Article

Efficacy of Low-Cost Sensor Networks at Detecting Fine-Scale Variations in Particulate Matter in Urban Environments

Asrah Heintzelman [1,2], Gabriel M. Filippelli [1,2,*], Max J. Moreno-Madriñan [3], Jeffrey S. Wilson [4], Lixin Wang [1], Gregory K. Druschel [1] and Vijay O. Lulla [5]

1 Department of Earth Sciences, Indiana University-Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA
2 Environmental Resilience Institute, Indiana University, Bloomington, IN 47408, USA
3 Department of Global Health, DePauw University, Greencastle, IN 46135, USA
4 Department of Geography, Indiana University-Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA
5 Independent Researcher, Indianapolis, IN 46214, USA
* Correspondence: gfilippe@iupui.edu

Abstract: The negative health impacts of air pollution are well documented. Not as well-documented, however, is how particulate matter varies at the hyper-local scale, and the role that proximal sources play in influencing neighborhood-scale patterns. We examined $PM_{2.5}$ variations in one airshed within Indianapolis (Indianapolis, IN, USA) by utilizing data from 25 active PurpleAir (PA) sensors involving citizen scientists who hosted all but one unit (the control), as well as one EPA monitor. PA sensors report live measurements of $PM_{2.5}$ on a crowd sourced map. After calibrating the data utilizing relative humidity and testing it against a mobile air-quality unit and an EPA monitor, we analyzed $PM_{2.5}$ with meteorological data, tree canopy coverage, land use, and various census variables. Greater proximal tree canopy coverage was related to lower $PM_{2.5}$ concentrations, which translates to greater health benefits. A 1% increase in tree canopy at the census tract level, a boundary delineated by the US Census Bureau, results in a ~0.12 µg/m^3 decrease in $PM_{2.5}$, and a 1% increase in "heavy industry" results in a 0.07 µg/m^3 increase in $PM_{2.5}$ concentrations. Although the overall results from these 25 sites are within the annual ranges established by the EPA, they reveal substantial variations that reinforce the value of hyper-local sensing technologies as a powerful surveillance tool.

Keywords: $PM_{2.5}$; citizen science; low-cost sensor; PA; tree canopy coverage

1. Introduction

Particulate matter 2.5 ($PM_{2.5}$), defined as particle mass with an aerodynamic diameter that is less than 2.5 µm, is regulated by the United States Clean Air Act and reported as micrograms per meter cubed (µg m^{-3}), which is a mass-based concentration [1]. $PM_{2.5}$ is typically found in higher concentrations in densely populated regions [2], and in combination with other air pollutants such as nitrogen dioxide (NO_2) and O_3 (Ozone), has been associated with serious health effects and an increased risk of mortality [3,4]. $PM_{2.5}$ was ranked #1 as a risk factor for deaths in China and #6 as a risk factor for deaths in the US [3]. Using an atmospheric chemical transport model, Anenberg et al. (2010) estimated that global anthropogenic $PM_{2.5}$ was associated with millions of cases of cardiopulmonary mortality and thousands of cases of lung cancer mortalities, but when reduced to a lower concentration of 5.8 µg m^{-3}, the mortality estimates were reduced by approximately 30% [5]. Between 2010 and 2012, roadway air pollution alone in 42 New York City neighborhoods resulted in 320 deaths and 870 hospitalizations and emergency room visits due to $PM_{2.5}$ [6]. The overlap of high particulate matter, high population, and poor health quality in the Midwestern region of the US results in a high incident of premature mortality [7].

Due to the serious human health impacts, the US Environmental Protection Agency (EPA) requires states to monitor exposure to $PM_{2.5}$ and other gases. The EPA has set exposure limits of $PM_{2.5}$ at 35 $\mu g/m^3$ in a 24-h cycle that is averaged over three years, and a 12 $\mu g/m^3$ annual standard limit. However, many studies indicate that chronic human exposure to levels of air pollution below the EPA standards results in a positive association with adverse health effects including shaping DNA methylation through epigenetic mechanisms, which can have multi-generational effects. Humans are most susceptible to such impacts during pregnancy, childhood, and the elderly stages of life [8–13]. Even though anthropogenic pollution from vehicles and industry result in increased $PM_{2.5}$, there is also a correlation between $PM_{2.5}$ and meteorological variables, which can explain up to 50% of the daily variations in some regions. For instance, on stagnant days, the average $PM_{2.5}$ concentrations tend to be 2.6 $\mu g/m^3$ higher [14,15]. Independent of geography, age, or gender, increasing long-term $PM_{2.5}$ by 10 $\mu g/m^3$ of $PM_{2.5}$ increases non-accidental mortality by 6% [16]. A limitation of the studies that use $PM_{2.5}$ in relation to human health outcomes is that due to the high cost and maintenance of regulatory monitors, most US cities have few monitors to gauge air quality and health burdens, sometimes just one or two sensors for millions of people. Enhanced spatial and temporal detail in monitoring is critical as a host of studies show the important role that local sources play in driving air pollution [14,17,18].

Scarcity of regulatory monitors has led to using satellite data to extend spatial coverage [19,20]. The deployment of low-cost in situ sensors in air quality studies has also increased with the development of more accurate monitoring devices, which has enhanced analysis of fine-scale variations [21–23]. Although low-cost devices have analytical drawbacks, they have been used effectively in research [24,25]. Low-cost PA sensors in particular perform well with respect to EPA regulatory monitors and have been used in several air quality studies [26] as a cheaper alternative to regulatory monitors to examine indoor and outdoor $PM_{2.5}$ [24,27]. The interest in low-cost sensors grew as the need to assess and evaluate personal exposure to airborne pollutants and their impacts on humans and communities became clearer [1,28,29]. For example PA sensors have been found to have high self-consistency and can be used to fill in the gaps from sparse coverage available from regulatory grade sensors [30]. These sensors cost under USD 300, and have been used in cities such as Ft. Collins and Pittsburgh to examine air pollution as well as in a unique study to examine the association of the deployment and demographics in parts of California [30–32].

The goal of this study was to measure the spatio-temporal distribution and variability of outdoor air quality using 25 PA units that record and report real-time $PM_{2.5}$, temperature, and humidity in Marion County, Indiana (Figure 1). Sensors were deployed from August 2018 through November 2019 and the resulting data were used to analyze the relationships between $PM_{2.5}$ and various meteorological, land-use, and census variables in the Pleasant Run airshed in Indianapolis, Indiana. We examined: (1) the factors associated with daily averages of $PM_{2.5}$ values that exceeded World Health Organization (WHO) guidelines of 25 $\mu g/m^3$ (in 2021, this was changed to 15 $\mu g/m^3$) [33], (2) locations with the highest odds ratio of exceeding the daily average of 25 $\mu g/m^3$, and (3) the impact of tree canopy percentage on $PM_{2.5}$ averages at the census tract level.
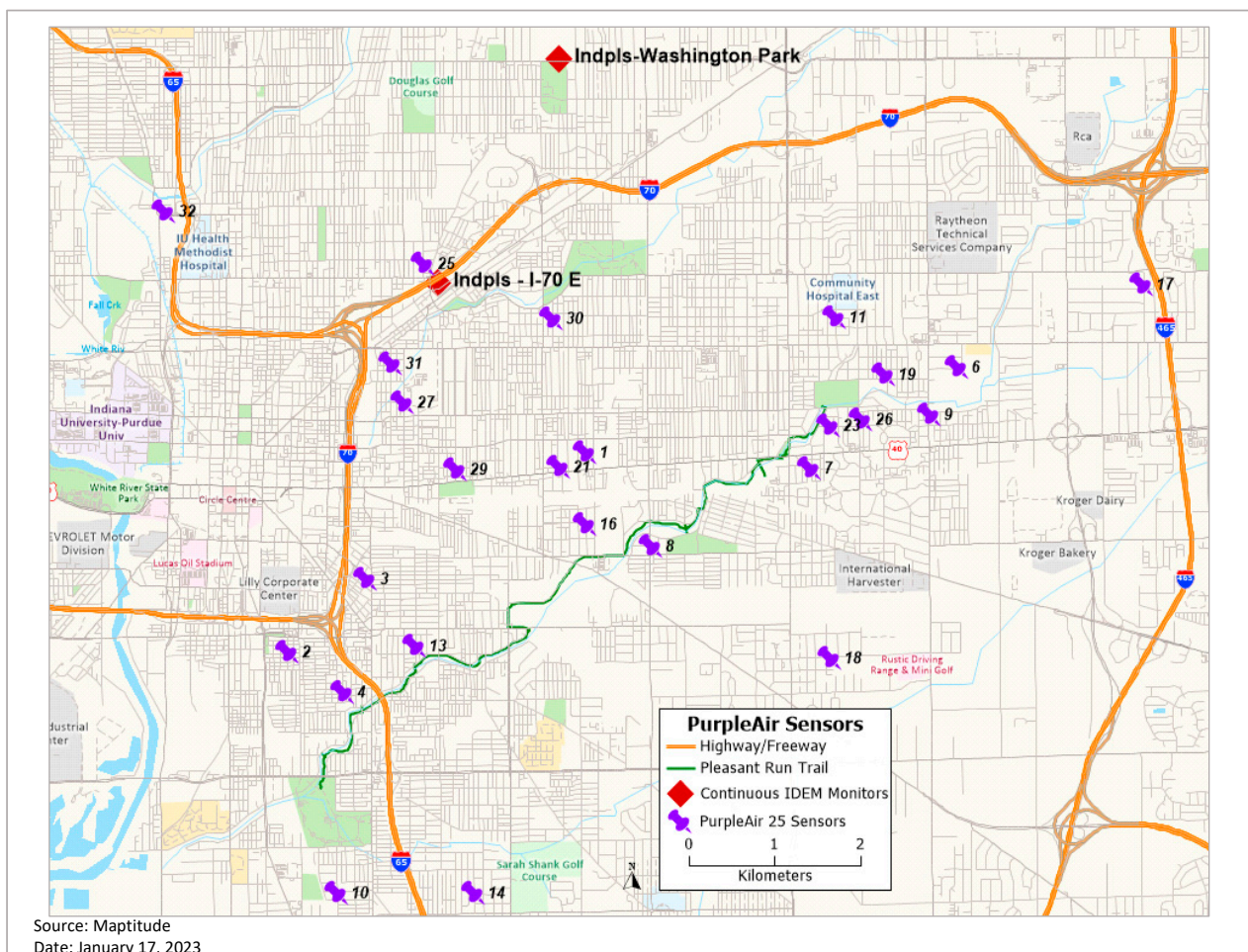
Figure 1. Locations of 25 PA sensors in the study area.

2. Materials and Methods

2.1. Sensor Network

We conducted this research in Indianapolis, partly because it has the characteristics of many large Midwest US cities (e.g., flat topography, multiple sources of PM), but also because it is local, allowing easier access for monitor deployment, QA/QC analysis, and the troubleshooting of sensors. Moreover, our research partner, Keep Indianapolis Beautiful (KIB) had access to citizen science networks from which we recruited volunteers to host the sensors. KIB is a local community-based nonprofit organization with a focus on improving the environment. We identified a list of individual households to contact and recruited 32 citizen scientists within the study area. We installed 32 PA-II-SD sensors, with Wi-Fi capabilities, over a roughly 96 (12×8) km^2 study area in the eastern part of Indianapolis in Marion County, Indiana, to collect PM$_{2.5}$ data (Figure 1). One of the sensors was installed on a three-story balcony. We are assuming that at such a height there may only be a small monotonic gradient difference between the ground and that level for PM$_{2.5}$, based on previous research [34] and our own assessment of monitor performance in comparison to other monitors installed on the first floor of structures. All the sensors, as per the recommendations of the manufacturer, were installed under at least an overhang to provide some protection against the weather. The installation height of all but one sensor ranged from approximately 4 to 8 feet (1.2–2.4 m) above the ground, based on the availability of power outlets and overhead coverage at each location.

PA sensors are designed with a fan that draws a sample of air past its two independent laser counters labeled 'Channel A' and 'Channel B'. Light from a particle is reflected to a detection plate where it is measured by a pulse. Particle size is determined by the length of each pulse, and particle count is determined by the number of pulses. Airborne particulate matter can include organic particles, inorganic particles, smoke, or dust. Particle sizes of 0.3, 0.5, 1, 2.5, 5, and 10 μm are counted and used to calculate mass concentrations (μg/m^3) of $PM_{1.0}$, $PM_{2.5}$, and PM_{10} using an algorithm that was developed by Plantower for the PMS5003 sensor, a factory-calibrated instrument that estimates the number of suspended air particles based on the method outlined above. This compensates for the varying densities of various sources of $PM_{2.5}$. $PM_{2.5}$ readings are averaged and reported every 120 s on an interactive website from which they can also be downloaded. A chip in the sensors (ESP8266) provides the Wi-Fi capabilities enabling the uploading of its collected data to storage in the cloud provided by ThingSpeak. As needed, PA utilizes the sensor's firmware to remotely check the devices and modify its software. An additional BME280 sensor is also incorporated into the PA unit that measures temperature, humidity, and pressure. The correction factor that the PA utilizes to adjust temperature and humidity before reporting it on their interactive map cannot be used universally. This technology, which uses a 5 v USB power source, offers a cheaper option than federal reference monitors to report $PM_{2.5}$. Additionally, PA sensors also report relative humidity and temperature readings [35,36].

The PA sensors evaluated by South Coast AQMD against Federal Equivalent Method instruments costing over USD 20,000 found that laser b (Channel B) reported slightly higher $PM_{2.5}$ mass concentrations; however, the two independent laser counters had a coefficient of determination of over 0.90 (R-square) tested over a range from 0 to 250 μg/m^3, even though the maximum range is ≥1000 μg/m^3 based on the technical documentation. The accuracy of the PA sensors with FEM between 0 and 250 μg/m^3 was between 54.3 and 96.1% (moderate to good accuracy). The maximum consistency error for $PM_{2.5}$ was found at ±10% at 100–500 μg/m^3, and ±10 between 1 and 100 μg/m^3 [36,37]. Due to a much higher number of rows of missing data (9197), we used the output from Channel B instead of Channel A in this study.

2.2. Processing $PM_{2.5}$ Data

The PA sensors were deployed from August 2018 through November 2019, covering 20 census tracts on the east side of Indianapolis. The US Census Bureau (https://www.census.gov/ (accessed on 9 January 2023)) delineates census tracts, the geographical boundaries within a county of each state, with an ideal population of 4000 to gather and report statistical data (https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13 (accessed on 9 January 2023)). Data from 25 of the 32 PA sensors (not in sequential order), deployed over an 11 month period from November 2018 to October 2019, was used in this study. A recall of power cords by the manufacturer in February 2019 resulted in approximately one lost data month as the sensors were disconnected and new cords were delivered and installed in all units. Hourly data from each sensor collected over the study period was downloaded from the PA website and we utilized Base-R in open-source software R-Studio version 4.0.3 (Posit, Boston, MA, USA), an open-source statistical computing software program, for data processing and analysis. Additionally, to validate the output from the PA, at deployment we compared raw PA online readings with an EPA-grade portable sensor (Thermo Scientific MIE pDR-1500).

PA units do not adjust for humidity, which impacts the hygroscopic growth of particles and particle count [30,38]. To reduce this humidity effect and account for the hygroscopic growth, the $PM_{2.5}$ data were corrected using the formula in Equation (1) [32]:

$$\text{C-PM}_{2.5} = \frac{\text{PM}_{2.5}}{1 + \frac{0.25\text{RH}^2}{(1-\text{RH})}} \quad (1)$$

Int. J. Environ. Res. Public Health **2023**, 20, 1934

5 of 18

In Equation (1), C-PM$_{2.5}$ is the corrected PM$_{2.5}$ that is reported by PA in µg/m^3, PM$_{2.5}$ is the raw PM$_{2.5}$ value reported by PA (µg/m^3), and RH is relative humidity retrieved from the Indiana Department of Environmental Management's (IDEM) regulatory air quality monitor at Washington Park in Indianapolis. After this adjustment, the correlation between the data of the three closest sensors (Sensor 25, 30, and 31) to the two IDEM monitors (Washington Park monitor and I-70 monitor) was also examined. The Washington Park (WP) location is the only location that has a photochemical assessment monitoring station (PAMS), providing hourly samples of data that are available via the IDEM website from 1999 to present day [39]. Additionally, the I-70 monitor had more missing data than the WP monitor. Utilizing the WP monitor instead of the PA data in the above equation does introduce some uncertainty into this calibration process. However, since this is a very small region of the city close to the WP monitor, we believe that it is more appropriate to make this adjustment than either not calibrate the data or utilize PA relative humidity values, which are not accurate.

2.3. Meteorological Data

Some abnormal fluctuations in the temperature and humidity data were detected in the PA sensors in our study area. A study in Pittsburg (USA) concluded that due to the shape of the PA units, they can trap heat and increase the inside temperature by an average of 2.7 °C (36.86 °F) and lower the humidity by an average of 9.7% versus outside the unit [30]. This necessitated downloading meteorological data (temperature, relative humidity, windspeed) from the IDEM website reporting data from the air quality monitor at WP (https://www.in.gov/idem/airquality/ (accessed on 20 December 2020)). They have recently changed their reporting system, which now requires a request to the air monitoring department of IDEM for data needed for analysis. WP data are continuously monitored by a photochemical assessment monitoring station and electronically transmitted to IDEM for processing. Meteorological factors such as relative humidity, windspeed, and drought conditions play a role in PM$_{2.5}$ concentrations [14,40–42].

2.4. Land Use and Land Cover

The estimates of tree canopy cover (TC) in the study region were based on a 1 m resolution land cover raster produced by our partner KIB and the University of Vermont Spatial Analysis Laboratory. The land cover data were derived from the classification of National Agricultural Imagery Program (NAIP) data acquired in 2013 and LiDAR data acquired in 2009. This dataset was used to quantify the tree canopy cover in the study region census tracts and within buffers at varying distances from the PA sensor locations.

Additionally, the distribution of "Heavy Industrial" land use was modeled using a vector-based GIS dataset from the City of Indianapolis's data portal [43] and processed using the mapping software Maptitude 2020 by Caliper$^®$. This polygon layer, with land use information, is utilized by the city of Indianapolis for zoning and planning. This particular variable was incorporated into the analysis to capture its impact on this study area [44].

2.5. American Community Survey

The most recent American Community Survey (ACS) 5-year estimates (2014–2018), at the census tract level, were used for socioeconomic and demographic variables [45]. Census tract data have been used in the past in conjunction with PA to examine environmental justice implications [46].

Statistical analysis including correlation, logistic regressions, and stepwise regression incorporating all of the variables discussed above (Appendices A and B) was conducted using the open-source software RStudio version 4.0.3. The binary response variables for logistic regression represent the success of days at sensor locations that exceeded the WHO daily limit (1) versus days that did not exceed the limit (0). The probability of an outcome varies based on the values of the explanatory variables.

Int. J. Environ. Res. Public Health **2023**, 20, 1934

6 of 18

3. Results

Sparse networks of regulatory monitors established by states are inadequate in providing us with local information, limiting the use of these data to make informed decisions on zoning and the infrastructure required to build robust communities. A dense network of sensors, as established in this study, enables us to examine local impacts that can then be utilized to make future informed decisions. The temporal array for each sensor spans thousands of data points, but the spatial array is limited. Due to a limited dataset, we did not examine a spatial relationship among the sensor arrays.

One clear example of local monitoring to assess event-based air quality variations involves smoke emissions from US Independence Day celebrations on 4 July 2021. As a result, sensor 4, as shown in Figure A1, and other sensors revealed higher daily $PM_{2.5}$ averages due to local fireworks than the values reported later that month (second highest peak) that resulted from massive forest fires in the Pacific Northwest and southern Canada [47–49].

3.1. PA Data and Portable EPA Grade Sensor

At deployment, PA validation with the portable EPA grade sensor indicated that, on average, the values for 21 of the 25 sensors deployed were 13.56% higher than the EPA grade sensor, which is within a relative accuracy of $+/-20\%$ [30]. It is important to note that the PA sensors' factory calibration is based on specific ambient aerosol, which may not be identical to our study region and could have contributed to the observed differences. The calibration of raw data has been shown to reduce errors by as much as 25% for extreme cases and by 10% for typical cases. Additionally, there is a systematic bias between instruments that can be compounded with the variation of particle composition and the sensors' performance in the field, which can be balanced by longer term (one year) averaging, thus significantly reducing the error [30]. Our field validation was limited with the EPA grade portable sensor, typically lasting just ~1 h at the beginning of deployment.

3.2. Correlation of PA Data against IDEM Monitor Data

Spearman's correlation was used to examine the relationship between the $PM_{2.5}$ measurements collected at the two IDEM monitoring sites and the adjusted $PM_{2.5}$ data collected at the three closest PA sensors (labeled 25, 30, and 31; Figure 2). This calculation was based on data spanning 11 months with over 300 observations each for the three sensors (Table 1).
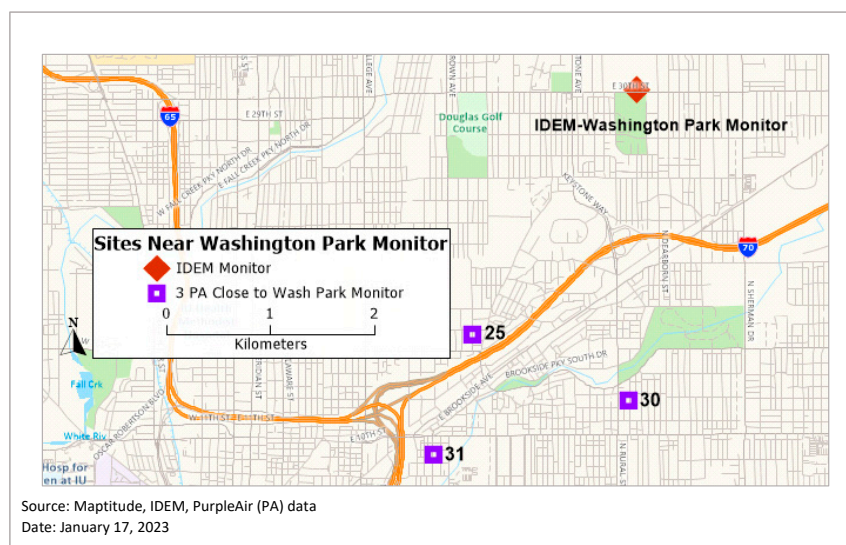

Figure 2. Sensors nearest to Indiana Department of Environmental Management (IDEM) regulatory monitors.

Table 1. Correlation of sensors 25, 30, and 31 with IDEM monitors and PM$_{2.5}$ data.

Sensor Number	IDEM Site	Correlation (Spearman)	*p*-Value
25	WP	0.75	$<2.2 \times 10^{-16}$
30	WP	0.74	$<2.2 \times 10^{-16}$
31	WP	0.72	$<2.2 \times 10^{-16}$
25	I-70	0.72	$<2.2 \times 10^{-16}$
30	I-70	0.71	$<2.2 \times 10^{-16}$
31	I-70	0.72	$<2.2 \times 10^{-16}$

The association between PM$_{2.5}$ readings at the three PA sensors was strongest with the IDEM Washington Park sensor. The range of raw PM$_{2.5}$ was 0.36–103.80 μg m^{-3} with RH ranging from 44 to 99%, and the average temperature was between −2.7 and 85.2 °F (−19 to 30 °C), based on the WP sensor data. After applying the relative humidity correction, this range adjusted to 0.08–80.14 μg m^{-3}, the mean values to 16.22–9.03 μg m^{-3}, and the median to 14.64–8.10 μg m^{-3}. This RH correction underestimation may be more pronounced due to the higher humidity in Indianapolis most of the year versus a less humid environment such as that examined in Colorado, where the median bias was reduced by −15% after the correction [32].

Table 2 shows that when the WP monitor's relative humidity values are used in Equation (1) to correct PM$_{2.5}$, the WP-adjusted values are less than 2% different from the WP data. This indicates that, based on the chemical composition of the pollutants in this region, an RH-based correction (Equation (1)) is adequate for calibration, which is what was used for all further analyses in this study. The mean value of PM$_{2.5}$ is lower in this network of 25 locations than is reported in a prior study in Indianapolis [50]. One possible explanation could be that the sensors in this study are predominantly in residential areas versus being near or on major thoroughfares.

Table 2. Results from applying Equation (1) to PA data.

Variable	Median (μg m^{-3})	% Diff from WP PM$_{2.5}$	Unit Diff from WP PM$_{2.5}$
Raw PM$_{2.5}$	12.874	36.45%	3.439
WP adjusted PM$_{2.5}$ (Equation (1))	9.616	1.92%	0.181
WP PM$_{2.5}$	9.435	0.00%	0.000

3.3. Driving Factors Impacting Daily Averages of PM$_{2.5}$ Exceeding WHO Limits

A total of 8124 counts of cumulative data by month were used in the regression analysis (Table 3). We recognize that the grouping of the months is subjective, but we grouped them, as outlined in Table 3, to best capture the seasonal variations in the study. To identify the potential factors impacting the days that PM$_{2.5}$ values exceed the WHO threshold of 25 μg/m^3, the independent variables listed in Table 4 (Appendices A and B) were used in setting up a logistic regression equation. The dependent binary variable is represented by 1 when the adjusted PM$_{2.5}$ values are greater than or equal to the WHO limit of 25 μg/m^3, and 0 when the adjusted PM$_{2.5}$ values are less than the WHO limit of 25 μg/m^3.

Logistic regression output and its significance based on *p*-values is represented in Table 4 for the days that exceed the WHO limit versus the days that do not. This model indicates that five variables (precipitation, windspeed, Month group 5, Tuesday, and Saturday) are significant in impacting the days exceeding the daily WHO limit (Table 4). Month groups 2–5 are in reference to Month group 1, indicating that all else being equal, Month group 5 is significant when compared to Month group 1. Saturday through Thursday are referenced against Friday. For ease of interpretation, the significant variables from

Table 4 are further extrapolated in Table 5 into the odds of a day exceeding the daily WHO limit. Any odds under 1 are represented by reduced odds and over 1 by increased odds of resulting in a high $PM_{2.5}$ concentration exceeding the WHO daily guidelines. Increased precipitation, increased windspeed, and Month group 5 (compared to Month group 1) have reduced odds of exceeding the WHO guidelines per day (Table 5). On Saturday, the odds of exceeding the WHO limit are sixfold compared to the reference day (Friday), and on Tuesday those odds decrease to threefold.

Table 3. Logistic regression observation breakdown by month for 25 deployed sensors.

Month	Month Group	Observations
2018		
November	1	728
December	2	769
2019		
January	2	739
March	3	753
April	3	726
May	3	773
June	4	702
July	4	752
August	4	751
September	5	733
October	5	698
Total		8124

Table 4. Logistic regression output. Negative values indicate a decrease, and positive values an increase, in the number of days with $PM_{2.5}$ averages above WHO guidelines.

Variable	Estimate	Pr (>|z|)
(Intercept)	−2.05301	0.0014 **
Precip cm	−34.137	0.0060 **
Windspeed kmh	−0.27462	0.0000 **
Temp C	−0.02154	0.2569
Month group 2	−0.58128	0.1487
Month group 3	0.10966	0.7550
Month group 4	−1.08057	0.0518
Month group 5	−1.1233	0.0214 *
Day (Monday)	0.5873	0.2822
Day (Saturday)	1.84268	0.0002 **
Day (Sunday)	1.01762	0.0509 .
Day (Thursday)	0.09561	0.8816
Day (Tuesday)	1.17912	0.0184 *
Day (Wednesday)	−1.07063	0.2030

"." signifies *p*-value < 0.1, "*" signifies *p*-value < 0.05, "**" signifies *p*-value < 0.01.

Table 5. Odds of the significant coefficients from the regression output.

Variable	Odds
Precip cm	0.0000
Windspeed_kmh	0.7599
Month group 5	0.3252
Day (Saturday)	6.3134
Day (Sunday)	2.7666
Day (Tuesday)	3.2515

3.4. Identify Sensors with Highest Odds Ratio of Exceeding the WHO Daily Limit

Odds ratio calculations were run to identify locations within the study area that are the likeliest to exceed the WHO daily limit for $PM_{2.5}$. There were only two sensors out of a total of 25 that had significant odds ratios for days that exceed the WHO daily limit of 25 $\mu g/m^3$. The odds for sensor 16 are 3.04 times that of other locations to meet or exceed 25 $\mu g/m^3$, and for sensor 13 it is 2.37 (Table 6). Interestingly, the two sites proximal to major interstate freeways in the area (sensors 32 and 25) had low odds ratios for exceeding the air quality parameters, perhaps owing to freeway turbulence, lack of idling, and modern emission controls for most vehicles.

Table 6. Odds output from significant sensors.

Sensor#	chi_sq (*p*-val)	Significant Status	OR
13	0.0243	Signif Relationship	2.37
16	0.0008	Signif Relationship	3.04

3.5. Analysis at the Census Tract Level

Correlations at the census tract level aggregation of $PM_{2.5}$ data, census data, tree canopy data, and GIS data revealed that four variables are significant (*p*-value 0.05), with one being almost significant at that level (Table 7, Appendix B). The two negative relationships with Tree Canopy % and White Only % indicate that they have an inverse relationship with $PM_{2.5}$ concentrations. The two positive correlations were found with 'Heavy Ind %' and 'Hwy Length km' with $PM_{2.5}$ concentrations.

Table 7. Correlations of average $PM_{2.5}$ at the census tract level.

Correlations with $PM_{2.5}$ Census Tract Level	Method	Value
Tree Canopy %	Spearman	−0.67 *
Heavy Ind %	Spearman	0.63 *
Hwy Length km	Spearman	0.50 *
Road Length km	Spearman	0.29
Pop25 + LT high school %	Spearman	0.44 [a]
Pop25 + Some College_Assoc %	Pearson	−0.10
Pop25+ Graduate_Prof Degree %	Spearman	−0.20
Hispanic Latino %	Spearman	0.24
Black One Race %	Spearman	0.30
White Only %	Spearman	−0.47 *
Median HH Inc.	Spearman	−0.41
Median Rent	Pearson	0.10

'*' represents *p*-value < 0.05, '[a]' represents almost significant at *p*-value 0.05.

Regression models were run with all of the variables (Table 8, Appendix B). Model 1 resulted in an adjusted r-square of 66.47% with four significant variables at a *p*-value < 0.05. However, this model resulted in a high variance inflation factor (VIF) of 176.17 for 'Black One Race %', which led to running Model 2 without this variable. The second model with an adjusted r-square of 57.94% resulted in a high VIF value of 12.57 for 'Road Length km'. Model 3 was run without 'Road Length km', which resulted in an adjusted r-square of 59.92%, higher that Model 2, but lower than Model 1, with all VIF values < 10. The output of Model 3 is detailed in Table 8, which shows four significant variables at *p*-value < 0.05, and population with some college and associates degree at *p*-value 0.1. As an additional step, stepwise regression model was run on Model 3, which resulted in an adjusted r-square of 62.79% with two significant variables at *p*-value < 0.01 and three variables significant at *p*-value < 0.05. All VIF values for the variables in the final model (stepwise model) are under 10, indicating that multicollinearity is not an issue in this model. We can make the following observations from the final model for our data points at the census tract level:

- A 1% increase in canopy coverage deceases average $PM_{2.5}$ at the census tract level by $0.12 \pm 0.03 \, \mu g/m^3$ (at the 95% confidence interval).
- A 1% increase in Heavy Industrial area increases $PM_{2.5}$ by $0.07 \, \mu g/m^3 \pm 0.02 \, \mu g/m^3$.
- As the percentage of population over 25 with some college and associates degree increases, it results in a proportional increase in $PM_{2.5}$ of $0.08 \, \mu g/m^3 \pm 0.03$.
- Hispanic Latino % has a proportional increase, indicating an increase in this population by one percent results in an increase in $PM_{2.5}$ of $0.06 \, \mu g/m^3 \pm 0.02$.
- Median Rent has an inverse relationship. An increase of USD 100 in median rent results in a decrease in $PM_{2.5}$ of $0.9 \, \mu g/m^3 \pm 0.03$.

Table 8. Stepwise regression output.

	Model 1 (Adj-R^2 = 66.47)	Model 3 (Adj-R^2 = 59.92)	Stepwise (Adj-R^2 = 62.79)
(Intercept)	32.1600 *	16.0000 **	16.66913 **
Tree Canopy %	−0.1840 *	−0.1165 *	−0.1243 **
Heavy Ind %	0.0796 *	0.0724 *	0.069337 *
Hwy Length km	−0.1284	0.0225	
Road Length km	0.0097		
Pop25 + LT high school %	−0.4331	−0.4173	−0.48426
Pop25 + Some College_Assoc %	0.1603 *	0.0790	0.080368 *
Pop25 + Grad Prof Degree %	−0.0131	0.0634	0.044515
Hispanic Latino %	−0.0831	0.0590 *	0.060178 *
Black One Race %	−0.1201		
White Only %	−0.0860	0.0167	0.01149
Median HH Inc	0.0000	0.0000	
Median Rent	−0.0153 *	−0.0083 *	−0.00873 **

"*" signifies *p*-value < 0.05, "**" signifies *p*-value < 0.01.

4. Discussion

Community engagement in meteorology may have its earliest examples dating back to the 1840s [51]. At any level, citizen science poses many challenges as well as opportunities [52]. By creating a link between researchers and the public, such engagement not only benefits the researchers, but also creates a more engaged citizenry [51]. Low-cost sensors such as PA are an inexpensive way for communities to use citizen science to participate in examining air quality at a fine scale [21].

The South Coast AQMD has tested PA sensors to evaluate their performance and found high field accuracy. Our findings generally corroborate this instrumental fidelity, with the correlation between the three sensors tested against the Washington Park sensor of >0.7. Over time, sensor fidelity does degrade [53], along with the added confounding effects of temperature and humidity. Thus, low-cost monitoring data need to be carefully examined over time to ensure that the calibration equation is adequately addressing any data deviations. To suppress the sensor from humidity [24], we calibrated the PA sensor output with humidity values (Equation (1)) from the Washington Park IDEM monitor [54]. This is a limitation of this study, since we were unable to utilize humidity data from each sensor location; however, we believe that, given the study area, utilizing accurate humidity data from the WP monitor was adequate. We also found that a relative humidity correction alone in the PR dataset was adequate (Table 2). We selected several census tracts for this comparison, as this analysis was used as an illustrative example of how AQ related to public health costs rather than a comprehensive analysis of this relationship, which is beyond the capabilities of this sensor array and our experimental approach.

In our study area, Saturday and Sunday were in the top three days for exceeding WHO $PM_{2.5}$ standards. This may be due to a higher volume of vehicle trips during these days on local roadways, with an increase in idling time as well as stop-and-go traffic. More vehicle traffic on residential roads proximal to the sensors during weekends would be captured by the sensors, as opposed to standard work commute days when the traffic would be more limited to traveling along major arterial roads to and from work. A previous study in Indianapolis (Sullivan & Pryor, 2014) found higher $PM_{2.5}$ during weekdays, in contrast to our results, but this is likely driven by the fact that this study used stationary sites from IDEM monitoring sites that are intentionally placed distant from local air pollution sources whereas the previous study sampled air quality using transects that were proximal to freeways and major arterials as opposed to neighborhood-based sites.

We expected that sensors placed near major highways (sensors at location 32 and 25), would have high odds of exceeding WHO limits, but we found instead that sensors 13 and 16 had the highest odds of exceeding daily limits. We expect that the exhaust from traffic on highways creates a temperature gradient between it and the ambient air, which, due to thermal buoyancy causes the plume to rise, which is then impacted by windspeed and direction. Lower windspeeds versus high windspeed on highways result in traffic exhaust plumes dispersing more slowly, thus resulting in higher measurements that are detected for longer periods of time [55]. However, since sensor 32 is upwind from I-65, it is impacted by local versus highway traffic. Sensor 16, on the other hand, along with vehicles, is also impacted by the frequent use of a wood burning stove by its neighbor, thus having a unique local impact, resulting in consistently higher values at that sensor location.

Urban greening initiatives in census tracts with low tree canopy cover can positively impact air quality and reduce related health disparities. Tree canopy coverage will become increasingly important, as urban land is projected to increase to 8.1% in 2050 versus 3.1% in 2000. Indiana's urban land is projected to increase from 8.8% in 2000 to 16.7% by 2050. In 1990, 2.6% of Indiana's tree canopy coverage was in urban areas. This percentage increased to 3.6% in 2000 and is expected to increase to 12.1% by 2050. By 2050, 8.8% of Indiana's tree canopy coverage outside of urban areas will be subsumed by urban growth. This projection equates to 1500 km^2 of land being subsumed by urbanization between 2000 to 2050. Since forests and trees are critical in enhancing human and environmental health, urban canopy cover should be prioritized [56]. Not only does reduced $PM_{2.5}$ improve air quality, but it also reduces expenses. A 10-city study in the USA found that mortality related to $PM_{2.5}$ ranged between 1 and 7.6 people/year, and the average value per mortality incidence was USD 7.8 million. Additionally, the average health benefits were USD 1600 per hectare of tree cover, with an average of USD 1.6 billion in health benefits per $1\mu g/m^3$ reduction [57]. According to this study, the health savings in our 20 census tracts ranged from approximately USD 36,300 (census tract 18097355900) to USD 1,121,585 (census tract 18097361400). This was calculated by multiplying the canopy coverage in each census tract by cost savings of USD 1600 as reported by Nowak et al. (2013). When this calculation is normalized by area, the health savings in census tract 18097361400 is 1.2 times that of the former.

Planting trees in the census tract with lower canopy coverage in conjunction with dividing the study area into low-emission zones (LEZs), where there are restrictions placed on high-polluting vehicles from entering, could give us direct control over minimizing the impacts from vehicle pollution [58] and combatting it through nature simultaneously.

5. Conclusions

This air quality study in Indianapolis is a community-based attempt to examine local impacts of $PM_{2.5}$ over a dense network of 25 sensors. This network was established by collaborating with local partners through community engagement and enhancing the examination of the impacts of canopy coverage along with land use and other variables on $PM_{2.5}$ concentrations in the city.

We found that the relative humidity correction, as captured in Equation (1), is adequate for calibrating raw $PM_{2.5}$ data that is used in all the analysis. Increased percentage of tree canopy coverage at the census tract level was associated with lower $PM_{2.5}$ concentrations. A 1% increase in canopy coverage at the census tract level resulted in lowering $PM_{2.5}$ by approximately 0.12 μg/m^3. Based on research by Nowak et al. (2013), we extrapolate further that the canopy coverage in our study region provides between USD 36,000 and USD 1,121,585 in health savings. Additionally, a 1% increase in the Heavy Industrial area classification in the census tract resulted in increasing $PM_{2.5}$ by a modest 0.07 μg/m^3.

In our logistic regression analysis, we found that increased windspeed and precipitation was associated with lower $PM_{2.5}$ concentrations. We also found that local impacts, as witnessed by the wood burning stove near sensor 16could exceed concentrations measured at locations near major highways. Lastly, because our sensors were deployed in predominantly residential areas, our finding that Saturdays had the highest accumulations of these particles may be due to increased neighborhood traffic on the weekends. Wind direction, which impacts pollutant concentrations, was not modeled in this paper, but should be included in future research.

Air pollutants such as $PM_{2.5}$ can vary both spatially and temporally based on the source(s), atmospheric conditions, and built environment features [59]. Integrating data from dense sensor networks like the one established in our study along with health outcomes data can facilitate understanding of spatial variability in air quality. It can also serve as a tool to facilitate partnership between the community and government entities to quantify and improve air quality as well as inform policies that contribute to creating more resilient communities.

Data Availability Statement: The data presented in this study are openly available at ScholarWorks at IUPUI (https://scholarworks.iupui.edu/handle/1805/30643 (accessed on 1 December 2022)).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix contains spatial patterns and variable explanations used in the $PM_{2.5}$ data and the writing.
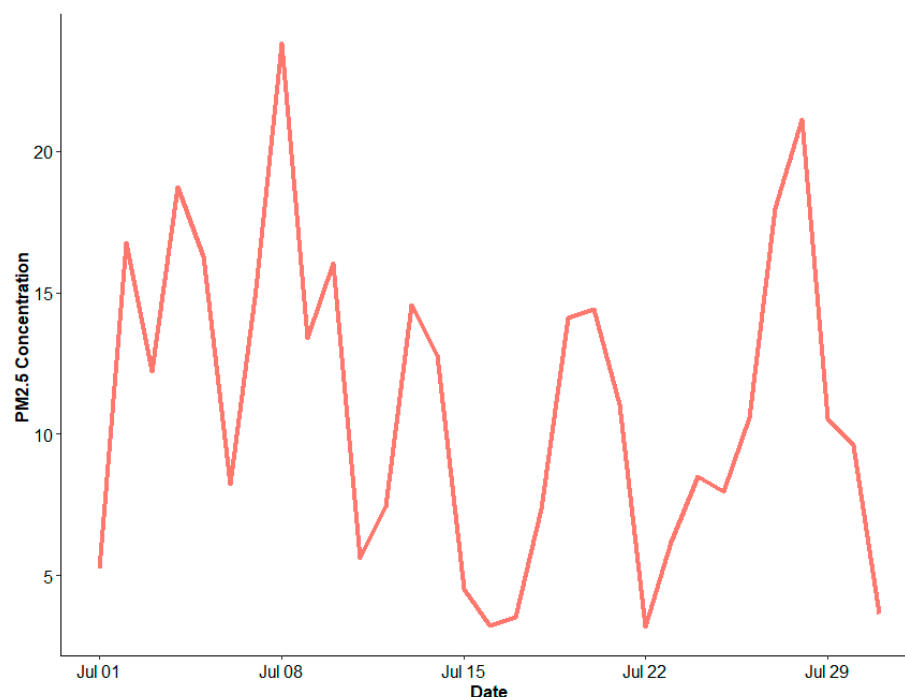
Figure A1. PM$_{2.5}$ from sensor 19 highlighting the first highest peak from 4 July celebrations and the second one from the fires in the northwest of US in 2021.

PA sensor data in micrograms per cubic meter air (μg/m^3):
PM$_{2.5}$ (μg/m^3) = pm25_adj_wp = Relative humidity corrected PM$_{2.5}$.

Washington Park (WP) IDEM monitor data and other data:
Temp C = wp_avg_Temp_C = Temperature in Celsius.
Precip cm = WP_avg_precip_cum_cm = Precipitation in centimeters.
Windspeed kmh = wp_avg_wndspd_kmh_new = Windspeed in kilometer per hour.
Day = day = consists of seven categories for the days of the week.
Month group = paper_mth_grp_rank *
* This variable consists of the following groups:

- Month group 1 = Fall 1 (Nov 2018): paper_mth_grp_rank = "1".
- Month group 2 =Winter (Dec 2018–Jan 2019): paper_mth_grp_rank:= "2".
- Month group 3 = Spring (Mar 2019–May 2019): paper_mth_grp_rank:= "3".
- Month group 4 = Summer (June 2019–August 2019): paper_mth_grp_rank:= "4".
- Month group 5 = Fall 2 (September 2019–October 2019): paper_mth_grp_rank:= "5".

American Community Survey 5-year 2018, at census tract level:
Hispanic Latino % = pop_pct_hispanic_latino = percent of Hispanic and Latino population.
Black One Race % = pop_pct_onerace_black = percent of black one race population.
White only % = pop_pct_whitealone = percent white only population.
Median HH Inc = med_hh_inc = median household income.
Median Rent = median_rent = median rent.
Pop25 + LT high school % = plus25LTHS_ovr25pop = % 25+ population with less than high school.
Pop25 + Some College_Assoc % = plus25_somecollege_assoc= % 25+ population with some college and an associate's degree.
Pop25+ Grad_Prof Degree % = plus25_graduate_prof_deg = % 25+ population with a graduate or a professional degree.

Int. J. Environ. Res. Public Health **2023**, 20, 1934

14 of 18

Indianapolis land use data from open data portal by census tract:
Heavy Industry % = HeavyInd_pct = percentage of heavy industrial parcels.
Measurements (GIS software) in the census tract:
Hwy Length km = hwy_length_km = highway length in kilometers.
Road Length km = rd_length_km = road length in kilometers.
Tree Canopy % = canopy_pct_ct_km = tree canopy percentage.

Appendix B

This appendix contains R code with some explanation for the analysis run for this study in R-Studio.

Meteorology Conversions
d3 <- d1[, ':=' (wp_avg_wndspd_kmh_new=WP_wind_spd*1.6090, WP_avg_precip_cum_cm=WP_precip_cum*2.54,wp_avg_Temp_C=((wp_avg_temp-32)*5/9))]

Logistic regression equation and R-Studio model
Logistic model equation: Log odds(of days when $PM_{2.5}$ exceeds WHO limit) = α + β1(precipitation accumulation)+ β2(windspeed)+ β3(month group)+ β4(day of the week)
summary (mod20 <- glm(who1_pm25_adj_wp_ge25 ~ WP_avg_precip_cum_cm+wp_avg_wndspd_kmh_new
+wp_avg_Temp_C+paper_mth_grp_rank+day,
family=binomial(link="logit"),
na.action=na.omit, data=d2)) # in paper 9/22/22
exp(mod20$coefficients[−1])

Table 6
Following is an example of what was run in R-Studio for each of the 25 sensors in the study:
d1[, sen_13:= as.factor(ifelse(Sensor == '13','1', '2'))]
t_sen13 <- d1[,table(sen_13,who_adj_GE25)]
colnames(t_sen13)<- c("GE25","Other")
rownames(t_sen13)<- c("Sen13", "Other")
chi_13 <- chisq.test(t_sen13)
fisher.test(t_sen13, conf.int = T, conf.level = 0.95)

Section 3.5
Code for Correlation:

```
cor.test(acs4$pm_adj_25Sen_11mt,acs4$canopy_pct_ct_km,          method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$HeavyInd_pct,             method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$hwy_length_km,            method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$rd_length_km,            method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$plus25LTHS_ovr25pop,     method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$plus25_somecollege_assoc, method=c("pearson"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$plus25_graduate_prof_deg, method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$pop_pct_hispanic_latino,  method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$pop_pct_whitealone,      method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$med_hh_inc,             method=c("spearman"))
cor.test(acs4$pm_adj_25Sen_11mt,acs4$median_rent,            method=c("pearson"))
```
Correlation plot (using library(corrplot)) in R-Studio:

Int. J. Environ. Res. Public Health **2023**, 20, 1934
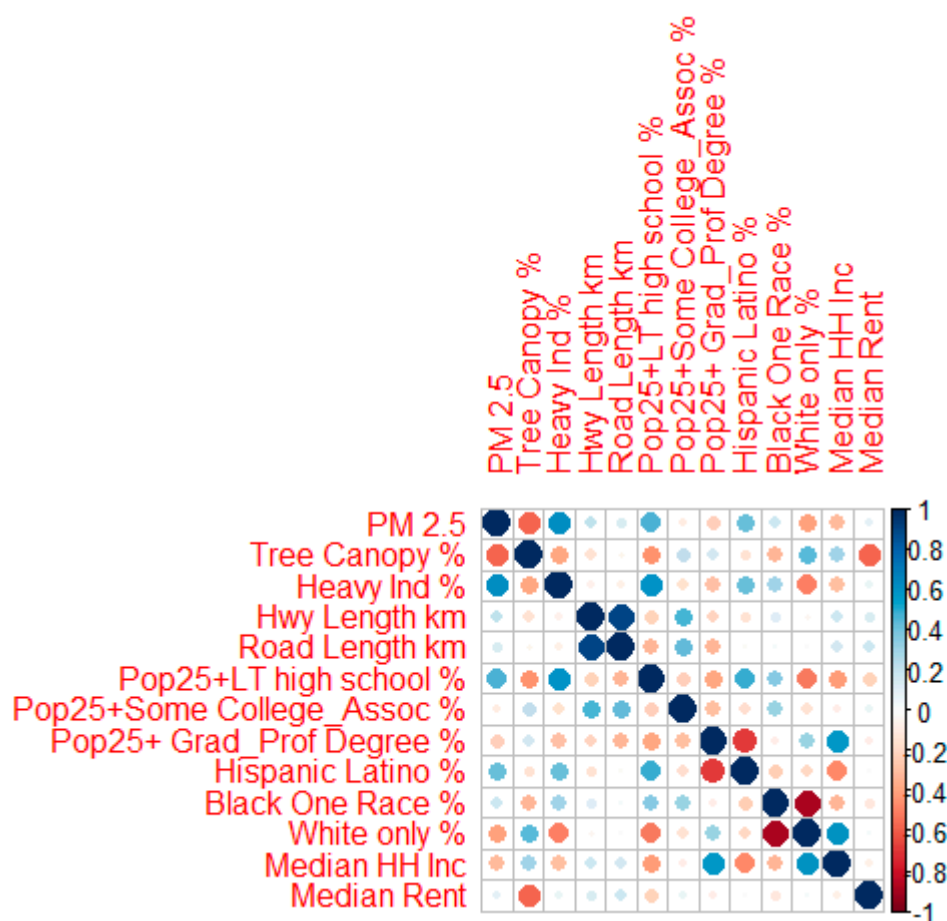
15 of 18

Table 8

Code for Regression:

```
mod1 <- lm(pm_adj_25Sen_11mth ~ canopy_pct_ct_km
+HeavyInd_pct
+hwy_length_km +rd_length_km
+plus25LTHS_ovr25pop+plus25_somecollege_assoc+plus25_graduate_prof_deg
+pop_pct_hispanic_latino
+pop_pct_onerace_black+pop_pct_whitealone
+med_hh_inc+median_rent,data=acs4)
mod2 <- lm(pm_adj_25Sen_11mth ~ canopy_pct_ct_km
+HeavyInd_pct
+hwy_length_km +rd_length_km
+plus25LTHS_ovr25pop+plus25_somecollege_assoc+plus25_graduate_prof_deg
+pop_pct_hispanic_latino
+pop_pct_whitealone
+med_hh_inc+median_rent,data=acs4)
mod3 <- lm(pm_adj_25Sen_11mth ~ canopy_pct_ct_km
+HeavyInd_pct
+hwy_length_km
+plus25LTHS_ovr25pop+plus25_somecollege_assoc+plus25_graduate_prof_deg
+pop_pct_hispanic_latino
+pop_pct_whitealone
+med_hh_inc+median_rent,data=acs4)
```

References

1. Zimmerman, N. Tutorial: Guidelines for implementing low-cost sensor networks for aerosol monitoring. *J. Aerosol Sci.* **2021**, *159*, 105872. [CrossRef]

2. Ji, X.; Yao, Y.; Long, X. What causes PM$_{2.5}$ pollution? Cross-economy empirical analysis from socioeconomic perspective. *Energy Policy* **2018**, *119*, 458–472. [CrossRef]

3. Cohen, A.J.; Brauer, M.; Burnett, R.; Anderson, H.R.; Frostad, J.; Estep, K.; Balakrishnan, K.; Brunekreef, B.; Dandona, L.; Dandona, R.; et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: An analysis of data from the Global Burden of Diseases Study 2015. *Lancet* **2017**, *389*, 1907–1918. [CrossRef]

4. Kasdagli, M.-I.; Katsouyanni, K.; de Hoogh, K.; Lagiou, P.; Samoli, E. Associations of air pollution and greenness with mortality in Greece: An ecological study. *Environ. Res.* **2021**, *196*, 110348. [CrossRef] [PubMed]

5. Anenberg, S.C.; Horowitz, L.W.; Tong, D.Q.; West, J.J. An Estimate of the Global Burden of Anthropogenic Ozone and Fine Particulate Matter on Premature Human Mortality Using Atmospheric Modeling. *Environ. Health Perspect.* **2010**, *118*, 1189–1195. [CrossRef]

6. Kheirbek, I.; Haney, J.; Douglas, S.; Ito, K.; Matte, T. The contribution of motor vehicle emissions to ambient fine particulate matter public health impacts in New York City: A health burden assessment. *Environ. Health* **2016**, *15*, 89. [CrossRef] [PubMed]

7. Fann, N.; Lamson, A.D.; Anenberg, S.C.; Wesson, K.; Risley, D.; Hubbell, B.J. Estimating the National Public Health Burden Associated with Exposure to Ambient PM$_{2.5}$ and Ozone. *Risk Anal.* **2011**, *32*, 81–95. [CrossRef] [PubMed]

8. Christidis, T.; Erickson, A.C.; Pappin, A.; Crouse, D.L.; Pinault, L.L.; Weichenthal, S.A.; Brook, J.R.; Van Donkelaar, A.; Hystad, P.; Martin, R.V.; et al. Low concentrations of fine particle air pollution and mortality in the Canadian Community Health Survey cohort. *Environ. Health* **2019**, *18*, 84. [CrossRef]

9. Chu, M.; Sun, C.; Chen, W.; Jin, G.; Gong, J.; Zhu, M.; Yuan, J.; Dai, J.; Wang, M.; Pan, Y.; et al. Personal exposure to PM$_{2.5}$, genetic variants and DNA damage: A multi-center population-based study in Chinese. *Toxicol. Lett.* **2015**, *235*, 172–178. [CrossRef]

10. Ferrari, L.; Carugno, M.; Bollati, V. Particulate matter exposure shapes DNA methylation through the lifespan. *Clin. Epigenetics* **2019**, *11*, 129. [CrossRef]

11. Shi, L.; Zanobetti, A.; Kloog, I.; Coull, B.A.; Koutrakis, P.; Melly, S.J.; Schwartz, J.D. Low-Concentration PM$_{2.5}$ and Mortality: Estimating Acute and Chronic Effects in a Population-Based Study. *Environ. Health Perspect.* **2016**, *124*, 46–52. [CrossRef] [PubMed]

12. Tan, Y.; Wang, Y.; Zou, Y.; Zhou, C.; Yi, Y.; Ling, Y.; Liao, F.; Jiang, Y.; Peng, X. LncRNA LOC101927514 regulates PM$_{2.5}$-driven inflammation in human bronchial epithelial cells through binding p-STAT3 protein. *Toxicol. Lett.* **2020**, *319*, 119–128. [CrossRef] [PubMed]

13. You, D.; Qin, N.; Zhang, M.; Dai, J.; Du, M.; Wei, Y.; Zhang, R.; Hu, Z.; Christiani, D.C.; Zhao, Y.; et al. Identification of genetic features associated with fine particulate matter (PM$_{2.5}$) modulated DNA damage using improved random forest analysis. *Gene* **2020**, *740*, 144570. [CrossRef] [PubMed]

14. Tai, A.P.K.; Mickley, L.J.; Jacob, D.J. Correlations between fine particulate matter (PM$_{2.5}$) and meteorological variables in the United States: Implications for the sensitivity of PM$_{2.5}$ to climate change. *Atmos. Environ.* **2010**, *44*, 3976–3984. [CrossRef]

15. Wang, P.; Guo, H.; Hu, J.; Kota, S.H.; Ying, Q.; Zhang, H. Responses of PM$_{2.5}$ and O$_3$ concentrations to changes of meteorology and emissions in China. *Sci. Total. Environ.* **2019**, *662*, 297–306. [CrossRef]

16. Chen, H.; Goldberg, M.S.; Viilenueve, P.J. A systematic review of the relation between long-term exposure to ambient air pollution and chronic diseases. *Rev. Environ. Health* **2008**, *23*, 243–297.

17. Li, X.; Feng, Y.J.; Liang, H.Y. The Impact of Meteorological Factors on PM$_{2.5}$ Variations in Hong Kong. *IOP Conf. Ser. Earth Environ. Sci.* **2017**, *78*, 012003. [CrossRef]

18. Heintzelman, A.; Filippelli, G.; Lulla, V. Substantial Decreases in U.S. Cities' Ground-Based NO$_2$ Concentrations during COVID-19 from Reduced Transportation. *Sustainability* **2021**, *13*, 9030. [CrossRef]

19. Bi, J.; Belle, J.H.; Wang, Y.; Lyapustin, A.I.; Wildani, A.; Liu, Y. Impacts of snow and cloud covers on satellite-derived PM$_{2.5}$ levels. *Remote Sens. Environ.* **2019**, *221*, 665–674. [CrossRef]

20. Xiao, Q.; Wang, Y.; Chang, H.H.; Meng, X.; Geng, G.; Lyapustin, A.; Liu, Y. Full-coverage high-resolution daily PM$_{2.5}$ estimation using MAIAC AOD in the Yangtze River Delta of China. *Remote. Sens. Environ.* **2017**, *199*, 437–446. [CrossRef]

21. English, P.B.; Olmedo, L.; Bejarano, E.; Lugo, H.; Murillo, E.; Seto, E.; Wong, M.; King, G.; Wilkie, A.; Meltzer, D.; et al. The Imperial County Community Air Monitoring Network: A Model for Community-based Environmental Monitoring for Public Health Action. *Environ. Health Perspect.* **2017**, *125*, 074501. [CrossRef] [PubMed]

22. Pope, F.D.; Gatari, M.; Ng'Ang'A, D.; Poynter, A.; Blake, R. Airborne particulate matter monitoring in Kenya using calibrated low-cost sensors. *Atmos. Meas. Tech.* **2018**, *18*, 15403–15418. [CrossRef]

23. Tanzer-Gruener, R.; Li, J.; Eilenberg, S.R.; Robinson, A.L.; Presto, A.A. Impacts of Modifiable Factors on Ambient Air Pollution: A Case Study of COVID-19 Shutdowns. *Environ. Sci. Technol. Lett.* **2020**, *7*, 554–559. [CrossRef]

24. Bi, J.; Wildani, A.; Chang, H.H.; Liu, Y. Incorporating Low-Cost Sensor Measurements into High-Resolution PM$_{2.5}$ Modeling at a Large Spatial Scale. *Environ. Sci. Technol.* **2020**, *54*, 2152–2162. [CrossRef]

25. Snyder, E.G.; Watkins, T.H.; Solomon, P.A.; Thoma, E.D.; Williams, R.W.; Hagler, G.S.W.; Shelow, D.; Hindin, D.A.; Kilaru, V.J.; Preuss, P.W. The Changing Paradigm of Air Pollution Monitoring. *Environ. Sci. Technol.* **2013**, *47*, 11369–11377. [CrossRef]

Int. J. Environ. Res. Public Health **2023**, *20*, 1934

17 of 18

26. Mousavi, A.; Wu, J. Indoor-Generated PM2.5 During COVID-19 Shutdowns Across California: Application of the PurpleAir Indoor–Outdoor Low-Cost Sensor Network. *Environ. Sci. Technol.* **2021**, *55*, 5648–5656. [CrossRef]

27. Gupta, P.; Doraiswamy, P.; Levy, R.; Pikelnaya, O.; Maibach, J.; Feenstra, B.; Polidori, A.; Kiros, F.; Mills, K.C. Impact of California Fires on Local and Regional Air Quality: The Role of a Low-Cost Sensor Network and Satellite Observations. *Geohealth* **2018**, *2*, 172–181. [CrossRef]

28. Koehler, K.A.; Peters, T.M. New Methods for Personal Exposure Monitoring for Airborne Particles. *Curr. Environ. Health Rep.* **2015**, *2*, 399–411. [CrossRef]

29. Liang, L. Calibrating low-cost sensors for ambient air monitoring: Techniques, trends, and challenges. *Environ. Res.* **2021**, *197*, 111163. [CrossRef] [PubMed]

30. Malings, C.; Tanzer, R.; Hauryliuk, A.; Saha, P.K.; Robinson, A.L.; Presto, A.A.; Subramanian, R. Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. *Aerosol Sci. Technol.* **2020**, *54*, 160–174. [CrossRef]

31. deSouza, P.; Kinney, P.L. On the distribution of low-cost PM2.5 sensors in the US: Demographic and air quality associations. *J. Expo. Sci. Environ. Epidemiol.* **2021**, *31*, 514–524. [CrossRef] [PubMed]

32. Tryner, J.; L'Orange, C.; Mehaffy, J.; Miller-Lionberg, D.; Hofstetter, J.C.; Wilson, A.; Volckens, J. Laboratory evaluation of low-cost PurpleAir PM monitors and in-field correction using co-located portable filter samplers. *Atmos. Environ.* **2020**, *220*, 117067. [CrossRef]

33. Krzyzanowski, M.; Cohen, A. Update of WHO air quality guidelines. *Air Qual. Atmos. Health* **2008**, *1*, 7–13. [CrossRef]

34. Sajani, S.Z.; Marchesi, S.; Trentini, A.; Bacco, D.; Zigola, C.; Rovelli, S.; Ricciardelli, I.; Maccone, C.; Lauriola, P.; Cavallo, D.M.; et al. Vertical variation of PM2.5 mass and chemical composition, particle size distribution, NO2, and BTEX at a high rise building. *Environ. Pollut.* **2018**, *235*, 339–349. [CrossRef] [PubMed]

35. Map—PurpleAir. Available online: https://www.purpleair.com/gmap?&zoom=3&lat=39.51596757727815&lng=-99.35539180755613&clustersize=45&sensorsActive2=10080&orderby=L&nwlat=74.04797179959134&selat=-25.54831832561498&nwlng=178.9063269424439&selng=-17.617110557556135 (accessed on 12 February 2021).

36. PurpleAir | Real Time Air Quality Monitoring. Available online: https://www2.purpleair.com/ (accessed on 12 February 2021).

37. PurpleAir PA-II. Available online: https://www.aqmd.gov/aq-spec/product/purpleair-pa-ii (accessed on 12 February 2021).

38. Crilley, L.R.; Shaw, M.; Pound, R.; Kramer, L.J.; Price, R.; Young, S.; Lewis, A.C.; Pope, F.D. Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring. *Atmos. Meas. Tech.* **2018**, *11*, 709–720. [CrossRef]

39. Monitoring, A. Air Toxics Data Map. Available online: https://www.in.gov/idem/airmonitoring/air-quality-data/air-toxics-data-map (accessed on 6 July 2020).

40. Chaloulakou, A.; Kassomenos, P.; Spyrellis, N.; Demokritou, P.; Koutrakis, P. Measurements of PM10 and PM2.5 particle concentrations in Athens, Greece. *Atmos. Environ.* **2003**, *37*, 649–660. [CrossRef]

41. Hart, R.; Liang, L.; Dong, P. Monitoring, Mapping, and Modeling Spatial–Temporal Patterns of PM2.5 for Improved Understanding of Air Pollution Dynamics Using Portable Sensing Technologies. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4914. [CrossRef]

42. Zhang, H.; Wang, Y.; Park, T.-W.; Deng, Y. Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmos. Res.* **2017**, *188*, 64–79. [CrossRef]

43. Open Indy Data Portal. Available online: https://data.indy.gov/ (accessed on 12 February 2021).

44. Harper, A.; Baker, P.N.; Xia, Y.; Kuang, T.; Zhang, H.; Chen, Y.; Han, T.-L.; Gulliver, J. Development of spatiotemporal land use regression models for PM2.5 and NO2 in Chongqing, China, and exposure assessment for the CLIMB study. *Atmos. Pollut. Res.* **2021**, *12*, 101096. [CrossRef]

45. U.S. Census Bureau. American Community Survey (ACS). Available online: https://data.census.gov/cedsci/ (accessed on 6 July 2020).

46. Mullen, C.; Flores, A.; Grineski, S.; Collins, T. Exploring the distributional environmental justice implications of an air quality monitoring network in Los Angeles County. *Environ. Res.* **2021**, *206*, 112612. [CrossRef]

47. Jaipuriar, R. Residents Told to Stay Indoors as Wildfire Smoke Crosses State. *Indianapolis Star*, 2021.

48. Smoke Across North America. Available online: https://earthobservatory.nasa.gov/images/148610/smoke-across-north-america (accessed on 12 February 2021).

49. Wildfire Smoke Pouring into Mid-Atlantic Prompts Air-Quality Alert for D.C. and Baltimore. Available online: https://www.washingtonpost.com/weather/2021/07/20/wildfire-smoke-air-quality-dc/ (accessed on 12 February 2021).

50. Sullivan, R.; Pryor, S. Quantifying spatiotemporal variability of fine particles in an urban environment using combined fixed and mobile measurements. *Atmos. Environ.* **2014**, *89*, 664–671. [CrossRef]

51. Rosner, H. Data on Wings. *Sci. Am.* **2013**, *308*, 68–73. [CrossRef] [PubMed]

52. Broeder, L.D.; Devilee, J.; Van Oers, H.; Schuit, A.J.; Wagemakers, A. Citizen Science for public health. *Health Promot. Int.* **2016**, *33*, 505–514. [CrossRef]

53. Masson, N.; Piedrahita, R.; Hannigan, M. Quantification Method for Electrolytic Sensors in Long-Term Monitoring of Ambient Air Quality. *Sensors* **2015**, *15*, 27283–27302. [CrossRef] [PubMed]

54. Jiao, W.; Hagler, G.; Williams, R.; Sharpe, R.; Brown, R.; Garver, D.; Judge, R.; Caudill, M.; Rickard, J.; Davis, M.; et al. Community Air Sensor Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos. Meas. Tech.* **2016**, *9*, 5281–5292. [CrossRef]

Int. J. Environ. Res. Public Health **2023**, *20*, 1934

18 of 18

55. He, M.; Dhaniyala, S. Vertical and horizontal concentration distributions of ultrafine particles near a highway. *Atmos. Environ.* **2012**, *46*, 225–236. [CrossRef]
56. Nowak, D.J.; Walton, J.T. Projected Urban Growth (2000–2050) and Its Estimated Impact on the US Forest Resource. *J. For.* **2005**, *103*, 383–389.
57. Nowak, D.J.; Hirabayashi, S.; Bodine, A.; Hoehn, R. Modeled $PM_{2.5}$ removal by trees in ten U.S. cities and associated health effects. *Environ. Pollut.* **2013**, *178*, 395–402. [CrossRef]
58. Morfeld, P.; Groneberg, D.A.; Spallek, M.F. Effectiveness of Low Emission Zones: Large Scale Analysis of Changes in Environmental NO2, NO and NOx Concentrations in 17 German Cities. *PLoS ONE* **2014**, *9*, e102999. [CrossRef]
59. Zimmerman, N.; Li, H.Z.; Ellis, A.; Hauryliuk, A.; Robinson, E.S.; Gu, P.; Shah, R.; Ye, Q.; Snell, L.; Subramanian, R.; et al. Improving Correlations between Land Use and Air Pollutant Concentrations Using Wavelet Analysis: Insights from a Low-cost Sensor Network. *Aerosol Air Qual. Res.* **2020**, *20*, 314–328. [CrossRef]