Standard Operating Procedure for Air Quality Sensor Deployments

Prepared by

Hilary Hafner
Levi Stanton
Alan Chan

Sonoma Technology, Inc.
1450 N. McDowell Blvd., Suite 200
Petaluma, CA 94954-6515 USA
Ph 707.665.9900 | F 707.665.9800
sonomatech.com

STI-917067-6882

April 25, 2018

Contents

Figures

1. Overview

For air quality monitoring deployments requiring a very high level of quality of data, such as data to track National Ambient Air Quality Standards (NAAQS) in the United States, instruments that meet Federal Reference Method (FRM) or Federal Equivalent Method (FEM) requirements are used. This regulatory monitoring generally requires sophisticated instrumentation to meet measurement accuracy and precision requirements. In addition, an extensive set of procedures regarding calibration, maintenance, and audits is followed to ensure that data quality objectives are met. These procedures are documented by the U.S. Environmental Protection Agency (EPA) in the U.S. Code of Federal Regulations (40 CFR Parts 50, 53, 58) and the *Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II*.[1] Instrumentation, infrastructure, and staffing requirements are costly for these regulatory monitoring networks.

Low-cost sensors are providing new opportunities for measuring air quality for applications that require data of modest quality as compared to FRM or FEM data. These sensors are suitable for deployment in a wide range of environments because of their relatively small size and portability. While generally easier to use than FRM or FEM instruments, operation of low-cost sensors still requires procedures to ensure that data quality objectives are met. Sonoma Technology, Inc. (STI) is providing technical support for a sensor network pilot field study to be carried out in India. This document provides standard operating procedure for low-cost sensor deployments including network design and sensor siting, data validation with performance objectives, deployment, and sensor maintenance.

2. Monitoring Network Design

2.1 Components of a Monitoring Network Design

The first step in designing an air quality monitoring network is to **set monitoring objectives**. Monitoring objectives can be short-term, long-term, or both, depending on the scope of the project (e.g., number of instruments available to deploy). Objectives could include:

- Assess population exposure
- Determine highest concentrations expected to occur in an area
- Identify air pollution hot spots (i.e., locations with higher concentrations than surrounding areas)
- Observe spatial distribution of pollutants
- Track specific sources or source types (e.g., near-road sources)

[1] EPA-454/B-17-001, available at https://www3.epa.gov/ttnamti1/qalist.html.

- Characterize pollutant transport or background concentrations
- Track progress over time in reducing emissions
- Compare concentrations between cities

Secondly, based on network objectives, consider the **zone of representation** (also called the measurement scale) for the network:

- *Microscale:* 10 to 100 m (e.g., near-road sensors represent the environment of about 10 to 100 meters)
- *Middle-scale:* 100 to 500 m (e.g., this scale is typical of assessing pollution gradients from urban ground-level sources)
- *Neighborhood scale:* 500 m to 4 km (e.g., this scale is typical for understanding health impacts on communities)
- *Urban scale:* 4 to 100 km (e.g., this scale is suitable for understanding exposure to general populations and providing data for air quality forecasting and alerting)
- *Regional scale:* 100 to 1000 km (e.g., at this scale, measurements are useful for understanding pollutant transport into or out of an area or, if monitoring sites are far removed from sources, for quantifying background concentrations)

Some combination of sites at different measurement scales may be needed in a network design to meet all the objectives.

In addition to setting objectives and considering what measurement scale(s) to use to meet the objectives, consider the following questions:

- **Where to monitor?** A useful first step in determining where to monitor is to develop a regional description of, for example, topography, climate, population, major emission sources, size of the area to be assessed, and, if available, location of current air quality measurements (and observed concentrations). Then consider the purpose of the monitoring:
  - To assess *population exposure*, find locations of high population density. Sensors located where people live, work, and play are important for addressing exposure and protecting public health.
  - To *track emissions* from a particular source, develop emission reduction strategies, or track changes due to controls, place sensors close to specific source hot spots (locations with high pollutant concentrations).
  - To monitor the area of *maximum pollutant concentration*, place monitors downwind of maximum emissions.
  - To understand *pollutant gradients*, deploy a network with a high density of sensors.

  When deploying a few sensors in a region that lacks a robust monitoring network with FEM or FRM instrumentation, a first priority could be to characterize the typical concentrations experienced by the population. For assessing population exposure, we recommend following the EPA's siting guidance[2] for regulatory air quality monitoring. Unless specifically assessing sources, do not place sensors near potential major sources such as roadways or industrial facilities.

- **When to monitor?** For establishing concentrations in an area where there is little or no previous monitoring, it may be necessary to monitor for a year in order to cover all seasons. To investigate peak concentrations, make measurements in the season when concentrations are expected to be highest. Since most low-cost sensors provide continuous operations, all times of day and days of the week can be monitored. Ensure that sufficient sampling occurs during the time periods of interest to be able to address project objectives. This means, collecting enough data to be able to understand spatial or temporal differences with statistical confidence.
- **What other data are needed?** To enhance the data usefulness and to aid in interpretation, meteorological and other pollutant data are highly valuable. Collocation of a sensor with other air quality or meteorological instruments is desirable.

2.2 Checklist for Sensor Siting

Siting considerations for sensors include many of those important for siting traditional, regulatory-grade monitors; these include the measurement scale of interest, exposure to the environment, and security. The following table provides a checklist of siting considerations.

[2] U.S. Code of Federal Regulations, Part 58, Appendix D and E, may be useful for such deployments (www.law.cornell.edu/cfr/text/40/appendix-D_to_part_58, www.law.cornell.edu/cfr/text/40/appendix-E_to_part_58).

Action	Rationale
Use meteorological information to aid in network design.	A pre-project analysis of winds can help determine where to sample relative to the sources of interest. Use wind roses to identify the predominant wind direction. Measure in that direction downwind of the source. Depending on local meteorology and terrain, "downwind" may encompass multiple directions.
Allow proper airflow to the sensor inlet	Blocked or restricted airflow can result in measurement artifacts and/or biased data. Ideally, the inlet should be unobstructed for 360 degrees around it, but some situations, such as pole mounting, prevent this. In these situations, unobstructed airflow 270 degrees around the inlet is acceptable if the unobstructed direction is in the direction of the prevailing wind.
Place sensor at a height consistent with the monitoring objectives	For assessing population exposure, place the inlet at about 2 m above the nearest horizontal obstruction (i.e., 2 m above a roof top).
Ensure sensor system is placed sufficiently far from nearby structures	Channeling of air around a structure near the sensor system can impact concentration readings. • Be more than 20 m from trees or walls • Be above the parapet if on a rooftop • The distance from an obstacle should be at least twice the obstacle's height
Select sampling locations that meet study objectives	Be downwind of emission sources, in areas of high population density, etc., as appropriate.
Build collocation with reference instruments into the deployment plan	Collocation aids in quality assessment of the sensor and data interpretation. Data will be used to assess whether sensor readings need correction.
Collocate with reference instruments to understand sensor accuracy	Set up the sensor within about 2 m of the reference instrument. Locate the air intake of the sensor system within about 1 m of the air intake of the reference instrument, at the same height. Ensure the time standards are synchronized between the sensor and the instrument data collection systems.
Build in collocation among the sensors	Collocation aids in data interpretation and provides understanding of inter-comparability of sensors. Data will be used to make corrections between instruments if necessary.
Consider sensor system needs and security	Ensure the sensor has adequate power and space around it. Ensure sensors are secure and that technicians are safe. Assess communication needs (wifi access or cellular connections).
Take photos of each sensor deployment	Photos aid in data interpretation.
Record the following during initial deployment: • Height above ground • Latitude and longitude • Nearby sources and distance to them	This information aids in data interpretation.
Other site logistics	Ensure the sensor is accessible in all weather conditions, has a stable connection for transmitting data, and has a reliable power connection.

Key performance goals include sensitivity, accuracy, precision, data recovery, and sensor uptime.

- *Range* is the system's capability of measuring low and high concentrations.
- *Sensitivity* is the ability of the system to distinguish between different pollution levels.
- *Accuracy* is how well the system reproduces the same values as the reference method.
- *Precision* is how well concentrations measured by two or more systems compare. This can be addressed by collocating all instrumentation, with air intakes located as close to one another as possible.
- *Data recovery* is percent of hourly (or other interval) valid data values that were collected, divided by the total number of expected data intervals in the date range.
- *Sensor uptime* is the percent of time the sensor systems were in operation, divided by the total expected amount of time. This is closely related to the data recovery statistic.

In addition to these performance goals, *climate susceptibility* is an important consideration, as some sensors are susceptible to extreme temperature or humidity. Some manufacturers have carefully characterized this effect and provide correction algorithms; however, these algorithms should be verified in the location where sampling will occur.

3. Data Validation

Data validation is the process of determining the quality and validity of observations. The purpose of data validation is to verify that data meet quality control criteria, and to detect and label any data values that may not represent the actual physical and chemical conditions at the sampling station before the data are used in analysis. For ongoing quality control during the sensor deployment, it is vital to:

- Follow guidelines on sensor use
- Calibrate sensors before, during, and at the end of sampling
- Keep track of changes and activities near the sampling location, and review data regularly throughout the sensor deployment

3.1 Performance Metrics

For study success, it is important to set, and hopefully meet, performance metrics. Metrics for sensitivity, accuracy, precision, data recovery, sensor uptime, climate susceptibility, and repair time are provided in the following table.

Assessment	Criteria
Range	1-999 µg/m^{3}for $PM_{2.5}$ and PM_{10}
Sensitivity	±1 µg/m^3 for $PM_{2.5}$ and PM_{10}
Accuracy	Sensor vs. Reference R^2 > 0.70 Exact reproduction of reference data is not needed for many applications, but comparable response should be expected.
Precision	Sensor vs. Sensor R^2 > 0.85 Precision is necessary so that intersite comparisons can be made. Sensor-to-sensor bias can be accounted for by collocation.
Data recovery	The percent of valid data available after collection is greater than 75% completeness (by hour and hours in day)
Sensor uptime	Sensor system is operational; collecting data greater than 90% is desirable.
Climate susceptibility	The sensor measurements should not be impacted by the environmental conditions at the sampling location.
Repair time	Repairs are performed less than a day after notification of issue.

3.2 Validation Steps

3.2.1 Validation Prior to Sampling

Take the following data validation steps before sampling begins:

- Gather and document metadata such as concentration units, time stamp, detection limits, and sampling interval.
  - Determine whether the reported time in the raw data files corresponds to the beginning of an averaging period or the end of an averaging period. For example, for begin-time, data collected by the instrument during the 8:00-9:00 a.m. hour are reported as collected at 8:00 a.m. local standard time (LST). For end-time, the same data are reported as 9:00 a.m. LST. End-time reporting makes the data look more current to the public and is consistent with data reporting on the U.S. EPA AirNow website (airnow.gov).
  - If comparing to other sensors or regulatory instruments, ensure the time stamps are consistent.
  - Ensure units are consistent with common practices. For PM, use µg/m^3.

- Document instrument configurations; verify that they are valid and align with data collection objectives.
- Identify quality objectives—such as data completeness, accuracy, or spatial coverage—that align with study goals.
- Set up methods to create averages of the raw data. Typical averaging periods include 1-minute (from 1-second data), 15-minute, 1-hour, and 24-hour. All averages should include 75% data completeness (i.e., ≥45 seconds for 1-minute, ≥11 minutes for 15-minute, etc.).

3.2.2 Validation During Operations

The process of data validation includes visually reviewing the data. When looking at collected data during operations, consider the following:

- Look at data immediately at the beginning of the study and often thereafter to correct problems before the project is over. Check that all sensors are reporting data, that concentrations change from one sampling interval to the next, and that the lowest concentrations (baseline) stay consistent over time and do not show drift.
- Inspect time series of all sensor data in the study. Look at how the concentrations vary among sites, with the weather, and with the typical activities around the sensors. Windy conditions can lead to higher dust levels in the air. Sensors deployed in locations with different emission sources may be expected to exhibit very different concentrations. However, if collocated sensors are showing different concentrations, this indicates that something may be wrong with one of the sensors.
- For collocated sensors, inspect scatter plots. Compute the linear regression statistics including correlation coefficient and slope. A higher correlation coefficient is better (with an R^2 of 1 being perfect agreement). A slope closer to 1 is preferred, but if there is a linear relationship between the measurements of two sensors, a slope offset from one can be corrected for.
- Periodically, generate and review summary statistics (e.g., minimum, maximum, median, average, percent completeness). As more data are collected, statistics can be produced for time of day, day of week, and season to explore how the data vary over these time periods.
- Review data. For example:
  - Check for several minutes or more of exactly the same (non-zero) concentration. Because the ambient atmosphere is very dynamic, repeated values of the same concentration are not reasonable. The sensor may be malfunctioning.
  - For high PM_{10} concentrations, look at whether the wind speed increased. Dust can be stirred up and transported in high winds.
  - Track temperatures and relative humidity along with the PM concentrations. Some PM sensors using optical methods, for example, are affected by high humidity. The sensor readings may become very erratic or noisy, and this can be seen in a time series. Very high or very low temperatures may also affect sensor output, also resulting in noisy data that can be seen in a time series. If FEM and sensor data are

collocated, plot concentration data from the FEM and sensor, with the data color-coded by temperature or relative humidity ranges, and look for deviations from a 1:1 relationship that may be explained by the weather variables. A scatter plot of concentration versus relative humidity is also useful.

- In a network of sensors, if one sensor reading over a period of time is very different from another, it is important to investigate possible reasons. If one of the sensors is located near a source, then higher concentrations might be expected.

3.2.3 Validation After the Measurement Campaign Is Complete

- Assess collocation results to understand any differences between sensors (bias) or differences from the reference instrument (accuracy). If sensors are collocated at the beginning and end of the campaign, sensor drift can also be assessed. Bias is assessed using linear regression between sensors to provide correlation and slope information. Similar information can be obtained by comparing each sensor to the reference instrument.
- Look at minimum and maximum concentrations to find anomalous values. Anomalous values can be very high or very low concentrations relative to the overall data set, relative to nearby measurements, or relative to expectations.
- Look for negative values. Some sensors may report slightly negative values, which is permissible, but values less than -10 $\mu g/m^3$ should be treated as suspect, investigated, and possibly invalidated. Negative values may indicate that a sensor is failing.
- Inspect data for expected changes in concentration corresponding with the following parameters:
  - *Wind direction.* A pollution rose is useful for this assessment. A scatter plot or polar plot of wind direction and concentration can also be used.
  - *Other weather conditions such as wind speed, temperature, relative humidity, or rainfall.* Time series and scatter plots of concentrations with these variables are useful. Look for trends in or relationships between concentration and the weather parameters. Higher wind speeds may be associated with dustier conditions—and thus, higher PM concentrations—than low wind speeds. Rainfall may suppress dust and be associated with low concentrations. Cold temperatures may be associated with increased use of home heating devices, such as wood burning, that can lead to higher concentrations. High humidity may cause noisy or erratic data.
  - *Time of day or daytime versus nighttime.* Box whisker and time series plots are useful to investigate how concentrations vary with time. Some emission sources may be most active during the day (e.g., heavy duty trucks) and others during the night (e.g., home heating). Differences in concentrations over these time periods may be related to the different emission sources.
  - *Day-of-week patterns.* As with time-of-day patterns, there can be significant variation in emissions with the day of the week, such as reduced commuter traffic on non-work days. Pollutant concentrations may be lower on non-work days if traffic is a dominant source. Box whisker plots are useful for this assessment.

- *Season and holidays*. There may be different emissions activities on holidays, such as fireworks displays, that can be observed in the ambient data. Time series plots are useful for this assessment. For seasonal variations, both emissions activities and weather influence pollutant concentrations. Statistical summaries or box whisker plots can help elucidate the differences.

- Inspect baseline concentrations (the lowest values) for stability over time. It is useful to create a time series of concentrations looking at a month or more of data in one graphic, focusing on the lowest values. If the lowest values seem to change over time (trending higher or lower, or showing sudden shifts), drift may be occurring. If the sensor is compared to a reference instrument at the beginning, and during or after deployment, a scatter plot of the sensor and reference instrument is useful to identify drift (**Figure 1**). Drift can be corrected for in the final data set if the sensor is collocated with a reference instrument or if the sensor is collocated with a reference for short periods during the deployment.

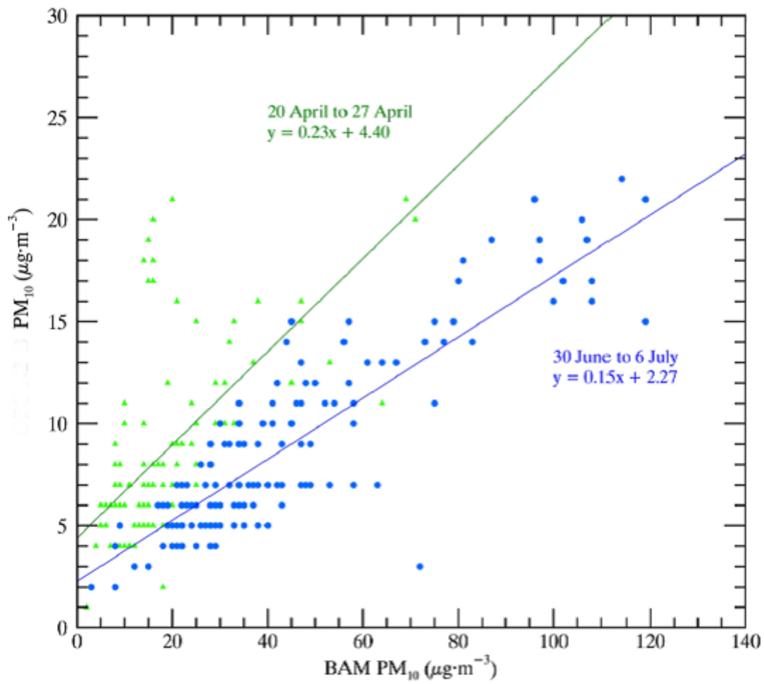

Figure 1. PM_{10} concentrations measured by a sensor showed significant drift (65% change in slope between periods) over the study period (green is the beginning of the study, blue is the end) relative to the reference instrument.

- Identify and investigate unusual values. An unusual value may be a very high concentration at a sensor when other sensors in the network did not show high concentrations. If there are no indications that the sensor was malfunctioning, investigate what potential sources may be near that sensor using local knowledge, satellite maps, and emission inventory information.

Inspect weather information as well to assess whether conditions at that particular site were different from conditions at other sites and may have influenced the measurements.

- Look for values that normally follow a predictable spatial or temporal pattern. In many cities, there are certain hours during which workers commute to their jobs, and often the pollutant concentrations are highest during those commute hours. When this pattern is typically observed, it is useful to investigate when the pattern is not observed. Temporal and spatial patterns can be observed using time series plots.
- Inspect data for odd patterns, systematic decreases in concentration over time (which may indicate a sensor problem), and other unusual features. Examples include high concentrations occurring repeatedly at a specific time of day (there may be an emission source near the sensor), frequent baseline shifts that may indicate a sensor problem, and excessive noise in the data that may indicate a sensor problem, The data would be valid if an emission source is impacting the site, but the sensor may need to be moved if the siting is no longer meeting criteria. If sensor problems are found, the data are invalidated.
- Look for data with several consecutive equal data values (stuck or repeated values) above the sensor detection limit, which may indicate a problem with the sensor. In ambient air, concentrations fluctuate and are rarely constant. This check can be visual (using a time series plot) or can be automated (using a database search). The repeated values are invalidated. Some sensors may reach a maximum limit and could have repeated high values indicating the sensor has exceeded its measurement limit. These data may be deemed suspect rather than invalid because the sensor is indicating high concentrations (although of unknown values), and this information may be useful.

3.2.4 Identifying Bad Data

When data are identified as erroneous through the investigations discussed in Section 3.2.3, it is best practice to assign a quality control (QC) flag to the data that indicates they are invalid or suspect. If data are being displayed to the public, these invalid or suspect data may be omitted from display. Similarly, invalid data are omitted from further analysis. Some indications of invalid data are:

- Values “sticking” or reporting the same value for an unrealistic length of time such as more than four hourly values. This check should only be applied to data above the detection limit of the sensor.
- Negative data. Slightly negative values are sometimes experienced due to calibration, but exceptionally negative values (i.e., less than -10 µg/m^3 for PM_{10}) should be invalidated.
- Extremely high values. If investigation indicates no clear reason for one sensor to show very high concentrations that are not measured by any other nearby sensors, the high concentrations may be suspect or invalid.
- Unsatisfactory correlation, e.g., R^2 less than 0.60, with collocated or nearby instruments. The acceptable correlation coefficient will depend on project objectives.
- Data points around data gaps, especially if they look unusual. Data collected just before a sensor begins having a problem and potentially immediately after it comes back online need

review. There may be a stabilization period during which data are not representative of real conditions.

- Abnormally noisy data (see **Figure 2**).
- Changes in the data rate. (i.e., if you expect 5-second data but get it every 15 seconds instead, investigation of the cause is needed).

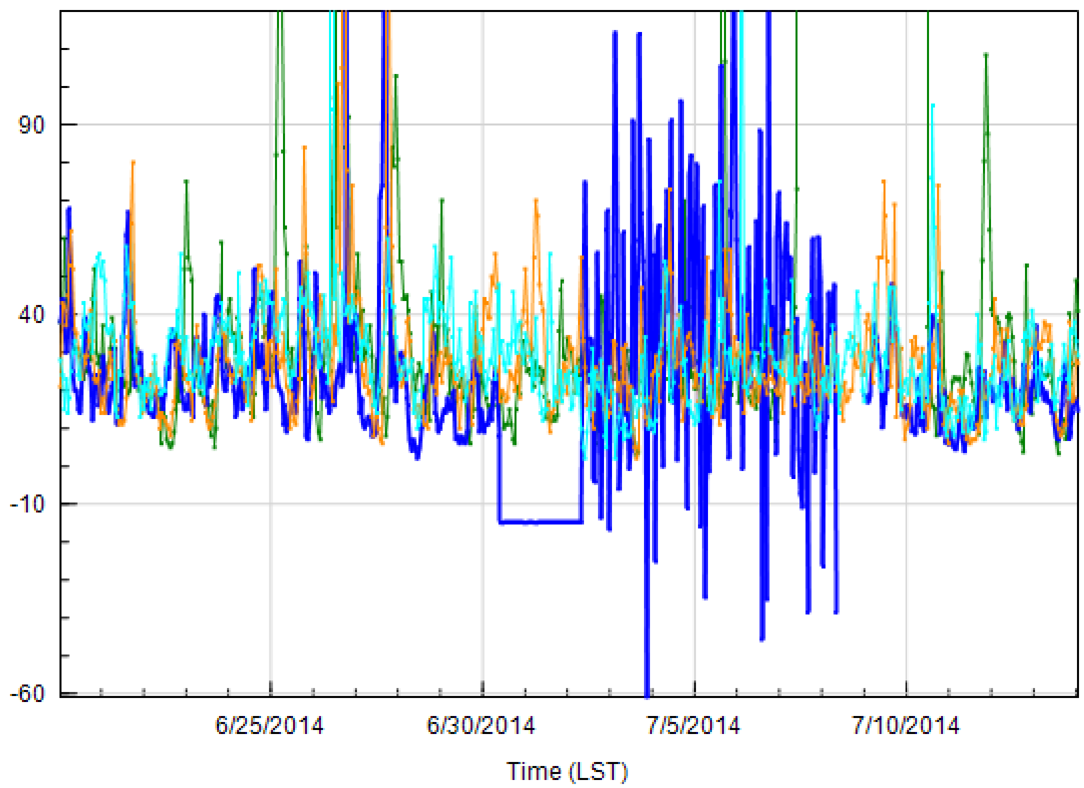

Figure 2. PM_{10} concentrations at several locations in a network. Note the flat line, negative concentrations from one sensor (in blue) followed by wild concentration swings and noisy data. The sensor was found to be malfunctioning.

4. Deployment

4.1 Pre-Deployment Checklist

- Charge batteries or test line power
- Confirm sensor system is operating
- Confirm measurements are reasonable (ideally, in comparison with a reference instrument)
- Confirm communications with server

4.2 Deployment Checklist

- Mount system in a secure area without any major obstructions. If collocating with another monitor, ensure that inlets are at the same height, and that they are between 1 and 4 meters apart.
- Attach any antennas, such as for cellular, Wi-Fi, or GPS
- If using solar power, align the panel to the south in the Northern Hemisphere, or to the north in the Southern Hemisphere. Consult this online solar angle calculator for the optimal angle for the location and time of year http://solarelectricityhandbook.com/solar-angle-calculator.html
- Ensure all electrical connections are watertight
- Power on system
    - Confirm system is operating as expected
    - Confirm measurements are reasonable (ideally, in comparison with a reference instrument)
    - Confirm communications with server
- Document installation with photographs
- Document installation location, height above ground level, and any nearby sources

5. Maintenance

5.1 Overview

Maintaining sensor performance over the project period may require some action, depending on the sensor system in use and the problems encountered. For example, degradation in sensor accuracy could require the following actions:

- Cleaning a part of the system (such as an inlet or fan blades); a fan with a large amount of dust built up on it can spin more slowly, which may change the flow rate through the system
- Replacing a battery; when battery voltage falls below the specifications of the system, the sensor performance can be adversely impacted

In another example, incomplete data can be caused by these problems that should be addressed:

- Poor communications; some sensor systems do not log data locally, which means that any loss in communications causes a loss in data
- Variable quality of power

It may occasionally be necessary to update sensor firmware, as manufacturers sometimes develop algorithms that improve sensor performance or make updates to improve the system's security, resilience, or stability. Whenever a firmware upgrade is performed, the version numbers should be tracked. Additionally, it is best to back up the version of firmware being used currently before making any updates, in case the new firmware has a bug or is incompatible with the hardware. Finally, simply turning the system off and on can solve many issues, but the root cause should be investigated to prevent future problems.

5.2 Routine Monitoring Checklist

The following maintenance should be performed at least once every three months, during every site visit, or when issues have been encountered, in order to ensure proper sensor operation. It may be necessary to check sensors more frequently when pollution concentrations are consistently high.

- Carefully clean any spiderwebs, bugs, or other debris from the inlet
- Ensure all electrical connections are secure
- Ensure that there are no new obstructions that could impact measurements or block solar radiation (i.e., new construction or tree growth)
- Note any new nearby emission sources
- Document site condition, especially if there are any new obstructions or sources identified
- Clean off solar panel if applicable
- Ensure the solar panel is angled optimally for the current season (see http://solarelectricityhandbook.com/solar-angle-calculator.html)
- Routinely back up data

6. Calibration Techniques

Low-cost sensors, like reference-grade instrumentation, require routine adjustments to their measurements to ensure accurate and consistent data. Calibrations can account for baseline drift and change in sensor response over time, but they do not necessarily account for biases caused by other pollutants or changing environmental conditions. There are a number of methods used to perform "calibrations" or adjustments of sensor data, so that the sensors produce more accurate data when compared to a reference instrument. Approaches that involve physical collocation are described below, but machine learning methodologies are becoming more widespread.

6.1 Overview

By collocating a sensor with a reference instrument, one can develop a correction factor that can be applied to the sensor data. There are several approaches to collocation, but at the minimum, a sensor should be collocated before it is ever deployed. **Figure 3** shows a collocation of three $PM_{2.5}$ sensors with reference instrumentation; it is important that there not be any nearby obstructions, and that the inlet heights are similar. Other approaches include performing subsequent collocations once every season by removing the sensor from the field (but this approach cannot account for local variables that may be influencing measurements), collocating a "golden sensor" (that is normally kept in collocation with a reference instrument) with another sensor in situ in the field, or by outfitting a mobile vehicle with reference grade instrumentation and visiting individual sensor nodes for a period of time.

Figure 3. Three $PM_{2.5}$ sensors collocated with FRM and FEM instrumentation in Sacramento, CA, USA.

6.2 Collocation Analysis

Once the collocation has been completed, a scatter plot can be made that compares the reference instrument to the sensor. It can be useful on the plots to color individual measurements by other parameters (such as meteorological parameters) to investigate the impact of different variables (such as temperature and humidity). For instance, Figure 4 demonstrates that the sensor had a varying response when the dew point was different.

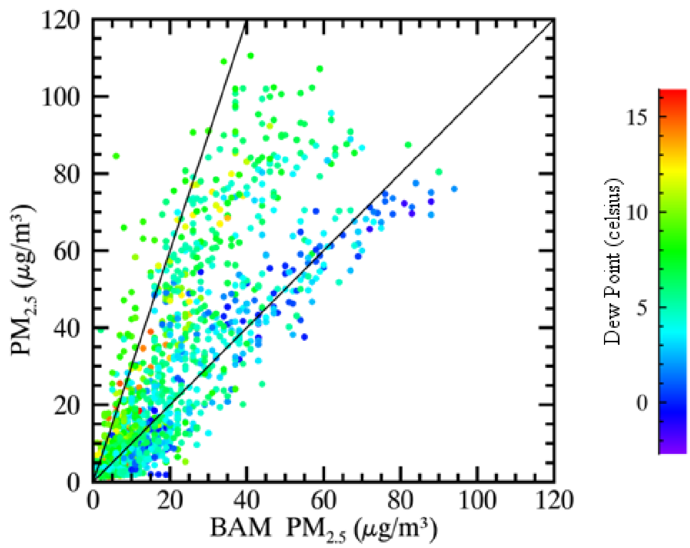

Figure 4. Sensor data compared to an FEM instrument and colored by dew point

A linear regression can be performed on the data, such as is shown in Figure 5. If the R^2 meets the data quality objectives of the application, the linear regression equation can be applied to the data to adjust the sensor measurement to better align with the reference measurement.

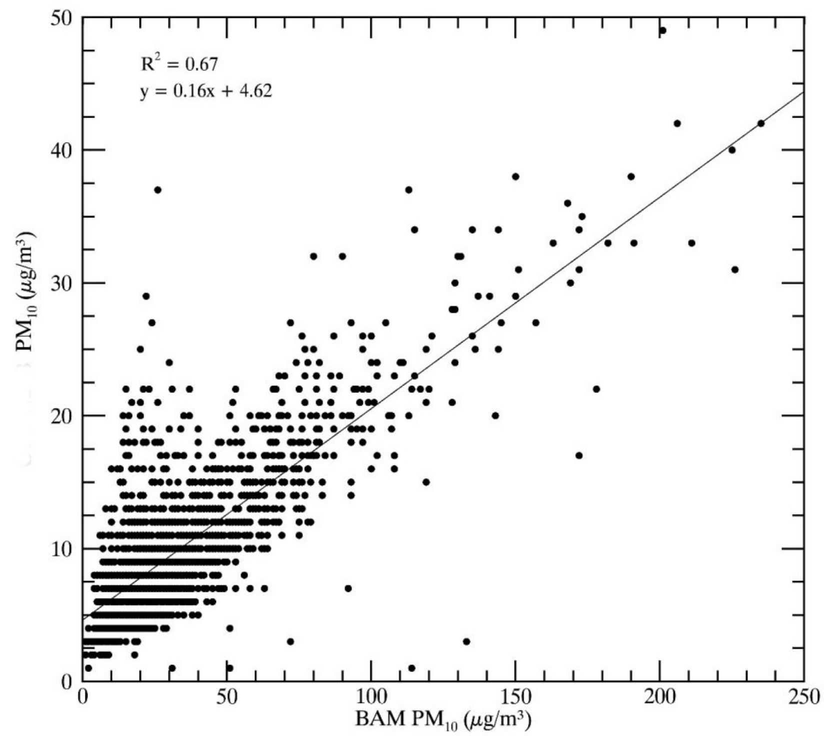

Figure 5. Sensor data compared to an FEM instrument with a linear regression equation and R^2.

7. Data Reporting

When reporting data to other (non-public) users such as other governmental organizations, universities, etc., or when using data from other entities, the following information should be known and shared:

- **Study design/deployment.** Describe the purpose of the measurements, where and for how long sensors were deployed, the siting criteria used, and photos of the deployment. Also describe whether regulatory grade monitors were present in the area and whether those data are available.
- **Sensor selection.** Describe what sensors were used; how the sensors were evaluated prior to the study and the results of the evaluation; sensor make, model, serial number, and when acquired; the specificity of the sensor to the pollutant of interest; sensor detection limits; the sampling frequency and duration; how the sensor was packaged (e.g., used as part of a complete sensing package or a custom-built system); how the sensors were powered; and how the data were stored and transmitted.

- **Sensor calibration.** Describe how sensors were calibrated prior to, during, and at the end of deployment and whether those data are available. Describe sensors' collocation with each other and/or with a regulatory grade monitor.
- **Sensor deployment.** Describe training provided to the sensor operators; describe maintenance performed on the sensors; provide a summary of sensor performance, data recovery, and uptime.
- **Supplemental data.** Describe other data collected as part of the deployment, such as meteorological data, including what was measured and where it was in relation to the sensors. At a minimum, temperature and relative humidity are needed. Describe whether any co-pollutants or potential interferents were measured. Provide other documentation such as operating procedures, study plan, etc. Other information that may be useful includes traffic counts, patterns, and fleet information; source inventories, locations; and exceptional events (fire, dust storm, etc.) that may have occurred during sampling.
- **Data format.** Data should be provided in a well-documented and consistent format. Data values should be consistently labeled (pollutant name, units, date/time stamp, time zone). Describe how missing data, data gaps, and data below detection or beyond the sampler's upper limits were handled. Describe whether/how data quality flags were used and provide any definitions. Describe data transformations (such as units, signal processing) that were made. If data are from a mobile platform, describe how location was reported and how the sensor system and location information were synched. If there is software associated with the sensor, document the version used and whether any updates were made during the measurement campaign.
- **Data quality.** Describe how often data were reviewed during the project and how data quality decisions were made (e.g., what rules were applied). Describe how the sensors compared with each other and with a regulatory monitor. Document data recovery/completeness and how it was calculated. Document bias, precision, and accuracy and how the parameters were calculated. If possible, describe how uncertainty was defined and calculated.

Thorough documentation can lead to increased confidence in—or at least understanding of—the data collected by a network of sensors.