University of Birmingham
Research at Birmingham

Implications for ozone control by understanding the survivor bias in observed ozone-volatile organic compounds system

Wang, Zhenyu; Shi, Zongbo; Wang, Feng; Liang, Weiqing; Shi, Guoliang; Wang, Weichao; Chen, Da; Liang, Danni; Feng, Yinchang; Russell, Armistead G.

DOI:
10.1038/s41612-022-00261-7

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Wang, Z, Shi, Z, Wang, F, Liang, W, Shi, G, Wang, W, Chen, D, Liang, D, Feng, Y & Russell, AG 2022, 'Implications for ozone control by understanding the survivor bias in observed ozone-volatile organic compounds system', *npj Climate and Atmospheric Science*, vol. 5, no. 1, 39. https://doi.org/10.1038/s41612-022-00261-7

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 20. Apr. 2024

npj | climate and atmospheric science

www.nature.com/npjclimatsci

ARTICLE OPEN

Implications for ozone control by understanding the survivor bias in observed ozone-volatile organic compounds system

Zhenyu Wang[1,2], Zongbo Shi[3], Feng Wang[1,2], Weiqing Liang[1,2], Guoliang Shi[1,2 ✉], Weichao Wang[4], Da Chen[5], Danni Liang[1,2], Yinchang Feng[1,2] and Armistead G. Russell[6]

Tropospheric ozone (O_3) affects Earth's climate and human health. Volatile organic compounds (VOCs), major contributors to O_3 formation, are of particular interest. Generally, the measured concentrations of VOCs (M-VOCs) and O_3 show nonlinear or even opposite time serial-trend. We attributed the phenomenon to survivor bias: lack of insight of the photochemically consumed VOCs (C-VOCs) which emitted from sources to ambient and devote to forming O_3, while excessive concern on the measured VOCs (M-VOCs) at observation site. Both observational and model results provide evidence that C-VOCs are the key to O_3 formation. We proposed an improved model to quantify the source contributions of C-VOCs (biogenic emissions, gasoline evaporation, industry, etc.) and their impacts on the formation of O_3, successfully avoiding the misidentification of dominant VOCs sources originated from the survivor bias in observational data. The survivor bias found in this study highlights that focusing of M-VOCs directly is insufficient and demonstrates the necessity of capture the sources of C-VOCs which contribute to O_3 formation.

npj Climate and Atmospheric Science (2022)5:39 ; https://doi.org/10.1038/s41612-022-00261-7

INTRODUCTION

Tropospheric ozone (O_3) is an important ambient air pollutant that poses serious public health risks[1]. Epidemiological studies have shown a clear link between high ambient O_3 concentrations and the risks of respiratory and cardiovascular mortality[2]. Severe O_3 pollution has been alleviated in many cities of developed countries, but it has gradually turned into a hot spot in South and East Asia due to rapid industrialization and urbanization[3–5]. More recently, summertime O_3 pollution has become an emerging concern in China[6,7].

Globally, continuous efforts are being made to capture the dynamic pollution of tropospheric O_3 and its causes, sources, and impacts. As a secondary pollutant, O_3 in the troposphere can accumulate due to the photochemical oxidation of volatile organic compounds (VOCs) reacting in the presence of nitrogen oxides (NO_x)[8], which has been proved expensive to control[9]. Rather, as crucial precursors for sustained formation of O_3 and active participant in the cycling of radicals, VOCs are directly emitted by anthropogenic and biogenic sources, and have been a focus of air quality management[7,10,11]. Studies have shown that although high VOCs concentrations could be more conducive to O_3 accumulation[12,13], in fact, O_3 formation rate has a nonlinear response to the (relative) changes in observed VOCs concentrations[14,15]. In this work, the concentration of VOCs and O_3 even exhibited an opposite trend in time series (Fig. 1). This phenomenon has also been found in previous studies. For example, Zou et al. manifested that the seasonal variation in the concentration of VOCs was opposite to that of O_3 concentration in Guangzhou[16], and in Wuhan, the seasonal variation of VOCs concentration was negatively correlated with O_3 concentration[17]. However, most studies focus more on the description of observational data, and the actual reasons behind the phenomenon remain unclear which require further research. The effective prevention and control of O_3 pollution therefore requires the reasons for this nonlinear relationship, as well as a detailed description of the sources and fates of VOCs.

The key to impeding the development of this problem is that some VOCs species readily react with atmospheric oxidants such as hydroxyl radicals (OH), nitrate radicals (NO_3), chlorine atoms, and O_3[10,17,18]. But which species is more reactive remains a matter of debate[19–21]. Complex chemical reactions lead to highly reactive VOCs species being consumed to varying degrees from sources to the observation sites. If we focus on the high concentration of observational VOCs species, can it play a vital role in O_3 prevention and control? Undoubtedly, such photochemical consumptions of different species add difficulties to reveal the real sources of VOCs and their impacts on O_3 formation.

Quantifying the photochemical consumption of VOCs and their sources profoundly is the key to capture the real source of VOCs' effect on O_3. However, previous studies could not directly provide the sources of consumed VOCs (C-VOCs)[21–23] which reduced the intuitive perception of VOCs contribution to O_3. Recently, the global VOCs online monitoring sites are gradually increasing, which can obtain high time resolution VOCs data. This provides a key data base and advantages for VOCs traceability. In this work, we demonstrated that C-VOCs are the direct causes of the nonlinear relationship between O_3 and VOCs, and it is the C-VOCs that control O_3 formation, by using chemical kinetics, environmental models, and theoretical calculations. Therefore, we developed an improved quantitative method of the source impacts of C-VOCs based on the observation VOCs dataset, which will help us to develop more effective strategies to control atmospheric O_3 pollution. At the same time, we also quantified the

---

[1] State Environmental Protection Key Laboratory of Urban Ambient Air Particulate Matter Pollution Prevention and Control, Tianjin Key Laboratory of Urban Transport Emission Research, College of Environmental Science and Engineering, Nankai University, Tianjin 300350, P.R. China. [2] CMA-NKU Cooperative Laboratory for Atmospheric Environment-Health Research (CLAER), College of Environmental Science and Engineering, Nankai University, Tianjin 300350, China. [3] School of Geography Earth and Environment Sciences, University of Birmingham, Birmingham B15 2TT, UK. [4] Department of Electronics and Tianjin Key Laboratory of Photo-Electronic Thin Film Device and Technology, Nankai University, Tianjin 300071, China. [5] Key Laboratory of Civil Aviation Thermal Hazards Prevention and Emergency Response, Civil Aviation University of China, Tianjin 300300, China. [6] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0512, USA. ✉email: nksgl@nankai.edu.cn

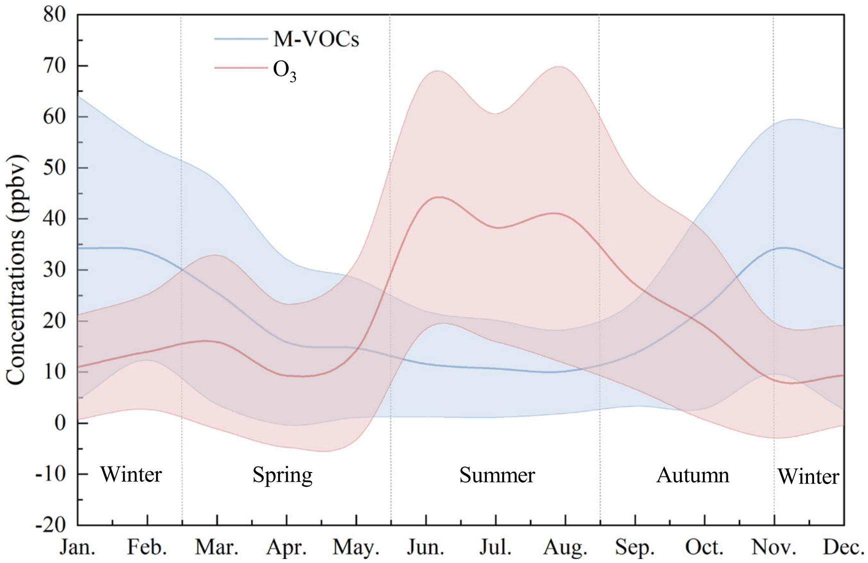

Fig. 1 The concentrations of measured VOCs (M-VOCs) and O_3 from January to December in 2018. Bold solid lines and shaded sections show the observed mean and standard deviation, respectively. From summer to winter, the variations of VOCs and O_3 show clear opposite trends. Annual mean concentration of M-VOCs was about 21.4 ± 26.8 ppbv, varying between 10.8 ± 9.3 ppbv in summer and 32.6 ± 26.1 ppbv in winter. The annual mean concentration of O_3 was about 20.4 ± 0.1 ppbv, ranging from 11.4 ± 10.4 ppbv (winter) to 40.7 ± 25.3 ppbv (summer). The hourly concentrations of measured VOCs, four VOCs categories and O_3 are shown in Supplementary Fig. 1.

source impacts of C-VOCs that contribute to O_3 production. This new method for source apportionment of C-VOCs will be a useful tool to better understand the VOCs sources and their impacts on O_3 formation, and thus aiding in developing O_3 pollution control strategies in different locations.

RESULTS

Opposite trends between measured VOCs and O_3

The temporal concentration variations of measured VOCs (M-VOCs, specific species in Supplementary Table 1 and Supplementary Fig. 1) and O_3 based on hourly online observations conducted in Tianjin during 2018 are illustrated in Fig. 1. The sampling site location is in a community located about 200 meters away from a major road, and more information is presented in the Methods. Interestingly, M-VOCs and O_3 exhibited an opposite trend during different seasons. M-VOCs concentrations were much lower in summer (10.8 ± 9.3 ppbv) than in winter (32.6 ± 26.1 ppbv). By contrast, seasonal O_3 concentrations variations were contrary to M-VOCs, which were high in summer (40.7 ± 25.3 ppbv) and low in winter (11.4 ± 10.4 ppbv). This inverse correlation between M-VOCs and O_3 agrees with earlier findings[16,24]. Generally, high VOCs may facilitate the production of O_3, but here low M-VOCs concentrations were accompanied by high O_3 concentrations. Studies have shown that the O_3 formation regime depends on the ratio of NO_x/VOCs[25–27]. In Supplementary Note 1, the O_3-NO_x-VOCs sensitivity is discussed by using the empirical kinetic modeling approach and NO_x/VOCs ratio. The simulated O_3 showed a clear dependence on VOCs, indicating that O_3 production at our site was in the VOCs-limited regime (Supplementary Fig. 2), which is consistent with previous studies[28,29]. Thus, this paper mainly discusses the relationship between VOCs and O_3. To explain this phenomenon, we need an in-depth understanding of the characteristics and reaction processes of VOCs, as well as the impact of VOCs consumption on O_3 formation. As key precursors of O_3, the photochemical reactions of VOCs are crucial in the formation of ambient O_3[30].

Theoretically, the generation, transport and diffusion of pollutants are jointly affected by physical and chemical processes, which eventually result in the removal or transformation of pollutants from the atmosphere[31,32]. For example, changes in boundary layer evolution can affect the diffusion dilution of pollutants in the vertical direction. Here, the monthly variation concentration of M-VOCs is negatively correlated with boundary layer evolution, as well as O_3 and the evolution of the boundary layer (correlation coefficients (r) are 0.06 and −0.12, see Supplementary Fig. 1c). But it is recognized that the physical dilution is far less than the chemical reaction[10]. For this reason, in this study, we mainly focus on consumed VOCs (C-VOCs) caused by the VOCs photochemical reaction to produce O_3 pathway. OH radical is the dominant atmospheric oxidant that reacts with VOCs to generate O_3 (e.g., substitution and addition reaction, details are shown in Supplementary Note 2)[10,33], and plays a leading role in the consumption of VOCs. This makes it reasonable to hypothesize that the observed inverse correlations between M-VOCs and O_3 (Fig. 1) can be adequately explained by OH consuming VOCs to form O_3, as described in Eq. (1):

$$\mathrm{VOCs} + \mathrm{OH} \rightarrow \mathrm{RO_2} + \mathrm{HO_2} \tag{1}$$

Affected by environmental factors such as solar radiation and ambient temperature (T), the reaction rates of various VOCs species with OH radicals are drastically different[34]. In addition, as photochemical age increased, the oxidation process of VOCs may be conducive to accelerating O_3 accumulation[13,34]. These variables are frequently observed to be correlated with the photochemical loss process of VOCs, and can be viewed as powerful indicators of chemical kinetics. Therefore, the C-VOCs are not only related to the reactant (OH), but also inseparable from other parameters of chemical kinetics.

Survivor bias in O_3-VOCs system

To understand the opposite trend between O_3 and M-VOCs, we should deeply analyze the relationship among C-VOCs, M-VOCs and O_3. In this study, VOCs are classified into three categories to better distinguish different types of VOCs reactions: initial VOCs (In-VOCs, fresh VOCs emitted from sources without chemical consumption), C-VOCs and M-VOCs. Detailed definitions of the three are presented in Methods. Then, (1) starting from the perspective of chemical kinetics, a method for estimating C-VOCs

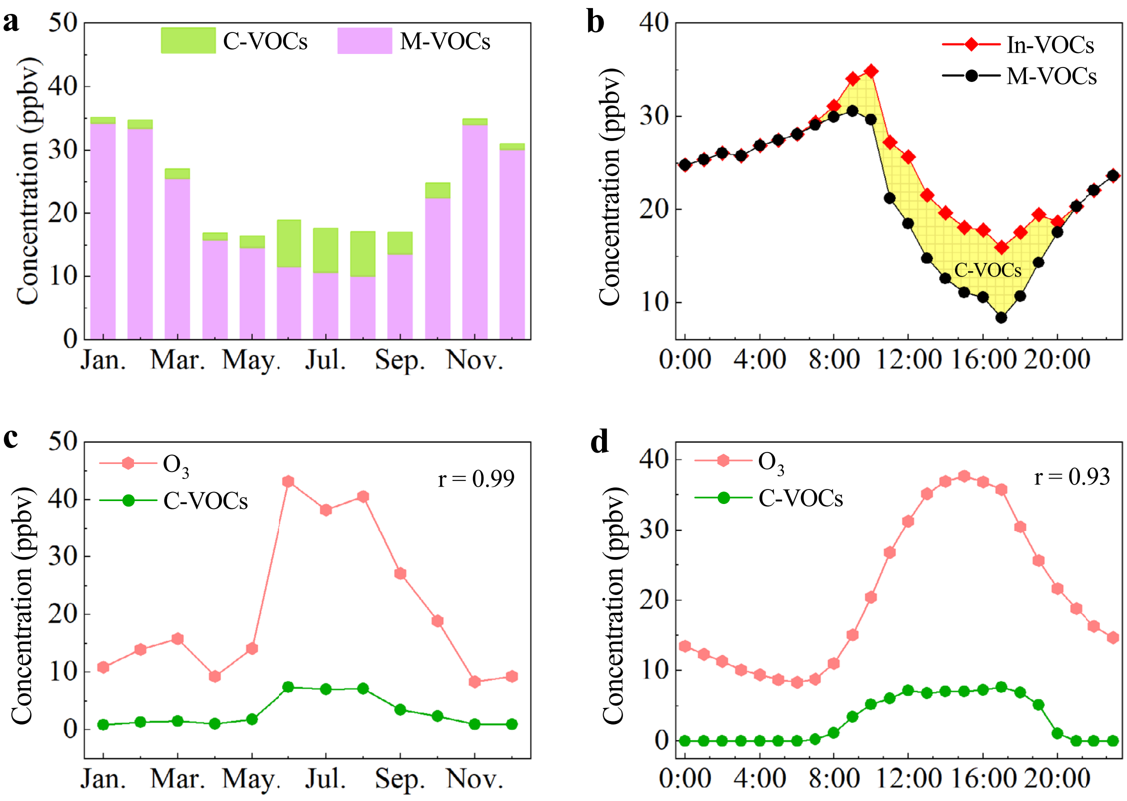

Fig. 2 Variation trends of VOCs concentrations in different reaction types and the correlation between O_3 and consumed VOCs (C-VOCs). Monthly variations of **a** C-VOCs and measured VOCs (M-VOCs) (average annual concentration was 21.4 ppbv), the sum of which were initial VOCs (In-VOCs) (average annual concentration was 24.3 ppbv); **b** average diurnal variation of In-VOCs and M-VOCs; **c** monthly variation of C-VOCs and O_3; and **d** diurnal trends of C-VOCs and O_3. **a**, **b** show that C-VOCs are consumed VOCs and M-VOCs are the survivable VOCs after consumption. In **c**, **d**, the high time-series consistency between C-VOCs and O_3 confirms the key role of C-VOCs in O_3 formation.

was established by theoretical calculations involving reaction rate constant (*k*), concentration of OH radical and the photochemical consumption time (Δt, details in Methods) on VOCs loss in the atmosphere; (2) subsequently, the relationship between C-VOCs and O_3 was analyzed to evaluate the contributions of C-VOCs on O_3 formation potentials.

We applied kinetic equations (Eqs. (3)–(6) in Methods) to calculate VOCs consumption. For accuracy, k_{OH} for different species is calculated by Eqs. (5) and (6)[10,35] based on real atmospheric conditions. Furthermore, a series of simulations using a box model (a model widely used in radical calculation[36,37]), Framework for 0-D Atmospheric Modeling[38], were performed to estimate the possible levels of OH, while observations and other methods were used to verify the reliability of the results (Supplementary Note 3.1). We also estimated the Δt for all detected VOCs species by applying a new method (more details are available in Supplementary Note 3.2) which takes into account both transport, chemical reaction processes and daytime duration (Supplementary Table 2, based on solar radiation in Supplementary Fig. 3). Calculation results in Supplementary Table 3 indicated that the photochemical age (t_a) of any given species varied among different months. Based on the results, Δt of each species (Supplementary Table 4) was estimated. As shown in Supplementary Fig. 4, OH radical was strongly influenced by light intensity, and its concentration is higher in summer and lower in winter (Supplementary Table 5). Such seasonal trend is in contrary to that of M-VOCs. These results suggest that photochemical aging is mainly dominated by the reaction with OH, which is a key factor leading to the evolution of VOCs and its contribution to O_3, especially in summertime.

After calculation, the annual average concentration of C-VOCs was 2.9 ppbv. From Fig. 2a, C-VOCs accounted for a significant fraction of In-VOCs, and exhibited an apparent seasonal variation: the average level of C-VOCs peaked at 7.4 ppbv in June and was lowest in January at 0.9 ppbv. Based on the observed concentration variation of C-VOCs, M-VOCs and In-VOCs, we conclude that source emissions would be greatly underestimated if substantial chemical losses were ignored, especially in summer and autumn (from June to September). During the day, with the increase in solar radiation and *T*, photochemical reactions were enhanced[13], C-VOCs (the gap between In-VOCs and M-VOCs) rose gradually, reaching their peak of 7.6 ppbv at 17:00 and then slowly decrease afterward (Fig. 2b). From 11:00 to 18:00, the concentrations of In-VOCs were 1.3–1.9 times higher than M-VOCs, which due primarily to active photochemical reactions in daytime. During night (21:00–6:00), C-VOCs dropped to zero, which is due to the slow reaction rate with NO_3 and Cl (several orders of magnitude lower than OH)[10]. Therefore, the estimation of C-VOCs at nighttime had been evaluated as negligible.

We released that C-VOCs were significantly correlated with O_3 in both monthly and diurnal variations (correlation coefficients were 0.99 and 0.93 in Fig. 2c, d) for the whole dataset, while M-VOCs were negatively related to O_3 (Supplementary Fig. 5). The excellent agreement between C-VOCs and O_3 suggests that, in ambient air, M-VOCs are residues of photochemical reactions rather than participants in O_3 formation. In addition, we estimated the ozone formation potentials (OFP) of C-VOCs ($OFP_{C\text{-}VOCs}$) according to Eq. (10). The results showed that correlation coefficient between daily or monthly average $OFP_{C\text{-}VOCs\ (total)}$ and O_3 concentration was 0.57 and 0.98 respectively, which far exceeded the correlation between $OFP_{M\text{-}VOCs}$ (calculated based on

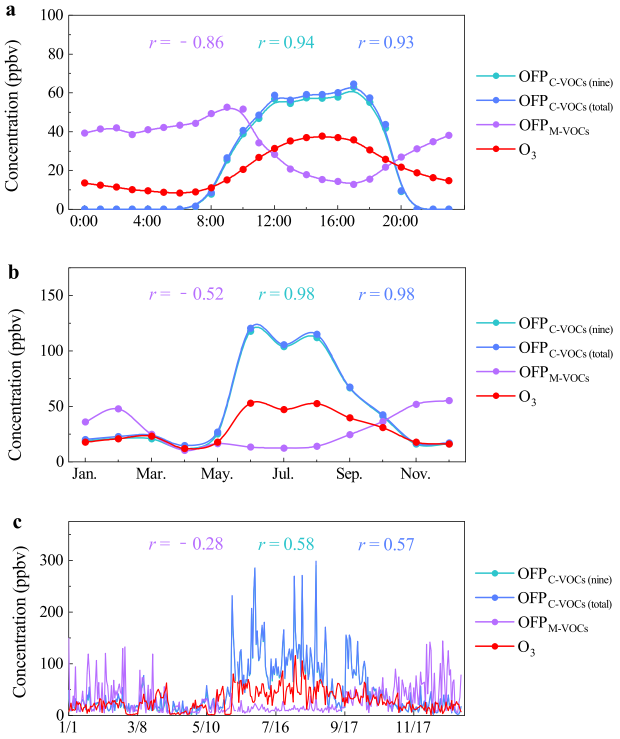

Fig. 3 The correlation of $\mathrm{OFP}_{\mathrm{C\text{-}VOCs\ (nine)}}$, $\mathrm{OFP}_{\mathrm{C\text{-}VOCs\ (total)}}$, $\mathrm{OFP}_{\mathrm{M\text{-}VOCs}}$ and O_3. Time series of **a** diurnal variation, **b** monthly concentration, and **c** daily concentration of ozone formation potential (OFP) and O_3. Note that the $\mathrm{OFP}_{\mathrm{C\text{-}VOCs\ (nine)}}$ and $\mathrm{OFP}_{\mathrm{C\text{-}VOCs\ (total)}}$ (OFP calculated based on the top nine or all C-VOCs) are higher than the observed O_3 especially in the summer. This may be because $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ calculates maximum ozone formation potential, though O_3 consumption during the summer is substantial. Or due to the uncertainty of maximum incremental reactivity (MIR) when applied to the airshed conditions in 2018 in Tianjin, China. And the influence of physical processes (boundary layer height, etc.). Here, we mainly focus on the change trend of OFP and O_3. The correlation between $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ and O_3 is obviously higher than that between $\mathrm{OFP}_{\mathrm{M\text{-}VOCs}}$ and O_3, which further illustrates the important role of C-VOCs in O_3 formation.

M-VOCs) and O_3 (Fig. 3). Thus, it can be further indicated that C-VOCs, instead of M-VOCs, are shown to significantly affect O_3 levels. In order to understand the impacts of regional transport, we also simulated in situ O_3 production using a box model[36,37]. The good correlation of in situ O_3 production with C-VOCs as well as $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ ($r = 0.68$ and 0.66) suggests that it is appropriate to use C-VOCs to evaluate their key role in O_3 production (Supplementary Fig. 5). C-VOCs are largely consumed (from source to ambient observation site) to produce high O_3, resulting in low survivable level (M-VOCs) at observation site in summer, which makes high-O_3 and low-VOCs (M-VOCs) were observed in summer, and oppositely in winter. Through the above analysis, we have concluded that C-VOCs explain the opposite seasonal trend between M-VOCs and O_3, and are also likely to be the direct cause of O_3 generation in ambient atmosphere. In this regard, we interpret this fact as survivor bias in O_3-VOCs system.

Role of C-VOCs on ozone formation

We further investigated the impact of C-VOCs on O_3 formation. Considering that different C-VOCs species have wildly varying molecular structure and chemical reaction activity, their impacts on O_3 generation are expected to be highly uneven. It is necessary to further investigate which are the key species in C-VOCs that contribute significantly to O_3 formation potentials, as well as the overall effects of C-VOCs on O_3 in the transport process from source to receptor. The relative contribution of the four categories of VOCs showed different levels of temporal variations in daytime (Supplementary Fig. 6). Alkenes were the dominant C-VOCs, contributing to ~85.2% of total C-VOCs. Such results are consistent with previous studies[21,23]. The remaining C-VOCs were alkanes, aromatics and alkynes, accounting for 7.8%, 6.5% and 0.5%, respectively. Then, Eq. (10) was used to evaluate key reactive components of atmospheric C-VOCs during the study period.

The calculated $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ of all detected VOCs species are presented in Fig. 4a. In addition, the results are also validated by using L_{OH} method (OH radical loss rate, see Supplementary Note 4 and Supplementary Fig. 7 for details), which provides a simple indicator for studying the relative contribution of specific VOCs to daytime photochemical reactions. In terms of C-VOCs, nine species (isoprene, cis-2-butene, trans-2-butene, propylene, ethylene, 1-butene, m/p-xylene, cis-2-pentene, and styrene) were identified as the key reactive species (Fig. 4b), which contributed to 96% of

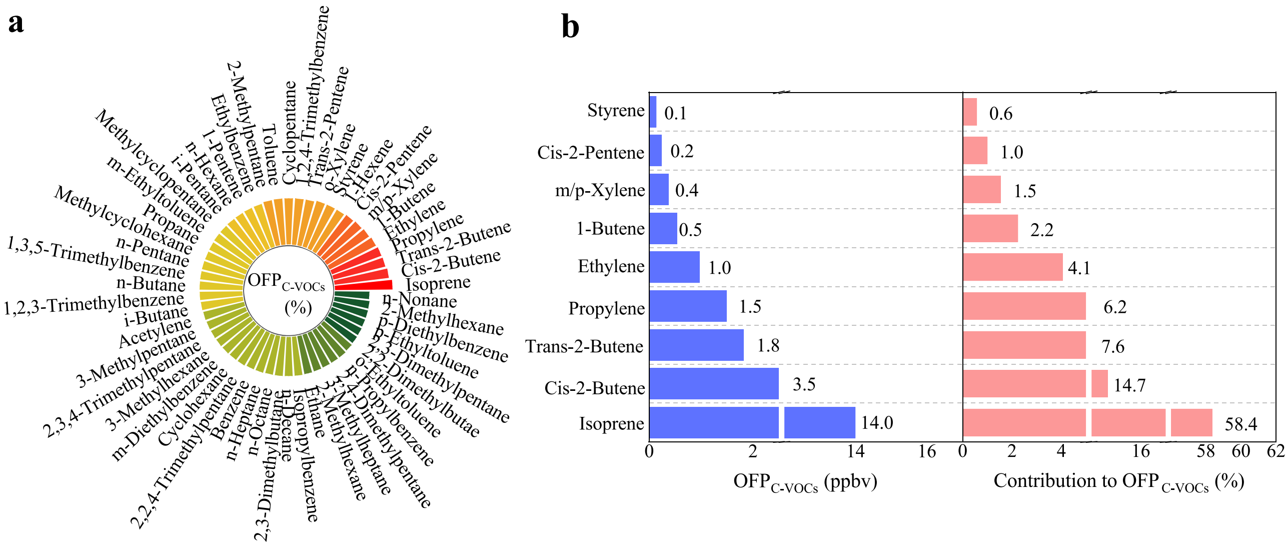

Fig. 4 $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ of different species were calculated. a $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ (OFP calculated based on C-VOCs) importance ranking of all volatile organic compounds; **b** $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$ of the top nine reactive species and their percentage contributions to $\mathrm{OFP}_{\mathrm{C\text{-}VOCs}}$. This figure illustrates that reactive VOCs species should be valued for their participation in chemical reactions.

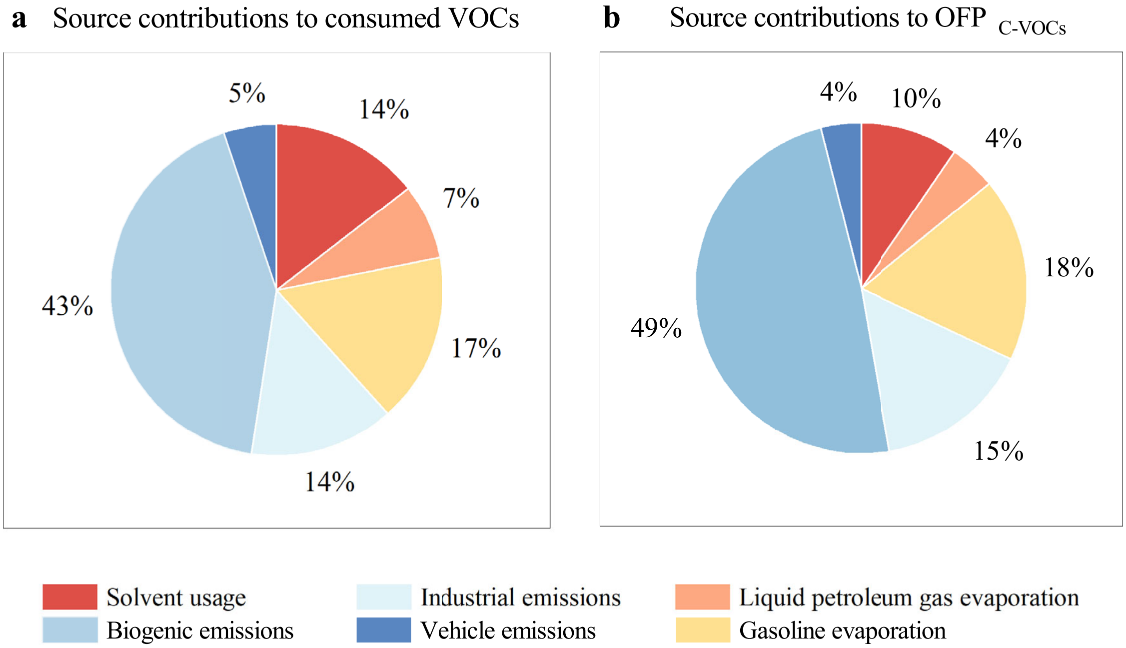

Fig. 5 **The main source contributions to VOCs and their potential impacts on O_3 formation.** **a** Source apportionment to consumed VOCs (C-VOCs); and **b** contributions of different pollution sources to $OFP_{C\text{-}VOCs}$ (based on nine reactive species). The highest contributor to C-VOCs is different from the highest contributors to M-VOCs and In-VOCs (Supplementary Fig. 8). This figure shows that focusing on the sources of C-VOCs can capture the real source of VOCs' effect on O_3.

total $OFP_{C\text{-}VOCs}$. These reactive species are important precursors that should be the focus for controlling atmospheric O_3 pollution. On the other hand, for the above mentioned nine species, their $OFP_{C\text{-}VOCs}$ showed a significant correlation with O_3 (Fig. 3), suggesting that the depletion (reaction) of key species in VOCs plays an important role in the formation of O_3. Therefore, for O_3 control, more attention needs to be paid to the emission sources of these key C-VOCs species.

C-VOCs sources play roles in O_3 formation

Given the chemical loss of VOCs, estimating their source contribution still has some challenges. To intuitively understand the actual source impacts of C-VOCs on O_3 formation, we proposed a new hybrid chemical kinetics and positive matrix factorization (CK-PMF) model (see Methods). The method is executed in two parts to account for sources that slowly release VOCs or far away from observation sites. This method is not designed for source appointment at receptors located close to sources, in which case VOCs may be transported to the receptor location with negligible loss (detailed description is presented in Methods).

We identified six sources of C-VOCs and their contributions (Fig. 5a): biogenic emissions (BE, 43%), gasoline evaporation (GE, 17%), industrial emissions (IE, 14%), solvent usage (SU, 14%), liquid petroleum gas evaporation (LPG, 7%) and vehicle emissions (VE, 5%). BE was the largest source of C-VOCs, which is consistent with the fact that the most reactive species, isoprene, is used as the marker species for this source. GE was greatly affected by VOCs photochemical consumption, while VE had the least impact, which was also related to the amount of reactive species released by the source. Besides, we also calculated the source contributions of M-VOCs and In-VOCs (Supplementary Fig. 8) for comparison. It should be noticed that the highest contributor to C-VOCs (BE) is different from the highest contributors to M-VOCs (GE, 28%) and In-VOCs (IE, 35%) (Supplementary Fig. 8), indicating that if we use VOCs data at diverse reaction types for source apportionment, various results of source contributions can be obtained. And as discussed above, C-VOCs play a vital role in O_3 formation, thus, their sources should be the real sources that contribute to O_3 formation. Additionally, for an individual key reactive species, the sources that contribute significantly to consumed concentration are summarized in Supplementary Fig. 9. In this work, isoprene was almost entirely from BE (99%). This finding is consistent with the fact that vegetation is abundant around the observation site. Ethylene mainly comes from IE (84%) with minor contribution from VE (7%); styrene is abundant in SU (64%) and IE (18%). Therefore, high reactive C-VOCs species (such as isoprene, ethylene, styrene, etc.) should be more focused to explore their real VOCs sources which play an important role in O_3 formation. However, the measured concentrations of these species are not very high in all M-VOCs. If we directly use the observed survivable VOCs data (M-VOCs) to track the sources of VOCs, it will cause a certain degree of deviation (survivor bias).

We further calculated the impact of C-VOCs sources on O_3 formation based upon C-VOCs source apportionment results and $OFP_{C\text{-}VOCs}$. Figure 5b showed that BE (49%) contributed the most to $OFP_{C\text{-}VOCs}$, followed by GE (18%), IE (15%), SU (10%), LPG (4%) and VE (4%). We note that among the contributions of C-VOCs and $OFP_{C\text{-}VOCs}$, the contribution of BE is the most significant (although its contribution to In-VOCs is small). This result indicates that VOCs emitted from vegetation significantly impact on O_3 formation, which is consistent with previous views[39–42]. But more importantly, we further quantified the source contribution of VOCs emitted by vegetation to O_3 pollution. Isoprene is a key species of BE, so changing its other conversion pathways (such as inhibiting the activity of isoprene synthase[43]) may provide new insights for future O_3 control.

Therefore, we suggest that both source emissions, key species and the chemical activity of VOCs should be taken into account in investigating the contribution of VOCs to O_3 formation. Supplementary Fig. 10 explores the diel variations of the concentration of different sources to C-VOCs. Note that the concentration of C-VOCs from VE does not follow the typical road traffic volume. This is because the light intensity is low in the morning, making the chemical reaction of VOCs slow and C-VOCs gradually peak at noon. By simulating the C-VOCs produced by different species

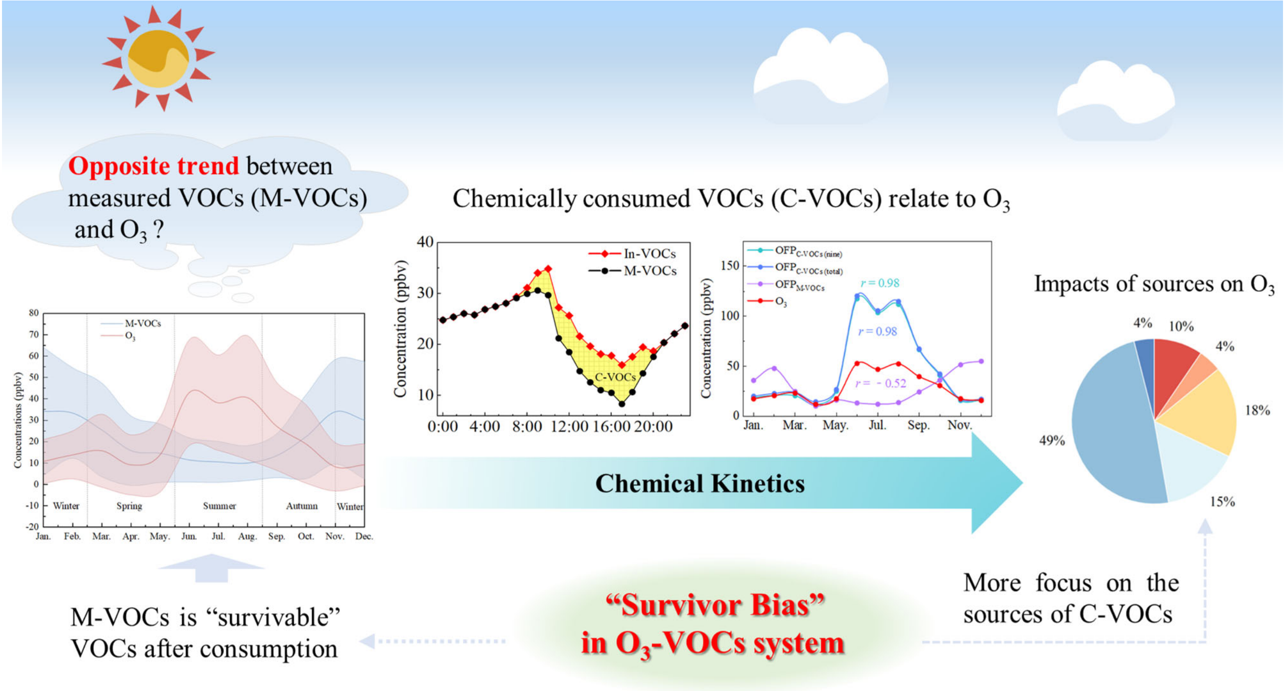

Fig. 6 **Schematic diagram.** This diagram explains the survivor bias for the O_3-VOCs system. At the observation site, low-VOCs concentrations in summer correspond to high O_3 concentrations, while high VOCs concentrations in winter correspond to low O_3 concentrations. With the enhancement of photochemical reaction (in summer), VOCs with high reactivity will undergo photochemical reaction and participate in the accumulation of O_3, resulting in less survivable M-VOCs. In contrast, a small number of VOCs were lost in winter, and the survivable M-VOCs gradually increased. Therefore, from the perspective of O_3 pollution control, we should focus more on the C-VOCs sources.

under different Δ*t* (Supplementary Note 5 and Supplementary Fig. 11), we further illustrated that the C-VOCs are related to the concentration and Δ*t* of species. Thus, prioritizing the source control of key reactive VOCs species within given time will help to reduce O_3 pollution more effectively. For the mitigation of surface O_3 pollution, VOCs control strategy inferred from the results of CK-PMF model will be effective.

DISCUSSION

This study proposes a new method to quantify the source contributions of C-VOCs (e.g., BE, industry, SU) and their impacts on O_3 formation, by using a combination of ground-based observations and model calculations. This finding not only offers new insights into the prioritization of VOCs control for mitigating O_3 pollution, but also better reduces the misjudgment of major VOCs sources (contributing to O_3 formation) caused by survivor bias in observational data. In summary, the processes for ambient VOCs to generate O_3 are complex due to VOCs reactivity and the behavior of source emissions. Figure 6 illuminates the concept map for survivor bias in O_3-VOCs system. We found that C-VOCs are the main precursors for O_3 production, and M-VOCs are actually the survivable VOCs after photochemical reactions. In summer, the highly reactive C-VOCs species elapsed a lot to produce O_3, leaving fewer survivable M-VOCs; in winter, with the weakening of photochemical reactions, a small part of C-VOCs elapsed, and the survivable M-VOCs gradually enriched. For example, in summer, the observed concentration of survivable ethylene (one of the reactive species) is low, but its contribution to O_3 is relatively great of all species. The idea of survivor bias answers the question that why the concentration between M-VOCs and O_3 exhibits an opposite trend, and clarifies the cause of this environmental problem.

Understanding the survivor bias can help us better capture the VOC-sources that actually impact O_3 formation. This work provides evidence that sources of C-VOCs play an important role on O_3 formation. Here we calculated the main sources of C-VOCs by the comprehensive application of chemical kinetics and source apportionment technology. The differences in source contribution results between C-VOCs and M-VOCs suggest that if we only focus on the sources of observed species with high concentrations, it will lead to a bias in O_3 pollution control strategies. Of course, to reduce VOCs concentration, focus on the main emission sources of M-VOCs is necessary. However, from the perspective of O_3 pollution control, we should focus more on the VOCs sources that contributing the most O_3 formation, that is, the emission sources of C-VOCs. The results suggested that: to better prevent and control O_3 precursors, it is necessary to control the C-VOCs from sources, rather than relying solely on observational data. The findings here will support the development of more efficient O_3 control strategies through VOCs chemical kinetics and photochemical depletion processes, and will also inform management of O_3-VOCs systems by assessing the source impacts of C-VOCs on O_3 formation. Future air control strategies should pay more attention to C-VOCs and their sources.

In addition, the understanding of this mechanism of survivor bias in O_3-VOCs system provided by this study is also applicable to other locations, and may be extended beyond air pollution to the management of volatile or highly active pollutants in a variety of environmental issues (e.g., water pollution), but may have limitations in the prevention and control of stable substances.

METHODS

Measurements and instruments

Online VOCs and meteorological measurements with 1 h time resolution were conducted in Tianjin, a megacity in northern China, from January 1 to December 31, 2018. The sampling site is located at a rural area surrounded by rich vegetation. The site is about 200 m away from a major roadway with high automobile traffic volume. Multiple industrial parks and industrial furnaces are located about 40 km to the northwest and northeast direction. The industrial facilities are engaged in rubber manufacturing, plastics production and metal smelting. The gas chromatograph (GC955 611/811, Syntech Spectras Inc., Holland) with two detectors was used to measure ambient VOCs, including 27 alkanes, 10 alkenes, 1 alkyne (acetylene) and 16 aromatic hydrocarbons (Supplementary Table 1). These VOCs are included in the list of important species that the USEPA recommends photochemical assessment monitoring stations should monitor. Hourly O_3 concentrations were measured by ultraviolet absorption instrument (Focused Photonics Inc., China). Meteorological parameters, such as T, were measured using an automatic meteorological observation system (LUFFT Inc., Germany). Solar radiation was recorded with a sun photometer (Kipp & Zonen Inc., the Netherlands). More details, including monitoring instruments and quality control, are described in Supplementary Note 6.

Receptor models

For the purpose of calculating the source of C-VOCs, the receptor model in source apportionment technology is introduced. Receptor models generally assume that the chemical composition of pollutants is relatively stable and do not take into account dispersion in rough terrain, long-range transport, wet and dry deposition[44]. The PMF model is a widely used receptor model for air pollutants source apportionment[27,45–48]. The principle of can be expressed by Eq. (2)[49–51]:

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \tag{2}$$

where x_{ij} (ppbv) is the concentration of the jth species in ith sample, g_{ik} (ppbv) is the contribution of the kth source to the ith sample, f_{kj} (ppbv/ppbv) is the fraction of the jth species from the kth source; e_{ij} (ppbv) are the residuals, and p is the number of factors.

Major VOCs species were selected (the principles for species selection are described in Supplementary Note 7.2) to put into factorization models. Here, the PMF/ME2-SR (PMF/multilinear engine 2-species ratios) model was used to estimate source contributions by adding the ratios of VOCs characteristic species components (Supplementary Note 7.3) in the extraction process[52,53]. The advantage of this model is that the extracted factors have more physical significance. To evaluate the reliability of the source apportionment results of PMF/ME2-SR, we also applied partial target transformation-positive matrix factorization (PTT-PMF) to calculate the sources of VOCs. It establishes a normalized target source profile based on measured source profiles, selects and fixes the source markers of all sources, and finally realizes source apportionment[54]. Details on PMF/ME2-SR and PTT-PMF models are provided in Supplementary Note 7.

Quantitative calculation of C-VOCs

The calculation of C-VOCs is the key step of the proposed source apportionment method. Existing VOCs source apportionment methods are based on a common assumption: all VOC species have a common photochemical age, that is, all species have the same transport time from the source to the receptor. In fact, the concentration of VOCs in the atmosphere is the result of the mixture of disparate pollutant air masses discharged in different time and space, and the photochemical age of VOCs emitted by various pollution sources may also be significantly different. Here, the proposed method considers the diversity of VOCs species and estimates unique reaction rate constants, radical concentrations, and photochemical consumption times for each VOC species through chemical kinetics. Therefore, compared to previous methods, the source emission concentration of VOCs by this method is more representative of the actual atmospheric conditions.

Based on the conservation of species, the consumption of VOCs is related to the reaction and deposition process. However, this study only refers to the consumption caused by the chemical reaction since it is combined with the PMF model which without considering diffusion for subsequent analysis. The photochemical consumption of VOCs (without considering dispersion) from the source to receptors can be described by Eq. (3):

$$\text{C–VOCs} = \text{In–VOCs} - \text{M–VOCs} \tag{3}$$

In this study, C-VOCs are defined as the photochemical consumption of VOCs in reaction with OH radicals during the transport from the source to observation site in the pathway of O_3 formation (Eq. (3), unit: ppbv); the concentration of VOCs in the air released from pollution sources (or, effectively, at the receptor assuming no photochemical consumption) as initial VOCs concentrations, hereafter defined as In-VOCs (see Eq. (4), unit: ppbv); and the measured concentration at the observation site (with photochemical loss) as M-VOCs (unit: ppbv). Their relationship can be better described in Supplementary Fig. 12.

The dominant chemical reactions of VOCs in the troposphere are OH radical reactions during the daytime[10], other removal paths including deposition and reacting with NO_3 radicals are assumed to be negligible. As shown in Eq. (4)[55], In-VOCs are calculated by assuming all VOCs come from pollution sources and there is no dispersion:

$$[\mathrm{VOC}_i]_{\mathrm{In}} = [\mathrm{VOC}_i]_{\mathrm{M}} \times \exp(k_i[OH]\Delta t) \tag{4}$$

where $[\mathrm{VOC_i}]_{\mathrm{In}}$ denotes the concentration of VOCs species i, as if no photochemical aging is taken place (unit: ppbv); k_i represents the rate constant of the reaction between species i and OH radicals (unit: $\mathrm{cm^3{\cdot}molecule^{-1}{\cdot}s^{-1}}$) and is estimated based on Eqs. (5)–(6); [OH] is the concentration of OH radicals (the estimation methods are in Supplementary Note 3.1, unit: $\mathrm{molecule{\cdot}cm^{-3}}$); and Δt is the photochemical consumption time of VOCs species i reacting with OH radicals from sources to the receptor (the estimation method is in Supplementary Note 3.2, unit: s):

$$k_{OH} = \mathrm{A}T^n \mathrm{e}^{-B/T} \tag{5}$$

$$k_{OH} = A\mathrm{e}^{-B/T} \tag{6}$$

where A is the Arrhenius constant (unit: $\mathrm{cm^3{\cdot}molecule^{-1}{\cdot}s^{-1}}$); B is the ratio of apparent activation energy (E_a) to molar gas constant (R) (unit: K); n is a coefficient, generally $n = 2$[10]. For alkanes, the recommended temperature-dependent expression is a three-parameter expression (Eq. (5)); while for alkenes, alkynes, and aromatics, Eq. (6) is used[56]. The values of A, B, n, and k_{OH} are provided in Supplementary Table 6 of the Supplementary Information.

Source apportionment of C-VOCs

Considering the reactivity of VOCs, using PMF/ME2-SR for directly analyzing the sources of M-VOCs is not appropriate because PMF is based on linear fitting (assuming no chemical loss of species). Here, CK-PMF model was established by combining chemical kinetics with PMF model to obtain the consumption sources and contributions of VOCs.

Stage 1. To calculate the sources of C-VOCs, we should quantify the source contributions of In-VOCs. The chemical kinetics method was used to calculate relevant parameters to obtain In-VOCs. Then the PMF/ME2-SR model was adopted to determine sources categories of typical VOCs (31 species in this study), with In-VOCs as input data. As shown in Eq. (7):

$$ES_{In-VOCs\,ij} = ES_{1\,ij} + ES_{2\,ij} + \cdots + ES_{n\,ij} \tag{7}$$

where $ES_{\mathrm{In-VOCs}\,ij}$ represents the total contribution of ith In-VOCs species emitted from all sources at jth hour; $\mathrm{ES}_{1\ ij}$, $\mathrm{ES}_{2\ ij}$, and $\mathrm{ES}_{n\ ij}$ represent the contributions of ith In-VOCs species emissions from source 1, source 2 and up to source n at jth hour, respectively.

For PMF/ME2-SR modeling, we applied species compositions (ratios) in VOCs source profiles (Supplementary Table 7) as constraints to make the extracted factors approach actual VOCs source profile. Six factors were identified (detail source identification procedures are described in Supplementary Note 8), including IE (35%), VE (21%), SU (18%), BE (13%), GE (7%) and LPG (6%) (Supplementary Fig. 8). The results were consistent with independent modeling performed using the partial target transformation-PMF (PTT-PMF) model (Supplementary Fig. 13).

Stage 2. Based on chemical kinetics, Eq. (8) was derived from Eqs. (3) and (4), and it was used to predict sources and their contributions to C-VOCs (chemical consumptions in daytime), based on the above results:

$$CS_n = ES_n - ES_n \cdot e^{-k_{ij}\cdot[OH]_{ij}\cdot\Delta t_{ij}} = ES_n \cdot \left(1 - e^{-k_{ij}\cdot[OH]_{ij}\cdot\Delta t_{ij}}\right) \tag{8}$$

where CS_n is the contribution of ith C-VOCs species from the nth source at jth time of the observation site.

Ozone formation potential

The OFP is used to reflect the ability of different VOCs species that participate in atmospheric chemical reactions[34,57,58]. It's estimated here using the maximum incremental reactivity (MIR) method[9,57]. The OFP of individual VOC is calculated as follows:

$$\mathrm{OFP}_i = [\mathrm{VOCs}]_i \times MIR_i \qquad (9)$$

where $[\mathrm{VOCs}]_i$ represents the concentration of VOC species i (units: ppbv) and MIR_i is the MIR coefficient of VOC species i (dimensionless), indicating the reference scale of the O_3 forming ability. $\mathrm{OFP}_{\text{C-VOCs}}$ represents the maximum contributions of C-VOCs species to O_3 formation under idealized photochemical reaction conditions:

$$\mathrm{OFP}_{i\,\text{C-VOCs}} = \left([\mathrm{VOC}_i]_{\mathrm{In}} - [\mathrm{VOC}_i]_{\mathrm{M}}\right) \times MIR_i \qquad (10)$$

The MIR coefficients were obtained from Carter et al.[9] (Supplementary Table 6). Additionally, the method of OH radical loss rate (L_{OH}) was also estimated for verification purposes, as shown in Supplementary Note 4.

DATA AVAILABILITY

The data included in this study could be downloaded at https://figshare.com/ (https://doi.org/10.6084/m9.figshare.19561609). The data of back trajectories are openly obtained from https://www.ready.noaa.gov/HYSPLIT.php. Additional data that support the plots and the findings of this study are available from the corresponding author (G.S.) upon reasonable request.

CODE AVAILABILITY

The code to carry out the current analyses is available from the corresponding author upon request.

Received: 29 September 2021; Accepted: 14 April 2022;

REFERENCES

1. Jaffe, D. Relationship between surface and free tropospheric ozone in the western U.S. *Environ. Sci. Technol.* **45**, 432–438 (2011).
2. Lu, X. et al. Severe surface ozone pollution in China: a global perspective. *Environ. Sci. Technol. Lett.* **5**, 487–494 (2018).
3. Chen, Y. et al. Mitigation of PM2.5 and ozone pollution in Delhi: a sensitivity study during the pre-monsoon period. *Atmos. Chem. Phys.* **20**, 499–514 (2020).
4. Shi, Z. et al. Abrupt but smaller than expected changes in surface air quality attributable to COVID-19 lockdowns. *Sci. Adv.* **7**, eabd6696 (2021).
5. Schnell, J. L. & Prather, M. J. Co-occurrence of extremes in surface ozone, particulate matter, and temperature over eastern North America. *Proc. Natl Acad. Sci. USA* **114**, 2854–2859 (2017).
6. Lu, K. et al. Exploring atmospheric free-radical chemistry in China: the self-cleansing capacity and the formation of secondary air pollution. *Natl Sci. Rev.* **6**, 579–594 (2019).
7. Li, K. et al. A two-pollutant strategy for improving ozone and particulate air quality in China. *Nat. Geosci.* **12**, 906–910 (2019).
8. Nussbaumer, C. M. & Cohen, R. C. The role of temperature and NO_x in ozone trends in the Los Angeles Basin. *Environ. Sci. Technol.* **54**, 15652–15659 (2020).
9. Carter, W. P. L. Development of ozone reactivity scales for volatile organic compounds. *Air Waste* **44**, 881–899 (1994).
10. Atkinson, R. & Arey, J. Atmospheric degradation of volatile organic compounds. *Chem. Rev.* **103**, 4605–4638 (2003).
11. Wang, T. et al. Ozone pollution in China: a review of concentrations, meteorological influences, chemical precursors, and effects. *Sci. Total Environ.* **575**, 1582–1596 (2017).
12. Li, Q. et al. "New" reactive nitrogen chemistry reshapes the relationship of ozone to its precursors. *Environ. Sci. Technol.* **52**, 2810–2818 (2018).
13. Chen, T. et al. Measurement report: effects of photochemical aging on the formation and evolution of summertime secondary aerosol in Beijing. *Atmos. Chem. Phys.* **21**, 1341–1356 (2021).
14. Cohan, D., Hakami, A., Hu, Y. & Russell, A. Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis. *Environ. Sci. Technol.* **39**, 6739–6748 (2005).
15. Tan, Z. et al. Explicit diagnosis of the local ozone production rate and the ozone-NO_x-VOC sensitivities. *Sci. Bull.* **63**, 1067–1076 (2018).
16. Zou, Y. et al. Characteristics of 1 year of observational data of VOCs, NO_x and O_3 at a suburban site in Guangzhou, China. *Atmos. Chem. Phys.* **15**, 6625–6636 (2015).
17. Wang, Y. et al. Formation of highly oxygenated organic molecules from chlorine-atom-initiated oxidation of alpha-pinene. *Atmos. Chem. Phys.* **20**, 5145–5155 (2020).
18. Heard, D. E. & Pilling, M. J. Measurement of OH and HO_2 in the troposphere. *Chem. Rev.* **103**, 5163–5198 (2003).
19. Huang, S., Shao, M., Lu, S. & Liu, Y. Reactivity of ambient volatile organic compounds (VOCs) in summer of 2004 in Beijing. *Chin. Chem. Lett.* **19**, 573–576 (2008).
20. Blake, D. & Rowland, F. Urban leakage of liquefied petroleum gas and its impact on Mexico city air quality. *Science* **269**, 953–956 (1995).
21. Gao, J. et al. Comparative study of volatile organic compounds in ambient air using observed mixing ratios and initial mixing ratios taking chemical loss into account—a case study in a typical urban area in Beijing. *Sci. Total Environ.* **628–629**, 791–804 (2018).
22. Makar, P. A. et al. Chemical processing of biogenic hydrocarbons within and above a temperate deciduous forest. *J. Geophys. Res. Atmos.* **104**, 3581–3603 (1999).
23. Min, S. et al. Effects of Beijing Olympics control measures on reducing reactive hydrocarbon species. *Environ. Sci. Technol.* **45**, 514–519 (2011).
24. Zheng, H. et al. Monitoring of volatile organic compounds (VOCs) from an oil and gas station in northwest China for 1 year. *Atmos. Chem. Phys.* **18**, 4567–4595 (2018).
25. Kleinman, L. I. et al. Dependence of ozone production on NO and hydrocarbons in the troposphere. *Geophys. Res. Lett.* **24**, 2299–2302 (1997).
26. Seinfeld, J. Urban Air Pollution: State of the Science. *Science* **243**, 745–752 (1989).
27. Liu, Y. et al. Characterization and sources of volatile organic compounds (VOCs) and their related changes during ozone pollution days in 2016 in Beijing, China. *Environ. Pollut.* **257**, 113599 (2020).
28. Wang, X. et al. Sensitivities of ozone air pollution in the Beijing-Tianjin-Hebei area to local and upwind precursor emissions using adjoint modeling. *Environ. Sci. Technol.* **55**, 5752–5762 (2021).
29. Wu, R. & Xie, S. Spatial distribution of ozone formation in China derived from emissions of speciated volatile organic compounds. *Environ. Sci. Technol.* **51**, 2574–2583 (2017).
30. Chen, W. et al. Understanding primary and secondary sources of ambient carbonyl compounds in Beijing using the PMF model. *Atmos. Chem. Phys. Discuss.* **13**, 15749–15781 (2013).
31. Bidleman, T. F. Atmospheric processes. *Environ. Sci. Technol.* **22**, 361–367 (1988).
32. Wang, Y., Raihala, T. S., Jackman, A. P. & St. John, R. Use of tedlar bags in VOC testing and storage: evidence of significant VOC losses. *Environ. Sci. Technol.* **30**, 3115–3117 (1996).
33. He, C. et al. Recent advances in the catalytic oxidation of volatile organic compounds: a review based on pollutant sorts and sources. *Chem. Rev.* **119**, 4471–4568 (2019).
34. Song, M. et al. Source apportionment and secondary transformation of atmospheric nonmethane hydrocarbons in Chengdu, Southwest China. *J. Geophys. Res.: Atmos.* **123**, 9741–9763 (2018).
35. Rangaiah, G. P. Estimation of frequency factor and activation energy in the Arrhenius expression. *Chem. Eng. J.* **29**, 159–166 (1984).
36. Qu, H. et al. Chemical production of oxygenated volatile organic compounds strongly enhances boundary-layer oxidation chemistry and ozone production. *Environ. Sci. Technol.* **55**, 13718–13727 (2021).
37. Slater, E. J. et al. Elevated levels of OH observed in haze events during wintertime in central Beijing. *Atmos. Chem. Phys.* **20**, 14847–14871 (2020).
38. Wolfe, G. M. et al. The framework for 0-D atmospheric modeling (F0AM) v3.1. *Geosci. Model Dev.* **9**, 3309–3319 (2016).
39. Park, J.-H. et al. Active atmosphere-ecosystem exchange of the vast majority of detected volatile organic compounds. *Science* **341**, 643–647 (2013).
40. Wennberg, P. O. et al. Gas-phase reactions of isoprene and its major oxidation products. *Chem. Rev.* **118**, 3337–3390 (2018).
41. Ryerson, T. B. et al. Observations of ozone formation in power plant plumes and implications for ozone control strategies. *Science* **292**, 719–723 (2001).
42. Trainer, M. et al. Models and observations of the impact of natural hydrocarbons on rural ozone. *Nature* **329**, 705–707 (1987).
43. Kuzma, J., Nemecek-Marshall, M., Pollock, W. H. & Fall, R. Bacteria produce the volatile hydrocarbon isoprene. *Curr. Microbiol.* **30**, 97–103 (1995).
44. Henry, R. C., Lewis, C. W., Hopke, P. K. & Williamson, H. J. Review of receptor model fundamentals. *Atmos. Environ.* **18**, 1507–1515 (1984).
45. Leuchner, M. & Rappenglück, B. VOC source-receptor relationships in Houston during TexAQS-II. *Atmos. Environ.* **44**, 4056–4067 (2010).
46. Yang, Y. et al. Ambient volatile organic compounds in a suburban site between Beijing and Tianjin: concentration levels, source apportionment and health risk assessment. *Sci. Total Environ.* **695**, 133889 (2019).
47. Ling, Z. H., Guo, H., Cheng, H. R. & Yu, Y. F. Sources of ambient volatile organic compounds and their contributions to photochemical ozone formation at a site in the Pearl River Delta, southern China. *Environ. Pollut.* **159**, 2310–2319 (2011).
48. Zhang, X. et al. Characteristics, reactivity and source apportionment of ambient volatile organic compounds (VOCs) in a typical tourist city. *Atmos. Environ.* **215**, 116898 (2019).

49. Paatero, P. & Tapper, U. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **5**, 111–126 (1994).
50. Paatero, P. Least squares formulation of robust non-negative factor analysis. *Chemom. Intell. Lab.* **37**, 23–35 (1997).
51. Jaars, K. et al. Receptor modelling and risk assessment of volatile organic compounds measured at a regional background site in South Africa. *Atmos. Environ.* **172**, 133–148 (2018).
52. Paatero, P. The multilinear engine–a table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model. *J. Comput. Graph. Stat.* **8**, 854–888 (1999).
53. Sturtz, T. M., Adar, S. D., Gould, T. & Larson, T. V. Constrained source apportionment of coarse particulate matter and selected trace elements in three cities from the multi-ethnic study of atherosclerosis. *Atmos. Environ.* **84**, 65–77 (2014).
54. Gao, J. et al. Source apportionment for online dataset at a megacity in China using a new PTT-PMF model. *Atmos. Environ.* **229**, 117457 (2020).
55. McKeen, S. A. et al. Hydrocarbon ratios during PEM-WEST A: a model perspective. *J. Geophys. Res.: Atmos.* **101**, 2087–2109 (1996).
56. Atkinson, R. Kinetics of the gas-phase reactions of OH radicals with alkanes and cycloalkanes. *Atmos. Chem. Phys.* **3**, 2233–2307 (2003).
57. Zhang, Y. et al. Development of ozone reactivity scales for volatile organic compounds in a Chinese megacity. *Atmos. Chem. Phys.* **21**, 11053–11068 (2021).
58. Zhang, K. et al. Precursors and potential sources of ground-level ozone in suburban Shanghai. *Front. Environ. Sci. Eng.* **14**, 92 (2020).

ACKNOWLEDGEMENTS

This study was supported by National Natural Science Foundation of China (42077191), Fundamental Research Funds for the Central Universities (63213072), and the Blue Sky Foundation. This work is a contribution from State Environmental Protection Key Laboratory of Urban Ambient Air Particulate Matter Pollution Prevention and Control. We would like to thank Dr. Haofei Yu (University of Central Florida) for editing and linguistic assistance.

AUTHOR CONTRIBUTIONS

Z.W. led the writing and carried out most of the data analysis. G.S. designed the structure of the paper and developed the writing and editing. Z.S, F.W., W.W., D.C., Y.F., and A.G.R. contributed significant comments and editing of the paper. W.L. performed on checking the format specification. D.L. helped in the analysis of the data. All authors contributed to the paper and approved the submitted version.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-022-00261-7.

Correspondence and requests for materials should be addressed to Guoliang Shi.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.