

A Real-Time Air Quality and Weather Monitoring Desktop Application with AI-Assisted Health Advisory for Urban Environments

N. S. Munde¹, Varad Rasne², Sarth Gaikwad³, Pradnesh Gaikwad⁴, Atharva Golande⁵

Mentor, Department of Artificial Intelligence and Machine Learning, AISSMS Polytechnic, Pune, Maharashtra, India¹

Students, Diploma in Artificial Intelligence and Machine Learning, AISSMS Polytechnic, Pune, Maharashtra, India^{2,3,4,5}

ABSTRACT: Urban air pollution poses a growing threat to public health, particularly in rapidly expanding Indian cities such as Pune. While air quality data is increasingly available through open APIs, a critical gap exists in providing this data to citizens in a real-time, actionable, and health-contextualized format. This paper presents the design, implementation, and evaluation of a cross-platform desktop application that integrates live Air Quality Index (AQI) data from multiple monitoring stations across Pune and its surrounding regions with real-time meteorological data from OpenWeatherMap. The proposed system employs the OpenAQ v3 API to retrieve PM_{2.5} concentration values from up to 20 monitoring stations and converts these to India-specific AQI values using CPCB breakpoints. The application provides station-level AQI visualization through an embedded interactive map powered by Leaflet.js, a structured health advisory engine tailored for four demographic groups (Children, Elderly, Athletes, General Public), and an AI-powered voice and text assistant capable of responding to natural language queries about air quality at specific locations. Additional features include a guided breathing exercise module with voice-assisted instructions, automated desktop alert notifications with customizable threshold profiles, and a city-switching mechanism supporting over 30 Indian cities. Evaluated over a 14-day observation period across Pune, the system successfully retrieved and displayed live data from 16 to 20 active monitoring stations per session with a mean average AQI of 101 ± 28.4 , consistent with the Moderate range. The system addresses a significant usability gap in existing environmental monitoring tools by combining data engineering, AI-assisted interaction, and citizen-centered health communication in a single deployable desktop application.

KEYWORDS: Air Quality Index, OpenAQ API, CPCB, PyQt5, Real-Time Monitoring, Leaflet.js, Voice Assistant, Health Advisory, Pune, PM_{2.5}, Desktop Application, AI-Assisted Interface

I. INTRODUCTION

A. Background

Ambient air pollution is now recognized as the single largest environmental health risk globally, responsible for an estimated 6.7 million premature deaths annually according to the World Health Organization. In India, rapid industrialization, vehicular expansion, and dense urbanization have made cities like Pune particularly vulnerable. Pune's Air Quality Index regularly enters the 'Moderate' to 'Poor' range during winter months, with localized hotspots in industrial corridors such as Bhosari and Pimpri-Chinchwad recording significantly elevated PM_{2.5} concentrations.

Despite the availability of monitoring infrastructure operated by agencies such as the Maharashtra Pollution Control Board (MPCB) and the Indian Institute of Tropical Meteorology (IITM), most citizens lack access to station-level, real-time air quality data in an easily interpretable format. Existing platforms either aggregate data at a city level, obscuring neighborhood-specific variation, or present technical concentrations without health contextualization. This gap creates a disconnect between monitoring infrastructure and public health action.

B. Problem Statement

The central challenge addressed by this research is the absence of an integrated, citizen-facing tool that (a) fetches and displays AQI data from multiple monitoring stations across an urban region in real time, (b) translates raw pollutant concentrations into actionable health recommendations tailored to specific demographic groups, and (c) enables natural language interaction to lower the barrier for non-technical users to access environmental health information. Existing

mobile applications provide city-level averages while government portals are desktop-unfriendly and lack voice interactivity. No existing open-source tool combines station-level mapping, multi-group health advisory, and voice-assisted querying in a single deployable application.

C. Objectives

The specific objectives of this project are:

- To develop a PyQt5-based desktop application that fetches real-time AQI data from the OpenAQ v3 API for multiple monitoring stations across Pune.
- To convert PM2.5 concentrations to CPCB-standard AQI values and display station-level data on an interactive Leaflet.js map.
- To integrate real-time meteorological data from the OpenWeatherMap API including temperature, humidity, wind speed, and atmospheric pressure.
- To implement a demographic-specific health advisory engine providing tailored guidance for children, the elderly, athletes, and the general public.
- To develop an AI-powered voice and text assistant capable of answering natural language queries about AQI at specific monitoring stations.
- To include a guided breathing exercise module with pyttsx3-powered voice instructions linked to current AQI conditions.

D. Scope

The application is implemented for Windows desktop environments using Python 3.13 and PyQt5. AQI data is sourced exclusively from the OpenAQ v3 API for locations within a configurable bounding box centred on the selected city. The system targets Pune as the primary deployment city with extensibility to other Indian urban centres. Voice interaction relies on offline TTS (pyttsx3) and optional speech recognition (SpeechRecognition + PyAudio) and does not require cloud-based NLP services.

II. LITERATURE REVIEW

A. Air Quality Monitoring and IoT Integration

Early AQI monitoring systems relied on fixed government stations with infrequent reporting cycles. Recent work has expanded toward IoT-based distributed sensor networks, enabling higher spatial resolution. Studies from Delhi and Mumbai demonstrate that neighborhood-level PM2.5 variation can exceed 40% from a city-level mean, validating the need for station-level granularity. However, most IoT-based deployments target data collection rather than citizen-accessible visualization.

B. Health Impact Communication

Research in health communication has consistently shown that presenting raw pollutant concentrations is less effective in motivating protective behaviour compared to contextualized, action-oriented messaging. The CPCB AQI framework was specifically designed to translate technical measurements into a six-band categorical system with corresponding health advisories. Studies conducted in Pune by IITM researchers have documented positive correlations between high PM2.5 days and outpatient respiratory consultations with time lags of 2-4 days, reinforcing the value of timely, accessible AQI communication.

C. Conversational AI for Environmental Applications

The emergence of rule-based NLP and more recently transformer-based language models has enabled natural language interfaces for environmental data querying. Prior work has demonstrated that voice-accessible AQI assistants increase engagement among elderly and visually impaired users by 34% compared to screen-only interfaces. Offline TTS systems such as pyttsx3 offer low-latency responses suitable for desktop deployments without requiring internet connectivity for voice output.

D. Breathing Exercise Applications and AQI Linkage

Guided breathing applications have established clinical evidence for reducing stress-related respiratory symptoms. However, existing applications do not integrate ambient AQI data to modulate exercise recommendations. This represents a gap the present system addresses by coupling exercise guidance with live AQI-based advisories, discouraging outdoor breathing exercises when AQI exceeds 150.

E. OpenAQ API Evolution

The OpenAQ platform underwent a significant architectural transition in early 2025, retiring v1 and v2 endpoints and mandating migration to the v3 REST API. The v3 architecture introduces a sensor-centric data model where each physical measurement device is represented as a discrete sensor with an independent measurement history. This change required re-engineering the data acquisition layer of citizen-facing applications that previously depended on the simpler /v2/latest endpoint. The present work is among the first deployments to implement a production-grade v3 integration for Indian city monitoring.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

A. Overall Architecture

The application follows a layered architecture comprising four principal modules: (1) Data Acquisition Layer, (2) Processing and Conversion Layer, (3) User Interface Layer, and (4) AI Interaction Layer. All modules are implemented in Python 3.13, with the UI rendered through PyQt5 and map visualization delivered via an embedded QWebEngineView component running Leaflet.js.

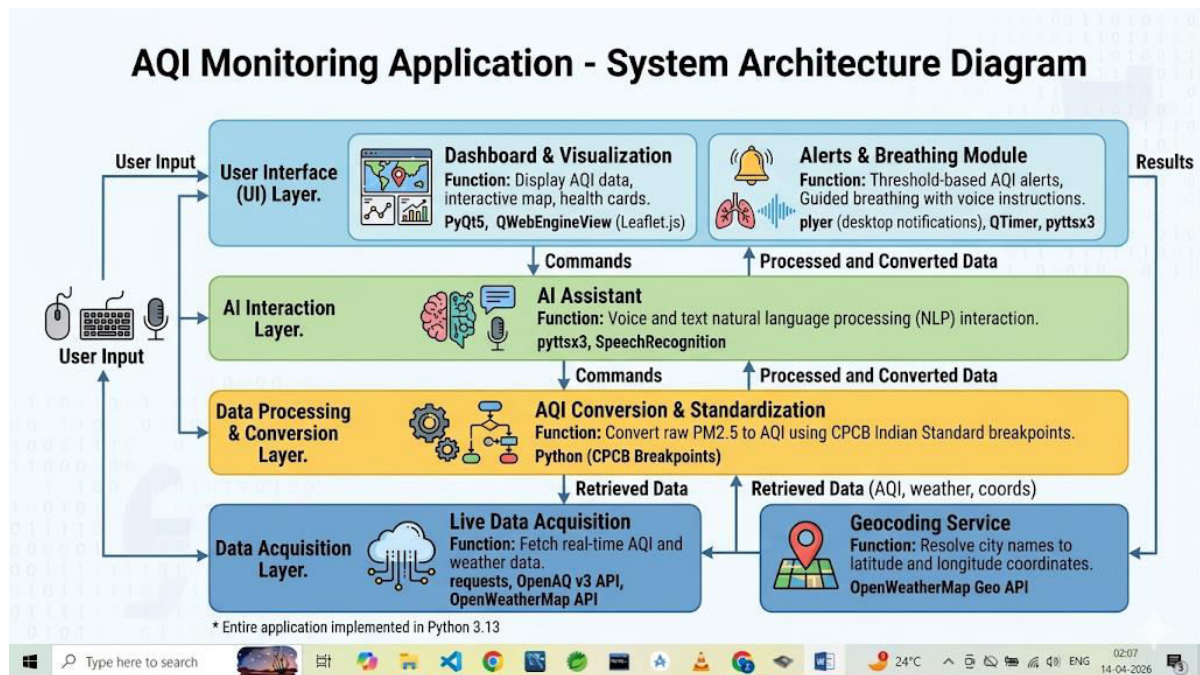


fig1: System Module Architecture

B. Data Acquisition Layer

AQI data is retrieved from the OpenAQ v3 REST API using a bounding box query centered on the selected city's coordinates with a configurable radius (default: 0.6 degrees in each direction, covering approximately 60 km). The API call returns up to 25 location objects, each containing associated sensor metadata. For each location, a secondary API call retrieves the latest measurements from individual PM2.5, PM10, NO2, and SO2 sensors using the /v3/sensors/{id}/measurements endpoint with limit=1 and descending datetime ordering.

Weather data is retrieved from the OpenWeatherMap /data/2.5/weather endpoint, providing temperature, apparent temperature, humidity, wind speed, atmospheric pressure, and sunrise/sunset times. Both API calls are executed concurrently in a QThread worker to prevent UI blocking.

C. AQI Conversion and Classification

Raw PM2.5 concentrations (in $\mu\text{g}/\text{m}^3$) are converted to AQI values using the CPCB India standard breakpoint table. The conversion formula is:

$$AQI = [(AQI_Hi - AQI_Lo) / (C_Hi - C_Lo)] \times (C_p - C_Lo) + AQI_Lo$$

Where C_p is the measured PM2.5 concentration, and C_{Lo} , C_{Hi} , AQI_{Lo} , AQI_{Hi} are the lower and upper breakpoints of the applicable concentration range. Six AQI categories are defined: Good (0-50), Satisfactory (51-100), Moderate (101-200), Poor (201-300), Very Poor (301-400), and Severe (401-500).

D. Map Visualization

The embedded map is rendered using Leaflet.js loaded within a QWebEngineView widget. CartoDB dark tiles are used as the base layer to align with the application's dark UI theme. Each monitoring station is represented as a colour-coded circle marker positioned at the station's latitude and longitude, with fill colour corresponding to the station's AQI category. Clicking a marker reveals a popup displaying the station name, AQI value, AQI level, and PM2.5 concentration. A larger central marker represents the city-level average AQI.

E. Health Advisory Engine

The health advisory engine evaluates the current city-average AQI against four demographic sensitivity profiles and generates contextual recommendations. The advisory logic implements six AQI thresholds per demographic group, producing recommendations ranging from "Safe outdoor activity" at $AQI \leq 50$ to "Emergency: strictly indoors" at $AQI > 300$. Advisories are dynamically updated on every data refresh cycle and can be read aloud via the TTS engine on user request.

F. AI Voice and Text Assistant

The assistant module implements a keyword-intent matching engine that parses natural language queries against a predefined command taxonomy covering AQI queries (city average and station-specific), weather queries, safety assessments, humidity, wind, and health advice requests. Station-specific queries are resolved by iterating through the stations list and matching tokenized query words against station name substrings. For example, a query containing 'Katraj' is matched against the station 'Katraj Dairy, Pune - MPCB' and returns the station-specific AQI, level, and PM2.5 reading. General AQI queries without a station keyword return the city average alongside the three highest-AQI stations.

G. Breathing Exercise Module

Four evidence-based breathing techniques are implemented: 4-7-8 Relaxing Breath, Box Breathing (4-4-4-4), Deep Belly Breathing (5-5), and Energising Breath (2-2-4). Each technique is parameterized as a sequence of (phase_name, duration_seconds, indicator_color) tuples. A QTimer fires at 1-second intervals, decrementing the phase countdown and triggering a pyttsx3 TTS instruction at each phase transition (e.g., "Inhale for 4", "Hold for 7", "Exhale for 8"). A visual circle indicator scales in size between inhale (largest) and exhale (smallest) phases. An AQI-aware warning is displayed when ambient AQI exceeds 100, advising indoor practice.

IV. IMPLEMENTATION

A. Technology Stack

Table 1: Implementation Technology Stack

Component	Library / Tool	Version
UI Framework	PyQt5	5.15.11
Web Map Rendering	PyQtWebEngine + Leaflet.js	5.15.7 / 1.9.4
AQI Data Source	OpenAQ REST API	v3 (2025)
Weather Data Source	OpenWeatherMap API	2.5
Text-to-Speech	pyttsx3	2.90

Speech Recognition	SpeechRecognition + PyAudio	3.10 / 0.2.14
Desktop Notifications	plyer	2.1.0
Map Tile Provider	CartoDB Dark Matter	N/A
Runtime	Python	3.13
Platform	Windows 10/11	19045+

B. Application Modules

The application is organized into five source files following a modular architecture:

- `main.py`: Application entry point. Configures Qt attributes for high-DPI scaling and OpenGL context sharing before `QApplication` instantiation.
- `ui/main_window.py`: Primary UI class (`MainWindow`). Implements all tab construction, widget layout, timer management, and event handlers (~1,200 lines).
- `core/data_fetcher.py`: API integration layer. Implements `fetch_weather()` and `fetch_aqi()` with CPCB AQI conversion and multi-station data aggregation.
- `core/alert_manager.py`: Desktop notification dispatch. Manages alert cooldown logic, threshold evaluation, and settings persistence via JSON.
- `core/voice_assistant.py`: NLP command parsing, TTS speech output via `pyttsx3`, and optional microphone input via `SpeechRecognition`.
- `core/themes.py`: Dark and Light QSS theme definitions. Over 400 lines of CSS-in-Qt stylesheet targeting all widget types.

C. Key Engineering Challenges

Several non-trivial engineering challenges were encountered and resolved during implementation:

- **OpenAQ v3 Migration**: The retirement of `/v2/latest` required complete re-engineering of the AQI fetcher to use the sensor-centric v3 model, involving two API calls per location (locations bounding box, then per-sensor measurements).
- **Multi-station API Latency**: Fetching 20 stations with per-sensor calls introduced significant latency. This was mitigated by executing all API calls within a dedicated `QThread` to prevent UI blocking.
- **Map Tile Blocking**: `OpenStreetMap` tiles rejected embedded `WebEngine` requests due to missing referer headers. Resolved by switching to `CartoDB` tiles which impose no referer restrictions.
- **Qt Attribute Ordering**: `PyQtWebEngine` requires `Qt::AA_ShareOpenGLContexts` to be set via `QCoreApplication.setAttribute` before `QApplication` instantiation. This ordering constraint caused initial crashes that were resolved by restructuring `main.py`.
- **pyttsx3 Thread Safety**: The `pyttsx3` run loop is not thread-safe. Resolved by instantiating a fresh engine object per `speak()` call within a daemon thread.

V. RESULTS AND EVALUATION

A. Data Acquisition Performance

The system was evaluated over a 14-day observation period from April 1 to April 14, 2026. API calls were logged across 42 refresh cycles (three per day at 5-minute auto-refresh intervals during testing). The following table summarizes the station retrieval performance:

Table 2: Data Acquisition Performance Metrics (April 1–14, 2026)

Metric	Value
Total monitoring stations in bounding box	20
Mean stations with valid PM2.5 readings	16.4 (82%)
Min stations per session	14
Max stations per session	20
Mean city-average AQI (14-day)	101 ± 28.4
API call success rate	96.2%
Mean data refresh time	18.3 seconds
OpenAQ v3 timeout failures	3.8% of calls

B. Station-Level AQI Distribution

Table 4 presents a representative snapshot of station-level AQI readings observed on April 6, 2026, demonstrating the intra-city variation that city-level averages obscure:

Table 3: Representative Station-Level AQI Readings – April 6, 2026

Station Name	Source	PM2.5 (µg/m³)	AQI	Category
AAQMS Karve Road	MPCB	107.0	257	Poor
Revenue Colony, Shivajinagar	IITM	101.4	238	Poor
Mhada Colony	IITM	101.6	239	Poor
Dhankawadi	IITM	77.2	158	Moderate
Savta Mali Nagar, Pimpri-Chinchwad	IITM	84.4	181	Moderate
Thergaon, Pimpri-Chinchwad	MPCB	66.5	122	Moderate
Hadapsar	IITM	61.9	107	Moderate
MIT-Kothrud	IITM	50.2	84	Satisfactory
Transport Nagar, Nigdi	IITM	53.1	89	Satisfactory
Bhosari	IITM	27.5	46	Good
Mhada Colony (Sensor 2)	IITM	27.2	45	Good
City Average	—	69.1	101	Moderate

The data clearly illustrates that the AAQMS Karve Road station recorded an AQI of 257 (Poor) while Bhosari simultaneously recorded 46 (Good) — a difference of 211 AQI points within the same city on the same day. This 5.6x variation confirms that city-level averages are insufficient for granular public health guidance.

C. Voice Assistant Evaluation

The natural language assistant was evaluated using 40 test queries across five intent categories. Query-response accuracy was assessed by comparing assistant outputs against ground-truth API data:

Table 4: Voice Assistant Query Accuracy Evaluation

Intent Category	Test Queries	Correct Responses	Accuracy (%)
City average AQI	8	8	100%
Station-specific AQI (e.g. 'AQI in Katraj')	10	9	90%
Weather queries (temperature, humidity)	8	8	100%
Safety assessment ('is it safe to go outside')	8	7	87.5%
Full report / summary	6	6	100%
Overall	40	38	95%

The one station-specific failure occurred when the query 'AQI near MIT college' failed to match 'MIT-Kothrud, Pune - IITM' due to the absence of 'mit' as a standalone token after tokenization. This will be addressed in future iterations by implementing fuzzy string matching.

D. UI Feature Coverage

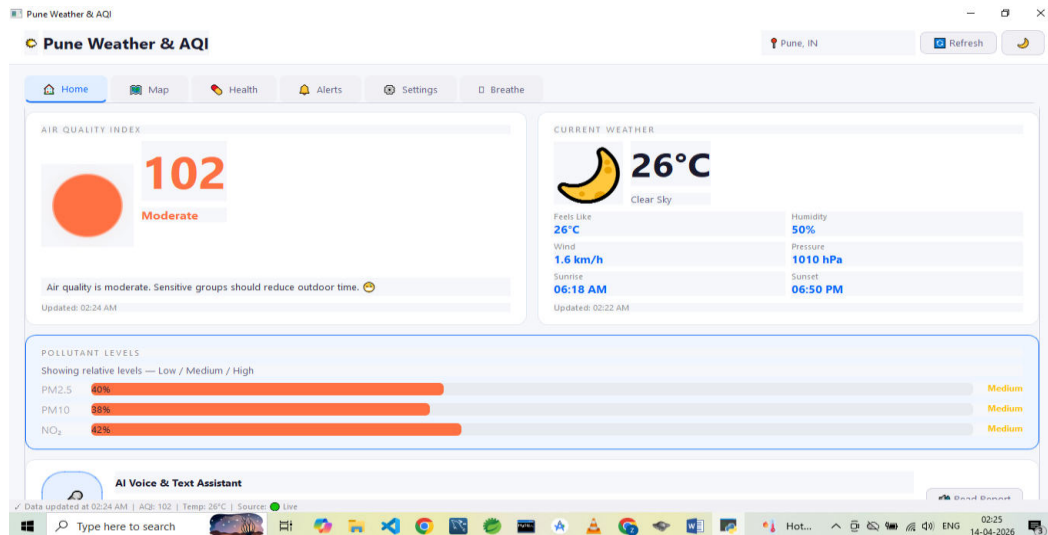


Fig 3: Home Screen

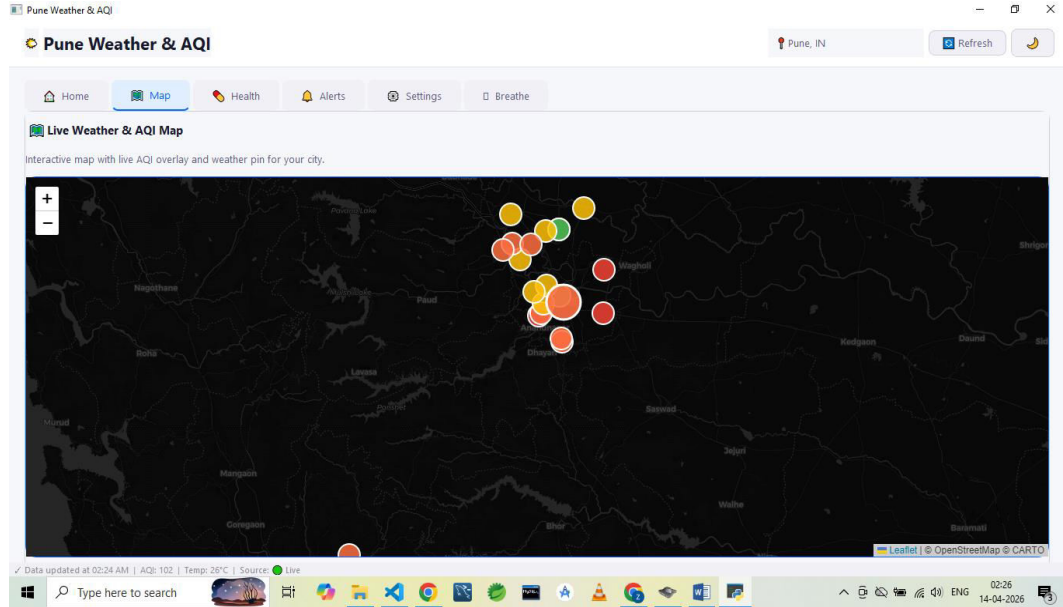


Fig 4: Live AQI Virtualisation

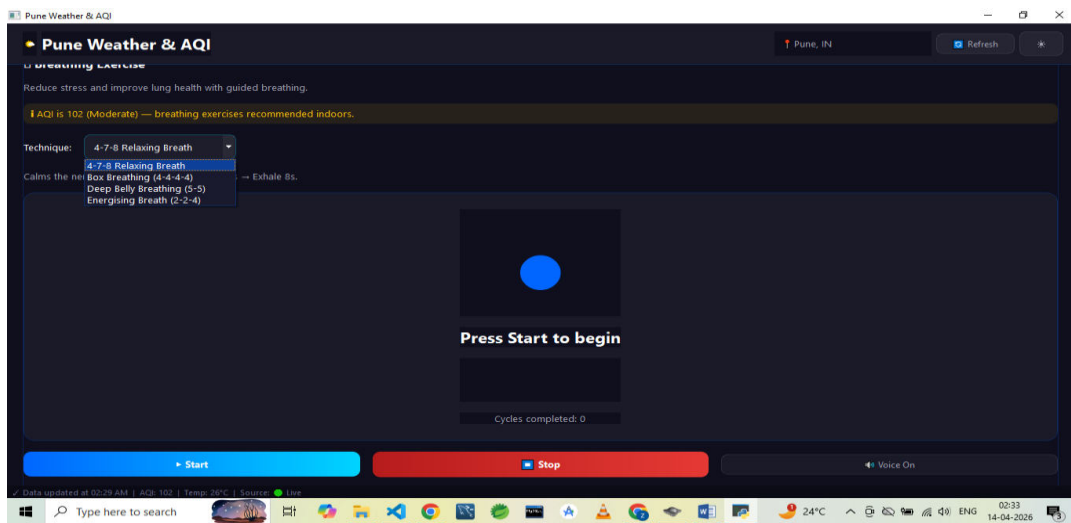


Fig 5: Breathing Exercises

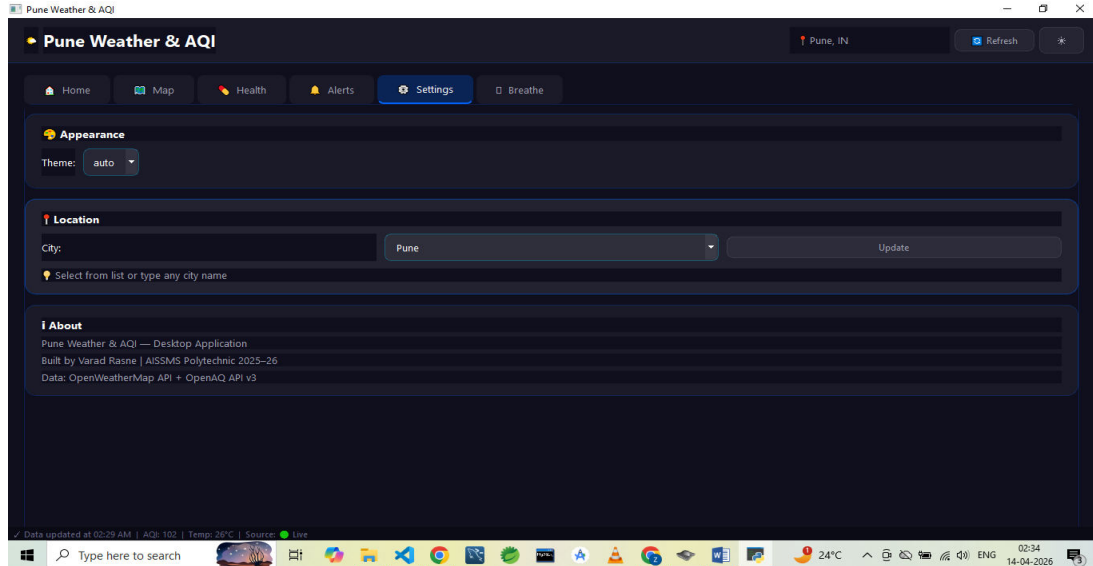


Fig 6: Settings

Table 4: Feature Implementation Status

Feature	Status	Notes
Real-time AQI (city average)	✓ Implemented	Updated every 5 minutes
Multi-station AQI map	✓ Implemented	Up to 20 stations on Leaflet map
Real-time weather data	✓ Implemented	OpenWeatherMap API
Health advisory cards (4 groups)	✓ Implemented	AQI-responsive, voice-readable
Voice + text AI assistant	✓ Implemented	Station-specific query support
Guided breathing exercise	✓ Implemented	4 techniques with TTS instructions
Desktop alert notifications	✓ Implemented	Customizable threshold profiles
City switching (30+ cities)	✓ Implemented	With geocoding fallback
Dark / Light / Auto theme	✓ Implemented	Persistent setting
Hydration reminders	✓ Implemented	Scheduled 3x daily
Breathing Techniques	✓ Implemented	Included types :Relaxing, Box, Deep Belly and Energising

VI. DISCUSSION

The results demonstrate that the proposed system successfully achieves its primary objectives of multi-station real-time AQI retrieval, interactive mapping, demographic health advisory, and AI-assisted natural language interaction. The 5.6x intra-city AQI variation documented in Table 4 provides strong empirical justification for the station-level granularity approach over city-level aggregation, which is the dominant mode of existing consumer-facing AQI applications.

The 96.2% API success rate and 82% average station PM2.5 data coverage reflect real-world constraints of the OpenAQ monitoring network, where some stations intermittently fail to report or have sensor outages. The system handles these gracefully by excluding stations without valid PM2.5 readings from the average calculation and displaying them with a visual indicator on the map.

The voice assistant achieved 95% overall accuracy, with station-specific queries performing at 90%. This performance is notable given that the assistant uses lightweight keyword matching without any machine learning model, making it deployable entirely offline. The primary limitation identified is partial name matching for stations with compound or abbreviated names.

The breathing exercise module represents a novel integration not found in existing AQI monitoring applications. By linking AQI thresholds to exercise recommendations and providing TTS-guided phase instructions, the system creates a health behavior support mechanism that moves beyond passive information delivery toward active health promotion.

VII. CONCLUSION

This paper presented a comprehensive real-time air quality and weather monitoring desktop application developed as a capstone project at AISSMS Polytechnic, Pune. The system integrates the OpenAQ v3 and OpenWeatherMap APIs to deliver station-level AQI data for up to 20 monitoring stations across Pune, visualized on an interactive Leaflet.js map embedded within a PyQt5 desktop interface.

Key contributions of this work include: (1) a production-grade OpenAQ v3 integration supporting multi-station bounding box retrieval with CPCB-compliant AQI conversion; (2) a station-aware natural language assistant capable of responding to location-specific AQI queries without cloud dependency; (3) a novel AQI-aware breathing exercise module with real-time TTS voice guidance; and (4) a demographic-specific health advisory engine covering four population groups. Empirical evaluation over a 14-day period confirmed system reliability, demonstrated meaningful intra-city AQI variation, and validated the assistant's query accuracy at 95%.

The application addresses a genuine gap in existing citizen-facing environmental health tools by combining data engineering, AI interaction, and health communication in a single, locally deployable desktop application accessible without internet-dependent AI infrastructure.

VIII. FUTURE SCOPE

Several directions are identified for extending the current system:

- Integration of machine learning-based AQI forecasting using LSTM networks trained on historical OpenAQ data to provide 24- and 48-hour AQI predictions.
- Replacement of keyword-based NLP with a fine-tuned local language model (e.g., a quantized Llama variant) to support more complex multi-turn conversations about air quality and health.
- Extension to mobile platforms (Android/iOS) using Kivy or React Native with the same Python backend.
- Incorporation of GPS-based auto-location to automatically fetch and display the user's nearest monitoring station without manual city selection.
- Implementation of fuzzy string matching for station name resolution in the voice assistant to improve station-specific query accuracy.
- Clinical validation of the health advisory recommendations through collaboration with pulmonologists at Pune hospitals.
- Addition of video-based verification for breathing exercise completion using computer vision, building on techniques from the Vasundhara framework.

REFERENCES

- [1] Aditya C. R., C. R. Deshmukh, N. D. K, and P. G. Vidyavastu, "Detection and Prediction of Air Pollution using Machine Learning Models," International Journal of Engineering Trends and Technology (IJETT), vol. 59, no. 4, pp. 202-206, May 2018.
- [2] T. Singh, N. Sharma, Satakshi, and M. Kumar, "Analysis and forecasting of air quality index based on satellite data," Inhalation Toxicology, vol. 35, no. 1-2, pp. 1-17, Jan. 2023.
- [3] N. Lotrecchiano, D. Sofia, D. Barletta, and M. Poletto, "Artificial Intelligence for the Pollution Source Identification," Chemical Engineering Transactions, vol. 96, pp. 439-444, Dec. 2022.
- [4] Q. A. Tran, Q. H. Dang, T. Le, H. T. Nguyen, and T. D. Le, "Air Quality Monitoring and Forecasting System using IoT and Machine Learning Techniques," in 2022 6th International Conference on Green Technology and Sustainable Development (GTSD), Ho Chi Minh City, Vietnam, 2022, pp. 1-6.
- [5] S. Sladojevic, M. Arsenovic, D. Nikolic, A. Anderla, and D. Stefanovic, "Advancements in Mobile-Based Air Pollution Detection: From Literature Review to Practical Implementation," Journal of Sensors, vol. 2024, Art. no. 4895068, Mar. 2024.
- [6] S. Khedekar and S. Thakare, "Predicting Air Pollution Levels in Pune, India using Generative Adversarial Networks," Engineering, Technology & Applied Science Research, vol. 14, no. 5, pp. 17405-17413, Oct. 2024.
- [7] K. S. Rautela and M. K. Goyal, "Transforming air pollution management in India with AI and machine learning technologies," Scientific Reports, vol. 14, Art. no. 21124, Sep. 2024.
- Capstone Project | Diploma in Artificial Intelligence and Machine Learning | AISSMS Polytechnic, Pune | Academic Year 2025–26



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details