Integrated Use of Low-Cost Sensors to Strengthen Air Quality Management in Indian Cities

Vital Strategies

Bloomberg Philanthropies

Acknowledgements

Coauthors and Main Contributors

Meenakshi Kushwaha, ILK Labs, India
Sumi Mehta, Vital Strategies, United States
Pooja Arora, The Energy Research Institute, India
Timothy Dye, TD Environmental Services
Thomas Matte, Vital Strategies, United States

External Reviewers

Joshua Apte, University of Texas, USA
Solomon Teffera, SolTeff Environmental Consulting, USA
Sachchida Tripathi, IIT Kanpur, India

The authors and contributors to this guide are grateful to government and civil society stakeholders and partners who provided invaluable input and insight during a two-day virtual technical exchange held in December 2020. The full list of participants is provided in the Appendices. Those interactions were used to frame the content of this compendium.

This publication was made possible with financial support from Bloomberg Philanthropies.

Vital Strategies. Integrated Use of Low-Cost Sensors to Strengthen Air Quality Management in Indian Cities. Meenakshi Kushwaha, Sumi Mehta, Pooja Arora, Timothy Dye, Thomas Matte. New York, NY, 2022. Available from: https://www.vitalstrategies.org/resources/integrated-use-of-low-cost-sensors-to-strengthen-air-quality-management-in-indian-cities/


SVP

Contents

Acronyms

API	Application Programming Interface
AQ	Air Quality
AQM	Air Quality Management
AQSPEC	Air Quality Sensor Performance Evaluation Center
ASTM	American Society for Testing and Materials
BAM	Beta Attenuated Monitor
CAAQMS	Continuous Ambient Air Quality Monitoring Stations
CSV	Comma-Separated Values
EPA	Environmental Protection Agency
GBD	Global Burden of Disease Study
CPCB	Central Pollution Control Board
LCS	Low-Cost Sensor
MLR	Multiple Linear Regression
MOEFCC	Ministry of Environment, Forest and Climate Change
NCAP	National Clean Air Programme
NOx	Nitrogen Oxides
QA/QC	Quality Assurance and Quality Control
PCB	Pollution Control Board
PM	Particulate Matter
RMSE	Root Mean Square Error
RH	Relative Humidity
SLR	Simple Linear Regression
SOW	Statement of Work
UNEP	United Nations Environment Programme
WHO	World Health Organization

Background and Overview

Air pollution continues to be the deadliest global environmental health risk, causing nearly 5 million deaths each year, mainly from exposure to fine particles ($PM_{2.5}$). The burden of air pollution is greatest and increasing in countries like India with rapid economic development and urbanisation, along with a proliferation of emissions from industry, electric power generation, and motorised transport. In countries with limited or no air quality regulation, this causes steep increases in harmful pollution. Compounding the air pollution challenge in many of these countries is the persistence of pre-industrial pollution sources such as burning of household solid fuels, crop waste and forests for land clearing as well as open trash burning.

For many city governments in many low- and middle-income countries, the complexity and cost of understanding and controlling air pollution have been barriers to initiating or sustaining effective clean air action. A new approach in air quality management that combines conventional solutions with innovations in monitoring, assessment, data use, and organisation can accelerate clean air action, especially in cities with presently limited technical capacity.

India's National Clean Air Programme (NCAP) has indicated that alternative technologies for real-time air quality monitoring such as low-cost sensors (LCS) should be explored and promoted to fill current gaps in air quality monitoring. At the same time, given the novelty of LCS, their broader application to inform air quality management and clean air policies requires careful consideration. In December 2020, more than 50 international and national experts were brought together in a virtual technical exchange to share international insight, best practices, and examples of sensor use to more fully explore their capabilities, limitations, technical specifications and how they should be selected and deployed. Here, we integrate these learnings into the broader monitoring framework provided in Accelerating City Progress on Clean Air: Innovation and Action Guide[1], a playbook for fast-tracking proven approaches and innovations to improve air quality.

In brief, this compendium provides pragmatic guidance for low-cost sensor (LCS) use based on applied research and field experience in India and internationally, including identifying monitoring goals and questions for which LCS may be fit to purpose, technical specifications, key quality assurance/quality control issues (i.e. validation and reliability of precise, reproducible, and consistent results). A key goal is to support state pollution control boards and city governments in their efforts to develop bid documents to procure the comprehensive set of services needed for LCS campaign planning, development, deployment, analysis, integration with complementary air quality data, and results communication and management.

The focus of this compendium is on use of LCS for measurement of $PM_{2.5}$ mass concentration, the pollutant for which current LCS technology is most suitable and the most important indicator of air quality for public health.

This Compendium Addresses the Following Topics:

- Using LCS to address the specific issues of NCAP cities and to collect actionable data suitable for filling gaps in official, reference monitor networks;
- Measuring and producing data useful for permanent reference monitor placement;

Q: Why is measuring and controlling fine particulate matter ($PM_{2.5}$) a priority?

A: It is the most harmful pollutant to health. $PM_{2.5}$ is an indicator of a pollution mixture that causes the most serious illness and death in India and globally. In addtion to being a proven cause of serious illness and death from cardiovascular and respiratory diseases, cancer, and diabetes, which are included in the global burden of disease estimates, $PM_{2.5}$ also impacts birth outcomes and child health, potentially impairing well-being and productivity across the life span.

- Evaluating geographically targeted pollution abatement measures (e.g. vehicle-free zones, neighborhood-scale adoption of clean household energy);

- Identifying hot spots and fence line monitoring of facilities to identify potential violations and local impacts of episodic sources of pollution (e.g. solid waste dumps, industrial sites), and addressing emergency situations (e.g. major structural fires and/or wildfires);
- Understanding important limitations of current LCS, including: distinguishing local PM sources from high, regional background pollution, inability to measure PM composition for source apportionment and inability to reliably measure gaseous pollutants important for air quality management (AQM);
- Guidance to inform tender or vendor scope of work for LCS services including objectives, planning, spatial allocation and site selection, device selection, calibration, deployment, maintenance, data acquisition, transfer, quality assurance and quality control, and analysis;
- Best practices for making data accessible and meaningful for a range of technical and non-technical stakeholders, necessary to increase accountability and support for clean air actions.

Simple decision tools and templates are provided to help shape the development of tenders and guide the evaluation of applications received. Two case studies describing the application of LCS in NCAP cities are also included.

Application of Low-Cost Sensors (LCS) within Comprehensive Air Quality Management Programs and Integrated, Robust Monitoring Systems

Air pollution monitoring, including low-cost sensors and other innovative methods, is only one aspect of air quality management. In addition to monitoring data, estimates on pollutant emissions from major sources—already available for a number of NCAP cities[2]—can and should be used to inform clean air action plans to reduce harmful emissions, even as robust monitoring networks are being established and enhanced.

A combination of monitoring approaches can inform a robust air quality management program, support the needs of local, regional and national air quality management, and provide data for research and public information[3]. There are a range of technologies for monitoring air quality from the ground and from satellites in space. On the ground, the most accurate are reference-grade instruments that produce high-quality data. These are, however, expensive to purchase and operate. While LCS offer the promise of less expensive monitoring, accuracy varies greatly depending on the sensing technology and pollutant. Satellite remote sensing has been a critical source of global air quality information4, especially in locations without any ground monitoring. This approach can also fill data gaps in areas with extensive ground monitoring[5-8]. When combined with chemical transport models and available surface measurements to relate atmospheric column measurements to surface concentrations, satellite-based estimates for $PM_{2.5}$ and some gaseous pollutants are available at varying spatial resolution (~1-10 km), but do not provide highly time-resolved, local, neighborhood- scale measurements.

For India, where many sizable cities lack reference-grade surface monitors, no single technology can provide comprehensive monitoring. Instead, a combination of technologies, including LCS, can provide cost-effective solutions to establishing air quality monitoring and enhancing it in a phased, sustainable manner that leads to an integrated system (Figure 1). LCS campaigns and networks can fit within a broader air quality monitoring system using integrated data from satellite remote sensing methods, traditional and advanced surface monitors, and periodic monitoring with one or more high spatial resolution approaches: land use regression, mobile monitoring or LCS campaigns and networks. Such a system can assess air pollution variation at different spatial and temporal resolutions, inform additional reference monitor placements and incorporate new innovations over time. The cost of deploying and operating such a system will vary greatly depending on local circumstances, but for filling monitoring gaps in India, such a hybrid system would likely be much less costly than a conventional regulatory network with many reference monitors[9]. It should be noted that such a hybrid system, in addition to combining different monitoring technologies as shown, can integrate credible data from monitoring conducted by different governmental and nongovernmental entities.

CURRENT LEVEL OF DATA AND CAPACITY

INCREASING MONITORING DEVICES AND NETWORK DENSITY

Levels	Available Data
1 **Limited or none** No sustained official reference $PM_{2.5}$ monitoring in place	• Satellite-based estimates • Non-official reference $PM_{2.5}$ monitoring • Land-use regression, low-cost sensor, or mobile monitoring studies
2 **Basic monitoring to support initial actions** At least one official reference $PM_{2.5}$ monitor in place with ongoing data collection and use (at a minimum for public information)	**Phase 1+** • One or more fixed reference $PM_{2.5}$ monitors
3 **Comprehensive monitoring for sustained actions** A network of several reference $PM_{2.5}$ monitors with at least one advanced monitoring station collecting $PM_{2.5}$ sampled for chemical composition and to measure gaseous pollutants. Data have been used in policy development.	**Phase 2+** • Advanced surface particle monitoring station • One or more reference monitoring stations for gaseous pollutants
4 **Advanced integrated system** Phase 3 monitoring plus periodic high-spatial—resolution monitoring	**Phase 3+** • Periodic land-use regression models or mobile monitoring campaigns • Low-cost sensor network

Figure 2. Application of Low-Cost Sensor (LCS) Networks by Current Level of Data and Capacity

Priority Questions

- LCS Is the air quality hazardous to health in the urban/metro area?
- Why is it important to have reliable official monitoring?
- LCS Where should initial reference monitors be placed?

- What is the baseline $PM_{2.5}$ level and trend as clean air actions are launched?
- LCS Where should reference grade monitors be located?
- Is local air quality compliant with local standards?
- When are short-term air pollution episodes occurring?
- LCS What do we know about local patterns in air quality and/or exposure to air pollution?

- What are important sources of urban/metro area air pollution?
- Are control measures improving air quality in the urban/metro area?
- LCS Where should additional reference grade monitors be located?
- LCS Are local sources (e.g. trash burning, biomass energy use) impacting nearby $PM_{2.5}$ levels?
- LCS Are there exposure hot spots?

- LCS What are the local neighborhood hot spots and sources?
- Are control measures improving local-neighborhood air quality?
- LCS Are geographically-focused air pollution control measures improving neighborhood air quality?

LCS = May be addressed through application of LCS networks

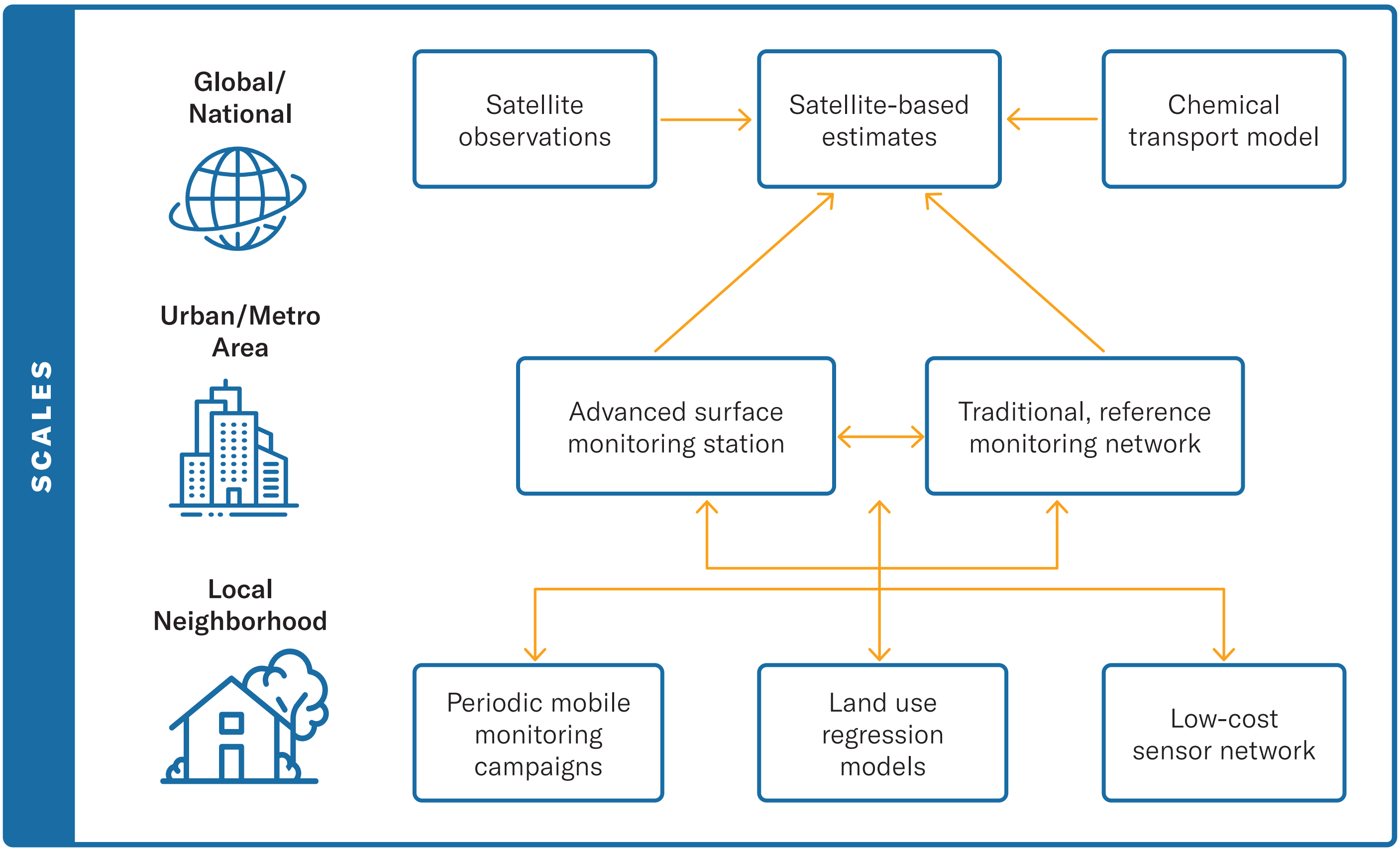


Figure 1. Integrated Air Quality Monitoring System

Possible LCS Applications

The preceding section provides an overview of how LCS can complement other conventional and innovative monitoring methods as part of a comprehensive monitoring system. Here we consider different applications of LCS, depending on the current level of monitoring capacity and the key questions to be addressed (Figure 2). The pragmatic application of LCS should start with an assessment of the current air quality monitoring data and capacity in a city or other jurisdiction where LCS are currently being considered, determining key questions about air quality relevant to informing or evaluating decisions. In addition, any available, credible monitoring and data on pollution levels should be inventoried to help assess how LCS data might fill gaps and be integrated with other air monitoring data.

For example, in NCAP cities with limited or no monitoring capacity, LCS networks may be used to ground truth or verify that $PM_{2.5}$ levels estimated from satellite-based remote sensing methods show a substantial exceedance of health-based standards across a city and surrounding region. The networks may also provide initial actionable data to inform the building of an integrated monitoring system of complementary approaches, such as by scoping sites for initial or additional permanent reference monitors. In cities with at least some established reference monitors, LCS offer a way to complement the existing monitoring, fill in spatial gaps in the monitoring network, identify hot spots, and measure the effectiveness of geographically targeted clean air actions. Potential applications of LCS depend on the current level of data and capacity in the city and key data gaps and questions about air quality. See Figure 2.

Project Goals and Design

This compendium focuses on four practical applications of LCS to inform air quality management in NCAP cities. These applications include:

1. **SITING** Characterizing the spatial patterns of PM2.5 levels in the absence of reference grade monitors to inform the siting of an initial reference grade monitor or additional monitor(s) to fill gaps in a sparse network.

Practical Questions to Be Considered Before Planning an LCS Monitoring Project

- What is the current status and capacity of official reference air pollution monitoring (Figure 2)?
- Given available official and unofficial data on air pollution levels, what are key data gaps and questions about air pollution levels that LCS might help answer?
- What are the limitations of currently available LCS devices and advantages and disadvantages compared to other monitoring approaches? Are LCS fit-to-purpose for answering key questions?
- What supporting data (e.g. on PM2.5 sources, emissions and their spatial distribution) are available or need to be assembled to help plan an LCS project and add value to LCS data for informing or evaluating air pollution control?

2. **MAPPING AIR QUALITY** Characterizing spatial and temporal exposure gradients across the city to evaluate emissions data and forecasts, provide exposure estimates for health studies, and inform expansion of the existing network of reference grade monitors.
3. **IDENTIFYING HOTSPOTS** Identifying areas with higher concentrations of air pollution due to proximity to sources, to prioritize local action or to raise public awareness about particular sources.
4. **EVALUATING** Assessing the impact of geographically targeted actions taken to control specific sources and/or reduce exposures to pollution in populations living close to specific sources.

The application of LCS sensor networks, and the priority questions to be addressed, are dependent on the current level of data and capacity for air quality management (Figure 2).

All four applications assume that data from LCS would be used together with conventional air quality management approaches, along with other innovative methodologies and measurement approaches, including land use regression and remote sensing.

- LCS should not be relied on to track long-term trends in urban-scale PM2.5 levels. That should be done by establishing at least one or more reference grade monitoring sites that can provide consistent data over time.
- LCS can help select location(s) but do not replace reference monitors. In the absence of reference monitoring, satellite-based estimates are an alternative for tracking long-term trends.

Deciding on adequate metrics for sensor performance depends on the intended application. Qualitative or indicative measurements could be useful at single sites or for educational purposes. For example, seeking to understand general changes in air pollution at a school or community center may not require the highest degree of data quality. However, higher data quality will be needed when comparing sensor data to other sensors or to standards.

The Limitations of Low-Cost Sensors

Awareness of LCS limitations can help ensure they are used only when fit-to-purpose and avoid data quality issues and other pitfalls by following the guidance in the subsequent sections of this compendium.

- Sensor accuracy and precision have been major limitations. One important step in reducing data quality problem is choosing a device that has performed well when tested by an established testing and evaluation programs as described in this compendium.
- Aside from LCS for measuring $PM_{2.5}$ current LCS have not been demonstrated to perform well in measuring several key pollutants for air quality management, including particle composition and nitrogen oxides and sulfur dioxides, and ozone. Other methods, such as lower cost filter based pm samplers, and passive gashes pollutant samplers are available.
- While the low unit cost of LCS is appealing, personnel and other resource costs of device deployment maintenance calibration and replacement, along with data network management and maintenance, must be considered.
- Finally, LCS networks may generate data that appear concerning, but are not actionable, for example, a short spike seconds to minutes in levels measured by a single sensor can indicate a malfunctioning sensor, or a transient increase in highly localized air pollution. The latter has little relevance for air quality management or public health, unless it affects a large population or occurs regularly. Either type of spike can distract limited government air quality personnel and resources from efforts to control identified important pollution sources.

Selecting/Specifying an Air Sensor

Air sensor performance is a measure of the accuracy, precision and reliability of the device. Performance can vary significantly from one manufacturer to another due to several factors (e.g., weather conditions, pollution levels). To help users and buyers select air sensors, organizations have developed evaluation centers and testing protocols to determine which air sensors perform best. These organizations include government agencies, academic institutions and associations which have developed ways to test air sensors with credible, independent methods. This section provides an overview of the evaluation methods, resources and other criteria for identifying which air sensors are best for your application. Standards and testing groups have recently developed testing methods[10,11]. The air sensor methods typically include evaluating sensors both in laboratory and field settings. Laboratory evaluations test air sensors alongside reference instruments in a monitoring chamber that controls all the test parameters—temperature, humidity, and pollutant concentrations. Laboratory evaluations cannot fully simulate actual conditions as field evaluations can. While laboratory-based evaluations allow us to investigate device performance systematically, the full range of real-world conditions is challenging to simulate in a controlled environment. Field testing provides the user the opportunity to challenge the sensor and determine the overall device performance. Field testing is conducted by colocating multiple (at least three) air sensors with a reference-grade instrument.

Performance Metrics

Performance metrics are parameters used to describe how well an air sensor performs relative to standard (typically a reference instrument). Many metrics can help you understand the performance of an air sensor and determine which sensor is best for your application. The metrics are briefly summarized below and in Figure 4, and are generally applicable to both particle and gaseous pollutants.

Accuracy and Precision

Accuracy and precision together refer to the ability of sensors to measure the correct value reliably each time (Figure 3). Precision refers to measuring the same thing every time and consistently. Accuracy refers to whether the sensor is measuring the correct or true value. For example, if a reference-grade instrument measures 60 µg/m3 and three colocated air sensors measure 40, 40.5, 42 µg/m3, the test devices are precise (values are close together) but not accurate (values are not accurate or different from the reference-grade instrument). See the Appendix on specific metrics to determine precision and bias.

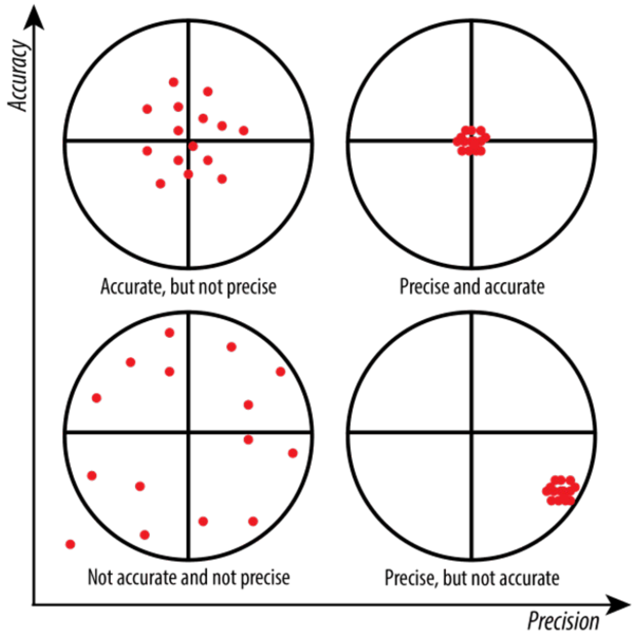


Figure 3: Sample data illustrating different levels of accuracy vs precision. Image source: https://wp.stolaf.edu/it/gis-precision-accuracy/

Correlation

Correlation measures how air sensor data relates to data from a reference instrument, which can be assessed for measurements over time or space. Temporal correlation measures how much measurements from different monitors vary together over time. Spatial correlation measures how much pairs of measurements, such as weekly averages, at multiple locations, vary together. A high correlation (>0.80) indicates that the sensor detects similar trends to the reference data. However, sensor measurements can have a high R^2,

Assessing Performance: Laboratory and Field Evaluations

Laboratory-Based Evaluation

Laboratory based validation helps systematically evaluate the performance of air sensors under a range of temperature and humidity conditions. Evaluations are typically done in a monitoring chamber that can control, maintain and monitor all the test parameters—temperature, relative humidity, aerosol concentrations. While laboratory evaluations are helpful in determining sensor performance under specific, regulated conditions, they cannot fully simulate actual conditions as in field evaluation. The difference between field and laboratory conditions may be due to factors like change in weather conditions, and difference in pollutant properties.

Field-Based Evaluation

While laboratory-based evaluations allow us to investigate device performance systematically, real-world conditions are difficult to simulate in a controlled environment. Field testing demonstrates how a device will perform under real-world conditions. Results of field testing, including colocation with a reference grade monitor, must be reviewed when planning a LCS project. Field evaluation should be conducted under conditions similar to conditions expected, with performance compared to reference grade devices. Established pollution control board monitoring stations with reference grade $PM_{2.5}$ devices may be used as colocation sites for field evaluation.

Key resources on air sensor evaluations

- U.S. EPA – Performance Targets: https://www.epa.gov/air-sensor-toolbox/air-sensor-performance-targets-and-testing-protocols
- Air Quality Sensor Performance Evaluation Center (AQSPEC): http://www.aqmd.gov/aq-spec
- Air Parif: http://www.airlab.solutions/en

but differ significantly from reference measurements due to bias (e.g., an R^2 of 0.99 and significantly over-or underestimating the actual concentration).

Another useful metric for such comparison is Normalized Root Mean Squared Error or Normalized RMSE. Lower values of this metric indicate a better agreement with the reference grade instrument. See appendix on how to calculate R^2 and RMSE values.

Intra-Model Comparisons

When using multiple air sensor devices for a project, it is important to ensure that readings from different devices agree with each other. To ensure this, intra-model comparisons, using at least three devices of the same make, model and firmware version for the test are recommended.

Performance Degradation

Air sensor performance (or drift) may decline with time due to various reasons—malfunctioning mechanical parts, particle deposition on internal sensing units, or other environmental factors. Instrument drift may be positive or negative and lead us to wrongly conclude that concentrations are decreasing or increasing with long-term projects, lasting more than six months. The manufacturer should provide the expected lifespan of the sensor and the methods to detect and fix drift of the air sensors.

Sensor Response to High Concentrations

Given that ambient conditions in Indian cities can exceed by an order of magnitude or more the typical levels under which LCS have been evaluated, it is important to determine sensor response under extreme pollutant concentrations (e.g., concentrations similar to Diwali and smog events in Delhi). Note that very good performance in a region with lower concentrations may not translate to other areas.

Data Completeness

Data completeness indicates the reliability of your air sensor, meaning whether the sensor network is consistently operational and produces a complete dataset needed to draw conclusions about the state of air quality. Expected seasonal pollutant concentration variations would not be reliable if the sensors did not last at least a year. Spare sensors that have been tested to compare well in colocated tests should be set aside.

Innovative Approaches for Low-Resource Settings

For deployment sites that do not have reliable access to electricity and Wi-Fi connection, air sensors with built-in battery, and cellular modules are more suitable. Alternatively, some research groups use a combination of portable battery packs along with a Wi-Fi module in conjunction with the air sensor. While this configuration allows for flexibility of deployment in low-resource settings, because of additional parts it will not be a plug-and-play setup and require additional personnel time for monitoring and troubleshooting. Added costs of purchasing and maintenance of battery pack and Wi-Fi module should also be considered.

Effect of Environmental Parameters

It is important to ascertain whether devices record temperature and humidity data. If not, make plans to record humidity and temperature data separately, or access these data from a nearby meteorological station. (Among many LCS models, humidity and temperature are parameters that impact performance and thus the importance of recording these values.)

Calibration Factors Can Change with Time and Space

Calibration can really be a moving target for most LCS measuring $PM_{2.5}$. Given that LCS operate on optical scattering principle, changes in optical properties of particulate matter impact LCS response. These properties (like size distribution, humidity-response, etc.) change frequently in urban areas and with different seasons. Thus, calibration factor can differ not only with location but also with time. For example, calibrations in Delhi don't hold in Bengaluru, and calibrations done in summer may not hold in monsoon or winter. To address this issue, it is advisable to maintain one LCS colocated with the reference monitor (eg. a Beta Attenuated Monitor (BAM) within the general study area.

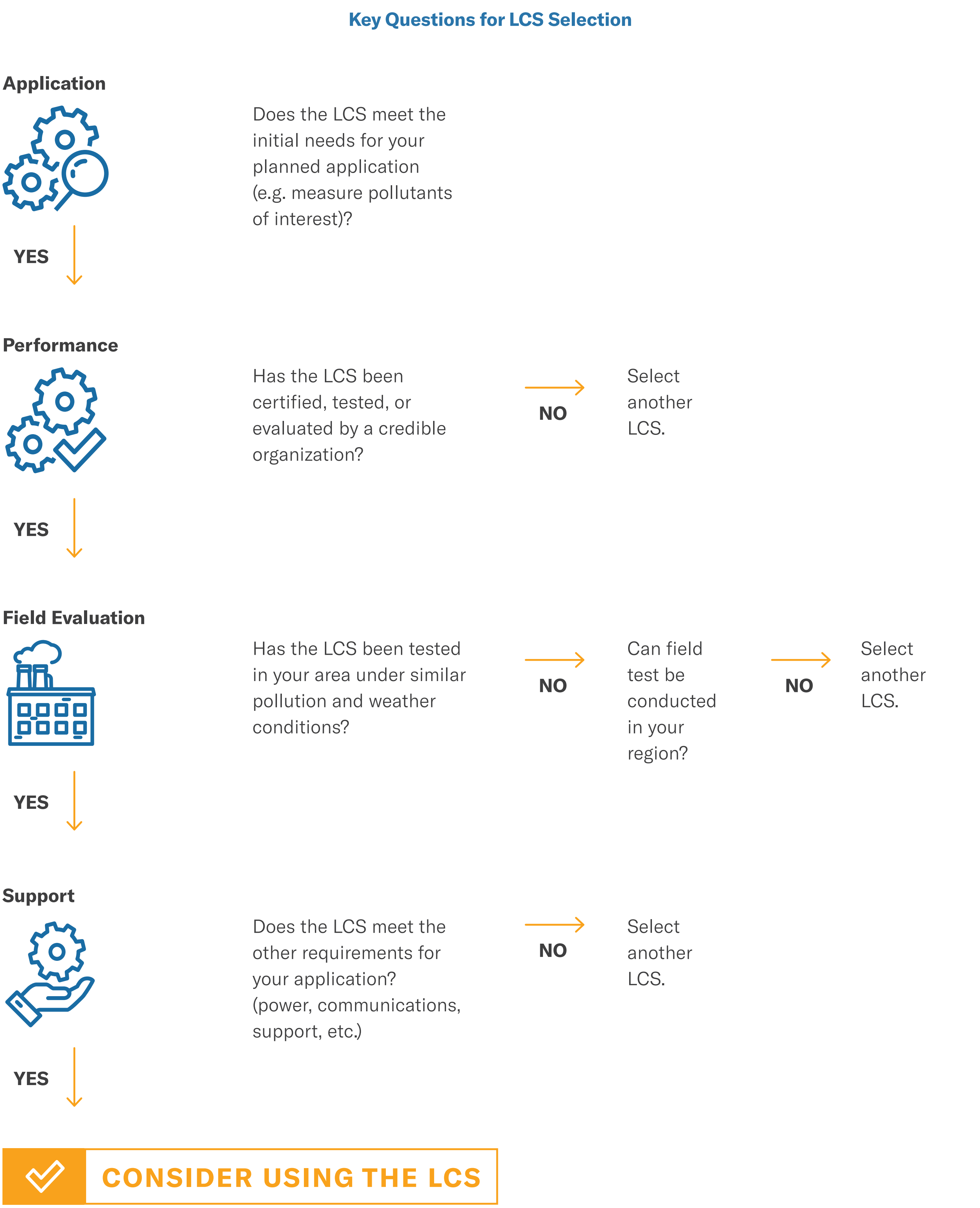


Figure 4. Key questions for LCS selection

Other Factors to Consider During Air Sensor Selection Process

Many other factors are essential to selecting an air sensor to meet stated goals and objectives. Some of these factors can significantly affect the performance of the air sensor, whereas others can affect the initial and ongoing costs of air sensors. For example, an unstable power supply can produce an incomplete dataset and limit analysis.

Power

This is an important consideration when designing networks for cities with frequent power outages or challenges getting power. Disruption in power supply translates to poor data quality. Several options exist for powering an LCS, including mains power, solar, and battery power. A backup power supply can improve reliability. With solar, ensure that you have sufficient sunlight. Plan ahead to ensure that your site has power or adequate sunlight for solar.

Expected Project Duration	Key Considerations
Short-term	Performance, ease of installation, maintenance
Mid-term (up to a year)	Repeated calibration, sensor drift, spare parts and replacement
Long-term (> 1 year)	Shelf life, extended warranty

Table 1. Key considerations for LCS selection depending on project duration.

Key questions to inform device selection, site location, and study implementation

- What are the main project goals and objectives?
- What is the current level of air quality management capacity?
- What data are currently available to inform study design and address key questions?
- What is the proposed time horizon to achieve results?
- How will data be analyzed and visualized?

Data Transmission

Typically, data from the air sensors is transmitted to a cloud in real time. This transmission happens either via a Wi-Fi network or cellular module (sim card) in the sensor. As a backup, data can be stored on internal memory (e.g., S.D. card) that can be accessed physically. Wifi can be unstable or require additional security protocols, which reduces the reliability of the air sensors. A cellular module may require monthly costs that need to be considered.

Calibration

Air sensors, like all instruments, will require periodic calibration and correction to the data. Sensor response may change with seasons, pollutant properties, and sensor age. It is recommended that air sensors be calibrated on-site before deployment and periodically after deployment. See the calibration section for further details. Ask the LCS manufacturer about the protocols, frequency, and additional cost to keep the sensor calibrated.

Considering Project Duration

How long you plan to monitor with LCS may affect some of the features and functions you select for your sensor. Short-term projects of a few months and those that do not span more than one season may not require extensive siting logistics or recalibration efforts. For long-term projects, especially beyond a year, the air sensor's shelf life and associated warranties become important. Given that, typically, air sensors last for about one or two years, you will need to plan for replacements and additional maintenance. Other costs to consider add up with time—repeat calibrations to capture seasonality, sensor-drift, network and data management fees, etc. (Table 1).

Selecting Data Services

Air sensors produce lots of data that must be managed, controlled for quality, and made accessible for the success of your project. It's critical to understand what data services are provided (or not provided) by an air sensor company. Alternatively, pollution control board staff may consider performing some of these services within if equipped with necessary experience and tools.

Data Management and Access

Manufacturers should provide details on where the data are stored and how it can be accessed. Cloud-based systems are ideal, and facilitate easy storing, tracking, and quality assurance of data.

A few questions to consider include:

- Are the data available publicly or password protected?
- Do users have access to raw data?
- Who owns the data?
- How long will the data be retained after the study?
- Are there additional costs for mobile app/website?
- Is there a limit on how many people can access the websites/app?

Data Review

It is important to decide early on whether Pollution Control Board staff or the vendor will be responsible for reviewing data and addressing any problems identified. If this service is provided, manufacturers should describe the process and frequency of reviewing data on a daily basis to identify problems early, course correct, and help create a high quality and complete dataset.

Customer Support

You may need additional help from the vendor with installing, operating, and using air sensors, so consider these questions: How long does customer support last? What support is included with the warranty? What support is not included? Does the vendor have support in India or nearby regions?

Pricing Models

There are two types of pricing models in the market: 1) purchasing air sensors; and 2) sensor subscription or leasing plans. When purchasing hardware, additional costs should be considered, for example: related accessories, spare parts, service and shipping for repairs and maintenance, ongoing operation costs, data access costs, etc. A request from the vendor can be solicited for cost of items needed and their cost for the first year of service. The subscription model typically includes hardware plus services. The annual subscription cost may consist of maintenance, repair/replacement, data access, data QA/QC, etc. The specific slew of services will depend on the manufacturer.

Logistics

LCS projects with multiple sensors placed in different locations often demand flexibility and triaging capability in order to make the best use of available resources. Beyond the setting-up stage, there are several small tasks that can drain time and resources. Travel times from one location to another to troubleshoot for setup/troubleshooting/repair/replacement can be long, and it helps to have a skilled field project manager who can prioritise and plan network maintenance in the most efficient way.

Comprehensive Cost Considerations

While the low unit cost of sensor devices may be appealing, personnel and other resource costs of device deployment, maintenance, calibration, and replacement along with data network management and maintenance[12] could offset much or all of the savings expected from instrument purchase. Key cost considerations (Figure 5) include:

- Purchase price of monitoring equipment;
- Supplies, parts, service and shipping for repairs and maintenance;
- Consumables, shipping, laboratory analysis for filter-based samples;
- Structures, supporting infrastructure, security;
- Property leases, if applicable;
- Utilities, including reliable electric power and wireless communications;
- Personnel costs for deployment, calibration and maintenance; and

Figure 5 Key costs to consider for LCS selection

- Data management equipment, fees and personnel.
- LCS network design and sensor placement

LCS Network Design and Sensor Placement

The deployment of networks of low-cost sensors involves first designing a network and then selecting places to install the air sensors. Significant planning goes into creating a network to determine the general location of sites, determine the number of sensors you'll need, and select and work on logistics to install the sensors. All these tasks can be done by Pollution Control Board staff, consultants, or in some cases, by the vendor who may provide this service. Tenders should specify the specific services needed, and responses should clearly state who will deliver these services.

Designing a network of sensors can be challenging because the choice of locations and number of sensors deployed depend on the application, geographic coverage, budget, and more. Several considerations are critical when designing a network and are described below, by intended application:

Siting

For siting new locations for reference monitoring, air sensors can be installed at many locations to monitor air quality gradients (i.e., differences in air quality across a region). The duration of monitoring would depend on the objective of the reference monitoring but can range from a short-term or long-term basis (across different seasons).

Mapping Air Quality

Sensor installation sites for mapping air quality of a city can be done at different locations across the city to capture gradient in pollutant concentrations, leading to corresponding population exposures. The appropriate locations would be near sources like industries, heavy traffic roads/intersections, and background sources such as pristine areas or rural sites. Certain sites that are likely to be affected during episodic events like crop residue burning can also be selected under this application.

Identification of Hot Spots

For identifications of hot spots, the air sensor devices should be installed at locations that are close to the source, such as near a high-traffic intersection, industrial cluster, open biomass burning, or a power plant where the concentrations are potentially high.

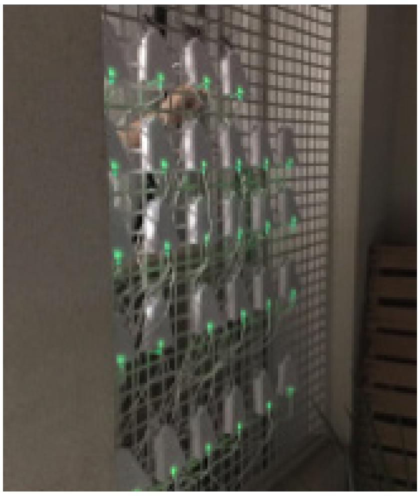
Evaluation

For evaluating the impact of policy interventions to curb air pollution in a specific geographic location, installation sites should be close to the identified sources where there is an anticipated impact on the human population.

Further Considerations

The location and number of sensors deployed should also take the following considerations into account[13]:

- **Need for backup sensors:** Budget for several spare sensors to replace air sensors that fail during the study.
- **Meteorological conditions:** Weather and wind conditions affect pollution concentration and distribution, so you may want to locate sites up-wind and downwind to monitor pollution flowing into and out of an area.
- **Colocation of sensors:** Plan to permanently operate an air sensor next to reference stations to continuously monitor sensor performance.
- **Access:** It is important to have access to the site to provide maintenance, replacement, and other activities throughout the study. A formal access agreement may be needed that lists the terms and procedures for accessing the site.
- **Power:** Air sensors may need to be plugged in, may have solar panels, or offer both options. If power is required, ensure that there will be access to reliable power. If solar is utilized, make sure there is sufficient exposure to sunlight (i.e., solar panels not blocked by buildings or other structures).
- **Security:** Air sensors and equipment are subject to tampering and theft. Look for a secure location overhead out of arm's reach, in an inconspicuous place, or behind a locked gate or fence.



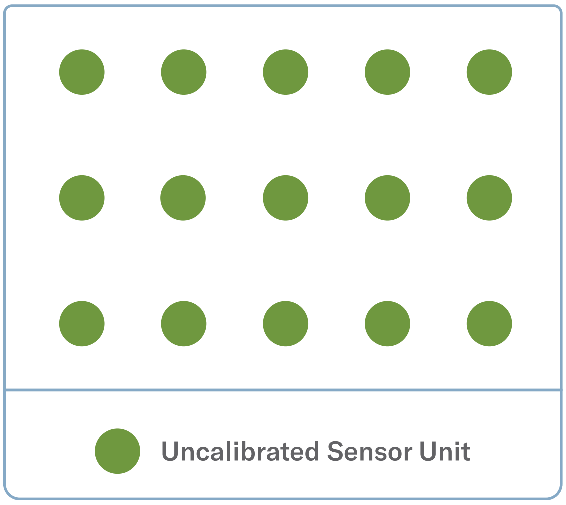


Figure 6. Example of sensor colocation for comparability testing

- **Height of sensor:** Try to locate air sensors about 1-2 meters above the ground or rooftop. Avoid areas close to local pollution sources (e.g., chimneys, cooking vents) or pollution sinks (e.g., trees). Place sensors to allow for unrestricted airflow to the sensor to provide representative measurements.

Quality Assurance/Quality Control

This section describes some key quality assurance/quality control approaches to maintaining quality throughout the project.

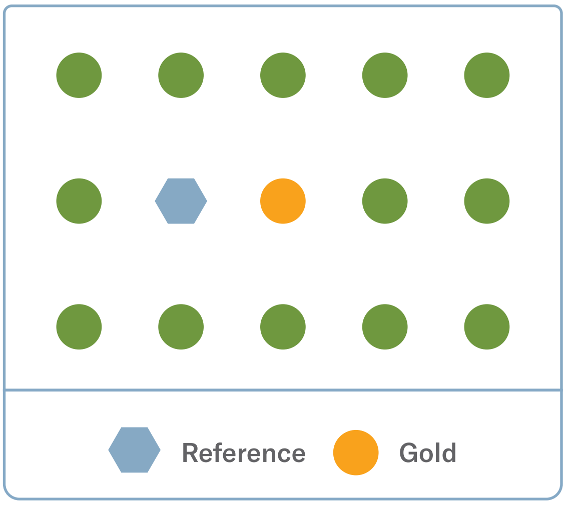
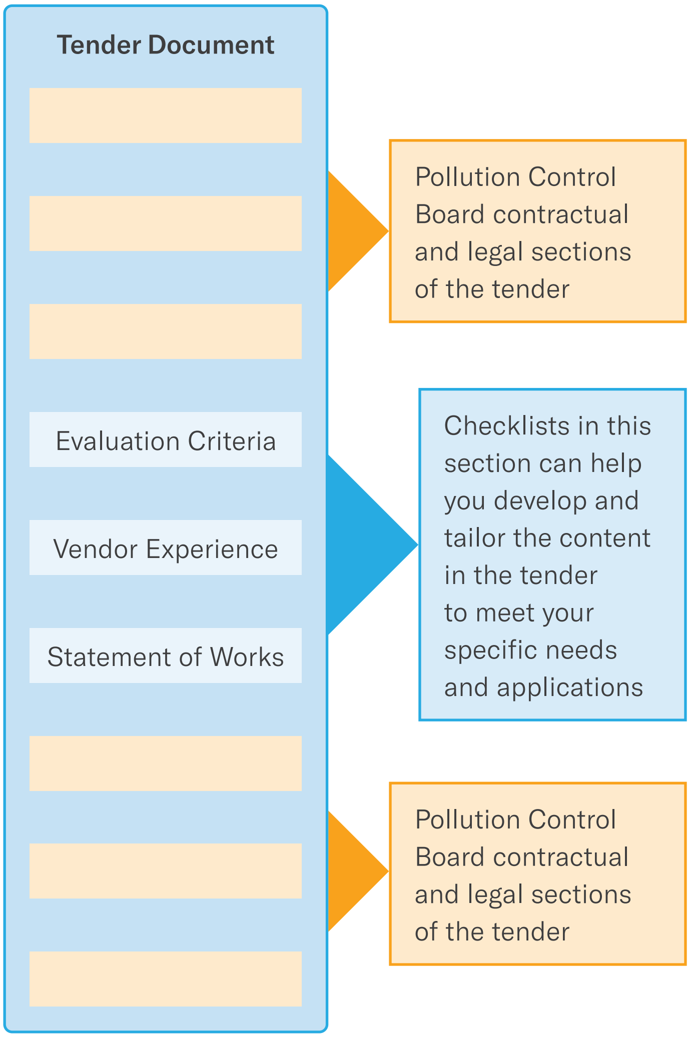


Figure 7. Calibrating air sensor devices by colocating with reference monitor

Colocation	Process of side-by-side comparison with a reference grade instrument. The air sensors and reference monitors are located next to each other.
Calibration	Process of adjusting the air sensor data with respect to a reference standard. Air sensors are calibrated by the manufacturer (factory calibration) and also in the field (by colocating with reference monitors) before a measurement campaign. Since sensor response may change with seasons, repeated calibrations are required for mid- to long-term projects. Some vendors offer “remote” or “cloud” calibration feature that may be misleading. See Figure 9 in Appendix to learn more.
Data-correction	Process of adjusting the data for other factors that may influence the air sensor response—for example, temperature, humidity, etc.

Table 2. Definitions of commonly used terms in low-cost sensor quality assurance/quality control context: colocation, calibration and data correction.

Pre-installation



Calibration/Colocation

Air sensors need to be calibrated to provide high-quality data. Some manufacturers may provide a calibration factor or may provide calibration services once the sensors are deployed. While the air sensors may be calibrated when installed, they should be recalibrated periodically before, during, and after the study. Calibrations are determined using data collected during a colocation. A colocation is a side-by-side comparison of air sensors to a known standard or reference. Determining calibrations is an active area of research for academic institutions and manufacturers who are developing new methods. There are several important recommendations regarding calibrations: 1) Understand the method used to calibrate the sensor data, 2) Determine how frequently an air sensor will need to be calibrated and who will calibrate it, and 3) Ensure the raw and calibrated data are archived. It is important to create a formal LCS calibration or colocated comparison guide so all personnel follow the same procedure. More details on formulations used estimate calibration factor are discussed in the Appendix.

Colocation to Assess Sensor Comparability

All sensor units to be deployed should be colocated and tested to determine any variation in measurements among the different units (also called intra-model variability). These colocations should be preferably done in the field where the sensors can be in close vicinity of each other so that the sampled air has similar pollutant concentrations as depicted in Figure 7. This method is of critical importance when considering applications involving measurements for determining spatial gradients (e.g., hotspot identification and air quality mapping). Such colocations are typically conducted for several days to several weeks. It is also important to evaluate performance under the expected range of pollution and weather conditions while colocations are conducted.

Colocation with Reference Monitor

Once the sensors have been evaluated for intra-model variability, at least one sensor should be calibrated in the field through colocation of sensors adjacent to a reference monitor to confirm that the air sampled has similar pollutant concentrations, as illustrated in Figure 6. These calibrations help to determine accuracy in relation to the accuracy of the known reference and a proof that they are working as claimed. Colocation with the reference monitor should be done for all sensors for several weeks at the beginning of the study (if possible). At the end of the study, all sensors should be colocated again at the reference site. Alternatively, there are other methods for calibration being used by researchers, such as the “golden” sensor method where the calibrated, or “gold” sensor is used to calibrate other sensors in the field.

Post-Installation Maintenance

Calibration: Once the comparability and accuracy criteria are established during the predeployment stage, the deployed sensors can be calibrated periodically (weekly or monthly) by colocating the “gold sensor” with the rest of the sensors in the field. This is called "transfer calibration." Alternatively, each of the deployed sensors can be brought back to the reference site and can be calibrated in rotation.

Accuracy: The data generated from sensors also need to reviewed periodically. Any anomaly needs to be documented and corrected from time to time in order to avoid hassles created while handling voluminous data. For estimation methods of accuracy, bias and precision refer to Appendix.

Data Validation: Data validation is an important step before analyzing the data to get definitive inferences. It requires visual interpretation of the raw data in order to identify unusual patterns, outliers, drifts and delay in sensor response. Data validation should be performed as quickly as possible, on a daily or weekly basis. A validation scheme should be determined for a quick assessment of data quality. This is especially important for cities with no historical data for comparison.

Some of the other specific tasks for maintenance include:

- Ensuring proper placement of air sensor devices;
- Cleaning of the device;
- Replacing consumables like filters, batteries and other accessories (if applicable);
- Replacing sensors generating faulty data or those which have been deployed beyond their lifespan;
- Keeping track of any physical obstruction like trees and buildings that might arise as the monitoring progresses;
- Creating (during installation), maintaining and validating a protocol so a record of actions is kept and a systematic methodology is followed.

Summary Guidance to Inform Tender Development and Evaluation

This section is intended to inform the development of evaluation and technical areas of a tender for an air sensor network, to be included along with other contractual and legal sections required. It provides a sample framework and sample criteria needed to request and evaluate vendor qualifications, experience, and responses received. Included are checklists for tasks, services, and features that a Pollution Control Board may want to consider requesting in the tender as relevant. The sample tender template may be used as a starting point to be adapted according to project objectives and context.

Proposal Evaluation Criteria

The following criteria may be used to evaluate submitted proposals

1. Extent of covered services and support;
2. Demonstrated performance and experience in relevant context;
3. Track record of service in India;
4. Warranty period and turnaround time for repairs/replacements;
5. Data quality and ease of access; and
6. Overall costs.

Vendor Qualifications and Experience

Past Experience

1. List past projects done in similar capacity
2. For each project include in the proposal:
3. Short description

    a. Name and contact information of client organization
    b. Dates of service
    c. Summary reports or publications
    d. Permission to contact the client ? ○Yes ○No

Demonstration of Performance

Applicants may include the following information as a demonstration of past performance:

- ○ Peer-reviewed publications
- ○ Reports or white papers by independent organizations
- ○ Certification of sensor technology based on performance standards issued by government agencies and associations
- ○ Evidence of sensor performance in similar pollution and weather conditions (as described in SOW)
- ○ Evidence of sensor performance evaluation by a credible organization
  - Does the evaluation include field evaluation? ○Yes ○No
  - List the name(s) of the evaluation organization?
  - Specify the date and time period of each evaluation?

Personnel Experience

Describe in the proposal, personnel experience for each level of support service offered.

Customer support locations

List locations and contact information for local and regional offices along with expected turnaround time for complaints and replacements.

Potential Inputs to Statement of Work

(To be completed by organization issuing the tender)

Background and Objectives

- ○ General location of study _______
- ○ Purpose of the network ___________
- ○ Study duration _________

  Note: Typically 12 to 24 months is the minimum duration needed to measure a range of conditions over different seasons. You might want to consider options to extend the study.
- ○ Pollutants to be measured ____________
- ○ Number of air sensors __________
- ○ Location of air sensor sites (if available) ____________
- ○ Ancillary data needed (weather, emissions, etc.) ___________

Sensor Specifications

Check box(es) for which supporting information is provided/service is included and provide additional details in the proposal.

- ○ Sensor Performance
  - Specify bias, precision, correlation and root mean square error (RMSE) and provide supporting data as part of the proposal

    Bias ____

    Precision ____

    Minimum correlation coefficient (with reference grade) ______

    RMSE _____
  - Note: Requirements for above parameters are as follows (to be filled by the Pollution Control Board based on project requirements):

    Max Bias ____

    Min Precision ____

    Min correlation coefficient (with ref grade) ______

    Max RMSE _____
- ○ Sensor Uptime
  - Specify the average sensor data completeness _____

    Provide supporting data in the proposal
  - Note: Minimum data completeness requirement is 75% i.e. air sensor working and producing valid data at least 75% of the time
- ○ Data Calibration
  - Describe in the proposal how data are calibrated ________
  - Specify frequency and process of data calibration ________
- ○ Power and telecommunication
- ○ How is the air sensor powered? (Check all that apply)
  - ○ Solar _____
  - ○ Mains _____
  - ○ Battery _____
  - For internal battery, include time required for a full charge ____

How is the data transmitted (Check all that apply)

- Wi-Fi _____
  Cost covered by vendor ○Yes ○No
- Cellular _____
  Cost covered by vendor ○Yes ○No

○ Other specifications

- Specify time resolution of data ______
  Note: Minimum requirement is 1-minute frequency
- Specify format of data stored ______
  Note: Data should be in a machine-readable format (eg. csv)
- All supporting infrastructure i.e. installation, mounting, support hardware should be included in the quote.
- Air sensor should be weatherproof and able to operate in hold, cold, and rainy conditions.
- Vendor should include a current user guide and manual covering installation, operation, data access, and repair.

Installation Services

Check appropriate box(es) for the services provided

○ Siting locations: Vendor will select sites based on project specifications and will be responsible for setting up power, supporting infrastructure, developing access agreements, and ensuring safety and secure access. Selected site should be representative of air quality in the area with minimal interference from the location's surroundings.

○ Colocation with reference site: Vendor will set up and operate the air sensors next to a reference site before the project begins. Any sensors that do not meet specifications will be replaced by the vendor.

○ On-site installation: Vendor will set up and install sensors at selected site. Vendor will be responsible for the infrastructure needed for installing on site. Vendor will confirm that each sensor is operating correctly after on-site installation and also documentation the site conditions, location, and elevation.

Key Considerations for Tender Development and Evaluation

- Ensure that a vendor's proposal delineates what support services are included (or not included) in their price.
- Vendors should also provide cost of replacement and services if the project outlasts the warranty period.
- Prefer vendors that have customer support locally and ensure faster turnaround time for replacement and repairs (to ensure data quality).
- Inquire about additional services that vendors can provide (eg. siting, colocation with PCB reference site and site installation of sensors).
- Vendors should provide evidence of verification/evaluation of their LCS product by independent and credible organization(s).
- Vendors should provide data from evaluations done under comparable settings to support field performance claims. This is to ensure that there is evidence of LCS performance under similar pollution and weather conditions.
- Vendors should not submit one-of-a-kind or experimental models.

Maintenance and Warranty Services

○ Specify the warranty period ________ (minimum 12 months required)

○ Specify the cost of additional warranty ________

Check boxes for services included within warranty period

○ Periodic on-site maintenance.
Specify how frequent the maintenance will be performed ---------------

○ Sensor replacement costs (shipping, installation, etc.)
Specify the turnaround time for repairs and replacement ___________
Are the shipping costs included in the warranty? ________

Data Services

Check appropriate box(es) for the services included

Note: Data should be in a machine readable format (e.g. CSV)

- ○ Data management system
  Describe the data management system in the proposal ________
- ○ Web-based data portal
  - Specify the maximum (if any) number of users ________
  - The portal is: ○ public ○private ○both options included
- ○ Data calibration services
  - Specify the frequency of calibration __________
  - Describe how raw and calibrated data will be stored/accessed (in the proposal) ________
- ○ Mobile app
  - The app is: ○public ○private ○both options included
  - The mobile app runs on: ○Android ○iOS ○both
- ○ Data retention
  - Specify for how long the data will be stored ______
  - Describe how the data will be stored/accessed (in the proposal)
- ○ Quality control checks (required)
  - Describe protocol for quality control checks (in the proposal)
- ○ Alarms and alerts
  - Describe types of alerts available (eg. emails/ text messages when the sensor goes down or there are very high concentrations).
- ○ Application Program Interface

Training Services

- ○ Specify if the training is remote or in-person
  How much time is needed for the training ________
- ○ Is there any travel required

Check appropriate box(es) for the trainings provided:

- ○ Air sensor installation, operations, and maintenance
- ○ Data access
- ○ Web portal features and functions
- ○ Calibration procedures and methods
- ○ Colocation protocols
- ○ Other. Specify ____________________

Customer Support

Check appropriate box(es) for the customer support provided:

- ○ On-call support
  Provide contact information ________________
- ○ After-hours support
  Provide contact information________________
- ○ Data review and analysis

Case Studies: Examples of Application of LCS in India

Innovative Calibration of Low-Cost Sensor Network in Maharashtra

Air pollution has emerged as a major environmental threat in India with a significant burden of disease attributed to it. The majority of human health risks due to air pollution are dominated by particulate matter (PM). Monitoring air pollution (PM) is challenging in a resource constraint situation as in India. The continuous ambient air quality monitoring stations (CAAQMS) that are used by the government agencies to monitor the concentration levels of PM are expensive (20 Lakh INR). Therefore, CAAQMS cannot meet the neighborhood monitoring needs of air pollution.

Low-cost sensors (LCS) have emerged as a potential cost-effective option for dense air pollution monitoring networks. However, they require in-field calibration to improve their precision and accuracy in comparison with CAAQMS.

Professors Sachchida Nand Tripathi and Vipul Arora from I.I.T. Kanpur (IITK) collaborated with the Maharashtra Pollution Control Board (MPCB) to evaluate PM sensors, with support from Bloomberg Philanthropies. The project aimed to develop a robust machine learning based method for the calibration of PM2.5 low-cost sensors in the 15 identified locations of the larger Mumbai Metropolitan Region (MMR). The 15 identified locations are namely: Airport, Sion, Nerul, Vileparle, Mahape, Vasai, Borivali, Kalyan, Powai, Colaba, Mulund, Bandra, Kandivali, Worli, and Kurla. The sensors' deployment and data collection started on Nov. 1, 2020 and continued till May 31, 2021.

Four startups—Airveda Technologies Private Limited, Personal Air Quality Systems (PAQS) Private Limited, Respirer Living Sciences Private Limited, and Oizom Instruments Private Limited—were shortlisted to participate in the experiment. In total, 40 LCS were deployed across 15 locations in the Mumbai Metropolitan Region. The low-cost sensors used by each start-up device are as follows: Respirer (PlanTower), Airveda (Nova Fitness), Oizom (Nova Fitness) and PAQS (Telaire Dust Sensor).

The frequency of data streaming varied from 1-60 minutes for the different start-ups. The dataset averaged for 60 minutes was considered in this work. The uptime, or time during which the sensors were operational, on a monthly basis for LCS was found to be comparable or superior (> 90%) to the CAAQMS monitors and required the least manual assistance during operation over the 7-month-long deployment.

When the sensor readings with the CAAQMS measurements were compared, researchers found quite a mismatch between the two[14]. Hence, they used machine-learning-based regression techniques to process the raw LCS readings and calibrate results to reduce the mismatch between calibrated values and CAAQMS measurements. The machine-learning-based calibration models were trained with the help of colocated and simultaneous data from LCS and CAAQMS, then deployed for the same LCS at the same location. Calibrated values are found to have less mismatch with respect to the CAAQMS measurements[15].

However, this present calibration method has its limitations. It requires longer duration training data, meaning longer deployment of LCS colocated simultaneously with CAAQMS monitors. This is cumbersome and not a cost-effective approach. To overcome this limitation, researchers are developing an adaptation method to minimize the requirement of colocated simultaneous CAAQMS data at all deployment sites. LCS are being installed next to CAAQMS at certain locations, called source locations, where lots of data may be collected for training. Other locations, called target locations, do not have CAAQMS installed. If mobile CAAQMS can be stationed at such target sites for a short duration, data may be used to reliably adapt calibration models.

Under this adaptation paradigm, the calibration model is being developed in two phases:

- In phase one, a base model is developed at a source location with a large training dataset. The results are obtained for the data collected from November 1, 2020 to January 31, 2021. The coefficient of determination (R2) is in the range of 0.60 to 0.92 for 11 out of 15 locations and 0.28 to 0.6 for the remaining four locations. The mean absolute percentage error (MAPE) is in the range of 7% to 16% for 11 out of 15 locations and 20% to 30% for the remaining four locations.
- In phase two, the base model is adapted to the target locations with a much smaller training dataset[14,15]. The adapted calibration model also shows good agreement with CAAQMS with the R2 ranges between 0.6 to 0.9 and MAPE ranges between 9% to 15% for 11 out of 14 locations. For the remaining three locations the adapted model shows the R2 ranges between -0.2 to 0.5 and MAPE ranges between 20% to 40%.

In the third phase, we are working on a generic model that can also work for the seasonal variations.

So far, results obtained exhibited a good correlation between the calibrated sensors data and Maharashtra Pollution Control Board monitors. The obtained results for the base calibration model are comparable with the other studies[16,17]. The developed model also tested on different sites other than the training site, and the model shows poor performance on the other locations[15]. However, the proposed calibration model exhibits good performance also for the other locations.

These results substantiate the potential of ML based calibration for rapid and mass deployment of LCS. The algorithms developed provide for quick calibration of new devices and their adaptation to new locations[14,15].

This project has already gained interest and support from government agencies.

Case Studies: Examples of Application of LCS in India

Application of Low-Cost Sensors to Provide Real-Time Spatio-Temporal Air Quality Data in Delhi and Bengaluru

India experiences a wide range of weather and environmental conditions that can impact the performance of low-cost sensors (LCS). Even at a single location, humidity can change dramatically from wet to dry season. Particle concentrations vary between clean conditions (rural environments, monsoon conditions) to some of the highest measured in the world (biomass burning in North India). Dr. Joshua Apte from the University of California has been working with colleagues to evaluate the precision (LCS measures the same thing every time) and accuracy (LCS produces correct results) of PurpleAir devices. The former is evaluated by comparing the devices of the same type and the latter is evaluated by comparing the LCS with a reference grade sensor.

Delhi

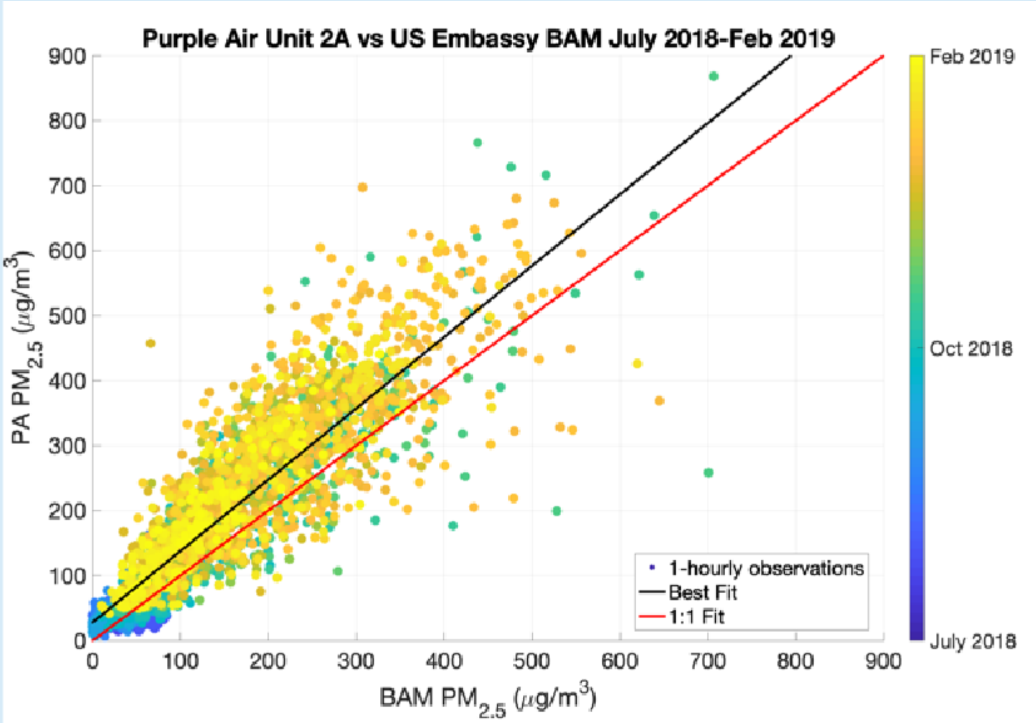
Dr. Apte's group evaluated two identical Purple Air sensors (PAII-SD) at the US Embassy in Chanakyapuri, New Delhi[18]. The Purple Air sensors operated continuously from July 2018 to February 2019 and were compared against the measurements from the U.S. Embassy's Environmental Protection Agency's Federal Equivalent Method Beta Attenuation Monitor (BAM), specifically the MetOne Model BAM-1020. Each Purple Air device has duplicate sensing units within (Plantower sensors). During the study period the Plantower sensors correlated well with other sensor outputs both within device and between devices. The hourly average measurements ranged from 0 to 750 ug/m3. The pairwise relationship between individual Plantower sensors varied by 10-15%. While some sensors read consistently higher or lower than other sensors, they all increased and decreased together, implying that the PurpleAir devices are highly precise.

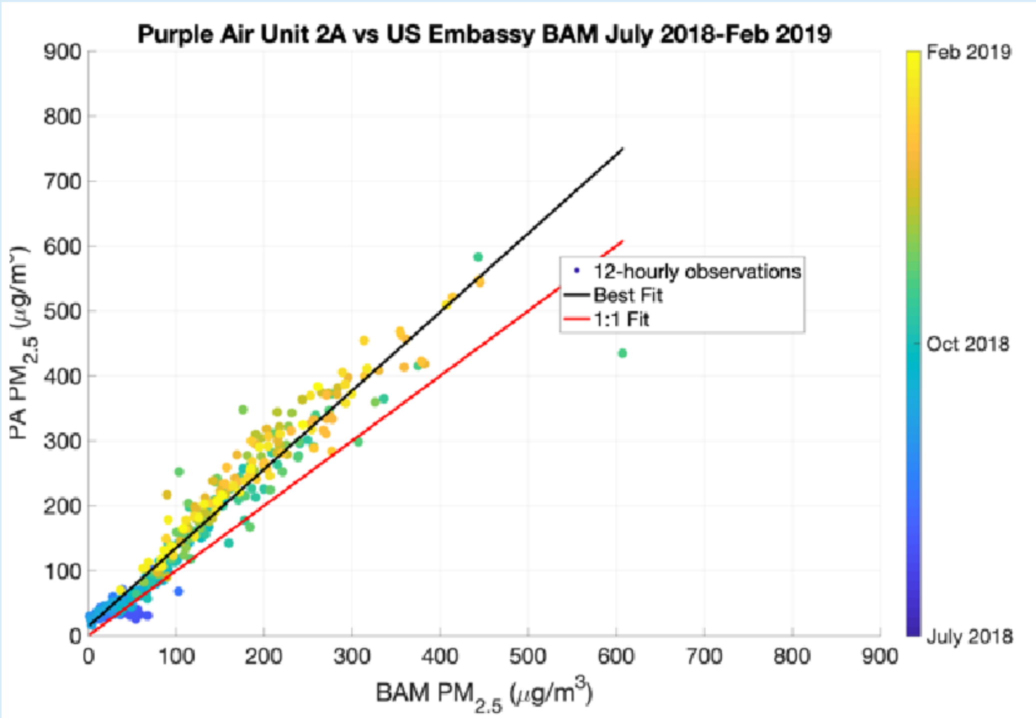
In order to check for accuracy, PurpleAir measurements were compared with those from BAM. The relationship between BAM and PurpleAir sensors was not consistent with time. PurpleAir over-reported values during the autumn seasons while underreporting during the monsoon season. During extreme pollution episodes, PurpleAir reported values that differed from BAM by ~150-200ug/m3 and in either direction. Furthermore, as seen in Figure A.2 the degree of agreement with BAM is dramatically higher for a larger averaging time (12-hour vs 1-hour). In conclusion, the PurpleAir devices can measure PM2.5 with strong precision, but variable accuracy. The findings also imply that the calibration factor of any LCS may significantly vary with time.

Bengaluru—a City Wide PurpleAir Network

To examine spatial variations in PurpleAir measurements (along with temporal variations), a network of ~40 sensors was set up in the city of Bengaluru by Professor Apte's group and partners. Established in summer of 2019, the network is up and running to date with ~80% data completeness. Real-time data from the network are publicly available

Figure 8. Hourly (upper row) and 12-hourly (bottom) average concentrations for PurpleAir sensor vs U.S. Embassy BAM reference monitor, July 2018 - February 2019.





on the PurpleAir website. The sensors are deployed at a mix of installation sites—schools, residential buildings, and offices. Prior to deployment all the sensors were colocated for calibration and quality-checked at an ambient site with a BAM-1022. In order to maintain data quality, in light of power/Wi-Fi disruptions, each PurpleAir sensor in the field is accompanied by a power backup and a W-Fi hotspot device which helps the sensor connect to an online server continuously.

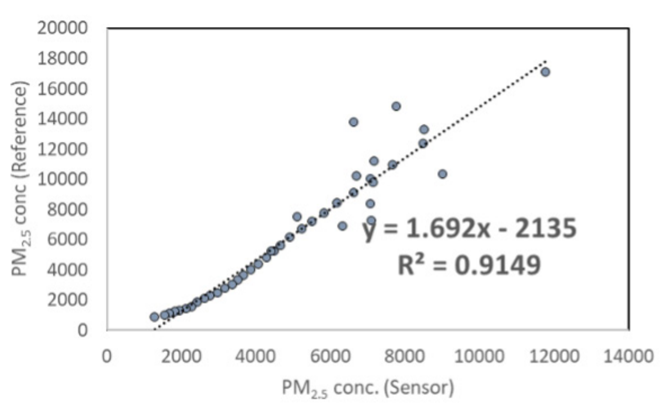
References

1. Vital Strategies 2020, Accelerating City Progress on Clean Air, Technical Guide. Viewed 19th August 2021, https://www.vitalstrategies.org/resources/accelerating-city-progress-on-clean-air-innovation-and-action-guide/
2. Guttikunda, S.K., Nishadh, K.A. and Jawahar, P., 2019. Air pollution knowledge assessments (APnA) for 20 Indian cities. Urban Climate, 27, pp.124-141.
3. Martin RV, Brauer M, van Donkelaar A, Shaddick G, Narain U, Dey S. No one knows which city has the highest concentration of fine particulate matter. Atmospheric Environ X. 2019 Jun 24;100040.
4. GBD Compare | IHME Viz Hub [Internet]. [cited 2018 Jun 27]. Available from: http:// vizhub.healthdata.org/gbd-compare
5. World Health Organization. Global Urban Ambient Air Pollution Database (update 2016) [Internet]. WHO. 2016 [cited 2016 Oct 21]. Available from: http://www.who.int/ phe/health_topics/outdoorair/databases/ cities/en/
6. WHO | Air quality guidelines - global update 2005 [Internet]. WHO. [cited 2016 Sep 14]. Available from: http://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/
7. United Nations Environment Program (UNEP). Actions on Air Quality: Policies & Programmes for improving Air Quality Around the World [Internet]. 2015 [cited 2019 Sep 10]. Available from: https://wedocs.unep.org/ bitstream/handle/20.500.11822/17203/ AQ_GlobalReport_Summary. pdf?sequence=1&isAllowed=y
8. Chow JC, Watson JG, Feldman HJ, Nolen JE, Wallerstein B, Hidy GM, et al. Will the Circle Be Unbroken: A History of the U.S. National Ambient Air Quality Standards. J Air Waste Manag Assoc. 2007 Oct;57(10):1151–63.
9. Brauer, M., Guttikunda, S.K., Nishad, K.A., Dey, S., Tripathi, S.N., Weagle, C. and Martin, R.V., 2019. Examination of monitoring approaches for ambient air pollution: A case study for India. Atmospheric Environment, 216, p.116940.
10. Duvall, R., A. Clements, G. Hagler, A. Kamal, Vasu Kilaru, L. Goodman, S. Frederick, K. Johnson Barkjohn, I. VonWald, D. Greene, AND T. Dye. Performance Testing Protocols, Metrics, and Target Values for Fine Particulate Matter Air Sensors: Use in Ambient, Outdoor, Fixed Site, Non-Regulatory Supplemental and Informational Monitoring Applications. U.S. EPA Office of Research and Development, Washington, DC, EPA/600/R-20/280, 2021.
11. ASTM WK64899, New Practice for Performance Evaluation of Ambient Air Quality Sensors and Other Sensor-Based Instruments, ASTM International, West Conshohocken, PA, 2018, https://www.astm.org
12. Bachmann J. Managing Air Quality: Lessons from the American Experience. Collaborative Clean Air Policy Centre. 2019 Available from: https://ccapc.org.in/policy-briefs/2019/us-experience-bachmann:14.
13. EPA (2021). A Guide to Siting and Installing Air Sensors. Available at: https://www.epa.gov/air-sensor-toolbox/guide-siting-and-installing-air-sensors (Accessed: 7th September, 2021)
14. S. K. Jha, et al. (2021): Domain adaptation based deep calibration of low-cost PM2.5 sensors. TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.15276186.v1
15. K. Yadav, et al. Few-shot calibration of low-cost air pollution (PM2. 5) sensors using meta-learning. arXiv preprint arXiv:2108.00640 (2021).
16. Zheng, T., Bergin, M.H., Johnson, K.K., Tripathi, S.N., Shirodkar, S., Landis, M.S., Sutaria, R. and Carlson, D.E., 2018. Field evaluation of low-cost particulate matter sensors in high-and low-concentration environments. Atmospheric Measurement Techniques, 11(8), pp.4823-4846.
17. M. A. Zaidan et al., Intelligent calibration and virtual sensing for integrated low-cost air quality sensors. IEEE Sensors Journal, vol. 20, no. 22, pp. 13 638–13 652, 2020.
18. Apte J. "State of Play of Low-Cost Sensors for Monitoring PM2.5," 2020. Available from https://www.edf.org/sites/default/files/documents/State-of-Play-Low-Cost-Sensors-Monitoring-PM2.5-India.pdf

Appendix: Approaches to Measuring Bias, Accuracy, and Precision

Calibration Curves

Calibration Curves may be generated using the data from colocation of sensors with reference monitors. A Simple Linear Regression (SLR) analysis, as presented in Figure A1 using the raw data readings for exactly the same monitoring duration, can be used to develop calibration factors. The calibration factors derived from the mathematical equations can further be used to correct the sensor data. For more robust calibration, Multiple Linear Regression (MLR) allows for the inclusion of other factors that may influence pollutant concentration measurements, such as humidity and temperature. Note that the calibration factor may change with environmental conditions and necessitate repeat calibrations.



Calibration Equation

y=1.692x-2135

Where '**y**' is reference concentration, '**x**' sensor concentration and '**1.692**' deviation in data **R2** is coefficient of determination which denotes closeness to slope intercept

Figure 9 Illustration of a calibration curve for data collected through colocation of sensor and reference monitor

Figure A1 Illustration of a calibration curve for data collected through colocation of sensor and reference monitor

Root Mean Square Error (RMSE)

RMSE is a measure of accuracy of calibration model that can be calculated using predicted and reference concentrations.

$$RMSE = \sqrt{\frac{\Sigma_i (p_i - y_i)^2}{n}}$$

Where p_i is individual predicted value, y_i is individual reference value and n is number of data points in the model

Bias

Bias in the sensor data result in concentrations that do not represent the true concentrations. It is a systematic error in measurement that results in either higher or lower concentrations at all the data points by some fixed deviation value. Bias should be calculated periodically, preferably whenever calibration is performed. Figure A1 shows a line graph representing bias in sensor data.

Bias can be estimated using the following formula:

$$B = \left({C}/{C_R} \right) - 1$$

where B is Bias, C is average pollutant concentration measured by sensor and CR is the average concentration of the same pollutant measured by the reference monitor.

Precision

Precision of the measured data shows the repeatability of concentrations when the sensor is used to collect data in the same conditions several times. Short measurement duration such as 1 sec interval might result in lower precision which can be corrected to some extent by grouping the data into 5 min averages.

Where P is precision, C_s is standard deviation of

$$P = \left({C_s}/{C_m} \right)$$

measurements and C_m is mean of the measurement at a fixed concentration.

Integrated Use of Low-Cost Sensors to
Strengthen Air Quality Management in Indian Cities