Source analyses of ambient VOCs considering reactive losses: methods of reducing loss effects, impacts of losses, and sources

Baoshuang Liu\textsuperscript{1,2*}, Yao Gu\textsuperscript{1,2}, Yutong Wu\textsuperscript{1,2}, Qili Dai\textsuperscript{1,2}, Shaojie Song\textsuperscript{1,2}, Yinchang Feng\textsuperscript{1,2*}, and Philip K. Hopke\textsuperscript{3,4}

\textsuperscript{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, China

\textsuperscript{2}CMA-NKT Cooperative Laboratory for Atmospheric Environment-Health Research, Tianjin 300350, China

\textsuperscript{3}Department of Public Health Sciences, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

\textsuperscript{4}Institute for a Sustainable Environment, Clarkson University, Potsdam, NY 13699, USA

Correspondence: Baoshuang Liu (lbsnankai@foxmail.com) and Yinchang Feng (fengyc@nankai.edu.cn)
Abstract

Chemical losses of ambient reactive volatile organic compounds (VOCs) is a long-term issue yet to be resolved in VOC source apportionments. These losses substantially reduce the concentrations of highly reactive species in the apportioned factor profiles and result in the underestimation of source contributions. This review assesses the common methods and existing issues in ways to reduce losses and loss impacts in source analyses and suggest research directions for improved VOC source apportionments. Positive Matrix Factorization (PMF) is now the main VOC source analysis method compared to other mathematical models. The issue in using any apportionment tool is the processing of the data to be analyzed to reduce the impacts of reactive losses. Estimating the initial concentrations of ambient VOCs based on photochemical age has become the primary approach to reduce reactive loss effects in PMF except for selecting low reactivity species or nighttime data into the analysis. Currently, the initial concentration method only considers daytime reactions with hydroxyl (•OH) radicals. However, the •OH rate constants vary with temperature and that has not been considered. Losses from reactions with O₃ and NO₃ radicals especially for alkene species remain to be included. Thus, the accuracy of the photochemical-age estimation is uncertain. Beyond developing accurate quantitative approaches for reactive losses, source analyses methods for the consumed VOCs and the accurate quantification of different source contributions to O₃ and secondary organic aerosols are important additional directions for future research.

Keywords: Reactive loss; Consumed VOCs; Source analyses; PMF; Initial concentrations
1 Introduction

Ambient VOCs are key precursors of ozone (O$_3$) and secondary organic aerosols (SOAs) formation (Li et al., 2016; Li et al., 2018a; Wu and Xie, 2018). Accurate apportionment of their sources can be important in developing effective prevention and control measures for atmospheric O$_3$ and secondary organic particulate matter pollution (Carrillo-Torres et al., 2017; Meng et al., 1997; Wang et al., 2022a). Current research on source analyses of atmospheric VOCs has been primarily conducted utilizing the species ratio method (Che et al., 2019; Zhang et al., 2021), the photochemical age-based parameterization method (Huang et al., 2020; Zhu et al., 2021), and receptor models (e.g., positive matrix factorization (PMF) (Gu et al., 2020; Liu et al., 2016), chemical mass balance (CMB) (Song et al., 2019), and principal component analysis/multiple linear regression (PCA/MLR) (Jia et al., 2016; Sanchez et al., 2008), etc.) based on the measured concentration data. Receptor models, especially PMF, are the most widely used source apportionment methods (Song et al., 2008; Vega et al., 2022; Yang et al., 2022b). However, compared to particulate matter, ambient VOC species can undergo rapid, complex chemical reactions with •OH radicals, NO$_3$ radicals, O$_3$, etc. (Atkinson and Arey, 2003), resulting in substantial chemical losses during their transport from their sources to the receptor site (Yang et al., 2022b; Yuan et al., 2012a). Therefore, the source apportioned results based on the measured VOC data have difficulty to reflect the actual impacts of emission sources on air quality (Wu et al., 2023b; Yang et al., 2022a).

To reduce the impact of reactive losses, many studies have selected low-reactivity VOC species to conduct source analyses when using PMF source analyses (Guan et al., 2020; Yang et al., 2022a). However, this method cannot fully solve the issues related to reactive losses and provide complete source apportionments since some highly active marker species such as isoprene (Tan et al., 2020) cannot be excluded from the PMF input species without a substantial loss of information (Liu et al., 2023a). Therefore, recent studies estimated the initial concentrations of ambient VOCs (i.e., the VOC concentrations in the fresh emissions before they can undergo chemical reactions) utilizing the photochemical age-based parameterization method and then performed source analyses with PMF (He et al., 2019; Zou et al., 2023). However, there could still be high uncertainties in the estimated photochemical age of VOC species (Parrish et al., 2007; Yuan et al., 2012b).

There were studies that only used nighttime data for source analyses to reduce the loss impacts
(Kim et al., 2005), but the representativeness of the apportioned results was likely limited since there could be daytime only sources that would not be observed at night (Buzcu and Fraser, 2006). Some studies applied decay factors to correct the impact of reactive losses in using CMB for VOC source analyses (Friedlander, 1981; Lin and Milford, 1994; Na and Pyo Kim, 2007). However, there were relatively few studies and the effectiveness of this method still needs to be assessed. In 2023, Liu et al. (2023a) systematically investigated the impact of VOCs photochemical losses on the PMF source apportioned results, and found that photochemical losses reduced the concentrations of highly reactive species in factor profiles resulting in the contributions of biogenic emissions and polymer production-related industrial sources being substantially underestimated. However, there has been little related research to assess the limitations of VOC apportionments.

With the substantial increase of O₃ concentrations in many locations worldwide in recent years (Li et al., 2020; Zhang et al., 2018; Zhao et al., 2021), accurate source apportionment approaches of the key precursor VOCs have been acquired increasing attention (Gu et al., 2022). Thus, a related issue is the determination of the VOCs consumed in the formation of the observed O₃ and SOAs (Gu et al., 2023). Although some studies have investigated the underestimated contributions of emission sources by comparing the apportioned results based on the initial and measured data (He et al., 2019; Wu et al., 2023a), there are few publications that conducted source analyses for the consumed VOCs and apportioned the contributions of different sources to the formation of O₃ and SOA. In 2023, Gu et al. (2023) attempted to develop a method for apportioning the emission sources of the consumed VOCs, and Wang et al. (2022b) conducted a similar study. Currently, although some studies have been conducted on these issues of VOC reactive losses and achieved some important results (Gu et al., 2023; Watson et al., 2001; Wu et al., 2023b), there remain unresolved issues. To better promote progress on the development of better VOCs source apportioned methods, the present work systematically investigated the main methods and shortcomings of those methods that are currently applied to resolve these issues by reviewing relevant papers, with the aim of identifying directions for the future developments and improvement of VOCs source apportionment methods.

2 Materials and methods

Relevant papers were collected by searching Science Direct (Elsevier), the Web of Science, and
the China National Knowledge Infrastructure (CNKI) with keywords: volatile organic compounds (VOCs), oxygenated VOCs (OVOCs), initial concentrations/mixing ratios, chemical/photochemical losses, source apportionment/analysis, positive matrix factorization (PMF), chemical mass balance (CMB), receptor model, and photochemical age-based parameterization method. The information extracted from each publication included methods for reducing the impacts of photochemical losses, the impacts of VOC photochemical losses on source analyses, the source apportioned methods and the results for the consumed VOCs in the atmosphere. This study identified a total of 151 papers, of which 62 were published in the past five years (i.e., 2019-2023), accounting for 41% of the total publications.

3 Results and discussion
3.1 Methods of reducing the impacts of reactive losses

Ambient VOCs can be substantially oxidized by O$_3$, hydroxyl (•OH), and nitrate (NO$_3$) radicals (Atkinson and Arey, 2003; Bey et al., 2001; Finlayson-Pitts and Pitts, 1997), especially oxidation by •OH radicals primarily during the daytime (Wang et al., 2013). Therefore, reducing the impacts of VOC reactive losses on source apportionment has long been an important but not easy issue to resolve in VOC source apportionments (Liu et al., 2023a). Ambient VOCs primarily include alkanes, alkenes, aromatic hydrocarbons, alkynes, OVOCs, and halogenated hydrocarbons, etc. Studies suggested that the Photochemical Assessment Monitoring Stations (PAMS) VOCs (i.e., alkanes, alkenes, aromatic hydrocarbons, and alkynes) accounted for ~58% of total VOCs (TVOCs), while OVOCs and halogenated hydrocarbons contributed ~8.5%-63% and ~0.1%-14% to the TVOCs, respectively (Table S1) (Li et al., 2016; Wu et al., 2023a; Zhu et al., 2021). The reported number of measured species ranged from 13 to 116, including 0-32 alkanes, 2-16 alkenes, 1 alkylene (only acetylene), 3-18 aromatic hydrocarbons, 4-21 OVOCs, and 28-35 halogenated hydrocarbons. There were substantial differences in the identified emission sources for the different types of VOCs (Mo et al., 2016). For example, PAMS VOCs mainly originate from primary anthropogenic sources (Chen et al., 2010), while OVOCs can also be formed by oxidation of PAMS VOCs in addition to primary source emissions (Chen et al., 2014; Seinfeld and Pandis, 1986). The contributions of secondary formation to some OVOC species (e.g., acetaldehyde and propionaldehyde) can exceed 50% (de
In addition to local emissions, the ambient concentrations of halogenated hydrocarbons can also be affected by the long-distance transport (Mintz and McWhinney, 2008). Therefore, utilizing only the same source analyses approach for multiple VOC species from different sources might produce results with high uncertainties in the apportionments.

Source analyses methods for ambient VOCs mainly included PMF, CMB, and photochemical age-based parameterization method (Table S2). PMF was the most commonly method used for source analyses (Yang et al., 2022b). Approximately 61% of the reviewed publications focused on source analyses of PAMS species using PMF, while OVOCs primarily used photochemical age-based parameterization for source apportionment but with only relatively few studies (only 7 papers) (Table S3). Additionally, the studies on data from simultaneous measurement of both PAMS and OVOC species utilized two methods to conduct source analyses: one method was to simultaneously input PAMS and some OVOC species into the receptor model (e.g., PMF) for source analyses and only separated primary and secondary source contributions of ambient VOCs (Li et al., 2023; Liu et al., 2023b; Tan et al., 2021) (Table S2). Another approach was to use the PMF and photochemical age-based parameterization to obtain the source apportioned results for ambient PAMS and OVOCs, respectively, and then combined the two apportioned results to obtain finally source resolution results for the ambient VOCs (Zhu et al., 2021). However, the utilization of this method has been limited.

3.1.1 Methods for reducing reactive loss in PMF source analyses

3.1.1.1 Selecting low activity species or incorporating night only data into PMF

The methods of reducing effects of reactive losses for source apportionments utilizing different models are shown in Fig. 1. At present, selecting the VOC species with lower reactivity to be input in PMF for apportionment is the most commonly used approach (Chen et al., 2019; Tan et al., 2020; Yang et al., 2022b). Many highly active alkene and aromatic hydrocarbon species were not included in the PMF calculations (Gu et al., 2023; Liu et al., 2023a). For example, Liu et al. (2023a) excluded highly active species such as 1-hexene, trans-2-butene, trans-2-pentene, cis-2-pentene, and cis-2-butene. In addition, some studies have utilized nighttime only data for source analyses (Buzcu-Guven and Fraser, 2008; Buzcu and Fraser, 2006; Kim et al., 2005; Xie and Berkowitz, 2006). However, the volatile emission sources commonly contribute substantially to VOCs during the
daytime (Gu et al., 2023). For example, Buzcu and Fraser (2006) used nighttime and all-day data to conduct VOC source apportionment in Houston, respectively. They found that the night only data did not identify biogenic or gasoline sources due to minimal emissions at night. Thus, the contributions of the nighttime emission sources substantially increased compared to the all-day results because of the missing source types. Meanwhile, the all-day VOC source analysis conducted by Zhao et al. (2004) in Houston during the same period also showed substantial differences from the results reported by Buzcu and Fraser (2006) utilizing night-only data. An additional consideration is that VOCs can react with NO$_3$ radicals at night causing reactive losses. Thus, this approach does not eliminate the effects of all possible loss mechanisms.

**Figure 1.** Methods for reducing the impacts of reactive losses for different source analyses models.
PMF represents positive matrix factorization, CMB represents chemical mass balance, and PAPM represents photochemical age-based parameterization method.

3.1.1.2 Incorporating VOC initial concentration estimated by photochemical age-based parameterization method into PMF

The photochemical age-based parameterization method is an approach to estimate the initial concentrations of ambient VOCs (Shao et al., 2011; Wang et al., 2013; Yuan et al., 2012b). This method assumes that the chemical loss of ambient VOCs mainly originates only from reactions of the VOCs with •OH radicals (Parrish et al., 1992; Sun et al., 2016; Wang et al., 2013). The photochemical ages (Δt) of VOC species are estimated, and then combined with the rate constants for reaction for •OH radicals with the VOC species (k_{OH}) to calculate their initial concentrations and quantified the VOC photochemical losses (Shao et al., 2011; Wang et al., 2013; Zhan et al., 2021).

The equation for the initial concentration calculation is shown in Eq. (1), where k_{OH} and Δt are the key parameters in this approach.

\[
[VOC]_t = [VOC]_0 \times \exp (-k_{\cdot OH} \times [\cdot OH] \Delta t)
\]  

(1)

where [VOC]_0 and [VOC]_t are the initial and measured concentrations of VOC at time t, [•OH] represents the concentration of •OH radical. However, many studies estimated [•OH] Δt because of the lack of •OH radical concentration data and defined the product as the •OH exposure (Shao et al., 2011; Wei et al., 2022; Yuan et al., 2012b).

The k_{OH} of VOC species are substantially affected by atmospheric temperature (Atkinson and Arey, 2003), which could be commonly measured in laboratory (or chamber) experiments or through the use of detailed chemical computer models (Atkinson and Arey, 2003). According to the reviewed papers regarding the estimation of photochemical age, it was found that approximately 46% of the publications used the Atkinson and Arey’s 2003 summary of k_{OH} values at 298° K (Atkinson and Arey, 2003), and approximately 9% used Carter’s 2010 summary of k_{OH} values at 300° K (Carter, 2010). The detailed k_{OH} values for different VOC species summarized in these two publications are provided in Table S4. However, other relevant studies cited k_{OH} values from other papers by Atkinson and/or Carter (Atkinson, 1991, 2007; Atkinson et al., 2006; Talukdar et al., 1994). The rate constant for a bimolecular reaction between a hydroxyl radical and a VOC molecule has a roughly exponential dependence on temperature in the Arrhenius equation (Liu et al., 2023a). Therefore,
these published papers only used $k_{\cdot \text{OH}}$ values at specific temperatures, which made it difficult to characterize the actual reaction rates at the actual temperatures in the atmosphere. To address this issue, Wang et al. (2022b) and Wang et al. (2023) utilized the Arrhenius expression to locally optimize the $k_{\cdot \text{OH}}$ values of VOC species based on the actual temperature conditions to further improve the accuracy of estimating the VOC initial data. In addition, some studies used the $k_{\cdot \text{OH}}$ of m,p-xylene instead of values of more reactive species to reduce high-outliers when estimating the initial concentration (Wang et al., 2013). However, the validity of this method was difficult to assess.

There are two main methods (as shown in Table S5) for estimating $\Delta t$, namely, the species ratio method (Roberts et al., 1984; Shao et al., 2011; Wan et al., 2022) and the sequential reaction model (Gong et al., 2018; Shao et al., 2011; Xie et al., 2008). The specific details are as follows:

1. **Estimation of $\Delta t$ based on the species ratio method**

\[
\Delta t = \frac{1}{(k_A - k_B) \times [\cdot \text{OH}]} \times \left[ \ln \left( \frac{[A]_0}{[B]_0} \right) - \ln \left( \frac{[A]_t}{[B]_t} \right) \right]
\]

where $k_A$ and $k_B$ are the reaction rate constants with $\cdot \text{OH}$ radicals for species A and B of highly homologous species with substantial differences in reaction activity, respectively, $\frac{[A]_0}{[B]_0}$ is the initial concentration ratio of A to B in fresh emissions, $\frac{[A]_t}{[B]_t}$ is the measured concentration ratio of A to B at time $t$. The determination of initial concentration ratio of A to B was extremely critical for estimating the $\Delta t$. The $\Delta t$ is used to then calculate the initial concentrations of the VOCs of interest using Eq. (1). The initial concentrations can then be the input variables to PMF (Liu et al., 2023a).

This ratio method was first proposed by Roberts et al. (1984). The method has been commonly applied to PAMS species emitted by anthropogenic sources. There have been differences in the species used to calculate the initial ratio in different papers (Table S6). Generally aromatic hydrocarbons (e.g., benzene, toluene, ethylbenzene, or m,p-xylene) were used as the selected species. The substantial differences in activity and high homology are generally the main basis for selecting species-ratio types (Lin et al., 2011; Shao et al., 2009; Zou et al., 2021; Zou et al., 2023). Most publications used the initial ratios of ethylbenzene/m,p-xylene (E/X) or m,p-xylene/ethylbenzene (X/E) to calculate $\Delta t$. These two species accounted for ~36.8% of the total reviewed publications in this study (Table S6), respectively. Approximately 15.8% of the papers used toluene/benzene (T/B), while m,p-xylene/benzene (X/B) and ethylbenzene/o-xylene (E/O) had limited use, accounting for...
~5.3% and ~10.5% of the reviewed publications.

The methods for obtaining the initial species ratio in the reviewed papers mainly included referring to other similar studies (Wang et al., 2016), source emission inventory values (Wang et al., 2013), and estimation based on the observed concentration data at the receptor sites at times when low reactivity was expected (Fang et al., 2021; Han et al., 2019; Wu et al., 2023b). There were several papers using the first two methods to obtain the initial species ratios (Table S6). The initial species ratios obtained by these methods had difficulty to accurately characterize the comprehensive impacts of multiple sources in the study area. Most studies were based on the observed data of low photochemical reaction periods at night (~ 19:00 to 07:00 LT) to calculate the initial species ratios (Table S6). This method assumes that VOC species emitted by different sources during the certain periods of night would not undergo oxidative reaction, and the concentration ratio of two homologous species had no the change during the transport process from source emissions to the receptor sites (Liu et al., 2023a; Sun et al., 2016; Yuan et al., 2012b). Since there could be residual reactions with nitrate radical and multiple sources of the indicator species, there remains uncertainty in the results. For example, McKeen and Liu (1993) and McKeen et al. (1996) found that the species ratios can be substantially influenced by the oxidative chemistry and atmospheric mixing. Comparing one species versus another, both the model results and the observations were consistent with an average rate of dilution roughly equivalent to n-butane oxidation (McKeen et al., 1996). This result has negative implications for the use of hydrocarbon ratios as chemical reaction clocks (McKeen et al., 1996). Parrish et al. (2007) also suggested that there were uncertainties in the determination of $\Delta t$ by the initial species ratios because of influence of fresh emissions along the transport path that perturb the results. However, this approach can still provide a useful measure of chemical processing in the atmosphere (Parrish et al., 2007).

There were differences in the estimation details of the initial species ratio in different reports (Liu et al., 2023a; Shao et al., 2011; Yuan et al., 2012b). For example, Yuan et al. (2012b) utilized the m,p-xylene and benzene data measured between 0:00 and 5:00 LT to conduct a linear fit. The regression line was then extrapolated to the highest benzene concentration during this period so that the initial ratio could be estimated from the m,p-xylene to benzene concentration ratios at this point. Wu et al. (2023b) and Yang et al. (2022a) used time periods (00:00-04:00 LT and 20:00-05:00 LT,
respectively) with the lowest and the highest E/X and T/B ratios at night to calculate the average values of corresponding E/X ratios less than various percentiles or the corresponding T/B values above various percentiles, and then determined the initial ratio by the minimum average ratio (i.e., E/X) or the maximum average ratio (i.e., T/B) of the two species with the highest homology (i.e., highest correlation) below different percentiles. Wu et al. (2023a) reported that the maximum value of X/E at 01:00-06:00 LT was considered as the initial species ratio. In addition, Liu et al. (2023a) utilized the slope of the linear fit of the scatterplots of corresponding ethylbenzene and m,p-xylene below the 10th percentile of E/X ratios during 00:00-05:00 LT to determine the initial ratio. Overall, the E/X range in the reviewed papers was between 0.22 and 0.75, T/B was between 3.14 and 4.25, X/E was between 0.39 and 2.47, E/O was between 1.30 and 1.32, and X/B was 2.2 (Table S6).

Zhang et al. (2020) attempted to confirm the feasibility and rationality of estimating \( \Delta t \) based on T/B utilizing the regressions of benzene and toluene versus CO and \( \Delta t \). Zou et al. (2023) demonstrated the rationality of selecting the ratio of E and X to estimate the \( \Delta t \) through high correlation of these two species having substantial activity differences. However, there were few studies to systematically assess how to choose the most suitable species ratios and the calculated method.

Multiple types of initial species ratios were used for estimating \( \Delta t \) of VOCs (Table S6), and the results estimated by the different ratios could vary substantially. Shao et al. (2011) suggested that the differences between the initial concentrations calculated using the ratios of three pairs of compound ratios (i.e., T/B, X/E, and i-butene/propene (iB/P)) were generally within 50%, and the X/E and iB/P results were in good agreement. Zou et al. (2021) examined the sensitivity of estimating initial concentration based on E/X and E/O initial ratios. They found that the relative variation range of the initial concentrations of PAMS species was between 0.41% and 68.06% for an initial E/O of 1.3 when compared with an initial E/X of 0.5.

Additionally, in the ratio method, the same \( \Delta t \) was calculated for different VOCs for each sample, and the paired species for estimating the ratio were aromatic hydrocarbons (Table S6), which are mainly emitted from anthropogenic emission sources such as solvent use and petrochemical enterprises (Mo et al., 2015; Na et al., 2004; Yuan et al., 2010; Zhang et al., 2016). Therefore, the \( \Delta t \) obtained based on these ratios may only reflect the chemical aging of VOC species emitted from
these specific sources. It is difficult to accurately characterize the $\Delta t$ for these species from other sources. In the future, different types of species ratios (not just aromatic hydrocarbons) need to be selected based on VOC species from different sources. The $\Delta t$ derived from different types of VOC species should be calculated to improve traditional ratio methods.

An additional issue is that this method only considers the reactions between VOCs and $\cdot$OH radicals, while chemical reactions with O$_3$ during the day or NO$_3$ radicals at night were generally excluded. Liu et al. (2023a) attempted to estimate the reactive losses between 19:00 and 23:00 LT, but they considered the gradual loss of $\cdot$OH radicals after sunset and did not consider the losses caused by NO$_3$ radicals reactions that may be present at night since they are not photolyzed as occurs in the daylight. Additionally, alkenes can directly react with O$_3$ in the atmosphere causing losses. However, there are currently no reports on estimation methods for these reactive losses. In future studies, the ratio of two alkene species with substantial differences in reaction rates with O$_3$ could be used as the base ratio as in Eq. (2), and combined with measured O$_3$ reaction rate constants (Atkinson and Arey, 2003) to estimate the chemical reaction time ($\Delta t$) of alkene species with O$_3$. Then, chemical losses of alkene species in the reaction with O$_3$ could be estimated using Eqs. (1) and (9) and the combined effects of O$_3$ and $\cdot$OH could be estimated.

(2) Estimation of $\Delta t$ based on the sequential reaction model

It has been found that using the ratio method to calculate the $\Delta t$ of isoprene could result in overestimation (Wu et al., 2023a; Yang et al., 2022b). An alternative is the sequential reaction method. It is based on measurements of the reaction products of these reactions.

Isoprene + $\cdot$ OH $\rightarrow$ 0.23MACR + 0.32MVK  $k_1$=1.0×10$^{-10}$ cm$^3$/molecule·s) (3)

MACR + $\cdot$ OH $\rightarrow$ products  $k_2$=3.3×10$^{-11}$ cm$^3$/molecule·s) (4)

MVK + $\cdot$ OH $\rightarrow$ products  $k_3$=1.9×10$^{-11}$ cm$^3$/molecule·s) (5)

\[
\frac{[\text{MACR}]}{[\text{Isoprene}]}_t = \frac{0.23k_3}{k_2-k_3}(1 - e^{(k_1-k_2)[\cdot \text{OH}][\Delta t]})
\] (6)

\[
\frac{[\text{MVK}]}{[\text{Isoprene}]}_t = \frac{0.32k_3}{k_2-k_3}(1 - e^{(k_1-k_3)[\cdot \text{OH}][\Delta t]})
\] (7)

where MVK is methyl vinyl ketone, MACR is methacrolein; $k_1$, $k_2$, and $k_3$ are the rate constants
reacting with •OH of isoprene, MACR, and MVK, respectively; [MVK]$_t$, [MACR]$_t$, and [isoprene]$_t$ are the measured ambient concentrations at time $t$.

This sequential reaction method was first proposed by Bertman et al. (1995) and was commonly used to estimate the $\Delta t$ of isoprene emitted primarily from biogenic emissions (Gong et al., 2018; Roberts et al., 2006; Wu et al., 2023a). Assuming the •OH-driven isoprene oxidation mechanism (Eqs. (3)-(5)) from the laboratory chamber studies of Carter and Atkinson (1996), Stroud et al. (2001) quantified the $\Delta t$ of isoprene by the reaction relationship between isoprene and its oxidation products (Eqs. (6)-(7)) (de Gouw et al., 2005; Wu et al., 2023a; Xie et al., 2008). The synchronous measurement of MVK and MACR is critical to obtain $\Delta t$ utilizing this method. Numerous studies failed in estimating the isoprene $\Delta t$ due to the lack of MVK and MACR (Gu et al., 2023; Liu et al., 2023a; Wang et al., 2023). Although some studies have used the average of $\Delta t$ obtained from both MVK and MACR as the final $\Delta t$ (Xie et al., 2008), there has been little assessment of which method was more suitable.

Additionally, there were studies using the isotopic hydrocarbon clock method (Kornilova et al., 2016; Rudolph and Czuba, 2000; Saito et al., 2009) to estimate the species $\Delta t$, that was calculated by the decay of isotopes in the emissions to the ambient receptor site (Table S5). However, the studies only calculated the $\Delta t$ without estimating the species initial concentrations. Our study found that approximately 76% of the publications that calculated the $\Delta t$ used the species ratio method, approximately 10% used both the ratio method and the sequential reaction model, while a few used other methods (Table S5).

### 3.1.2 Reducing the losses impacts in OVOC source analyses based on photochemical age

Ambient OVOCs are an important fraction of VOCs, which primarily included aldehydes, ketones, ethers, and alcohols (Mellouki et al., 2015) since OVOC photolysis is one of the main sources of •OH radicals (Li et al., 2018b). The OVOCs can be directly emitted from biogenic and anthropogenic sources (Huang et al., 2019; Huang et al., 2020; Tanimoto et al., 2014) and can also be formed by oxidation with •OH radicals, O$_3$, and NO$_3$ radicals (Legreid et al., 2008; Sahu et al., 2016; Tanimoto et al., 2014). OVOCs were commonly apportioned by the photochemical age-based
parameterization method as shown in Eq. (8).

\[
[OVOC] = ER_{OVOC} \times [tracer] \times \exp(- (k_{OVOC} - k_{tracer})[\cdot OH] \Delta t) + ER_{precursor} \times \\
[tracer] \times \frac{k_{precursor}}{k_{OVOC} - k_{precursor}} \times \frac{\exp(-k_{precursor}[\cdot OH] \Delta t) - \exp(-k_{OVOC}[\cdot OH] \Delta t)}{\exp(-k_{tracer}[\cdot OH] \Delta t)} + ER_{biogenic} \times \\
(\text{isoprene source}) + [\text{background}]
\]  

(8)

where [OVOC], [tracer], and [background] are the concentrations of measured ambient OVOCs, the tracer (e.g., benzene, acetylene: C₂H₃, or carbon monoxide: CO, Table S7) from primary anthropogenic sources, and background OVOC concentration, respectively; \(k_{OVOC}\), \(k_{tracer}\), and \(k_{precursor}\) are the \(\cdot OH\) rate constants of the OVOCs, tracer, and precursor, respectively, and \(k_{OVOC}\) and \(k_{tracer}\) are commonly obtained from the related publications (Atkinson and Arey, 2003); \(ER_{OVOC}\) and \(ER_{precursor}\) are the emission ratios of OVOCs and precursors relative to the tracer, respectively, \(ER_{biogenic}\) is the emission ratio of OVOCs to the isoprene concentration emitted from biogenic sources (i.e., \(\text{isoprene source}\)), and \(ER_{OVOC}\), \(ER_{precursor}\), \(ER_{biogenic}\), \(k_{precursor}\), and [background] can be determined from a linear least-squares fit that minimizes the difference between the measured [OVOC] and those calculated from Eq. (8), ER values in the reviewed papers are listed in Table S8; \([\cdot OH] \Delta t\) represents the exposure of \(\cdot OH\) radicals, which can be estimated by Eq. (2); and \(\text{isoprene source}\) can be estimated by Eqs. (6)-(7) and (1) based on the measured concentrations of ambient isoprene and its photochemical products (i.e., MVK and MACR). The chemical removal of isoprene in the atmosphere is so rapid that it is impossible to differentiate between primary and secondary OVOC sources based on the measured data. Therefore, \(\text{isoprene source}\) represents both primary and secondary biogenic sources (de Gouw et al., 2005).

However, this method remains highly uncertain because of the source complexities of OVOCs (Mo et al., 2016; Schlundt et al., 2017). This method assumes that (1) anthropogenic emissions of OVOCs and their precursors are proportional to the selected primary tracer, (2) the removal process of OVOCs is dominated by reactions with \(\cdot OH\) radicals, (3) biogenic sources of OVOCs are proportional to the emission of isoprene, and (4) the photochemical age for a sampled air mass can be determined (de Gouw et al., 2005).

This method estimates the source contributions of primary and secondary anthropogenic sources, biogenic emissions, and background to different OVOC species, and then further obtained the contributions of the four types of sources to the ambient OVOCs (de Gouw et al., 2005; Yuan et al.,
2012b; Zhu et al., 2021). This approach was difficult to finely apportion the contributions of primary anthropogenic sources and to obtain the contributions of primary and secondary biogenic emissions. Therefore, to obtain refined source apportioned results for OVOCs, many studies incorporated both OVOC and PAMS species into the receptor model for source analyses (Guan et al., 2020; Yang et al., 2019; Zhou et al., 2022).

This method defines the measured concentrations of OVOCs as the sum of the concentrations after photochemical losses (i.e., losses caused by the formation of O$_3$ and SOAs via photochemical reactions) of OVOCs directly emitted by anthropogenic sources (i.e., the first term in Eq. (8)), the concentrations after photochemical losses of OVOCs formed by the conversion of precursors emitted from anthropogenic sources (i.e., the second term in Eq. (8)), the concentration from biogenic emissions (without considering losses because of small contributions of biogenic emissions) (i.e., the third term in Eq. (8)), and the background concentration (i.e., the fourth term in Eq. (8)). Therefore, although this approach considered the influences of photochemical losses in the calculation process, the final results only reflected the contributions of four types of sources to the measured OVOCs and could not characterize the impacts of photochemical losses. In addition, to compensate for the photolytic losses in OVOCs in Eq. (8), which was not considered by de Gouw et al. (2005), Huang et al. (2020) and Zhu et al. (2021) introduced a modification coefficient (m) before the $k_{OVOC}$ to modify it, assuming that the photolysis rate is proportional to the •OH reaction rate (de Gouw et al., 2005; Wang et al., 2017). The m value depends on the relative rate of photolysis versus the •OH reaction for an OVOC species. However, related studies remain limited at present.

3.1.3 Methods for reducing loss effects in CMB source apportionments

The chemical mass balance method uses known profiles measured at the source. Thus, any reactions in transit will result in a change in that source’s profile and difficulty in fitting the data at the receptor site. Thus, corrections to the receptor site data are needed to make them comparable to the measured source profiles. As early as the 1980s and 1990s, studies had been conducted to reduce the impacts of different VOC species reactions on CMB estimation results, including minimization of reaction effects by limiting source profiles to VOCs with similar reaction rates (Harley et al., 1992; Lewis et al., 1993; Nelson and Quigley, 1983; Wadden et al., 1986) and VOC samples obtained.
during winter (Aronian et al., 1989) or early in the morning (Scheff and Klevs, 1987). In 1983, Nelson and Quigley (1983) estimated the reactions extent by the changes in the ratios of xylene to ethylbenzene at the receptor and the release site and then obtained decay factors of other VOC species. However, it was found that the concentrations adjusted by decay factors at the receptor site had little impact on their estimated source contributions (Nelson and Quigley, 1983). Since the emissions varied with time, especially during periods of intense photochemical reactions (Lin and Milford, 1994), some studies tried to use decay factors to adjust the fractions of VOC species in the source profiles at different times, making them more compatible with the data measured at the receptor site (Lin and Milford, 1994; Na and Pyo Kim, 2007). However, this approach has not been used other than in these two instances.

There are two methods to obtain the decay factors: one method was to considering an urban airshed as a continuous stirred tank reactor and relating the decay factor for a given species to its first-order reaction rate constant (Friedlander, 1981). The other method was to conduct estimation based on the reaction rate constants of specific VOC species and “aging coefficients” (Junninen et al., 2006; Lin and Milford, 1994). The details and limitations of the two methods are provided in Text S1. However, this method of adjusting source profiles cannot truly address the issue of reactive losses affecting the CMB apportioned results. Since the receptor measured data and adjusted source profiles input to CMB were both data after reactive losses. With the progress of VOC source analyses studies in recent years, this method has not been widely applied and further developed in CMB source analyses. This change might be because CMB itself required input from VOC source profiles, but there were relatively few locally measured VOC source profiles due to high costs and difficult to obtain access to the source facilities (Yang et al., 2022b).

Alternatively, with the monitoring of highly time resolved VOCs data in recent years, it became infeasible to use CMB to conduct source analyses, because CMB requires analysis of one sample (one hour) at a time resulting in needing thousands of separate analyses (Yang et al., 2022b). Additionally, the weights in the analyses are generally based only on the measurement uncertainties of the measurement samples from specific sources at a few specific times. Thus, CMB does not account for variability in the source profiles that is included in the uncertainty values used to weight the data in PMF (Yang et al., 2022b). To reduce the impact of reactive losses, the initial concentration
data estimated by photochemical age-based parameterization method has been used as input into CMB for source analyses (Shao et al., 2011). Additionally, the methods for reducing the impacts of reactive losses in PMF and CMB primarily focused on the PAMS species, and few considerations for OVOCs.

3.2 Effects of reactive losses on source analyses

According to the review of relevant publications, it has been found that most of the current studies analyzed the impacts of VOC photochemical losses on the source contributions by comparing the PMF apportioned results based on the measured and initial concentrations estimated by a photochemical age-based parameterization method (Gao et al., 2018; Kong et al., 2023; Li et al., 2023; Liu et al., 2023c; Zou et al., 2023). For example, Wang et al. (2013) found that the contributions of biogenic and industrial emissions were underestimated by 30% and 10%, respectively, due to photochemical losses. Liu et al. (2023a) suggested that biogenic emissions and polymer production-related industrial sources in Tianjin were underestimated by 73% and 50%, respectively. Wu et al. (2023a) also reported that the underestimation of the contribution of biogenic emissions was substantially higher than other sources. In 2023, Liu et al. (2023a) systematically investigated the impacts of photochemical losses of VOCs on the PMF apportioned factor profiles and contributions by comparing the estimated results based on the initial and measured concentrations, and found that the photochemical losses reduced the concentrations of these highly reactive species in the apportioned factor profiles based on the measured data. Meanwhile, Gu et al. (2023) also obtained similar results. However, due to the uncertainty of initial concentration estimation, the reliability of these results needs further validation and evaluation. Thus, further research is required.

3.3 Estimation methods of VOC reactive losses

There were two main methods for estimating the reactive losses of VOCs in the atmosphere. The first was to estimate the initial concentrations of VOCs utilizing the photochemical age-based parameterization method (Liu et al., 2023a; Wu et al., 2023b), and the differences between the initial and the measured concentrations were considered to the photochemical losses of VOCs (Table S9) (as shown in Eq. (9)) (Wang et al., 2023; Wu et al., 2023b). Due to the uncertainty in the initial
concentration estimation (as shown in section 3.1.1.2), there were also uncertainties in the estimated
photochemical losses.

\[
[\text{VOC}]_{C,t} = [\text{VOC}]_t \times \left( \exp(k_{\text{VOC}}[\cdot \text{OH}]\Delta t) - 1 \right)
\]  

(9)

where \([\text{VOC}]_C\) represents the photochemical loss of VOC at time \(t\), \(K_{\text{VOC}}\) represents the reaction rate
constants with •OH radicals of VOC, and \([\cdot \text{OH}]\Delta t\) represents the exposure of •OH radicals.

The second method applies to other VOC species without isoprene in which the losses could be estimated utilizing isoprene conversion (Wiedinmyer et al., 2001). In this study, it was defined as the isoprene loss reference method (Table S9). Its principle is to first use the photochemical-age parameter method based on the sequential reaction model as shown in Eqs. (1) and (6)-(7), to estimate the photochemical loss of isoprene (i.e., \(\Delta \text{isoprene}\)), and then the photochemical losses of other VOC species can be estimated by proportion relationship \(\frac{k_{\text{VOC}}}{k_{\text{isoprene}}} \times \frac{[\text{VOC}]_t}{[\text{isoprene}]_t}\) between \(\Delta \text{isoprene}\) and other species losses (Wiedinmyer et al., 2001; Xie et al., 2008) (as shown in Eq. (10)).

This method assumes that the relative source strengths of VOCs are constant in an immediate area surrounding the site, that atmospheric transport and dispersion are non-limiting factors compared with chemistry (Xie et al., 2008), and the photochemical losses of other VOC species can be calculated by Eq. (10).

\[
[\text{VOC}]_{C,t} = \Delta \text{isoprene} \times \frac{k_{\text{VOC}}}{k_{\text{isoprene}}} \times \frac{[\text{VOC}]_t}{[\text{isoprene}]_t}
\]  

(10)

where \(k_{\text{isoprene}}\) represents the reaction rate constants with •OH radicals of isoprene, and \([\text{isoprene}]_t\)
represents the measured concentration of isoprene at time \(t\).

Because isoprene mainly originates from natural sources (Fu et al., 2008; Kuhn et al., 2004; Lu et al., 2019), while many other VOC species are primarily emitted from anthropogenic sources (Li et al., 2021; Seinfeld and Pandis, 1986; Wei et al., 2011). Therefore, the assumption of this method itself has an obvious issue, which resulted in high uncertainties in the estimations. This defect substantially limits its application, only 2 publications have utilized this method to quantify chemical losses of VOCs in this study (Wiedinmyer et al., 2001; Xie et al., 2008), and most studies still used the first method (Table S9).

### 3.4 Spatiotemporal variation of VOC reactive losses

According to the papers reviewed in this study, we found that studies on the VOC
photochemical losses were mainly conducted in Chinese cities. The VOC losses in different cities are provided in Figs. 2-3 and S1 and Table S10. Due to relatively limited number of studies, the representativeness of the analyses of the spatiotemporal distributions of photochemical losses of VOCs might also be limited. In addition, this study converted the unit of ppbC in some papers (Chen et al., 2023) to ppbv for better comparative analyses.

Figure 2. The measured, initial, and reactive loss concentrations of ambient VOCs in Beijing (Gao et al., 2018; Ma et al., 2022; Zhan et al., 2021), Tianjin (Liu et al., 2023a; Wang et al., 2023), Shanghai (Wang et al., 2013), Chengdu (Kong et al., 2023), Guangzhou (Wang et al., 2023), Qingdao (Gu et al., 2023), and Taipei (Chen et al., 2023). The data in Beijing was the average from three published papers data. The horizontal axis represents longitude (°E), and the vertical axis represents latitude (°N). The base map is from Natural Earth.

The photochemical loss of VOCs in Qingdao was the highest (45.1 ppbv), followed by Tianjin (2.9-17.8 ppbv), Shanghai (9.00-9.60 ppbv), Taipei (0.76-8.29 ppbv), Beijing (3.40-7.72 ppbv), Chengdu (2.30-6.60 ppbv), and Guangzhou (5.20 ppbv). The chemical loss rate (i.e., chemical loss/initial concentration) in Qingdao (69.1%) was the highest, and the lowest was in Taipei (2.97%-
Compared to other VOC groups, alkenes had the highest loss (Figs. 3 and S1), accounting for 55.4%-93.3% of the total losses, followed by aromatic hydrocarbons (3.81%-38.7%), and alkanes (2.33%-22.7%) (Table S11). There were substantial differences in VOC species with high losses in different cities (Fig. 3). The losses of ethene, propene, and isoprene in most cities were relatively higher than those of other species (Fig. 3), likely closely related to their high reactivities (Table S4).

The reactive losses of 1-butene and trans-2-butene in Qingdao were 10.96 and 9.96 ppbv, substantially higher than other species, while the loss of 1-hexene (9.50 ppbv) in Tianjin was remarkably higher compared to other species. However, the losses of isoprene in Beijing, Chengdu, and Taipei were the highest (Fig. 3).

**Figure 3.** The photochemical losses and percentages of the main VOCs in Beijing (a: cited from publication (Gao et al., 2018) and b: cited from publication (Zhan et al., 2021)), Qingdao (Gu et al., 2023), Taipei (Chen et al., 2023), Tianjin (Liu et al., 2023a), and Chengdu (Kong et al., 2023).
In addition, the VOC reactive losses in spring and summer were substantially higher than those in autumn and winter. Atmospheric oxidation and meteorological factors can have impacts on the VOC losses. The trend of the reactive losses of VOCs was often consistent with the changes in the concentrations of $\text{O}_x$ ($\text{O}_3+\text{NO}_2$) in the atmosphere (Wang et al., 2013). In addition, studies also found that both solar radiation and temperature can have a substantial impact on the chemical reactions of VOCs (Yadav et al., 2016). However, compared to the solar radiation, temperature directly affects the reaction rates of VOC species (Atkinson and Arey, 2003). Liu et al. (2023a) suggested that the reactive losses of ambient VOCs were highly correlated with temperature; when the temperature was above 25 ℃, the losses of VOCs increased most substantially.

3.5 Source analyses of VOC reactive losses in the atmosphere

At present, source apportionments utilizing the measured and initial VOC concentrations do not consider the complementary issues of $\text{O}_3$ and SOA formation (Gu et al., 2023; Wang et al., 2022b). The chemical losses of VOCs by reaction (i.e., reactive VOCs forming SOA and $\text{O}_3$) were real contributors to these pollutants (Ma et al., 2022; Wang et al., 2013). Thus, it is important to apportion the reactive losses of VOCs to provide the input needed to identify the responsible source types and thereby allow effective control of $\text{O}_3$ and secondary aerosol pollution. In 2023, Gu et al. (2023) developed a source analyses method for consumed VOCs (i.e., the reactive losses of VOCs) in the atmosphere and conducted an applied study. This method was that the measured and initial data were first used to conduct source analyses by PMF, and then the difference in the contributions of the same factors from the paired analyses were considered to be the source contribution of consumed VOCs in the atmosphere. The ozone formation potential (OFP) and SOA formation potential (SOAFP) of the consumed VOCs from the different sources were estimated. To reduce the impacts of atmospheric dispersion on the apportioned results, dispersion-normalized PMF (DN-PMF) was applied for source analyses in this method, and its principle is provided by Dai et al. (2020) and Gu et al. (2022). The method flowchart is shown in Fig. 4. Similar studies were reported by Wang et al. (2022b) and Wang et al. (2023). However, the current method for apportioning the sources of consumed VOCs in the atmosphere is still imperfect. The shortcomings included: the uncertainty in estimating the initial concentrations of VOCs (see section 3.1.1.2); the issue of factor identification, even if identified as the same type of factor, there were substantial differences in their factor profiles.
and the current studies focused on the PAMS species, and apportionments of OVOC losses have not yet been conducted. There were numerous technical issues, for example, there is still a lack of rational methods for estimating the OVOC initial concentrations.

The source analyses of the consumed VOCs conducted by Gu et al. (2023) in Qingdao found that biogenic emissions (56.3%), vehicle emissions (17.2%), and gasoline evaporation (9.37%) were the main sources of the consumed VOCs. However, the apportioned results from Wang et al. (2023) in Tianjin and Guangzhou suggested that biogenic emissions (43% and 35%, respectively), solvent usage (14% and 18%, respectively), and industrial sources (14% and 22%, respectively) were the major contributors of the consumed VOCs. Therefore, biogenic emissions may be an important source of reactive losses of VOCs. However, due to the limited research currently available, the representativeness of the results is still insufficient at present.

**Figure 4.** The method flowchart of source analyses for consumed VOCs in the atmosphere.
4 Conclusions

This study systematically reviewed the major methods of reducing the impacts of reactive loss and the existing critical issues in the current VOC source analyses research. The purpose was to clarify the future research directions needed to improve the accuracy of VOC source apportionments, and provided a potential supporting role in completely solving the issue of VOC chemical conversion. PMF is currently the most useful tool in treating highly time-resolved data compared to other receptor models. Estimating the initial concentrations of ambient VOCs based on photochemical age is the primary approach to reduce reactive loss effects in PMF. However, due to the shortcomings existing in the photochemical-age estimation method, such as only considering the photochemical reactions with \( \cdot \text{OH} \) radicals during the day and not considering the reactions with \( O_3 \) (especially for alkene species) and \( \text{NO}_3 \) radicals, difficulty in dynamically adjusting the required \( \cdot \text{OH} \)-radical reaction rate constants with temperature changes, high uncertainty in the estimation of photochemical age, and difficulty in quantifying the initial concentrations of OVOCs, etc., resulting in substantial uncertainty in the PMF apportioned results based on the initial concentration data. Furthermore, both the measured and initial VOCs data do not match the ambient \( O_3 \) or SOAs measured synchronously, the VOCs consumed by reactions are real contributors to these pollutants. Source analyses of the consumed VOCs can effectively guide the prevention and control of \( O_3 \) and SOAs pollution in the atmosphere. Therefore, in addition to the research into more accurate quantitative approaches for ambient VOC reactive losses, the source analyses of consumed VOCs and the accurate quantification of their separate contributions to \( O_3 \) and SOAs should also be important directions for future research.

Author contributions

BSL: Writing–original draft, Writing–review & editing. YG: Data curation, Writing–original draft. YTW: Data curation. QLD and SJS: Investigation and Supervision. YCF: Supervision. PKH: Supervision–review & editing.

Competing interests

The author declares that they have no competing interests.
Acknowledgment

This study was financially supported by National Key R & D Program of China (No. 2023YFC3705801) and National Natural Science Foundation of China (No. 42177085).

References


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
isoprene observed at a forested mountaintop site in southern China: implications for strong
regional atmospheric oxidative capacity, Atmos. Chem. Phys., 18, 14417-14432,
Gu, Y., Liu, B. S., Li, Y. F., Zhang, Y. F., Bi, X. H., Wu, J. H., Song, C. B., Dai, Q. L., Han, Y., Ren,
G., and Feng, Y. C.: Multi-scale volatile organic compound (VOC) source apportionment in
Tianjin, China, using a receptor model coupled with 1-hr resolution data, Environ. Pollut., 265,
improved positive matrix factorization for source apportionment of volatile organic compounds
during the COVID-19 shutdown in Tianjin, China, Environ. Int., 158, 106979,
apportionment of consumed volatile organic compounds in the atmosphere, J. Haz. Mat., 459,
Temporal variations and source apportionment of volatile organic compounds at an urban site in
Han, Y., Huang, X. F., Wang, C., Zhu, B., and He, L. Y.: Characterizing oxygenated volatile organic
compounds and their sources in rural atmospheres in China, J. Environ. Sci., 81, 148-155,
detection of excess unburned gasoline in the atmosphere, Environ. Sci. Technol., 26, 2395-2408,
He, Z. R., Wang, X. M., Ling, Z. H., Zhao, J., Guo, H., Shao, M., and Wang, Z.: Contributions of
different anthropogenic volatile organic compound sources to ozone formation at a receptor site in
the Pearl River Delta region and its policy implications, Atmos. Chem. Phys., 19, 8801-8816,
Huang, X.-F., Wang, C., Zhu, B., Lin, L.-L., and He, L.-Y.: Exploration of sources of OVOCs in
various atmospheres in southern China, Environ. Pollut., 249, 831-842,
Huang, X.-F., Zhang, B., Xia, S.-Y., Han, Y., Wang, C., Yu, G.-H., and Feng, N.: Sources of
oxygenated volatile organic compounds (OVOCs) in urban atmospheres in North and South China,


Xie, Y. L., and Berkowitz, C. M.: The use of positive matrix factorization with conditional probability functions in air quality studies: An application to hydrocarbon emissions in Houston,


Zhang, W. J., Lin, S., Hopke, P. K., Thurston, S. W., van Wijngaarden, E., Croft, D., Squizzato, S.,


