Analysis of Traffic-Related Air Pollution in an Urban Area Using Advanced Statistical Methods

Sandamini Senaratne
Dept of Statistics & Data Science
University of Central Florida
Orlando, FL, USA

***Abstract*—This study is about traffic-related air pollution in a major urban area of a developing nation, focusing on carbon monoxide (CO) and particulate matter (PM2.5 and PM10) pollutants. The air pollutant measurements were taken at 30 different sites during three periods of the day: early morning, morning rush hour, and evening rush hour. Using statistical methods including missing value imputation, hierarchical clustering, principal component analysis (PCA), and mixed-effects modeling, the relationship between environmental variables and pollutant concentrations are analysed. The results of the LSD test indicate that particulate matter levels (PM2.5 and PM10) do not vary significantly across different times of day, suggesting stable levels of particulate matter regardless of traffic conditions. However, CO levels were significantly lower during the early morning compared to rush hours, reflecting higher traffic-related emissions during peak periods. Hierarchical clustering identified five distinct clusters of variables, and PCA highlighted that traffic related variables in Cluster 1 consistently showed significant associations with higher concentrations of pollutants, especially during rush hours. The mixed-effects model further revealed that certain site characteristics, particularly traffic related variables are associated with elevated pollutant levels. Wind related variables were also found to be have a slight association with pollutant levels.**

***Index Terms*—Air Pollution, Carbon Monoxide, Particulate Matter, Mixed-effects Model**

I. Introduction

Air pollution is contamination of the indoor or outdoor environment by any chemical, physical, or biological agent that notifies the natural characteristics of the atmosphere [5]. According to WHO, 99% of the global population is exposed to air that contains high concentrations of pollutants which exceeds the WHO guideline limits. The low and middle income countries suffer in this aspect majorly.

Specifically, air pollution in urban areas is a huge challenge to public health and environmental sustainability. In cities around the world, the quality of air has deteriorated due to various sources of pollution, with traffic-related emissions being a significant contributor. Urbanization and increase in vehicular activity in developing countries have worsened this issue, leading to elevated levels of pollutants that impact both human health and the environment.

Among the pollutants the study focuses on carbon monoxide (CO) and particulate matter (PM). Carbon monoxide is a colorless, odorless gas that can interfere with the body's ability to transport oxygen, leading to a range of health issues, from headaches and dizziness to more severe effects like impaired cognitive function and cardiovascular complications [3]. Particulate matter, classified by its size, is also hazardous. Particulate matter less than 2.5 microns in diameter (PM2.5) can penetrate deep into the lungs and enter the bloodstream, causing respiratory and cardiovascular diseases. Particulate matter less than 10 microns in diameter (PM10) poses similar risks but can also affect the upper respiratory tract [1]. The combined effects of these pollutants are linked to increased morbidity and mortality rates, particularly among vulnerable populations such as children, the elderly, and those with pre-existing health conditions [2] [4] [8].

This study aims to explore the levels and variations of traffic-related air pollutants in a major urban area within a developing nation. The research focuses on three key pollutants: carbon monoxide (CO), particulate matter less than 2.5 microns in diameter (PM2.5), and particulate matter less than 10 microns in diameter (PM10). These pollutants were measured across multiple sites in the city, with data collected at different times of the day: early morning, morning rush hour, and evening rush hour. This gives the researchers the chance to study the relationship between the concentrations of the pollutants with the different periods of the day, given a set of environmental variables related to the area.

The primary objective of this study is to explore the relationships between the physical characteristics of the measurement sites and the concentrations of these pollutants. By analyzing factors such as road width, traffic flow, and closeness to sources of pollution, the research seeks to identify the key variables that influence air quality. Furthermore, understanding how these factors vary across different times of the day can provide insights into the dynamics of pollution distribution and exposure.

To achieve these objectives, the study employs a comprehensive approach to data analysis. The data collection process and the variables involved are described in detail, providing a foundation for the subsequent analytical methods. These methods include missing value imputation to handle incomplete data, hierarchical clustering to identify patterns within the data, principal component analysis to reduce dimensionality and highlight key factors, and mixed-effects modeling to account for the variability across different measurement sites and times.

The results of these analyses are presented and discussed with a focus on their implications for urban air quality management. By identifying the critical factors that influence pollution levels and understanding their variations over time, this research aims to provide valuable insights for urban planners and policymakers. This results in improving public health, and enhancing the overall sustainability of urban environments.

In this paper, we first describe the data collection process and the variables involved in the study. This is followed by an explanation of the statistical methods used to analyze the data, including missing value imputation, hierarchical clustering, principal component analysis, and mixed-effects modeling. Then the results of these analyses are presented and discussed. Finally, the conclusion and discussion are stated.

II. METHODOLOGY

A. Data Collection

The study is focused on the traffic-related air pollution levels in a developing nation. The air pollution is measured by the concentrations of airborne pollutants carbon monoxide (CO) and particulate matter (PM). Altogether, three measurements of carbon monoxide (CO), particulate matter of a diameter less than 2.5 microns (PM2.5), and particulate matter of a diameter less than 10 microns (PM10.5) are taken three times a day, early morning, morning rush hour, and evening rush hour at 30 different sites. Each site is located at the edge of a major urban street or avenue. Each pollutant was measured for a period of 15 minutes and the average pollutant concentration during the the 15 minute period is taken as the response variable to be analyzed. The measurements have been taken on weekdays during a 9 day period in December.

B. Variables

There are sixteen variables taken in the 30 sites. Since three measurements are taken at each site, there are 90 observations present in the dataset. Out of the sixteen variables there are six variables that are dynamic. These variables depends on what was happening at the intersection at the time the measurements were taken. Namely, the the number of gasoline powered vehicles passing by, the the number of gasoline powered vehicles passing by, the number of automobiles/ motorcycles passing by, the number of people crossing, the wind direction, and the wind strength. Moreover, there are ten variables are fixed characteristics of the intersection under study. Those are, the average number of building stories present, whether there are any traffic stops, the traffic flow, the presence of pollution source, number of pollution sources, whether there is a hill present, whether the hill is near or not, the road width and the number of road lanes. Further more, each of the site is signified by a unique site id and the period in which the measurement is taken are categorized as 0 for early morning, 1 for morning rush hour, and 2 for evening rush hour.

C. Statistical Analysis

1) Missing Value Imputation: There were 37 missing values spread across six variables. The variables that contains missing values are $People\_min$, $Wind\_dir$, $Wind\_strength$, $Upwind\_source$, $Upwind\_num$, and $Road\_width$. The continuous variables were imputed using predictive mean matching (PMM) [6], and the categorical variables were imputed using knn algorithm [7]. PMM selects observed values that are closest in predicted value to the missing data point, thus preserving the natural variability and relationships within the dataset. By using the most similar observations to fill in missing categories, k-NN respects the categorical nature of the data and helps maintain consistency in the dataset.

2) Response Variables: In order to check how the response variables PM 2.5, PM 10, and CO behave during 3 periods, early morning, morning rush hour, and evening rush hour, boxplots were drawn.

3) LSD test: The difference in the mean of the pollutants were measured using a LSD test with respect to the period of time and the site location. Since there are only three levels under the variable period, LSD test provides a straightforward and powerful method to compare the means. This does not require conservative adjustments for multiple comparisons that are needed with larger numbers of groups or post hoc tests like Tukey's HSD.

LSD test is a statistical method used to compare the means of different groups to determine if there are significant differences between them. An ANOVA test is conducted to test if there is a significant difference among the group means. If the ANOVA is significant, the LSD value is computed using the formula:

$$\text{LSD} = t_{\alpha/2,\,\text{df}} \cdot \sqrt{2 \cdot \frac{\text{MSE}}{n}}$$

where, $t_{\alpha/2,\,\text{df}}$ is the critical value of the t-distribution at a significance level α with degrees of freedom df, MSE is the Mean Square Error from the ANOVA, and n is the number of observations per group. The absolute difference between the means of two groups is compared to the LSD value. If the difference exceeds the LSD value, the difference between the groups is said to be statistically significant.

4) Hierarchical Clustering and Principal Component Analysis: Since there were a lot of variables that correlated to each other, hierarchical clustering was used to cluster the similar variables together. The $ClustOfVar$ package in R was used in this case.

Hierarchical clustering is a method of cluster analysis that seeks to build a hierarchy of clusters. This method can be either agglomerative (bottom-up) or divisive (top-down). The result of hierarchical clustering is represented as a dendrogram, a tree-like diagram that illustrates the arrangement of the clusters at each level of hierarchy. The height of the branches in the dendrogram reflects the distance or dissimilarity between clusters.

Subsequently, principal component analysis was done to the clustered variables. The first principal component for each of the clusters were extracted for the purpose of further analysis and modeling. The choice of clusters directly influences the interpretation of PCA results, as it determines the grouping of variables that share similar patterns, thereby allowing the principal components to capture the most meaningful variance within each cluster [6].

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms a dataset into a smaller set of uncorrelated variables called principal components. These components capture the most significant variance in the data. The process involves standardizing the data, computing the covariance matrix, and calculating eigenvalues and eigenvectors. The data is then projected onto the principal components, creating a new set of variables that explain the most variance.

5) Mixed Effects Model: Two mixed effects models were fitted using the data provided. One was with all the variables before clustering (model 1), and the other model was after clustering (model 2). For both these models, period and site id were considered as the random effects.

Model 1:

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{16j}X_{16ij} + b_i + \epsilon_{ij}$$

where y_{ij} is the dependent variable for the i^{th} site in the j^{th} measurement period. β_{0j}, β_{1j}, β_{2j}, $\ldots$, β_{16j} are the fixed effects coefficients for the predictors. X_{1ij}, X_{2ij}, $\ldots$, X_{16ij} are the variables related to the fixed effects. b_i represent the random effects (period and site id), and ϵ_{ij} is the residual error with zero mean and constant variance. For $i = 1 \ldots 30$ and $j = 1, 2, 3$

Model 2:

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{5j}X_{5ij} + b_i + \epsilon_{ij}$$

where y_{ij} is the dependent variable for the i^{th} site in the j^{th} measurement period. β_{0j}, β_{1j}, β_{2j}, $\ldots$, β_{16j} are the principal component coefficients for the predictors. X_{1ij}, X_{2ij}, $\ldots$, X_{16ij} are the principal components with respect to the clusters. b_i represent the random effects (period and site id), and ϵ_{ij} is the residual error with zero mean and constant variance. For $i = 1 \ldots 30$ and $j = 1, 2, 3$

III. Results of the Statistical Analysis

A. Behaviour of Response Variables

According to the boxplots displayed in Fig.1, both pollutants PM2.5 and PM10 show slight variations across the periods, with the highest averages in Period 0 and Period 2 for PM2.5 and in Period 1 for PM10. The distributions are right-skewed, indicating that higher values are more common in a few instances. With respect to CO levels, highest values are observed in Periods 1 and 2. The data is right-skewed across all periods, but Period 2 shows a more symmetric distribution.

B. LSD test

For particulate matter with diameter less than 2.5 microns (PM 2.5), the mean differences between any pair of periods are all less than the LSD value (0.0706), indicating that none of the pairwise comparisons are statistically significant (TABLE I). Therefore, all three periods (0, 1, and 2) are in the same group

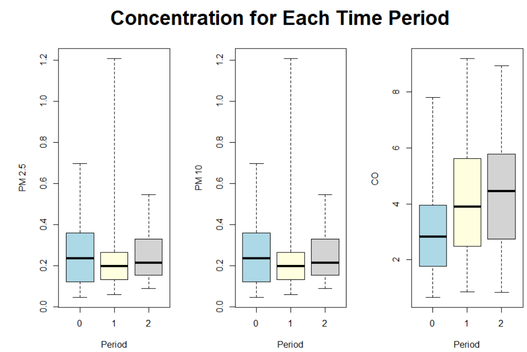

Fig. 1. PM 2.5, PM 10, CO pollutant variation during each period: early morning, morning rush hour, evening rush hour

(a), which indicates that there are no significant differences in PM2.5 concentrations between any of the periods.

For particulate matter with diameter less than 10 microns (PM 10), the mean differences between any pair of periods are all less than the LSD value (0.0911), which means none of the pairwise comparisons show statistically significant differences (TABLE I). Hence, all three periods (0, 1, and 2) are in the same group (a). This shows that there are no significant differences in PM 10 concentrations between any of the periods.

For carbon monoxide (CO) (TABLE I), the grouping indicates that Period 0 (early morning) is significantly different from Periods 1 and 2 (morning and afternoon rush hours). That is Period 0 (early morning) has significantly lower CO concentrations compared to Periods 1 (morning rush hour) and 2 (afternoon rush hour). Periods 1 and 2 are not significantly different from each other, as they fall into the same group.

TABLE I
LSD test of each concentration with respect to the Period

Pollutant / Period	Mean	group	LSD value
PM 2.5			0.07058075
Period 0	0.227809	a	
Period 1	0.2025332	a	
Period 2	0.2220019	a	
PM 10			0.09110742
Period 0	0.2461428	a	
Period 1	0.2571858	a	
Period 2	0.2440212	a	
CO			1.048848
Period 0	3.026914	a	
Period 1	4.137809	a	
Period 2	4.487308	b	

C. Clustering and PCA

The variables was clustered using hierarchical clustering. The cut-off point was decided by the large vertical distances between merges, which suggested distinct clusters. According to the dendrogram there were 5 distinct clusters Fig.2. These clusters consisted of the variables, Cluster 1: Gas_min, $Diesel\_min$, $Autos\_min$, $People\_min$, $Stop\_traffic$, $Traffic\_flow$

Cluster 2: $Wind\_dir$, $Wind\_strength$
Cluster 3: Avg_story, $Hill$, $Hill\_near$
Cluster 4: $Median$, $Road\_width$, $Road\_lanes$
Cluster 5: $Upwind\_source$, $Upwind\_num$
Then a principal component analysis was conducted for the clustered data, where the first principal component that explains majority of the variability of the variables was extracted for further analysis.

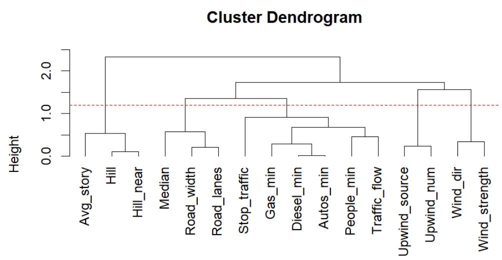

Fig. 2. Hierarchical Clustering of Variables

D. Model Analysis

Model 1

The model with respect to the PM 2.5 and PM 10 concentration, suggests that $Wind\_strength$ significantly impacts the response variable, while other predictors don't show strong evidence of a significant effect (TABLE V, TABLE VI). The model with respect to the carbon monoxide concentration indicates that certain physical site characteristics and environmental conditions (like Gas_min, $Wind\_strength$, and $Hill\_near$) significantly impact the pollutant concentration (TABLE VIII). The response variables were considered in their logarithmic scale due to them being heteroscedastic. Although initially, the model was analyzed with the two random effects mentioned in the methodology section, the matrix was found out to be singular resulting variance of the period random effect to be zero. Hence, this random effect was dropped and only the site id random effect was considered in the modeling. The results of the random effects model are attached in the appendix section .

Model 2

More importance was given for the model with the clustered variables and their principal components. This model is given more importance because it allows for a more nuanced analysis by capturing the underlying relationships within groups of related variables, leading to a more accurate and interpretable model of the data. In this sense, considering only the three periods were resulted in dropping the site id. Therefore, model 2 was considered as a fixed effect model for each of the periods.

In TABLE II, only cluster 1 was found out to be significant. The estimate for Cluster 1 is 0.54463 in logarithmic terms. This suggests that there is a 72.39% increase in the PM 2.5 concentration with respect to the variables under Cluster. This indicates that during the morning hour, variables in Cluster 1 are associated with significantly higher PM 2.5 levels. No significant predictors were identified for PM 2.5 during the morning rush hour period, suggesting that the variables tested did not show any significant association with PM 2.5 levels. During the evening rush hour there is a 17.88% increase in the PM 2.5 concentration. This result indicates that in the evening rush hour, only the variables in Cluster 1 are associated with higher PM 2.5 levels, similar to the morning hour, although the effect is smaller in magnitude.

With respect to the TABLE III Cluster 1 was the only significant variable during the early morning period with a 59.28% increase in PM 10 concentration. This indicates higher PM 10 concentrations in the variables in Cluster 1. No significant predictors were found for PM 10 during the morning rush hour, indicating that the variables tested did not significantly affect PM 10 concentrations in the morning rush hour. During the evening rush hour, cluster 1 was found to be significant with an estimate of 0.16081, a standard error of 0.07102 and a p-value of 0.0329. This denotes that Cluster 1 is associated with higher PM 10 levels during the evening rush hour, with a statistically significant effect.

According to the TABLE IV, Cluster 1 and Cluster 2 are significantly associated with the CO concentrations during the morning hour with a 42.61% increase in CO concentration and a 14.47% decrease in CO concentration for cluster 1 and cluster 2 respectively. During the morning rush hour period, the estimate for Cluster 2 is -0.27019, with a standard error of 0.12541 and a p-value of 0.0415. This result indicates that Cluster 2 is significantly associated with lower CO concentrations during the morning rush hour, with a stronger negative effect compared to Period 0. Cluser 1 is the only significant variable for evening rush hour with an estimate of 0.18481. This shows a significant positive association between Cluster 1 and CO levels during the evening rush hour, indicating higher CO concentrations at these sites.

TABLE II
MODEL 2 RESULTS FOR LOG(PM 2.5)

Period	Parameter	Estimate	Std Error	p-value
0	Cluster 1	0.54463	0.21483	0.0182
1	None	-	-	-
2	Cluster 1	0.16449	0.07139	0.0302

TABLE III
MODEL 2 RESULTS FOR LOG(PM 10)

Period	Parameter	Estimate	Std Error	p-value
0	Cluster 1	0.46549	0.19340	0.0241
1	None	-	-	-
2	Cluster 1	0.16081	0.07102	0.0329

TABLE IV
MODEL 2 RESULTS FOR LOG(CO)

Period	Parameter	Estimate	Std Error	p-value
0	Cluster 1	0.35493	0.14065	0.0186
	Cluster 2	-0.15632	0.06208	0.0189
1	Cluster 2	-0.27019	0.12541	0.0415
2	Cluster 1	0.18481	0.08039	0.0305

IV. DISCUSSION

A. LSD test

The LSD test results indicate no significant differences in PM2.5 and PM10 concentrations across different times of day, suggesting that particulate matter levels remain stable throughout the day, regardless of traffic conditions. This may imply that factors beyond traffic, such as background pollution or non-traffic-related sources, play a larger role in determining particulate matter levels in this urban environment. In contrast, carbon monoxide (CO) concentrations are significantly lower in the early morning compared to the morning and afternoon rush hours. This finding aligns with the expectation that CO emissions are higher during peak traffic periods, reflecting the direct impact of vehicular traffic on CO pollution levels. Studies also shows that high CO emission is detected from vehicles in high traffic roads [9] [10]. These results suggest that while CO pollution may benefit from targeted interventions during rush hours, reducing particulate matter pollution likely requires broader, city-wide strategies.

B. Mixed Effects model

Cluster 1 contains variables related to traffic variables and cluster 2 contains variables related to wind. Therefore, we could further analyse the relationships of these traffic and wind related variables have on pollutant concentrations.

With respect to the mixed effects model results, cluster 1 consistently shows significant positive associations with higher concentrations of PM 2.5, PM 10, and CO across multiple periods, particularly during the morning and evening rush hours. This suggests that variables in Cluster 1 generally experience higher pollution levels. Cluster 2 is associated with lower CO levels during both the morning hour and morning rush hour. However, it does not show significant effects on PM 2.5 or PM 10 concentrations, indicating that its influence might be more specific to CO during these periods.

Period 1 (morning rush hour) lacks significant associations for PM 2.5 and PM 10, suggesting that these pollutants might be influenced by different factors during this time, or that the clusters used do not capture the variation in these pollutants effectively during rush hour. However, CO levels are still significantly impacted by Cluster 2. These results provide insight into how site characteristics influence pollutant concentrations at different times of the day. Therefore, it can be mentioned that certain site characteristics like vehicle fuel type, vehicle type, number of traffic stops, traffic flow, number of people crossing, wind details have a stronger influence on air pollutant levels at specific times of the day [6] [11].

C. Limitations and Future Research

The study is based on data from 30 intersections in a specific developing nation, and the findings may not be generalizable to other urban contexts or countries with different traffic patterns or environmental conditions. Moreover, the data has been gathered only for a period of 9 days. Therefore, seasonal effect could not be quantified, where we could have compared with the yearly pollutant concentrations.

Future research could explore the impact of additional variables, such as weather conditions or seasonal variations, on air pollutant levels. Additionally, investigating the role of specific interventions (e.g., traffic light timing, road modifications) at high-pollution sites could provide more insights.

V. CONCLUSION

The patterns observed from the LSD underscore the complexity of air pollution dynamics in urban settings, highlighting the need for further research into the sources and controls of these pollutants. Despite the site location CO concentrations were found to be high specially during the rush hours. Also, significant PM concentrations were found regardless of the site and period of the time. The study demonstrates the importance of understanding the relationship between urban site characteristics and air pollutant levels, with significant temporal variability observed in these relationships. By identifying key factors that influence pollutant concentrations during different times of the day, the research provides valuable insights for developing targeted strategies to mitigate traffic-related air pollution in urban areas.

REFERENCES

[1] Morakinyo, O.M.; Mokgobu, M.I.; Mukhola, M.S.; Hunter, R.P. Health Outcomes of Exposure to Biological and Chemical Components of Inhalable and Respirable Particulate Matter. Int. J. Environ. Res. Public Health 2016, 13, 592. https://doi.org/10.3390/ijerph13060592

[2] Thangavel, P.; Park, D.; Lee, Y.-C. Recent Insights into Particulate Matter (PM2.5)-Mediated Toxicity in Humans: An Overview. Int. J. Environ. Res. Public Health 2022, 19, 7511. https://doi.org/10.3390/ijerph19127511

[3] Wilbur S, Williams M, Williams R, et al. Toxicological Profile for Carbon Monoxide. Atlanta (GA): Agency for Toxic Substances and Disease Registry (US); 2012 Jun. 3, HEALTH EFFECTS. Available from: https://www.ncbi.nlm.nih.gov/books/NBK153687/

[4] Manisalidis I, Stavropoulou E, Stavropoulos A, Bezirtzoglou E. Environmental and Health Impacts of Air Pollution: A Review. Front Public Health. 2020 Feb 20;8:14. doi: 10.3389/fpubh.2020.00014. PMID: 32154200; PMCID: PMC7044178.

[5] World Health Organization. (2021). Air pollution. https://www.who.int/health-topics/air-pollution#tab=tab_1

[6] Brandon E. Cassidy, Mary Anne Alabanza-Akers, Timothy A. Akers, Daniel B. Hall, P. Barry Ryan, Charlene W. Bayer, Luke P. Naeher, Particulate matter and carbon monoxide multiple regression models using environmental characteristics in a high diesel-use area of Baguio City, Philippines, Science of The Total Environment, Volume 381, Issues 1–3, 2007, Pages 47-58, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2007.03.010.

[7] Beretta, L., Santaniello, A. Nearest neighbor imputation algorithms: a critical evaluation. BMC Med Inform Decis Mak 16 (Suppl 3), 74 (2016). https://doi.org/10.1186/s12911-016-0318-z

[8] WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide https://www.who.int/publications/i/item/9789240034228

[9] Vladimir Lozhkin, Olga Lozhkina, Viktor Dobromirov, A study of air pollution by exhaust gases from cars in well courtyards of Saint Petersburg, Transportation Research Procedia, Volume 36, 2018, Pages 453-458, ISSN 2352-1465, https://doi.org/10.1016/j.trpro.2018.12.124. (https://www.sciencedirect.com/science/article/pii/S2352146518304769)

[10] Shuhaili, A.F.A. & Ihsan, Sany & Faris, Waleed. (2013). Air pollution study of vehicles emission in high volume traffic: Selangor, Malaysia as a case study. WSEAS Transactions on Systems. 12. 67-84.

[11] Opeoluwa Oluwanifemi Akomolafe, Tolulope Olorunsogo, Evangel Chinyere Anyanwu, Femi Osasona, Jane Osareme Ogugua, & Onyeka Henry Daraojimba. (2024). AIR QUALITY AND PUBLIC HEALTH: A REVIEW OF URBAN POLLUTION SOURCES AND MITIGATION MEASURES. Engineering Science & Technology Journal, 5(2), 259-271. https://doi.org/10.51594/estj.v5i2.751

VI. Appendix

TABLE V
MODEL 2 RESULTS FOR LOG(PM 2.5)

Effect	Estimate	Std. Error	t value
Gas_min	0.015538	0.030977	0.502
Diesel_min	0.017700	0.014867	1.191
Peopel_min	-0.003075	0.003097	-0.993
Wind_dir	0.149644	0.077368	1.934
Wind_strength	-0.407143	0.126066	-3.230
Avg_story	-0.115612	0.103714	-1.1115
Median	0.065596	0.235374	0.279
Stop_traffic	0.020756	0.166623	0.125
Traffic_flow	0.030467	0.123561	0.247
Upwind_source	-0.162317	0.229684	-0.707
Upwind_num	0.160527	0.094112	1.706
Hill	0.083664	0.378564	0.221
Hill_near	0.249956	0.249214	1.003
Road_width	0.053402	0.030121	1.773
Road_lanes	-0.158494 0.111820	-1.417	

TABLE VI
MODEL 2 RESULTS FOR LOG(PM 10)

Effect	Estimate	Std. Error	t value
Gas_min	0.013177	0.030504	0.432
Diesel_min	0.020304	0.014568	1.394
Peopel_min	-0.002181	0.003043	-0.717
Wind_dir	0.155768	0.076252	2.043
Wind_strength	-0.373427	0.123938	-3.013
Avg_story	-0.039527	0.100559	-0.393
Median	0.056526	0.228042	0.248
Stop_traffic	0.002887	0.162244	0.018
Traffic_flow	-0.026296	0.121739	-0.216
Upwind_source	-0.136651	0.225632	-0.606
Upwind_num	0.110251	0.092424	1.193
Hill	-0.035592	0.366573	-0.097
Hill_near	0.269978	0.241067	1.120
Road_width	0.039974	0.029136	1.372
Road_lanes	-0.133385	0.108190	-1.233

TABLE VII
FIXED EFFECTS

The respective codes are attached from the next page onwards.

TABLE VIII
MODEL 2 RESULTS FOR LOG(CO)

Effect	Estimate	Std. Error	t value
Gas_min	0.052365	0.025889	2.023
Diesel_min	0.001047	0.012386	0.085
Peopel_min	0.001351	0.002584	0.523
Wind_dir	0.055832	0.064698	0.863
Wind_strength	-0.290176	0.105251	-2.757
Avg_story	-0.024406	0.85803	-0.284
Median	0.248709	0.194630	1.278
Stop_traffic	0.231103	0.138235	1.672
Traffic_flow	0.097310	0.103303	0.942
Upwind_source	-0.079577	0.191661	-0.415
Upwind_num	0.056299	0.078517	0.717
Hill	0.338093	0.312922	1.080
Hill_near	0.063593	0.205860	0.309
Road_width	0.051100	0.024881	2.054
Road_lanes	-0.184129	0.092382	-1.993