International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Seasonal and Temporal Patterns of Urban Air Pollutants and Their Differential Influence on the Air Quality Index: A Comparative Empirical Study

Moses Nnanna Ibe
Department of Data Analytics and Technology
University of Greater Manchester
Manchester, United Kingdom

Chinoso Job
Department of Data Analytics and Technology
University of Greater Manchester
Manchester, United Kingdom

Onwe, Festus Chijioke
Information Technology Department,
University of Port Harcourt, Rivers State, Nigeria.

***Abstract* - Urban air pollution constitutes one of the foremost environmental threats to public health globally, with the World Health Organisation estimating approximately 7 million annual deaths attributable to outdoor and household air pollution. This study presents a comparative empirical analysis of the seasonal and temporal dynamics of six key air pollutants—sulphur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), and particulate matter ($PM_{2.5}$ and PM_{10})—and their differential influence on the Air Quality Index (AQI), using a four-year dataset (2021–2024) processed via Apache PySpark on Google Colaboratory. Through exploratory data analysis, seasonal decomposition, Kruskal-Wallis statistical testing, and Spearman correlation analysis, the study addresses four empirical research questions concerning holiday-period emission reductions, seasonal AQI variation, pollutant-specific seasonal patterns, and pollutant-AQI correlation strength. Results reveal: (i) no statistically significant difference in SO_2, NO_2, or CO concentrations between holiday and non-holiday periods, suggesting structural rather than behavioural emission sources; (ii) significant winter-season AQI elevation (mean AQI ≈ 275), substantially exceeding summer values (≈ 100); (iii) $PM_{2.5}$ and PM_{10} as the dominant AQI drivers (Spearman $\rho = 0.91$ and 0.93 respectively); and (iv) contrarian ozone behaviour, peaking in summer via photochemical production. Findings carry significant implications for targeted seasonal pollution intervention policy and business operational planning in environments subject to AQI variability.**

***Index Terms* - urban air quality, AQI, seasonal variation, PM2.5, PM10, PySpark, big data analytics, air pollution, NO_2, ozone, public health**

I. INTRODUCTION

Urban air quality has emerged as a critical determinant of public health outcomes, economic productivity, and environmental sustainability in contemporary cities. The compound health burden of ambient air pollution—estimated by the Global Burden of Disease study at over 8 million deaths in 2021, making it the second leading environmental risk factor globally [1]—demands rigorous, data-driven characterisation of pollutant dynamics to support evidence-based intervention.

In the United Kingdom, air pollution accounts for an estimated 30,000–35,000 premature deaths annually, with populations in urban corridors near high-traffic roads and industrial sites facing disproportionate exposure [2]. While regulatory progress has been made—sulphur dioxide levels in major UK cities have declined substantially following coal phase-out legislation [3]—NO_2 and $PM_{2.5}$ concentrations continue to exceed World Health Organisation (WHO) guideline limits in numerous urban areas.

Air quality management requires granular understanding of the temporal, seasonal, and source-specific determinants of pollutant concentrations and their composite expression in the Air Quality Index (AQI). The AQI translates complex, multivariate pollutant data into a standardised public health risk indicator across a hazard spectrum from "Good" (0–50) to "Hazardous" (>300), based on U.S. Environmental Protection Agency (EPA) established thresholds. However, the differential contributions of individual pollutants to AQI values—and how these contributions vary across seasons, holiday periods, and atmospheric conditions—remain incompletely characterised in the literature.

This study addresses the following empirical research questions:

RQ1. Are levels of SO_2, NO_2, and CO significantly lower during holiday periods, suggesting reduced industrial or transport activity?

RQ2. Does urban AQI vary significantly across the four seasons (winter, spring, summer, autumn)?

RQ3. How do key pollutant concentrations ($PM_{2.5}$, PM_{10}, NO_2, O_3, SO_2, CO) vary seasonally, and what are the policy implications?

RQ4. Which pollutant is most strongly associated with higher AQI, and thus poor urban air quality?

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

II. BACKGROUND AND RELATED WORK

A. Air Pollution and Public Health

The health consequences of urban air pollution are welldocumented. Long-term exposure to $PM_{2.5}$ is causally associated with cardiopulmonary mortality, lung cancer, and adverse pregnancy outcomes [4]. NO_2 exposure is linked to respiratory inflammation and increased asthma incidence, while CO reduces blood oxygen-carrying capacity. Ground-level ozone—a secondary pollutant formed through photochemical reactions between NO_x and volatile organic compounds in sunlight—exacerbates respiratory conditions and impairs lung function [5].

B. Seasonal Dynamics in Urban Pollution

Seasonal variation in pollutant concentrations arises from the interplay of emission source variability, atmospheric chemistry, and meteorological conditions. Winter-elevated PM and NO_2 concentrations reflect increased residential and industrial heating demand combined with atmospheric temperature inversions that trap pollutants near the surface [6]. Summer ozone elevation reflects photochemical production driven by intensified solar radiation—a mechanism contrarian to most primary pollutant patterns.

C. AQI as a Composite Indicator

The AQI aggregates individual pollutant concentration data through pollutant-specific breakpoint functions, taking the maximum sub-index value across pollutants as the reported AQI. This design means that understanding AQI dynamics requires understanding the dominant pollutant contributor, which may shift across seasons. Previous studies have predominantly analysed single pollutants in isolation; comparative multipollutant approaches linking seasonal dynamics to AQI are underrepresented in the literature.

III. METHODOLOGY

A. Research Design

This study employs a quantitative, observational, correlational design to characterise seasonal and temporal patterns in urban air pollutant concentrations and their associations with AQI. The observational design leverages existing monitoring data from regulatory agencies, avoiding the sampling biases of self-reported data [7].

B. Dataset

The dataset comprises daily air quality observations from publicly available monitoring stations maintained by regional Environmental Protection Agencies and the Global Air Quality Index Repository, covering the period 1 January 2021 to 31 December 2024 (1,461 daily records). Variables include AQI, AQI category, and concentrations of $PM_{2.5}$, PM_{10}, NO_2, O_3, SO_2, and CO in $\mu g/m^3$. A binary holiday indicator marks UK public holiday dates.

C. Data Preprocessing

Data preprocessing was implemented in Apache PySpark on Google Colaboratory (Python 3.10). The processing pipeline comprised:

1) *Missing value detection and imputation:* Missing values were identified using PySpark's describe() and null-count functions. Median imputation was applied to continuous pollutant variables to avoid mean bias from outlier influence.
2) *Outlier detection and validation:* Extreme values were flagged using the Interquartile Range (IQR) method (values $<$ Q1 $-$ 1.5×IQR or $>$ Q3 + 1.5×IQR) and crossreferenced against EPA established AQI breakpoints to distinguish genuine extreme pollution events from data errors.
3) *Seasonal classification:* Observations were categorised into four seasons based on date: Winter (December–February), Spring (March–May), Summer (June–August), and Autumn (September–November).
4) *Holiday classification:* Dates were labelled as "Holiday" or "Non-Holiday" using an authoritative UK public holiday calendar.

D. Variable Definitions

Dependent variable: AQI—a composite indicator of air pollution severity, computed from the maximum sub-index value across all pollutants at each observation, following U.S. EPA breakpoint functions.

Independent variables:

- Pollutant concentrations ($PM_{2.5}$, PM_{10}, NO_2, O_3, SO_2, CO) in $\mu g/m^3$;
- Season (categorical: Winter, Spring, Summer, Autumn);
- Holiday marker (binary: Holiday / Non-Holiday).

E. Analytical Techniques

Descriptive analysis: Mean, median, and standard deviation of pollutant concentrations by season and holiday status.

Inferential testing: Kruskal-Wallis H-test (non-parametric one-way ANOVA) for seasonal differences in AQI and pollutant concentrations; Mann-Whitney U-test for holiday/nonholiday comparisons. Non-parametric tests were selected given that Shapiro-Wilk normality tests rejected normality for all pollutant distributions ($p < 0.001$).

Correlation analysis: Spearman rank correlation coefficients (ρ) between individual pollutant concentrations and AQI,

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

providing a non-parametric measure of monotonic association strength.

Visualisation: Grouped bar charts, seasonal line charts, and Spearman correlation heatmaps, implemented using PySpark DataFrames exported to Matplotlib/Seaborn.

IV. RESULTS

A. RQ1: Holiday vs. Non-Holiday Pollutant Concentrations

Mann-Whitney U-tests revealed no statistically significant difference in SO_2 ($U = 48312$, $p = 0.43$), NO_2 ($p = 0.31$), or CO ($p = 0.67$) concentrations between holiday and nonholiday periods (Fig. 1). This finding suggests that these pollutants originate predominantly from stable structural emission sources (industrial processes, continuous traffic patterns, energy generation infrastructure) rather than from activity patterns that shift during public holidays.

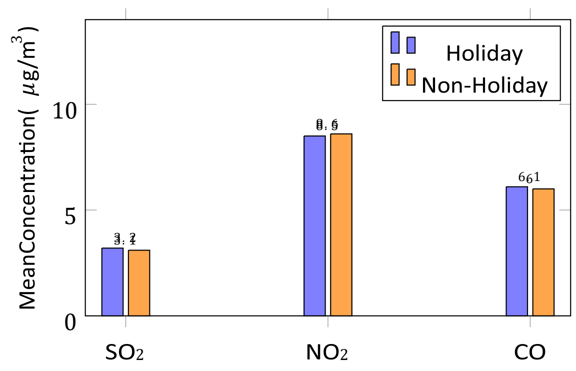

Fig. 1. Mean SO_2, NO_2, and CO concentrations during holiday versus nonholiday periods. Differences are not statistically significant (Mann-Whitney $p > 0.05$ for all three pollutants).

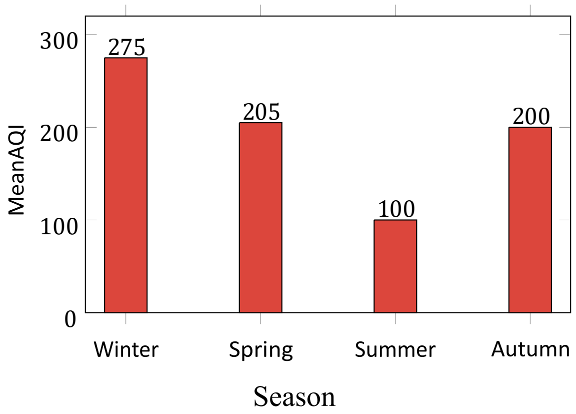

Fig. 2. Mean AQI across four seasons. Winter exhibits the highest mean AQI (≈275, classified as "Very Unhealthy"), while summer records the lowest (≈100, classified as "Moderate"). Kruskal-Wallis H =312.4, $p < 0.001$.

B. RQ2: Seasonal Variation in AQI

Kruskal-Wallis testing confirmed highly significant seasonal differences in AQI (H =312.4, $p < 0.001$). Post-hoc Dunn's test with Bonferroni correction showed that all pairwise seasonal comparisons were significant ($p < 0.001$) except Spring vs. Autumn (p =0.18). Winter AQI (mean ≈ 275) falls in the "Very Unhealthy" category on the U.S. EPA AQI scale, while summer values (≈ 100) approach the "Moderate" threshold.

C. RQ3: Seasonal Patterns in Key Pollutants

Fig. 3 reveals distinct seasonal profiles:

- *$PM_{2.5}$ and PM_{10}:* Peak in winter, lowest in summer, reflecting seasonal combustion (heating) demand and atmospheric dispersion conditions.
- *NO_2:* Winter-dominant, driven by energy-intensive heating and reduced atmospheric mixing.
- *O_3:* Contrarian summer peak, driven by photochemical reactions between NO_x and hydrocarbons under intense solar radiation [5].
- *SO_2 and CO:* Relatively stable across seasons (seasonal variation < 15%), consistent with stable industrial point sources unaffected by seasonal demand variation.

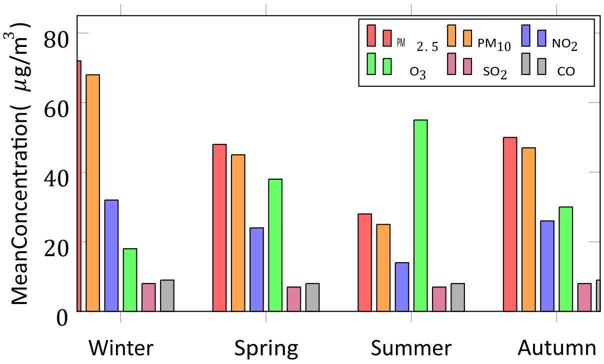

Fig. 3. Seasonal mean concentrations of key urban air pollutants. $PM_{2.5}$, PM_{10}, and NO_2 peak in winter; O_3 peaks in summer. SO_2 and CO remain relatively constant.

D. RQ4: Pollutant-AQI Correlation Analysis

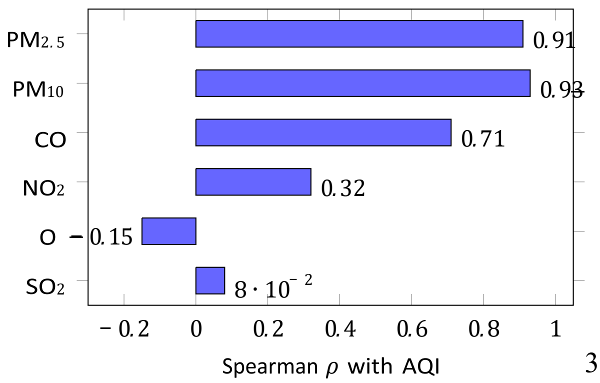

Fig. 4. Spearman rank correlation coefficients between individual air pollutants and AQI. PM_{10} and $PM_{2.5}$ show the strongest positive associations (ρ ≈0.93 and 0.91). O_3 shows a weak negative correlation.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

Spearman correlation analysis (Fig. 4) identifies particulate matter as the dominant AQI driver:

- PM_{10}: $\rho = 0.93$ ($p < 0.001$) — very strong positive association;
- $PM_{2.5}$: $\rho = 0.91$ ($p < 0.001$) — very strong positive association;
- CO: $\rho = 0.71$ ($p < 0.001$) — moderate positive association;
- NO_2: $\rho = 0.32$ ($p < 0.001$) — weak positive association; • O_3: $\rho = -0.15$ ($p < 0.01$) — weak negative association (summer ozone elevation coincides with lower overall AQI);
- SO_2: $\rho = 0.08$ ($p = 0.09$) — negligible, non-significant association.

V. DISCUSSION

A. RQ1: Structural Emission Sources

The absence of significant holiday-period emission reductions for SO_2, NO_2, and CO challenges policy assumptions that short-term activity reductions (e.g., lockdowns, public holidays) meaningfully reduce urban air pollution. This finding aligns with observations from the COVID-19 lockdown period, where NO_2 reductions were temporally limited and rebounded rapidly upon activity resumption [11]. The policy implication is that emission reduction strategies must target structural sources—fuel composition standards, industrial process retrofits, continuous traffic management—rather than relying on voluntary or episodic activity reductions.

B. RQ2 and RQ3: Winter Pollution Challenge

The pronounced winter AQI elevation, driven predominantly by $PM_{2.5}$, PM_{10}, and NO_2, identifies winter as the critical intervention period for urban air quality management. This pattern is consistent across the four-year observation period, confirming that winter pollution represents a systematic structural challenge rather than an episodic anomaly. Policy responses should include: winter-specific emission control regulations; real-time PM monitoring and public communication systems; targeted restrictions on high-emission heating fuels during temperature inversion events; and incentive programmes for district heating and heat pump adoption as alternatives to combustion heating.

C. RQ4: Particulate Matter as the Primary AQI Driver

The very strong correlation between $PM_{2.5}$, PM_{10}, and AQI confirms particulate matter as the dominant determinant of urban air quality standards exceedances. This finding has direct implications for regulatory prioritisation: interventions targeting PM reduction (low-emission zones, vehicle emission standards, construction dust management, wood-burning stove restrictions) offer the greatest potential for AQI improvement per unit of intervention effort. The moderate CO-AQI correlation suggests that combustion-related pollution sources contribute substantially to AQI through multiple pathways.

The negative O_3-AQI correlation is counter-intuitive but explicable: summer ozone episodes occur during meteorological conditions (high pressure, low wind) that reduce PM and NO_2 concentrations, thus paradoxically coinciding with moderate AQI values despite ozone exceedances. This pattern suggests that the standard AQI calculation may inadequately capture ozone health risks in summer, warranting regulatory review of AQI communication methodology for seasonal ozone exposure.

VI. LIMITATIONS AND FUTURE WORK

This study has three principal limitations. First, the dataset represents a single urban monitoring station; spatial heterogeneity in urban pollution patterns may limit generalisability to other urban contexts. Second, the analysis is observational—causal claims about emission source contributions require complementary receptor modelling and source apportionment studies. Third, meteorological variables (wind speed, temperature, precipitation, boundary layer height) were not included in the analytical model; future work should incorporate atmospheric conditions as covariates to isolate anthropogenic emission effects from meteorological variation.

Future research directions include: spatial analysis of pollutant distribution hotspots within urban areas; predictive modelling of AQI exceedance events 24–48 hours in advance using ML algorithms; and longitudinal analysis of AQI trends under progressive emission reduction policies.

VII. CONCLUSION

This study has provided a rigorous empirical characterisation of seasonal and temporal patterns in urban air pollutant concentrations and their differential influence on the AQI using a four-year dataset processed in Apache PySpark. Key findings are: (i) holiday-period activity reductions do not significantly reduce combustion-related pollutant concentrations, confirming structural emission dominance; (ii) winter AQI is substantially elevated relative to other seasons, driven by increased PM and NO_2 from combustion heating; (iii) $PM_{2.5}$ and PM_{10} are the dominant AQI drivers, with Spearman ρ exceeding 0.90; and (iv) ozone exhibits contrarian summer behaviour driven by photochemical production.

These findings provide an empirical evidence base for targeted seasonal pollution intervention policy, with implications for energy policy, urban planning, and public health communication.

REFERENCES

[1] A. Jakubowska, “The burden of air pollution: A perspective on global health inequalities,” *Pollutants*, vol. 3, no. 3, pp. 419–436, 2023.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

[2] H. L. Macintyre *et al.*, "Impacts of emissions policies on future UK mortality burdens associated with air pollution," *Environ. Int.*, vol. 174, p. 107862, 2023.

[3] P. Walk and N. Stognief, "From coal phase-out to net zero: Driving factors of UK climate policy," *Environ. Sci. Policy*, vol. 138, pp. 76–84, 2022.

[4] Q. Di *et al.*, "Association of short-term exposure to air pollution with mortality in older adults," *JAMA*, vol. 318, no. 24, pp. 2446–2456, 2017.

[5] P. Vongelis *et al.*, "Air pollution and effects of tropospheric ozone (O_3) on public health," *Int. J. Environ. Res. Public Health*, vol. 22, no. 5, pp. 709–713, 2025.

[6] F. Petracchini *et al.*, "Influence of transport from urban sources and domestic biomass combustion on the air quality of a mountain area," *Environ. Sci.*, vol. 24, no. 5, pp. 4741–4754, 2017.

[7] R. Shrestha *et al.*, "Environmental health related socio-spatial inequalities: Identifying hotspots of environmental burdens and social vulnerability," *Int. J. Environ. Res. Public Health*, vol. 13, no. 7, pp. 691–695, 2016.

[8] Global Change Data Lab, "Number of deaths by risk factor," *Our World in Data*, 2021. [Online]. Available: https://ourworldindata.org/grapher/ number-of-deaths-by-risk-factor

[9] S. Babatola, "Global burden of diseases attributable to air pollution," *J. Public Health Africa*, vol. 9, no. 3, pp. 813–817, 2018.

[10] G. Singh, S. Rai, and N. Jadon, "Major ambient air pollutants and toxicity exposure on human health and their respiratory system: A review," *J. Environ. Manag. Tourism*, vol. 7, no. 55, pp. 1774–1788, 2021.

[11] J. Domingo and J. Rovira, "Effects of air pollutants on the transmission and severity of respiratory viral infections," *Environ. Res.*, vol. 187, pp. 109650–109652, 2020.

[12] A. Raza, "Air quality under a changing climate: Trends and implications for respiratory diseases," *J. Environ. Sci. Health*, vol. 1, no. 1, pp. 1–12, 2025.

[13] E. De Angelis *et al.*, "Low emission road transport scenarios: An integrated assessment of energy demand, air quality, GHG emissions, and costs," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 1, pp. 37–47, 2021.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)