Analyzing Air Quality Data from Low-Cost Sensors in the Vista Hermosa Community

Hao Hu, Tara Goller, Melanie Saavedra

ENE 428 - Course Final Project Report

Word Count: 1939

Executive Summary

This project report analyzes multi-pollutant, low-cost air quality sensor data in the Vista Hermosa community of downtown Los Angeles. This metropolitan region has a complex air pollutant profile from industrial, vehicle, and wildfire emissions, making it crucial to accurately monitor air quality. Low-cost sensors have the ability to supplement regulatory sensor data and measure localized air pollutant trends. The primary objective of this study is to visualize trends from the data and identify potential air quality impacts through the low cost sensors deployed in the Vista Hermosa community.

All datasets were analyzed with Python where raw data were cleaned and calibrated before AQI conversion and figure visualization. The overall results illustrated that air quality for the Vista Hermosa community stayed within “Good” and “Moderate” AQI. Primary analysis focused on $PM_{2.5}$ and ozone readings given their potency. $PM_{2.5}$ was the most present contaminant and typically peaked in the morning and evening hours. Ozone had elevated readings during the warmer seasons during the months August and September. There was little spatial variation with pollutant patterns between the two sensors.

Our findings emphasize the importance of air quality awareness with vulnerable populations in the Vista Hermosa community. The temporal peaks with $PM_{2.5}$ and ozone suggest that outdoor activity should be limited during these periods. Other community-based strategies include investing in low-emission activities (i.e. public transit and green spaces) and staying updated with local AQI. Low-cost sensor networks have the ability to identify localized air quality trends and improve community awareness on pollution events.

Introduction

The basin geography and metropolitan infrastructure of Los Angeles (LA) has led to a historical record of poor air quality. Ground-level ozone and particulate matter (PMs) are the most prevalent air pollutants in LA as they continue to exceed National Ambient Air Quality Standards (NAAQS) (AQMD). These contaminants are two out of the six criteria air pollutants regulated by the Environmental Protection Agency’s (EPA). The criteria pollutants are primarily emitted on a large, collective scale from the transportation and industrial sectors which contribute 60% of the emissions in California according to 2023 data (CARB). These elevated pollutants put urban residents at risk for lethal respiratory diseases. One study estimated that ozone and $PM_{2.5}$ exposure are responsible for approximately 3.9-5.5% of deaths in Southern California (Stewart et al.). This data does not account for all health impacts as many communities suffer from respiratory irritation and non-lethal cancers.

Air quality monitoring technologies are essential in tracking contaminants and identifying these health risks, especially within vulnerable communities located at metropolitan interfaces. However, current large-scale regulatory monitors are sparse and limited in capturing an accurate air quality profile particularly in urban dense areas. Low-cost air quality sensors are an emerging technology that "have the potential to provide preliminary data to inform more targeted studies, supplement existing monitoring networks, and aid in the quicker detection of pollution hotspots" (Snyder et al.). These sensors are typically smaller and more affordable, equipping communities with air quality data and immediate health risks.

Literature Review

Low-cost sensor networks continue to expand through their ability to increase spatial coverage of air quality and supplement data from federal and regulatory monitoring stations. Past research has shown that low-cost sensors are "particularly suitable for providing real-time warnings of high pollution episodes" such as wildfires and contaminant leaks (Carotenuto et al.). The spatial density of these sensors have a critical role in monitoring air quality events as these episodes are often sudden and rapidly evolve. Given the variable urban pollutants and wildfire risk in LA, investigating the effectiveness of low-cost sensor implementation is essential for improving responses to public health threats.

There are several studies involving low-cost air sensors, specifically in Los Angeles county neighborhoods, to profile and identify contaminant trends. One downtown LA study developed a model that could accurately capture $PM_{2.5}$ concentration patterns using data from the PurpleAir sensor network (Lu et al.). It illustrated the effectiveness and reliability of low-costs sensors in profiling $PM_{2.5}$ trends in LA. A case study investigating possible emissions from local oil and gas operations deployed 15 low-cost sensing systems measuring gas phase pollutants such as volatile organic compounds (VOCs) and methane (Collier-Oxandale et al.). Their sensors were able to identify short-term trends and sudden contaminant peaks in local air quality, providing valuable data that regulatory monitoring stations failed to capture.

There is currently limited research on the performance of low cost sensors monitoring several contaminants simultaneously specifically in downtown LA. Addressing the lack of multi-pollutant data with spatial and temporal analysis is critical for quantifying downtown LA's complex air quality profile that can have disproportionate impacts. Our project aims to analyze $PM_{2.5}$, carbon monoxide (CO), nitrogen dioxide (NO_2), and ozone (O_3) from two small-scale and low-cost air quality monitoring sensors located in downtown LA.

Project Background and Objectives

The main objective of this project is to analyze data from the two sensors deployed in Vista Hermosa and identify the impact of observed air quality trends on the community. Vista Hermosa is located in downtown LA, south of Echo Park and bordered by 110 and 101 highway intersection. It is composed of mostly community spaces, schools, and residential buildings with roughly 33.5% of its population being Hispanic (Data USA). The community selected the AirSENCE (Standard AS400X) from the Air Quality Sensor Performance Evaluation Center (AQ-SPEC) library based on their needs and budget. AirSENCE is a small-scale sensor that delivers real-time data on PM_1, $PM_{2.5}$, PM_{10}, CO, NO_2, and O_3 (AirSENCE). There are currently two sensors deployed with one at Firmin Ct and the other at the LA Department of Water and Power (LADWP).

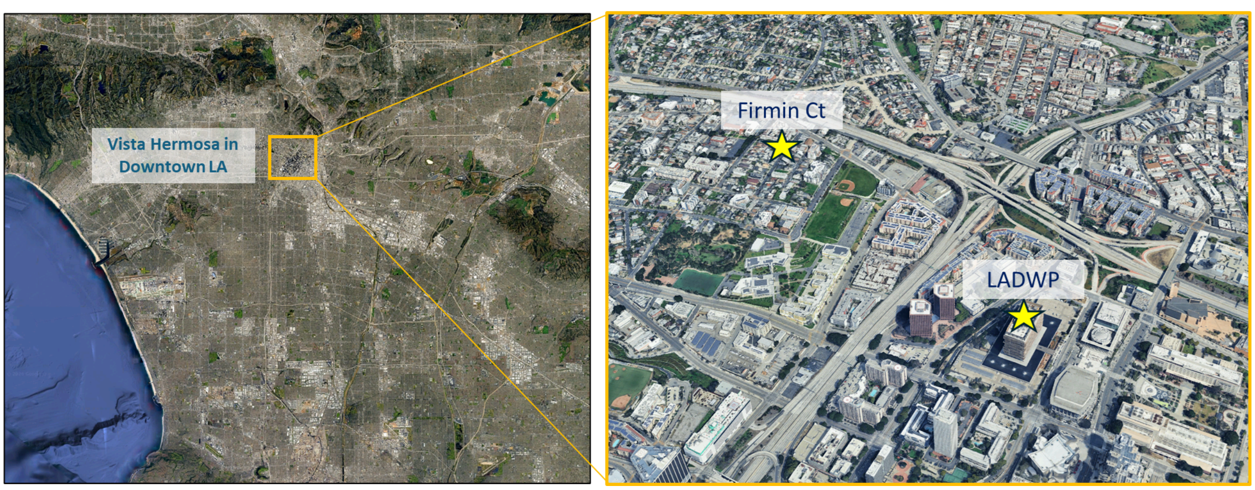

Figure 1. Satellite image of sensor locations in downtown LA

Our role is to take the raw pollutant concentrations measured by these sensors and transform it into accessible figures that highlight important trends and potential air quality concerns. We aim to form community-based strategies to reduce emission exposure through the empirical findings.

Methodology

Description of the observational data

The dataset includes near-minute-resolution air-quality measurements from the two monitoring units. Firmin Ct station covers the period from April 1, 2025 to October 23, 2025, while LADWP station covers April 1, 2025 to November 30, 2025. The original records are collected at approximately 1-minute resolution and include AQI-related pollutant variables, including $PM_{2.5}$, O_3, NO_2, and CO, along with meteorological variables such as temperature, relative humidity, and pressure.

Data preprocessing and calibration

The raw observation data were first screened and cleaned to ensure that all pollutant and meteorological variables were stored in a valid numeric format. Invalid records, including missing values and measurements outside the manufacturer-recommended bounds (Table 1), were removed before further processing.

Table 1. Max/Min threshold for sensors.

Pollutant	Unit	Invalid Min	Invalid Max
PM2.5	µg/m³	0	6000
Ozone / O3	ppm	0	9000
NO2	ppb	0	9000
CO	ppm	0	8000

We then applied unit-specific calibration equations (Table 2) to correct the raw sensor observations for each monitoring site and pollutant.

Table 2. Calibration equations applied to raw sensor observations at Firmin Ct and LADWP.

Pollutant	Station	Calibration Equation
NO2	Firmin Ct	y = (0.287924 * NO2 + 20.5815) + (0.00951 * CO) + (-0.07099 * Ozone) + (0.105212639 * PM1) + (-0.23567 * Humidity)
Ozone	Firmin Ct	y = 0.411448x + 6.815147
PM2.5	Firmin Ct	y = (1.398813 * PM2.5 + 18.79961) + (-.24294 * Humidity) + (-0.59602 * PM1) + (0.009518 * CO) + (-0.22486 * Temperature)
CO	Firmin Ct	Use raw CO data; no correction needed
NO2	LADWP	y = (0.218991 * NO2 + 9.53585) + (-0.05947 * Ozone) + (0.007387 * CO) + (0.440557 * PM1) + (-0.2453 * PM10)
Ozone	LADWP	y = 0.407887x + 8.492655
PM2.5	LADWP	y = (0.902479 * PM2.5 + 12.55986) + (-0.13155 * Humidity) + (0.08933 * NO2) + (-0.20993 * PM1) + (-0.13158 * Temperature)
CO	LADWP	y = (0.720171 * CO + .1364003) + (-0.68366 * Ozone) + (-1.88482 * NO2) + (9.35634 * PM1) + (-6.21643 * PM10)

After calibration, the 1-min records were aggregated to hourly values using a 75% data-completeness criterion, meaning that an hourly value was retained only when at least 45 valid 1-min observations were available within that hour. This preprocessing procedure helped ensure that the subsequent analysis was based on calibrated and sufficiently complete observations.

AQI conversion and classification

After calibration and hourly aggregation, pollutant concentrations were converted to Air Quality Index (AQI) values to provide a more interpretable metric for community-facing analysis. The conversion followed the U.S. EPA AQI breakpoint table which defines pollutant-specific concentration breakpoints and corresponding AQI ranges for each health category (EPA). $PM_{2.5}$ was converted using the 24-hour breakpoint, ozone and CO were converted using the 8-hour running-average breakpoints, and NO_2 was converted using the

1-hour breakpoint. For each pollutant, the AQI was calculated using linear interpolation within the corresponding breakpoint interval:

$$AQI = \frac{I_{\text{high}} - I_{\text{low}}}{C_{\text{high}} - C_{\text{low}}} \times (C - C_{\text{low}}) + I_{\text{low}}$$

where *C* is the observed concentration after applying the required averaging time, *C_low* and *C_high* are the lower and upper concentration breakpoints, and *I_low* and *I_high* are the corresponding AQI breakpoints. The resulting AQI values were then classified into six health categories: *Good, Moderate, Unhealthy for Sensitive Groups, Unhealthy, Very Unhealthy, and Hazardous.*

Coding platform: Python

All data preprocessing, calibration, AQI conversion, statistical analysis, and figure generation were conducted using Python.

Results

Air quality data from the Firmin Ct and LADWP monitoring sites were analyzed for $PM_{2.5}$, O_3, NO_2, and CO across multiple temporal scales, including time series trends, AQI category distributions, daily AQI maximum, and diurnal cycles.

I. Data Quality and Cleaning

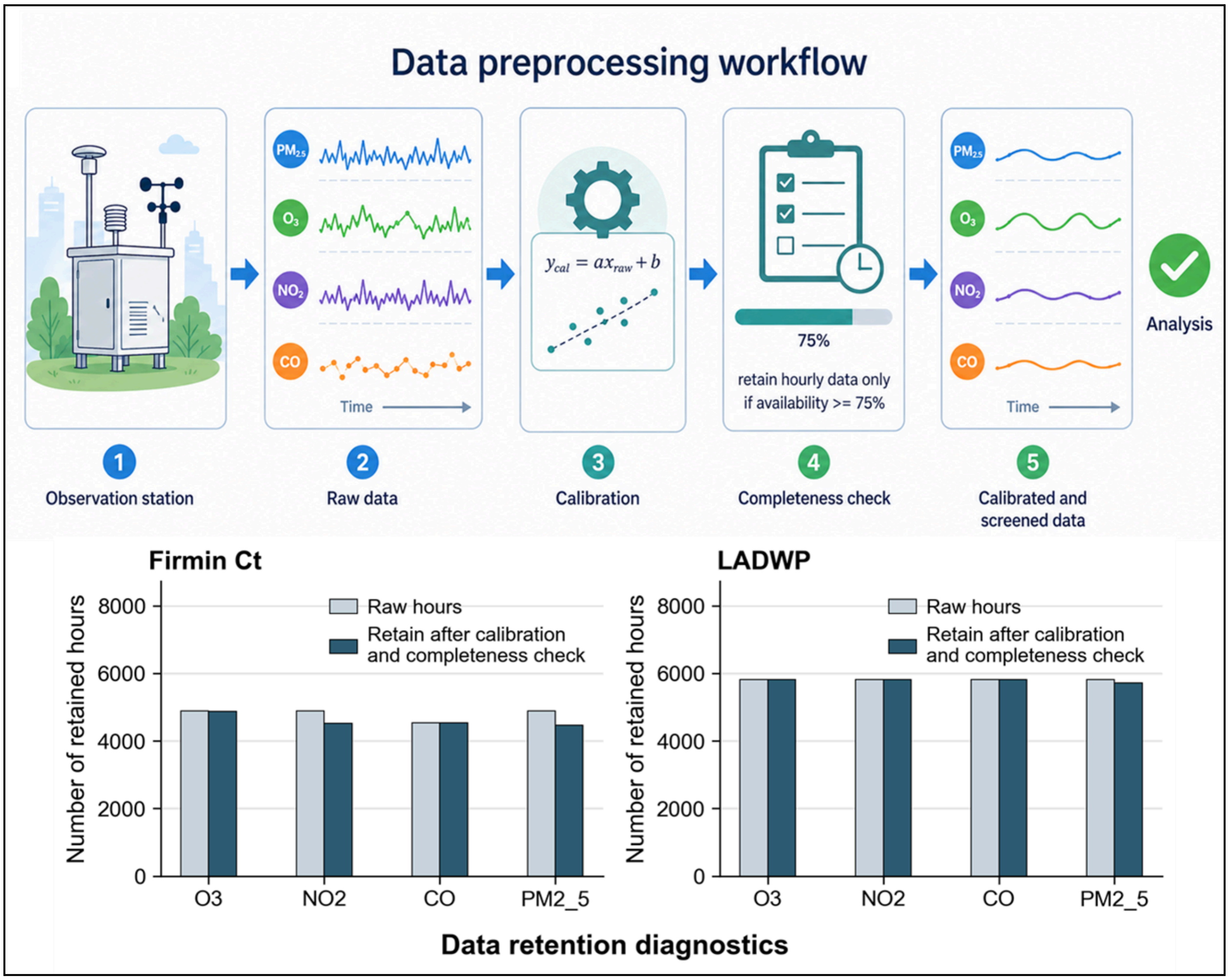

Figure 1.1. Data preprocessing workflow and retention diagnostics.

Note. Raw observations were first corrected using calibration equations and then screened using a 75% hourly completeness check. The bar charts show the number of hourly records retained after preprocessing at Firmin Ct and LADWP.

II. Time Series Trends

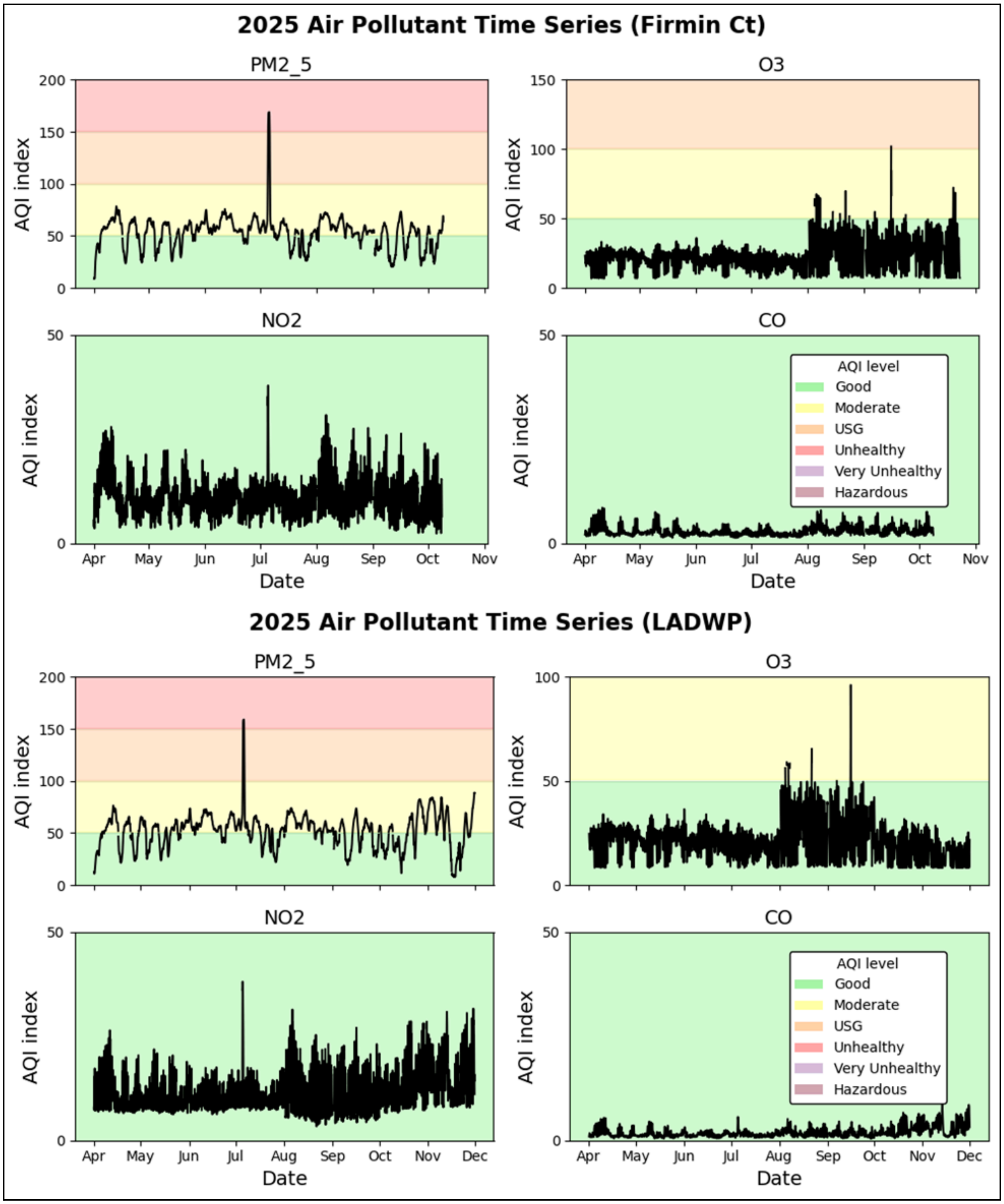

Figure 2.1. Times Series for Multiple Air Pollutants 2025 (Firmin Ct Sensor and LAWDP Sensor)

Note. The colored background indicates EPA AQI health categories and the black lines represent hourly AQI values after data preprocessing and calibration.

Ⅲ. Differences Between Two Observations

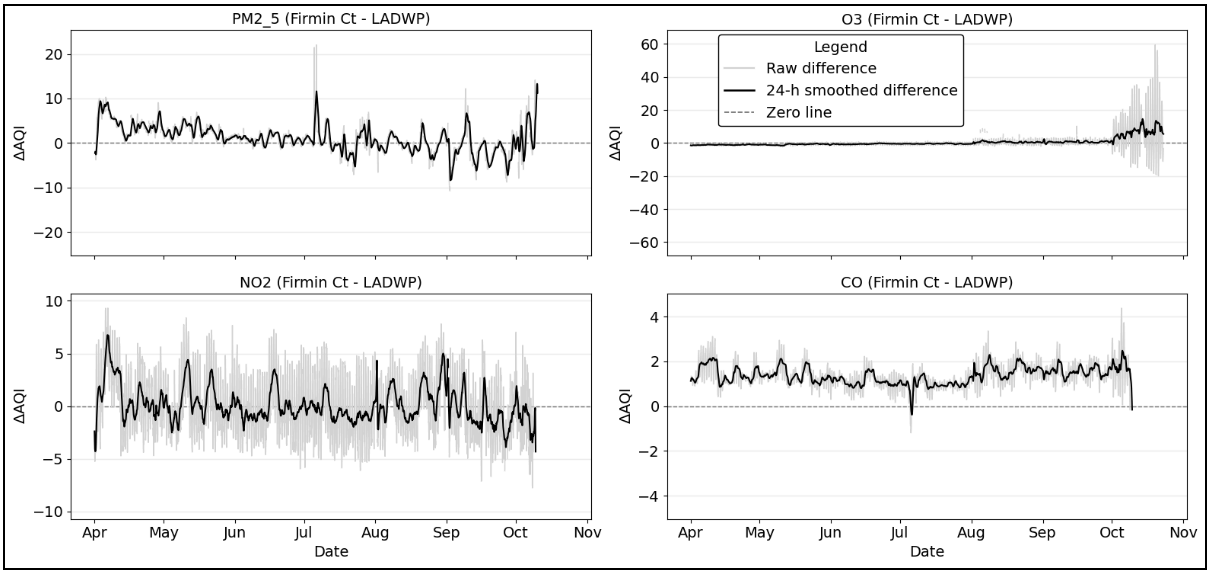

Figure 3.1. Differences in hourly AQI between Firmin Ct and LADWP for $PM_{2.5}$, O_3, NO_2, and CO in 2025.

Note. The gray lines show raw hourly AQI differences, the black lines show 24-hour smoothed differences, and the dashed horizontal lines indicate zero difference between the two sites.

Ⅳ. Statistic of Firmin Ct and LADWP Health Level

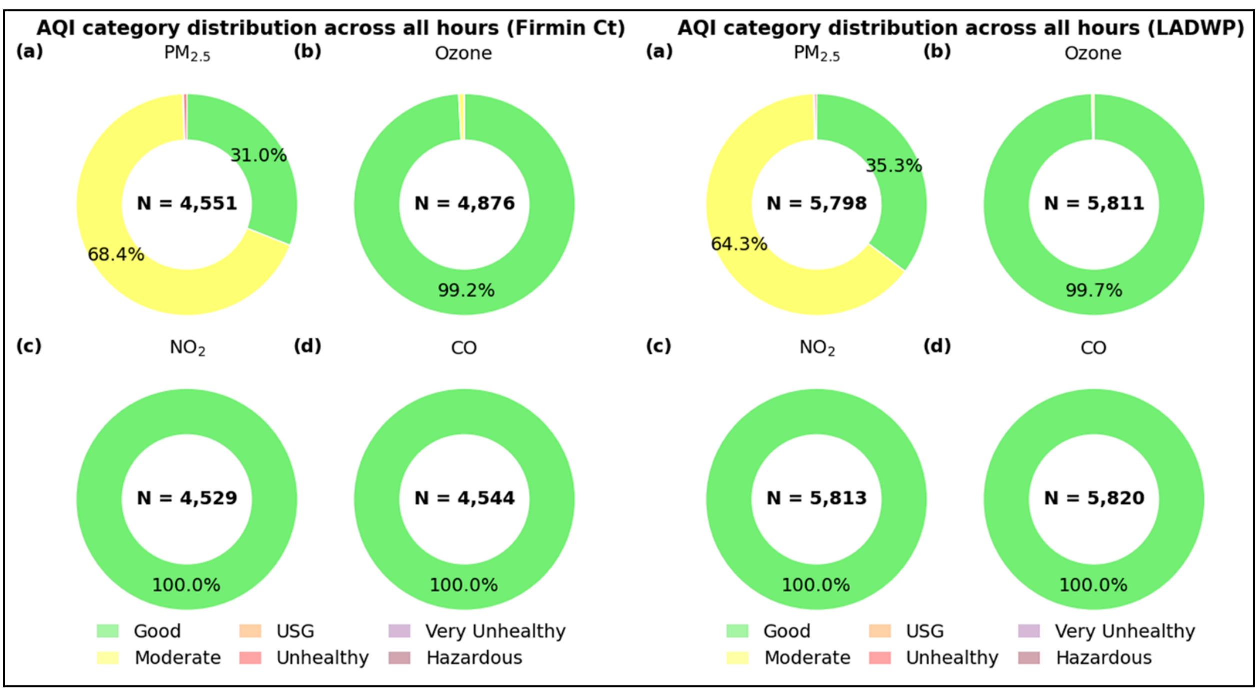

Figure 4.1. Distribution of hourly AQI categories for $PM_{2.5}$, O_3, NO_2, and CO at Firmin Ct and LADWP in 2025.

Note. Donut charts show the fraction of valid hourly AQI records in each EPA AQI category, with the total number of valid records shown at the center of each panel.

Ⅴ. Calendar Maps of Daily Maximum AQI (PM2.5 and Ozone)

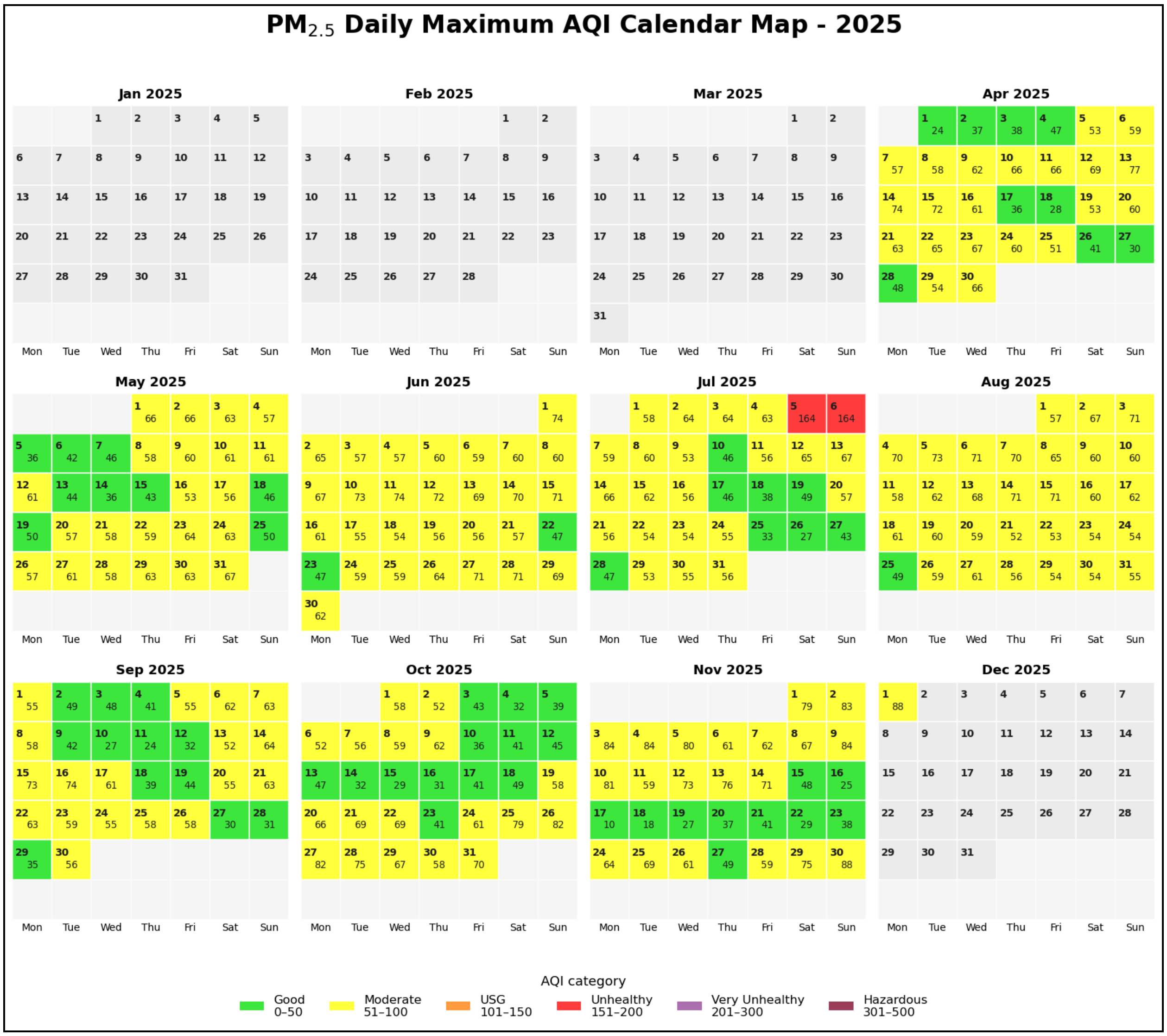

Figure 5.1. Daily maximum $PM_{2.5}$ AQI calendar map in 2025.

Note. Each calendar cell represents the maximum AQI value for that day, with background colors indicating the corresponding EPA AQI category.

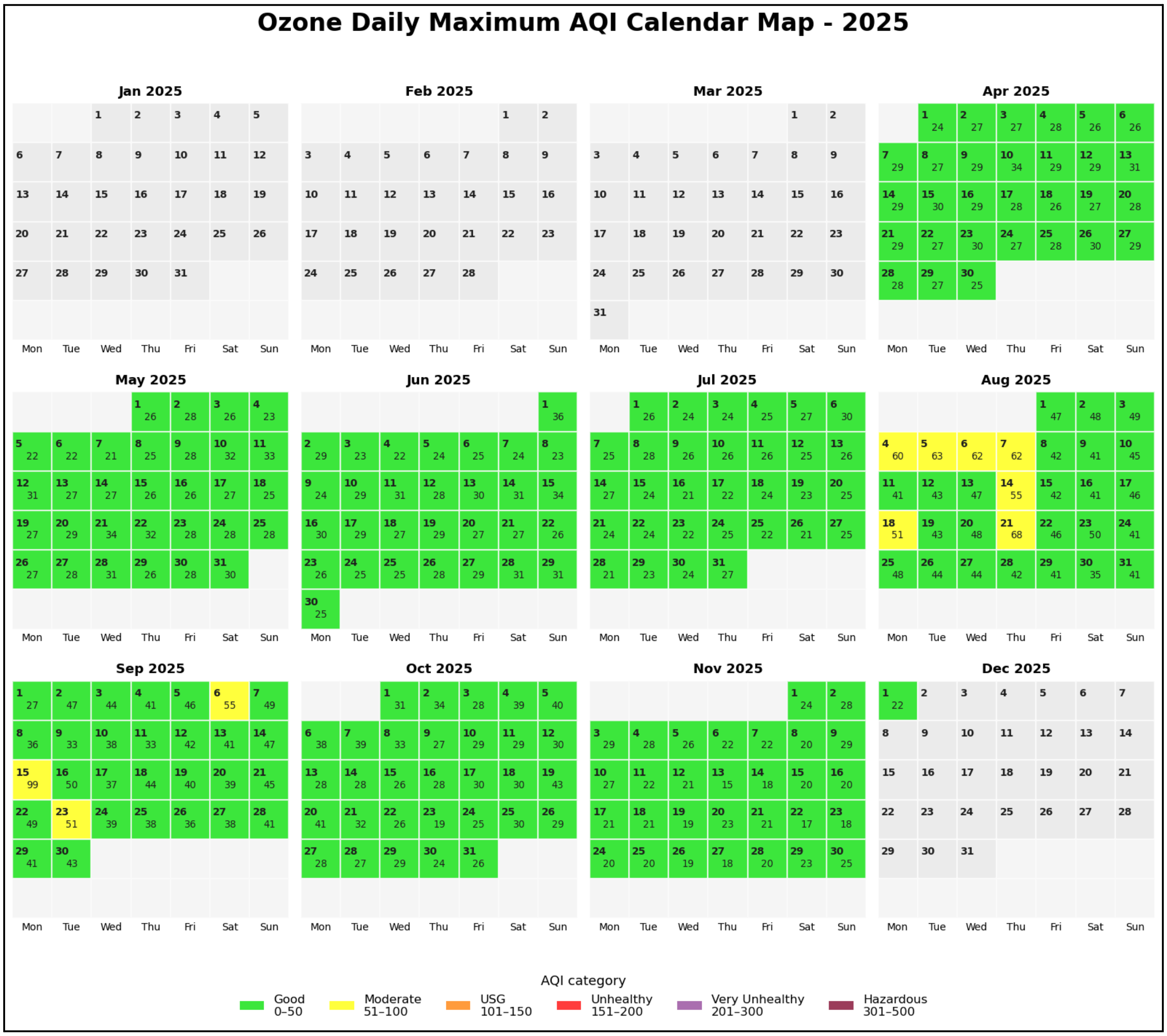

Figure 5.2. Daily maximum ozone AQI calendar map in 2025.

Note. Each calendar cell represents the maximum AQI value for that day, with background colors indicating the corresponding EPA AQI category.

Ⅵ. Diurnal Cycle on weekends/weekdays of different months (PM2.5 and Ozone)

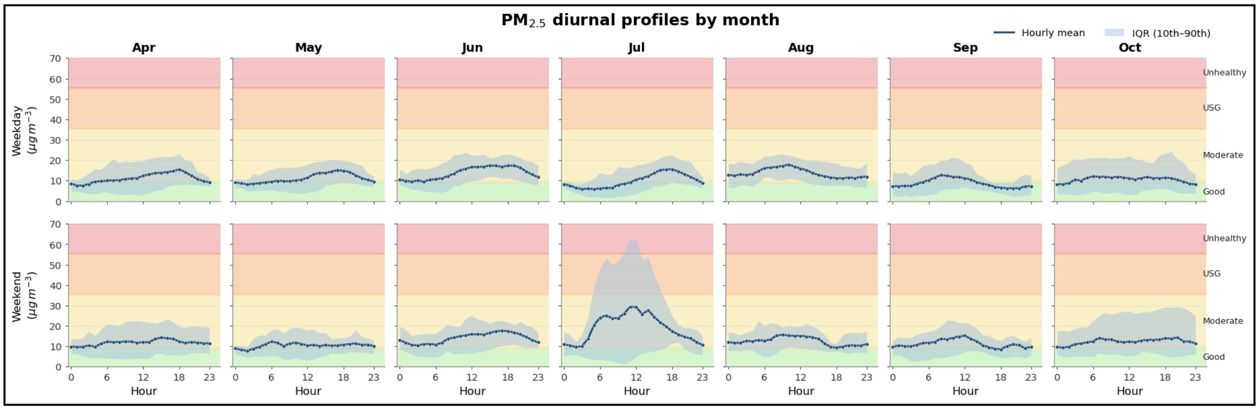

Figure 6.1. Monthly diurnal profiles of PM*₂.₅ *AQI.

Note. Lines represent hourly mean values, and shaded areas indicate the 10th–90th percentile range for weekday and weekend conditions from April to October 2025.

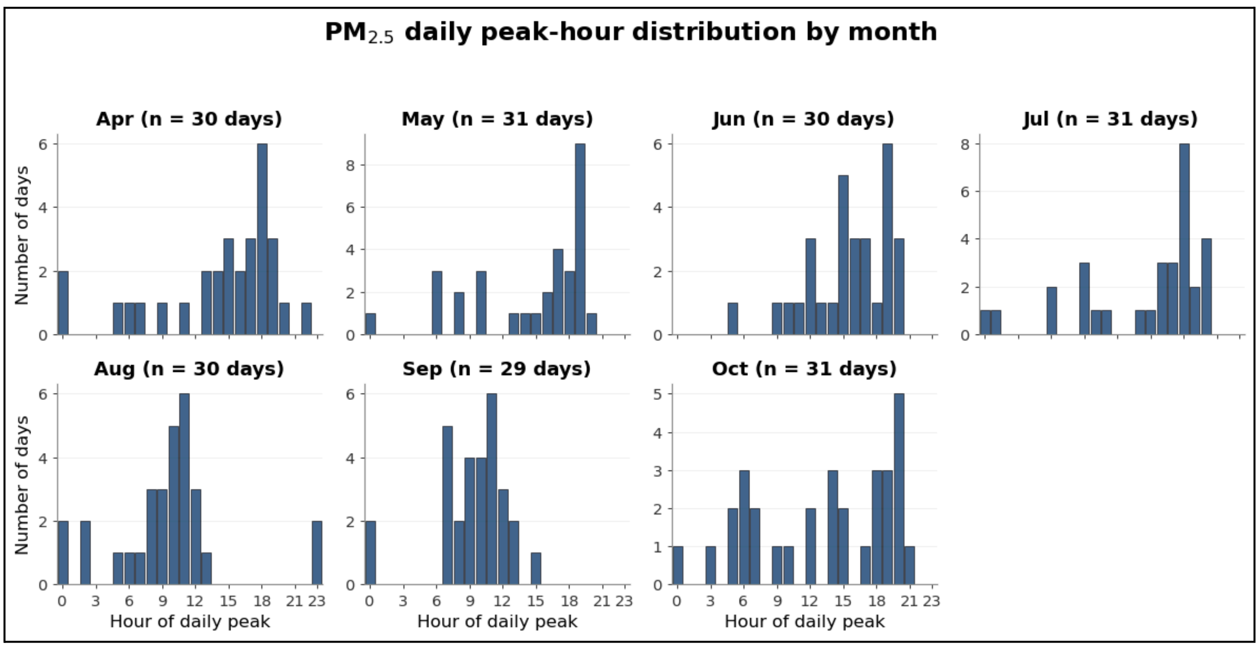

Figure 6.2. Monthly distribution of the daily PM*₂.₅ *AQI peak hour.

Note. Histograms show the number of days when the daily maximum PM₂.₅ AQI occurred at each hour from April to October 2025.

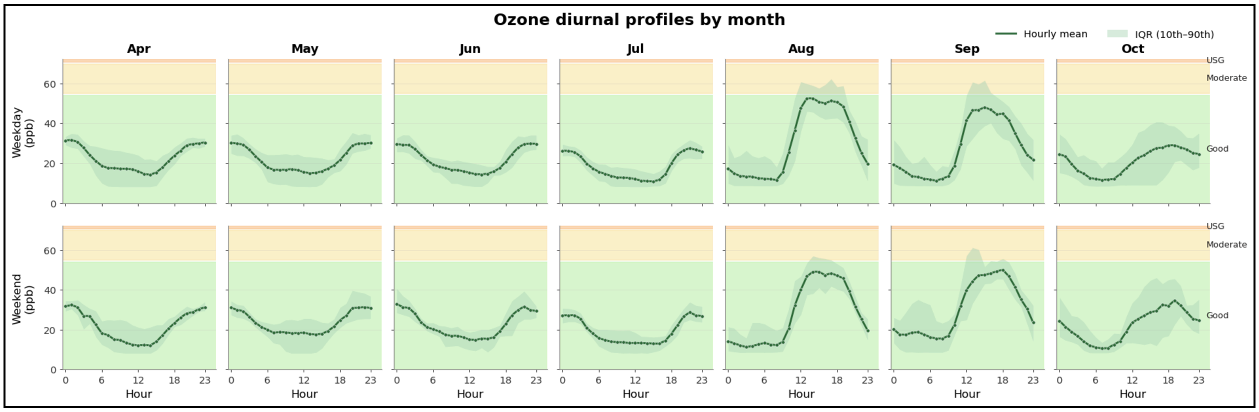

Figure 6.3. Monthly diurnal profiles of ozone AQI.
Note. Lines represent hourly mean values, and shaded areas indicate the 10th–90th percentile range for weekday and weekend conditions from April to October 2025.

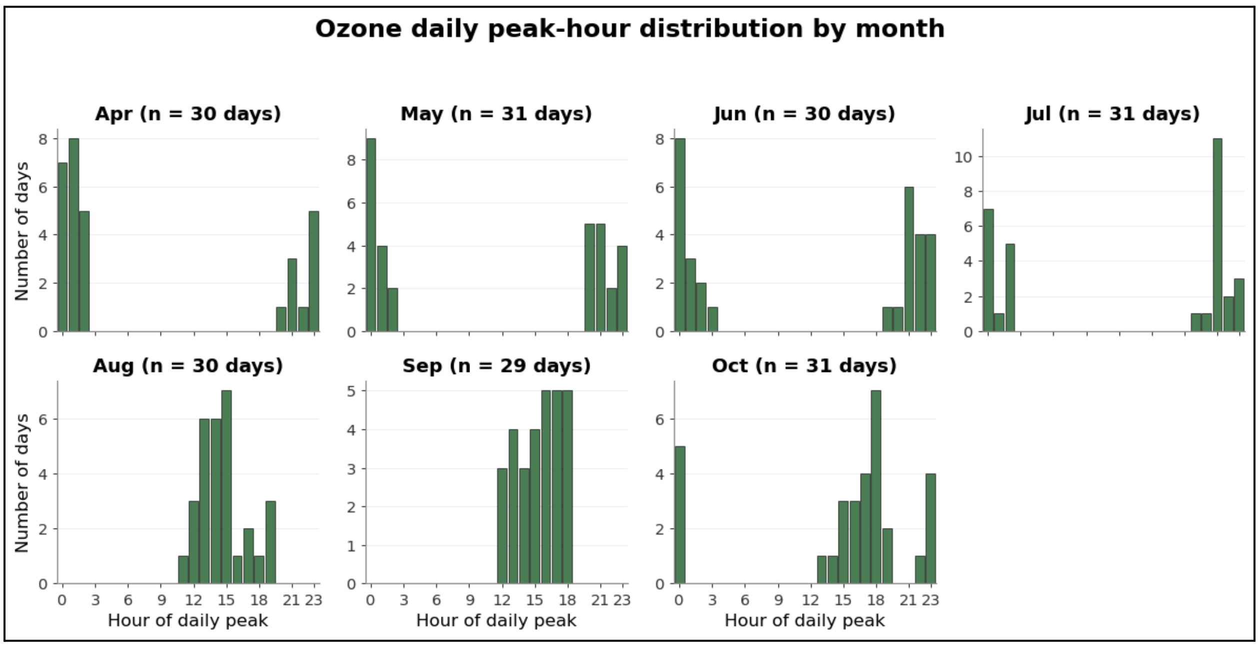

Figure 6.4. Monthly distribution of the daily ozone AQI peak hour.
Note. Histograms show the number of days when the daily maximum ozone AQI occurred at each hour from April to October 2025.

Discussion

The results indicate that air quality in the Vista Hermosa community is generally within the “Good” to “Moderate” AQI categories throughout the monitoring period. Although severe pollution events were uncommon, the analysis revealed several temporal and seasonal trends that may still impact sensitive populations within the community.

$PM_{2.5}$ concentration remained relatively stable overall but experienced occasional short-term spikes throughout the monitoring period. These elevated events may be associated with traffic emissions, construction activity, or regional wildfire smoke, all of which commonly affect air quality in Los Angeles. While many observations remained within acceptable AQI ranges, even short-term exposure to elevated $PM_{2.5}$ concentrations can contribute to respiratory irritation and cardiovascular stress, particularly among individuals with asthma or other respiratory conditions.

The diurnal cycle analysis further showed that $PM_{2.5}$ concentrations tended to peak during morning and evening hours. These patterns are consistent with increased commuter traffic and roadway activity during rush-hour periods. The recurring nature of these peaks suggests that transportation-related emissions remain an important contributor to particulate pollution within the study area.

Ozone demonstrated greater seasonal variability compared to the other pollutants, with elevated concentrations occurring more frequently during summer and early fall months. This trend is likely influenced by increased sunlight and warmer temperatures, which promote photochemical ozone formation. The ozone diurnal profiles also showed peak concentrations during midday and afternoon hours when solar radiation is strongest. Although ozone levels generally remained within “Good” or “Moderate” AQI categories, repeated exposure during elevated periods may still negatively impact respiratory health for children, older adults, and other sensitive individuals.

Comparisons between Firmin Ct and LADWP showed relatively small differences across pollutants, suggesting that air quality conditions were generally consistent throughout the study area. This may indicate that pollution levels within the Vista Hermosa community are influenced more strongly by broader regional urban emissions rather than isolated local sources. The similarity between the two monitoring locations also suggests that pollutant exposure patterns are relatively uniform across the surrounding area.

The AQI distribution and calendar map analyses demonstrated that elevated pollution events were intermittent rather than continuous. Most observations remained within the “Good” category, while fewer observations reached “Moderate” or “Unhealthy for Sensitive Groups” levels. Seasonal patterns were most noticeable for ozone, which showed more frequent elevated days during warmer months. PM2.5 spikes appeared more sporadic throughout the monitoring period.

Conclusion

These findings emphasize the importance of community awareness regarding periods of elevated pollution exposure. Vista Hermosa contains residential areas, schools, and community spaces located near major freeway corridors, which may increase exposure risks for vulnerable populations. Community-based strategies such as reducing vehicle idling, encouraging public transportation, increasing green space, and limiting outdoor activity during peak pollution periods may help reduce overall exposure within the area.

This study contains several limitations. The use of low-cost sensors introduces potential measurement uncertainty when compared to regulatory-grade monitoring equipment, despite calibration efforts. Additionally, only two monitoring locations were included in the analysis, which may not fully represent spatial variability across the broader community. Future work should incorporate additional monitoring locations and longer-term datasets to better evaluate pollutant trends and exposure variability within the Vista Hermosa area.

Overall, this study demonstrates the value of low-cost sensor networks in identifying localized air quality trends and improving community awareness of pollution exposure. These findings support the use of accessible air quality monitoring tools to better inform public health decisions and future environmental planning efforts in urban communities.

Author Contributions

Hao Hu	Methodology, Results
Tara Goller	Executive Summary, Introduction, Literature Review, Project Background and Objectives
Melanie Saavedra	Discussion, Conclusion

References

AirSENCE. https://airsence.com/.

California Air Resources Board (CARB). "Current California GHG Emission Inventory Data." 2025, https://ww2.arb.ca.gov/ghg-inventory-data.

Carotenuto et al. "Low-cost air quality monitoring networks for long-term field campaigns: A review." *Meteorological Applications,* Vol 30, No 6, Royal Meteorological Society, Nov 2023. https://doi.org/10.1002/met.2161.

Collier-Oxandale et al. "Using gas-phase air quality sensors to disentangle potential sources in a Los Angeles neighborhood." *Atmospheric Environment,* Vol 233, Elsevier, Jul 2020, pp. 117519. https://doi.org/10.1016/j.atmosenv.2020.117519.

Data USA. "Los Angeles County--LA City (East Central/Silver Lake, Echo Park & Westlake) PUMA, CA" https://datausa.io/profile/geo/los-angeles-county-la-city-east-centralsilver-lake-echo-park-westlake-puma-ca#demographics.

EPA. "AQI Breakpoints." Environmental Protection Agency, May 2026, https://aqs.epa.gov/aqsweb/documents/codetables/aqi_breakpoints.html.

Lu et al. "Estimating hourly $PM_{2.5}$ concentrations at the neighborhood scale using a low-cost air sensor network: A Los Angeles case study." *Environmental Research,* Vol 195, Elsevier, Apr 2021, pp.110653. https://doi.org/10.1016/j.envres.2020.110653.

South Coast AQMD. "Historical Air Quality Data." 2025, https://www.aqmd.gov/home/air-quality/historical-air-quality-data.

Snyder et al. "The Changing Paradigm of Air Pollution Monitoring." *Environmental Science and Technology,* Vol 47, No 20, American Chemical Society, Aug 2013, pp. 11369-11377. https://pubs.acs.org/doi/10.1021/es4022602.

Stewart et al. "Linking Air Quality and Human Health Effects Models: An Application to the Los Angeles Air Basin." *Environmental Health Insights,* Vol 11, The Author (s), Sept 2017, pp. 1-13. *Sage Pub*, https://journals.sagepub.com/doi/pdf/10.1177/1178630217737551.