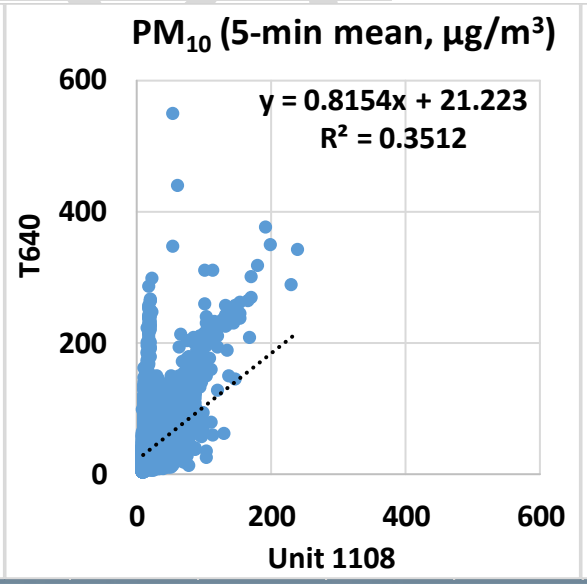


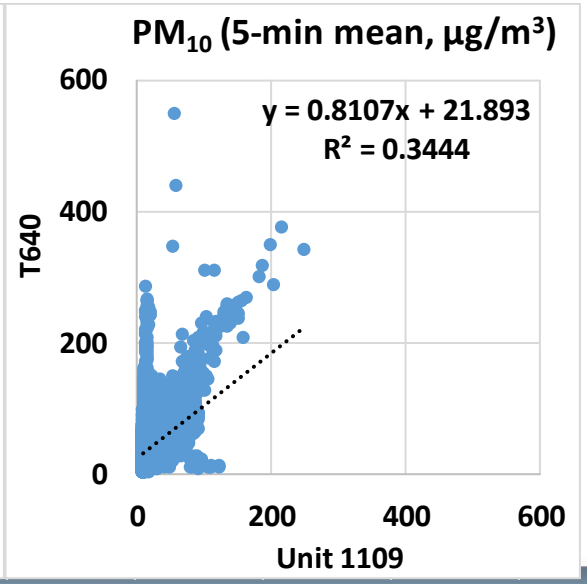
Airly vs T640 (PM_{10}; 5-min mean)



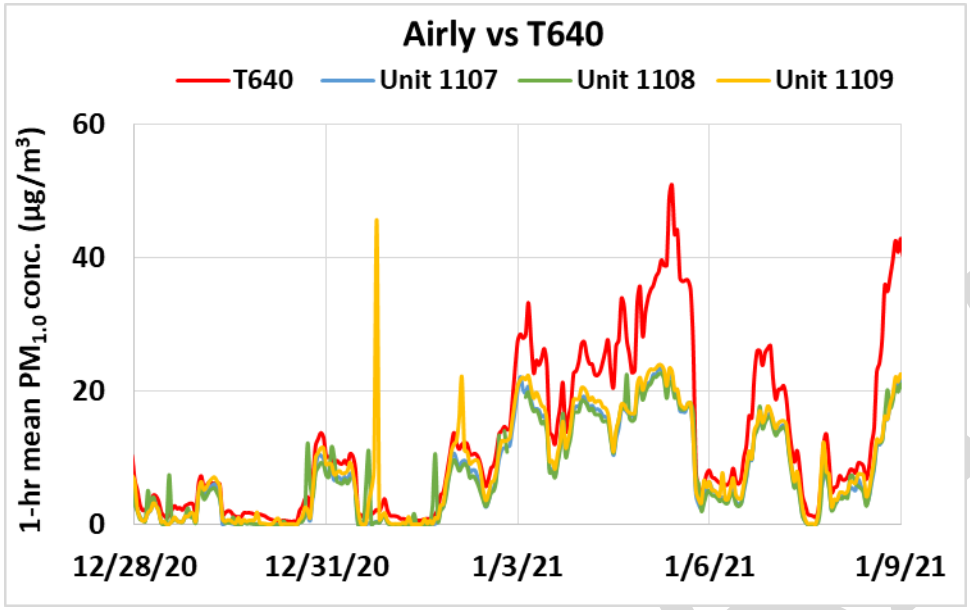
- Airly sensors showed weak correlations with the corresponding T640 data ($0.34 < R^2 < 0.37$)
- Overall, the Airly sensors underestimated the PM_{10} mass concentration as measured by the T640
- The Airly sensors did not seem to track the diurnal PM_{10} variations as recorded by the T640



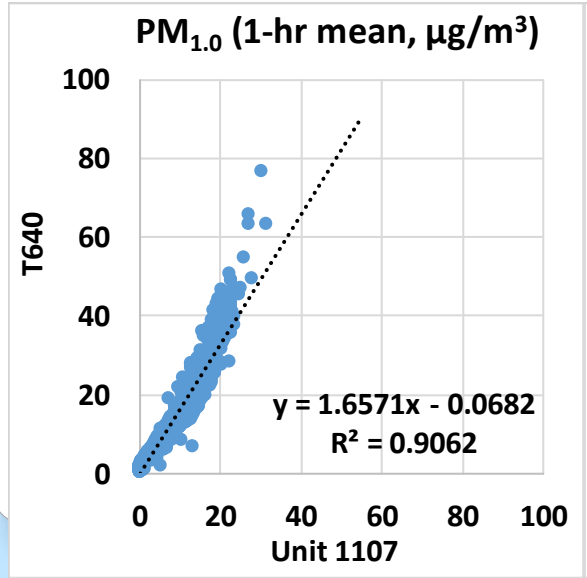


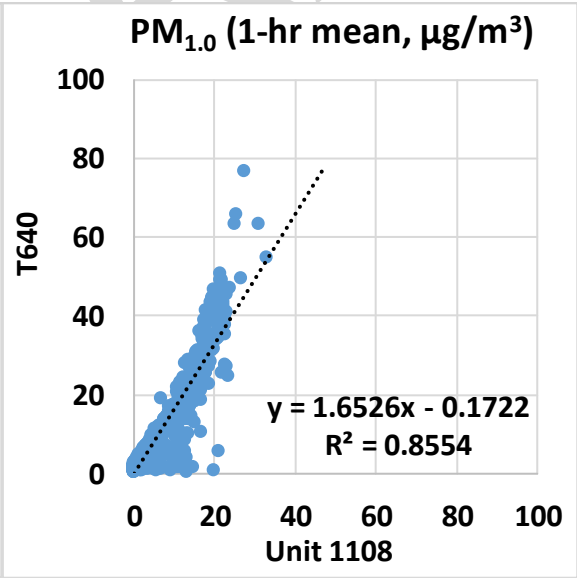


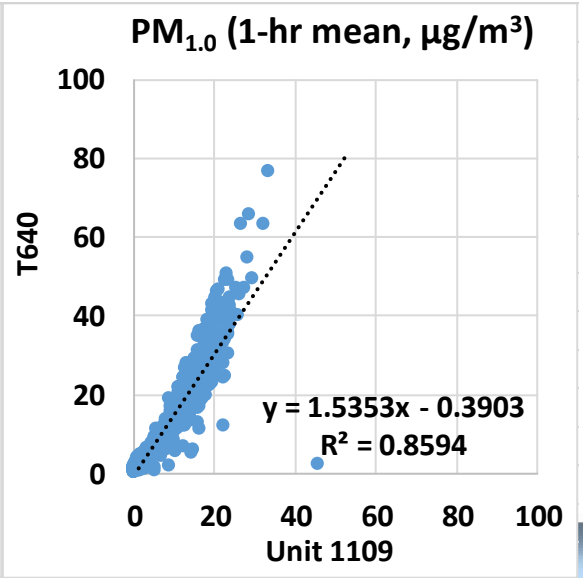
Airly vs T640 ($PM_{1.0}$; 1-hr mean)



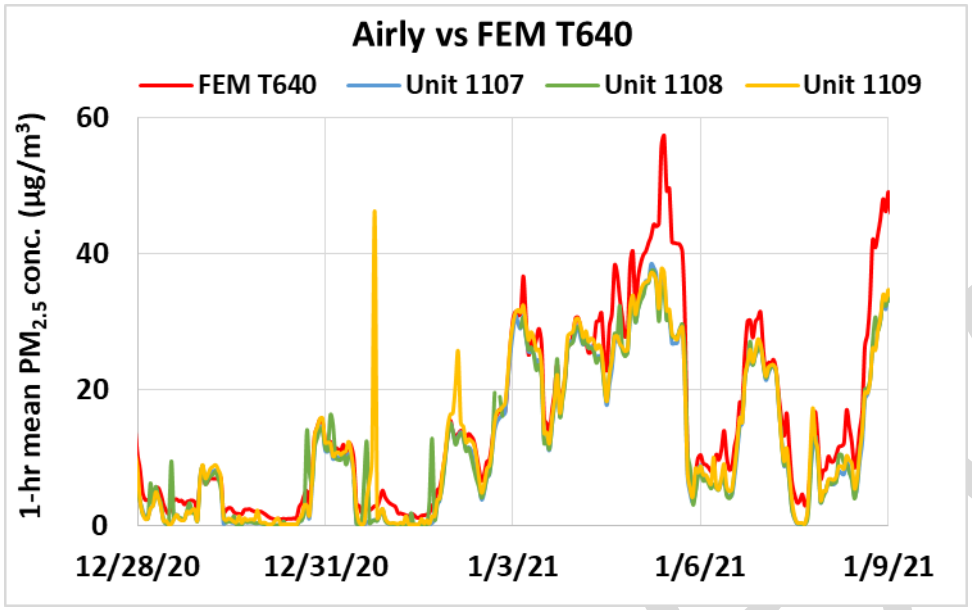
- Airly sensors showed strong to very strong correlations with the corresponding T640 data ($0.85 < R^2 < 0.91$)
- Overall, the Airly sensors underestimated the $PM_{1.0}$ mass concentration as measured by the T640
- The Airly sensors seemed to track the diurnal $PM_{1.0}$ variations as recorded by the T640



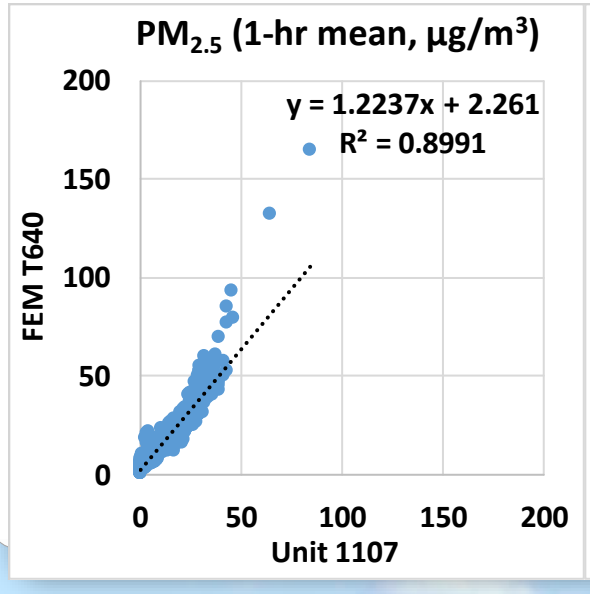


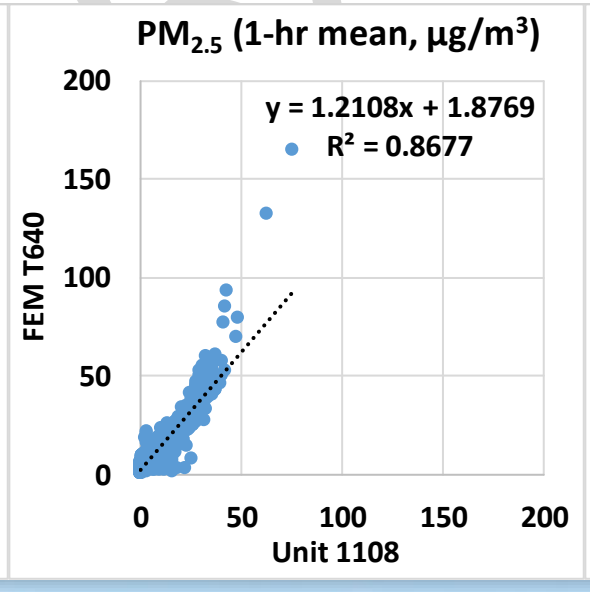


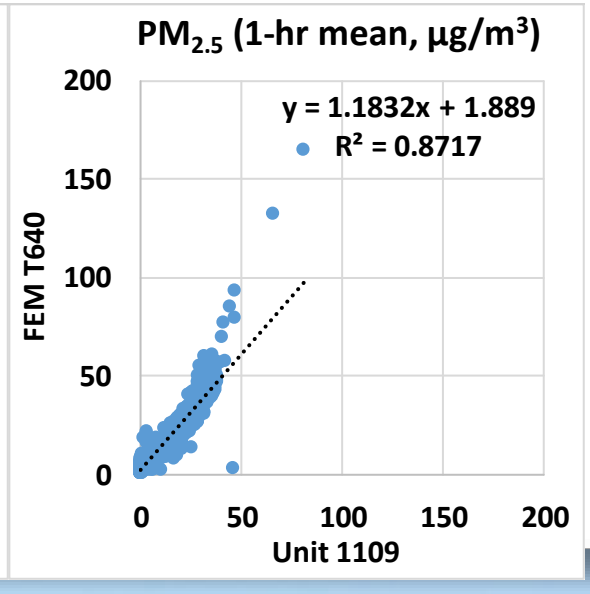
Airly vs FEM T640 ($PM_{2.5}$; 1-hr mean)



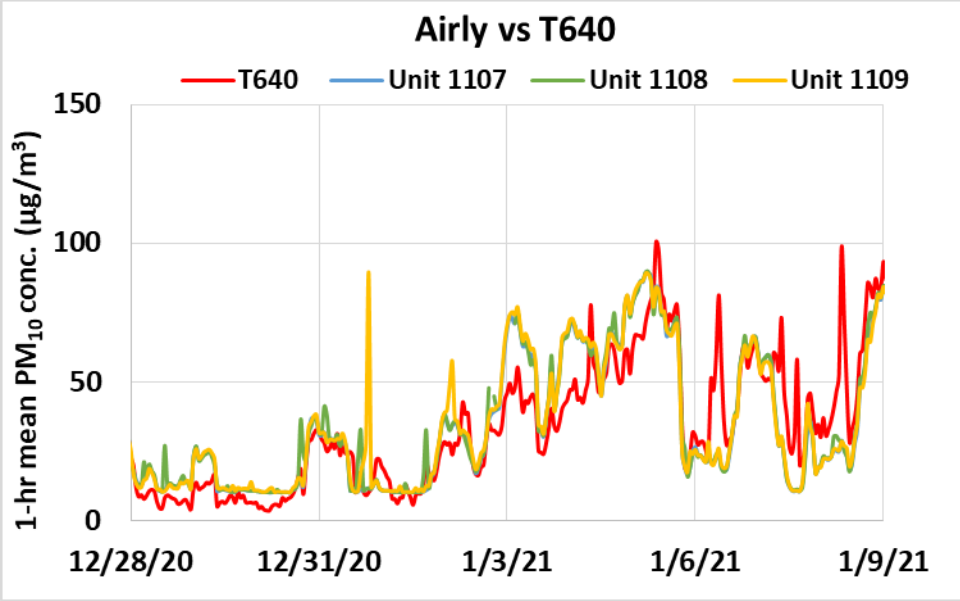
- Airly sensors showed strong correlations with the corresponding FEM T640 data ($0.86 < R^2 < 0.90$)
- Overall, the Airly sensors underestimated the $PM_{2.5}$ mass concentration as measured by the FEM T640
- The Airly sensors seemed to track the diurnal $PM_{2.5}$ variations as recorded by the FEM T640



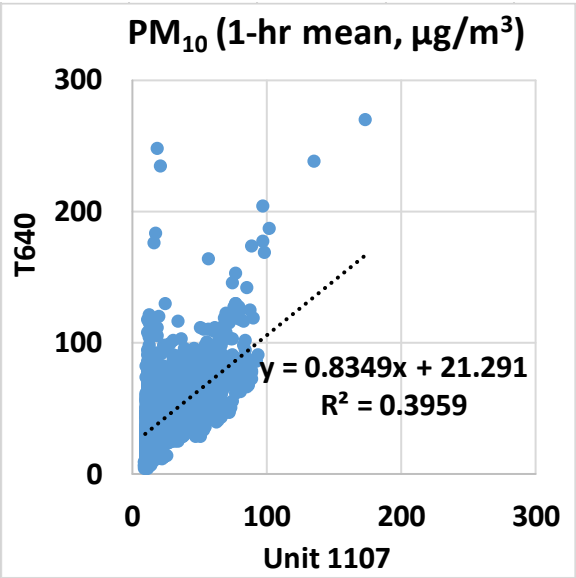


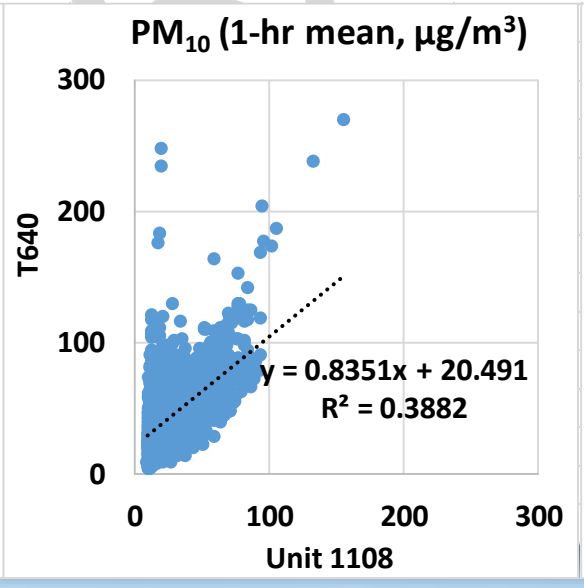


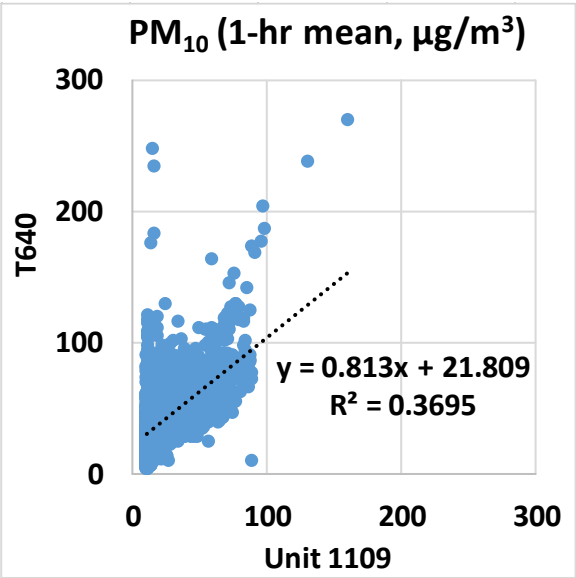
Airly vs T640 (PM_{10}; 1-hr mean)



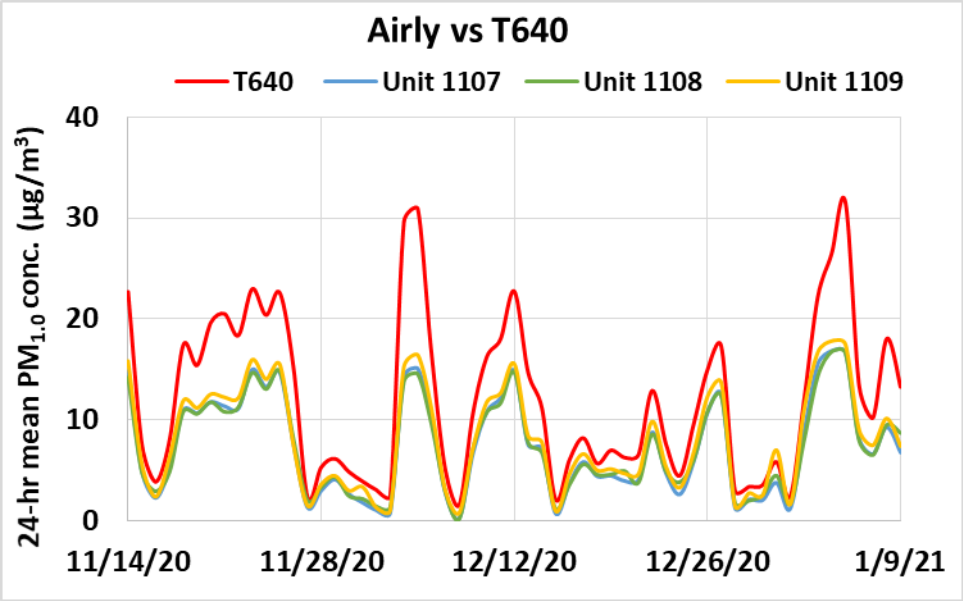
- Airly sensors showed weak correlations with the corresponding T640 data ($0.36 < R^2 < 0.40$)
- Overall, the Airly sensors underestimated the PM_{10} mass concentration as measured by the T640
- The Airly sensors did not seem to track the diurnal PM_{10} variations as recorded by the T640



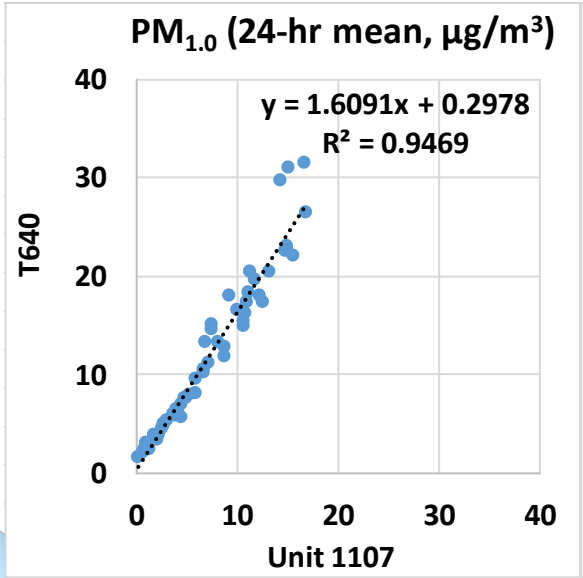


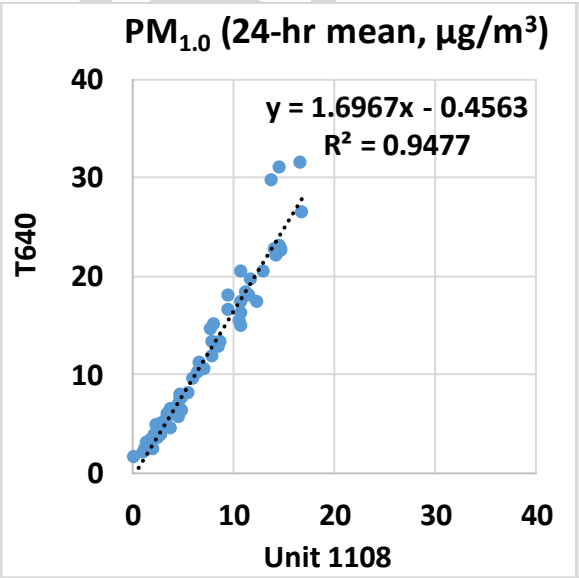


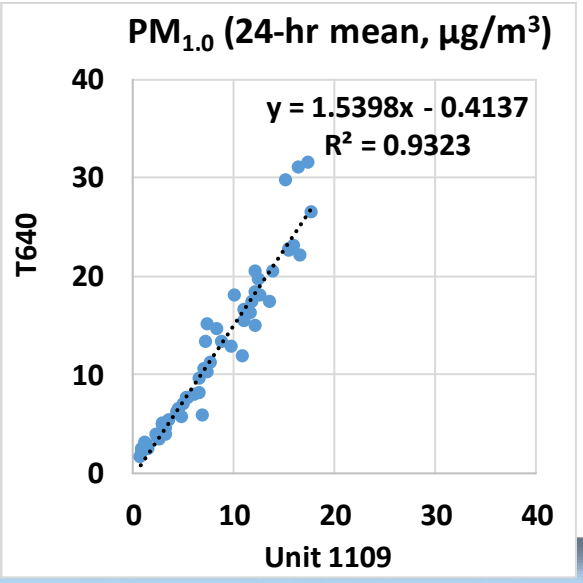
Airly vs T640 ($PM_{1.0}$; 24-hr mean)



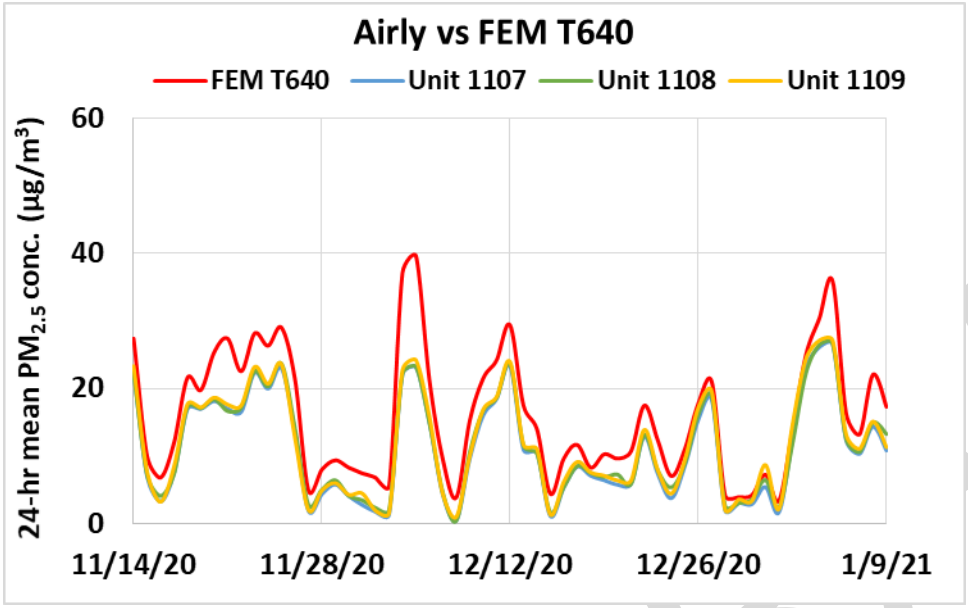
- Airly sensors showed very strong correlations with the corresponding T640 data ($0.93 < R^2 < 0.95$)
- Overall, the Airly sensors underestimated the $PM_{1.0}$ mass concentration as measured by the T640
- The Airly sensors seemed to track the diurnal $PM_{1.0}$ variations as recorded by the T640



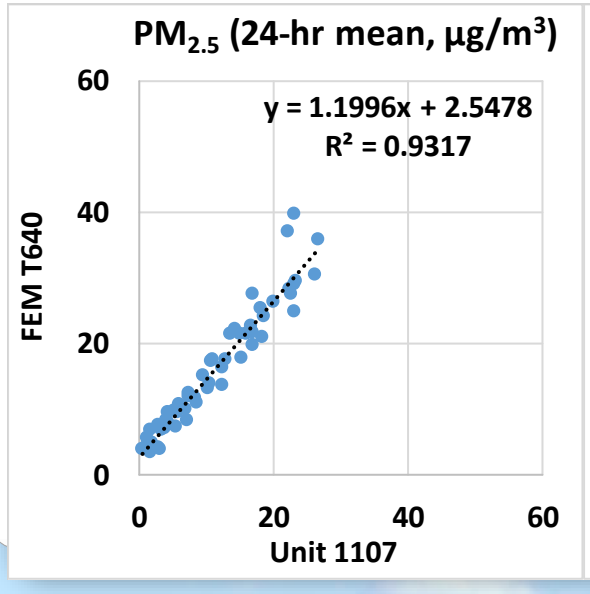


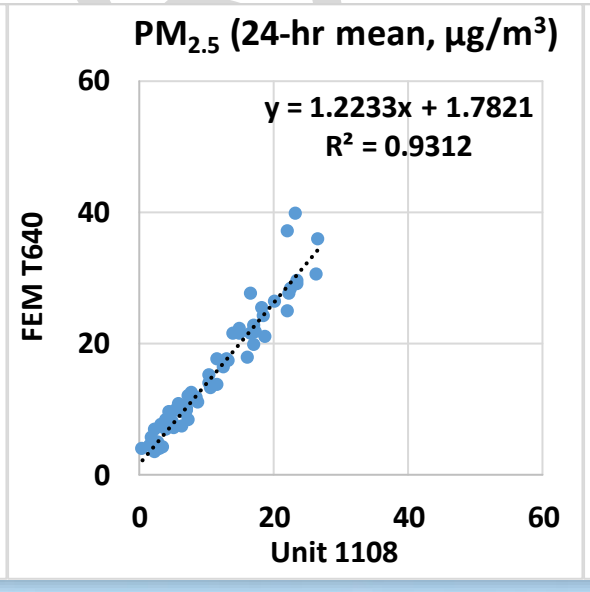


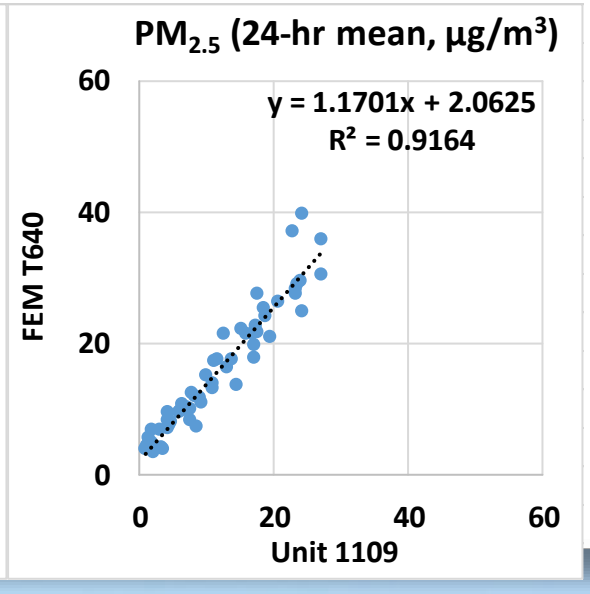
Airly vs FEM T640 ($PM_{2.5}$; 24-hr mean)



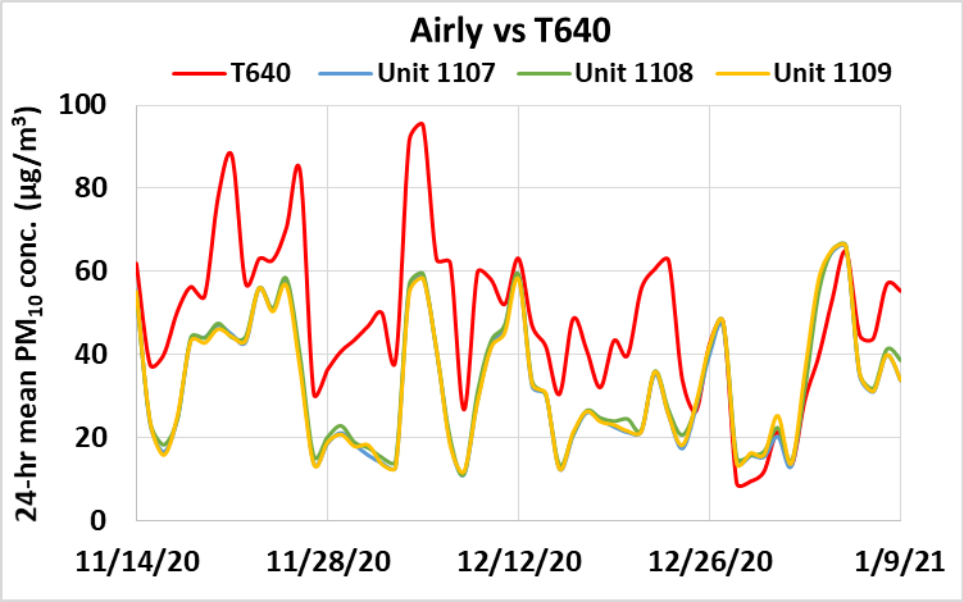
- Airly sensors showed very strong correlations with the corresponding FEM T640 data ($0.91 < R^2 < 0.94$)
- Overall, the Airly sensors underestimated the $PM_{2.5}$ mass concentration as measured by the FEM T640
- The Airly sensors seemed to track the diurnal $PM_{2.5}$ variations as recorded by the FEM T640



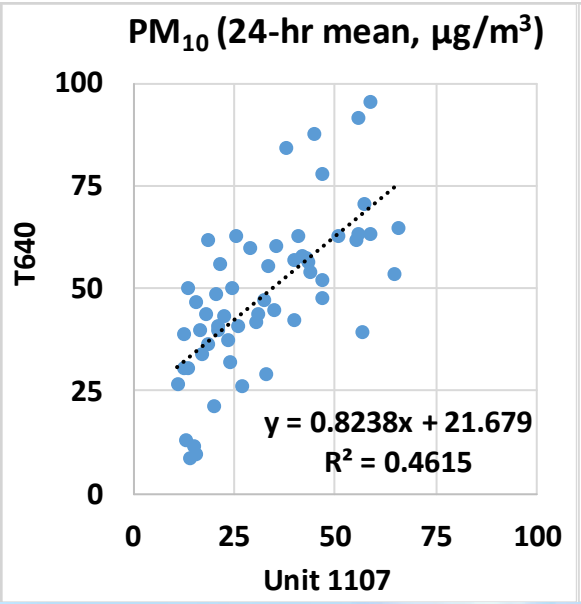


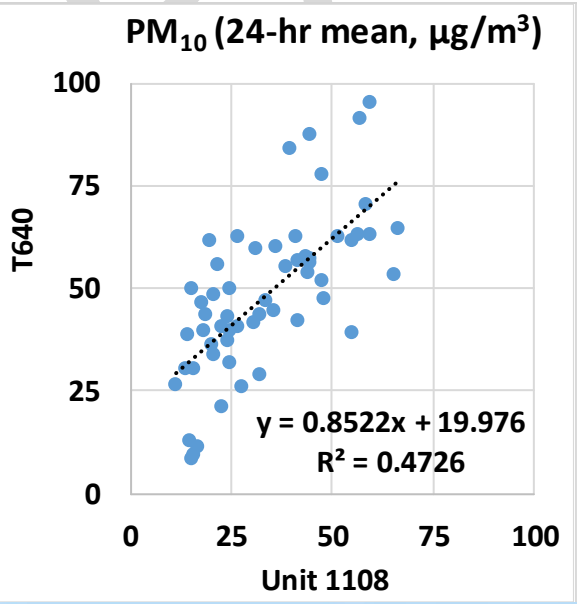


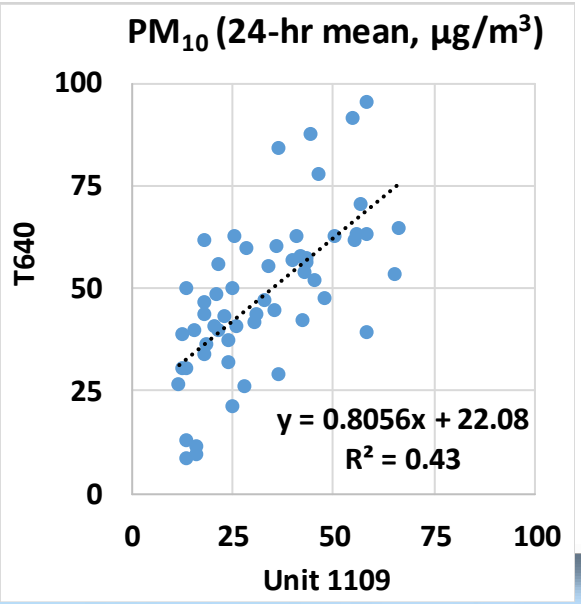
Airly vs T640 (PM_{10}; 24-hr mean)



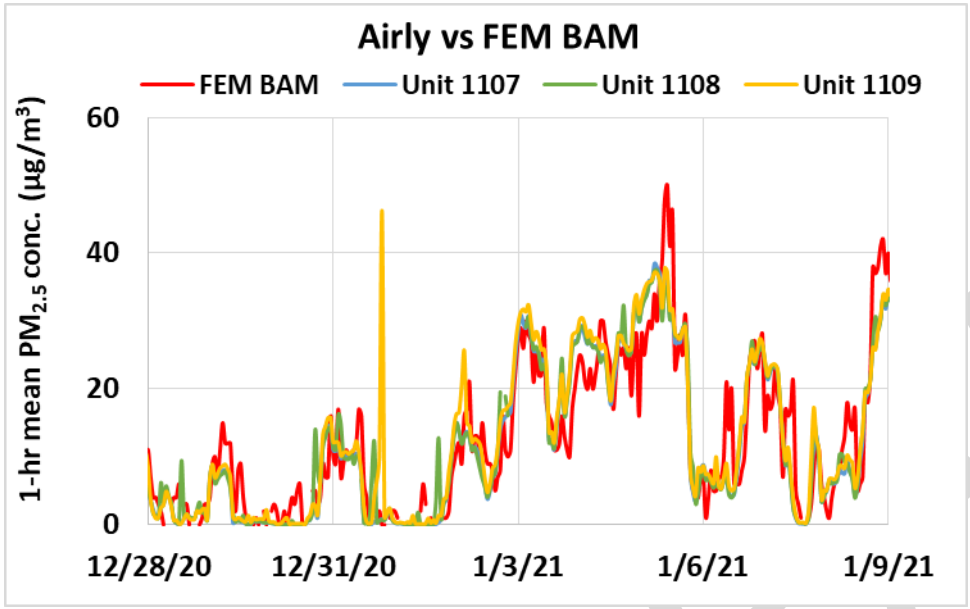
- Airly sensors showed weak correlations with the corresponding T640 data ($0.43 < R^2 < 0.48$)
- Overall, the Airly sensors underestimated the PM_{10} mass concentration as measured by the T640
- The Airly sensors did not seem to track the diurnal PM_{10} variations as recorded by the T640



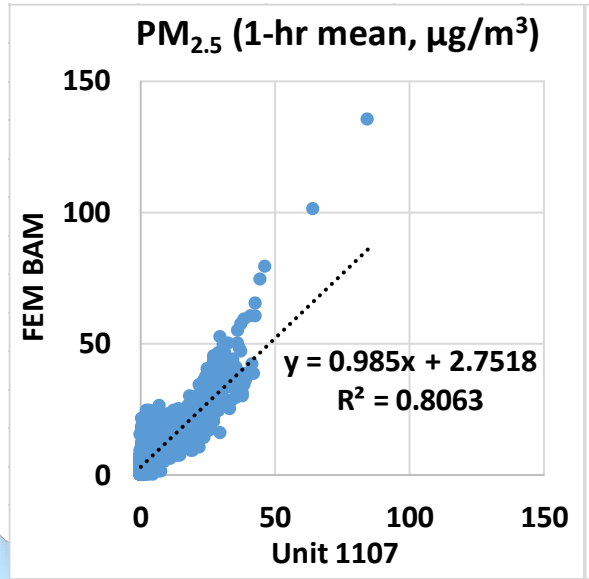


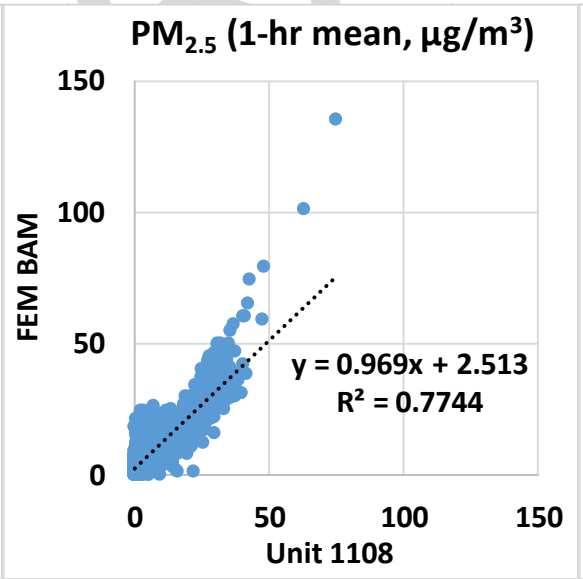


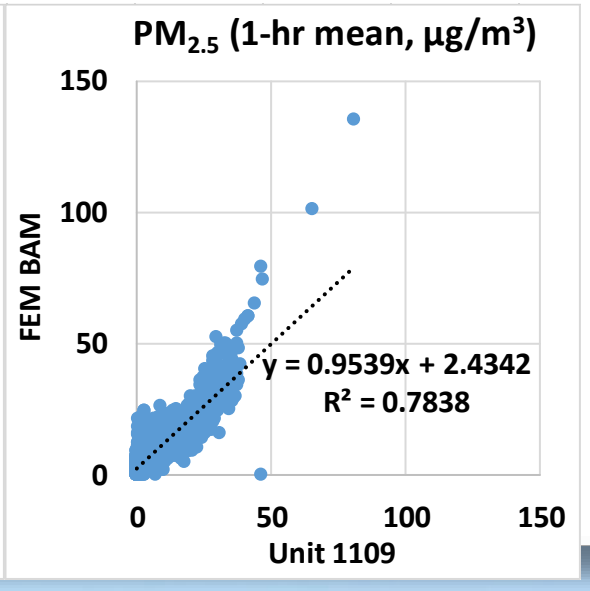
Airly vs FEM BAM ($PM_{2.5}$; 1-hr mean)



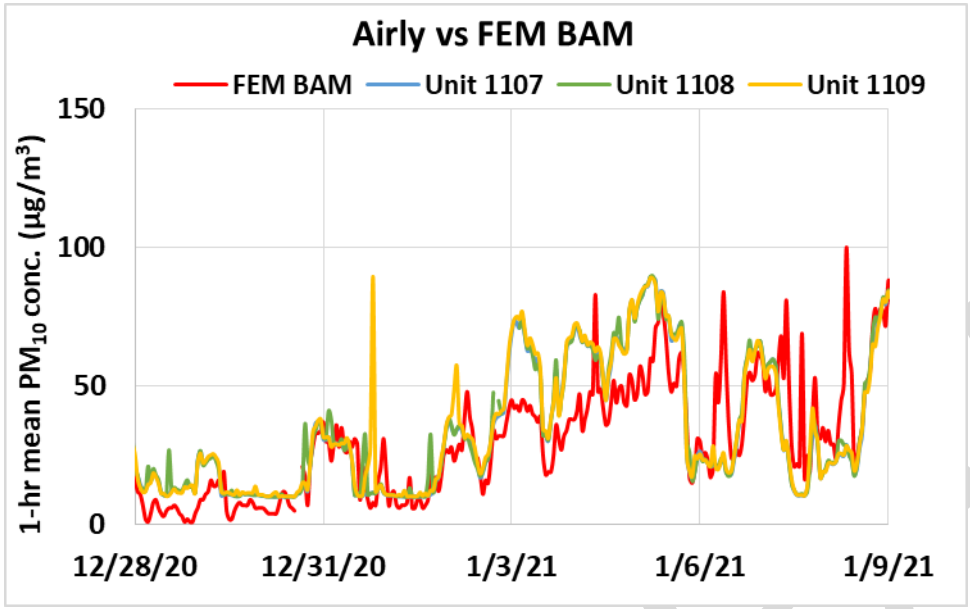
- Airly sensors showed strong correlations with the corresponding FEM BAM data ($0.77 < R^2 < 0.81$)
- Overall, the Airly sensors underestimated the $PM_{2.5}$ mass concentration as measured by the FEM BAM
- The Airly sensors seemed to track the diurnal $PM_{2.5}$ variations as recorded by the FEM BAM



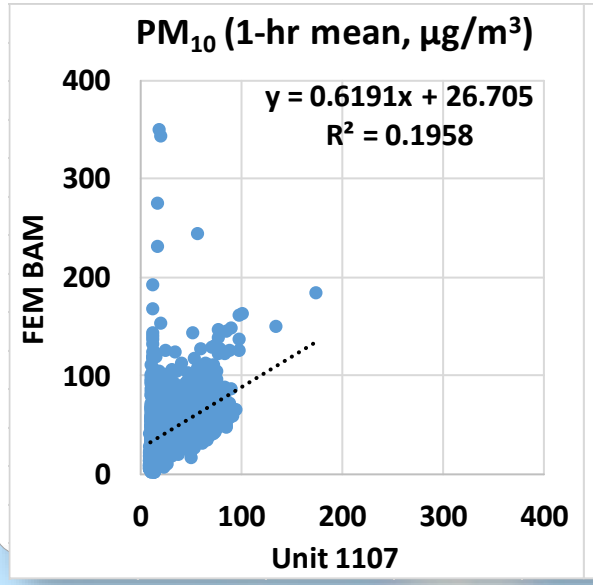


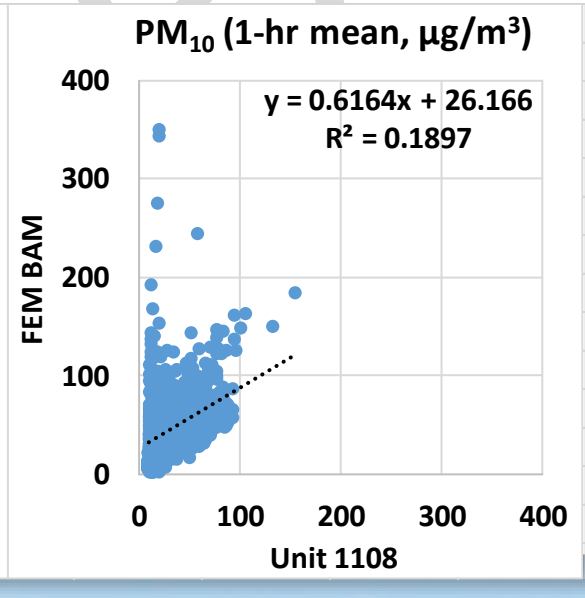


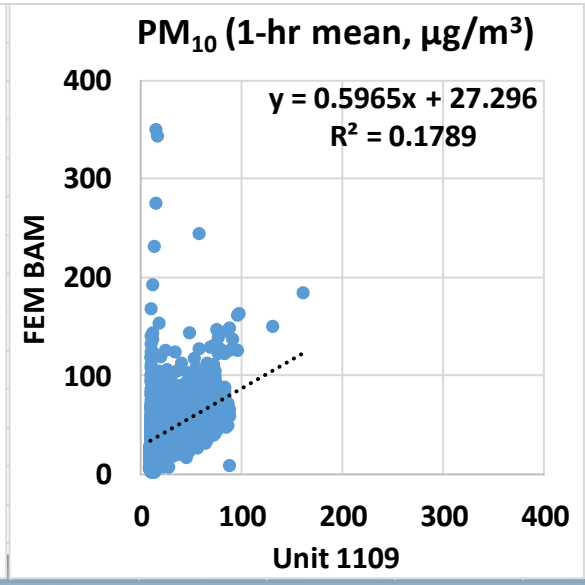
Airly vs FEM BAM (PM_{10}; 1-hr mean)



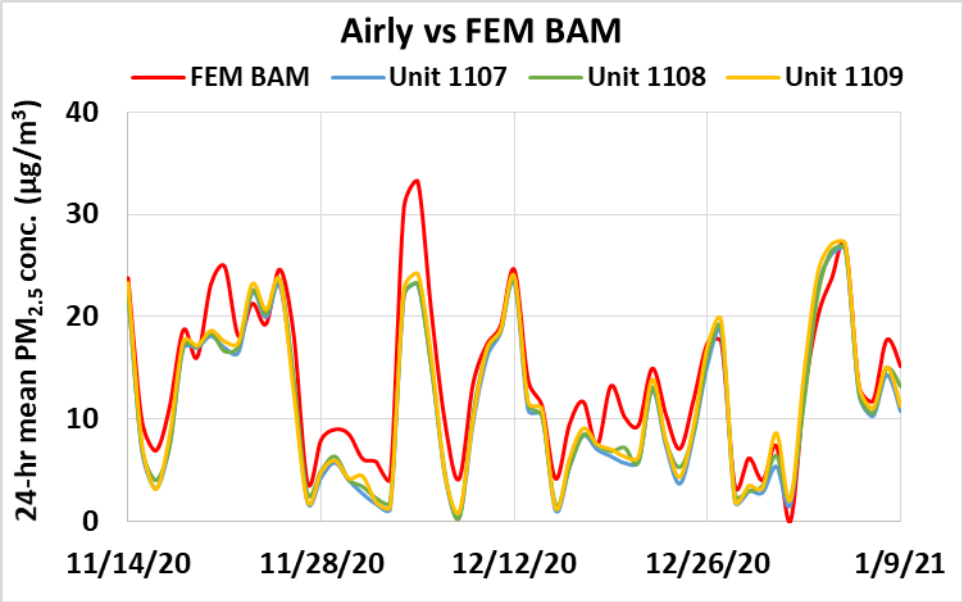
- Airly sensors showed very weak correlations with the corresponding FEM BAM data ($0.17 < R^2 < 0.20$)
- Overall, the Airly sensors underestimated the PM_{10} mass concentration as measured by the FEM BAM
- The Airly sensors did not seem to track the diurnal PM_{10} variations as recorded by the FEM BAM



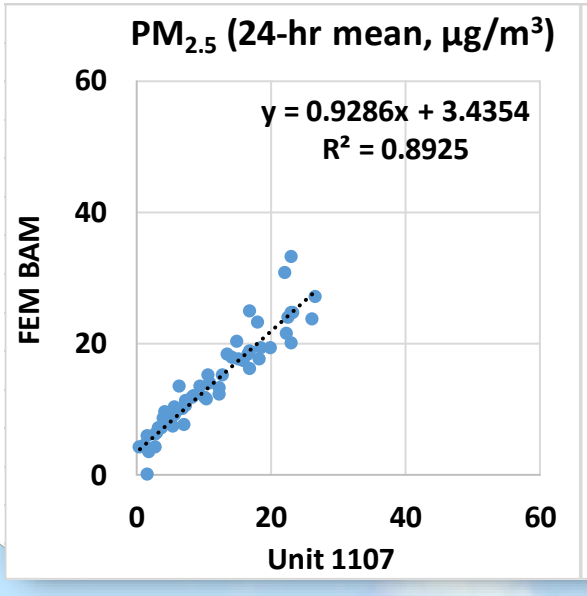


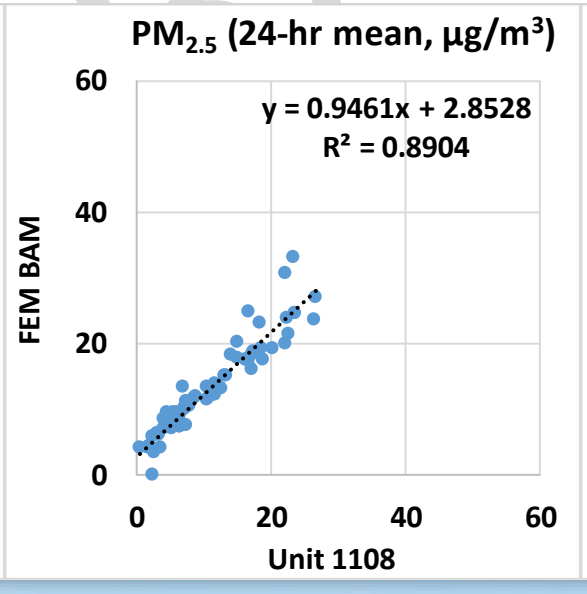


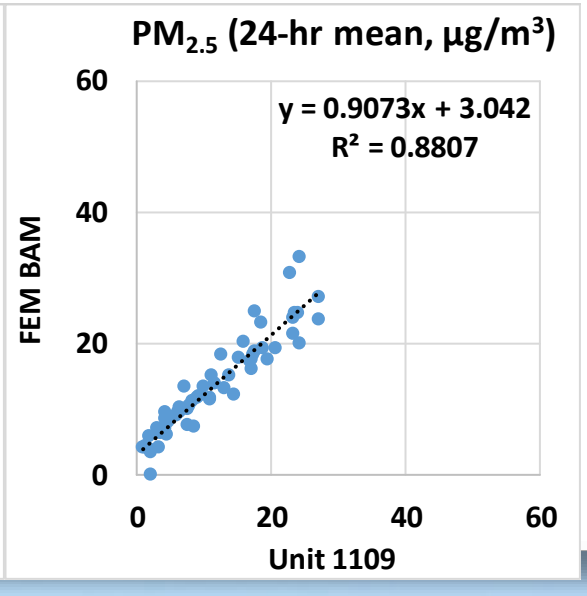
Airly vs FEM BAM ($PM_{2.5}$; 24-hr mean)



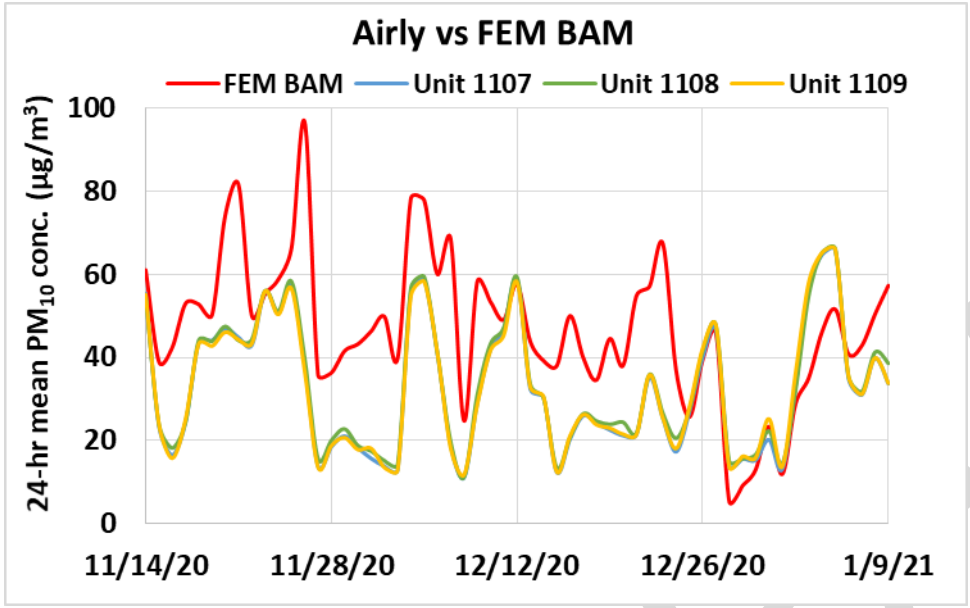
- Airly sensors showed strong correlations with the corresponding FEM BAM data ($0.88 < R^2 < 0.90$)
- Overall, the Airly sensors underestimated the $PM_{2.5}$ mass concentration as measured by the FEM BAM
- The Airly sensors seemed to track the diurnal $PM_{2.5}$ variations as recorded by the FEM BAM



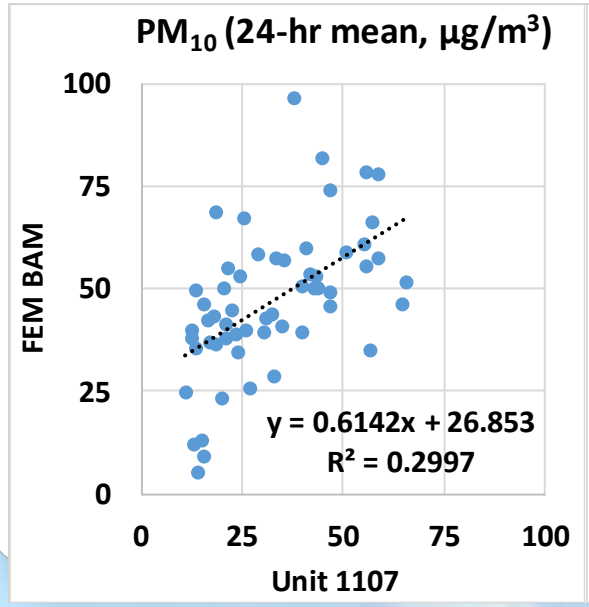


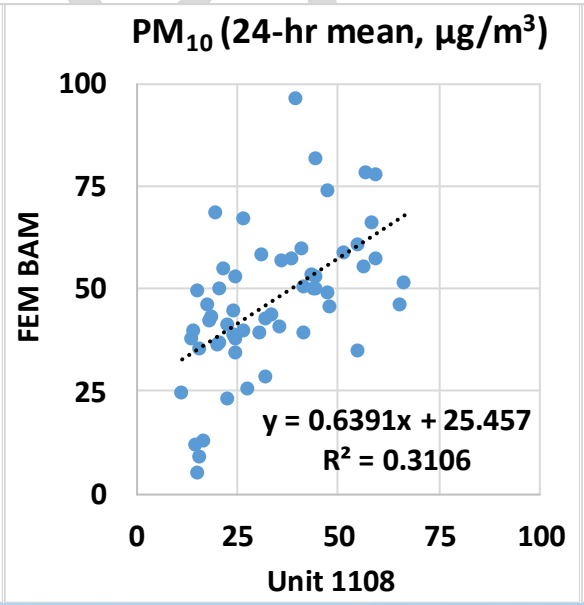


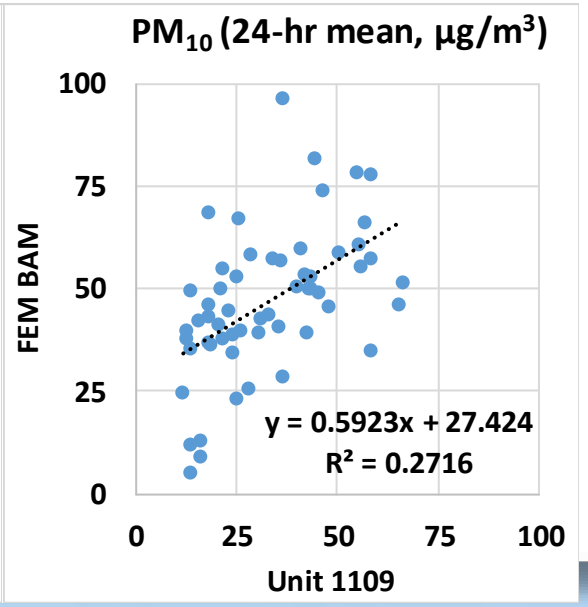
Airly vs FEM BAM (PM_{10}; 24-hr mean)



- Airly sensors showed very weak to weak correlations with the corresponding FEM BAM data ($0.27 < R^2 < 0.32$)
- Overall, the Airly sensors underestimated the PM_{10} mass concentration as measured by the FEM BAM
- The Airly sensors did not seem to track the diurnal PM_{10} variations as recorded by the FEM BAM







Summary: PM

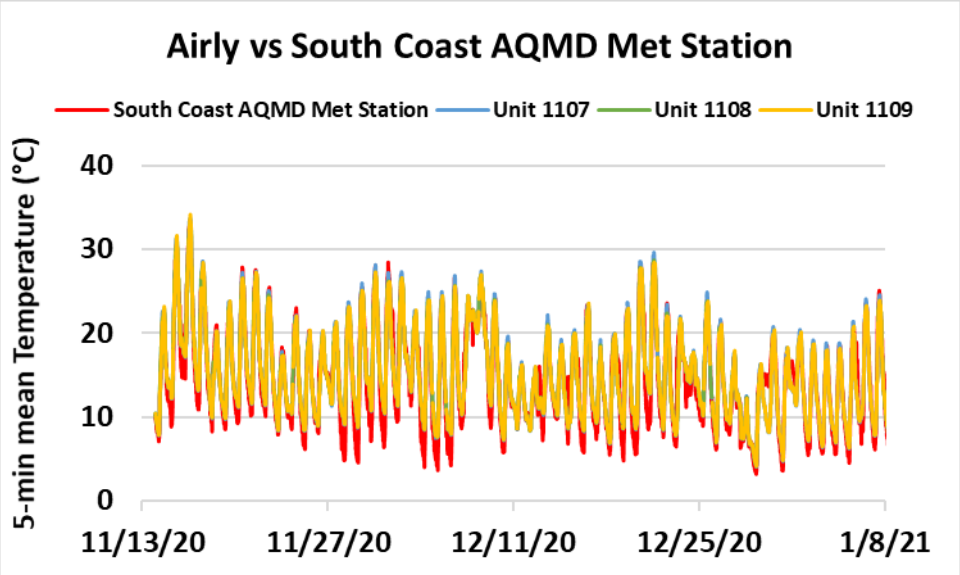
	Average of 3 Sensors, $PM_{1.0}$		Airly vs T640, $PM_{1.0}$						T640 ($PM_{1.0}$, $\mu g/m^3$)		
	Average ($\mu g/m^3$)	SD ($\mu g/m^3$)	R^2	Slope	Intercept	MBE[1] ($\mu g/m^3$)	MAE[2] ($\mu g/m^3$)	RMSE[3] ($\mu g/m^3$)	Ref. Average	Ref. SD	Range during the field evaluation
5-min	7.7	7.2	0.79 to 0.89	1.51 to 1.65	-0.17 to 0.6	-4.8 to -4.0	4.2 to 5.3	11.9 to 13.2	12.1	12.3	0.3 to 217.0
1-hr	7.6	7.0	0.86 to 0.91	1.54 to 1.66	-0.4 to -0.07	-4.8 to -4.0	4.2 to 5.1	7.1 to 7.9	12.1	12.0	0.4 to 63.2
24-hr	7.6	4.9	0.93 to 0.95	1.54 to 1.70	-0.4 to 0.3	-4.8 to -4.0	4.0 to 4.8	14.1 to 14.7	12.1	8.1	1.5 to 31.6
	Average of 3 Sensors, $PM_{2.5}$		Airly vs FEM BAM & FEM T640, $PM_{2.5}$						FEM BAM and FEM T640 ($PM_{2.5}$, $\mu g/m^3$)		
	Average ($\mu g/m^3$)	SD ($\mu g/m^3$)	R^2	Slope	Intercept	MBE[1] ($\mu g/m^3$)	MAE[2] ($\mu g/m^3$)	RMSE[3] ($\mu g/m^3$)	Ref. Average	Ref. SD	Range during the field evaluation
5-min	11.5	10.9	0.83 to 0.89	1.17 to 1.22	2.0 to 2.3	-4.8 to -4.1	4.5 to 5.0	11.1 to 12.0	15.9	14.0	0.8 to 239.7
1-hr	11.5	10.7	0.77 to 0.90	0.95 to 1.22	1.9 to 2.8	-4.8 to -1.9	4.0 to 4.8	5.7 to 7.0	14.0 to 15.9	11.7 to 13.7	0 to 165.1
24-hr	11.5	7.5	0.88 to 0.93	0.91 to 1.22	1.8 to 3.4	-4.6 to -2.0	2.7 to 4.8	3.3 to 5.6	14.0 to 15.9	7.2 to 9.4	3.4 to 39.7
	Average of 3 Sensors, PM_{10}		Airly vs FEM BAM & T640, PM_{10}						FEM BAM and T640 (PM_{10}, $\mu g/m^3$)		
	Average ($\mu g/m^3$)	SD ($\mu g/m^3$)	R^2	Slope	Intercept	MBE[1] ($\mu g/m^3$)	MAE[2] ($\mu g/m^3$)	RMSE[3] ($\mu g/m^3$)	Ref. Average	Ref. SD	Range during the field evaluation
5-min	33.0	22.4	0.34 to 0.37	0.81 to 0.84	21.2 to 21.9	-15.9 to -15.0	19.3 to 19.7	47.0 to 47.6	48.5	30.8	1.3 to 547.2
1-hr	33.0	21.9	0.18 to 0.40	0.60 to 0.84	20.5 to 27.3	-15.9 to -15.0	18.7 to 19.4	27.6 to 32.5	46.9 to 48.5	29.3 to 30.9	1 to 349
24-hr	33.0	15.4	0.27 to 0.47	0.59 to 0.85	20.0 to 27.4	-15.7 to -13.1	16.4 to 17.8	20.5 to 21.4	46.9 to 48.5	17.5 to 18.9	5.4 to 96.5

[1] Mean Bias Error (MBE): the difference between the sensors and the reference instruments. MBE indicates the tendency of the sensors to underestimate (negative MBE values) or overestimate (positive MBE values).

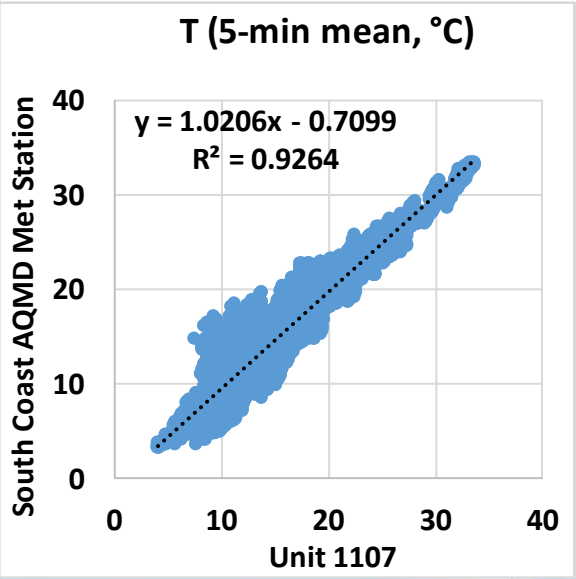
[2] Mean Absolute Error (MAE): the absolute difference between the sensors and the reference instruments. The larger MAE values, the higher measurement errors as compared to the reference instruments.

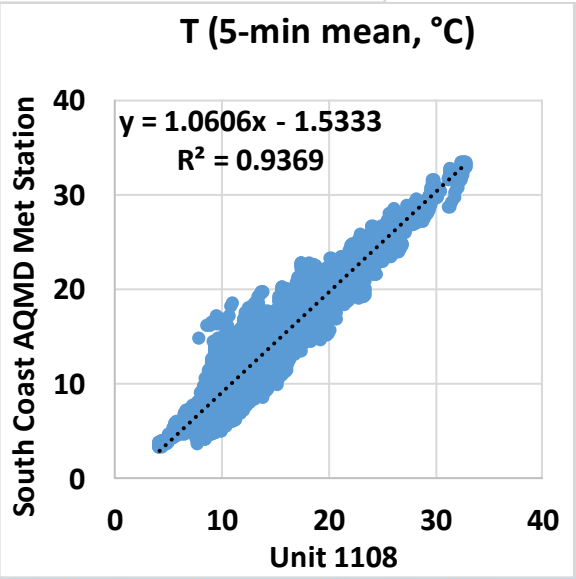
[3] Root Mean Square Error (RMSE): another metric to calculate measurement errors.

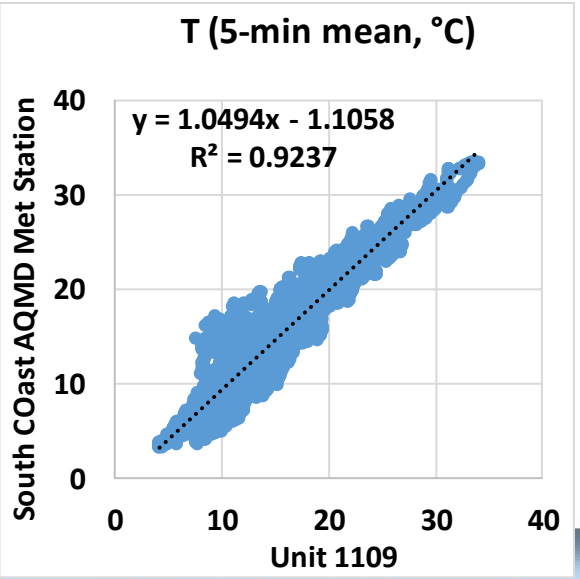
Airly vs South Coast AQMD Met Station (Temp; 5-min mean)



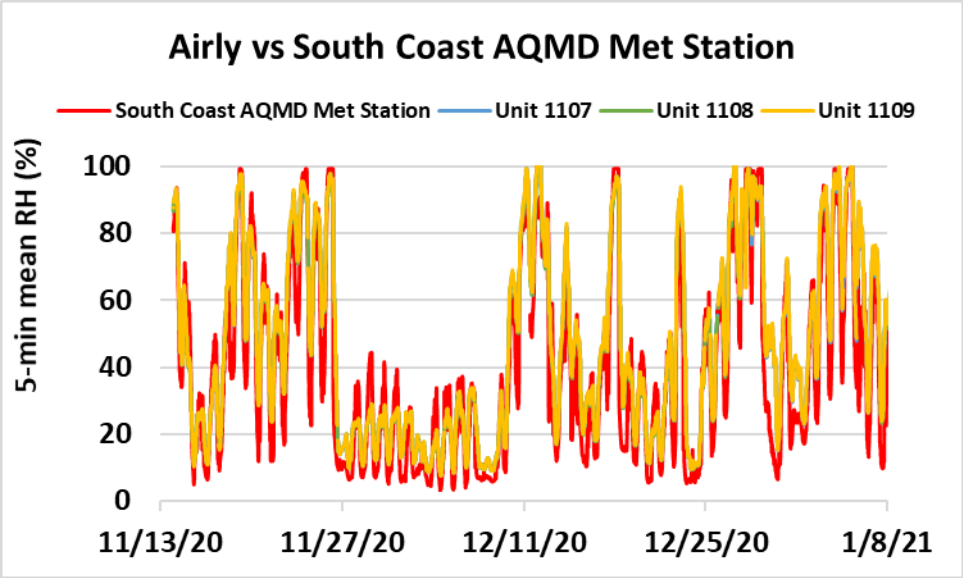
- Airly sensors showed very strong correlations with the corresponding South Coast AQMD Met Station data ($0.92 < R^2 < 0.94$)
- Overall, the Airly sensors overestimated the temperature measurement as recorded by South Coast AQMD Met Station
- The Airly sensors seemed to track the diurnal temperature variations as recorded by South Coast AQMD Met Station



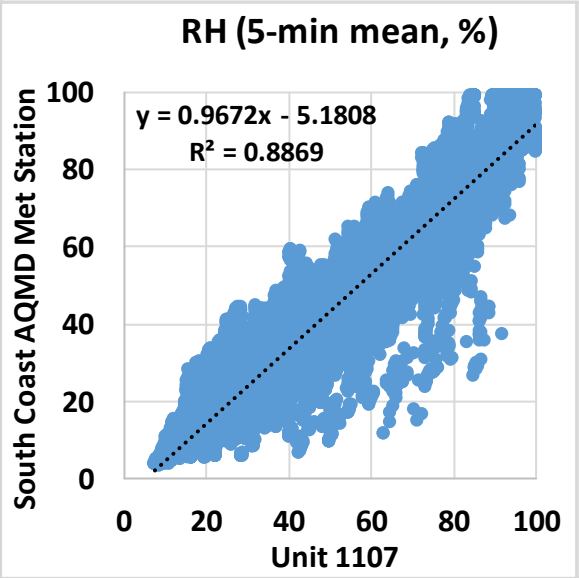


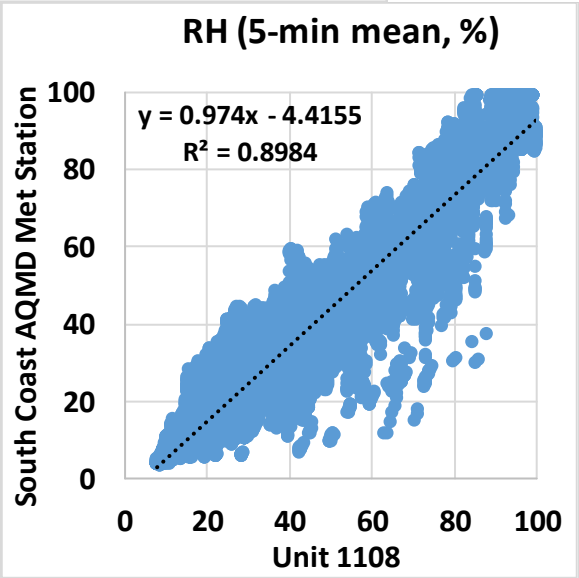


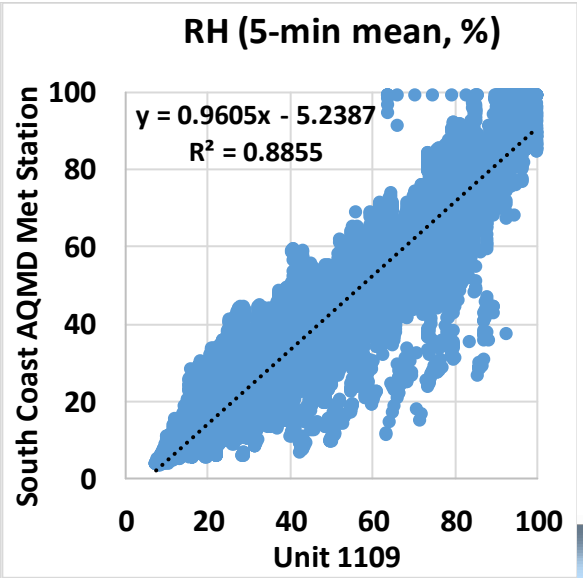
Airly vs South Coast AQMD Met Station (RH; 5-min mean)



- Airly sensors showed strong correlations with the corresponding South Coast AQMD Met Station data (R^2 ~ 0.89)
- Overall, the Airly sensors underestimated the RH measurement as recorded by South Coast AQMD Met Station
- The Airly sensors seemed to track the diurnal RH variations as recorded by South Coast AQMD Met Station







Discussion

- The three **Airly** sensors' average data recovery for ozone, NO_2 and PM was ~ 93%, ~ 93% and ~100%; respectively.
- The absolute intra-model variability was 1.3 ppb for ozone, 4.3 ppb for NO_2, and 0.36, 0.29 and 0.41 µg/m^3 for $PM_{1.0}$, $PM_{2.5}$ and PM_{10}, respectively.
- The reference instruments (FEM BAM and FEM T640) showed very strong and strong correlations with each other for $PM_{2.5}$ and PM_{10} mass concentration measurements (R^2 ~ 0.91 and R^2 ~ 0.88, 1-hr mean), respectively.
- During the entire field deployment testing period:
  - Ozone sensors showed very strong correlations with the FEM instrument ($0.90 < R^2 < 0.94$, 5-min mean) and overestimated the corresponding FEM data
  - NO_2 sensors showed moderate to strong correlations with the FRM instrument ($0.53 < R^2 < 0.81$, 5-min mean) and overestimated the corresponding FRM data
  - The sensors showed strong to very strong correlations with the corresponding $PM_{1.0}$ data ($0.85 < R^2 < 0.91$, 1-hr mean); and showed strong correlations with the corresponding $PM_{2.5}$ data ($0.77 < R^2 < 0.90$, 1-hr mean) and very weak to weak correlations with the corresponding PM_{10} data ($0.17 < R^2 < 0.40$ respectively, 1-hr mean). The Airly sensors underestimated the corresponding $PM_{1.0}$, $PM_{2.5}$ and PM_{10} data.
  - Temperature and relative humidity sensors showed very strong and strong correlations with the South Coast AQMD Met Station data, respectively (T: R^2 ~ 0.93 and RH: R^2 ~ 0.89) and overestimated the T data and underestimated the RH data as recorded by the South Coast AQMD Met Station
- No sensor calibration was performed by AQ-SPEC prior to the beginning of this field testing
- Laboratory chamber testing is necessary to fully evaluate the performance of these sensors under controlled T and RH conditions, and known target and interferent pollutants concentrations.
- These results are still preliminary