IoT-Based Low-Cost Intelligent Air-Quality Monitoring and Self-Calibrated Autonomous Purification System

Md. Saurob Bhuyan
Department of Computer Science and Engineering
Jatiya Kabi Kazi Nazrul Islam University
Trishal, Mymensingh, Bangladesh
mdsaurobbhuyan@gmail.com

F. R. M Mohiul Haque Bhuiyan
Department of Computer Science and Engineering
Jatiya Kabi Kazi Nazrul Islam University
Trishal, Mymensingh, Bangladesh
mohiulhaquerabby1432@gmail.com

Md. Sujan Ali
Department of Computer Science and Engineering
Jatiya Kabi Kazi Nazrul Islam University
Trishal, Mymensingh, Bangladesh
sujan_cse@jkkniu.edu.bd

***Abstract*—Air pollution is a leading cause of premature death worldwide, yet many low-cost IoT air-quality systems still handle monitoring and purification separately, depend on constant cloud access, and rarely say how reliable their own sensors are. This paper presents a solar-powered, ESP32-S3-based system that brings monitoring, sensor-health tracking, and autonomous purification together in one transparent, equation-based design. It measures $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3, cleans the raw readings through simple preprocessing and self-calibration, and computes the Air Quality Index (AQI) using the standard EPA method. A lightweight Sensor Health Indicator (SHI), built from the same anomaly statistics, continuously flags sensors that start to drift. Flagged sensors are removed only from a separate control AQI that drives the purifier, while the AQI reported to the cloud always uses all six pollutants. A simple threshold-based controller then uses this control AQI to run a multi-stage purifier, supported by a solar-first power system with MPPT and battery backup, and remote data is shared through ThingSpeak without the system depending on it to function. Using a real hourly Bangladesh air-quality dataset covering 103 cities and over 3.19 million records from 2000–2026, the preprocessing, AQI, and sensor-health logic were validated, correctly flagging SO_2 and NO_2 as the least reliable sensors, while an estimated bill of materials of USD 68–97 keeps the design affordable. Together, these results point toward a practical, self-aware and low-cost path to continuous air-quality management for pollution-prone regions.**

***Index Terms*—Air pollution, Air quality index (AQI), Air quality monitoring, Anomaly detection, Calibration, ESP32-S3, Internet of Things (IoT), Sensor health indicator, Solar power**

I. Introduction

Air pollution is a major global environmental and public health challenge with consequences for human health, quality of life and economic productivity. It is the second leading risk factor for noncommunicable diseases (NCDs) after tobacco use, with an estimated 4.2 million premature deaths annually, mainly from coronary heart disease, stroke, lung cancer, and chronic obstructive pulmonary disease (COPD) [1] as estimated by the World Health Organization (WHO). It affects not only health but also the participation in work, education, and social life [2]. This emphasizes the importance of continuous air quality monitoring, accurate Air Quality Index (AQI) assessment, and prompt purification.

The main pollutants for air quality assessment from the WHO Global Air Quality Guidelines [3] were $PM_{2.5}$, PM_{10}, O_3, NO_2, SO_2, and CO. Several volatile organic compounds (VOCs) are also known to be carcinogenic [4]. Recent IoT-based monitoring systems (Section II) have demonstrated the feasibility of low-cost, sensor-driven platforms, but they are still constrained by heavy dependence on the cloud [8], difficulties in sensor calibration [9], inadequate solar energy management [10], and limited automated purification [7]. These gaps necessitate a robust, edge-based, energy-efficient framework for accurate AQI assessment, continuous sensor-health tracking and autonomous air quality management.

To this end, this paper proposes a solar-powered, ESP32-S3-based IoT system that unifies air quality monitoring, sensor-health tracking, and autonomous purification within a single deterministic, fully auditable framework, with contributions summarized in Section II-B.

The remainder of the paper is organized as follows: Section II reviews related work and identifies the research gaps addressed in this study; Section III describes the proposed system design; Section IV reports experimental results; and Section V concludes the paper.

II. Literature Review and Research Gaps

A. Related Work

IoT-based air quality monitoring has enabled continuous, cost-effective measurement of pollutants indoors and outdoors. Imankulova et al. [8] developed a Raspberry Pi–Arduino indoor monitor with a Django-based web interface; Chandrakala et al. [11] used Arduino/Raspberry Pi with PM and ozone sensors to monitor PM_1, $PM_{2.5}$, PM_{10}, and O_3; and Stanko et al. [12] introduced fog-based processing for lower-latency, edge-level decisions. Among ESP32-class platforms closer to the one adopted here, Sinambela et al. [9] implemented an ESP32-based monitor but left sensor drift and calibration as open challenges with no automated correction. El Anshori et al.

979-8-3195-2804-9/26/$31.00 ©2026 IEEE

[10] built a real-time ESP32-based AQI system for policy support, but it relies on grid power with no solar-based energy management. Wadkar et al. [7] proposed an ESP32 alert system for Indian metropolises that notifies users of high pollution but stops at the alert, with no automated purification response.

AI has also entered air-quality prediction: Kumar and Doss [13] proposed AIRO, combining low-cost sensing with a CNN–Bi-LSTM model for real-time monitoring and short-term prediction; Poulose et al. [5] combined sensor data with an ML model for indoor monitoring and purification; Abubeker et al. [6] integrated LoRaWAN into an automated indoor system for passenger safety; and Pujitha and Rama Krishna [14] combined IoT sensing with ML for AQI prediction. Such approaches are generally based on heavier, cloud-connected models, and rarely report sensor reliability over time. Sankurantripati and Duchaine [15] reviewed the state-of-the-art in UV-based purifiers and called for more realistic evaluation, Gurung et al. [16] designed an IoT purifier that activates when pollutant levels go above a threshold, and Attia et al. [17] reported better removal with nano-carbon filters. Anitha and Kumar [24] combined multi-pollutant sensing with multi-stage filtration but reported raw concentrations rather than a standardized AQI, with no sensor-reliability tracking. Afrial et al. [20] integrated a relay-controlled purifier with real-time logging, but activation is a simple on/off decision with no sensor-health awareness or graded, AQI-driven control. Further related work explores IoT-based monitoring, sensing, and filtration [18], [19], [21].

Despite this progress, several gaps remain: most systems address either monitoring or purification, not both, within a single autonomous framework [8], [11], [12], [15]–[17]; many depend on continuous cloud connectivity, hurting latency and reliability during outages [8], [12], [13]; and few report a continuous, auditable sensor-health indicator alongside the AQI [8], [11], [12], [16], even where AI or statistical models are used [13], [14]. Limited pollutant coverage, sensor drift, and insufficient solar-first energy management add to these gaps. An integrated ESP32-S3 framework combining equation-based AQI assessment, self-calibration, sensor-health monitoring, autonomous purification, and solar-assisted operation therefore remains a valuable research direction.

B. Research Contributions

The main contributions of this work are summarized as follows:

- A **low-cost ESP32-S3-based system** is developed for air quality monitoring, AQI calculation, sensor-health tracking, and automatic purification.
- A **simple, equation-based processing method** is used for data preprocessing, anomaly detection, and sensor self-calibration before the AQI is calculated.
- The current AQI is calculated using the **standard AQI equation**, supported by a lightweight, equation-based **Sensor Health Indicator (SHI)** that continuously tracks per-sensor reliability.
- A **simple, AQI-based purification control** scheme is developed, where sensors flagged as unreliable by the SHI are excluded from the control AQI that drives the purifier, while the AQI reported for remote visualization always reflects all six sensor channels.
- A **solar-first power system** with a solar panel, MPPT, battery management, and DC regulation is integrated to support the ESP32-S3 and the purification unit.
- The system supports **multiple pollutant measurements** and local edge operation, while ThingSpeak provides remote monitoring of air quality and system status.

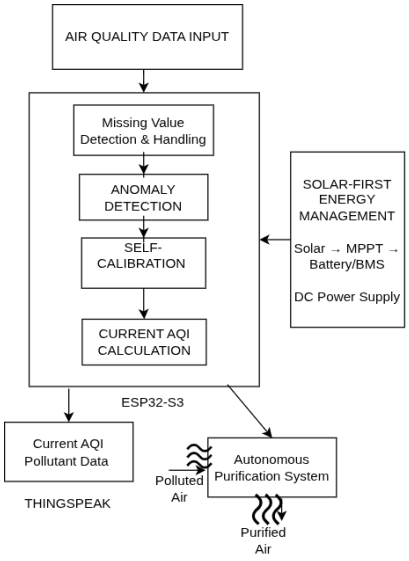


Fig. 1. Overall architecture of the proposed solar-first air-quality monitoring and autonomous purification system, showing the preprocessing/self-calibration stage feeding the two parallel AQI and SHI computations, the deterministic controller that derives the control AQI, and the solar-first power subsystem.

III. System Design and Implementation

A. System Overview

Fig. 1 shows the overall architecture. Raw readings for $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and O_3 pass through preprocessing and self-calibration (Section III-D), then feed two parallel computations: the AQI (Section III-E), maintained in two named variants (AQI_{rep} for reporting and AQI_{ctrl} for actuation), and a per-sensor Sensor Health Indicator (SHI, Section III-F) from the same anomaly statistics. Both feed a deterministic controller (Section III-G) that drives the purifier using AQI_{ctrl}, while a solar-first power subsystem (Section III-H) supplies the node and ThingSpeak reports AQI_{rep} (Section III-J) without being required for core operation. All thresholds below are design specifications, validated empirically in Section IV. The system is a proposed design, validated offline on historical data; hardware fabrication and field testing are future work.

B. Dataset

For offline development and AQI-equation validation, this work uses the public Bangladesh AQI dataset [22]: over 3.19 million hourly records from **103 cities** (2000–2026), covering $PM_{2.5}$, PM_{10}, CO, CO_2, NO_2, SO_2, O_3, and AQI. Coverage is uneven across cities and years since stations were added over time. The hardware suite (Section III-C) uses six of these

pollutants ($PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3). The dataset serves only as historical/offline data to validate the AQI and sensor-health logic, not as a live sensor simulation.

C. Hardware Sensor Suite

The edge node is based on an **ESP32-S3** microcontroller, which was chosen for its dual-core processor, adequate GPIO/ADC channels, and low power draw for solar-first operation. The sensing suite consists of laser-scattering particulate sensor (PMS7003/SDS011 class) for $PM_{2.5}$/PM_{10}, electrochemical or low-power infrared CO sensor, and electrochemical gas sensors for NO_2, SO_2 and O_3. Readings are sampled every 60 s and buffered locally before preprocessing.

D. Data Preprocessing and Self-Calibration

Each raw pollutant stream $x_p[n]$ (pollutant p, sample index n) passes through range clamping, missing-value handling, and anomaly flagging.

Non-negativity clamp. Since concentrations cannot be physically negative, each raw sample is first clamped at zero, independently of the statistical test below:

$$x_p[n] \leftarrow \max\big(0,\, x_p[n]\big) \tag{1}$$

This removes small negative artifacts occasionally produced by electrochemical gas sensors near zero (e.g., NO_2 and O_3 in Table III), which are physically impossible but not always large enough to trip the z-score test in Eq. (2).

Missing-value handling. A missing sample is replaced by the most recent valid reading if the gap is below a bound T_{max}; longer gaps are left missing and excluded from AQI computation for that interval.

Anomaly detection and self-calibration. A rolling z-score is computed over a window of length W:

$$z_p[n] = \frac{x_p[n] - \mu_p[n]}{\sigma_p[n]} \tag{2}$$

where $\mu_p[n]$ and $\sigma_p[n]$ are the mean and standard deviation over the trailing window $[n-W, n)$. A sample is flagged anomalous if $|z_p[n]| > z_{th}$ ($z_{th} = 3$) and replaced by the window median rather than discarded, so a single glitch does not create a data gap. This equation-based method is deliberately not a learned model, keeping it fully auditable. "Self-calibration" here means only this online, anomaly-driven correction — sensor gain/offset remain fixed after initial bench calibration. A cumulative daily flag count per sensor feeds the *Sensor Health Indicator* (Section III-F).

E. Current AQI Calculation

The current AQI uses the standard piecewise-linear breakpoint formula [23]:

$$I_p = \frac{I_{Hi} - I_{Lo}}{BP_{Hi} - BP_{Lo}}\,(C_p - BP_{Lo}) + I_{Lo} \tag{3}$$

where C_p is the preprocessed concentration of pollutant p (Section III-D), BP_{Lo}/BP_{Hi} are the bracketing breakpoints, and I_{Lo}/I_{Hi} the corresponding index values. Applied per pollutant over a sensor set $\mathcal{S}[n]$, the dominant-pollutant AQI is

$$\mathrm{AQI}(\mathcal{S}[n]) = \max_{p \in \mathcal{S}[n]} I_p \tag{4}$$

matching regulatory AQI reporting while staying deterministic and model-free.

Two named instances of Eq. (4), differing only in $\mathcal{S}[n]$, are maintained:

$$\mathrm{AQI}_{\mathrm{rep}}[n] = \mathrm{AQI}(\mathcal{P})[n], \tag{5}$$

$$\mathrm{AQI}_{\mathrm{ctrl}}[n] = \mathrm{AQI}(\mathcal{H}[n])[n], \tag{6}$$

where $\mathcal{P} = \{\mathrm{PM}_{2.5}, \mathrm{PM}_{10}, \mathrm{CO}, \mathrm{NO}_2, \mathrm{SO}_2, \mathrm{O}_3\}$ is the full pollutant set, and $\mathcal{H}[n] \subseteq \mathcal{P}$ is the subset currently rated Healthy or Watch by the SHI (Section III-F). $\mathrm{AQI}_{\mathrm{rep}}[n]$ is the *reported* AQI — always all six channels, logged to ThingSpeak, and used throughout Section IV. $\mathrm{AQI}_{\mathrm{ctrl}}[n]$ is the *control* AQI, used only by the purification controller (Section III-G), and equals $\mathrm{AQI}_{\mathrm{rep}}[n]$ whenever every sensor is Healthy or Watch. Hereafter, "AQI" means $\mathrm{AQI}_{\mathrm{rep}}$; only the controller uses $\mathrm{AQI}_{\mathrm{ctrl}}$.

F. Sensor Health Monitoring

Beyond correcting the pollutant stream itself (Section III-D), the system tracks a per-pollutant *Sensor Health Indicator* (SHI) so gradual degradation (dust-sensor fouling, electrode aging) can be flagged for maintenance. The SHI comes directly from the anomaly-flag count of the rolling z-score stage (Eq. (2)):

$$\mathrm{SHI}_p[d] = 100 \times \left(1 - \frac{F_p[d]}{N_p[d]}\right) \tag{7}$$

where $F_p[d]$ and $N_p[d]$ are the flagged and total samples for pollutant p on day d. $\mathrm{SHI}_p[d] \in [0, 100]$, with values near 100 indicating a stable sensor.

Each sensor is placed into one of three bands by its flag rate $1 - \mathrm{SHI}_p[d]/100$: *Healthy* ($< 0.5\%$), *Watch* (0.5–1.0%), and *Attention* ($> 1.0\%$, a dashboard maintenance trigger). The SHI never alters the reported AQI: $\mathrm{AQI}_{\mathrm{rep}}[n]$ (Eq. (5)) always uses all six channels, so the value logged to ThingSpeak is never silently changed by a maintenance flag. Instead, SHI status determines membership in $\mathcal{H}[n]$ (Section III-E): an Attention-rated sensor is excluded from $\mathcal{H}[n]$ and therefore from the control AQI, $\mathrm{AQI}_{\mathrm{ctrl}}[n]$ (Eq. (6)), which is what the purification controller (Section III-G) actually consults. SHI uses W=24 h, z_{th}=3, and the flag-rate bands above; the SO_2/NO_2 Attention result comes from the historical dataset only, not from hardware drift.

G. Purification Control Logic

Fan/stage activation is kept deterministic, being safety-relevant. A threshold rule maps the control AQI to a fan duty level:

$$d[n] = \begin{cases} 0, & \mathrm{AQI}_{\mathrm{ctrl}}[n] \le \tau_1 \\ d_{low}, & \tau_1 < \mathrm{AQI}_{\mathrm{ctrl}}[n] \le \tau_2 \\ d_{med}, & \tau_2 < \mathrm{AQI}_{\mathrm{ctrl}}[n] \le \tau_3 \\ d_{high}, & \mathrm{AQI}_{\mathrm{ctrl}}[n] > \tau_3 \end{cases} \tag{8}$$

where $d[n]$ is the fan duty level and τ_1, τ_2, τ_3 follow the standard AQI boundaries (50, 100, 150), adopted from the EPA AQI categories [23], not empirically tuned, as no field data exist yet. Since $\mathrm{AQI}_{\mathrm{ctrl}}[n]$ (Eq. (6)) uses only Healthy/Watch sensors,

TABLE I
DESIGN-STAGE POWER BUDGET (COMPONENT RATINGS)

Component	Rated Power	Duty Assumption
ESP32-S3 (active/Wi-Fi)	∼0.5–0.7 W	Continuous
Sensor suite (all sensors)	∼0.3–0.5 W	Continuous
Purifier fan (per stage level)	3–12 W	Duty-cycled per Eq. (8)
UV-C stage (optional)	3–5 W	Duty-cycled

TABLE II
APPROXIMATE BILL OF MATERIALS (SINGLE-UNIT, INDICATIVE)

Component	Approx. Cost (USD)
ESP32-S3 development board	7–9
PM sensor (laser scattering)	9–13
Gas sensor set (CO, NO_2, SO_2, O_3)	17–24
Solar panel + MPPT + BMS + battery	17–23
Purifier (fan, HEPA, activated carbon)	13–18
Enclosure, wiring, misc.	5–10
Total (approx.)	**68–97**

a sensor entering Attention is automatically dropped from $\mathcal{H}[n]$ and cannot silently bias actuation. The purification stack uses conventional, well-characterized filtration — washable pre-filter, HEPA for $PM_{2.5}$/PM_{10}, activated carbon for gases/odor, and a shielded optional UV-C stage to avoid ozone generation. Exotic, unvalidated pathways (e.g., MOF-based CO_2 sorption, photocatalytic oxidation) are excluded, because the regeneration and byproduct chemistry are not characterizable in this work.

H. Solar-First Power Management

Table I lists the design-stage (nameplate/estimated) power budget (component ratings and targets, not field measurements); field-measured values will be reported from the deployed prototype. The power subsystem consists of a solar panel with an MPPT charge controller, a protected Li-ion battery with BMS, and a DC-DC regulator to power the 3.3 V/5 V rails; an AC adapter serves as a backup when solar/battery power is insufficient.

I. Approximate Bill of Materials

Table II gives a rough, indicative (single-unit) component-level cost breakdown that helps to justify the "low-cost" framing used elsewhere in the paper.

J. IoT Connectivity

Processed AQI values ($\mathrm{AQI}_{\mathrm{rep}}[n]$), fan-duty state, and per-sensor SHI values are periodically pushed to ThingSpeak for remote visualization. Since AQI calculation, sensor-health monitoring, and control decisions (Sections III-E–III-G) all run locally on the ESP32-S3, a connectivity loss only impacts remote visualization, not essential monitoring/purification functionality.

TABLE III
SUMMARY STATISTICS OF POLLUTANT CONCENTRATIONS (POST-CLAMP)

Pollutant	Mean	Std.	Min	Max
PM_{10}	63.51	52.59	0.01	957.4
$PM_{2.5}$	45.34	36.03	0.10	670.6
CO	384.20	267.71	0.01	6857.0
NO_2	14.33	16.05	0.00	178.1
SO_2	8.09	8.01	0.00	173.4
O_3	77.90	42.57	0.00	703.2

TABLE IV
ANOMALIES FLAGGED AND REPLACED BY SELF-CALIBRATION

Pollutant	Flagged Samples	% of Records
PM_{10}	10,013	0.31%
$PM_{2.5}$	9,741	0.31%
CO	18,645	0.58%
NO_2	34,424	1.08%
SO_2	42,227	1.32%
O_3	7,261	0.23%

IV. EXPERIMENTAL RESULTS

In this section, we present results of running the preprocessing, self-calibration and AQI modules of Section III on the dataset of Section III-B; all figures and numbers are computed directly from this dataset. The prototype is not yet fabricated, so this section validates the algorithmic layer only; hardware metrics are future work (Section V). Sensor-health results appear in Section IV-F; full purification-control validation, including control-AQI behavior under an Attention-flagged sensor, is left for the deployed hardware.

A. Dataset Characteristics

Table III presents the summary statistics of the six pollutant channels after the non-negativity clamp (Section III-D, Eq. (1)); all minimums are then ≥ 0, with the near-zero negative artifacts seen before in NO_2 and O_3 removed in this stage. The missing values are negligible (below 0.1% for each pollutant) and thus no interval was excluded under the T_{max} gap rule. The $PM_{2.5}$ and PM_{10} have a large dynamic range (up to 670.6 and 957.4 μg/m^3).

B. Preprocessing and Self-Calibration Results

The rolling z-score procedure (Eq. (2), $W = 24$ h, $z_{th} = 3$) was applied to each pollutant stream following the clamp of Eq. (1). The samples that were flagged and median-replaced are shown in Table IV. The highest flagging rates are observed for SO_2 and NO_2 (1.32% and 1.08%, respectively), in line with their higher variance (Table III), and the lowest for O_3 and $PM_{2.5}$. Fig. 2 shows this for SO_2: isolated spikes are suppressed but the extended post-2022 elevated period is retained, demonstrating the method can discriminate glitches from real pollution events. These flag counts feed the SHI in Section IV-F.

C. Pollutant Correlation Structure

Fig. 3 shows the Pearson correlation matrix of the six calibrated pollutants. $PM_{2.5}$ and PM_{10} are strongly correlated

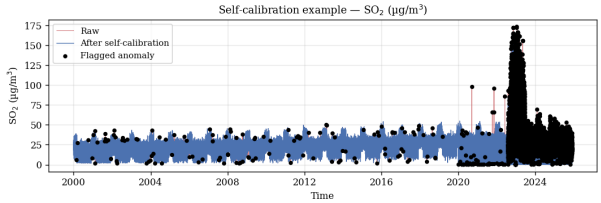


Fig. 2. Self-calibration example for SO_2: raw readings, output after rolling z-score self-calibration, and flagged anomalous samples.

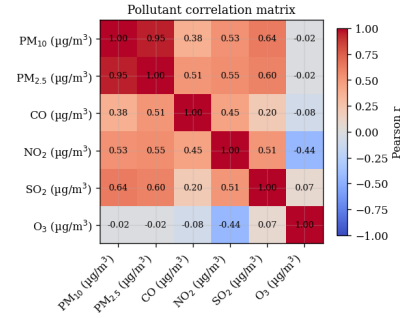


Fig. 3. Pearson correlation matrix of the six calibrated pollutant concentrations.

($r = 0.95$), as expected from shared particulate sources, while SO_2 correlates moderately with both ($r = 0.60$–0.64). O_3 is weakly or negatively correlated with the others, notably NO_2 ($r = -0.44$), consistent with the ozone–NO_x titration effect. This supports the dominant-pollutant AQI formulation (Eq. (4)) over a single-pollutant proxy, since no one pollutant captures overall air quality across all periods.

D. Current AQI Computation Results

The reported AQI, $\mathrm{AQI}_{\mathrm{rep}}[n]$ (Eq. (5)), was computed for every timestamp from all six self-calibrated channels (Eq. (3)). Since the offline dataset has no live SHI-triggered exclusions, all figures here refer to $\mathrm{AQI}_{\mathrm{rep}}[n]$; control-AQI behavior under an Attention-flagged sensor is exercised only by the field-deployed controller (Section III-G). $PM_{2.5}$ dominated the large majority of hourly records, followed by O_3, while PM_{10}, NO_2, and SO_2 rarely did. A substantial share of hours fall in the Moderate and Unhealthy (151–200) ranges, with Good and USG also well represented, and only a small fraction in Very Unhealthy/Hazardous, concentrated post-2022. This confirms the AQI is sensitive across the full severity range rather than saturating, as needed for the control logic of Eq. (8).

E. Temporal Variation of AQI

Fig. 4 shows $\mathrm{AQI}_{\mathrm{rep}}[n]$ by hour of day and day of week. The median stays in a narrow band ($\sim$100–130) across all 24 hours, with a mild late-afternoon/evening rise, and a similarly narrow spread across weekdays. The absence of a day-of-week effect implies that pollution in this place is mainly driven by ambient/traffic factors.

F. Sensor Health Monitoring Results

Using Table IV, the per-sensor SHI (Eq. (7)) was computed for the full dataset. Fig. 5 shows the resulting flag rates and SHI

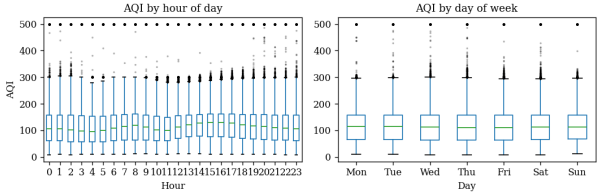


Fig. 4. Distribution of the reported AQI ($\mathrm{AQI}_{\mathrm{rep}}$) by hour of day (left) and by day of week (right).

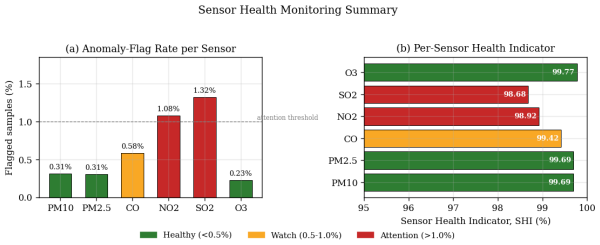


Fig. 5. Sensor health monitoring summary: (a) anomaly-flag rate per pollutant sensor and (b) resulting per-sensor Sensor Health Indicator (SHI), with deterministic Healthy/Watch/Attention bands.

values: SO_2 and NO_2 fall into the Attention band ($> 1.0\%$), CO into Watch, and PM_{10}, $PM_{2.5}$, and O_3 remain Healthy throughout. This confirms the equation-based SHI separates high-variance sensors from stable channels without a learned model, and directly identifies SO_2/NO_2 as needing maintenance; in a live deployment these two would accordingly be excluded from $\mathcal{H}[n]$ and thus from $\mathrm{AQI}_{\mathrm{ctrl}}[n]$ (Section III-E).

G. Comparison with Related Systems

Table V compares the proposed system with four recent systems. All four rely on mains power and none tracks sensor health, and their AQI handling ranges from concentration-only reporting or threshold switching to ML-based estimation, whereas the proposed system combines an equation-based AQI, a sensor-health indicator, AQI-driven multi-stage purification, and solar-first power in one node. Its estimated cost of USD 68–97 (Table II) is also below the one reported figure of about USD 187 [24]. The comparison is feature-based, and the proposed system's cost is an indicative estimate, not a measured result.

V. Conclusion

This paper presented a unified, edge-based framework for air-quality monitoring, sensor-health tracking, and autonomous purification on a low-cost, solar-powered ESP32-S3 platform, without cloud dependence or opaque learned models. Deterministic preprocessing and self-calibration feed both the AQI and a lightweight SHI: unreliable sensors are excluded from a dedicated control AQI driving the purifier, while the reported AQI continues to reflect every channel. Validation on a large, real-world hourly Bangladesh dataset spanning 103 cities confirmed the preprocessing, AQI, and SHI modules, correctly identifying SO_2 and NO_2 as the least reliable channels, with a bill of materials of USD 68–97 supporting the low-cost objective. This study targets the algorithmic layer of

TABLE V
COMPARISON WITH RECENT IOT-BASED AIR-QUALITY MONITORING AND PURIFICATION SYSTEMS

Feature/System	MCU	Pollutants Monitored	AQI Calculation	Sensor Health Tracking	Autonomous Purification	Power Source	Remote Visualization	Approx. Cost (USD)
Imankulova et al. [8]	Raspberry Pi 3	CO_2, NH_3, benzene, smoke, temp., humidity	Yes	No	No	Mains	Django Web App	Not reported
Anitha and Kumar [24]	Raspberry Pi 3 B+	PM_1, $PM_{2.5}$, PM_{10}, CO, CO_2, NH_4, LPG, smoke	No (concentration only)	No	Yes (multi-filter)	Mains	ThingSpeak	∼187
Afrial et al. [20]	ESP32	CO, CH_4, $PM_{2.5}$, temp., humidity	No (threshold-based)	No	Yes (relay-based)	Mains	ThingSpeak	Not reported
Poulose et al. [5]	Not specified	$PM_{2.5}$, CO_2, VOCs, temp., humidity	Yes (ML-based)	No	Yes (AI-driven)	Mains	Not specified	Not reported
Proposed System	**ESP32-S3**	**$PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, O_3**	**Yes (EPA equation-based)**	**Yes (equation-based SHI)**	**Yes (AQI-driven, multi-stage)**	**Solar-first (MPPT + BMS)**	**ThingSpeak (edge-first)**	**68–97**

a proposed design: the system was evaluated on historical data rather than live conditions, as the preprocessing, AQI, and SHI logic are hardware-independent. The dataset cannot substitute for physical sensor behavior, so field testing is the planned next stage. Future work will first fabricate the prototype and validate it in the field: accuracy/error against a reference-grade monitor, SHI false-positive/false-negative rates via fault injection, and response time with before/after purification performance. Building on this, we will target short-term forecasting and lightweight edge-AI models for early pollution warnings, adaptive learning for sensor drift, and integrating forecasting, AQI, SHI, and purification into one closed-loop system.

REFERENCES

[1] World Health Organization, "Estimating the morbidity from air pollution and its economic costs." Accessed: Oct. 8, 2026. [Online]. Available: https://www.who.int/health-topics/environmental-health/estimating-the-morbidity-from-air-pollution-and-its-economic-costs

[2] World Health Organization, "Air quality and health." Accessed: Oct. 8, 2026. [Online]. Available: https://iris.who.int/server/api/core/bitstreams/c3fce484-3841-4f49-894f-3dc41e4d5d7e/content

[3] World Health Organization, "WHO global air quality guidelines." Accessed: Oct. 8, 2026. [Online]. Available: https://www.who.int/news-room/questions-and-answers/item/who-global-air-quality-guidelines

[4] World Health Organization, "Types of pollutants." Accessed: Oct. 8, 2026. [Online]. Available: https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants

[5] A. Poulose *et al.*, "Real-time indoor air quality monitoring and purification system using sensor data and machine learning technique," in *Proc. 2nd Int. Conf. Sustain. Comput. Smart Syst.*, 2024, pp. 1271–1279, doi: 10.1109/ICSCSS60660.2024.10624835.

[6] K. M. Abubeker *et al.*, "Internet of Things and LoRaWAN-assisted real-time indoor air quality monitoring and automation system for enhancing passenger safety," *Int. J. Low-Carbon Technol.*, vol. 20, pp. 1427–1439, 2025.

[7] V. S. Wadkar *et al.*, "IoT-based air quality monitoring and alert system: Addressing pollution challenges in Indian metropolises," in *Proc. 6th Int. Conf. Emerg. Technol.*, 2025, pp. 1–6.

[8] B. Imankulova, A. Zhakipova, and A. Razaque, "IoT-based real-time indoor air quality monitoring and web server management system using Raspberry Pi," in *Proc. 9th Int. Conf. Digit. Technol. Educ. Sci. Ind.*, Almaty, Kazakhstan, Oct. 2024.

[9] M. Sinambela *et al.*, "IoT-based air quality monitoring system design and development using ESP32," in *Proc. 9th Int. Conf. Inform. Comput.*, 2024, pp. 1–6, doi: 10.1109/ICIC64337.2024.10956737.

[10] Y. T. El Anshori, R. M. Kunda, and F. Manuhutu, "Design and construction of a real-time air quality monitoring system using IoT-based ESP32 to strengthen environmental policies," *J. Penelitian Pendidikan IPA*, vol. 11, no. 2, pp. 145–152, 2025, doi: 10.29303/jppipa.v11i2.9820.

[11] M. Chandrakala, M. V. Lakshmaiah, and P. Thimmaiah, "Advanced air quality monitoring: PM and ozone with Raspberry Pi and Arduino," *Discov. Appl. Sci.*, vol. 7, no. 5, Art. no. 436, 2025, doi: 10.1007/s42452-025-06724-9.

[12] A. Stanko *et al.*, "Real-time air quality management: Integrating IoT and fog computing for effective urban monitoring," in *Proc. 2nd Int. Workshop Comput. Inf. Technol. Ind. 4.0*, Ternopil, Ukraine, Jun. 2024.

[13] T. Kumar and A. Doss, "AIRO: Development of an intelligent IoT-based air quality monitoring solution for urban areas," *Procedia Comput. Sci.*, vol. 218, pp. 262–273, 2023, doi: 10.1016/j.procs.2023.01.008.

[14] S. Pujitha and M. Rama Krishna, "IoT and machine learning based air quality monitoring and AQI prediction system," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 15, no. 4, pp. 739–748, Apr. 2026, doi: 10.17148/IJARCCE.2026.15496.

[15] S. Sankurantripati and F. Duchaine, "Indoor air quality control for airborne diseases: A review on portable UV air purifiers," *Fluids*, vol. 9, no. 12, Art. no. 281, 2024, doi: 10.3390/fluids9120281.

[16] M. Gurung *et al.*, "IoT based automatic air pollution monitoring and purification system," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 9, pp. 1378–1383, Sep. 2021, doi: 10.22214/ijraset.2021.38039.

[17] Y. A. Attia *et al.*, "Nano carbon-modified air purification filters for removal and detection of particulate matters from ambient air," *Sci. Rep.*, vol. 14, Art. no. 621, 2024, doi: 10.1038/s41598-023-50902-x.

[18] N. M. Zain *et al.*, "The development of an IoT-based air quality monitoring system using the Blynk application," *J. Comput. Res. Innov.*, vol. 9, no. 1, pp. 158–166, 2024, doi: 10.24191/jcrinn.v9i1.426.

[19] M. Abdullah *et al.*, "Development of cellular IoT-based, portable outdoor air quality monitoring system for pollution mapping," *Eng. Proc.*, vol. 118, no. 1, Art. no. 30, 2025, doi: 10.3390/ECSA-12-26529.

[20] M. Afrial *et al.*, "Multi-sensor indoor air quality monitoring with real-time logging and air purifier integration," *Mater. Proc.*, vol. 23, no. 1, Art. no. 12, 2025, doi: 10.3390/materproc2025023012.

[21] A. Jose *et al.*, "IoT based solar powered air purifier with air quality monitoring system," *E3S Web Conf.*, vol. 529, Art. no. 04016, 2024, doi: 10.1051/e3sconf/202452904016.

[22] S. M. Shakil Ahmed, "Hourly air quality index (AQI) of Bangladesh," Kaggle, 2025. [Online]. Available: https://doi.org/10.34740/KAGGLE/DSV/16464601

[23] U.S. Environmental Protection Agency, "Technical assistance document for the reporting of daily air quality—The air quality index (AQI)," Tech. Rep. EPA-454/B-24-002, May 2024. [Online]. Available: https://document.airnow.gov/technical-assistance-document-for-the-reporting-of-daily-air-quailty.pdf

[24] M. Anitha and L. S. Kumar, "Development of an IoT-enabled air pollution monitoring and air purifier system," *Mapan*, vol. 38, no. 3, pp. 669–688, 2023.