pscleanair.org
Puget Sound Clean Air Agency

*Working together for clean air*

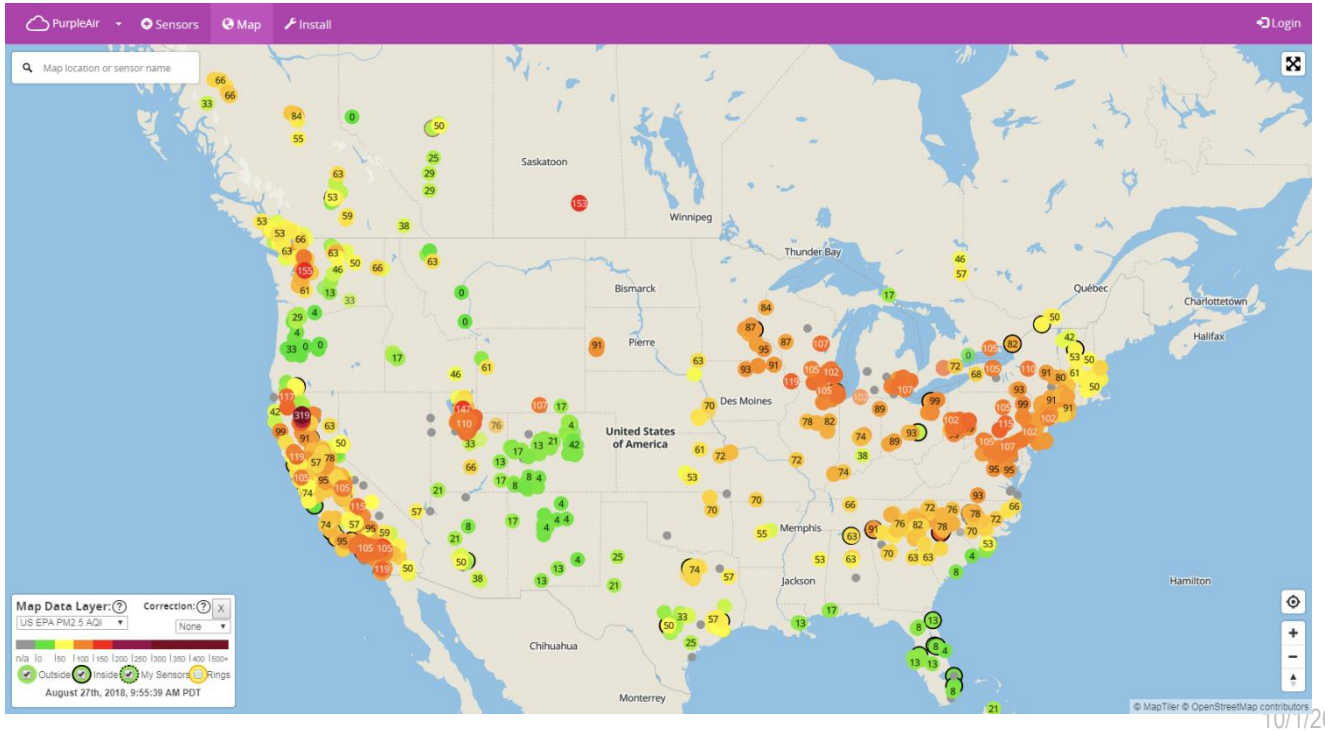
# Purple Air Calibration

## Effect of space, time and environment

Graeme Carvlin, PhD

# Overview

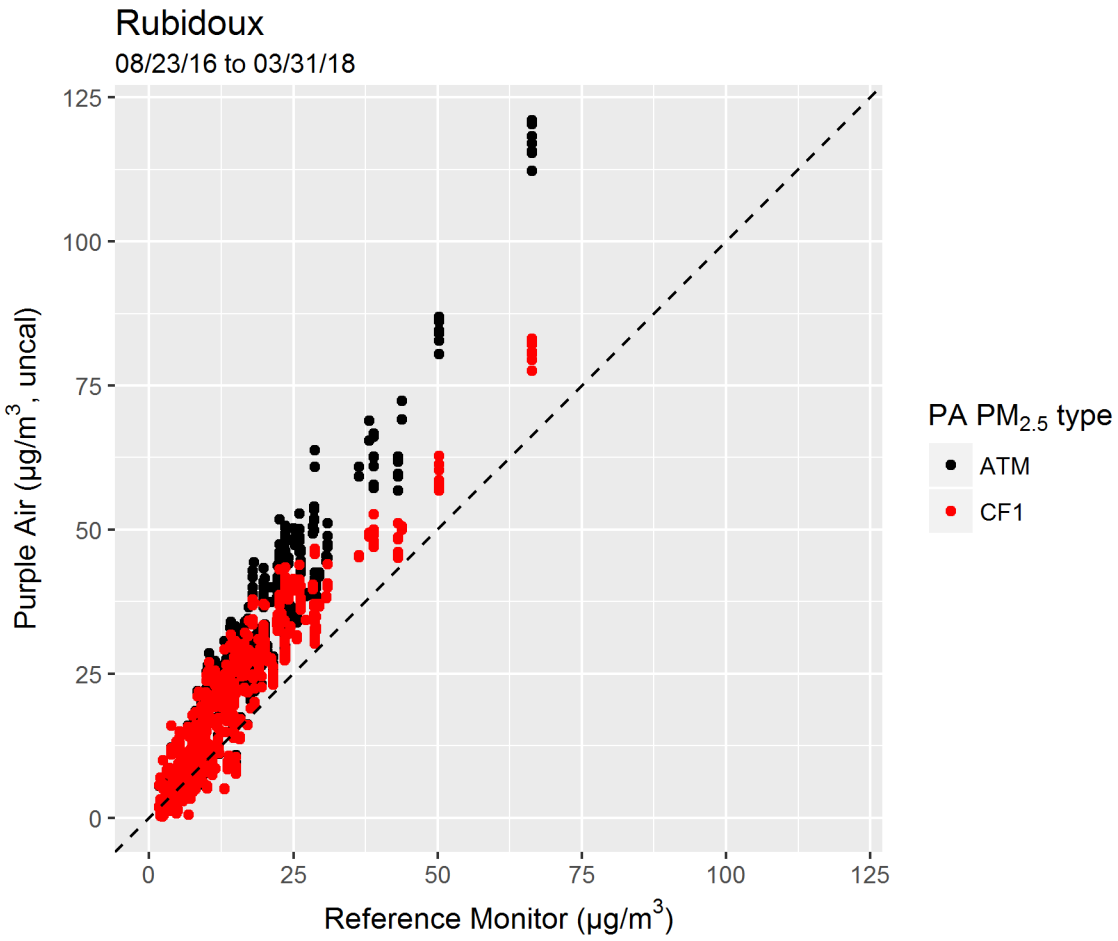
- **Analysis of Purple Airs collocated with reference monitors in the US and British Columbia**
- **How much does the calibration equation change over space, time, and environmental conditions?**

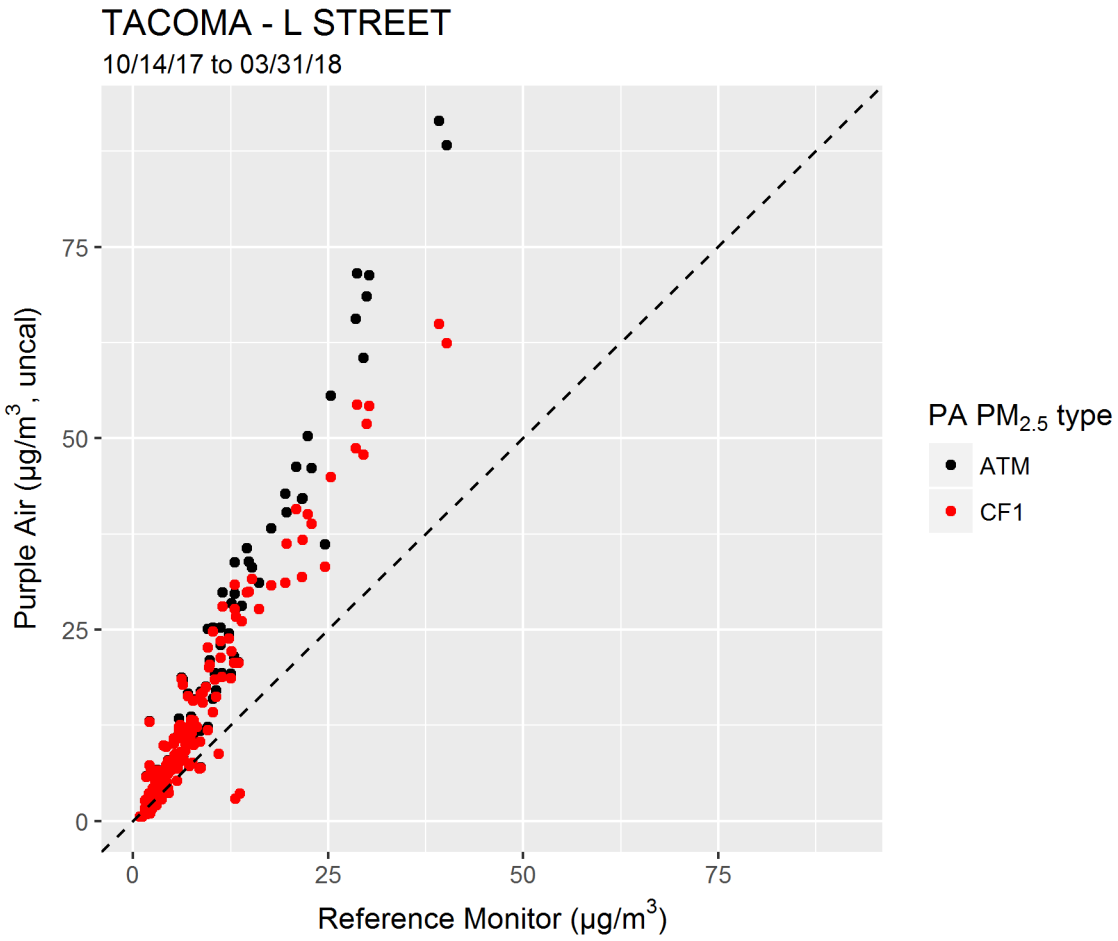


# Data collection

- Data collected August 2018 from Thingspeak, lat/lon from Purple Air (PA)
- Reference data from US EPA's AQS database and Environment Canada
- Selected PAs within 150m of a regulatory site (19 sites)
- QC
  - Data completeness (75%) for averaging; 14 days of data; <1000 ug/m3
  - One sensor reading <5 and the other reading > 15 (one sensor is dead)
  - One sensor is >50 away from other sensor (1+ sensor is bad)
  - Timeseries manual analysis (were PAs at that site the whole time? – need historical lat/lon!)
  - No manual filtering on daily data; manual filtering of hourly data (outliers)
- Averaging regulatory monitors at a single site, keep PAs separate

# Scatterplots (Daily)




Uncalibrated data

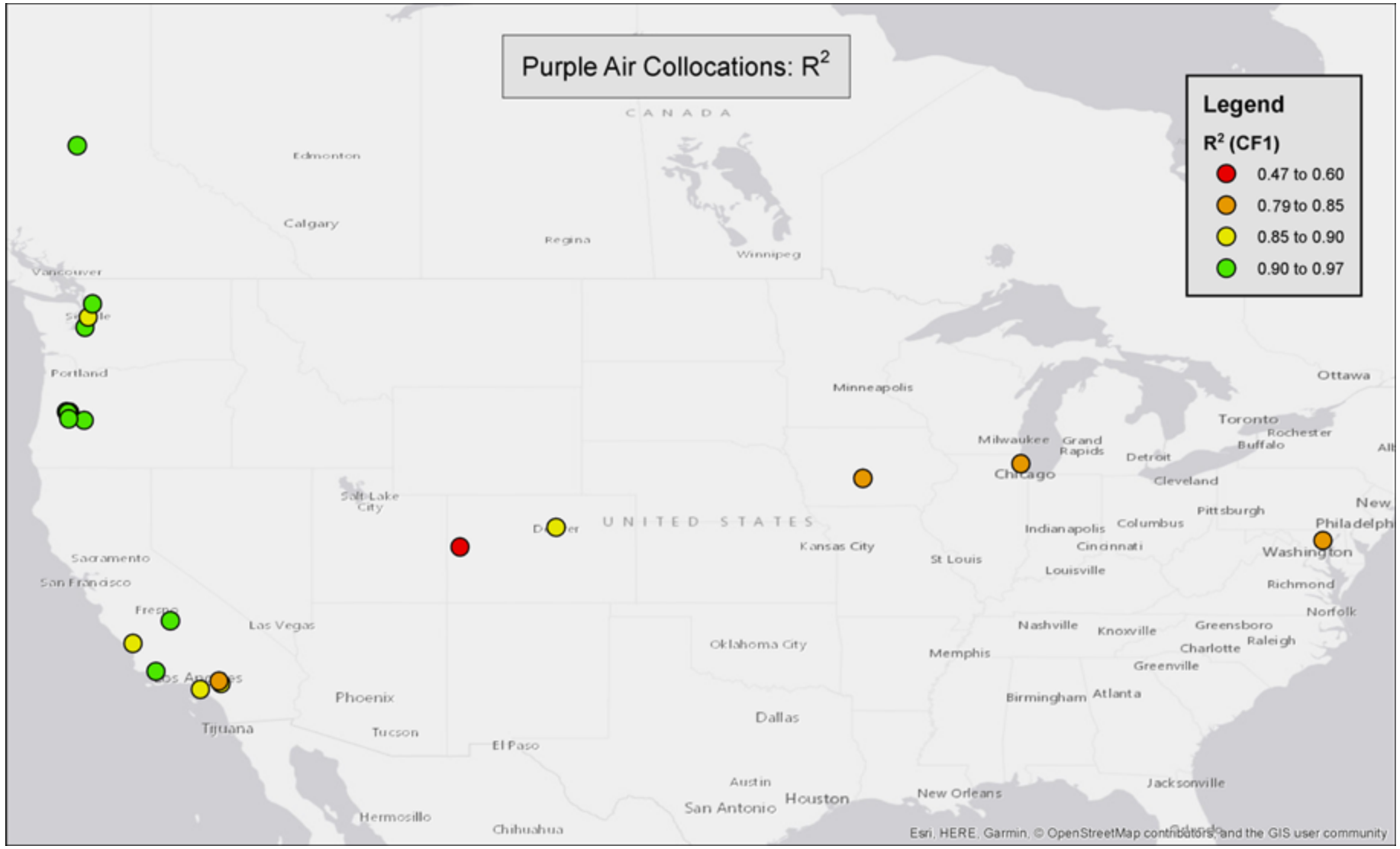
# Comparison

- **Should I use the CF1 or ATM mass concentration value?**

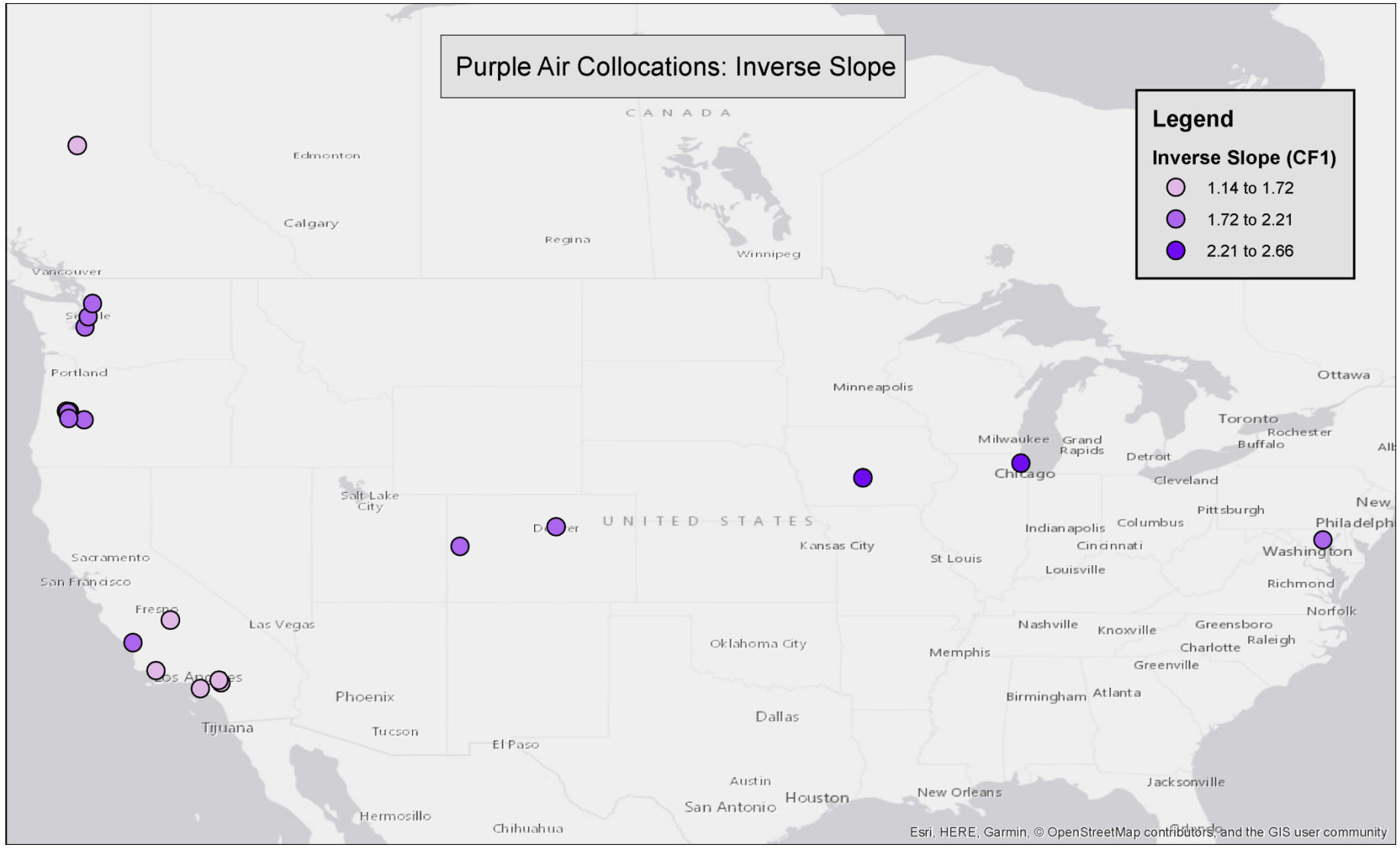
| Data Type | $R^2$ | RMSE | Slope | Inverse Slope | Intercept |
|---|---|---|---|---|---|
| ATM | 0.86 (0.47 to 0.99) | 11.79 (1.99 to 31.34) | 0.45 (0.37 to 0.58) | 2.27 (1.71 to 2.71) | 2.07 (-0.92 to 3.52) |
| CF1 | 0.84 (0.44 to 0.97) | 8.05 (2.03 to 13.61) | 0.55 (0.38 to 0.88) | 1.90 (1.14 to 2.66) | 1.07 (-6.26 to 3.03) |

Notes: Stats are average (min to max) across all 19 sites

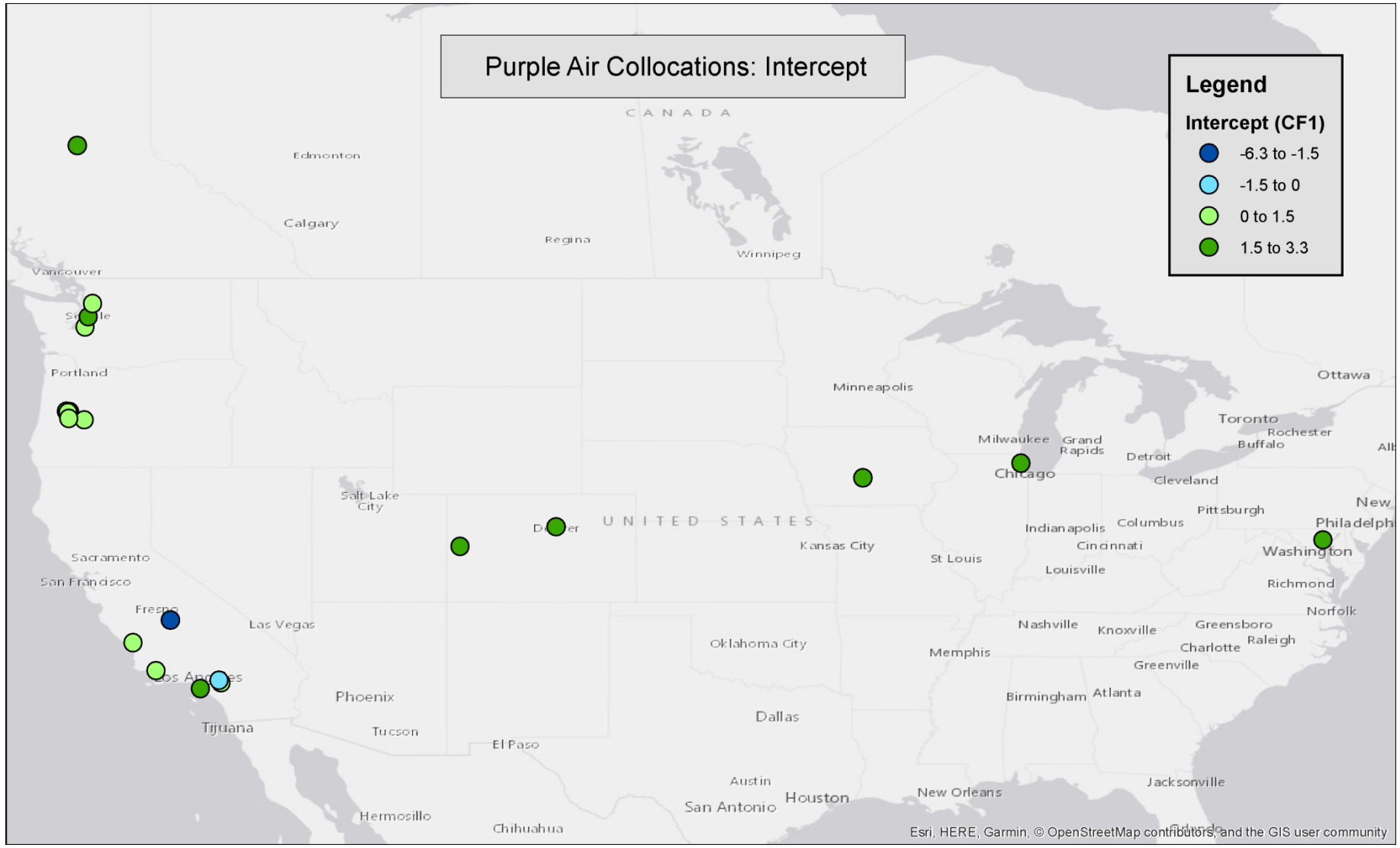
# Calibration Equation Map (Daily): $R^2$



# Calibration Equation Map (Daily): Inverse Slope

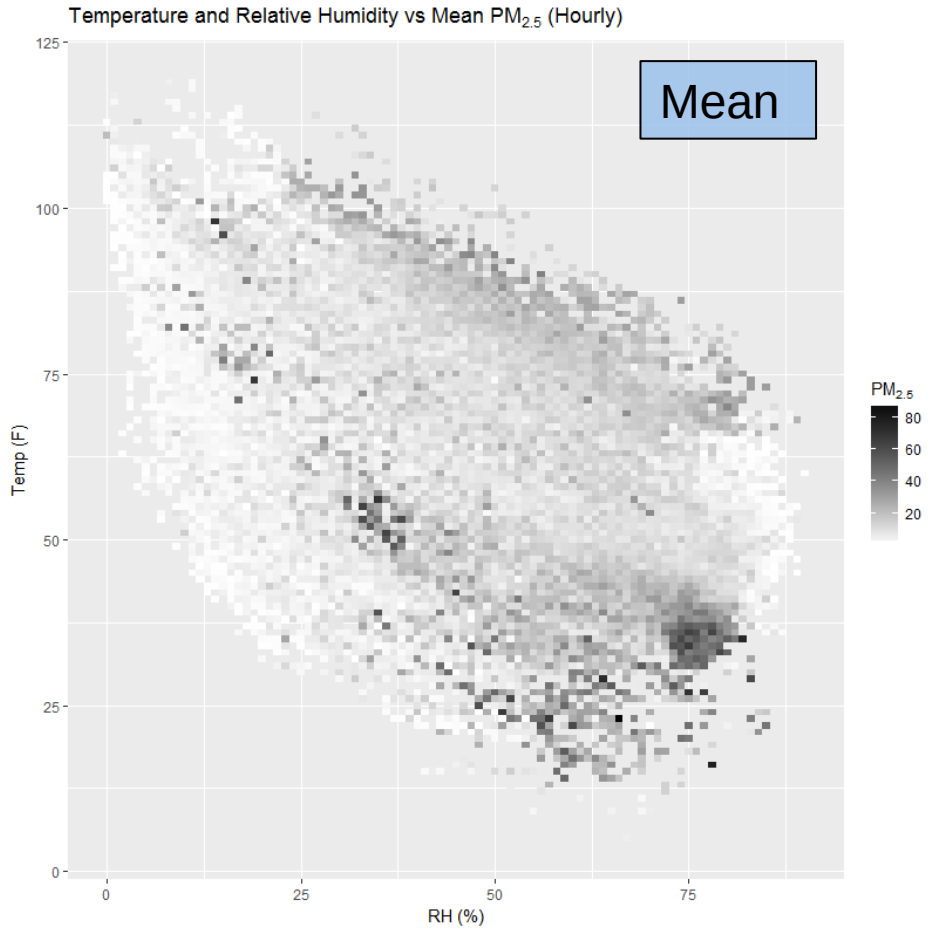


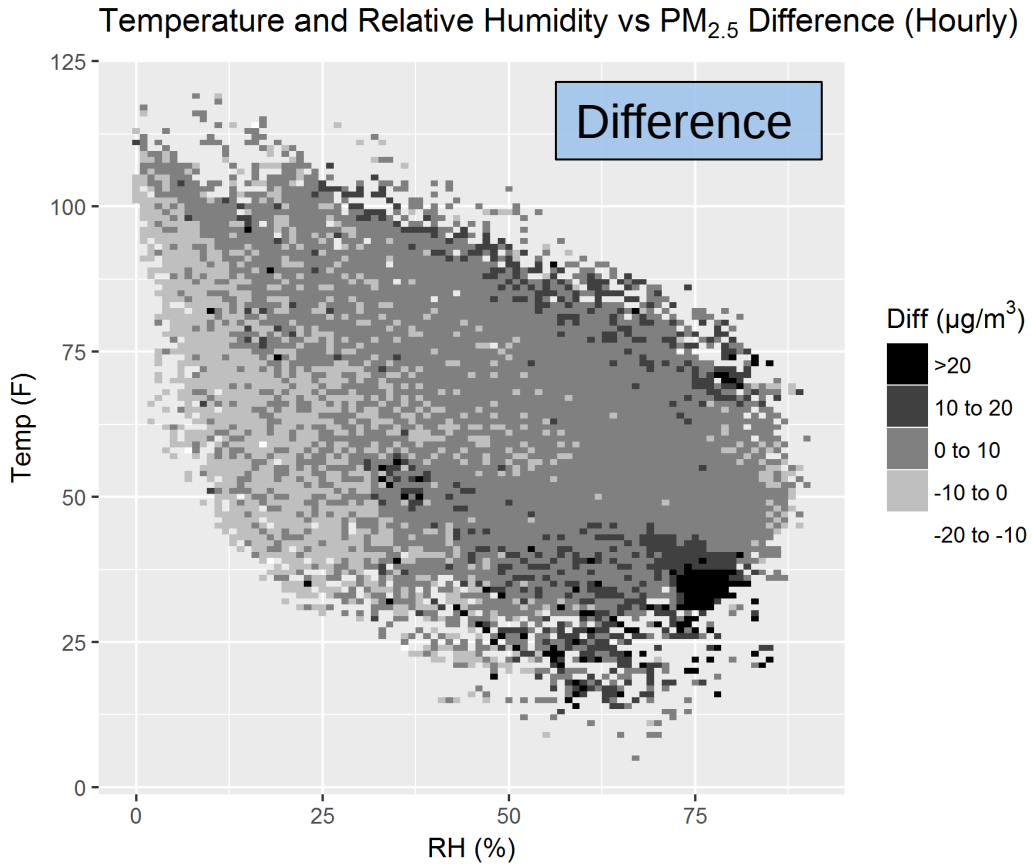
# Calibration Equation Map (Daily): Intercept



# Temperature and Relative Humidity

- **High PM and more bias at high RH and low temp**
  - **Winter (NW)**

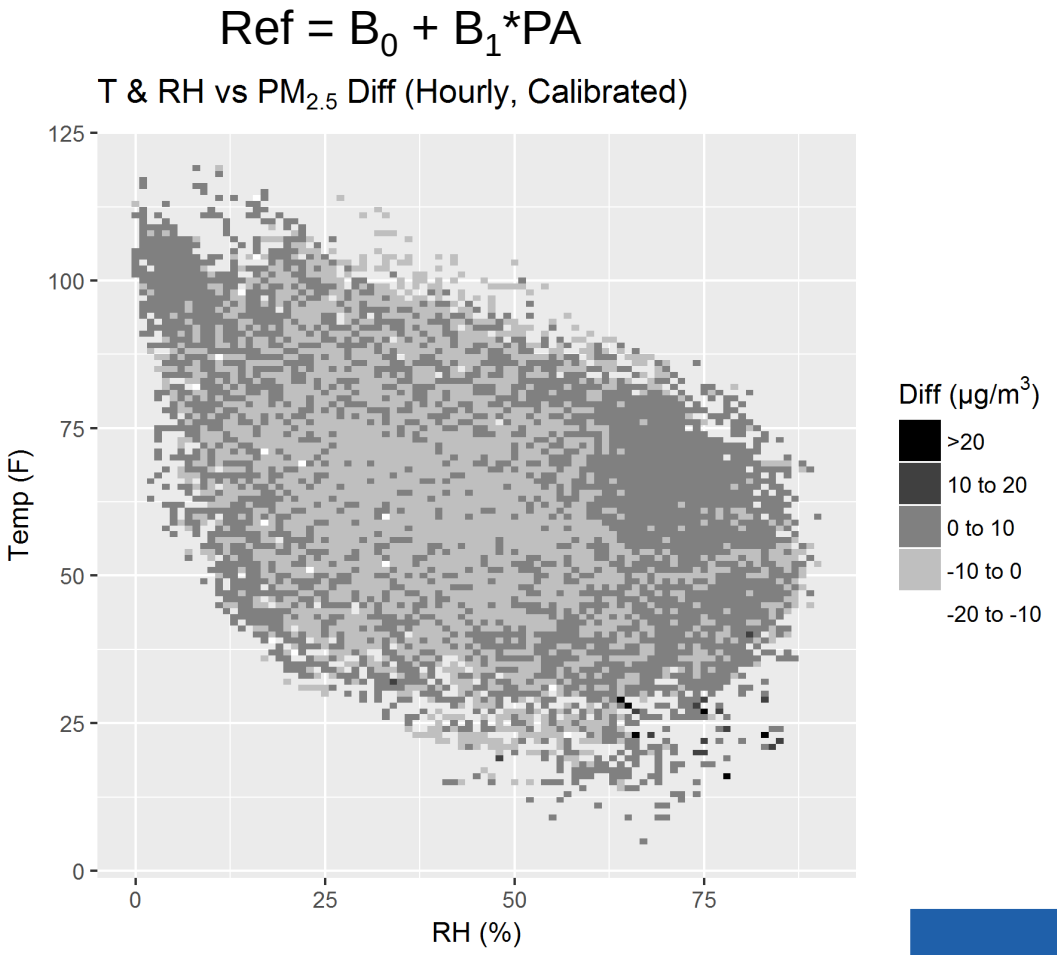


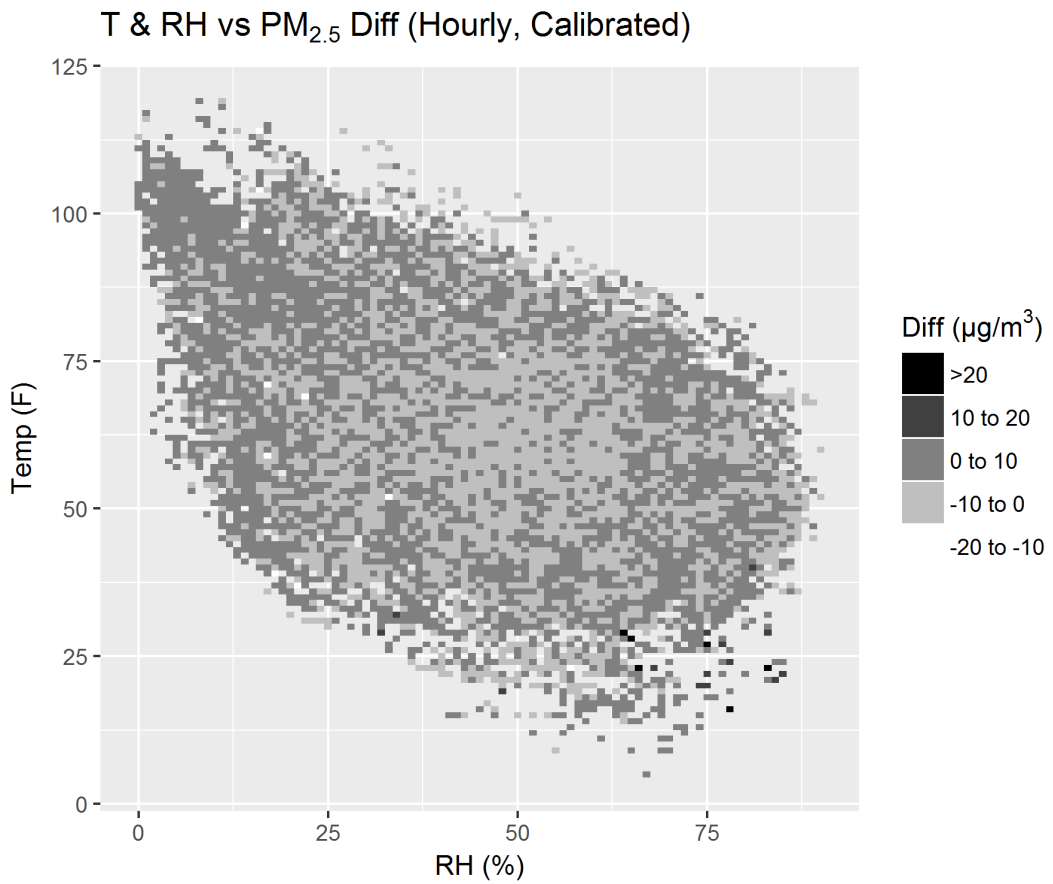

Uncalibrated data

# Calibrated Temp/RH

$$Ref = B_0 + B_1*PA$$




$$Ref = B_0 + B_1*PA + B_2*Temp + B_3*RH$$

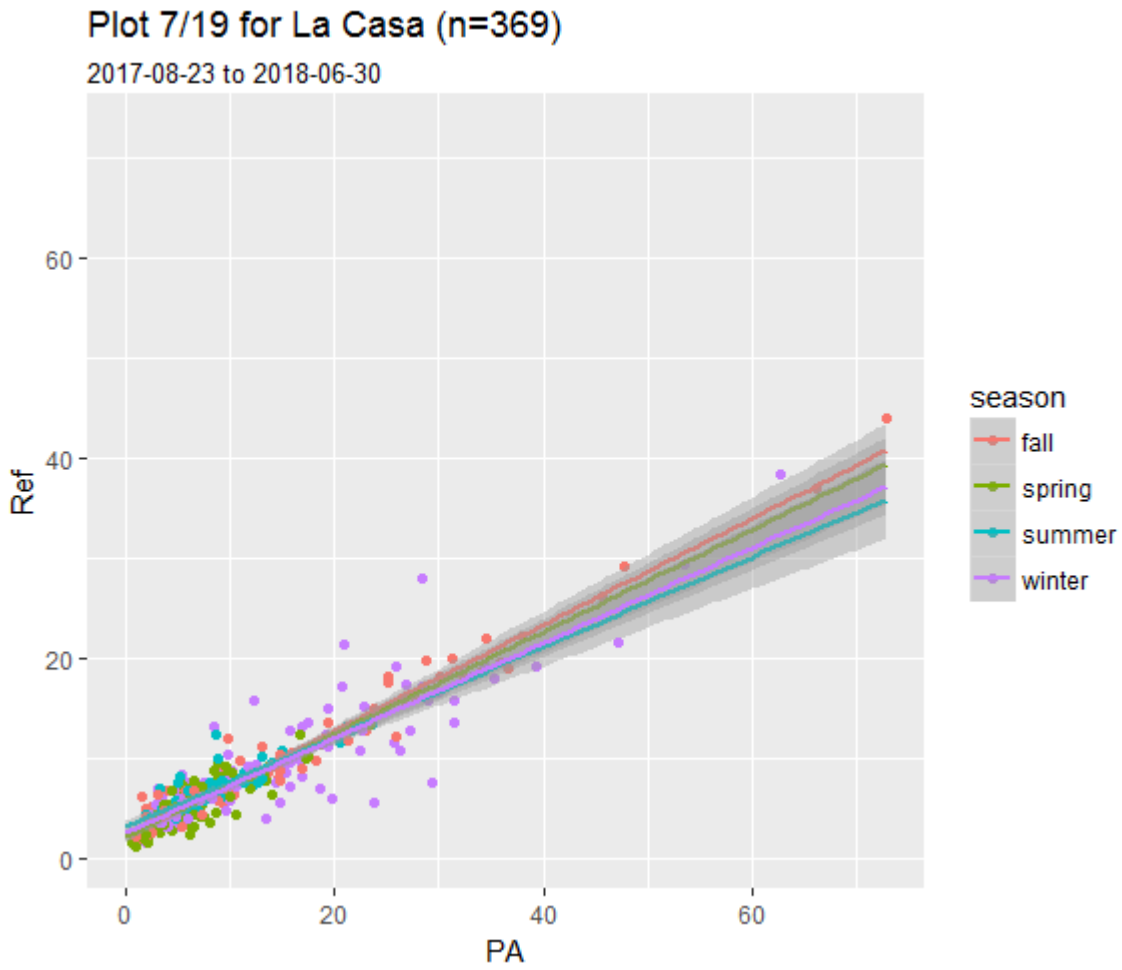


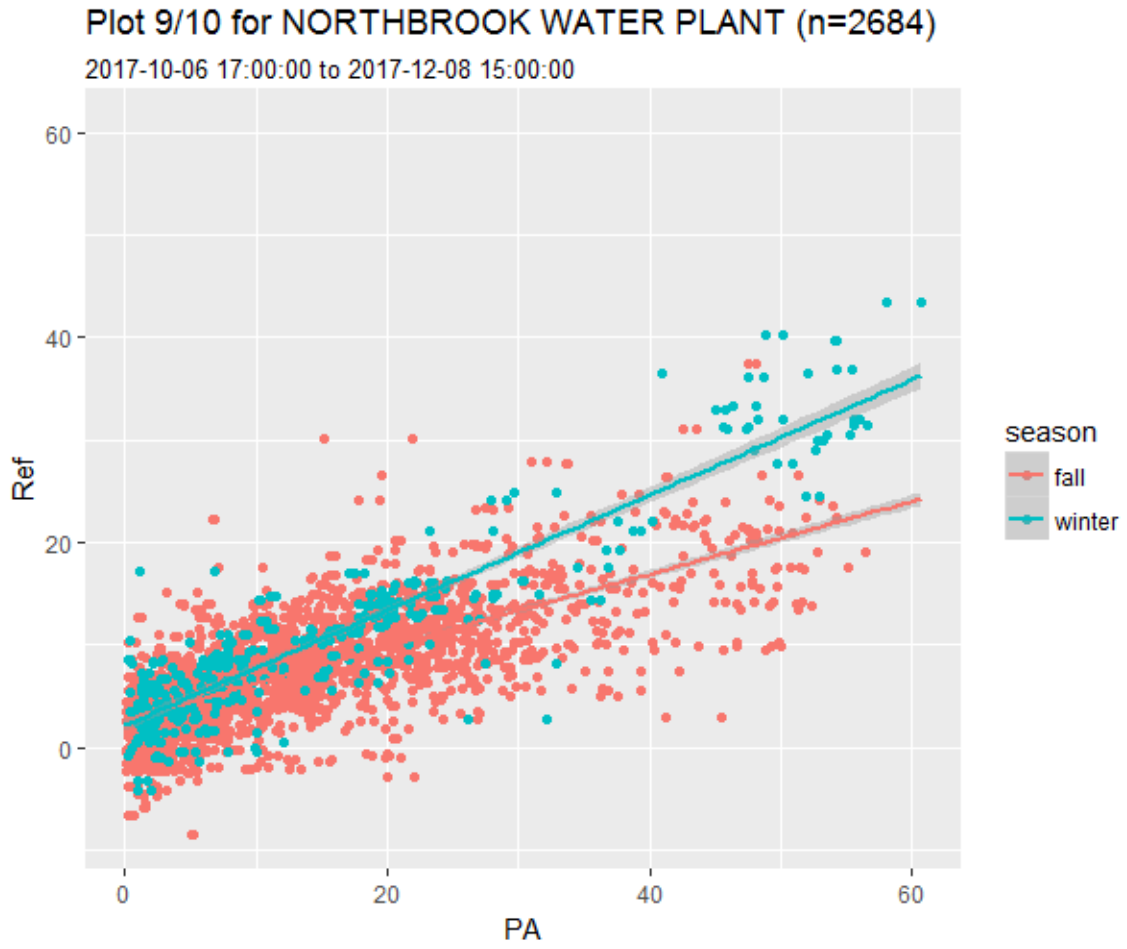

| | Ref = PA | Ref = PA + RH | Ref = PA + T + RH |
|---|---|---|---|
| $R^2$ | 0.68 (0.32 to 0.90) | 0.70 (0.32 to 0.92) | 0.70 (0.36 to 0.92) |
| RMSE | 3.7 (2.1 to 4.8) | 3.6 (2.1 to 4.7) | 3.6 (2.1 to 4.6) |

Calibrated data

# Seasonality

- **Most sites have little seasonal variability, some have a significant amount**



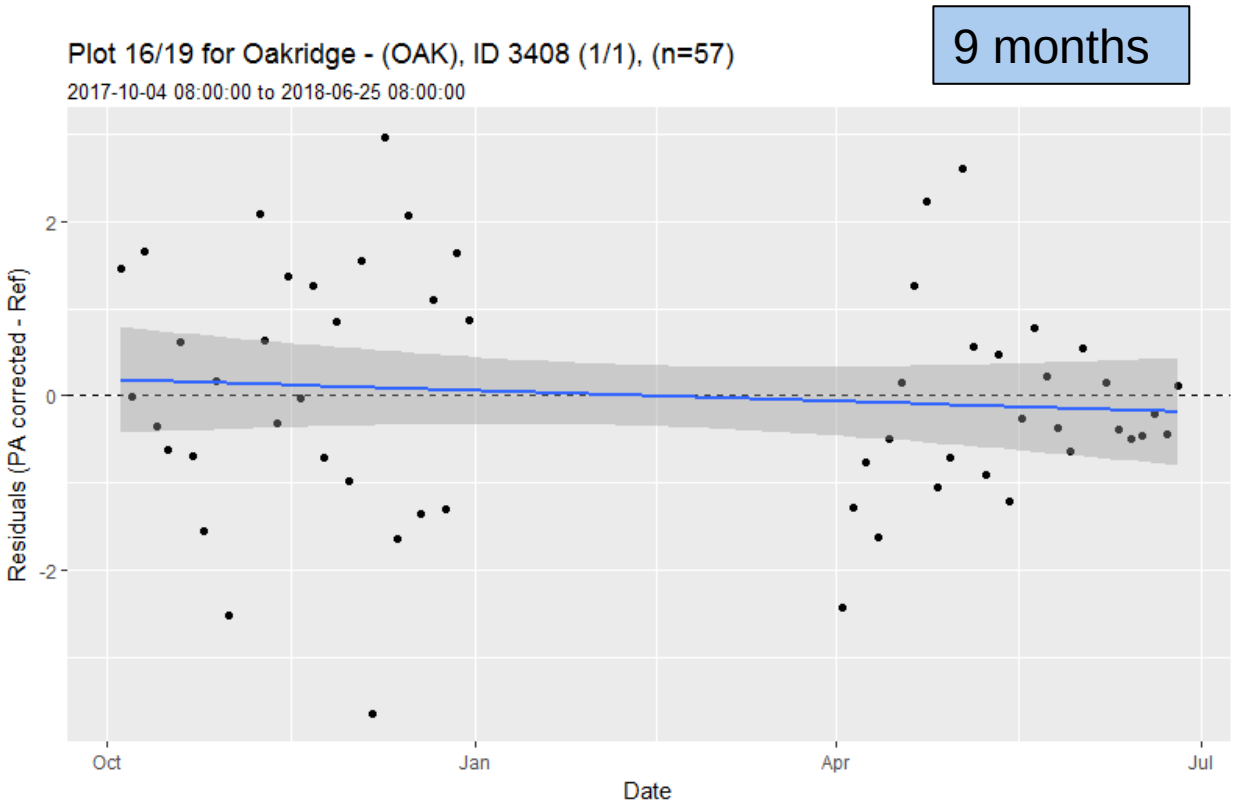

Uncalibrated data

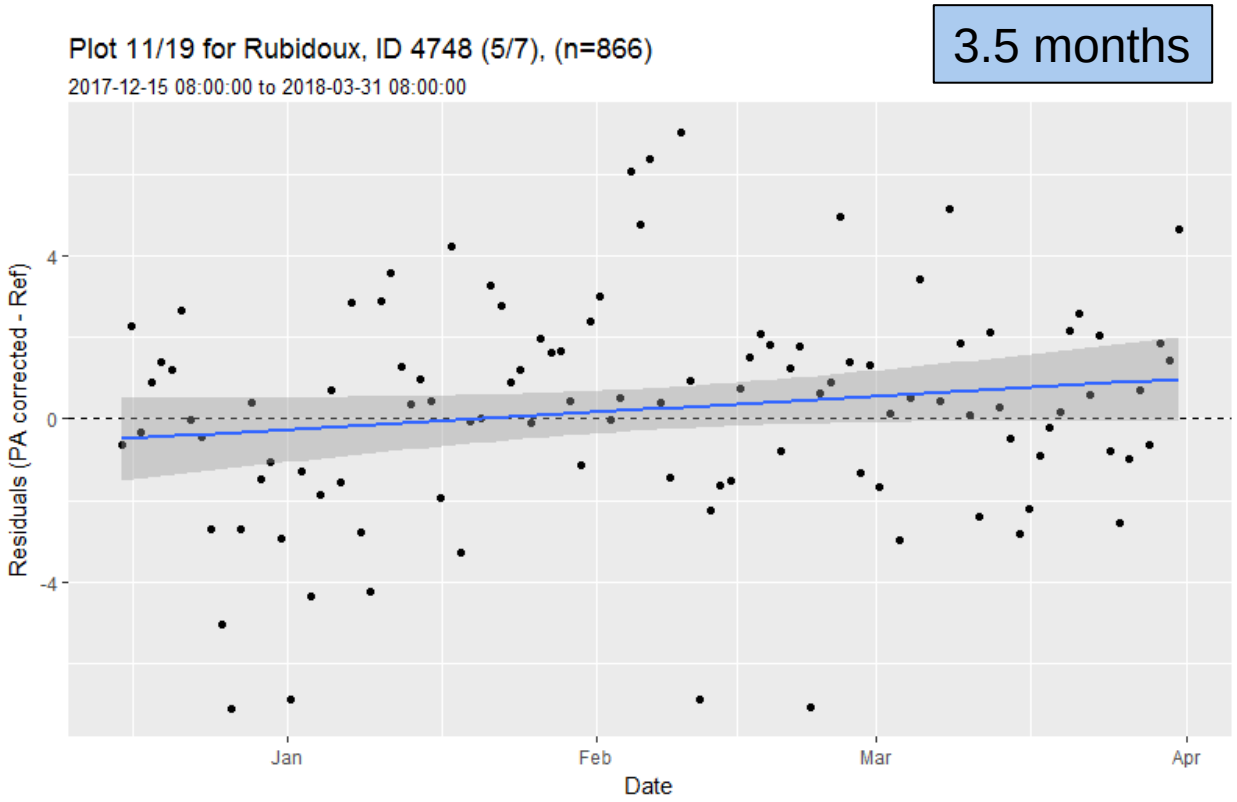
Daily data

Hourly data to help show difference

# Residuals

- **Some Purple Airs become biased low over time, other become biased high**




Calibrated data (Ref = PA + T + RH)

# Performance Criteria

## EPA Air Sensor Toolbox: Suggested Performance Goals for Air Sensors

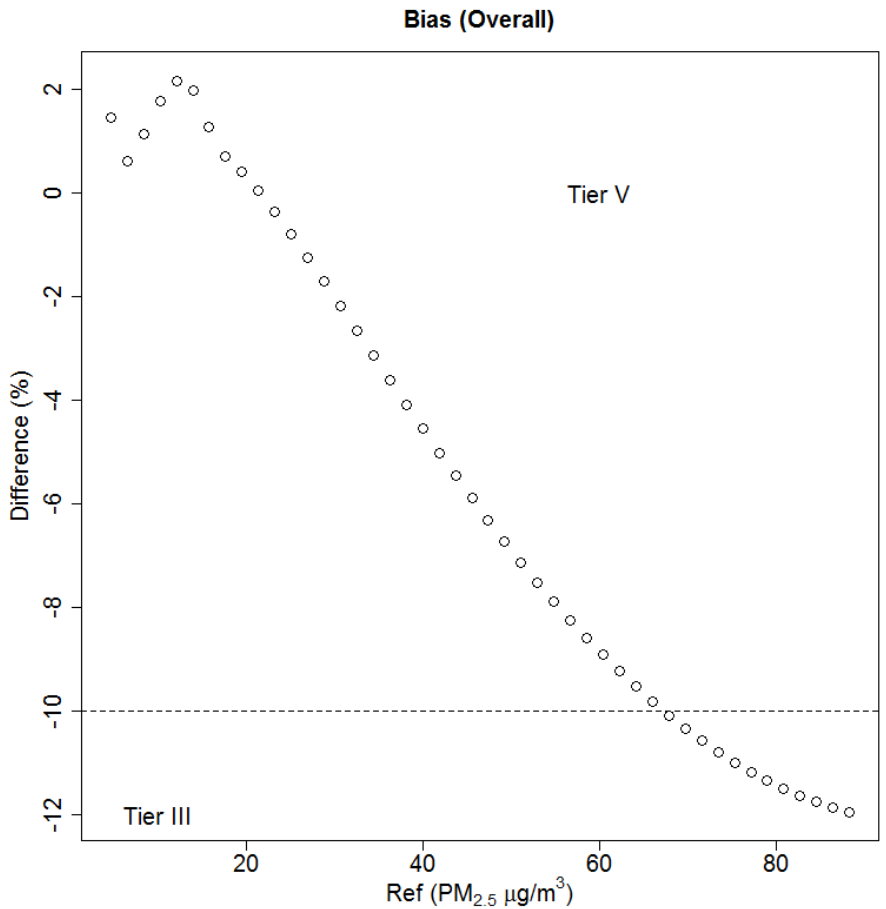
| Tier | Name | Precision/Bias % | Data Completeness % |
|---|---|---|---|
| I | Educational | 50 | 50 |
| II | Hotspot identification | 30 | 75 |
| IV | Personal monitoring | 30 | 80 |
| III | Supplemental monitoring* | 20 | 75 |
| V | Regulatory monitoring | 10 | 75 |

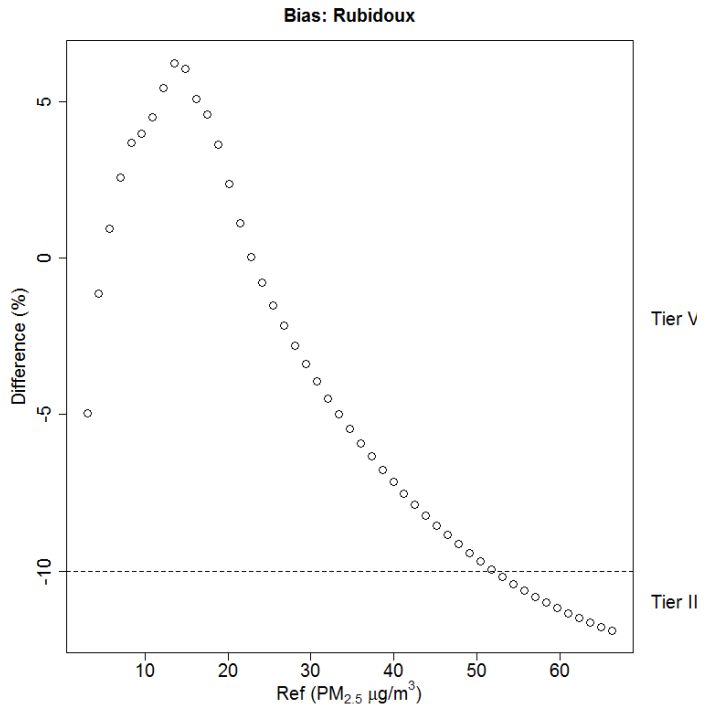
* "Supplemental monitoring might have value in potentially providing additional air quality data to **complement existing monitors**. To be useful in providing such complementary data, it must be of sufficient quality to ensure that the additional information is helping to **"fill in" monitoring gaps rather than making the situation less understood."**

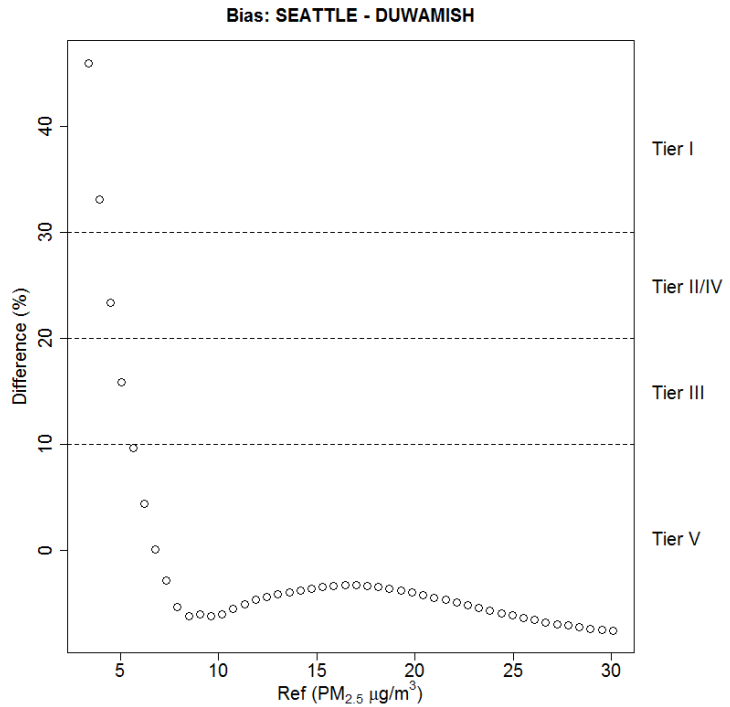
Link: https://www.epa.gov/air-sensor-toolbox/how-use-air-sensors-air-sensor-guidebook

# Bias

- **On average, Tier III or V Bias, 2% to -12%**
- **Most sites are biased low at the high end**



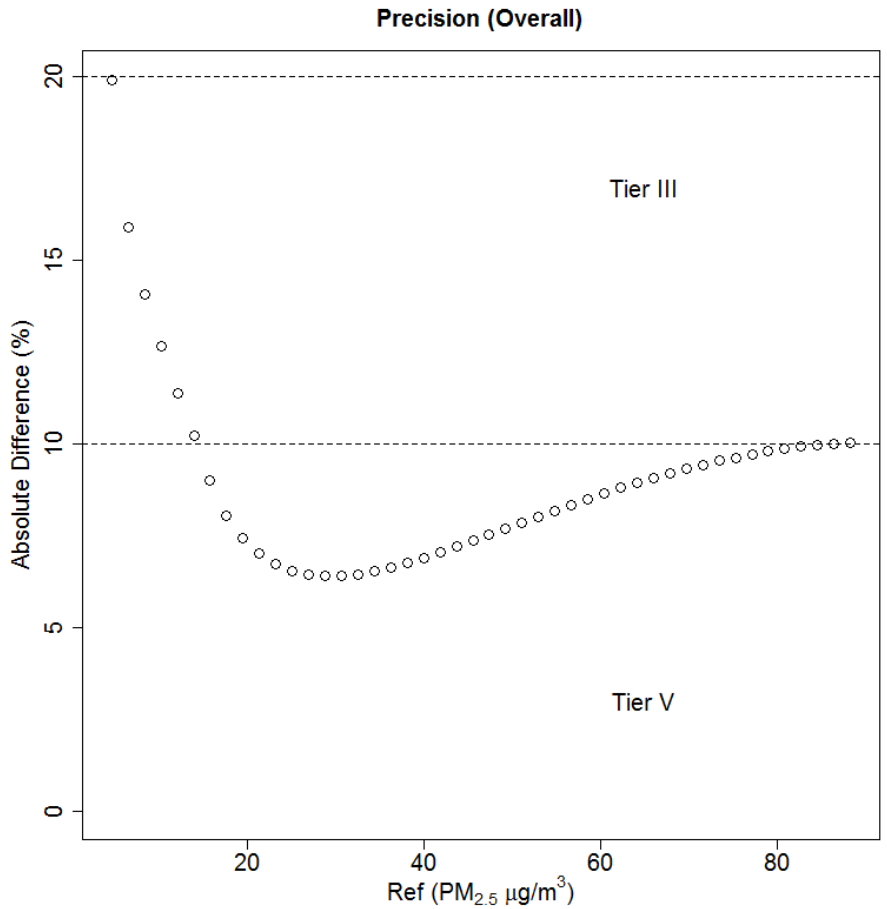



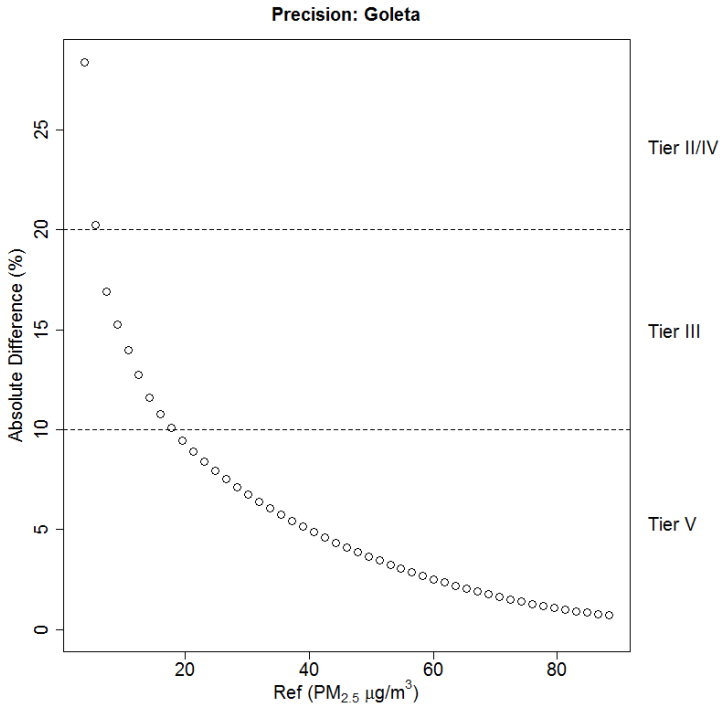



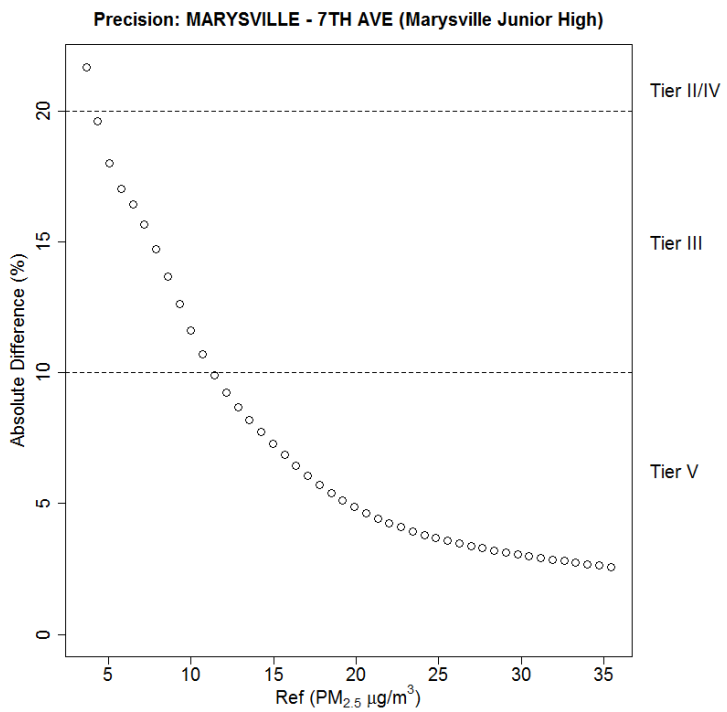

Calibrated data (Ref = PA + T + RH)

# Precision

- **On average, Tier III or V Precision, 7-15%**
- **Many sites get more precise the higher the concentration, some reach a max precision then get less precise**



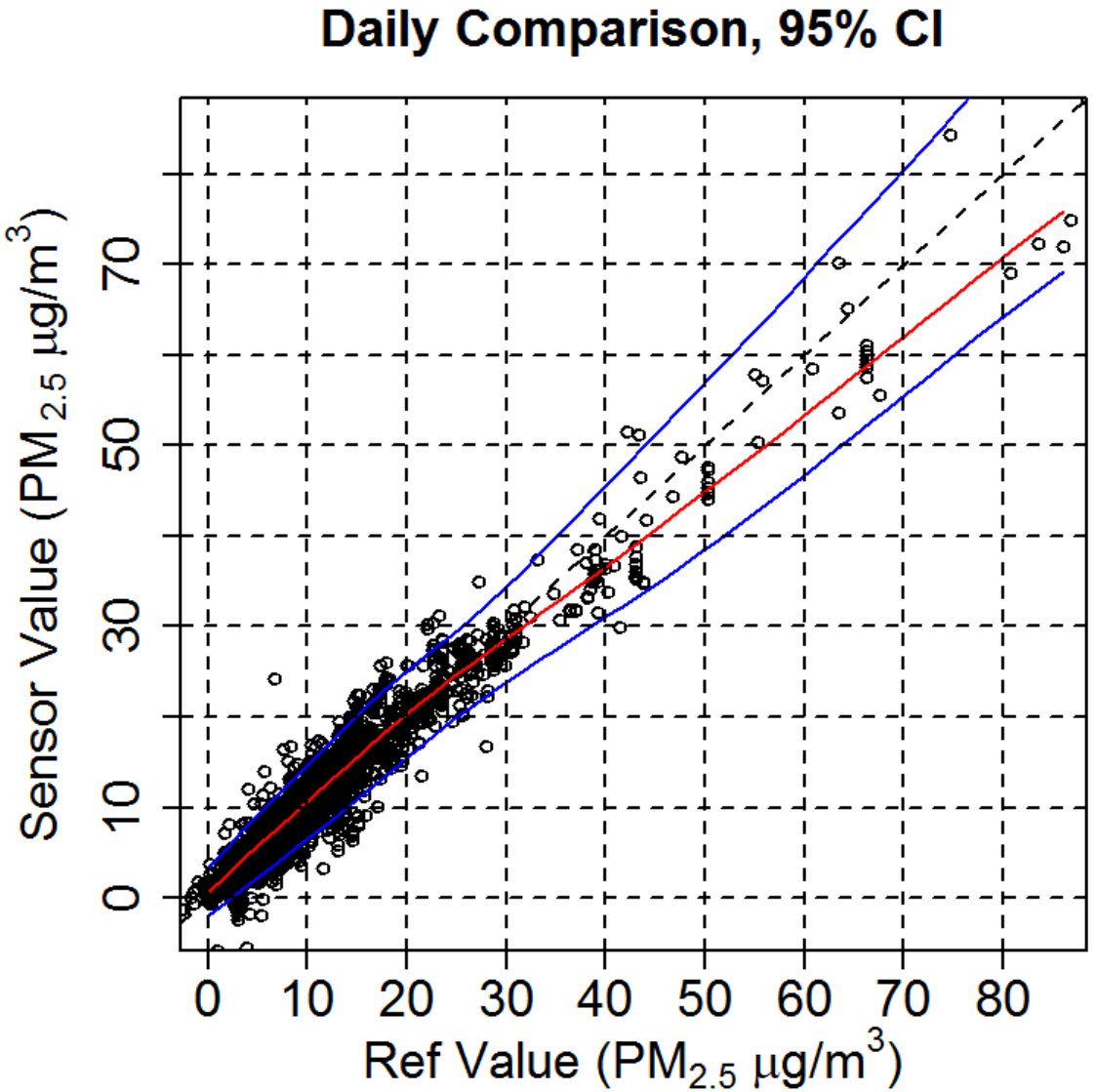




Calibrated data (Ref = PA + T + RH)

# Comparison

- **Biased low at high concentrations**
  - **Difference between the dashed black line and red smooth line**
- **When the Purple Air reads 30 µg/m³ the reference instrument is mostly likely reading 25-40 µg/m³**
- **When the Purple Air reads 50 µg/m³ the reference instrument is mostly likely reading 45-63 µg/m³**




Calibrated data (Ref = PA + T + RH)

# Summary

- **Useful tools for Supplemental Monitoring *after calibration***
- **Some spatial differences in intercept, slope, $R^2$**
  - **Difference in aerosol across the country/continent**
- **Susceptible to environmental effects**
  - **Calibration to reference monitor accounts for most, but not all, of this (see Extra Slides)**
- **No obvious decrease in performance over time in this dataset (max ~1 year). However, anecdotal evidence with other laser particle counters points to coating of the sensor over time and the necessity of regular cleaning (especially in dusty environments).**

# Thank You!

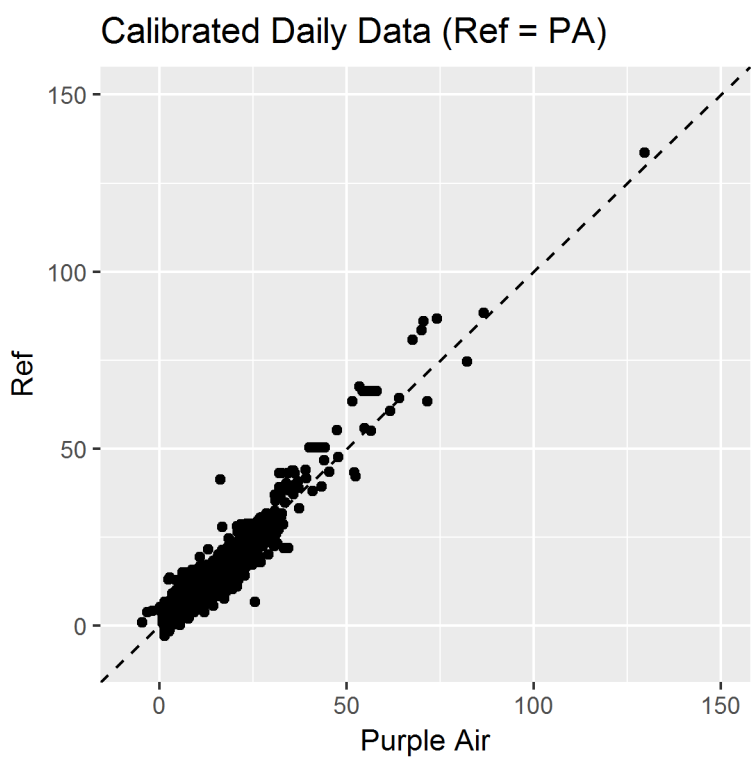
pscleanair.org
Puget Sound Clean Air Agency

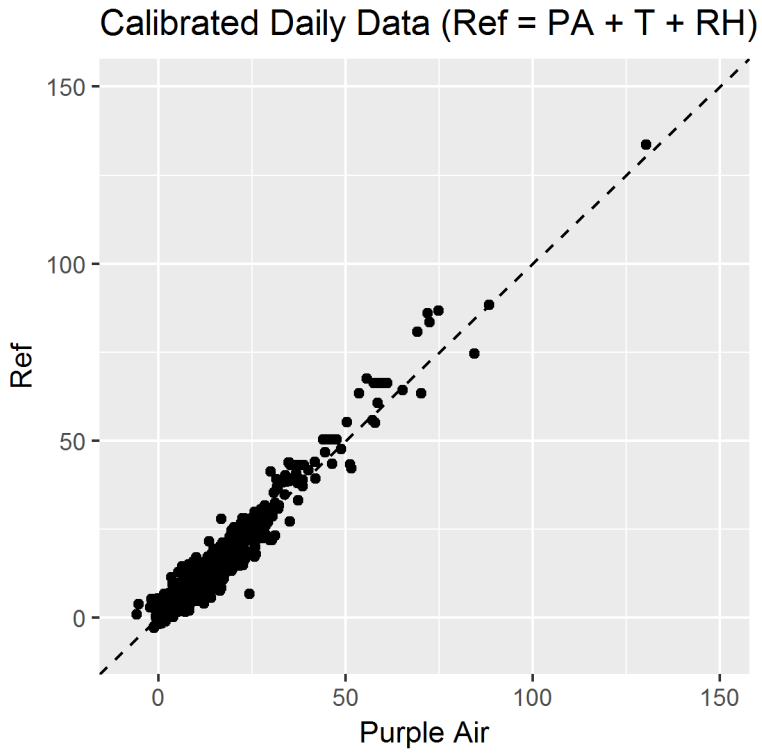
- **Do you have Purple Air data/analyses?  Want to talk about air sensors?**
  **Email: GraemeC@pscleanair.org**

# Extra Slides

pscleanair.org
Puget Sound Clean Air Agency

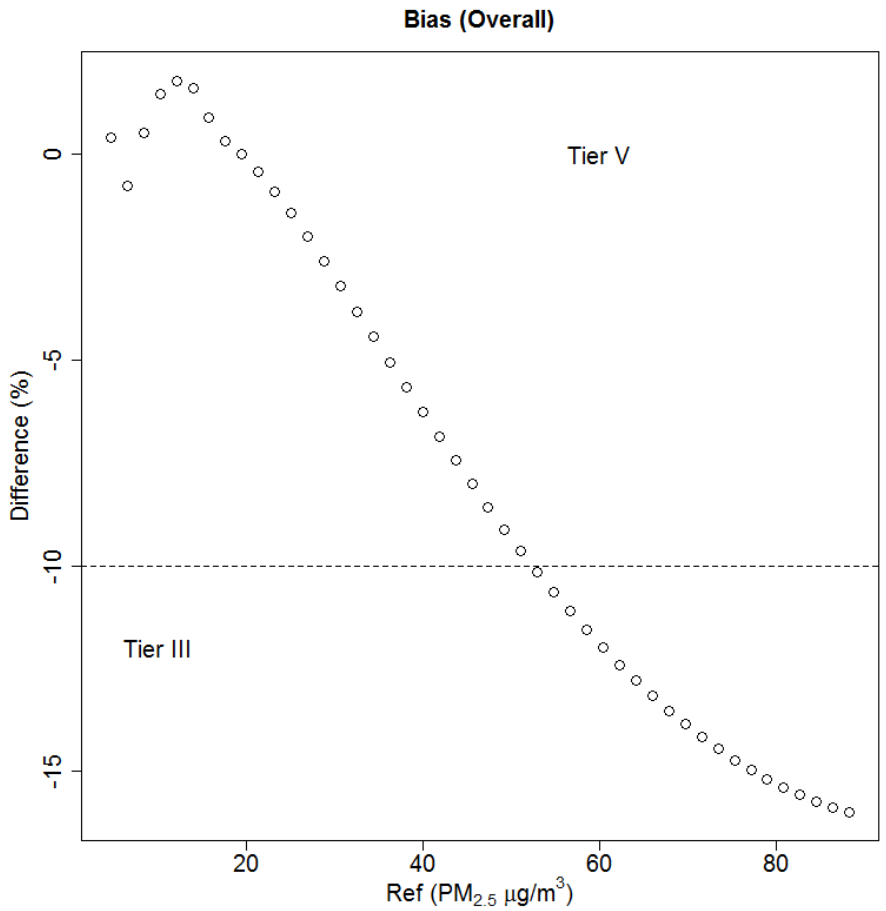
# Calibration models

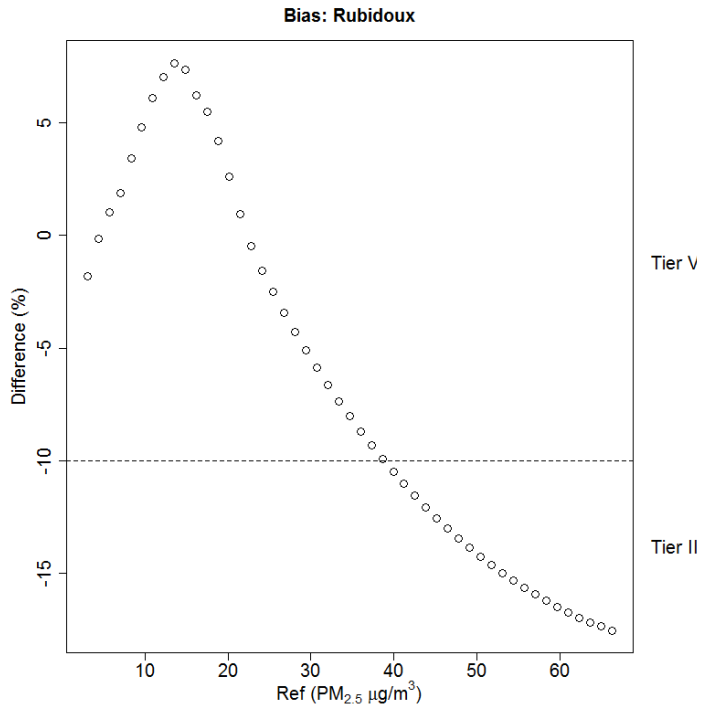


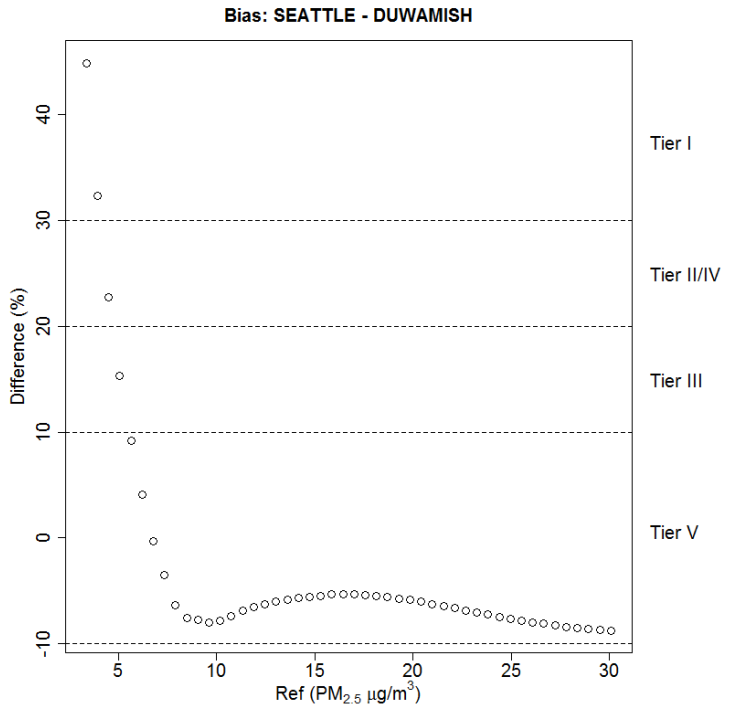



# Bias (Ref = $B_0$ + $B_1$*PA)

- **In general, Tier III or V Bias, 0% to -15%**
- **Most sites are biased low at the high end**



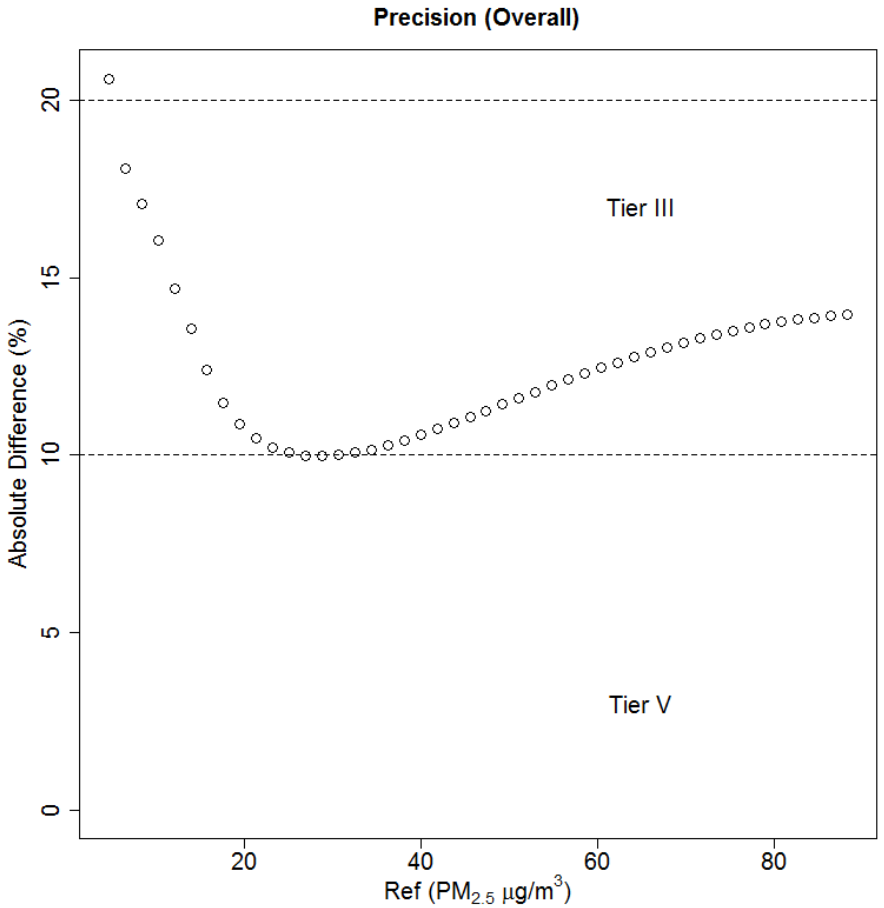



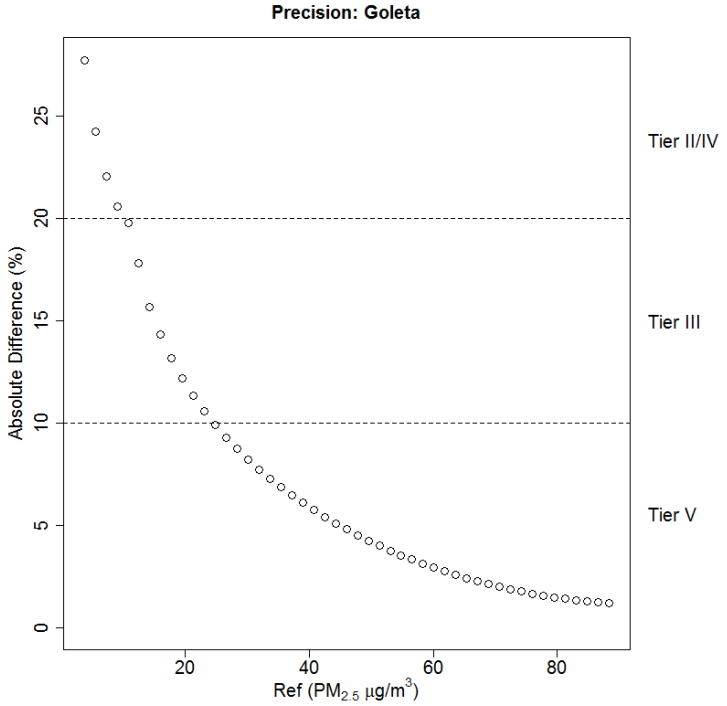



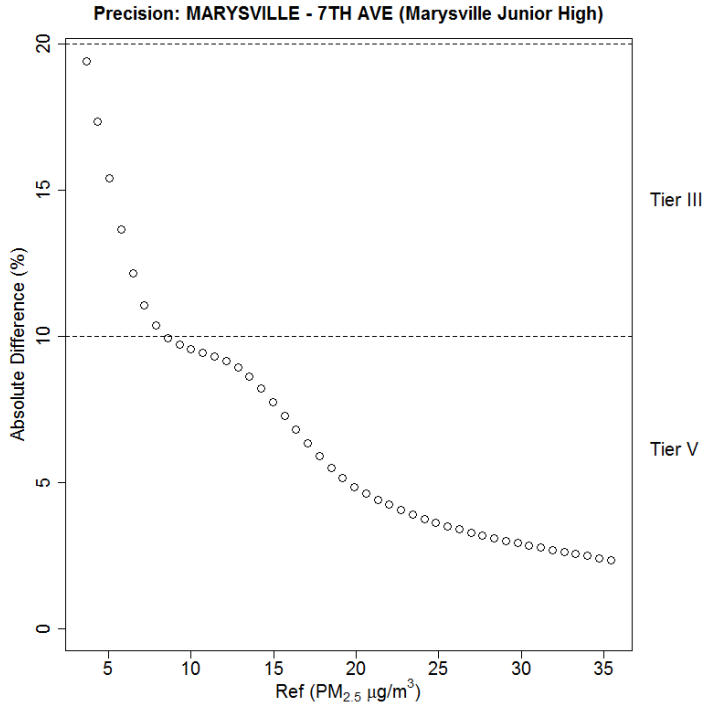

Calibrated data (Ref = PA)

# Precision (Ref = $B_0$ + $B_1$*PA)

- **In general, Tier III Precision, 10-15%**
- **Many sites get more precise the higher the concentration, some reach a max precision then get less precise**



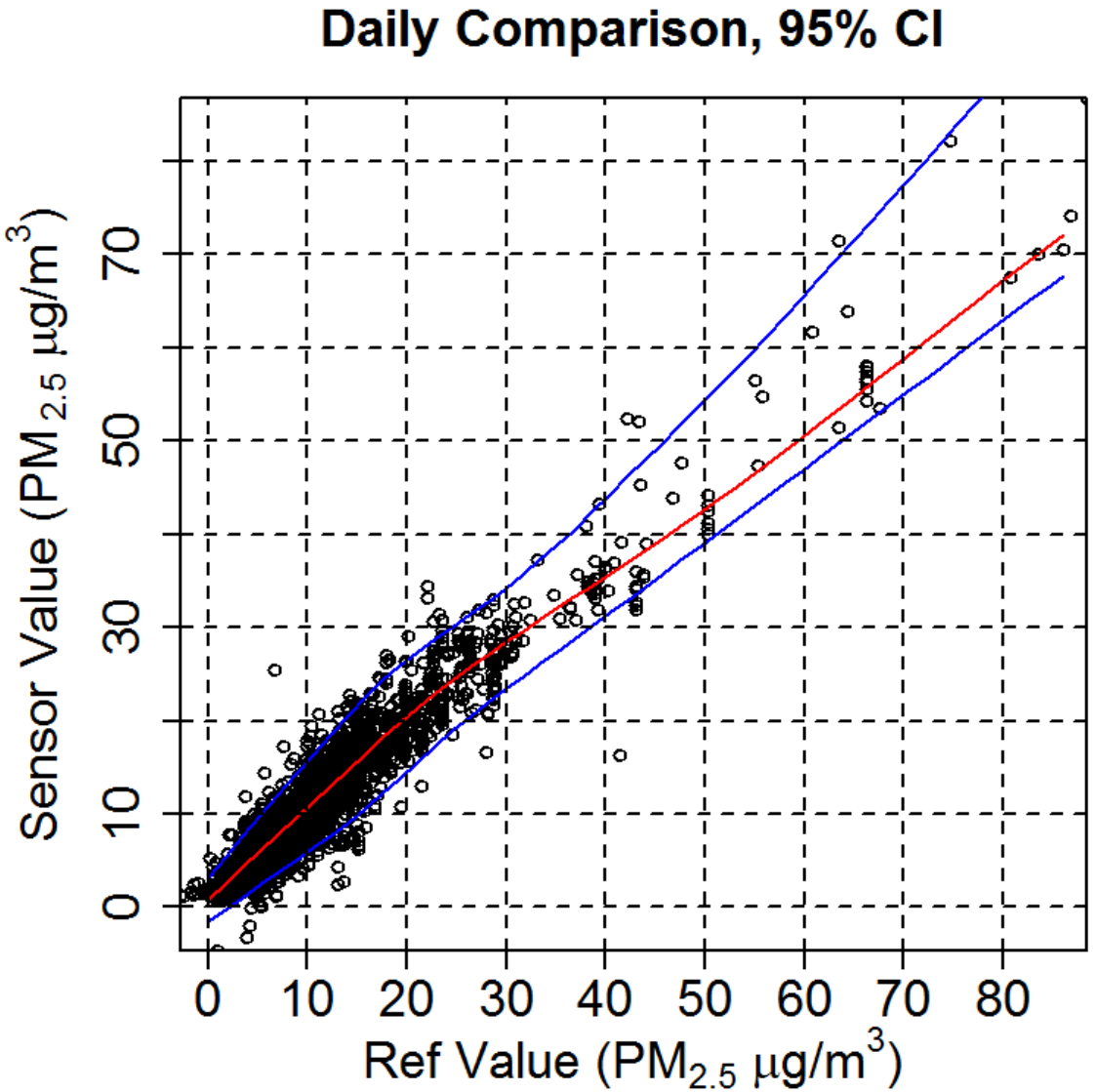




Calibrated data (Ref = PA)

# Comparison (Ref = $B_0$ + $B_1$*PA)

- Biased low at high concentrations
- When the Purple Air reads 30 µg/m³ the reference instrument is mostly likely reading 25-40 µg/m³
- When the Purple Air reads 50 µg/m³ the reference instrument is mostly likely reading 45-63 µg/m³




Calibrated data (Ref = PA)