

AirGuide AI: An Agentic Outdoor Activity Planner

Vishal Sivakumar

St. Francis High School, 1885 Miramonte Ave, Mountain View, CA 94040, USA

Email: vishalsivakumar1010@gmail.com

ABSTRACT

Outdoor activities such as walking, running, and cycling provide significant health benefits, but planning these activities can be challenging due to multiple air pollutants, rapidly changing conditions, and individual health sensitivities. Existing air-quality applications primarily provide city-level monitoring information and lack any personalized guidance for outdoor activities. This study presents AirGuide AI, an intelligent decision-support system that integrates multi-pollutant environmental assessment, hyperlocal PM_{2.5} forecasting, EPA-aligned health guidance, and agentic reasoning to generate personalized outdoor-activity recommendations. Hyperlocal forecasting models were developed using 30 months of data from 20 PurpleAir sensors in Fremont, California. XGBoost models combined recent local PM_{2.5} history, neighboring-sensor conditions, and meteorological information to predict concentrations one and three hours ahead, with persistence forecasting incorporated into the final ensemble. On a forward holdout dataset of 84,006 observations, the system achieved MAEs of 0.760 and 1.278 $\mu\text{g}/\text{m}^3$ at one-hour and three-hour horizons. Scenario-based evaluation demonstrated how AirGuide AI adapts recommendations based on pollutant type, forecast conditions, hyperlocal variation, and user health sensitivity, including changes in activity timing or intensity, cleaner nearby locations, and lower-exposure routes. These results demonstrate how hyperlocal environmental sensing, machine-learning forecasting, and personalized reasoning can be integrated to translate air-quality information into actionable guidance for outdoor-activity planning.

KEYWORDS

Air Quality; PM_{2.5}; Hyperlocal Forecasting; Machine Learning; Outdoor Activity Planning; Agentic AI

INTRODUCTION

Air pollution remains a major environmental and public health challenge worldwide, contributing to millions of deaths and substantial economic burden each year [1,2]. The United States Environmental Protection Agency (EPA) monitors several pollutants known to affect human health, including fine particulate matter (PM_{2.5}), coarse particulate matter (PM₁₀), ozone (O₃), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), and carbon monoxide (CO). The Air Quality Index (AQI) translates pollutant concentrations into a standardized scale representing potential health risk [3]. AQI values are widely reported by government agencies, weather services, and mobile applications to help individuals make informed decisions about outdoor activities such as walking, running, and cycling.

While these systems have significantly improved public awareness of air quality conditions, several limitations remain. First, reported values are often derived from monitoring stations that may be located far from a user's actual location. Second, AQI values are based on pollutant-specific averaging periods (e.g., 8-hr or 24-hr averages) that may not fully reflect rapidly changing environmental conditions. Third, air quality information from these sources is typically reported at city or regional scales, which can mask substantial neighborhood-level variations in pollutants such as PM_{2.5} caused by traffic patterns, localized emissions, and meteorological conditions. Finally, most existing platforms provide the same information to all users without considering individual health sensitivity, planned activity, or personal risk factors.

Recent advances in low-cost Internet of Things (IoT) sensors have significantly improved air quality monitoring at fine spatial resolution. Dense PurpleAir sensor networks provide neighborhood-scale air quality measurements that complement traditional regulatory monitoring stations and enable much denser environmental coverage. This difference in monitoring density is illustrated in Figure 1.

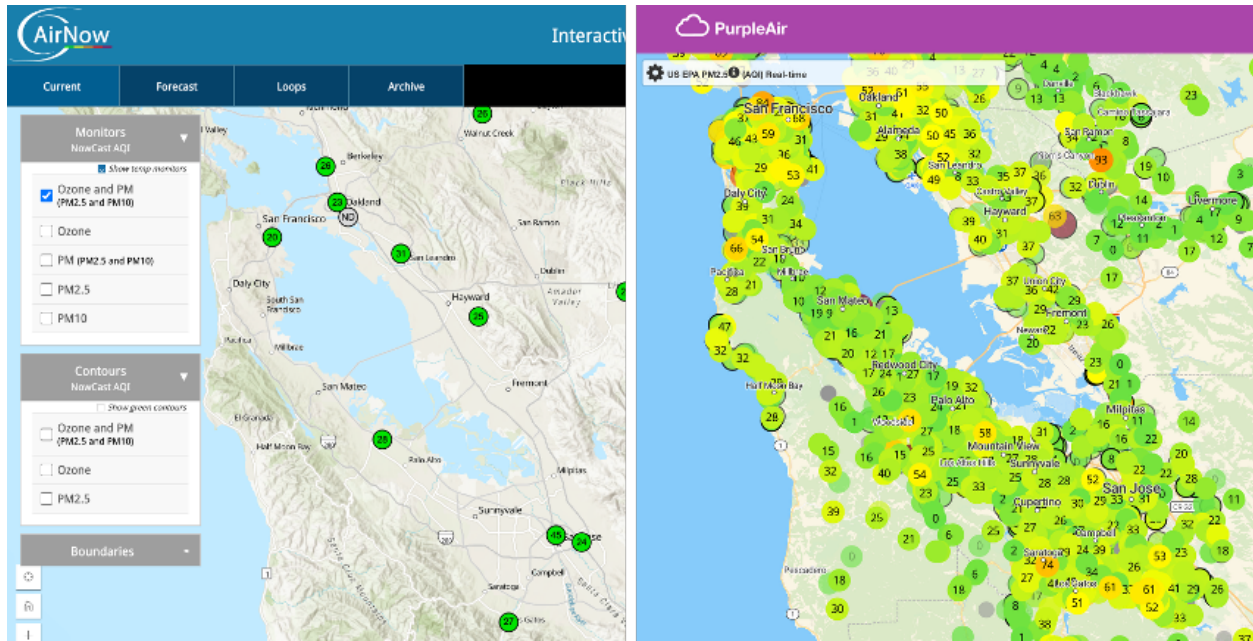


Figure 1 Comparison of regulatory air-quality monitoring stations and PurpleAir sensors across the same geographic area of the San Francisco Bay Area.

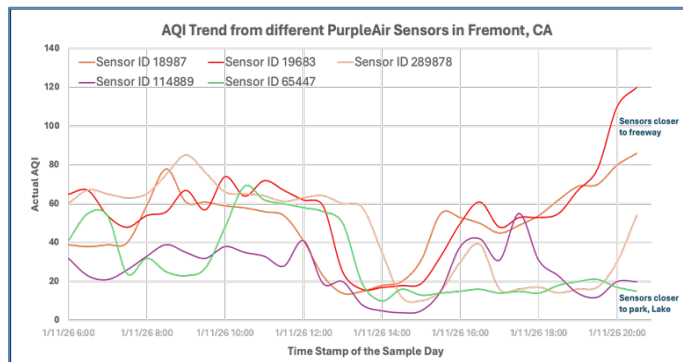


Figure 2 PM2.5 measurements from five sensors across Fremont, California, for one day.

This increased sensor density is particularly valuable for monitoring PM2.5, one of the most significant air pollutants affecting human health [4]. Neighborhood-scale variability was examined using representative PurpleAir sensors across Fremont, California (Figure 2). Although all sensors were located within the same city, PM2.5 AQI varied substantially throughout the day, demonstrating that air quality can differ significantly between neighborhoods.

Recent work has demonstrated the growing use of machine learning and deep learning for air-quality prediction. Trinkl [5] compared XGBoost, LightGBM, and support vector machines using pollutant and meteorological variables and found that ensemble-based methods, particularly XGBoost, performed strongly for AQI prediction. Rosca et al. [6] evaluated a broad set of machine-learning and deep-learning approaches using pollutant, meteorological, and traffic data, and incorporated the resulting models into an IoT-enabled air-quality application. Jayaraman et al. [7] proposed a time-based spatial forecasting framework that combined spatial information from multiple cities with temporal pollutant patterns, demonstrating the value of integrating both spatial and temporal dependencies in AQI forecasting. Janarthanan et al. [8] applied deep-learning methods to metropolitan AQI prediction, further illustrating the use of neural-network approaches for urban air-quality forecasting. Together, these studies show that

pollutant history, meteorological context, and spatial information can all contribute useful predictive signal for short-term air-quality forecasting.

Beyond forecasting, prior work has explored how artificial intelligence can support more personalized or context-aware decision making. López-Barreiro et al. [9] reviewed AI-powered recommender systems for promoting healthy habits and physical activity, highlighting the growing use of personalization to adapt recommendations to individual characteristics. Martínez-España et al. [10] proposed an AI-driven decision-support framework for air-quality forecasting in which model selection adapts to factors such as pollutant type and forecast horizon. Sutton et al. [11], in the broader field of clinical decision support, described how computerized systems can combine individual-specific information with structured knowledge to generate targeted assessments and recommendations.

While these studies demonstrate advances in air-quality forecasting and intelligent decision support, these capabilities have largely been studied separately. Relatively little work has explored how hyperlocal air-quality information, short-term forecasting, multi-pollutant assessment, personalized health sensitivity, activity context, and intelligent reasoning can be integrated to provide actionable guidance for everyday outdoor activities.

To address this gap, this study presents AirGuide AI, an end-to-end intelligent decision-support system for personalized outdoor-activity planning. The key contributions of this work are:

1. An integrated outdoor-activity decision-support system that combines environmental risk assessment, hyperlocal PM2.5 forecasting, and agentic reasoning to provide real-time personalized guidance.
2. A multi-pollutant environmental risk assessment framework that considers current and forecasted air quality, EPA-aligned health guidance, user health sensitivity, activity type, and location.
3. A machine-learning-based hyperlocal PM2.5 forecasting model that predicts neighborhood-scale conditions one and three hours into the future to support proactive activity planning.
4. An agentic recommendation framework that transforms environmental assessments and user context into actionable guidance, including activity modification, timing changes, cleaner nearby locations, and lower-exposure route options.

METHODS AND MATERIALS

System Architecture

Figure 3 presents the overall system architecture and data flow. Users provide a location, health sensitivity profile, and a natural-language question related to an outdoor activity. Three EPA-aligned health sensitivity profiles are supported: Normal, Sensitive, and Unusually Sensitive Group. The system supports walking, running, and biking, with recommendations generated for current conditions or one-hour and three-hour forecast horizons. The natural-language input is interpreted to determine the activity and time horizon, while the provided location is used to retrieve environmental conditions and support location- and route-based analysis.

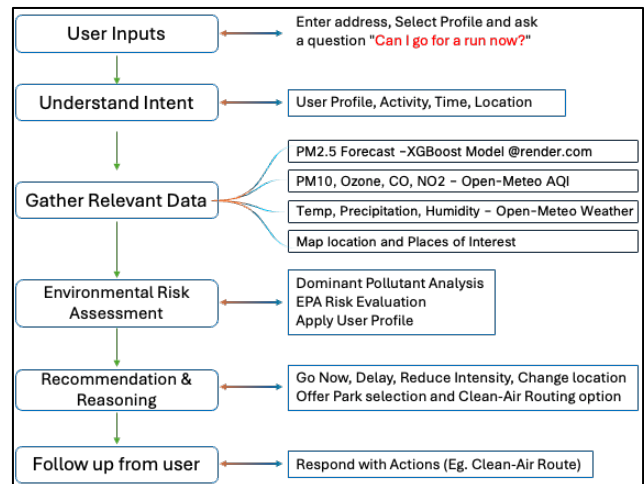


Figure 3 Overall Architecture and workflow of AirGuide AI

Environmental and geospatial data are retrieved from multiple external sources and combined with hyperlocal PM2.5 forecasts generated by the machine-learning forecasting service. The Agentic Reasoning Engine evaluates current and forecasted pollutant conditions, identifies the dominant pollutant, and combines this assessment with the user’s health sensitivity and activity context. The resulting assessment is used to select an intervention strategy and generate the final recommendation.

Data Sources

AirGuide AI integrates multiple environmental, meteorological, regulatory, and geospatial data sources to support air quality assessment, forecasting, and recommendation generation. Table 1 summarizes the primary data sources used by the system and their respective roles.

Table 1 Data Sources used by AirGuide AI and their purposes

Data Source	What is gathered?	Purpose
PurpleAir [12]	Historical and real-time PM2.5 measurements	PM2.5 assessment and ML-based PM2.5 forecasting model
Open-Meteo Air Quality API [13]	PM10, O ₃ , NO ₂ , SO ₂ , CO concentrations	Multi-pollutant risk assessment
Open-Meteo Weather API [13]	Temperature, humidity, wind, weather variables	Meteorological feature evaluation and forecasting
EPA AQI Guidance	AQI breakpoints and health recommendations	Risk assessment and recommendations
Mapbox	Geocoding and routing services	Location-aware recommendations and clean-air routing

Hyperlocal PM2.5 Forecasting Model

Short-term PM2.5 forecasting was developed using machine-learning models to support proactive outdoor-activity planning based on neighborhood-scale air-quality conditions. PM2.5 was selected because it exhibited substantial neighborhood-level variability and was frequently the dominant pollutant. The forecasting methodology included dataset preparation and quality control, analysis of temporal and spatial behavior, feature engineering, model development, feature-ablation analysis, ensemble forecasting, and forward holdout evaluation.

Dataset Collection and Quality Control

Historical PM2.5 observations were collected from 25 outdoor PurpleAir sensors distributed across Fremont, California, at 30-minute intervals from January 2024 through June 2026. For each observation, the reported PM2.5 average and measurements from the sensor’s independent Channels A and B were collected, resulting in 955,535 raw observations.

The dataset was screened for sensor failures, disagreement between Channels A and B, insufficient data coverage, and isolated measurements inconsistent with surrounding observations and nearby sensors. Sensors with insufficient coverage or persistent data-quality concerns were excluded. The final quality-controlled dataset contained 760,702 observations from 20 sensors. A chronological experimental design was used to prevent temporal leakage:

- Training (January 2024–December 2025): 592,832 observations
- Validation (January–March 2026): 83,864 observations
- Holdout (April–June 2026): 84,006 observations

The training period was used for model fitting, the validation period for model evaluation and ensemble calibration, and the forward holdout period only for final performance evaluation.

Hyperlocal Forecasting Framework

PM_{2.5} concentrations across the Fremont sensor network varied substantially over time. Monthly network median and 90th-percentile (P90) concentrations showed seasonal and episodic variation (Figure 4). Winter months generally had higher concentrations and greater separation between the median and P90, while lower and more stable concentrations were more common during spring and early summer. Short-duration elevated periods also occurred within otherwise lower-concentration months.

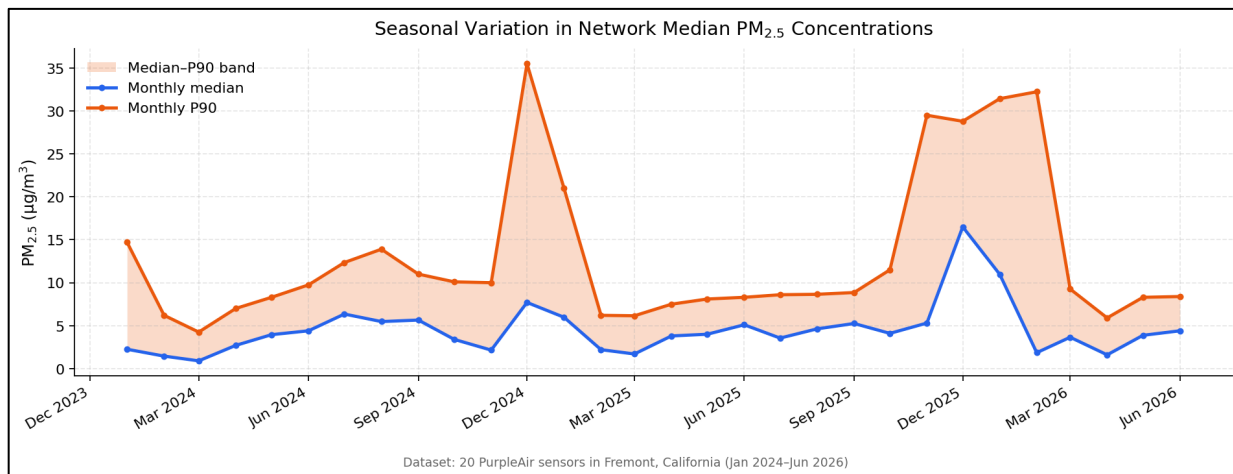


Figure 4 Monthly median and 90th-percentile PM_{2.5} concentrations across the Fremont sensor network during the study period.

Analysis across individual sensors also showed spatial structure. Pairwise correlations in PM_{2.5} levels were typically high ($\rho \approx 0.85$ – 0.92), indicating that major changes often occurred across the sensor network. However, correlations in short-term rates of change were considerably weaker ($\rho \approx 0.18$ – 0.42), indicating that locations generally shared the direction of pollution events but not necessarily the magnitude of short-term changes. These observations motivated a forecasting framework combining recent conditions at the target sensor with surrounding-sensor and meteorological information.

Feature Engineering

Thirteen features were engineered to represent local PM_{2.5} history, neighboring-sensor conditions, and meteorological context. Only information available at the time of prediction was used; no future pollutant or weather observations were included as forecasting inputs.

Local PM_{2.5} history consisted of eight features: the current concentration; concentrations at 30-minute, 1-hour, 1.5-hour, 2-hour, 3-hour, and 6-hour lags; and the change during the most recent 30-minute interval. The shorter lags represented recent persistence and movement, while the longer lags provided context about how current conditions developed over the preceding several hours.

Neighboring-sensor context consisted of two features derived from the five geographically nearest sensors: their median PM_{2.5} concentration and the 30-minute change in that median. The median was used to characterize surrounding pollution conditions while reducing sensitivity to unusually high or low measurements from an individual neighboring sensor.

Meteorological context consisted of wind speed and two directional wind components derived from Open-Meteo observations. Wind direction was represented as east-west and north-south components rather than as an angular value, avoiding the discontinuity between values such as 359° and 1°. These variables provided information about atmospheric transport and dispersion.

Additional candidate features representing longer pollutant histories, alternative neighboring-sensor summaries, temporal patterns, and additional meteorological variables were evaluated but did not provide consistent improvement and were not retained. The final 13-feature set is summarized in Table 2.

Table 2 - Final Feature Set for PM2.5 Forecasting

Feature group	Features	Count
Local PM2.5 history	Current PM2.5; 30-min, 1-h, 1.5-h, 2-h, 3-h, 6-h lags; 30-min change	8
Neighboring-sensor context	Median PM2.5 of 5 nearest sensors; 30-min change in nearby median	2
Meteorological context	Wind speed; east-west and north-south wind components	3
Total		13

Model Development and Training

Separate XGBoost regression models were developed for the one-hour and three-hour forecasting horizons. Both models used the same 13 input features but different prediction targets. The one-hour model predicted PM2.5 two 30-minute intervals into the future, while the three-hour model predicted six intervals ahead. PM2.5 concentration was predicted directly in $\mu\text{g}/\text{m}^3$ and converted to AQI only after forecasting.

XGBoost [14] was selected because gradient-boosted decision trees can capture nonlinear relationships and interactions among recent pollutant history, neighboring-sensor behavior, and meteorological conditions without requiring feature scaling. Observations from all 20 sensors were pooled into a common training dataset, allowing the models to learn patterns shared across the Fremont sensor network rather than training separate models for individual locations.

Both models used a squared-error regression objective with 200 trees, a maximum tree depth of 6, learning rate of 0.1, row subsampling of 0.8, feature subsampling of 0.8, minimum child weight of 3, gamma of 0.1, L1 regularization of 0, and L2 regularization of 1. Histogram-based tree construction and a fixed random seed of 42 were used. The training configuration was identical for both forecast horizons; only the prediction target differed. All 200 trees were trained without early stopping. The same model architecture and training configuration were maintained during feature-ablation experiments, allowing differences among candidate feature sets to be attributed to the information provided by the features rather than changes in the learning algorithm or training procedure.

Feature Selection and Ablation Analysis

Feature-ablation experiments were conducted to evaluate the contribution of each information source and determine the final forecasting feature set. The analysis began with a model using only the target sensor's recent PM2.5 history. Neighboring-sensor information was then added to determine whether measurements from surrounding locations provided useful forecasting signal beyond the target sensor's own history. Several neighboring-sensor representations were evaluated, with the median PM2.5 concentration of the five geographically nearest sensors and its 30-minute change retained in the final model.

Wind speed and directional components were subsequently added to the local-history and neighboring-sensor feature set. These features provided a smaller additional improvement and were retained, resulting

in the final 13-feature model. Feature-group contributions were evaluated at both forecast horizons, with the results reported in the Results section.

Ensemble Forecasting and Evaluation

We also compared the XGBoost forecast with persistence forecasting, a simple method that assumes the current PM2.5 concentration will remain unchanged over the forecast period. Candidate ensemble strategies were evaluated using the validation dataset, including multi-tier and simplified AQI-based weighting rules. A two-tier rule was selected to provide a simple and robust blending strategy while maintaining competitive validation performance. When current AQI was 100 or below, the final forecast used 50% XGBoost and 50% persistence. When AQI was above 100, it used 40% XGBoost and 60% persistence. The same weighting rule was applied at both forecast horizons.

Forecast performance was evaluated using mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2) at the one-hour and three-hour forecast horizons.

Environmental Risk Assessment

AirGuide AI evaluates current and forecasted air quality conditions to assess environmental risk and support personalized outdoor activity recommendations. Assessment is performed for the current time, one-hour and three-hour forecast horizons. Because pollutant concentrations are reported using different units and scales, AirGuide AI converts each pollutant concentration into a common Instant Air Quality Index (IAQI) scale. For pollutants whose native units differ from EPA breakpoint units, appropriate unit conversions are applied prior to IAQI calculation as shown in Table 3.

Table 3 Pollutant Concentration Conversions

Pollutant	Source	Native Unit	EPA Breakpoint unit	Conversion
PM2.5	AQI Forecasting Service	$\mu\text{g}/\text{m}^3$	$\mu\text{g}/\text{m}^3$	None
PM10	Open-Meteo Air Quality API	$\mu\text{g}/\text{m}^3$	$\mu\text{g}/\text{m}^3$	None
O ₃		$\mu\text{g}/\text{m}^3$	ppb	$\mu\text{g}/\text{m}^3 \div 1.96$
NO ₂		$\mu\text{g}/\text{m}^3$	ppb	$\mu\text{g}/\text{m}^3 \div 1.88$
CO		$\mu\text{g}/\text{m}^3$	ppm	$\mu\text{g}/\text{m}^3 \div 1145$

IAQI values for each pollutant are computed using the standard EPA breakpoint formula.

$$\text{IAQI} = \text{round} \left(\left((I_{\text{high}} - I_{\text{low}}) / (C_{\text{high}} - C_{\text{low}}) * (C - C_{\text{low}}) + I_{\text{low}} \right) \right)$$

where C is the pollutant concentration, $[C_{\text{low}}, C_{\text{high}}]$ is the EPA concentration breakpoint interval containing C , and $[I_{\text{low}}, I_{\text{high}}]$ is the corresponding AQI breakpoint interval.

Unlike official EPA AQI values, which are based on pollutant-specific averaging periods, IAQI is calculated directly from current or forecasted pollutant concentrations. This approach provides a more responsive representation of environmental conditions that are relevant to outdoor activity planning. The pollutant with the highest IAQI value is identified as the dominant pollutant, and the corresponding EPA health risk category is determined. Because pollutant conditions vary over time, the dominant pollutant may change across the current, one-hour, and three-hour assessment horizons. AirGuide AI then applies an EPA-aligned assessment framework that combines the environmental risk category with the user's health sensitivity profile and planned activity. The framework, summarized in Table 4, provides decision guardrails for interpreting environmental risk before personalized recommendations are generated.

Table 4 EPA aligned recommendation framework used by AI Agent - By AQI category and Profile

AQI Category	Normal Group	Sensitive Group	Unusually Sensitive
Good (0–50)	Go ahead with outdoor activity	Go ahead with outdoor activity	Go ahead with outdoor activity
Moderate (51–100)	Go ahead	Go ahead, but watch for symptoms	Reduce intensity or shorten duration; monitor symptoms
Unhealthy for Sensitive Groups (101–150)	Use caution for prolonged or intense activity	Make activities shorter and less intense; take more breaks	Avoid prolonged outdoor activity; low-intensity activity better
Unhealthy (151–200)	Reduce long or intense activity	Avoid long or intense outdoor activity; consider indoors	Avoid outdoor activity when possible
Very Unhealthy (201–300)	Avoid long or intense outdoor activity	Avoid outdoor activity	Avoid all outdoor physical activity
Hazardous (301–500)	Avoid outdoor activity	Avoid outdoor activity; remain indoors if possible	Remain indoors and keep activity levels low

The resulting environmental assessment consists of the IAQI values, dominant pollutant, EPA health-risk category, user profile, activity type, and forecast horizon. These parameters are provided to the Agentic Reasoning Engine for intervention selection and personalized recommendation generation.

Agentic Recommendation Generation

The personalized recommendation engine transforms assessment outputs into actionable guidance for outdoor activity planning. Based on the environmental assessment, user characteristics, and activity goals, the engine selects an appropriate intervention strategy and generates personalized recommendations. This may include proceeding with the planned activity, reducing activity intensity or duration, changing activity timing, selecting a cleaner nearby location, or avoiding outdoor activity altogether. The recommendation process is designed to balance health risk reduction with the user’s desired outdoor activity goals.

AirGuide AI incorporates pollutant-specific intervention strategies that account for the characteristics of the dominant pollutant. For example, elevated PM_{2.5} levels may trigger recommendations to relocate activities to cleaner nearby areas or utilize lower-exposure routes. Elevated ozone conditions may result in recommendations to reduce exercise intensity or shift activities to periods with lower ozone concentrations. Traffic-related pollutants such as nitrogen dioxide and carbon monoxide may trigger avoidance strategies focused on reducing exposure near major roadways and congested areas.

Unlike conventional monitoring systems, AirGuide AI’s recommendation engine evaluates both present and forecasted conditions and can provide guidance based on anticipated changes. For example, a user may be advised to delay an outdoor activity when current conditions are unfavorable but forecast to improve, or to perform the activity sooner when favorable conditions are expected to deteriorate over the next several hours. This forecast-aware approach enables AirGuide AI to provide more proactive and personalized guidance for outdoor activity planning.

Clean-Air Route Planning

AirGuide AI includes a route planning capability designed to reduce pollutant exposure during outdoor activities. Route optimization is triggered when PM_{2.5} is identified as the dominant pollutant because PM_{2.5} can exhibit substantial spatial variability. When route planning is requested, the system evaluates multiple candidate routes between the origin and destination using PM_{2.5} conditions along each route. The system presents both the fastest route and the route with the lowest estimated pollutant exposure to the user for selection. By integrating route-aware environmental intelligence into the recommendation

framework, AirGuide AI extends beyond traditional air quality reporting and provides users with practical options for reducing pollutant exposure during outdoor activities.

System Implementation

AirGuide AI was implemented and deployed as a web application using the Lovable development platform. [15] The application uses Google Gemini 2.5 Flash-Lite to support natural-language interaction and agentic recommendation generation. The XGBoost hyperlocal PM2.5 forecasting models were developed using Cursor and deployed on Render.com as REST API services that provide real-time predictions to the application. The forecasting codebase is publicly available through the project repository [16].

The application interface integrates user location and health-profile inputs, natural-language activity queries, and a map-based interface for nearby-location and route visualization. (Figure 5)

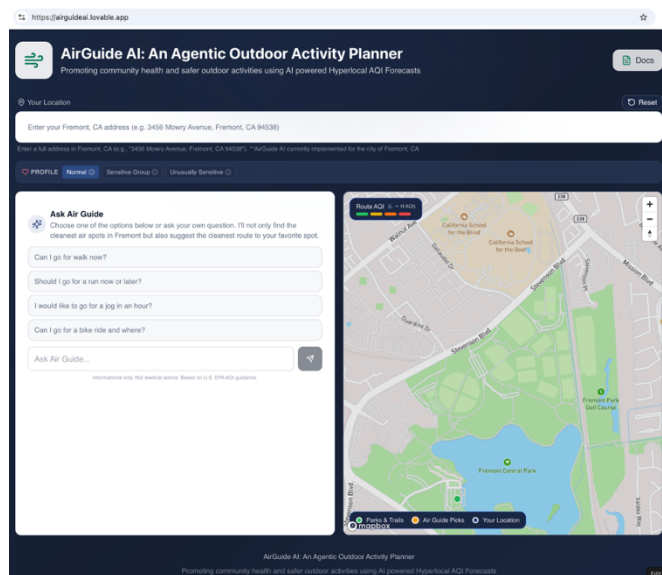


Figure 5 AirGuide AI application interface showing user inputs, conversational activity planning, and map-based location and route visualization.

RESULTS

Hyperlocal PM2.5 Forecasting Performance

The final forecasting system was evaluated across the chronological training, validation, and forward holdout periods at both forecast horizons (Table 5). On the April–June 2026 holdout dataset, the one-hour forecast achieved an MAE of 0.760 $\mu\text{g}/\text{m}^3$, RMSE of 1.626 $\mu\text{g}/\text{m}^3$, and R^2 of 0.804. At the three-hour horizon, MAE increased to 1.278 $\mu\text{g}/\text{m}^3$ and RMSE to 2.201 $\mu\text{g}/\text{m}^3$, while R^2 decreased to 0.641, indicating the expected reduction in predictive performance at the longer forecast horizon.

Table 5 Performance of the final PM2.5 forecasting model across the training, validation, and holdout datasets for 1-hour and 3-hour forecast models. (All data in PM2.5 ($\mu\text{g}/\text{m}^3$))

One-hour forecasting model performance							
Dataset	Period	n	Mean	SD	MAE	RMSE	R^2
Training	Jan 2024 – Dec 2025	592,832	5.74	7.27	1.219	2.532	0.879
Validation	Jan 2026 – Mar 2026	83,864	8.20	10.75	1.628	3.358	0.902
Holdout	Apr 2026 – Jun 2026	84,006	3.79	3.68	0.760	1.626	0.804

Three-hour forecasting model performance							
Dataset	Period	n	Mean	SD	MAE	RMSE	R^2
Training	Jan 2024 – Dec 2025	592,832	5.74	7.27	2.021	3.657	0.747
Validation	Jan 2026 – Mar 2026	83,864	8.20	10.75	2.763	5.070	0.777
Holdout	Apr 2026 – Jun 2026	84,006	3.79	3.68	1.278	2.201	0.641

Performance varied across the chronological periods in part because the underlying PM2.5 distributions differed substantially. The January–March validation period had the highest variability, with a mean

concentration of $8.20 \mu\text{g}/\text{m}^3$ and standard deviation of $10.75 \mu\text{g}/\text{m}^3$, compared with $3.79 \mu\text{g}/\text{m}^3$ and $3.68 \mu\text{g}/\text{m}^3$ in the April–June holdout. In this context, validation produced higher absolute errors but also higher R^2 values. For example, the one-hour validation MAE was $1.628 \mu\text{g}/\text{m}^3$ compared with $0.760 \mu\text{g}/\text{m}^3$ on the holdout, while R^2 was 0.902 compared with 0.804 . This reflects the dependence of absolute-error and variance-based metrics on the concentration range of the evaluation period.

Observed versus Predicted PM_{2.5}

The observed-versus-predicted comparison provides a visual assessment of forecast accuracy across the holdout dataset (Figure 6).

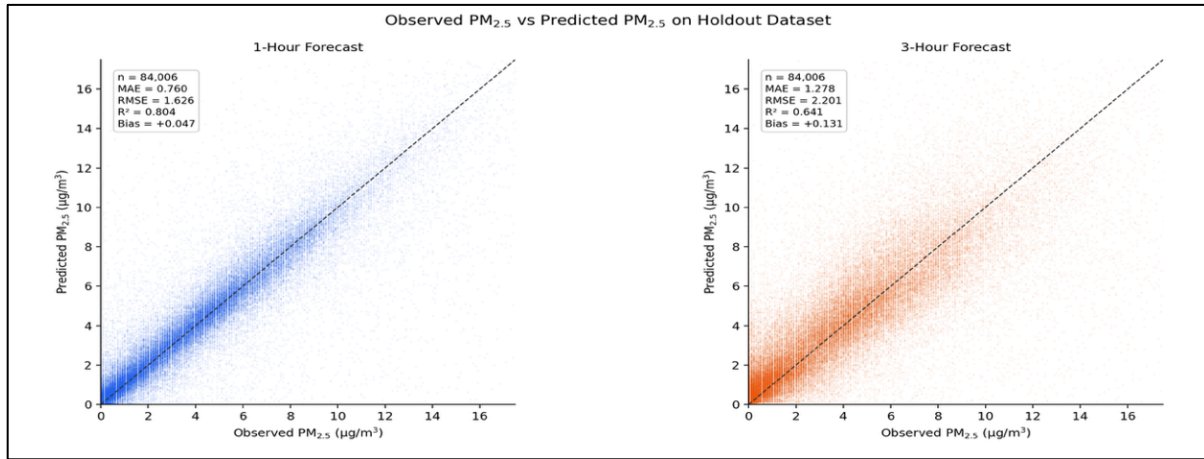


Figure 6 Observed vs predicted PM_{2.5} forecasts for the one-hour and three-hour horizon on the holdout dataset.

One-hour predictions were more tightly concentrated around the line of agreement, consistent with the holdout R^2 of 0.804 . Greater dispersion was observed for the three-hour predictions, consistent with the lower R^2 and higher MAE and RMSE at the longer forecast horizon.

Feature Ablation and Ensemble Performance

Feature-ablation analysis showed that neighboring-sensor information provided the largest improvement beyond local PM_{2.5} history, while wind information provided a smaller additional benefit (Table 6). The local-history-only model did not outperform the persistence baseline at either forecast horizon. Adding neighboring-sensor context reduced MAE below persistence at both horizons, demonstrating that measurements from surrounding locations provided useful predictive information beyond the recent history of the target sensor. Wind information produced a further, smaller reduction in error.

Table 6 Feature-ablation and ensemble performance on the holdout dataset.

Model / Feature Set	1-hour MAE ($\mu\text{g}/\text{m}^3$)	3-hour MAE ($\mu\text{g}/\text{m}^3$)
Persistence baseline	0.796	1.362
Local PM _{2.5} history	0.806	1.390
+ Neighboring-sensor context	0.777	1.318
+ Wind information	0.773	1.302
Final ensemble	0.760	1.278

Combining the final 13-feature XGBoost model with persistence forecasting produced the lowest holdout MAE at both horizons. The ensemble reduced MAE from 0.773 to $0.760 \mu\text{g}/\text{m}^3$ at one hour and from 1.302 to $1.278 \mu\text{g}/\text{m}^3$ at three hours. During model development, alternative ensemble weights produced

only marginal differences in validation performance, with the best tested alternatives improving MAE by approximately $0.004 \mu\text{g}/\text{m}^3$ at both forecast horizons. The simpler two-tier weighting rule described in the earlier section was therefore retained for the final forecasting system.

Model Interpretation Using SHAP

SHAP analysis was used to examine how individual features influenced the one-hour and three-hour PM_{2.5} forecasts. Across both horizons, current PM_{2.5} concentration was the most influential individual feature, followed by recent local PM_{2.5} history and neighboring-sensor conditions (Figure 7). Higher current and recent PM_{2.5} concentrations generally increased predicted future concentrations, while lower values generally reduced the forecasts.

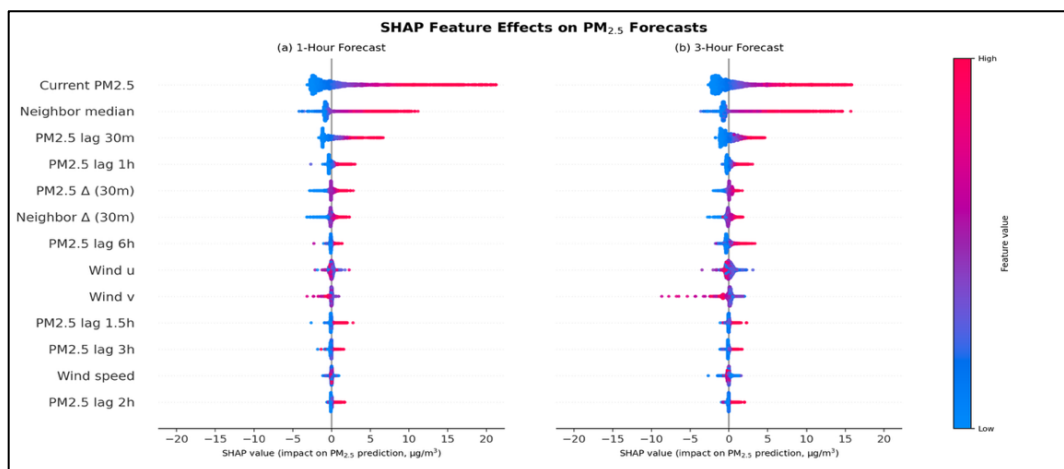


Figure 7 SHAP feature effects for one-hour and three-hour PM_{2.5} forecasting.

Grouped SHAP importance showed that local PM_{2.5} history accounted for the largest share of model influence at both forecast horizons, decreasing from 73.1% at one hour to 63.3% at three hours (Table 7). Neighboring-sensor contribution remained substantial, while the contribution of wind increased from 4.9% to 12.1% at the longer horizon.

Table 7 Grouped SHAP feature importance for the one-hour and three-hour PM_{2.5} forecasting models.

Feature Group	1-hour Forecast	3-hour Forecast
Local PM _{2.5} history	73.1%	63.3%
Neighboring-sensor context	22.0%	24.5%
Meteorological (wind) features	4.9%	12.1%
Total	100.0%	99.9%*

Temporal Forecast Behavior

Figure 8 shows observed PM_{2.5} concentrations and the corresponding one-hour and three-hour forecasts over a representative 30-day period from the April–June 2026 holdout dataset. Forecasts are aligned to their target timestamps; therefore, the one-hour and three-hour predictions shown at each timestamp were generated one and three hours earlier, respectively. Both forecast horizons captured the overall temporal pattern of PM_{2.5} concentrations. The one-hour forecast generally tracked short-term increases and decreases more closely, while the three-hour forecast showed greater deviation during sharper changes in concentration. During relatively stable periods, both forecasts remained close to the observed measurements.

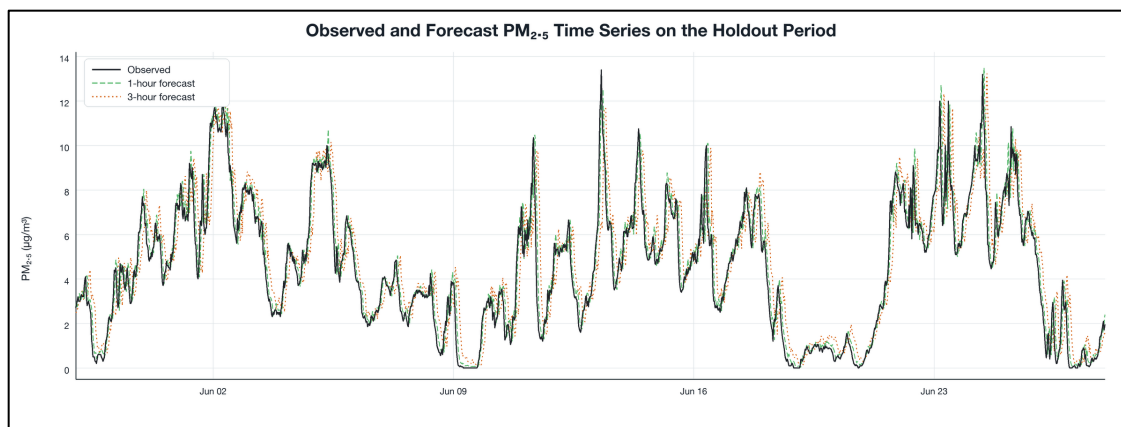


Figure 8 Observed and forecast PM2.5 time series during a representative 30-day period from the holdout dataset.

Scenario-Based Evaluation of Personalized Recommendations

AirGuide AI was evaluated using four representative application scenarios designed to exercise distinct components of the recommendation framework. The scenarios examined pollutant-specific and forecast-aware guidance, hyperlocal location selection, adaptation to user health sensitivity, and exposure-aware route planning. Table 8 summarizes the environmental and user context, capability evaluated, and resulting system response.

Table 8 Representative scenarios used to evaluate AirGuide AI recommendation behavior.

#	Environmental and User Context	Capability Evaluated	AirGuide AI Response
1	Ozone dominant; IAQI 55; Sensitive profile; forecast improves to IAQI 45 in 3 h	Pollutant-specific and forecast-aware guidance	Recommended delaying the run until conditions improved
2	PM2.5 dominant; current location IAQI 75; nearby parks IAQI 18–26	Hyperlocal location selection	Recommended relocating the walk to a cleaner nearby park
3	PM2.5 dominant; IAQI 112; Unusually Sensitive profile; forecast improves to IAQI 80 in 3 h	Health-sensitive and forecast-aware guidance	Recommended avoiding prolonged outdoor activity because conditions remained unsuitable for the user's profile
4	PM2.5 dominant; cleaner nearby destination identified and selected	Exposure-aware route planning	Presented fastest and lowest-exposure routes to the selected destination

Scenario 1 – Pollutant-Specific and Forecast-Aware Recommendation. For a Sensitive user planning a run, ozone was the dominant pollutant at IAQI 55. The three-hour forecast indicated improvement to Good conditions (IAQI 45). AirGuide AI recommended delaying the run until conditions improved rather than changing location, consistent with the treatment of ozone as a regional pollutant.

Scenario 2 – Hyperlocal Location Recommendation. At the user's current location, PM2.5 was the dominant pollutant with an IAQI of 75. AirGuide AI identified nearby parks approximately 1–2 miles away with PM2.5 IAQI values of 18–26 and recommended relocating the walk to one of these cleaner locations.

Scenario 3 – Forecast-Aware Recommendation Based on Health Sensitivity. For an Unusually Sensitive user, PM2.5 conditions were initially Unhealthy for Sensitive Groups (IAQI 112) and were forecast to improve to Moderate (IAQI 80) within three hours. Because the forecast conditions remained unsuitable for the user's health-sensitivity profile, AirGuide AI recommended avoiding prolonged outdoor activity and suggested indoor activity instead.

Scenario 4 – Exposure-Aware Route Planning. After identifying a nearby park with lower PM_{2.5} and receiving the user’s destination selection, AirGuide AI generated both the fastest route and the route with the lowest estimated PM_{2.5} exposure. The application displayed the corresponding air-quality information for both alternatives, allowing the user to select a route.

Complete application interactions for the four scenarios are provided in Supplement 1.

DISCUSSION

The results demonstrate the value of combining environmental information across pollutants, locations, and time rather than relying on a single value. AirGuide AI integrates multi-pollutant assessment, hyperlocal PM_{2.5} conditions and forecasts, user health sensitivity, activity context, and agentic reasoning within a single decision-support framework. Importantly, the appropriate intervention depends not only on the severity of air-quality conditions, but also on the pollutant driving the risk, its spatial behavior, expected changes over time, and the characteristics of the individual performing the activity.

The scenario-based evaluation illustrates these differences. When ozone was the dominant pollutant, changing location offered limited benefit because ozone was treated as a regional pollutant; forecast improvement instead supported delaying the activity. In contrast, neighborhood-scale variation in PM_{2.5} allowed the system to identify cleaner nearby locations rather than simply recommending that outdoor activity be avoided. Health sensitivity further modified the recommendation: improving forecast conditions did not necessarily make an activity appropriate when pollutant levels remained above the threshold for an unusually sensitive user. Exposure-aware routing extended this approach by incorporating environmental conditions into the journey to a selected destination, allowing users to compare travel efficiency with estimated PM_{2.5} exposure. These examples show why similar air-quality conditions need not result in the same recommendation when environmental and individual context differ.

The PM_{2.5} forecasting experiments also demonstrate the importance of combining temporal and spatial information. Local PM_{2.5} history provided the strongest forecasting signal but did not outperform persistence when used alone. Adding neighboring-sensor context improved performance at both forecast horizons, while wind information provided a smaller additional benefit. SHAP analysis supported the same general pattern, with local history contributing the greatest model influence, followed by neighboring-sensor conditions and meteorological information. Together, the ablation and SHAP analyses indicate that recent conditions at the target location provide the primary predictive signal, while surrounding sensors and meteorological conditions contribute complementary information.

Forecast performance decreased as the prediction horizon increased, with the one-hour model tracking short-term changes more closely than the three-hour model. The persistence component also provided a modest improvement when combined with XGBoost at both horizons. This suggests that persistence remains useful for short-term PM_{2.5} prediction even when machine-learning models incorporate local history, surrounding-sensor information, and meteorological features.

Taken together, these capabilities shift the role of an air-quality application from environmental reporting toward environmental decision support. Hyperlocal measurements help determine where conditions differ; forecasting provides information about when they may change; multi-pollutant assessment identifies what is driving environmental risk; user and activity context establish who is affected and what they intend to do; and agentic reasoning translates these inputs into an appropriate course of action. The contribution of AirGuide AI is therefore not any individual model or data source, but their integration into an interactive system that converts complex environmental information into practical outdoor-activity decisions.

LIMITATIONS AND FUTURE WORK

This study has several limitations. First, the hyperlocal PM_{2.5} forecasting system was developed and evaluated using a sensor network in Fremont, California. Although the chronological holdout dataset provided evaluation on future unseen observations, the results do not establish generalizability to locations with different geography, meteorology, pollution sources, or sensor-network density. In addition, the April–June 2026 holdout period contained predominantly lower and relatively stable PM_{2.5} concentrations, with relatively few rapidly changing or highly elevated pollution events. Model performance under these more challenging conditions therefore requires further evaluation.

Second, machine-learning-based forecasting was developed only for PM_{2.5}. Other EPA criteria pollutants are incorporated into the environmental risk assessment, but pollutant-specific forecasting models were not developed for ozone, NO₂, SO₂, CO, or PM₁₀. Extending short-term forecasting to additional pollutants would enable more comprehensive forecast-aware recommendations.

Finally, the scenario-based evaluation demonstrated the behavior of the integrated recommendation framework under selected environmental and user conditions, but it did not evaluate recommendation effectiveness through real-world user studies. Future work could include evaluation across broader geographic and environmental conditions and assessment of recommendations with users. Additional extensions include forecasting for other pollutants, integration with wearable health devices, personalized exposure modeling, and route-planning methods that account for multiple pollutants simultaneously.

CONCLUSION

This study presents AirGuide AI, an integrated decision-support system that combines multi-pollutant air-quality assessment, hyperlocal PM_{2.5} forecasting, health and activity context, and agentic reasoning to support personalized outdoor-activity planning. The forecasting results showed that local PM_{2.5} history provides the strongest predictive signal, while neighboring-sensor and meteorological information contribute additional value for short-term forecasting. Scenario-based evaluation further demonstrated how pollutant type, forecast conditions, spatial variation, and user health sensitivity can lead to different recommendations.

By integrating these capabilities within an interactive application, AirGuide AI moves beyond reporting environmental conditions toward translating them into practical decisions about when, where, and how to perform outdoor activities. The results provide a foundation for broader geographic evaluation, additional pollutant forecasting, and future real-world assessment of personalized environmental decision support.

ACKNOWLEDGEMENTS

This research was conducted with the guidance and mentorship of Ivan Felipe Rodriguez from Inspirit AI and the author is extremely grateful for his support. The author also acknowledges PurpleAir, Open-Meteo, Mapbox, and EPA for providing publicly accessible data that made this work possible.

AI ACKNOWLEDGEMENTS

During the preparation of this work, the author used ChatGPT and Google AI Mode to assist with brainstorming, literature review organization, language refinement, and manuscript editing. All technical design decisions, implementation, analysis, and conclusions were independently reviewed and validated by the author, who assumes full responsibility for the content of this publication.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

1. World Bank. *The Global Health Cost of PM_{2.5} Air Pollution: A Case for Action Beyond 2021*. Washington, DC: World Bank; 2022. ISBN: 978-1-4648-1816-5. doi:10.1596/978-1-4648-1816-5.
2. World Health Organization. *WHO Global Air Quality Guidelines: Particulate Matter (PM_{2.5} and PM₁₀), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. Geneva: World Health Organization; 2021. ISBN: 978-92-4-003422-8.
3. U.S. Environmental Protection Agency. *Technical Assistance Document for the Reporting of Daily Air Quality—The Air Quality Index (AQI)*. EPA-454/B-18-007. Available from: <https://document.airnow.gov/technical-assistance-document-for-the-reporting-of-daily-air-quality.pdf> (accessed on 2026-9-24).
4. U.S. Environmental Protection Agency. Particulate Matter (PM) Pollution. Available from: <https://www.epa.gov/pm-pollution> (accessed on 2026-9-24).
5. Tırınk S. Machine learning-based forecasting of air quality index under long-term environmental patterns: a comparative approach with XGBoost, LightGBM, and SVM. *PLoS One*. 2025;20(10):e0334252. doi:10.1371/journal.pone.0334252.
6. Rosca CM, Carbureanu M, Stancu A. Data-driven approaches for predicting and forecasting air quality in urban areas. *Appl Sci*. 2025;15(8):4390. doi:10.3390/app15084390.
7. Jayaraman S, Nathezhtha T, Abirami S, Sakthivel G. Enhancing urban air quality prediction using time-based-spatial forecasting framework. *Sci Rep*. 2025;15:4139. doi:10.1038/s41598-024-83248-z.
8. Janarthanan R, Partheeban P, Somasundaram K, Navin Elamparithi P. A deep learning approach for prediction of air quality index in a metropolitan city. *Sustain Cities Soc*. 2021;67:102720. doi:10.1016/j.scs.2021.102720.
9. López-Barreiro J, Garcia-Soidan JL, Alvarez-Sabucedo L, Santos-Gago JM. Artificial intelligence-powered recommender systems for promoting healthy habits and active aging: a systematic review. *Appl Sci*. 2024;14(22):10220. doi:10.3390/app142210220.
10. Martínez-España R, Balderas-Díaz S, Guerrero-Contreras G, Terroso-Sáenz F, Morales-García J. An AI-driven decision support system for air quality forecasting with adaptive model selection. *SSRN*; 2026. doi:10.2139/ssrn.6981431.
11. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med*. 2020;3:17. doi:10.1038/s41746-020-0221-y.
12. PurpleAir. PurpleAir sensor network and API documentation. Available from: <https://www.purpleair.com> (accessed on 2026-9-24).
13. Open-Meteo. Open-Meteo Weather API documentation. Available from: <https://open-meteo.com> (accessed on 2026-9-24).
14. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery; 2016. p. 785–794. doi:10.1145/2939672.2939785.
15. Web App url: <https://airguideai.lovable.app/> (accessed on 2026-9-24).
16. Sivakumar V. AirGuide AI PM_{2.5} Forecasting API. GitHub repository. Available from: <https://github.com/vishalsivakumar1010/aqiprediction-api> (accessed on 2026-9-24).

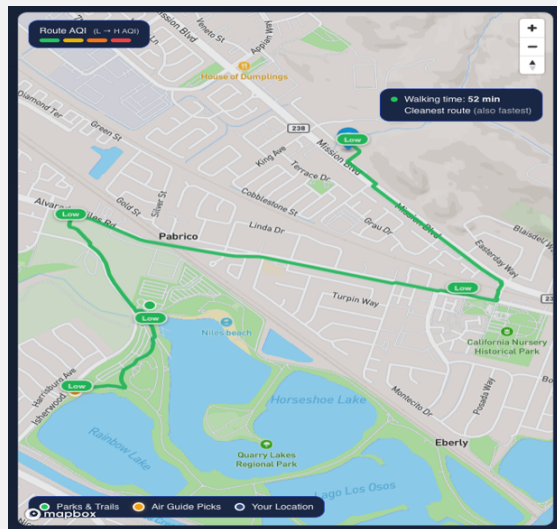
SUPPLEMENT 1 – SCENARIO EXAMPLES

Scenario 1:

Forecast-Aware Pollutant-Specific Recommendation	
Address	4500 Stevenson Boulevard, Fremont, California 94538, United States
Profile	Sensitive Group
Question	Should I go for a run now or later?
AirGuide AI response	
<p>🕒 Make activities shorter and less intense; take more breaks. Air quality — Moderate (IAQI 55, ozone) 🌫️ Forecast: +1h Moderate (IAQI 57, ozone) · +3h Good (IAQI 45, ozone) 📈 Trend: IAQI improves from 55 now to 45 in 3 hours (Δ -10). 🧠 Ozone may cause a cough and throat irritation, so shorten and lessen the intensity of your running. ⌚ Best time: Wait until in 3 hours — improves to Good (IAQI 45).</p> <p>🌿 Nearby clean-air parks:</p> <ul style="list-style-type: none"> • Blacow Park — Good (PM2.5 IAQI 50) (0.7 mi) • Irvington Park — Moderate (PM2.5 IAQI 52) (0.9 mi) • Central Park — Moderate (PM2.5 IAQI 55) (0.9 mi) <p>Note: Ozone is a regional pollutant and is likely similar across these parks — PM2.5 is what varies by location.</p>	

Scenario 2:

Hyperlocal Location Recommendation Using Neighborhood PM2.5 Variability	
Address	35490 Mission Boulevard, Fremont, California 94536, United States
Profile	Sensitive Group
Question	Can I go for a walk now?
AirGuide AI response	
<p>Skip the walk near home — but Alameda Creek Regional Trail - Isherwood Staging Area is in the Good range, so you can walk there. Currently, the air quality is Moderate with an IAQI of 75, driven by PM2.5. Near home, walking would expose you to more PM2.5 than is comfortable for your group, but at Alameda Creek Regional Trail - Isherwood Staging Area the air is clean enough to walk. A short trip out can make this an easy yes.</p> <p>🌿 Nearby clean-air parks:</p> <ul style="list-style-type: none"> • Alameda Creek Regional Trail - Isherwood Staging Area — Good (PM2.5 IAQI 18) (1.1 mi) • Centerville Community Park — Good (PM2.5 IAQI 25) (2.3 mi) • Niles Community Park — Good (PM2.5 IAQI 26) (1.3 mi) <p>Would you like a clean-air route to Alameda Creek Regional Trail - Isherwood Staging Area? Reply yes or type route to Alameda Creek Regional Trail - Isherwood Staging Area.</p>	



Scenario 3:

Forecast-Aware Recommendation Based on User Health Sensitivity	
Address	39550 Liberty Street, Fremont, California 94538, United States
Profile	Unusually Sensitive Group
Question	Can I go for a jog now or later?
<p>AirGuide AI response</p> <p>🕒 Avoid prolonged outdoor activity; prefer low-intensity activity. Air quality — Unhealthy for Sensitive Groups (IAQI 112, PM2.5) 🌫️ Forecast: +1h Unhealthy for Sensitive Groups (IAQI 105, PM2.5) · +3h Moderate (IAQI 80, PM2.5) 📉 Trend: IAQI improves from 112 now to 80 in 3 hours (Δ -32). 🧠 PM2.5 levels are high, so avoid prolonged outdoor activity and prefer low-intensity exercise. 🕒 Best time: No window clears your threshold in the next 3 hours — consider indoor activity (cleanest is in 3 hours, IAQI 80).</p> <p>🌿 Nearby clean-air parks:</p> <ul style="list-style-type: none"> Central Park — Moderate (PM2.5 IAQI 100) (0.6 mi) Centerville Community Park — Unhealthy for Sensitive Groups (PM2.5 IAQI 105) (0.9 mi) Shinn Historical Park and Arboretum — Unhealthy for Sensitive Groups (PM2.5 IAQI 110) (1.1 mi) <p>Would you like a clean-air route to Central Park? Reply yes or type route to Central Park.</p>	

Scenario 4: Route Planner Output Example -

Hyperlocal Cleaner Route Recommendation Using Neighborhood PM2.5 Variability	
Address	43330 Mission Boulevard, Fremont, California 94539, United States
Profile	Sensitive Group
Question	Can I go for a walk now?
<p>AirGuide AI response</p> <p>Skip the walk near home — but Mission San José Community Park is in the Good range, so you can walk there. Currently, the air quality is Moderate with an IAQI of 57, driven by PM2.5. PM2.5 levels are high at home, so it's better to walk at the cleaner nearby location instead of skipping your walk. Sensitive groups may experience more PM2.5 exposure walking here. A cleaner nearby option offers much less exposure. A small detour pays off today.</p> <p>🌿 Nearby clean-air parks:</p> <ul style="list-style-type: none"> Mission San José Community Park — Good (PM2.5 IAQI 38) (1.4 mi) Sabercat Historical Park — Good (PM2.5 IAQI 42) (0.6 mi) Old Mission Community Park — Good (PM2.5 IAQI 46) (0.6 mi) <p>Would you like a clean-air route to Mission San José Community Park? Reply yes or type route to Mission San José Community Park.</p>	

