

A WEATHER-DEPENDENT ANALYTICAL FRAMEWORK FOR MULTI-POLLUTANT URBAN AIR POLLUTION EXPOSURE IN SMALL AND MEDIUM TOWNS OF THE UK (Haider et al., 2026)

- Syed Mustafain Haider¹,
- Richard Hodgkins¹, Dave Ryves¹, Helen Glanville¹
- ¹Department of Geography and Environment, Loughborough University, Loughborough United Kingdom LE11 3TU,

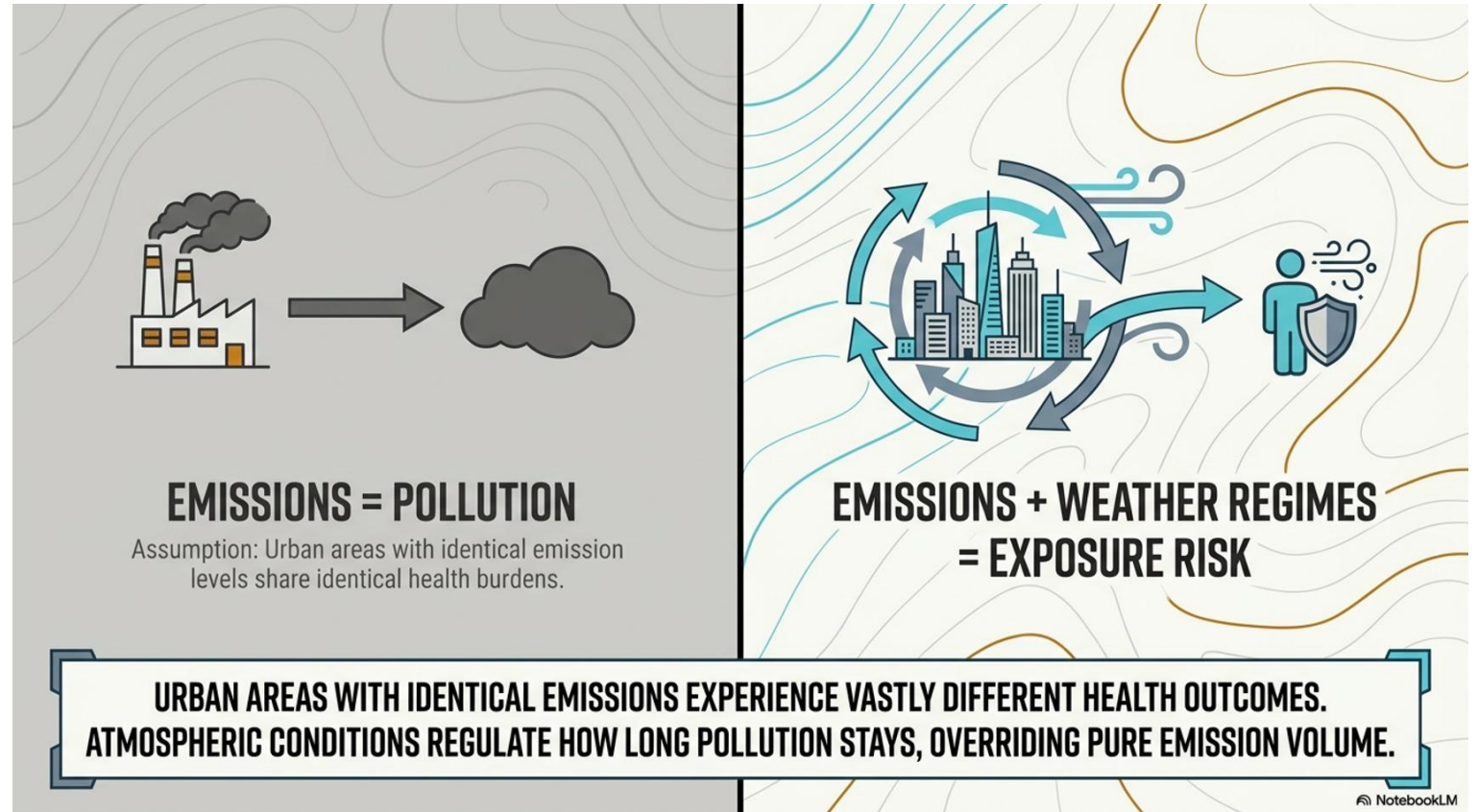


Why Study Weather and Air Pollution Together?

- 4.14 million premature deaths linked to outdoor air pollution (2019)^[1]
- Urban population expected to reach ~70% by 2050^[1]
- **Exposure to:**
 - NO₂
 - PM_{2.5}
 - O₃
 - remains a major public health concern^[2]



Key Challenge



Research Gap

- **Existing Research**

- ✓ Large cities
- ✓ National-scale studies
- ✓ Synoptic weather classifications
- Lamb Weather Types (LWTs) ^[4]
- North Atlantic Oscillation (NAO) ^[6]

- **Limited Understanding**

- Small and medium-sized UK towns
- Multi-pollutant behaviour
- Pollution persistence under different weather conditions

Aim & Objectives

Aim

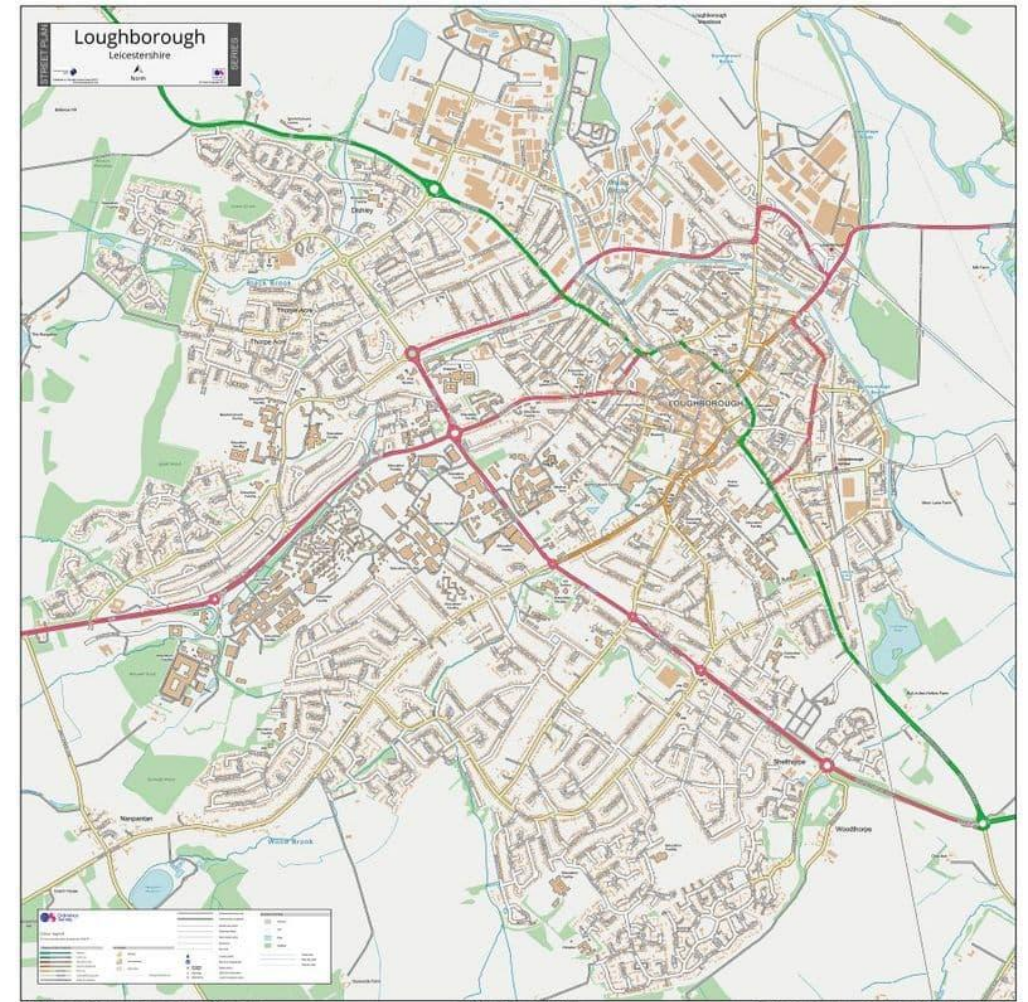
- To investigate how meteorological conditions influence multi-pollutant urban air pollution behaviour in a medium-sized UK town.

Objectives

- Assess whether meteorology can predict pollution states.
- Identify key meteorological drivers of pollution variability.
- Develop machine-learning informed weather regimes.
- Evaluate pollution burden and persistence under each regime.
- Quantify exposure risk associated with different atmospheric conditions.

Study Area

- **Loughborough, UK**
- Medium-sized university town
- Population ~103,000 ^[12]
- East Midlands, United Kingdom
- To investigate pollution behaviour within the town, Loughborough University campus was used as an urban analogue...



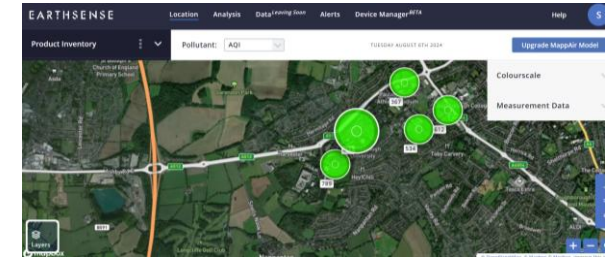
Why Loughborough University?

- Embedded within the urban environment
 - Multiple air-quality monitoring stations
 - Continuous meteorological observations
 - Represents mixed urban activities
 - Academic buildings
 - Residential areas
 - Sports facilities
 - Traffic movement



Data Collection

- The pollution data is usually collected on daily basis for the period of 24 hours from the air sampling units installed at strategic locations in campus.
- Pollutants under study:
 - Nitrogen Oxide (NO)
 - Nitrogen Dioxide (NO₂)
 - Particulate Matter (1, 2.5 and 10)
 - Ozone (O₃)
- The metrological conditions are being monitored and reported by the campus weather station. Weather parameters under study are:
 - Temperature,
 - Humidity,
 - Rainfall
 - Wind Speed
 - Absolute Pressure



METHODOLOGY

A weather-dependent analytical framework



1. Data Assimilation

Hourly multi-pollutant & meteorological data converted to daily means.



2. Dynamic Thresholds

P33/P66 percentiles calculated for initial severity baseline.



3. Weather-Only Classification

ML models (Random Forest, Logistic Regression) test weather's predictive power.



4. Regime Identification

Clustering variables (K-means) to define distinct physical weather states.



5. Regime-Specific Recalculation

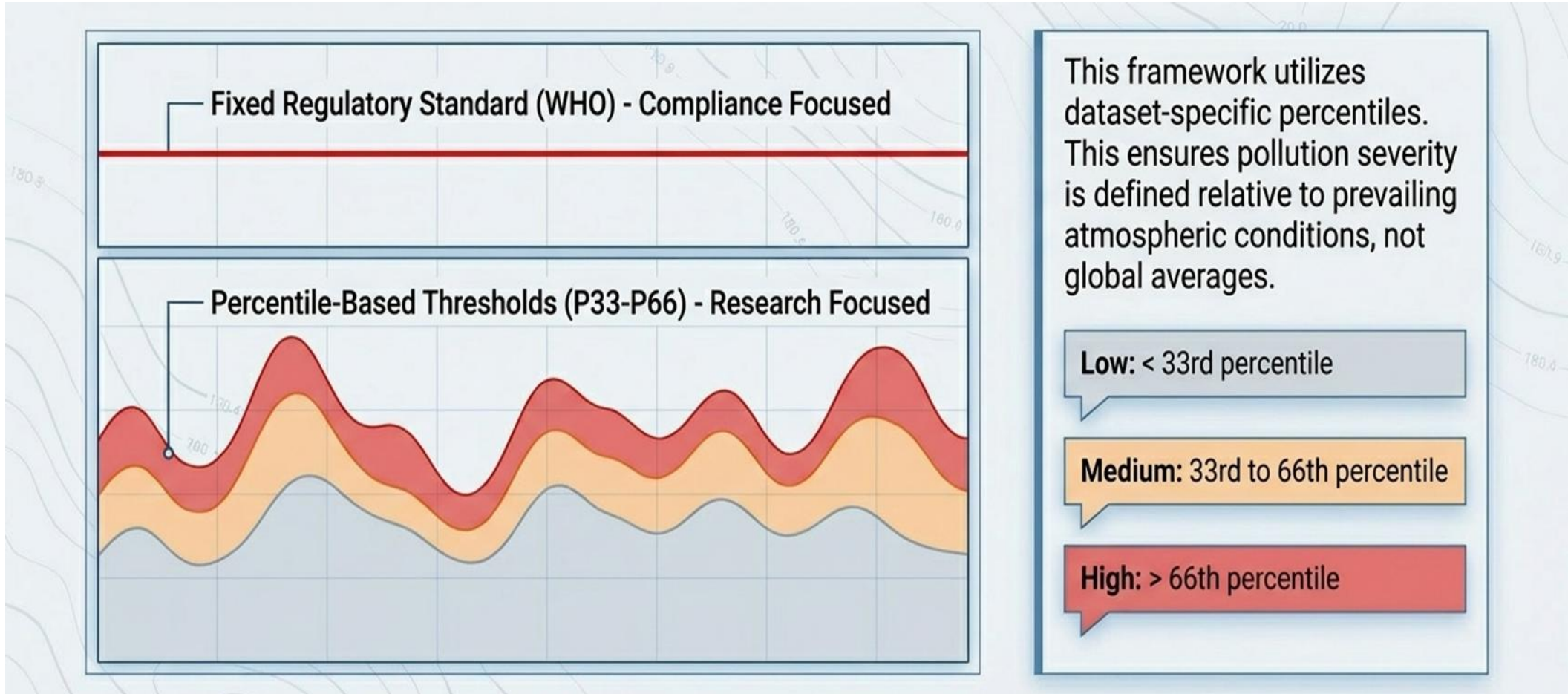
Pollution thresholds recalculated relative to each newly defined weather state.



6. Risk Synthesis

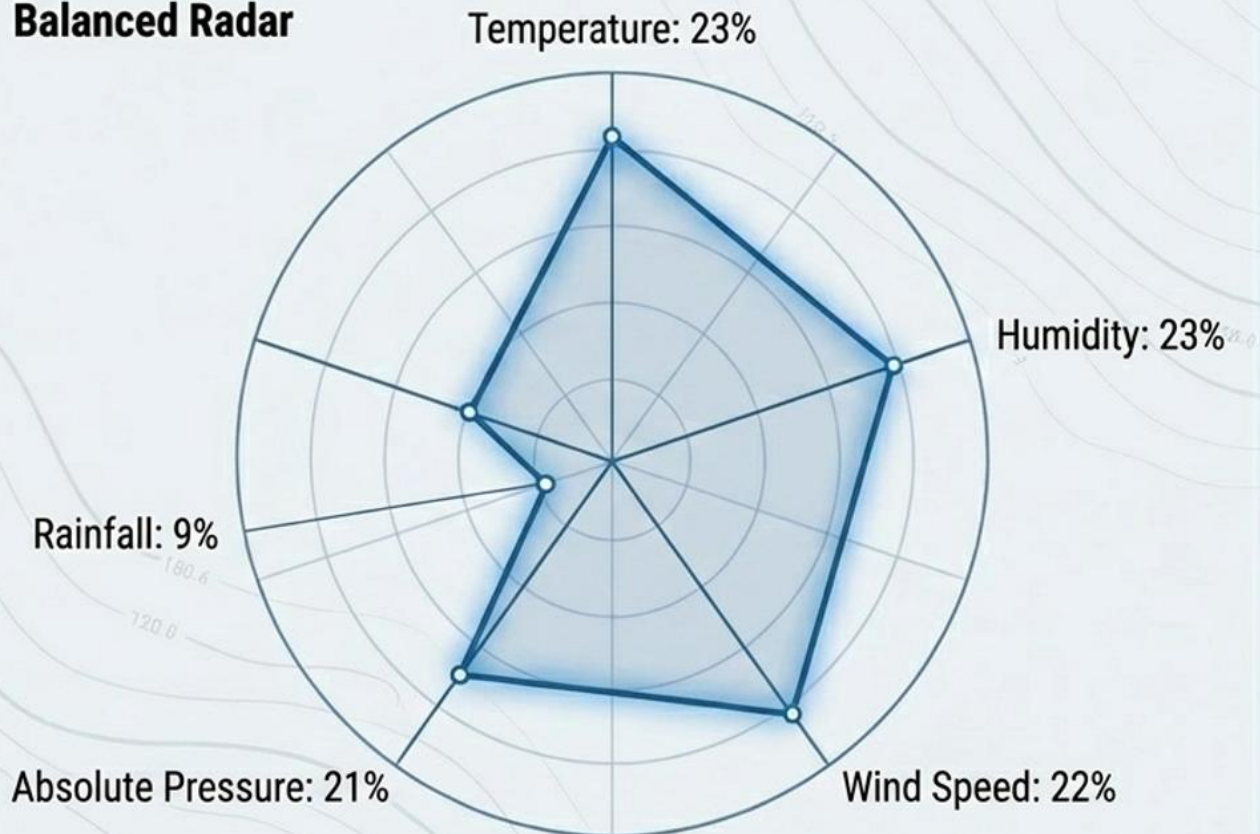
Quantifying burden and pollution persistence (spell length) per regime.

Redefining High Pollution Through Dynamic Thresholds



The distributed drivers of atmospheric pollution

Balanced Radar



Classification Baseline

Random Forest and Logistic Regression achieved ~0.49 accuracy (significantly above the random 3-class assignment baseline).

Feature Exclusion

Rainfall (9%) was excluded from further regime construction due to low impact.

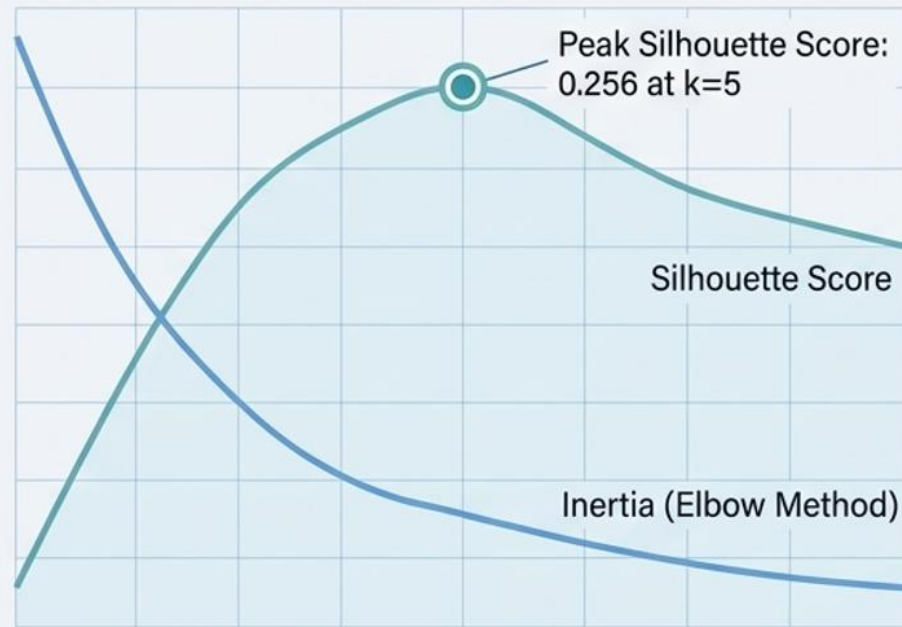
Key Takeaway

No single variable dominates. Pollution states are governed by a complex, multivariate balance of atmospheric conditions.

Clustering the atmosphere into distinct states

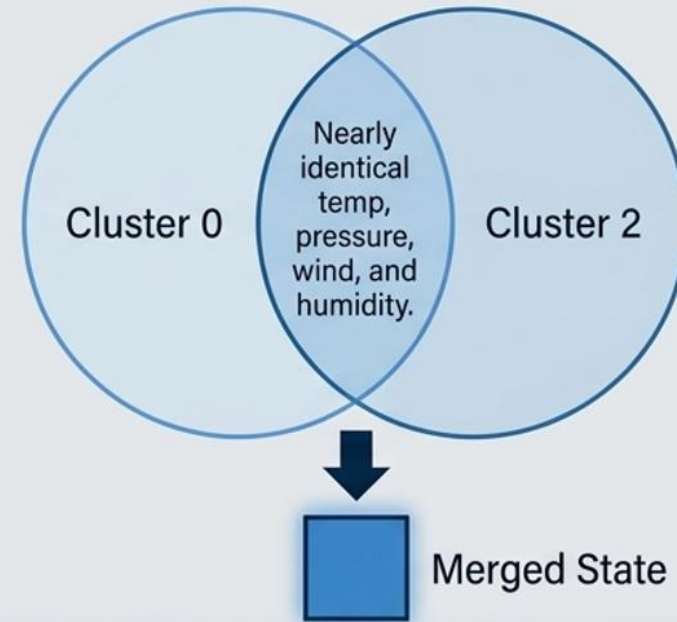
Standardized influential variables (z-score normalization) were processed via K-means clustering to find optimal physical weather states.

Diagnostic Validation



Silhouette scores peaked at k=5, indicating the optimal balance of cohesion and separation.

Regime Consolidation

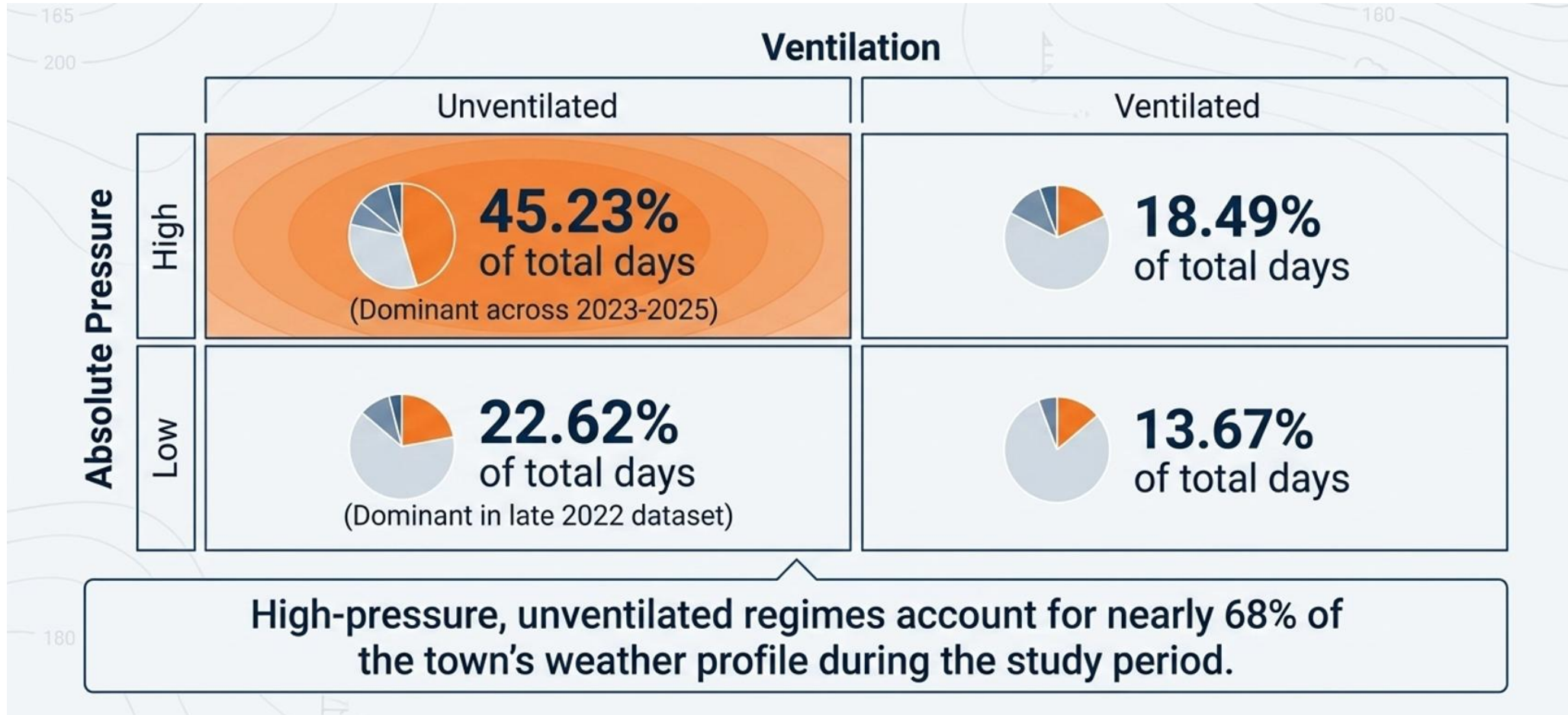


Result: Merging overlapping states resulted in 4 distinct, physically coherent weather regimes.

The Core 4 Weather Regimes

Regime	Temperature (°C)	Humidity (%)	Wind Speed (m/s)	Absolute Pressure (hPa)	Regime Name
1	15.74	78.35	0.95	1004.83	Warm-High Pressure-Unventilated
2	8.61	81.08	2.48	1004.00	Cold-Low Pressure-Ventilated
3	6.60	88.99	1.19	1014.59	Cold-High Pressure-Unventilated
4	8.16	91.24	1.51	988.35	Cold-Low Pressure-Unventilated

Regime Occupancy: Which skies dominate Loughborough



Pollution Burden of Weather Regimes

- To quantify how frequently each regime is associated with polluted conditions, a pollution severity score (S_r) was computed by converting daily pollution states (Safe, Medium and Polluted) into weighted values (Low=0, Medium= 0.5, High = 1) and computing a weighted average based on their proportion within each regime.

The Severity Illusion: Pollution peaks in all weather

One might assume ventilated weather guarantees clean air, but daily pollution classification reveals a **baseline pollution signal** insensitive to meteorology.



Key Takeaway: Short-term pollution presence and severity are broadly comparable across all weather regimes. Local emissions ensure no regime is entirely pollution-free.

Calculating the cumulative Pollution Burden

By weighing how often a regime occurs (**Occupancy**) against its typical pollution levels (**Severity**), we can quantify its long-term exposure relevance.

$$\text{Severity Score } (S_r) \times \text{Occupancy } (\%) = \text{Pollution Burden}$$

Warm-High-Unventilated

Burden = 43.35 (Massive cumulative impact driven by frequency)

Cold-High-Unventilated

Burden = 21.93

Cold-Low-Unventilated

Burden = 17.80

Cold-Low-Ventilated

Burden = 11.60

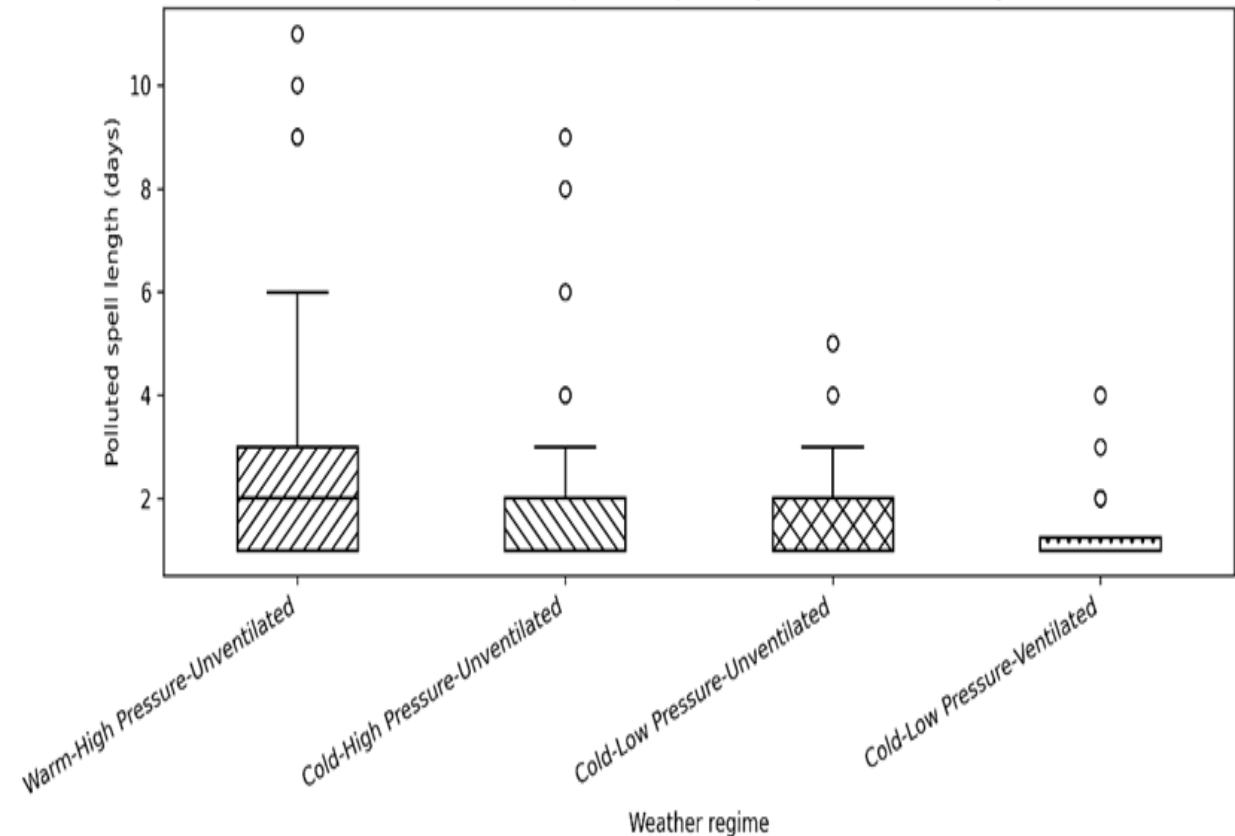
Pollution Persistence and Spell Characteristics

- The polluted spells database was calculated by:

Spell length (L) = number of consecutive polluted days

- The mean polluted spell length was calculated by:

$$\text{Mean Spell Length}(r) = \frac{1}{m} \sum L_j$$



The hidden variable: Atmospheric Persistence

If severity is similar, what differentiates the health risk? Spell length.

Persistence Timeline

Cold-Low-Ventilated
(Mean Spell Length: 1.34 days)



Episodic Spikes (Short, easily dispersed episodes)

Warm-High-Unventilated
(Mean Spell Length: 2.49 days)



Persistent Exposure (Prolonged, unbroken exposure spells)

The Mechanism: Stable, unventilated high-pressure regimes create weak atmospheric mixing and low wind speeds, effectively acting as a lid that traps emissions over the town for days at a time.

The Final Exposure Risk Matrix

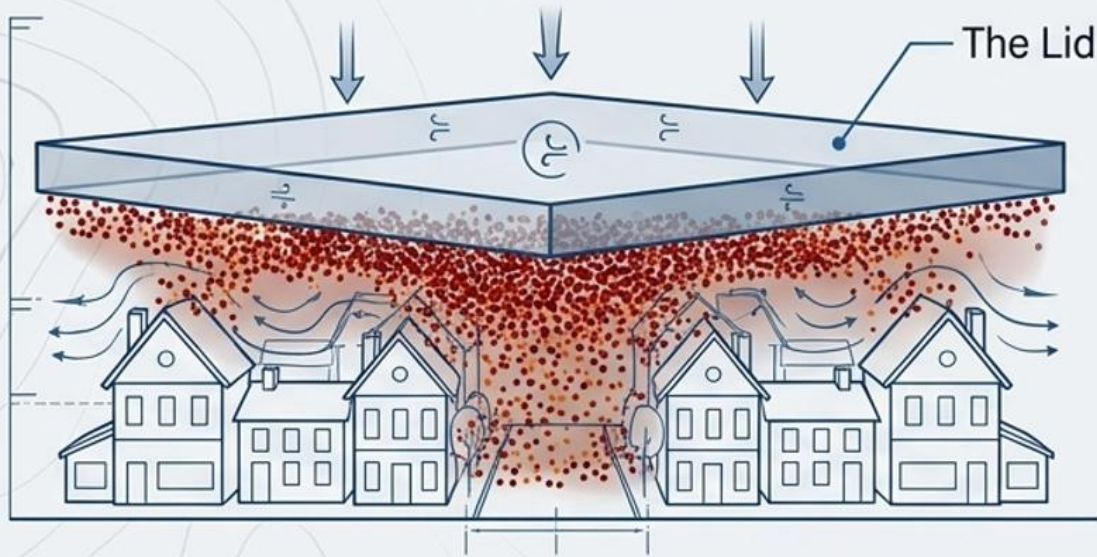
$$\text{Exposure Risk} = \text{Severity} \times \text{Occupancy} \times \text{Mean Spell Length}$$

Absolute Pressure	High	Warm-High-Unventilated High Risk (108.03) The ultimate driver of public health burden.	Cold-Low-Unventilated Moderate Risk (28.03)
	Low	Cold-High-Unventilated High Risk (41.12)	Cold-Low-Ventilated Low Risk (15.60) Frequent atmospheric flushing.
		Unventilated	Ventilated
		Ventilation	

Presence vs. Persistence

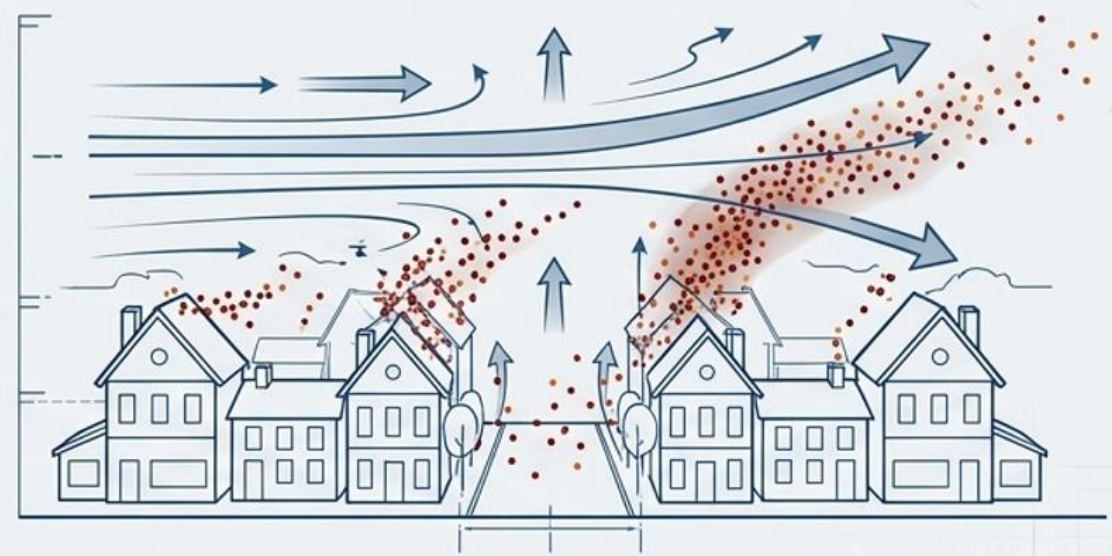
Meteorology regulates how pollution behaves and persists in the atmosphere, rather than fully determining its mere presence.

High Pressure - Unventilated



Because local emissions create a persistent baseline signal, urban air quality strategies cannot simply rely on waiting for "good weather."

Low Pressure - Ventilated



However, public health responses must be calibrated to weather regimes, because high-pressure systems exponentially multiply the duration of exposure.

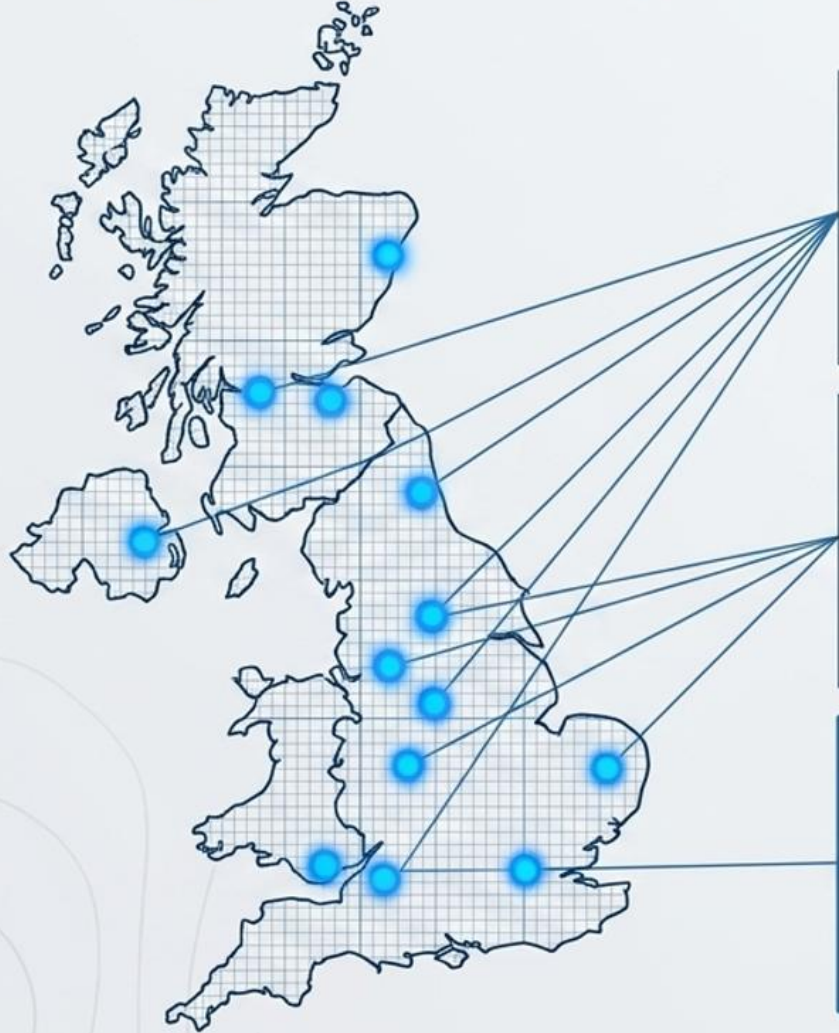
Conclusion

“

Meteorology regulates how pollution behaves and and persists in the atmosphere, rather than fully determining its presence.

The 5-stage weather-dependent framework is fully data-driven. It provides a scalable, repeatable blueprint for small and medium-sized towns worldwide to finally quantify and predict their true exposure risk based on the sky above them.

Scaling the analytical framework



Scalability

The framework relies strictly on commonly available daily meteorological and air quality observations, requiring no bespoke, location-specific parameters.

Adaptability

It can be immediately ported to other small and medium-sized urban settlements to map their unique atmospheric risk quadrants.

Policy Impact

Shifts public health advisories from reactive (measuring high pollution after it happens) to predictive (anticipating unventilated high-pressure regimes that guarantee prolonged exposure spells).

References

1. Sicard, P. *et al.* (2023) “Trends in urban air pollution over the last two decades: A global perspective,” *Science of the Total Environment*. Elsevier B.V. Available at: <https://doi.org/10.1016/j.scitotenv.2022.160064>.
2. Anenberg, S C *et al.* (2022) Long-term trends in urban NO₂ concentrations and associated paediatric asthma incidence: estimates from global datasets, *Articles Lancet Planet Health*. Available at: www.thelancet.com/.
3. Grange, S.K. and Carslaw, D.C. (2019) ‘Using meteorological normalisation to detect interventions in air quality time series’, *Science of The Total Environment*, 653, pp. 578–588. doi: 10.1016/j.scitotenv.2018.10.344.
4. Pope, R.J. *et al.* (2015) “The influence of synoptic weather regimes on UK air quality: Regional model studies of tropospheric column NO₂,” *Atmospheric Chemistry and Physics*, 15(19), pp. 11201–11215. Available at: <https://doi.org/10.5194/acp-15-11201-2015>.
5. Jones, P.D., Harpham, C., and Briffa, K.R. (2012) ‘Lamb weather types derived from reanalysis products’, *International Journal of Climatology*, 33(5), pp. 1129–1139. doi:10.1002/joc.3498.
6. Jones, P.D., Jonsson, T. and Wheeler, D. (1997) ‘Extension to the North Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west iceland’, *International Journal of Climatology*, 17(13), pp. 1433–1450. doi:10.1002/(sici)1097-0088(19971115)17:13<1433:aid-joc203>3.0.co;2-p.
7. Osborn, T.J. (2006), Recent variations in the winter North Atlantic Oscillation. *Weather*, 61: 353-355. <https://doi.org/10.1256/wea.190.06>
8. Jacob, D.J. and Winner, D.A. (2009) “Effect of climate change on air quality,” *Atmospheric Environment*, 43(1), pp. 51–63. Available at: <https://doi.org/10.1016/j.atmosenv.2008.09.051>
9. Sá, E. *et al.* (2015) “Meteorological driven changes on air quality over Portugal: A KZ filter application,” *Atmospheric Pollution Research*, 6(6), pp. 979–989. Available at: <https://doi.org/10.1016/j.apr.2015.05.003>.
10. Fenger, J. (2009) “Air pollution in the last 50 years - From local to global,” *Atmospheric Environment*, 43(1), pp. 13–22. Available at: <https://doi.org/10.1016/j.atmosenv.2008.09.061>.

11. Graham, A.M. *et al.* (2020) "Impact of weather types on UK ambient particulate matter concentrations," *Atmospheric Environment: X*, 5. Available at: <https://doi.org/10.1016/j.aeaoa.2019.100061>.
12. Office for National Statistics (2025) Population estimates for England and Wales: mid-2024. (online) ONS. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationestimatesforenglandandwales/mid2024> (Accessed 13 February 2026).
13. Hodgkins, R. 2026. "A 14-Year Meteorological Dataset From a University Campus in the East Midlands of the UK." *Geoscience Data Journal* 13, no. 1: e70048. <https://doi.org/10.1002/gdj3.70048>.
14. World Health Organization (2021) *WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulphur dioxide and carbon monoxide*. Geneva: World Health Organization. Available at: <https://iris.who.int/handle/10665/345329> (Accessed: 13 February 2026).
15. Donat, M.G. *et al.* (2013) "Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset," *Journal of Geophysical Research Atmospheres*, 118(5), pp. 2098–2118. Available at: <https://doi.org/10.1002/jgrd.50150>.
16. Hosmer, D.W., Lemeshow, S. and Sturdivant, R.X., 2013. *Applied logistic regression*. 3rd ed. New York: Wiley.
17. Breiman, L., 2001. Random forests. *Machine Learning*, 45(1), pp.5–32.
18. Sokolova, M., and Lapalme, G. (2009) 'A systematic analysis of performance measures for classification tasks', *Information Processing & Management*, 45(4), pp. 427–437. doi: 10.1016/j.ipm.2009.03.002.
19. Rousseeuw, P.J. (1987) 'Silhouettes: A graphical aid to the interpretation and validation of cluster analysis', *Journal of Computational and Applied Mathematics*, 20, pp. 53–65. doi:10.1016/0377-0427(87)90125-7.
20. di Bernardino, A. *et al.* (2022) "Classification of synoptic and local-scale wind patterns using k-means clustering in a Tyrrhenian coastal area (Italy)," *Meteorology and Atmospheric Physics*, 134(2). Available at: <https://doi.org/10.1007/s00703-022-00871-z>.
21. Hodgson, E.C. and Phillips, I.D. (2021) "Seasonal variations in the synoptic climatology of air pollution in Birmingham, UK," *Theoretical and Applied Climatology*, 146(3–4), pp. 1421–1439. Available at: <https://doi.org/10.1007/s00704-021-03779-7>.



Any Questions / Suggestions?