CSTEP

Performance Evaluation of Multi-pollutant Air Quality Sensors at Indi-SET, Bengaluru, India – First Edition

A Short- and Medium-Term Evaluation

Performance Evaluation of Multi-pollutant Air Quality Sensors at Indi-SET, Bengaluru, India – First Edition

A Short- and Medium-Term Evaluation

Dr Emil Varghese

Dr Nirav Lekinwala

Srinivas Sridharan

Kavyashree N Kalkura

Vinod S

Yashwant Pratap Singh Yadav

Dr R Subramanian

Center for Study of Science, Technology and Policy (CSTEP)

April 2026

Edited and Designed by CSTEP

Disclaimer

Every effort has been made to ensure the correctness of data and information used in this document. However, the authors or CSTEP does not accept any legal liability for the accuracy or inferences of the material contained in this document and for any consequences arising from the use of this material.

Generative AI declaration

Generative AI tools were used for language enhancement and for research and analysis. All AI-assisted outputs were reviewed and validated by the authors, who retain full responsibility for the accuracy, originality, and integrity of this document.

Suggested citation: CSTEP. 2026. *Performance evaluation of multi-pollutant air quality sensors at Indi-SET, Bengaluru, India, First edition, A short- and medium-term evaluation*. (CSTEP-RR-2026-02).

April 2026

Editor: Veena P

Designer: Alok Kumar Saha

Bengaluru

No. 18, 10th Cross, Mayura Street

Papanna Layout, Nagashettyhalli

RMV Stage 2, Bengaluru 560094

Karnataka (India)

Noida

1st Floor, Tower-A

Smartworks Corporate Park

Sector 125, Noida 201303

Uttar Pradesh (India)

Tel.: +91 (80) 6690 2500

Email: cpe@cstep.in

Acknowledgements

We are deeply thankful to Google LLC and Coefficient Giving (formerly known as Open Philanthropy) for the financial support to establish the India Air Quality Sensor Evaluation and Training (Indi-SET) facility at the Center for Study of Science, Technology and Policy (CSTEP), Bengaluru, and to carry out this sensor evaluation study. We also acknowledge Rainmatter Foundation for their core support.

We express our gratitude to the sensor manufacturers—Aeron Systems Pvt Ltd, Airveda, Airvoice Global, Aurassure Pvt Ltd, Respirer Living Sciences, and Sensit Technologies Pvt Ltd (the RAMP model was obtained from Pollution Protection Systems Mumbai Pvt Ltd)—for their collaboration in the sensor evaluation process and commitment to publicly sharing their sensor performance data.

We thank our project advisors—Professor Albert Presto (Carnegie Mellon University), Dr TS Panwar (Air Quality Advisor, CSTEP), Divya Narayanan (Jhatkaa.org), and Kartik Kompella (Purposeful Brands)—for their guidance and insights throughout the project. We also acknowledge valuable contributions by Nisar Ali Baig, Srishti, Dr Deepthi Yaparla, and Dr Nidhi Malik at various stages of the project.

We appreciate the continuous support provided by our Artificial Intelligence & Digital Platforms (AI&DP) team, particularly Ramanaidu Venkat, Vedant Kanaji, Rakshita Anand K, Shobitha Cherian, Ashwin K R, and Rajesh Shenoy (Sector Head, AI&DP), in developing the air quality data management system. We are also grateful to Rajeshree Menon (Chief Financial Officer), Manjunatha Eshwar, and our Finance team for their coordination in setting up the Indi-SET facility, as well as to Veena P and the Communications and Policy Engagement team for their meticulous editing and design of this report.

Finally, we express our gratitude for the motivation and guidance provided by the CSTEP leadership team, including Dr Jai Asundi (Executive Director), Munish Sapra (Director - Strategy and Corpus Generation), Dr Piyush Bhardwaj (Research Scientist, Group Head, Air Quality Observations and Modelling), Swagata Dey (Policy Specialist, Group Head, Air Quality Policy and Outreach), Dr Indu K Murthy (interim Sector Head, Air Quality), and Dr Prakash Doraiswamy (Sector Head, Air Quality).

Executive Summary

Air quality sensors, also called low-cost sensors (LCSs) or microsensors, are sensitive to meteorological conditions, ambient pollutant concentrations, and cross-interference from non-target pollutants. Hence, the sensor performance needs to be evaluated against reference monitors and the data corrected before deployment across a city for air pollution monitoring studies. In this context, a sensor evaluation study was conducted at the first-of-its-kind facility in India—the India Air Quality Sensor Evaluation and Training (Indi-SET) facility, at the Center for Study of Science, Technology and Policy's (CSTEP's) Bengaluru campus—prior to deployment of a multi-pollutant sensor network across urban Bengaluru.

The evaluation was performed by collocating all the sensor nodes (or sensor devices), which would be eventually part of the network, with a reference-grade air quality monitoring station (AQMS) at Indi-SET in two phases: (i) a short-term collocation study for 4–6 weeks, from January 20, 2024, to March 19, 2024, and (ii) a medium-term collocation study for ~8 months, from January 20, 2024, to September 30, 2024. This report presents results from 48 multi-pollutant (particulate matter [$PM_{2.5}$ and PM_{10}], nitrogen dioxide [NO_2], ozone [O_3], and carbon monoxide [CO]) sensor nodes from six manufacturers that supply sensors commercially in India.

Evaluation methodology

The sensors evaluated included 25 nodes from Aurassure Pvt Ltd; 5 each from Aeron Systems Pvt Ltd, Airveda, Respirer Living Sciences, and Sensit Technologies Pvt Ltd; and 3 from Airvoice Global. Each node incorporated LCSs that use light-scattering principles and electrochemical reactions to measure PM and gaseous pollutants, respectively.

During the short-term collocation, the intra-model precision (precision among sensor nodes from the same manufacturer) and the accuracy of sensor nodes relative to reference instruments were evaluated. The Pearson correlation coefficient (r) was used to assess the intra-model precision of the sensor nodes for each pollutant. The accuracy of manufacturer-reported sensor data relative to the AQMS data (reference instrument data) was evaluated using the Pearson correlation coefficient (r), coefficient of determination (R^2), mean absolute error (MAE), coefficient of variation in MAE (CvMAE), and root mean square error (RMSE). During the medium-term evaluation, all the nodes were assessed for sensor health.

Evaluation results of manufacturer-reported data

Overall, the initial evaluation indicated that manufacturer-reported sensor data can vary substantially between devices and across manufacturers, particularly when compared with data from reference instruments. To measure $PM_{2.5}$ and PM_{10}, the sensor nodes from Aeron, Respirer, and Sensit were integrated with two models of PM sensors (optical particle counter–based [OPC-based] and non-OPC-based), while the sensor nodes from Airveda, Airvoice, and Aurassure were integrated with a single model of PM sensor (non-OPC-based). In the intra-model precision assessment, Sensit demonstrated the highest precision across both $PM_{2.5}$ and PM_{10}. None of the sensor manufacturers met the precision targets for NO_2 and O_3. However, Aeron, Airveda, and Airvoice achieved acceptable precision for CO across nodes. When comparing sensor-reported data with reference-instrument data across multiple performance metrics, no manufacturer met all criteria for all pollutants. Suboptimal sensor performance in manufacturer-reported data may be due to manufacturers developing correction factors under different environmental conditions and/or at different pollution levels than the evaluation/deployment site conditions.

In the medium-term evaluation, several issues were identified with the EC Sense gas sensors integrated into Aurassure nodes. The NO_2 sensors consistently capped their maximum readings at 17 ppb, while most O_3 sensors reported fixed values of either 2 or 4 ppb for extended periods. Similarly, most CO sensors recorded 0 ppm throughout. In addition, within 6 months of use, many PM sensors from different manufacturers stopped functioning. These findings highlight the importance of periodic testing to ensure reliable sensor performance during use.

Improving sensor data quality with localised calibration models

To improve the sensor performance (before deployment across Bengaluru), localised calibration models were developed using 80% of the data from the short-term collocation period. Two types of correction models were developed: sensor-wise calibration models and generalised calibration models. The sensor-wise calibration model was built for each sensor node using data from that node. In contrast, the generalised calibration model was built as a common model for all sensor nodes from a particular manufacturer, using their median sensor response. After developing the models, the best-performing model for each pollutant/manufacturer was identified by comparing the held-out 20% data of the collocation period with the corresponding AQMS data.

A generalised model was adopted for the Aurassure sensor nodes, as we had 25 nodes with moderate precision. Due to the smaller number of devices (three to five nodes) from Aeron, Airveda, Respirer, and Sensit, sensor-wise models were selected, as a generalised model would be less representative and could perform worse. These correction models were particularly effective for NO_2 and O_3, with correlation coefficients (r) between the corrected sensor data and the AQMS reference data exceeding 0.7, indicating strong agreement with the reference monitors.

Recommendations for sensor-based air quality monitoring

Based on the short- and medium-term assessments of the 48 sensor nodes, the following recommendations are proposed.

- **Develop site-specific calibration models**

It is recommended that the raw sensor data (PM mass in µg/m^3 derived from nephelometers or OPCs for PM sensors or voltage signals from electrochemical gas sensors) be evaluated and that site-specific calibration models be developed to account for local environmental conditions. To address sensor limitations related to meteorological influences and pollutant cross-sensitivities, these parameters should also be included as inputs during model development.

- **Establish more calibration facilities across India**

India's unique climatic conditions, varied land-use patterns, and differing pollution loads necessitate regional, if not localised, calibration approaches. To support manufacturers in developing region-appropriate calibration models, additional calibration facilities should be established across India's diverse regions.

- **Assess long-term sensor and calibration model performance**

Sensor performance degrades over time due to ageing, and calibration models trained on data from a specific season may fail when exposed to out-of-range conditions. Therefore, long-term performance evaluation is essential for both the sensors and their calibration models. It is recommended that each sensor node undergo periodic performance checks every 3–6 months by collocating with reference instruments and assessing both raw and calibrated data.

Table of Contents

Figures

Tables

1. Introduction

Traditionally, air quality measurements are performed using in-situ reference-grade instruments and remote sensing. Such measurements reveal the current state of air quality, show how it evolves over time, and assess the impact of mitigation measures. However, over the last decade, sensor-based monitoring using low-cost sensors (LCSs) has been increasingly used to complement these measurements and to identify hyperlocal hotspots across several cities worldwide. In practice, most sensor networks only focus on particulate matter (PM) measurements. Therefore, this study deployed a network of 40 multi-pollutant sensor nodes integrated with different sensor models to measure both gaseous (nitrogen dioxide [NO_2] and ozone [O_3]) and particulate ($PM_{2.5}$ and PM_{10}) air pollutants in Bengaluru.

Further, LCSs are sensitive to meteorological conditions, ambient pollutant concentrations, and cross-interference from other pollutants; therefore, they need to be evaluated and corrected before deployment (Malings, et al., 2019a; Malings., et al., 2019b). To guide such evaluations, the United States Environmental Protection Agency (US EPA) has provided various protocols for evaluating the performance of LCSs (EPA/600/R-20/279, 2021; EPA/600/R-20/280, 2021; EPA/600/R-23/145, 2023; EPA/600/R-23/146, 2024). Complementing these protocol-based efforts, several sensor evaluation facilities operate globally to conduct independent testing of air quality sensors. These include the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) and the Air Quality Sensor Evaluation and Training Centre for West Africa (Afri-SET).

In the current study, 48 sensor nodes were evaluated at the first-of-its-kind facility in India–the India Air Quality Sensor Evaluation and Training (Indi-SET) facility at the Center for Study of Science, Technology and Policy's (CSTEP's) Bengaluru campus. The sensor nodes that were evaluated included 25 from Aurassure Pvt Ltd (purchased by the funder), 3–5 each from five manufacturers (Aeron Systems Pvt Ltd, Airveda, Airvoice Global, Respirer Living Sciences, and Sensit Technologies Pvt Ltd; whose performance is not publicly available, especially in the Indian context), and 3 from Airvoice Global (who voluntarily participated in the initial evaluation period). An initial 6-week collocation of the 48 sensor nodes with reference-grade instruments was conducted to evaluate sensor performance across different manufacturers. Following this initial collocation, one node from each of the five manufacturers (Aeron, Airveda, Aurassure, Respirer, and Sensit) was retained at the collocation site for long-term performance evaluation. The 3 nodes from Airvoice were returned after evaluation, and the remaining 40 sensor nodes from the 5 manufacturers were deployed across Bengaluru.

This report compiles the results of the initial evaluation of 48 multi-pollutant sensor nodes measuring NO_2, O_3, CO, $PM_{2.5}$, and PM_{10}. The sensor performance results for SO_2 are not provided as the levels of this pollutant at the test facility were far below the sensor detection limits. During the preliminary analysis, it was observed that the manufacturer-provided data were of inferior quality; therefore, new calibration models were developed by CSTEP for each sensor node by using the collocation data set. Hence, the improvement in performance from the localised collocation (as compared with the manufacturer-calibrated data) is also reported here. Further, the preliminary results of the medium-term performance of these sensors are presented.

2. Multi-pollutant Sensor Calibration Facility at CSTEP

The Indi-SET facility established at the CSTEP Bengaluru campus (Figure 1) is approximately 10 meters above the ground level and 150 meters from a major road. Indi-SET houses a real-time air quality monitoring station (AQMS) equipped with reference-grade instruments that monitor multiple air pollutants, including PM ($PM_{2.5}$ and PM_{10}) and key pollutant gases (O_3, CO, NO_2, and SO_2). The facility also features a weather station that measures temperature, relative humidity, ambient pressure, and precipitation. To evaluate the performance of sensor nodes from various manufacturers and develop calibration or correction models, 48 multi-pollutant LCS nodes were collocated with the AQMS at Indi-SET for 4–6 weeks, from January 20, 2024, to March 19, 2024.

***Figure 1**: Indi-SET facility established at CSTEP.*

(a) Gas analysers for O_3, NO_2, SO_2, and CO are installed, operated, and maintained in an air-conditioned room. The inlets to the gas analysers are in ambient conditions.

(b) The PM monitors for $PM_{2.5}$ and PM_{10} are placed outdoors in weatherproof enclosures next to the sensor test bed that accommodates up to 60 sensor nodes.

2.1. Instrumentation

The AQMS at Indi-SET utilises US-EPA-certified reference-grade monitors to measure criteria air pollutants. These monitors were supplied and installed by Vasthi Instruments Pvt Ltd and are operated and maintained by CSTEP. The gas analysers were installed in an air-conditioned room, as shown in Figure 1(a), while the PM monitors were installed outdoors, as shown in Figure 1(b). To measure ambient PM_{10}, the Vasthi Instruments' Vair-9009 PM_{10} ambient PM monitor was used. For ambient $PM_{2.5}$, the same model coupled with a $PM_{2.5}$ cyclone was employed. Two stainless steel sampling lines were used for PM sampling, with inlets placed 1 meter apart. A separate inlet and Teflon sampling line, located 10 meters from the PM monitors, were used to sample pollutant gases. The gases SO_2, NO_x, CO, and O_3 were measured using Mezus 110, 210, 310, and 410 analysers (Kentek Co. Ltd), respectively. Instrument specifications are provided in Annexure A.

2.2. Operation and maintenance of the AQMS

The temperature of the room housing the reference gas analysers was maintained at 24–26 °C throughout the study period. Data from the AQMS were reported at 1-minute averages for gas analysers and 1-hour averages for PM monitors. A rigorous operation and maintenance protocol (Annexure A) was developed to ensure the quality and reliability of the AQMS data. This protocol adhered to the guidelines and recommendations outlined by the US EPA in the Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II (EPA-454/B-17-001), as well as to the standards set by the Central Pollution Control Board (CPCB) in India.

2.3. Ambient air quality during the evaluation period

During the 6-week collocation period, the ambient temperature varied from 17.4 °C to 39.0 °C, averaging 26.9 °C, while the relative humidity varied between 13.3% and 93.8%, with a mean of 49.7%. Table 1 presents the observed range of pollutant concentrations measured by the reference-grade instruments during the collocation period at the Indi-SET facility.

***Table 1**: Pollutant concentration range observed in reference-grade instruments during the 6-week collocation period*

Pollutant	Average Interval	Min.	Max.	Mean
O_3 (ppb)	15 minute	0.50	105.10	36.94
NO_2 (ppb)	15 minute	0.00	120.34	10.95
$PM_{2.5}$ (µg/m^3)	1 hour	8.80	329.94	66.37
PM_{10} (µg/m^3)	1 hour	33.00	391.50	82.09

3. Performance Evaluation of the Sensors

In the current evaluation, 43 sensors from 5 different Indian manufacturers (Aeron, Airveda, Airvoice, Aurassure, and Respirer) were assessed over a period of 6 weeks, from January 20, 2024, to March 19, 2024, and 5 sensors from Sensit were evaluated from February 21, 2024, to March 19, 2024 (due to a delay in shipment). The evaluation results for SO_2 sensor performance are not provided as the levels of this pollutant at the test facility were far below the sensor detection limits.

Further, the reference analyser used for CO measurements had an operating range of 0–100 ppm, resulting in limited sensitivity at the low ambient CO concentrations typically observed in Bengaluru. Consequently, while CO sensors were included in the evaluation, their data were not used to assess accuracy against the reference instrument. Instead, CO sensor measurements were used to develop calibration models for NO_2 and O_3 sensors, and the precision of CO sensors within each manufacturer was evaluated. Due to the limitations of the reference analyser, an accuracy assessment for CO sensors is not included in this report. For more recent evaluations at Indi-SET, a different reference instrument with an operating range of 0–10 ppm, which is better suited to Bengaluru background concentrations, is being used and will be reported in a future report.

3.1. Sensor specifications

Each sensor node had varying configurations for sensing and reporting the multi-pollutant concentrations. Table 2 outlines the specifications of the various sensor nodes evaluated. All gas sensors in this study are electrochemical sensors, either from EC Sense (integrated in Aurassure) or Alphasense (integrated in all the other manufacturers). PM sensors employ light scattering techniques.

Table 2: *Specifications of the sensor nodes as reported by the manufacturer*

Manufacturer and Sensor Node Model	Pollutant	Sensor Model Integrated	Working Principle	Measurement Range
Aeron - LAMINAR AQM21	PM_{10}	Alphasense OPC-N3	Electrochemical reaction	0–2,000 µg/m^3
	$PM_{2.5}$	Alphasense OPC-N3		0–2,000 µg/m^3
	PM_{10}	Tera Sensor NextPM	Light scattering	0–2,000 µg/m^3
	$PM_{2.5}$	Tera Sensor NextPM		0–2,000 µg/m^3
	PM_1	Tera Sensor NextPM		0–2,000 µg/m^3
	CO	Alphasense CO-B4	Electrochemical reaction	0–1,000 ppm
	NO_2	Alphasense NO2-B43F		0–20 ppm
	SO_2	Alphasense SO2-B4		0–100 ppm
	O_3	Alphasense OX-B431		0–20 ppm
Airveda - PM2510TH-SNOZCO-WPGSM-EYE	PM_{10}	Plantower PMS7003	Light scattering	0–1,999 µg/m^3
	$PM_{2.5}$	Plantower PMS7004		0–999 µg/m^3
	CO	Alphasense CO-A4F	Electrochemical reaction	0–9,999 ppm
	SO_2	Alphasense SO2-A4F		0–999 ppb
	NO_2	Alphasense NO2-A43F		0–9,999 ppb
	O_3	Alphasense OX-A431		0–999 ppb
Aurassure Infra	PM_{10}	Sensirion SPS30	Light scattering	0–1,000 µg/m^3
	$PM_{2.5}$	Sensirion SPS30		0–1,000 µg/m^3
	CO	EC Sense TB600	Electrochemical reaction	0–10 ppm
	SO_2	EC Sense TB600		0–5 ppm
	NO_2	EC Sense TB600		0–2 ppm
	O_3	EC Sense TB600		0–5 ppm

Manufacturer and Sensor Node Model	Pollutant	Sensor Model Integrated	Working Principle	Measurement Range
Respirer Atmos	PM	Alphasense OPC-R2	Light scattering	0–2,000 µg/m^3
		Plantower PMS7003		0–1,999 µg/m^3
	CO	Alphasense CO-B4	Electrochemical reaction	0–1,000 ppm
	SO_2	Alphasense SO2-B4		0–100 ppm
	NO_2	Alphasense NO2-B43F		0–20 ppm
	O3	Alphasense OX-B431		0–20 ppm
Sensit RAMP	PM_1	Alphasense OPC-R2	Light scattering	0–2,000 µg/m^3
	$PM_{2.5}$	Alphasense OPC-R2		0–2,000 µg/m^3
	PM_{10}	Alphasense OPC-R2		0–2,000 µg/m^3
	PM_1	Plantower PMS5003T		0–1,000 µg/m^3
	$PM_{2.5}$	Plantower PMS5003T		0–1,000 µg/m^3
	PM_{10}	Plantower PMS5003T		0–1,000 µg/m^3
	CO	Alphasense CO-B4	Electrochemical reaction	0–1,000 ppm
	SO_2	Alphasense SO2-B4		0–100 ppm
	NO_2	Alphasense NO2-B43F		0–20 ppm
	O_3	Alphasense OX-B431		0–20 ppm
Airvoice	$PM_{2.5}$	Plantower	Light scattering	0–1,000 µg/m^3
	PM_{10}	Plantower		0–1,000 µg/m^3
	CO	Alphasense	Electrochemical reaction	0–55,000 µg/m^3
	SO_2	Alphasense		0–8,000 µg/m^3
	NO_2	Alphasense		0–5,000 µg/m^3
	O_3	Alphasense		0–4,000 µg/m^3

Gas sensors

Alphasense gas sensors: This sensor model operates in the amperometric mode with three solid electrodes and sulfuric acid (H_2SO_4, concentration range 3 to 7 molar) as the electrolyte. The auxiliary electrode of the sensor is not exposed to the target analyte gas, while the working electrode is gas sensitive. The counter electrode balances the oxidation or reduction reaction of the working electrode and the overall cell reaction. The corrected signal response to the target gas concentration is derived from the difference between the signals from the auxiliary electrode and the working electrode (Alphasense Application Note: AAN 104; Baron and Saffell, 2017).

EC Sense gas sensors: This sensor model uses solid polymer electrochemical technology. The oxidation of the gas at the working electrode generates a current proportional to the gas concentration (EC Sense Ozone Gas Sensor Module TB600B Datasheet).

PM sensors

Laser scattering method

- **Nephelometric method:** Utilised by Plantower, Tera Sensor, and Sensirion. This method is less effective for PM_{10} and larger particles due to truncation of the forward-scattering coefficient and the potential inability to aspirate larger particles into the device (Ouimette, et al., 2022).
- **Optical particle counter (OPC):** Used by Alphasense OPC. It operates at a high flow rate and allows particle counting in 24-size bins ranging from 0.35 to 40 μm. This method behaves similarly to aerosol spectrometers and effectively measures dust events or PM when the ratio of coarse PM ($PM_{2.5–10}$) to fine PM ($PM_{2.5}$) is higher (Kaur and Kelly, 2023).

The following sections provide detailed performance evaluation results of the 48 sensor nodes from six manufacturers, integrated with the various sensor models for monitoring PM and gaseous pollutants.

3.2. Performance metrics

Several statistical metrics were used to evaluate sensor performance. These metrics help in understanding the accuracy and reliability of the sensor measurements compared with the reference instruments.

Linearity: The linearity of the sensor-reported values compared with the AQMS-reported values was assessed using the Pearson correlation coefficient (r) and the coefficient of determination (R^2). The Pearson correlation coefficient (r) measures the strength and direction of the linear relationship between the sensor and AQMS values. The coefficient of determination (R^2) indicates the proportion of variance in the sensor data that the AQMS data can explain.

Error metrics: The error in the sensor measurements relative to that in the reference-grade instruments was quantified using the mean absolute error (MAE), the root mean square error (RMSE), and the coefficient of variation in MAE (CvMAE).

- **Mean absolute error:** $\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y}_i|$

where y_i represents the AQMS-reported value, $\widehat{y}_i$ represents the sensor value, and n is the number of observations. This metric provides the average magnitude of errors in a set of predictions, without considering their direction.

- **Coefficient of variation in MAE:** $\text{CvMAE} = \frac{\text{MAE}}{\sum_{i=1}^{n} y_i}$

This metric gives the MAE divided by the mean of the reference values.

- **Root mean squared error:** $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}$

This metric measures the square root of the average squared differences between sensor- and AQMS-reported values, providing insight into the magnitude of errors, with a greater penalty for larger errors.

3.3. Testing conditions

The manufacturers provided data as either 'raw' or 'corrected' values. Raw values refer to particle counts or electrochemical sensor response (mV). In this study, we use only the electrochemical sensor response for gases. The corrected values are in PM mass concentration (µg/m^3) or gas mixing ratio (ppb or ppm) and are based on a factory calibration applied by the manufacturer; here, we call them 'manufacturer-calibrated values'. These are the values reported by the manufacturer.

The performance evaluation of the air quality sensors was conducted in real-world field conditions over 4–6 weeks, from January 20, 2024, to March 19, 2024. CO and SO_2 sensors were not evaluated due to problems with the reference instrument, which limits the usefulness of the data set for this period. To ensure accurate comparisons between sensor data and reference-grade instrument data at the AQMS, sensor data from six manufacturers were standardised to Indian Standard Time (IST).

Data averaging: To minimise the influence of short-term spikes and better represent ambient air quality, the sensor and reference data were averaged prior to comparison. The starting time of the averaging process was used as a time stamp. For example, the average PM data logged between 10:00 AM and 11:00 AM was reported as the PM mass concentration at 10:00 AM. An average was only reported if at least 75% of the data points were available for the aggregation period. Data were averaged over 15-minute intervals for gas measurements as the reference monitor reports at 1-minute intervals. This approach avoids the influence of short-term spikes, allowing for the capture of dynamic changes in gas concentrations and providing a more granular assessment of sensor performance. For particulate measurements, data were averaged over 1-hour intervals, as the reference monitors reported data at hourly intervals.

3.4. Intra-model precision

Intra-model precision among devices from the six manufacturers was assessed by evaluating the correlation of sensors to each other within each manufacturer using the Pearson correlation coefficient (r). Annexure B presents the correlation matrix plots for each device for each pollutant. The manufacturer-specific concentration ranges reported by each sensor during the initial evaluation period are provided in Annexure C. These comparisons were used to evaluate the intra-model precision among sensors.

The manufacturers whose sensors showed a Pearson correlation coefficient (r) greater than 0.9 (when comparing all sensors) were classified as having high precision, indicating strong agreement among their sensors. Sensors from manufacturers with correlation values between 0.7 and 0.9 were classified as having moderate precision, while those with r values below 0.7 were considered to have poor precision. For the manufacturer-corrected values, a negative correlation was interpreted as very poor performance.

In contrast, for the raw sensor values, the absolute Pearson correlation (|r|)[1] was used. Table 3 summarises the intra-model precision range among sensors from the same manufacturer.

Table 3*[2]*: Summary of intra-model precision: Pearson correlation (r) for raw and manufacturer-calibrated (mfr cal) values among the nodes from the same manufacturer

Manufacturer	Aeron	Airveda	Airvoice	Aurassure	Respirer	Sensit
$PM_{2.5}$ raw	-	0.84 to 0.99	-	-	-	-
$PM_{2.5}$ mfr cal	0.93 to 1.00	0.92 to 0.99	0.97 to 0.99	0.06 to 1.00	0.99 to 1.00	1
OPC $PM_{2.5}$ raw	-	-	-	-	-	-
OPC $PM_{2.5}$ mfr cal	0.77 to 1.00	-	-	-	0.88 to 0.99	0.98 to 0.99
PM_{10} raw	-	0.96 to 0.99	-	-	-	-
PM_{10} mfr cal	0.95 to 1.00	-	0.96 to 0.99	0.10 to 1.00	0.99 to 10.00	1
OPC PM_{10} raw	-	-	-	-	-	-
OPC PM_{10} mfr cal	0.70 to 0.99	-	-	-	0.83 to 0.99	0.97 to 0.99
NO_2 raw	0.68 to 0.97	0.08 to 0.94	0.80 to 0.98	-	0.50 to 0.96	0.94 to 0.99
NO_2 mfr cal	−0.25 to 0.98	0.84 to 0.98	0.80 to 0.98	0.73 to 0.98	-	0.86 to 0.99

1 Raw data from the electrochemical gas sensors are defined as the voltage difference between the working electrode, which responds to the target gas, and the auxiliary electrode, which is not exposed to pollutants. However, we have noticed that some manufacturers interchange the working and auxiliary electrode assignments, even within the same sensor batch. Therefore, raw data are evaluated using Pearson correlation without regard to response direction. For consistency, the same approach is applied to the raw values of PM sensors.

2 In Table 3, - denotes unavailability of the mentioned sensor data from the manufacturers for evaluation.

Manufacturer	Aeron	Airveda	Airvoice	Aurassure	Respirer	Sensit
O_3 raw	0.18 to 0.66	0.18 to 0.70	0.25 to 0.69	-	0.85 to 0.98	0.20 to 0.98
O_3 mfr cal	0.52 to 0.99	0.76 to 0.99	−0.69 to 0.65	0.10 to 0.96	-	0.40 to 0.92
CO raw	0.98 to 1.00	0.99 to 1.00	0.94 to 0.99	-	0.99 to 1.00	0.98 to 1.00
CO mfr cal	0.91 to 0.99	0.96 to 0.98	0.94 to 0.100	−0.13 to 0.99	-	0.74 to 0.99

3.5. Evaluating the accuracy of manufacturer-calibrated pollutant values against the AQMS values

The metrics used for this evaluation included the Pearson correlation coefficient (r), the coefficient of determination (R^2), MAE, the coefficient of variation in MAE (CvMAE), and RMSE. Scatter plots illustrating the performance of each sensor against the AQMS-reported values for various pollutants are provided in Annexure D. Table 4 provides a detailed overview of the performance metric ranges for $PM_{2.5}$, PM_{10}, NO_2, and O_3 for both raw and manufacturer-corrected (or factory-calibrated) values, compared with AQMS-reported values.

Table 4[3]: Evaluation of raw and manufacturer-calibrated values against the AQMS reference data

Pollutant	Sensor	r		R^2		MAE	RMSE
		Raw	Manufacturer-Calibrated	Raw	Manufacturer-Calibrated	Manufacturer-Calibrated	Manufacturer-Calibrated
$PM_{2.5}$ ($\mu g/m^3$)	Aeron	NA	0.70 to 0.77 (Tera Sensor)	NA	0.49 to 0.60 (Tera Sensor)	8.44 to 10.25 (Tera Sensor)	13.83 to 15.62 (Tera Sensor)
		NA	0.47 to 0.57 (OPC)	NA	0.22 to 0.33 (OPC)	12.36 to 13.87 (OPC)	17.87 to 19.55 (OPC)
	Airveda	NA	0.63 to 0.75	NA	0.40 to 0.56	8.92 to 10.97	14.21 to 16.50
	Airvoice	NA	0.76 to 0.77	NA	0.58 to 0.59	6.28 to 6.75	9.70 to 10.67
	Aurassure	NA	0.22 to 0.74	NA	0.07 to 0.56	9.71 to 15.17	14.65 to 21.65
	Respirer	NA	0.72 to 0.74 (Plantower)	NA	0.52 to 0.55 (Plantower)	9.15 to 9.36 (Plantower)	14.43 to 14.94 (Plantower)
		NA	0.57 to 0.67 (OPC)	NA	0.33 to 0.45 (OPC)	10.63 to 12.11 (OPC)	15.97 to 17.72 (OPC)
	Sensit	NA	0.81 to 0.86 (Plantower)	NA	0.65 to 0.75 (Plantower)	7.86 to 8.87 (Plantower)	10.29 to 12.32 (Plantower)
		NA	0.64 to 0.74 (OPC)	NA	0.40 to 0.55 (OPC)	10.31 to 11.58 (OPC)	13.76 to 16.26 (OPC)

3 In Table 4, 'NA' denotes 'Not Available'. The evaluation metrics are not available as the mentioned sensor data were not available from the manufacturer.
AI tools were used to improve code for data analysis and visualisation and to help fix errors and bugs. The AI-generated code was thoroughly validated by the authors.

Pollutant	Sensor	r		R^2		MAE	RMSE
		Raw	Manufacturer-Calibrated	Raw	Manufacturer-Calibrated	Manufacturer-Calibrated	Manufacturer-Calibrated
PM_{10} (µg/m³)	Aeron	NA	0.61 to 0.70 (Tera Sensor)	NA	0.38 to 0.49 (Tera Sensor)	16.77 to 18.99 (Tera Sensor)	26.12 to 28.77 (Tera Sensor)
		NA	0.50 to 0.62 (OPC)	NA	0.25 to 0.38 (OPC)	18.95 to 22.05 (OPC)	28.41 to 32.36 (OPC)
	Airveda	NA	0.55 to 0.59	NA	0.30 to 0.35	19.04 to 19.76	29.26 to 30.26
	Airvoice	NA	0.64 to 0.67	NA	0.41 to 0.45	9.91 to 10.43	13.56 to 14.90
	Aurassure	NA	0.08 to 0.52	NA	0.01 to 0.31	23.35 to 28.16	32.68 to 39.52
	Respirer	NA	0.55 to 0.57 (Plantower)	NA	0.30 to 0.32 (Plantower)	19.00 to 19.30 (Plantower)	29.71 to 30.28 (Plantower)
		NA	0.56 to 0.70 (OPC)	NA	0.32 to 0.49 (OPC)	16.58 to 19.79 (OPC)	25.94 to 29.82 (OPC)
	Sensit	NA	0.61 to 0.68 (Plantower)	NA	0.37 to 0.46 (Plantower)	16.59 to 18.76 (Plantower)	22.21 to 26.17 (Plantower)
		NA	0.60 to 0.69 (OPC)	NA	0.37 to 0.48 (OPC)	16.20 to 18.50 (OPC)	21.80 to 26.33 (OPC)

Pollutant	Sensor	r		R^2		MAE	RMSE
		Raw	Manufacturer-Calibrated	Raw	Manufacturer-Calibrated	Manufacturer-Calibrated	Manufacturer-Calibrated
NO_2 (ppb)	Aeron	0.69 to 0.87	0.49 to 0.68	0.47 to 0.75	0.24 to 0.46	7.01 to 7.92	10.03 to 11.48
	Airveda	0.06 to 0.71	−0.33 to −0.26	0 to 0.50	0.07 to 0.11	7.84 to 8.74	12.95 to 14.38
	Airvoice	0.32 to 0.39	−0.07 to 0.07	0.10 to 0.16	0	43.16 to 81.83	64.30 to 101.06
	Aurassure	NA	0.09 to 0.56	NA	0.01 to 0.31	8.01 to 8.96	11.15 to 13.77
	Respirer	0.52 to 0.79	NA	0.27 to 0.62	NA	NA	NA
	Sensit	0.61 to 0.71	0.41 to 0.62	0.37 to 0.51	0.17 to 0.38	6.60 to 7.52	10.57 to 12.52
O_3 (ppb)	Aeron	0.03 to 0.55	0.66 to 0.87	0 to 0.31	0.44 to 0.76	6.18 to 8.37	8.10 to 12.69
	Airveda	0.27 to 0.68	0.7 to 0.84	0 to 0.46	0.48 to 0.71	6.93 to 9.30	9.15 to 12.54
	Airvoice	0.54 to 0.68	−0.76 to 0.41	0.29 to 0.36	0.17 to 0.57	4.74 to 5.37	6.04 to 6.67
	Aurassure	NA	−0.15 to 0.73	NA	0 to 0.54	8.85 to 14.39	12.04 to 18.46
	Respirer	0.36 to 0.73	NA	0.13 to 0.53	NA	NA	NA
	Sensit	0.21 to 0.75	0.49 to 0.88	0 to 0.56	0.24 to 0.77	6.60 to 14.54	9.27 to 17.37

3.6. Need for localised collocation and calibration

The observations during the collocation period (evaluation of precision among devices and accuracy relative to the AQMS) revealed that the manufacturer-calibrated sensor values vary within devices and manufacturers. This could be because the correction factors applied to each sensor by the manufacturer were developed under different environmental conditions and exposed to a different source mix, or even the use of correction algorithms that may not transfer well to a new city. Previous studies have shown that the performance of PM and electrochemical gas sensors is affected by temperature, relative humidity, and cross-sensitivity from other gases (Zimmerman, et al., 2018). Hence, to improve the sensor performance for deployment in Bengaluru, a localised calibration was conducted.

4. Sensor Calibration

To improve data reliability and accuracy, sensor data collected during the collocation period from January 20, 2024, to March 19, 2024, were used to develop localised calibration and correction models. For the manufacturer Sensit, the collocation period spanned from February 21, 2024, to March 19, 2024, due to a shipment delay. Therefore, the calibration models were trained using sensor data and AQMS data from the collocation period, incorporating various model inputs to account for these factors. Table 5 details the model inputs for each device to build correction models.

***Table 5**: Model inputs used to build various calibration models*

Vendor	Model Name	Model Input
Aeron	PM2.5	PM2.5_tera,RH_sensor
	PM2.5_hybrid	PM1_tera,(PM2.5 – PM1)_OPC,RH_sensor
	PM2.5_OPC	PM2.5_OPC,RH_sensor
	PM10	PM10_tera,RH_sensor
	PM10_hybrid	PM1_tera,(PM2.5 – PM1)_OPC,(PM10 – PM2.5)_OPC,RH_sensor
	PM10_OPC	PM10_OPC,RH_sensor
	NO2	NO2_netvolt_sensor,O3_netvolt_sensor,CO_netvolt_sensor, AT_sensor
	O3	NO2_netvolt_sensor,O3_netvolt_sensor,CO_netvolt_sensor, AT_sensor
Airveda	PM2.5	PM2.5_plantower,RH_sensor
	PM10	PM10_plantower,RH_sensor
	NO2	NO2_netvolt_sensor,O3_netvolt_sensor,CO_netvolt_sensor, AT_sensor
	O3	NO2_netvolt_sensor, O3_netvolt_sensor,CO_netvolt_sensor, AT_sensor
Aurassure	PM2.5	PM2.5_sensirion,RH_sensor
	PM10	PM10_sensirion,RH_sensor
	NO2	NO2_sensor,O3_sensor,CO_sensor,AT_sensor
	O3	NO2_sensor,O3_sensor,CO_sensor,AT_sensor

Vendor	Model Name	Model Input
Respirer	PM2.5	PM2.5_plantower,RH_sensor
	PM2.5_OPC	PM2.5_OPC,RH_sensor
	PM10	PM10_plantower,RH_sensor
	PM10_hybrid	PM2.5_plantower,(PM10 – PM2.5)_OPC,RH_sensor
	PM10_OPC	PM10_OPC,RH_sensor
	NO2	NO2_netvolt_sensor,O3_netvolt_sensor,CO_netvolt_sensor, AT_sensor
	O3	NO2_netvolt_sensor, O3_netvolt_sensor,CO_netvolt_sensor,AT_sensor
Sensit	PM2.5	PM2.5_plantower,RH_sensor
	PM2.5_hybrid	PM1_plantower,(PM2.5 – PM1)_OPC,RH_sensor
	PM2.5_OPC	PM2.5_OPC,RH_sensor
	PM10	PM10_plantower,RH_sensor
	PM10_hybrid	PM1_plantower, (PM2.5 – PM1)_OPC,(PM10 – PM2.5)_OPC,RH_sensor
	PM10_OPC	PM10_OPC,RH_sensor
	NO2	NO2_netvolt_sensor,O3_netvolt_sensor,CO_netvolt_sensor,AT_sensor
	O3	NO2_netvolt_sensor,O3_netvolt_sensor,CO_netvolt_sensor,AT_sensor

RH_sensor: *Relative humidity reported by the sensor*
AT_sensor: *Ambient temperature reported by the sensor*
PM1_plantower/tera/sensirion: *PM_1 value reported by Plantower or similar type sensors such as Tera Sensor or Sensirion*
PM2.5_plantower/tera/sensirion: *$PM_{2.5}$ value reported by Plantower or similar type sensors such as Tera Sensor or Sensirion*
PM10_plantower/tera/sensirion: *PM_{10} value reported by Plantower or similar type sensors such as Tera Sensor or Sensirion*
PM2.5_OPC: *$PM_{2.5}$ value reported by Alphasense OPC*
PM10_OPC: *PM_{10} value reported by Alphasense OPC*
(PM2.5 – PM1)_OPC: *The difference between $PM_{2.5}$ and PM_1 value reported by Alphasense OPC sensors*
(PM10 – PM2.5)_OPC: *The difference between PM_{10} and $PM_{2.5}$ value reported by Alphasense OPC sensors*
Gas_netvolt_sensor: *Difference of auxiliary and working electrode signals reported by the Alphasense electrochemical gas sensor*
Gas_sensor: *The mixing ratio reported by the EC Sense gas sensor*

4.1. Development of calibration models

Various calibration models were built using the model inputs listed in Table 5. Two types of calibration models were developed using two methods: sensor-wise calibration models and generalised calibration models. A sensor-wise calibration model was built for each sensor node using data from that node. For sensor-wise calibration models, the training data consisted of 80% of the collocation data, selected randomly. The model was tested on the remaining 20% data. The generalised calibration model was built as a common model for sensor nodes from each manufacturer. This was built by considering the median values measured across all sensor nodes of that manufacturer at the same time stamp. For generalised models, the training data included 100% of the median data set of the devices. The testing was performed on each sensor node using 100% of the data from that node.

Five algorithms were used for each of the above-mentioned model methods: multiple linear regression [LR], quadratic regression [QR], support vector regression [SVR], random forest regression [RFR], and XGBoost regression [XGB]. Each algorithm has its advantages and disadvantages, briefly discussed in Annexure E. A grid search approach was employed to test parameters and calculate R^2, facilitating the identification of optimal model parameters for each algorithm and model method. In the QR model, the model features were further refined by eliminating features with p-values > 0.05.

To evaluate model performance, various metrics were calculated, including the Pearson correlation coefficient (r), bias-corrected MAE (b-MAE), and bias-corrected RMSE (b-RMSE). The following expressions were used to compute these metrics:

Bias: $$\text{bias} = \frac{1}{n}\sum_{i=1}^{n}\widehat{y}_i - \frac{1}{n}\sum_{i=1}^{n} y_i$$

Bias-corrected predictions: $$\widehat{y}_i^{\text{corrected}} = \widehat{y}_i - \text{bias}$$

b-MAE: $$\text{b-MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y}_i^{\text{corrected}}\right|$$

b-RMSE: $$\text{b-RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i^{\text{corrected}}\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i^{\text{corrected}}\right)\right)^2}$$

Models from different categories were ranked using a weighted combination of these metrics. The following expression is used to obtain a ranking score:

$$\text{Score} = 0.25\left[\frac{b\text{-MAE}_{m,s} - \min(b\text{-MAE}_s)}{\min(b\text{-MAE}_s)}\right] + 0.25\left[\frac{b\text{-RMSE}_{m,s} - \min(b\text{-RMSE}_s)}{\min(b\text{-RMSE}_s)}\right] + 0.5\left[\frac{\max(\text{R}_s) - \text{R}_{m,s}}{\max(\text{R}_s)}\right]$$

A model with a score of zero indicated the best-performing model, with low error and high correlation between AQMS-measured and model-predicted pollutant concentrations. Models with a difference of less than 10% (0.1) were considered similar.

Table 6 lists the selected models for each device on the basis of these metrics. A generalised model was adopted for Aurassure, which has 25 devices with moderate precision. For Aeron, Airveda, Respirer, and Sensit, sensor-wise models were selected due to the smaller number of devices (three to five nodes) per manufacturer, making a generalised model less representative.

***Table 6**: Calibration models selected for the various sensor nodes*

Manufacturer	Device	$PM_{2.5}$	PM_{10}	O_3	NO_2
Aurassure	12031	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12032	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12033	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12034	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12036	Sensor-wise SVR	Sensor-wise SVR	Generalised RFR	Generalised XGB
Aurassure	12037	Generalised QR	Generalised QR	Generalised RFR	Sensor-wise XGB
Aurassure	12048	Generalised QR	Generalised QR	Generalised RFR	Sensor-wise XGB
Aurassure	12058	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12062	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12065	Sensor-wise SVR	Sensor-wise SVR	Generalised RFR	Generalised XGB
Aurassure	12066	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12094	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12095	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12096	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12097	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB

Manufacturer	Device	$PM_{2.5}$	PM_{10}	O_3	NO_2
Aurassure	12098	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12100	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12104	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12105	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12106	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12107	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12116	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12194	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12200	Generalised QR	Generalised QR	Generalised RFR	Generalised XGB
Aurassure	12201	Generalised QR	Generalised QR	Generalised RFR	Sensor-wise XGB
Respirer	002	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Respirer	2FA	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Respirer	814	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Respirer	DC1	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Respirer	FF5	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Airveda	117	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Airveda	136	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Airveda	137	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB

Manufacturer	Device	$PM_{2.5}$	PM_{10}	O_3	NO_2
Airveda	138	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Airveda	139	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Sensit	1168	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Sensit	1169	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Sensit	1170	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Sensit	1171	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Sensit	1172	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Aeron	3515	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Aeron	5515	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Aeron	6904	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Aeron	9123	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB
Aeron	9662	Sensor-wise QR	Sensor-wise QR	Sensor-wise XGB	Sensor-wise XGB

It was observed that XGB was the best model for NO_2 across all manufacturers. For O_3, the best model was RFR for Aurassure devices with EC Sense electrochemical sensors and XGB for other devices with Alphasense electrochemical sensors. For PM, SVR, followed by QR, performed better than other model algorithms. Because linear and quadratic regression algorithms can extrapolate beyond the training data range, QR was adopted to calibrate $PM_{2.5}$ and PM_{10}.

4.2. Performance evaluation with CSTEP-developed calibration models

The test data set was evaluated on the basis of the above models developed for various pollutants and sensors, and the performance was analysed using various metrics, such as the Pearson correlation coefficient (r), MAE, the coefficient of variation in MAE (CvMAE), and RMSE. Table 7 provides the range of performance metrics reported after applying correction with CSTEP models. For $PM_{2.5}$ and PM_{10}, three model categories were considered: (i) PM models, which used data from non-OPC sensors (Tera Sensor, Sensirion, or Plantower); (ii) PM_OPC models, which relied exclusively on measurements from the Alphasense OPC; and (iii) PM_hybrid models, which were developed using PM data from both an OPC sensor and a non-OPC sensor. All models additionally used relative humidity measured by the sensor as an input variable. For NO_2 and O_3, potential cross-interference from non-target pollutants was accounted for by including sensor responses to NO_2, O_3, and CO, along with sensor-measured ambient temperature as model input variables. The detailed list of model inputs for each sensor manufacturer is listed in Table 5.

Figures 2, 3, and 4 compare sensor performance with manufacturer calibration and localised CSTEP calibration. The figures are arranged such that the lower-left corner denotes 'better' performance (CvMAE close to 0 and r close to 1). The figures show that the overall performance of the sensors from the various manufacturers has improved (high Pearson r and low CvMAE) after applying localised CSTEP calibration.

Table 7[4]*: Evaluation of CSTEP calibration data with AQMS-reported data (test data set)*

Pollutant	Sensor	r	MAE	RMSE
$PM_{2.5}$	Aeron	0.6 to 0.72 (OPC)	10.37 to 12.13 (OPC)	14.25 to 16.88 (OPC)
		0.71 to 0.78 (Tera Sensor)	8.87 to 10.12 (Tera Sensor)	12.9 to 15.72 (Tera Sensor)
		0.71 to 0.80 (Hybrid)	8.65 to 9.84 (Hybrid)	12.43 to 15.53 (Hybrid)
	Airveda	0.77 to 0.83 (Plantower)	8.18 to 9.51 (Plantower)	12.26 to 13.8 (Plantower)
	Aurassure	0.72 to 0.84 (Sensirion)	7.90 to 9.92 (Sensirion)	10.81 to 16.58 (Sensirion)
	Respirer	0.65 to 0.86 (Plantower)	7.68 to 11.90 (Plantower)	10.48 to 17.64 (Plantower)
		0.57 to 0.77 (OPC)	9.95 to 12.39 (OPC)	13.71 to 17.63 (OPC)
	Sensit	0.8 to 0.86 (Plantower)	7.32 to 8.307 (Plantower)	9.84 to 13.06 (Plantower)
		0.7 to 0.78 (OPC)	9.37 to 11.6 (OPC)	13.31 to 14.55 (OPC)
		0.82 to 0.86 (Hybrid)	7.35 to 8.71 (Hybrid)	9.58 to 12.60 (Hybrid)

4 AI tools were used to improve code for data analysis and visualisation and to help fix errors and bugs. The AI-generated code was thoroughly validated by the authors.

Pollutant	Sensor	r	MAE	RMSE
PM_{10}	Aeron	0.51 to 0.71 (OPC)	20.65 to 24.44 (OPC)	29 to 35.41 (OPC)
		0.6 to 0.71 (Tera Sensor)	19.37 to 21.90 (Tera Sensor)	29.10 to 33.62 (Tera Sensor)
		0.6 to 0.75 (Hybrid)	18.62 to 21.64 (Hybrid)	27.09 to 33.82 (Hybrid)
	Respirer	0.71 to 0.80 (Hybrid)	16.56 to 21.15 (Hybrid)	23.15 to 30.12 (Hybrid)
		0.64 to 0.74 (Plantower)	18 to 22.26 (Plantower)	25.82 to 32.84 (Plantower)
		0.61 to 0.70 (OPC)	20.83 to 22.13 (OPC)	27.57 to 32.6 (OPC)
	Sensit	0.65 to 0.75 (Plantower)	16.86 to 22.16 (Plantower)	21.64 to 32.06 (Plantower)
		0.48 to 0.62 (OPC)	24.33 to 17.31 (OPC)	29.62 to 33.82 (OPC)
		0.62 to 0.76 (Hybrid)	17.74 to 21.64 (Hybrid)	23.83 to 30.53 (Hybrid)
NO_2	Aeron	0.80 to 0.94	2.93 to 4.84	4.37 to 8.19
	Airveda	0.89 to 0.93	3.50 to 3.76	5.75 to 7.92
	Aurassure	0.21 to 0.86	4.24 to 14.23	6.69 to 19.36
	Respirer	0.88 to 0.92	3.24 to 3.70	5.26 to 6.81
	Sensit	0.92 to 0.96	2.32 to 2.75	3.90 to 5.25
O_3	Aeron	0.96 to 0.98	2.21 to 3.58	3.4 to 5.024
	Airveda	0.90 to 0.95	4.27 to 5.70	5.75 to 7.92
	Aurassure	0.76 to 0.89	6.32 to 9.71	8.32 to 12.51
	Respirer	0.97 to 0.98	2.65 to 3.21	3.81 to 4.37
	Sensit	0.94 to 0.97	3.23 to 4.74	4.40 to 6.96

***Figure 2**: Comparative performance of $PM_{2.5}$ values reported by various manufacturers against AQMS-reported values with manufacturer calibration and CSTEP calibration (localised calibration).*

The performance shown is based on the entire collocation data set for manufacturer-calibrated values, while the test data set of the calibration data is used for the CSTEP-calibrated values. Proximity to the lower-left corner of each figure indicates better performance.

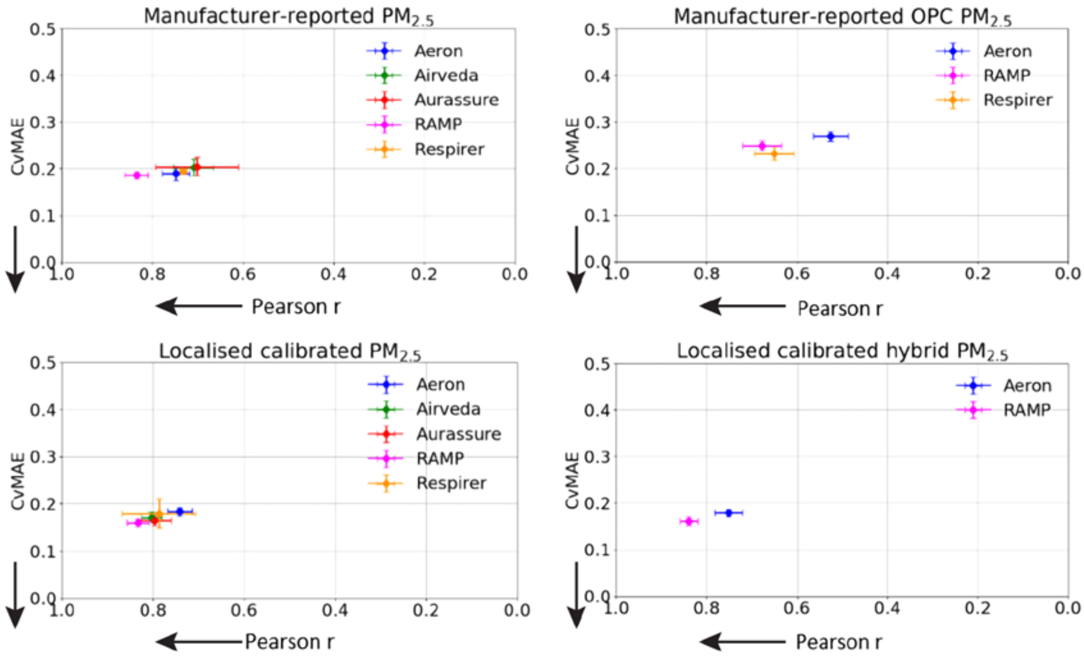


***Figure 3**: Comparative performance of PM_{10} values reported by various manufacturers against AQMS-reported values with manufacturer calibration and CSTEP calibration (localised calibration).*

The performance shown is based on the entire collocation data set for manufacturer-calibrated values, while the test data set of the calibration data is used for the CSTEP-calibrated values. Proximity to the lower-left corner of each figure indicates better performance.

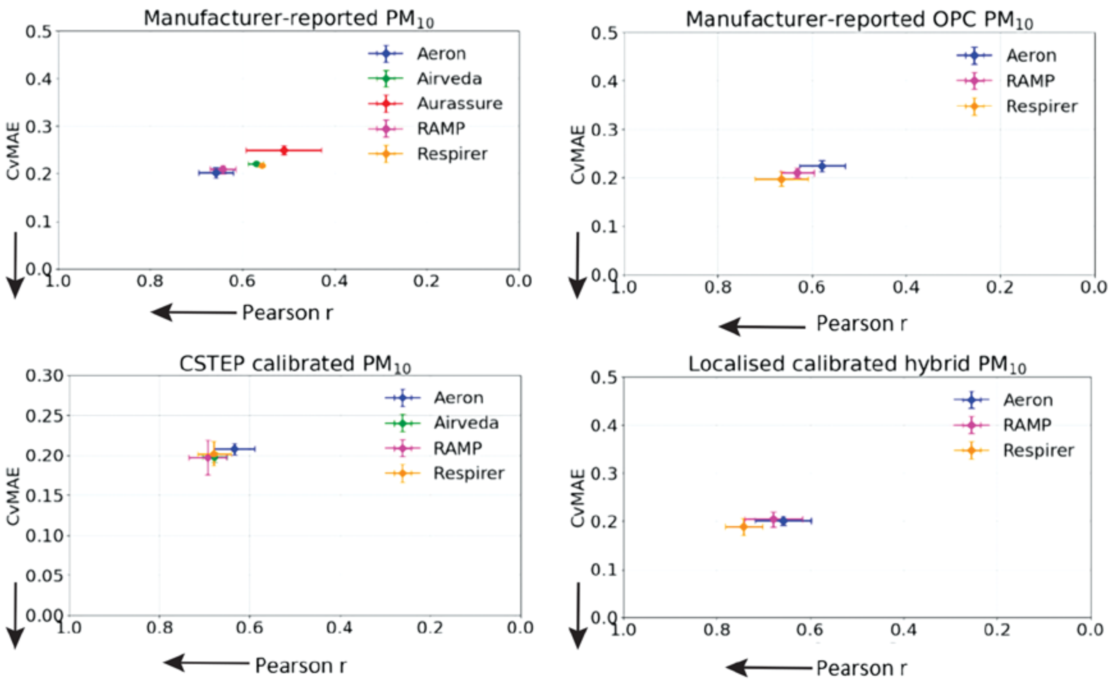
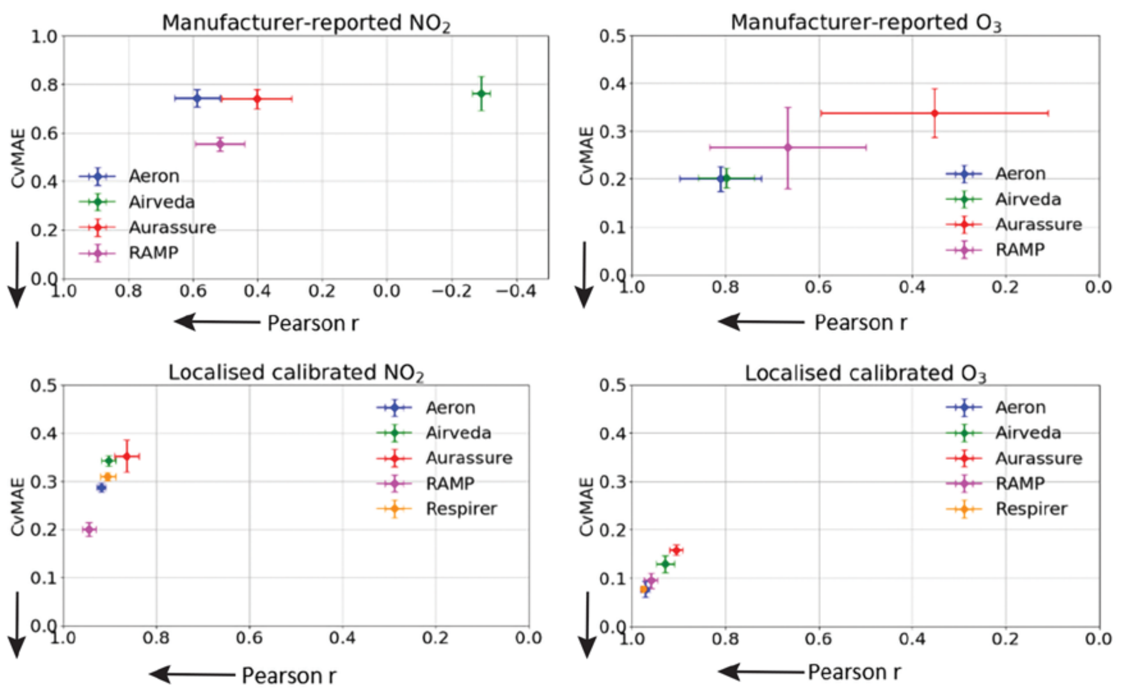
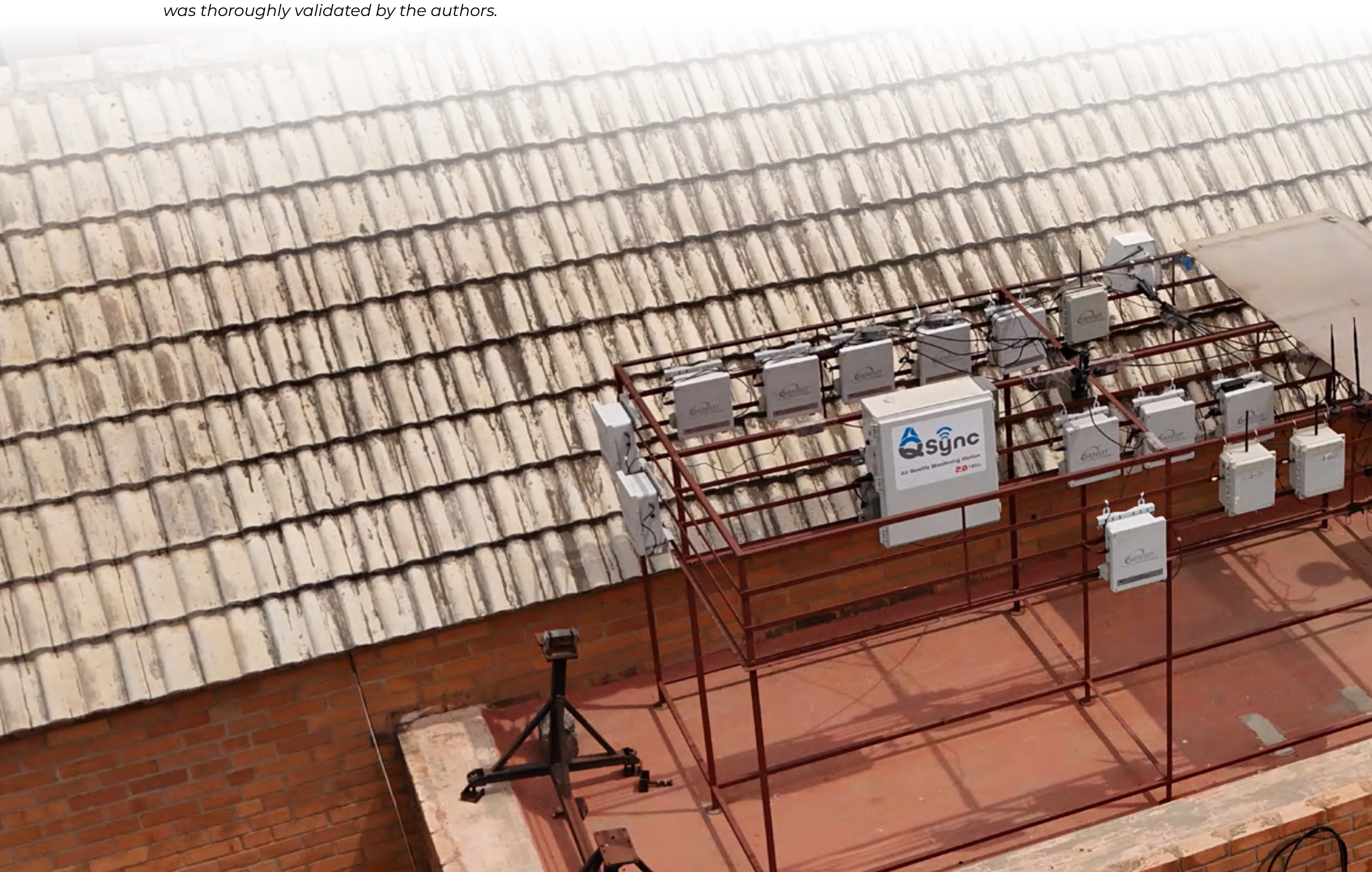


Figure 4[5]*: Comparative performance of NO_2 and O_3 values reported by various manufacturers against AQMS-reported values with manufacturer calibration and CSTEP calibration (localised calibration).*

The performance shown is based on the entire collocation data set for manufacturer-calibrated values, while the test data set of the calibration data is used for the CSTEP-calibrated values. Proximity to the lower-left corner of each figure indicates better performance.



5 *AI tools were used to improve code for data analysis and visualisation and to help fix errors and bugs. The AI-generated code was thoroughly validated by the authors.*



5. Summary: The Short-term Collocation

The short-term collocation study showed that precision can be poor even in raw values among sensor nodes from the same manufacturer. This could be due to variations in sensor sensitivity across sensor batches manufactured. The preliminary evaluation demonstrates that factory calibration, irrespective of the manufacturer, is insufficient for improving sensor performance due to the influence of environmental factors, such as different pollution loads and meteorological conditions. Hence, systematic collocation at the target site is required to ensure the reliability and accuracy of sensor data. In the short-term evaluation, machine learning methods such as XGB and RFR provided robust calibration models for different pollutant types. These methods were particularly effective for NO_2 and O_3, yielding correlation coefficients exceeding 0.8 and indicating strong agreement with reference data.

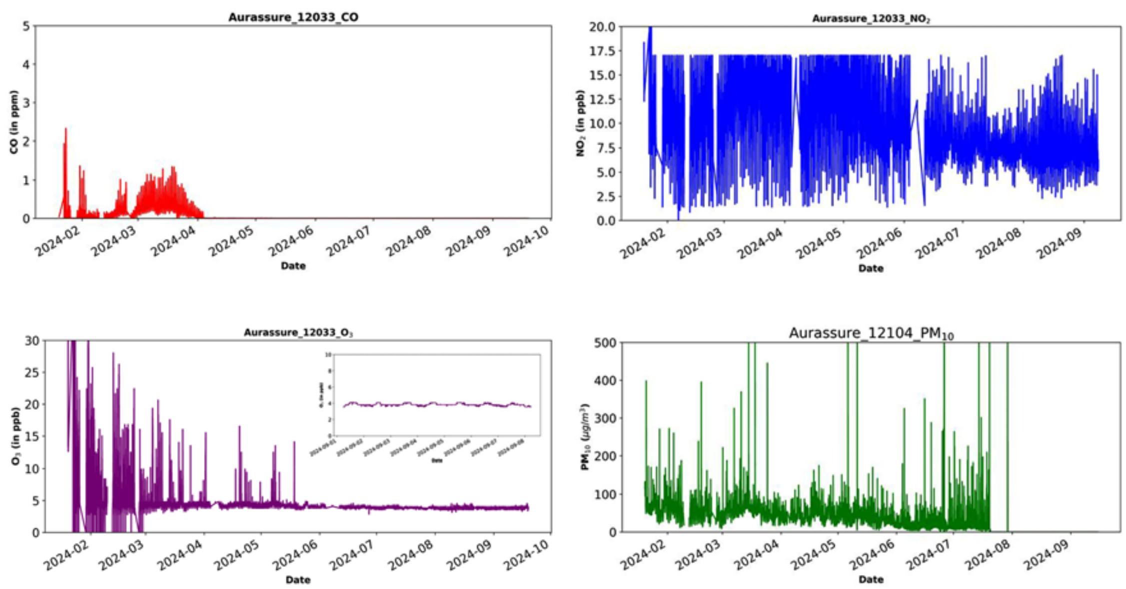
6. Medium-Term Performance Evaluation

To extend the current study and assess the long-term performance of sensors and the built calibration models, one sensor node from each manufacturer was deployed and maintained in collocation with the AQMS at Indi-SET. Moreover, 40 sensor nodes were installed across Bengaluru to continue the study's research objectives. The long-term performance evaluation helps assess the overall sensor performance from different manufacturers. Here, we report the medium-term performance of selected sensors over time. The detailed performance assessment is ongoing. The results presented here are based on preliminary assessments conducted using data as of September 30, 2024 (~8 months of operation).

In the medium-term evaluation, the overall performance of the Aurassure sensors deteriorated, especially for their EC Sense gas sensors. Figure 5 shows the performance of a few Aurassure nodes. We faced multiple issues with the sensors. Regarding CO sensors in Aurassure nodes, 16 of the 25 nodes reported 0 ppm after the initial months. An example of this issue is shown in Figure 5. In the case of NO_2 sensors, it was noticed that in one of the nodes, the maximum values reported by the sensor were capped at 17 ppb. Among the O_3 sensors, 11 of the 25 Aurassure nodes reported either 2 or 4 ppb, or the same value, for most of the time since July 2024. One of the PM sensors (Sensirion) also stopped working, reporting 0 µg/m^3. In June and July, it was also observed that the RH sensor of six Aurassure nodes reported 100% most of the time.

In the case of the other 20 nodes from the four manufacturers, a few PM sensors required replacement. In Aeron and Airveda, two PM sensors each (Tera Sensor in Aeron and Plantower in Airveda) stopped working in June 2024. One PM sensor (Alphasense OPC) from Respirer and Sensit stopped working in September. Further, three of five Respirer nodes have been reporting 100% RH since May.

Figure 5: *Examples of performance deterioration of Aurassure sensors over time*



7. Recommendations

The preliminary medium-term evaluation revealed that sensor performance varies over time and that a few do not function properly; therefore, the sensors will need to be replaced. These replaced sensors need to undergo further collocation, performance evaluation, and localised calibration model development. As a way forward, a detailed evaluation of the sensor performance is being conducted for the sensor nodes collocated with the AQMS at Indi-SET. This is to evaluate the long-term performance of sensors and the CSTEP-developed calibration models, as meteorological conditions and pollutant concentrations vary across seasons. For example, the monsoon period is characterised by high relative humidity and low pollution levels.

Further, for evaluating sensors in ambient conditions, we recommend that the target values mentioned in Table 8 should be applied, using hourly averages for PM and 15-minute averages for gaseous pollutants (EPA/600/R-20/279, 2021; EPA/600/R-20/280, 2021; EPA/600/R-23/145, 2023; EPA/600/R-23/146, 2024).

Table 8: Recommended target metrics for sensor evaluation (adapted from US-EPA recommendations)

Performance Metric		Target Value					
		$PM_{2.5}$	PM_{10}	O_3	NO_2	CO	SO_2
Precision	Pearson correlation coefficient (r)	≥ 0.9	≥ 0.9	≥ 0.9	≥ 0.9	≥ 0.9	≥ 0.9
	Coefficient of determination (R^2)	≥ 0.8	≥ 0.8	≥ 0.8	≥ 0.8	≥ 0.8	≥ 0.8
	Standard deviation (SD)	≤5 μg/m³	≤ 5 μg/m³	≤5 ppb	≤ 5 ppb	≤0.02 ppm	≤ 5 ppb
	Coefficient of variation (CV)	≤ 30%	≤ 30%	≤ 30%	≤ 30%	≤ 30%	≤ 30%
Bias	Slope (m)	1 ± 0.35	1 ± 0.35	1 ± 0.2	1 ± 0.35	1 ± 0.2	1 ± 0.35
	Intercept (c)	−5 ≤ c ≤ 5 μg/m³	−10 ≤ c ≤ 10 μg/m³	−5 ≤ c ≤ 5 ppb	−5 ≤ c ≤ 5 ppb	−0.05 ≤ c ≤ 0.05 ppm	−5 ≤ c ≤ 5 ppb
Linearity	Pearson correlation coefficient (r)	≥ 0.8	≥ 0.8	≥ 0.8	≥ 0.8	≥ 0.8	≥ 0.8
	Coefficient of determination (R^2)	≥ 0.7	≥ 0.7	≥ 0.8	≥ 0.7	≥ 0.8	≥ 0.7
Error	RMSE	≤ 7 μg/m³	≤14 μg/m³	≤ 5 ppb	≤ 15 ppb	≤0.15 ppm	≤ 15 ppb
	Normalised RMSE (NRMSE)	≤ 30%	≤ 30%	≤ 30%	≤ 30%	≤ 30%	≤ 30%
	Coefficient of variation in MAE (CvMAE)	≤ 30%	≤ 30%	≤ 30%	≤ 30%	≤ 30%	≤ 30%

8. References

Duvall, R., Clements, A., Barkjohn, K., Kumar, M., Greene, D., Dye, T., Papapostolou, V., Mui, W., & Kuang, M. (2024). *NO_2, CO, and SO_2 supplement to the 2021 report on performance testing protocols, metrics, and target values for ozone air sensors* (EPA/600/R-23/146). United States Environmental Protection Agency. https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=360578&Lab=CEMM

Duvall, R., Clements, A., Hagler, G., Kamal, A., & Kilaru, V. (2021). *Performance testing protocols, metrics, and target values for ozone air sensors: Use in ambient, outdoor, fixed-site, non-regulatory supplemental and informational monitoring applications* (EPA/600/R-20/279). United States Environmental Protection Agency .https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=350784&Lab=CEMM

Duvall, R., Clements, A., Hagler, G., Kamal, A., Kilaru, V., Goodman, L., Frederick, S., Johnson Barkjohn, K., VonWald, I., Greene, D., & Dye, T. (2021). *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* (EPA/600/R-20/280). United States Environmental Protection Agency. https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=350785&Lab=CEMM.

Duvall, R., Clements, A., Barkjohn, K., Kumar, M., Greene, D., Dye, T., Papapostolou, V., Mui, W., & Kuang, M. (2023). *PM_{10} supplement to the 2021 report on performance testing protocols, metrics, and target values for fine particulate matter air sensors* (EPA/600/R-23/145). United States Environmental Protection Agency. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=360419&Lab=CEMM.

Ouimette, J., Arnott, W. P., Laven, P., Whitwell, R., Radhakrishnan, N., Dhaniyala, S., Sandink, M., Tryner, J., & Volckens, J. (2023). Fundamentals of low-cost aerosol sensor design and operation. *Aerosol Science and Technology, 58*(1), 1–15. https://doi.org/10.1080/02786826.2023.2285935

United States Environmental Protection Agency (2017). *Quality assurance handbook for air pollution measurement systems, Volume II*. EPA-454/B-17-001. https://www.epa.gov/sites/default/files/2020-10/documents/final_handbook_document_1_17.pdf

Alphasense (n.d.) How electrochemical gas sensors work; Application Note AAN 104. https://www.alphasense.com/-/media/project/oneweb/oneweb/alphasense/products/application-notes/aan_104_app-note_v0.pdf?revision=217c547e-3447-4c26-b001-a473f75a950e

Baron, R. & Safell, J. (2017). Amperometric gas sensors as a low cost emerging technology platform for air quality monitoring applications: A review. *ACS Sensors*, *2*(11), 1553–1566. https://doi.org/10.1021/acssensors.7b00620

EC Sense Easy Gas Sensors (n.d.) *Ozone gas sensor module TB600B datasheet*. https://ecsense.com/wp-content/uploads/2024/03/TB600B-O_3_Datasheet_V3.0_20250904.pdf

Kaur, K. & Kelly, E. K. (2023) Performance evaluation of the Alphasense OPC-N3 and Plantower PMS5003 sensor in measuring dust events in the Salt Lake Valley, Utah. *Atmos. Meas. Tech.*, 16, 2455–2470. https://doi.org/10.5194/amt-16-2455-2023

Zimmerman, N., Presto, A. A., Kumar, S. P.N., Gu, J., Hauryliuk, A., Robinson, E. S., Robinson, A. L., & Subramanian, R. (2018). A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmos. Meas. Tech.*, 11, 291–313. https://doi.org/10.5194/amt-11-291-2018

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P. K., Robinson, A. L., Presto, A. A., & Subramanian, R. (2019a). Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. *Aerosol Science and Technology, 54*(2), 160–174. https://doi.org/10.1080/02786826.2019.1623863

Malings, C., Tanzer, R., Hauryliuk, A., Kumar, S. P. N., Zimmerman, N., Kara, L. B., Presto, A. A., & Subramanian, R. (2019b). Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring, *Atmos. Meas. Tech., 12*, 903–920, https://doi.org/10.5194/amt-12-903-2019

The Authors

Dr Emil Varghese

Dr Emil Varghese is a Senior Associate at CSTEP, with expertise in air quality monitoring using low-cost sensors and advanced aerosol and gaseous instruments. She served as the Principal Investigator for this project, spearheaded sensor evaluations at Indi-SET, and oversaw the operation of the multi-pollutant sensor network across Bengaluru. She also led the writing and data analysis in this report.

Dr Nirav Lekinwala

Dr Nirav Lekinwala is a Senior Associate at CSTEP, with expertise in air quality modelling, data analysis, and machine learning. He led the development of the localised calibration models for the sensors evaluated in this report.

Srinivas Sridharan

Srinivas Sridharan is a Lead Engineer at CSTEP, working in data science research across various domains such as renewable energy, air quality, and sustainability. He helped with exploratory data analysis and the development of machine learning models for calibrating low-cost air quality sensors.

Kavyashree N Kalkura

Kavyashree N Kalkura was an Analyst at CSTEP, with expertise in aerosol chemical characterisation, low-cost sensors, and the evolution of secondary organic aerosols. She supported the operation and maintenance of reference-grade instruments for Indi-SET and conducted the data analysis for the short-term sensor evaluation reported here.

Vinod S

Vinod S is a Senior Analyst at CSTEP, with expertise in the operation and maintenance of multi-pollutant low-cost sensors and reference-grade air quality instruments. He manages the air quality sensor network across Bengaluru and has led sensor troubleshooting efforts for Indi-SET, including validating data from reference-grade instruments.

Yashwant Pratap Singh Yadav

Yashwant Pratap Singh Yadav is an Analyst at CSTEP specialising in data analytics, emission inventory development, and air quality modelling. He supported the maintenance of sensors and reference instruments and contributed to the medium-term evaluation of sensor performance in this report.

Dr R Subramanian

Dr R Subramanian, former Sector Head of Air Quality at CSTEP, conceived of and acquired funding for the Indi-SET project. He also oversaw project progress, contributed to writing, and reviewed this final report.

CSTEP

Center for Study of Science, Technology & Policy

Bengaluru

#18, 10th Cross, Mayura Street, Papanna Layout,
Nagashettyhalli (RMV II Stage),
Bengaluru – 560094, Karnataka, India

Noida

1st Floor, Tower-A, Smartworks
Corporate Park, Sector-125,
Noida – 201303, Uttar Pradesh, India

+91-8066902500 www.cstep.in cpe@cstep.in @cstep_India