ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

INCOMING AIR-MASS PATHWAYS, FIRE ACTIVITY AND $PM_{2.5}$ IN LAHORE, PAKISTAN: ASSOCIATIONS AND WEEKLY FORECASTING

Mian Shaayan Akbar*[1], Sardar Zayaan Akhtar Nakai*[1], Robia Shaheen[2]
[1]*Lahore Grammar School, Senior Boys Campus, Johar Town, Lahore, Pakistan*
Email: mianshaayanakbar@gmail.com, zayaanakai@gmail.com
[2]*Lecturer & PhD Scholar, Institute of Geography, University of the Punjab, Lahore, Pakistan*
Email: robia.res.geog@pu.edu.pk

Abstract
Every autumn and winter, Lahore's air turns thick with fine particulate matter ($PM_{2.5}$), and the explanation depends on who you ask. Some blame smoke from crop fires in Punjab, on both sides of the border. Others blame the city's own traffic and industry, or the cold, still air that lets pollution pile up. This study asks one question with three parts: which incoming air-mass pathways and weather conditions go with high $PM_{2.5}$ in Lahore, how much of that is linked to ventilation and upwind fire activity, and whether this information helps forecast the following week's mean $PM_{2.5}$. We rebuilt almost six years of hourly $PM_{2.5}$ from the U.S. Department of State monitor in Lahore (May 2019 to February 2025; preliminary AirNow data) and paired it with 2,161 daily HYSPLIT back-trajectories, ERA5 meteorology, NASA FIRMS fire detections and simple ventilation features from a one-box model. Three recurring pathways emerged, most often slow air from the south-east (60% of days). Within seasons, $PM_{2.5}$ was up to 30% higher on days with fast north-westerly air, but these differences largely disappeared once weather was taken into account. Poor ventilation was the clearest link: in winter and October to November, a 10% drop in ventilation went with about 5% higher $PM_{2.5}$, whatever the pathway. In October and November, more fire activity along the air's path went with modestly higher $PM_{2.5}$ (about 6% per doubling of exposure), beyond what seasonal timing alone explains. For forecasting, every week was predicted using only information available before it began. Over 138 test weeks, a Random Forest had a mean absolute error of 28.3 $\mu g\, m^{-3}$, beating persistence (35.5) and seasonal climatology (32.6) but not reliably better than a simple model of past $PM_{2.5}$ (29.7). Each added group of inputs (weather, fire, ventilation and trajectories) changed the error by less than 1 $\mu g\, m^{-3}$; only the fire inputs gave a gain narrowly larger than chance, and every model underestimated the worst smog weeks. When the data were assumed to arrive a week late, errors rose by 3 to 5 $\mu g\, m^{-3}$ and the Random Forest was no better than climatology. The extra inputs told the model little that last week's $PM_{2.5}$ and the time of year did not already. When one year of trajectories was recomputed with finer (1°) meteorology, the paths and several pathway results changed, but the fire association did not, so the pathway findings depend on the trajectory data used. Because the study uses one monitor and no chemical data, it describes associations and forecasting skill, not where Lahore's pollution comes from.

Keywords: Lahore, $PM_{2.5}$, air-mass trajectories, HYSPLIT, satellite fire detections, meteorology, Random Forest, air-quality forecasting, walk-forward validation

1. Introduction

1.1 Lahore and $PM_{2.5}$ pollution

$PM_{2.5}$ refers to airborne particles with an aerodynamic diameter of 2.5 micrometres or less. Measurements in Lahore have repeatedly found very high concentrations. Filter samples collected every sixth day from January 2007 to January 2008 gave an annual mean $PM_{2.5}$ of 194 ± 94 µg m^{-3} (Stone et al., 2010). Monitoring between 2005 and 2007 found winter $PM_{2.5}$ loads roughly four times higher than in other seasons (Lodhi et al., 2009), and data from the Pak-

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

istan Environmental Protection Agencies gave a mean of 150.4 ± 87.9 μg m^{-3} for Lahore in autumn 2007 (Rasheed et al., 2015). During the November 2024 haze event, Lahore recorded a daily $PM_{2.5}$ concentration of 621 μg m^{-3}, the highest in the five years examined by Dahal et al. (2026). For comparison, Rasheed et al. (2015) cite the Pakistan National Environmental Quality Standards in force at the time of their study as 25 μg m^{-3} for the annual mean and 40 μg m^{-3} for the 24-hour mean, and a 2021 Punjab EPA bulletin lists a Punjab standard of 35 μg m^{-3} for 24-hour $PM_{2.5}$.

1.2 Local emissions, transport and stagnant conditions

Studies of particle composition in Lahore point to several sources. Chemical mass balance modelling of the 2007 samples showed that fine particles were dominated by carbonaceous material, with organic matter and elemental carbon making up 61 ± 17% of $PM_{2.5}$ mass on average (Stone et al., 2010). Positive Matrix Factorization of winter 2005 to 2006 samples identified soil and road dust, industrial emissions, vehicle emissions and secondary aerosols as the main contributors. The same study suggested that secondary sulfate, possibly from outside the city, raises winter concentrations when the mixing height is low (Lodhi et al., 2009). Electron microscopy and spectroscopy of $PM_{2.5}$ collected during haze episodes in late 2019 identified particle types consistent with vehicle exhaust, crop-residue and other biomass burning, construction, soil dust and industry. That study did not perform quantitative source apportionment (Mohyuddin et al., 2024).

Weather controls how much of this pollution accumulates. Two persistent winter fog and smog events in Lahore in 2016 and 2017 coincided with strong low-level anticyclonic circulation and wind speeds below 1.5 m s^{-1}, conditions that favour temperature inversions (Bilal et al., 2022). Across the Indo-Gangetic Plain, chemical transport modelling found that the change from the post-monsoon season to winter brings a shallower boundary layer and weaker winds, lowering the ventilation coefficient by a factor of about four (Ojha et al., 2020). Across the Indo-Gangetic Plain more widely, days with stagnant air (little rain, weak surface wind and a strong low-level temperature inversion) had $PM_{2.5}$ about 45 μg m^{-3} (45%) higher than other days in winter (Zhou et al., 2024). These studies suggest that high $PM_{2.5}$ in Lahore can arise from local emissions trapped in a shallow, stagnant boundary layer, from pollution transported into the city, or from both at once.

1.3 Incoming air-mass pathways and fire activity

Each October and November, crop residues are burned across the north-western Indo-Gangetic Plain. In north-western India, hotspot districts accounted for about 80% of fire counts in Haryana and about 50% in Punjab, and modelling attributed roughly 50 to 70% of the $PM_{2.5}$ increase caused by crop-residue burning in downwind cities to these hotspots (Saharan et al., 2024). Lahore lies close to this region, and several studies of the city have used backward trajectories to describe where its air comes from. Karim and Rappenglück (2023) used HYSPLIT trajectory clustering at the U.S. diplomatic monitoring site and concluded that local emissions and regional transport from the east most strongly affected the city's air quality. Majeed et al. (2024) combined HYSPLIT trajectories with satellite fire counts and argued that agricultural fires have a strong effect on regional air quality. Tariq et al. (2016) studied an intense haze event in October 2013, when aerosol optical depth over Lahore reached 3.70, and linked it primarily to large-scale crop-residue burning.

Trajectory and fire evidence of this kind shows that air arriving in Lahore may have passed over burning areas. It does not by itself show how much fire smoke reached the monitor. Studies of Delhi, where more quantitative work has been done, show how uncertain this step is. One

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

modelling study attributed between 7% and 78% of maximum observed $PM_{2.5}$ enhancements in Delhi to fires, depending on the year and the emission inventory used (Cusworth et al., 2018). Another estimated that fire emissions contributed about 20% of post-monsoon $PM_{2.5}$ on average, rising to 50 to 75% during pollution events (Kulkarni et al., 2020). No comparable quantitative estimate for Lahore was found in the literature reviewed for this study.

1.4 Forecasting motivation

Advance warning of high-pollution periods could help schools, hospitals and city authorities plan. Existing forecasting studies that include Lahore have mostly predicted daily $PM_{2.5}$ one day ahead using only the past $PM_{2.5}$ series (Iftikhar et al., 2024; Qureshi et al., 2025) or have used time-series methods such as SARIMA (Bhatti et al., 2021). A recent study used deep learning with meteorological inputs (Rashid & Atif, 2026). This project instead targets the mean $PM_{2.5}$ of the following week. A weekly horizon is long enough to be useful for planning, and it smooths out some of the day-to-day noise in a single-monitor record. It also changes which predictors can be known when the forecast is issued, a point that is central to the design (Section 5.6).

1.5 Research questions and study aims

The primary research question is:

Which incoming air-mass pathways and meteorological conditions are associated with elevated $PM_{2.5}$ in Lahore, and do HYSPLIT trajectory clusters, upwind fire-exposure indicators and one-box mass-balance features improve forecasts of the following week's mean $PM_{2.5}$ beyond historical-pollution and local-meteorological baselines?

The study has three connected investigations:

1. **Pathway analysis.** Identify recurring incoming air-mass pathways using HYSPLIT backward trajectories and clustering, and compare their seasonal frequency and associated $PM_{2.5}$ levels, and map the regions that air passed over before Lahore's most polluted days.
2. **Physical interpretation.** Test whether differences between pathways are associated with local meteorology, ventilation, boundary-layer height, rainfall and satellite-detected fires along trajectories, and try to distinguish possible transport from local accumulation.
3. **Forecasting.** Develop a Random Forest regression model and test whether trajectory, fire-exposure and physically motivated features improve out-of-sample forecasts of next-week mean $PM_{2.5}$ beyond simpler baselines.

Supporting questions are listed with the hypotheses in Section 3. None of the questions is assumed to have a positive answer. The study is limited to associations and predictive skill. With one monitor, no emission inventory, no measured inflow concentrations and no chemical speciation for 2019 to 2025, it cannot separate the contributions of local emissions, regional transport, fire smoke and secondary aerosol formation to the $PM_{2.5}$ measured in Lahore.

2. Literature Review and Research Gap

Sources were found through keyword searches of scholarly databases and publisher websites, and through the reference lists of key papers. Every work cited was checked against its DOI record. Appendix A records, for each work, whether the full text, the abstract only or only the bibliographic record was reviewed. Peer-reviewed articles are distinguished there from conference papers, preprints, reports and official documents. Where a quantitative finding is taken only from an abstract, Appendix A says so, and such findings must be checked against the full text before submission.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

2.1 Lahore and Punjab pollution sources

Source studies in Lahore span more than fifteen years but are short and based on samples. Lodhi et al. (2009) collected 310 $PM_{2.5}$ samples between November 2005 and December 2007 and applied Positive Matrix Factorization to the winter subset. Stone et al. (2010) used organic tracers and chemical mass balance on a year of samples collected every sixth day. Mohyuddin et al. (2024) characterised 45 $PM_{2.5}$ samples collected on alternate days between October and December 2019 using SEM-EDX, X-ray diffraction and Raman spectroscopy. They state that their study did not examine transport mechanisms and that long-term monitoring with source apportionment is still needed. Together these studies establish that combustion, dust, industry and secondary aerosol all contribute. Because they used different methods in different years, they cannot provide source shares that apply to any particular week of the 2019 to 2025 period studied here.

2.2 Regional transport and crop burning

The influence of crop-residue burning on air quality has been studied most intensively for Delhi. Using satellite data, ground measurements and back trajectories, Jethva et al. (2018) found that more than 80% of near-surface north-westerly trajectory clusters reaching New Delhi had crossed the burning region during the post-monsoon season. Liu et al. (2018) defined daily upwind "airsheds" from HYSPLIT trajectories and summed satellite fire radiative power (FRP) within them. They found that, in the post-monsoon season, each additional 1,000 MW of FRP in Delhi's airshed was associated with an increase of 16.34 $\mu g\, m^{-3}$ in PM_{10}. Their method used 72-hour HYSPLIT back trajectories driven by GDAS 0.5° meteorology (with a 500 m arrival height for post-monsoon Delhi), a 10 km buffer around each trajectory, and daily MODIS Collection 6 FRP summed within that buffer. It is the closest methodological precedent found for the trajectory-weighted fire-exposure indicator proposed here. Chemical transport models give a range of estimates (Kulkarni et al., 2020; Saharan et al., 2024), and a Lagrangian modelling study found that the attribution range depends strongly on the fire emission inventory used (Cusworth et al., 2018). Cusworth et al. also noted that thick smoke may interfere with satellite fire detection, and their model reproduced up to 25% of the weekly variability in observed $PM_{2.5}$ in Delhi, a useful reference point because the present study also works at weekly resolution.

For Lahore, the evidence is mainly descriptive. Rasheed et al. (2015) reported that high-pollution trajectories reaching Pakistani cities in 2007 to 2011 (2007 to 2008 for Lahore) commonly originated over western India, especially in summer, or over eastern Afghanistan. Using MODIS imagery and prevailing winds for 2007 to 2011, Khanum et al. (2017) described transport from India towards Pakistan from February to October and from Pakistan towards India from November to January. Back trajectories used to interpret satellite absorbing-aerosol data over Lahore came 36% from the north, 33% from the east and 21% from western India, although these shares describe column aerosol rather than surface $PM_{2.5}$ and are not split by season (Nawaz et al., 2023). A summertime concentration-weighted trajectory analysis found local sources dominant, with an additional contribution from western India (Ahmad et al., 2020). An earlier source-apportionment study attributed about 15% of $PM_{2.5}$ to biomass burning and warned that potential source contribution function analysis is unsuitable when high concentrations are dominated by local sources and local meteorology (Raja et al., 2010). Delhi studies place Lahore at the upwind end of the main winter pathway: a concentration-weighted trajectory study for Delhi described a winter cluster that begins north-west of the city over industrial Lahore (Ghosh et al., 2014). This suggests that north-westerly flow in the burning

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

season tends to carry Indian Punjab smoke away from Lahore, so that transport towards Lahore should occur mainly on days with easterly or south-easterly winds. Majeed et al. (2024) used an Analytical Hierarchy Process alongside HYSPLIT and satellite data and assigned transboundary agricultural fires moderate importance relative to vehicle emissions. Dahal et al. (2026) studied the November 2024 haze event in both Delhi and Lahore. They found $PM_{2.5}$ correlated with boundary-layer height, lower-tropospheric stability, relative humidity and FRP, with back trajectories overlapping the stubble-burning region. They also reported that peak fire activity had shifted from afternoon to evening, as seen by a geostationary satellite. This matters for the present study because polar-orbiting fire products may under-sample evening burning. Gautam et al. (2023) argued that strengthening lower-tropospheric stability over northern India from 1980 to 2019 has contributed to more poor-visibility days. Together these studies support transport and stagnation as competing hypotheses for Lahore rather than settled explanations.

2.3 HYSPLIT trajectory clustering

HYSPLIT is NOAA's Lagrangian trajectory and dispersion model (Stein et al., 2015). A backward trajectory estimates the path an air parcel followed before reaching a receptor, calculated from gridded wind fields. Reviews by Fleming et al. (2012) and Stohl (1998) describe how trajectories are calculated, how accurate they are, and why the path the air has taken says more about its pollution than the local wind direction alone. In general, uncertainty in a trajectory grows with travel time and depends on the resolution of the meteorological data, so a 72-hour trajectory should be read as a probable pathway, not an exact route.

Clustering groups many trajectories into a few representative pathways. Cui et al. (2021) compared K-means, hierarchical clustering and self-organising maps for 72-hour trajectories arriving at Qingdao, China. They found that height information added little, that K-means and self-organising maps scored better than hierarchical clustering and the standard HYSPLIT clustering on internal indices, and that the Davies-Bouldin index was useful for choosing the number of clusters. Citing earlier work, they also note that K-means was the method least affected by the choice of arrival height. Because this comparison was made for a single coastal city in China, it does not show which method is best for Lahore. Closer to Lahore, Ghosh et al. (2014) grouped back trajectories arriving in Delhi (72-hour paths at 500 m and 48-hour paths at 100 m) and weighted them by the $PM_{2.5}$ measured on arrival, a method known as concentration-weighted trajectory analysis. They reported that regional sources, including areas in Pakistan and the Indian states of Punjab, Haryana and Uttar Pradesh, contributed to high $PM_{2.5}$ in Delhi. Karim and Rappenglück (2023) ran 72-hour HYSPLIT back trajectories from the Lahore diplomatic monitor at 200, 500, 1,000 and 4,000 m above ground. They used GDAS meteorology, which they describe as 0.5° and 3-hourly, combined with cluster, residence-time and concentration-weighted trajectory analyses. Their study establishes that this monitor has already been used as a HYSPLIT receptor. Its aim, however, was to compare lockdown and reference periods, not to build forecast features.

2.4 Boundary-layer height and ventilation

The planetary boundary layer is the lowest part of the atmosphere, where surface emissions are mixed. A shallower layer means emissions are diluted into a smaller volume. Using lidar-derived boundary-layer height at about 1,600 Chinese stations, Su et al. (2018) found a generally negative correlation between particulate matter and boundary-layer height that varied with location and season. Miao et al. (2019) review the two-way interaction between the boundary layer and $PM_{2.5}$ in Chinese megacities. In the Indo-Gangetic Plain, Ojha et al. (2020) linked the winter enhancement of $PM_{2.5}$ to a ventilation coefficient about four times lower than in the

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

post-monsoon season. In their simulation for October to December 2016, ventilation coefficients of about 600 to 1,000 $m^2\ s^{-1}$ over the central plain in October fell below 400 $m^2\ s^{-1}$ over most of the plain by December. Guttikunda and Gurjar (2011) used a Lagrangian tracer model to show the role of meteorology in Delhi's seasonal pollution cycle. ERA5, the reanalysis used in this study, provides hourly fields, including boundary-layer height, at about 31 km horizontal resolution (Hersbach et al., 2020). Reanalysis boundary-layer heights are model estimates, not observations, and their accuracy over Lahore has not yet been checked in this project.

2.5 Machine-learning $PM_{2.5}$ forecasting

Random Forests, ensembles of decision trees trained on bootstrap samples (Breiman, 2001), are among the most widely used machine-learning methods for air quality because they capture non-linear relationships with little tuning (Rybarczyk & Zalakeviciute, 2018; Méndez et al., 2023; see also Zhang et al., 2012). Tree-based models have been used with meteorology and boundary-layer height to explain daily PM_{10} in Switzerland (Grange et al., 2018) and, together with SHAP, to describe how meteorology drives PM_1 (Stirnberg et al., 2021).

Few studies have used trajectory information as a model input. Le et al. (2021) included HYSPLIT cluster labels as predictors in a Random Forest for Hanoi, but to adjust concentrations for weather rather than to forecast. Feng et al. (2015) added a trajectory-based indicator to a neural network forecasting $PM_{2.5}$ two days ahead in the Beijing-Tianjin-Hebei region and reported up to 40% lower RMSE. Because wavelet decomposition was added at the same time, the share of that improvement due to trajectories is unknown.

Studies that include Lahore are summarised in Table 1. They differ from the present design in forecast horizon, inputs and data source, so their error values cannot be compared directly with this study's.

Time-series forecasts must be evaluated on data that come after the training period (Tashman, 2000; Bergmeir & Benítez, 2012), using error measures chosen with care (Hyndman & Koehler, 2006). A particular risk is leakage: the use of information that would not have been available when the forecast was made, which can make a model look far more skilful than it is (Kaufman et al., 2012). A recent systematic review of air-quality forecasting found that no single model is best once evaluation is leakage-safe and rolling (Chianese & Riccio, 2026).

SHAP (SHapley Additive exPlanations) assigns each prediction's deviation from the average prediction to individual input features (S. Lundberg & Lee, 2017). An efficient exact version exists for tree ensembles (S. M. Lundberg et al., 2020). SHAP values describe how a fitted model uses its inputs. They do not show that a feature causes pollution, and they can divide credit unpredictably between correlated features.

2.6 Physical and hybrid models

Full chemical transport models such as WRF-Chem simulate emissions, transport, chemistry and deposition. They underpin the Delhi fire-attribution estimates above (Kulkarni et al., 2020; Saharan et al., 2024), but they need emission inventories and computing resources beyond this project. At the other extreme, a one-box model treats a city's air as a single well-mixed volume whose concentration changes through emission, inflow, outflow and removal (Jacob, 1999).

Hybrid approaches use physical knowledge to guide data-driven models. Examples include theory-guided data science (Karpatne et al., 2017), the combination of process understanding with machine learning in Earth system science (Reichstein et al., 2019), and the work of Di et al. (2016), who gave a neural network the $PM_{2.5}$ values simulated by a chemical transport model, together with satellite and land-use data, as inputs for estimating daily $PM_{2.5}$ across the United States. In this project the box model is used only to derive a small number of physically

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Table 1

Forecasting studies that include Lahore, as verified for this study.

Study	Data and period	Horizon and inputs	Notes relevant to this study
Bhatti et al. (2021)	Federal and provincial EPA monitoring data for Lahore, 2014 to 2019	SARIMA with factor analysis; HYSPLIT back trajectories (GFS meteorology), December 2018 to December 2019	Full text checked in part. Reports trajectories originating over Afghanistan; validation design not checked.
Iftikhar et al. (2024)	Daily $PM_{2.5}$ from U.S. embassy monitors, four cities; the start date is 1 January 2019 in the abstract and 1 June 2019 in the methods, to 31 May 2023	Short-term; linear and non-linear time-series models in weighted ensembles	The abstract reports "mean errors" of 0.81% to 13.52% for Lahore, but these cannot be matched to the paper's own tables, which give a MAPE of 21.67% for Lahore with the same model.
Qureshi et al. (2025)	Daily $PM_{2.5}$, 2019 to 2023, four cities; data availability statement cites IQAir	One day ahead; STL decomposition with ensembles; $PM_{2.5}$ history only, no meteorology	Missing values imputed with MICE. Best Lahore RMSE reported as 12.475. A statement that imputation and decomposition were fitted only on training data was not found.
Rashid and Atif (2026)	Lahore (details not checked)	Deep learning with meteorological inputs	DOI registered (Atmospheric Environment 381, 122110); content not yet checked.
Haseeb et al. (2026)	Daily $PM_{2.5}$ in Lahore from satellite, reanalysis and ground data; trained and tested on 2019 to 2023	Daily prediction; ensemble machine learning including Random Forest	Chronological split with an independent 2024 validation year (Random Forest R^2 0.77, RMSE 24.75 µg m^{-3}); hyperparameters tuned with 10-fold cross-validation.

motivated features. It does not estimate source contributions.

2.7 Synthesis of the research gap

The literature reviewed shows that Lahore's $PM_{2.5}$ has been characterised chemically, that HYSPLIT trajectories and satellite fire data have been used to describe transport into the city, and that forecasting studies including Lahore exist. In the studies found, trajectory and fire information for Lahore was used descriptively or diagnostically, and the forecasting studies used $PM_{2.5}$ history, sometimes with meteorology. This search did not find a study that tests whether Lahore-specific trajectory clusters, trajectory-weighted fire exposure or box-model features add predictive skill beyond historical-pollution and meteorological baselines under leakage-free walk-forward evaluation.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

The proposed contribution is therefore a transferable evaluation framework, applied to Lahore: a test of whether air-mass history and fire-exposure information add forecast skill beyond strong persistence, climatology, historical-$PM_{2.5}$ and weather-only baselines, under leakage-free walk-forward evaluation with realistic data latency, and with forecasting kept separate from retrospective diagnostic analysis. A null result, in which these features add no skill, would also be informative. This is a proposed contribution, not a claim that the combination is unprecedented. The search covered English-language records indexed by OpenAlex and publisher sites up to 18 October 2026 and may have missed relevant work, particularly theses, reports and studies in Urdu.

3. Research Questions and Hypotheses

The hypotheses below are stated so that each could be shown false. H1 to H3 address the first two parts of the research question (pathways and their physical interpretation), and H4 and H5 address the third (forecasting).

Table 2

Hypotheses. H1 to H3 concern associations and cannot establish causation.

Question	Hypothesis	What would count against it
Which trajectory clusters are associated with the highest $PM_{2.5}$?	H1: Daily $PM_{2.5}$ distributions differ between dominant trajectory clusters within the same season.	No difference in $PM_{2.5}$ between clusters once season is accounted for, or differences that disappear under alternative arrival heights or numbers of clusters.
How do pathway and pollution relationships vary with season and ventilation?	H2: Within a season, high-$PM_{2.5}$ days are associated with lower ventilation (wind speed multiplied by boundary-layer height), regardless of pathway.	No association with ventilation, or associations explained entirely by pathway membership.
Are high-pollution days associated with more fire exposure along incoming pathways?	H3: In October and November, days with higher trajectory-weighted fire exposure have higher $PM_{2.5}$ after accounting for ventilation and seasonal timing, and exposure computed from the actual fire record outperforms exposure computed from time-shifted fire records.	No association after accounting for ventilation and season, or an association no stronger than that obtained with time-shifted (null) fire records.
Do these features improve out-of-sample forecasts?	H4: Adding trajectory and fire-exposure features lowers walk-forward MAE relative to the model without them on the same forecast weeks (tested with no data embargo and with a one-week embargo; Section 5.6).	No reduction in MAE, or a reduction whose block-bootstrap interval includes zero or that does not survive the pre-specified sensitivity analyses.
Which features influence the predictions, and are results robust?	H5: SHAP rankings and skill differences are stable across the sensitivity analyses in Section 5.11.	Rankings or conclusions that change with reasonable alternative choices.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

4. Study Area and Data

4.1 Monitoring site

The study uses the single $PM_{2.5}$ monitor operated by the U.S. Department of State in Lahore. Its identifiers and metadata, as verified in the project's Stage 1 reports, are shown in Table 3.

Table 3
Monitor metadata.

Item	Value	Evidence
AirNow AQS identifier	586PK3010003, "U.S. Department of State Pakistan - Lahore"	AirNow site file, 15 September 2021
OpenAQ location and sensor	Location 8664, "US Diplomatic Post: Lahore"; sensor 25135	OpenAQ Explorer and archive files
Coordinates	31.560078° N, 74.33589° E	Identical in AirNow site file and every OpenAQ row
Instrument	BAM 1020 beta attenuation instrument (Met One)	Reported by Karim and Rappenglück (2023); not stated in AirNow or OpenAQ metadata
Parameter and units	$PM_{2.5}$, µg m^{-3}, 1-hour values	Archive files
Time zone	Pakistan Standard Time (PKT, UTC+05:00, no daylight saving)	AirNow GMT offset +5.00; OpenAQ timestamps carry +05:00
Record	AirNow daily values 22 May 2019 to 17 February 2025; last valid OpenAQ hour 18 February 2025, 18:00 PKT	Stage 1C report

OpenAQ location 8664 is treated as the same monitor as AirNow 586PK3010003, based on identical coordinates, matching start dates, 99.7% agreement between 2,553 overlapping hourly pairs (allowing for one-hour label shifts) and daily-mean correlations of at least 0.997. A second OpenAQ identifier, location 6506, has also been associated with this post. It could not be resolved: its Explorer page does not exist, the archive holds no files for it, and checking it would need a registered OpenAQ API key. Any link between 6506 and 8664 is an assumption until checked.

A single monitor cannot represent the whole city. Its readings reflect local sources near the site as well as citywide and regional conditions, and how representative it is of Lahore as a whole has not been established.

4.2 $PM_{2.5}$ data and quality control

Sources examined. The official State Department archive (dosairnowdata.org) no longer resolves, Punjab EPA publishes only daily PDF bulletins, and IQAir offers no documented multi-year archive. Two machine-readable copies of the monitor's record were found: the OpenAQ public archive and the AirNow file archive.

OpenAQ copy. All 1,909 daily OpenAQ files (23 May 2019 to 4 March 2025) were downloaded and checked for completeness; 232 of them contain only missing values.

Problems found by comparison with the original source. Hourly OpenAQ values were compared with original AirNow hourly files for 18 eight-day sample windows, and three problems were found:

- A constant value of 67.0 µg m^{-3} from 31 May to 5 June 2024 (a 7-hour and a 120-hour run) does not appear in AirNow, which has no Lahore record for those hours. These

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

values are not measurements.

- OpenAQ timestamps were shifted by one hour in some periods and inconsistently aligned in others, so that some values are skipped and others duplicated.
- Despite this, complete-day OpenAQ means agree closely with AirNow's own daily values (r of at least 0.997 in every year), so the problems matter mainly for hourly analysis.

Timestamp conventions. OpenAQ labels each hour by its end and AirNow hourly files label it by its start, in GMT. Both are converted to hour-ending Pakistan Standard Time and assigned to local days.

Reconstruction of the original record. Because of these findings, the hourly record was rebuilt directly from the AirNow archive, keeping only the Lahore lines of each size-checked hourly file. The rebuilt record covers 22 May 2019 to 18 February 2025 (Section 6.1). AirNow describes these data as preliminary and not fully validated, and asks that analyses say so and do not use the data to establish trends.

Causes of the main gaps. The AirNow Data Management Center explained the three longest gaps (M. Hylton, AirNow Data Management Center, personal communication, 23 October 2026). From 21 October to 7 December 2021, AirNow received no data from the monitor, for reasons it does not know. Data for 6 September to 1 October 2022 had been sent to AirNow late, so they were missing from the archived hourly files; AirNow regenerated that month's file, and its values were added to the record for hours with no archive value (Section 5.1). From 26 May to 29 June 2024, some hours were never received and the rest were withheld after AirNow's quality checks because the values did not change from hour to hour, probably because of a monitor fault. This is consistent with the constant block of 67.0 $\mu g\,m^{-3}$ found in the OpenAQ copy, and it supports the repeated-value rule used here.

Quality-control rules. A daily mean is computed only for days with at least 18 valid hourly values (75% of 24), and no values are imputed. Runs of 7 or more identical hourly values are treated as invalid (rule S7); on the rebuilt record this removed 9 hours. A week is eligible if it has at least 6 valid days. Coverage statistics for the OpenAQ copy are given in Appendix G.

4.3 ERA5 meteorology

Downloaded. ERA5 reanalysis (Hersbach et al., 2020) was obtained from the Copernicus Climate Data Store as an hourly point time series for the grid point nearest the monitor (31.50° N, 74.25° E; Copernicus Climate Change Service, 2025). The record has 19 variables and 51,864 hours with no gaps. Times are in UTC and were converted to PKT before daily aggregation, so that meteorological days match $PM_{2.5}$ days; hourly accumulated variables, such as precipitation and radiation, were assigned to the PKT day of the hour they cover. Units were converted to °C, hPa, mm and $W\,m^{-2}$, wind speed was derived from its components at 10 m and 100 m, and relative humidity from temperature and dewpoint. For each forecast week, 21 features were computed from the seven days of the previous week: mean, mean daily minimum and mean daily maximum 2 m temperature; mean dewpoint and relative humidity; mean eastward and northward wind, mean wind speed at 10 m and 100 m, maximum gust and the number of calm hours (wind below 1 $m\,s^{-1}$); mean, mean daily minimum and mean daily maximum boundary-layer height; mean surface and sea-level pressure; mean cloud cover; total rainfall and the number of days with at least 1 mm; and mean downward solar radiation and surface sensible heat flux. ERA5 is published several days after real time, so its use for the previous week stands in for operational weather analyses (Section 5.6).

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

4.4 HYSPLIT trajectories

Computed. One 72-hour backward trajectory was computed for each day from 1 April 2019 to 28 February 2025 (2,161 trajectories; no day failed) with NOAA's HYSPLIT model, version 5.4.2 (Stein et al., 2015). Each trajectory arrives at the monitor's coordinates (31.560078° N, 74.33589° E) at 06:00 UTC (11:00 PKT), 500 m above ground level, and its position is recorded every hour, giving 73 points. Vertical motion was taken from the meteorological data, and all other settings were HYSPLIT defaults. The driving meteorology is the NCEP/NCAR Reanalysis (Kalnay et al., 1996) on a 2.5° global grid every 6 hours; all 72 monthly files were checked against NOAA's published checksums. Each trajectory is assigned to the PKT calendar day on which it arrives. Trajectories arriving at 100 and 250 m were later computed with the same meteorology and settings (4,322 trajectories, none failed) for a declared sensitivity analysis. For a declared one-year comparison, the trajectories arriving from 1 March 2024 to 28 February 2025 were also recomputed with NOAA's GDAS 1° meteorology (Section 6.10); the main analysis uses only the 2.5° trajectories. Other arrival times and durations were not run. A 2.5° grid cell is roughly 240 to 280 km across, which is coarse near the Himalaya and the Sulaiman Range and for a 500 m arrival height, so the trajectories describe synoptic-scale transport. They show the probable recent path of air arriving at the monitor. They do not identify the origin of the pollution measured.

4.5 NASA FIRMS detections

Downloaded. Active-fire detections from the VIIRS 375 m product (Schroeder et al., 2014) on the Suomi NPP satellite, in standard processing, were downloaded from NASA FIRMS for 69 to 78° E and 27 to 35° N. Only detections classed as presumed vegetation fires with nominal or high confidence were kept (740,436 detections); low-confidence detections, static land sources and offshore detections were removed. Each detection was assigned to its PKT day and located by its great-circle distance and bearing from the monitor. The same sensor is used for the whole record: NOAA-20 VIIRS was used only as a cross-check (weekly correlation 0.97 to 0.99), and MODIS (Giglio et al., 2016) and NOAA-21 detections are not used. For each forecast week, 10 features were computed from the previous week: detection counts and total fire radiative power (FRP) in three distance rings (0 to 100, 100 to 300 and 300 to 500 km) and in two direction sectors within 500 km (east/south-east, bearing 45° to 180°, and west/south, 180° to 315°; Figure 1). The download box cuts off the 300 to 500 km ring to the north and east. A detection shows a thermal anomaly at the time of a satellite overpass. It does not identify the burning material (crop residue, other biomass or industrial heat sources), it can miss fires obscured by smoke or cloud, and it can miss burning between overpasses.

4.6 Variable definitions and dataset alignment

Weeks run from Monday to Sunday in PKT, matching ISO-8601 weeks. The target for week T is the mean of its valid daily means, and a week is eligible only if at least 6 of its 7 days are valid. Each week is assigned to a season by its Thursday, as in ISO-8601 week numbering. Forecasting results are reported for October-to-January weeks (the smog season) and for the other months. A four-season scheme (winter, December to February; pre-monsoon, March to May; monsoon, June to September; post-monsoon, October and November) is used for the association analysis, which compares days within seasons (Section 5.14).

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

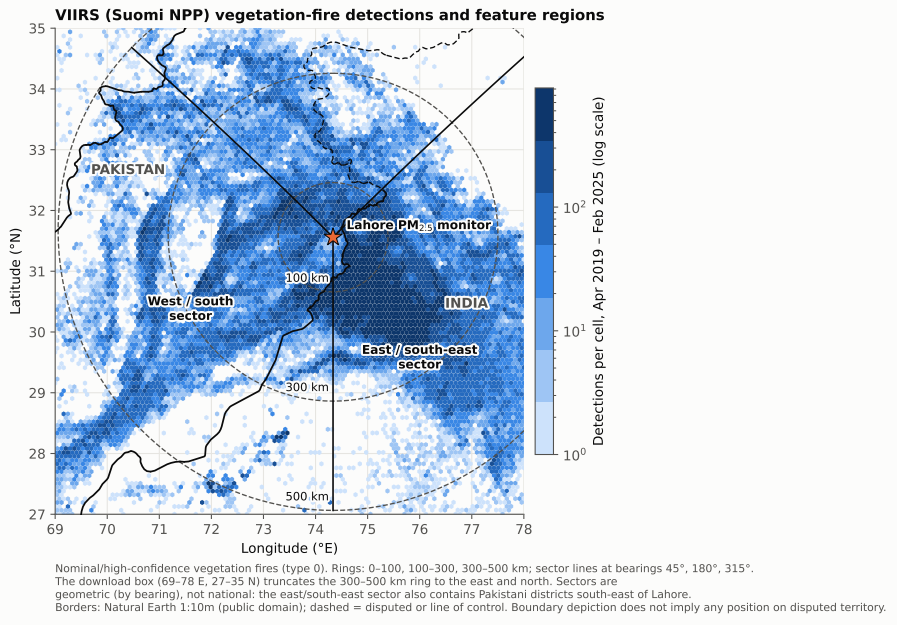


Figure 1

Density of VIIRS (Suomi NPP) vegetation-fire detections from April 2019 to February 2025 around the Lahore $PM_{2.5}$ monitor (star). Dashed circles mark 100, 300 and 500 km; solid lines separate the east/south-east sector (bearing 45° to 180°) from the west/south sector (180° to 315°) used for the fire features. Sectors are defined by direction, not country. Borders: Natural Earth (public domain); dashed borders are disputed boundaries or the line of control, and their depiction implies no position on disputed territory.

Table 4

Time bases and alignment.

Dataset	Native time base	Native averaging	Alignment rule
AirNow hourly $PM_{2.5}$	GMT, hour-beginning	1 hour	Convert to hour-ending PKT; assign to PKT days
OpenAQ $PM_{2.5}$	PKT (+05:00), documented as hour-ending	1 hour	Used only for comparison
AirNow daily $PM_{2.5}$	Local standard time, midnight to midnight	24 hours	Used only to check the rebuilt daily means
ERA5	UTC	Hourly (instantaneous or accumulated by variable)	Convert to PKT, then aggregate to PKT days and weeks
HYSPLIT	UTC start times	Hourly trajectory points	Label each trajectory by its PKT arrival day
FIRMS	UTC acquisition time	Instantaneous overpass	Assigned to PKT days; summed by distance ring and direction sector for the previous week

5. Methodology

5.1 Data quality control

Quality control is applied in code to copies of the data; raw files are never edited. On the rebuilt AirNow hourly record, negative values are treated as invalid, and every hour of a run of 7 or

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

more identical consecutive valid values is invalid. This removed 9 hours, all in one run on 3 May 2020. The run rule is applied before the daily completeness rule. The rebuilt daily means were checked against AirNow's own daily values (Section 6.1). Data for September 2022 that AirNow received late, and that are therefore missing from the archived hourly files, were taken from a file AirNow regenerated for that month. They were used only for hours with no archive value; on the hours present in both, all values were identical, which also confirmed that the file's local times are hour-ending PKT. This change was declared before the record was rebuilt and every result recomputed.

5.2 Daily and weekly target construction

A daily mean requires at least 18 valid hours out of 24 and a finite, non-negative value; no values are imputed. The record is reindexed to every calendar day, so absent days count as missing rather than being skipped. A Monday-to-Sunday week is eligible if at least 6 of its 7 days are valid, and its target is the unweighted mean of the valid days. The 16 eligible weeks with only 6 valid days are flagged, because a 6-day mean is not necessarily the true 7-day mean. The code is unit-tested on synthetic data.

Removing days at random from the 145 complete OpenAQ weeks (every possible 1-day and 2-day subset) gave the following typical errors: a median absolute error of about 3% with one day missing, and 4 to 5% with two days missing. The 95th percentiles were about 10% and 16% respectively (Appendix G). This test assumes that days go missing at random. Real outages may cluster in particular conditions, which has not been tested.

5.3 Trajectory generation and clustering

Trajectory generation is described in Section 4.4. For the forecasts, trajectories were grouped by shape. Each trajectory was represented by 24 numbers: its position 6, 12, …, 72 hours before arrival, expressed in kilometres east and north of the monitor (equirectangular projection). Because all 24 values share one unit, no scaling was fitted. Clusters were found with k-means ($k = 6$, 20 random starts, seed 20261021). To avoid leakage, the clustering was refitted for every forecast week T using only trajectories that arrived more than seven days before T, which are the only ones that can enter training features, and at least 365 trajectories were required. The seven trajectories of week $T - 1$, which form the forecast's inputs, were only assigned to the nearest fitted centre. A cluster number can therefore refer to different pathways in different weeks. This does not affect the forecasts, because each model is trained and applied within one clustering, but it means the cluster numbers are not interpreted.

Each week was then described by nine features computed from its seven trajectories: the share of trajectories in each of the six clusters; the mean great-circle distance of the 72-hour endpoint from Lahore, which measures how far, and so how fast, the air travelled; the share of hourly endpoints within 500 km that lay in the east/south-east sector (bearing 45° to 180°, the same sector as the fire features); and the mean endpoint height above ground. Features were left missing if fewer than six of the seven trajectories existed, which never happened. Solutions with 4 and 8 clusters were run as sensitivity checks. This design was declared before any trajectory feature was computed (Section 5.13).

Of the robustness checks in the original study proposal, other arrival heights (Section 6.8) and a one-year comparison with GDAS 1° meteorology (Section 6.10) were run; bootstrap stability of cluster assignments and ERA5-driven trajectories were not. The descriptive pathway analysis (Section 5.14) is not a forecast and uses a separate, single clustering of all trajectories.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

5.4 Fire-exposure feature engineering

The forecasts use simple fire features (Section 4.5): detection counts and FRP in distance rings and direction sectors around Lahore, from the week before each forecast. The trajectory-weighted index described below has not been computed for the forecasts; a simplified, unweighted version is used in the association analysis (Section 5.14). The sectors are defined by direction, not country, so the east and south-east sector also contains Pakistani districts. A candidate indicator sums the FRP of detections near each trajectory, weighted by distance from the trajectory and by the time difference between detection and the air parcel's passage:

$$\mathrm{FE}(\text{week}) = \sum_{\text{trajectories}} \sum_{\text{points}} \sum_{\text{fires}} \mathrm{FRP} \times w_d(\text{distance}) \times w_t(\text{time difference}) \tag{1}$$

Here w_d and w_t are weights that decline with distance and time (for example, zero beyond a fixed buffer radius or time window). Their forms and parameters will be fixed before results are examined and then varied in the sensitivity analysis. A simpler version, total FRP in a fixed upwind region, will be computed for comparison. This follows the airshed approach of Liu et al. (2018). The indicator measures how much detected fire activity lay near the probable path of incoming air. It is not an estimate of smoke delivered to Lahore. Because weeks differ in the number of valid trajectories, the weekly value will be divided by the number of trajectories, so that it is a mean per trajectory rather than a total. Detections from different satellites will not be pooled freely. VIIRS detections from NOAA-21 were added to FIRMS in February 2024, so pooling every available sensor would raise detection counts during the final test period without any change in burning, and the same fire seen by two satellites a short time apart would be counted twice. The primary index will therefore use one sensor, or one fixed set of sensors, for the whole record, and MODIS and VIIRS FRP will not be mixed in one index.

Because fire activity and $PM_{2.5}$ both peak in October and November, a positive association could reflect seasonal timing rather than smoke reaching Lahore. Four checks address this. First, the weather-enhanced comparison model already includes calendar and climatology features, so fire exposure must add skill beyond the seasonal cycle. Second, fire-exposure skill will be evaluated separately for October and November weeks. Third, a null test will recompute exposure using fire records shifted by whole years, which keeps each fire's location and time of season but breaks its link with the actual trajectories; the real exposure must outperform these null versions. Fourth, results will be reported with and without 2020 and 2021, when COVID-19 restrictions changed activity patterns (Karim & Rappenglück, 2023).

5.5 Ventilation and one-box mass balance

Five weekly features based on the ventilation coefficient were computed from hourly ERA5 data and used in model 4 (Section 5.7). The one-box equation that motivates them is:

$$\frac{dC}{dt} = \frac{E}{H} + \frac{U}{L}(C_{\text{in}} - C) - kC \tag{2}$$

Here C is $PM_{2.5}$ concentration ($\mu g\,m^{-3}$), E is emission per unit area ($\mu g\,m^{-2}\,s^{-1}$), H is mixing-layer depth (m), U is a representative wind speed ($m\,s^{-1}$), L is an airshed length scale (m), C_{in} is the concentration of incoming air ($\mu g\,m^{-3}$) and k is a first-order removal rate (s^{-1}). Every term has units of $\mu g\,m^{-3}\,s^{-1}$ (Appendix C). Hourly ERA5 inputs in $m\,s^{-1}$ must be combined with time steps in seconds, or converted consistently, before the equation is used.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

If conditions change slowly compared with the flushing time L/U, the equation approaches the steady state:

$$C^* = \frac{E/H + (U/L)\, C_{\text{in}}}{U/L + k} \tag{3}$$

With $L = 20$ km, used here only as an illustration (roughly the square root of the city's quoted area of 404 km^2 in Mohyuddin et al., 2024), the flushing time is about 5.6 hours at $U = 1$ m s^{-1}. That is short compared with a week, so the steady-state form may be reasonable for weekly averages; this has not been checked against the data. The features use $L = 30$ km, an approximate extent of Lahore, which only scales the flushing-time feature.

E, C_{in} and k are not observed for Lahore. The model therefore cannot produce a trustworthy concentration or source split, and it is not used to do so. What it provides is a set of quantities computed from observed meteorology. For each forecast week, five were computed from the hourly ERA5 data of the previous week:

- the mean ventilation coefficient VC $= U \times H$, with U the 10 m wind speed and H the boundary-layer height (m^2 s^{-1});
- the mean ventilation coefficient at night (18:00 to 05:59 PKT);
- an accumulation index, the mean of $1/\text{VC}$ scaled by 1,000, with VC floored at 10 m^2 s^{-1} to avoid dividing by near-zero values;
- the median flushing time L/U in hours, with U floored at 0.1 m s^{-1};
- the number of hours with VC below 1,000 m^2 s^{-1}, a project-defined low-ventilation threshold.

These quantities are transformations of the ERA5 wind and boundary-layer variables already in model 2, so they were expected to be largely redundant. They were added as a layer of their own (model 4), so that their contribution could be measured once the raw ERA5 variables were present, and they did not reliably improve the forecasts (Section 6.6). Weekly values are means or medians of hourly indices rather than indices of weekly means, because the relationships are non-linear. The model is therefore not described as physically informed.

The box model can also be connected to the observations through a diagnostic test that needs no values for E, C_{in} or k. If inflow is small and removal is slow compared with ventilation, the steady state reduces to $C \approx EL/(UH)$, so $PM_{2.5}$ should scale with $1/(UH)$. Within a season, the slope of log weekly $PM_{2.5}$ against log weekly ventilation coefficient would then be close to -1. A slope nearer 0 would suggest that incoming pollution, removal or changing emissions also matter. This is a consistency check on the physical reasoning, not an estimate of source contributions, and it depends on how well ERA5 wind and boundary-layer height represent conditions at the monitor. It was run on weekly data within each season (Section 6.4); a daily version is part of the association analysis (H2).

5.6 Forecast-time feature availability and leakage prevention

A forecast for week T is issued at 00:00 PKT on Monday at the start of week T. Only weeks that ended before that time may be used, and among them only eligible weeks. Every model enforces this rule, and the observed target is joined only after prediction, for scoring. The main forecasts use no embargo: their inputs come from week $T - 1$. A one-week embargo was also run (see below).

Two kinds of experiment will be kept strictly separate:

1. **Forecast experiments** use only information that would have existed at issue time: lagged

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

$PM_{2.5}$, meteorology, trajectory statistics and fire exposure from earlier weeks, and calendar variables. If forecast meteorology or forecast trajectories are used, their issue times and archive source must be documented. Only these experiments can support claims about forecasting skill.

2. **Diagnostic experiments** may use reanalysis weather, trajectories and fire detections from the target week itself, and the explanatory model in Section 5.14 uses same-day conditions. They show how much of a week's $PM_{2.5}$ is associated with that week's conditions, an upper bound on what better weather forecasts could offer. They are not forecasts and will be labelled as such.

Any preprocessing that is fitted, such as the Ridge model's scaling or the trajectory clustering, is fitted only on the training portion of each walk-forward step.

Real data are not available instantly. ERA5 reanalysis, for example, is published some days after real time, and the previous week's $PM_{2.5}$, trajectories and fire detections may not all be complete by 00:00 on Monday. The main results therefore describe an idealised forecast that assumes zero data latency. A stricter version was also run, with a one-week embargo: every input had to end at least seven days before week T began. This applied to $PM_{2.5}$ history, persistence, climatology and the training targets, and the weather, fire, mass-balance and trajectory features came from week $T-2$. The trajectory clustering was then fitted only on trajectories that arrived more than 14 days before T. A per-source latency rule, which would lag each input by its documented delay, was not run: those delays were not compiled for every source, and the FIRMS standard-processing product used here is not released in near real time. The one-week embargo covers inputs published within a week, such as ERA5's preliminary release, which appears about five days after real time. If trajectory or fire features helped only without an embargo, their practical forecasting value would be limited.

5.7 Baseline models and Random Forest

Three reference baselines and a Random Forest built up in five layers were compared on the same forecast weeks.

1. **Persistence:** the most recent eligible weekly mean before the forecast cutoff. No maximum age was imposed: of 238 persistence forecasts for eligible weeks, 210 used the previous week, 25 used a week 2 to 5 weeks old and 3 used a week 7 to 9 weeks old.
2. **Seasonal climatology:** the median of eligible weekly means within ±2 weeks of the same calendar date in earlier years, requiring at least 4 such weeks from at least 2 earlier years. These settings were fixed in advance and not tuned. Because the record starts in May 2019, climatology is available only from mid-2021.
3. **Ridge $PM_{2.5}$-history model:** a linear model of log weekly $PM_{2.5}$ on the most recent eligible weekly mean (log) and its age in weeks, the mean of the eligible weeks among the previous four (log) and their number, and two annual sine and cosine harmonics of the target date. Features were standardised inside each training window, the penalty was fixed at $\alpha = 1$, and forecasts were back-transformed with a smearing correction (Duan, 1983).

The Random Forest (300 trees, minimum leaf size 5, a third of the features tried at each split, seed 20261021) was built up in layers, each adding one group of inputs from the week before the forecast: model 1, $PM_{2.5}$ history and calendar terms (8 features, the same as the Ridge model); model 2, plus ERA5 meteorology (29 features); model 3, plus FIRMS fire features (39); model 4, plus one-box mass-balance features (44); and model 5, plus the nine trajectory

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

features (53). It was trained on weekly mean $PM_{2.5}$ without transformation, imputation or fitted preprocessing, and refitted from scratch for every forecast week. Its settings were chosen once, before any results, and were not tuned. As a secondary comparison, the trajectory features were also added to model 1 alone. Two follow-up models were declared before fitting on 21 October 2026, to test whether the under-prediction of high weeks could be reduced: a Random Forest on log $PM_{2.5}$ with out-of-bag smearing, and a Ridge regression on log $PM_{2.5}$ using all inputs. Models 2 to 5 use 29 to 53 features, more than the 10 to 15 in the original study proposal (Section 5.13). With as few as 52 training weeks at the start of the walk-forward period and strongly correlated inputs, this raises the risk of overfitting. It is one reason the layer comparisons are judged by bootstrap intervals rather than by differences in MAE alone.

5.8 Walk-forward validation

Evaluation uses an expanding-window, rolling-origin design: for each forecast week, models are trained on all eligible weeks that ended before the cutoff and then predict that week. Weeks are never shuffled, and models are refitted from scratch every week. Forecasting starts once 52 training weeks are available. Every eligible week from mid-2021 to February 2025 was forecast in this way, with no separate held-back test period (Section 5.13). Models are compared on the 138 weeks (17 May 2021 to 10 February 2025, 53 of them in October to January) for which every method produced a forecast. Weeks drop out where climatology lacks enough earlier years, where no eligible week exists in the previous four weeks, or where too few earlier trajectories exist for clustering. Because consecutive weeks are autocorrelated and winter episodes cluster, the effective number of independent test weeks is smaller than 138.

5.9 Metrics

For each model, the following are reported on the shared forecast weeks: mean absolute error (MAE), root mean square error (RMSE), out-of-sample R^2 (1 minus the ratio of the sum of squared errors to the total sum of squares about the test-period mean, which can be negative), and mean bias (predicted minus observed). A skill score relative to persistence, $1 - \mathrm{MAE}_{\mathrm{model}}/\mathrm{MAE}_{\mathrm{persistence}}$, is also given. MAE is the primary metric; the others are secondary. Differences between models will be assessed from paired errors on the same weeks, using a moving-block bootstrap of the paired weekly error differences, which allows for autocorrelation (4-week blocks and 2,000 resamples with a fixed seed, with 2- and 8-week blocks as checks). Results are also reported separately for October-to-January weeks, a group known when each forecast is issued. Scores for observed high-pollution weeks alone are not reported, because choosing weeks by their observed value favours forecasts that always predict high values.

5.10 SHAP interpretation

SHAP values (S. Lundberg & Lee, 2017) were computed for the full model (model 5) with the exact TreeSHAP algorithm for tree ensembles (S. M. Lundberg et al., 2020; `shap` version 0.54.1). Each of the 232 walk-forward forecasts was explained by the model trained for that week, so no model explained data it had been trained on. Two checks were applied. The refitted models had to reproduce the saved forecasts to within 10^{-6} µg m^{-3} (largest difference 5.0×10^{-7}), and the SHAP values of each forecast had to add up to its prediction (largest error 4.3×10^{-13}). Features were grouped into six families: $PM_{2.5}$ history, calendar, ERA5 weather, fire, mass balance and trajectories. A family's contribution to a forecast is the sum of its features' SHAP values. Results are reported as the mean absolute contribution of each family and its share of the total, over the 138 common weeks and separately for October to January. The ten most-used individual features are also listed, leaving out the cluster shares, whose meaning changes from week to week (Section 5.3). Inputs such as boundary-layer height, temperature,

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

season and ventilation are correlated, so SHAP may divide credit among them unpredictably, and family totals are given more weight than individual ranks. The design was declared before any SHAP value was computed (Section 5.13). SHAP values describe what the model relied on: they are associations learned by the model, not causes. The analysis was repeated in the embargo run and the four data-processing sensitivity runs (Section 5.11).

5.11 Sensitivity analysis

Sensitivity analyses are limited to a pre-specified set, because trying many combinations on about 150 to 200 weeks could make some version look successful by chance. Eight were run:

- the number of trajectory clusters (4 and 8 against 6);
- the trajectory arrival height (100 and 250 m against 500 m), which was also used to repeat the association tests;
- the bootstrap block length (2 and 8 weeks against 4);
- the one-week embargo (Section 5.6);
- the $PM_{2.5}$ source (AirNow's own daily values against the rebuilt daily means);
- the weekly completeness threshold (7 of 7 against 6 of 7 days);
- the repeated-value run rule (6 against 7 hours);
- the exclusion of 2020 and 2021, when COVID-19 restrictions changed activity, as training and evaluation weeks.

For each of the last four, a run was defined in advance to agree with the main result if three conditions held: the trajectory layer's interval against model 4 includes zero; the full model beats persistence with an interval that excludes zero; and its interval against the Ridge model includes zero. The SHAP analysis was repeated in the embargo run and in each of these four runs. The rankings were judged stable if the two largest feature families stayed the same and the Spearman correlation of the six family shares with the main run was at least 0.8. Three checks proposed earlier could not be run with the available data:

- other trajectory durations, and finer trajectory meteorology for the whole record (a one-year comparison with GDAS 1° was declared separately and is reported in Section 6.10);
- an average of neighbouring ERA5 cells, because only the nearest grid point was downloaded;
- the fire-exposure buffer distance for the forecasts, which use distance rings rather than trajectory buffers (buffer sensitivity for the association analysis is reported in Section 6.4).

Any other variation will be reported as exploratory.

5.12 AQI reporting

The primary target is weekly mean $PM_{2.5}$ concentration. No AQI is computed in this study, and no AQI standard has been chosen. Under the US EPA's current rules, the $PM_{2.5}$ AQI uses 24-hour concentrations truncated to one decimal place, with breakpoints of 0.0 to 9.0 $\mu g\,m^{-3}$ (Good), 9.1 to 35.4 (Moderate), 35.5 to 55.4 (Unhealthy for Sensitive Groups), 55.5 to 125.4 (Unhealthy), 125.5 to 225.4 (Very Unhealthy) and 225.5 or more (Hazardous), effective 6 May 2024 (US EPA, 2024; 40 CFR Part 58, Appendix G). Punjab EPA bulletins use a different scale, and two bulletins checked in this project used different category bands. A 2021 bulletin used "Smog Policy 2017" bands, while a November 2024 bulletin listed bands from 0 to 50 "Good" to 401 to 500 "Hazardous" without stating the averaging period or method. Because every candidate standard is defined for 24-hour (or shorter) averages, a weekly mean cannot be converted into an official AQI. If an AQI-related output is reported later, it will be labelled with

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

the standard used, derived explicitly, and kept separate from the primary $PM_{2.5}$ target.

5.13 Pre-analysis plan

No single plan for the whole study was written before the analysis began. Instead, the forecasting analysis was fixed in stages. Each stage's design was set before its results were computed, and from the follow-up models onwards each design was written into a dated declaration file before any model was fitted. The declarations for the trajectory layer and the SHAP analysis were also committed to the project's version history before their results existed.

- **Reference baselines and the Ridge $PM_{2.5}$-history model** (20 October 2026): settings recorded in the project's design notes before results were computed; not tuned.
- **Random Forest models 1 to 4** (21 October 2026): design fixed before results were computed and recorded in the results report, without a separate declaration file.
- **Follow-up models** (21 October 2026): declared in a separate file before fitting.
- **Trajectory features and model 5** (23 October 2026): declared before any trajectory feature was computed.
- **SHAP analysis** (23 October 2026): declared before any SHAP value was computed.
- **Backfilled September 2022 data** (25 October 2026): the decision to add AirNow's regenerated file and rerun every step was declared before the record was rebuilt; all results in this draft come from that rerun.
- **Association analysis** (25 October 2026): the pathway, ventilation and fire-exposure tests of Section 5.14 were declared, approved and committed before any of their results were computed.
- **Embargo, sensitivity, SHAP-stability, diagnostic and weekly ventilation checks** (25 October 2026): declared, approved and committed before any of their code was run (Sections 6.4 to 6.9).
- **Boundary-layer check** (25 October 2026): the share of days on which ERA5 boundary-layer height was below the trajectory arrival height was declared and approved before it was computed (Section 7.2).
- **Trajectories at 100 and 250 m** (25 October 2026): their analysis was declared, approved and committed before the trajectories existed (Sections 6.3, 6.4, 6.8 and 7.2).

The settings used throughout were: the rebuilt AirNow hourly record; daily means from at least 18 of 24 valid hours, after removing runs of 7 or more identical hourly values; Monday-to-Sunday (PKT) weeks with at least 6 of 7 valid days; forecasts issued at 00:00 PKT on Monday with inputs from the previous week; an expanding training window with at least 52 training weeks; MAE as the primary metric, with RMSE, bias, out-of-sample R^2 and skill relative to persistence as secondary metrics; and a paired moving-block bootstrap of weekly errors (4-week blocks, 2,000 resamples, fixed seed), with 2- and 8-week blocks as checks. Results are also reported for October-to-January weeks.

Because the forecasting runs were carried out before this pre-analysis plan was written, they depart from the original study proposal in the following ways. No setting was changed after its results had been seen; each stage report records this.

- **Embargo.** The main forecasts use inputs from the week immediately before the forecast, which assumes that all data are available without delay. A one-week embargo was run as a declared check (Section 6.5); the per-source latency rule was not run (Section 5.6).
- **No separate final test period.** Every week from 17 May 2021 to 10 February 2025 was forecast with the same expanding walk-forward design, so the proposed final test period

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

(October 2023 to February 2025) was not held back as an unseen test.

- **No high-pollution threshold.** Results are instead reported for October-to-January weeks, a group that is known when a forecast is issued.
- **Larger models.** Models 2 to 5 use 29 to 53 features rather than 10 to 15, and the Random Forest settings were chosen once rather than from a small grid (Section 5.7).
- **Fire features.** The forecasts use detection counts and fire radiative power in distance rings and direction sectors (Section 5.4). The trajectory-weighted fire exposure has not been computed.
- **Number of clusters.** It was fixed at 6, with 4 and 8 as sensitivity checks, rather than chosen with the Davies–Bouldin index (Section 5.3).

5.14 Pathway and association analysis (H1 to H3)

The association analysis was declared before any of its results were computed (Section 5.13), and all of its results are reported. It works with days rather than weeks, because a week usually mixes several pathways and averaging would blur the differences between them. It uses the 1,801 days with a complete daily mean and a trajectory arriving that day. Each trajectory arrives at 06:00 UTC (11:00 PKT), so the daily mean covers about 11 hours before arrival and 13 hours after it. The outcome is log daily $PM_{2.5}$, so effects are ratios of geometric means. All tests are made within the four seasons of Section 4.6 (winter 475 days, pre-monsoon 385, monsoon 633, post-monsoon 308), with October to January as a secondary split.

Clustering. All 2,161 trajectories were clustered once, with Ward's hierarchical method on the same 24-number shape vector as in Section 5.3. This clustering is descriptive and is not used in any forecast. The number of clusters was fixed in advance by HYSPLIT's rule of thumb: the largest k between 2 and 10 at which merging into $k - 1$ clusters would raise the total within-cluster variance by at least 30%. The comparisons were repeated with $k - 1$ and $k + 1$ clusters.

Pathway, weather and fire variables. Each day was described by its cluster; its 72-hour transport distance and that distance's tercile within the season; daily means of ERA5 10 m wind speed, boundary-layer height, ventilation coefficient (the hourly product of the two), 2 m temperature and relative humidity; and whether at least 1 mm of rain fell. Fire exposure is the total FRP of the unique VIIRS detections that lie within 50 km of any hourly trajectory point below 1,500 m above ground and were detected on that point's PKT day or the day before. It is a simplified, unweighted version of the indicator in Section 5.4 and is the project's own construction. Days on which Suomi NPP recorded no detections within 500 km while NOAA-20 recorded at least 100 were treated as missing (22 days).

Tests. For H1, log $PM_{2.5}$ was regressed on cluster within each season; clusters with fewer than 30 days in a season were described but not tested. The same comparison was repeated on anomalies from a day-of-year climatology and after adjusting for ventilation, wind, temperature, humidity and rain (a "confounding ladder"). A secondary test, H1b, compared the third of days with the shortest transport distance with the third with the longest. For H2, log $PM_{2.5}$ was regressed on log ventilation coefficient within each season, with and without cluster. For H3, on October and November days, log $PM_{2.5}$ was regressed on log ventilation coefficient, cluster, day of year, year and log(1 + fire exposure). The fire coefficient was then compared with the coefficients obtained when the fire record was shifted by one to five years, which keeps each fire's location and time of season but breaks its link with the actual trajectories. Because pollution and pathways on neighbouring days are correlated, confidence intervals come from a moving-block bootstrap of 7-day blocks within each season and year (2,000 resamples; 14- and 28-day blocks as checks), and p-values from a block permutation test, with Holm's correction

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

across seasons. Sensitivity checks used $k - 1$ and $k + 1$ clusters, fire buffers of 25 and 100 km, and no height limit.

Source-region maps. Concentration-weighted trajectory (CWT) maps, as used at this monitor by Karim and Rappenglück (2023) and for Delhi by Ghosh et al. (2014), give each 1° grid cell the mean $PM_{2.5}$ of the days whose trajectory points fell in it, down-weighted where few points fell. A 1° grid replaces the 0.5° grid proposed earlier, because the 2.5° meteorology cannot resolve finer detail. Maps are shown for the post-monsoon and winter seasons beside maps of trajectory-point frequency. They show where air passed before polluted days, not where pollution came from.

The same-day "explanatory model" in the original study proposal has not been run.

6. Results

Sections 6.1 to 6.4 report descriptive results and associations, which on their own say nothing about forecasting skill. Sections 6.5 to 6.8 report out-of-sample predictive results under the pre-analysis plan (Section 5.13). Section 6.9 reports a diagnostic experiment that uses target-week information (Section 5.6); it is not a forecast and is labelled as such.

6.1 Dataset coverage

The rebuilt AirNow record covers 22 May 2019 to 18 February 2025 (2,100 local days and 50,400 hours), of which 43,750 hours are valid, including 599 hours from AirNow's regenerated September 2022 file (Section 4.2). After the S7 repeated-value rule, which removed 9 hours, 1,801 days meet the 18-of-24-hour rule. Of 299 full Monday-to-Sunday weeks, 239 are eligible under the rule of at least 6 valid days (223 with all 7 days and 16 with 6). Rebuilt daily means agree with AirNow's own daily values to within 0.5 $\mu g\,m^{-3}$ on 99.5% of 1,778 days. Data availability is shown in Figure 2, the weekly series in Figure 3, and the monthly cycle of $PM_{2.5}$, fire activity, ventilation and rainfall in Figure 4. Assigning each week to the year of its Thursday, the number of eligible weeks was 30 in 2019 (from late May), 51 in 2020, 33 in 2021, 33 in 2022, 44 in 2023, 41 in 2024 and 7 in 2025 (to mid-February). The low counts for 2021 and 2022 come from long outages of unknown cause, such as 21 October to 7 December 2021. Of the 239 eligible weeks, 88 fall in October to January. In the walk-forward evaluation, the Ridge history model produced 232 forecasts, and 138 weeks (53 in October to January) had forecasts from every method (Section 5.8). The preliminary audit of the OpenAQ copy is a data-quality check, not a result, and is reported in Appendix G.

6.2 Trajectory clusters

The rule chose three clusters: merging three clusters into two would raise the within-cluster variance by 39%, whereas merging four into three would raise it by only 24%. Three clusters also had the highest silhouette score (0.42), and k-means gave a similar grouping (adjusted Rand index 0.82). The clusters differ mainly in speed (Figure 5). The most common, C1 (60% of days), is slow air that arrived from the south-east after travelling a mean of 375 km in 72 hours, at a mean height of 1,064 m above ground. C2 (37%) is faster air from the north-west (1,171 km; 2,295 m), and C3 (3%) is fast, high air from far to the west (3,632 km; 3,298 m). C1 dominates the monsoon months (77 to 90% of days in July to September); in the other months C1 and C2 are about equally common, and C3 occurs mainly in February and March. Each cluster mixes days with very different weather, so the clusters are better read as broad flow regimes than as source regions.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

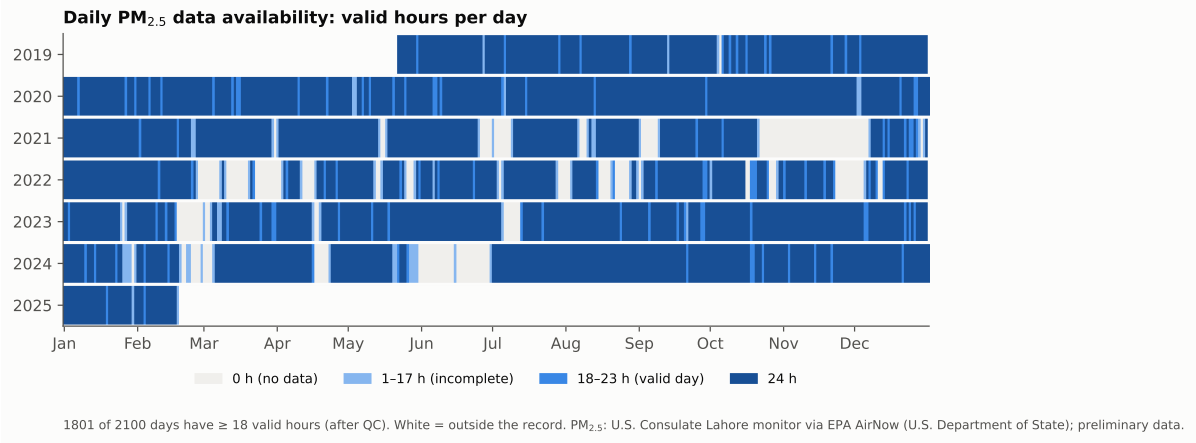


Figure 2

Daily $PM_{2.5}$ data availability: valid hours per day, May 2019 to February 2025. Of 2,100 days, 1,801 have at least 18 valid hours after quality control. $PM_{2.5}$: U.S. Department of State monitor in Lahore, via EPA AirNow; preliminary data.

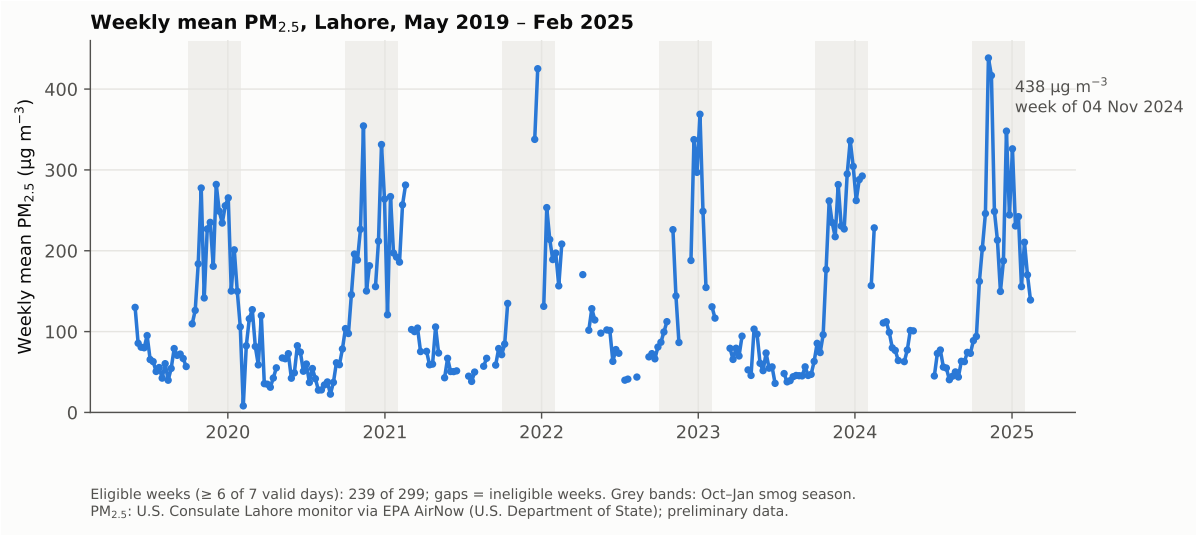


Figure 3

Weekly mean $PM_{2.5}$ (Monday to Sunday, PKT) for the 239 eligible weeks with at least 6 of 7 valid days; gaps are ineligible weeks, and grey bands mark October to January. The highest week, 438 µg m^{-3}, began on 4 November 2024. Preliminary AirNow data.

6.3 $PM_{2.5}$ differences across pathways

Figure 6 shows daily $PM_{2.5}$ by cluster within each season, and Table 5 summarises the tests. C3 had fewer than 30 days in every season and was not tested. Fast north-westerly days (C2) had higher $PM_{2.5}$ than slow south-easterly days (C1) in every season: by 16% in winter (ratio of geometric means 1.16, 95% CI 1.00 to 1.28), 30% in the pre-monsoon season (1.30, 1.18 to 1.49), 9% in the monsoon (1.09, 1.00 to 1.19) and 7% in the post-monsoon season (1.07, 0.91 to 1.34). Only the pre-monsoon difference was clear in the block permutation test after Holm's correction (p = 0.006, against 0.21 in winter and the monsoon and 0.52 in the post-monsoon season). The pattern was the same with two or four clusters and with longer bootstrap

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

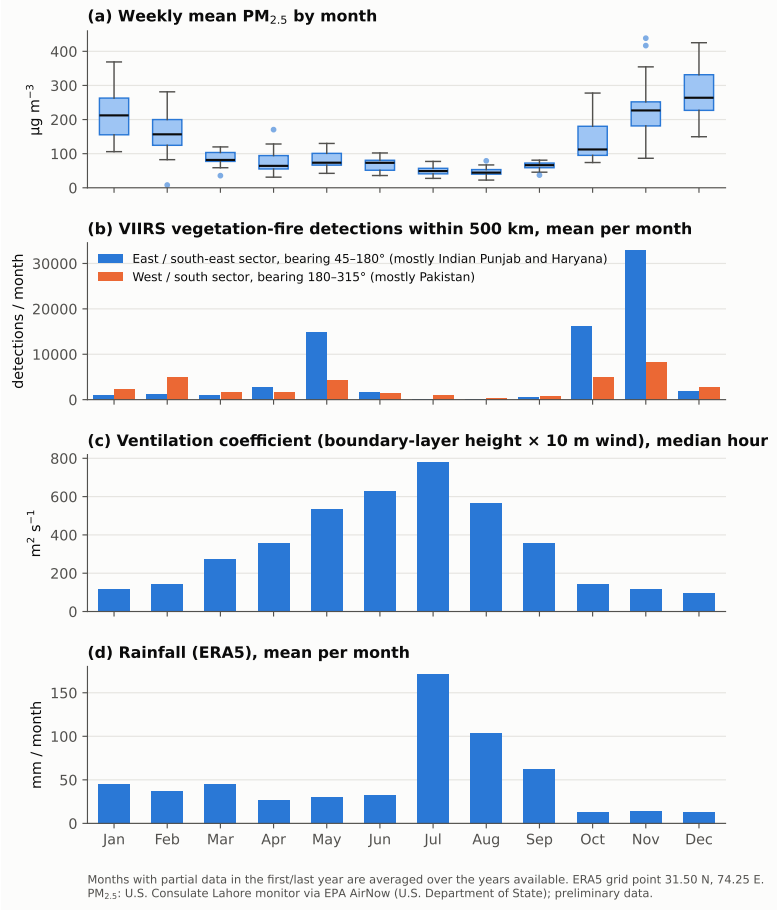


Figure 4

Monthly cycle of (a) weekly mean $PM_{2.5}$, (b) VIIRS vegetation-fire detections within 500 km by direction sector, (c) median hourly ventilation coefficient (ERA5 boundary-layer height × 10 m wind speed) and (d) ERA5 rainfall. $PM_{2.5}$ is highest from November to January, when ventilation is lowest; fire detections peak in October and November, with a smaller peak in May, mostly in the east/south-east sector. Preliminary AirNow data.

blocks. Accounting for weather removed these differences. With ventilation, wind, temperature, humidity and rain in the model, the ratios became 0.98 in winter, 1.07 in the pre-monsoon season, 0.94 in the monsoon and 1.04 in the post-monsoon season, all with intervals that include 1. The differences between pathways therefore largely go with the weather that accompanies each pathway.

Slow air was not more polluted (Figure 7). On the third of days with the shortest 72-hour transport distance, $PM_{2.5}$ was lower than on the third with the longest in winter (ratio 0.86, 0.68 to 1.03), the pre-monsoon season (0.80, 0.70 to 0.90) and the post-monsoon season (0.84, 0.67 to 1.02), and higher only in the monsoon (1.18, 1.05 to 1.31). The secondary hypothesis H1b is therefore not supported. Trajectories recomputed with arrival heights of 100 and 250 m (declared in advance; same meteorology) gave the same pattern. The C2/C1 ratios were 1.06 to 1.34 raw and 0.95 to 1.12 after weather adjustment. The shortest-to-longest distance ratios were below 1 in every season except the monsoon, with a winter ratio of 0.90 (95% CI 0.75 to 1.03) at 100 m (Section 7.2).

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

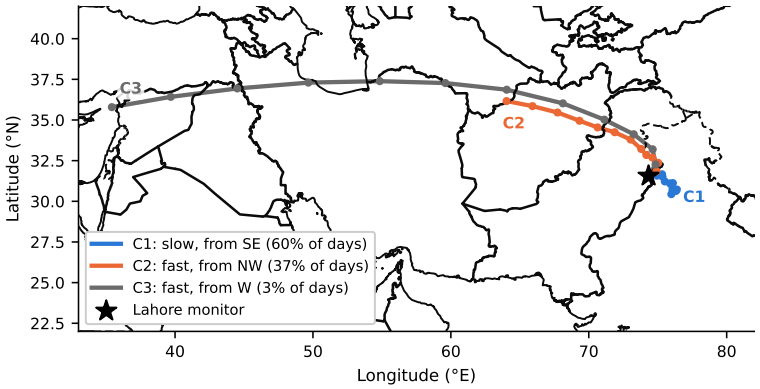


Figure 5

Mean 72-hour back-trajectory paths of the three clusters, with the share of all 2,161 days in each (dots every 6 hours). HYSPLIT v5.4.2 with NCEP/NCAR Reanalysis 2.5° meteorology, arrival at 06:00 UTC and 500 m above ground. Borders: Natural Earth; no position on disputed territory is implied.

Table 5

Pathway associations within seasons. Ratios of geometric mean daily $PM_{2.5}$ with 95% block-bootstrap intervals; C1 is slow air from the south-east and C2 fast air from the north-west. The weather-adjusted ratio controls for ventilation, wind speed, temperature, relative humidity and rain. The ventilation slope is from log $PM_{2.5}$ on log ventilation coefficient with cluster in the model. Preliminary AirNow data.

Season	Days	C2/C1, raw	C2/C1, adjusted for weather	Permutation p (Holm)	Shortest/longest distance tercile	Ventilation slope
Winter (Dec–Feb)	475	1.16 (1.00 to 1.28)	0.98 (0.85 to 1.09)	0.21	0.86 (0.68 to 1.03)	−0.47 (−0.57 to −0.39)
Pre-monsoon (Mar–May)	385	1.30 (1.18 to 1.49)	1.07 (0.96 to 1.22)	0.006	0.80 (0.70 to 0.90)	−0.16 (−0.24 to −0.02)
Monsoon (Jun–Sep)	633	1.09 (1.00 to 1.19)	0.94 (0.87 to 1.00)	0.21	1.18 (1.05 to 1.31)	−0.07 (−0.19 to +0.03)
Post-monsoon (Oct–Nov)	308	1.07 (0.91 to 1.34)	1.04 (0.92 to 1.26)	0.52	0.84 (0.67 to 1.02)	−0.47 (−0.61 to −0.37)

The CWT maps (Figure 8) show trajectory points concentrated near Lahore in both seasons. In the post-monsoon season the highest weighted values lie in cells south-east of Lahore (around 29 to 30° N, 77 to 79° E) and north-west of it (around 32 to 33° N, 70 to 71° E); in winter, high values cover a broad area around and north-west of the city. Single cells are not interpreted.

6.4 Weather and fire-exposure associations

Ventilation was strongly associated with $PM_{2.5}$ in the colder seasons (Table 5). Within winter and the post-monsoon season, the slope of log $PM_{2.5}$ on log ventilation coefficient was −0.47 (95% CI −0.57 to −0.39 in winter and −0.61 to −0.37 in the post-monsoon season, with cluster in the model). A 10% lower ventilation coefficient therefore went with about 5% higher $PM_{2.5}$, whatever the pathway. The slope was smaller in the pre-monsoon season (−0.16, −0.24 to −0.02) and indistinguishable from zero in the monsoon (−0.07, −0.19 to +0.03), and adding

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

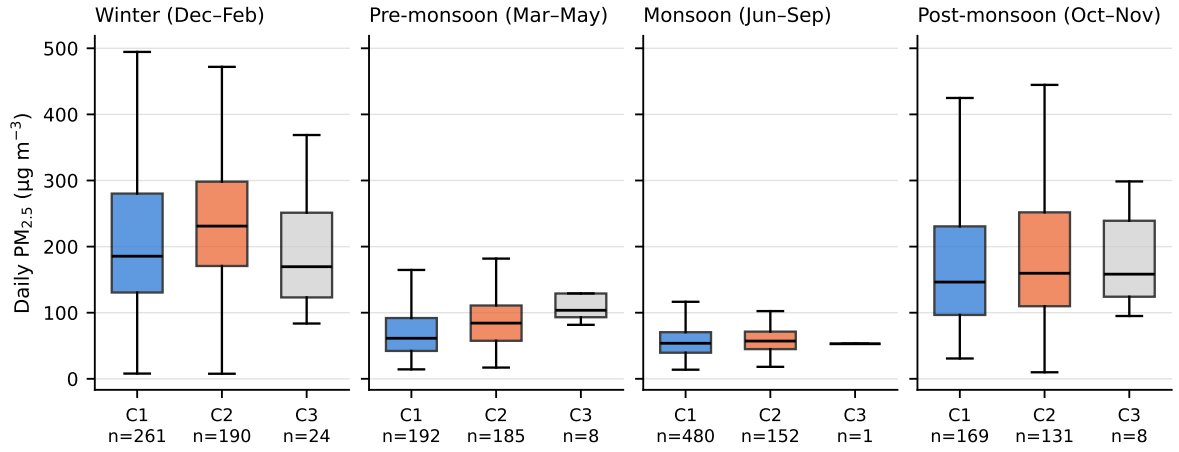


Figure 6

Daily mean $PM_{2.5}$ by trajectory cluster within each season (boxes: median and interquartile range; whiskers: 1.5 times the interquartile range; outliers not shown). Grey boxes have fewer than 30 days and were not tested. Preliminary AirNow data.

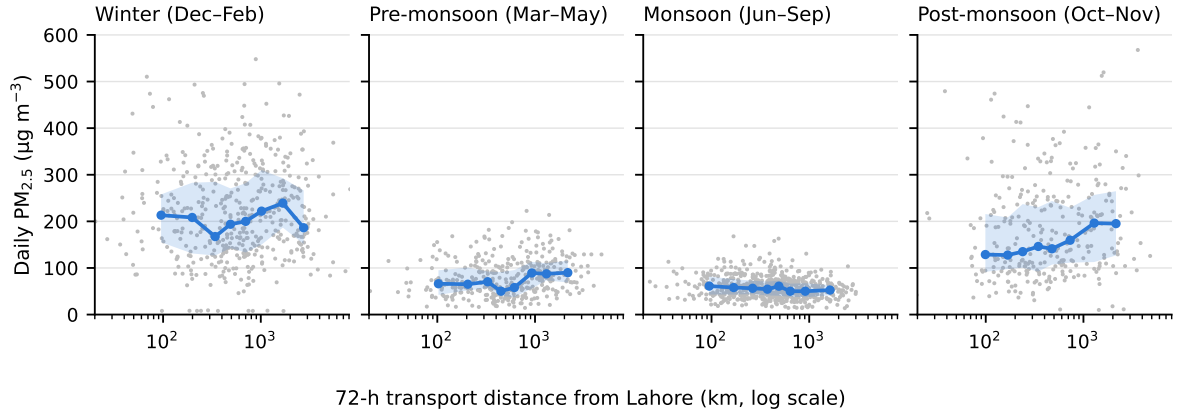


Figure 7

Daily mean $PM_{2.5}$ against the 72-hour transport distance of the day's trajectory, by season. Grey dots are days; the line and band show the median and interquartile range within eight distance bins. Short distances mean slow-moving air. Preliminary AirNow data.

cluster barely changed any slope. H2 is supported in three of the four seasons. The slopes are about half the value of −1 expected if $PM_{2.5}$ simply scaled with 1/ventilation (Section 5.5), which suggests that emissions, inflow or removal also vary from day to day.

At weekly resolution, the same check (Section 5.5) gave slopes closer to −1: −0.58 in winter (95% CI −0.76 to −0.41; 61 weeks) and −0.85 in the post-monsoon season (−1.15 to −0.63; 43 weeks), whose interval includes −1. There was no association in the pre-monsoon season (+0.09, −0.28 to +0.39) or the monsoon (+0.12, −0.14 to +0.27). Averaging over a week smooths out day-to-day changes in emissions and inflow, so in the polluted seasons weekly $PM_{2.5}$ behaves more like the simple box model predicts. This is a same-week association, not a forecast or a source estimate.

On October and November days, trajectory-weighted fire exposure was positively associated

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

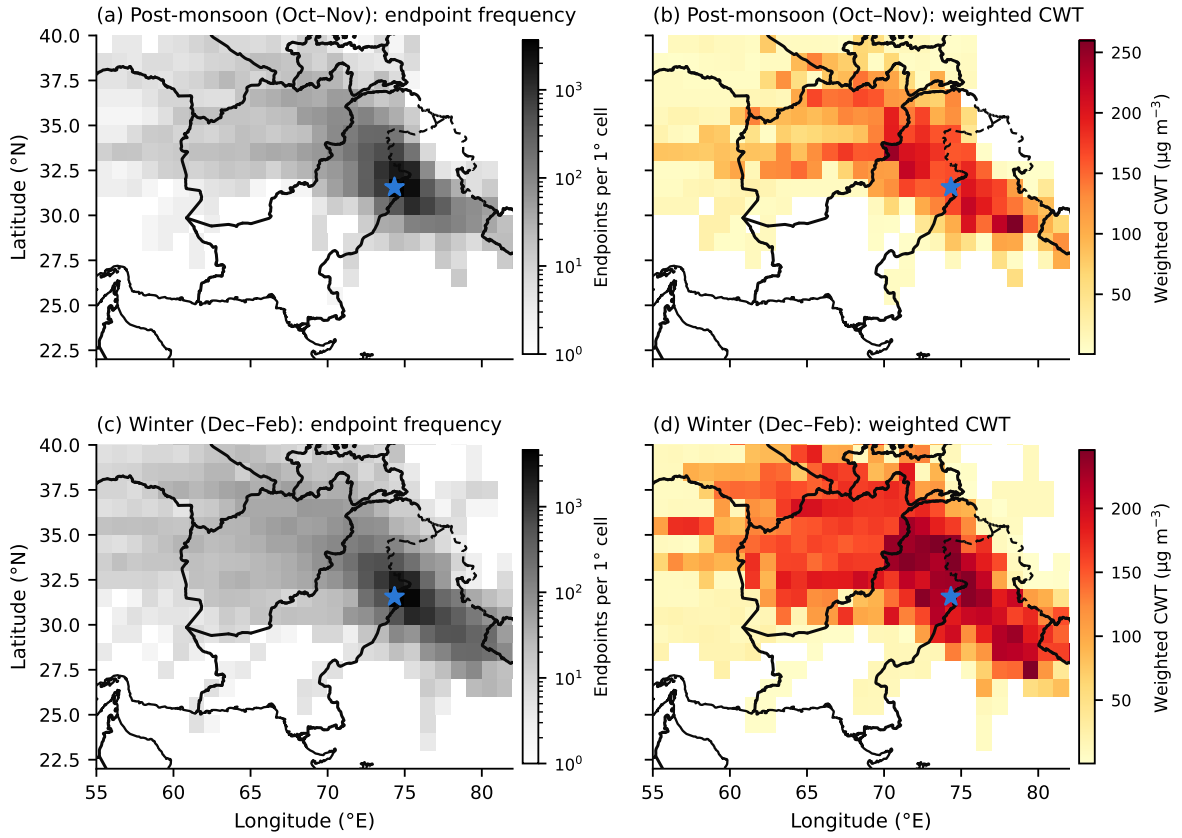


Figure 8

Trajectory-point frequency (left) and weighted concentration-weighted trajectory (CWT) values (right) on a 1° grid, for the post-monsoon (top) and winter (bottom) seasons. CWT gives each cell the mean daily $PM_{2.5}$ of the days whose trajectories passed through it, down-weighted where few points fell. The maps show where air passed before polluted days, not where pollution came from; single cells are not interpreted. Star: Lahore monitor. Preliminary AirNow data.

with $PM_{2.5}$ after accounting for ventilation, pathway, time of season and year (Figure 9). The coefficient on log(1 + exposure) was 0.081 (95% CI 0.028 to 0.118; 296 days), so doubling exposure went with about 6% higher $PM_{2.5}$. It exceeded all nine coefficients obtained from year-shifted fire records (0.011 to 0.071). The result held with a 100 km buffer (0.131, 0.057 to 0.194), without a height limit (0.092, 0.046 to 0.128) and with 14- and 28-day blocks, but was borderline with a 25 km buffer (0.045, –0.000 to 0.079). H3 is therefore supported, with a modest effect. The year-shifted coefficients were themselves positive, so part of the raw link between fire activity and $PM_{2.5}$ reflects their shared seasonal timing. Fire exposure was above zero on 96% of October and November days. With trajectories arriving at 100 or 250 m, the coefficient was similar (0.099, CI 0.057 to 0.133; and 0.096, CI 0.046 to 0.133). However, one and two of the nine year-shifted records respectively gave coefficients at least as large (up to 0.110), so at these heights the real fire record no longer beat every shifted record. The comparison with shifted records, and with it the third part of H3, is therefore sensitive to the arrival height.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

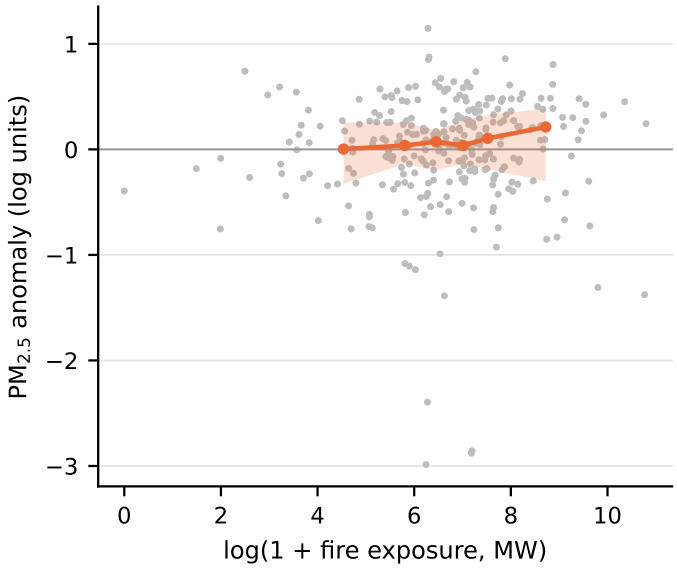


Figure 9

Trajectory-weighted fire exposure against the $PM_{2.5}$ anomaly (log daily $PM_{2.5}$ minus its day-of-year average) on October and November days. Grey dots are days; the line and band show the median and interquartile range within six exposure bins. Fire exposure: total FRP of VIIRS detections within 50 km of the trajectory below 1,500 m. Preliminary AirNow data.

6.5 Forecasting performance

Set-up. Every method was scored on the same 138 eligible weeks from 17 May 2021 to 10 February 2025, 53 of them in October to January. Forecasts are walk-forward, start once 52 training weeks are available, and use no data embargo, with inputs from week $T - 1$. Random Forest models 1 to 4 were run before the pre-analysis plan was written (Section 5.13); the trajectory layer (model 5) and the SHAP analysis were declared in dated files committed on 23 October 2026, before their results existed. The numbers come from the rerun of 25 October 2026, after AirNow's backfilled September 2022 data were added (Section 5.1); before that change the comparison set had 133 weeks and every MAE was 0.6 to 1.3 µg m^{-3} higher. $PM_{2.5}$ observations are preliminary AirNow data.

Reading the scores. Table 6 gives five scores for each method. MAE, the primary score, is the average size of a forecast's error in µg m^{-3}; lower is better. RMSE is similar but penalises large misses more. R^2 is the share of the week-to-week variation in $PM_{2.5}$ that the forecasts reproduce (1 would be perfect). Bias is the average of forecast minus observed, so a negative bias means the method tends to forecast too low. Skill is the fraction of persistence's MAE that a method removes: a skill of 0.20 means errors 20% smaller than simply repeating last week's value. Numbers in brackets are 95% confidence intervals; when an interval for a difference excludes zero, the difference is unlikely to be chance. All values in this section are in µg m^{-3} unless stated otherwise.

Baselines. Repeating the last eligible week (persistence) gave an MAE of 35.5. The seasonal climatology was 8% better (MAE 32.6), and a Ridge regression on $PM_{2.5}$ history, the strongest baseline, was 16% better (MAE 29.7).

Random Forest. Adding input groups one at a time changed the MAE only slightly. It was 29.2 with $PM_{2.5}$ history alone (model 1), 28.9 after adding ERA5 meteorology (model 2), 28.1

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

after adding fire features (model 3, the lowest of any method), 28.7 after adding mass-balance features (model 4), and 28.3 for the full model with HYSPLIT trajectory features (model 5; skill 0.20). With 4 or 8 trajectory clusters instead of 6, the full model's MAE was 28.3 and 28.2. The two follow-up models declared before fitting did not do better: a Random Forest on log-transformed $PM_{2.5}$ had an MAE of 28.9, and a Ridge regression with all inputs 32.7.

Table 6

Forecast skill of every method on the 138 common weeks, with errors for October to January (53 weeks) and the other months (85 weeks). MAE, RMSE and bias are in μg m^{-3}; bias is forecast minus observed, and skill is $1 - \mathrm{MAE}/\mathrm{MAE}_{\mathrm{persistence}}$. *Bold marks the best value in each column except bias. Walk-forward forecasts with no data embargo; preliminary AirNow data. * Follow-up models declared before fitting (21 October 2026).*

Method	MAE	RMSE	R^2	Bias	Skill	MAE Oct–Jan	Bias Oct–Jan	MAE other months
Persistence (last eligible week)	35.5	56.9	0.64	0.0	0.00	63.8	−6.7	17.9
Seasonal climatology	32.6	49.3	0.73	−11.1	0.08	54.8	−23.2	18.7
Ridge, $PM_{2.5}$ history	29.7	**46.1**	**0.77**	−4.7	0.16	54.1	−11.1	14.5
RF 1: $PM_{2.5}$ history	29.2	47.5	0.75	−8.5	0.18	52.3	−29.0	14.8
RF 2: + ERA5 meteorology	28.9	47.3	0.75	−7.8	0.19	50.3	−26.8	15.5
RF 3: + fire	**28.1**	47.3	0.75	−7.7	**0.21**	**50.1**	−26.9	14.4
RF 4: + mass balance	28.7	48.0	0.75	−6.9	0.19	50.7	−25.9	15.0
RF 5: + HYSPLIT trajectories (full model)	28.3	47.1	0.76	−6.2	0.20	50.4	−23.9	14.6
RF on log $PM_{2.5}$, all inputs*	28.9	48.2	0.74	−7.7	0.19	52.3	−25.5	**14.3**
Ridge, all inputs*	32.7	53.2	0.69	−0.9	0.08	56.8	+0.3	17.7

Full model against the baselines. Table 7 gives the paired differences in MAE, from a 4-week moving-block bootstrap (2,000 resamples). Over all weeks, the full model was reliably better than persistence (7.1 lower) and climatology (4.2 lower), but not reliably better than the Ridge history model (1.3 lower, with an interval that includes zero). In October to January, the season that matters most, its MAE was 50.4, against 63.8 for persistence, 54.8 for climatology and 54.1 for the Ridge model, and only the gain over persistence was reliable.

Table 7

Difference in MAE (μg m^{-3}), full Random Forest minus each baseline, with 95% moving-block bootstrap intervals (4-week blocks, 2,000 resamples). Negative values favour the full model; intervals that include zero indicate no reliable difference.

Baseline	All weeks (n = 138)	Oct–Jan (n = 53)
Persistence	−7.1 (−12.7 to −2.5)	−13.4 (−24.4 to −3.7)
Climatology	−4.2 (−8.2 to −1.2)	−4.5 (−10.8 to +1.0)
Ridge, $PM_{2.5}$ history	−1.3 (−5.4 to +1.7)	−3.7 (−11.3 to +3.5)

High-pollution weeks. Every method under-predicted the most polluted weeks. In October to January the full model's bias was −23.9, meaning its forecasts were on average about 24 μg m^{-3} too low. Figure 10 compares the errors of all methods, Figure 11 shows the forecasts week by week, and Figure 12 plots each forecast against the observed value.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

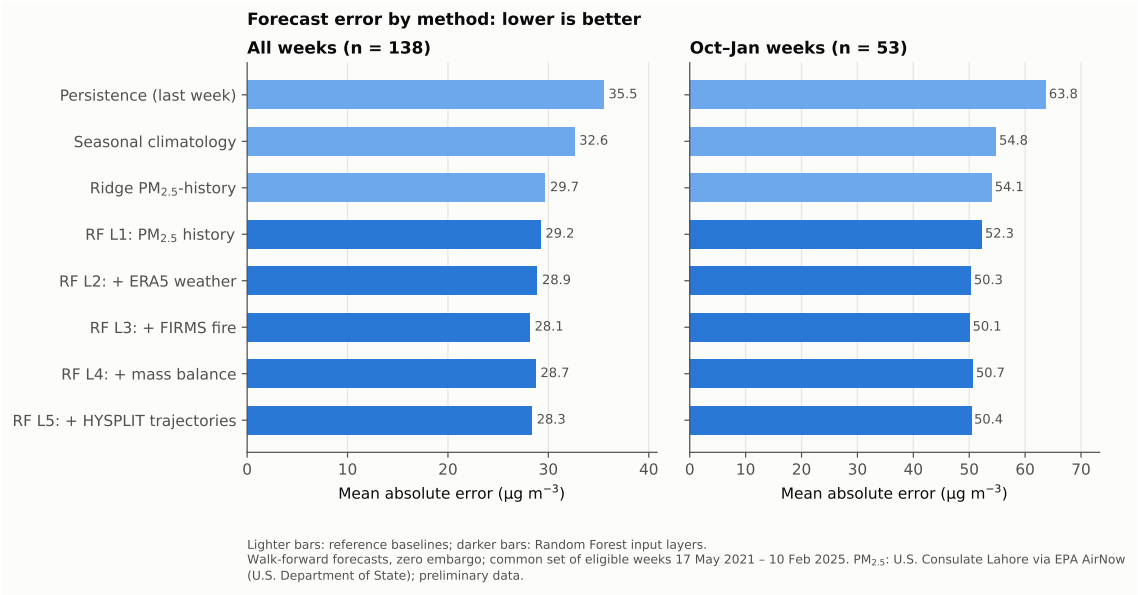


Figure 10

Mean absolute error of each forecasting method over all 138 common weeks (left) and the 53 October-to-January weeks (right). Lighter bars are reference baselines; darker bars are Random Forest models with inputs added layer by layer. Walk-forward forecasts with no data embargo.

With a one-week embargo. When every input had to end a week before the forecast week began (Section 5.6), every method except climatology lost accuracy (Table 8). Compared with the same model without an embargo, the full model's MAE rose by 5.1 (95% CI 2.3 to 8.3). The other Random Forest models' MAE rose by 3.2 to 5.1, all with intervals above zero. Persistence lost 10.5 and the Ridge model 3.5, while climatology, which uses only earlier years, was unchanged. No added input group improved the embargoed forecasts. The weather, fire and trajectory layers changed MAE by less than 1, with intervals that include zero, and the mass-balance layer made it slightly but reliably worse (+1.3, CI 0.4 to 2.2). The full model still beat persistence clearly (−13.8, CI −23.1 to −6.3; −28.0 in October to January), but it was no better than climatology (+0.7, CI −2.8 to +3.8) or the Ridge model (−0.2, CI −5.5 to +3.5). The best embargoed models were the Random Forests on $PM_{2.5}$ history, with or without trajectories (MAE 30.5 and 30.3), barely better than climatology (31.5). With realistic data delays, the extra inputs therefore add nothing measurable, and a week-old $PM_{2.5}$ record is only slightly more useful than the seasonal average. The per-source latency rule was not run (Section 5.6).

6.6 Incremental benefit of transport features

Table 9 shows the change in MAE as each feature group was added, with 95% intervals from the same 4-week block bootstrap. Negative values mean the larger model had the lower error. Only one added group reduced MAE by an amount whose interval excluded zero: the fire features lowered MAE by 0.7 over all weeks (CI −1.9 to −0.01). That interval only just excludes zero, the gain disappeared in October to January (−0.2, CI −1.8 to +1.8), and the next layer (mass balance) raised MAE by a similar amount; with this many comparisons, one narrow result is weak evidence. The trajectory layer, whose comparison with model 4 was declared as primary before it was computed, lowered MAE by 0.4 (CI −1.3 to +0.8), and the result was the

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

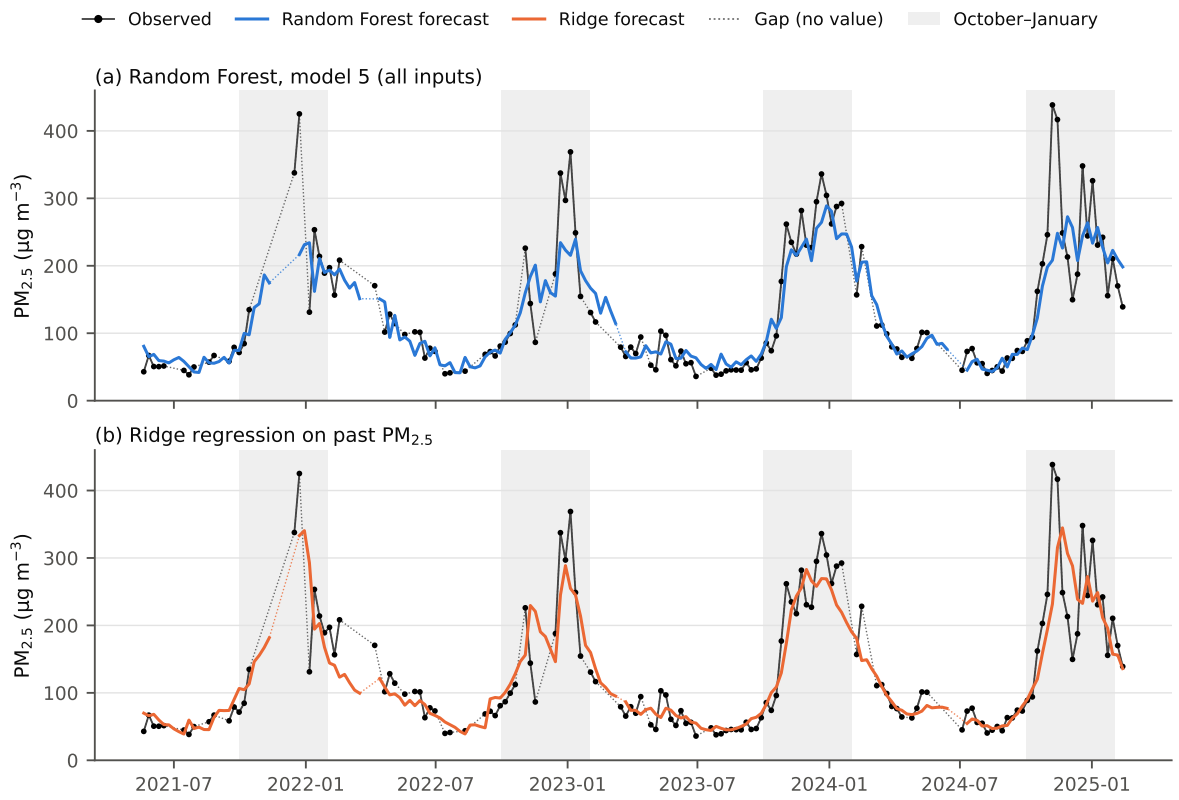


Figure 11

Observed weekly mean $PM_{2.5}$ (black dots) and one-week-ahead forecasts from (a) the full Random Forest (model 5) and (b) the Ridge $PM_{2.5}$-history model, 17 May 2021 to 10 February 2025. The figure shows every eligible observed week (142) and every forecast each model made, whereas the error scores use only the 138 weeks with a forecast from every method. Solid lines join consecutive weeks; faint dotted lines cross weeks with no value (no eligible observation or no forecast), which are not filled in. Grey bands mark October to January. Both models miss the sharpest peaks, such as November 2024. Walk-forward forecasts with no data embargo; preliminary AirNow data.

same with 2-week and 8-week blocks. Added to $PM_{2.5}$ history alone, the trajectory features lowered MAE by 1.1 (CI −2.8 to +0.6): 89% of bootstrap resamples favoured them, but the interval still includes zero, so the result is inconclusive. In the full model, much of the same information may already be carried by the ERA5 wind and ventilation features. Figure 13 shows these intervals together with the comparisons of the full model against the three baselines. Under the one-week embargo, no added group reduced MAE reliably (Table 8). The trajectory layer changed it by −0.6 (CI −1.2 to +0.2), trajectories added to $PM_{2.5}$ history alone by −0.1 (CI −1.0 to +1.0), and the fire layer's narrow gain disappeared (+0.3, CI −0.7 to +1.0).

6.7 SHAP analysis

SHAP values were computed for each of the 232 walk-forward forecasts of the full model, using the model trained for that week, so every explanation is out-of-sample (Figure 14). Over the 138 common weeks, the model relied most on $PM_{2.5}$ history (38.0% of total mean absolute SHAP) and the calendar terms that represent the season (25.9%), followed by ERA5 meteorology (15.6%), mass-balance features (8.6%), fire features (8.3%) and trajectory features (3.7%).

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

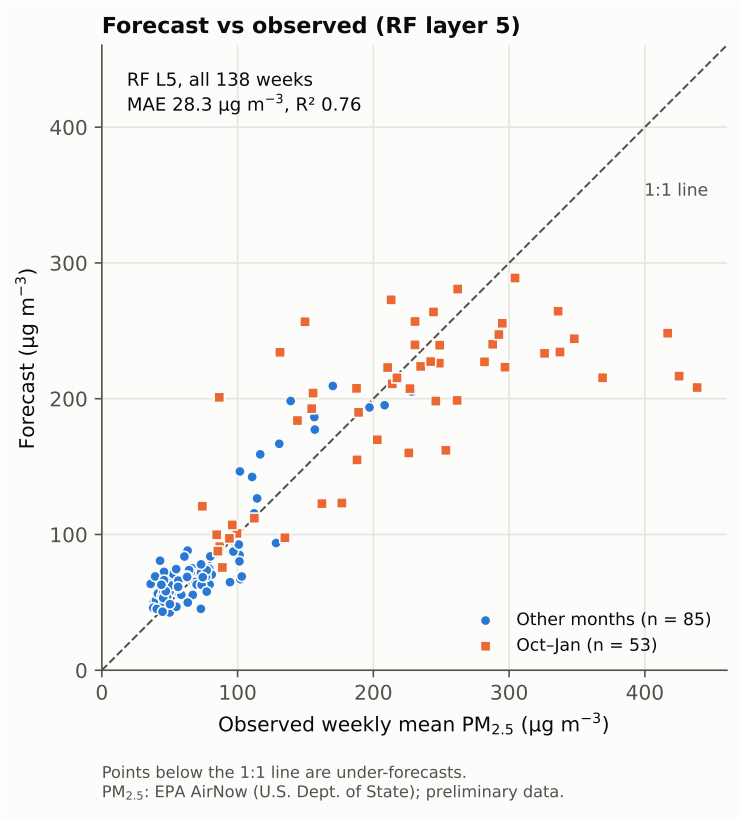


Figure 12

Forecast against observed weekly mean $PM_{2.5}$ for the full Random Forest (MAE 28.3 µg m^{-3}, R^2 0.76). Points below the dashed 1:1 line are under-forecasts; most of the highest October-to-January weeks fall below it.

Table 8

Forecast skill with a one-week data embargo, on the 136 weeks (52 in October to January) with an embargoed forecast from every method. MAE and bias are in µg m^{-3}. The last column is each method's MAE with the embargo minus its MAE without it, on all weeks where both versions made a forecast (142 to 237 weeks), with 95% moving-block bootstrap intervals. Bold marks the lowest MAE in each column. Preliminary AirNow data.

Method	MAE	R^2	Bias	MAE Oct–Jan	Change from no embargo
Persistence	46.1	0.37	+1.6	83.5	+10.5 (+5.8 to +15.6)
Seasonal climatology	31.5	0.74	−9.8	**52.4**	0.0
Ridge, $PM_{2.5}$ history	32.4	0.69	−6.6	59.5	+3.5 (+1.2 to +5.7)
RF 1: $PM_{2.5}$ history	30.5	0.72	−10.1	55.1	+3.2 (+1.1 to +5.6)
RF 2: + ERA5 meteorology	31.2	0.71	−9.7	54.1	+3.7 (+1.6 to +6.0)
RF 3: + fire	31.5	0.71	−10.0	54.6	+4.6 (+2.6 to +6.8)
RF 4: + mass balance	32.8	0.70	−9.6	55.7	+5.1 (+2.7 to +7.7)
RF 5: + HYSPLIT trajectories (full model)	32.2	0.71	−9.9	55.5	+5.1 (+2.3 to +8.3)
RF 1 + trajectories	**30.3**	0.72	−10.9	54.7	+4.5 (+1.8 to +7.5)

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Table 9

Change in MAE (μg m^{-3}) when each feature group is added, with 95% moving-block bootstrap intervals (4-week blocks).

Feature group added	**All weeks** (n = 138)	**October to January** (n = 53)
ERA5 meteorology (model 2 vs model 1)	−0.3 (−2.2 to +1.3)	−2.0 (−5.8 to +1.7)
Fire (model 3 vs model 2)	−0.7 (−1.9 to −0.01)	−0.2 (−1.8 to +1.8)
Mass balance (model 4 vs model 3)	+0.6 (−0.04 to +1.2)	+0.6 (−1.0 to +1.5)
HYSPLIT trajectories (model 5 vs model 4)	−0.4 (−1.3 to +0.8)	−0.3 (−2.8 to +1.9)
Trajectories added to $PM_{2.5}$ history only (model 1 + trajectories vs model 1)	−1.1 (−2.8 to +0.6)	−1.7 (−4.9 to +2.5)

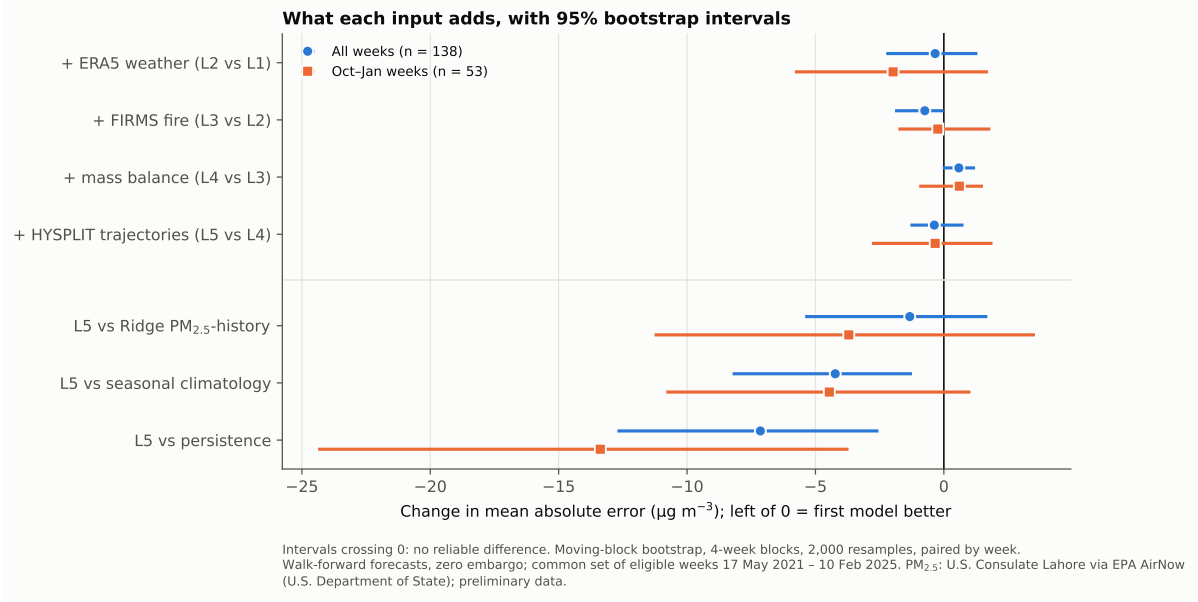


Figure 13

Change in mean absolute error when each input group is added (top) and for the full model against the three baselines (bottom), with 95% moving-block bootstrap intervals (4-week blocks, 2,000 resamples). Negative values favour the first model named; intervals that cross zero indicate no reliable difference.

In October to January the trajectory share rose to 4.9%, mostly through mean trajectory height. After $PM_{2.5}$ history and season, the most-used single inputs were hours of low ventilation, boundary-layer height and surface pressure, all of which describe stagnant conditions. Several inputs are strongly correlated, so SHAP may divide credit among them unpredictably, and group totals are more reliable than individual ranks. These values describe what the model relied on to make its forecasts. They are not estimates of how much each factor contributes to $PM_{2.5}$ in Lahore. Reliance is also not the same as benefit: the model used the trajectory features, but they did not reliably improve its accuracy, and the fire features gave only a narrow gain (Section 6.6). The SHAP analysis was repeated in the embargo run and the four data-

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

processing runs (Section 6.8), and the family ranking met the pre-specified stability criterion in four of the five. With AirNow daily values, 7-of-7 weeks or the stricter run rule, no family's share changed by more than 1.6 percentage points. Under the embargo, the calendar became the largest family (37.1%) and $PM_{2.5}$ history fell to 19.1%. The two largest families were still the same, and the rank correlation with the main run was 0.89. Without 2020 and 2021, however, ERA5 weather became the largest family (41.2%) and $PM_{2.5}$ history fell to 17.8%. The rank correlation was 0.77, so the criterion was not met. Trajectories were the smallest family in every run (1.7 to 3.7%). H5 is therefore only partly supported: which inputs the model relies on depends on the years it is trained and tested on.

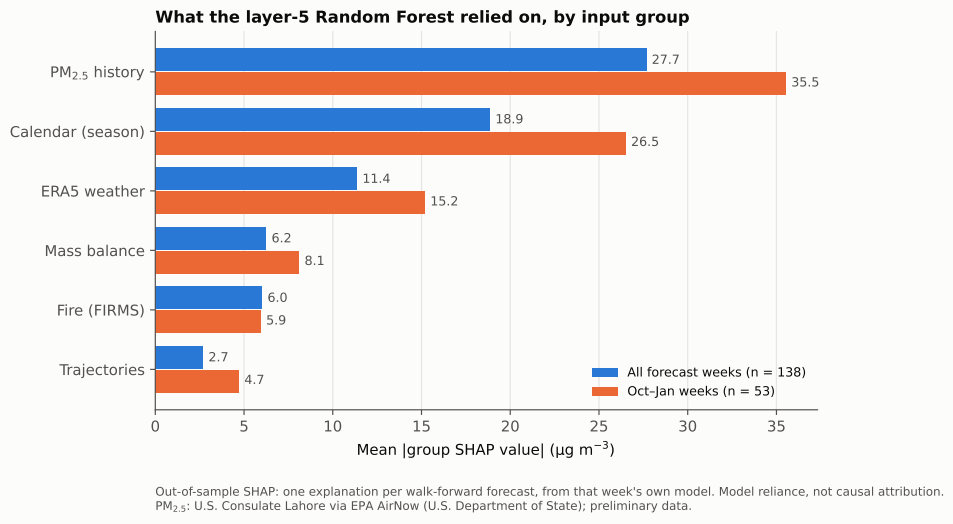


Figure 14

Mean absolute SHAP value of each input group in the full Random Forest, for all 138 forecast weeks and for the 53 October-to-January weeks. Each forecast is explained by the model trained for that week. SHAP values show what the model relied on, not what causes $PM_{2.5}$.

6.8 Sensitivity analysis

The first two checks concern the forecasting set-up. With 4 or 8 trajectory clusters instead of 6, the full model's MAE was 28.3 and 28.2 µg m^{-3}, against 28.3 with 6 clusters. With trajectory features from 100 m or 250 m instead of 500 m, the full model's MAE was 28.3 and 28.4, and the trajectory layer again gave no reliable gain (−0.4, CI −1.4 to +0.7; and −0.3, CI −1.3 to +0.9). Changing the bootstrap block length from 4 weeks to 2 or 8 weeks left the trajectory-layer comparison unchanged: −0.4 µg m^{-3}, with 95% intervals of −1.5 to +0.7 (2-week blocks) and −1.1 to +0.6 (8-week blocks). The follow-up Random Forest on log $PM_{2.5}$, declared to test whether the under-prediction of high weeks could be reduced, barely changed it: its October-to-January bias was −25.5, against −25.9 for the comparable model 4. Four further runs each changed one data-processing choice (Table 10). Using AirNow's own daily values, requiring all 7 days in a week, or a 6-hour repeated-value rule left every conclusion unchanged: the trajectory layer gave no reliable gain, the full model beat persistence, and it was not reliably better than the Ridge model. The 6-hour rule removed 15 hours instead of 9 and changed no week's eligibility. Excluding 2020 and 2021 left 99 common weeks, all from 2022 onwards, and raised every method's MAE (full model 35.5, persistence 37.8). The trajectory and Ridge comparisons were unchanged, but the full model's advantage over persistence was no longer

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

reliable (−2.3, CI −9.9 to +4.0). The main forecasting conclusions therefore hold under the data-processing choices tested, but the advantage over persistence depends partly on the 2020 and 2021 weeks.

Table 10

Sensitivity runs, each changing one data-processing choice (no embargo). Differences in MAE ($\mu g\, m^{-3}$) are full model minus the comparison, with 95% moving-block bootstrap intervals (4-week blocks); each run is scored on its own common weeks. A run agrees with the main result if the first and second intervals include zero and the third excludes it (criteria fixed in advance).

Run	Weeks	MAE	vs model 4	vs Ridge	vs persistence	Agrees
Main analysis	138	28.3	−0.4 (−1.3 to +0.8)	−1.3 (−5.4 to +1.7)	−7.1 (−12.7 to −2.5)	–
AirNow daily values	131	29.1	+0.1 (−0.8 to +1.3)	−0.8 (−5.3 to +2.2)	−5.9 (−11.8 to −0.6)	Yes
All 7 days valid	124	29.0	−0.3 (−1.4 to +0.9)	−2.2 (−6.7 to +0.9)	−7.0 (−13.3 to −2.0)	Yes
6-hour repeated-value rule	138	28.4	−0.4 (−1.3 to +0.8)	−1.3 (−5.4 to +1.7)	−7.1 (−12.7 to −2.5)	Yes
2020 and 2021 excluded	99	35.5	−0.4 (−1.1 to +0.4)	+0.3 (−4.3 to +3.8)	−2.3 (−9.9 to +4.0)	No

6.9 Diagnostic experiment (not a forecast)

To see how much of the remaining error better weather forecasts could remove, the Random Forest was refitted with the ERA5, fire, mass-balance and trajectory features of the target week itself, keeping $PM_{2.5}$ history as in the forecasts (Section 5.6). This is a diagnostic, not a forecast, because these conditions are not known when a forecast is issued. It was compared with the forecasting model and persistence on the 182 weeks where all three had predictions (June 2020 to February 2025). It had an MAE of 24.0, against 29.9 for the forecasting model on the same weeks (difference −5.9, 95% CI −9.7 to −2.2) and 36.7 for persistence. In October to January the gain was larger (41.1 against 51.0; −9.9, CI −18.1 to −2.8), but the error was still large and the forecasts were still too low on average (bias −16.0). Target-week trajectories added nothing beyond target-week weather (MAE 24.0 with them, 23.8 without). Perfect knowledge of the coming week's conditions would therefore remove about a fifth of the forecast error, which bounds what weather forecasts could add to this model.

6.10 Comparison with sharper trajectory meteorology (one year)

All trajectory results above use NCEP/NCAR meteorology on a 2.5° grid. To test whether finer meteorology would change them, one year (arrivals from 1 March 2024 to 28 February 2025, including the November 2024 episode) was recomputed with NOAA's GDAS 1° meteorology. Everything else was unchanged: HYSPLIT v5.4.2, 72 hours, arrival at 06:00 UTC, and heights of 100, 250 and 500 m. All 1,095 trajectories ran without failure, and all 61 weekly GDAS files matched NOAA's published checksums. The comparison and a decision rule were declared before any GDAS file was downloaded. The finer meteorology would count as changing the results substantially if any of three conditions held at 500 m: fewer than 80% of days assigned to the same cluster (nearest of the cluster centres in Section 6.2); a rank correlation of the 72-hour transport distance below 0.8; or any of five key estimates moving outside the 2.5° estimate's 95% interval or to the other side of no effect. The GDAS trajectories were used only for this comparison and were not mixed into the main analysis.

The two meteorologies gave clearly different paths. At 500 m, only 70% of days fell in the same cluster, and the rank correlation of transport distance was 0.22. After 24 hours the two versions of a trajectory were typically 254 km apart (interquartile range 157 to 363 km), and after 72 hours 713 km apart, about as far as the air had travelled. The GDAS trajectories stayed

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

lower (mean endpoint height 1,215 m against 1,817 m) and, in winter and October to November, travelled shorter distances (winter median 481 against 607 km). Agreement was no better at 100 m and 250 m (65% and 70% of days in the same cluster; rank correlation 0.15 at both).

Three of the five estimates moved (Table 11). The fire-exposure coefficient was almost unchanged. In both seasons the ratio between clusters 2 and 1 fell below 1 with GDAS, but every interval included 1, so neither version shows a clear difference. The clearest change was in winter. With 2.5° meteorology, slow air was again no more polluted than fast air (ratio 0.94). With GDAS 1°, days in the shortest third of transport distances had about 35% higher geometric mean $PM_{2.5}$ than days in the longest third (1.35, 95% CI 1.04 to 1.53). All three pre-declared conditions therefore held: in this year, the finer meteorology changed the trajectory results substantially. Only one year was recomputed, so these estimates are imprecise, and whether the same differences would appear over the whole 2019 to 2025 record is not known. The comparison shows that the trajectory-based results in Sections 6.2 to 6.4 depend on the meteorology used, and that with finer meteorology at least one winter showed the stagnation association that the 2.5° trajectories did not. This is an association in one season, not evidence about where the pollution came from.

Table 11

One-year comparison of trajectory meteorology (arrivals 1 March 2024 to 28 February 2025, 500 m). Ratios of geometric mean $PM_{2.5}$ (no effect = 1) and the fire-exposure coefficient c (no effect = 0), each with a 95% moving-block bootstrap interval (7-day blocks, 2,000 resamples). The 2.5° estimates use the cluster labels and fire exposure of Sections 6.2 to 6.4; the GDAS estimates use GDAS trajectories assigned to the same cluster centres. Winter is December 2024 to February 2025; post-monsoon is October to November 2024.

Estimate	NCEP/NCAR 2.5°	GDAS 1°	Outside 2.5° interval	Other side of no effect
Cluster 2 / cluster 1, winter	1.07 (0.80 to 1.40)	0.82 (0.70 to 1.03)	No	Yes
Cluster 2 / cluster 1, post-monsoon	1.17 (0.72 to 2.13)	0.93 (0.62 to 1.19)	No	Yes
Shortest / longest distance third, winter	0.94 (0.73 to 1.26)	1.35 (1.04 to 1.53)	Yes	Yes
Shortest / longest distance third, post-monsoon	0.82 (0.46 to 1.39)	0.93 (0.76 to 1.29)	No	No
Fire coefficient *c*, October to November	0.144 (–0.071 to 0.218)	0.133 (–0.101 to 0.188)	No	No

7. Discussion

7.1 Main findings

The association analysis answers the first two parts of the research question. Three recurring airmass pathways reach Lahore, most often slow air from the south-east. $PM_{2.5}$ differs modestly between them within seasons, but these differences largely disappear once weather is taken into account. Ventilation, which describes how well air near the ground is mixed and moved, was the strongest and most consistent association: in winter and October to November, a 10% drop in ventilation went with about 5% higher $PM_{2.5}$, whatever the pathway. Fire activity along the air's path was positively but modestly associated with October and November $PM_{2.5}$, beyond what their shared seasonal timing produces, although this comparison with shifted fire records weakened when trajectories arriving at 100 or 250 m were used.

The forecasting results answer the third part of the research question. A Random Forest using $PM_{2.5}$ history, ERA5 weather, fire, mass-balance and trajectory features forecast weekly mean $PM_{2.5}$ with an MAE of 28.3 $\mu g\,m^{-3}$. That is clearly better than persistence and seasonal climatology, but not reliably better than a simple regression on past $PM_{2.5}$. Each added input group changed the error by less than 1 $\mu g\,m^{-3}$, and only the fire features gave a gain narrowly

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

larger than chance; the model relied mostly on recent $PM_{2.5}$ and the time of year, and every model under-predicted the worst smog weeks. With a one-week delay in the data, every model lost 3 to 5 $\mu g\,m^{-3}$ of accuracy, the extra inputs added nothing measurable, and the full model was no better than climatology. A second finding concerns the data: a widely used secondary copy of the Lahore diplomatic monitor's record contains timing misalignments and at least one long block of values that do not exist in the original source. These problems leave daily and weekly means largely intact but affect hourly analysis and completeness counts. Any study that uses hourly OpenAQ values for this monitor, including trajectory matching at hourly resolution, should check them against the original archive.

7.2 Transport versus local accumulation

The results point more to local accumulation than to long-range transport as the main association with high $PM_{2.5}$, with fire smoke as a smaller, additional signal in October and November. Three findings support this. First, ventilation, a local measure, was the clearest association in the polluted seasons, and its link with $PM_{2.5}$ did not depend on the pathway. Second, the differences between pathways vanished once weather was accounted for. Third, slow trajectories were not more polluted. This suggests that a 72-hour trajectory arriving at 500 m on a 2.5° grid describes the regional flow above Lahore rather than the stagnant air near the ground where $PM_{2.5}$ accumulates. In winter, 500 m may lie above a shallow mixed layer, and ERA5 supports this. A descriptive check, declared in advance, found that at the trajectories' 11:00 PKT arrival time the boundary layer was below 500 m on 78% of winter days (median height 368 m). This rose to 92% on the most polluted third of winter days, against 62% on the cleanest third. By contrast, the boundary layer was below 500 m on 21% of post-monsoon days and on fewer than 4% of pre-monsoon and monsoon days. On most winter days, especially polluted ones, the 500 m trajectories therefore describe the flow above the layer the monitor samples. Recomputing the trajectories to arrive at 100 and 250 m, inside ERA5's boundary layer on most winter days, did not change the result. Slow air was still no more polluted than fast air: the winter ratio of shortest to longest transport distance was 0.90 (95% CI 0.75 to 1.03) at 100 m and 0.88 at 250 m, against 0.86 at 500 m. The arrival height therefore does not explain the result, at least with 2.5° meteorology. On that grid the three heights largely follow the same synoptic flow: the rank correlation of their transport distances with those at 500 m was 0.85 to 0.95. The meteorology itself may matter more. When one year was recomputed with GDAS 1° meteorology (Section 6.10), the paths were lower and, in winter, shorter, and in that winter the days with the shortest paths had about 35% higher $PM_{2.5}$ than those with the longest (95% CI 1.04 to 1.53), an association the 2.5° trajectories did not show. The finding that slow air was no more polluted therefore rests on 2.5° meteorology and may not hold with finer data; this was not tested beyond one year. Fire exposure did add information in October and November, but its effect was modest (about 6% per doubling of exposure), and part of the raw link between fires and $PM_{2.5}$ reflects their shared timing. With lower trajectories, fires at the same time of season in other years gave links almost as strong. None of this separates transported from local $PM_{2.5}$ at the monitor: stagnant conditions can both trap local emissions and coincide with slow transport from upwind fires, as Dahal et al. (2026) describe for November 2024. Separating the two would need chemical tracers or a transport model.

7.3 Comparison with published literature

The pathway results agree with earlier Lahore work that found local emissions and regional transport from the east most important (Karim & Rappenglück, 2023), and with Majeed et al. (2024), who gave crop fires moderate importance relative to vehicle emissions. The modest fire

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

signal in October and November fits the timing of crop burning and the episode-scale transport documented for November 2024 (Dahal et al., 2026), and it is consistent with the conflicting and mostly unquantified evidence on the fire share of Lahore's $PM_{2.5}$ (Section 2.2). Across the Indo-Gangetic Plain, stagnation days carry about 45 µg m^{-3} more $PM_{2.5}$ in winter (Zhou et al., 2024); the strong ventilation association found here is the Lahore counterpart of that result, measured with ventilation rather than trajectory speed.

The forecast results can already be placed in context, although only qualitatively, because the published Lahore studies predict daily values from different data and with different validation designs (Iftikhar et al., 2024; Qureshi et al., 2025; Haseeb et al., 2026). Under-predicting the highest episodes is common even for physics-based systems: Delhi's high-resolution (400 m) $PM_{2.5}$ forecasting system struggled with episodes above 350 µg m^{-3} in an earlier version (Jena et al., 2021). Small and uncertain gains from fire and trajectory inputs are also consistent with Cusworth et al. (2018), whose fire-transport model reproduced at most 25% of the weekly variability of $PM_{2.5}$ in Delhi.

7.4 Forecasting implications

The monitor has not reported since February 2025, and no currently reporting reference-grade replacement in Lahore has been verified. Any model developed here therefore supports retrospective (hindcast) evaluation only. The embargo results also show that the forecasts depend on fresh data: a one-week delay in the inputs raised errors by 3 to 5 µg m^{-3}, so an operational system would need $PM_{2.5}$ and weather data within days. Operational use would need a new data source with a demonstrated overlap period, so that differences in bias and scale could be measured. A standard Random Forest also does not learn from new data on its own: the surrounding code would have to store new observations and retrain the model on a documented schedule, which is planned only after the forecasting workflow has been evaluated.

7.5 Limitations

- One monitor represents the city. Its readings mix citywide and regional conditions with sources close to the site, and how well it represents Lahore as a whole has not been established, so results describe this site rather than the whole city.
- The data are preliminary AirNow data, not fully validated; some periods (notably 2022) are poorly covered.
- The weekly series is short (239 eligible weeks in the rebuilt record), which limits model complexity and the precision of skill estimates, especially for the smog season.
- Trajectories carry uncertainty that grows with travel time and depends on the meteorological data, and cluster labels simplify continuous variation. They use 2.5° meteorology, one arrival height (500 m) and one arrival time (06:00 UTC). On stagnant days the direction of a short trajectory is poorly defined, so claims about which side of the border air came from cannot be supported at this resolution. Transport distance at 500 m also did not capture surface stagnation. In winter, the 500 m arrival point lay above ERA5's boundary layer on most days (Section 7.2), so winter trajectories describe the flow above the surface layer; trajectories arriving at 100 and 250 m, computed to test this, gave the same results (Section 7.2), but all three heights use the same 2.5° meteorology. When one year was recomputed with GDAS 1° meteorology, the paths differed substantially (only 70% of days in the same cluster; rank correlation of transport distance 0.22), and three of five key estimates moved, including the winter stagnation result (Section 6.10). The pathway and stagnation results therefore depend on the trajectory meteorology, and the whole record was not recomputed with the finer data. The fire-exposure association

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

did not change in that year. Only one clustering method was used.

- The main forecasts assume that inputs are available immediately. With a one-week embargo, errors rose by 3 to 5 $\mu g\,m^{-3}$ and the full model was no better than climatology (Section 6.5); the per-source latency rule was not run. The advantage over persistence also became unreliable when 2020 and 2021 were excluded (Section 6.8).
- Satellite fire detections miss fires obscured by cloud or thick smoke and burning between overpasses, and do not identify the burning material. Because the smokiest days are among those most likely to be under-counted, fire exposure may be biased low when $PM_{2.5}$ is highest, which would tend to weaken associations rather than create them. Burning that has shifted to the evening, after the early-afternoon VIIRS overpass, is undercounted (Dahal et al., 2026), so if the timing of burning changed during 2019 to 2025, fire counts may have changed for reasons unrelated to the amount burned.
- ERA5 boundary-layer height is a model estimate on a grid of about 31 km (Hersbach et al., 2020), so one grid cell covers much of the city and its surroundings and does not resolve urban effects on the boundary layer. Its accuracy over Lahore has not been checked, and a comparison of the monitor's grid cell with an average of neighbouring cells could not be made with the downloaded data (Section 5.11). Boundary-layer height is also not purely a cause of high $PM_{2.5}$: heavy aerosol loading can itself lower the boundary layer (Miao et al., 2019), so adjusting for it may partly remove a real transport effect in the association analysis.
- The box model's emission, inflow and removal terms are unobserved, so it cannot apportion sources.
- SHAP describes model behaviour, not causation.
- Data terms: the EPA AirNow program was notified of this manuscript on 27 October 2026, as its guidelines request; whether the U.S. Department of State should be notified separately has been asked but not yet answered, and the OpenAQ terms on download authorisation have not yet been resolved.

7.6 Future research

Possible extensions include recomputing the whole trajectory record with GDAS 1° meteorology, which changed the pathway and stagnation results in the one year tested; using archived forecast meteorology to make trajectory features genuine forecast inputs; adding geostationary fire products to capture evening burning; comparing several Lahore monitors once suitable records exist; and validating the box-model features against observed vertical profiles where available.

8. Conclusion

This study asks which incoming air-mass pathways and weather conditions go with high $PM_{2.5}$ in Lahore, and whether that information improves week-ahead forecasts. Three recurring pathways reach the city, and $PM_{2.5}$ differs modestly between them, but those differences largely disappear once the weather that accompanies each pathway is taken into account. Poor ventilation was the clearest and most consistent association with high $PM_{2.5}$ in winter and October to November, and fire activity along the air's path added a smaller, positive association in October and November. The pathway results depend on the trajectory meteorology: in one year recomputed with finer (1°) meteorology, the paths and several pathway results changed, while the fire association did not. These are associations, not measures of how much $PM_{2.5}$ is local or transported. For forecasting, a Random Forest beat persistence and seasonal climatology under

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

a strict walk-forward design but not a simple model of past $PM_{2.5}$, and adding ERA5 weather, satellite fire activity, ventilation features and HYSPLIT trajectories changed the error by less than 1 $\mu g\,m^{-3}$ each, with only the fire features giving a gain narrowly larger than chance. With a one-week delay in the data, every model lost accuracy and the extra inputs added nothing measurable. Building the data also showed that the source and checking of a $PM_{2.5}$ record matter: the hourly record was rebuilt from the original AirNow archive because a widely used copy contains timing errors and values that are not in the original.

Declarations

Acknowledgements: $PM_{2.5}$ data were collected by the U.S. Department of State and distributed through the US EPA AirNow program and OpenAQ. AirNow data are preliminary and have not been fully verified or validated.

Author contributions (CRediT): Mian Shaayan Akbar: Conceptualization, Software, Data curation, Validation, Resources, Writing – original draft. **Sardar Zayaan Akhtar Nakai:** Conceptualization, Methodology, Software, Investigation, Formal analysis, Visualization, Project administration, Writing – review & editing.

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The authors live in Lahore and, like its other residents, are exposed to its air pollution; this motivated the study but gives them no financial or other interest in its results.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Data availability: $PM_{2.5}$ data come from the public AirNow file archive (U.S. Consulate Lahore, AQS 586PK3010003; preliminary data from the U.S. Department of State and the EPA AirNow program), with the OpenAQ archive (location 8664) used only for comparison. ERA5 data are available from the Copernicus Climate Data Store, fire detections from NASA FIRMS, and the NCEP/NCAR Reanalysis and GDAS meteorology from NOAA's Air Resources Laboratory. The processed $PM_{2.5}$ series, trajectory endpoints, derived features, result tables and download manifests with file sizes and checksums are available at `https://github.com/coolagegoon-hub/lahore-air-quality`.

Code availability: The Python code that reproduces every table and figure, with fixed package versions and random seeds, is available at `https://github.com/coolagegoon-hub/lahore-air-quality`.

Data terms: AirNow's data-use guidelines ask that analyses state that the data are preliminary, credit the U.S. Department of State and the EPA AirNow program, and notify the program of publications; the AirNow program was notified of this manuscript on 27 October 2026, and its signed guidelines form has not been returned. OpenAQ requires attribution, and whether anonymous archive download counts as authorised use has not yet been confirmed.

References

Ahmad, M., Cheng, S., Yu, Q., Qin, W., Zhang, Y., & Chen, J. (2020). Chemical and source characterization of PM2.5 in summertime in severely polluted Lahore, Pakistan. *Atmospheric Research*, *234*, 104715. doi: 10.1016/j.atmosres.2019.104715

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, *191*, 192–213. doi: 10.1016/j.ins.2011.12.028

Bhatti, U. A., Yan, Y., Zhou, M., Ali, S., Hussain, A., Qingsong, H., … Yuan, L. (2021). Time Series Analysis and Forecasting of Air Pollution Particulate Matter ($PM_{2.5}$): An

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

SARIMA and Factor Analysis Approach. *IEEE Access*, *9*, 41019–41031. doi: 10.1109/access.2021.3060744

Bilal, M., Hassan, M., Tahir, D. B. T., Iqbal, M. S., & Shahid, I. (2022). Understanding the role of atmospheric circulations and dispersion of air pollution associated with extreme smog events over South Asian megacity. *Environmental Monitoring and Assessment*, *194*(2). doi: 10.1007/s10661-021-09674-y

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. doi: 10.1023/a:1010933404324

Chianese, E., & Riccio, A. (2026). Air-Quality Forecasting Across Monitoring, Predictor, and Validation Regimes: A Systematic Mapping Review and Decision Framework. *Forecasting*, *8*(4), 72. doi: 10.3390/forecast8040072

Copernicus Climate Change Service. (2025). *ERA5 hourly time-series data on single levels from 1940 to present.* ECMWF. doi: 10.24381/1CF1AD76

Cui, L., Song, X., & Zhong, G. (2021). Comparative Analysis of Three Methods for HYSPLIT Atmospheric Trajectories Clustering. *Atmosphere*, *12*(6), 698. doi: 10.3390/atmos12060698

Cusworth, D. H., Mickley, L. J., Sulprizio, M. P., Liu, T., Marlier, M. E., DeFries, R. S., … Gupta, P. (2018). Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India. *Environmental Research Letters*, *13*(4), 044018. doi: 10.1088/1748-9326/aab303

Dahal, S., Sharma, N., Dey, S., Bokhari, S. A. A., Boomsma, J., & Holmes, H. A. (2026). Meteorological modulation of ambient air quality in Delhi and Lahore during the November 2024 haze event. *Scientific Reports*. doi: 10.1038/s41598-026-63281-w

Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., & Schwartz, J. (2016). Assessing $PM_{2.5}$ Exposures with High Spatiotemporal Resolution across the Continental United States. *Environmental Science & Technology*, *50*(9), 4712–4721. doi: 10.1021/acs.est.5b06121

Duan, N. (1983). Smearing Estimate: A Nonparametric Retransformation Method. *Journal of the American Statistical Association*, *78*(383), 605–610. doi: 10.1080/01621459.1983.10478017

Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L., & Wang, J. (2015). Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. *Atmospheric Environment*, *107*, 118–128. doi: 10.1016/j.atmosenv.2015.02.030

Fleming, Z. L., Monks, P. S., & Manning, A. J. (2012). Review: Untangling the influence of airmass history in interpreting observed atmospheric composition. *Atmospheric Research*, *104-105*, 1–39. doi: 10.1016/j.atmosres.2011.09.009

Gautam, R., Patel, P. N., Singh, M. K., Liu, T., Mickley, L. J., Jethva, H., & DeFries, R. S. (2023). Extreme Smog Challenge of India Intensified by Increasing Lower Tropospheric Stability. *Geophysical Research Letters*, *50*(11). doi: 10.1029/2023gl103105

Ghosh, S., Biswas, J., Guttikunda, S., Roychowdhury, S., & Nayak, M. (2014). An investigation of potential regional and local source regions affecting fine particulate matter concentrations in Delhi, India. *Journal of the Air & Waste Management Association*, *65*(2), 218–231. doi: 10.1080/10962247.2014.982772

Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, *178*, 31–41. doi: 10.1016/j.rse.2016.02.054

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Grange, S. K., Carslaw, D. C., Lewis, A. C., Boleti, E., & Hueglin, C. (2018). Random forest meteorological normalisation models for Swiss PM $_{10}$ trend analysis. *Atmospheric Chemistry and Physics*, *18*(9), 6223–6239. doi: 10.5194/acp-18-6223-2018

Guttikunda, S. K., & Gurjar, B. R. (2011). Role of meteorology in seasonality of air pollution in megacity Delhi, India. *Environmental Monitoring and Assessment*, *184*(5), 3199–3211. doi: 10.1007/s10661-011-2182-8

Haseeb, M., Tahir, Z., Mehmood, S. A., Arif, H., Kousar, S., Ghafoor, S., & Mehmood, K. (2026). Ensemble Artificial Intelligence Fusing Satellite, Reanalysis, and Ground Observations for Improved PM2.5 Prediction. *Atmosphere*, *17*(4), 411. doi: 10.3390/atmos17040411

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679–688. doi: 10.1016/j.ijforecast.2006.03.001

Iftikhar, H., Qureshi, M., Zywiołek, J., López-Gonzales, J. L., & Albalawi, O. (2024). Short-term PM2.5 forecasting using a unique ensemble technique for proactive environmental management initiatives. *Frontiers in Environmental Science*, *12*. doi: 10.3389/fenvs.2024.1442644

Jacob, D. J. (1999). *Introduction to Atmospheric Chemistry*. Princeton, NJ: Princeton University Press.

Jena, C., Ghude, S. D., Kumar, R., Debnath, S., Govardhan, G., Soni, V. K., … Rajeevan, M. (2021). Performance of high resolution (400 m) PM2.5 forecast over Delhi. *Scientific Reports*, *11*(1). doi: 10.1038/s41598-021-83467-8

Jethva, H., Chand, D., Torres, O., Gupta, P., Lyapustin, A., & Patadia, F. (2018). Agricultural Burning and Air Quality over Northern India: A Synergistic Analysis using NASA's A-train Satellite Data and Ground Measurements. *Aerosol and Air Quality Research*, *18*(7), 1756–1773. doi: 10.4209/aaqr.2017.12.0583

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., … Joseph, D. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. doi: 10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2

Karim, I., & Rappenglück, B. (2023). Impact of Covid-19 lockdown regulations on PM2.5 and trace gases (NO2, SO2, CH4, HCHO, C2H2O2 and O3) over Lahore, Pakistan. *Atmospheric Environment*, *303*, 119746. doi: 10.1016/j.atmosenv.2023.119746

Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., … Kumar, V. (2017). Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. *IEEE Transactions on Knowledge and Data Engineering*, *29*(10), 2318–2331. doi: 10.1109/tkde.2017.2720168

Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data*, *6*(4), 1–21. doi: 10.1145/2382577.2382579

Khanum, F., Chaudhry, M. N., & Kumar, P. (2017). Characterization of five-year observation data of fine particulate matter in the metropolitan area of Lahore. *Air Quality, Atmosphere & Health*, *10*(6), 725–736. doi: 10.1007/s11869-017-0464-1

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Kulkarni, S. H., Ghude, S. D., Jena, C., Karumuri, R. K., Sinha, B., Sinha, V., … Khare, M. (2020). How Much Does Large-Scale Crop Residue Burning Affect the Air Quality in Delhi? *Environmental Science & Technology*, *54*(8), 4790–4799. doi: 10.1021/acs.est.0c00329

Le, N. H., Ly, B.-T., Thai, P. K., Pham, G.-H., Ngo, I.-H., Do, V.-N., … Vu, T. V. (2021). Assessing the Impact of Traffic Emissions on Fine Particulate Matter and Carbon Monoxide Levels in Hanoi through COVID-19 Social Distancing Periods. *Aerosol and Air Quality Research*, *21*(10), 210081. doi: 10.4209/aaqr.210081

Liu, T., Marlier, M. E., DeFries, R. S., Westervelt, D. M., Xia, K. R., Fiore, A. M., … Milly, G. (2018). Seasonal impact of regional outdoor biomass burning on air pollution in three Indian cities: Delhi, Bengaluru, and Pune. *Atmospheric Environment*, *172*, 83–92. doi: 10.1016/j.atmosenv.2017.10.024

Lodhi, A., Ghauri, B., Khan, M. R., Rahman, S., & Shafique, S. (2009). Particulate matter (PM2.5) concentration and source apportionment in lahore. *Journal of the Brazilian Chemical Society*, *20*(10), 1811–1820. doi: 10.1590/s0103-50532009001000007

Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions.* arXiv. doi: 10.48550/ARXIV.1705.07874

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., … Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, *2*(1), 56–67. doi: 10.1038/s42256-019-0138-9

Majeed, R., Anjum, M. S., Imad-ud din, M., Malik, S., Anwar, M. N., Anwar, B., & Khokhar, M. F. (2024). Solving the mysteries of Lahore smog: the fifth season in the country. *Frontiers in Sustainable Cities*, *5*. doi: 10.3389/frsc.2023.1314426

Méndez, M., Merayo, M. G., & Núñez, M. (2023). Machine learning algorithms to forecast air quality: a survey. *Artificial Intelligence Review*, *56*(9), 10031–10066. doi: 10.1007/s10462-023-10424-4

Miao, Y., Li, J., Miao, S., Che, H., Wang, Y., Zhang, X., … Liu, S. (2019). Interaction Between Planetary Boundary Layer and PM2.5 Pollution in Megacities in China: a Review. *Current Pollution Reports*, *5*(4), 261–271. doi: 10.1007/s40726-019-00124-5

Mohyuddin, S., Alam, K., Zeb, B., Khokhar, M. F., Mir, K. A., Wexler, A. S., … Shahid, I. (2024). Characterization and source identification of PM2.5 during intense haze episodes in an urban environment of Lahore. *Atmospheric Environment: X*, *23*, 100276. doi: 10.1016/j.aeaoa.2024.100276

Nawaz, H., Tariq, S., Haq, Z. u., & Mehmood, U. (2023). Identifying the natural and anthropogenic drivers of absorbing aerosols using OMI data and HYSPLIT model over South Asia. *Air Quality, Atmosphere & Health*, *16*(12), 2553–2577. doi: 10.1007/s11869-023-01423-z

Ojha, N., Sharma, A., Kumar, M., Girach, I., Ansari, T. U., Sharma, S. K., … Gunthe, S. S. (2020). On the widespread enhancement in fine particulate matter across the Indo-Gangetic Plain towards winter. *Scientific Reports*, *10*(1). doi: 10.1038/s41598-020-62710-8

Qureshi, M., Hashem, A. F., Iftikhar, H., & Rodrigues, P. C. (2025). A Hybrid STL-Based Ensemble Model for PM2.5 Forecasting in Pakistani Cities. *Symmetry*, *17*(11), 1827. doi: 10.3390/sym17111827

Raja, S., Biswas, K. F., Husain, L., & Hopke, P. K. (2010). Source Apportionment of the Atmospheric Aerosol in Lahore, Pakistan. *Water, Air, and Soil Pollution*, *208*(1-4), 43–

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

57. doi: 10.1007/s11270-009-0148-z

Rasheed, A., Aneja, V. P., Aiyyer, A., & Rafique, U. (2015). Measurement and Analysis of Fine Particulate Matter (PM2.5) in Urban Areas of Pakistan. *Aerosol and Air Quality Research*, *15*(2), 426–439. doi: 10.4209/aaqr.2014.10.0269

Rashid, A., & Atif, R. M. (2026). Assessing meteorological influences and predicting PM2.5 using deep learning approach: A case of Asian metropolitan city, Lahore. *Atmospheric Environment*, *381*, 122110. doi: 10.1016/j.atmosenv.2026.122110

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. doi: 10.1038/s41586-019-0912-1

Rybarczyk, Y., & Zalakeviciute, R. (2018). Machine Learning Approaches for Outdoor Air Quality Modelling: A Systematic Review. *Applied Sciences*, *8*(12), 2570. doi: 10.3390/app8122570

Saharan, U. S., Kumar, R., Singh, S., Mandal, T. K., Sateesh, M., Verma, S., & Srivastava, A. (2024). Hotspot driven air pollution during crop residue burning season in the Indo-Gangetic Plain, India. *Environmental Pollution*, *350*, 124013. doi: 10.1016/j.envpol.2024.124013

Schroeder, W., Oliva, P., Giglio, L., & Csiszar, I. A. (2014). The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, *143*, 85–96. doi: 10.1016/j.rse.2013.12.008

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., & Ngan, F. (2015). NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society*, *96*(12), 2059–2077. doi: 10.1175/bams-d-14-00110.1

Stirnberg, R., Cermak, J., Kotthaus, S., Haeffelin, M., Andersen, H., Fuchs, J., … Favez, O. (2021). Meteorology-driven variability of air pollution (PM_1) revealed with explainable machine learning. *Atmospheric Chemistry and Physics*, *21*(5), 3919–3948. doi: 10.5194/acp-21-3919-2021

Stohl, A. (1998). Computation, accuracy and applications of trajectories—A review and bibliography. *Atmospheric Environment*, *32*(6), 947–966. doi: 10.1016/s1352-2310(97)00457-3

Stone, E., Schauer, J., Quraishi, T. A., & Mahmood, A. (2010). Chemical characterization and source apportionment of fine and coarse particulate matter in Lahore, Pakistan. *Atmospheric Environment*, *44*(8), 1062–1070. doi: 10.1016/j.atmosenv.2009.12.015

Su, T., Li, Z., & Kahn, R. (2018). Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors. *Atmospheric Chemistry and Physics*, *18*(21), 15921–15935. doi: 10.5194/acp-18-15921-2018

Tariq, S., Zia, u.-H., & Ali, M. (2016). Satellite and ground-based remote sensing of aerosols during intense haze event of October 2013 over lahore, Pakistan. *Asia-Pacific Journal of Atmospheric Sciences*, *52*(1), 25–33. doi: 10.1007/s13143-015-0084-3

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, *16*(4), 437–450. doi: 10.1016/s0169-2070(00)00065-0

US EPA. (2024). *Fact sheet on the revised PM2.5 Air Quality Index.* United States Environmental Protection Agency, `https://www.epa.gov/system/files/documents/`

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

2024-02/pm-naaqs-air-quality-index-fact-sheet.pdf.

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., & Baklanov, A. (2012). Real-time air quality forecasting, part I: History, techniques, and current status. *Atmospheric Environment*, *60*, 632–655. doi: 10.1016/j.atmosenv.2012.06.031

Zhou, M., Xie, Y., Wang, C., Shen, L., & Mauzerall, D. L. (2024). Impacts of current and climate induced changes in atmospheric stagnation on Indian surface PM2.5 pollution. *Nature Communications*, *15*(1). doi: 10.1038/s41467-024-51462-y

Appendix A. Literature Matrix

Access level: **Full text** means relevant sections of the article were read; **Abstract** means only the abstract and bibliographic record were checked; **Record** means only the DOI or bibliographic record was checked. All DOIs were checked against Crossref, OpenAlex or the publisher on 18 October 2026. All entries are peer-reviewed journal articles unless stated otherwise.

Reference	Access	Verified content used here	Limitation for this study
Karim and Rappenglück (2023)	Full text (sections)	US consulate Lahore monitor (BAM 1020) via AirNow; 72 h HYSPLIT 4.9, GDAS, 200 to 4,000 m AGL; clusters, residence time, CWT; local and eastern transport; $PM_{2.5}$ down about 38% in lockdown	Lockdown comparison, not forecasting; describes the site as both "consulate" and "embassy"
Mohyuddin et al. (2024)	Full text (sections)	45 $PM_{2.5}$ samples, October to December 2019; SEM-EDX, XRD, Raman; qualitative source types	No quantitative apportionment; no transport analysis; sampling site is not the diplomatic monitor
Saharan et al. (2024)	Abstract and highlights	WRF-Chem, EDGAR v5, FINN v2.4; hotspots about 80% (Haryana) and 50% (Punjab) of fire counts; hotspots about 50 to 70% of CRB-induced $PM_{2.5}$ increase	Indian domain; does not quantify Lahore; numbers from abstract, to verify before submission
Qureshi et al. (2025)	Full text (sections)	IQAir daily data 2019 to 2023; MICE; STL with ensembles; one day ahead; Lahore RMSE 12.475; no meteorology	Different target and source; training-only fitting of preprocessing not stated

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Reference	Access	Verified content used here	Limitation for this study
Cui et al. (2021)	Full text (sections)	Qingdao 2015 to 2018, GDAS 1°, 72 h; K-means and SOM beat hierarchical and HYSPLIT clustering; DBI preferred	Different city and climate
Lodhi et al. (2009)	Abstract	310 samples, 2005 to 2007; winter about 4 times other seasons; PMF sources	Older period; sampling site differs
Stone et al. (2010)	Abstract	2007 annual $PM_{2.5}$ 194 ± 94 μg m^{-3}; carbonaceous 61 ± 17%; CMB with organic tracers	One year; one-in-six-day sampling
Rasheed et al. (2015)	Abstract	Lahore autumn 2007 mean 150.4 μg m^{-3}; HYSPLIT high-pollution origins; NEQS values cited	2007 to 2008 for Lahore
Tariq et al. (2016)	Abstract (partial)	October 2013 haze; AOD up to 3.70; linked to crop-residue burning	Single event; column aerosol, not surface $PM_{2.5}$
Bilal et al. (2022)	Abstract (partial)	2016 to 2017 Lahore fog events; ERA5 and HYSPLIT; anticyclonic flow, winds below 1.5 m s^{-1}	Two events
Majeed et al. (2024)	Abstract	Ground and satellite data, fire counts, HYSPLIT, AHP; fires moderate importance relative to vehicles	AHP weights involve judgement
Dahal et al. (2026)	Abstract (published article page checked 25 October 2026)	November 2024 haze; Lahore daily maximum 621 μg m^{-3}; links to stagnation, PBLH, FRP; evening fire shift	Single event

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Reference	Access	Verified content used here	Limitation for this study
Cusworth et al. (2018)	Full text (sections)	STILT with five fire inventories; 7 to 78% of maximum enhancements in Delhi attributed to fires; up to 25% of weekly $PM_{2.5}$ variability reproduced	Delhi
Kulkarni et al. (2020)	Abstract	WRF-Chem, FINN v1.5; about 20% average, 50 to 75% in events	Delhi; numbers from abstract, to verify before submission
Liu et al. (2018)	Full text (sections)	72 h HYSPLIT back trajectories (GDAS 0.5°, 500 m for post-monsoon Delhi); 10 km trajectory buffer; daily MODIS C6 FRP; +16.34 μg m^{-3} PM_{10} per 1,000 MW in post-monsoon Delhi	PM_{10}; Indian cities
Jethva et al. (2018)	Abstract (full-text PDF not read)	More than 80% of north-westerly clusters crossed burning region before Delhi	Delhi; numbers from abstract, to verify before submission
Ojha et al. (2020)	Full text (sections)	WRF-Chem, October to December 2016; ventilation coefficient about 4 times lower in winter (about 600 to 1,000 m^2 s^{-1} in October, below 400 m^2 s^{-1} by December)	Model study
Gautam et al. (2023)	Abstract	Increasing lower-tropospheric stability over northern India, 1980 to 2019	India-focused
Ghosh et al. (2014)	Abstract	Delhi trajectory clustering and CWT; sources include Pakistan and Indian Punjab	Delhi, 2008 to 2010
Stein et al. (2015)	Abstract	HYSPLIT description	Methods reference
Fleming et al. (2012)	Abstract	Review of air-mass history methods	Methods reference

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Reference	Access	Verified content used here	Limitation for this study
Stohl (1998)	Record	Review of trajectory computation and accuracy	Content not re-read in this session
Su et al. (2018)	Abstract	Negative PM and PBLH correlation varying by place and season (China)	China
Miao et al. (2019)	Abstract	Review of PBL and $PM_{2.5}$ interaction	China
Guttikunda and Gurjar (2011)	Abstract (partial)	Meteorology and Delhi seasonality; Lagrangian tracer model	Delhi
Hersbach et al. (2020)	Abstract	ERA5: hourly, 31 km	Methods reference
Giglio et al. (2016); Schroeder et al. (2014)	Abstract; Record	MODIS C6 and VIIRS 375 m fire products	Product versions in FIRMS to be confirmed
Breiman (2001)	Record	Random Forests	Methods reference
Grange et al. (2018)	Abstract	RF with PBLH for Swiss PM_{10}	Normalisation, not forecasting
Stirnberg et al. (2021)	Abstract	Tree model and SHAP for meteorology and PM1	Explanation, not forecasting
Le et al. (2021)	Full text (sections)	HYSPLIT cluster labels as categorical RF predictors (Hanoi)	Weather normalisation; no cluster ablation seen
Feng et al. (2015)	Abstract	Trajectory-based indicator, wavelet, ANN; RMSE reduced up to 40%	Trajectory effect not isolated
Iftikhar et al. (2024)	Abstract	Daily $PM_{2.5}$ 2019 to 2023; ensembles; Lahore errors 0.81 to 13.52%	U.S. embassy monitor data; start date differs between abstract and methods; the 0.81 to 13.52% range cannot be matched to the paper's tables
Bhatti et al. (2021)	Full text (sections)	EPA monitoring data 2014 to 2019; SARIMA and factor analysis; HYSPLIT with GFS meteorology	Validation design not checked
Rashid and Atif (2026)	Record (DOI registered; checked 24 October 2026)	Deep learning with meteorology, Lahore	Content not checked

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Reference	Access	Verified content used here	Limitation for this study
Zhou et al. (2024)	Abstract and publisher page	Indo-Gangetic Plain stagnation days: winter $PM_{2.5}$ about 45 µg m^{-3} (45%) higher; stagnation defined by little rain, weak wind and a strong inversion	Regional, not Lahore-specific
Khanum et al. (2017)	Publisher page	Lahore 2007 to 2011; transport India to Pakistan in February to October and Pakistan to India in November to January, from MODIS and prevailing winds	No trajectories
Nawaz et al. (2023)	Publisher page	Lahore air masses 36% north, 33% east, 21% western India, for OMI absorbing aerosol	Column aerosol, not $PM_{2.5}$; no seasonal split
Raja et al. (2010)	Publisher page	Lahore source apportionment; biomass burning about 15%; PSCF unsuitable when local sources dominate	Older period
Ahmad et al. (2020)	Publisher page	Summertime Lahore CWT: local sources dominant, additional contribution from western India	Summer only
Ghosh et al. (2014) (additional detail)	Publisher page	Winter Delhi cluster beginning north-west over industrial Lahore	Delhi
Jena et al. (2021)	Publisher page	Delhi 400 m $PM_{2.5}$ forecast; earlier version struggled with episodes above 350 µg m^{-3}	Delhi; daily forecasts
Haseeb et al. (2026)	Publisher page	Lahore daily $PM_{2.5}$; chronological split with independent 2024 validation; Random Forest R^2 0.77, RMSE 24.75 µg m^{-3}	Daily, not weekly; no persistence benchmark

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Reference	Access	Verified content used here	Limitation for this study
Chianese and Riccio (2026)	Publisher page	Systematic review: no universal best model under leakage-safe rolling evaluation	Review
Rybarczyk and Zalakeviciute (2018); Méndez et al. (2023); Zhang et al. (2012)	Abstract; Abstract; Record	Reviews of ML and real-time air-quality forecasting	Background
Tashman (2000); Bergmeir and Benítez (2012); Hyndman and Koehler (2006)	Record	Forecast evaluation and error measures	Methods references
Kaufman et al. (2012)	Abstract	Definition of leakage	Methods reference
S. Lundberg and Lee (2017)	Record	SHAP (conference paper, NeurIPS)	Not a journal article
S. M. Lundberg et al. (2020)	Record	TreeSHAP	Methods reference
Karpatne et al. (2017); Reichstein et al. (2019); Di et al. (2016)	Abstract; Record; Abstract	Physics-guided and hybrid modelling	Background
Jacob (1999)	Record (book)	Box-model framework	Textbook, not peer-reviewed article

Appendix B. Research Decisions Register

Decision	Options	Evidence	Status
Primary hourly $PM_{2.5}$ source	Rebuilt AirNow hourly; OpenAQ 8664; AirNow daily values	OpenAQ has label shifts and spurious values; AirNow is the original and has more coverage	Settled: rebuilt AirNow hourly record (complete)
Repeated-value run rule	S6 (6 h or more) or S7 (7 h or more)	Both remove the confirmed 2024 artefact; differ on one unverified 2020 run	Settled: S7 (removed 9 hours)
Weekly completeness threshold	7/7, at least 6/7, at least 5/7, or weighting	Table G3 errors; 145, 178 or 200 weeks	Settled: at least 6/7 primary, 7/7 as sensitivity
Week boundary	Monday to Sunday PKT or alternative	Matches ISO weeks and Stage 1C tables	Settled: Monday to Sunday PKT (20 October 2026)
Partial weeks at record edges	Keep if eligible, or drop	Code flags them	Settled: only full weeks are used (299 full weeks)

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Decision	Options	Evidence	Status
Persistence maximum age	None, or a set number of weeks	Gaps up to 120 days exist	Settled: no maximum age; source ages reported (Section 5.7)
Climatology parameters	Window width, minimum years and samples, median or mean	Required arguments in code	Settled before results: median, ±2 weeks, at least 4 weeks from at least 2 earlier years
Data latency and embargo	One-week embargo or none	ERA5 and other inputs are released with delays; exact latencies not documented	Used: none for the main results; one-week embargo run as a check (Section 6.5); per-source latency not run (Section 5.6)
Development and final test periods	Proposal: development May 2019 to September 2023; final test October 2023 to February 2025	Best-covered smog seasons are 2023-24 and 2024-25	Not used: every week forecast by the same expanding walk-forward design (Section 5.13)
High-pollution threshold	Fixed concentration or training-period percentile	None	Not used: October-to-January weeks reported instead (Section 5.9)
Trajectory meteorology and settings	GDAS or ERA5; 48, 72 or 96 h; 100, 500 or 1,000 m	Karim and Rappenglück (2023) used GDAS and 72 h	Settled: NCEP/NCAR 2.5°, 72 h, 500 m, 06:00 UTC; 100 and 250 m run as a check; GDAS 1° run for one year as a comparison; other settings not run
Fire product and weights	MODIS or VIIRS; confidence filter; buffer and time window	Liu et al. (2018) airshed precedent	Settled: VIIRS Suomi NPP standard processing, nominal or high confidence, vegetation fires; ring and sector features used; trajectory weights open
Box-model length scale L	Illustrative 20 km or alternatives	Unobserved	Used: 30 km (only scales the flushing-time feature)
AQI standard	US EPA (2024), Punjab EPA scale, or none	Punjab bulletins use differing bands; weekly mean is not an official AQI	Not chosen; no AQI computed
OpenAQ location 6506	Resolve with API key, or drop	No page, no archive files	Not resolved; not used

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Decision	Options	Evidence	Status
Data terms	Register with OpenAQ; submit AirNow guidelines form	Terms reviewed in Stage 1C	AirNow notified 27 October 2026; U.S. Department of State contact asked; OpenAQ terms open

Appendix C. Box-Model Dimensional Check

Term	Quantities and units	Resulting units
dC/dt	$\mu g\,m^{-3}$ per s	$\mu g\,m^{-3}\,s^{-1}$
E/H	$(\mu g\,m^{-2}\,s^{-1})$ / (m)	$\mu g\,m^{-3}\,s^{-1}$
$(U/L)(C_{in} - C)$	$(m\,s^{-1}$ / m) × $\mu g\,m^{-3}$	$\mu g\,m^{-3}\,s^{-1}$
kC	s^{-1} × $\mu g\,m^{-3}$	$\mu g\,m^{-3}\,s^{-1}$
Ventilation coefficient $U \times H$	$m\,s^{-1}$ × m	$m^2\,s^{-1}$
Accumulation index $L/(U \times H)$	m / $(m^2\,s^{-1})$	$s\,m^{-1}$ (multiplied by E gives $\mu g\,m^{-3}$)

Appendix D. Figures

All figures appear in the text.

Figure	Content	Section	Status
1	Density of VIIRS fire detections around Lahore, with distance rings and direction sectors	4.5	Made
2	Daily $PM_{2.5}$ data availability	6.1	Made
3	Weekly mean $PM_{2.5}$ series	6.1	Made
4	Monthly cycle of $PM_{2.5}$, fire detections, ventilation and rainfall	6.1	Made
5	Mean paths of the trajectory clusters	6.2	Made
6	$PM_{2.5}$ by cluster within each season	6.3	Made
7	$PM_{2.5}$ against transport distance, by season	6.3	Made
8	Trajectory frequency and CWT maps, post-monsoon and winter	6.3	Made
9	Fire exposure against $PM_{2.5}$ anomaly, October and November	6.4	Made
10	Mean absolute error of each forecasting method	6.5	Made
11	Observed weekly $PM_{2.5}$ and one-week-ahead forecasts	6.5	Made
12	Forecast against observed weekly $PM_{2.5}$ for the full model	6.5	Made

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Figure	Content	Section	Status
13	Change in error for each added input group, with bootstrap intervals	6.6	Made
14	Model reliance on each input group (SHAP)	6.7	Made
–	Conceptual diagram of the one-box model	5.5	Not made

Appendix E. Software and Test Status

The analysis code is in the project's src folder. It covers data download and quality control (download_openaq, airnow_archive, airnow_bulk, airnow_hourly_series, pm25_qc, run_openaq_coverage, stage1c_analysis), weekly targets and baselines (weekly_targets, baselines, stage2_weekly, history_model), input features (download_era5, era5_features, download_firms, firms_features, mass_balance_features, trajectory_features), models (main_model, followup_models, stage4_trajectory_model, stage8_robustness), interpretation (shap_analysis) and figures (figures_data, figures_results). Embargo, sensitivity, diagnostic and weekly ventilation checks are in stage8_robustness. The boundary-layer check is in stage9a_blh_check, the 100 and 250 m trajectory analysis in stage9b_multiheight, and the one-year GDAS 1° comparison in stage9c_gdas_comparison. The full test suite (151 tests) passed on 27 October 2026. All tests use synthetic data: they check that the code behaves as specified and say nothing about Lahore's air quality. Exact package versions are recorded in the project's requirements-lock.txt, and rerunning the pipeline with them reproduces the committed results exactly.

Appendix F. Verification Log

Project status was determined on 18 October 2026 by reading the project's rules file, README.md, PROGRESS_LOG.md, the Stage 1A, 1B and 1C reports, the forecasting design note, the source code, the download manifests and the test cache. No data, code or model files were changed. The progress log ends at Stage 1C and does not yet record the weekly-target and baseline modules or the AirNow bulk reconstruction; the README still describes Stage 1B. Bibliographic details were checked against Crossref, OpenAlex and DataCite records, and against publisher pages (ScienceDirect, MDPI, Springer and AAQR) where the full text or abstract was available. One Perplexity query was used to find leads; every lead was checked independently, and one Perplexity description that went beyond what its source showed was not relied on. Relevant sections of the full texts of Karim and Rappenglück (2023), Mohyuddin et al. (2024), Qureshi et al. (2025), Le et al. (2021), Liu et al. (2018), Cusworth et al. (2018), Ojha et al. (2020), Cui et al. (2021) and Bhatti et al. (2021) were read; the other works were checked at the level recorded in Appendix A.

Update, 24 and 25 October 2026. Every reference in this version was looked up by its DOI at doi.org (Crossref and DataCite), which also showed that Dahal et al. (2026) and Rashid and Atif (2026) have been published. The 621 $\mu g\ m^{-3}$ figure from Dahal et al. (2026) was checked against the published abstract on 25 October 2026. Sources added from a separate literature review (Zhou et al., 2024; Khanum et al., 2017; Nawaz et al., 2023; Raja et al., 2010; Ahmad et al., 2020; Jena et al., 2021; Haseeb et al., 2026; Chianese & Riccio, 2026) were checked against the publisher's page for the claims used here. Stone et al. (2010) has been checked only at the level of its abstract and DOI record. The methods and results in Sections 4 to 6 were checked

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

against the project's stage reports, code and saved outputs.

Appendix G. Preliminary Audit of the OpenAQ Copy

These figures describe the OpenAQ copy of the record (S0, no exclusions) as reported in the project's Stage 1B and 1C analyses. They are data-quality checks, not results about Lahore's air quality, and all will be recomputed on the rebuilt AirNow record.

Across 2,113 calendar days, 1,464 (69.3%) met the 18-of-24-hour rule, and 36,409 of 50,703 hours (71.8%) were valid.

Table G.1

Daily completeness of the OpenAQ copy by year.

Year	Days in record	Complete days	% complete	% hours valid
2019 (from 23 May)	223	208	93.3	94.7
2020	366	331	90.4	92.5
2021	365	226	61.9	66.2
2022	365	99	27.1	30.0
2023	365	290	79.5	79.1
2024	366	263	71.9	76.5
2025 (to 4 March)	63	47	74.6	74.9

Table G.2

Smog-season coverage. Week counts are for the OpenAQ copy; AirNow daily values are from daily_data_v2 files.

Smog season (Oct to Jan)	Full weeks	7/7	At least 6/7	At least 5/7	AirNow daily values	OpenAQ complete days
2019-20	18	14	16	16	121	114
2020-21	18	12	15	17	121	113
2021-22	17	7	8	8	72	70
2022-23	17	0	2	4	97	37
2023-24	17	16	16	16	120	117
2024-25	18	17	18	18	122	122
Total	105	66	75	79		

Across the whole record there were 301 full weeks: 145 met the 7/7 rule, 178 at least 6/7 and 200 at least 5/7. AirNow's own daily files hold 1,778 daily values, including 318 days on which the OpenAQ copy is incomplete. The 2022-23 smog season shows the largest difference (97 AirNow daily values against 37 complete OpenAQ days), which suggests that the rebuilt record may recover much of the missing coverage. AirNow's own daily completeness rule has not been found in its documentation, so these daily values cannot yet be assumed to meet the 18-hour rule.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.04 No.01 (2026)

Table G.3

Error in a weekly mean when days are removed from complete weeks (all possible subsets; assumes days are missing at random).

Days removed	Season	Weeks	Mean absolute error ($\mu g\, m^{-3}$)	Median absolute % error	95th percentile % error
1	Non-smog	79	2.83	2.97	10.33
1	Smog	66	7.93	2.83	10.15
2	Non-smog	79	4.44	4.94	16.47
2	Smog	66	12.00	4.14	16.26