https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Measurement report: Characterizing O_3-NO_x-VOC sensitivity and O_3 formation in a heavily polluted central China megacity using multi-methods during 2019–2021 warm seasons

Shijie Yu [a], Hongyu Liu [a], Hui Wang [a]*, Fangcheng Su [b,c], Minghao Yuan [d], Ruiqin Zhang [b,c]**

a. Department of Environmental Engineering, Henan University of Science and Technology, Luoyang 471023, China

b. Institute of Environmental Sciences, Zhengzhou University, Zhengzhou 450001, China

c. School of Ecology and Environment, Zhengzhou University, Zhengzhou 450001, China

d. Environmental Protection Monitoring Center Station of Zhengzhou, Zhengzhou 450007, China

*Correspondence to: H. Wang, Department of Environmental Engineering, Henan University of Science and Technology, Luoyang, Henan, PR China, 471023

**Correspondence to: R. Zhang, Research Institute of Environmental Science, School of Ecology and Environment, Zhengzhou University, Zhengzhou, Henan, PR China, 450001

E-mail address: wanghui79@haust.edu.cn (H. Wang), rqzhang@zzu.edu.cn (R. Zhang)

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Abstract

This study investigated the high ozone pollution in Zhengzhou City from 2019 to 2021 using observational data and model simulations, focusing on volatile organic compound (VOC) pollution and its impact on ozone formation. Using online VOC data and statistical analyses, the results showed that VOC concentration increased with ozone pollution level, with average values of 84.7±51.0, 96.6±53.4 and 105.3±59.4 μg/m³ for non-pollution, mildly polluted and moderately polluted periods, respectively. Source apportionment of ozone and its precursor VOCs was performed using CMAQ and PMF models. The results demonstrated that reducing vehicle emissions should be prioritized to mitigate ozone pollution in Zhengzhou, as transportation emissions accounted for 64% and 31% of ozone and VOC emissions, respectively. In addition, local ozone production rates and HOx base budgets were calculated using an observation-based model (OBM). The ozone production rates on non-pollution, mildly polluted, and moderately polluted days were respectively 2.0, 4.5, and 6.9 ppbv/h on average. The HOx radical concentration on polluted days was 1.5-6.4 times higher than that on non-pollution days, which is indicative of more efficient radical cycling during photochemical pollution. The O3-NOx-VOC sensitivity was analyzed using the OBM model, CMAQ model and ratio method. The results showed that ozone generation in Zhengzhou was mainly limited by VOCs, suggesting that the reduction of VOCs should be focused on aromatic hydrocarbons and olefins. The optimal reduction ratio of anthropogenic VOCs to NOx was about 2.9:1. This study will offer deeper insights for formulating effective ozone pollution prevention and control strategies.

Keywords: Volatile organic compounds; the Observation-based model; the Community multiscale air quality; the Source apportionment; the O_3–NO_x–VOCs sensitivity.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

1. Introduction

Tropospheric ozone (O_3) is a secondary pollutant generated from precursors such as volatile organic compounds (VOCs) and nitrogen oxides (NOx) through complex photochemical reactions driven by solar radiation (Meng et al., 2023;Ye et al., 2023). Its strong oxidizing properties can trigger multi-dimensional ecological health risks such as human respiratory damage, oxidative stress in plant leaves, and enhanced greenhouse effects in the troposphere (Jia et al., 2024; Kittipornkul et al., 2023; Zhao et al., 2023). The distinct nonlinear characteristics of ozone generation have posed significant challenges to the analysis of relevant pollution causes and the formulation of corresponding prevention and control strategies (Wang et al., 2023). However, it remains unclear about the contribution of motor vehicle exhaust, industrial emissions, and other anthropogenic sources of pollution to ozone generation and their driving mechanisms in the context of accelerated urbanization. Therefore, there is an urgent demand for systematic investigation to resolve the key issue in the field of atmospheric environment.

Recent research has advanced our understanding of VOCs, key precursors to ozone. As revealed by global monitoring data, there are substantial geographical differences in the concentration and composition of VOCs. For example, alkanes dominate accounts for over 80% of VOCs in Colorado, USA; while oxygenated VOCs (OVOCs), particularly acetone and acetic acid, are predominant in Athens, Greece (Abeleira et al., 2017; Kaltsonoudis et al., 2016). In China, VOCs pollution shows complex spatial and temporal patterns, with concentrations ranging from 27.0 ppbv in Nanjing to 92.0 ppbv in Tianjin during the summer (An et al., 2017; Han et al., 2015). Besides, alkanes take a sizeable proportion in urban regions like the North China Plain and Yangtze River Delta, though aromatic hydrocarbons and OVOCs exhibit higher photochemical activity in some areas (Mozaffar et al., 2020). Some regional differences have been highlighted in VOCs source analysis. Specifically, in Paris, VOCs largely come from liquefied petroleum gas (LPG) and solvents, while in Punjab, India, biomass burning is the major contributor to VOCs (Baudic et al., 2016; Pallavi et al., 2019). In the North China Plain, fossil fuel combustion, solvent use, and

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

LPG account for significant VOCs emissions; while in the Yangtze River Delta, the petrochemical industry is a major source of VOCs emissions (Mozaffar et al., 2020). Despite extensive research on VOCs sources, most studies focus on short-term pollution events and lack continuous tracking of VOCs component dynamics across varying ozone pollution levels. Furthermore, there is a lack of coordinated traceability research on ozone and its precursors. Traditional studies are commonly conducted based on single models, which may result in uncertainty in emission reduction strategies and hinder precise ozone pollution control.

Sensitivity studies of O_3 and its precursors are essential for developing effective control policies (Liao et al., 2024; Liu et al., 2021); however, existing methods have notable limitations. For example, the empirical kinetic modeling approach (EKMA, a model that establishes the relationship between reaction rates and concentrations based on experimental data) and air quality models simplify photochemical mechanisms, which may induce bias-related risks (Liu et al., 2021). Besides, uncertainties in emission inventories and aerosol-ozone interactions pose challenges to air quality models (Sharma et al., 2017). Additionally, observational methods like photochemical indicators lack universality across regions with differing photochemical conditions (Sillman, 1995). Moreover, the fluctuations in emission sources during events like the COVID-19 epidemic and the complexity of aerosol-ozone interactions further highlight the demand for more comprehensive approaches (Liu et al., 2020). Facing these methodological dilemmas, ozone formation mechanisms vary significantly in different regions. Ozone formation is typically categorized into NOx-controlled, VOCs-controlled, and transition zones (Sillman, 2021). Urban areas, such as New York (Tran et al., 2023), London (Tudor, 2022), Mexico City (Santiago et al., 2024), Beijing (He et al., 2022), Shanghai (Liu et al., 2021), Guangzhou (Hong et al., 2022), and Chengdu (Tan et al., 2018), are predominantly classified into VOCs-controlled zones. Contrarily, remote rural areas, such as Okinawa, Japan (Martin et al., 2004), Taian in Shandong Province (Li et al., 2024), and Wangdu in Hebei Province (Ran et al., 2011), are primarily classified into Nox-controlled zones. Temporal variations also add complexity, with seasonal and

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

diurnal shifts in O_3–NOx–VOCs sensitivity (Liu et al., 2010; Pan et al., 2015). Given these spatial and temporal variations, a multi-method, data fusion approach is required for better capturing the dynamics of O_3 generation, thus providing more policy guidance for pollution control.

To address these scientific challenges, Zhengzhou, a megacity in central China, is selected in this study. As of 2023, there are over 13 million residents and 5 million registered vehicles in the metropolitan area of Zhengzhou, facing severe haze and photochemical pollution from intensive anthropogenic emissions (Wang et al., 2021; Zhang et al., 2023). Although the Central Plains urban cluster centered on Zhengzhou has achieved significant progress in controlling primary pollutants, the paradoxical situation of persistently rising ozone levels still exists in this area, necessitating urgent investigation into its formation mechanisms (Jia et al., 2024; Li et al., 2020; Min et al., 2022; Yu et al., 2021). In this study, based on VOCs monitoring data (2019 - 2021), some advanced modeling approaches including the Ozone Box Model (OBM) and Community Multiscale Air Quality (CMAQ) model are employed to: (1) analyze the pollution characteristics of atmospheric VOCs in Zhengzhou and clarify the differences in their temporal distributions; (2) identify the sources of O3 and VOCs; (3) investigate the local O3 photochemical generation and removal pathways. Furthermore, the O3 isoconcentration curve in Zhengzhou is plotted to elucidate the relationship between O3 generation, VOCs, and NOX. On that basis, targeted O3 pollution control measures are proposed. The findings of this study can help Zhengzhou and other similar cities formulate targeted O_3 pollution control measures and also contribute to regional and global atmospheric environment research and pollution control theory.

2. Observation and methodology

2.1 Monitoring stations and instruments

Zhengzhou, the capital of Henan Province, is located in central China. The region faces severe air pollution, with its air quality ranking among the bottom twenty of major cities in China from January to September 2024. Although significant progress

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

has been made in air pollution control, resulting in a noticeable reduction in pollutant concentrations, the region still faces the dual pressure of winter smog and summer ozone pollution, making it a prominent high-pollution area (Wang et al., 2021; Zhang et al., 2023; Yu et al., 2022).

The municipal environmental monitoring station (MEM; 113°36′E, 34°45′N) was selected as the study site to obtain real-time online data, covering May to September during the period from 2019 to 2021 (Fig. S1). Located on the roof of a four-story building at the municipal environmental monitoring station, the area of the sampling site is predominantly commercial and residential, with no significant industrial sources nearby. The station is situated 300 m west of Qinling Road and 200 m south of Zhongyuan Road, both of which experience heavy traffic. Thus, mobile sources may significantly contribute to the VOCs concentration of the site. The MEM station is part of the air monitoring network operated by the Zhengzhou Environmental Monitoring Center. Simultaneously observed meteorological parameters include temperature, relative humidity, atmospheric pressure, wind direction, wind speed, and trace gases, such as O_3, NO, and NO_X.

Here, VOCs data with a temporal resolution of 1 h were collected using a Wuhan Tianhong online monitoring system (TH-300B), comprising two main modules (a cryogenic preconcentration system and gas chromatography/mass spectrometry (GC/MS) system). The cryogenic preconcentration unit employs an electronic refrigeration technique, achieving an extreme internal temperature of −150°C in the cold trap, effectively capturing the target compounds. This low-temperature empty tube trapping method is advantageous over traditional techniques because it minimizes the disadvantages of adsorbent adsorption, reduces VOCs loss, and enhances data accuracy. Teflon tubes were used to prevent chemical interference from adsorbents. Prior to air sample collection, a water removal device was used to eliminate excess moisture, preventing VOCs loss during low-temperature preconcentration. A particulate removal device was installed at the inlet of the sampling tube to filter out airborne particulate matter.

The complete workflow of the monitoring system includes sample collection, freeze trapping, thermal desorption, GC-flame ionization detector (FID)/MS analysis, and heating. To ensure the accuracy of the data obtained during the observation, rigorous quality assurance and quality control measures were implemented (Wang et al., 2022). Prior to analytical testing, the GC-FID/MS system was periodically calibrated using

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

an external standard gas across five concentration gradients (0.8, 1.6, 2.4, 4.0, and 8.0 ppbv), generating five-point calibration curves for each analyte. Four internal standard gases (bromochloromethane, 1,4-dichlorobenzene, chlorobenzene, and fluorobromobenzene) were used to ensure instrument stability. The instrument quantified a total of 108 VOCs species, including PAMs and TO-15 (Linde Spectra Environment Gases, USA). The linear correlation coefficients (R^2) for the VOCs measured using the instrument exceeded 0.99, and the method detection limits (MDLs) were in the range of 0.003–0.121 ppbv (Yu et al., 2021; Wang et al., 2022). Notably, 90% of the target compounds exhibited a quantification accuracy within 25%, and the measurement precision, as indicated by the relative standard deviation of the peak area, was maintained below 5%.

2.2 Analytical model for O_3 formation mechanism

2.2.1 Relative incremental reactivity

Here, we employed the OBM, commonly used in transformation studies of atmospheric VOCs to investigate the formation of O_3, free radicals, and intermediates (Niu et al., 2024; Zhou et al., 2024). The atmospheric chemistry framework employed here was based on the Master Chemical Mechanism (MCM) v3.3.1 (http://mcm.leeds.ac.uk/mcm/), which describes the degradation processes of methane and 142 non-methane VOCs, encompassing over 17000 reactions and 5800 substances and radicals (Chen et al., 2023; Fu et al., 2024).

The model inputs include 61 VOCs and 8 oxygenated VOCs (including acrolein, acetone, 2-butanone, 4-methyl-2-pentanone, 2-hexanone, 2-propanol, ethyl acetate, methyl tert-butyl ether (MTBE), and 1,4-dioxane), along with inorganic trace gases (NO_x, SO_2, and CO) and relevant meteorological factors (temperature, barometric pressure, and relative humidity) to constrain the model. Owing to the unavailability of measured photolysis rate parameters (*j* values), we simulated these parameters using the Tropospheric Ultraviolet and Visible (TUV) model (TUVv5.2, http://cprm.acom.ucar.edu/models/TUV), which is widely employed for such applications. The simulation period was set to 05:00–19:00 within the observation time frame.

The relative importance ratios (RIRs) were computed using the OBM to assess the relationship between O_3 precursors (Chai et al., 2023; Hu et al., 2023), as follows:

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

$$RIR^S(X) = \frac{\left[P_{O_3}^S(X) - P_{O_3}^S(X - \Delta X)\right] / P_{O_3}^S(X)}{\Delta S(X) / S(X)} \quad (1)$$

where X is a specific precursor, S(X) is the actual concentration of substance X, and ΔS(X) is the theoretical change in S(X). $P_{O3}^S(X)$ and $P_{O3}^S(X-\Delta X)$ refer to the simulated O_3 yields based on varying the concentration of species X in baseline and theoretical scenarios, respectively.

Net O_3 production was simulated based on the OBM model. The net O_3 production rate (P_{O3}^S) is the difference between the gross O_3 production rate (G_{O3}^S) and destruction rate (D_{O3}^S).

$$P_{O_3}^S = G_{O_3}^S - D_{O_3}^S . \quad (2)$$

G_{O3}^S was calculated by accumulating the oxidation rates of NO by HO_2 and RO_2.

$$G_{O_3}^S = k_{HO_2+NO}[HO_2][NO] + \sum k_{RO_{2i}+NO}[RO_{2i}][NO] \quad (3)$$

In addition, D_{O3}^S was calculated based on O_3 photolysis, reactions between HO_2 and olefins, and reactions between O_3 and OH and between NO_2 and OH.

$$D_{O_3}^S = k_{O(^1D)+H_2O}\left[O(^1D)\right][H_2O] + k_{OH+NO_2}[OH][NO_2] + k_{OH+O_3}[OH][O_3] + k_{HO_2+O_3}[HO_2][O_3] + k_{alkenes+O_3}[alkenes][O_3] \quad (4)$$

The values of the intermediates and radicals were obtained from the output of the OBM model. Constants *k* in Eqs. (3) and (4) are the rate coefficients for the matching reactions, respectively.

2.2.2 Empirical kinetic modeling approach

The empirical kinetic modeling approach (EKMA) was developed based on OBM calculations, commonly employed to assess the sensitivities of O_3 to NO_x and VOCs (Liang et al., 2023; Liu et al., 2022). This approach was employed to characterize the nonlinear relationship between O_3 and its precursors. Considering that the mixing ratio of VOCs does not accurately reflect the amount of O_3 produced, the VOCs concentration was substituted with the total OH reactivity of anthropogenic hydrocarbons in generating the EKMA curves. The reactivity was calculated using the hydroxyl radical (OH) reaction constants corresponding to the concentrations of VOCs and NO_x in the model. By varying the concentrations and reactivities of the

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

VOCs and NO_x, the precursors were identified as a function of the O_3 production rate P_{O3}^S, leading to the generation of the EKMA curves. The net O_3 production rate P_{O3}^S, O_3 generation rate G_{O3}^S, and O_3 depletion rate D_{O3}^S were calculated using Eqs. (2)–(4), respectively.

2.2.3 Decoupled direct method

The decoupled direct method (DDM) is simulated using the WRF/CMAQ model, with detailed configuration information provided in the papers published by our research group (Su et al., 2021). Additionally, the WRF/CMAQ setup is summarized in Table S1. In brief, the WRF/CMAQ model was configured with four nested domains: 36 km for East Asia, 12 km for central and eastern China, 4 km for Henan Province, and 1 km for Zhengzhou (Fig S2). WRF provided meteorological inputs for CMAQ, using 6-h FNL global reanalysis data for initial and boundary conditions. The CMAQ model utilized the SAPRC-99 gas-phase photochemical mechanism and AERO6 aerosol module, with modified heterogeneous chemistry for SO2 to sulfate and NO2 to nitrate conversion (Hu et al., 2014). The 36-km simulation used clean continental IC/BC, and nested domain IC/BC were derived from the parent domains. The first five days of CMAQ output were discarded to minimize IC influence. Anthropogenic emissions for China were based on the 2016 MEIC inventory (0.25° × 0.25° resolution), while emissions from other regions used the REAS2 inventory. Emissions from Henan (4-km domain) were based on local data from Bai et al. (2020). Biogenic emissions for all domains were generated using MEGAN version 2.10, with windblown dust emissions generated online in CMAQ simulations.

The decoupled direct method (DDM) can be employed to analyze the sensitivity of O_3 to its precursors. By directly solving the sensitivity equations of the air quality model, various sensitivity coefficients can be obtained, enabling comprehensive sensitivity analyses. The primary objective of this study was to calculate the semi-normalized first- and second-order sensitivities of O_3 concentration with respect to anthropogenic VOCs and NO_x emissions.

$$S_V = \frac{\partial C_{O_3}}{\partial V}. \tag{5}$$

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

$$S_N = \frac{\partial C_{O_3}}{\partial N}. \tag{6}$$

$$S_{VV} = \frac{\partial^2 C_{O_3}}{\partial V^2}. \tag{7}$$

$$S_{NN} = \frac{\partial^2 C_{O_3}}{\partial N^2}. \tag{8}$$

$$S_{VN} = \frac{\partial^2 C_{O_3}}{\partial V \partial N}. \tag{9}$$

In the model, V, N, and O_3 represent VOCs, NO_x, and O_3, respectively. C_{O3} is the O_3 concentration, whereas ε_V and ε_N are the relative perturbations in total anthropogenic VOCs and NO_x emissions from sources in Henan Province. The first-order sensitivities of the O_3 concentration to NO_x and VOCs emissions are denoted by S_N and S_V, respectively. S_{NN} and S_{VV} denote the second-order sensitivities.

For model calibration using CMAQ-DDM, the pollutants considered for validation included routinely observed meteorological parameters and pollutant concentration data. Table S2 shows that the simulated $PM_{2.5}$ concentrations were slightly overestimated, with an overall overestimation of approximately 20%, closely aligning with the actual air quality conditions. The simulation results for O_3 revealed a better performance than those for $PM_{2.5}$, achieving an overall correlation (R) of 0.74, which met the requirements for targeted research on O_3 during the study period. However, the O_3 concentrations were slightly underestimated, with an overall underestimation of 17%. On the contrary, the simulation of nitrogen dioxide is generally more accurate, as NO_2 is directly emitted, leading to better model performance in simulating NO_2.. This consistent underestimation of O_3 suggests that VOCs emission sources may be underestimated to an extent, although NO_2 simulations are largely accurate.

Overall, the simulation results effectively reproduced the spatial and temporal distribution characteristics of air pollution in Henan Province and Zhengzhou City, providing a solid foundation for research (Su et al., 2021).

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

2.3 Pollutant source attribution

2.3.1 O_3 source apportionment

Traceability analysis is employed to uniquely label or add tracer ions to single emission substances from different regions or industries within the emission source inventory (Xian et al., 2024; Zhang et al., 2023). Thus, our CMAQ (Source Oriented CMAQ v5.3.2) model, which includes a source apportionment function, independently computed the scientific processes affecting pollutants marked with unique industry or regional labels (Su et al., 2023). By analyzing the concentration results of these labeled species, we determined the contribution of each pollutant.

In the regional traceability simulation, we quantified the contributions of various emission sources in Zhengzhou through industry traceability (as shown in Fig. S2). Specifically, we simplified the emission sources (originally categorized into 16 types) into seven categories based on a refined 1-km inventory of Zhengzhou City. Thus, the contributions of local emissions to O_3 formation from different industries were determined, facilitating more accurate industry and regional control.

2.3.2 VOCs source apportionment

The Positive Matrix Factorization (PMF) model (version 5.0) developed by the U.S. Environmental Protection Agency was used for the source apportionment of VOCs (Farhat et al., 2024; Frischmon et al., 2024). PMF is a multivariate factor analysis tool that decomposes measurement data into source profile and source contribution matrices. As shown in Eq. (10), the concentration of species can be determined using the contribution of the source to the target source and the species distribution of each source:

$$X_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \,. \tag{10}$$

X_{ij} is the concentration of the jth substance measured in the ith sample, g_{ik} is the contribution of the kth source to the ith sample, f_{kj} is the proportion of the jth substance in the kth source, and e_{ij} is the residual amount of the jth substance in the jth sample.

The results obtained using Eq. 10 typically present uncertainty (UNC) because of the error fraction of the species concentration and MDL. PMF analysis relies on the objective function (Q) to minimize residuals and uncertainties.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\frac{x_{ij} - \sum_{k=1}^{p} g_{ik} f_{kj}}{u_{ij}} \right]^2 \tag{11}$$

where n and m are the numbers of species and samples, respectively, and u_{ij} is the UNC of the jth species in the i-th sample. Q (true) is the fitting parameter calculated when all the data are included, whereas Q (robust) is calculated when the model excludes inappropriate data. Q can be used to select the best mathematical result.

The calculation method for UNC related to the PMF model is as follows:

If the sample concentration is less than or equal to the MDL, the UNC is calculated using Eq. (12):

$$unc = \frac{5}{6} \times MDL\ . \tag{12}$$

If the concentration exceeds the MDL, UNC is calculated using Eq. (13):

$$unc = \sqrt{\left(ErrorFration \times Conc.\right)^2 + \left(0.5 \times MDL\right)^2} \tag{13}$$

where unc refers to the UNC of species in the sample, MDL refers to the minimum limit of the detection method, and ErrorFraction refers to the sample error (typically 10%–20%, set to 10% here).

The species selection of PMF is performed based on the following principles: (1) prioritizing VOC species with higher concentrations for effective signal separation; (2) selecting emission tracers like isoprene to accurately trace pollution sources; (3) excluding species with over 25% missing data or concentrations below the method detection limit (MDL) to avoid bias; (4) excluding species with a signal-to-noise ratio (S/N) below 0.5 to ensure adequate signal strength.The final selection comprised 37 VOCs species. The total concentrations of these VOCs accounted for 82% of the overall VOCs concentration, indicating that they effectively represented the main VOCs situation. In this study, a six-factor solution was chosen in the PMF analysis based on two parameters (Ulbrichetal., 2009): (1) Qtrue/Qrobust values and (2) Qtrue/Qtheoretical values (Fig. S3). As shown in Table S3, the stability assessment of the six-factor solution of PMF by Bootstrap resampling (BS) and parametric displacement test (DISP) shows that the solution has high statistical robustness.

2.4 Machine Learning and SHAP Analysis

To accurately capture the complex nonlinear relationship among meteorological

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

variables, pollutant concentrations, and O_3 levels, two mainstream ensemble learning algorithms, XGBoost and Random Forest (RF), were selected in this study to construct a prediction model. The strong robustness and high prediction accuracy of these two methods in dealing with nonlinear regression problems make them an ideal choice for environmental pollutant concentration prediction (Fan et al., 2025).

XGBoost, as a scalable tree-based boosting algorithm, has the core advantage of optimizing the differentiable loss function through gradient descent and introducing regularization term to effectively control the complexity of the model, significantly reducing the risk of over fitting and enhancing the generalization ability of the model. The algorithm can also automatically process missing values, support parallel computing and efficiently prune trees, especially for structured table data processing in this study (Li et al., 2025). With its strong ability to deal with the complex correlation between ozone concentration and multiple characteristics such as temperature, humidity, nitrogen dioxide (NO_2), and fine particulate matter ($PM_{2.5}$), XGBoost has become the main modeling framework of this study with its excellent performance and computational efficiency (Tabo et al., 2025).

RF is constructed using the bagging method, which introduces sample randomness through bootstrap sampling. Moreover, it incorporates a feature selection mechanism during the node splitting process of each decision tree to reduce inter-tree correlation effectively, enhancing model stability and robustness against interference. In regression tasks, the model consolidates predictions from multiple decision trees to generate the final output, striking a balance between predictive performance and practicality (Wang et al., 2024).

Both models are implemented using tools from the Python ecosystem: the XGBoost model is built using the XGBoost library, and the RF model is developed with the scikit-learn library. To achieve optimal performance, a grid search strategy combined with cross-validation is used to systematically optimize the models' hyperparameters. By traversing predefined parameter combinations and assessing the models' generalization ability, the optimal parameter configuration is ultimately identified to ensure the reliability of the prediction results (Khan et al., 2025).

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

2.4.2 Interpretability Analysis Using SHAP

To overcome the "black box" limitation of machine learning, it is necessary to quantitatively calculate and understand the importance of the role of each influence factor on the prediction of ozone concentration, as well as its influencing direction and size (Liu et al., 2024). Therefore, SHAP (Shapley Additive exPlanations) value analysis method was used for interpretation of machine learning model in this study. This method uses Shapley value theorem of cooperative game theory, to judge which factors or variable values play important and fair roles both globally (the sum of all samples' predicting results) and individually (every sample's single prediction), making up the deficiencies of the traditional parameter analysis in dealing with complex nonlinear relationship between variables (Takefuji 2025).

In the specific analysis process, SHAP values are first calculated for all observed samples. By taking the absolute values of the SHAP values and averaging them, the global importance ranking of meteorological elements, pollutants, and other features affecting ozone concentration is determined (Zhang et al., 2024). Meanwhile, based on the sign of the SHAP values (positive or negative), the direction of each feature's influence on ozone concentration is clarified: positive values indicate that the feature promotes ozone concentration increase, while negative values suggest an inhibitory effect (Yao et al., 2024). This analytical framework provides data-driven support for identifying key driving factors behind ozone concentration changes and elucidating their mechanisms of action.

2.4.3 Model Performance Evaluation

To comprehensively evaluate model fitting and generalization performance, three standard regression metrics were used: the coefficient of determination (R^2), mean squared error (MSE), and mean absolute error (MAE). Both XGBoost and RF models were assessed on training and testing sets (Liu et al., 2024).

In terms of overall performance metrics (Table S4), the XGBoost model significantly outperforms the RF model: Regarding fitting capability, XGBoost achieves a higher training set R^2 (0.93) than RF (0.91), along with lower MSE (0.07) and MAE (0.20), indicating more precise capture of data patterns and reduced training error. In terms of

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

generalization capability, the XGBoost model achieved a higher test set R^2 (0.86) than the random forest model (0.84), along with lower MSE (0.13) and MAE (0.27), demonstrating superior prediction stability on new data. Furthermore, the XGBoost model exhibits a smaller gap between training and test metrics, indicating more effective control over overfitting. In summary, the XGBoost model demonstrates superior performance across fitting accuracy, generalization capability, and error control (Yan et al., 2025).

Beyond its superior simulation performance, XGBoost showed slight advantages in scalability, computational speed, and ability to capture complex feature interactions. Therefore, the XGBoost model was selected as the core method for subsequent factor identification and SHAP-based interpretability analysis (Takefuji et al., 2025).

3. Results and discussion

3.1 General characteristics

3.1.1 Levels of air pollutants and meteorological parameters

The results indicate severe photochemical pollution in the period from May to September during the years 2019-2021 (in Fig. 1). According to the GB 3095-2012 standard, the maximum daily average 8 h (MDA8) O_3 concentrations exceeding 160 160 μg/m^3 and 160 μg/m^3 are categorized as light pollution and moderate pollution, respectively. The proportion of days with O_3 concentrations exceeding 160 μg/m^3 was as high as 45%, with days classified as moderate or higher pollution accounting for 7%. The MDA8 was recorded on June 6, 2021 (285 μg/m^3), with a severe pollution level. The proportion of MDA8 O_3 concentrations exceeding the standard during sampling periods from 2019 to 2021 was 53%, 37%, and 36%, showing a downward trend but still indicating severe O_3 pollution.

O_3 concentrations are significantly influenced by meteorological factors. As shown in Fig. S3, Pearson correlation coefficients between O_3 concentrations and various meteorological factors reveal that O_3 concentrations are positively correlated with temperature and wind speed, with correlation coefficients of 0.66 and 0.43 ($p < 0.01$), respectively. It is negatively correlated with relative humidity ($p < 0.01$), with a

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

correlation coefficient of −0.23. The generation of ozone is intricately tied to the emissions of its precursor gases. As illustrated in Fig. S3, O_3 concentrations show significant correlations with its precursors—VOCs, NO, and NO_2—with correlation coefficients (r) of -0.28, -0.30, and -0.57, respectively. In contrast, $PM_{2.5}$ demonstrates notable positive correlations with these same three precursors, exhibiting r values of 0.36, 0.17, and 0.38, respectively. This positive relationship between $PM_{2.5}$ and its precursor emissions suggests that regulating these emissions could effectively mitigate particulate matter concentrations (Shao et al., 2024).

Since the Pearson correlation coefficient primarily measures linear relationships, it is sensitive to outliers and cannot detect non-linear associations. To address this limitation, Spearman and Kendall statistical methods were also employed in this study to assess potential non-linear correlations (see Tables S5 and S6), and the results were consistent with those obtained using the Pearson correlation coefficient. Furthermore, when analyzing the relationship between O_3 and different VOCs components, a significant positive correlation was found between O_3 and OVOCs ($r = 0.22$, $p < 0.01$), suggesting that OVOCs concentrations in Zhengzhou are influenced by secondary formation processes. In contrast, significant negative correlations were observed with other VOCs components, with the strongest correlation found between O_3 and reactive aromatic hydrocarbons ($r = 0.32$, $p < 0.01$).

In summary, the formation of ozone is highly complex and closely related to the concentration of precursor substances as well as meteorological factors. Unlike $PM_{2.5}$ pollution, reducing precursor emissions may not necessarily alleviate photochemical pollution, which presents a greater challenge for control efforts and highlights the complexity of managing ozone pollution (Wang et al., 2024). Therefore, this paper will provide a comprehensive analysis of the formation mechanisms and sensitivity of O_3-NOx-VOCs.

3.1.2 Quantitative evaluation of influencing factors via machine learning and SHAP analysis

To move beyond the limitations of traditional correlation analysis, which only measures the strength and direction of linear associations, this study used XGBoost to

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

model the nonlinear relationships between O_3 concentrations and meteorological factors (e.g., temperature, humidity, wind speed) as well as pollutants (e.g., NO_x, VOC species, $PM_{2.5}$) (Razavi et al., 2024). SHAP values were employed to quantify the contribution and directional influence of each factor (Zhang et al., 2024). The XGBoost model demonstrated strong performance, with R^2 values of 0.93 on the training set and 0.86 on the test set, indicating high predictive accuracy and generalization ability.

As illustrated in the corresponding Fig. 2a, meteorological factors and pollutants contributed nearly equally to ozone formation, accounting for 52% and 48% of the total influence, respectively. This suggests that meteorological conditions play a dominant role in ozone pollution episodes. Specifically, temperature was the most influential factor (21.93%), primarily promoting photochemical reactions through enhanced solar radiation and radical activity. Humidity (17.74%) generally inhibited ozone formation by suppressing radical generation and affecting boundary layer structure. Wind speed and direction contributed 5.01% and 4.71%, respectively, by regulating pollutant transport and dispersion.

Among pollutants, NO_x was the most significant precursor, with a total contribution of 24.77%. NO_2 accounted for 17.57%, largely due to its photolysis being a direct source of ozone. NO contributed 7.20%; although it can titrate ozone, its net effect during the study period was promotive via oxidation to NO_2. The total contribution of VOCs was 13.04%, with substantial variation among species: highly reactive aromatics (4.10%) and alkenes (1.88%) strongly promoted ozone formation by facilitating NO-to-NO_2 conversion. Low-reactivity species such as alkanes and alkynes each contributed less than 1%. OVOCs, halogenated hydrocarbons, and sulfides showed limited influence due to their secondary reaction pathways. $PM_{2.5}$ contributed 7.35%, indirectly promoting ozone formation through heterogeneous reactions on particle surfaces that enhance NO conversion and through radiative warming effects.

SHAP analysis further revealed complex nonlinear behaviors (Fig. 2b): high temperature and low humidity generally promoted ozone formation. NO_2 exhibited concentration-dependent effects, promoting ozone at low levels but inhibiting it at

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

high concentrations due to titration. $PM_{2.5}$ also showed a promotive effect at elevated concentrations. Highly reactive VOCs consistently correlated with increased ozone.

By integrating machine learning, this study overcomes the constraints of traditional statistics, enabling quantitative identification of key drivers and elucidating the mechanisms of ozone formation under multi-factor synergy and antagonism (Garbagna et al., 2025). The findings emphasize the need for coordinated control of NO_x and highly reactive VOCs, as well as integrated management of $PM_{2.5}$ and ozone, providing a scientific basis for targeted air pollution policies.

3.1.3 O_3 vs. non-O_3 episode days

Meteorological conditions are a significant factor in O_3 pollution. High temperatures, low relative humidity, and weak winds facilitate photochemical pollution (Xu et al., 2023). Apart from meteorological factors, O_3 precursors significantly affect O_3 formation. Table 1 shows that NO_x concentrations during pollution periods significantly exceeded those during non-pollution periods, and the concentration increased as pollution intensified. PM_{10} exhibited a pattern similar to that of NOx, suggesting that during photochemical pollution periods, the emission intensity of atmospheric pollutants is relatively high (Saqer et al., 2024). Daily average $PM_{2.5}$ concentrations on ozone-polluted days exceeded those on non-polluted days. During the sampling period, the average $PM_{2.5}$ concentration on moderate pollution days was 35.5 ± 16.4 μg/m³, closed to the annual average ambient air quality standard (GB 3095-2012) Grade II standard of 35 μg/m³ and 1.3 times higher than that during non-pollution periods. NO concentration decreased with increasing pollution severity. On moderately polluted days, the average NO was 1.3 μg/m³, which was lower than 7% of the levels observed on non-polluted days. This may be attributed to the fact that, under O_3-polluted conditions, atmospheric oxidation capacity could potentially be enhanced, with possibly elevated concentrations of both O_3 and free radicals that might accelerate NO consumption. (Shao et al., 2024).

The average VOCs concentrations for non-pollution, light pollution, and moderate pollution periods were 84.7 ± 51.0, 96.6 ± 53.4 and 105.3 ± 59.4 μg/m³, respectively.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Considering the numerous VOCs types and sources, the top 20 substances for the three stages were analyzed. Table 2 and Fig. S4 show that during the observation period, higher concentrations of small-molecule hydrocarbons, such as ethane and propane, suggest a significant influence of LPG/natural gas (NG) at the monitoring site (Derwent et al., 2017). The acetylene and 1,2-dichloroethane concentrations increased as pollution intensified, indicating a substantial impact from combustion, particularly during photochemical pollution (Zuo et al., 2024). The concentrations of C4–C5 alkanes and benzene series compounds were high, suggesting an association with vehicle emissions (Han et al., 2024). Furthermore, on moderate pollution days, vehicle tracer substances were more concentrated. The concentrations of n-hexane, dichloromethane, trichloromethane, tetrachloroethylene, and ethyl acetate were high, indicating emissions from solvent use. During pollution periods, the isoprene, 2-butanone, and 2-hexanone concentrations exceeded that during non-pollution periods, indicating a significant impact of plants and more photochemical secondary products during high O_3 periods.

Table S7 presents a comparison between the concentration characteristics of VOCs in this study and those reported in domestic and international literature. The concentration of VOCs in Zhengzhou (90.3 ± 52.8 μg/m³; 2019-2021) is higher than that of most international urban regions, aligning with other regions in China, such as Beijing (101.5 ± 65.2 μg/m³; 2016) (Liu et al., 2021). In contrast, cities like Istanbul (40-60 μg/m³), Athens (50.3 μg/m³; 2016-2017), and Vitória, Brazil (24.1 ± 29.6 μg/m³; 2022-2024) show significantly lower levels in the concentration of VOCs, reflecting regional disparities in emission control and industrial development (Thera et al., 2019; Panopoulou et al., 2021; Galvão et al., 2025). The VOCs in Zhengzhou are mainly composed of dichloromethane, acetone, ethane, isopentane, and n-hexane, indicating mixed sources from solvents and vehicular emissions, different from other cities where industrial and traffic emissions are more specialized. This finding suggests that there is an urgent need for targeted abatement efforts to reduce the concentration of VOCs and alleviate ozone pollution in a rapidly growing city like Zhengzhou.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

3.1.4 Diurnal variation

Fig. S5 illustrates the O_3 diurnal cycle divided into the suppression phase (P1) of O_3 by midnight and early morning NO emissions, photochemical generation phase (P2) of O_3, and titration phase (P3) of O_3 by precursor substances prior to the evening peak (Du et al., 2024). To reflect the photochemical processes, we examined the ratios of compounds with different reaction rates of K_{OH} radicals but similar sources. Fig. S5 shows the relationship between ethylbenzene and xylene, revealing their homogeneity ($R^2 > 0.9$). Ethylbenzene has a lifetime of 3 days, whereas the lifetime of xylene is 1 day. During the observation period, the diurnal variation trends of O_3 and the ethylbenzene/xylene ratio were similar. The strong correlation between the age indicator (the ratio of both compounds) and O_3 provides strong evidence linking O_3 with the photochemical processes of non-methane hydrocarbons (Hui et al., 2019).

The mean O_3 concentrations on non-polluted, lightly polluted, and moderately polluted midnights are 82.9±50.3, 121.4±78.2, and 149.4±80.7 μg/m³, respectively. As shown in Fig. S5, on polluted days, high fresh NO emissions at midnight significantly reduced the O_3 levels. The O_3 concentration decreased to a minimum in the early morning because of the high fresh NO emissions during the morning rush hour. Contrarily, on non-polluted days, the NO concentration was lower at night, leading to a weaker titration effect. Thus, the O_3 concentration remained relatively stable at midnight and decreased to its minimum with the morning peak. Fig. 3 demonstrate the daily variation trends of characteristic VOCs. In stage P1, the concentrations of NO, NO_2, and n-pentane increase on polluted days, whereas these pollutants remain relatively stable on non-polluted days. Furthermore, benzene series compounds (toluene, ethylbenzene, and meta/para-xylene) exhibit similar patterns. Thus, we can infer that on polluted days, nighttime emissions are significantly influenced by motor vehicle emissions (Song et al., 2023).

Stage P2 is the accumulation phase of O_3 (08:00–16:00). Photoreactions generate many peroxy radicals (such as HO_2 and RO_2), converting sufficient NO into NO_2 (Fittschen, 2019). With increasing solar radiation, a large amount of NO_2 is photolyzed to generate O_3, causing peak O_3 levels in this phase. The increase in O_3

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

during polluted periods significantly exceeded that during non-polluted periods. For example, at 08:00 in non-polluted periods, the O_3 concentration was 47.2 μg/m³, increasing to 59.3 μg/m³ the next moment. During lightly and moderately polluted periods, the O_3 concentration ranges were 56.4–81.5 and 70.8–104.2 μg/m³, respectively. This is because higher NO concentrations at night during polluted periods contributed to the formation of more OH and peroxy radicals. Furthermore, O_3 began to accumulate with increasing light intensity. Owing to the higher radical content, the photochemical reaction activity was strong, and the high concentration of peroxy radicals further oxidized NO to NO_2, leading to higher O_3 generation efficiency during polluted periods. The minimum value of O_3 during moderately polluted periods was observed at 07:00, and although the NO concentration remained high (13.6 μg/m³), the O_3 concentration had already accumulated and rapidly increased at 09:00. This indicated that the radical concentration was high, rendering it advantageous in competing with NO for O_3, leading to an increase in the O_3 concentration. As the NO titration weakened and photochemical activity increased, O_3 rapidly accumulated. Throughout the P2 stage, the concentrations of O_3 precursors decreased because of the consumption in photochemical reactions. As shown in Fig. 3, the concentrations of certain VOCs (particularly benzene series compounds) and NO were lower on polluted days than on non-polluted days in this stage. It should be noted that isoprene peaks around noon, owing to temperature- and light-dependent emission rates. For the sensitivity of daily variation patterns, the ratio gradually increased in P2, suggesting that the increase in the surrounding biogenic VOCs (BVOCs) may shift the O_3 generation mechanism from VOCs to transition zones at noon (Chen et al., 2022).

With the arrival of the evening peak and a gradual decrease in light intensity, the O_3 sources were less than the sinks, resulting in a continuous decrease in O_3 concentration in P3. The NO concentration during P3 was significantly lower than that during P1. Owing to the effective NO titration, the O_3 concentration during P3 on polluted days exceeded that on non-polluted days. For example, at 20:00 on a moderately polluted day, the average O_3 concentration remained high at 176.3 μg/m³,

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

decreasing to 131.5 μg/m³ by 23:00. This resulted in stronger atmospheric oxidation at midnight and higher radical reaction activity (An et al., 2024). This influenced the O_3 generation the next day, contributing to consecutive O_3 pollution days.

3.2 Source apportionment of O_3 and VOCs

3.2.1 Source contributions to O_3

Fig. 4 shows the spatial distribution of anthropogenic O_3 emissions in Zhengzhou. As shown in Fig. 4a, local anthropogenic emissions contribute significantly to O_3 levels in Zhengzhou, with the overall MDA8-O_3 (8-hour maximum daily average ozone concentration) exceeding 28 ppbv (55 μg/m³), and concentrations exceeding 40 ppbv observed in urban areas. Fig. 4b shows the anthropogenic contributions in other regions of Henan Province, revealing that areas surrounding Zhengzhou also have relatively high O_3 concentrations, reflecting the regional nature of ozone pollution. This phenomenon suggests that controlling ozone pollution is challenging and requires cross-regional mitigation measures.

To identify the main sources of the significant increase in MDA8-O_3, we categorized the total contributions into six sectors: industries, solvents, transportation, electricity, residential areas, and other sources (Fig. 4c–h). The results indicate that transportation is the largest contributor to O_3 formation in Zhengzhou, with the highest concentrations exceeding 30 ppbv in the eastern part of the city. This area is characterized by a dense road network, including several national expressways, which suggests a close relationship between transportation emissions and high ozone levels (Gu et al., 2019). Industrial emissions are the second largest source, with high ozone concentrations primarily found in the northern and northwestern parts of Zhengzhou, aligning with the city's industrial layout. The power sector also contributes to ozone formation in Zhengzhou, with a concentration peak observed in the southwestern part of the city.

Fig. S6 shows the contributions of various sources to O_3 formation during both polluted and non-polluted periods of the observation. Transportation accounted for 63.6% of total contributions, but this decreased to 57.4% on polluted days. Traffic

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

sources are the largest contributors to O_3, consistent with previous research findings (Cheng et al., 2019; Su et al., 2023). Industrial sources contributed 30.4% on average and 26.4% on polluted days, indicating a slight reduction during pollution events. Electricity contribution significantly increased on polluted days, reaching 3.3 times the average. Simulation results suggest that more aggressive control measures are required in the transportation and industrial sectors during the summer. Attention should also be given to the electricity due to its increased emissions on polluted days.

3.2.2. Source apportionment of ambient VOCs

Fig. 5 shows the source apportionment factor spectrum for high O_3 pollution periods. Six factors were identified.

Factor 1 is characterized by dominant species, including acetylene (63%), chloromethane (25%), benzene (15%), and certain lower-carbon hydrocarbons (isobutane, n-pentane, ethylene, propylene, and trans-2-pentene, etc.). Acetylene, ethylene, and chloromethane are important indicators of fossil fuel and biomass combustion (Liu et al., 2008; Wu et al., 2016). Fixed combustion sources are major sources of C2–C3 lower-carbon alkanes and benzene (Li et al., 2024); thus, Factor 1 is identified as a combustion source.

Factor 2 mainly comprises C2–C5 alkanes, including ethane (56%), propane (48%), n-butane (37%), isobutane (31%), n-pentane (15%), and isopentane (19%). These substances are tracers for fuel evaporation (gasoline and LPG/NG) (Zhang et al., 2019). Pentane is one of the most abundant VOCs species associated with gasoline evaporation (Zhang et al., 2019)., and butane has been reported as a tracer for LPG (Liu et al., 2008; Shen et al., 2018). Furthermore, the aromatic content in this source is extremely low; thus, this source is identified as LPG/NG.

Factor 3 is characterized by high MTBE levels (54%), small-molecule hydrocarbons (C2–C6), and benzene series compounds. C2–C6 alkanes, alkenes, and benzene series compounds are typical tracers of motor vehicle exhaust (Xiao et al., 2023; Wu et al., 2023). MTBE is commonly used as an additive in gasoline, which improves the octane rating, enhances engine performance, and reduces exhaust emissions, making

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

it a tracer for motor vehicle exhaust (Schifter et al., 2017). Thus, Factor 3 is determined to be a motor vehicle emission source.

Factor 4 is characterized by high C6–C8 alkane levels, such as n-hexane (68%), 3-methylpentane (29%), 2-methylpentane (30%), and n-heptane (26%). This factor features high levels of acetone (59%), dichloromethane (45%), chloroform (31%), carbon tetrachloride (69%), and benzene series compounds. Studies have shown that benzene series compounds commonly originate from solvent-use emissions (Zhou et al. 2019, Wang et al. 2021). Carbon tetrachloride, n-hexane, and dichloromethane are commonly used chemical reagents. The content of highly volatile small-molecule hydrocarbons in Factor 4 is low. Thus, Factor 4 is identified as a solvent source.

Factor 5 is characterized by pollutants, mainly comprising small-molecule hydrocarbons (propane, butane, ethylene, and propylene, etc.), BTEX (VOCs group including benzene, toluene, ethylbenzene, and xylene), carbon disulfide, and halogenated hydrocarbons (carbon tetrachloride, dichloromethane, chloromethane, and 1,2-dichloroethane, etc.). These substances are widely used in manufacturing, furniture, shoe, and rubber industries (Hui et al., 2018; Yu et al., 2021); thus, Factor 5 is identified as an industrial source.

Factor 6 has the highest proportion of isoprene (89%), which is a marker of plant emissions (Cheun et al., 2014; Khruengsai et al., 2024;); thus, Factor 6 is identified as a plant source.

Fig. 6 shows the source apportionment of VOCs at different O_3 pollution levels from June to September. The results indicate that the motor vehicle emissions gradually increased as O_3 pollution intensified. On moderately polluted days, this source accounted for up to 35%; thus, during high O_3 pollution periods, it is crucial to enhance the control of motor vehicle emissions. The proportion of combustion sources is significantly higher on moderately polluted days compared with other periods. Therefore, combustion sources must be controlled during O_3 pollution periods. On moderately polluted days, the proportion of plant sources is relatively high, closely related to high temperatures and strong radiation. During the observation period, the proportion of solvent-use sources was high but decreased with increasing

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

pollution. It is speculated that solvent emissions, which include highly active aromatic hydrocarbons, are consumed more because of the high photochemical activity during O_3 pollution periods. Summarily, during high O_3 pollution periods, attention should be focused on controlling motor vehicle, solvent use, and combustion sources.

Table S8 presents a comparison of source apportionment between Zhengzhou and other cities. During the observation period, the main sources of VOCs in Zhengzhou comprise vehicle emissions (31%), solvent use (24%), and industrial processes (21%), collectively accounting for 76% of the total pollution, highlighting the dominant role of traffic and industrial pollution. In contrast, cities like Paris and Turkey have significantly lower proportions of vehicle emissions (15% and 15.8%, respectively) (Baudic et al., 2016; Thera et al., 2019). The proportion of solvent use in Zhengzhou (21%) is similar to that in the cities of the Yangtze River Delta but higher than that in Turkey and other regions (Zhang et al., 2025; Thera et al., 2019). The severity of biogenic pollution in Zhengzhou is lower, a phenomenon particularly evident in cities with richer vegetation cover, such as Athens (Kaltsonoudis et al., 2016). The source apportionment structure in Zhengzhou reflects its typical characteristics as an industrial city, with significant pressure from traffic emissions, as well as notable contributions from solvent use and industrial processes, while biogenic sources contribute relatively little.

3.3 Differences in photochemical reactivity

3.3.1 In situ net O_3 production

As shown in Fig. 7, the mean of net $P(O_3)$ during daytime (05:00–19:00) in Zhengzhou City during the observation period was 3.1 ppbv/h. This was lower than that of Beijing (5.8 ppbv/h) (Xue et al., 2014), Wuhan (6.2 ppbv/h) (Lu et al., 2017), and Taishan (4.2 ± 0.9 ppbv/h) (Kanaya et al., 2009), and higher than that of Shanghai (Liu et al., 2021) (2.8 ± 0.7 ppbv/h), etc. Fig. 7 shows the daily $P(O_3)$ trends for different pollution periods. The average net $P(O_3)$ during the daytime averages were 2.0 (non-polluted), 4.5 (mildly polluted), and 6.9 ppbv/h (moderately and highly polluted), which were converted into the O_3 daytime chemical accumulation, with

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

values of 30, 67, and 103 ppbv, respectively. To elucidate the local O_3 photochemical generation and removal pathways, the source–sink pathways and their corresponding shares in the O_3 generation process were investigated.

As shown in Table 3, the mean F(O_3) in Zhengzhou was 3.8 ppbv/h, mainly controlled by three pathways, with contributions of 84% (RO_2+NO), 16% (HO_2+NO), and <1% (MO_2+NO), respectively. The mean O_3 removal rate D(O_3) was 0.7 ppbv/h; the contribution of OH+NO_2 and O_3+alkenes contributed the most to the D(O_3) with 56% and 33%, respectively. The distribution of O_3 generation and removal pathways in different periods and years were statistically analyzed, and the results showed that RO_2+NO and OH+NO_2 dominated local O_3 generation and removal, respectively. This shows that atmospheric free radicals play a key role in localized O_3 generation in Zhengzhou City (Wang et al., 2022). The next subsection describes the free radical chemistry for an in-depth investigation.

3.3.2 HO_x budget

The OH concentration was calculated using the OBM. The simulation results showed (Fig. S6) that the mean daily peak values of OH and HO_2 radicals in Zhengzhou were 5.6×10^6 and 3.8×10^8 molecule/cm^3, respectively. A comparison of the results with those of Tan et al. (2019) showed that the OH concentration in Zhengzhou was slightly lower than those in Beijing and Shanghai and higher than those in Chongqing and Guangzhou, However, the concentration of HO_2 radicals was lower than that in Chongqing, consistent with the results of Beijing, Shanghai, and Guangzhou. Previous experiments based on comprehensive field observations in China have shown that OH concentrations may be underestimated under low NO_x conditions (Fuchs et al., 2017; Hofzumahaus et al., 2009). Here, the NO_x concentration was high; thus, the model reproduced the OH concentration relatively well (Rohrer et al., 2014). The prominent feature of the high NO_x state is the underestimation of HO_2 (Tan et al., 2017), which has been observed at urban sites outside China (Dusanter et al., 2009; Kanaya et al., 2007), partially explaining the low HO_2 radicals in Zhengzhou City.

Fig. 8 demonstrates the daily trends of HOx radicals under different O_3 pollution

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

conditions. The results showed that the HOx concentration significantly increased with an increase in photochemical pollution. The average daily peak concentration of OH radicals on non-polluted days was 4.2×10^6 molecule/cm^3, with peaks increasing 1.5 and 2.7 times under mild and moderate pollution conditions, respectively. For the HO_2 results, the peaks increased 3.6 and 6.4 times, respectively. The aforementioned phenomena indicate a more active radical cycle during high O_3 periods (Zhu et al., 2020), and the HO_x radical source–sink cycle was investigated.

HO_X radicals trigger VOCs oxidation and promote O_3 formation. Fig. S7 illustrates the formation and loss pathways of OH radicals during the observation period. For OH, the generation pathways are mainly HO_2 + NO and O_3 + VOCs, with 43% and 56%, respectively. The removal pathways are mainly based on OH + VOCs. Although the air pollution problems are visually extremely similar, the free radical chemistry, particularly the primary radical sources, significantly varies across different metropolitan areas. For example, Lanzhou had a higher OH contribution from O_3 + VOCs (32%) (Jia et al., 2018), whereas Wuhan had a higher contribution of HO_2 + NO (Zhu et al., 2020). O_3 photolysis is the main source of OH in Nashville (Martinez et al., 2003). Nitrous acid (HONO) photolysis plays a dominant role in New York City (Ren et al., 2003), Paris (Michoud et al., 2012), and Wangdu, China (Tan et al., 2017). Formaldehyde photolysis is an important source of OH in Milan (Alicke et al., 2002).

Fig. S7 shows the simulated average generation and loss rates of OH for the three periods. The OH formation or loss rate increased with increased photochemical pollution, implying a higher efficiency of free radical cycling during photochemical pollution. The situation of the source–sink pathways of OH in different pollution periods was similar to that in the observation period, and VOCs and NO_x significantly impacted the HO_x free radical cycling.

HONO is an important source of HO_X, playing a crucial role in atmospheric chemistry (Xue et al., 2014). Considering that we did not measure the HONO mixing ratio, the results may be underestimated. Thus, the current OBM-MCM results may have underestimated the daytime HO_X concentration to an extent, and supplemental HONO is required to better determine the HO_x balance.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

3.4 O_3–NO_x–VOCs sensitivity

3.4.1 VOCs/NO_x ratio

The influence of O_3 precursors on O_3 formation can be defined as the VOCs and NO_x control zones, which are critical in developing effective strategies for reducing regional O_3 pollution. The VOCs/NO_x ratio has been widely used to determine the state of O_3 formation. Generally, at a VOCs/NO_x ratio below 10 (ppbC/ppbv), a VOCs-sensitive zone is observed. However, when the ratio exceeds 20, it is in a NO_x-sensitive state. At a ratio between 10 and 20, the reduction of VOCs and NO_x can effectively reduce O_3 concentrations (Hanna et al., 1996; Sillman et al., 1999).

MEM was selected as the study site to investigate the VOCs/NO_x (ppbC/ppbv) during the high O_3 hours from 2019 to 2021. As shown in Table S9, the VOCs/NO_x at this site was 6.2 ± 7.1 during the observation period. The ratio increased with an increase in the photochemical pollution levels, i.e., 5.9 ± 7.3 for non-polluted days, 7.0 ± 6.6 for mildly polluted days, and 7.3 ± 6.7 for moderately and highly polluted days. In addition, the fraction of O_3-polluted days with VOCs/NO_x > 10 increased. The proportions of mildly and moderately/highly polluted days were 15% and 18%, respectively, indicating that the proportion of O_3 generation in Zhengzhou subject to the transition zone increased with increasing photochemical pollution (Zhu et al., 2020).

Fig. 9(b) shows the daily trends of the VOCs/NO_x ratios for different photochemical pollution periods. The ratios for the three periods exhibited similar daily variations. Higher ratios were observed at midnight (1:00–6:00), after which the ratios rapidly decreased, indicating that NO_x concentrations increased more rapidly than VOCs in terms of the effect of vehicle emissions (Gu et al., 2019). Thereafter, the VOCs/NO_x ratio increased with the O_3 concentrations. In the afternoon (12:00–16:00), at a high O_3 concentration, the VOCs/NO_x was high in all the periods (moderate and high pollution > light pollution > non-pollution). During moderate and high O_3 pollution, the ratio of VOCs/NO_x exceeded 10, characterized by the transition zone. synergistic emission reduction of VOCs and NO_x to effectively mitigate photochemical pollution.

The VOCs/NO_x ratios are only a preliminary determination of O_3 sensitivity and were

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

subsequently validated by the CMAQ model and OBM.

3.4.2 Relative importance ratio and empirical kinetic modeling approach using the box model

RIR is a key parameter for determining the relationship between O_3 and its precursors. Thus, RIR values are important for developing science-based O_3 pollution control strategies. Higher positive RIR values indicate that the precursors are more sensitive to O_3 production, whereas substances with negative RIR values play a negative role in O_3 formation (Niu et al., 2024; Zhang et al., 2024). Here, we quantified the RIR values of NO_x, CO, and different fractions of VOCs and further classified anthropogenic VOCs (AHCs) into aromatic hydrocarbons, olefins, and alkanes to better understand the effects of different sources on O_3.

The city monitoring station was selected as the target site. The acquired observation data for the high O_3 period (May–September) for 2019–2021 were used to determine the RIR in Zhengzhou (Fig. 9a). The results showed that the RIR of AHCs was larger during the observation period, indicating that anthropogenic sources significantly contribute to local O_3 generation, and the reduction of anthropogenic sources of VOCs can effectively mitigate local O_3 pollution. The contribution of BVOCs to local O_3 generation was high owing to the high reactivity of BVOCs and the higher emission intensity caused by the high temperature and strong radiation in May–September. The RIR of CO was low, indicating that the mitigation of O_3 pollution through CO reduction was ineffective. The negative RIR for NO_x indicated that reducing NO_x contrarily promoted O_3 production.

Fig. 9a illustrates the distribution of RIR on O_3 non-pollution, mild pollution, and moderate and heavy pollution days. RIR_{AHC} exhibited high values in the three periods; therefore, VOCs control must be strengthened in the region, particularly olefins and aromatic hydrocarbons during the O_3 pollution hours. The RIR values of BVOCs were high and tended to increase with an increase in pollution. The values of O_3 non-pollution, mild pollution, and moderate and heavy pollution days were 0.4, 0.5, and 0.7, respectively. The RIR_{NOx} values were negative on O_3 non-pollution and mild

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

pollution days, indicating that it was in the VOCs control zone at this time. The RIR_{NOx} value became positive (0.4) with an increase in pollution levels. Thus, the synergistic control of NO_x and VOCs effectively reduced the photochemical pollution on high O_3 days.

Owing to the use of the reactivity concept, EKMA can be employed as a standardized framework for investigating the sensitivity of regional O_3 production to VOCs and NO_x (Liu et al., 2023; Wang et al., 2022). Thus, based on the study period, when the photochemical pollution was more severe, the pollutant information and values of meteorological factors from Zhengzhou monitoring stations were inputted into the OBM. As shown in Fig. 10, the O_3 contours show the local maximum concentration of O_3 as a function of the initial NO_x and VOCs concentrations. The relationship between O_3 and its precursors was highly nonlinear. At low NO_x concentrations, the O_3 concentration increased almost linearly with increasing NO_x concentration. The increase in the O_3 concentration gradually slowed with an increase in the NO_x concentration, reaching a local peak. The line connecting the localized peaks in O_3 concentration is called a "ridge" (Fig. 10). The ridge divides the O_3 formation into two photochemical states. Below the ridge is the NO_x control zone, and the VOCs control zone is above the ridge.

Based on online data from the Zhengzhou monitoring station, an EKMA curve was plotted (Fig. 10), and the results were consistent with the RIR. Zhengzhou was in the VOCs control zone on O_3 non-pollution and mild pollution days. The local O_3 susceptibility was transformed into the transition zone as pollution increased. This indicates that the summer O_3 pollution in the urban area of Zhengzhou was mainly in the VOCs-sensitive zone, and reducing the VOCs concentration facilitated O_3 pollution control. As shown by the slope of the ridge in Fig. 10, the optimal reduction ratio of VOCs to nitrogen oxides is 2.9:1, and it is recommended not to be lower than 2:1.

3.4.3 Regional distribution of O_3 sensitivity based on CMAQ-DDM

Based on the DDM method, the sensitivity of O_3 to its precursors was assessed across

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

different regions of Henan Province (Su et al., 2023; Yu et al., 2021). The province was categorized into various sensitivity regions according to the ratios of O_3 precursor concentration sensitivities (Fig. 11). The 3-year simulation results revealed that the VOCs-sensitive region in Henan Province encompassed the Anyang-Zhengzhou area along the Taihang Mountains, including three cities north of the Yellow River, and areas south of the river, such as Zhengzhou, Xuchang, and parts of Luoyang, Kaifeng, and Luohe. This aligns with conventional knowledge, as these areas experience high NO_x emissions; thus, VOCs concentration is critical in O_3 generation (Su et al., 2023). In Zhengzhou City, subregional sensitivity analysis indicated that all areas fell within the VOCs control zone, suggesting that O_3 management during summer should prioritize VOCs control measures. A comparison of sensitivities in 2021 with those from the previous 2 years showed a northward shift in the VOCs-sensitive area. This was largely attributed to a significant reduction in anthropogenic emissions owing to heavy precipitation in Henan in July 2021, which led to a 25% decrease in anthropogenic emissions in the affected areas. Natural VOCs emissions from plants remained largely unaffected, resulting in a marked increase in the VOCs/NO_x mass ratio. Consequently, the NO_x control area in Henan Province shifted northward, and the VOCs-sensitive area decreased relative to July 2021. However, in Zhengzhou City, the reduction in anthropogenic emissions did not impact VOCs sensitivity. This is because natural VOCs emissions from plants were not the dominant factor, even in the context of significant reductions in anthropogenic emissions in northern Henan. Therefore, the alterationin the VOCs/NOx ratio resulting from the heavy rainfall was inadequate to alter O_3 sensitivity.

4 Summary and conclusions

The pollution characteristics of atmospheric VOCs in Zhengzhou City were analyzed using real-time VOCs data from May to September in the period of 2019-2021. The sources of atmospheric VOCs and O_3 were determined using PMF and CMAQ models. Factors affecting free radical equilibrium were investigated to highlight the major factors driving local ozone generation. The main conclusions are summarized

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

as follows:

1. Pollution characteristics and source distribution of VOCs and ozone

The study clarified the pollution characteristics and sources of ozone and its precursor VOCs in Zhengzhou City. The city suffered severe photochemical pollution during the observation period, with ozone concentration exceeding the standard value by 45% and average VOCs concentration of 90.3 ± 52.8 μg/m³. Moreover, VOCs concentration increased with the enhancement of ozone pollution. Their concentrations during the non-pollution, mildly polluted, and moderately polluted periods were 84.7 ± 51.0, 96.6 ± 53.4 and 105.3 ± 59.4 μg/m³, respectively. The PMF modeling yielded six factors contributing to VOCs emissions, namely motor vehicle exhaust, solvent use, industrial emissions, liquefied petroleum gas (LPG)/natural gas (NG), stationary combustion, and biogenic sources, among which motor vehicle exhaust was the largest source of VOCs. As ozone pollution became more intense, the contribution of motor vehicles to VOCs was 30%, 31% and 35%, respectively. Industrial emissions were the second largest source of VOCs, accounting for 21%. Ultimately, the source apportionment of ozone was performed based on the CMAQ model. The results showed that ozone formation in Zhengzhou was mainly attributed to local anthropogenic emissions, with motor vehicle exhaust and industrial emissions being the two largest sources.

2. Mechanism and sensitivity of ozone formation

The local ozone formation rate and its generation and removal pathways were revealed, and the O_3-NOx-VOCs sensitivity of Zhengzhou City was comprehensively evaluated. According to the OBM model analysis, the average $P(O_3)$ on non-pollution, mildly polluted, and moderately/highly polluted days was 2.0, 4.5, and 6.9 ppbv/h, respectively. The contribution of RO_2+NO to local ozone generation was more than 80%, indicating that atmospheric free radicals had a significant effect on local ozone formation. The HOx radical concentration increased 1.5-6.4 times on polluted days compared with non-pollution days. The results obtained by employing VOCs/NOx ratio, RIR, EKMA and DDM methods indicated that Zhengzhou was located in the control zone of VOCs and was shifting to the transition zone with the increase in the

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

intensity of ozone pollution. AHCs contributed greatly to local ozone formation. In particular, reducing aromatic hydrocarbons and olefins helped to effectively mitigate ozone pollution. The optimal reduction ratio of VOCs to NOx was determined to be 2.9:1, which is recommended not to be lower than 2:1.

3. Scientific contribution and policy implications

Three years of observational data were combined with advanced modeling techniques, such as the CMAQ and OBM model in this study to comprehensively explore ozone pollution dynamics in Zhengzhou City. The significant impact of vehicle emissions on ozone and its precursors is consistent with the results of other urban studies (Song et al., 2018; Yu et al., 2022), reinforcing the key role of controlling mobile sources in mitigating photochemical pollution. While PMF-based VOCs source apportionment is widely used in urban studies (Farhat et al., 2024; Frischmon et al., 2024), the inclusion of CMAQ in ozone source tracking provides new insights into the role of anthropogenic emissions in ozone formation.

In addition, this study provided a comprehensive assessment of ozone sensitivity using multiple diagnostic methods. The findings confirmed that Zhengzhou City was located in the VOCs control zone, consistent with the results of other urban studies (He et al., 2022; Santiago et al., 2024; Tran et al., 2023; Tudor, 2022). However, notably, different diagnostic methods used for analyzing ozone formation sensitivity have their inherent advantages and limitations. To accurately determine the O_3-NOx-VOCs sensitivity, high-resolution observations were combined with multiple methods to ensure reliable, scientifically sound conclusions that help to identify key factors controlling local ozone formation and provide actionable insights for mitigating photochemical pollution.

In this study, the ozone sensitivity was evaluated using multiple methods, resulting in a reliable understanding of the complex interactions between O_3, NOx, and VOCs, emphasizing the need for balanced emission control strategies. The proposed optimal VOCs/NOx reduction ratio was identified to be 2.9:1, which provides a practical framework for the effective control of ozone pollution. This approach addresses the limitations of single-pollutant abatement measures and ensures that emission control

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

policies are both scientific and practically feasible.

The results of this study are particularly relevant to rapidly urbanizing regions such as Zhengzhou, where industrialization and motorization are driving significant changes in air quality. By highlighting the importance of controlling transportation emissions and optimizing the VOCs/NOx emission reduction ratio, this study provides a solid scientific basis for air quality management. And this study also emphasizes the need for integrated emission control strategies that take into account the unique sources and interactions of pollutants in such environments. These insights are of guiding reference for the development of targeted policies to address photochemical pollution, ultimately contributing to the long-term improvement of air quality in rapidly growing urban areas.

4. Limitations and future research directions

This study has certain limitations. The accuracy of the OBM model depends on high-quality input data, and the lack of measured HONO data potentially lead to underestimated HOx radical concentration. Future studies are advised to incorporate measured HONO data and explore advanced techniques such as machine learning to improve data quality and reduce model uncertainty. In addition, long-term monitoring and modeling efforts are required to capture seasonal and inter-annual variations in ozone formation mechanism.

Author contributions

YSJ: Writing-original draft, Methodology, Data curation, Investigation, Visualization, Validation, Software, Formal analysis. LHY: Formal analysis, Data curation, Investigation. WH: Supervision, Resources, Project administration, Funding acquisition. SFC: Methodology,Software. YMH: Data curation, Resources. ZRQ: Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

Competing Interests

The contact author has declared that none of the authors has any competing interests.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Acknowledgments

This work was supported by Key Research and Development Program of Henan (No. 241111320300) and National Key Research and Development Program of China (No. 2017YFC0212403).

Data availability

All the data presented in this article can be accessed through https://doi.org/10.5281/zenodo.17214861 (Yu et al., 2025).

References

Alicke, B., Platt, U., and Stutz, J.: Impact of nitrous acid photolysis on the total hydroxyl radical budget during the Limitation of Oxidant Production/Pianura Padana Produzione di Ozono study in Milan, J. Geophys. Res-Atmos., 107(D22), LOP-9, 2002.

An, C., Li, H., Ji, Y. Y., Chu, W. H., Yan, X. Y., and Chai, F. H.: A review on nocturnal surface ozone enhancement: Characterization, formation causes, and atmospheric chemical effects, Sci. Total Environ., 921, 170731, 2024.

Barletta, B., Meinardi, S., Rowland, F. S., Chan, C. Y., Wang, X. M., Zou, S. C., Chan, L.Y., and Blake, D. R.: Volatile organic compounds in 43 Chinese cities, Atmos. Environ., 39(32), 5979-5990, 2005.

Chai, W. X., Wang, M., Li, J. Y., Tang, G. G., Zhang, G. H., and Chen, W. T.: Pollution characteristics, sources, and photochemical roles of ambient carbonyl compounds in summer of Beijing, China, Environ. Pollut., 336, 122403, 2023.

Chen, J., Liu, T., Gong, D. C., Li, J. Y., Chen, X., Li, Q.Q., Liao, T., Zhou, Y., Zhang, T., Wang, Y., Wang, H., and Wang, B. G.: Insight into decreased ozone formation across the Chinese National Day Holidays at a regional background site in the Pearl River Delta, Atmos. Environ., 315, 120142, 2023.

Chen, W. H., Guenther, A. B., Jia, S. G., Mao, J. Y., Yan, F. H., Wang, X. M., and Shao, M.: Synergistic effects of biogenic volatile organic compounds and soil nitric oxide emissions on summertime ozone formation in China, Sci. Total Environ., 828, 154218, 2022.

Cheng, J., Zhang, Y. S., Wang, T., Xu, H., Norris, P., and Pan, W. P.: Emission of

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

volatile organic compounds (VOCs) during coal combustion at different heating rates, Fuel, 225, 554-562, 2018.

Cheung, K., Guo, H., Ou, J. M., Simpson, I. J., Barletta, B., Meinardi, S., and Blake, D. R.: Diurnal profiles of isoprene, methacrolein and methyl vinyl ketone at an urban site in Hong Kong, Atmos. Environ., 84, 323-331, 2014.

Derwent, R. G., Field, R. A., Dumitrean, P., Murrells, T. P., and Telling, S. P.: Origins and trends in ethane and propane in the United Kingdom from 1993 to 2012, Atmos. Environ., 156, 15-23, 2017.

Du, J. Y., Wang, X. Y., and Zhou, S. Y.: Dominant mechanism underlying the explosive growth of summer surface O_3 concentrations in the Beijing-Tianjin-Hebei Region, China, Atmos. Environ., 333, 120658, 2024.

Dusanter, S., Vimal, D., Stevens, P. S., Volkamer, R., Molina, L. T., Baker, A., Meinardi, S., Blake, D., Sheehy, P., Merten, A., Zhang, R., Zheng, J., Fortner, E. C., Junkermann, W., Dubey, M., Rahn, T., Eichinger, B., Lewandowski, P., Prueger, J., and Holder, H.: Measurements of OH and HO_2 concentrations during the MCMA-2006 field campaign–Part 2: Model comparison and radical budget, Atmos. Chem. Phys., 9(18), 6655-6675, 2009.

Fan, X. L., Gao, P., Wang, J. G., Zhang, M. L., Hong, S., Qin, S. Z., Zhang, Q., Ma, L. L., Zhang, L. F., Zhang, Z., and Lv, X.: Assessment and prediction of soil heavy metal pollution in cotton fields by multi-source data feature fusion combined with machine learning, Ind. Crops Prod., 233, 121447, 2025.

Fu, N., Cao, L. M., Xia, S. Y., Zeng, L. W., He, L., He, L. Y., and Huang, X. F.: Sensitivity of atmospheric peroxyacetyl nitrate (PAN) formation and its impact on ozone pollution in a coastal city, Atmos. Environ., 330, 120545, 2024.

Fuchs, H., Novelli, A., Rolletter, M., Hofzumahaus, A., Pfannerstill, E. Y., Kessel, S., Edtbauer, A., Williams, J., Michoud, V., Dusanter, S., Locoge, N., Zannoni, N., Gros, V., Truong, F., Sarda-Esteve, R., Cryer, D. R., Brumby, C. A., Whalley, L. K., Stone, D., Seakins, P. W., Heard, D. E., Schoemaecker, C., Blocquet, M., Coudert, S., Batut, S., Fittschen, C., Thames, A. B., Brune, W. H., Ernest, C., Harder, H., Muller, J. B. A., Elste, T., Kubistin, D., Andres, S., Bohn, B., Hohaus, T., Holland,

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

F., Li, X., Rohrer, F., Kiendler-Scharr, A., Tillmann, R., Wegener, R., Yu, Z.J., Zou, Q., and Wahner, A.: Comparison of OH reactivity measurements in the atmospheric simulation chamber SAPHIR, Atmos. Meas. Tech., 10(10), 4023-4053, 2017.

Garbagna, L., Saheer, L. B., and Oghaz, M. M. D.: AI-driven approaches for air pollution modeling: A comprehensive systematic review, Environ. Pollut., 373, 125937, 2025.

Gu, X. K., Yin, S. S., Lu, X., Zhang, H., Wang, L. L., Bai, L., Wang, C., Zhang, R. Q., and Yuan, M. H.: Recent development of a refined multiple air pollutant emission inventory of vehicles in the Central Plains of China, J. Environ. Sci., 84, 80-96, 2019.

Han, J., Liu, Z., Hu, B., Zhu, W., Tang, G., Liu, Q., Ji, D., and Wang, Y.: Observations and explicit modeling of summer and autumn ozone formation in urban Beijing: identification of key precursor species and sources, Atmos. Environ., 309, 119932, 2023.

Hanna, S. R., Moore, G. E., and Fernau, M. E.: Evaluation of photochemical grid models (UAM-IV, UAM-V, and the ROM/UAM-IV couple) using data from the Lake Michigan Ozone Study (LMOS), Atmos. Environ., 30(19), 3265-3279, 1996.

He, Z. M., Liu, P. F., Zhao, X. X., He, X. W., Liu, J. F., and Mu, Y. J.: Responses of surface O_3 and $PM_{2.5}$ trends to changes of anthropogenic emissions in summer over Beijing during 2014–2019: A study based on multiple linear regression and WRF-Chem, Sci. Total Environ., 807, 150792, 2022.

Hofzumahaus, A., Rohrer, F., Lu, K., Bohn, B., Brauers, T., Chang, C. C., Fuchs, H., Holland, F., Kita, K., Kondo, Y., Li, X., Lou, S. R., Shao, M., Zeng, L. M., Wahner, A., and Zhang, Y. H.: Amplified trace gas removal in the troposphere, Sci., 324(5935), 1702-1704, 2009.

Hong, Q. Q., Zhu, L. B., Xing, C. Z., Hu, Q. H., Lin, H., Zhang, C. X., Zhao, C. H., Liu, T., Su, W. J., and Liu, C.: Inferring vertical variability and diurnal evolution of O_3 formation sensitivity based on the vertical distribution of summertime HCHO and NO_2 in Guangzhou, China, Sci. Total Environ., 827, 154045, 2022.

Hu, B. Y., Chen, G. J., Chen, J. S., Xu, L. L., Fan, X. L., Hong, Y. W., Li, M. R., Lin, Z. Y., Huang, M. Q., Zhang, F. W., and Wang, H.: The effect of nitrous acid (HONO)

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

on ozone formation during pollution episodes in southeastern China: Results from model improvement and mechanism insights, Sci. Total Environ., 891, 164477, 2023.

Huang, Y. S., Hsieh, C. C.: Ambient volatile organic compound presence in the highly urbanized city: source apportionment and emission position, Atmos. Environ., 206, 45-59, 2019.

Hui, L. R., Liu, X. G., Tan, Q. W., Feng, M., An, J. L., Qu, Y., Zhang, Y. H., and Jiang, M. Q.: Characteristics, source apportionment and contribution of VOCs to ozone formation in Wuhan, Central China, Atmos. Environ., 192, 55-71, 2018.

Hui, L. R., Liu, X. G., Tan, Q. W., Feng, M., An, J. L., Qu, Y., Zhang, Y. H., and Cheng, N. L.: VOC characteristics, sources and contributions to SOA formation during haze events in Wuhan, Central China, Sci. Total Environ., 650, 2624-2639, 2019.

Jia, C. H., Mao, X. X., Huang, T., Liang, X. X., Wang, Y. N., Shen, Y. J., Jiang, W. Y. H., Wang, H. Q., Bai, Z. L., Ma, M. Q., Yu, Z. S., Ma, J. M., and Gao, H.:Non-methane hydrocarbons (NMHCs) and their contribution to ozone formation potential in a petrochemical industrialized city, Northwest China, Atmos. Res., 169, 225-236, 2016.

Jia, C. H., Wang, Y. N., Li, Y. J., Huang, T., Mao, X. X., Mo, J. Y., Li, J. X., Jiang, W. Y. H., Liang, X. X., Gao, H., and Ma, J. M.: Oxidative capacity and radical chemistry in a semi-arid and petrochemical-industrialized city, Northwest China, Aerosol Air Qual. Res., 18(6), 1391-1404, 2018.

Jia, J., Wang, J., Jin, W. Y., Yu, N. N., Gong, S. L., Ni, J. W., Zhang, X., and Zhou, L. L.: Inter-annual variability and health risk assessment of summer VOCs in a Plain City of China, Atmos. Environ., 337, 120790, 2024.

Kanaya, Y., Cao, R., Akimoto, H., Fukuda, M., Komazaki, Y., Yokouchi, Y., Koike, M., Tanimoto, H., Takegawa, N., and Kondo, Y.: Urban photochemistry in central Tokyo: 1. Observed and modeled OH and HO2 radical concentrations during the winter and summer of 2004, J. Geophys. Res., 112(D21), 2007.

Kanaya, Y., Pochanart, P., Liu, Y., Li, J., Tanimoto, H., Kato, S., Suthawaree, J., Inomata, S., Taketani, F., Okuzawa, K., Kawamura, K., Akimoto, H., and Wang, Z. F.:

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Rates and regimes of photochemical ozone production over Central East China in June 2006: a box model analysis using comprehensive measurements of ozone precursors, Atmos. Chem. Phys., 9(20), 7711-7723, 2009.

Khan, S., Juvigny-Khenafou, N.P., Dalu, T., Milham, P.J., Hamid, Y., Eltohamy, K.M., Ullah, H., Amiri, B.J., Chen, H., and Wu, N. C.: The influence of pH, and temperature on benthic chlorophyll-a: Insights from SHAP-XGBoost and random forest models, Ecol. Inform., 91, 103355, 2025.

Khruengsai, S., Sivapornnukul, P., Janta, R., Phonrung, N., Sripahco, T., Meesang, W., Aiyathiti, C., Prabamroong, T., Mahatheeranont, S., Pripdeevech, P., Poshyachind, S., and Pongpiachan, S.: Seasonal and height dynamics of volatile organic compounds in rubber plantation: Impacts on ozone and secondary organic aerosol formation, Sci. Total Environ., 945, 173984, 2024.

Kittipornkul, P., Thiravetyan, P., Hoshika, Y., Sorrentino, B., Popa, I., Leca, S., Sicard, P., Paoletti, E., and De Marco, A.: Surface ozone risk to human health and vegetation in tropical region: The case of Thailand, Environ. Res., 234, 116566, 2023.

Li, C. Q., Guo, S. C., Huang, Q., Mu, H. W., Yuan, B., Xia, Z. L., Fang, H., Zhang, W., Tang, P. F., and Du, P. J.: Dynamic monitoring of fine-grained ecological vulnerability in dryland urban agglomeration integrating novel remote sensing index and explainable machine learning, GISci. Remote Sens., 62, 2528302, 2025.

Li, P. Z., Chen, C., Liu, D., Lian, J., Li, W., Fan, C. Y., Yan, L. Y., Gao, Y., Wang, M., Liu, H., Pan, X. L., and Mao, J.: Characteristics and source apportionment of ambient volatile organic compounds and ozone generation sensitivity in urban Jiaozuo, China, J. Environ. Sci., 138, 607-625, 2024.

Li, Y. D., Yin, S. S., Yu, S. J., Yuan, M. H., Dong, Z., Zhang, D., Yang, L. M., and Zhang, R. Q.: Characteristics, source apportionment and health risks of ambient VOCs during high ozone period at an urban site in central plain, China, Chem., 250, 126283, 2020.

Li, Y., Ye, C., Ma, X. F., Tan, Z. F., Yang, X. P., Zhai, T. Y., Liu, Y. H., Lu, K. D., and Zhang, Y. H.: Radical chemistry and VOCs-NOx-O3-nitrate sensitivity in the polluted atmosphere of a suburban site in the North China Plain, Sci. Total Environ., 947,

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

174405, 2024.

Liang, S. Y., Gao, S., Wang, S., Chai, W. X., Chen, W. T., and Tang, G. G.: Characteristics, sources of volatile organic compounds, and their contributions to secondary air pollution during different periods in Beijing, China, Sci. Total Environ., 858, 159831, 2023.

Liao, Z. H., Jia, X. C., Qiu, Y. L., Quan, J. N., Pan, Y. B., Ma, P. K., Cheng, Z.G., and Wang, Q. Q.: Synoptic controls on warm-season O_3 pollution in eastern China: A focus on O3-NOx-VOC chemistry, Atmos. Res., 311, 107660, 2024.

Liu, B. H., Lin, H., Chen, Y. F., and Yang, C. Y.: Prediction of rock unloading strength based on PSO-XGBoost hybrid models, Materials, 17, 4214, 2024.

Liu, C. Q., Shi, K.: A review on methodology in O_3-NO_x-VOC sensitivity study, Environ. Pollut., 291, 118249, 2021.

Liu, C., Lu, B. Q., Wang, Q., Zhang, Z. K., Meng, X., Huo, J. T., Herrmann, H., and Li, X.: High-level HONO exacerbates double high pollution of O_3 and $PM_{2.5}$ in China, Sci. Total Environ., 174066, 2024.

Liu, M. C., Guo, C. H., and Xu, L. C.: An interpretable automated feature engineering framework for improving logistic regression, Appl. Soft Comput., 153, 111269, 2024.

Liu, X. F., Guo, H., Zeng, L. W., Lyu, X., Wang, Y., Zeren, Y., Yang, J., Zhang, L. Y., Zhao, S. Z., Li, J., and Zhang, G.: Photochemical ozone pollution in five Chinese megacities in summer , Sci. Total Environ., 801, 149603, 2018.

Liu, X. H., Zhang, Y., Xing, J., Zhang, Q., Wang, K., Streets, D. G., Jang, C., Wang, W. X., and Hao, J. M.: Understanding of regional air pollution over China using CMAQ, part II. Process analysis and sensitivity of ozone and particulate matter to precursor emissions, Atmos. Environ., 44(30), 3719-3727, 2010.

Liu, Y. F., Qiu, P. P., Li, C. L., Li, X. K., Ma, W., Yin, S. J., Yu, Q., Li, J. F., and Liu, X. G.: Evolution and variations of atmospheric VOCs and O_3 photochemistry during a summer O_3 event in a county-level city, Southern China, Atmos. Environ., 272, 118942, 2022.

Liu, Y. F., Qiu, P. P., Xu, K., Li, C. L., Yin, S. J., Zhang, Y. J., Ding, Y., Zhang, C., Wang, Z., Zhai, R. X., Deng, Y J., Yan, F. Y., Zhang, W. J., Xue, Z. G., Sun, Y.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

L., Ji, D. S., Li, J., Chen, J., Tian, H. Z.., Liu, X. G., and Zhang, Y.: Analysis of VOC emissions and O_3 control strategies in the Fenhe Plain cities, China, J. Environ. Manage., 325, 116534, 2023.

Liu, Y., Shao, M., Fu, L. L., Lu, S. H., Zeng, L. M., and Tang, D. G.: Source profiles of volatile organic compounds (VOCs) measured in China: Part I, Atmos. Environ., 42(25), 6247-6260, 2008.

Lu, X. C., Chen, N., Wang, Y. H., Cao, W. X., Zhu, B., Yao, T., Jimmy C. H., Fung and Lau, A. K. H.: Radical budget and ozone chemistry during autumn in the atmosphere of an urban site in central China, J. Geophys. Res. Atmos., 122(6), 3672-3685, 2017.

Mandal, T. K., Yadav, P., Kumar, M., Lal, S., Soni, K., Yadav, L., Saharan, U. S., and Sharma, S. K.: Characteristics of volatile organic compounds (VOCs) at an urban site of Delhi, India: Diurnal and seasonal variation, sources apportionment, Urban Clim., 49, 101545, 2023.

Martin, R. V., Fiore, A. M., and Donkelaar, A. V.: Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions, Geophys. Res. Lett., 31(6), 2004.

Martinez, M., Harder, H., Kovacs, T. A., Simpas, J. B., Bassis, J., Lesher, R., Brune, W. H., Frost, G. J., Williams, E. J., Stroud, C. A., Jobson, B. T., Roberts, J. M., Hall, S. R., Shetter, R. E., Wert, B., Fried, A., Alicke, B., Stutz, J., Young, V. L., White, A. B., and Zamora, R. J.: OH and HO_2 concentrations, sources, and loss rates during the Southern Oxidants Study in Nashville, Tennessee, summer 1999, J. Geophys. Res-Atmos., 108(D19), 2003.

Michoud, V., Kukui, A., Camredon, M., Colomb, A., Borbon, A., Miet, K., Aumont, B., Beekmann, M., Durand-Jolibois, R., Perrier, S., Zapf, P., Siour, G., Ait-Helal, W., Locoge, N., Sauvage, S., Afif, C., Gros, V., Furger, M., Ancellet, G., and Doussin, J. F.: Radical budget analysis in a suburban European site during the MEGAPOLI summer field campaign, Atmos. Chem. Phys., 12(24), 11951-11974, 2012.

Min, R. Q., Wang, F., Wang, Y. B., Song, G. X., Zheng, H., Zhang, H. P., Ru, X. T., and Song, H. Q.: Contribution of local and surrounding area anthropogenic emissions

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

to a high ozone episode in Zhengzhou, China. Environ. Res., 212, 113440, 2022.

Niu, Y. Y., Yan, Y. L., Xing, Y. R., Duan, X. L., Yue, K., Dong, J. Q., Hu, D. M., Wang, Y. H., and Peng, L.: Analyzing ozone formation sensitivity in a typical industrial city in China: Implications for effective source control in the chemical transition regime, Sci. Total Environ., 919, 170559, 2024.

Pan, X., Kanaya, Y., Tanimoto, H., Inomata, S., Wang, Z., Kudo, S., and Uno, I.: Examining the major contributors of ozone pollution in a rural area of the Yangtze River Delta region during harvest season, Atmos. Chem. Phys., 15(11), 6101-6111, 2015.

Ran, L., Zhao, C. S., Xu, W. Y., Lu, X. Q., Han, M., Lin, W. L., Yan, P., Xu, X. B., Deng, Z. Z., Ma, N., Liu, P. F., Yu, J., Liang, W. D., and Chen, L. L.: VOC reactivity and its effect on ozone production during the HaChi summer campaign. Atmospheric Chemistry and Physics, 11(10), 4657-4667, 2011.

Razavi-Termeh, S. V., Sadeghi-Niaraki, A., Ali, F., Naqvi, R. A., and Choi, S. M.: Spatio-temporal modeling of asthma-prone areas: Exploring the influence of urban climate factors with explainable artificial intelligence (XAI), Sustain. Cities Soc., 116, 105889, 2024.

Ren, X. R., Harder, H., Martinez, M., Lesher, R. L., Oliger, A., Shirley, T., Adams, J., Simpas, J. B., and Brune, W. H.: HO_x concentrations and OH reactivity observations in New York City during PMTACS-NY2001, Atmos. Environ., 37(26), 3627-3637, 2003.

Rohrer, F., Lu, K. D., Hofzumahaus, A., Bohn, B., Brauers, T., Chang, C. C., Fuchs, H., Häseler, R., Holland, F., Hu, M., Kita, K., Kondo, Y., Li, X., Lou, S. R., Oebel, A., Shao, M., Zeng, L. M., Zhu, T., Zhang, Y. H., and Wahner, A.: Maximum efficiency in the hydroxyl-radical-based self-cleansing of the troposphere, Nat. Geosci., 7(8), 559-563, 2014.

Santiago, J. V., Palomera, M. J., Martinez, C. R., Matamoros, A. H., Hata, H., Inoue, K., and Tonokura, K.: Ozone responses to reduced precursor emissions: A modeling analysis on how attainable goals can improve air quality in the Mexico City Metropolitan Area, Sci. Total Environ., 912, 169180, 2024.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Saqer, R., Issa, S., and Saleous, N.: Spatio-Temporal Characterization of PM_{10} Concentration across Abu Dhabi Emirate (UAE), Heliyon, 10(12), e32812, 2024.

Schifter, I., González, U., Díaz, L., González-Macías, C., and Mejía-Centeno, I.: Experimental and vehicle (on road) test investigations of spark-ignited engine performance and emissions using high concentration of MTBE as oxygenated additive, Fuel, 187, 276-284, 2017.

Shao, B., Cui, Y., He, Q. S., Guo, L. L., Gao, J., Zhao, J. R., and Wang, X. M.: Parameterized atmospheric oxidation capacity during summer at an urban site in Taiyuan and implications for O_3 pollution control, Atmos. Pollut. Res., 15(8), 102181, 2024.

Shao, M., Lv, S., Wei, Y. J., and Zhu, J. L.: The various synergistic impacts of precursor emission reduction on $PM_{2.5}$ and O_3 in a typical industrial city with complex distributions of emissions, Sci. Total Environ., 173497, 2024.

Shen, L. J., Xiang, P., Liang, S. W., Chen, W. T., Wang, M., Lu, S. H., and Wang, Z. W.: Sources profiles of volatile organic compounds (VOCs) measured in a typical industrial process in Wuhan, Central China, Atmosphere-Baseal, 9(8), 297, 2018.

Sillman, S.: The relation between ozone, NOx and hydrocarbons in urban and polluted rural environments, Atmos. Environ., 33(12), 1821-1845, 1999.

Song, M., Kim, E., Lee, Y., Oh, S. H., Yu, G. H., Choe, S., Park, G., Lee, T., and Bae, M. S.: Seasonal vehicle emission rate of chemical compounds related to fuel type from on-road tunnel measurement, Atmos. Environ., 305, 119777, 2023.

Song, M., Tan, Q., Feng, M., Qu, Y., Liu, X. G., An, J. L., and Zhang, Y. H.: Source apportionment and secondary transformation of atmospheric nonmethane hydrocarbons in Chengdu, Southwest China, J. Geophys. Res-Atmos., 123(17), 9741-9763, 2018.

Su, F. C., Xu, Q. X., Wang, K., Yin, S. S., Wang, S. B., Zhang, R. Q., Tang, Y., and Ying, Q.: On the effectiveness of short-term intensive emission controls on ozone and particulate matter in a heavily polluted megacity in central China, Atmos. Environ., 246, 118111, 2021.

Su, F. C., Xu, Q. X., Yin, S. S., Wang, K., Liu, G. J., Wang, P., Kang, M. J., Zhang, R.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Q., and Ying, Q.: Contributions of local emissions and regional background to summertime ozone in central China, J. Environ. Manage., 338, 117778, 2023.

Tabo, Z., Wangalwa, R., Rwibutso, M., Breuer, L., and Albrecht, C.: Future climate and demographic changes will almost double the risk of schistosomiasis transmission in the Lake Victoria Basin, One Health, 21, 101148, 2025.

Takefuji, Y.: Beyond XGBoost and SHAP: Unveiling true feature importance, J. Hazard. Mater., 488, 137382, 2025.

Takefuji, Y.: Reevaluating feature importance in machine learning: concerns regarding SHAP interpretations in the context of the EU artificial intelligence act, Water Res., 280, 123514, 2025.

Tan, Z. F., Fuchs, H., Lu, K. D., Hofzumahaus, A., Bohn, B., Broch, S., Dong, H., Gomm, S., Häseler, R., He, L. Y., Holland, F., Li, X., Liu, Y., Lu, S. H., Rohrer, F., Shao, M., Wang, B.L., Wang, M., Wu, Y. S., Zeng, L. M., Zhang, Y. S., Wahner, A., and Zhang, Y. H.: 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(1), 663-690, 2017.

Tan, Z. F., Lu, K. D., Jiang, M. Q., Su, R., Dong, H. B., Zeng, L. M., Xie, S. D., Tan, Q. W., and Zhang, Y. H.: 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, 2018.

Tan, Z. F., Lu, K. D., Jiang, M. Q., Su, R., Wang, H. L., Lou, S. R., Fu, Q. Y., Zhai, C. Z., Tan, Q. W., Yue, D. L., Chen, D. H., Wang, Z. S., Xie, S. D., Zeng, L. M., and Zhang, Y.: Daytime atmospheric oxidation capacity in four Chinese megacities during the photochemically polluted season: a case study based on box model simulation, Atmos. Chem. Phys., 19(6), 3493-3513, 2019.

Tran, T., Kumar, N., and Knipping, E.: Investigating sensitivity of ozone to emission reductions in the New York City (NYC) metropolitan and downwind areas, Atmos. Environ., 301, 119675, 2023.

Tudor, C.: Ozone pollution in London and Edinburgh: spatiotemporal characteristics, trends, transport and the impact of COVID-19 control measures, Hel., 8(11), 2022.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Vestenius, M., Hopke, P. K., Lehtipalo, K., Petäjä, T., Hakola, H., and Hellén, H.: Assessing volatile organic compound sources in a boreal forest using positive matrix factorization (PMF), Atmos. Environ., 259, 118503, 2021.

Wang, F., Zhang, C., Ge, Y., Zhang, Z., Shi, G. L., and Feng, Y. C.: Multi-scale analysis of the chemical and physical pollution evolution process from pre-co-pollution day to $PM_{2.5}$ and O_3 co-pollution day, Sci. Total Environ., 945, 173729, 2024.

Wang, H. J., Liang, Q. X., Hancock, J. T., and Khoshgoftaar, T. M.: Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods, J. Big Data, 11, 44, 2024.

Wang, M., Lu, S. H., Shao, M., Zeng, L. M., Zheng, J., Xie, F. J., Lin, H. T., Hu, K., and Lu, X.D.: Impact of COVID-19 lockdown on ambient levels and sources of volatile organic compounds (VOCs) in Nanjing, China, Sci. Total Environ., 757, 143823, 2021.

Wang, R. P., Duan, W. J., Cheng, S. Y., and Wang, X. Q.: Nonlinear and lagged effects of VOCs on SOA and O_3 and multi-model validated control strategy for VOC sources, Sci. Total Environ., 887, 164114, 2023.

Wang, S. B., Wang, L. L., Wang, N., Ma, S. L., Su, F. C., and Zhang, R. Q.: Formation of droplet-mode secondary inorganic aerosol dominated the increased $PM_{2.5}$ during both local and transport haze episodes in Zhengzhou, China, Chem., 269, 128744, 2021.

Wang, S.Y., Zhao, Y. L., Han, Y., Li, R., Fu, H. B., Gao, S., Duan, Y. S., Zhang, L. W., and Chen, J. M.: Spatiotemporal variation, source and secondary transformation potential of volatile organic compounds (VOCs) during the winter days in Shanghai, China, Atmos. Environ., 286, 119203, 2022.

Wang, X. D., Yin, S. S., Zhang, R. Q., Yuan, M. H., and Ying, Q.: Assessment of summertime O_3 formation and the O_3-NO_x-VOC sensitivity in Zhengzhou, China using an observation-based model, Sci. Total Environ., 813, 152449, 2022.

Wu, F. K., Yu, Y., Sun, J., Zhang, J. K., Wang, J., Tang, G. Q., and Wang, Y. S.: Characteristics, source apportionment and reactivity of ambient volatile organic

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

compounds at Dinghu Mountain in Guangdong Province, China, Sci. Total Environ., 548, 347-359, 2016.

Wu, Y. J., Liu, Y., Liu, P. J., Sun, L. N., Song, P. F., Peng, J. F., Li, R. K., Wei, N., Wu, L., Wang, T., Zhang, L. N., Yang, N., and Mao, H. J.: Evaluating vehicular exhaust and evaporative emissions via VOC measurement in an underground parking garage, Environ. Pollut., 333, 122022, 2023.

Wu, Y. T., Liu, B. S., Meng, H., Wang, F. Q., Li, S., Xu, M., Shi, L. Y., Zhang, S. F., Feng, Y. C., and Hopke, P. K.: Unexpected changes in source apportioned results derived from different ambient VOC metrics, Environ. Int., 190, 108910, 2024.

Xian, Y. H., Zhang, Y., Liu, Z. H., Wang, H. F., Wang, J. J., and Tang, C.: Source apportionment and formation of warm season ozone pollution in Chengdu based on CMAQ-ISAM, Urban Clim., 56, 102017, 2024.

Xiao, S. X., Zhang, Y. L., Zhang, Z., Song, W., Pei, C. L., Chen, D. H., and Wang, X. M.: The contributions of non-methane hydrocarbon emissions by different fuel type on-road vehicles based on tests in a heavily trafficked urban tunnel, Sci. Total Environ., 873, 162432, 2023.

Xiao, Z., Yang, X. R., Gu, H. M., Hu, J. L., Zhang, T. G., Chen, J. N., Pan, X. K., Xiu, G. L., Zhang, W., and Lin, M. Y.:Characterization and sources of volatile organic compounds (VOCs) during 2022 summer ozone pollution control in Shanghai, China, Atmos. Environ., 327, 120464, 2024.

Xu, Y. F., Shen, A., Jin, Y. B., Liu, Y. M., Lu, X., Fan, S. J., Hong, Y. Y., and Fan, Q.: A quantitative assessment and process analysis of the contribution from meteorological conditions in an O_3 pollution episode in Guangzhou, China, Atmos. Environ., 303, 119757, 2023.

Xue, L. K., Wang, T., Gao, J., Ding, A. J., Zhou, X. H., Blake, D. R., Wang, X. F., Saunders, S. M., Fan, S. J., Zuo, H. C., Zhang, Q. Z., and Wang, W. X.: Ground-level ozone in four Chinese cities: precursors, regional transport and heterogeneous processes, Atmos. Chem. Phys., 14(23), 13175-13188, 2014.

Yan, Y. B., and Yang, Y.: Revealing the synergistic spatial effects in soil heavy metal pollution with explainable machine learning models, J. Hazard. Mater., 482, 136578,

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

2025.

Yao, T. E., Lu, S. H., Wang, Y. Q., Li, X. H., Ye, H. X., Duan, Y. S., Fu, Q. Y., 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.

Yu, S. J., Su, F. C., Yin, S. S., Wang, S. B., Xu, R. X., He, B., Fan, X. G., Yuan, M. H., and Zhang, R. Q.: Characterization of ambient volatile organic compounds, source apportionment, and the ozone-NO_x-VOC sensitivities in a heavily polluted megacity of central China: Effect of sporting events and the emission reductions, Atmos. Chem. Phys., 2021, 1-61, 2021.

Yu, S. J., Wang, S. B., Xu, R. X., Zhang, D., Zhang, M., Su, F. C., Lu, X., Li, X., Zhang, R. Q., and Wang, L. L.: Measurement report: Intra-and interannual variability and source apportionment of volatile organic compounds during 2018–2020 in Zhengzhou, central China, Atmos. Chem. Phys., 22(22), 14859-14878, 2022.

Zhang, C. W., Xie, Y. M., Shao, M., and Wang, Q. G.: Application of machine learning to analyze ozone sensitivity to influencing factors: A case study in Nanjing, China, Sci. Total Environ., 929, 172544, 2024.

Zhang, H., Zhang, C., Liu, S. S., Yin, S. J., Zhang, S. Q., Zhu, H. G., Yan, F. Y., Yang, H., Ru, X. N., and Liu, X. G.: Insights into the source characterization, risk assessment and ozone formation sensitivity of ambient VOCs at an urban site in the Fenwei Plain, China, J. Hazard. Mater., 136721, 2024.

Zhang, L. H., Li, H., Wu, Z. H., Zhang, W. Q., Liu, K. K., Cheng, X., Zhang, Y. J., Li, B., and Chen, Y. Z.: Characteristics of atmospheric volatile organic compounds in urban area of Beijing: Variations, photochemical reactivity and source apportionment, J. Environ. Sci., 95, 190-200, 2020.

Zhang, L., Wang, L. L., Ji, D., Xia, Z., Nan, P. F., Zhang, J. X., Li, K., Qi, B., Du, R. G., Sun, Y., Wang, Y. S., 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.

Zhang, S. X., Zhang, Z. Z., Li, Y., Du, X. H., Qu, L. L., Tang, W., Xu, J., and Meng, F.:

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Formation processes and source contributions of ground-level ozone in urban and suburban Beijing using the WRF-CMAQ modelling system, J. Environ. Sci., 127, 753-766, 2023.

Zhang, X. F., Yin, Y. Y., Wen, J. H., Huang, S. L., Han, D. M., Chen, X. j., and Cheng, J. P.: Characteristics, reactivity and source apportionment of ambient volatile organic compounds (VOCs) in a typical tourist city, Atmos. Environ., 215, 116898, 2019.

Zhang, Y., Chen, Y., Jiang, N., Wang, S. S., Zhang, R. Q., Lv, Z. Q., Hao, X. X., and Wei, Y. F.: Chemical-composition characteristics of PM_1 and $PM_{2.5}$ and effects on pH and light-extinction coefficients under different pollution levels in Zhengzhou, China, J. Clean. Prod., 409, 137274, 2023.

Zhou, M., Liu, Y. H., Lu, K. D. Yu, D., Li, C. M., Zhai, T. Y., Yang, S. D., Tan, Z. F., Ma, X. F., Li, X., Dong, H. B., Zeng, L. M., Chen, S. Y., Chen, J. H., Tan, Q. W., Song, D. L., Zhang, X. L., and Zhang, Y. H.: Ozone production sensitivity analysis for the Chengdu Plain Urban Agglomeration based on a muti-site and two-episode observation, Sci. Total Environ., 950, 175068, 2024.

Zhou, X., Li, Z. Q., Zhang, T. J., Wang, F. L., Wang, F. T., Tao, Y., Zhang, X., Wang, F. L., and Huang, J.: Volatile organic compounds in a typical petrochemical industrialized valley city of northwest China based on high-resolution PTR-MS measurements: Characterization, sources and chemical effects, Sci. Total Environ., 671, 883-896, 2019.

Zhu, J. X., Cheng, H. R., Peng, J., Zeng, P., Wang, Z. W., Lyu, X. P., and Guo, H.: O_3 photochemistry on O_3 episode days and non-O_3 episode days in Wuhan, Central China, Atmos. Environ., 223, 117236, 2020.

Zuo, H. F., Jiang, Y. C., Yuan, J., Wang, Z. Q., Zhang, P. Z., Guo, C., Wang, Z. Z., Chen, Y., Qing, W., Wei, Y. J., and Li, X. Q.: Pollution characteristics and source differences of VOCs before and after COVID-19 in Beijing, Sci. Total Environ., 907, 167694, 2024.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Figure list

Fig.1 Smoothing the time series of pollutants. Savitzky-Golay smoothing denoising method was employed to facilitate a clearer and more intuitive observation of pollutant trends, with the window size set to 50 points and the polynomial order configured to 1.

Fig. 2 (a) Impacts of driving factors on O_3 formation derived from SHAP analysis during the entire period; (b) Summary plots of SHAP interaction matrix values for O_3 in the entire episode.

Fig. 3 Diurnal variations in concentrations of some reactive VOCs species in Zhengzhou under different pollution levels.

Fig. 4 O_3 sector contribution distribution in Zhengzhou:(a) local anthropogenic emissions in Zhengzhou, (b) anthropogenic contributions from other regions in Henan Province, (c) industry, (d) solvents, (e) transportation, (f) electricity, (g) residential, (h) others.

Fig. 5 Source profiles of six factors derived from PMF modeling.

Fig. 6 Contributions (%) for the six sources identified by PMF model during the sampling period.

Fig. 7 Diurnal patterns of O_3 production and destruction rates under different pollution levels.

Fig. 8 Diurnal variation distribution of HOx radicals under different pollution levels.

Fig. 9 Distribution of O_3-NOx-VOCs sensitivity in Zhengzhou under different pollution levels.

Fig. 10 The nonlinear relationship between the local ozone production rate and the activities of VOCs and NOx under different pollution levels. The x-axis represents the OH reaction activity of AHCs, while the y-axis represents the OH reaction activity of NOx. The black straight line indicates the ridge line, and the black contour lines represent the local ozone production rate, measured in ppbv/h. The green pentagram, orange four-pointed star, and red triangle correspond to non-pollution days, mild pollution days, and moderate pollution days, respectively.

Fig. 11 Spatial comparison of O_3-NOx-VOCs sensitive regime from 2019 to 2021 in Zhengzhou.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

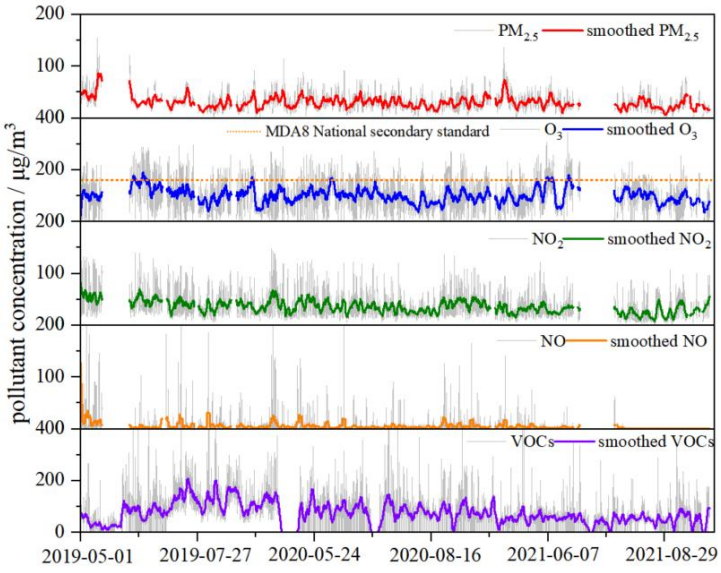

Fig.1 Smoothing the Time Series of Pollutants. Savitzky-Golay smoothing denoising method was employed to facilitate a clearer and more intuitive observation of pollutant trends, with the window size set to 50 points and the polynomial order configured to 1.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

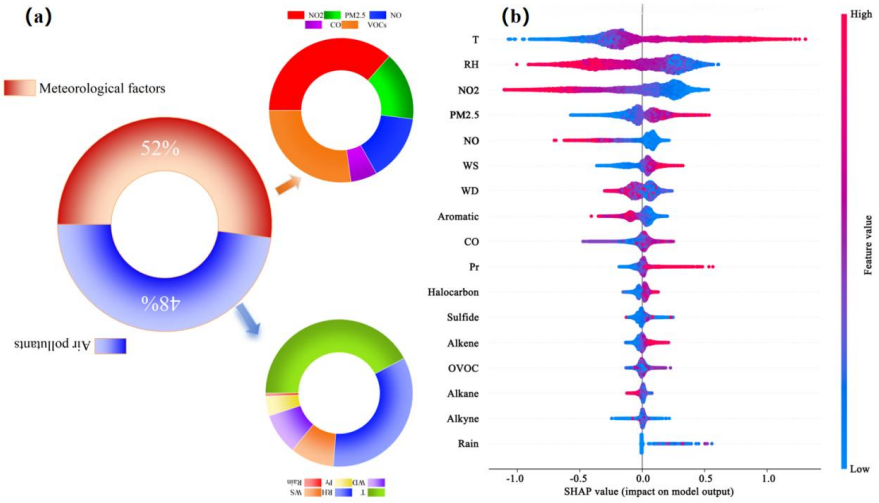

Fig. 2 (a) Impacts of driving factors on O_3 formation derived from SHAP analysis during the entire period; (b) Summary plots of SHAP interaction matrix values for O_3 in the entire episode.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

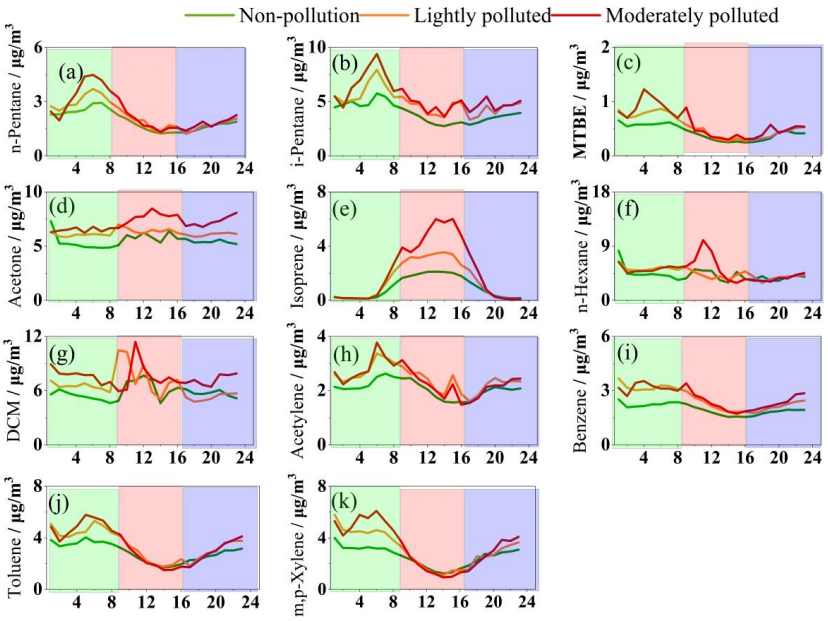

Fig. 3 Diurnal variations in concentrations of some reactive VOCs species in Zhengzhou under different pollution levels.The light blue, light red, and light green shadows represent stages P1, P2, and P3, respectively.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

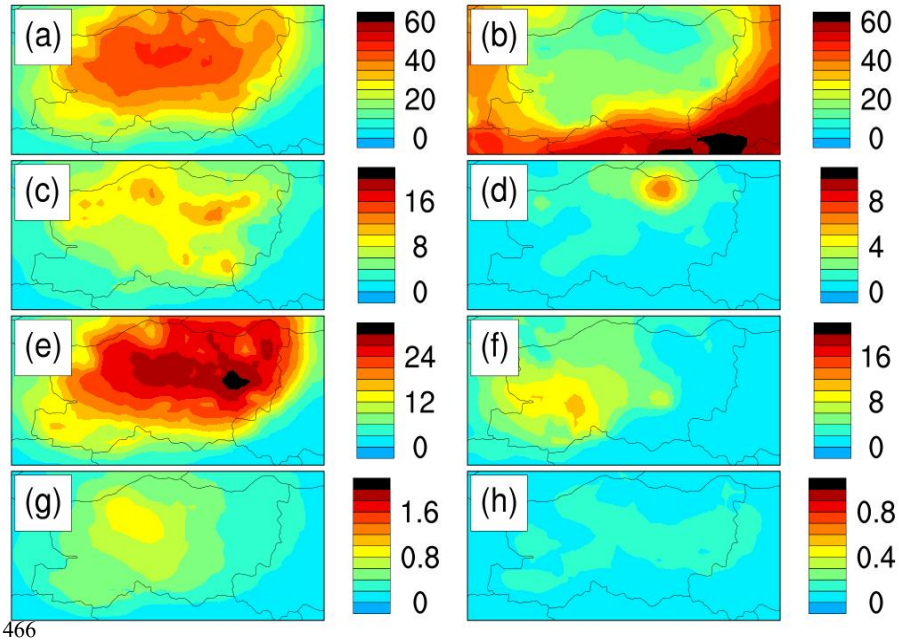

Fig. 4 O_3 sector contribution distribution in Zhengzhou:(a) local anthropogenic emissions in Zhengzhou, (b) anthropogenic contributions from other regions in Henan Province, (c) industry, (d) solvents, (e) transportation, (f) electricity, (g) residential, (h) others.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

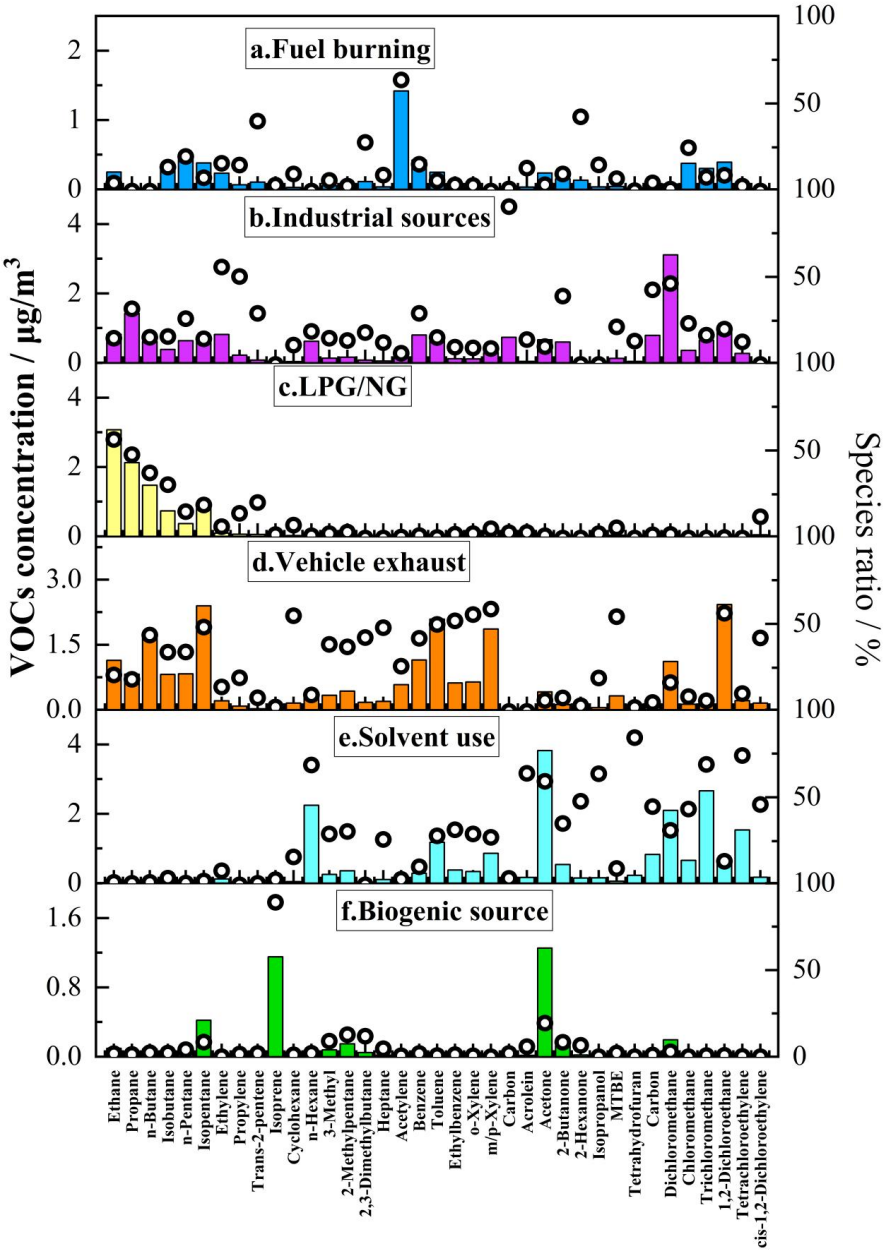

Fig. 5 Source profiles of six factors derived from PMF modeling.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

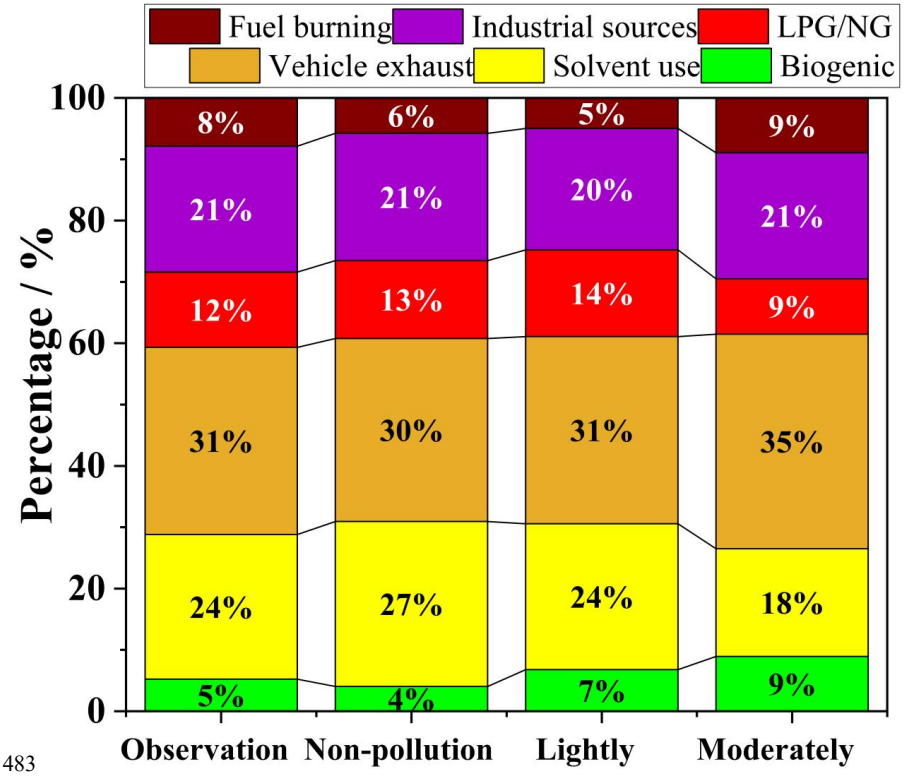

Fig. 6 Contributions (%) for the six sources identified by PMF model during the sampling period.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

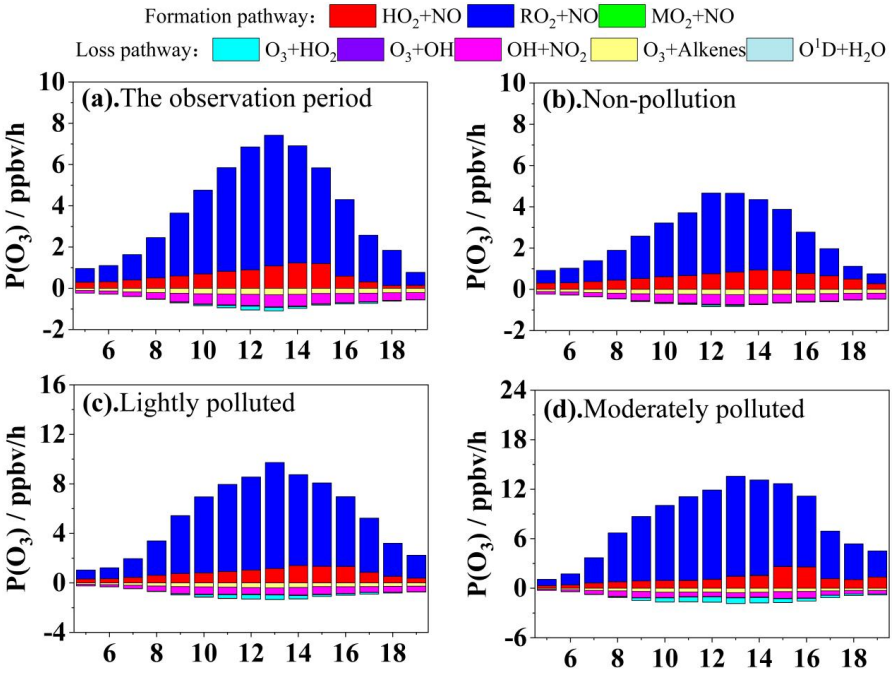

Fig. 7 Diurnal patterns of O_3 production and destruction rates under different pollution levels.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

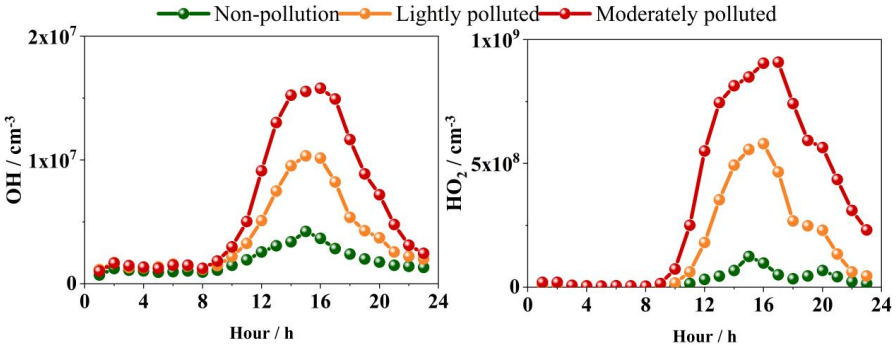

Fig. 8 Diurnal variation distribution of HOx radicals under different pollution levels.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

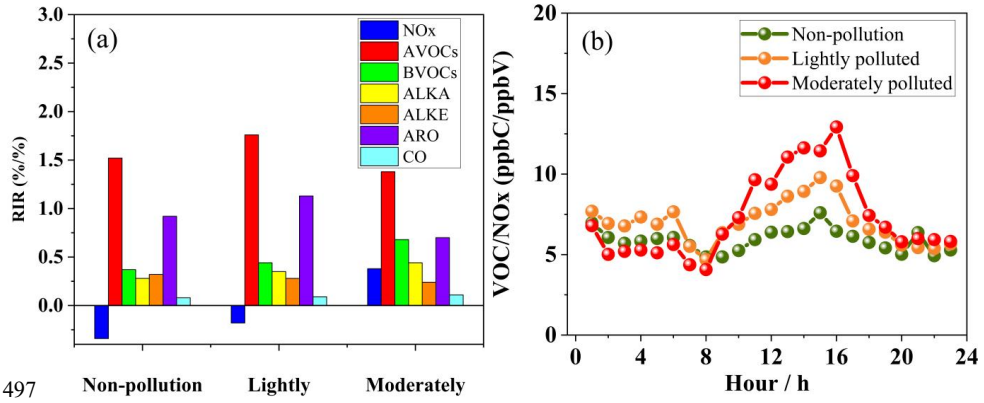

Fig. 9 Distribution of O_3-NOx-VOCs sensitivity in Zhengzhou under different pollution levels.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

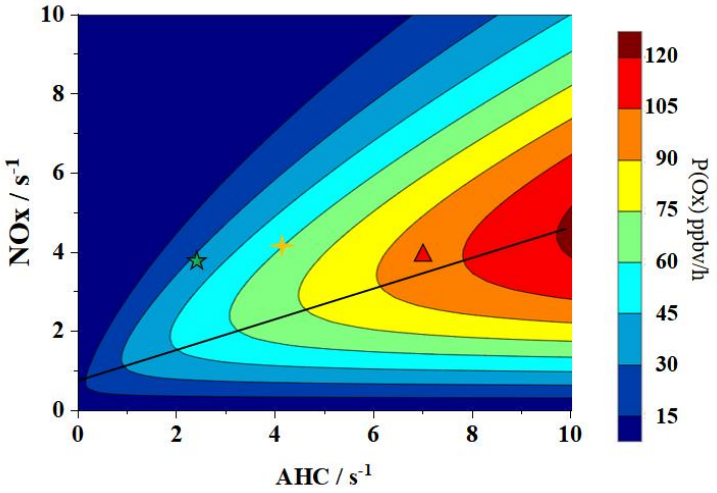

Fig. 10 The nonlinear relationship between the local ozone production rate and the activities of VOCs and NOx under different pollution levels. The x-axis represents the OH reaction activity of AHCs, while the y-axis represents the OH reaction activity of NOx. The black straight line indicates the ridge line, and the black contour lines represent the local ozone production rate, measured in ppbv/h. The green pentagram, orange four-pointed star, and red triangle correspond to non-pollution days, mild pollution days, and moderate pollution days, respectively.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

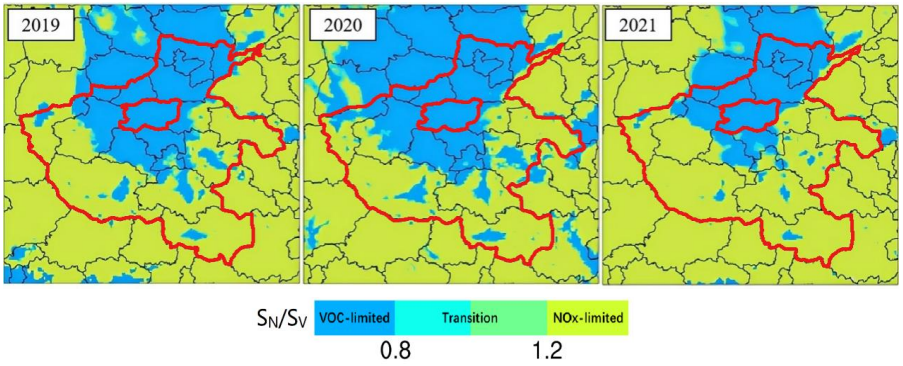

Fig. 11 Spatial comparison of O_3-NOx-VOCs sensitive regime from 2019 to 2021 in Zhengzhou.

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Table list:

Table 1 Meteorological factors and pollutant concentrations under different pollution levels.

Table 2 Concentrations and standard deviations of the Top 20 VOCs by concentration during different pollution periods（μg/m^3）.

Table 3 Contribution proportions of O_3 formation and removal pathways during the observation Period (%).

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Table 1 Meteorological factors and pollutant concentrations under different pollution levels.

	Unit	Non-pollution	Lightly	Moderately	Average
Percentage of polluted days	%	48	45	7	--
WS	m/s	1.6±1.1	1.8±1.2	1.7±1.1	1.7±1.2
RH	%	74.9±37.6	62.6±30	56.7±26.6	69.4±34.8
T	℃	25±11.2	27.9±11.4	29.6±12.1	26.3±11.5
VOCs	μg/m^3	84.7±51	96.6±53.4	105.3±59.4	90.3±52.8
NO	μg/m^3	5.7±18.3	6.2±18.1	5.3±13	5.8±17.9
NO_2	μg/m^3	32.7±25.6	38.7±28.8	40.7±25.9	35.3±27
PM_{10}	μg/m^3	69.8±61.9	88.3±58.1	100.1±54.5	78.3±61.1
$PM_{2.5}$	μg/m^3	27.1±20.9	32.8±20.4	35.5±16.4	29.7±20.7

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Table 2 Concentrations and standard deviations of the Top 20 VOCs by concentration during different pollution periods（μg/m^3）.

Non-pollution	Ave ± SD	Lightly	Ave ± SD	Moderately	Ave ± SD
Dichloromethane	6.7 ± 12.6	Dichloromethane	6.4 ± 15.6	Dichloromethane	7.6 ± 7.7
Ethane	5.5 ± 4	Acetone	6.2 ± 4.5	Acetone	7.2 ± 4.7
Acetone	5.1 ± 5.9	Ethane	5.4 ± 2.8	n-Hexane	6 ± 27.3
Propane	4.4 ± 3.2	Isopentane	4.9 ± 4.7	Ethane	5.5 ± 2.9
n-Hexane	4.1 ± 10.6	n-Hexane	4.5 ± 7.7	Isopentane	5.4 ± 4
Isopentane	4 ± 3.6	Propane	4.1 ± 3.5	1,2-Dichloroethane	4 ± 3.2
n-Butane	4 ± 3.5	n-Butane	3.8 ± 3.4	Propane	3.7 ± 2.2
1,2-Dichloroethane	3.6 ± 5.2	1,2-Dichloroethane	3.7 ± 3.4	Toluene	3.4 ± 2.9
Toluene	3.5 ± 3.7	Toluene	3.4 ± 3.5	n-Butane	3.3 ± 2.5
m/p-Xylene	3.2 ± 4.3	m/p-Xylene	3.1 ± 4	m/p-Xylene	3.3 ± 3.7
Trichloromethane	2.9 ± 5.3	Trichloromethane	2.8 ± 3.2	Trichloromethane	2.8 ± 1.8
Naphthalene	2.4 ± 4.7	Tetrachloroethylene	2.7 ± 3.6	Benzene	2.6 ± 1.8
Benzene	2.3 ± 1.7	Acetylene	2.5 ± 3	Tetrachloroethylene	2.6 ± 3
Acetylene	2.3 ± 1.9	Benzene	2.5 ± 2	n-Pentane	2.4 ± 1.9
Isobutane	2.2 ± 1.9	n-Pentane	2.3 ± 2	Acetylene	2.4 ± 1.5
n-Pentane	2.1 ± 1.7	Vinyl acetate	2.2 ± 3.3	Isoprene	2.3 ± 3.1
Ethylene	1.8 ± 1.2	Isobutane	2.2 ± 1.6	Vinyl acetate	2.2 ± 3.1
Tetrachloroethylene	1.8 ± 3.3	Carbon tetrachloride	1.8 ± 2.3	Isobutane	2.1 ± 1.1
Vinyl acetate	1.6 ± 4.1	Freon 12	1.6 ± 1.8	2-Methylpentane	1.9 ± 2.6
Freon 11	1.6 ± 0.7	Ethylene	1.6 ± 1.4	2-Butanone	1.8 ± 1

https://doi.org/10.5194/egusphere-2025-4519
Preprint. Discussion started: 23 October 2025
© Author(s) 2025. CC BY 4.0 License.

Table 3 Contribution proportions of O_3 formation and removal pathways during the observation Period (%).

	The observation period	Non-pollution	Lightly	Moderately
HO_2+NO	16	23	15	14
MO_2+NO	<1	<1	<1	<1
RO_2+NO	84	77	85	85
O_3+alkenes	33	35	31	28
OH+NO_2	56	58	51	43
O_3+OH	2	2	3	3
O_3+HO_2	8	4	15	25
O^1D+H_2O	<1	1	1	1