Deciphering the impacts of meteorology on surface ozone variability in eastern China using explainable machine learning models

Xingpei Ye[1,2], Lin Zhang[1,3,4*], Xiaolin Wang[1], Ni Lu[1], Sebastian Hickman[2], Guo Luo[1], and Alex T. Archibald[2,5]

[1]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China

[2]Centre for Atmospheric Science, Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, United Kingdom

[3]Institute of Carbon Neutrality, Peking University, Beijing, China

[4]Center for Environment and Health, Peking University, Beijing, China

[5]National Centre for Atmospheric Science, University of Cambridge, Cambridge, United Kingdom

*Corresponding to: Lin Zhang (zhanglg@pku.edu.cn)

Abstract. Understanding how meteorology influences surface ozone variability is critical for interpreting trends and designing effective air quality policies. This study employs explainable machine learning (XML) with SHapley Additive exPlanations (SHAP) to interpret daily ozone variations from 2013 to 2023 across three major regions in eastern China: North China Plain (NCP), Yangtze River Delta (YRD), and Pearl River Delta (PRD). An ensemble of five machine learning models (LightGBM, XGBoost, CatBoost, Random Forest, and Extra Trees) is trained using 14 meteorological variables and two temporal indicators. XML reveals nonlinear, region-specific ozone-meteorology relationships that are broadly consistent with physical understanding, while differences in SHAP attributions across algorithms highlight structural uncertainty arising from multicollinearity among input variables. We use SHAP-derived contributions to attribute warm-season ozone trends to meteorological versus non-meteorological drivers. Before 2019, ozone increases are mainly associated with the temporal proxy for non-meteorological influences (e.g., emission changes), whereas after 2019 meteorological variability dominates regional ozone trends. Exploiting the additive nature of SHAP, we develop a de-weathering framework that partitions daily ozone into a SHAP-based climatological baseline and a meteorology-induced ozone anomaly (MOA). Across all three regions, the magnitude of positive MOA events increases over 2013–2023, while their frequency and duration show no significant trends, indicating that meteorological conditions increasingly amplify the intensity of ozone pollution episodes, without a corresponding increase in their frequency or duration. Our results demonstrate both the utility and limitations of XML for disentangling meteorological drivers of ozone pollution and provide new constraints on how meteorology shapes surface ozone under China's clean air actions.

Keywords: surface ozone, air quality, explainable machine learning, SHAP, de-weathering

1. Introduction

Surface ozone is a major air pollutant that threatens human health and ecosystems (Cohen et al., 2017; Fowler et al., 2009; Mills et al., 2018; Zhang et al., 2019). It is formed mainly through photochemical reactions involving precursors such as carbon monoxide (CO), nitrogen oxides ($NO_x = NO + NO_2$), and non-methane volatile organic compounds (NMVOCs) under sunlight. In China, nationwide monitoring has revealed a widespread and rapid increase in surface ozone levels since 2013, drawing growing attention from researchers and policymakers (Lu et al., 2020; Wang et al., 2017). The Chinese government has implemented a series of clean air actions since 2013 to reduce anthropogenic emissions of air pollutants (State Council of the People's Republic of China, 2013; 2018). Early emission control efforts (2013−2017) primarily targeted $PM_{2.5}$ and NO_x reductions, which may have contributed to ozone increases under the VOC-limited photochemical conditions prevalent in eastern Chinese cities. In the subsequent phase (2018–2020), integrated VOC and NO_x controls were implemented to address ozone pollution (Liu et al., 2023; Wang et al., 2023).

Ozone formation is highly sensitive not only to chemical precursors but also to meteorological conditions. Temperature and solar radiation control photochemical production rates, while wind speed, boundary-layer height, and large-scale circulation determine the dispersion, accumulation, and transport of ozone and its precursors (Fiore et al., 2012; Lu et al., 2019). Relative humidity, cloud cover, and precipitation further modulate ozone lifetimes and vertical mixing (Fu and Tian, 2019; Lu et al., 2019). Recent studies have shown that meteorological variability can explain a substantial fraction of interannual ozone changes in China and can even mask or amplify the effects of emission controls in some years (Liu and Wang, 2020; Liu et al., 2023; Weng et al., 2022). Thus, explicitly isolating and quantifying the meteorological contribution to surface ozone variability is critical for correctly interpreting observed ozone trends and evaluating the effectiveness of clean air policies.

Conventional approaches for quantifying meteorological impacts on surface ozone include statistical models, such as multiple linear regression, generalized additive models (Bloomer et al., 2009; Gong et al., 2017), and numerical atmospheric models that perturb selected variables to attribute ozone changes to specific processes (e.g., Liu et al., 2023; Liu and Wang, 2020). These approaches suggest that worsening meteorological conditions—particularly higher temperatures and lower humidity—have contributed significantly to the rising surface ozone levels in eastern China since 2012, with impacts comparable to those of anthropogenic emissions (Liu et al., 2023; Liu and Wang, 2020). Temperature plays a particularly important role in northern China by directly enhancing ozone formation and increasing anthropogenic and biogenic precursor emissions (Dang et al., 2021; Wu et al., 2024), whereas humidity is found to be especially influential in central and southern China (Han et al., 2020; Li et al., 2020). However, both statistical and numerical models face key limitations in quantifying meteorological contributions. Linear statistical models can approximate the dominant relationships and meteorology-driven trends in ozone at seasonal or annual scales (e.g., Weng et al., 2022), yet they have limited flexibility to represent complex nonlinear interactions. Process-based models, on the other hand, are computationally intensive and often exhibit biases relative to observed ozone concentrations (Travis and Jacob, 2019; Ye et al., 2022). Moreover, they rely heavily on emission inventories, which are typically updated on multi-year timescales, limiting their ability to reflect recent changes in emissions and to provide a comprehensive understanding of their impacts on ozone chemistry.

Machine learning (ML) and emerging explainable ML (XML) techniques provide a powerful framework for capturing the complex, nonlinear relationships between observed meteorological variables and surface ozone concentrations, thereby addressing some of the limitations of traditional statistical and process-based methods. XML information can be broadly categorized into global and local explainability (Flora et al., 2024). Global explainability characterizes how input features influence model predictions across the dataset, often by ranking input features in terms of importance, for example using feature-importance outputs from Random Forest, which have been widely applied to identify key meteorological drivers of surface ozone in China (Ma et al., 2021; Weng et al., 2022; Zhan et al., 2018). In contrast, local explainability attributes individual predictions to specific input features, offering more detailed insights into input–output dependencies at the data-point level. Recently, the

SHapley Additive exPlanation (SHAP) technique has been applied to ML models to separate meteorological, chemical, and source-related effects on ozone in individual regions, such as Hangzhou (Yao et al., 2024; Zhang et al., 2024), Shenzhen (Mai et al., 2024), and Taipei (Li et al., 2025). These works demonstrate the potential of XML for ozone attribution at the city or regional scale.

Despite recent progress, the application of XML techniques to systematically understand meteorological influences on ozone remains limited. Key challenges include assessing XML's ability to capture ozone–meteorology relationships across spatiotemporal scales and evaluating the consistency of explanations across different ML algorithms. In this study, we address these challenges by combining an ensemble of five SHAP-interpreted ML models with a new de-weathering framework to analyze daily surface ozone across three regions in eastern China from 2013 to 2023. Specifically, we (1) quantify nonlinear ozone–meteorology relationships and evaluate the robustness of XML explanations across five tree-based models; (2) use SHAP values to attribute long-term warm-season ozone trends to meteorological versus non-meteorological factors; and (3) develop a SHAP-based de-weathering approach that partitions daily ozone into a climatological baseline and meteorology-induced ozone anomalies (MOA), allowing trends in the intensity and occurrence of meteorology-driven ozone events to be diagnosed. These advances enable a more comprehensive assessment of how meteorology shapes ozone variability under China's clean air actions.

2. Data and Methods

2.1 Surface ozone measurements and meteorology reanalysis

We use nearly 11 years of surface ozone measurements (from 1 January 2013 to 30 September 2023, 3925 days in total) collected by the national air-quality monitoring network operated by the Chinese Ministry of Ecology and Environment (https://air.cnemc.cn:18007, last access: 4 October 2024). The ozone exposure metric is the daily maximum 8-hour average (MDA8), calculated for each site on each day. Data quality control procedures are applied to remove outliers, following Lu et al. (2018). Specifically, we discard extremely high hourly ozone concentrations, require a sufficient number of valid hourly measurements to compute daily MDA8, and exclude sites with substantial data gaps over the study period (Lu et al., 2018). The measurement sites are then averaged by each city to account for air quality condition at the city level. The number of monitoring sites per city varies from 3 to 27. Here, we focus our analyses on three populous regions in eastern China: the North China Plain (NCP), the Yangtze River Delta (YRD), and the Pearl River Delta (PRD). The locations of the cities and the regional mean MDA8 ozone concentrations are shown in Figure 1.

In this study, we use 14 meteorological variables as input features: 2-m temperature (T2M), surface incoming shortwave solar radiation (SWGDN), northward wind at 850 hPa (V850), eastward wind at 850 hPa (U850), 10-m wind speed (W10M), geopotential heights at 500 hPa (H500) and 850 hPa (H850), surface pressure (PS), surface evaporation (EVAP), surface relative humidity (RH), total surface precipitation flux (PRECTOT), planetary boundary-layer height (PBLH), total cloud fraction (CLDTOT), and surface albedo (ALBEDO). These variables are selected based on their strong correlations with daily surface ozone variations, as established in previous modelling and statistical studies (Liu and Wang, 2020; Liu et al., 2020; Mo et al., 2020; Requia et al., 2020).

The meteorological data are obtained from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) reanalysis (Gelaro et al., 2017), which provides hourly global fields at $0.5° \times 0.625°$ resolution. To match each ozone monitoring site, we use its latitude–longitude coordinates to identify the nearest MERRA-2 grid cell and extract the corresponding 24-hour mean meteorological variables from that grid cell for each day. When multiple sites fall within the same MERRA-2 grid cell, they share the same meteorological inputs at the grid scale. Meteorological variables are also averaged at the city level to ensure consistency with the ozone observations used for model training. Using city-level ozone and meteorological conditions helps reduce representativeness errors between point measurements and coarse-resolution reanalysis data. In addition, sensitivity tests show that using only one site per city does not materially affect model performance or the subsequent analyses.

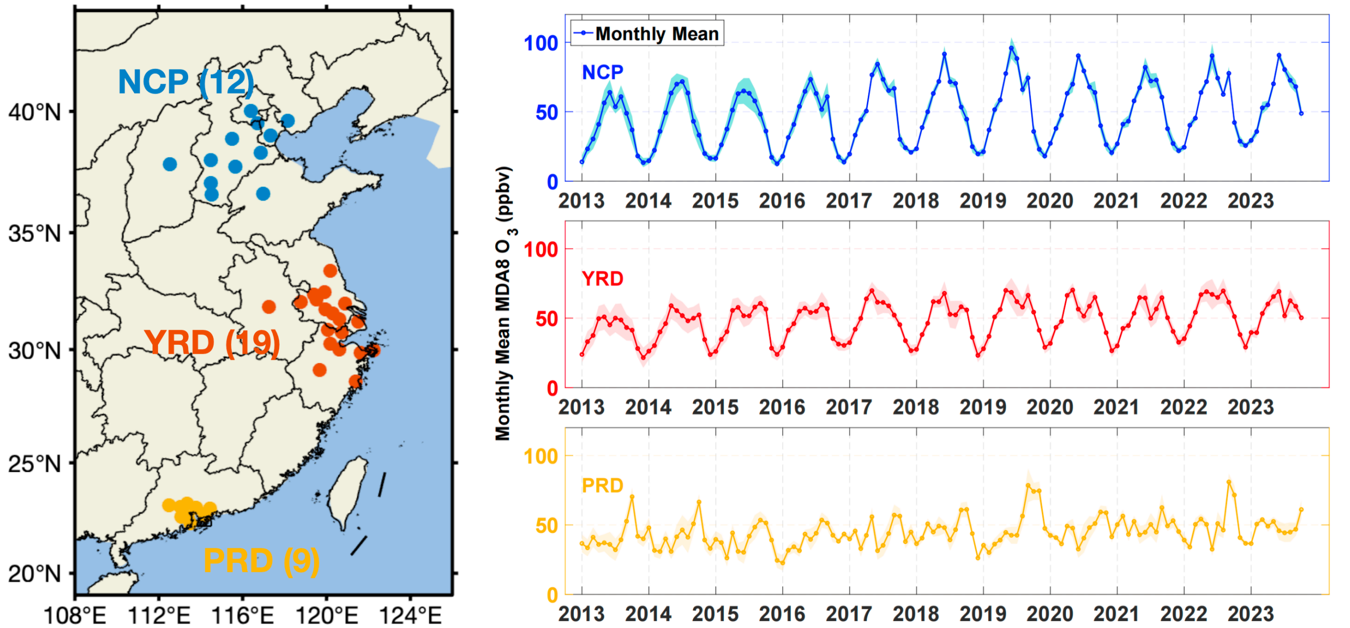

Figure 1. Locations of ozone monitoring cities and regional averaged monthly MDA8 ozone concentrations analyzed in this study. The three regions are North China Plain (NCP, blue), Yangtze River Delta (YRD, Red), and Pearl River Delta (PRD, yellow). Numbers in brackets denote the number of cities. The right panels show monthly mean ozone concentrations in the three regions over January 2013 to September 2023. Shaded areas represent ±1 standard deviation across cities within each region.

2.2 ML models and explainability

We apply five ML algorithms: LightGBM (Ke et al., 2017), XGBoost (Chen and Guestrin, 2016), CatBoost (Prokhorenkova et al., 2018), Random Forest (Breiman, 2001), and Extremely Randomized Trees (Extra Trees; Geurts et al., 2006), to predict daily variations in surface ozone. All five are tree-based ensemble models that utilize decision trees as their basic structure. LightGBM, XGBoost, and CatBoost are gradient-boosting methods, which iteratively train trees to correct errors from previous iterations. In contrast, Random Forest and Extra Trees employ bagging, where multiple trees are trained independently on random subsets of the data and input features, and their predictions are then averaged to produce the final output. We select these tree-based models because their structure is highly compatible with SHAP-based explainability methods, which can efficiently compute feature contributions by exploiting decision-tree ensembles.

The output variable for all models is the daily MDA8 ozone concentration. To construct the ML models, in addition to the 14 meteorological variables introduced in Section 2.1, two temporal features are included: the day of year (DOY), which ranges from 1 to 365 (or 366 in leap years) and captures the seasonal ozone cycle, and a UNIX time variable, defined as a continuous, non-repeating count of days since 1 January 2013. The UNIX term serves as a proxy for slowly varying, non-meteorological influences such as long-term changes in anthropogenic emissions and background concentrations (Vu et al., 2019). While 14 meteorological predictors may appear numerous, the large sample sizes (~35,000–74,000 samples per region) provide sufficient statistical power to support model training without overfitting. In addition, tree-based ensemble methods inherently mitigate multicollinearity through bootstrap aggregation and random feature subsampling (Breiman, 2001, Chen and Guestrin, 2016). We examine the correlations among all input variables (14 meteorological predictors, DOY, and UNIX) using a Pearson correlation matrix (Figure S1). As expected, several variables are moderately to strongly correlated (for example, between surface temperature and solar radiation, or among wind-related variables), reflecting the coupled nature of boundary-layer dynamics, radiation, and circulation. We retain these correlated variables in the models because they represent physically distinct but linked processes that can each modulate ozone formation and transport.

We train separate ML models for each region using all available monitoring sites to ensure sufficient data coverage. The datasets are split randomly into 90% for training and 10% for testing. Hyperparameters are optimized using five-fold cross-validation based on the coefficient of determination (R^2). For explainability analysis, we use the training data rather than the test data, as the goal is to understand how the model relies on input features rather than its generalization performance (Flora et al., 2024; Lakshmanan et al., 2015). Therefore, after model evaluation, we retrain each model on the full dataset before applying XML techniques. Hyperparameters for ML models are listed in Table S1.

We use the SHAP method (Lundberg and Lee, 2017) to quantify the contribution of each input feature to ozone variability. SHAP is grounded in cooperative game theory and provides local explainability by assigning each feature a contribution value for individual predictions. For a given ML model, the predicted MDA8 ozone concentration on the d-th day in year y in a specific region ($O_3ori_{y,d}$) can be expressed as,

$$O_3ori_{y,d} = C + \sum_{i=1}^{2} Tshap_{y,d,i} + \sum_{i=1}^{14} Mshap_{y,d,i} \tag{1}$$

where C is the base SHAP value, $Tshap$ denotes SHAP values from the two temporal terms (DOY and UNIX), and $Mshap$ denotes SHAP values from 14 meteorological factors. SHAP values are estimated at the daily level and are specific to each feature, thus providing local explainability. Averaging the absolute SHAP values across all days for feature i yields its global importance G_i for the entire model (Eq. (2)).

$$G_i = \frac{1}{N}\sum_{y}\sum_{d}\left|shap_{y,d,i}\right|, \qquad N = 3925\ (days) \tag{2}$$

This dual capability, local and global interpretability, makes SHAP particularly suitable for analyzing complex ozone–meteorology relationships.

2.3 Attributable de-weathering approach

Leveraging the additive property of SHAP values, we propose an attributable de-weathering approach to partition daily surface ozone into two components: a climatological baseline and a meteorology-induced ozone anomaly (MOA). This framework, presented in Eqs. (3)–(6), enables attribution of ozone variability to specific meteorological factors on a daily basis.

$$\overline{Mshap}_{y,d,i} = \frac{1}{11}\sum_{p=1}^{11}\frac{1}{15}\sum_{t=d-7}^{t=d+7} Mshap_{p,t,i} \tag{3}$$

$$O_3dw_{y,d} = C + \sum_{i=1}^{2} Tshap_{y,d,i} + \sum_{i=1}^{14} \overline{Mshap}_{y,d,i} \tag{4}$$

$$\sum_{i=1}^{14} Mshap'_{y,d,i} = O_3ori_{y,d} - O_3dw_{y,d} = \sum_{i=1}^{14}(Mshap_{y,d,i} - \overline{Mshap}_{y,d,i}) \tag{5}$$

$$G'_i = \frac{1}{N}\sum_y \sum_d \left|Mshap'_{y,d,i}\right|,\ N = 3925\ (days) \tag{6}$$

We first compute the climatological mean SHAP values for each meteorological variable ($\overline{Mshap}$) by averaging their daily SHAP values across an 11-year period within a 15-day moving window (Eq. (3)). This averaging is performed across years for the same calendar period, aiming to quantify how a meteorological variable affects ozone levels under typical conditions for that time of year. We adopt a 15-day running window to define this climatological component. This window is long enough to smooth over individual pollution episodes and day-to-day fluctuations, while short enough to preserve sub-seasonal variations in the meteorological background. We then construct the de-weathered ozone estimate (O_3dw) by retaining temporal SHAP values (associated with DOY and UNIX) while replacing the meteorological SHAP values ($Mshap$) with their climatological means (Eq. (4)). The resulting O_3dw represents the ozone level under climatologically mean meteorological conditions.

The difference between O_3ori and O_3dw, denoted as the $Mshap'$ in Eq. (5), then defines the MOA. Additionally, we compute the global importance of MOA (G'_i in Eq. (6)). Compared with $Mshap$ and G_i, $Mshap'$ and G'_i quantify how individual meteorological features contribute to anomalies relative to climatology. Compared with previous de-weathering methods that typically estimate only total meteorological contributions or MOA magnitudes (Vu et al., 2019; Grange and Carslaw, 2019), this framework provides feature-resolved attribution of daily anomalies to individual meteorological drivers.

3. Results

3.1 Ozone–meteorology relationships revealed by XML

We evaluate the performance of the five ML models using the test dataset for each region, employing the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) as performance metrics (Figure S2). All models reproduce daily MDA8 ozone variability well, with testing R^2 values of 0.80-0.88 in NCP, 0.76-0.84 in YRD, and 0.78-0.86 in PRD, and with RMSE and MAE in the range of a few ppbv. Three gradient boosting models (LightGBM, XGBoost, and CatBoost) perform slightly better ($R^2 > 0.8$, MAE < 7.1 ppbv, and RMSE < 10 ppbv) than Random Forest and Extra Trees. The NCP yields the highest R^2 (0.86 to 0.88), while the YRD shows the lowest R^2 (0.76 to 0.81). In addition, we assess model outputs on the full training dataset to ensure internal consistency among different ML algorithms (Figure S3). In most cases, the training R^2 exceeds 0.95 (with some exceptions in YRD). As expected for supervised learning, the training R^2 is higher than the testing R^2 because the models are optimized on the training subset. However, the differences between training and testing R^2 are modest and consistent across regions and algorithms.

We compute SHAP values from each ML model across the three regions and average the absolute values for each feature to obtain the global feature importance (G_i from Eq. (2)). Figure 2a–c shows the ranked feature importance across all five models. In all regions, temperature (T2M), shortwave solar radiation (SWGDN), and relative humidity (RH) emerge as the dominant meteorological drivers, followed by circulation and stability-related variables such as V850, H500, and PBLH. DOY and UNIX also contribute substantially, reflecting the recurring seasonal cycle and long-term non-meteorological changes. These identified drivers are broadly consistent with previous ML-based studies of ozone variability in China (Weng et al., 2022; Liu et al., 2020; Li et al., 2023).

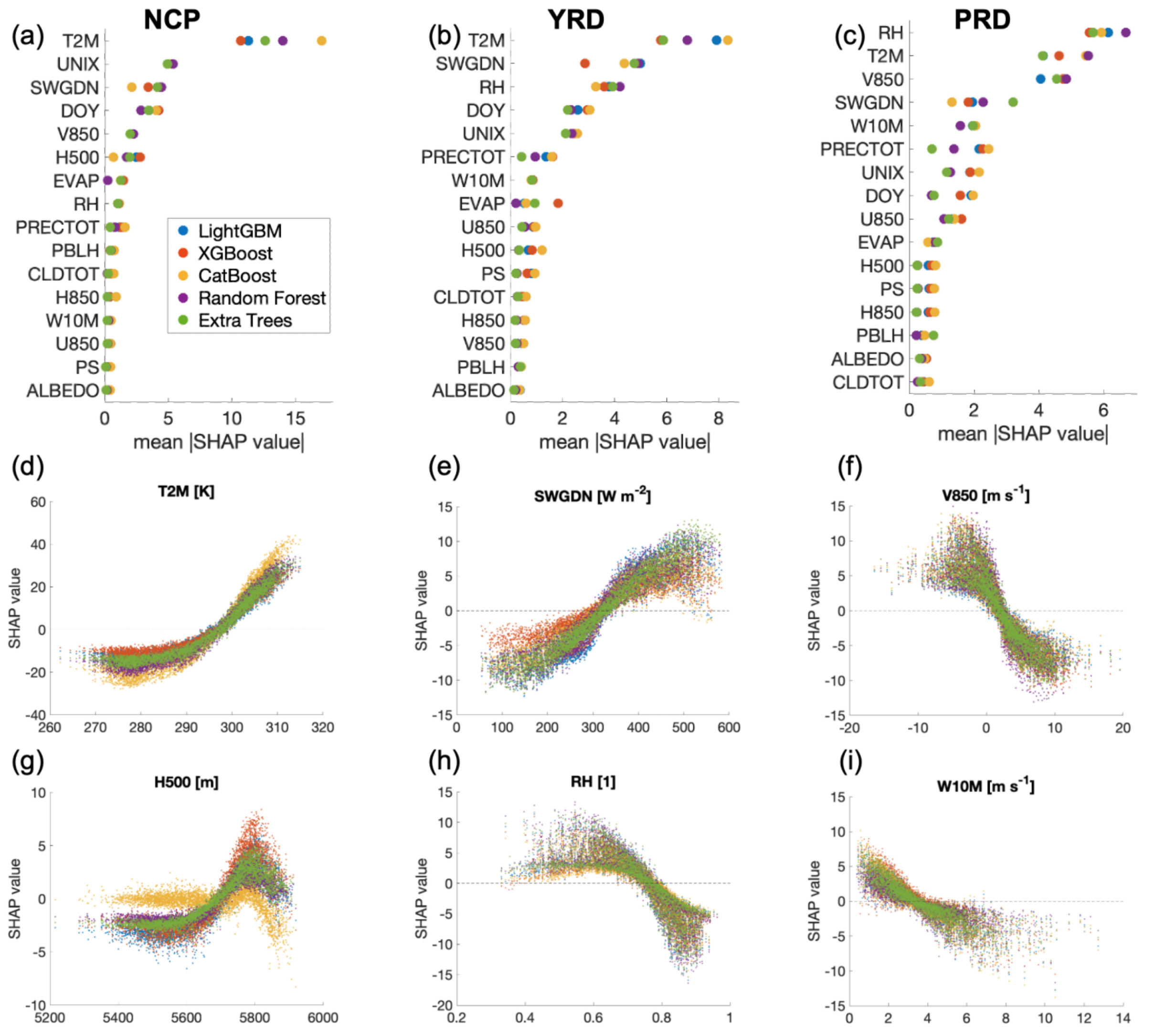

Figure 2. (a)–(c): Global feature importance for the three regions: North China Plain (NCP, left column), Yangtze River Delta (YRD, middle column), and Pearl River Delta (PRD, right column), calculated by averaging the absolute SHAP values across all days. (d–i): Scatter plots of selected meteorological variables versus their SHAP values, illustrating local explainability. Each point represents a daily observation, and different colors correspond to different ML models.

Figure 2d–i shows scatter plots of individual meteorological variables against their SHAP values, providing local explainability insights for each region. Here, positive (negative) SHAP values indicate that the feature value increases (decreases) the predicted ozone concentration relative to the baseline expectation (*C* in Eq. (1)). Two key variables are highlighted per region, with additional examples shown in Figure S4–S6. Overall, the ozone–meteorology relationships revealed by XML are consistent with established physical understanding. For instance, surface ozone levels are positively correlated with T2M and SWGDN, and negatively correlated with RH and W10M (Wang et al., 2017; Lu et al., 2019; Archibald et al., 2020; Fu and Tian, 2019). V850 exhibits opposite effects in NCP and PRD (Figure S4 vs. S6). In NCP, southerly winds (V850 > 0) tend to transport polluted air masses northward, worsening ozone pollution, whereas in PRD, they typically bring cleaner marine air, mitigating ozone levels. The SHAP results also uncover important nonlinear relationships. For example, the O_3–T2M relationship in NCP flattens at lower temperatures, indicating that temperature has a weaker effect on ozone variability in cold seasons than in summer. High 500 hPa geopotential height (H500) is generally associated with elevated ozone due to stable atmospheric conditions; however, beyond 5800 m, its contribution diminishes, likely due to enhanced precipitation and humidity (Figure S7). Furthermore, the SHAP-derived O_3–V850 (Figure 2f) and O_3–RH (Figure 2h) relationships exhibit a two-mode pattern that varies with temperature (Figure S8), suggesting distinct meteorological influences in summer vs. winter.

We observe considerable discrepancies in feature attribution across the five ML models. As shown in Figure 2, although the ranking of key predictors is broadly consistent (e.g., T2M is the leading predictor in NCP), the magnitudes of SHAP values vary substantially across models, leading to uncertainty in their quantitative importance. One illustrative example is the sensitivity of ozone to temperature (dO_3/dT), a key metric for assessing the ozone climate penalty (Bloomer et al., 2009; Jaffe and Zhang, 2017). We estimate dO_3/dT by fitting a linear regression between summertime SHAP values and T2M values (Figure 2d). Across the five models, dO_3/dT estimates range from 1.74 to 2.70 ppbv/K in NCP, 0.81 to 1.27 ppbv/K in YRD, and 0.80 to 1.43 ppbv/K in PRD (Table S2). These inter-model variations are comparable to the spread observed in previous process-based modeling studies (Chen et al., 2024; Porter and Heald, 2019; Zhang et al., 2022), underscoring the importance of model choice when interpreting climate sensitivity estimates.

These discrepancies in explainability are likely driven by multicollinearity among input features. As shown in the correlation heatmap (Figure S1), T2M is highly correlated with SWGDN, H500, surface pressure (PS), and evaporation (EVAP) in all three regions. Different ML algorithms assign varying importance to these correlated features. For example, in NCP, CatBoost attributes the highest global importance to T2M and the lowest to SWGDN and H500 compared with other algorithms (Figure 2a). This is also reflected in CatBoost's SHAP values for H500, which are near zero when H500 is below 5700 m (Figure 2g). We further conduct a sensitivity experiment in which SWGDN, H500, PS, and EVAP are excluded from the input features (Figure S9). However, inter-model discrepancies persist, primarily due to strong collinearity between T2M and DOY (Figure S1). Removing DOY reduces these discrepancies but degrades performance across all five models. Since our goal is to use XML to reveal underlying physical relationships, we prioritize retaining a comprehensive set of relevant features, even if some are correlated and may lead to overfits, so that subtle but important processes can be kept (e.g., ozone responses to H500 or SWGDN) (Jiang et al., 2024).

3.2 The impacts of meteorology on surface ozone trends over 2013–2023

We assess the contribution of meteorology to long-term ozone trends by calculating linear trends in warm-season (April-September) SHAP values for each input variable using a parametric regression method (Lu et al., 2020). We divide the study period into two phases (pre-2019 and post-2019) since 2019 marks a turning point in the ozone trend, transitioning from rapid increases to stabilization or decline (Figure 3). SHAP values enable attribution of observed ozone trends to individual meteorological variables and two temporal features. DOY captures the seasonal cycle, and UNIX reflects non-meteorological influences such as emission changes (see Section 2.3). The observed warm-season ozone trends (dashed line in Figure 3) are 3.5, 2.1, and 1.8 ppbv/yr for NCP, YRD, and PRD, respectively, before 2019, and -0.66, 0.14, and 0.06 ppbv/yr after 2019. The five ML models reproduce these trends reasonably well, with some discrepancies, particularly in PRD, where Random Forest and Extra Trees underestimate the pre-2019 increase and overestimate the post-2019 trend.

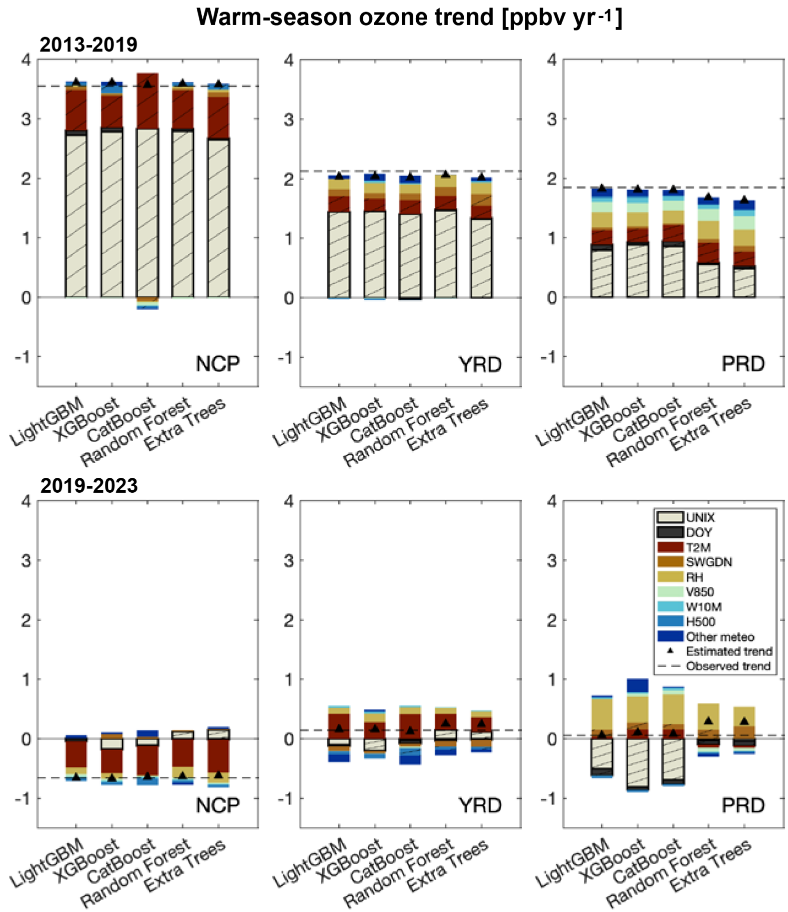

Figure 3. Attribution of warm-season (April to September) surface ozone trends to meteorological and temporal variables for the periods 2013–2019 (top panels) and 2019–2023 (bottom panels) in three Chinese regions (NCP, YRD, and PRD). Results from the five ML models (LightGBM, XGBoost, CatBoost, Random Forest, and Extra Trees) are presented as individual bars, with colors representing contributions from specific meteorological and temporal (UNIX and DOY) variables. Hatches denote statistically certain trends ($p<0.05$).

We find that meteorological and non-meteorological factors show contrasting roles in shaping surface ozone trends before and after 2019. In NCP and YRD, all five ML models indicate that the UNIX variable accounts for over half of the 2013–2019 ozone increases, whereas its influence is weak during 2019–2023, when meteorological factors become the primary drivers. In PRD, results from LightGBM, XGBoost, and CatBoost indicate that the UNIX term explains nearly 46% of the pre-2019 ozone increase but shifts to a negative influence after 2019, offsetting the positive effects of meteorology. By contrast, Random Forest and Extra Trees attribute only ~28% of the pre-2019 ozone rise to the UNIX term and show almost no influence over 2019–2023, reflecting weaker ozone–UNIX relationships in these models (Figure S6). These findings align with previous studies showing that China's clean air actions contributed to rising ozone levels during Phase 1 (2013–2017) but led to slight declines during Phase 2 (2018–2020) as coordinated controls of NMVOCs and NO_x were implemented (Liu et al., 2023; Wang et al., 2023). Our XML results support these findings, demonstrating that the UNIX variable, a proxy for non-meteorological influences such as emissions, plays distinct roles in the evolution of ozone concentrations before and after 2019.

We analyze the influence of specific meteorological variables on ozone trends using SHAP values. Temperature (T2M) is the most critical factor before 2019, with all five ML models consistently identifying its significant contribution to ozone trends ($p < 0.05$) across the three regions. In NCP, temperature is particularly important, contributing to ozone increases of 0.53-0.93 ppbv/yr across models during 2013–2019, followed by ozone declines after 2019. Such shifts in meteorological influences around 2019 in NCP have also been reported in a recent study that attributed them to reversed changes in meteorological variables (Wang et al., 2024). In YRD, the combined effects of temperature, solar radiation, and humidity contribute to ozone increases of ~0.54 ppbv/yr during 2013–2019. After 2019, although changes in temperature and humidity moderately enhance ozone levels, their effects are largely offset by other meteorological variables. In PRD, temperature, humidity, and wind conditions (V850 and W10M), along with other meteorological factors, collectively enhance surface ozone before 2019, whereas RH becomes the primary driver of ozone increases afterward.

3.3 Meteorology-induced ozone anomaly

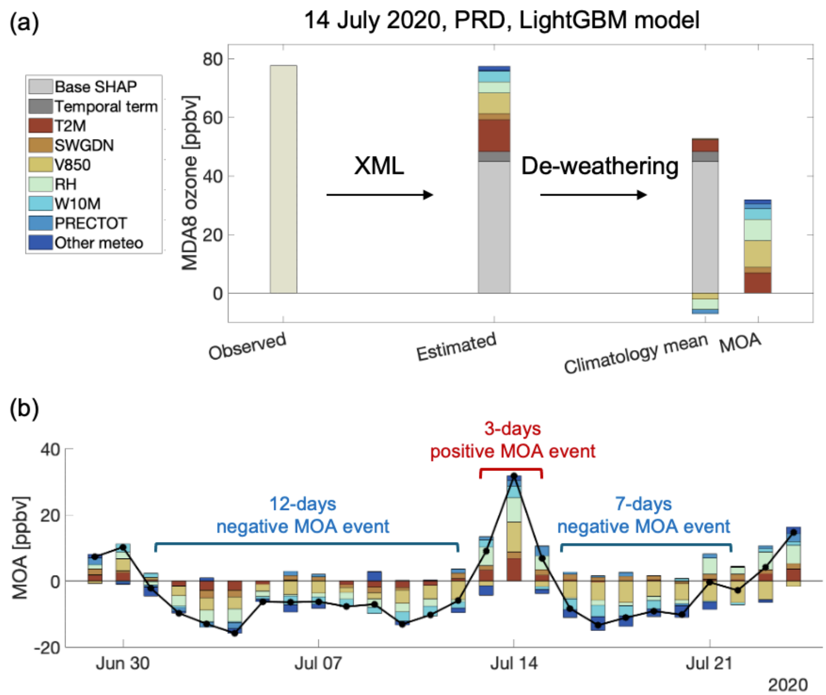

Figure 4. Schematic diagram of the attributable de-weathering approach, using an event of 14 July 2020 in the PRD region as diagnosed by the LightGBM model. (a) Steps for deriving the climatological mean ozone and the meteorology-induced ozone anomaly (MOA) on 14 July 2020. (b) MOA values in PRD from 29 June to 24 July 2020, with identified MOA events highlighted.

We now apply the attributable de-weathering approach described in Section 2.3 to decompose daily ozone into a climatological mean and meteorology-induced ozone anomaly (MOA). Figure 4a shows a pollution event in PRD, with regional mean surface ozone peaking on 14 July 2020. The observed MDA8 ozone reached 77.6 ppbv on 14 July 2020, while the five ML models estimated similar values ranging from 76.4 to 77.5 ppbv (only the LightGBM result is shown for clarity). After de-weathering, the climatological ozone level is estimated to be ~45 ppbv. Several meteorological variables, such as RH, V850, and PRECTOT, negatively contribute to this climatology, reflecting the influence of 15-day running averages (Eq. (3)) in mid-July. Subtracting it from the total ozone yields the MOA being as high as 32 ppbv on that day. All five models identify V850 as the most influential meteorological factor, accounting for about one-third of the MOA. Additional important factors include RH, T2M, PRECTOT, SWGDN, and 10-m wind speed (W10M), all enhancing the MOA during this pollution event (Figure 4b). We note that extreme ozone events may also be influenced by non-meteorological factors (e.g., emission pulses or regional transport of precursors), which are captured by the UNIX term in our framework but are not included in the MOA calculation. Therefore, MOA should be interpreted as the meteorologically driven component of ozone anomalies, rather than a comprehensive representation of all extreme ozone events.

We focus on positive MOA events during warm seasons, and calculate the average magnitude and duration for each event (e.g., Figure 4b). Figure 5 shows their interannual variations and trends from 2013 to 2023. Averaged across five ML models, the mean magnitudes of positive MOA values are 11.7 ppbv in NCP, 11.5 ppbv in YRD, and 15.9 ppbv in PRD (Table S3). Our analysis reveals significant increasing trends ($p < 0.05$) in MOA magnitude across all three regions, with annual rates of 0.36, 0.28, and 0.43 ppbv/year in NCP, YRD, and PRD, respectively. After 2019, the magnitude of MOA events becomes more variable. For instance, in 2021, MOA values in NCP were notably low, mainly due to the reduced T2M influences. In contrast, PRD experienced a sharp peak in MOA in 2022, reaching nearly 20 ppbv, largely driven by decreased RH. Figure S10 compares the global importance of original SHAP values with those based on anomalies. In NCP, although T2M is less dominant in anomaly-based importance compared to the original ranking, it remains the most influential factor in explaining ozone anomalies. In YRD, the importance of temperature drops from first in the original ranking to third in the anomaly-based ranking, where it contributes comparably to RH and SWGDN. Moreover, the relative importance of V850 increases in both NCP and PRD, indicating enhanced influence of synoptic-scale transport on daily ozone variability.

In contrast to the increasing trend in MOA magnitude, none of the five ML models indicates significant trends in the duration of MOA events over the study period (Figure 5). However, the MOA duration in PRD shows notable interannual variability. On average, positive MOA events last 3.1 days in NCP, 3.4 days in YRD, and 4.3 days in

PRD. Notably, a prolonged positive MOA event lasting over 30 days was recorded in PRD from 27 August to 26 September 2022. XML results suggest that this event was mainly driven by sustained anomalies in RH, T2M, and V850 (Figure S11). The extended duration of positive MOA in PRD relative to NCP and YRD may be attributed to persistent weather systems such as the subtropical high and typhoons (Chen et al., 2022; Wang et al., 2018). We also analyze the annual counts of positive MOA events and find no significant trends in any of the three regions (Figure 5). These findings suggest that the intensified meteorological influence on surface ozone in China from 2013 to 2023 is primarily due to increasing MOA magnitudes, rather than more frequent or longer-lasting episodes of unfavorable meteorological conditions.

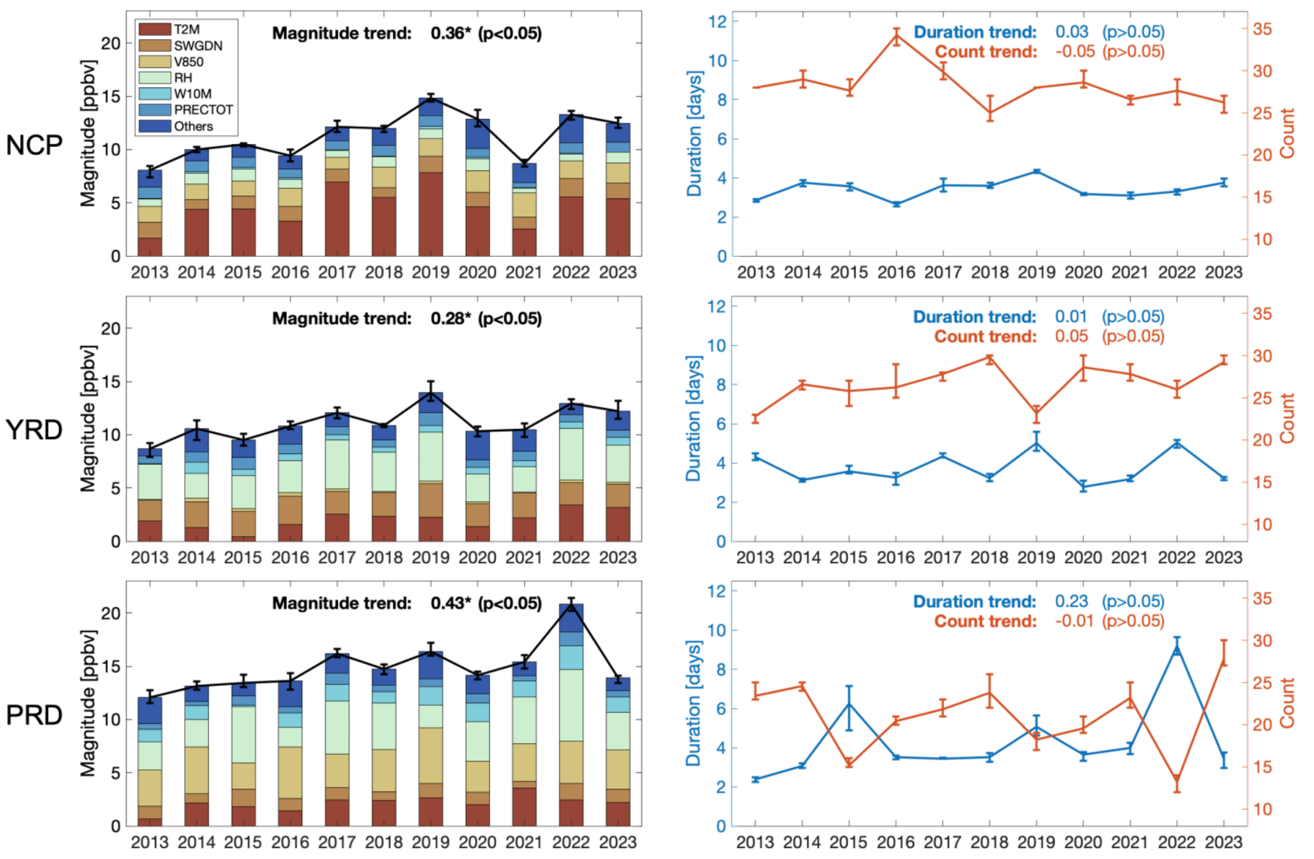

Figure 5. Left column: Magnitude of positive MOA events during warm seasons from 2013 to 2023 for NCP (top), YRD (middle), and PRD (bottom). Bars represent contributions from individual meteorological variables. Trends in

unit of ppbv/year are shown inset. Right column: Duration and count of positive MOA events over the same period. Results are averaged across five ML models, with error bars indicating the range among models.

4. Conclusions and uncertainty analyses

In summary, we have applied the SHAP technique to the outputs of five ML models to analyze the impact of meteorology on surface ozone in three Chinese regions. All five models consistently identify important meteorological drivers of ozone variability, including temperature, solar radiation, humidity, wind, and geopotential height. The XML results also reveal their nonlinear relationships with ozone, which generally align well with established physical understanding. Our results highlight a temporal shift in the dominant drivers of ozone trends over the past decade. During the early phase of China's clean air actions (prior to 2019), rising ozone levels were largely attributed to emission changes, while in the later phase (post-2019), meteorological variability become prominent. Among meteorological factors, temperature is identified as the most influential variable, with significant contributions to pre-2019 ozone trends in all three regions across all five models. Furthermore, the SHAP-based de-weathering approach shows a clear increase in the magnitude of positive meteorology-induced ozone anomaly (MOA) events in recent years.

Our work highlights not only the valuable insights that XML can offer but also the potential challenges associated with its application to ozone attribution and broader environmental analysis. One major challenge is the presence of multicollinearity among meteorological variables, which complicates the interpretation of individual variable contributions. Machine learning models may arbitrarily assign higher importance to one correlated variable over another, leading to potentially misleading conclusions. This issue is particularly relevant when using XML outputs for policy-relevant assessments, such as estimating the ozone climate penalty (i.e., dO_3/dT). More broadly, the collinearity among input variables can introduce uncertainty in explaining the relative importance of different drivers. In our case, only the UNIX and DOY terms are included as non-meteorological indicators due to their weak correlations with meteorological variables (Figure S8). However, incorporating additional inputs, such as concentrations of other pollutants, would likely increase the model's complexity and hinder interpretability. It is also noted that systematic differences in meteorological reanalysis products for certain variables may influence the relative importance assigned to predictors. For instance, Lu et al. (2025) reported discrepancies in surface shortwave radiation between MERRA-2 and ERA5 over eastern China, potentially affecting radiation-related attributions. While our conclusions regarding the dominant role of temperature remain robust, future studies could benefit from multi-reanalysis ensembles to quantify this uncertainty.

Second, inconsistencies in XML results may arise from two sources: variations in feature attributions across different models and differences in the models' predictive performance. Disentangling these two sources of uncertainty is inherently difficult. For instance, in our study, the contribution of the UNIX variable to ozone trends in PRD, as identified by Random Forest and Extra Trees, differs significantly from that identified by the other algorithms (Figure 3). This discrepancy could be due to weaker modeled relationships between ozone and UNIX in the two models, larger biases compared to observational data, or a combination of both factors. In practice, models with higher predictive performance tend to provide more reliable explanations identifying the key features.

Finally, our analyses rely exclusively on the SHAP technique due to its ability to provide local explainability information. However, the use of alternative XML methods may introduce additional layers of uncertainty in interpretation. For example, Table S4 compares global feature importance rankings generated by three different XML approaches, SHAP, Gain, and Permutation, for the LightGBM model. The inconsistencies among these importance rankings suggest that only general features, such as which variables are influential and which are not, can be reliably inferred. This highlights the need for future work to develop standardized evaluation metrics that can help distinguish more reliable and consistent explanations from the diverse outputs of multiple XML methods, such as those evaluated by Krishna et al. (2022).

Data availability

Surface ozone observations from the national air quality monitoring network are operated by the China Ministry of Ecology and Environment (https://air.cnemc.cn). MERRA-2 meteorological reanalysis data are available from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) at https://disc.gsfc.nasa.gov/. All derived data products used in this study are available from the corresponding author upon request.

Code availability

The machine-learning and SHAP analyses were implemented using standard open-source libraries (https://github.com/microsoft/FLAML; https://github.com/shap/shap). The scripts used in this study are available from the corresponding author upon request.

Author contributions

L.Z. designed the study. X.Y. performed the research. X.W., N.L., S.H. and A.T.A. helped with results interpretation. X.Y. and L.Z. wrote the paper with inputs from all the co-authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (grant no. 42275106) and the National Key Research and Development Program of China (2023YFC3706104).

References

Archibald, A. T., Turnock, S. T., Griffiths, P. T., Cox, T., Derwent, R. G., Knote, C., and Shin, M.: On the changes in surface ozone over the twenty-first century: sensitivity to changes in surface temperature and chemical mechanisms, Philos. Trans. A. Math. Phys. Eng. Sci., 378 (2183), 20190329, 2020.

Bloomer, B. J., Stehr, J. W., Piety, C. A., Salawitch, R. J., and Dickerson, R. R.: Observed relationships of ozone air pollution with temperature and emissions, Geophys. Res. Lett., 36, L09803, 2009.

Breiman, L.: Random forests, Mach. Learn., 45, 5–32, 2001.

Chen, T. and Guestrin, C.: XGBoost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), Association for Computing Machinery, New York, NY, USA, 785–794, 2016.

Chen, X., Wang, N., Wang, G., Wang, Z., Chen, H., Cheng, C., Li, M., Zheng, L., Wu, L., Zhang, Q., Tang, M., Huang, B., Wang, X., and Zhou, Z.: The influence of synoptic weather patterns on spatiotemporal characteristics of ozone pollution across Pearl River Delta of southern China, J. Geophys. Res.-Atmos., 127, e2022JD037121, 2022.

Chen, B., Zhen, L., Wang, L., Zhong, H., Lin, C., Yang, L., Xu, W., and Huang, R.-J.: Revisiting the impact of temperature on ground-level ozone: a causal inference approach, Sci. Total Environ., 953, 176062, 2024.

Cohen, A. J., Brauer, M., Burnett, R., Anderson, H. R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C. A., Shin, H., Straif, K., Shaddick, G., Thomas, M., Van Dingenen, R., Van Donkelaar, A., Vos, T., Murray, C. J. L., and Forouzanfar, M. H.: Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, Lancet, 389, 1907–1918, 2017.

Dang, R., Liao, H., and Fu, Y.: Quantifying the anthropogenic and meteorological influences on summertime surface ozone in China over 2012–2017, Sci. Total Environ., 754, 142394, 2021.

Fiore, A. M., Naik, V., Spracklen, D. V., Steiner, A., Unger, N., Prather, M., Bergmann, D., Cameron-Smith, P. J., Cionni, I., Collins, W. J., Dalsøren, S., Eyring, V., Folberth, G. A., Ginoux, P., Horowitz, L. W., Josse, B., Lamarque, J. F., MacKenzie, I. A., Nagashima, T., O'Connor, F. M., Righi, M., Rumbold, S. T., Shindell, D. T., Skeie, R. B., Sudo, K., Szopa, S., Takemura, T., and Zeng, G.: Global air quality and climate, Chem. Soc. Rev., 41, 6663–6683, 2012.

Flora, M. L., Potvin, C. K., McGovern, A., and Handler, S.: A machine learning explainability tutorial for atmospheric sciences, Artif. Intell. Earth Syst., 3, e230018, 2024.

Fowler, D., Pilegaard, K., Sutton, M. A., Ambus, P., Raivonen, M., Duyzer, J., Simpson, D., Fagerli, H., Fuzzi, S., Schjoerring, J. K., Granier, C., Neftel, A., Isaksen, I. S. A., Laj, P., Maione, M., Monks, P. S., Burkhardt, J., Daemmgen, U., Neirynck, J., Personne, E., Wichink-Kruit, R., Butterbach-Bahl, K., Flechard, C., Tuovinen, J. P., Coyle, M., Gerosa, G., Loubet, B., Altimir, N., Gruenhage, L., Ammann, C., Cieslik, S., Paoletti, E., Mikkelsen, T. N., Ro-Poulsen, H., Cellier, P., Cape, J. N., Horváth, L., Loreto, F., Niinemets, Ü., Palmer, P. I., Rinne, J., Misztal, P., Nemitz, E., Nilsson, D., Pryor, S., Gallagher, M. W., Vesala, T., Skiba, U., Brüggemann, N., Zechmeister-Boltenstern, S., Williams, J., O'Dowd, C., Facchini, M. C., de Leeuw, G., Flossman, A., Chaumerliac, N., and Erisman, J. W.: Atmospheric composition change: ecosystems–atmosphere interactions, Atmos. Environ., 43, 5193–5267, 2009.

Fu, T. M. and Tian, H.: Climate change penalty to ozone air quality: review of current understandings and knowledge gaps, Curr. Pollut. Rep., 5, 159–171, 2019.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Silva, A. M. da, Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Clim., 30, 5419–5454, 2017.

Geurts, P., Ernst, D., and Wehenkel, L.: Extremely randomized trees, Mach. Learn., 63, 3–42, 2006.

Gong, X., Kaulfus, A., Nair, U., and Jaffe, D. A.: Quantifying O_3 impacts in urban areas due to wildfires using a generalized additive model, Environ. Sci. Technol., 51, 13216–13223, 2017.

Grange, S. K. and Carslaw, D. C.: Using meteorological normalisation to detect interventions in air quality time series, Sci. Total Environ., 653, 578–588, 2019.

Han, H., Liu, J., Shu, L., Wang, T., and Yuan, H.: Local and synoptic meteorological influences on daily variability in summertime surface ozone in eastern China, Atmos. Chem. Phys., 20, 203–222, 2020.

Jaffe, D. A. and Zhang, L.: Meteorological anomalies lead to elevated O_3 in the western US in June 2015, Geophys. Res. Lett., 44, 1990–1997, 2017.

Jiang, S., Sweet, L., Blougouras, G., Brenning, A., Li, W., Reichstein, M., Denzler, J., Shangguan, W., Yu, G., Huang, F., and Zscheischler, J.: How interpretable machine learning can benefit process understanding in the geosciences, Earth's Future, 12, e2024EF004540, 2024.

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, in: Advances in Neural Information Processing Systems 30 (NeurIPS 2017), Curran Associates, Inc., 2017.

Krishna, S., Han, T., Gu, A., Pombra, J., Jabbari, S., Wu, S., and Lakkaraju, H.: The disagreement problem in explainable machine learning: a practitioner's perspective, arXiv, 8 February 2022, available at: http://arxiv.org/abs/2202.01602 (last access: 10 May 2024), 2022.

Lakshmanan, V., Karstens, C., Krause, J., Elmore, K., Ryzhkov, A., and Berkseth, S.: Which polarimetric variables are important for weather/no-weather discrimination? J. Atmos. Ocean. Technol., 32, 1209–1223, 2015.

Li, K., Jacob, D., Shen, L., Lu, X., De Smedt, I., and Liao, H.: 2013–2019 increases of surface ozone pollution in China: anthropogenic and meteorological influences, Atmos. Chem. Phys., 20, 11423–11440, 2020.

Li, T., Lu, Y., Deng, X., and Zhan, Y.: Spatiotemporal variations in meteorological influences on ambient ozone in China: a machine learning approach, Atmos. Pollut. Res., 14, 101720, 2023.

Li, Z., Wang, Y., Liu, J., and Xian, J.: Using machine learning to unravel chemical and meteorological effects on ground-level ozone: Insights for ozone–climate control strategies, Environ. Int., 201, 109567, 2025.

Liu, R., Ma, Z., Liu, Y., Shao, Y., Zhao, W., and Bi, J.: Spatiotemporal distributions of surface ozone levels in China from 2005 to 2017: a machine learning approach, Environ. Int., 142, 105823, 2020.

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, 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, 2023.

Lu, X., Hong, J., Zhang, L., Cooper, O. R., Schultz, M. G., Xu, X., Wang, T., Gao, M., Zhao, Y., and Zhang, Y.: Severe

surface ozone pollution in China: a global perspective, Environ. Sci. Technol. Lett., 5, 487–494, 2018.
Lu, X., Zhang, L., and Shen, L.: Meteorology and climate influences on tropospheric ozone: a review of natural sources, chemistry, and transport patterns, Curr. Pollut. Rep., 5, 238–260, 2019.
Lu, X., Zhang, L., Wang, X., Gao, M., Li, K., Zhang, Y., Yue, X., and Zhang, Y.: Rapid increases in warm-season surface ozone and resulting health impact in China since 2013, Environ. Sci. Technol. Lett., 7, 240–247, 2020.
Lu, X., Liu, Y., Su, J., Weng, X., Ansari, T., Zhang, Y., He, G., Zhu, Y., Wang, H., Zeng, G., Li J., He, C., Li S., Amnuaylojaroen, T., Butler, T., Fan Q., Fan S., Forster, G. L., Gao, M., Hu, J., Kanaya, Y., Latif, M. T., Lu, K., Nédélec, P., Nowack, P., Sauvage, B., Xu X., Zhang, L., Li, K., Koo, J., and Nagashima T.: Tropospheric ozone trends and attributions over East and Southeast Asia in 1995–2019: an integrated assessment using statistical methods, machine learning models, and multiple chemical transport models. Atmos. Chem. and Phys., 25, 7991–8028, 2025.
Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, in: Advances in Neural Information Processing Systems 30 (NeurIPS 2017), Curran Associates, Inc., 2017.
Ma, R., Ban, J., Wang, Q., Zhang, Y., Yang, Y., He, M. Z., Li, S., Shi, W., and Li, T.: Random Forest model based fine scale spatiotemporal O_3 trends in the Beijing-Tianjin-Hebei region in China, 2010 to 2017, Environ. Pollut., 276, 116635, 2021.
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, 2024.
Mills, G., Pleijel, H., Malley, C. S., Sinha, B., Cooper, O. R., Schultz, M. G., Neufeld, H. S., Simpson, D., Sharps, K., Feng, Z., Gerosa, G., Harmens, H., Kobayashi, K., Saxena, P., Paoletti, E., Sinha, V., and Xu, X.: Tropospheric ozone assessment report: present-day tropospheric ozone distribution and trends relevant to vegetation, Elem. Sci. Anth., 6, 47, 2018.
Mo, Y., Li, Q., Karimian, H., Zhang, S., Kong, X., Fang, S., and Tang, B.: Daily spatiotemporal prediction of surface ozone at the national level in China: an improvement of CAMS ozone product, Atmos. Pollut. Res., 12, 391–402, 2020.
Porter, W. C. and Heald, C. L.: The mechanisms and meteorological drivers of the summertime ozone–temperature relationship, Atmos. Chem. Phys., 19, 13367–13381, 2019.
Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A.: CatBoost: unbiased boosting with categorical features, in: Advances in Neural Information Processing Systems 31 (NeurIPS 2018), Curran Associates, Inc., 2018.
Requia, W. J., Di, Q., Silvern, R., Kelly, J. T., Koutrakis, P., Mickley, L. J., Sulprizio, M. P., Amini, H., Shi, L., and Schwartz, J.: An ensemble learning approach for estimating high spatiotemporal resolution of ground-level ozone in the contiguous United States, Environ. Sci. Technol., 54, 11037–11047, 2020.
State Council of the People's Republic of China: Notice of the General Office of the State Council on issuing the Air Pollution Prevention and Control Action Plan (in Chinese), available at: https://www.gov.cn/zwgk/2013-09/12/content_2486773.htm (last accessed: 30 October 2025), 2013.
State Council of the People's Republic of China: Notice of the General Office of the State Council on issuing the Three-Year Action Plan for Winning the Blue Sky Defense Battle (in Chinese), available at: https://www.gov.cn/zhengce/content/2018-07/03/content_5303158.htm (last accessed: 30 October 2025), 2018.

Travis, K. R. and Jacob, D. J.: Systematic bias in evaluating chemical transport models with maximum daily 8 h average (MDA8) surface ozone for air quality applications: a case study with GEOS-Chem v9.02, Geosci. Model Dev., 12, 3641–3648, 2019.

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, 2019.

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, 2017.

Wang, H., Lyu, X., Guo, H., Wang, Y., Zou, S., Ling, Z., Wang, X., Jiang, F., Zeren, Y., Pan, W., Huang, X., and Shen, J.: Ozone pollution around a coastal region of South China Sea: interaction between marine and continental air, Atmos. Chem. Phys., 18, 4277–4295, 2018.

Wang, Y., Zhao, Y, Liu, Y., Jiang, Y., Zheng, B., Xing, J., Liu, Y., Wang, S., and Nielsen, C. P.: Sustained emission reductions have restrained the ozone pollution over China, Nat. Geosci., 16, 967–974, 2023.

Wang, M., Chen, X., Jiang, Z., He, T.-L., Jones, D., Liu, J., and Shen, Y.: Meteorological and anthropogenic drivers of surface ozone change in the North China Plain in 2015–2021, Sci. Total Environ., 906, 167763, 2024.

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, 2022.

Wu, W., Fu, T.-M., Arnold, S. R., Spracklen, D. V., Zhang, A. X., Tao, W., Wang, X. L., Hou, Y., Mo, J., Chen, J., Li, Y., Feng, X., Lin, H., Huang, Z., Zheng, J. Y., Shen, H., Zhu, L., Wang, C., Ye, J., and Yang, X.: Temperature-dependent evaporative anthropogenic VOC emissions significantly exacerbate regional ozone pollution, Environ. Sci. Technol., 58, 5430–5441, 2024.

Ye, X., Wang, X., and Zhang, L.: Diagnosing the model bias in simulating daily surface ozone variability using a machine learning method: the effects of dry deposition and cloud optical depth, Environ. Sci. Technol., 56, 16665–16675, 2022.

Yao, T., Lu, S., Wang, Y., Li, X., Ye, H., Duan, Y., Fu, Q., and Li, J.: Revealing the drivers of surface ozone pollution by explainable machine learning and satellite observations in Hangzhou Bay, China, J. Clean. Prod., 440, 140938, 2024.

Zhan, Y., Luo, Y., Deng, X., Grieneisen, M. L., Zhang, M., and Di, B.: Spatiotemporal prediction of daily ambient ozone levels across China using random forest for human exposure assessment, Environ. Pollut., 233, 464–473, 2018.

Zhang, J. J., Wei, Y., and Fang, Z.: Ozone pollution: a major health hazard worldwide, Front. Immunol., 10, 2518, 2019.

Zhang, X., Waugh, D. W., Kerr, G. H., and Miller, S. M.: Surface ozone–temperature relationship: the meridional gradient ratio approximation, Geophys. Res. Lett., 49, e2022GL098680, 2022.

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, 2024.