Supplement of Atmos. Chem. Phys., 26, 5553–5566, 2026
https://doi.org/10.5194/acp-26-5553-2026-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Impact of the Chinese Spring Festival on $PM_{2.5}$ air quality in the Beijing-Tianjin-Hebei and surrounding region: a machine learning-based counterfactual modeling approach

Yuan Li et al.

Correspondence to: Qili Dai (daiql@nankai.edu.cn) and Yinchang Feng (fengyc@nankai.edu.cn)

List of tables

List of texts

List of figures

0 representing Chinese New Year's Eve and 1 representing the first day of the Lunar New Year.

Table S1. Analytical instruments/methods, method detection limits (MDLs), and uncertainty (error fractions) for $PM_{2.5}$ chemical species used as PMF inputs. Instrument calibration and routine QA/QC procedures (multi-point calibration using standards, field/laboratory blanks, and replicate analyses) were performed to ensure analytical accuracy. The MDLs and error fractions listed in this table were used as PMF model inputs.

Chemical Composition	Chemical Species	Analytical Instruments and Methods	Method Detection Limit (ug/cm^2)	Uncertainty (ErrorFraction)
Elemental species	Na	iCAP 7000 Series ICP-OES spectrometer (Thermo Scientific, USA)	0.00017584	0.1
	Mg		0.00020096	0.1
	Al		0.00001	0.1
	Si		0.000271296	0.1
	S		0	0.1
	Cl		0.00211008	0.1
	K		0.002841072	0.1
	Ca		0.003062128	0.1
	Ti		0.00001	0.1
	V		0.0054636	0.1
	Cr		0.002755664	0.1
	Mn		0.00145696	0.1
	Fe		0.003225408	0.1
	Cu		0.00236128	0.1
	Co		0.004255328	0.1
	Zn		0.00001	0.1
	As		0.002512	0.1
	Pb		0.010306736	0.1
	Ba		0.00001	0.1
	Ni		0.00105504	0.1
	Cd		0.010007808	0.1
	Se		0.00001	0.1
Water-soluble ions	Na^+	ion chromatography (IC)	0.0001	0.1
	NH_4^+		0.1	0.0764
	K^+		0.05	0.1
	Mg^{2+}		0.01	0.15
	Ca^{2+}		0.0001	0.1
	Cl^-		0.05	0.13
	NO_3^-		0.1	0.07
	SO_4^{2-}		0.08	0.0655
Carbonaceous components	OC	DRI 2001A Thermal/Optical OC/EC analyzer (IMPROVE_A protocol)	0.5	0.1
	EC		0.5	0.2

Table S2. Summary of the meteorological factor variables downloaded from ERA5. Note: a positive zonal wind means westerly; positive meridional wind means southerly; positive vertical velocity means downward motion.

Abbreviations	Names of variables	Units of variables
U10	Zonal wind at 10 m	m/s
V10	Meridional wind at 10 m	m/s
T2	Temperature at 2 m	K
D2	Dewpoint temperature at 2 m	K
SSR	Surface net solar radiation	J/m^2
TP	Total precipitation	m
BLH	Boundary layer height	m
TCC	Total cloud cover	%
SP	Surface pressure	Pa

Table S3. Comparison of bootstrap diagnostics for PMF solutions with different factor numbers (6-10 factors). The 9-factor solution shows the highest bootstrap mapping stability (96-100%) with no unmapped factors. In contrast, solutions with fewer factors tend to mix multiple sources, whereas the 10-factor solution shows reduced stability and limited improvement in model performance. These results indicate that the 9-factor solution is the most stable and interpretable among the tested solutions.

BS diagnostics	6-fatcor	7-factor	8-factor	9-factor	10-factor
Lowest %BS mapping	91	94	94	96	94
Highest %BS mapping	100	100	100	100	97
Highest unmapped	1	0	0	0	0

S1. Preprocessing of downloaded ERA5 meteorological variables

Surface meteorological variables, including 10-meter zonal wind (U10) and meridional wind (V10), 2-meter temperature (T2) and dew point temperature (D2), boundary layer height (BLH), total cloud cover (TCC), total precipitation (TP), surface pressure (SP), and surface net solar radiation (SSR), were obtained from the ERA5 hourly reanalysis on single levels (Copernicus Climate Data Store, https://cds.climate.copernicus.eu/), spanning 1940 to present with a temporal resolution of 1 hour and a horizontal resolution of 0.25° × 0.25°. The ERA5 total precipitation (TP) variable represents the accumulated rainfall within each hour (m h^{-1}); for aggregated analyses, hourly TP values were summed to daily totals. The 10 m surface wind speed (WS) and direction (WD) were derived from the zonal (U10) and meridional (V10) wind components using vector synthesis and trigonometric relationships. Relative humidity (RH) was calculated using Emanuel's (1994) empirical saturation vapor pressure formula, expressed as:

$$In\ E = 53.67957 - \frac{6743.769}{T} - 4.8451\ In\ T$$

Where E is the vapor pressure (hPa) and T is the absolute temperature (K). The formula was applied separately to the 2-m air temperature (T2) and dew point temperature (D2) to obtain the actual vapor pressure (E_a) and saturation vapor pressure (E_s), respectively. Relative humidity was then computed as RH = (E_a / E_s) × 100%.

S2. Representativeness of ERA5 meteorological data

To evaluate the representativeness of ERA5 meteorological data during the study period (20 December 2023-16 February 2024), we compared hourly ERA5 variables at the grid cell centered over downtown Hangzhou (30.2084° N, 120.212° E), which served as the meteorological inputs for our case study of the 2024 Hangzhou New Year haze episode, with ground-based observations from the Hangzhou Xiaoshan International Airport meteorological monitoring site (30.23397° N, 120.43144° E), located approximately 20 km southeast of the city center. ERA5 shows small mean biases for near-surface temperature and wind speed (≈ 0.36 °C and 0.40 m/s, respectively), while wind direction exhibits a modest mean bias (≈ 17°), consistent with the expected spatial smoothing of a 0.25° gridded reanalysis product.

Importantly, the strength of ERA5 in the context of this study lies in its ability to provide a spatially and temporally consistent representation of synoptic-scale meteorological conditions. In contrast, ground-based meteorological observations are often highly localized, contain missing or interrupted records, and many key variables are available only at 3-hour intervals, which restricts their direct use for machine-learning-based modeling at an hourly resolution; moreover, ERA5 provides a more comprehensive set of meteorological variables.

The reliability of ERA5 as a primary driver for atmospheric modeling is well documented. According to the comprehensive global evaluation by Hersbach et al. (2020), ERA5 significantly improves the representation of near-surface meteorological states and atmospheric dynamics compared to previous reanalysis products, benefiting from a high-resolution 4D-Var data assimilation system. While localized point-to-grid comparisons may exhibit high-frequency variability due to spatial smoothing, the core strength of ERA5 lies in its internal physical consistency and its ability to accurately capture synoptic-scale transitions. These features are critical for ML-based $PM_{2.5}$ predictions because they ensure that the regional meteorological background, including essential vertical mixing parameters, is represented with high fidelity.

Our counterfactual modeling framework aims to estimate relative changes in regional $PM_{2.5}$ concentrations driven by large-scale meteorological patterns, rather than to capture hyper-local or sub-grid meteorological extremes. The observed discrepancies, while inherent to reanalysis data, do not compromise ERA5's capability to characterize the regional-scale meteorological forcing that underpins our analysis. Therefore, ERA5

offers a suitable and robust meteorological foundation for the counterfactual modeling framework employed in this study.

S3. Modeling details of the ML-Based counterfactual $PM_{2.5}$ air quality prediction framework

For the Hangzhou 2024 Spring Festival modeling dataset, model hyperparameters were optimized using Bayesian optimization (Optuna, version 4.5.0) with a 5-fold cross-validation strategy applied to the training set. The objective function maximized the coefficient of determination (R^2) obtained from cross-validation.

The optimal hyperparameters were as follows:
n_estimators = 246, *learning_rate* = 0.067, *max_depth* = 8, *min_child_weight* = 2, *subsample* = 0.76, *colsample_bytree* = 0.77, *reg_alpha* = 0.94, and *reg_lambda* = 0.56.

The mean R^2 across the cross-validation folds was 0.956.

On the independent test set, the model achieved an R^2 of 0.969 with an RMSE of 6.72 μg m^{-3}.

Figure S1 shows the learning curve of the Bayesian-optimized XGBoost model. The training R^2 (red line) remains close to 1.0 across all training sizes, while the cross-validation R^2 (green line) increases steadily with more training data and stabilizes around 0.95. This indicates that the model achieves strong generalization without significant overfitting, and that the available dataset (~1000 samples) is sufficient for reliable model training.

Figure S2 illustrates the observed $PM_{2.5}$ concentrations together with the ensemble-based counterfactual predictions, their 95% prediction intervals, and the deviations (Δ = observed - predicted). To quantify predictive uncertainty, we implemented an ensemble-based bootstrap approach. Fifty XGBoost models with identical hyperparameters were trained on bootstrap-resampled versions of the training set, generating an ensemble of counterfactual BAU predictions. The ensemble mean was used as the point estimate, and the 2.5 th-97.5 th percentile range of the ensemble distribution was taken as the 95% prediction interval. The ensemble-mean predictions achieved an R^2 of approximately 0.95 on the test set, slightly lower than the single-model performance but providing well-calibrated uncertainty estimates.

Figure S3 illustrates the SHAP (SHapley Additive exPlanations) summary combining a beeswarm distribution and mean absolute SHAP bar plot. The horizontal axis represents each feature's SHAP value contribution to model prediction, while the color gradient (blue-red) denotes the corresponding feature value (low-high).

Among all predictors, date_unix and length exerted the most substantial influences on model behavior, but with distinct physical interpretations. The feature date_unix, representing the long-term temporal trend associated with emission dynamics, exhibited predominantly negative SHAP values as time progressed throughout the study period. This pattern indicates that emissions during the Chinese New Year period

were generally lower than those observed before the festival, reflecting reduced anthropogenic activity.

Similarly, length, denoting the trajectory length, showed an overall negative correlation with its SHAP contribution, implying that shorter back-trajectories were linked to higher predicted concentrations. This suggests that during the Spring Festival period, local emissions within Hangzhou dominated the observed pollution levels, while the influence of surrounding regional transport was comparatively minor. The categorical features cluster C1-C6, which represent six typical air-mass transport clusters along with their relative contributions, further support this interpretation. Among them, clusters C5 and C4, which exerted the strongest influence on local conditions, were associated with negative SHAP contributions, indicating that these trajectories likely brought cleaner air masses into Hangzhou.

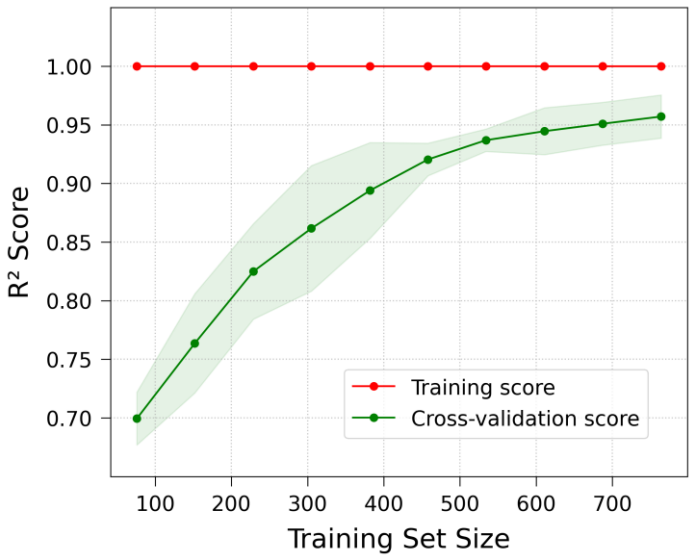

Figure S1. Learning curve of the ML-Based counterfactual $PM_{2.5}$ prediction framework.

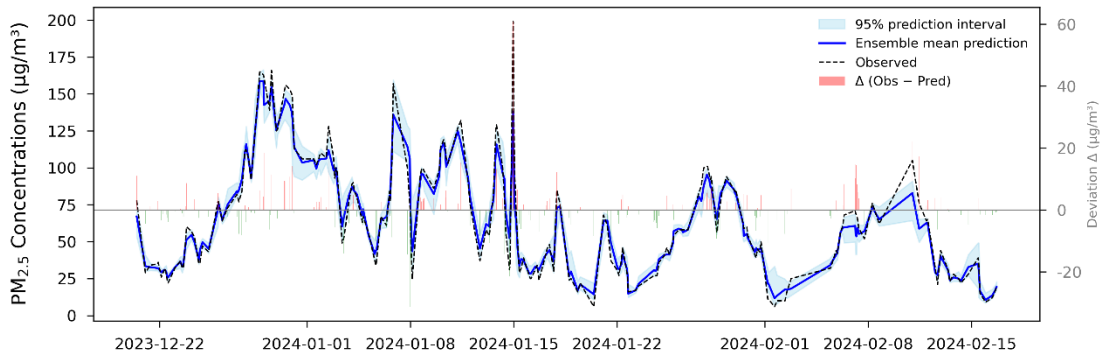

Figure S2. Observed $PM_{2.5}$ concentrations and ensemble-based counterfactual predictions from the XGBoost model, together with their 95% prediction intervals and the corresponding deviations (Δ = observed - predicted).

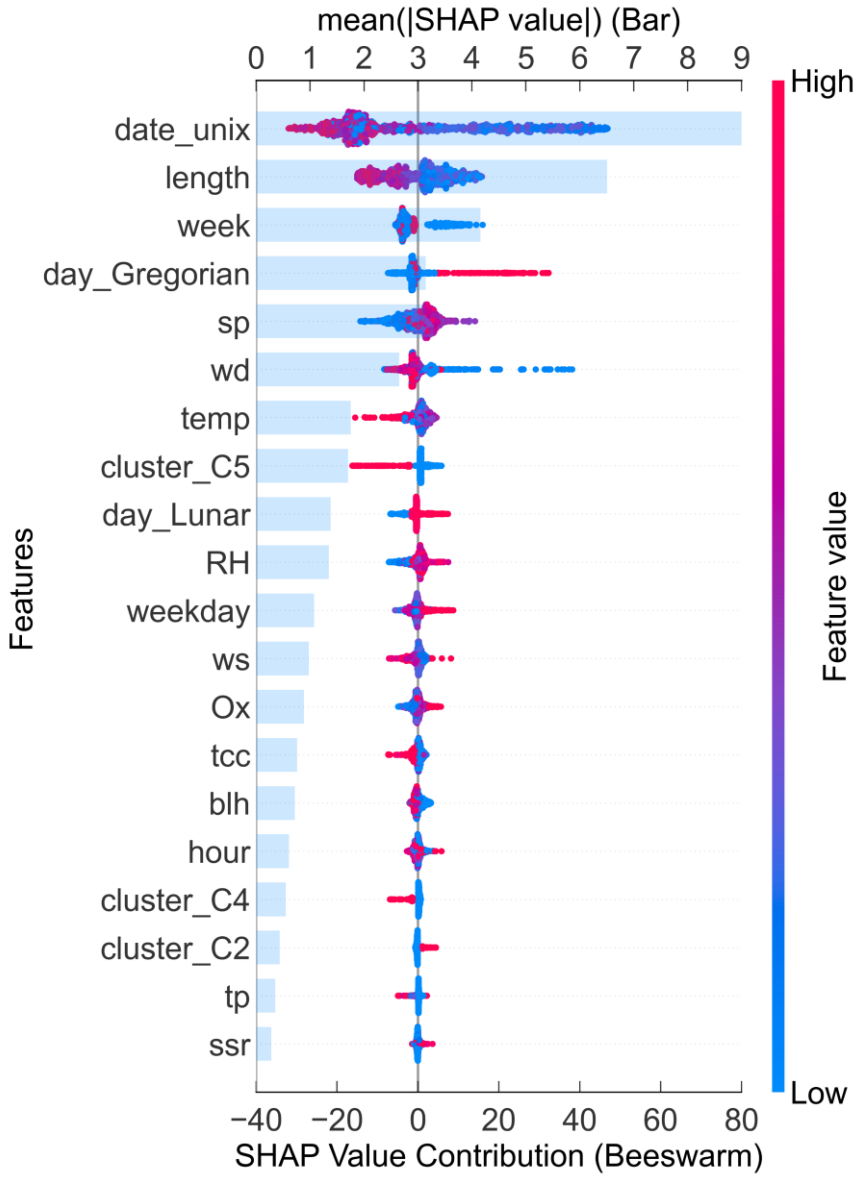

Figure S3. SHAP summary plots of feature contributions to the ML counterfactual $PM_{2.5}$ model (upper: feature importance; lower: feature contributions).

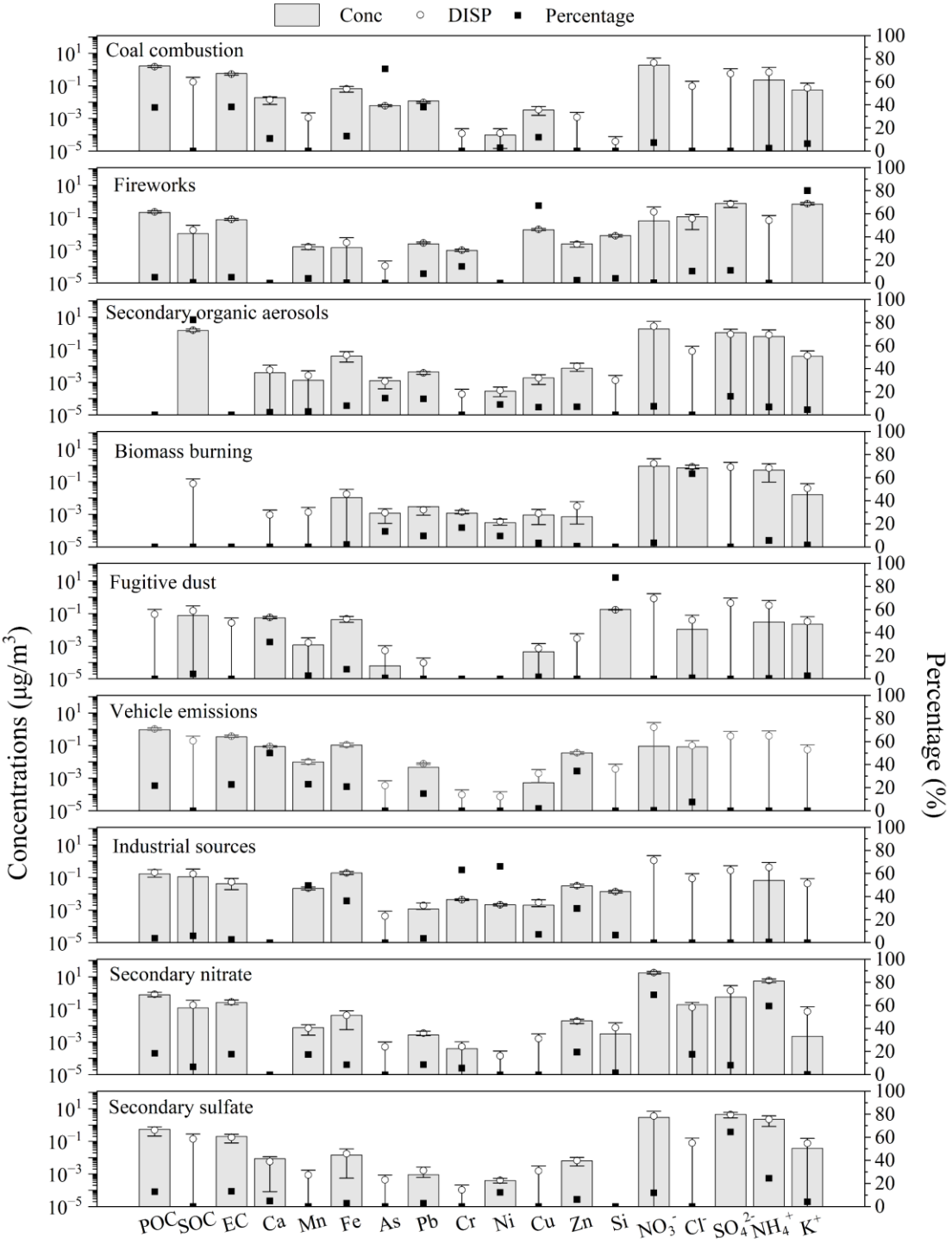

Figure S4. Source profiles of the nine factors resolved by the DN-PMF analysis for the $PM_{2.5}$ pollution episode during New Year's Eve 2024 in Hangzhou. The gray bars represent the species concentrations apportioned to each factor (μg m^{-3}), open circles denote the DISP-derived uncertainties, and black squares indicate the relative contributions (%) of each species to the corresponding factor. The identified factors include: coal combustion, fireworks, secondary organic aerosols (SOA), biomass burning, fugitive dust, vehicle emissions, industrial sources, secondary nitrate, and secondary sulfate.

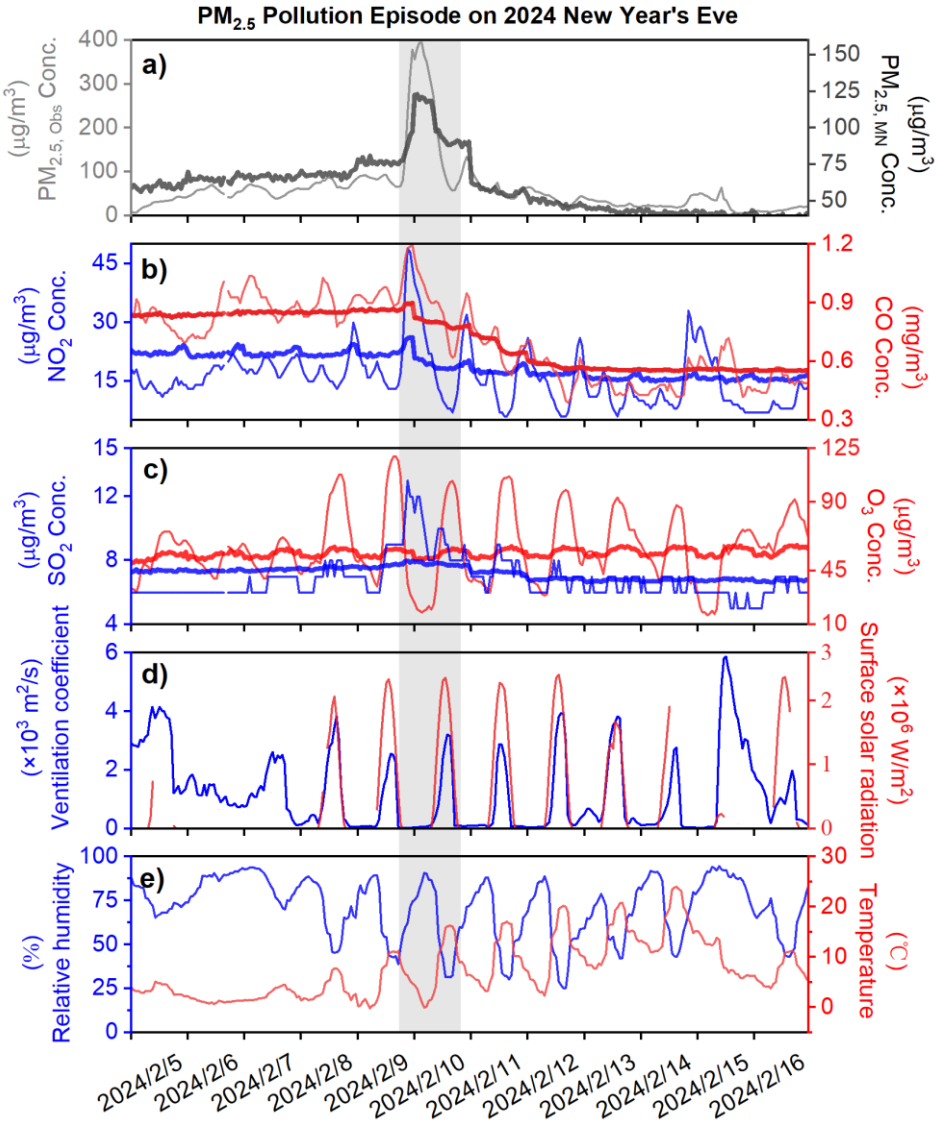

Figure S5. Temporal variations of air quality, meteorologically normalized concentrations, and meteorological factors around the 2024 Spring Festival. In subplots a, b, and c, the thin lines represent the observed values of pollutants, while the thick lines indicate the meteorologically normalized concentrations.

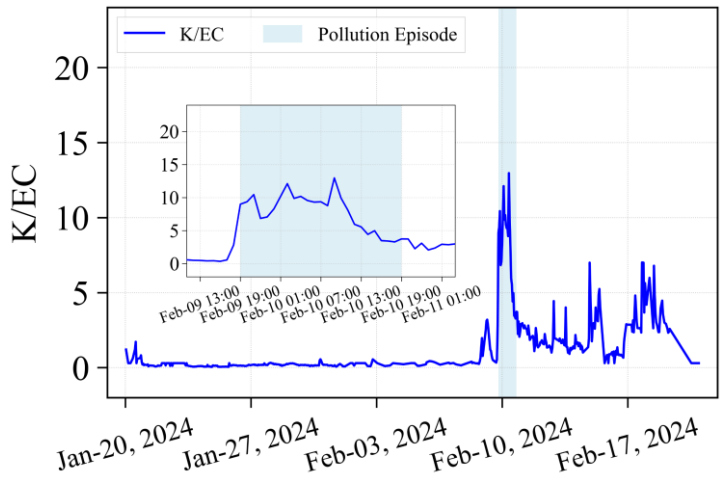

Figure S6. A dramatic surge in the K/EC ratio of $PM_{2.5}$ during the 2024 New Year's Eve: Evidence for firework emissions in Hangzhou.

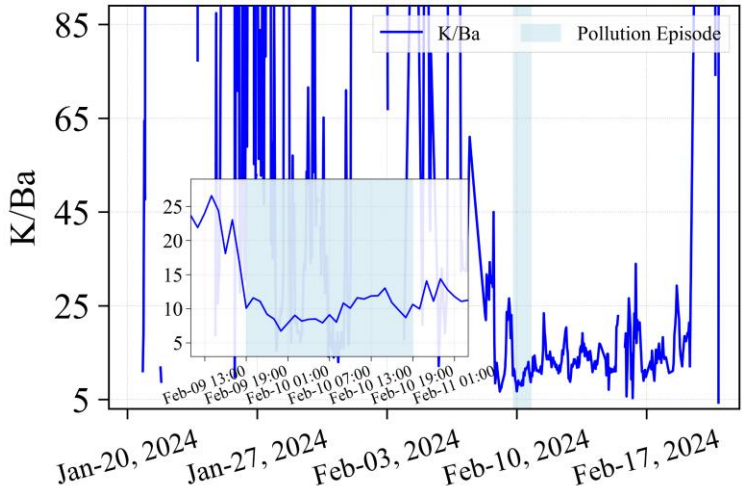

Figure S7. Temporal variation of the K/Ba ratio in $PM_{2.5}$ during the 2024 New Year's Eve haze episode in Hangzhou.

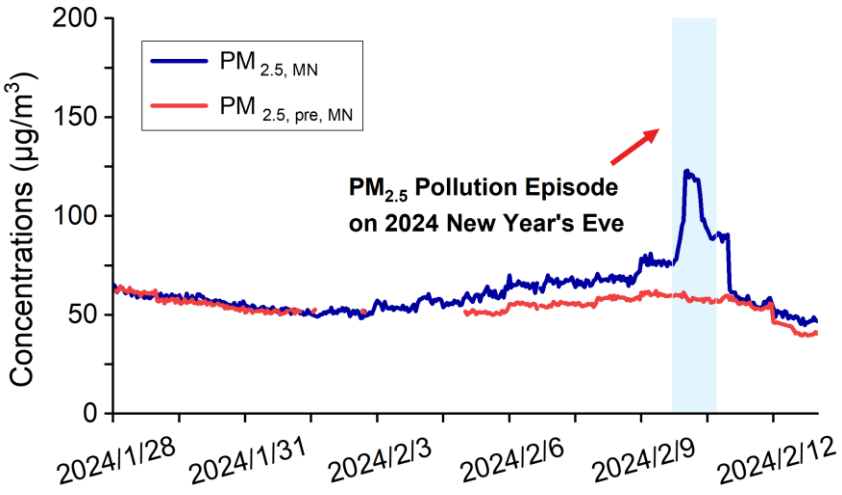

Figure S8. Temporal variations of $PM_{2.5,\ MN}$ and $PM_{2.5,\ pre,\ MN}$ around the 2024 Lunar New Year's Eve in Hangzhou.

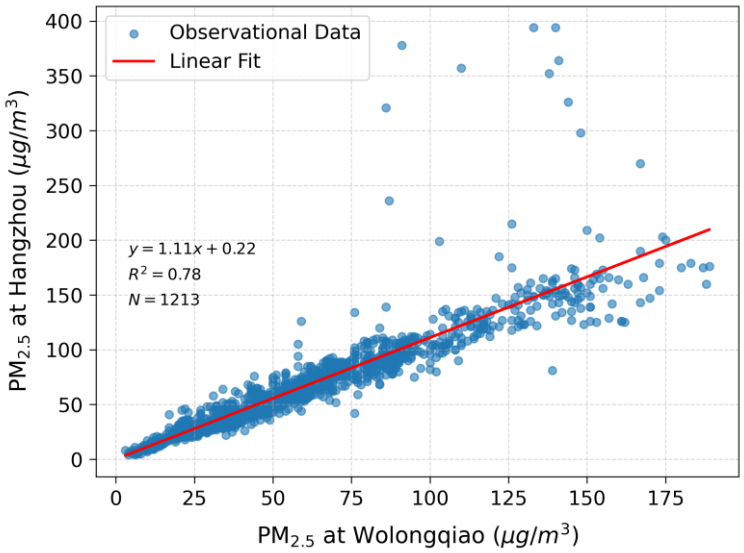

Figure S9. Linear regression analysis of hourly $PM_{2.5}$ concentrations between the Wolongqiao site (x-axis) and the Hangzhou city-wide average (14 sites, y-axis) during the study period (2023/12/20 - 2024/2/20). The strong correlation (R^2 = 0.78) supports the representativeness of the Wolongqiao site in capturing city-scale pollution trends during the observation period.

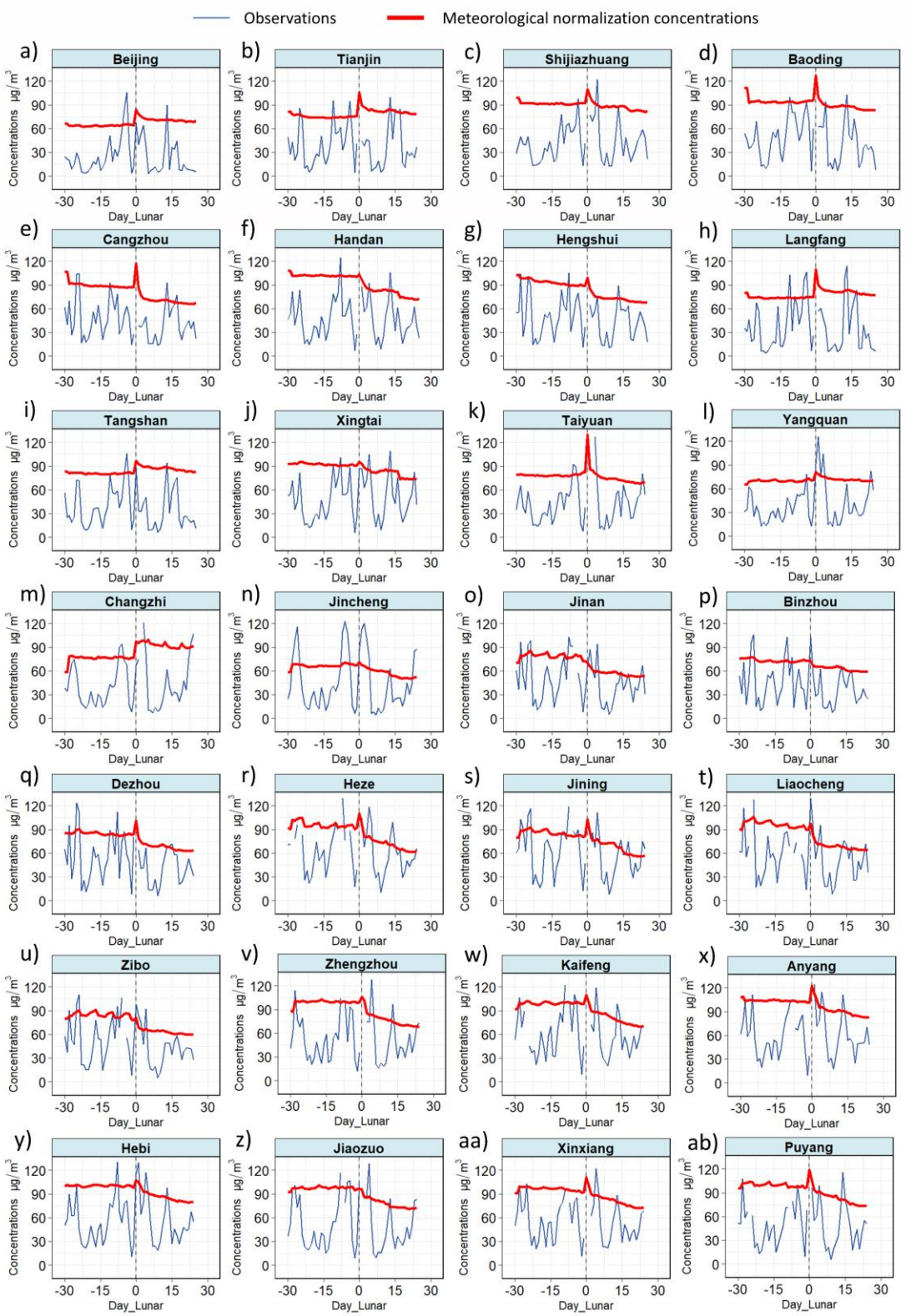

Figure S10. Temporal variations of observed and meteorologically normalized $PM_{2.5}$ concentrations during the 2025 Spring Festival in the "2+26" cities of northern China. The thick red line represents the meteorologically normalized concentration of $PM_{2.5}$, while the thin blue line indicates the observed concentration. The x-axis shows the lunar calendar days, with 0 representing Chinese New Year's Eve and 1 representing the first day of the Lunar New Year.

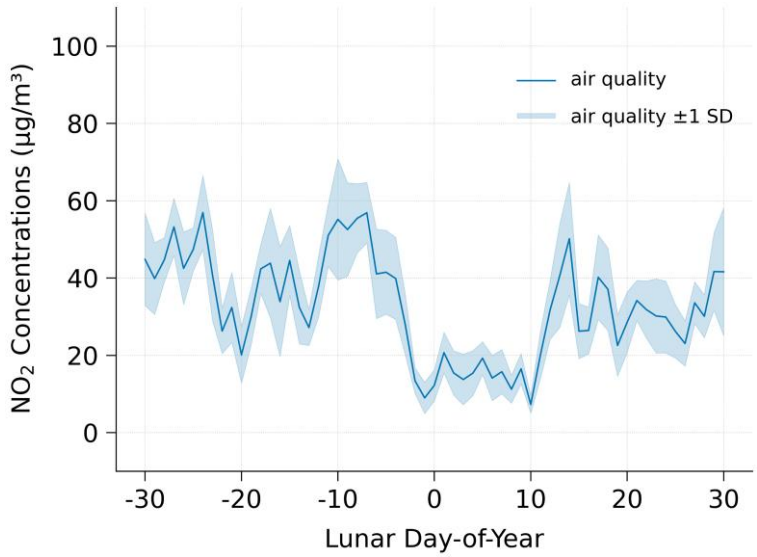

Figure S11. Temporal variation of NO_2 concentrations in the "2+26" cities of northern China around the 2025 spring festival. The x-axis shows the lunar calendar days, with 0 representing Chinese New Year's Eve and 1 representing the first day of the Lunar New Year.

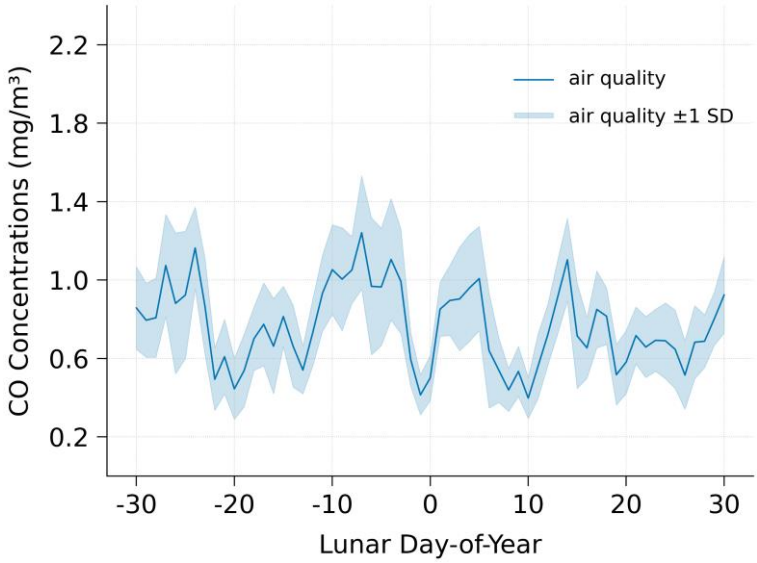

Figure S12. Temporal variation of CO concentrations in the "2+26" cities of northern China around the 2025 spring festival. The x-axis shows the lunar calendar days, with 0 representing Chinese New Year's Eve and 1 representing the first day of the Lunar New Year.

Reference:

Emanuel, K.A., 1994. Atmospheric convection. Oxford University Press.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Q. J. Roy. Meteorol. Soc., 146, 1999-2049, https://doi.org/10.1002/qj.3803, 2020.