Atmospheric
Chemistry
and Physics

Insights into ozone pollution control in urban areas by decoupling meteorological factors based on machine learning

Yuqing Qiu[1], Xin Li[1,2], Wenxuan Chai[3], Yi Liu[1], Mengdi Song[1], Xudong Tian[4], Qiaoli Zou[4], Wenjun Lou[5], Wangyao Zhang[5], Juan Li[5], and Yuanhang Zhang[1]

[1]College of Environmental Sciences and Engineering, Peking University, Beijing 100871, China
[2]Institute of Carbon Neutrality, Peking University, Beijing 100871, China
[3]China National Environmental Monitoring Center, Beijing 100012, China
[4]Zhejiang Ecological and Environmental Monitoring Center, Hangzhou 310012, China
[5]Jinhua Ecological and Environmental Monitoring Center, Jinhua 321015, China

Correspondence: Xin Li (li_xin@pku.edu.cn) and Wenxuan Chai (chaiwx@cnemc.cn)

Abstract. Ozone (O_3) pollution is posing significant challenges to urban air quality improvement in China. The formation of surface O_3 is intricately linked to chemical reactions which are influenced by both meteorological conditions and local emissions of precursors (i.e., NO_x and volatile organic compounds, VOCs). When meteorological conditions deteriorate, the atmosphere's capacity to cleanse pollutants decreases, leading to the accumulation of air pollutants. Although a series of emission reduction measures have been implemented in urban areas, the effectiveness of O_3 pollution control proves inadequate. Primarily due to adverse changes in meteorological conditions, the effects of emission reduction are masked. In this study, we integrated a machine learning model, an observation-based model, and a positive matrix factorization model based on 4 years of continuous observation data from a typical urban site. We found that transport and dispersion impact the distribution of O_3 concentration. During the warm season, positive contributions of dispersion and transport to O_3 concentration ranged from 12.9 % to 24.0 %. After meteorological normalization, the sensitivity of O_3 formation and the source apportionment of VOCs changed. The sensitivity of O_3 formation shifted towards the transition regime between VOC- and NO_x-limited regimes during the O_3 pollution event. Vehicle exhaust became the primary source of VOC emissions after "removing" the effect of dispersion, contributing 41.8 % to VOCs during the pollution periods. On the contrary, the contribution of combustion to VOCs decreased from 33.7 % to 25.1 %. Our results provided new recommendations and insights for implementing O_3 pollution control measures and evaluating the effectiveness of emission reduction in urban areas.

1 Introduction

Ozone (O_3) plays a significant role in atmospheric oxidation and global climate. It is also considered one of the major atmospheric pollutants. A high concentration of surface O_3 is harmful to human health, such as causing respiratory diseases and even cancer (Cohen et al., 2017; Monks et al., 2015). In recent years, China has been in a stage of rapid economic development, accompanied by the emergence of various air pollution problems due to industrialization and urbanization (Zhang et al., 2012). In order to deal with the air pollution, the Chinese government has issued some control policies, such as the Clean Air Action Plan in 2013 (Chinese State Council, 2013) and the Blue Sky Protection Campaign in 2018 (Chinese State Council, 2018). These policies have resulted in reductions in the concentrations of particulate matter (PM), nitrogen dioxide (NO_2), and sulfur dioxide (SO_2) (Zheng et al., 2018). However, O_3 pollution has be-

come increasingly serious, especially in typical urban clusters such as Beijing–Tianjin–Hebei (BTH), the Yangtze River Delta (YRD), and the Fenwei Plain (FWP). In 2022, the 90th percentile of maximum daily 8 h average (MDA8) O_3 was $179\,\mu g\,m^{-3}$ in the BTH, $162\,\mu g\,m^{-3}$ in the YRD, and $167\,\mu g\,m^{-3}$ in the FWP, 4.7 %, 7.3 %, and 1.2 % higher than that in 2021, respectively (Ministry of Ecology and Environment of China, https://www.mee.gov.cn/hjzl/dqhj/qgkqzlzk/, last access: 31 January 2025). Frequent O_3 pollution events have attracted the attention of the public and the government. Surface O_3 is mainly formed by the photochemical reactions of volatile organic compounds (VOCs) and nitrogen oxides ($NO_x = NO + NO_2$) (Atkinson, 2000). The emissions of precursors effectively affect the change in O_3 concentration (Tan et al., 2018). The sources of VOCs are complex and widespread, making it challenging to control emissions (Miller et al., 2002; Borbon et al., 2013). Meteorological conditions can directly or indirectly affect O_3 concentration (Liu and Wang, 2020; Zhang et al., 2015; Gao et al., 2021; Han et al., 2020). Wind and boundary layer height influence the diffusion of the concentrations of O_3 and its precursor. Poor dispersion can result in a decrease in atmospheric environmental capacity, making O_3 pollution events more likely to occur even with low precursor emissions. High ultraviolet radiation and temperature promote photochemical reactions of O_3 formation (Yang et al., 2019). In addition, O_3 can be transported over long distances due to its the long atmospheric lifetime, which can cause regional O_3 problems (Han et al., 2019). In short, the O_3 concentration is nonlinearly affected by meteorological conditions, emissions of precursors, and chemical reactions (Fu et al., 2019; Hu et al., 2021; Zhang et al., 2022, 2024).

Li et al. (2020) discovered that approximately one-third of the growth of O_3 concentration in summer in China was attributed to meteorological conditions. This indicated that the reduction of air pollutant concentrations due to the control policies may be offset by the deterioration of meteorological conditions. Therefore, decoupling meteorological factors from temporal concentration series of atmospheric pollutants is helpful to assess the impact of clean air actions. At present, many mathematical statistical methods have been developed to "remove" the influences of meteorological factors. The technique for predicting air pollutant concentrations under randomly selected meteorological parameters was first introduced by Grange et al. (2018). Weng et al. (2022) found that the temperature near the surface 2 m, the downward radiation flux of the surface and the relative humidity were the most important meteorological factors to affect O_3 concentration in China by applying two machine learning algorithms (ridge regression and random forest regression). Mousavinezhad et al. (2021) employed the Kolmogorov–Zurbenko (KZ) filter method and found that meteorological factors played the dominant role in O_3 formation in four typical urban agglomerations in China. Guo et al. (2022) used the random forest method to obtain the characteristics of air pollution in

12 megacities in China from 2013 to 2020 and carried out a comprehensive assessment of the actual impact of the national clean air action. Compared to traditional statistical methods, machine learning models perform better at removing meteorological effects from concentration data.

In response to severe O_3 pollution, a series of emission reduction measures targeting O_3 precursors were implemented in urban areas. However, the effectiveness of controlling O_3 pollution fell short of expectations. According to previous studies, O_3 formation in urban areas was more sensitive to VOCs (Feng et al., 2019; Wang et al., 2023), with anthropogenic emissions of VOCs playing a dominant role (Ahmad et al., 2017). Understanding the sensitivity of O_3 formation and the source characteristics of VOCs is helpful to design effective strategies to control O_3 pollution. The observation-based model (OBM), positive matrix factorization model (PMF), and air quality model are commonly used to analyze the causes of O_3 pollution and provide theoretical support for reducing O_3 precursors. However, the results of the OBM and PMF, which rely on observed data, may be influenced by fluctuations in meteorological conditions, potentially introducing bias. Wu et al. (2023) developed initial concentration-dispersion normalized PMF (ICDN-PMF) to reflect the changes in source emissions of VOCs in Qingdao. The results proved that the contribution of solvent use was overestimated due to air dispersion during O_3 pollution. Additionally, the actual effectiveness of emission reduction measures can also be obscured by unfavorable meteorological conditions. In this study, we applied the random forest (RF) method proposed by Grange et al. (2018) to remove the dispersion and transport effects on O_3 concentration, as well as the dispersion effect on precursors in Hangzhou from 2019 to 2022. After meteorological normalization, the concentrations of VOCs were imported into the OBM and PMF to obtain the sensitivity of O_3 formation and the contributions of emission sources, providing more accurate results. The interplay of meteorological and local factors on O_3 pollution can be evaluated effectively and comprehensively using this method. Our results emphasized the importance of decoupling the meteorological effects of transport and dispersion for understanding the mechanisms of local O_3 formation and devising appropriate emission reduction measures.

2 Methods

2.1 Observation data

The online hourly observation data from 2019 to 2022 were measured by the Zhejiang Ecological and Environmental Monitoring Center (30.29° N, 120.13° E). This station was located in the urban area of Hangzhou, Zhejiang Province, surrounded by residential and commercial areas. The dataset of air pollutants included O_3; NO_2; and 98 different kinds of VOCs detected by the gas chromatography system, including 29 alkanes, 11 alkenes, 1 alkyne, 16 aromatics,

28 halohydrocarbons, 12 oxygenated VOCs (OVOCs), and 1 acetonitrile. The online gas chromatography system was equipped with a mass spectrometer and flame ionization detector (GC-MS/FID) (ZF-PKU-VOC1007, Beijing Pengyu Changya Environmental Technology Co. Ltd., China), which used a dual gas path separation method. VOCs with low carbon numbers (C_2–C_5) were measured by a FID, while VOCs with high carbon numbers (C_5–C_{10}) were detected by a MS. NO_2 and O_3 were measured by a commercial instrument (Model 42i/42iTL and Model 49i, Thermo Scientific, USA). The meteorological parameters measured included temperature (T), relative humidity (RH), atmospheric pressure (P), wind speed (WS), and wind direction (WD), which were measured by the WS500-UMB instrument manufactured by Lufft Corporation. In addition, we used the meteorological data from the ERA5 reanalysis product (Hersbach et al., 2020), such as boundary layer height (BLH) and ultraviolet radiation B (UVB). The EAR5 meteorological data are spatial grid data with a resolution of $0.25° \times 0.25°$, available at https://cds.climate.copernicus.eu/datasets (last access: 31 January 2025). The back trajectories were calculated backwards in time for 24 h and started 500 m above ground level using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Stein et al., 2015). The meteorological data from the Global Data Assimilation System (GDAS) with a horizontal resolution of 1° longitude $\times$ 1° latitude were adopted in the trajectory model. The back trajectories were then clustered into five clusters using the Euclidian distance. Clusters of backward trajectories have been widely employed to represent the main directions of air masses at monitoring sites (Song et al., 2021).

2.2 Meteorological normalization method

Random forest is a versatile classifier that comprises multiple decision trees, applicable to classification, regression, and dimension reduction problems. When constructing each tree in the RF model, a dataset of the same size is selected for training, potentially containing duplicates. This sampling method, which involves putting instances back into the dataset, is referred to as bootstrap. At each node, the optimal segmentation is calculated by randomly selecting a subset of features from the entire set. The RF model describes the relationship between the time series of atmospheric pollutant concentrations and their corresponding feature. We constructed the RF model based on original datasets, which contained air pollutant variables (O_3, NO_2, total non-methane hydrocarbon compounds (NMHCs), and 98 VOC species), time variables (trend, hour, weekday, month, and day of year), and meteorological variables (T, RH, P, WS, WD, UVB, BLH, and cluster). In the RF model, the air pollutants were the response variables, while the explanatory variables included time variables representing source emissions and meteorological variables representing physical and chemical processes. Time variables such as day of year, month, weekday, and hour

are used to indicate the seasonal, weekly, and daily cycles of emission intensity (Dai et al., 2023a, b; Vu et al., 2019). The parameter "trend" can indicate the long-term changes in air pollutant concentrations resulting from the implementation of policy measures (Vu et al., 2019). Environmental regulations and policies aimed at reducing pollutant emissions were implemented during specific periods, and their effects became apparent in the long-term trends. Therefore, the trend not only reflected changes in emission sources closely related to activity levels but also represented the long-term variations in air pollutants caused by the enforcement of policies and regulations. The parameter trend was calculated using Eq. (1):

$$\text{trend} = \text{year}_i + \frac{t_{JD} - 1}{N_i} + \frac{t_H}{24 N_i}, \tag{1}$$

where N_i is the number of days in the year$_i$ (year$_i$ is from 2019 to 2022), t_H is hour time (0–23, referring to 00:00–23:00), and t_{JD} is the day of the year (1–365) (Carslaw and Taylor, 2009).

Temperature was a key factor influencing the rate of chemical reactions, with higher temperatures typically promoting the photochemical reactions that generate O_3. UVB served as the driving force for the photochemical reactions, directly impacting O_3 formation. Additionally, humidity played an important role in the chemical processes involved in O_3 formation. Therefore, T, RH, and UVB were identified as the key features associated with atmospheric photochemical reactions. WS influences the dispersion of atmospheric pollutants. At high wind speeds, air pollutants tended to be dispersed, while low wind speeds resulted in local pollutant accumulation, leading to increased concentrations. WD determined the dispersion path of atmospheric pollutants. BLH was a critical factor affecting the vertical dispersion of pollutants. A higher boundary layer allowed pollutants to disperse more effectively into the upper atmosphere, reducing surface concentrations, whereas a lower boundary layer resulted in pollutant accumulation near the ground. Thus, WS, WD, and BLH were regarded as the features of atmospheric physical dispersion on a local scale. The cluster can serve as a feature of transport from remote regions.

There are approximately 32 856 valid data with a time resolution of 1 h. The RF model was trained using a forest of 1000 trees. Training the datasets of the RF model was conducted using 80 % of the original datasets, and the remaining 20 % were selected for testing. Correlation coefficients (r^2), root-mean-square error (RMSE), FAC2 (fraction of predictions with a factor of 2), mean bias (MB), mean gross error (MGE), normalized mean bias (NMB), normalized mean gross error (NMGE), coefficient of efficiency (COE), and index of agreement (IOA) were used to evaluate model performance (Table S2 in the Supplement). Based on previous related research, these statistical measures indicated that the model performed well (Emery et al., 2017; Henneman et al., 2017; Vu et al., 2019).

The process of meteorological normalization involved replacing the original meteorological variables with those randomly resampled from the observation dataset and using the established RF model to predict atmospheric pollutant concentrations under different meteorological conditions. The resampling of meteorological variables was conducted over the 2-week period before and after the selected date, with the resampled hours remaining constant. This approach effectively preserved the seasonal and diurnal variations in the response variables (Vu et al., 2019). The resampling and prediction process was repeated 1000 times to generate 1000 predicted pollutant concentrations. The average values were taken as the final meteorologically normalized concentrations. In the meteorologically normalization process of O_3 concentration, meteorological variables such as WS, WD, BLH, and the cluster, which signify dispersion and transport, were randomly sampled. In the case of O_3 precursors, namely NO_2 and NMHCs, resampling was exclusively applied to WS, WD, and BLH. NO_2 and NMHCs have short atmospheric lifetimes, making them less susceptible to the influence of regional transport over large scales (Wang et al., 2023). To take into consideration that some NMHCs have relatively long lifetimes (such as acetylene), the cluster was incorporated as an explanatory variable in the RF model. For NMHCs with different lifetimes, the feature importance of the cluster was relatively low (around 1 %). Therefore, it can be approximated that NMHCs were primarily influenced by dispersion effects within the uncertainty. Feature importance was used to reflect the overall significance of explanatory variables in the RF model (Yang et al., 2023). The importance was typically represented as an array, where each value corresponded to the importance score of a specific feature. These scores usually range from 0 to 1. The higher importance score indicated that the feature had a stronger predictive capability for the response variable. The RF model was constructed using R "deweather" packages developed by Carslaw (https://github.com/openair-project/deweather, last access: 1 February 2025).

2.3 Observation-based model

An observation-based model is used in this study to simulate the formation of O_3. The model is based on Regional Atmospheric Chemical Mechanisms version 2 (RACM2) updated with a detailed isoprene oxidation mechanism (Goliff et al., 2013). As a 0-D model, this model incorporates dilution mixing within the boundary layer. However, vertical or horizontal transport of the air mass is not considered in this model. Detail of the observation-based box model can be found in Tan et al. (2017). The photolysis frequencies (J values) were calculated using the Tropospheric Ultraviolet and Visible (TUV) model (Wolfe et al., 2016). Model calculations were constrained to measured trace gases, including inorganic species (NO_2 and O_3) and organic species (VOCs). In addition, physical parameters like J values, tempera-

ture, pressure, and relative humidity were also constrained to measured values. The empirical kinetic modeling approach (EKMA) serves as a sensitivity test for the OBM. The EKMA curve offers a means to quantify intricate nonlinear relationships among O_3, NO_x, and VOCs, which can be used as a theoretical basis for designing O_3 pollution reduction strategies (Tan et al., 2018). In this study, a total of 30 emission scenarios were established for both NO_x and anthropogenic VOCs. Subsequently, O_3 concentrations resulting from changes in these precursor emissions were simulated across 900 scenarios. The EKMA curve was plotted according to the O_3 formation rate under different VOCs and NO_x conditions.

2.4 Positive matrix factorization

The positive matrix factorization model is based on a large number of data to estimate the compositions and contributions of emission sources (Paatero and Tapper, 1994). The PMF model is widely used for VOC source apportionment (Song et al., 2021; Yuan et al., 2010). In the PMF model, it is assumed that the pollutant concentrations measured at the receptor point can be represented as a linear sum of components emitted by different sources. Indeed, the temporal variation in atmospheric pollutants is influenced not only by emissions but also by dispersion (Dai et al., 2020). Direct PMF analysis based on observed data may lead to the loss of real information regarding emission sources. In this study, the observed and meteorologically normalized VOC concentrations were fed into US EPA PMF v5.0 to identify and quantify major emission sources of VOCs. In contrast to the PMF results based on observation, examining the alterations in contributions of emission sources after meteorological normalization can reveal the impact of dispersion on VOC sources. The RF model for meteorological normalization was a nonlinear machine learning algorithm. To satisfy the fundamental mathematical requirement of the PMF model, which stated that the total concentration was a linear combination of contributions from individual sources, the RF model was applied for meteorological normalization of individual VOC species and total VOCs in this study. This ensured that the sum of the meteorologically normalized VOC species remained linearly correlated with total VOCs (Fig. S4), indicating that the nonlinear processing did not significantly alter the overall structure of total VOC concentrations. With this approach, the results obtained by inputting the meteorologically normalized data into the PMF model were reasonable.

3 Results and discussion

3.1 Temporal variations of O_3 and its precursors

3.1.1 Long-term variations

Figure 1 displayed the time series of air pollutant concentrations based on observation and meteorological normaliza-

tion from 2019 to 2022. After meteorological normalization, the concentrations of O_3 and its precursors were primarily affected by local factors, including precursor emission and chemical reactions. From a long-term perspective, the trends of air pollutant concentrations after meteorological normalization were consistent with those based on observation. After meteorological normalization, MDA8 O_3 significantly decreased in 2020, followed by a slight increase in 2021 and 2022. The observed annual variation in MDA8 O_3 exhibited a similar trend. The meteorologically normalized annual mean MDA8 O_3 in 2020 decreased by 10 % compared to 2019, which aligned with the observed change of -8.7 %. Based on both meteorologically normalized and observed results, the concentrations of NO_2 and NMHCs showed declining trends, with a significant decrease in 2022. Compared to 2019, the meteorologically normalized concentrations of NO_2 and NMHCs in 2022 decreased by 46.1 % and 24 %, respectively, while the observed concentrations of NO_2 and NMHCs decreased by 45.7 % and 16 %, respectively. This indicated that the variation in O_3 concentration in Hangzhou was mainly driven by precursor emissions and chemical formation in the long term.

From the diurnal variation in NO_2 and NMHC concentrations, it is shown that the observed concentrations were lower during the day and higher at night, which was contrary to the daily trends of WS and BLH (Fig. S1). Stable WS and low BLH at night were not conducive to the diffusion of air pollutants, resulting in the accumulation of pollutant concentrations, while the situation was opposite during the day (Song et al., 2018). After the dispersion effect was removed, the precursor concentrations decreased at night and increased significantly during the day. The diurnal variation in the MDA8 O_3 concentration showed a typical single-peak structure before and after meteorological normalization. Different from the change in the concentrations of precursors, the MDA8 O_3 concentration increased at night and decreased during the day after meteorological normalization. At night, the titration reaction of NO_x and the horizontal transport reduced the O_3 concentration (Li et al., 2022). The NO_x concentration decreased after meteorological normalization, and the weakening of titration resulted in the increase in O_3 concentration at night. In addition, the decrease in horizontal transport at night also resulted in the increase in O_3 concentration after normalization. During the day, the destruction of the stable boundary layer strengthened the vertical mixing effect of the atmosphere, so that the O_3 in the upper atmosphere mixed with the O_3 generated near the surface, increasing the O_3 concentration (Lei et al., 2023). When the effect of transport was removed, the daytime MDA8 O_3 concentration decreased. It can be seen from the diurnal variations that meteorological factors directly affected the concentrations of precursors through dispersion. Meteorological factors not only directly affected the O_3 concentration through horizontal and vertical transport but also indirectly changed

O_3 concentration by influencing precursor concentration and titration reaction.

Figure 2 showed the importance of the different features in the RF model. The time variables can represent anthropogenic emissions to some extent. Time variables were closely related to the periodic changes in human activities. For example, weekdays versus weekends and peak versus non-peak hours corresponded to different levels of anthropogenic emissions. Anthropogenic emissions influenced the seasonal variations of atmospheric pollutants, as seen in winter heating effects. Previous studies also used time variables to represent anthropogenic emissions (Dai et al., 2023a, b; Song et al., 2023; Vu et al., 2019). The chemical reaction of O_3 formation was affected by meteorological factors such as UVB, T, and RH. Local dispersion of O_3 and its precursors was mainly affected by WS, WD and BLH, and long-distance transport of O_3 was characterized by the cluster. The importance of local chemical reactions to O_3 was 83.9 %. UVB, influencing photochemical reactions, emerged as the most crucial factor for O_3 concentration, with an importance of 25.9 %. This is consistent with the findings by Weng et al. (2022) in the same region. Additionally, the importance of RH and T to O_3 was also evident, with the importance of 18.2 % and 11.3 % respectively. Higher relative humidity was usually associated with a higher cloud cover, and relative humidity was generally negatively correlated with O_3 (Liu et al., 2023).

High temperatures increased the rate of most chemical reactions in the atmosphere, especially photochemical reactions that lead to O_3 formation (Li et al., 2020). In addition, elevated temperature enhanced the emission of biogenic VOCs (Lu et al., 2019). Hence, some O_3 pollution events were associated with high temperature (Dang et al., 2021). Ding et al. (2023) found that temperature was the dominant factor affecting O_3 concentration in Tianjin. Wind and BLH also played significant roles in O_3 concentration (16.1 %), mainly through vertical diffusion, vertical convection, and horizontal convection (Li et al., 2012).

Different from O_3, BLH exerted the most significant impact on NO_2 and NMHC variation, with importance values of 26.1 % and 20 %, respectively. Turbulent mixing in the active boundary layer facilitated the dispersion of air pollutants, whereas the stable boundary layer attenuated vertical diffusion, thereby intensifying the accumulation of air pollutants near the ground (Huang et al., 2020). The importance of dispersion to NO_2 and NMHCs was 34.2 % and 30.7, respectively. Consequently, unfavorable meteorological dispersion conditions can result in the accumulation of precursors, causing O_3 pollution even in scenarios with low emissions. Temporal variables representing emissions, such as month and day of year, also occupied important positions. The importance of month to NO_2 and NMHCs exceeded 18 %, which represented the significant influences of seasonal anthropogenic emissions on the concentrations of precursors. The importance of local emission, production, and
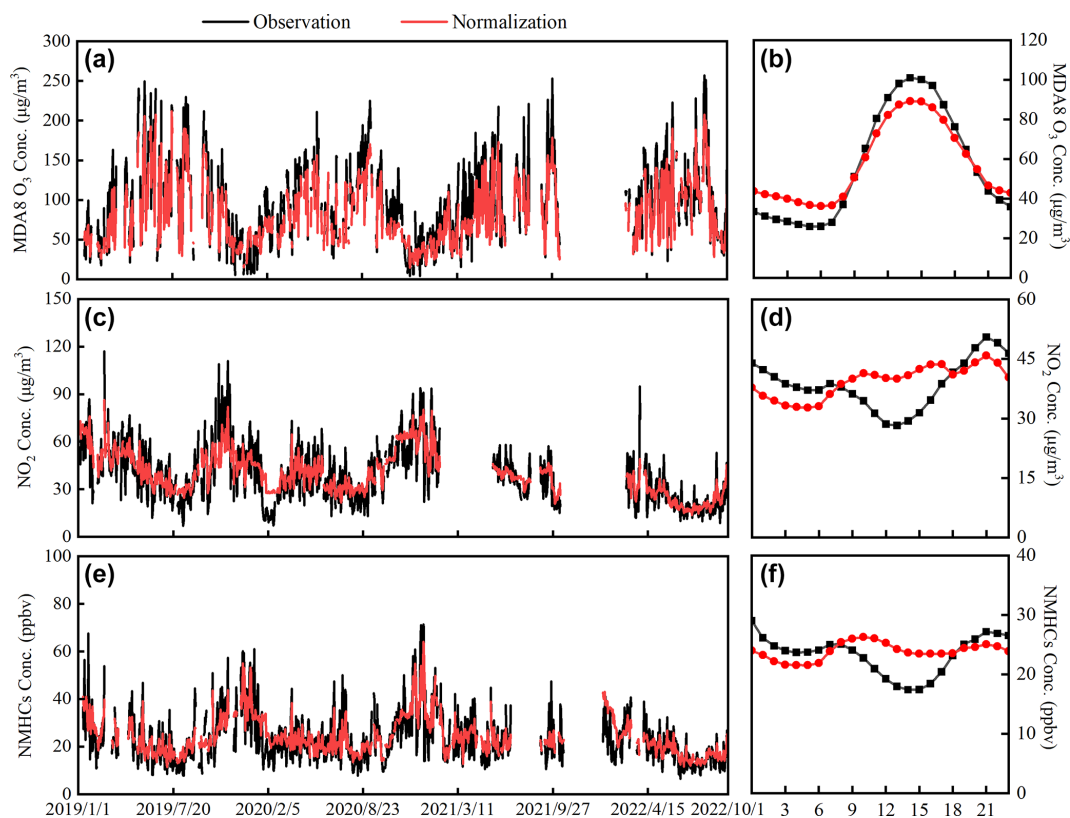
Figure 1. Long-term trends of daily average concentrations of air pollutants **(a, c, e)** and mean diurnal variations of air pollutant concentrations **(b, d, f)** based on observation and meteorological normalization from 2019 to 2022.
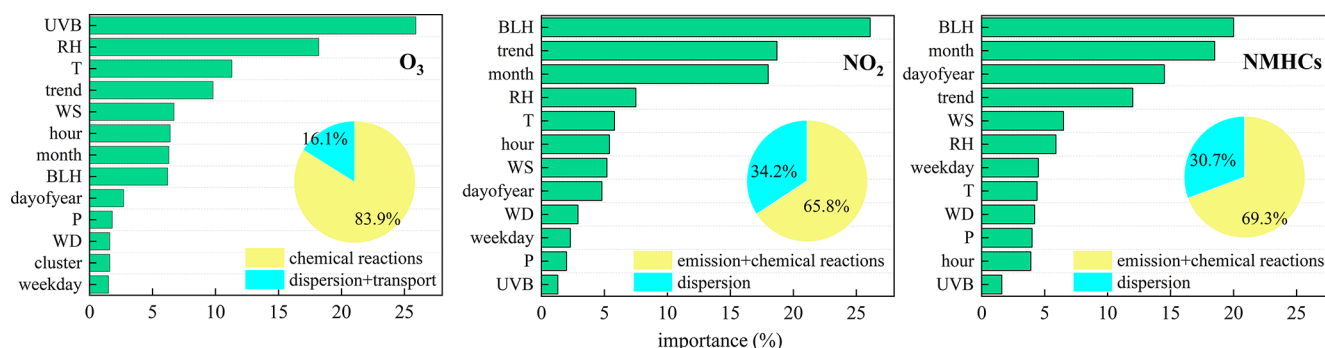

Figure 2. The importance of each feature to O_3, NO_2, and NMHCs in the RF model.

consumption to NO_2 and NMHCs was 65.8 % and 69.3 %, respectively (Fig. 2).

3.1.2 Comparison between pollution and non-pollution periods

O_3 pollution occurs frequently between May and September each year. In order to evaluate the influences of meteorological conditions on the concentrations of O_3 and its precursors, the relative change in air pollutant concentrations caused by meteorological factors during O_3 pollution and non-pollution periods in the warm season (From May to September) from 2019 to 2022 was analyzed (Fig. 3). In the non-pollution periods, the negative effect of dispersion on the concentrations of NO_2 and NMHCs was apparent, with average relative changes ranging from −9.3 % to −27.98 % for NO_2 and −10.5 % to −22.8 % for NMHCs. Dispersion and transport have less influence on the MDA8 O_3 concentrations, with the average relative change ranging from −0.1 % to 8.1 %. During the pollution periods, the positive effects of dispersion and transport on O_3 became evident (from 12.9 % to 24.0 %). Simultaneously, the negative effect of dispersion on
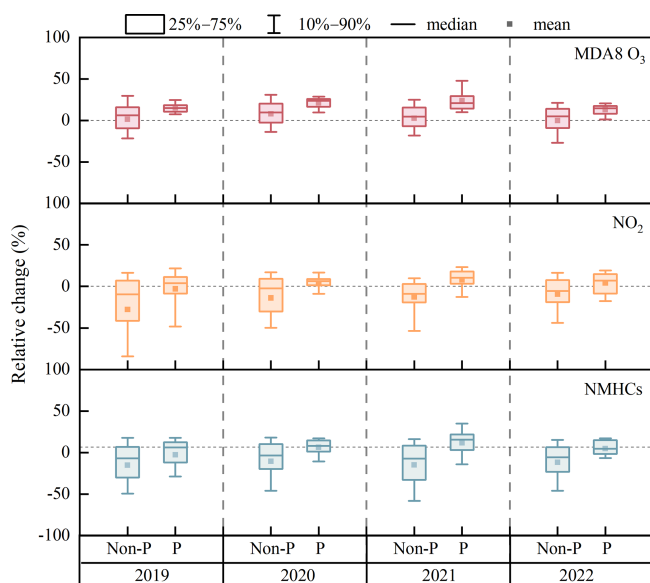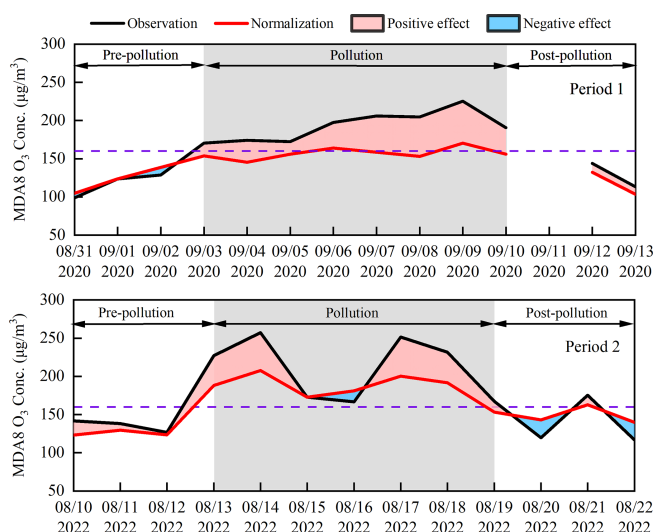
Figure 3. Relative change caused by meteorological factors during O_3 pollution (P) and non-pollution (Non-P) periods in the warm season from 2019 to 2022. Relative change = the observed concentrations − the meteorologically normalized concentrations / the observed concentrations.

Figure 4. The MDA8 O_3 concentration based on observation and meteorological normalization and the contributions of dispersion and transport to the MDA8 O_3 during pre-pollution, pollution, and post-pollution periods in Period 1 and Period 2 (red: positive contribution, blue: negative contribution).

the concentrations of precursors decreased and even transformed into a positive effect. In 2021 in particular, dispersion had a significant positive effect on NO_2 and NMHCs, with an average relative change of 7.8 % and 11.8 %, respectively. O_3 concentration was affected by the long-distance transport as well as the deterioration of diffusion conditions in the pollution periods. Therefore, the influence of meteorological factors on O_3 was more obvious than that of its precursors during pollution periods in the warm season.

3.1.3 Variations during short-term pollution events

In order to explore the effects of meteorological dispersion and transport on O_3 concentration in the short term, we selected two typical pollution periods from 2019 to 2022 (Fig. 4). During Period 1 (31 August to 13 September in 2020), the average MDA8 O_3 in Hangzhou was 193 µg m^{-3} in the pollution period, exceeding the national air quality standard (> 160 µg m^{-3}; GB 3095-2012, MEE, 2018). At the same time, other cities in the YRD regions such as Shanghai, Nanjing, Wuxi, Changzhou, Suzhou, and Jiaxing also experienced O_3 pollution (Fig. S2). Period 1 represented a large-scale regional pollution event. During the pre-pollution period (31 August to 2 September in 2020), dispersion and transport had negative effects on MDA8 O_3. In the pollution period (3 to 10 September in 2020), the concentration of locally generated O_3 (depicted by the red line) remained below the limit, with an average concentration of 157 µg m^{-3}, with only slight exceedances recorded on 6 and 9 Septem-

ber. Locally generated O_3 was produced in the atmosphere through photochemical reactions involving VOCs and nitrogen oxides (NO_x; Song et al., 2021). However, the actual observed O_3 concentration was much higher than the standard, and the O_3 concentration was about 200 µg m^{-3} from 6 to 10 September. The positive contribution of dispersion and transport was significant (depicted by the red area) in the pollution periods, resulting in an 18.7 % increase in the MDA8 O_3 concentration. During the post-pollution period, contributions of dispersion and transport decreased significantly.

In Period 2 (10 to 22 August in 2022), the average MDA8 O_3 concentration in Hangzhou was as high as 211 µg m^{-3} during the pollution period, while the concentration of MDA8 O_3 in most surrounding cities was less than 160 µg m^{-3}. Thus the O_3 pollution in Period 2 was influenced by both local formation and transport. During the pollution period (13 to 19 August in 2022), locally generated O_3 basically exceeded the standard, and the MDA8 O_3 concentration was greater than 180 µg m^{-3} on most days, with an average concentration of 185 µg m^{-3}. On 16 August, the meteorological negative contribution (−14.4 %) appeared, exerting dilution effects on the O_3 concentration, but the MDA8 O_3 on that day still exceeded 160 µg m^{-3}, indicating intense local O_3 production. The positive contributions of dispersion and transport to O_3 were significant during the pollution periods; the contributions ranged from 8.5 % to 20.4 %. For precursors, the concentration of NMHCs increased between 17 and 19 August (Fig. S3). The positive contribution of dispersion to NO_2 and NMHCs ranged from 4.4 % to 13.7 % and from 0.6 % to 8.5 % during pollution. During the post-pollution period (20 to 22 August in 2022), the contributions of dis-

persion and transport turned negative, indicating that meteorological diffusion conditions favored the elimination of O_3 pollution.

3.2 VOC–NO_x–O_3 sensitivity

Unfavorable meteorological conditions can cause the accumulation of O_3, making it essential to have a clear understanding of local O_3 formation pathways for effective control of O_3 pollution. The relationship between O_3 and NO_2 under long-term trends was analyzed based on the observed (left) and meteorologically normalized (right) data (Fig. 5). The dotted red line shows the turning point of the relationship between O_3 and NO_2 concentrations. The blue triangle represented the average value of the MDA8 O_3 during the warm season each year. On the left side of the dotted red line, O_3 concentration elevated with the increase in NO_2 concentration. At this point, controlling the emission of NO_2 was conducive to limiting the formation of O_3, suggesting that the sensitivity of O_3 formation was limited by NO_x. On the right side of the dotted red line, O_3 concentration decreased with the increase in NO_2 concentration. At this point, the inhibition effect of NO_x emission reduction on O_3 formation was not significant, and it is necessary to control the emission of VOCs, indicating that the sensitivity of O_3 formation was limited by VOCs (Kong et al., 2024). After meteorological normalization, the NO_2 concentration in the turning point increased from 9 to $19\,\mu g\,m^{-3}$, suggesting that when NO_2 concentration was at a higher level, O_3 concentration decreased with the increase in NO_2 concentration. In other words, a higher NO_2 value at the turning point suggested a greater likelihood that the actual NO_x concentration was below that value, indicating a higher probability of being in a NO_x-limited regime. In addition, based on average results in warm season each year, the sensitivity of O_3 formation before and after meteorological normalization was also shown in Fig. 5. Whether based on observed or meteorologically normalized data, the O_3 formation from 2019 to 2021 was located in the VOC-limited regime, while O_3 production entered the transition regime between VOC- and NO_x-limited regimes in 2022. The transition regime referred to the region near the turning point, where O_3 formation was sensitive to changes in both VOCs and NO_x.

The OBM was used to analyze the sensitivity of O_3 formation. The OBM is zero-dimensional, meaning it excludes the processes of atmospheric transport and dispersion. Therefore, it is reasonable to remove the influences of transport and dispersion when using the OBM. The VOC–NO_x–O_3 sensitivity and the net ozone production rate ($P(O_3)$) exhibited significant differences before and after meteorological normalization in the short-term pollution events (Fig. 6). The O_3 concentration in Period 2 was affected by both transport and local formation. The concentration of local precursors increased after removing the effect of dispersion, resulting in the change in the sensitivity of O_3 formation. Based on

the observation results, the O_3 formation in pollution was located in the strict NO_x-limited regime. After meteorological normalization, O_3 formation shifted towards the transition regime between VOC- and NO_x-limited regimes. The limitation of O_3 formation by NO_x concentration was weakened. After removing the influence of dispersion and transport on O_3 concentrations, the value of $P(O_3)$ increased, indicating that the $P(O_3)$ calculated based on observation was likely underestimated. Therefore, when the OBM was used to analyze the VOC–NO_x–O_3 sensitivity, removing the influences of dispersion and transport was beneficial to accurately identify the limited regime of O_3 formation.

3.3 VOC source apportionment

The PMF method was further used for VOC source analysis. The optimal solution was selected by examining the interpretability of factors and the distribution of scale residuals. Based on observed and meteorologically normalized concentrations, seven possible emission sources of VOC from May to September in 2022 were extracted using the US EPA PMF v5.0. The possible emission sources of VOC included combustion, industrial source, vehicle exhaust, fuel evaporation, secondary and aging source, biogenic source, and solvent use. The differences in the source profiles resolved from the observed and normalized concentrations are illustrated in Fig. S5.

The combustion source was characterized by high concentrations of ethane, propane, and acetylene. Low carbon alkanes and alkenes were likely to be the products of incomplete combustion (Wang et al., 2015). Acetylene was a typical tracer of combustion. Toluene and some halohydrocarbons, such as chloromethane, were also released from combustion (Liu et al., 2008). Additionally, the proportion of acetonitrile was also high, which was an important product of biomass combustion (de Gouw et al., 2003). Biomass combustion emission was relatively intense in the YRD. The industrial source was characterized by halohydrocarbons (Sun et al., 2016), and 1,2-dichloroethane accounted for nearly 80 % of this factor in both PMF results. Vehicle exhaust was featured by high concentrations of ethane, propane, isobutane, n-butane, isopentane, ethylene, and toluene(Cai et al., 2010; Liu et al., 2008). Fuel evaporation was characterized by the high concentration and proportion of isopentane, isobutane, n-butane, and n-pentane, while the concentration of acetylene was minimal in this factor. The secondary and aging source was characterized by halohydrocarbons and oxygenated VOCs (OVOCs). Methacrolein (MACR) and methyl vinyl ketone (MVK) were products of the oxidation of isoprene (Mo et al., 2018). OVOCs and halohydrocarbons have long lifetimes in the atmosphere and can serve as important tracers for aging sources (Yang et al., 2021b). The biogenic source was featured by highest concentration of isoprene, primarily emitted by plants (Gong et al., 2018). Additionally, the oxidation products of isoprene (MACR and MVK) also
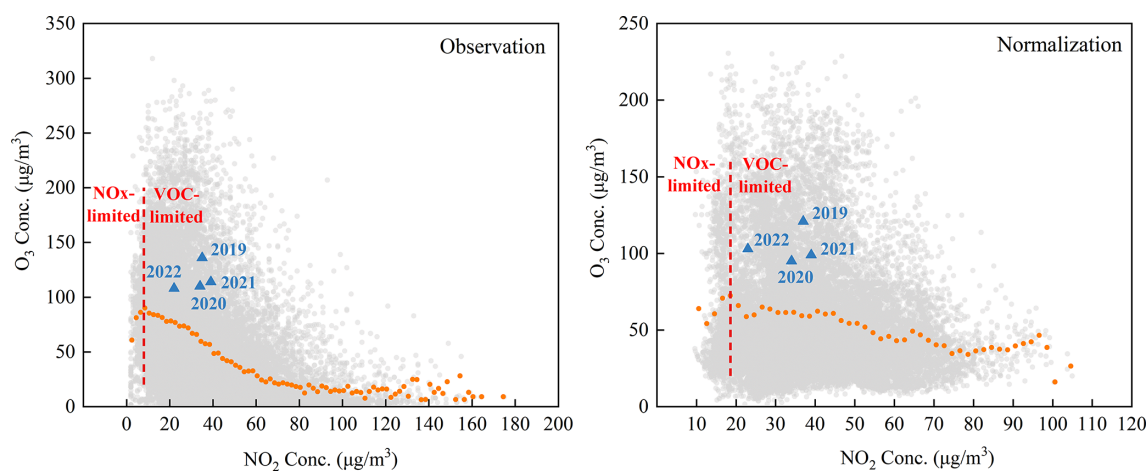
Figure 5. The changes in O_3 concentration and NO_2 concentration from 2019 to 2022. The light-gray circles represent the hourly O_3 concentration. The orange circles represent the average value of O_3 concentration in each interval ($2\,\mu g\,m^{-3}$) of NO_2. The blue triangle represents the average value of the MDA8 O_3 during the warm season each year.
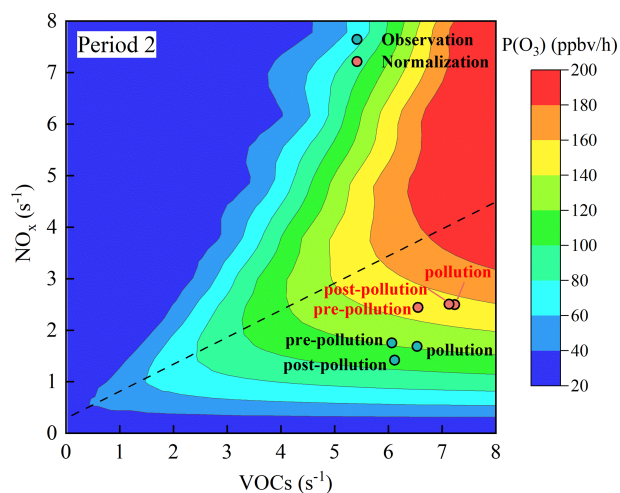

Figure 6. The O_3 isopleth diagram versus NO_x and anthropogenic VOCs using the EKMA. The circles represent the average concentrations of NO_x and VOC during pre-pollution, pollution, and post-pollution periods in Period 2.

contributed to this factor. The solvent source was characterized by high concentrations of aromatics. Toluene, ethylbenzene, *m*-xylene, and *o*-xylene are commonly used as materials in solvents (Song et al., 2021).

After smoothing out the effect of dispersion, the absolute contribution of emission sources to VOCs changed. The mean absolute contribution of vehicle exhaust to VOCs increased most significantly, from 3.97 to 6.72 ppbv during the non-pollution periods and from 6.84 to 9.76 ppbv during the pollution periods. The mean absolute contribution of combustion decreased by 1.55 and 2.09 to 2.86 and 5.85 ppbv during the non-pollution and pollution periods, respectively. Dispersion caused overestimation of the contribution of com-

bustion to VOCs, which indicates that the reduction in VOC concentration by abatement measures can be offset by the effect of dispersion. Therefore, the impact of dispersion should be taken into account when evaluating the effectiveness of emission reduction measures on VOC emission sources. The normalized contributions of solvent use and industrial source in the pollution were comparable, with an average absolute contribution of 2.78 and 2.57 ppbv. In comparison to the results based on the observations, the absolute contribution of fuel evaporation decreased from 1.94 to 1.33 ppbv after meteorological normalization during the pollution periods. After meteorological normalization, the contributions of biogenic source and secondary and aging source to VOCs during the pollution period were relatively low, with absolute contributions of 0.54 ppbv.

Figure 8 showed the proportion of VOC sources before and after meteorological normalization during the non-pollution periods and pollution periods. The pie charts for normalized source contributions illustrate the relative contribution of each source to the total VOC concentration after removing the effects of dispersion. According to the results of the observations, combustion and vehicle exhaust were the largest contributors to VOCs, accounting for 27.1 % and 24.3 % in the non-pollution periods. And the proportion of combustion and vehicle exhaust increased to 33.7 % and 29 % in the pollution periods. During the pollution periods, the proportion of solvent use and fuel evaporation also increased, accounting for 15.9 % and 8.2 %, respectively. After the normalization of dispersion, vehicle exhaust became the predominant emission source of VOCs (37 % in the non-pollution periods and 41.8 % in the pollution periods), much higher than the proportion of other emission sources. According to the motor vehicle data released by the Zhejiang Public Security Department in 2022, the number of motor vehicles reached 23.29 million. During the non-
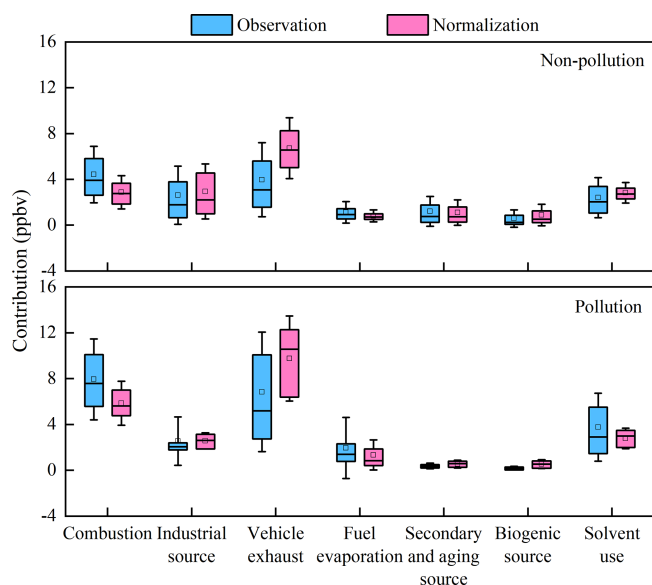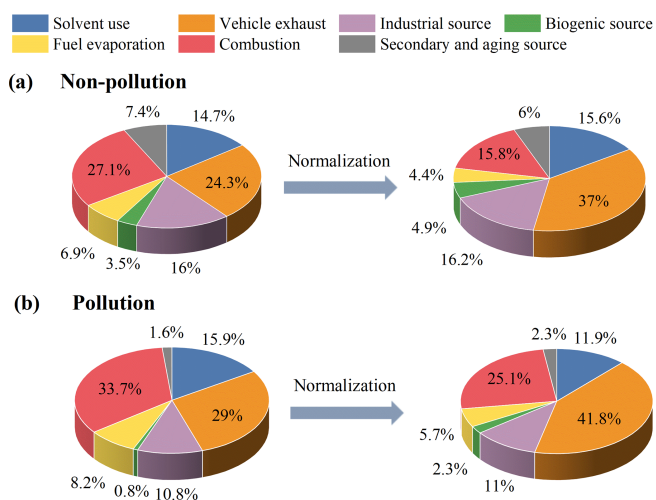
Figure 7. The absolute contributions of emission sources to VOCs based on observation and meteorological normalization during the non-pollution periods and pollution periods in the warm season in 2022.
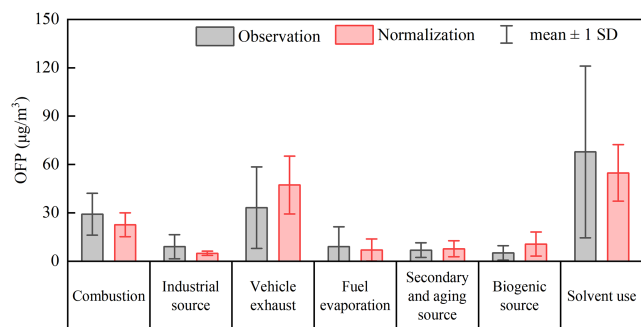

Figure 8. Comparison of VOC source proportions before and after meteorological normalization during the non-pollution periods and pollution periods in the warm season in 2022.

pollution periods, the contributions of solvent use, industrial source, and combustion were comparable, accounting for the proportions ranging from 15.6 % to 16.2 %. Compared to the non-pollution periods, the influence of combustion on VOCs increased (25.1 %), while the proportions of industrial source and solvent use decreased during the pollution periods (11 % and 11.9 %). Straw burning occurs frequently in Zhejiang Province. According to the remote sensing monitoring of straw burning announced by the Ecological Environment Monitoring Center of Zhejiang Province, a total of 135 straw-burning points in the province were monitored by satellite remote sensing from January to October 2022. The proportion of industrial emission and solvent use decreased during the pollution periods, and the VOC concentrations from these two sources also declined (Fig. 7), indicating that the implementation of shutdown or off-peak production measures at the time of pollution warning was effective.

The O_3 formation potential (OFP) is used to assess VOC photochemical activity (Carter, 2010), and it can be calculated using Eq. (2):

$$OFP_i = MIR_i \times [VOC_i], \tag{2}$$

where MIR_i represents the maximum incremental reactivity for VOC species i, and $[VOC]_i$ represents the concentration of VOC species i ($\mu g\,m^{-3}$). MIR values for each VOC species were taken from the updated Carter research results (https://intra.engr.ucr.edu/~carter/SAPRC/, last access: 1 February 2025). The contributions of emission sources to OFP were further analyzed and are shown in Fig. 9. Based on the results of the observations, the emission

sources that contribute the most to OFP were solvent use ($67.79\,\mu g\,m^{-3}$), vehicle exhaust ($33.16\,\mu g\,m^{-3}$), and combustion ($29.16\,\mu g\,m^{-3}$) during the pollution periods in the warm season in 2022. After removing the effect of dispersion, the contribution of vehicle exhaust to OPF increased to $47.25\,\mu g\,m^{-3}$, while the contribution of solvent use and combustion to OFP decreased to 54.77 and $22.58\,\mu g\,m^{-3}$, respectively. The actual contributions of combustion and solvent use to O_3 formation were larger under the dispersion effect. Thus, it was necessary to consider the cumulative effect of dispersion and enhance emission reduction measures for specific emission sources. For Period 2 mentioned in Sect. 3.1.3, we also found that the contributions of VOC emission sources changed after meteorological normalization (Figs. S6 and S7). After removing the dispersion effect, the contributions of solvent use and vehicle exhaust to OFP increased during the pollution periods, while the contribution of combustion and secondary and aging source decreased. From 17 to 19 August, the normalized contribution of solvent source to OFP was significant, with an average OFP of $105.81\,\mu g\,m^{-3}$, indicating that the emission of solvent source was enhanced on these days. The dispersion effect of meteorological conditions on precursors can conceal the real information of emission sources and misjudge the formation process of O_3.

4 Conclusion

In this paper, a RF model was established based on the hourly data of 4 years of continuous observation, and some meteorological effects on the concentration time series of air pollutants were "removed". Transport and dispersion effects were removed for O_3, and the dispersion effect was removed for its precursors. In the process of building the RF model,

Figure 9. The contributions of emission sources to OFP based on observation and meteorological normalization during the pollution periods in the warm season in 2022.

UVB, RH, and T were found to be the most important factors affecting O_3 concentration, with the importance of 25.9 %, 18.2 %, and 11.3 %, respectively. Local influences, including precursor emissions and secondary photochemical reactions, occupied 83.9 % of the importance to O_3 concentration. To understand the mechanisms of local O_3 formation, the meteorological effects were analyzed in long-term trends, pollution and non-pollution periods in the warm season, as well as short-term pollution events. After decoupling meteorological effects, the concentration trends of O_3 were consistent with those observed in the long term, indicating that O_3 concentration was mainly driven by precursor emissions and local chemical reactions. During the pollution periods in the warm season from 2019 to 2022, the positive contributions of dispersion and transport to the MDA8 O_3 ranged from 12.9 % to 24.0 %. The effects of dispersion and transport were further analyzed for different types of O_3 pollution events. For transmission-type O_3 pollution (Period 1), dispersion and transport contributed 18.7 % to the MDA8 O_3 concentration, increasing the mean MDA8 O_3 concentration from 157 to 193 µg m^{-3}. For local and transmission-type O_3 pollution (Period 2), the average locally generated MDA8 O_3 concentration was 185 µg m^{-3}. Under the influences of dispersion and transport, the average MDA8 O_3 concentration increased to 211 µg m^{-3}, and the positive contributions of dispersion and transport ranged from 8.5 % to 20.4 %. BLH, as a parameter of dispersion, was of the highest importance for NO_2 and NMHCs, accounting for 34.2 % to NO_2 and 30.7 % to NMHCs. Therefore, precursor concentrations were accumulated even in the case of low emissions when the dispersion condition was poor, promoting the photochemical production of O_3. This also corresponds to the fact that even with the implementation of precursor emission reduction policies, O_3 concentrations in urban areas remain persistently high.

By decoupling the influences of meteorological conditions, it was observed that the sensitivity of local O_3 formation and the apportionment of VOC emission sources have changed. From the EKMA of short-term pollution events, the sensitivity of O_3 formation in Period 2 shifted towards the transition regime between VOC- and NO_x-limited regimes after meteorological normalization. Based on the PMF model, the changes in VOC emission sources after removing the dispersion effect during the warm season in 2022 were further analyzed. After removing the effect of dispersion, the absolute contribution of vehicle exhaust to VOCs during the pollution was 9.76 ppbv, accounting for 41.8 %, and the contribution of vehicle exhaust to OFP was 47.25 µg m^{-3}. The contribution of vehicle exhaust to VOCs was underestimated due the dispersion effect. After meteorological normalization, combustion remained an important source of VOCs, contributing 25.1 % during the pollution period, which was overestimated by 8.6 %. The normalized contribution of solvent use to VOCs decreased to 11.9 %, but it is undeniable that solvent use was still a crucial contributor to OFP, contributing 54.77 µg m^{-3}. Neglecting the influences of meteorology can lead to a deviation in emission reduction priorities, and the effectiveness of emission reduction may be masked by unfavorable meteorological conditions. The conclusion of this research suggested that meteorological fluctuations can interfere with the results of the OBM and PMF. Decoupling meteorological effects before using traditional models was beneficial for deepening the understanding of local O_3 formation and improving the rationality of precursor emission reduction measures.

Author contributions. XL, WC, and YZ conceived and designed this study and revised the article critically. YQ and XL analyzed and interpreted data, drafted the article, and revised the article critically. YL and MS contributed to the modeling of the data. XT, QZ, WL, WZ, and JL acquired the field observation data.

References

Ahmad, W., Coeur, C., Tomas, A., Fagniez, T., Brubach, J.-B., and Cuisset, A.: Infrared spectroscopy of secondary organic aerosol precursors and investigation of the hygroscopicity of SOA formed from the OH reaction with guaiacol and syringol, Appl. Opt., 56, E116–E122, https://doi.org/10.1364/ao.56.00e116, 2017.

Atkinson, R.: Atmospheric chemistry of VOCs and NO_x, Atmos. Environ., 34, 2063–2101, https://doi.org/10.1016/s1352-2310(99)00460-4, 2000.

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D., and de Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris, J. Geophys. Res.-Atmos., 118, 2041–2057, https://doi.org/10.1002/jgrd.50059, 2013.

Cai, C., Geng, F., Tie, X., Yu, Q., and An, J.: Characteristics and source apportionment of VOCs measured in Shanghai, China, Atmos. Environ., 44, 5005–5014, https://doi.org/10.1016/j.atmosenv.2010.07.059, 2010.

Carslaw, D. C. and Taylor, P. J.: Analysis of air pollution data at a mixed source location using boosted regression trees, Atmos. Environ., 43, 3563–3570, https://doi.org/10.1016/j.atmosenv.2009.04.001, 2009.

Carter, W. P. L.: Development of the SAPRC-07 chemical mechanism, Atmos. Environ., 44, 5324–5335, https://doi.org/10.1016/j.atmosenv.2010.01.026, 2010.

Chinese State Council: Action Plan on Air Pollution Prevention and Control, Chinese State Council, https://www.gov.cn/gongbao/content/2013/content_2496394.htm (last access: 28 March 2024), 2013 (in Chinese).

Chinese State Council: Three-Year Action Plan on Defending the Blue Sky, Chinese State Council, http://www.gov.cn/zhengce/content/2018-07/03/content_5303158.htm (last access: 28 March 2024), 2018 (in Chinese).

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 III, 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, https://doi.org/10.1016/s0140-6736(17)30505-6, 2017.

Dai, Q., Liu, B., Bi, X., Wu, J., Liang, D., Zhang, Y., Feng, Y., and Hopke, P. K.: Dispersion Normalized PMF Provides Insights into the Significant Changes in Source Contributions to $PM_{2.5}$ after the COVID-19 Outbreak, Environ. Sci. Technol., 54, 9917–9927, https://doi.org/10.1021/acs.est.0c02776, 2020.

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, 2023a.

Dai, T., Dai, Q., Ding, J., Liu, B., Bi, X., Wu, J., Zhang, Y., and Feng, Y.: Measuring the Emission Changes and Meteorological Dependence of Source-Specific BC Aerosol Using Factor Analysis Coupled With Machine Learning, J. Geophys. Res.-Atmos., 128, e2023JD038696, https://doi.org/10.1029/2023jd038696, 2023b.

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, https://doi.org/10.1016/j.scitotenv.2020.142394, 2021.

de Gouw, J. A., Warneke, C., Parrish, D. D., Holloway, J. S., Trainer, M., and Fehsenfeld, F. C.: Emission sources and ocean uptake of acetonitrile (CH_3CN) in the atmosphere, J. Geophys. Res.-Atmos., 108, D114329, https://doi.org/10.1029/2002jd002897, 2003.

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.

Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G., and Kumar, N.: Recommendations on statistics and benchmarks to assess photochemical model performance, J. Air Waste Manag. Assoc., 67, 582–598, https://doi.org/10.1080/10962247.2016.1265027, 2017.

Feng, R., Zheng, H.-J., Zhang, A.-R., Huang, C., Gao, H., and Ma, Y.-C.: Unveiling tropospheric ozone by the traditional atmospheric model and machine learning, and their comparison: A case study in hangzhou, China, Environ. Pollut., 252, 366–378, https://doi.org/10.1016/j.envpol.2019.05.101, 2019.

Fu, Y., Liao, H., and Yang, Y.: Interannual and Decadal Changes in Tropospheric Ozone in China and the Associated Chemistry-Climate Interactions: A Review, Adv. Atmos. Sci., 36, 975–993, https://doi.org/10.1007/s00376-019-8216-9, 2019.

Gao, D., Xie, M., Liu, J., Wang, T., Ma, C., Bai, H., Chen, X., Li, M., Zhuang, B., and Li, S.: Ozone variability induced by synoptic weather patterns in warm seasons of 2014–2018 over the Yangtze River Delta region, China, Atmos. Chem. Phys., 21, 5847–5864, https://doi.org/10.5194/acp-21-5847-2021, 2021.

Goliff, W. S., Stockwell, W. R., and Lawson, C. V.: The regional atmospheric chemistry mechanism, version 2, Atmos. Environ., 68, 174–185, https://doi.org/10.1016/j.atmosenv.2012.11.038, 2013.

Gong, D., Wang, H., Zhang, S., Wang, Y., Liu, S. C., Guo, H., Shao, M., He, C., Chen, D., He, L., Zhou, L., Morawska, L., Zhang, Y., and Wang, B.: Low-level summertime isoprene observed at a forested mountaintop site in southern China: implications for strong regional atmospheric oxidative capacity, Atmos. Chem. Phys., 18, 14417–14432, https://doi.org/10.5194/acp-18-14417-2018, 2018.

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.

Han, H., Liu, J., Yuan, H., Wang, T., Zhuang, B., and Zhang, X.: Foreign influences on tropospheric ozone over East Asia through global atmospheric transport, Atmos. Chem. Phys., 19, 12495–12514, https://doi.org/10.5194/acp-19-12495-2019, 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, https://doi.org/10.5194/acp-20-203-2020, 2020.

Henneman, L. R. F., Liu, C., Hu, Y., Mulholland, J. A., and Russell, A. G.: Air quality modeling for accountability research: Operational, dynamic, and diagnostic evaluation, Atmos. Environ., 166, 551–565, https://doi.org/10.1016/j.atmosenv.2017.07.049, 2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hu, C., Kang, P., Jaffe, D. A., Li, C., Zhang, X., Wu, K., and Zhou, M.: Understanding the impact of meteorology on ozone in 334 cities of China, Atmos. Environ., 248, 118211, https://doi.org/10.1016/j.atmosenv.2021.118221, 2021.

Huang, X., Huang, J., Ren, C., Wang, J., Wang, H., Wang, J., Yu, H., Chen, J., Gao, J., and Ding, A.: Chemical Boundary Layer and Its Impact on Air Pollution in Northern China, Environ. Sci. Technol. Lett., 7, 826–832, https://doi.org/10.1021/acs.estlett.0c00755, 2020.

Kong, L., Song, M., Li, X., Liu, Y., Lu, S., Zeng, L., and Zhang, Y.: Analysis of China's PM$_{2.5}$ and ozone coordinated control strategy based on the observation data from 2015 to 2020, J. Environ. Sci., 138, 385–394, https://doi.org/10.1016/j.jes.2023.03.030, 2024.

Lei, Y., Wu, K., Zhang, X., Kang, P., Du, Y., Yang, F., Fan, J., and Hou, J.: Role of meteorology-driven regional transport on O$_3$ pollution over the Chengdu Plain, southwestern China, Atmos. Res., 285, 106619, https://doi.org/10.1016/j.atmosres.2023.106619, 2023.

Li, C., Zhu, Q., Jin, X., and Cohen, R. C.: Elucidating Contributions of Anthropogenic Volatile Organic Compounds and Particulate Matter to Ozone Trends over China, Environ. Sci. Technol., 56, 12906–12916, https://doi.org/10.1021/acs.est.2c03315, 2022.

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, L., Chen, C. H., Huang, C., Huang, H. Y., Zhang, G. F., Wang, Y. J., Wang, H. L., Lou, S. R., Qiao, L. P., Zhou, M., Chen, M. H., Chen, Y. R., Streets, D. G., Fu, J. S., and Jang, C. J.: Process analysis of regional ozone formation over the Yangtze River Delta, China using the Community Multi-scale Air Quality modeling system, Atmos. Chem. Phys., 12, 10971–10987, https://doi.org/10.5194/acp-12-10971-2012, 2012.

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., Shao, M., Fu, L., Lu, S., Zeng, L., and Tang, D.: Source profiles of volatile organic compounds (VOCs) measured in China: Part I, Atmos. Environ., 42, 6247–6260, https://doi.org/10.1016/j.atmosenv.2008.01.070, 2008.

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., 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, https://doi.org/10.1007/s40726-019-00118-3, 2019.

Miller, S. L., Anderson, M. J., Daly, E. P., and Milford, J. B.: Source apportionment of exposures to volatile organic compounds. I. Evaluation of receptor models using simulated exposure data, Atmos. Environ., 36, 3629–3641, https://doi.org/10.1016/s1352-2310(02)00279-0, 2002.

Ministry of Ecology and Environment (MEE): Revision of the Ambien air quality standards (GB 3095-2012), Ministry of Ecology and Environment, https://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/201808/t20180815_629602.html (last access: 28 March 2022), 2018 (in Chinese).

Mo, Z., Shao, M., Wang, W., Liu, Y., Wang, M., and Lu, S.: Evaluation of biogenic isoprene emissions and their contribution to ozone formation by ground-based measurements in Beijing, China, Sci. Total Environ., 627, 1485–1494, https://doi.org/10.1016/j.scitotenv.2018.01.336, 2018.

Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., and Williams, M. L.: Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer, Atmos. Chem. Phys., 15, 8889–8973, https://doi.org/10.5194/acp-15-8889-2015, 2015.

Mousavinezhad, S., Choi, Y., Pouyaei, A., Ghahremanloo, M., and Nelson, D. L.: A comprehensive investigation of surface ozone

pollution in China, 2015-2019: Separating the contributions from meteorology and precursor emissions, Atmos. Res., 257, 105599, https://doi.org/10.1016/j.atmosres.2021.105599, 2021.

Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111–126, https://doi.org/10.1002/env.3170050203, 1994.

Song, C., Liu, B., Cheng, K., Cole, M. A., Dai, Q., Elliott, R. J. R., and Shi, Z.: Attribution of Air Quality Benefits to Clean Winter Heating Policies in China: Combining Machine Learning with Causal Inference, Environ. Sci. Technol., 57, 17707–17717, https://doi.org/10.1021/acs.est.2c06800, 2023.

Song, M., Tan, Q., Feng, M., Qu, Y., Liu, X., An, J., and Zhang, Y.: Source Apportionment and Secondary Transformation of Atmospheric Nonmethane Hydrocarbons in Chengdu, Southwest China, J. Geophys. Res.-Atmos., 123, 9741–9763, https://doi.org/10.1029/2018jd028479, 2018.

Song, M., Li, X., Yang, S., Yu, X., Zhou, S., Yang, Y., Chen, S., Dong, H., Liao, K., Chen, Q., Lu, K., Zhang, N., Cao, J., Zeng, L., and Zhang, Y.: Spatiotemporal variation, sources, and secondary transformation potential of volatile organic compounds in Xi'an, China, Atmos. Chem. Phys., 21, 4939–4958, https://doi.org/10.5194/acp-21-4939-2021, 2021.

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., and Ngan, F.: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System, B. Am. Meteorol. Soc., 96, 2059–2077, https://doi.org/10.1175/bams-d-14-00110.1, 2015.

Sun, J., Wu, F., Hu, B., Tang, G., Zhang, J., and Wang, Y.: VOC characteristics, emissions and contributions to SOA formation during hazy episodes, Atmos. Environ., 141, 560–570, https://doi.org/10.1016/j.atmosenv.2016.06.060, 2016.

Tan, Z., Fuchs, H., Lu, K., Hofzumahaus, A., Bohn, B., Broch, S., Dong, H., Gomm, S., Häseler, R., He, L., Holland, F., Li, X., Liu, Y., Lu, S., Rohrer, F., Shao, M., Wang, B., Wang, M., Wu, Y., Zeng, L., Zhang, Y., Wahner, A., and Zhang, Y.: Radical chemistry at a rural site (Wangdu) in the North China Plain: observation and model calculations of OH, HO_2 and RO_2 radicals, Atmos. Chem. Phys., 17, 663–690, https://doi.org/10.5194/acp-17-663-2017, 2017.

Tan, Z., Lu, K., Jiang, M., Su, R., Dong, H., Zeng, L., Xie, S., Tan, Q., and Zhang, Y.: Exploring ozone pollution in Chengdu, southwestern China: A case study from radical chemistry to O_3-VOC-NO_x sensitivity, Sci. Total Environ., 636, 775–786, https://doi.org/10.1016/j.scitotenv.2018.04.286, 2018.

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, M., Shao, M., Chen, W., Lu, S., Liu, Y., Yuan, B., Zhang, Q., Zhang, Q., Chang, C.-C., Wang, B., Zeng, L., Hu, M., Yang, Y., and Li, Y.: Trends of non-methane hydrocarbons (NMHC) emissions in Beijing during 2002–2013, Atmos. Chem. Phys., 15, 1489–1502, https://doi.org/10.5194/acp-15-1489-2015, 2015.

Wang, Y., Jiang, S., Huang, L., Lu, G., Kasemsan, M., Yaluk, E. A., Liu, H., Liao, J., Bian, J., Zhang, K., Chen, H., and Li, L.: Differences between VOCs and NO_x transport contributions, their impacts on O_3, and implications for O_3 pollution mitigation based on CMAQ simulation over the

Yangtze River Delta, China, Sci. Total Environ., 872, 162118, https://doi.org/10.1016/j.scitotenv.2023.162118, 2023.

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.

Wolfe, G. M., Kaiser, J., Hanisco, T. F., Keutsch, F. N., de Gouw, J. A., Gilman, J. B., Graus, M., Hatch, C. D., Holloway, J., Horowitz, L. W., Lee, B. H., Lerner, B. M., Lopez-Hilfiker, F., Mao, J., Marvin, M. R., Peischl, J., Pollack, I. B., Roberts, J. M., Ryerson, T. B., Thornton, J. A., Veres, P. R., and Warneke, C.: Formaldehyde production from isoprene oxidation across NO_x regimes, Atmos. Chem. Phys., 16, 2597–2610, https://doi.org/10.5194/acp-16-2597-2016, 2016.

Wu, Y., Liu, B., Meng, H., Dai, Q., Shi, L., Song, S., Feng, Y., and Hopke, P. K.: Changes in source apportioned VOCs during high O_3 periods using initial VOC-concentration-dispersion normalized PMF, Sci. Total Environ., 896, 165182, https://doi.org/10.1016/j.scitotenv.2023.165182, 2023.

Yang, C., Dong, H., Chen, Y., Wang, Y., Fan, X., Tham, Y. J., Chen, G., Xu, L., Lin, Z., Li, M., Hong, Y., and Chen, J.: Machine Learning Reveals the Parameters Affecting the Gaseous Sulfuric Acid Distribution in a Coastal City: Model Construction and Interpretation, Environ. Sci. Technol. Lett., 10, 1045–1051, https://doi.org/10.1021/acs.estlett.3c00170, 2023.

Yang, J., Wen, Y., Wang, Y., Zhang, S., Pinto, J. P., Pennington, E. A., Wang, Z., Wu, Y., Sander, S. P., Jiang, J. H., Hao, J., Yung, Y. L., and Seinfeld, J. H.: From COVID-19 to future electrification: Assessing traffic impacts on air quality by a machine-learning model, P. Natl. Acad. Sci. USA, 118, e2102705118, https://doi.org/10.1073/pnas.2102705118, 2021a.

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, S., Li, X., Song, M., Liu, Y., Yu, X., Chen, S., Lu, S., Wang, W., Yang, Y., Zeng, L., and Zhang, Y.: Characteristics and sources of volatile organic compounds during pollution episodes and clean periods in the Beijing-Tianjin-Hebei region, Sci. Total Environ., 799, 149491, https://doi.org/10.1016/j.scitotenv.2021.149491, 2021b.

Yuan, B., Shao, M., Lu, S., and Wang, B.: Source profiles of volatile organic compounds associated with solvent use in Beijing, China, Atmos. Environ., 44, 1919–1926, https://doi.org/10.1016/j.atmosenv.2010.02.014, 2010.

Zhang, H., Wang, Y., Hu, J., Ying, Q., and Hu, X.-M.: Relationships between meteorological parameters and criteria air pollutants in three megacities in China, Environ. Res., 140, 242–254, https://doi.org/10.1016/j.envres.2015.04.004, 2015.

Zhang, K., Liu, Z., Zhang, X., Li, Q., Jensen, A., Tan, W., Huang, L., Wang, Y., de Gouw, J., and Li, L.: Insights into the significant increase in ozone during COVID-19 in a typical urban city of China, Atmos. Chem. Phys., 22, 4853–4866, https://doi.org/10.5194/acp-22-4853-2022, 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, https://doi.org/10.1016/j.scitotenv.2024.171295, 2024.

Zhang, Q., He, K., and Huo, H.: Cleaning China's air, Nature, 484, 161–162, https://doi.org/10.1038/484161a, 2012.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos. Chem. Phys., 18, 14095–14111, https://doi.org/10.5194/acp-18-14095-2018, 2018.