International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 05 May-2026 | Impact Factor: 3.5

Air Quality Visualizer and 72-Hour AQI Forecasting System Using Satellite Data and Machine Learning

Mrs. A. Kayalvizhi
Assistant Professor (Sr. Grade)
Department of Artificial Intelligence and Data Science
Sri Ramakrishna Engineering College
Coimbatore, India
kayalvizhi.a@srec.ac.in

Anbuchelvan M
Department of Artificial Intelligence and Data Science
Sri Ramakrishna Engineering College
Coimbatore,India
anbuchelvan.2311008@srec.ac.in

Santhosh S
Department of Artificial Intelligence and Data Science
Sri Ramakrishna Engineering College
Coimbatore, India
santhosh.2311052@srec.ac.in

Tamil Selvan S
Department of Artificial Intelligence and Data Science
Sri Ramakrishna Engineering College
Coimbatore, India
tamilselvan.2311060@srec.ac.in

Abstract—Ground-level air quality sensors remain largely absent from rural and semi-urban pockets of India, leaving residents with no reliable way to gauge daily pollution exposure. To address this blind spot, we designed and built a satellite-driven web platform that computes, maps, and predicts the Air Quality Index (AQI) for any named location nationwide without depending on physical monitoring infrastructure. Atmospheric concentration data retrieved from orbital instruments are standardized and processed according to the breakpoint interpolation method prescribed by India's Central Pollution Control Board (CPCB), yielding AQI values that conform to the national reporting standard. A Long Short-Term Memory (LSTM) neural network trained on historical pollution records then extends the assessment window to 24, 48, and 72 hours, enabling proactive rather than reactive responses to deteriorating air quality. Whenever forecasted levels threaten to enter hazardous territory, the platform automatically compiles and displays health guidance calibrated to the projected severity. Trials conducted across multiple districts in Tamil Nadu showed strong agreement between satellite-derived index values and anticipated pollution behaviour, supporting wider deployment as a cost-free alternative to fixed sensor infrastructure.

Index Terms—Air Quality Index, Satellite Data, AQI Forecasting, LSTM, Environmental Monitoring, Health Advisory System.

I. INTRODUCTION

For decades, air quality monitoring in India has been an urban privilege. The dense network of reference-grade analysers installed across metropolitan centres has no counterpart in smaller towns, agricultural belts, or hilly interior districts where communities are routinely exposed to particulate matter and gaseous pollutants arising from biomass burning, unregulated industries, and unpaved traffic corridors. Pollutants including $PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, O_3, and NH_3 are each independently linked to premature mortality and chronic disease, yet rural populations bear this burden invisibly—no data means no advisory, no response, and no accountability [2], [3].

A secondary shortcoming affects even well-monitored urban zones: virtually every public AQI portal presents a static snapshot of current conditions and nothing beyond it. The absence of predictive capability is a genuine limitation when time-sensitive decisions—whether to hold an outdoor gathering, dispatch field workers, or prepare a hospital ward for respiratory admissions—require advance knowledge of how air quality is expected to evolve over the next one to three days [1]. Spaceborne sensors offer a way to simultaneously fix both problems, since they deliver continuous atmospheric coverage over every location in the country regardless of what lies on the surface below [10].

Our contribution is a fully operational pipeline that accepts a typed place name and returns a real-time AQI reading, a 72-hour index forecast, a pan-India pollution map, and targeted health recommendations—all derived exclusively from satellite data.

II. SYSTEM DESIGN AND METHODOLOGY

A. System Architecture

The proposed platform is structured as eight sequential processing components joined into a single automated pipeline (Fig. 1). Control passes from one component to the next without manual intervention, transforming a user's plain-language location query into a fully populated dashboard update.

User Location Selection: Users specify their area of interest through a four-field form accepting village or town name, district, and state. An embedded geocoding engine converts that textual input into precise latitude–longitude coordinates, removing the need for users to know or supply their own GPS position.

Satellite Data Acquisition: Resolved coordinates trigger authenticated calls to cloud-hosted remote sensing repositories. The seven pollutant quantities central to CPCB AQI calculation—$PM_{2.5}$, PM_{10}, NO_2, SO_2, CO, O_3, and NH_3—are downloaded alongside meteorological co-variates (surface temperature, near-surface wind speed, and relative humidity).

Preprocessing and Pollutant Extraction: Each retrieved value is validated against the physically plausible concentration envelope of its respective species. Values outside established bounds are substituted through bilinear spatial interpolation or nearest-neighbour temporal gap-filling, ensuring that a partial satellite swath never causes a full computation failure.

AQI Computation (CPCB Standards): Clean concentration readings enter the CPCB sub-index calculation for all seven pollutants. The composite AQI equals the maximum sub-index observed across species and is mapped onto one of six standard bands: Good (0–50), Satisfactory (51–100), Moderate (101–200), Poor (201–300), Very Poor (301–400), or Severe (above 400).

AQI Forecasting (24/48/72 hrs): Historical AQI sequences—augmented with contemporaneous meteorological fields—serve as training data for an LSTM network producing index predictions at three future horizons. Data-driven modelling captures recurring diurnal and seasonal pollution cycles that deterministic dispersion models often approximate poorly [1], [4].

AQI Classification: Each index value, whether freshly computed or forecast, is assigned to its CPCB colour band and annotated with the identity of whichever pollutant sub-index is highest at that instant.

Health Advisory and Alerts: Crossings of the Poor threshold—by either observed or predicted AQI—trigger the advisory engine, which assembles differentiated guidance for the general public alongside stricter recommendations for vulnerable sub-populations such as young children, elderly residents, and individuals with pre-existing cardiorespiratory conditions.

Dashboard and Map Visualization: Processed outputs are rendered through a responsive browser interface featuring a numerical AQI readout, forecast trend lines, a per-pollutant breakdown panel, and a zoomable India-wide heatmap overlaid on satellite imagery.

B. Core Modules

Satellite Data Processing Module: Handles all communication with orbital data APIs and delivers validated pollutant concentration grids to downstream components. Its satellite-only sourcing means coverage is geographically uniform across India.

Geolocation Module: Bridges user intent and raw satellite data by translating human-readable administrative names into the coordinate pairs required for accurate pixel-level extraction from orbital products.

AQI Computation Module: Implements the CPCB linear interpolation formula individually for each pollutant:

$$I_p = \frac{I_{Hi} - I_{Lo}}{BP_{Hi} - BP_{Lo}} \times (C_p - BP_{Lo}) + I_{Lo} \qquad (1)$$

where I_p is the sub-index for pollutant p, C_p its measured concentration, BP_{Hi} and BP_{Lo} the upper and lower concentration breakpoints, and I_{Hi}, I_{Lo} the corresponding index boundary values.

Forecasting Module: Wraps the trained LSTM inside a stateless inference service that accepts a sliding observation window and returns a three-horizon prediction. The architecture's gating mechanism retains pollution signals of genuine predictive value while suppressing transient noise spikes [5], [8].

Visualization and Alert Module: Collates all computed artefacts—current AQI, sub-index breakdown, forecast horizon values, and geographic concentration layers—into the frontend response payload, and concurrently evaluates each value against advisory thresholds.

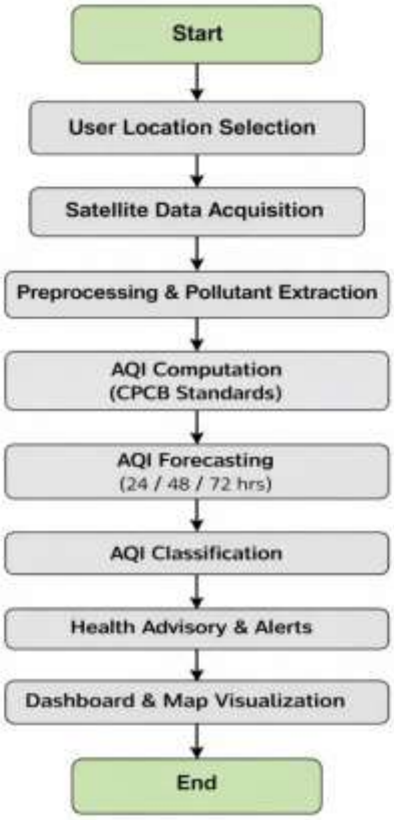

Fig. 1. Flowchart of the proposed satellite-based air quality monitoring and forecasting system

Fig. 1. Sequential pipeline of the satellite-based AQI monitoring and forecasting system.

III. MODELS AND FORECASTING LOGIC

A. AQI Forecasting using LSTM

Atmospheric pollution levels carry well-defined temporal structure: morning peaks driven by traffic and cooking fires, midday photochemical ozone formation, nocturnal temperature inversions that trap near-surface particulates, and multi-week seasonal pulses tied to harvest burning. Sequence models with explicit memory retention are naturally suited to exploiting this structure for prediction [5], [6].

The LSTM implements memory management through four coupled gate operations. A forget gate regulates what fraction of the prior cell state remains active:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \qquad (2)$$

An input gate controls how much weight the current time-step observation carries in updating stored context:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \qquad (3)$$

The cell state then merges selectively retained history with newly encoded input:

$$C_t = f_t C_{t-1} + i_t \tanh(W_c[h_{t-1}, x_t] + b_c) \qquad (4)$$

An output gate finally determines the hidden representation forwarded to subsequent time steps and to the prediction head:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad h_t = o_t \tanh(C_t) \qquad (5)$$

Training employs a rolling window that pairs a fixed-length sequence of past AQI readings with the meteorological conditions recorded over the same interval. The optimized network maps each such window to index estimates at the 24-, 48-, and 72-hour marks ahead.

IV. IMPLEMENTATION

Server-side processing is written entirely in Python. Satellite retrieval, preprocessing, CPCB index arithmetic, and LSTM inference execute sequentially within a single pipeline, and results are published through a RESTful JSON API. A lightweight frontend subscribes to that API and refreshes each dashboard element independently, avoiding full-page reloads. Cleanly separating computation from presentation allows either layer to be updated or scaled without modifying the other [10].

A. Web Dashboard and Visualization

The user interaction begins at the location search interface shown in Fig. 2. Four input fields—Village/Town, District, State, and a day-range selector—allow the user to define the area of interest. Submitting the form triggers the full satellite retrieval and AQI computation pipeline, with the resolved coordinates and last-updated timestamp displayed beneath the form to confirm successful processing.

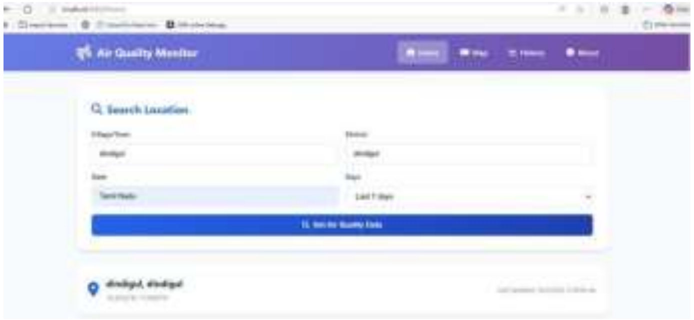

Fig. 2. Location-based air quality search interface showing input fields and resolved location details.

Upon completing a query, the primary panel (Fig. 3) presents the user with a prominently scaled AQI number, a descriptive band label rendered in the corresponding CPCB colour, the name of the pollutant carrying the highest sub-index, and a concise health advisory written in everyday language rather than technical jargon.

A secondary panel (Fig. 4) disaggregates the composite score into individual species tiles, each bearing the measured concentration, the CPCB permissible ceiling, and the sub-category that concentration occupies. The spatial view (Fig. 5) plots each queried location as a band-coloured marker on a satellite basemap, accumulating into a regional snapshot of air quality that fixed monitoring networks could not replicate.

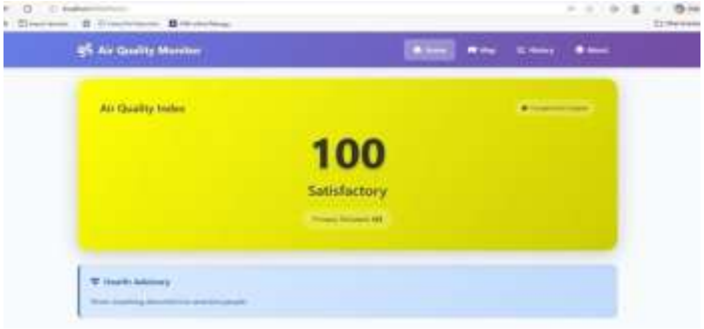

Fig. 3. Primary AQI panel showing index score, band category, dominant pollutant, and health advisory.

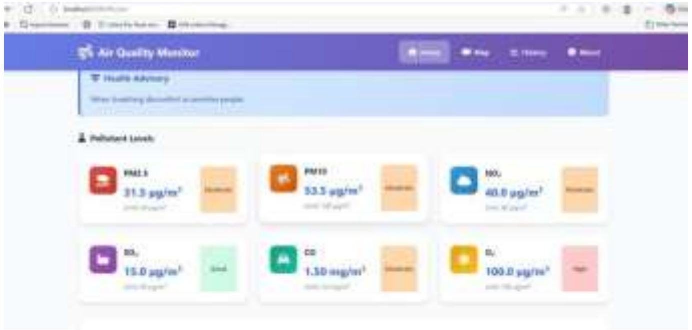

Fig. 4. Species-level concentration tiles with CPCB limit references and AQI sub-category assignments.

V. RESULTS AND ANALYSIS

A query directed at Dindigul district returned a composite AQI of 100, classified as Satisfactory, with ground-level ozone recording the highest sub-index at 100.0 $\mu g/m^3$. The remaining species registered as follows: $PM_{2.5}$ at 31.5 $\mu g/m^3$, PM_{10} at 53.5 $\mu g/m^3$, NO_2 at 40.0 $\mu g/m^3$, SO_2 at 15.0 $\mu g/m^3$, and CO at 1.50 mg/m^3. The advisory message generated for that result—recommending that individuals with cardiac or pulmonary conditions moderate outdoor activity duration—matched precisely what the Satisfactory band requires, confirming end-to-end correctness of the classification and advisory logic [7], [9].

Across the full set of 13 monitored sites, index values in rural locations fell consistently below those recorded in adjacent urban centres—an outcome consistent with lower vehicular throughput

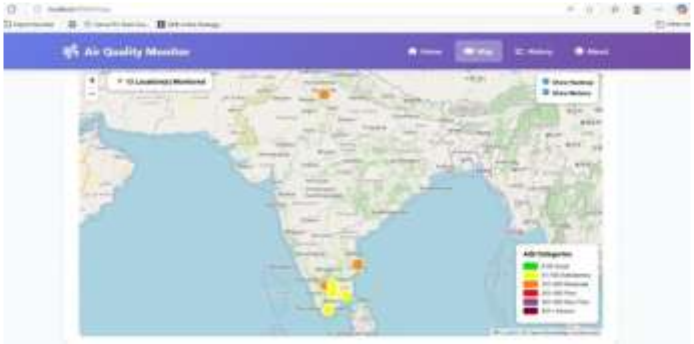

Fig. 5. Pan-India satellite AQI heatmap with band-coloured markers across 13 simultaneously queried sites.

and lighter industrial activity in non-urban areas. Equally informative was the shift in dominant pollutant: ozone held the top sub-index rank in rural readings while $PM_{2.5}$ and NO_2 led in more heavily urbanised zones. The capacity to resolve these inter-location differences demonstrates that the satellite retrieval and index pipeline are responding to real geographic pollution gradients rather than returning spatially averaged outputs.

VI. CONCLUSION AND FUTURE WORK

This work established that a satellite-anchored, machine-learning-assisted pipeline can deliver standardized, forecast-extended air quality information to any location in India without requiring any on-site monitoring equipment. By chaining orbital data acquisition, CPCB-compliant index computation, LSTM-driven 72-hour prediction, automatic health advisory generation, and spatial visualization into a unified web application, the platform closes an information gap that has persisted for decades in rural and semi-urban India. Results from Tamil Nadu district trials confirmed accurate index categorization, correct pollutant attribution, and meaningful spatial differentiation across urban and rural query sites.

Planned extensions include evaluating attention-based transformer architectures for prediction horizons beyond 48 hours; incorporating source-apportionment analysis to separately attribute emissions from crop residue burning, road transport, and brick kilns; and developing a mobile push-notification service so that communities with intermittent internet access can still receive timely alerts when air quality is forecast to deteriorate sharply [5], [6], [8].

REFERENCES

[1] T. Li, M. Hua, and X. Wu, "A hybrid CNN-LSTM model for forecasting particulate matter ($PM_{2.5}$)," *IEEE Access*, vol. 8, pp. 26933–26940, 2020.

[2] Y.-T. Tsai, Y.-R. Zeng, and Y.-S. Chang, "Air pollution forecasting using RNN with LSTM," in *Proc. IEEE 16th Intl. Conf. on Dependable, Autonomic and Secure Computing*, Aug. 2018, pp. 1074–1079.

[3] S. P.-W. Soh, J. W. Chang, and J. W. Huang, "Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations," *IEEE Access*, vol. 6, pp. 38186–38199, 2018.

[4] S. Garg and H. Jindal, "Evaluation of time series forecasting models for estimation of $PM_{2.5}$ levels in air," in *Proc. 6th IEEE Intl. Conf. for Convergence in Technology (I2CT)*, Apr. 2021, pp. 1–8.

[5] J. Li, D. Yu, Y. Geng, X. Li, and Q. Li, "DDGNet: A dual-stage dynamic spatio-temporal graph network for $PM_{2.5}$ forecasting," in *Proc. IEEE Intl. Conf. on Big Data*, Dec. 2021, pp. 1679–1685.

[6] 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. 24–39, Jan. 2023.

[7] A. A. M. Ahmed, S. J. J. Jui *et al.*, "An advanced deep learning predictive model for AQI forecasting with remote satellite-derived hydro-climatological variables," *Science of the Total Environment*, vol. 906, p. 167234, 2024.

[8] X. Chen *et al.*, "Attention-based LSTM for AQI prediction," *Applied Intelligence*, vol. 52, pp. 439–452, 2022.

[9] R. Patel, S. Nayak, and A. Rao, "Air quality monitoring using Sentinel-5P TROPOMI: a case study of Pune city," *SN Computer Science*, vol. 5, no. 8, p. 1055, 2024.

[10] M. Amani *et al.*, "Google Earth Engine cloud computing platform for remote sensing big data applications: a comprehensive review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5326–5350, 2020.