https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Driving Factors of Oxalic Acid and Enhanced Role of Gas-Phase Oxidation under Cleaner Conditions: Insights from 2007-2018 Field Observations in the Pearl River Delta

Yunfeng He [1, 4], Xiang Ding [1, 2, 3], Quanfu He [8], Yuqing Zhang [10], Metin Baykara [9], Duohong Chen [11], Tao Zhang [11], Kong Yang [1, 4], Junqi Wang [1, 4], Qian Cheng [1, 4], Hao Jiang [1, 4], Zirui Wang [1, 4], Ping Liu [1, 4], Xinming Wang [1, 2, 3], Michael Boy [5, 6, 7]

[1]State Key Laboratory of Advanced Environmental Technology, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China

[2]Guangdong–Hong Kong–Macao Joint Laboratory for Environmental Pollution and Control, Guangzhou Institute of Geochemistry, Chinese Academy of Science, Guangzhou 510640, China

[3]Guangdong Key Laboratory of Environmental Protection and Resources Utilization, Guangzhou Institute of Geochemistry, Guangzhou 510640, China

[4]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

[5]Institute for Atmospheric and Earth Systems Research, University of Helsinki, Helsinki 00014, Finland

[6]Atmospheric Modelling Centre – Lahti, Lahti University Campus, Lahti 15140, Finland

[7]School of Engineering Sciences, Lappeenranta-Lahti University of Technology, Lappeenranta 53850, Finland

[8]Thrust of Earth, Ocean and Atmospheric Sciences, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511453, China

[9]Climate and Marine Sciences Department, Eurasia Institute of Earth Sciences, Istanbul Technical University, Maslak, Istanbul 34469, Turkey

[10]School of Environment and Safety Engineering, North University of China, Taiyuan 030051, China

[11]State Environmental Protection Key Laboratory of Regional Air Quality Monitoring, Guangdong Environmental Monitoring Center, Guangzhou 510308, China

Correspondence to: Xiang Ding (xiangd@gig.ac.cn), Michael Boy (michael.boy@helsinki.fi)

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Abstract

Secondary organic aerosol (SOA) is a dominant constituent of fine particulate matter, exerting significant impacts on both climate and human health. Oxalic acid (C_2), a key end-product formed from the oxidation of volatile organic compounds, can provide insights into the formation mechanism of SOA. Thus, long-term measurements of C_2 and related compounds help understand the changes in SOA formation with decreasing pollutant levels. In this study, C_2 and its homologs, along with five primary anthropogenic source markers and three SOA markers, were measured in the Pearl River Delta (PRD) during 2007–2018. The concentrations of C_2 did not exhibit significant downward trends, despite substantial reductions in anthropogenic emissions, for example, biomass burning (-11% yr^{-1}), vehicle emissions (-17% yr^{-1}), and cooking emissions (-7% yr^{-1}). Correlation analysis revealed that aerosol liquid water content (ALWC) and O_x (O_3 + NO_2) were the main drivers of C_2 variation. Moreover, the relative contribution of biogenic SOA increased under cleaner conditions. A machine learning model was applied to quantify the contributions of anthropogenic precursors emission, biogenic precursors emission, aqueous-phase oxidation processes, and gas-phase oxidation processes to C_2 variability. As pollution levels declined, the contribution of gas-phase oxidation increased from 24% to 48%, whereas that of aqueous-phase oxidation declined from 35% to 20%. This shift indicated a transition from aqueous-phase to gas-phase pathways in C_2 and SOA formation. Our findings highlight the increasing importance of gas-phase oxidation under low-pollution conditions and underscore the need for effective ozone control strategies to further reduce SOA in the future.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

1 Introduction

Low molecular weight dicarboxylic acids (DCA) represent one of the most abundant classes of organic compounds in aerosols and are ubiquitously distributed in the lower troposphere (Kawamura and Bikkina, 2016). In urban areas, DCA and related compounds (e.g., glyoxal (Gly), methylglyoxal (mGly), glyoxylic acid (ωC2), and pyruvic acid (Pyr)) account for 0.2–6% of total carbon in aerosols (Kawamura and Ikushima, 1993; Kawamura and Watanabe, 2004), while this proportion could rise to 16% in remote marine (Kawamura and Sakaguchi, 1999; Wang et al., 2006). Due to their high solubility and hygroscopicity (Bilde et al., 2015), DCA can promote aerosol hygroscopic growth and facilitate the activation of cloud condensation nuclei (CCN). Thus, they play a crucial role in the Earth's climate and radiative forcing by scattering solar radiation and forming clouds (Kawamura et al., 2013a; Meng et al., 2018; 2023).

Previous studies have suggested that DCA can be directly emitted from biogenic sources (Rinaldi et al., 2011), fossil fuel combustion (Kawamura and Kaplan, 1987), biomass burning (Kawamura et al., 2013b; Cao et al., 2017), waste incineration (Jung et al., 2010), and cooking (He et al., 2004). However, increasing evidence has demonstrated that most DCA are generated through photochemical oxidation of volatile organic compounds (VOCs), such as olefins (Bikkina et al., 2021), cyclic alkanes (Kawamura and Usukura, 1993), and unsaturated fatty acids (Kawamura et al., 1996), followed by partitioning into wet aerosols. Therefore, DCA have been widely used as SOA indicators to investigate the secondary reactions and aging processes of VOCs in the atmosphere (Kawamura and Bikkina, 2016; Meng et al., 2023). As a key end-product of numerous formation pathways in aerosols, oxalic acid (C_2) is the most abundant compound among DCA homologs (58%–89%) (Ho et al., 2007; Xu et al., 2023). Extremely high concentrations of C_2 (~1000 ng m^{-3}) have been observed in polluted areas (Poore, 2000; Wang et al., 2012; Xu et al., 2023), suggesting fast photochemical oxidation of organic precursors. Thus, a growing number of studies have focused on identifying the sources and elucidating the formation mechanism of C_2 over the past decades.

Previous studies have shown that the concentrations of DCA and C_2 in urban areas are generally higher than those in coastal, mountain, and marine areas (Table S1). The elevated concentrations have been attributed to higher emissions of organic precursors in urban environments with intensive human activities. However, an unexpected increase in C_2 was reported during the COVID-19 period (Meng et al., 2023), a period characterized by a dramatic reduction in anthropogenic emissions. They identified higher oxidant levels and stronger solar radiation as drivers of this C_2 enhancement. These findings indicate that effective mitigation of C_2 and SOA requires the simultaneous control of both anthropogenic emissions and atmospheric oxidants. As anthropogenic emissions decline, the relative contribution of biogenic sources to C_2 is expected to increase. For example, Xu et al. (2022) found that biogenic sources accounted for only 22% of C_2 in continental outflows characterized by high anthropogenic emissions. In contrast, this proportion increased to 61% in oceanic air masses with minimal anthropogenic

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

influence. In addition, relatively low anthropogenic emissions and enhanced biogenic activity in summer led to a higher non-fossil contribution to C_2 (46%) compared to that in winter (33%) (Xu et al., 2023).

Numerous studies have reported a strong correlation between C_2 and sulfate across diverse regions, including East Asia (Yu et al., 2005; Wang et al., 2017; Zhang et al., 2022a; Meng et al., 2023), the USA (Hilario et al., 2021), the Pacific and Atlantic Oceans (Hilario et al., 2021), the Bay of Bengal (Bikkina et al., 2017) and Mongolia (Jung et al., 2010). This widespread correlation suggests that C_2 and sulfate may share common formation pathways across different atmospheric environments. Assuming that sulfate primarily originates from aqueous-phase processing, C_2 has been considered a secondary product formed via aqueous reactions (Yu et al., 2005; Hilario et al., 2021). However, recent research found that gas-phase oxidation also played an important role in C_2 formation. For instance, Ding et al. (2021) observed a strong correlation between C_2 and ozone (O_3) but not with sulfate during a haze event in the Yangtze River Delta, indicating C_2 was mainly produced via gas-phase oxidation. During COVID-19, lower ALWC and elevated O_3 shifted the dominant formation pathway of C_2 from aqueous-phase oxidation of ωC_2 and Pyr to gas-phase photochemical decomposition of longer-chain DCA (malonic (C_3) and succinic (C_4)). As a result, the contribution of gas-phase oxidation increased from 12.3% to 50.5% (Meng et al., 2023). In addition, Xu et al. (2023) found that the gas-phase photochemical aging dominated the formation of C_2 at five emission hotspots in China in summer, when pollutant levels were relatively low and solar radiation and oxidants concentrations were high. Recently, a laboratory experiment also proposed new reaction schemes for high levels of C_2 via gas-phase chemical reactions with O_3 and hydroxyl radicals, which could explain the high levels of C_2 in the free troposphere (Bikkina et al., 2021). Despite these findings, the formation mechanism of C_2 under different pollution conditions remains incompletely understood and requires further investigation.

Over the past decades, the implementation of stringent air pollution control policies has led to a significant improvement in air quality across the PRD region (Bian et al., 2019; Geng et al., 2019; Zhang et al., 2019; Yan et al., 2020). Although VOC emissions increased during this period (Bian et al., 2019; Guo et al., 2024), a reduction in SOA was observed (Yan et al., 2020), which was likely related to the changes in ALWC and atmospheric oxidation capacity. Long-term measurements of DCA can provide critical insights into the SOA formation processes from high-pollution to low-pollution conditions. However, until now, such measurements remain scarce in the PRD. In this study, we conducted long-term measurements of aliphatic DCA (C_2, C_4–C_{10}), primary anthropogenic sources markers, as well as SOA markers at a background station in the PRD during 2007–2018. Based on annual trends and correlation analysis, the driving factors of C_2 were identified. Then, a machine learning model was applied to quantify the contributions of anthropogenic precursors emission, biogenic precursors emission, aqueous-phase oxidation processes, and gas-phase oxidation processes to C_2 variability. Our results highlight an enhanced role of gas-phase oxidation in C_2 formation under cleaner conditions.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

2 Materials and methods

2.1 Field sampling

Particulate matter with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$) samples were collected on the rooftop of a five-storey building (approximately 30 m above the ground) in the campus of a middle school at Wanqingsha (WQS; 22.42°N, 113.32°E). The sampling site is in a small town in Guangzhou, which is located in the center of the PRD region (Fig. 1). The influence of local anthropogenic emissions on this site was limited due to the low traffic and less industrial emissions from surrounding areas. During the summertime, southwesterly winds bring clean and moist air masses from the ocean. In contrast, northerly winds transport polluted and dry air masses from urban clusters during fall and winter. This makes WQS an ideal background station for investigating regional air pollution.

$PM_{2.5}$ samples were collected using prebaked (450°C, 4h) quartz filters (8in. × 10in., QMA, Whatman, UK). Each sample lasted for 24h using a high-volume air sampler (HVPM2.5, Tisch Environmental Inc., USA) at an airflow rate of 1.1 m^3 min^{-1}. Field blank samples were also collected by mounting the blank filter onto the sampler for 10 min without turning on the sampler. In this study, a total of 453 $PM_{2.5}$ samples were collected mostly during the wintertime (October, November, and December) of each year from 2007 to 2018. After the collection, each filter was wrapped in an aluminum foil, zipped in Teflon bags, and stored in a freezer (-20°C) prior to analysis. Gaseous pollutants data (e.g., NO_2 and O_3) and meteorological parameters (e.g., temperature, relative humidity) were obtained from an air quality monitoring station operated at WQS. The station was under maintenance during 2012–2013, therefore data for this period are unavailable. The boundary layer height (BLH) data used in this study were obtained from the ERA5 reanalysis dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) via the Copernicus Climate Data Store (CDS, https://cds.climate.copernicus.eu/datasets/). The concentrations of $PM_{2.5}$ and its main components, as well as ALWC and pH, can be found in our previous study (He et al., 2025).

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

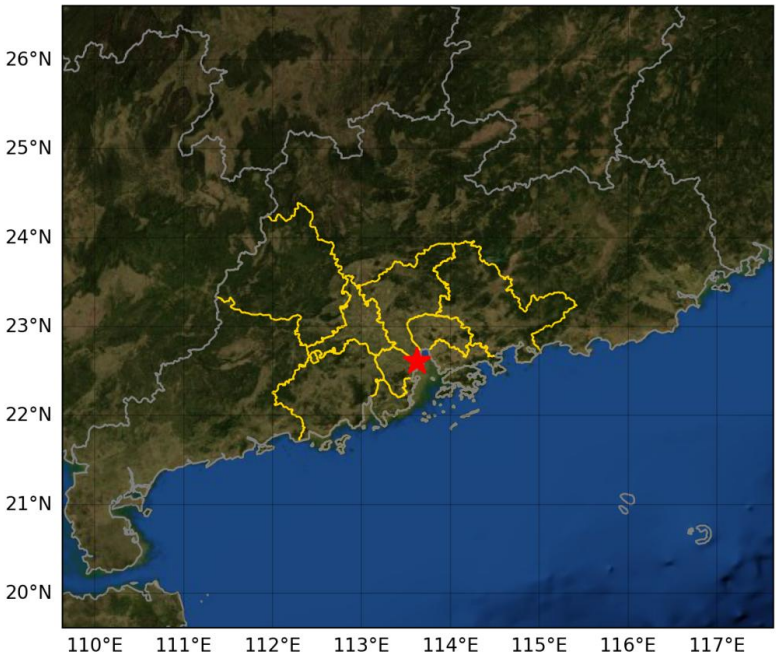


Figure 1. PRD region (golden line) consists of Guangzhou, Shenzhen, Zhuhai, Dongguan, Foshan, Huizhou, Zhongshan, Zhaoqing, and Jiangmen. The measurement station (red star) is located in the central area of the PRD. The map background is based on data from NASA's Goddard Space Flight Center and the Blue Marble project by Reto Stockli (NASA/GSFC).

2.2 Chemical analysis

For analysis of oxalate ($C_2O_4^{2-}$), a punch (5.06 cm^2) of the filters was extracted twice with 10 mL ultrapure Milli–Q water (18.2 MΩ cm / 25 °C) each for 15 min using an ultrasonic ice–water bath. The total water extracts (20 mL) were filtered through a 0.22 μm pore size filter and then stored in a pre-cleaned HDPE bottle. An ion-chromatography system (Metrohm, 883 Basic IC plus) was used to identify oxalate.

For organic tracers, a quarter of the filter was cut into small pieces and then mixed with 400 μL isotope-labeled internal standards, which included hexadecane-d_{24}, tetracosane-d_{50}, triacontane-d_{62}, naphthalene-d_8, acenaphthene-d_{10}, phenanthrene-d_{10}, chrysene-d_{12}, perylene-d_{12}, C_{27}aaa (20R) -cholestane-d_4, dodecanoic acid-d_{23}, hexadecanoic acid-d_{31}, docosanoic acid-d_{43}, levoglucosan-$^{13}C_6$, and phthalic acid-d_4. Samples were sonicated in an ice-water bath using mixed solvents, twice with a 1:1 (*v*/*v*) dichloromethane (DCM)/hexane solution, followed by twice with a 1:1 (*v*/*v*) DCM/methanol solution. The extraction solutions of each sample were combined, filtered, and concentrated to ~ 2 mL, and then split into two parts for silylation and methylation, respectively. All derivatized solutions were analyzed by an Agilent gas chromatography-mass spectrometry (GC/MS, 7890/5975C) in the scan mode with an HP-5MS capillary column (30m × 0.25 mm × 0.25μm). The detailed

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

information about sample treatment and GC-MS analysis was described by previous studies (Ding et al., 2012; 2014). In this study, we measured 9 dicarboxylic acids (succinic acid, glutaric acid, adipic acid, pimelic acid, suberic acid, azelaic acid, sebacic acid, phthalic acid, terephthalic acid), 5 hopanes (17α(H)-22,29,30-trisnorhopane, 17α(H),21β(H)-30-norhopane, 17α(H),21β(H)-30-hopane, 17α(H),21β(H)-22R-homohopane, and 17α(H),21β(H)-22S-homohopane), levoglucosan, octadecanoic acid, picene, 2,3-dihydroxy-4-oxopentanioic acid (DHOPA), and malic acid. Table S2 shows detailed information about target compounds and their internal standards. Due to the lack of a commercial standard, DHOPA was quantified using an isomer surrogate, citramalic acid. The analytical procedure of DHOPA is described elsewhere (Ding et al., 2017).

2.3 Quality assurance / quality control

We analyzed the field blank and found that oxalate, succinic acid, octadecanoic acid, phthalic acid, and terephthalic acid were detected in trace amounts. Thus, the data reported in this study were corrected by corresponding field blanks. We also analyzed the spiked samples (n = 7) to test the recoveries of analytical procedures. The results indicated that the recoveries were 95%–103% for oxalate, 100–121% for levoglucosan, 66%–118% for dicarboxylic acid and octadecanoic acid, 89%–113% for hopanes, 76%–97% for picene, 88%–94% for citramalic acid. The method detection limits (MDLs) were 0.04 μg m^{-3} for oxalate, 6.7–40.5 ng m^{-3} for dicarboxylic acids, 0.07 ng m^{-3} for levoglucosan, 0.010–0.017 ng m^{-3} for hopanes, 0.18 ng m^{-3} for octadecanoic acid, 0.11 ng m^{-3} for picene, and 0.11 ng m^{-3} for citramalic acid. Before data analysis, all data were manually inspected and the outliers (i.e., $X_{75\%} + 3(X_{75\%} - X_{25\%})$) were removed to rule out the influence of extreme concentrations on the overall trends. The change rates were calculated using the slopes derived from Theil–Sen regression and evaluated for statistical significance via the non-parametric Mann–Kendall test, providing a robust and reliable assessment of temporal variations.

2.4 Evaluate the contributions of various factors to oxalic acid variation by the machine learning model

Extreme Gradient Boosting (XGBoost), an advanced ensemble machine learning method based on gradient boosting decision trees, is known for its high computational efficiency, robust predictive performance (Chen et al., 2016) and thus has been applied in air pollutant research recently (Hou et al., 2022; Peng et al., 2023; Liu et al., 2025). In this study, XGBoost was employed to assess the relative contributions of various factors to oxalic acid variation. The implementation and Python package of XGBoost algorithm are publicly available online (https://github.com/dmlc/xgboost). A total of 11 variables were used as input features to train the model, including levoglucosan, hopanes, octadecanoic acid, picene, terephthalic acid (tPh), O_x, ALWC, pH, temperature (Temp), solar radiation (SR), and relative humidity (RH). To avoid redundant and confounding

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

explanations, the secondary organic molecular markers, such as DHOPA, phthalic acid (Ph), and malic acid, were excluded in the model training. Because they are influenced by VOC emissions and secondary oxidation processes, which are already represented by the factors mentioned above. Our results showed that there was a great agreement between the observations and simulations (r^2 = 0.86, Fig. S1), which indicated the model predictions were reliable.

To further explain the results of XGBoost, Python package of SHapley Additive ExPlanation (SHAP), a game theory-based approach proposed by Lundberg and Lee (2017), was used in this study (https://github.com/slundberg/shap). The SHAP approach can quantify the contribution of each feature to the model output while also accounting for the complex interdependencies among features. Thus, this approach provides deeper insights into the key drivers of oxalic acid formation. The fundamental principle of SHAP is to fairly allocate the prediction output among all input features by evaluating their marginal contributions across all possible feature subsets. For a given prediction $f(x_i)$ generated by the model, the SHAP decomposes the prediction into the sum of feature attributions (SHAP values) and a baseline expectation (equation 1):

$$f(x_i) = \emptyset_0(f,x) + \sum_{i=1}^{K} \emptyset_j(f,x_i) \quad (1)$$

where $\emptyset_j(f,x_i)$ is the SHAP value representing the impact of variable j on the prediction of the model for input x_i. The base value, $\emptyset_0(f,x) = E[f(x)]$, is the expected value of the model output over the data set. More details about this approach can be found in previous literature (Lundberg and Lee, 2017; Hou et al., 2022).

SHAP value represents the impact of a unit change in an individual variable on the predicted C_2 concentration, while holding all other variables constant. The sign of the SHAP value reflects the direction of the feature's effect on the target variable, while its magnitude quantifies the strength of that effect. For example, a positive (negative) SHAP value indicates that the change in a given feature increases (decreases) the predicted value relative to the model's baseline. The mean absolute SHAP (|SHAP|) values can remove the influence of directional effects that may cancel each other out when averaged, thereby providing a more robust measure of each feature's net contribution to model output. The use of |SHAP| values thus facilitates a comprehensive ranking of feature importance based solely on contribution intensity, regardless of whether the effect is positive or negative. We further calculated the impact factor (IF) to evaluate the contribution of individual factor (i) to oxalic acid variation as equation 2, where N is the number of samples:

$$IF\ (\%) = \frac{|\ SHAP\ value\ (i)\ |}{\sum_{j=1}^{N} |SHAP\ value\ (j)|} \quad (2)$$

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

3 Results and discussion

3.1 Annual trends of the molecular tracers and aliphatic DCA in the PRD

Annual variations in the organic molecular markers, aliphatic DCA and meteorological parameters are summarized in Table S3 and presented in Fig. 2. Five molecular markers representing different primary sources were analyzed in this study, including levoglucosan, hopanes, octadecanoic acid, picene, and terephthalic acid. These compounds are commonly used as tracers for biomass burning (Giannoni et al., 2012), vehicle emissions (Riddle et al., 2007), cooking (Zhao et al., 2007), coal combustion (Yu et al., 2020), and waste incineration (Simoneit et al., 2005), respectively. Between 2007 and 2018, the concentrations of levoglucosan, hopanes and octadecanoic acid exhibited significant declining trends (-11% yr^{-1}, -17% yr^{-1}, and -7% yr^{-1}, respectively; $p < 0.01$). In contrast, no clear trends were observed for picene and terephthalic acid, suggesting that control measures targeting biomass burning, vehicle emissions, and cooking were effective, while those for coal combustion and waste incineration may have been less effective or less implemented. Phthalic acid has been identified as a SOA tracer derived from naphthalene (Kleindienst et al., 2012), while DHOPA is a tracer for SOA formed from aromatic hydrocarbons (Ding et al., 2017). Given the substantial anthropogenic sources of naphthalene and aromatic hydrocarbons, phthalic acid and DHOPA can be used as anthropogenic SOA (ASOA) markers. Previous studies suggested that malic acid was the photodegradation product of the typical biogenic VOCs, such as monoterpenes and isoprene (Hu and Yu, 2013; Cheng et al., 2021). Thus, we used malic acid as a tracer for biogenic SOA (BSOA). Although phthalic acid, DHOPA, and malic acid decreased from 51.9 ± 14.9, 1.85 ± 1.35, and 24.2 ± 19.4 ng m^{-3} to 16.7 ± 5.7, 1.05 ± 0.88, 5.9 ± 4.9 ng m^{-3}, respectively, their declining trends were not statistically significant ($p > 0.05$). This indicated that the influence of reductions in emissions of anthropogenic organic precursors on SOA was limited.

Aliphatic DCA, a group of typical SOA molecular markers (Kawamura and Bikkina, 2016), were analyzed to investigate further the changes in SOA formation under decreasing pollutant levels. Although their concentrations in 2007 (864 ± 283 ng m^{-3}) were much higher than those in 2018 (307 ± 122 ng m^{-3}) (Fig. 3), this downward trend was not statistically significant ($p > 0.05$), which was similar to the trends of other SOA markers. Notably, a rebound in the concentrations of molecular markers and aliphatic DCA was observed in 2013, coinciding with a similar increase in $PM_{2.5}$ levels in the PRD (Yan et al., 2020). However, the reason for this was unclear.

C_2 was the most abundant compound among aliphatic DCA, accounting for 80%–91%, followed by C_4 (4%–13%), and C_9 (1%–4%). Therefore, the overall trend of aliphatic DCA was primarily driven by C_2 (Fig. S2), so subsequent discussions will focus on C_2. Its concentration declined from 692 ± 243 (2007) to 274 ± 114 (2018), but did not exhibit a clear trend ($p > 0.05$). To further explore the changes of SOA formation under different pollution conditions, our samples were divided into

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

five groups according to interim targets recommended by the Worle Health Organization (WHO) in 2021 (World Health Organization, 2021): IT0 ($PM_{2.5}$ > 75 μg m^{-3}), IT1 (75 μg m^{-3} > $PM_{2.5}$ > 50 μg m^{-3}), IT2 (50 μg m^{-3} > $PM_{2.5}$ > 37.5 μg m^{-3}), IT3 (37.5 μg m^{-3} > $PM_{2.5}$ > 25 μg m^{-3}), and IT4 (25 μg m^{-3} > $PM_{2.5}$). As presented in Table S4, the molecular markers and C_2 decreased significantly ($p < 0.01$) from IT0 to IT4. However, the ratio of C_2 to $PM_{2.5}$ ($C_2/PM_{2.5}$), which can avoid the effect of atmospheric dilution caused by meteorology and better reflect the relative contribution of SOA to $PM_{2.5}$, increased from 6.8 × 10^{-3} to 10.3× 10^{-3} ($p < 0.01$, Fig. S3). This indicated that SOA played a more important role under lower pollution conditions.

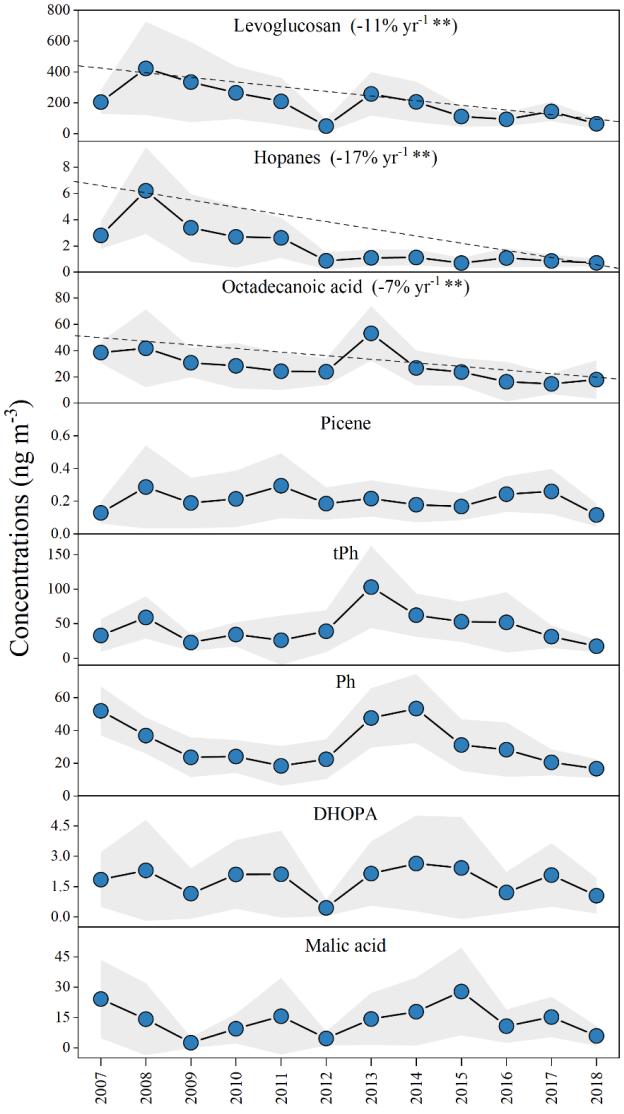


Figure 2. Annual variations in different molecular markers in the PRD during 2007 to 2018. Two asterisks denote p value < 0.01.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

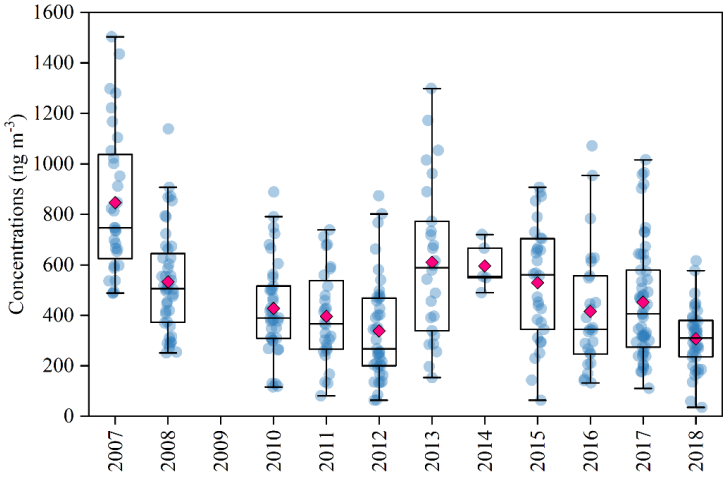


Figure 3. Annual variations in aliphatic DCA. The concentrations decreased from 864 ± 283 ng m^{-3} (2007) to 307 ± 122 ng m^{-3} (2018), but the trend was not statistically significant ($p > 0.05$). Due to the absence of oxalic acid measurements in 2009, the concentrations of aliphatic DCA for that year are not presented.

3.2 Sources and driving factors of oxalic acid

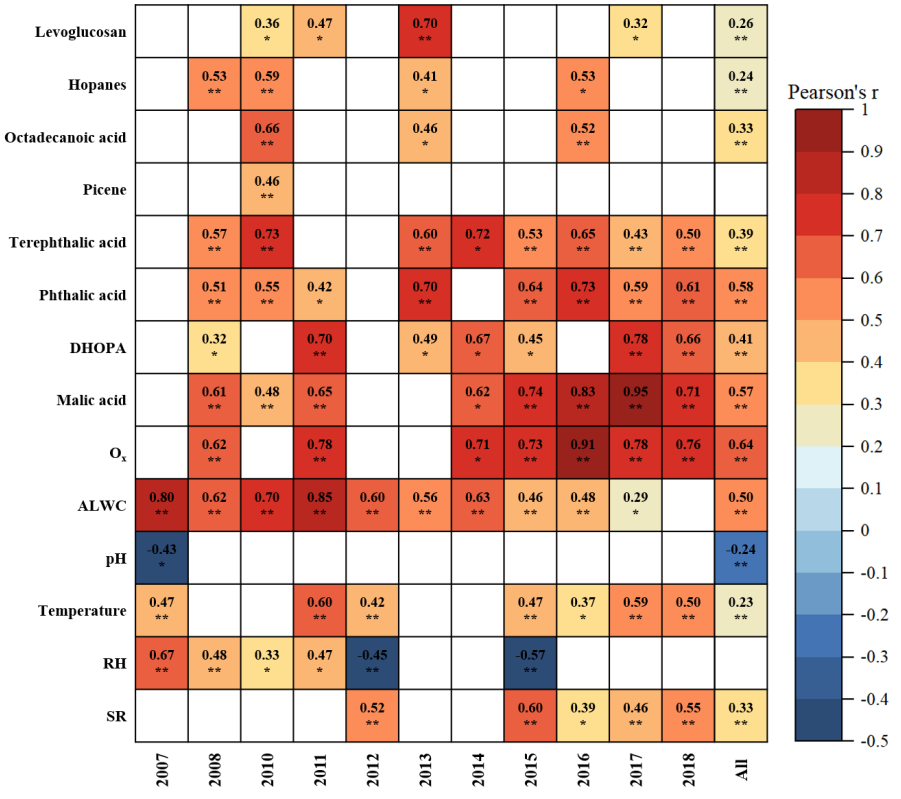
The ratio of C_2 to C_4 (C_2/C_4) has been used as an indicator for the aging degree of DCA and C_2 (Bikkina et al., 2015; Bikkina et al., 2017; Zhao et al., 2020; Meng et al., 2023), because the photochemical degradation of C_4 can lead to C_3, and C_3 can be photochemically oxidized into C_2 via intermediates (Kawamura and Bikkina, 2016). In this study, C_2/C_4 (21.9 ± 18.0) was significantly higher than those reported for primary sources, such as fossil fuel emission (4.1) (Kawamura and Kaplan, 1987) and biomass burning (1.6) (Cao et al., 2017). This indicated that C_2 was highly aged and mainly formed through secondary processes rather than being directly emitted from primary sources in this study. The ratio was comparable to those observed at a rural site of the PRD (15.4–21.0) (Xu et al., 2022), but substantially higher than those reported in urban areas of North China, such as Jinan (8.4 ± 3.4) (Meng et al., 2023), Tianjin (4.5–6.1) (Pavuluri et al., 2021), Xi'an (4.6–8.0) (Wang et al., 2012), and Beijing (4.7–8.6) (Zhao et al., 2018). This was likely related to elevated temperature, solar radiation, and relative humidity in the PRD, which led to a higher degree of aerosol aging.

The results of correlation analysis are presented in Fig. 4. Pearson's r values between C_2 and SOA tracers (phthalic acid, DHOPA, and malic acid) were relatively higher (r = 0.58, 0.41, and 0.51, respectively; $p < 0.01$), further supporting that C_2 was primarily formed via secondary oxidation processes. Notably, the correlation between C_2 and malic acid became progressively stronger with the reductions in anthropogenic emissions. This trend became more apparent when the data were categorized by pollution levels, with the Pearson's r value increased from 0.33 (IT0) to 0.72 (IT4) (Table 1). Meanwhile, the correlations between C_2 and ASOA markers became weaker. These results suggested that the relative contributions of biogenic

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

sources to SOA become more important under cleaner conditions. Although high Pearson's r values between C_2 and primary anthropogenic source markers were observed in certain individual years, their overall correlations remained weak. Similarly, a marked increase in C_2 during the COVID-19 pandemic (Meng et al., 2023), when anthropogenic emissions were substantially reduced, might also reflect the weak correlation between C_2 and its organic precursors. These findings implied that the influence of anthropogenic VOCs on C_2 formation was limited. Because our measurements were conducted in the same season, the meteorological variations were minimal. As a result, the correlations between C_2 and meteorological parameters, such as temperature, RH, and SR, were generally weak, indicating the meteorological influences on C_2 were also limited.

ALWC not only regulates the gas-particle partitioning of semi-volatile VOCs and their reaction rates by acting as a medium (Nenes et al., 2021), but also serves as a nucleophile that participates in reactive uptake of SOA intermedium (Zhang et al., 2022b). O_x, a proxy of atmospheric oxidants, facilitates secondary photochemical oxidation of VOCs. Therefore, ALWC and O_x play key roles in VOCs oxidation processes in aqueous and gas phase, respectively. In this study, C_2 showed strong correlations with ALWC ($r = 0.50$) and O_x ($r = 0.64$) ($p < 0.01$), suggesting they were the dominant drivers of C_2 variability between 2007 and 2018 (discussed in Sect. 3.3).



https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Figure 4. Correlations between C_2 and primary anthropogenic source markers, SOA markers, O_x, ALWC, pH, as well as meteorological parameters. Blank cells indicate no significant correlations. One asterisk, two asterisks denote p value < 0.05, 0.01, respectively. Due to the unavailability of O_x data in 2012 and 2013, correlation analysis was not conducted for these two years.

Table 1. The Pearson's r between C_2 and phthalic acid, DHOPA, as well as malic acid under different pollution conditions.

Pollution conditions	Phthalic acid	DHOPA	Malic acid
IT0 ($PM_{2.5}$ > 75 μg m^{-3})	0.63 ($p < 0.01$)	0.19 ($p > 0.05$)	0.33 ($p < 0.05$)
IT1 (75 μg m^{-3} > $PM_{2.5}$ > 50 μg m^{-3})	0.28 ($p < 0.01$)	0.49 ($p < 0.01$)	0.53 ($p < 0.01$)
IT2 (50 μg m^{-3} > $PM_{2.5}$ > 37.5 μg m^{-3})	0.44 ($p < 0.01$)	0.45 ($p < 0.01$)	0.66 ($p < 0.01$)
IT3 (37.5 μg m^{-3} > $PM_{2.5}$ > 25 μg m^{-3})	0.34 ($p < 0.01$)	0.42 ($p < 0.01$)	0.69 ($p < 0.01$)
IT4 (25 μg m^{-3} > $PM_{2.5}$)	0.31 ($p < 0.01$)	0.32 ($p < 0.01$)	0.72 ($p < 0.01$)

3.3 Gas-phase and aqueous-phase pathways of oxalic acid formation

Atmospheric oxidation of VOCs, such as aromatics, isoprene, and monoterpenes, with ozone and/or hydroxyl radicals produces semi-volatile gaseous organic precursors (e.g., mGly and Gly). These α-dicarbonyls partition into ALWC-enriched aerosols and the subsequent aqueous-phase reactions produce Pyr and ωC2, which are further oxidized to form C_2 (Ervens et al., 2008; Bikkina et al., 2017). In addition, several studies have also demonstrated that C_2 can be formed through the photochemical degradation of longer-chain DCA (e.g., C_3 and C_4) in the gas phase (Kawamura and Bikkina, 2016; Zhao et al., 2020; Meng et al., 2023). Here, O_x and ALWC are used as indicators of gas-phase and aqueous-phase pathways of C_2 formation, respectively. As mentioned above, they were the dominant drivers of C_2 in this study. Due to the reductions in sulfate and nitrate, the highly hygroscopic compounds in $PM_{2.5}$, ALWC exhibited a fast downward trend between 2007 and 2012 (Table S1, Fig. S4). This hindered the partitioning of semi-volatile gaseous organic precursors from the gas phase into the aqueous phase, thereby slowing the aqueous-phase production of C_2. This resulted in an obvious decline in C_2 concentrations during 2007–2012 (Fig S2). Because ALWC exhibited minor fluctuations between 2013 and 2018, the influence of aqueous-phase pathway on C_2 became weaker.

As presented in Fig. 4, the correlation between C_2 and ALWC weakened, whereas that between C_2 and O_x strengthened from 2007 to 2018. It is worth noting that no significant correlations were observed between C_2 and O_x in 2007 and 2010, when ALWC concentrations were high. In contrast, a strong correlation ($r = 0.62$, $p < 0.01$) emerged in 2008, coinciding with a sharp decrease in ALWC (Fig. S4). This suggested that the gas-phase formation pathway of C_2 was enhanced when ALWC was low. The pattern became more evident with decreasing pollution levels (Fig. 5). From IT0 to IT4, ALWC decreased from 20.9 ± 11.0 μg m^{-3} to 7.2 ± 3.0 μg m^{-3} (Table S4). The Pearson's r values between C_2 and ALWC dropped from 0.43 to 0.15,

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

while that between C_2 and O_x increased from 0.28 to 0.68. This opposite trend suggested a shift in the dominant C_2 formation pathway from aqueous-phase oxidation to gas-phase photochemical oxidation under lower pollution conditions.

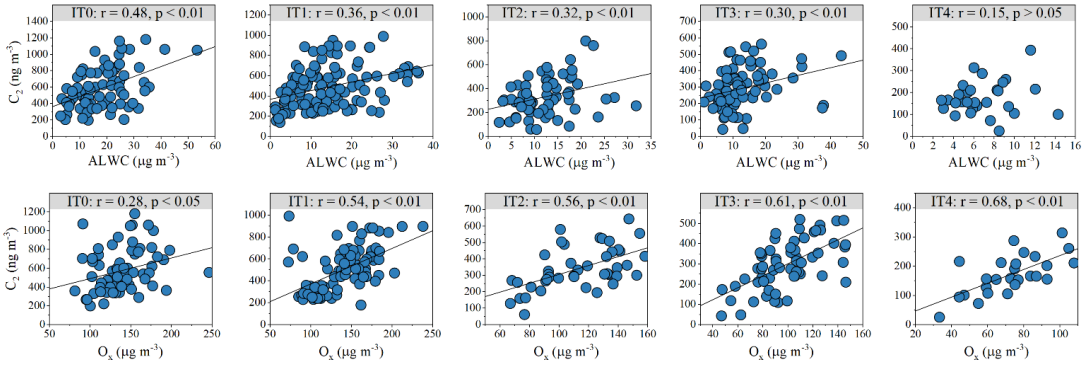


Figure 5. The correlation between C_2 and ALWC, as well as O_x. Pearson's r values between C_2 and ALWC decreased from 0.43 to 0.15, while those between C_2 and O_x increased from 0.28 to 0.68.

3.4 Evaluate the impact of different factors on oxalic acid formation

We identified ALWC and O_x as the dominant drivers for C_2 variation based on annual variations and linear correlation analysis. Although pH has been suggested to play an important role in heterogeneous uptake and aqueous-phase acid-catalyzed reactions (Xu et al., 2016; Cooke et al., 2024), it did not exhibit a strong correlation ($r = -0.24$, $p < 0.01$) with C_2. This was likely due to their non-linear relationship. In addition, correlation analysis is unable to quantify the contributions of each factor to C_2. Hence, we used a machine learning model to further evaluate the impact of individual factors on C_2 variation during our study period. The details about the model used in this study can be seen in Section 2.4 and elsewhere (Chen et al., 2016; Lundberg and Lee, 2017).

The feature importance is presented in Fig. 6a. O_x, ALWC, and pH, which represent secondary oxidation processes, exhibited the three highest |SHAP| values, indicating their dominant contributions to C_2 variation. In contrast, the values for hopanes, levoglucosan, octadecanoic acid, and picene were considerably lower, suggesting that the influence of anthropogenic emissions was relatively minor compared to that of secondary processes. As shown in Fig. 6b, O_x, ALWC, and temperature exhibited positive correlations with their SHAP values, indicating that higher values of these variables contributed to increases in C_2 concentrations. However, pH showed a negative correlation with its SHAP values, suggesting that lower pH levels were associated with higher C_2 concentrations. Notably, the influence of extremely low pH on C_2 formation appeared to be more pronounced.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

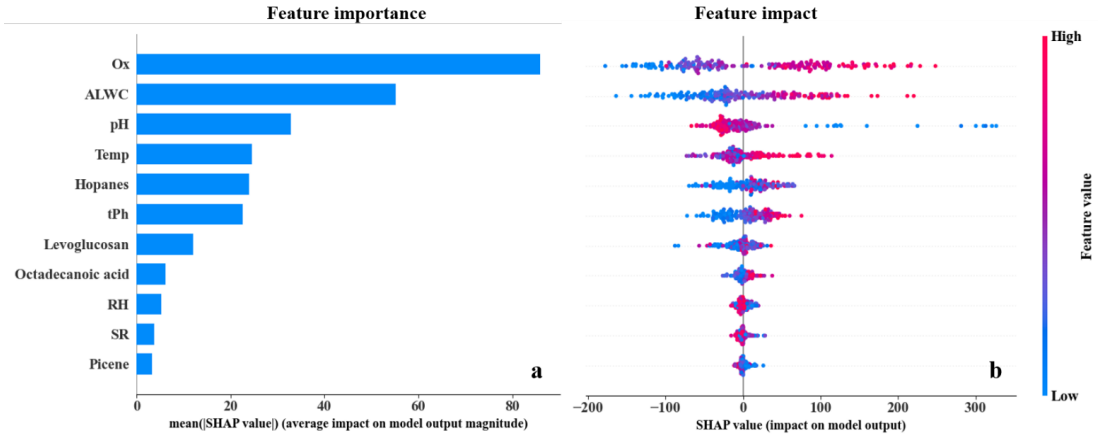


Figure 6. (a) Bar plot of the mean |SHAP| values representing the overall importance of each feature in predicting C_2 concentrations. (b) Beeswarm plot of individual SHAP values for each feature across all samples. Red (blue) represents high (low) value in each feature. Positive (negative) SHAP values indicate that the feature contributes to an increase (decrease) in the C_2 prediction.

To further quantify the impact of all factors on C_2 variation, IF (discussed in Section 2.4) was calculated and presented in Fig. 7. O_x accounted for the highest contribution (32%), followed by ALWC (21%) and pH (12%). All factors were then classified into four groups: (1) The five molecular markers (levoglucosan, hopanes, octadecanoic acid, picene, and tPh) can reflect primary emissions from different anthropogenic sources, they represent anthropogenic precursors emission. (2) Biogenic emissions are highly dependent on meteorological condition (Guenther et al., 1993; Guenther et al., 2006), Temp, SR, and RH represent biogenic precursors emission. (3) O_x can be used as a proxy of atmospheric oxidants, which determine reaction rate in gas phase. It represents gas-phase oxidation processes. (4) ALWC and pH have important impacts on SOA formation in aqueous phase, they represent aqueous-phase oxidation processes (Nguyen et al., 2015; Xu et al., 2016). Although anthropogenic precursors emission showed a clear decreasing trend over the study period, their contributions (24%) to C_2 variation were significantly lower than that of secondary oxidation processes. The contributions of aqueous-phase oxidation processes (33%) and gas-phase oxidation processes (32%) were comparable. They accounted for more than two-thirds of C_2 variation, underscoring the dominant role of secondary oxidation processes in C_2 formation. Since the measurements were primarily conducted in the same season, meteorological parameters remained relatively stable, and thus the contribution of biogenic precursors emission was minor (11%).

From IT0 to IT4, IF values for gas-phase oxidation processes increased from 24% to 48%, whereas those for aqueous-phase oxidation processes decreased from 35% to 20% (Fig. 7b). This indicated that the gas-phase oxidation pathway became increasingly important as pollution levels decreased, which was consistent with the result of correlation analysis. A possible explanation is that under cleaner conditions, lower ALWC levels favored the partitioning of semi-volatile C_2 precursors (e.g., Gly and mGly) from the particle-phase into the gas-phase. Thus, their aqueous-phase reaction pathway was hindered, and more

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

C_2 was formed via photochemical degradation of longer-chain DCA (Kawamura and Bikkina, 2016; Meng et al., 2023). This underscored the growing importance of gas-phase oxidation processes in the formation of C_2 and SOA under cleaner conditions. Although O_x did not exhibit a clear trend at our measurement station, Cao et al. (2024) reported a rapid increase in O_3 concentration across the PRD region over the past decade. This may promote SOA formation through enhanced gas-phase oxidation pathways. Therefore, coordinated control of VOCs and NO_x should be emphasized in the future to reduce ozone pollution (Wang et al., 2021) and further mitigate SOA formation. Under low pollution conditions, the reduction in anthropogenic emissions slowed down. As a result, their contributions to C_2 variation in IT4 (20%) were less than that in IT0 (27%).

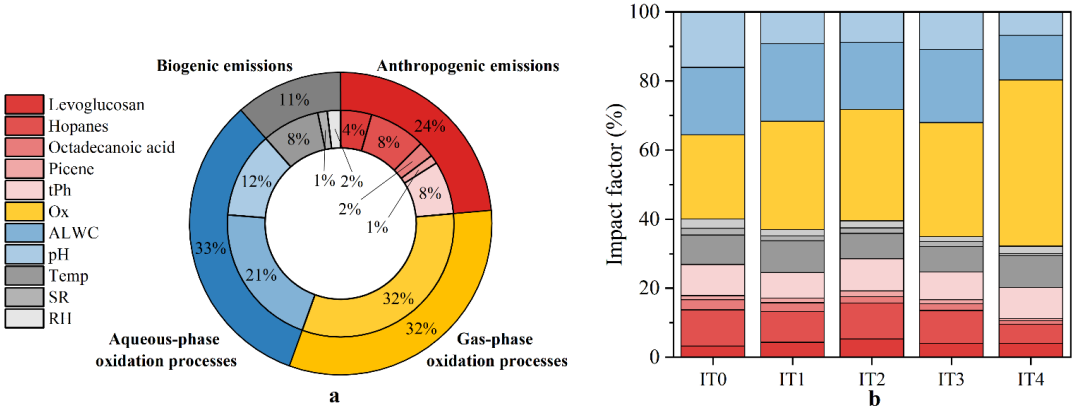


Figure 7. (a) Impact of individual variable on C_2 variation during the whole study period. (b) Impact factor of individual variable under different pollution conditions.

4 Conclusions

In this study, long-term measurements were conducted at a regional background site in the PRD region during 2007–2018. Primary and secondary molecular markers were measured, including five primary anthropogenic source molecular markers (levoglucosan, hopanes, octadecanoic acid, picene, and terephthalic acid), three SOA markers (phthalic acid, DHOPA, and malic acid), and a group of general SOA indicators (aliphatic DCA). C_2, the most abundant compound among aliphatic DCA, was investigated to identify the key drivers for SOA variability. Previous studies have attributed elevated C_2 to higher emissions of its organic precursors (Cao et al., 2017; Ding et al., 2021; Yu et al., 2021) and emphasized the dominance of aqueous-phase pathways in C_2 formation (Yu et al., 2005; Hilario et al., 2021). However, our results showed that reductions in anthropogenic organic precursors had a limited influence on C_2 formation. In addition, we found a growing contribution of gas-phase oxidation to C_2 and SOA formation under cleaner conditions.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Between 2007 and 2018, C_2 and other SOA molecular markers did not show clear trends, despite substantial reductions in anthropogenic emissions, for example, biomass burning (-11% yr^{-1}), vehicle emissions (-17% yr^{-1}), and cooking emissions (-7% yr^{-1}). Correlation analysis revealed that ALWC and O_x were the main drivers for C_2 variation. In addition, the correlation between BSOA tracer and C_2 became stronger from high to low pollution conditions (from IT0 to IT4), underscoring an increasing relative contribution of BSOA. We applied a machine learning model to quantify further the contributions of anthropogenic emissions, aqueous-phase oxidation processes, gas-phase oxidation processes, and meteorological conditions on C_2 variability. The model results showed that the contributions of aqueous-phase oxidation processes and gas-phase oxidation processes were comparable (33% and 32%, respectively) and higher than those of anthropogenic precursors emission (24%) and biogenic precursors emission (11%). With decreasing pollutant levels, the contribution of gas-phase oxidation processes increased from 24% to 48%, whereas that of aqueous-phase oxidation processes declined from 35% to 20%. This probably resulted from the low ALWC-favored partitioning of semi-volatile precursors (e.g., Gly and mGly) from particle-phase to gas-phase, thereby suppressing their aqueous-phase reactions and enhancing the gas-phase photochemical degradation of longer-chain DCA (e.g., C3 and C4).

Our results offer critical insights that can improve the accuracy of atmospheric modeling of C_2 and support the formulation of more targeted pollution mitigation strategies in the future. Meanwhile, we acknowledge there are several limitations. First, our measurements were mainly conducted in wintertime, which may not represent summertime conditions when photochemical activity is higher. Second, hydroxyl radical, an important gaseous oxidant, was not measured in this study. Using O_x as a proxy for gaseous oxidants may introduce uncertainties.

The continuous increase in O_3 over the past decade in the PRD poses a growing challenge to further improvements in particulate pollution. As gas-phase oxidation processes become more important in SOA formation under a cleaner environment, coordinated control of both NO_x and VOCs is essential for mitigating O_3 pollution and achieving further reductions in SOA in the PRD region.

Code and data availability. The experimental data in this study are available upon request to the corresponding author by email.

Author contributions. XD conceived the project and designed the study. Y-FH performed the data analysis and wrote the paper. Y-QZ, MB, D-HC, TZ, KY, J-QW, QC, and HJ arranged the sample collection and assisted with the data analysis. Z-RW and PL analyzed the samples. XD, Q-FH, X-MW, and MB performed the data interpretation and edited the paper. All authors contributed to the development of the final paper.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Financial support. This research was funded by the National Natural Science Foundation of China (42321003/42177090), Guangdong Foundation for Program of Science and Technology Research (2023B0303000007/2023B1212060049/2024A1515011181).

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

References

Bian, Y. H., Huang, Z. J., Ou, J. M., Zhong, Z. M., Xu, Y. Q., Zhang, Z. W., Xiao, X., Ye, X., Wu, Y. Q., Yin, X. H., Li, C., Chen, L. F., Shao, M., and Zheng, J. Y.: Evolution of anthropogenic air pollutant emissions in Guangdong Province, China, from 2006 to 2015, Atmos. Chem. Phys., 19, 11701-11719, https://doi.org/10.5194/acp-19-11701-2019, 2019.

Bikkina, S., Kawamura, K., and Sarin, M.: Secondary Organic Aerosol Formation over Coastal Ocean: Inferences from Atmospheric Water-Soluble Low Molecular Weight Organic Compounds, Environ. Sci. Technol., 51, 4347-4357, https://doi.org/10.1021/acs.est.6b05986, 2017.

Bikkina, S., Kawamura, K., Sakamoto, Y., and Hirokawa, J.: Low molecular weight dicarboxylic acids, oxocarboxylic acids and α-dicarbonyls as ozonolysis products of isoprene: Implication for the gaseous-phase formation of secondary organic aerosols, Sci. Total Environ., 769, 144472, https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.144472, 2021.

Bikkina, S., Kawamura, K., Imanishi, K., Boreddy, S. K. R., and Nojiri, Y.: Seasonal and longitudinal distributions of atmospheric water-soluble dicarboxylic acids, oxocarboxylic acids, and α-dicarbonyls over the North Pacific, J. Geophys. Res.: Atmos., 120, 5191-5213, https://doi.org/10.1002/2014JD022972, 2015.

Bilde, M., Barsanti, K., Booth, M., Cappa, C. D., Donahue, N. M., Emanuelsson, E. U., McFiggans, G., Krieger, U. K., Marcolli, C., Topping, D., Ziemann, P., Barley, M., Clegg, S., Dennis-Smither, B., Hallquist, M., Hallquist, Å. M., Khlystov, A., Kulmala, M., Mogensen, D., Percival, C. J., Pope, F., Reid, J. P., Ribeiro da Silva, M. A. V., Rosenoern, T., Salo, K., Soonsin, V. P., Yli-Juuti, T., Prisle, N. L., Pagels, J., Rarey, J., Zardini, A. A., and Riipinen, I.: Saturation Vapor Pressures and Transition Enthalpies of Low-Volatility Organic Molecules of Atmospheric Relevance: From Dicarboxylic Acids to Complex Mixtures, Chem. Rev., 115, 4115-4156, https://doi.org/10.1021/cr5005502, 2015.

Cao, F., Zhang, S.-C., Kawamura, K., Liu, X., Yang, C., Xu, Z., Fan, M., Zhang, W., Bao, M., Chang, Y., Song, W., Liu, S., Lee, X., Li, J., Zhang, G., and Zhang, Y.-L.: Chemical characteristics of dicarboxylic acids and related organic compounds in PM2.5 during biomass-burning and non-biomass-burning seasons at a rural site of Northeast China, Environ. Pollut., 231, 654-662, https://doi.org/https://doi.org/10.1016/j.envpol.2017.08.045, 2017.

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

Chen, T. Q., Guestrin, C., and Assoc Comp, M.: XGBoost: A Scalable Tree Boosting System, 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, Aug 13-17, WOS:000485529800092, 785-794, 10.1145/2939672.2939785, 2016.

Cheng, Y., Ma, Y., and Hu, D.: Tracer-based source apportioning of atmospheric organic carbon and the influence of anthropogenic emissions on secondary organic aerosol formation in Hong Kong, Atmos. Chem. Phys., 21, 10589-10608, https://doi.org/10.5194/acp-21-10589-2021, 2021.

Cooke, M. E., Armstrong, N. C., Fankhauser, A. M., Chen, Y., Lei, Z., Zhang, Y., Ledsky, I. R., Turpin, B. J., Zhang, Z., Gold, A., McNeill, V. F., Surratt, J. D., and Ault, A. P.: Decreases in Epoxide-Driven Secondary Organic Aerosol Production under Highly Acidic Conditions: The Importance of Acid–Base Equilibria, Environ. Sci. Technol., 58, 10675-10684, https://doi.org/10.1021/acs.est.3c10851, 2024.

Ding, X., He, Q. F., Shen, R. Q., Yu, Q. Q., and Wang, X. M.: Spatial distributions of secondary organic aerosols from isoprene, monoterpenes, beta-caryophyllene, and aromatics over China during summer, J. Geophys. Res.: Atmos., 119, 11877-11891, https://doi.org/10.1002/2014jd021748, 2014.

Ding, X., Wang, X.-M., Gao, B., Fu, X.-X., He, Q.-F., Zhao, X.-Y., Yu, J.-Z., and Zheng, M.: Tracer-based estimation of secondary organic carbon in the Pearl River Delta, south China, J. Geophys. Res.: Atmos., 117, https://doi.org/10.1029/2011JD016596, 2012.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Ding, X., Zhang, Y. Q., He, Q. F., Yu, Q. Q., Wang, J. Q., Shen, R. Q., Song, W., Wang, Y. S., and Wang, X. M.: Significant Increase of Aromatics-Derived Secondary Organic Aerosol during Fall to Winter in China, Environ. Sci. Technol., 51, 7432-7441, https://doi.org/10.1021/acs.est.6b06408, 2017.

Ding, Z., Du, W., Wu, C., Cheng, C., Meng, J., Li, D., Ho, K., Zhang, L., and Wang, G.: Summertime atmospheric dicarboxylic acids and related SOA in the background region of Yangtze River Delta, China: Implications for heterogeneous reaction of oxalic acid with sea salts, Sci. Total Environ., 757, 143741, https://doi.org/10.1016/j.scitotenv.2020.143741, 2021.

Ervens, B., Carlton, A. G., Turpin, B. J., Altieri, K. E., Kreidenweis, S. M., and Feingold, G.: Secondary organic aerosol yields from cloud-processing of isoprene oxidation products, Geophys. Res. Lett., 35, https://doi.org/https://doi.org/10.1029/2007GL031828, 2008.

Geng, G. N., Xiao, Q. Y., Zheng, Y. X., Tong, D., Zhang, Y. X., Zhang, X. Y., Zhang, Q., He, K. B., and Liu, Y.: Impact of China's Air Pollution Prevention and Control Action Plan on PM2.5 chemical composition over eastern China, Science China-Earth Sciences, 62, 1872-1884, https://doi.org/10.1007/s11430-018-9353-x, 2019.

Giannoni, M., Martellini, T., Del Bubba, M., Gambaro, A., Zangrando, R., Chiari, M., Lepri, L., and Cincinelli, A.: The use of levoglucosan for tracing biomass burning in PM2.5 samples in Tuscany (Italy), Environ. Pollut., 167, 7-15, https://doi.org/https://doi.org/10.1016/j.envpol.2012.03.016, 2012.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181-3210, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Guenther, A. B., Zimmerman, P. R., Harley, P. C., Monson, R. K., and Fall, R.: Isoprene and monoterpene emission rate variability: Model evaluations and sensitivity analyses, Journal of Geophysical Research-Atmospheres, 98, 12609-12617, https://doi.org/10.1029/93JD00527, 1993.

Guo, Q., Wang, Y., Zheng, J., Zhu, M., Sha, Q. e., and Huang, Z.: Temporal evolution of speciated volatile organic compound (VOC) emissions from solvent use sources in the Pearl River Delta Region, China (2006–2019), Sci. Total Environ., 933, 172888, https://doi.org/10.1016/j.scitotenv.2024.172888, 2024.

He, L.-Y., Hu, M., Huang, X.-F., Yu, B.-D., Zhang, Y.-H., and Liu, D.-Q.: Measurement of emissions of fine particulate organic matter from Chinese cooking, Atmos. Environ., 38, 6557-6564, https://doi.org/10.1016/j.atmosenv.2004.08.034, 2004.

He, Y., Ding, X., He, Q., Zhang, Y., Chen, D., Zhang, T., Yang, K., Wang, J., Cheng, Q., Jiang, H., Wang, Z., Liu, P., Wang, X., and Boy, M.: Long–term Trends in PM2.5 Chemical Composition and Its Impact on Aerosol Properties: Field Observations from 2007 to 2020 in Pearl River Delta, South China, EGUsphere, 2025, 1-25, https://doi.org/10.5194/egusphere-2025-2204, 2025.

Hilario, M. R. A., Crosbie, E., Bañaga, P. A., Betito, G., Braun, R. A., Cambaliza, M. O., Corral, A. F., Cruz, M. T., Dibb, J. E., Lorenzo, G. R., MacDonald, A. B., Robinson, C. E., Shook, M. A., Simpas, J. B., Stahl, C., Winstead, E., Ziemba, L. D., and Sorooshian, A.: Particulate Oxalate-To-Sulfate Ratio as an Aqueous Processing Marker: Similarity Across Field Campaigns and Limitations, Geophys. Res. Lett., 48, e2021GL096520, https://doi.org/10.1029/2021GL096520, 2021.

Ho, K. F., Cao, J. J., Lee, S. C., Kawamura, K., Zhang, R. J., Chow, J. C., and Watson, J. G.: Dicarboxylic acids, ketocarboxylic acids, and dicarbonyls in the urban atmosphere of China, J. Geophys. Res.: Atmos., 112, https://doi.org/10.1029/2006JD008011, 2007.

Hou, L. L., Dai, Q. L., Song, C. B., Liu, B. W., Guo, F. Z., Dai, T. J., Li, L. X., Liu, B. S., Bi, X. H., Zhang, Y. F., and Feng, Y. C.: Revealing Drivers of Haze Pollution by Explainable Machine Learning, Environ. Sci. Technol. Lett., 9, 112-119, https://doi.org/10.1021/acs.estlett.1c00865, 2022.

Hu, D. and Yu, J. Z.: Secondary organic aerosol tracers and malic acid in Hong Kong: seasonal trends and origins, Environ. Chem., 10, 381-394, https://doi.org/10.1071/en13104, 2013.

Jung, J., Tsatsral, B., Kim, Y. J., and Kawamura, K.: Organic and inorganic aerosol compositions in Ulaanbaatar, Mongolia, during the cold winter of 2007 to 2008: Dicarboxylic acids, ketocarboxylic acids, and α-dicarbonyls, J. Geophys. Res.: Atmos., 115, https://doi.org/10.1029/2010JD014339, 2010.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Kawamura, K. and Bikkina, S.: A review of dicarboxylic acids and related compounds in atmospheric aerosols: Molecular distributions, sources and transformation, Atmos. Res., 170, 140-160, https://doi.org/10.1016/j.atmosres.2015.11.018, 2016.

Kawamura, K. and Ikushima, K.: Seasonal changes in the distribution of dicarboxylic acids in the urban atmosphere, Environ. Sci. Technol., 27, 2227-2235, https://doi.org/10.1021/es00047a033, 1993.

Kawamura, K. and Kaplan, I. R.: Motor exhaust emissions as a primary source for dicarboxylic acids in Los Angeles ambient air, Environ. Sci. Technol., 21, 105-110, https://doi.org/10.1021/es00155a014, 1987.

Kawamura, K. and Sakaguchi, F.: Molecular distributions of water soluble dicarboxylic acids in marine aerosols over the Pacific Ocean including tropics, J. Geophys. Res.: Atmos., 104, 3501-3509, https://doi.org/10.1029/1998JD100041, 1999.

Kawamura, K. and Usukura, K.: Distributions of low molecular weight dicarboxylic acids in the North Pacific aerosol samples, J. Oceanogr., 49, 271-283, https://doi.org/10.1007/BF02269565, 1993.

Kawamura, K. and Watanabe, T.: Determination of Stable Carbon Isotopic Compositions of Low Molecular Weight Dicarboxylic Acids and Ketocarboxylic Acids in Atmospheric Aerosol and Snow Samples, Anal. Chem., 76, 5762-5768, https://doi.org/10.1021/ac049491m, 2004.

Kawamura, K., Sempéré, R., Imai, Y., Fujii, Y., and Hayashi, M.: Water soluble dicarboxylic acids and related compounds in Antarctic aerosols, J. Geophys. Res.: Atmos., 101, 18721-18728, https://doi.org/10.1029/96JD01541, 1996.

Kawamura, K., Okuzawa, K., Aggarwal, S. G., Irie, H., Kanaya, Y., and Wang, Z.: Determination of gaseous and particulate carbonyls (glycolaldehyde, hydroxyacetone, glyoxal, methylglyoxal, nonanal and decanal) in the atmosphere at Mt. Tai, Atmos. Chem. Phys., 13, 5369-5380, https://doi.org/10.5194/acp-13-5369-2013, 2013a.

Kawamura, K., Tachibana, E., Okuzawa, K., Aggarwal, S. G., Kanaya, Y., and Wang, Z. F.: High abundances of water-soluble dicarboxylic acids, ketocarboxylic acids and α-dicarbonyls in the mountaintop aerosols over the North China Plain during wheat burning season, Atmos. Chem. Phys., 13, 8285-8302, https://doi.org/10.5194/acp-13-8285-2013, 2013b.

Kleindienst, T. E., Jaoui, M., Lewandowski, M., Offenberg, J. H., and Docherty, K. S.: The formation of SOA and chemical tracer compounds from the photooxidation of naphthalene and its methyl analogs in the presence and absence of nitrogen oxides, Atmos. Chem. Phys., 12, 8711-8726, https://doi.org/10.5194/acp-12-8711-2012, 2012.

Liu, J., Chen, M., Chu, B., Chen, T., Ma, Q., Wang, Y., Zhang, P., Li, H., Zhao, B., Xie, R., Huang, Q., Wang, S., and He, H.: Assessing the Significance of Regional Transport in Ozone Pollution through Machine Learning: A Case Study of Hainan Island, ACS ES&T Air, https://doi.org/10.1021/acsestair.4c00297, 2025.

Lundberg, S. M. and Lee, S. I.: A Unified Approach to Interpreting Model Predictions, 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, Dec 04-09, WOS:000452649404081, 2017.

Meng, J., Wang, Y., Li, Y., Huang, T., Wang, Z., Wang, Y., Chen, M., Hou, Z., Zhou, H., Lu, K., Kawamura, K., and Fu, P.: Measurement Report: Investigation on the sources and formation processes of dicarboxylic acids and related species in urban aerosols before and during the COVID-19 lockdown in Jinan, East China, Atmos. Chem. Phys., 23, 14481-14503, https://doi.org/10.5194/acp-23-14481-2023, 2023.

Meng, J. J., Wang, G. H., Hou, Z. F., Liu, X. D., Wei, B. J., Wu, C., Cao, C., Wang, J. Y., Li, J. J., Cao, J. J., Zhang, E. X., Dong, J., Liu, J. Z., Ge, S. S., and Xie, Y. N.: Molecular distribution and stable carbon isotopic compositions of dicarboxylic acids and related SOA from biogenic sources in the summertime atmosphere of Mt. Tai in the North China Plain, Atmos. Chem. Phys., 18, 15069-15086, https://doi.org/10.5194/acp-18-15069-2018, 2018.

Nenes, A., Pandis, S. N., Kanakidou, M., Russell, A. G., Song, S., Vasilakos, P., and Weber, R. J.: Aerosol acidity and liquid water content regulate the dry deposition of inorganic reactive nitrogen, Atmos. Chem. Phys., 21, 6023-6033, https://doi.org/10.5194/acp-21-6023-2021, 2021.

Nguyen, T. K. V., Capps, S. L., and Carlton, A. G.: Decreasing Aerosol Water Is Consistent with OC Trends in the Southeast U.S, Environ. Sci. Technol., 49, 7843-7850, https://doi.org/10.1021/acs.est.5b00828, 2015.

Pavuluri, C. M., Wang, S., Fu, P. Q., Zhao, W., Xu, Z., and Liu, C.-Q.: Molecular Distributions of Diacids, Oxoacids, and α-Dicarbonyls in Summer- and Winter-Time Fine Aerosols From Tianjin, North China: Emissions From Combustion Sources and Aqueous Phase Secondary Formation, J. Geophys. Res.: Atmos., 126, e2020JD032961, https://doi.org/10.1029/2020JD032961, 2021.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Peng, X., Xie, T.-T., Tang, M.-X., Cheng, Y., Peng, Y., Wei, F.-H., Cao, L.-M., Yu, K., Du, K., He, L.-Y., and Huang, X.-F.: Critical Role of Secondary Organic Aerosol in Urban Atmospheric Visibility Improvement Identified by Machine Learning, Environ. Sci. Technol. Lett., https://doi.org/10.1021/acs.estlett.3c00084, 2023.

Poore, M. W.: Oxalic Acid in PM2.5 Particulate Matter in California, Journal of the Air & Waste Management Association, 50, 1874-1875, https://doi.org/10.1080/10473289.2000.10464226, 2000.

Riddle, S. G., Robert, M. A., Jakober, C. A., Hannigan, M. P., and Kleeman, M. J.: Size Distribution of Trace Organic Species Emitted from Light-Duty Gasoline Vehicles, Environ. Sci. Technol., 41, 7464-7471, https://doi.org/10.1021/es070153n, 2007.

Rinaldi, M., Decesari, S., Carbone, C., Finessi, E., Fuzzi, S., Ceburnis, D., O'Dowd, C. D., Sciare, J., Burrows, J. P., Vrekoussis, M., Ervens, B., Tsigaridis, K., and Facchini, M. C.: Evidence of a natural marine source of oxalic acid and a possible link to glyoxal, J. Geophys. Res.: Atmos., 116, https://doi.org/https://doi.org/10.1029/2011JD015659, 2011.

Simoneit, B. R. T., Medeiros, P. M., and Didyk, B. M.: Combustion Products of Plastics as Indicators for Refuse Burning in the Atmosphere, Environ. Sci. Technol., 39, 6961-6970, https://doi.org/10.1021/es050767x, 2005.

Wang, G., Kawamura, K., Cheng, C., Li, J., Cao, J., Zhang, R., Zhang, T., Liu, S., and Zhao, Z.: Molecular Distribution and Stable Carbon Isotopic Composition of Dicarboxylic Acids, Ketocarboxylic Acids, and α-Dicarbonyls in Size-Resolved Atmospheric Particles From Xi'an City, China, Environ. Sci. Technol., 46, 4783-4791, https://doi.org/10.1021/es204322c, 2012.

Wang, H., Kawamura, K., and Yamazaki, K.: Water-Soluble dicarboxylic acids, ketoacids and dicarbonyls in the atmospheric aerosols over the southern ocean and western pacific ocean, J. Atmos. Chem., 53, 43-61, https://doi.org/10.1007/s10874-006-1479-4, 2006.

Wang, J., Wang, G., Gao, J., Wang, H., Ren, Y., Li, J., Zhou, B., Wu, C., Zhang, L., Wang, S., and Chai, F.: Concentrations and stable carbon isotope compositions of oxalic acid and related SOA in Beijing before, during, and after the 2014 APEC, Atmos. Chem. Phys., 17, 981-992, https://doi.org/10.5194/acp-17-981-2017, 2017.

Wang, N., Xu, J. W., Pei, C. L., Tang, R., Zhou, D. R., Chen, Y. N., Li, M., Deng, X. J., Deng, T., Huang, X., and Ding, A. J.: Air Quality During COVID-19 Lockdown in the Yangtze River Delta and the Pearl River Delta: Two Different Responsive Mechanisms to Emission Reductions in China, Environ. Sci. Technol., 55, 5721-5730, https://doi.org/10.1021/acs.est.0c08383, 2021.

World Health Organization: . WHO Global Air Quality Guidelines in 2021, https://www.who.int/news-room/questions-and-answers/item/who-global-air-quality-guidelines, last access: 17 June 2024.

Xu, B., Tang, J., Tang, T., Zhao, S., Zhong, G., Zhu, S., Li, J., and Zhang, G.: Fates of secondary organic aerosols in the atmosphere identified from compound-specific dual-carbon isotope analysis of oxalic acid, Atmos. Chem. Phys., 23, 1565-1578, https://doi.org/10.5194/acp-23-1565-2023, 2023.

Xu, B., Zhang, G., Gustafsson, Ö., Kawamura, K., Li, J., Andersson, A., Bikkina, S., Kunwar, B., Pokhrel, A., Zhong, G., Zhao, S., Li, J., Huang, C., Cheng, Z., Zhu, S., Peng, P., and Sheng, G.: Large contribution of fossil-derived components to aqueous secondary organic aerosols in China, Nat. Commun., 13, 5115, https://doi.org/10.1038/s41467-022-32863-3, 2022.

Xu, L., Middlebrook, A. M., Liao, J., de Gouw, J. A., Guo, H., Weber, R. J., Nenes, A., Lopez-Hilfiker, F. D., Lee, B. H., Thornton, J. A., Brock, C. A., Neuman, J. A., Nowak, J. B., Pollack, I. B., Welti, A., Graus, M., Warneke, C., and Ng, N. L.: Enhanced formation of isoprene-derived organic aerosol in sulfur-rich power plant plumes during Southeast Nexus, Journal of Geophysical Research-Atmospheres, 121, 11,137-111,153, https://doi.org/10.1002/2016JD025156, 2016.

Yan, F. H., Chen, W. H., Jia, S. G., Zhong, B. Q., Yang, L. M., Mao, J. Y., Chang, M., Shao, M., Yuan, B., Situ, S., Wang, X. M., Chen, D. H., and Wang, X. M.: Stabilization for the secondary species contribution to PM2.5 in the Pearl River Delta (PRD) over the past decade, China: A meta-analysis, Atmos. Environ., 242, https://doi.org/10.1016/j.atmosenv.2020.117817, 2020.

Yu, J. Z., Huang, X.-F., Xu, J., and Hu, M.: When Aerosol Sulfate Goes Up, So Does Oxalate: Implication for the Formation Mechanisms of Oxalate, Environ. Sci. Technol., 39, 128-133, https://doi.org/10.1021/es049559f, 2005.

https://doi.org/10.5194/egusphere-2025-4624
Preprint. Discussion started: 5 October 2025

Yu, Q., Chen, J., Cheng, S., Qin, W., Zhang, Y., Sun, Y., and Ahmad, M.: Seasonal variation of dicarboxylic acids in PM2.5 in Beijing: Implications for the formation and aging processes of secondary organic aerosols, Sci. Total Environ., 763, 142964, https://doi.org/10.1016/j.scitotenv.2020.142964, 2021.

Yu, Q. Q., Ding, X., He, Q. F., Yang, W. Q., Zhu, M., Li, S., Zhang, R. Q., Shen, R. Q., Zhang, Y. L., Bi, X. H., Wang, Y. S., Ping, P. A., and Wang, X. M.: Nationwide increase of polycyclic aromatic hydrocarbons in ultrafine particles during winter over China revealed by size-segregated measurements, Atmos. Chem. Phys., 20, 14581-14595, https://doi.org/10.5194/acp-20-14581-2020, 2020.

Zhang, N., Fan, F., Feng, Y., Hu, M., Fu, Q., Chen, J., Wang, S., and Feng, J.: Insight into Source and Evolution of Oxalic Acid: Characterization of Particulate Organic Diacids in a Mega-City, Shanghai from 2008 to 2020, Atmosphere, 13, 1347, https://doi.org/doi:10.3390/atmos13091347, 2022a.

Zhang, Q., Zheng, Y. X., Tong, D., Shao, M., Wang, S. X., Zhang, Y. H., Xu, X. D., Wang, J. N., He, H., Liu, W. Q., Ding, Y. H., Lei, Y., Li, J. H., Wang, Z. F., Zhang, X. Y., Wang, Y. S., Cheng, J., Liu, Y., Shi, Q. R., Yan, L., Geng, G. N., Hong, C. P., Li, M., Liu, F., Zheng, B., Cao, J. J., Ding, A. J., Gao, J., Fu, Q. Y., Huo, J. T., Liu, B. X., Liu, Z. R., Yang, F. M., He, K. B., and Hao, J. M.: Drivers of improved PM2.5 air quality in China from 2013 to 2017, PNAS, 116, 24463-24469, https://doi.org/10.1073/pnas.1907956116, 2019.

Zhang, Y.-Q., Ding, X., He, Q.-F., Wen, T.-X., Wang, J.-Q., Yang, K., Jiang, H., Cheng, Q., Liu, P., Wang, Z.-R., He, Y.-F., Hu, W.-W., Wang, Q.-Y., Xin, J.-Y., Wang, Y.-S., and Wang, X.-M.: Observational Insights into Isoprene Secondary Organic Aerosol Formation through the Epoxide Pathway at Three Urban Sites from Northern to Southern China, Environ. Sci. Technol., 56, 4795-4805, https://doi.org/10.1021/acs.est.1c06974, 2022b.

Zhao, W., Kawamura, K., Yue, S., Wei, L., Ren, H., Yan, Y., Kang, M., Li, L., Ren, L., Lai, S., Li, J., Sun, Y., Wang, Z., and Fu, P.: Molecular distribution and compound-specific stable carbon isotopic composition of dicarboxylic acids, oxocarboxylic acids and α-dicarbonyls in PM2.5 from Beijing, China, Atmos. Chem. Phys., 18, 2749-2767, https://doi.org/10.5194/acp-18-2749-2018, 2018.

Zhao, W., Ren, H., Kawamura, K., Du, H., Chen, X., Yue, S., Xie, Q., Wei, L., Li, P., Zeng, X., Kong, S., Sun, Y., Wang, Z., and Fu, P.: Vertical distribution of particle-phase dicarboxylic acids, oxoacids and α-dicarbonyls in the urban boundary layer based on the 325 m tower in Beijing, Atmos. Chem. Phys., 20, 10331-10350, https://doi.org/10.5194/acp-20-10331-2020, 2020.

Zhao, Y., Hu, M., Slanina, S., and Zhang, Y.: Chemical Compositions of Fine Particulate Organic Matter Emitted from Chinese Cooking, Environ. Sci. Technol., 41, 99-105, https://doi.org/10.1021/es0614518, 2007.