https://doi.org/10.5194/egusphere-2026-2846
Preprint. Discussion started: 23 June 2026

Meteorological normalization of surface ozone variability across warm and cold seasons in China during 2015–2024

Zhenying Ye[1], Run Liu[1], Tingting Hu[2,3], Guoxin Chen[1], Min Wu[1]

[1]College of Environment and Climate, Institute for Environmental and Climate Research, Jinan University, Guangzhou, 511443, China
[2]Fujian Academy of Environmental Sciences, Fuzhou, 350013, China
[3]Fujian Key Laboratory of Environmental Engineering, Fuzhou, 350013, China

Correspondence to: Run Liu (liurun@jnu.edu.cn)

Abstract. Surface ozone (O_3) concentrations in China have continued to increase despite substantial reductions in primary air pollutants, but the relative roles of meteorological variability and longer-term non-meteorological changes remain uncertain across seasons and regions. Here, we applied a LightGBM-based meteorological normalization framework combined with SHapley Additive exPlanations (SHAP) to investigate interannual variations in maximum daily 8 h average O_3 (MDA8 O_3) across China during warm and cold seasons from 2015 to 2024. The model reproduced daily MDA8 O_3 variability reasonably well, with mean testing R^2 values of 0.71 and 0.77 in the warm and cold seasons, respectively. Observed national mean MDA8 O_3 increased by 2.1 $\mu g\ m^{-3}\ yr^{-1}$ in the warm season and 1.7 $\mu g\ m^{-3}\ yr^{-1}$ in the cold season. After meteorological normalization, MDA8 O_3 still increased at 1.2 $\mu g\ m^{-3}\ yr^{-1}$ in both seasons, indicating that the decadal O_3 increase was mainly associated with non-meteorological components. SHAP analysis revealed distinct seasonal and regional meteorological associations. During the warm season, temperature and solar radiation were more important in northern inland and basin regions, whereas relative humidity and wind-field variables were more important in southern coastal regions. During the cold season, solar radiation dominated across most regions, while relative humidity was more important in the Pearl River Delta. The net meteorological contribution shifted from generally negative during 2015–2019 to positive after 2019, indicating that the recent unfavorable meteorological conditions have amplified O_3 pollution in several regions.

1 Introduction

Surface ozone (O_3) has become an increasingly important air pollutant in many regions of the world, with adverse effects on human health, vegetation, and crop production (Feng et al., 2015; Chen et al., 2023; He et al., 2023). In China, substantial reductions in primary air pollutants, including SO_2 and $PM_{2.5}$, have been achieved since the implementation of the Air Pollution Prevention and Control Action Plan in 2013. However, surface O_3 concentrations have continued to increase in many regions, limiting further improvements in air quality (Wang et al., 2017; Li et al., 2020; Wang T. et al, 2022). This problem is particularly evident in densely populated and highly industrialized regions, such as the North China Plain (NCP), Yangtze River Delta (YRD), Pearl River Delta (PRD), Sichuan-Chongqing Basin (SCB), and Fenwei Plain (FWP), where O_3 pollution exhibits pronounced regional characteristics and complex temporal variability (Lu et al., 2019; Wang W. et al, 2022; Ren et al., 2023; Cao T., Wang H., Li L., et al, 2024).

The formation and variability of tropospheric O_3 are controlled by coupled effects of precursor emissions, atmospheric chemistry, and meteorological conditions (Li et al., 2016; Wang et al., 2017). Photochemical O_3 production depends nonlinearly on nitrogen oxides (NOx) and volatile organic compounds (VOCs), resulting in different O_3 sensitivity regimes across regions and seasons (Ren et al., 2022; Wang T. et al, 2022). Meteorological conditions further regulate O_3 formation, accumulation, transport, and removal. High temperature, strong solar radiation, low relative humidity, weak ventilation, and stagnant conditions generally favor photochemical production and near-surface accumulation of O_3, especially during the warm season (Liu et al., 2019; Wang, L. et al; Yang et al., 2024). In addition, boundary-layer dynamics and regional transport can modify the vertical mixing and horizontal redistribution of O_3 and its precursors (Li et al., 2016; Wang et al., 2017; Weng et al., 2022). These processes indicate that observed O_3 variability reflects both changes in precursor emissions and meteorological modulation, and that their relative importance may differ substantially between seasons and regions.

Quantifying the respective roles of meteorological variability and non-meteorological changes is essential for interpreting long-term O_3 trends and evaluating the effectiveness of air-quality management. Several approaches have been used for this purpose, including chemical transport models (CTMs), statistical decomposition methods, and machine-learning-based meteorological normalization (Yang and Zhao, 2023; Li et al., 2022; Guo et al., 2022; Liu et al., 2019; Zhou et al., 2022; Xue et al., 2023; Zhang et al., 2024). CTMs can explicitly represent emissions, chemistry, transport, and deposition processes, and are therefore powerful tools for process-level attribution. However, their results can be sensitive to emission inventories, chemical mechanisms, boundary conditions, and meteorological inputs, which introduces uncertainty in long-term and large-scale applications (Yang and Zhao, 2023; Li et al., 2022). Statistical methods, such as Kolmogorov-Zurbenko filtering and multiple linear regression, provide useful empirical estimates of meteorological and non-meteorological contributions, but they may have limited ability to represent nonlinear chemical responses and interactions among meteorological variables (Liu X. et al, 2023). Machine learning provides an alternative framework for capturing nonlinear relationships between O_3 and meteorological predictors and has been increasingly used to remove meteorological variability from air-pollution time series (Grange and Carslaw, 2019; Vu et al., 2019; Grange et al., 2018; Ding et al., 2023; Dai et al., 2023).

Among machine-learning approaches, tree-based ensemble models have been widely applied in air-quality studies because of their ability to handle nonlinear responses, variable interactions, and large observational datasets. LightGBM is a computationally efficient gradient-boosting framework that can provide robust predictive performance for large-scale environmental data (Ke et al., 2017). When combined with meteorological normalization, such models can estimate the O_3 concentrations expected under average meteorological conditions and thereby separate meteorologically driven variability from longer-term non-meteorological changes (Grange et al., 2018; Vu et al., 2019; Dai et al., 2023). In addition, SHapley Additive exPlanations (SHAP) can be used to interpret trained machine-learning models by attributing model predictions to individual input variables (Lundberg and Lee, 2017; Lundberg et al., 2020). This provides a useful way to identify the meteorological variables most strongly associated with O_3 variability in the model, and recent studies have demonstrated the value of interpretable machine-learning approaches for analyzing O_3 pollution and its meteorological drivers (Weng et al., 2022; Zhang et al., 2024; Tan et al., 2024; Lyu et al., 2025). However, such attributions should be interpreted as model-based statistical associations rather than direct evidence of physical causality.

Despite these advances, several gaps remain in understanding recent O_3 changes in China. First, many previous studies have focused on individual cities, selected polluted regions, or warm-season O_3 episodes, whereas a consistent national-scale assessment over the most recent decade remains limited (Li et al., 2020; Wang T. et al, 2022; Wu et al., 2022; Ding et al., 2023). Second, warm-season O_3 has received much attention because of strong photochemical production and frequent exceedances, but cold-season O_3 changes are also important. During the cold season, weaker photochemistry, reduced NO titration, changes in background O_3, boundary-layer processes, and regional transport may all affect surface O_3 levels (Li et al., 2021; Tong et al., 2023; Wang W. et al, 2022). A direct comparison between warm and cold seasons is therefore needed to clarify whether the drivers of O_3 changes differ between high-photochemical and low-photochemical environments. Third, the spatial heterogeneity of meteorological modulation across major polluted regions remains insufficiently quantified, particularly under the recent changes in both emissions and climate variability (Liu and Wang, 2020; Weng et al., 2022; Wang X. et al., 2025).

Here, we investigate the interannual variability of maximum daily 8 h average O_3 concentration (MDA8 O_3) across China during the warm and cold seasons from 2015 to 2024. We first characterize the spatial distribution and temporal trends of observed MDA8 O_3 at the national scale and across major polluted regions. We then develop a LightGBM-based meteorological normalization framework to estimate meteorologically normalized O_3 concentrations and to quantify the relative contributions of meteorological variability and non-meteorological components to interannual O_3 changes. The non-meteorological component is interpreted as a composite signal that is likely influenced by changes in precursor emissions, background O_3, regional transport, and other slowly varying processes. Finally, we apply SHAP analysis to identify the meteorological variables most strongly associated with O_3 variability across regions and seasons. By comparing warm- and cold-season O_3 changes, this study aims to provide a seasonally resolved understanding of the meteorological modulation and long-term evolution of surface O_3 pollution in China.

2 Data and methods

2.1 Data

This study used hourly surface O_3 observations from the China National Environmental Monitoring Centre (CNEMC, http://www.cnemc.cn/en, last access: 19 April 2026) for the period 2015–2024. MDA8 O_3 was used as the primary metric because it is widely applied in studies of surface O_3 pollution and is closely related to air-quality assessment. Quality control was first applied to the hourly observations. Monitoring sites with more than 10% missing hourly records during the study period were excluded. For the remaining sites, short gaps in hourly O_3 records were filled using the K-nearest-neighbour (KNN) algorithm, with n_neighbors = 10, metric = "euclidean", and weights = "distance". Previous studies have shown that KNN-based imputation can provide reasonable performance for missing air-pollutant data while preserving the statistical distribution of pollutant concentrations (Saeipourdizaj et al., 2021). The filled hourly records were then used to calculate MDA8 O_3. After quality control, 1173 monitoring sites were retained for the subsequent analysis.

Previous studies of O_3 pollution in China have focused mainly on the warm season, but recent evidence indicates that O_3 increases are not confined to photochemically active months and can also extend into the cold season. To examine the seasonal contrast in O_3 variability, we defined May–October as the warm season and November–April as the cold season. For the cross-year cold season, November and December were assigned to the following cold-season year together with January–April. For example, the 2016 cold season refers to November–December 2015 and January–April 2016. This definition separates the high-photochemical and low-photochemical periods and is consistent with previous studies of seasonal O_3 variability in China (Wang T. et al, 2022).

To examine regional differences, we grouped representative cities into five major polluted regions: NCP, YRD, PRD, FWP, and SCB. The regional classification follows previous work on surface O_3 pollution and source-meteorology influences in China (Lu et al., 2019). The detailed city lists for each region are provided in Table S1.

Meteorological fields were obtained from the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5, https://cds.climate.copernicus.eu/, last access: 19 April 2026), with a horizontal resolution of 0.25° × 0.25° and an hourly temporal resolution. We selected meteorological variables that are closely related to O_3 photochemical production, atmospheric transport, dilution, and boundary-layer processes. These variables include 2 m temperature (T2M), surface solar radiation downwards (SSRD), total cloud cover (TCC), boundary-layer height (BLH), 10 m zonal and meridional wind components (U10 and V10), surface pressure (SP), relative humidity (RH), and total precipitation (TP). The daily aggregation of meteorological variables followed the approach of Weng et al. (2022), with different averaging periods selected to represent the meteorological conditions most relevant to daytime O_3 formation and transport (Table S2).

For spatial matching, ERA5 meteorological variables were extracted for each monitoring site using the nearest grid cell. Site-level daily MDA8 O_3 and meteorological variables were first calculated for all valid monitoring sites. For city-level analysis, the arithmetic mean of all valid sites within each city was used. Regional and national mean values were then calculated from the corresponding city-level daily series. This aggregation procedure reduces the influence of uneven monitoring-site density

https://doi.org/10.5194/egusphere-2026-2846
Preprint. Discussion started: 23 June 2026

among cities and provides consistent city-, regional-, and national-scale time series for subsequent trend analysis, meteorological normalization, and SHAP interpretation.

2.2 LightGBM-based meteorological normalization method

In this study, we developed a LightGBM-based prediction model to estimate daily MDA8 O_3 concentrations. LightGBM is a gradient-boosting decision-tree algorithm that is well suited for large environmental datasets because it can capture nonlinear relationships and interactions among predictors while maintaining high computational efficiency (Ke et al., 2017). The target variable was daily MDA8 O_3, and the input predictors included both temporal and meteorological variables. Temporal variables included Unix time, day of year (DOY), and day of week (DOW), which were used to represent long-term changes and recurring seasonal or weekly patterns that are not explicitly explained by meteorological variability. Meteorological predictors included T2M, SSRD, TCC, BLH, U10, V10, SP, RH, and TP, representing thermal, radiative, dynamic, moisture, and boundary-layer conditions relevant to O_3 formation and dispersion.

Independent LightGBM models were trained for each monitoring site. For each site, the daily dataset was randomly divided into training and testing subsets at an 80:20 ratio. Hyperparameters were optimized using grid search with 5-fold cross-validation to reduce overfitting and improve generalization. Model performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). These evaluation metrics were calculated for both warm and cold seasons to assess whether the model could reproduce the seasonal variability of MDA8 O_3.

Meteorological normalization was then applied to separate the component of O_3 variability associated with meteorological fluctuations from the longer-term non-meteorological component. The method follows previous meteorological normalization frameworks used in air-quality studies (Grange et al., 2018; Vu et al., 2019; Dai et al., 2023). For each target date, the temporal variables were kept unchanged, while the meteorological variables were randomly sampled from the same calendar period in other years. Specifically, meteorological conditions were sampled from a ±14 d window around the target day of year across the full study period. For example, for a target date of 15 July, meteorological variables were randomly selected from the period between 1 July and 29 July across all available years. This sampling strategy preserves the seasonal characteristics of meteorology while reducing the influence of year-specific meteorological anomalies.

For each target date, the resampled meteorological variables and fixed temporal variables were input into the trained LightGBM model to predict MDA8 O_3. This procedure was repeated 1000 times, and the average of the 1000 predictions was defined as the meteorologically normalized MDA8 O_3 concentration for that date. The normalized concentration represents the O_3 level expected under average meteorological conditions for the corresponding season and calendar period. Therefore, the difference between the observed and normalized concentrations can be interpreted as the meteorological anomaly contribution, whereas the normalized time series reflects the long-term non-meteorological component. This non-meteorological component should not be interpreted as the isolated effect of anthropogenic emissions alone; rather, it represents a composite signal that may include changes in precursor emissions, background O_3, regional transport, natural emissions, and other slowly varying processes not explicitly included in the model.

https://doi.org/10.5194/egusphere-2026-2846
Preprint. Discussion started: 23 June 2026

Sensitivity tests were conducted to evaluate the influence of the normalization parameters, including the sampling window size and the number of iterations (Figs. S1 and S2). The interannual trends of normalized MDA8 O_3 remained stable when the sampling window varied from ±7 to ±28 d, indicating that the selected ±14 d window provided a reasonable balance between meteorological representativeness and sample size. The standard deviation of the normalized predictions decreased rapidly as the number of iterations increased and stabilized at approximately 1000 iterations. Therefore, 1000 iterations were used in the final meteorological normalization procedure.

The meteorological contribution was quantified as the difference between observed and meteorologically normalized MDA8 O_3:

$$M_i = O_i - N_i$$

where M_i is the meteorological contribution in year i, O_i is the observed seasonal mean MDA8 O_3 concentration in year i, and N_i is the corresponding meteorologically normalized concentration.

The long-term non-meteorological contribution was evaluated from the interannual trend or change in N_i. For a given period from year i to year j, the observed O_3 change can be decomposed as:

$$\Delta O = O_j - O_i$$

$$\Delta N = N_j - N_i$$

$$\Delta M = \Delta O - \Delta N$$

where ΔN represents the non-meteorological component of the O_3 change, and ΔM represents the meteorological contribution to the observed change. This decomposition was applied consistently at the national and regional scales for both warm and cold seasons.

2.3 SHapley Additive exPlanations (SHAP) approach

To interpret the trained LightGBM models and identify the meteorological variables most strongly associated with MDA8 O_3 variability, we applied the SHapley Additive exPlanations (SHAP) method (Lundberg and Lee, 2017; Lundberg et al., 2020). SHAP is based on cooperative game theory and decomposes each model prediction into the sum of a baseline value and the contributions assigned to individual input features. For a given prediction, the SHAP representation can be expressed as:

$$f(x) = \varphi_0 + \sum_{i=1}^{M} \varphi_i$$

where $f(x)$ is the LightGBM-predicted MDA8 O_3 concentration for sample x, φ_0 is the baseline model output, defined as the expected prediction over the background dataset, M is the number of input variables, and φ_i is the SHAP value of predictor i. A positive SHAP value indicates that the predictor increases the predicted MDA8 O_3 concentration relative to the baseline, whereas a negative value indicates a decrease. The absolute magnitude of the SHAP value indicates the strength of the model-attributed contribution of that predictor for a specific prediction.

In this study, SHAP values were calculated for the meteorological variables included in the LightGBM model, including T2M, SSRD, TCC, BLH, U10, V10, SP, RH, and TP. To assess the relative importance of meteorological variables, we used the mean absolute SHAP value for each variable. The variable with the largest mean absolute SHAP value at a site was identified as the dominant model-attributed meteorological factor for that site. The spatial distribution of the top-ranked meteorological variables was then analyzed separately for the warm and cold seasons.

For regional analysis, SHAP values were aggregated to quantify the relative contribution of each meteorological variable to MDA8 O_3 variability in each key region. Positive and negative SHAP values were retained when calculating annual net meteorological effects, allowing us to examine whether meteorological conditions in a given year tended to increase or decrease predicted MDA8 O_3 concentrations relative to the baseline. The regional mean SHAP values were calculated from city-level or site-level averages using the same aggregation framework as that used for O_3 concentrations.

It should be noted that SHAP values provide a model-based interpretation of statistical associations between predictors and predicted O_3 concentrations. They do not by themselves establish physical causality. Therefore, the SHAP results were interpreted together with known atmospheric processes, such as photochemical production, ventilation, humidity-related suppression of photochemistry, and boundary-layer mixing.

3 Results and discussion

3.1 Performance of the machine learning model

Figure 1 shows the performance of the LightGBM models in predicting daily MDA8 O_3 concentrations across China during the warm and cold seasons. The model performance was evaluated using independent testing datasets for each monitoring site, and the spatial distributions of R^2 are shown in Fig. 1a and b. The density scatter plots between observed and predicted MDA8 O_3 concentrations are shown in Fig. 1c and d. Overall, the LightGBM models reproduced the observed daily variability of MDA8 O_3 reasonably well in both seasons, although the model skill varied among regions and monitoring sites.

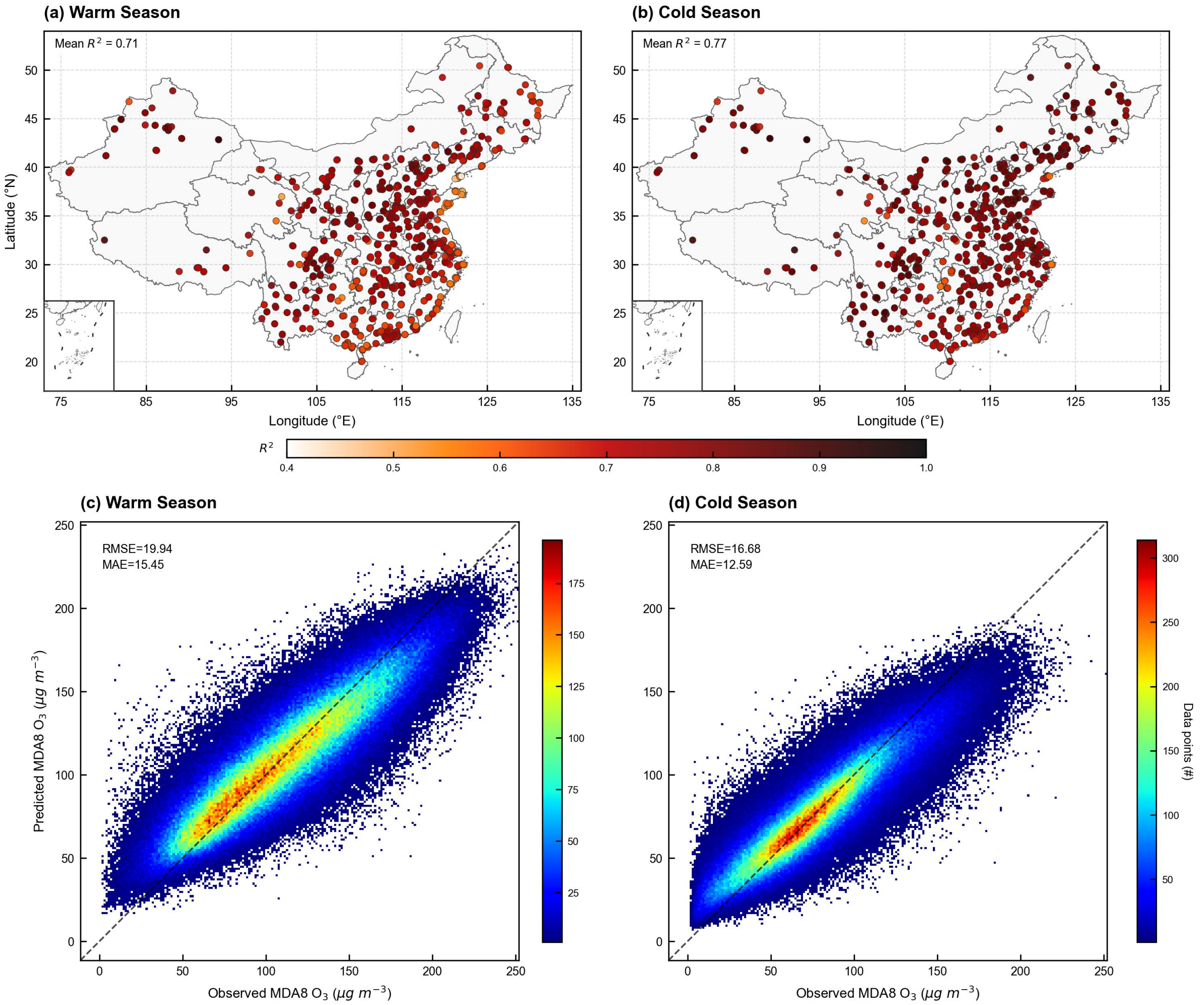

Figure 1: Performance of the LightGBM models for daily MDA8 O_3 prediction. (a, b) Spatial distributions of site-level testing R^2 during the warm and cold seasons, respectively. (c, d) Density scatter plots of observed versus predicted daily MDA8 O_3 concentrations during the warm and cold seasons, respectively. The dashed lines indicate the 1:1 line.

During the warm season, the mean testing R^2 was 0.71, with values ranging from 0.42 to 0.89 across monitoring sites. The corresponding RMSE and MAE were 19.9 and 15.5 μg m^{-3}, respectively. During the cold season, the model showed slightly better performance, with a mean testing R^2 of 0.77 and a site-level range of 0.54–0.92. The RMSE and MAE were 16.7 and 12.6 μg m^{-3}, respectively. The observed and predicted values were generally distributed around the 1:1 line, indicating that the model captured the major variations in daily MDA8 O_3 without strong systematic bias.

The better model performance in the cold season may be related to the weaker photochemical activity and relatively lower variability in O_3 concentrations compared with the warm season. In contrast, warm-season O_3 is more strongly influenced by nonlinear photochemical production, episodic high-temperature events, and regional transport, which may increase the difficulty of prediction. Nevertheless, the testing results indicate that the LightGBM models are able to represent the main temporal variability of MDA8 O_3 in both seasons. The model performance is comparable to that reported in recent machine-learning studies of O_3 variability in China (Mao et al., 2024; Wang Y., Gong C et al, 2025; Ye et al., 2026; Chen et al., 2026), providing a reasonable basis for subsequent meteorological normalization and SHAP-based interpretation.

3.2 Spatiotemporal patterns of MDA8 O_3 across China

Figure 2 shows the spatial distributions of mean MDA8 O_3 concentrations across China during the warm and cold seasons from 2015 to 2024. Clear seasonal differences were observed in both the magnitude and spatial pattern of surface O_3. During the warm season, the national mean MDA8 O_3 concentration was 110.1 ± 4.4 μg m^{-3}, indicating widespread enhancement of surface O_3 under photochemically active conditions. High MDA8 O_3 concentrations were mainly distributed over northern and eastern China, with pronounced regional clustering in the NCP, FWP, and YRD. The NCP showed the highest regional mean MDA8 O_3 concentration, reaching 135.4 ± 6.5 μg m^{-3}, followed by the FWP at 126.7 ± 7.8 μg m^{-3} and the YRD at 124.1 ± 5.7 μg m^{-3}. Several cities in Shandong Province, including Dezhou, Dongying, and Jinan, had warm-season mean MDA8 O_3 concentrations exceeding 150 μg m^{-3}. These high values are likely associated with strong photochemical production under high temperature and intense solar radiation, together with abundant anthropogenic precursor emissions in densely industrialized and urbanized areas.

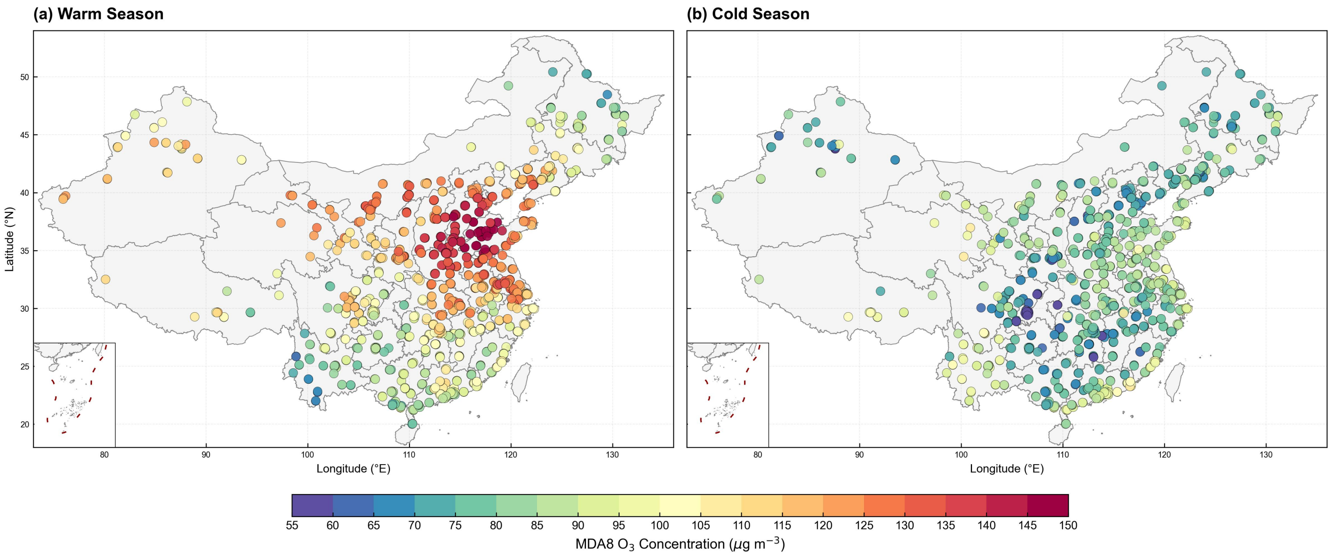

Figure 2: Spatial distributions of mean MDA8 O_3 concentrations across China during 2015–2024. (a) Warm season and (b) cold season. The warm season is defined as May–October, and the cold season is defined as November–April.

Lower warm-season MDA8 O_3 concentrations were mainly found in parts of Southwest and Northeast China. For example, Xishuangbanna and Pu'er in Yunnan Province had mean concentrations of 62.8 and 65.6 μg m^{-3}, respectively, while Yichun in Heilongjiang Province had a mean concentration of 70.0 μg m^{-3}. These relatively low values may be related to weaker photochemical conditions, lower anthropogenic precursor emissions, stronger ventilation, or more frequent cloudy and rainy conditions. Among the key regions, the SCB had a warm-season mean MDA8 O_3 concentration of 104.2 ± 5.0 μg m^{-3}, while the PRD showed the lowest warm-season level, at 100.3 ± 4.3 μg m^{-3}. The relatively lower warm-season O_3 in the PRD may reflect the combined effects of monsoonal ventilation, cloudy and rainy conditions, and regional transport processes, consistent with previous studies of O_3 variability in southern China (Li et al., 2020; Cao T., Wang H., Li L., et al, 2024).

During the cold season, the national mean MDA8 O_3 concentration decreased to 80.5 ± 3.7 μg m^{-3}, reflecting the seasonal weakening of photochemical O_3 production. In contrast to the warm-season pattern, high-O_3 areas shifted southward, and the PRD became the most prominent cold-season O_3 region among the five key regions, with a mean MDA8 O_3 concentration of 86.2 ± 7.4 μg m^{-3}. Shantou had a cold-season mean concentration of 105.1 μg m^{-3}. This spatial pattern is likely related to the relatively warm climate in southern China during winter, changes in monsoonal transport, and weaker NO titration under changing precursor emissions (Wang W. et al, 2022; Cao T., Wang H., Li L., et al, 2024). By comparison, the YRD, NCP, and FWP showed lower cold-season mean concentrations of 85.0 ± 4.5, 80.0 ± 3.8, and 74.3 ± 5.2 μg m^{-3}, respectively, while the SCB remained at a relatively low level of 70.2 ± 3.0 μg m^{-3}. These results indicate that surface O_3 pollution in China exhibits a clear seasonal redistribution, with warm-season high values concentrated mainly in northern and eastern polluted regions and cold-season enhancement becoming more evident in southern coastal regions.

Figure 3 further shows the interannual variations in MDA8 O_3 concentrations during the warm and cold seasons from 2015 to 2024. At the national scale, observed MDA8 O_3 concentrations increased significantly in both seasons. Approximately 89 % of monitoring sites showed increasing MDA8 O_3 trends in both the warm and cold seasons, and 41.0 % and 40.6 % of sites showed statistically significant increases in the warm and cold seasons, respectively ($p < 0.05$). These widespread increases indicate that surface O_3 pollution in China has continued to intensify during the past decade, not only during the photochemically active warm season but also during the cold season.

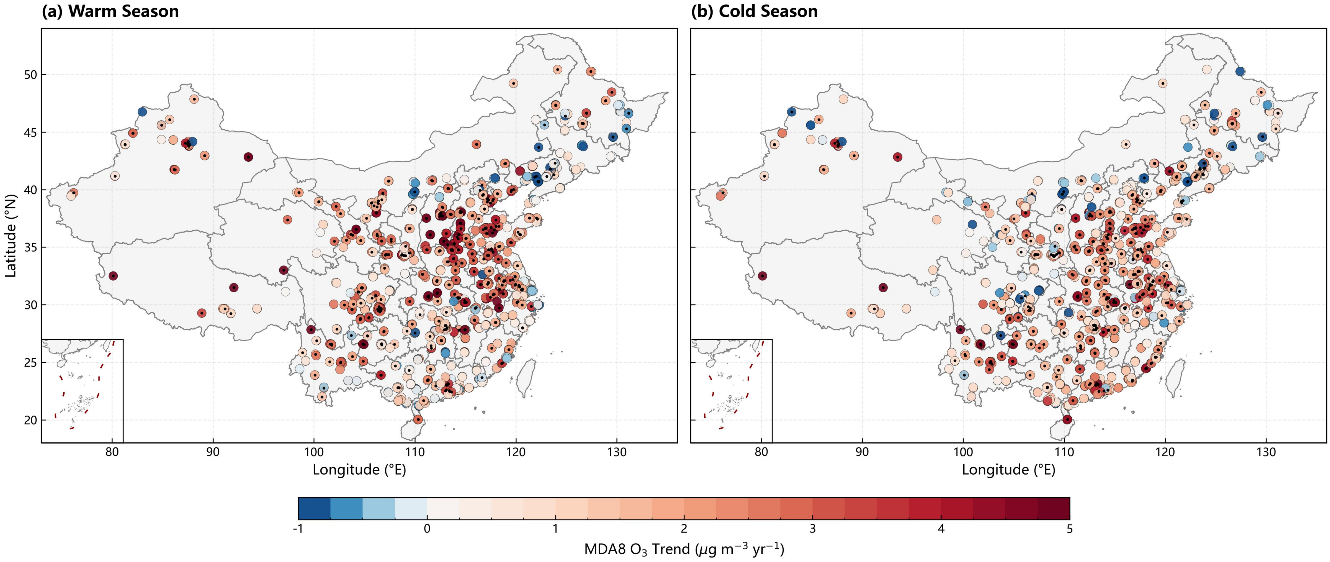

Figure 3: Interannual variations in MDA8 O_3 concentrations across China from 2015 to 2024,where dots denote statistical significance ($p < 0.05$). (a) Warm-season MDA8 O_3 and (b) cold-season MDA8 O_3.

During the warm season, the national mean MDA8 O_3 concentration increased at a rate of 2.1 μg m^{-3} yr^{-1}, rising from 97.6 ± 21.3 μg m^{-3} in 2015 to 116.8 ± 22.9 μg m^{-3} in 2024, corresponding to a cumulative increase of approximately 19.7 %. This increasing trend is consistent with previous studies reporting sustained O_3 enhancement in China under the combined influence of changing precursor emissions and meteorological variability (Dang and Liao, 2019; Ren et al., 2022; Fang et al., 2025; Goto et al., 2026). Regionally, the FWP showed the strongest warm-season increase, with a trend of 3.5 μg m^{-3} yr^{-1} and a cumulative increase of 42.6 %. By 2024, its regional mean MDA8 O_3 concentration reached 142.1 ± 2.2 μg m^{-3}, exceeding the level in the NCP. The NCP, YRD, and SCB also showed marked increases, with trends of 2.7, 2.4, and 2.1 μg m^{-3} yr^{-1}, respectively. In contrast, the PRD exhibited a more moderate warm-season increase of 1.4 μg m^{-3} yr^{-1} (Wang T. et al, 2022; Wu et al., 2022; Zhang et al., 2025). Overall, warm-season O_3 increased more rapidly in northern and inland regions than in southern coastal regions.

Cold-season MDA8 O_3 concentrations also showed a widespread increase, with a national mean trend of 1.7 μg m^{-3} yr^{-1}. Unlike the warm season, the most prominent cold-season increase occurred in the PRD, where MDA8 O_3 increased at a rate of 2.9 μg m^{-3} yr^{-1} and reached 89.6 ± 3.5 μg m^{-3} in 2024. Other key regions, including the NCP, YRD, FWP, and SCB, also showed sustained increases during the cold season. These results suggest that the recent rise in surface O_3 in China is not limited to the warm season. The cold-season increase may reflect the combined effects of changing precursor emissions, weakened NO titration, background O_3 changes, regional transport, and meteorological variability, highlighting the need to consider season-specific processes in O_3 pollution control (Li et al., 2021; Tong et al., 2023; Cao T., Wang H., Chen X., et al, 2024).

3.3 Interannual variations in meteorologically normalized MDA8 O_3

Figure 4 shows the spatial distributions of interannual trends in meteorologically normalized MDA8 O_3 during the warm and cold seasons from 2015 to 2024. Meteorological normalization was applied to reduce the influence of short-term and interannual meteorological fluctuations and to extract the long-term component of O_3 variability that cannot be explained by meteorology in the model framework. After normalization, MDA8 O_3 still showed increasing trends at the national scale and across most key regions in both seasons, indicating that the observed O_3 increase was not solely driven by meteorological variability.

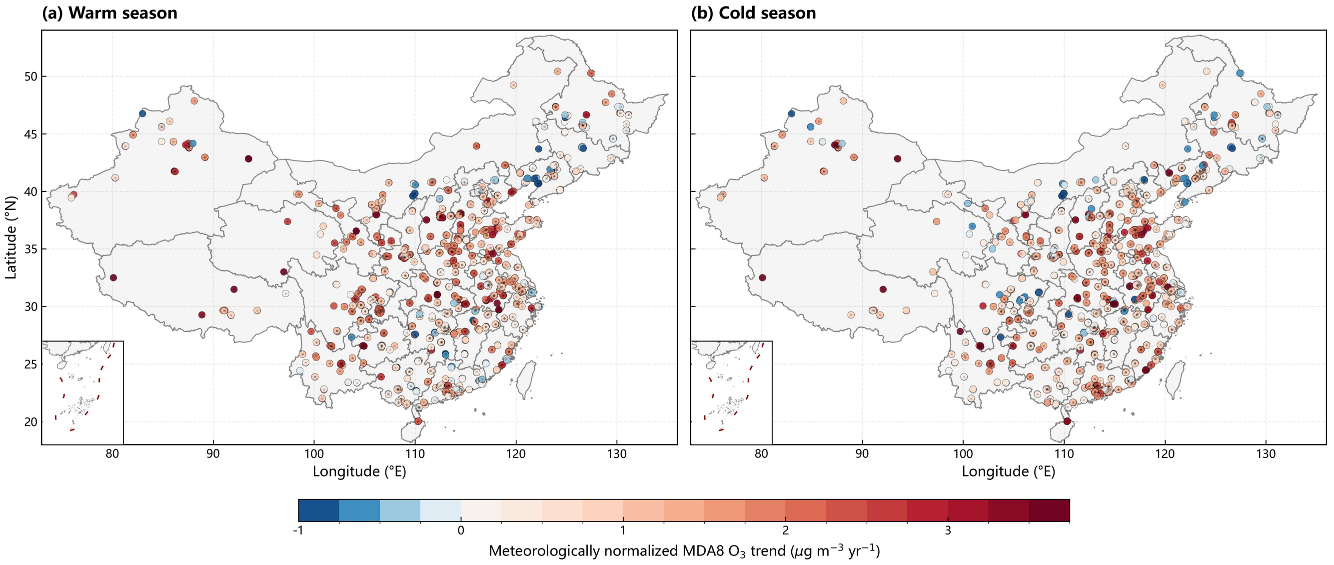

Figure 4: Spatial distributions of trends in meteorologically normalized MDA8 O_3 across China during 2015–2024, where dots denote statistical significance ($p < 0.05$). (a) Warm season and (b) cold season. Trends were calculated from the meteorologically normalized seasonal mean MDA8 O_3 time series at each monitoring site.

At the national scale, observed warm-season MDA8 O_3 increased from 97.6 ± 21.3 μg m^{-3} in 2015 to 116.8 ± 22.9 μg m^{-3} in 2024, corresponding to a trend of 2.1 μg m^{-3} yr^{-1} ($p < 0.01$). The meteorological normalization analysis indicates that the non-meteorological component contributed 16.1 μg m^{-3} to the cumulative increase, whereas meteorological fluctuations contributed approximately 3.1 μg m^{-3}. After removing the meteorological contribution, the normalized warm-season MDA8 O_3 still increased significantly, with a trend of 1.2 μg m^{-3} yr^{-1}. This result suggests that long-term non-meteorological changes played a dominant role in the warm-season O_3 increase, consistent with previous studies that emphasized the importance of precursor-emission changes and chemical-regime shifts in recent O_3 increases in China (Li et al., 2020; Liu et al., 2023b). However, the positive meteorological contribution also indicates that meteorological variability modulated the magnitude of the observed increase, particularly during years with more favorable conditions for photochemical O_3 production (Gao et al., 2023; Meng et al., 2023).

During the cold season, observed MDA8 O_3 also increased substantially, from 74.9 ± 1.8 μg m^{-3} in 2015 to 86.4 ± 1.1 μg m^{-3} in 2024. Among the cumulative cold-season increase, the non-meteorological component contributed 10.2 μg m^{-3}, while meteorological variability contributed approximately 1.3 μg m^{-3}. The meteorologically normalized cold-season MDA8 O_3 increased at the same rate as that in the warm season, reaching 1.2 μg m^{-3} yr^{-1}. This finding indicates that the cold-season O_3 increase was also largely associated with long-term non-meteorological changes, rather than being a purely meteorological phenomenon. The increase is consistent with a weakened NO titration effect under declining NOx emissions and with changes in background O_3 and regional transport, although these processes cannot be fully separated using meteorological normalization alone (Le et al., 2020; Li et al., 2021; Tong et al., 2023). Meteorological conditions may further influence cold-season O_3 through changes in solar radiation, boundary-layer stability, and ventilation, which affect photochemical production and near-surface accumulation under low-photochemical environments (Xiang et al., 2019).

Figure 5 further compares the observed, meteorological, and non-meteorological components of MDA8 O_3 in the five key regions. During the warm season, all key regions showed statistically significant increases in meteorologically normalized MDA8 O_3 ($p < 0.05$), although the magnitude of the increase differed among regions. The SCB showed the strongest normalized increase, with a trend of 1.4 μg m^{-3} yr^{-1}, followed by the FWP 1.4 μg m^{-3} yr^{-1} and the NCP at 1.2 μg m^{-3} yr^{-1}. The YRD and PRD showed more moderate normalized increases of 1.0 and 0.8 μg m^{-3} yr^{-1}, respectively. These regional differences suggest that the long-term non-meteorological component of warm-season O_3 increased more rapidly in northern inland and basin regions than in eastern and southern coastal regions. This pattern may be related to regional differences in precursor-emission changes, chemical sensitivity regimes, topographic constraints, and ventilation conditions (Yang et al., 2019; Liu Y. et al, 2023).

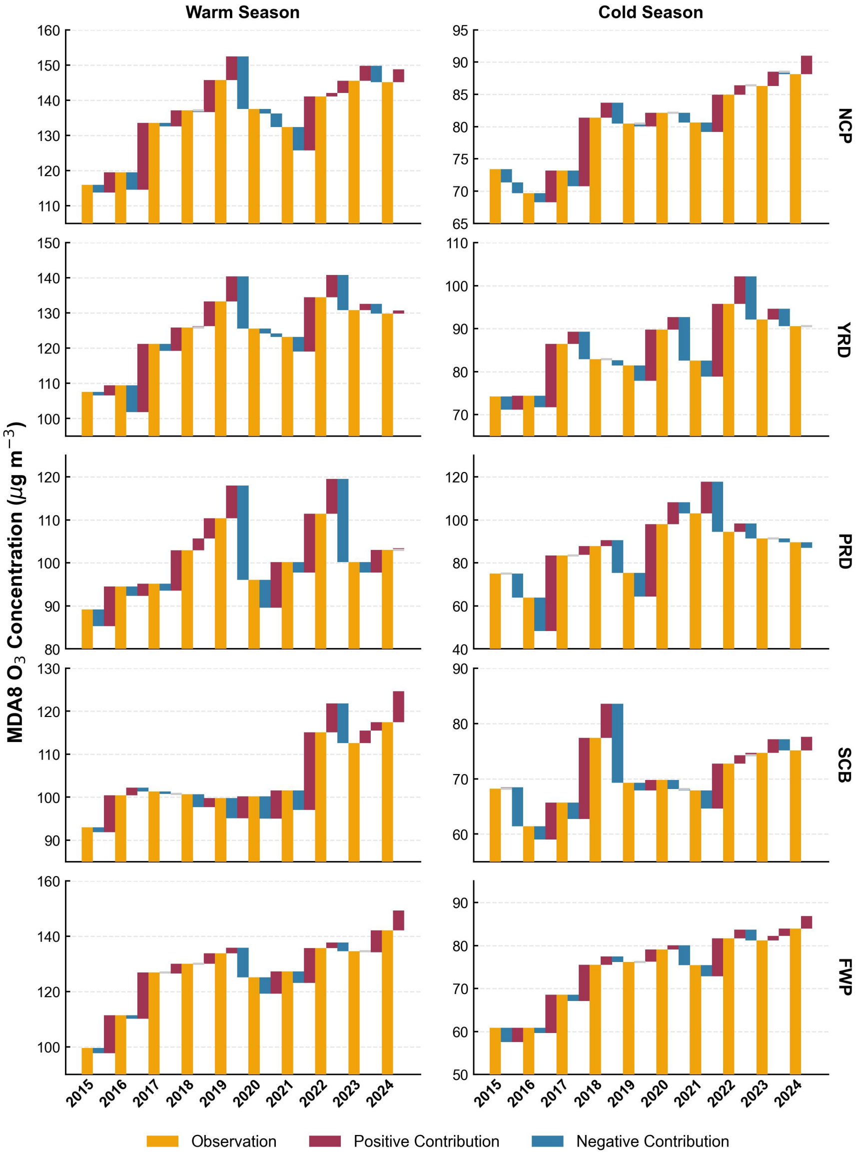

Figure 5: Decomposition of regional MDA8 O_3 variations into meteorological and non-meteorological components during 2015–2024. Stacked bars show the contributions of meteorological variability and the meteorologically normalized non-meteorological component to observed MDA8 O_3 in the five key regions during the warm and cold seasons.

The temporal evolution of normalized warm-season MDA8 O_3 showed three broad stages across the key regions: a rapid increase during 2015–2019, a temporary decline around 2020, and an oscillatory rebound during 2021–2024. The decline around 2020 may reflect short-term perturbations in emissions and atmospheric chemistry associated with the COVID-19 period, as well as interannual variability in meteorological and regional transport conditions. Because the normalized component still contains multiple non-meteorological processes, this stage should be interpreted as a change in the long-term residual component rather than as a direct measure of emission reductions alone.

During the cold season, meteorologically normalized MDA8 O_3 also increased in all key regions, but the regional pattern differed from that of the warm season. In several regions, the cold-season normalized trend exceeded the corresponding warm-season trend. The FWP showed the largest cold-season normalized increase, reaching 1.9 μg m^{-3} yr^{-1}, while the YRD and PRD increased by 1.4 and 1.5 μg m^{-3} yr^{-1}, respectively. These rates were higher than their warm-season normalized trends. In contrast, the NCP and SCB showed relatively weaker cold-season normalized increases of 1.1 and 0.5 μg m^{-3} yr^{-1}, respectively, lower than their warm-season trends. These differences indicate that the processes controlling long-term O_3 changes are seasonally dependent and regionally heterogeneous. In particular, the strong cold-season normalized increase in the PRD suggests that cold-season O_3 enhancement in southern coastal China deserves more attention in future O_3 control strategies.

Overall, the meteorological normalization results indicate that non-meteorological changes were the dominant contributor to the long-term increase in MDA8 O_3 during both warm and cold seasons. Meteorological variability contributed less to the decadal trend but played an important role in modulating the year-to-year magnitude of O_3 changes. These findings highlight the need to distinguish long-term non-meteorological trends from meteorological fluctuations when evaluating recent O_3 changes and the effectiveness of air-quality control measures.

3.4 SHAP-based interpretation of meteorological influences on O_3

The meteorological normalization analysis indicates that the long-term increase in MDA8 O_3 was mainly associated with the non-meteorological component, whereas meteorological variability modulated the magnitude and interannual fluctuations of O_3. To further examine which meteorological variables were most strongly associated with O_3 variability in the LightGBM model, we applied SHAP analysis to the trained models. The SHAP results provide a model-based interpretation of the meteorological contributions to predicted MDA8 O_3, rather than direct evidence of physical causality. Therefore, the following interpretation focuses on the consistency between the SHAP-derived variable importance and known atmospheric processes, including photochemical production, humidity-related suppression, ventilation, transport, and boundary-layer mixing.

Figure 6a shows the spatial distribution of the top-ranked meteorological variables for warm-season MDA8 O_3. At the national scale, SSRD, RH, and T2M were the most frequently identified dominant variables, accounting for 39.05 %, 37.17 %, and 17.48 % of monitoring sites, respectively. This indicates that warm-season O_3 variability in the model was mainly associated with radiative, moisture, and thermal conditions. The dominant variables exhibited clear spatial heterogeneity. In northern inland and basin regions, such as the NCP, FWP, and SCB, O_3 variability was more strongly associated with temperature and radiation. In the NCP, T2M was the dominant variable at 56.0 % of monitoring sites, with a mean absolute SHAP value of 15.2 μg m^{-3}. In the FWP and SCB, SSRD was identified as the dominant variable at 66.7 % and 97.2 % of monitoring sites, with mean absolute SHAP values of 16.0 and 17.0 μg m^{-3}, respectively.

https://doi.org/10.5194/egusphere-2026-2846
Preprint. Discussion started: 23 June 2026

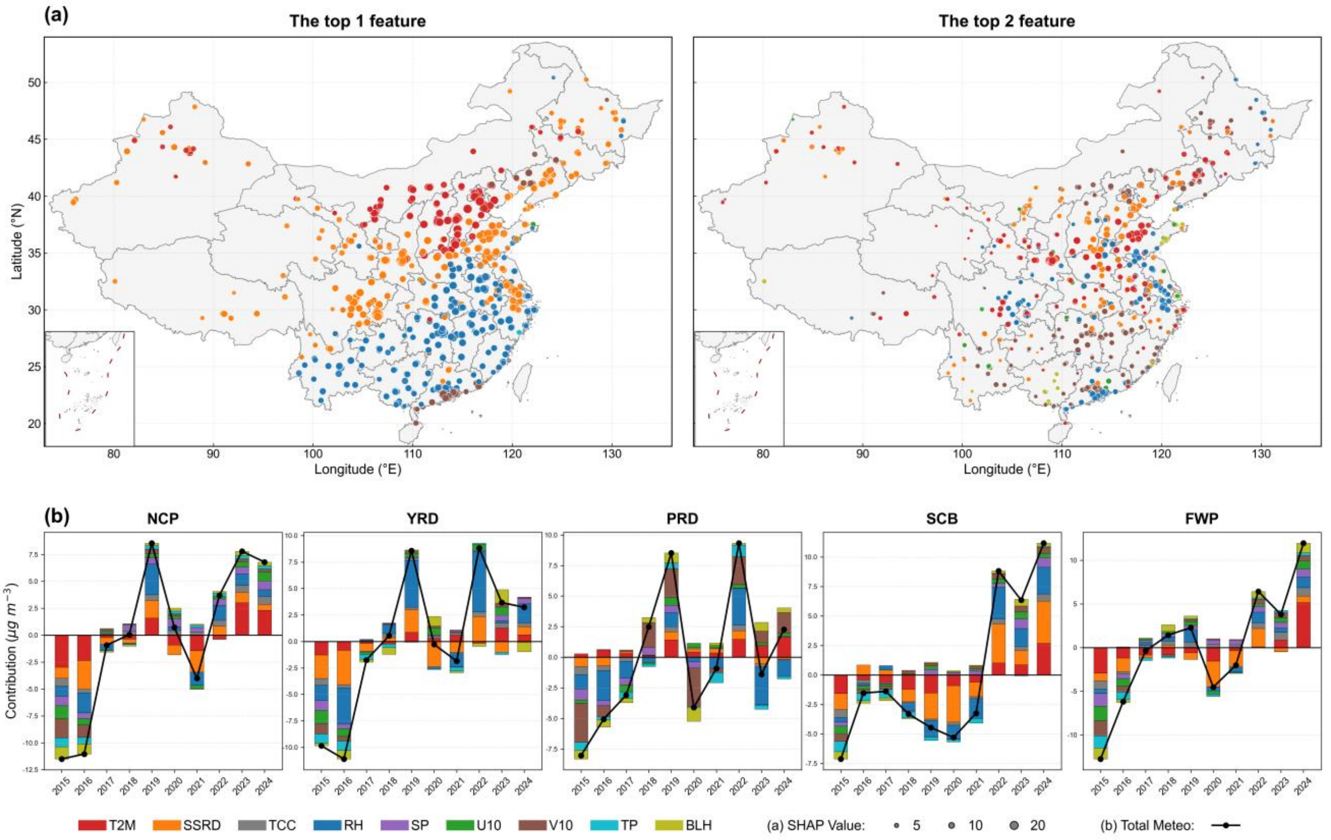

Figure 6: SHAP-based interpretation of warm-season meteorological influences on MDA8 O_3 during 2015–2024. (a) Spatial distributions of the top-ranked meteorological variables at monitoring sites. Colors indicate different meteorological variables, and marker sizes indicate the corresponding mean absolute SHAP values. (b) Annual SHAP-attributed contributions of meteorological variables to warm-season MDA8 O_3 in the five key regions. Colored bars represent individual meteorological variables, and black lines indicate the net meteorological contribution.

The dominance of T2M and SSRD in these northern and basin regions is consistent with the role of high temperature and strong solar radiation in enhancing photochemical O_3 production during the warm season. High temperatures can accelerate photochemical reaction rates and may also enhance biogenic VOC emissions, while strong solar radiation promotes photolysis reactions that initiate O_3 formation (Wang et al., 2024). In basin regions such as the SCB and FWP, topographic constraints and relatively stagnant conditions may further favor O_3 accumulation when photochemical production is strong. However, these SHAP-derived patterns should be interpreted as model-attributed associations. Additional chemical observations or process-based modeling would be needed to fully separate the effects of precursor emissions, chemistry, and transport.

In contrast, southern and coastal regions showed stronger associations with moisture and wind-field variables. In the YRD, SSRD and RH had comparable importance, with SSRD identified as the dominant variable at 52.0 % of monitoring sites and RH at 48.0 %. Their mean absolute SHAP values were also similar, at 12.0 and 12.1 μg m^{-3}, respectively. In the PRD, RH was the most frequently identified dominant variable, accounting for 51.0 % of monitoring sites, followed by V10, which accounted for 47.1 %. Their mean absolute SHAP values were 12.4 and 12.1 μg m^{-3}, respectively. This pattern suggests that warm-season O_3 variability in southern coastal China is strongly linked to humid monsoonal air masses, cloudiness, and regional ventilation. High RH may act as an indicator of moist and cloudy conditions that suppress photochemical O_3 production through reduced solar radiation and enhanced atmospheric mixing, while V10 reflects the meridional wind component associated with regional transport and ventilation. This interpretation is broadly consistent with previous studies emphasizing the role of monsoonal circulation and regional transport in southern China (Mai et al., 2024).

Figure 6b shows the annual contributions of individual meteorological variables and the net meteorological contribution to warm-season MDA8 O_3 in the five key regions. The net meteorological contribution exhibited a clear temporal transition during 2015−2024. During 2015−2017, meteorological conditions generally made negative contributions to warm-season MDA8 O_3 in most key regions. In 2015, the negative meteorological contributions in the NCP and FWP reached -11.5 and -12.8 μg m^{-3}, respectively, indicating that meteorological conditions in these regions were less favorable for O_3 enhancement relative to the model baseline. Around 2019, the net meteorological contribution increased markedly in several regions. In the NCP, YRD, and PRD, it rose to approximately 8.5 μg m^{-3}, while the FWP also showed a positive contribution of 2.3 μg m^{-3}. In contrast, the SCB retained a negative net meteorological contribution of -4.5 μg m^{-3} during the same year.

During 2022–2024, meteorological conditions became more favorable for warm-season O_3 enhancement in several key regions. The SCB showed the most pronounced shift, with the average meteorological contribution during 2022–2024 increasing by 12.3 μg m^{-3} relative to 2015–2019. By 2024, the net meteorological contributions in the FWP and SCB reached 11.9 and 11.2 μg m^{-3}, respectively, the highest values during the study period. These positive contributions were likely associated with the combined effects of high temperature, strong solar radiation, and unfavorable dispersion conditions, which can enhance photochemical production and near-surface accumulation of O_3 (Wang Y.,Wang Y. et al, 2025; Wang Z. et al, 2025). These results indicate that recent meteorological anomalies amplified warm-season O_3 levels in several regions and may have partly offset the apparent benefits of precursor-emission controls.

Figure 7a shows the dominant meteorological variables for cold-season MDA8 O_3. Compared with the warm season, the cold-season SHAP pattern was more spatially homogeneous. SSRD was identified as the dominant meteorological variable at 80.48 % of monitoring sites nationwide, while RH and T2M were dominant at only 9.97 % and 9.04 % of sites, respectively. In the NCP, YRD, SCB, and FWP, SSRD showed a dominant role in the model, with mean absolute SHAP values ranging from 16.5 to 17.7 μg m^{-3}. This result suggests that, under cold-season conditions with weaker photochemistry, variations in solar radiation can still strongly affect MDA8 O_3 by modulating photolysis rates and net photochemical production (Li et al., 2026). T2M generally acted as a secondary variable in these regions.

https://doi.org/10.5194/egusphere-2026-2846
Preprint. Discussion started: 23 June 2026

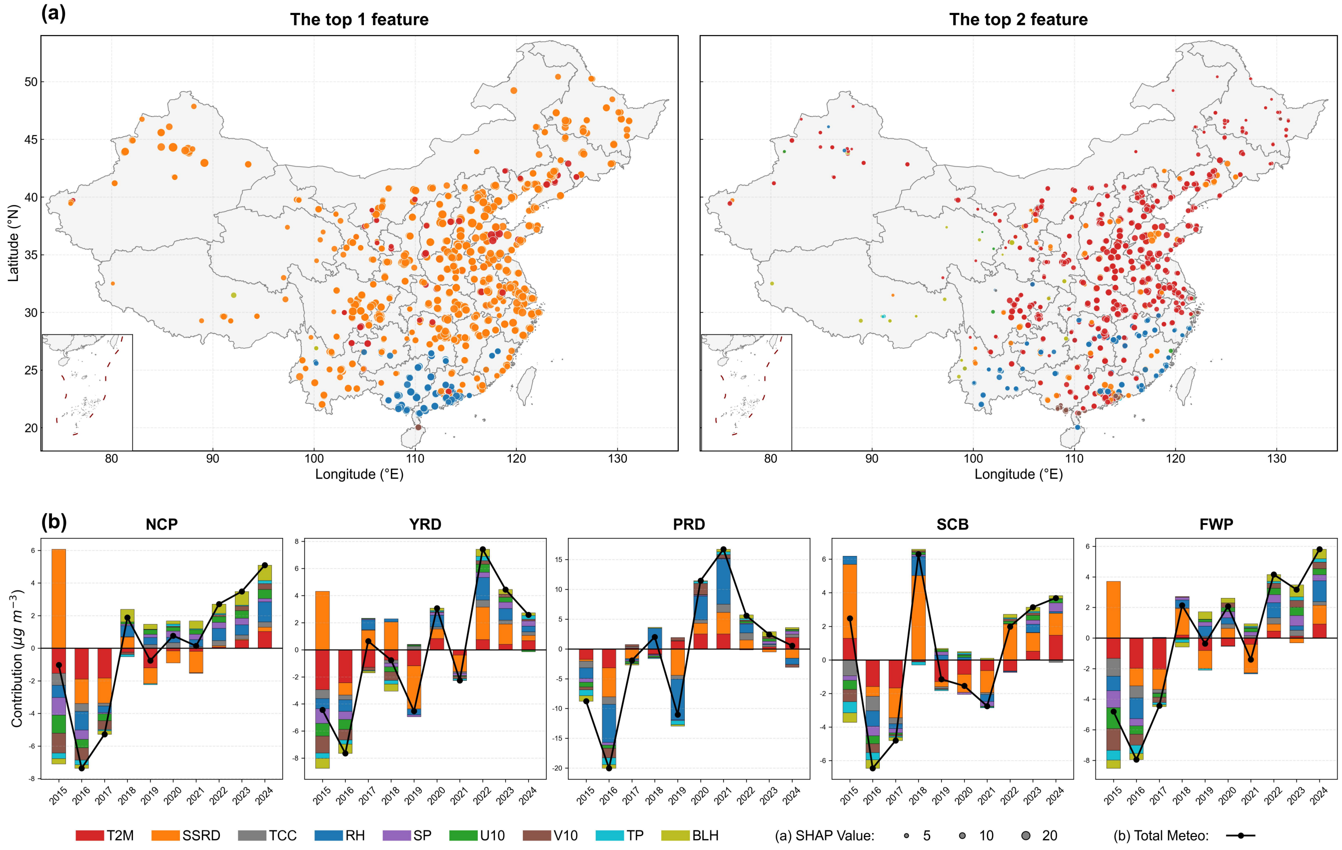

Figure 7: SHAP-based interpretation of cold-season meteorological influences on MDA8 O$_3$ during 2015–2024. (a) Spatial distributions of the top-ranked meteorological variables at monitoring sites. Colors indicate different meteorological variables, and marker sizes indicate the corresponding mean absolute SHAP values. (b) Annual SHAP-attributed contributions of meteorological variables to cold-season MDA8 O$_3$ in the five key regions. Colored bars represent individual meteorological variables, and black lines indicate the net meteorological contribution.

The PRD exhibited a different cold-season pattern. RH exceeded SSRD and became the dominant variable at 58.8 % of monitoring sites, with a mean absolute SHAP value of 12.8 μg m^{-3}. This suggests that cold-season O$_3$ variability in the PRD is closely associated with moisture-related meteorological conditions. In this region, RH may represent the combined influence of humid air masses, cloud cover, radiation reduction, and local or regional transport processes. Therefore, the RH contribution should not be interpreted as a purely direct chemical effect; rather, it likely reflects the integrated influence of moisture conditions and associated meteorological backgrounds on O$_3$ formation and accumulation.

Figure 7b shows that the cold-season net meteorological contribution also shifted from negative to positive during the study period. The national mean meteorological contribution was -3.3 μg m^{-3} during 2015–2019 but increased to 3.8 μg m^{-3} during 2022 – 2024. In northern and inland regions, including the NCP, YRD, SCB, and FWP, the interannual variation in net meteorological contribution showed relatively coherent temporal behavior, mainly associated with the contribution of SSRD. This is consistent with the cold-season dominance of SSRD shown in Fig. 7a. Because photochemical production is generally suppressed under low temperature and weak radiation, even moderate interannual changes in solar radiation may produce measurable differences in O_3 production and accumulation.

The PRD showed a more distinct cold-season meteorological pattern. The net meteorological contribution in this region varied widely, ranging from -20.0 to 16.7 μg m^{-3} during 2015–2024. RH was the dominant SHAP-attributed variable, with annual contributions ranging from -6.8 to 7.6 μg m^{-3}. This large variability suggests that cold-season O_3 in the PRD is particularly sensitive to moisture-related circulation and air-mass changes. High-RH conditions may be associated with enhanced cloudiness and reduced solar radiation, whereas drier conditions may coincide with stronger photochemical activity or different regional transport pathways. These results highlight the distinct meteorological sensitivity of cold-season O_3 in southern coastal China.

Overall, SHAP analysis reveals clear seasonal and regional differences in the meteorological variables associated with MDA8 O_3 variability. During the warm season, temperature and radiation were more important in northern and basin regions, whereas humidity and wind-field variables were more important in southern coastal regions. During the cold season, SSRD dominated most regions, while RH played a stronger role in the PRD. These results suggest that meteorological modulation of O_3 is highly season- and region-dependent, supporting the need for season-specific and region-specific O_3 control strategies.

3.5 Uncertainty analysis

Although the LightGBM-based meteorological normalization framework provides a useful approach for separating meteorologically driven variability from longer-term O_3 changes, several sources of uncertainty should be acknowledged. First, long-term and spatially consistent observations of O_3 precursors, especially VOCs and NOx, are not available at the national scale for the full study period. Therefore, precursor concentrations or emissions were not included as direct predictors in the model. As a result, the meteorologically normalized component should be interpreted as a non-meteorological residual rather than as the isolated effect of anthropogenic emission changes. This component may include the combined influences of anthropogenic precursor-emission changes, changes in chemical sensitivity regimes, biogenic VOC variability, background O_3, regional transport, and other slowly varying processes not explicitly represented in the model.

Second, uncertainties may arise from the meteorological input data. ERA5 provides spatially and temporally continuous meteorological fields, but its 0.25° × 0.25° spatial resolution may not fully capture local meteorological features near monitoring sites, especially in complex terrain, coastal areas, and highly urbanized regions. Local-scale processes such as urban heat-island effects, land-sea breezes, mountain-valley circulations, and boundary-layer heterogeneity may therefore be underrepresented. These uncertainties could influence the estimated meteorological contribution, particularly during short-lived or localized high-O_3 episodes.

Third, uncertainties are also associated with the machine-learning and interpretation framework. Although the LightGBM model captured the main variability in MDA8 O_3, meteorological normalization relies on the assumption that the trained model can adequately describe the relationships between O_3, temporal predictors, and meteorological variables. The use of temporal variables helps represent long-term and seasonal patterns, but these variables are proxies and cannot explicitly distinguish the effects of anthropogenic emissions, natural emissions, background O_3, and regional transport. In addition, SHAP analysis provides a model-based attribution of predicted O_3 variability to individual predictors. These SHAP values should be interpreted as statistical associations within the trained model rather than as direct physical or causal contributions. Correlations among meteorological variables, such as between RH, TCC, SSRD, and TP, may also affect the allocation of SHAP values among predictors.

Finally, this study focuses on observational and statistical attribution. While this framework is useful for diagnosing broad spatial and seasonal patterns, it cannot fully resolve detailed chemical mechanisms or source contributions. Future studies combining long-term precursor observations, updated emission inventories, chemical transport models, and causal inference methods would help further constrain the roles of precursor emissions, chemical sensitivity, regional transport, and meteorological variability in driving O_3 changes across different seasons and regions.

4 Conclusions

In this study, we developed a LightGBM-based meteorological normalization framework combined with SHAP analysis to examine the interannual variability of MDA8 O_3 across China during the warm and cold seasons from 2015 to 2024. The results show that surface O_3 increased widely across China in both seasons. The observed national mean MDA8 O_3 increased by 2.1 $\mu g\ m^{-3}\ yr^{-1}$ during the warm season and by 1.7 $\mu g\ m^{-3}\ yr^{-1}$ during the cold season, indicating that recent O_3 enhancement in China is not confined to photochemically active months.

Meteorological normalization further shows that the long-term increase in MDA8 O_3 was mainly associated with the non-meteorological component. After removing meteorological variability, the national mean MDA8 O_3 increased at the same rate of 1.2 $\mu g\ m^{-3}\ yr^{-1}$ in both the warm and cold seasons. This result suggests that non-meteorological changes played a dominant role in the decadal O_3 increase. This component likely reflects the combined influence of precursor-emission changes, background O_3, regional transport, natural emissions, and other slowly varying processes not explicitly represented in the model. Therefore, it should not be interpreted as the isolated effect of anthropogenic emission changes alone.

The normalized O_3 trends showed pronounced seasonal and regional differences. During the warm season, the strongest normalized increases occurred in the FWP, SCB, and NCP, indicating larger non-meteorological increases in northern inland and basin regions. During the cold season, the FWP, YRD, and PRD showed stronger normalized increases than their warm-season counterparts, while the NCP and SCB showed weaker cold-season increases. In particular, the strong cold-season increase in the PRD suggests that cold-season O_3 enhancement in southern coastal China deserves greater attention in future O_3 control strategies.

SHAP analysis revealed distinct seasonal and regional patterns in the meteorological variables associated with MDA8 O_3 variability. During the warm season, T2M and SSRD were the dominant model-attributed meteorological variables in northern inland and basin regions, including the NCP, FWP, and SCB, consistent with the importance of temperature and radiation for photochemical O_3 production. In contrast, RH and wind-field variables were more important in the YRD and PRD, suggesting stronger influences of moisture conditions, monsoonal transport, and ventilation in southern coastal regions. During the cold season, SSRD became the dominant meteorological variable across most regions, whereas RH played a more important role in the PRD. These findings indicate that meteorological modulation of O_3 is highly season- and region-dependent.

The net meteorological contribution also exhibited a clear temporal transition. During 2015–2019, meteorological conditions generally suppressed O_3 concentrations in many regions. After 2019, the meteorological contribution gradually shifted toward positive values. During 2022–2024, meteorological conditions became more favorable for O_3 enhancement in several key regions, particularly the SCB and FWP, where warm-season net meteorological contributions exceeded 11 $\mu g\ m^{-3}$ in 2024. These results suggest that unfavorable meteorological anomalies in recent years amplified O_3 pollution and may have partly offset the apparent benefits of precursor-emission controls.

Overall, this study demonstrates the importance of separating meteorological variability from long-term non-meteorological changes when interpreting recent O_3 trends in China. The results highlight that O_3 control strategies should account for both evolving emission structures and seasonally varying meteorological modulation. In particular, stronger attention should be paid to the cold-season O_3 increase and to region-specific meteorological sensitivities when designing future O_3 mitigation policies.

Data availability

Hourly surface ozone was obtained from the China National Environmental Centre (http://www.cnemc.cn/en, last access: 19 April 2026). The ERA5 data were acquired from European Centre for Medium-Range Weather Forecasts Reanalysis v5 dataset (https://cds.climate.copernicus.eu/, last access: 19 April 2026).

Author contributions

ZY and RL designed the research; ZY analyzed the data; ZY and RL wrote the original draft; ZY, RL, XC, MW, and TH reviewed and edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

Copernicus Publications adds a standard disclaimer: “Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While

Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher."

Acknowledgments

We acknowledge the China National Environmental Centre and the European Centre for Medium-Range Weather Forecasts for providing datasets that made this work possible. We also acknowledge the support of the College of Environment and Climate, Institute for Environmental and Climate Research, Jinan University.

Financial support

This research was supported by the National Key Research and Development Program of China (grant number 2023YFC3706204).

References

Cao, T., Wang, H., Chen, X., Li, L., Lu, X., Lu, K., and Fan, S.: Rapid increase in spring ozone in the Pearl River Delta, China during 2013–2022, npj Clim. Atmos. Sci., 7, 309, https://doi.org/10.1038/s41612-024-00847-3, 2024.

Cao, T., Wang, H., Li, L., Lu, X., Liu, Y., and Fan, S.: Fast spreading of surface ozone in both temporal and spatial scale in Pearl River Delta, J. Environ. Sci., 137, 540–552, https://doi.org/10.1016/j.jes.2023.02.025, 2024.

Chen, C., Li, T., Sun, Q., Shi, W., He, M. Z., Wang, J., Liu, J., Zhang, M., Jiang, Q., Wang, M., and Shi, X.: Short-term exposure to ozone and cause-specific mortality risks and thresholds in China: evidence from nationally representative data, 2013–2018, Environ. Int., 171, 107666, https://doi.org/10.1016/j.envint.2022.107666, 2023.

Chen, L., Lin, J., Xu, J., Kong, H., Weng, H., Zhang, Y., Yun, X., and Tao, S.: Chemically enhanced transboundary ozone pollution suppresses city-level emission control benefits, Environ. Res. Lett., 20, 054009, https://doi.org/10.1088/1748-9326/adc1e1, 2025.

Chen, Y., Dai, Z., Wang, X., Zhang, Q., Chen, H., Zhong, S., Lu, Q., and Yu, L.: Meteorological and socioeconomic impacts on ozone in China: past and future analysis, iScience, 29, 114335, https://doi.org/10.1016/j.isci.2025.114335, 2026.

Dai, Q., Dai, T., Hou, L., Li, L., Bi, X., Zhang, Y., and Feng, Y.: Quantifying the impacts of emissions and meteorology on the interannual variations of air pollutants in major Chinese cities from 2015 to 2021, Sci. China Earth Sci., 66, 1725–1737, https://doi.org/10.1007/s11430-022-1128-1, 2023.

Dang, R. and Liao, H.: Severe winter haze days in the Beijing-Tianjin-Hebei region from 1985 to 2017 and the roles of anthropogenic emissions and meteorology, Atmos. Chem. Phys., 19, 10801–10816, https://doi.org/10.5194/acp-19-10801-2019, 2019.

Ding, J., Dai, Q., Fan, W., Lu, M., Zhang, Y., Han, S., and Feng, Y.: Impacts of meteorology and precursor emission change on O_3 variation in Tianjin, China from 2015 to 2021, J. Environ. Sci., 126, 506–516, https://doi.org/10.1016/j.jes.2022.03.010, 2023.

Fang, J., Zhang, Y., Hauglustaine, D., Zheng, B., Wang, M., Li, J., Sun, Y., Li, H., Wang, J., Wu, Y., Yuan, B., Chen, M., and Ge, X.: Tracking surface ozone responses to clean air actions under a warming climate in China using machine learning, Atmos. Chem. Phys., 26, 851–867, https://doi.org/10.5194/acp-26-851-2026, 2026.

Feng, Z., Hu, E., Wang, X., Jiang, L., and Liu, X.: Ground-level O_3 pollution and its impacts on food crops in China: a review, Environ. Pollut., 199, 42–48, https://doi.org/10.1016/j.envpol.2015.01.016, 2015.

Gao, M., Wang, F., Ding, Y., Wu, Z., Xu, Y., Lu, X., Wang, Z., Carmichael, G. R., and McElroy, M. B.: Large-scale climate patterns offer preseasonal hints on the co-occurrence of heat wave and O_3 pollution in China, P. Natl. Acad. Sci. USA, 120, e2218274120, https://doi.org/10.1073/pnas.2218274120, 2023.

Gong, C., Yue, X., Liao, H., and Ma, Y.: A humidity-based exposure index representing ozone damage effects on vegetation, Environ. Res. Lett., 16, 044030, https://doi.org/10.1088/1748-9326/abecbb, 2021.

Goto, D., Yashiro, H., Yamashita, Y., Takigawa, M., Sudo, K., Uchida, J., Yumimoto, K., and Takemura, T.: Assessment of the impacts of anthropogenic ozone precursors on domestic and transboundary ozone concentrations using a high-resolution global chemistry model, Prog. Earth Planet. Sci., 13, 3, https://doi.org/10.1186/s40645-025-00791-7, 2026.

Grange, S. K. and Carslaw, D. C.: Using meteorological normalisation to detect interventions in air quality time series, Sci. Total Environ., 653, 578–588, https://doi.org/10.1016/j.scitotenv.2018.10.344, 2019.

Grange, S. K., Carslaw, D. C., Lewis, A. C., Boleti, E., and Hueglin, C.: Random forest meteorological normalisation models for Swiss PM_{10} trend analysis, Atmos. Chem. Phys., 18, 6223–6239, https://doi.org/10.5194/acp-18-6223-2018, 2018.

Guo, Y., Li, K., Zhao, B., Shen, J., Bloss, W. J., Azzi, M., and Zhang, Y.: Evaluating the real changes of air quality due to clean air actions using a machine learning technique: results from 12 Chinese mega-cities during 2013–2020, Chemosphere, 300, 134608, https://doi.org/10.1016/j.chemosphere.2022.134608, 2022.

He, C., Wu, Q., Li, B., Liu, J., Gong, X., and Lu, Z.: Surface ozone pollution in China: trends, exposure risks, and drivers, Front. Public Health, 11, 1131753, https://doi.org/10.3389/fpubh.2023.1131753, 2023.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y.: LightGBM: a highly efficient gradient boosting decision tree, Adv. Neural Inf. Process. Syst., 30, 3149–3157, https://dl.acm.org/doi/10.5555/3294996.3295074, 2017.

Le, T., Wang, Y., Liu, L., Yang, J., Yung, Y. L., Li, G., and Seinfeld, J. H.: Unexpected air pollution with marked emission reductions during the COVID-19 outbreak in China, Science, 369, 702–706, https://doi.org/10.1126/science.abb7431, 2020.

Li, J., Yang, W., Wang, Z., Chen, H., Hu, B., Li, J., Sun, Y., Fu, P., and Zhang, Y.: Modeling study of surface ozone source-receptor relationships in East Asia, Atmos. Res., 167, 77–88, https://doi.org/10.1016/j.atmosres.2015.07.010, 2016.

Li, J., Yu, S., Chen, X., Zhang, Y., Li, M., Li, Z., Song, Z., Liu, W., Li, P., Xie, M., and Xing, J.: Evaluation of the WRF-CMAQ model performances on air quality in China with the impacts of the observation nudging on meteorology, Aerosol Air Qual. Res., 22, 220023, https://doi.org/10.1016/j.4209/aaqr.220023, 2022.

Li, K., Jacob, D. J., Liao, H., Qiu, Y., Shen, L., Zhai, S., Bates, K. H., Sulprizio, M. P., Song, S., Lu, X., Zhang, Q., Zheng, B., Zhang, Y., Zhang, J., Lee, H. C., and Kuk, S. K.: Ozone pollution in the North China Plain spreading into the late-winter haze season, P. Natl. Acad. Sci. USA, 118, e2015797118, https://doi.org/10.1073/pnas.2015797118, 2021.

Li, K., Jacob, D. J., Shen, L., Lu, X., De Smedt, I., and Liao, H.: Increases in surface ozone pollution in China from 2013 to 2019: anthropogenic and meteorological influences, Atmos. Chem. Phys., 20, 11423–11433, https://doi.org/10.5194/acp-20-11423-2020, 2020.

Li, Y., Li, C., Li, Y., Wang, T., Li, M., Qu, Y., Wu, H., Xie, M., and Wang, Y.: Decadal evolution of aerosol-mediated ozone responses in eastern China under clean air actions and carbon neutrality policies, Atmos. Chem. Phys., 26, 1301–1319, https://doi.org/10.5194/acp-26-1301-2026, 2026.

Liu, J., Wang, L., Li, M., Liao, Z., Sun, Y., Song, T., Gao, W., Wang, Y., Li, Y., Ji, D., Hu, B., Kerminen, V.-M., Wang, Y., and Kulmala, M.: Quantifying the impact of synoptic circulation patterns on ozone variability in northern China from April to October 2013–2017, Atmos. Chem. Phys., 19, 14477–14492, https://doi.org/10.5194/acp-19-14477-2019, 2019.

Liu, Q. and Hang, T.: Seasonal synergistic management of urban heat island effect and $PM_{2.5}$ pollution: insights from interpretable LightGBM-SHAP machine learning model, Environ. Impact Assess. Rev., 116, 108129, https://doi.org/10.1016/j.eiar.2025.108129, 2026.

Liu, X., Gao, H., Zhang, X., Zhang, Y., Yan, J., Niu, J., and Chen, F.: Driving forces of meteorology and emission changes on surface ozone in the Huaihe River Basin, China, Water Air Soil Pollut., 234, 355, https://doi.org/10.1007/s11270-023-06345-1, 2023.

Liu, Y. and Wang, T.: Worsening urban ozone pollution in China from 2013 to 2017 – Part 1: the complex and varying roles of meteorology, Atmos. Chem. Phys., 20, 6305–6321, https://doi.org/10.5194/acp-20-6305-2020, 2020.

Liu, Y., Geng, G., Cheng, J., Liu, Y., Xiao, Q., Liu, L., Shi, Q., Tong, D., He, K., and Zhang, Q.: Drivers of increasing ozone during the two phases of Clean Air Actions in China 2013–2020, Environ. Sci. Technol., 57, 8954–8964, https://doi.org/10.1021/acs.est.3c00054, 2023.

Lu, X., Zhang, L., Chen, Y., Zhou, M., Zheng, B., Li, K., Liu, Y., Lin, J., Fu, T.-M., and Zhang, Q.: Exploring 2016–2017 surface ozone pollution over China: source contributions and meteorological influences, Atmos. Chem. Phys., 19, 8339–8361, https://doi.org/10.5194/acp-19-8339-2019, 2019.

Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, arXiv [preprint], arXiv: 1705.07874, 25 November, 2017.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell., 2, 56–67, https://doi.org/10.1038/s42256-019-0138-9, 2020.

Lyu, Y., Xu, H., Wu, H., Han, F., Lv, F., Kang, A., and Pang, X.: Spatiotemporal variations of $PM_{2.5}$ and ozone in urban agglomerations of China and meteorological drivers for ozone using explainable machine learning, Environ. Pollut., 365, 125380, https://doi.org/10.1016/j.envpol.2024.125380, 2025.

Mai, Z., Shen, H., Zhang, A., Sun, H. Z., Zheng, L., Guo, J., Liu, C., Chen, Y., Wang, C., Ye, J., Zhu, L., Fu, T.-M., Yang, X., and Tao, S.: Convolutional neural networks facilitate process understanding of megacity ozone temporal variability, Environ. Sci. Technol., 58, 15691–15701, https://doi.org/10.1021/acs.est.3c07907, 2024.

Mao, J., Tai, A. P. K., Yung, D. H. Y., Yuan, T., Chau, K. T., and Feng, Z.: Multidecadal ozone trends in China and implications for human health and crop yields: a hybrid approach combining a chemical transport model and machine learning, Atmos. Chem. Phys., 24, 345–366, https://doi.org/10.5194/acp-24-345-2024, 2024.

Meng, X., Jiang, J., Chen, T., Zhang, Z., Lu, B., Liu, C., Xue, L., Chen, J., Herrmann, H., and Li, X.: Chemical drivers of ozone change in extreme temperatures in eastern China, Sci. Total Environ., 874, 162424, https://doi.org/10.1016/j.scitotenv.2023.162424, 2023.

Ni, Y., Yang, Y., Wang, H., Wang, P., Li, K., Chen, L., Zhu, J., Li, B., and Liao, H.: Machine learning-based bias-corrected future projections of ozone concentrations from a chemistry-climate model, Environ. Sci. Technol., 60, 3135–3147, https://doi.org/10.1021/acs.est.5c11992, 2026.

Ren, H. H., Cheng, Y., Wu, F., Gu, Z. L., Cao, J. J., Huang, Y., Xue, Y. G., Cui, L., Zhang, Y. W., Chow, J. C., Watson, J. G., Zhang, R. J., Lee, S. C., Wang, Y. L., and Liu, S.: Spatiotemporal characteristics of ozone and the formation sensitivity over the Fenwei Plain, Sci. Total Environ., 881, 163369, https://doi.org/10.1016/j.scitotenv.2023.163369, 2023.

Ren, J., Guo, F., and Xie, S.: Diagnosing ozone-NOx-VOC sensitivity and revealing causes of ozone increases in China based on 2013–2021 satellite retrievals, Atmos. Chem. Phys., 22, 15035–15047, https://doi.org/10.5194/acp-22-15035-2022, 2022.

Saeipourdizaj, P., Sarbakhsh, P., and Gholampour, A.: Application of imputation methods for missing values of PM_{10} and O_3 data: interpolation, moving average and K-nearest neighbor methods, Environ. Health Eng. Manag., 8, 215–226, https://doi.org/10.34172/EHEM.2021.25, 2021.

Tan, Y., Zhang, Y., Wang, T., Chen, T., Mu, J., and Xue, L.: Dissecting drivers of ozone pollution during the 2022 multicity lockdowns in China sheds light on future control direction, Environ. Sci. Technol., 58, 6988–6997, https://doi.org/10.1021/acs.est.4c01197, 2024.

Tong, L., Liu, Y., Meng, Y., Dai, X., Huang, L., Luo, W., Yang, M., Pan, Y., Zheng, J., and Xiao, H.: Surface ozone changes during the COVID-19 outbreak in China: an insight into the pollution characteristics and formation regimes of ozone in the cold season, J. Atmos. Chem., 80, 103–120, https://doi.org/10.1007/s10874-022-09443-2, 2023.

Vu, T. V., Shi, Z., Cheng, J., Zhang, Q., He, K., Wang, S., and Harrison, R. M.: Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique, Atmos. Chem. Phys., 19, 11303–11314, https://doi.org/10.5194/acp-19-11303-2019, 2019.

Wang, H., Liu, X., Wu, C., and Lin, G.: Regional to global distributions, trends, and drivers of biogenic volatile organic compound emission from 2001 to 2020, Atmos. Chem. Phys., 24, 3309–3328, https://doi.org/10.5194/acp-24-3309-2024, 2024.

Wang, L., Chen, B., Ouyang, J., Mu, Y., Zhen, L., Yang, L., Xu, W., and Tang, L.: Causal-inference machine learning reveals the drivers of China's 2022 ozone rebound, Environ. Sci. Ecotechnology, 24, 100524, https://doi.org/10.1016/j.ese.2025.100524, 2025.

Wang, T., Xue, L., Brimblecombe, P., Lam, Y. F., Li, L., and Zhang, L.: Ozone pollution in China: a review of concentrations, meteorological influences, chemical precursors, and effects, Sci. Total Environ., 575, 1582–1596, https://doi.org/10.1016/j.scitotenv.2016.10.081, 2017.

Wang, T., Xue, L., Feng, Z., Dai, J., Zhang, Y., and Tan, Y.: Ground-level ozone pollution in China: a synthesis of recent findings on influencing factors and impacts, Environ. Res. Lett., 17, 063003, https://doi.org/10.1088/1748-9326/ac69fe, 2022.

Wang, W., Parrish, D. D., Wang, S., Bao, F., Ni, R., Li, X., Yang, S., Wang, H., Cheng, Y., and Su, H.: Long-term trend of ozone pollution in China during 2014–2020: distinct seasonal and spatial characteristics and ozone sensitivity, Atmos. Chem. Phys., 22, 8935–8949, https://doi.org/10.5194/acp-22-8935-2022, 2022.

Wang, X., Zhu, J., Jiao, G., Chen, X., Yang, Z., Chen, L., Jin, X., and Liao, H.: Meteorological influence on surface ozone trends in China: assessing uncertainties caused by multi-dataset and multi-method, Atmos. Chem. Phys., 25, 13863–13878, https://doi.org/10.5194/acp-25-13863-2025, 2025.

Wang, Y., Gong, C., Dong, L., and Huang, Y.: Estimation of near-surface high spatiotemporal resolution ozone concentration in China using Himawari-8 AOD, Remote Sens., 17, 528, https://doi.org/10.3390/rs17030528, 2025.

Wang, Y., Wang, Y., Li, Q., Chen, T., Tan, Y., and Wang, T.: Synergy of natural sources exacerbates ozone pollution in China during drought–heatwave extremes, Environ. Sci. Technol., 59, 22739–22748, https://doi.org/10.1021/acs.est.5c07239, 2025.

Wang, Z., Shi, C., Zhang, H., Ji, X., Zhu, Y., Xia, C., Sun, X., Lin, X., Yan, S., Wang, S., Zhou, Y., Xing, C., Chen, Y., and Liu, C.: Opposing trends in the peak and low ozone concentrations in eastern China: anthropogenic and meteorological influences, Atmos. Chem. Phys., 25, 347–366, https://doi.org/10.5194/acp-25-347-2025, 2025.

Weng, X., Forster, G. L., and Nowack, P.: A machine learning approach to quantify meteorological drivers of ozone pollution in China from 2015 to 2019, Atmos. Chem. Phys., 22, 8385–8402, https://doi.org/10.5194/acp-22-8385-2022, 2022.

Wu, K., Wang, Y., Qiao, Y., Liu, Y., Wang, S., Yang, X., Wang, H., Lu, Y., Zhang, X., and Lei, Y.: Drivers of 2013–2020 ozone trends in the Sichuan Basin, China: impacts of meteorology and precursor emission changes, Environ. Pollut., 300, 118914, https://doi.org/10.1016/j.envpol.2022.118914, 2022.

Xiang, Y., Zhang, T., Liu, J., Lv, L., Dong, Y., and Chen, Z.: Atmosphere boundary layer height and its effect on air pollutants in Beijing during winter heavy pollution, Atmos. Res., 215, 305–316, https://doi.org/10.1016/j.atmosres.2018.09.014, 2019.

Xue, J., Wang, F., Zhang, K., Zhai, H., Jin, D., Duan, Y., Yaluk, E., Wang, Y., Huang, L., Li, Y., Lei, T., Fu, Q., Fu, J. S., and Li, L.: Elucidate long-term changes of ozone in Shanghai based on an integrated machine learning method, Front. Environ. Sci. Eng., 17, 138, https://doi.org/10.1007/s11783-023-1738-5, 2023.

Yang, J. and Zhao, Y.: Performance and application of air quality models on ozone simulation in China – a review, Atmos. Environ., 293, 119446, https://doi.org/10.1016/j.atmosenv.2022.119446, 2023.

Yang, J., Wang, Y., Zhang, L., and Zhao, Y.: Investigating the response of China's surface ozone concentration to the future changes of multiple factors, Atmos. Chem. Phys., 25, 2649–2666, https://doi.org/10.5194/acp-25-2649-2025, 2025.

Yang, L., Luo, H., Yuan, Z., Zheng, J., Huang, Z., Li, C., Lin, X., Louie, P. K. K., Chen, D., and Bian, Y.: Quantitative impacts of meteorology and precursor emission changes on the long-term trend of ambient ozone over the Pearl River Delta, China, and implications for ozone control strategy, Atmos. Chem. Phys., 19, 12901–12916, https://doi.org/10.5194/acp-19-12901-2019, 2019.

Yang, Y., Zhou, Y., Wang, H., Li, M., Li, H., Wang, P., Yue, X., Li, K., Zhu, J., and Liao, H.: Meteorological characteristics of extreme ozone pollution events in China and their future predictions, Atmos. Chem. Phys., 24, 1177–1191, https://doi.org/10.5194/acp-24-1177-2024, 2024.

Ye, X., Zhang, L., Wang, X., Lu, N., Hickman, S., Luo, G., and Archibald, A. T.: Deciphering the impacts of meteorology on surface ozone variability in eastern China using explainable machine learning models, Atmos. Chem. Phys., 26, 6377–6390, https://doi.org/10.5194/acp-26-6377-2026, 2026.

Zhang, L., Wang, L., Ji, D., Xia, Z., Nan, P., Zhang, J., Li, K., Qi, B., Du, R., Sun, Y., Wang, Y., and Hu, B.: Explainable ensemble machine learning revealing the effect of meteorology and sources on ozone formation in megacity Hangzhou, China, Sci. Total Environ., 922, 171295, https://doi.org/10.1016/j.scitotenv.2024.171295, 2024.

Zhang, X., Zhang, W.-C., Wu, W., and Liu, H.-B.: Understanding ozone variability in spatial responses to emissions and meteorology in China using interpretable machine learning, iScience, 28, 113036, https://doi.org/10.1016/j.isci.2025.113036, 2025.

Zhou, W., Lei, L., Du, A., Zhang, Z., Li, Y., Yang, Y., Tang, G., Chen, C., Xu, W., Sun, J., Li, Z., Fu, P., Wang, Z., and Sun, Y.: Unexpected increases of severe haze pollution during the post COVID-19 period: effects of emissions, meteorology, and secondary production, J. Geophys. Res.-Atmos., 127, e2021JD035710, https://doi.org/10.1029/2021JD035710, 2022.