# Intermediate-Volatility Organic Compounds: Source Emission Profiles, Resolved Chemical Composition and Contribution to Secondary Organic Aerosol Formation

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Department of Mechanical Engineering

Quanyang Lu

B.E., Thermal Engineering, Tsinghua University
M.S., Mechanical Engineering, Carnegie Mellon University
M.S., Machine Learning, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

August, 2020

# Acknowledgements

I feel truly grateful for the last five years spent at Carnegie Mellon University. There have been ups and downs along the way, but life is all about the journey, and about to find the true self. Being around so many great people in this community, I have experienced the best interdisciplinary environment in the world.

First and foremost, I want to thank my advisor Allen Robinson with all my heart. I learned a ton from him on conducting research, scientific writing and presentation, and most importantly big-picture thinking. We have been working hard on state-of-the-art scientific questions and pushing the boundaries. He has been very supportive for the past five years, especially when I need the support the most. He has also supported me in pursuing my interest in machine learning, which has led to an important chapter in this thesis and my future career as a data scientist.

I would like to thank my committee members: Peter Adams, Albert Presto, Jeremy Michalek and Benjamin Murphy. They have been very helpful on providing many insightful comments on my thesis work. They also encouraged me to combine the best research of both experimental and simulation worlds which eventually lead to the completion of this work. I feel very lucky working with all committee members, and this thesis would not be possible without them.

I want to thank Benjamin Murphy and Havala Pye as my internship advisors at EPA. The three-month internship was a wonderful experience. Ben has been a great advisor and friend and helped me settle down in that area at Durham, NC. He has taught me how to conduct simulations from the basics, and we worked together on the exciting topics, which has also led to a chapter in my thesis. Havala is super helpful on a lot of the discussion topics, and has provided many helpful comments on this work.

I also want to thank Yunliang Zhao for his help and guidance in the first two years of my PhD study. He has taught me a lot about conducting experiments and data analysis, and always offered help whenever I have questions.

I want to thank Venkat Viswanathan and Jon Malen for the great TA experiences with them. They teach with great enthusiasm in student's learning and always try to engage us in the entire teaching experiences. I appreciate the opportunities chatting with Venkat and shared some of my thoughts on PhD study and future plan.

I would like to offer my thanks to the staff in the Mechanical Engineering department, Chris Hertz, Mi Kim, Kate Sencindiver, Melissa Brown, Meredith Blobner, Alex Caprara and Amy Carroll. I really appreciate all the help throughout my PhD study. I would like to thank my collaborators: Momei Qin, Chris Hennigan, Rongzhi Tang and Song Guo for their efforts and helpful comments. I want to thank Neil Donahue and Ryan Sullivan for their help and insightful comments on my thesis work.

I want to thank Xiang Li, Georges Saliba, Rose Eilenberg, Peishi Gu, Zhongju Li, Rishabh Shah, Penglin Ye, Qing Ye, Lydia Jahl, Ningxin Wang, Meredith Schervish, Tom Brubaker, Naser Mahfouz, Kerrigan Cain, Jiqiao Shi, Dana McGuffin, Katerina Liangou, Mingyi Wang, Fan Tong, Spiro Jorga and Jiayu Li for the help and accompany at CAPS. I want to thank Jack Johnson, Brian Dinkelacker, Luke Habib, Farrah Haeri, Jingzhi Wang and Jian Zhao for many chats in our PhD office. I want to thank all friends at CAPS, MechE and CMU for the time spent together.

Last but not least, I probably spent most of my free time in the gym, swimming pool and basketball courts. I am grateful for the nice sports facilities, and I really value the friends and the memorable times that we trained and played together. I also want to thank Bon Jovi's song "Livin' on a Prayer", its lyrics once gave me the courage: "Take my hand, we'll make

it, I swear. Woah-oh, livin' on a prayer. ”

My thesis work was funded by the Center for Air, Climate, and Energy Solutions (CACES), which is supported under Assistance Agreement No. RD83587301 awarded by the U.S. Environmental Protection Agency. Part of my thesis work was also supported by the Oak Ridge Institute for Science and Education (ORISE) Research Participation Program for the U.S. Environmental Protection Agency.

# Abstract

Atmospheric aerosols are tiny particles or droplets suspended in the air. Aerosols with diameters less than 2.5 microns are fine particulate matter, or $PM_{2.5}$. Human exposure to fine particulate matter has been associated with higher mortality, stroke, lung cancer and cardiovascular diseases. Organic aerosol (OA) contributes significantly to fine particulate matter mass globally. Intermediate-volatility organic compounds (IVOCs) are a group of organic compounds that has effective saturation concentration (C*) at 298K between $10^3$ and $10^6$ $\mu g/m^3$ (roughly equivalent to $C_{12}$ to $C_{22}$ *n*-alkanes). They are important precursors to secondary organic aerosol (SOA) formation.

To investigate the contribution of IVOCs to ambient SOA formation, it is crucial to correctly model them with emission, volatility distribution and chemical composition from direct measurements. Traditional emission inventories such as the National Emission Inventory only account for gas-phase organic emissions of volatile organic compounds (VOCs). There are large gaps in IVOCs and semi-volatile organic compounds (SVOCs) between traditional emission profiles and recent experiment results. Also, despite recent efforts to include IVOCs and SVOCs in inventories, most of the model assumptions are overly simplistic and scaled based on limited experimental data. For example, if they are included, IVOCs are often represented using a single surrogate species. No emission profiles or simulations have been developed to incorporate the most recent experimental data that measured comprehensive organic species in a wide range of volatility spectrum.

This thesis consists of three main parts: compiling comprehensive organic emission profiles; implementing SOA parameterization in a chemical transport model (CTM) and run simulations; and developing a positive matrix factorization (PMF) technique to resolve the IVOC chemical composition from traditional gas chromatography-mass spectrometry (GC-

MS) data.

Organic emission data on gasoline and diesel vehicles and aircraft emissions from previously published papers are compiled to create comprehensive emission profiles for use in CTMs. The organic emissions from all three source categories show tri-modal volatility distributions, and the volatility distributions are consistent across sources using the same fuel type. The traditional profiles dramatically underestimate IVOCs and SVOCs, which are important classes of SOA precursors. Accounting for IVOCs in gasoline exhaust almost doubles the predicted SOA production compared to the traditional profile. For gas-turbine and diesel sources, IVOCs and SVOC vapors combining contribute factors of 13 (gas-turbine) and 44 (diesel) more SOA than VOCs alone.

We developed a new parameterization to model the SOA formation from mobile source IVOC emissions designed for implementation in CTMs. The parameterization has six lumped IVOC species: two aromatics and four aliphatics, to account for the volatility and chemical composition of the IVOC emissions. Simulation results show that mobile sources contribute 2.7 μg $m^{-3}$ of IVOCs at Pasadena site, which is 43% of measured concentrations of hydrocarbon IVOCs. They also contribute ~1 μg $m^{-3}$ in daily peak SOA concentration, a 67% increase compared to the base case without IVOC emissions. Therefore, it is crucial to include mobile-source IVOC emissions in simulations. Results from the exploratory model runs suggest that additional 12% to 26.8% of non-mobile organic emissions are likely IVOCs.

We also developed a PMF-based technique to resolve the chemical composition of IVOCs in traditional GC-MS data. Evaluation on multiple datasets shows this technique can recover much of the chemical information compared to more sophisticated instruments. Source apportionment analysis on tunnel samples shows major contributions from diesel-

source IVOCs. PMF analysis on ambient samples in Pasadena, California shows more than 70% IVOCs are oxygenates indicating complex atmospheric oxidation. SOA modeling of gasoline vehicle emissions shows an 80% SOA yield increase under low-$NO_x$ conditions, highlights the need to include the IVOC aromatics in resolved chemical composition and CTM studies.

This thesis presents systematic efforts to better understand IVOCs and incorporate them into the model, from emission to chemical composition and SOA formation. To include IVOCs in CTM simulations, we compiled model-ready emission profiles for mobile-sources from direct measurements. To simulate the SOA formation from IVOCs and account for chemical composition, we developed a PMF technique to extract necessary information from traditional GC-MS data. And to evaluate the contribution of IVOCs to ambient OA, we implemented SOA parameterization and perform CTM simulations. This pipeline could be reused in the future study of other sources with IVOC emissions.

# Table of Contents

# List of Tables

# List of Figures

# Glossary

AMS: Aerosol Mass Spectrometer

BC: Black Carbon

CMAQ: The Community Multiscale Air Quality Modeling System

CMB: Chemical Mass Balance

CTM: Chemical Transport Model

DPF: Diesel Particulate Filter

EC: Elemental Carbon

EPA: Environmental Protection Agency

GC-MS: Gas Chromatography-Mass Spectrometry

HDDV: Heavy Duty Diesel Vehicle

HOA: Hydrocarbon-like Organic Aerosol

IVOC: Intermediate-Volatility Organic Compound

LDGV: Light Duty Gasoline Vehicle

LEV: Low Emission Vehicle

LVOC: Low-Volatile Organic Compound

NEI: National Emissions Inventory

NIST: National Institute of Standards and Technology

NMOG: Non-Methane Organic Gases

NVOC: Non-Volatile Organic Compound

OA: Organic Aerosol

OC: Organic Carbon

OOA: Oxygenated Organic Aerosol

PM: Particulate Matter

PMF: Positive Matrix Factorization

POA: Primary Organic Aerosol

SOA: Secondary Organic Aerosol

SORE: Small Off-Road Engine

SRA: Single-Ring Aromatics

SVOC: Semi-Volatile Organic Compound

UCM: Unresolved Complex Mixture

ULEV: Ultra Low Emission Vehicle

VOC: Volatile Organic Compound

# Chapter 1: Introduction

## 1.1 Motivation

Atmospheric aerosol are tiny particles or droplets suspended in the air. Aerosol with diameter less than 2.5 microns are classified as fine particulate matter, or $PM_{2.5}$. Human exposure to $PM_{2.5}$ has been associated with increased mortality, stroke, lung cancer and cardiovascular diseases (Apte et al., 2018; Burnett et al., 2018; Di et al., 2017) and influences climate (Kanakidou et al., 2005). Increases of 10 μg per cubic meter in $PM_{2.5}$ were associated with increases in all-cause mortality of 7.3%, and the adverse effect is still present with exposure to $PM_{2.5}$ of less than 12 μg per cubic meter (Di et al., 2017).

Organic aerosol (OA) contributes significantly (20-90%) to submicron atmospheric fine particulate matter mass globally (De Gouw and Jimenez, 2009; Jimenez et al., 2009; Kanakidou et al., 2005). Organic aerosol is classified as primary OA (POA), which is directly emitted by sources, such as from fossil fuel combustion, biomass burning, and other sources, and secondary OA (SOA), which is formed in the atmosphere through photo-oxidation gas-phase organics. A series of studies have shown that SOA accounts for a large fraction of the OA burden, even in urban areas (Hallquist et al., 2009; Jimenez et al., 2009; Zhang et al., 2007). Both POA and SOA concentrations depend on the gas-particle partitioning of a complex mixture of organics that span a wide range of volatility, including low-volatility organic compounds (LVOCs), semi-volatile organic compounds (SVOCs), intermediate-volatility organic compounds (IVOCs) and volatile organic compounds (VOCs) (Donahue et al., 2006; Goldstein and Galbally, 2007; Robinson et al., 2007).

A major challenge has been the underestimation of SOA in many chemical transport model (CTM) studies (Ensberg et al., 2014; Hodzic et al., 2010; Robinson et al., 2007). For example, simulation on OA formation during the 2010 CalNex campaign (Baker et al., 2015)

shows traditional VOC SOA precursors only explain 7% of the organic carbon at Pasadena, whereas aerosol mass spectrometer (AMS) measurements suggest 63% of organic carbon are secondary organic carbon (Hayes et al., 2013). Another simulation study (Woody et al., 2016) shows the modelled SOA concentrations were approximately 4 to 5.4 times lower than the AMS oxygenated organic aerosol (OOA), with the largest underestimation in the early afternoon peak. Robinson et al. (2007) proposed that potential contribution from IVOCs and SVOCs to SOA formation, and such SOA precursors need to be directly incorporated into CTM studies (Dzepina et al., 2009; Hodzic et al., 2010). This thesis examines the role that IVOCs potentially play in SOA formation.

IVOCs are important precursors to secondary organic aerosol (SOA) formation in chamber studies (Jathar et al., 2013; Presto et al., 2010; Tkacik et al., 2012; Zhao et al., 2018), oxidation reactor studies (Shah et al., 2020; Tkacik et al., 2014) and simulation studies (Hodzic et al., 2010; Jathar et al., 2014; Lu et al., 2019; Ma et al., 2017; Murphy et al., 2017; Pye and Pouliot, 2012; Robinson et al., 2007). They are organic compounds that have effective saturation concentration (C*) at 298K between $10^3$ and $10^6$ $\mu g/m^3$ (roughly equivalent to $C_{12}$ to $C_{22}$ *n*-alkanes) (Presto et al., 2012), are ubiquitous in the ambient atmosphere (Hunter et al., 2017; Zhao et al., 2014) and are emitted from anthropogenic and biogenic emission sources, such as mobile emissions (Drozd et al., 2019; Lu et al., 2018; Qi et al., 2019; Schauer et al., 1999; Zhao et al., 2015), coal combustion (Cai et al., 2019), biomass burning (Hatch et al., 2018), and volatile chemical products (VCPs) (Khare and Gentner, 2018; McDonald et al., 2018). Detailed emission profiles on volatility distribution and chemical composition of IVOCs is critical to accurately simulate atmospheric SOA and $PM_{2.5}$ concentrations.

Traditional emissions inventories such as the National Emission Inventory (NEI)

only account for gas-phase organic emissions of VOCs and POA as non-volatile particulate matter. Robinson et al. (2007) and Shrivastava et al. (2008) argued that this is an overly simplistic representation of organic emissions. Primary IVOCs and SVOCs have not been routinely implemented in models because of lack of the mass and chemical composition (Shrivastava et al., 2008). Mobile sources contribute about one-third of the anthropogenic organic emissions in the 2014 EPA NEI; they are an important source of POA and SOA precursor gases especially in urban environments (Gentner et al., 2012, 2017; McDonald et al., 2013). Recent studies have reported IVOCs, SVOCs and LVOCs emissions and gas-particle partitioning on POA emissions from mobile sources (Gordon et al., 2013; May et al., 2014; Presto et al., 2011; Zhao et al., 2015, 2016). Therefore, compiling comprehensive profiles for organic emissions including non-volatile organics (NV), LVOCs, SVOCs, IVOCs, VOCs for important sources are needed.

Previous studies have incorporated IVOCs in the chemical transport models (CTMs) with volatility basis set and aging mechanism (Hodzic et al., 2010; Koo et al., 2014; Murphy et al., 2017; Robinson et al., 2007; Shrivastava et al., 2008). For example, Robinson et al. (2007) shows oxidation of the SVOC and IVOC vapors produces a considerable amount of regional SOA, and improve the predicted SOA fraction and urban-to-regional OA ratios for four large cities. Murphy et al. (2017) shows semi-volatile POA and extra SOA precursors have improved the model performance on OA diurnal patterns compared with measurement data at four different sites.

Although accounting for IVOCs improves model performance, current approaches are likely overly simplistic and based on very limited experiment data (Hodzic and Jimenez, 2011; Pye and Seinfeld, 2010). For example, Hodzic and Jimenez (2011) proposed a simple anthropogenic emission surrogate emitted at the rate of 0.08 g per g of CO as SOA precursor

(including VOCs, IVOCs and SVOCs), Pye and Seinfeld (2010) used 66 times naphthalene as total IVOCs surrogate, and Woody et al. (2016) assumed 7.5 × POA based on previous estimations (Hodzic et al., 2010; Koo et al., 2014), applied to all emission source categories. While it may work for some emission sources, this approach did not account for differences in IVOC chemical composition, as well as its effect on SOA yield (Pye and Pouliot, 2012; Yee et al., 2013). A more sophisticated CTM implementation including IVOCs that account for variations by source class, volatility distribution and chemical composition is needed.

Another challenge is the need for chemical information on IVOC emissions. This is important to simulate the SOA formation because the SOA yield depends on IVOC volatility and molecular structure (Chacon-Madrid et al., 2010; Loza et al., 2014; Presto et al., 2010). Because of the exponential increase in isomers as carbon number increase, the vast majority of the IVOC/SVOC mass cannot be resolved at the molecular level using traditional gas chromatography-based (GC) techniques (Goldstein and Galbally, 2007). More than 80% of the measured IVOC mass is often left as unresolved complex mixture (UCM) (Presto et al., 2011; Zhao et al., 2014, 2015, 2016). In recent works, researchers have used limited mass spectral information to crudely classify the bulk UCM into one or two different chemical classes (Zhao et al., 2015, 2016). Alternatively, researchers used instruments such as 2D GC-MS or soft-ionization (GC-VUV-MS) to provide more insight into the UCM composition (Drozd et al., 2019; Hatch et al., 2018), but it comes at the cost of more sophisticated (i.e. expensive) instruments and more detailed analysis. A fast yet interpretable technique on traditional GC-MS data would be valuable to recover the chemical information in large amounts of traditional GC-MS data at a fraction of or no cost.

## 1.2 Objectives of Dissertation.

The objectives of this thesis are to improve the understanding of IVOCs and the tools to simulate the contribution of IVOCs to SOA formation in urban environments, with respect to compiling comprehensive source emission profiles, resolve the chemical composition of GC-MS measured IVOCs and simulate their contribution to organic aerosol formation at regional scales.

The first objective of the thesis is to develop model-ready comprehensive emission profiles for mobiles sources. We intend to do this by synthesizing recently measured VOC, IVOC, SVOC, LVOC and NVOC emission data. We intend to compare new profiles to the traditional SPECIATE profiles for similar sources. Finally, we intend to model the SOA yield of mobile sources exhaust using the new profiles, and compare the results with other estimating approaches.

The second objective is to develop a model-ready parameterization of SOA formation from IVOCs and to implement it in CTM simulations. Given the importance of molecular structure and volatility in SOA formation, we intend to develop a parameterization that efficiently (i.e. minimum number of species) accounts for these factors. We intend to include IVOCs using comprehensive emission profiles into CTM simulations and explore the SOA formation from mobile and non-mobile sources through difference cases. Proper chemical mechanism mapping needs to be developed to properly synthesize IVOC profiles into SOA-relevant model species.

The third objective is to develop a technique to resolve the chemical composition of IVOC UCMs measured by traditional GC-MS. We intend to evaluate the new technique by comparing resolved chemical composition with results from more sophisticated instruments. We intend to gain more insight in source apportionment using this analysis on tunnel and ambient IVOC datasets. Resolved chemical composition of IVOC UCMs could also provide

key information in SOA formation simulation described in the second objective.

All three objective represent the complete steps needed in incorporating IVOCs from direct measurements into CTM simulations for SOA formation studies.

### 1.3 Outline of Dissertation

Each chapter in this thesis corresponds to one objective mentioned above. Below, we provide a brief overview for each chapter.

In chapter 2, we complied all data including VOCs, IVOCs, SVOCs, LVOCs and NVOCs from previously published papers on gasoline and diesel vehicles and aircraft emissions (Gordon et al., 2013; May et al., 2014; Presto et al., 2011; Zhao et al., 2015, 2016) into comprehensive emission profiles in volatility basis set from nonvolatile to $C^* = 10^{11}$ $\mu g/m^3$ at 298K. Comparing new profiles to traditional SPECIATE profiles, we highlight the needs to includes IVOCs and SVOCs into comprehensive emission profiles, and also show the consistency in VOC composition. Different sets of IVOC parameterization (Gentner et al., 2012; Jathar et al., 2014; Murphy et al., 2017; Pye and Seinfeld, 2010; Robinson et al., 2007) are compared to our new profiles in a SOA formation box-model to highlight the need to use emission profiles compiled from measurement data in CTM studies.

In chapter 3, we developed a parameterization that has six lumped precursor species to account for both volatility and molecular structure (aromatic versus aliphatic). The IVOC emissions are added to the model scaled from non-methane organic gas (NMOG) according to the comprehensive emission profiles obtained in the first part of this thesis. The CTM simulations includes four cases to explore the contributions of both mobile and non-mobile sources to ambient IVOC level, as well as OA formation in Southern California during the CalNex campaign. Modeled ambient IVOC level is compared the literature value reported

in Zhao et al. (2014). Model OA outputs are evaluated using aerosol mass spectrometer (AMS) and filter measurements from CalNex campaign. All profiles used in this part have been released in SPECIATE 5.0.

In chapter 4, we developed a new positive matrix factorization (PMF) technique to resolve the IVOC chemical composition measured by traditional GC-MS. Evaluations on mobile sources and biomass burning datasets shows that resolved chemical composition is consistent with results reported using more sophisticated instruments on similar datasets (Drozd et al., 2019; Hatch et al., 2017). Most PMF resolved factors show strong correlation ($r > 0.9$) with the mass spectra of reference compounds in the NIST database. The technique is then used to analyze the compositional trend in various datasets. We also conducted source apportionment analysis on tunnel and ambient samples, showing large contribution from diesel vehicle hydrocarbon IVOCs and highly oxygenated IVOCs, respectively.

In chapter 5, we summarize the key findings and scientific contributions in this thesis and provide future work directions.

## 1.4 References

Apte, J. S., Brauer, M., Cohen, A. J., Ezzati, M. and Pope, C. A.: Ambient PM2.5 Reduces Global and Regional Life Expectancy, Environ. Sci. Technol. Lett., 5(9), 546–551, doi:10.1021/acs.estlett.8b00360, 2018.

Baker, K. R., Carlton, A. G., Kleindienst, T. E., Offenberg, J. H., Beaver, M. R., Gentner, D. R., Goldstein, A. H., Hayes, P. L., Jimenez, J. L., Gilman, J. B., De Gouw, J. A., Woody, M. C., Pye, H. O. T. T., Kelly, J. T., Lewandowski, M., Jaoui, M., Stevens, P. S., Brune, W. H., Lin, Y. H., Rubitschun, C. L. and Surratt, J. D.: Gas and aerosol carbon in California: Comparison of measurements and model predictions in Pasadena and Bakersfield, Atmos. Chem. Phys., 15(9), 5243–5258, doi:10.5194/acp-15-5243-2015, 2015.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S. S., Kan, H., Walker, K. D., Thurston, G. D., Hayes, R. B., Lim, C. C., Turner, M. C., Jerrett, M., Krewski, D., Gapstur, S. M., Diver, W. R., Ostro, B., Goldberg, D., Crouse, D. L., Martin, R. V, Peters, P., Pinault, L., Tjepkema, M., van Donkelaar, A., Villeneuve, P. J., Miller, A. B., Yin, P., Zhou, M., Wang, L., Janssen, N. A. H., Marra, M., Atkinson, R. W., Tsang, H., Quoc Thach, T., Cannon, J. B., Allen, R. T., Hart, J. E., Laden, F., Cesaroni, G., Forastiere, F., Weinmayr, G., Jaensch, A., Nagel, G., Concin,

H. and Spadaro, J. V: Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter, Proc. Natl. Acad. Sci., 115(38), 9592 LP – 9597, doi:10.1073/pnas.1803222115, 2018.

Cai, S., Zhu, L., Wang, S., Wisthaler, A., Li, Q., Jiang, J. and Hao, J.: Time-Resolved Intermediate-Volatility and Semivolatile Organic Compound Emissions from Household Coal Combustion in Northern China, Environ. Sci. Technol., 53(15), 9269–9278, doi:10.1021/acs.est.9b00734, 2019.

Chacon-Madrid, H. J., Presto, A. A. and Donahue, N. M.: Functionalization vs. fragmentation: n-aldehyde oxidation mechanisms and secondary organic aerosol formation, Phys. Chem. Chem. Phys., 12(42), 13975–13982, doi:10.1039/C0CP00200C, 2010.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F. and Schwartz, J. D.: Air Pollution and Mortality in the Medicare Population, N. Engl. J. Med., 376(26), 2513–2522, doi:10.1056/NEJMoa1702747, 2017.

Donahue, N. M., Robinson, A. L., Stanier, C. O. and Pandis, S. N.: Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics, Environ. Sci. Technol., 40(8), 2635–2643, doi:10.1021/es052297c, 2006.

Drozd, G. T., Zhao, Y., Saliba, G., Frodin, B., Maddox, C., Oliver Chang, M.-C., Maldonado, H., Sardar, S., Weber, R. J., Robinson, A. L. and Goldstein, A. H.: Detailed Speciation of Intermediate Volatility and Semivolatile Organic Compound Emissions from Gasoline Vehicles: Effects of Cold-Starts and Implications for Secondary Organic Aerosol Formation, Environ. Sci. Technol., 53(3), 1706–1714, doi:10.1021/acs.est.8b05600, 2019.

Dzepina, K., Volkamer, R. M. R., Madronich, S., Tulet, P., Ulbrich, I. M., Zhang, Q., Cappa, C. D., Ziemann, P. J. and Jimenez, J. L.: Evaluation of recently-proposed secondary organic aerosol models for a case study in Mexico City, Atmos. Chem. Phys., 9, 5681–5709, doi:10.5194/acp-9-5681-2009, 2009.

Ensberg, J. J., Hayes, P. L., Jimenez, J. L., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Holloway, J. S., Gordon, T. D., Jathar, S., Robinson, A. L. and Seinfeld, J. H.: Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation, Atmos. Chem. Phys., 14(5), 2383–2397, doi:10.5194/acp-14-2383-2014, 2014.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A. and Goldstein, A. H.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, Proc. Natl. Acad. Sci., 109(45), 18318–18323, doi:10.1073/pnas.1212272109/-
/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1212272109, 2012.

Gentner, D. R., Jathar, S. H., Gordon, T. D., Bahreini, R., Day, D. A., El Haddad, I., Hayes, P. L., Pieber, S. M., Platt, S. M., de Gouw, J., Goldstein, A. H., Harley, R. A., Jimenez, J. L., Prévôt, A. S. H. and Robinson, A. L.: Review of Urban Secondary Organic Aerosol Formation from Gasoline and Diesel Motor Vehicle Emissions, Environ. Sci. Technol., acs.est.6b04509, doi:10.1021/acs.est.6b04509, 2017.

Goldstein, A. H. and Galbally, I. E.: Known and Unexplored Organic Constituents in the Earth's Atmosphere, Environ. Sci. Technol., 41(5), 1514–1521, doi:10.1021/es072476p, 2007.

Gordon, T. D., Tkacik, D. S., Presto, A. A., Zhang, M. and Shantanu, H.: Primary Gas- and Particle-Phase Emissions and Secondary Organic Aerosol Production from Gasoline and Diesel Off-Road Engines, 2013.

De Gouw, J. and Jimenez, J. L.: Organic Aerosols in the Earth's Atmosphere, Environ. Sci. Technol., 43(20), 7614–7618, doi:10.1021/es9006004, 2009.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R. and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9(14), 5155–5236, doi:10.5194/acp-9-5155-2009, 2009.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J. and Barsanti, K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors, Atmos. Chem. Phys., 17(2), 1471–1489, doi:10.5194/acp-17-1471-2017, 2017.

Hatch, L. E., Rivas-Ubach, A., Jen, C. N., Lipton, M., Goldstein, A. H. and Barsanti, K. C.: Measurements of I/SVOCs in biomass-burning smoke using solid-phase extraction disks and two-dimensional gas chromatography, Atmos. Chem. Phys., 18(24), 17801–17817, doi:10.5194/acp-18-17801-2018, 2018.

Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W., Toohey, D. W., Flynn, J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglück, B., Taylor, J. W., Allan, J. D., Holloway, J. S., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Massoli, P., Zhang, X., Liu, J., Weber, R. J., Corrigan, A. L., Russell, L. M., Isaacman, G., Worton, D. R., Kreisberg, N. M., Goldstein, A. H., Thalman, R., Waxman, E. M., Volkamer, R., Lin, Y. H., Surratt, J. D., Kleindienst, T. E., Offenberg, J. H., Dusanter, S., Griffith, S., Stevens, P. S., Brioude, J., Angevine, W. M. and Jimenez, J. L.: Organic aerosol composition and sources in Pasadena, California, during the 2010 CalNex campaign, J. Geophys. Res. Atmos., 118(16), 9233–9257, doi:10.1002/jgrd.50530, 2013.

Hodzic, A. and Jimenez, J. L.: Modeling anthropogenically controlled secondary organic aerosols in a megacity: a simplified framework for global and climate models, Geosci. Model Dev., 4(4), 901–917, doi:10.5194/gmd-4-901-2011, 2011.

Hodzic, A., Jimenez, J. L., Madronich, S., Canagaratna, M. R., Decarlo, P. F., Kleinman, L. and Fast, J.: Modeling organic aerosols in a megacity: Potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation, Atmos. Chem. Phys., 10(12), 5491–5514, doi:10.5194/acp-10-5491-2010, 2010.

Hunter, J. F., Day, D. A., Palm, B. B., Yatavelli, R. L. N., Chan, A. W. H., Kaser, L., Cappellin, L., Hayes, P. L., Cross, E. S., Carrasquillo, A. J., Campuzano-Jost, P., Stark, H., Zhao, Y., Hohaus, T., Smith, J. N., Hansel, A., Karl, T., Goldstein, A. H., Guenther, A., Worsnop, D. R., Thornton, J. A., Heald, C. L., Jimenez, J. L. and Kroll, J. H.: Comprehensive characterization of atmospheric organic carbon at a forested site, Nat. Geosci., 10, 748 [online] Available from: http://dx.doi.org/10.1038/ngeo3018, 2017.

Jathar, S. H., Miracolo, M. A., Tkacik, D. S., Donahue, N. M., Adams, P. J. and Robinson, A. L.: Secondary organic aerosol formation from photo-oxidation of unburned fuel: experimental results and implications for aerosol formation from combustion emissions., Environ. Sci. Technol., 47(22), 12886–93, doi:10.1021/es403445q, 2013.

Jathar, S. H., Gordon, T. D., Hennigan, C. J., Pye, H. O. T., Pouliot, G., Adams, P. J., Donahue, N. M. and Robinson, A. L.: Unspeciated organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States, Proc. Natl. Acad. Sci., 111(29), 10473–10478, doi:10.1073/pnas.1323740111, 2014.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, J., Huffman, J. A., Onasch,

T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U. and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, Science (80-. )., 326(5959), 1525–1529 [online] Available from: http://science.sciencemag.org/content/326/5959/1525.abstract, 2009.

Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Van Dingenen, R., Ervens, B., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G. and Wilson, J.: Organic aerosol and global climate modelling: a review, Atmos. Chem. Phys., 5(35), 1053–1123, doi:10.5194/acp-5-1053-2005, 2005.

Khare, P. and Gentner, D. R.: Considering the future of anthropogenic gas-phase organic compound emissions and the increasing influence of non-combustion sources on urban air quality, Atmos. Chem. Phys., 18(8), 5391–5413, doi:10.5194/acp-18-5391-2018, 2018.

Koo, B., Knipping, E. and Yarwood, G.: 1.5-Dimensional volatility basis set approach for modeling organic aerosol in CAMx and CMAQ, Atmos. Environ., 95, 158–164, doi:10.1016/j.atmosenv.2014.06.031, 2014.

Loza, C. L., Craven, J. S., Yee, L. D., Coggon, M. M., Schwantes, R. H., Shiraiwa, M., Zhang, X., Schilling, K. A., Ng, N. L., Canagaratna, M. R., Ziemann, P. J., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol yields of 12-carbon alkanes, Atmos. Chem. Phys., 14(3), 1423–1439, doi:10.5194/acp-14-1423-2014, 2014.

Lu, Q., Zhao, Y. and Robinson, A. L.: Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions, Atmos. Chem. Phys. Discuss., 2018, 1–28, doi:10.5194/acp-2018-752, 2018.

Lu, Q., Murphy, B. N., Qin, M., Adams, P. J., Zhao, Y., Pye, H. O. T., Efstathiou, C., Allen, C. and Robinson, A. L.: Simulation of organic aerosol formation during the CalNex study: updated mobile emissions and simplified secondary organic aerosol parameterization for intermediate volatility organic compounds, Atmos. Chem. Phys. Discuss., 2019, 1–36, doi:10.5194/acp-2019-986, 2019.

Ma, P. K., Zhao, Y., Robinson, A. L., Worton, D. R., Goldstein, A. H., Ortega, A. M., Jimenez, J. L., Zotter, P., Prévôt, A. S. H., Szidat, S. and Hayes, P. L.: Evaluating the impact of new observational constraints on P-S/IVOC emissions, multi-generation oxidation, and chamber wall losses on SOA modeling for Los Angeles, CA, Atmos. Chem. Phys., 17(15), 9237–9259, doi:10.5194/acp-17-9237-2017, 2017.

May, A. A., Nguyen, N. T., Presto, A. A., Gordon, T. D., Lipsky, E. M., Karve, M., Gutierrez, A., Robertson, W. H., Zhang, M., Brandow, C., Chang, O., Chen, S., Cicero-Fernandez, P., Dinkins, L., Fuentes, M., Huang, S. M., Ling, R., Long, J., Maddox, C., Massetti, J., McCauley, E., Miguel, A., Na, K., Ong, R., Pang, Y., Rieger, P., Sax, T., Truong, T., Vo, T., Chattopadhyay, S., Maldonado, H., Maricq, M. M. and Robinson, A. L.: Gas- and particle-phase primary emissions from in-use, on-road gasoline and diesel vehicles, Atmos. Environ., 88, 247–260, doi:10.1016/j.atmosenv.2014.01.046, 2014.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B. and Trainer, M.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science (80-. )., 359(6377), 760–764, doi:10.1126/science.aaq0524, 2018.

McDonald, B. C. B., Gentner, D. R., Goldstein, A. H. and Harley, R. A.: Long-term trends in motor

vehicle emissions in U.S. urban areas, Environ. Sci. Technol., 47(17), 10022–10031, doi:10.1021/es401034z, 2013.

Murphy, B. N., Woody, M. C., Jimenez, J. L., Carlton, A. M. G., Hayes, P. L., Liu, S., Ng, N. L., Russell, L. M., Setyan, A., Xu, L., Young, J., Zaveri, R. A., Zhang, Q. and Pye, H. O. T.: Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: Impacts on source strength and partitioning, Atmos. Chem. Phys., 17(18), 11107–11133, doi:10.5194/acp-17-11107-2017, 2017.

Presto, A. a., Hennigan, C. J., Nguyen, N. T. and Robinson, A. L.: Determination of Volatility Distributions of Primary Organic Aerosol Emissions from Internal Combustion Engines Using Thermal Desorption Gas Chromatography Mass Spectrometry, Aerosol Sci. Technol., 46(10), 1129–1139, doi:10.1080/02786826.2012.700430, 2012.

Presto, A. A., Miracolo, M. A., Donahue, N. M. and Robinson, A. L.: Secondary Organic Aerosol Formation from High-NOx Photo-Oxidation of Low Volatility Precursors: n-Alkanes, Environ. Sci. Technol., 44(6), 2029–2034, doi:10.1021/es903712r, 2010.

Presto, A. A., Nguyen, N. T., Ranjan, M., Reeder, A. J., Lipsky, E. M., Hennigan, C. J., Miracolo, M. A., Riemer, D. D. and Robinson, A. L.: Fine particle and organic vapor emissions from staged tests of an in-use aircraft engine, Atmos. Environ., 45(21), 3603–3612, doi:10.1016/j.atmosenv.2011.03.061, 2011.

Pye, H. O. T. and Pouliot, G. A.: Modeling the Role of Alkanes, Polycyclic Aromatic Hydrocarbons, and Their Oligomers in Secondary Organic Aerosol Formation, Environ. Sci. Technol., 46(11), 6041–6047, doi:10.1021/es300409w, 2012.

Pye, H. O. T. and Seinfeld, J. H.: A global perspective on aerosol from low-volatility organic compounds, Atmos. Chem. Phys., 10(9), 4377–4401, doi:10.5194/acp-10-4377-2010, 2010.

Qi, L., Liu, H., Shen, X., Fu, M., Huang, F., Man, H., Deng, F., Shaikh, A. A., Wang, X., Dong, R., Song, C. and He, K.: Intermediate-Volatility Organic Compound Emissions from Nonroad Construction Machinery under Different Operation Modes, Environ. Sci. Technol., 53(23), 13832–13840, doi:10.1021/acs.est.9b01316, 2019.

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. a, Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R. and Pandis, S. N.: Rethinking Organic Aerosols :, Science (80-. )., 315(March), 1259–1262, doi:10.1126/science.1133061, 2007.

Schauer, J. J., Kleeman, M. J., Cass, G. R. and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 2. C1through C30Organic Compounds from Medium Duty Diesel Trucks, Environ. Sci. Technol., 33(10), 1578–1587, doi:10.1021/es980081n, 1999.

Shah, R. U., Coggon, M. M., Gkatzelis, G. I., McDonald, B. C., Tasoglou, A., Huber, H., Gilman, J., Warneke, C., Robinson, A. L. and Presto, A. A.: Urban Oxidation Flow Reactor Measurements Reveal Significant Secondary Organic Aerosol Contributions from Volatile Emissions of Emerging Importance, Environ. Sci. Technol., 54(2), 714–725, doi:10.1021/acs.est.9b06531, 2020.

Shrivastava, M. K., Lane, T. E., Donahue, N. M., Pandis, S. N. and Robinson, A. L.: Effects of gas particle partitioning and aging of primary emissions on urban and regional organic aerosol concentrations, J. Geophys. Res. Atmos., 113(18), 1–16, doi:10.1029/2007JD009735, 2008.

Tkacik, D. S., Presto, A. A., Donahue, N. M. and Robinson, A. L.: Secondary organic aerosol formation from intermediate-volatility organic compounds: Cyclic, linear, and branched alkanes, Environ. Sci. Technol., 46(16), 8773–8781, doi:10.1021/es301112c, 2012.

Tkacik, D. S., Lambe, A. T., Jathar, S., Li, X., Presto, A. a., Zhao, Y., Blake, D., Meinardi, S., Jayne, J. T., Croteau, P. L. and Robinson, A. L.: Secondary organic aerosol formation from in-use motor vehicle emissions using a potential aerosol mass reactor., Environ. Sci. Technol., 48(19), 11235–42, doi:10.1021/es502239v, 2014.

Woody, M. C., Baker, K. R., Hayes, P. L., Jimenez, J. L., Koo, B. and Pye, H. O. T. T.: Understanding sources of organic aerosol during CalNex-2010 using the CMAQ-VBS, Atmos. Chem. Phys., 16(6), 4081–4100, doi:10.5194/acp-16-4081-2016, 2016.

Yee, L. D., Craven, J. S., Loza, C. L., Schilling, K. A., Ng, N. L., Canagaratna, M. R., Ziemann, P. J., Flagan, R. C. and Seinfeld, J. H.: Effect of chemical structure on secondary organic aerosol formation from $C_{12}$ alkanes, Atmos. Chem. Phys., 13(21), 11121–11140, doi:10.5194/acp-13-11121-2013, 2013.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M. and Sun, Y. L.: Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes, Geophys. Res. Lett., 34(13), 2007.

Zhao, Y., Hennigan, C. J., May, A. A., Tkacik, D. S., De Gouw, J. A., Gilman, J. B., Kuster, W. C., Borbon, A. and Robinson, A. L.: Intermediate-volatility organic compounds: A large source of secondary organic aerosol, Environ. Sci. Technol., 48(23), 13743–13750, doi:10.1021/es5035188, 2014.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environ. Sci. Technol., 49(19), 11516–11526, doi:10.1021/acs.est.5b02841, 2015.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines, Environ. Sci. Technol., 50(8), 4554–4563, doi:10.1021/acs.est.5b06247, 2016.

Zhao, Y., Lambe, A. T., Saleh, R., Saliba, G. and Robinson, A. L.: Secondary Organic Aerosol Production from Gasoline Vehicle Exhaust: Effects of Engine Technology, Cold Start, and Emission Certification Standard, Environ. Sci. Technol., 52(3), 1253–1261, doi:10.1021/acs.est.7b05045, 2018.

# Chapter 2: Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions *


## Abstract

Emissions from mobile sources are important contributors to both primary and secondary organic aerosols (POA and SOA) in urban environments. We compiled recently published data to create comprehensive model-ready organic emission profiles for on- and off-road gasoline, gas-turbine, and diesel engines. The profiles span the entire volatility range, including volatile organic compounds (VOCs, effective saturation concentration $C^*=10^7$-$10^{11}$ μg/m$^3$), intermediate-volatile organic compounds (IVOCs, $C^*=10^3$-$10^6$ μg/m$^3$), semi-volatile organic compounds (SVOCs, $C^*=1$-$10^2$ μg/m$^3$), low-volatile organic compounds (LVOCs, $C^*\leq0.1$ μg/m$^3$) and non-volatile organic compounds (NVOCs). Although our profiles are comprehensive, this paper focuses on the IVOC and SVOC fractions to improve predictions of SOA formation. Organic emissions from all three source categories feature tri-modal volatility distributions ('by-product' mode, 'fuel' mode, and 'lubricant oil' mode). Despite wide variations in emission factors for total organics, the mass fractions of IVOCs and SVOCs are relatively consistent across sources using the same fuel type; for example, contributing 4.5% (2.4-9.6% as 10$^{th}$ to 90$^{th}$ percentile) and 1.1% (0.4-3.6%) for a diverse fleet of light duty gasoline vehicles tested over the cold-start unified cycle, respectively. This consistency indicates that limited number of profiles are needed to construct emissions


---


* Originally published as Lu, Q., Zhao, Y., and Robinson, A. L.: Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions, Atmos. Chem. Phys., 18, 17637–17654, https://doi.org/10.5194/acp-18-17637-2018, 2018.

inventories. We define five distinct profiles: (i) cold-start and off-road gasoline, (ii) hot-operation gasoline, (iii) gas turbine, (iv) traditional diesel and (v) diesel-particulate-filter equipped diesel. These profiles are designed to be directly implemented into chemical transport models and inventories. We compare emissions to unburned fuel; gasoline and gas-turbine emissions are enriched in IVOCs relative to unburned fuel. The new profiles predict that IVOCs and SVOC vapor contribute significantly to SOA production. We compare our new profiles to traditional source profiles and various scaling approach used previously to estimate IVOC emissions. These comparisons reveal large errors in these different approaches ranging from failure to account for IVOC emissions (traditional source profiles) to assuming source-invariant scaling ratios (most IVOC scaling approaches).

## 2.1 Introduction

Atmospheric particulate matter imposes health risks (Di et al., 2017) and influences climate (Kanakidou et al., 2005). Organic aerosol (OA) contributes 20-90% of submicron atmospheric fine particulate matter mass (Jimenez et al., 2009). OA is commonly classified as primary OA (POA), which is directly emitted by sources, or secondary OA (SOA), which is formed in the atmosphere through photo-oxidation gas-phase organics. Both POA and SOA concentrations depend on the gas-particle partitioning of a complex mixture of organics that span a broad range of volatility (Hallquist et al., 2009; Kroll and Seinfeld, 2008). Mobile sources contribute about one-third of the anthropogenic organic emissions in the 2014 EPA National Emission Inventory (NEI); they are an important source of POA and SOA precursor gases especially in urban environments (Gentner et al., 2017; USEPA-OAQPS, 2015).

Traditional emissions inventories such as the NEI account for emissions of gas-phase volatile organic compounds (VOCs, typically smaller than $C_{12}$) and non-volatile particulate

matter (PM). These emissions are speciated for use in chemical transport models using source-specific emission profiles. Robinson et al. (2007) and Shrivastava et al. (2008) argued that this is an overly simplistic representation of organic emissions.

First, multiple studies have demonstrated that a large fraction of POA is semi-volatile with dynamic gas-particle partitioning while traditional inventories and models treat it as non-volatile (Fujitani et al., 2012; Kuwayama et al., 2015; Li et al., 2016; May et al., 2013b, 2013a, 2013c; Robinson et al., 2007). Semi-volatile POA concentrations depend on the gas-particle partitioning of the emissions, which is determined by their volatility distribution and atmospheric conditions. In addition, source tests are often conducted at unrealistically high OA loading, which biases POA emission factor compared to more dilute, atmospheric conditions (Fujitani et al., 2012; Lipsky and Robinson, 2006). Second, most traditional inventories do not account for emissions of lower volatility organic gases, including intermediate-volatile organic compounds (IVOCs, effective saturation concentration $C^*=10^3$-$10^6$ $\mu g/m^3$) and semi-volatile organic compounds (SVOCs, $C^*=1$-$10^2$ $\mu g/m^3$). Laboratory experiments indicate that IVOCs and SVOCs form SOA efficiently (Chan et al., 2009; Presto et al., 2010), but quantifying their emissions requires sorbents which are not routinely used for source testing (Kishan et al., 2008). Neglecting SOA production from IVOCs and SVOCs can lead to substantial underprediction of atmospheric SOA production (Hodzic et al., 2010; Woody et al., 2016). The net effect of these two issues is to cause chemical transport models to overestimate POA emissions and underestimate SOA production, leading to errors in the predicted OA composition and concentrations (Baker et al., 2015; Ensberg et al., 2014; Woody et al., 2016). Accounting for these two issues improves model-measurement agreement (Jathar et al., 2017; Murphy et al., 2017; Woody et al., 2016).

IVOC and SVOC emissions have not been routinely implemented in models because of lack of the mass and chemical composition of total IVOCs and SVOCs (Shrivastava et al., 2008). Although many studies report emissions of individual IVOC and SVOC species (typically polycyclic aromatic hydrocarbon or *n*-alkanes) (Schauer et al., 1999a, 1999b, 2002; Siegl et al., 1999; Zielinska et al., 1996), the vast majority of the IVOC/SVOC mass cannot be resolved at the molecular level using traditional gas chromatography based techniques (Goldstein and Galbally, 2007; Zhao et al., 2014).

Recent studies have reported comprehensive IVOC, SVOC and/or low-volatile organic compound (LVOC, $C^* \leq 0.1$ $\mu g/m^3$) emissions and gas-particle partitioning on POA emissions from mobile sources (May et al., 2014; Presto et al., 2011; Zhao et al., 2015, 2016). Zhao et al. (2015, 2016) characterized the total emissions and chemical composition of IVOCs and SVOCs from a fleet of on- and off-road gasoline and diesel sources. Cross et al. (2013, 2015) reported total IVOC and/or SVOC emission from an aircraft and diesel engine. Presto et al. (2011) and Drozd et al. (2012) reported IVOC and SVOC emissions for two gas-turbine engines. Gentner et al. (2012) and Isaacman et al. (2012a) report molecular and mass spectrum information for IVOC and SVOC in liquid fuel and quartz filter samples. May et al. (2013a, 2013b), Kuwayama et al. (2015), and Li et al. (2016) also investigated the gas-particle partitioning of on-road vehicle POA in dynamometer and tunnel studies. However, only limited comparisons have been made between source categories and the data have not been compiled into model ready profiles.

In this paper, we report comprehensive organic emission profiles for mobile sources by integrating recently published data of organic emissions based on their volatility, including IVOCs and SVOCs, to improve model predictions of SOA formation. We compare our new profiles to traditional source profiles and unburned fuel, focusing on the volatility

distribution and SOA precursors. We then use the new profiles to evaluate different scaling approaches previously used to incorporate IVOC emissions into inventories and models. Finally, we present box model calculations of SOA formation to demonstrate the importance to implement the new profiles in SOA modelling.

## 2.2 Methods

### 2.2.1 Datasets

This paper combines previously published measurement data of organic emissions (Gordon et al., 2013; May et al., 2014; Presto et al., 2011; Zhao et al., 2015, 2016) from gasoline, gas-turbine and diesel engines to create comprehensive model-ready source profiles. All tests used the same procedures to characterize IVOC and SVOC emissions to create a self-consistent dataset for low-volatile organics, but slightly different sampling media (Tedlar bags and/or canisters) and level of speciation were used to characterize VOC emissions. In the results and discussion sections, we compare these data to other recently published measurements made using different techniques.

We present two types of data: (i) emission factors of total organics and (ii) speciation profiles. We present total organic emissions factors for all tested engines: 64 gasoline vehicles, 5 diesel trucks, 6 off-road gasoline engines, 1 off-road diesel engine and 1 gas-turbine engine. We define total organic emissions as the sum of non-methane organic gases (NMOG) measured by flame ionization detection plus 1.2 times organic carbon (OC) measured using thermal optical analysis of quartz filter sample (the factor of 1.2 is the organic-mass-to-organic-carbon ratio, which accounts for the contribution of non-carbonaceous species in the organic (Turpin and Lim, 2001)). We define the NMOG as THC (measured with FID) minus $CH_4$ plus carbonyls. We define POA as organics collected by a

bare quartz filter analyzed by thermal-optical analysis. We converted measured pollutant concentrations to fuel-based emission factors (EF, mg/kg-fuel) using the carbon-mass-balance approach and the measured mass fraction of carbon in fuel (0.82 for gasoline, 0.86 for jet fuel and 0.85 for diesel) (May et al., 2014; Presto et al., 2011).

We derive speciation profiles from gas-chromatography-based analyses of filter, adsorbent tubes and Tedlar bag/canister samples. Details on the analytical procedures are described by Zhao et al. (2015, 2016). The speciation profiles are based the subset of tests with complete data (all three media): VOCs, IVOCs, SVOCs, and LVOCs. This included 29 gasoline vehicles, 4 diesel trucks, 3 off-road gasoline engines, 1 off-road diesel engine and 1 gas-turbine engine (Table S1). A detailed description of experimental set-up, sampling and chemical analysis is provided in the original articles (Gordon et al., 2013; May et al., 2014; Presto et al., 2011; Zhao et al., 2015, 2016). Only a brief description is provided here.

Emissions samples were collected from diluted exhaust. Gasoline and diesel source emissions were collected from a constant volume sampler (CVS) that diluted the exhaust with ambient air treated by high-efficient particulate air (HEPA) filters (Gordon et al., 2013; May et al., 2014). Gas-turbine engine exhaust was sampled from a rake inlet installed 1-m downstream of the engine exit plane (Presto et al., 2011). Sources were tested using standard test cycles (Gordon et al., 2013; May et al., 2014; Presto et al., 2011). On-road gasoline vehicles were tested on both cold-start and hot-start unified cycles. On-road diesel vehicles were tested in both lower-speed (creep and idle) and high-speed operation modes. Gas-turbine engine was operated on 4% and 85% engine thrust. Off-road engines were operated on certification cycles.

A suite of complementary sampling media was employed to characterize emissions across the entire volatility range. Tedlar bags (for gasoline and diesel sources) or canisters

(for gas-turbine source) were collected and analyzed by GC-FID and GC-MS to determine $CH_4$ and VOC hydrocarbon emissions up to $C_{12}$ compounds (May et al., 2014; Presto et al., 2011). Carbonyls (up to $C_6$) were sampled using 2,4-dinitrophenylhydrazine (DNPH) impregnated cartridges and analyzed by high-performance liquid chromatography (HPLC) (May et al., 2014). Quartz filters followed by two Tenax TA adsorbent tubes collected low-volatility organics that were analyzed by GC/MS equipped with a thermal desorption and injection system (Gerstel) (Zhao et al., 2015, 2016). The filter samples were also analyzed using a Thermal/Optical Carbon Analyzer for total organic carbon (OC) (May et al., 2014). The adsorbent tubes collect IVOCs and some SVOCs; SVOCs and even lower volatility organics were collected on quartz filters (Zhao et al., 2015, 2016). Except for the gas-turbine engine tests, total hydrocarbon (THC) emissions were determined by FID analysis of Tedlar bag samples (Gordon et al., 2013; May et al., 2014).

All adsorbent tubes and quartz filters were analyzed following the same procedure. Total (speciated and unspeciated) mass of IVOCs, SVOCs and LVOCs was determined by Zhao et al. (2015, 2016). The analysis quantified 57 individual IVOCs, which together contributed less than 10% of the total IVOC mass. The residual IVOCs, SVOCs and LVOCs commonly appear as an unresolved complex mixture (UCM); they were quantified into 29 lumped group ($C_{12}$ – $C_{38}$) based on the retention time of *n*-alkanes (each group corresponds to the mass that elutes between two sequential *n*-alkanes). Each IVOC lumped group ($C_{12}$ – $C_{22}$) was further subdivided into two chemical classes (unspeciated branched and cyclic compounds) based on their mass spectra. NVOCs are determined as the difference between the thermal optical analysis (1.2*OC) and the GC/MS analysis (IVOC+SVOC+LVOC) of the quartz filter samples.

Different levels of speciation were performed on the canister or Tedlar bag samples,

depending on source category. The Tedlar bag samples of gasoline exhaust were analyzed for 192 individual VOCs and 10 IVOCs; gas turbine exhaust was analyzed for 81 individual VOCs and 5 IVOCs; diesel exhaust was analyzed for 47 individual VOCs, 2 IVOCs and 11 Kovats lumped groups in the VOC range (organics with a GC retention time between $n^{th}$ and $n+1^{th}$ *n*-alkanes). Given the different levels of VOC characterization, we supplemented our gas-turbine and diesel VOC data with existing speciation profiles (SPECIATE profiles #4674 and #5565). The method for combining the VOC data is described in Supporting Information.

### 2.2.2 Mapping organics into volatility basis sets

Gas-phase organic emissions must be speciated for use in chemical mechanisms such as SAPRC (Carter, 2010) or Carbon Bond (CB). These mechanisms typically group individual VOCs into a set of lumped compounds based on reactivity or other chemical properties. We compared gas-phase organic emissions using the lumping specified by the SAPRC mechanism; we also compare gas- and particle-phase emissions using the volatility basis set (VBS). The VBS framework lumps organics into logarithmically spaced bins of saturation concentrations (C*) at 298K. It is designed for representing the emissions and atmospheric evolution of lower volatility organics ($C_{12}$ and larger) in chemical transport models (Donahue et al., 2006). It is also useful visualizing and comparing emissions data across the entire volatility space; the VBS is not intended to replace chemical mechanisms used to represent VOCs in models. Figure S1 shows the overall processes of mapping speciated and unspeciated compounds data collected on sampling medias to volatility basis set (VBS).

To map emissions into the VBS, we assigned C* values to individual compounds and

lumped groups of unspeciated organics. For each speciated compound (i.e. individual VOCs and IVOCs), C* values are calculated as,

$$C_i^* = \frac{M_i 10^6 \zeta_i p_{L,i}^0}{760RT} \quad (1)$$

where $M_i$ is the molecular weight (g/mol), $\zeta_i$ is the activity coefficient of compound *i* in the condensed phase (assumed to be 1), and $p_{L,i}^0$ is the liquid vapor pressure (Torr) of compound *i*, R is the ideal gas constant ($8.206\times10^{-5}$ $m^3$ atm $mol^{-1}$ $K^{-1}$), T is temperature (K). $p_{L,i}^0$ values are from EPA Suite data at 298K (USEPA, 2012). Although experimental and/or predicted vapor pressure values are uncertain (Komkoua Mbienda et al., 2013), the factor of 10 spacing of the volatility bins in the VBS reduces the chance of misclassification errors.

For unspeciated organics, C* values were assigned to lumped groups using the retention time of *n*-alkanes as reference species. In the VOC range, Kovats groups are assigned the mean of log C* value of the two *n*-alkanes in each group (Presto et al., 2012). For IVOCs, SVOCs and LVOCs, the C* value of the *n*-alkane in each bin is used to represent the UCM that elutes around that *n*-alkane. IVOCs, SVOCs and LVOCs correspond to the retention time range of $C_{12}$ to $C_{22}$, $C_{23}$ to $C_{32}$, and $C_{33}$ to $C_{36}$ *n*-alkanes, respectively. Although calibrating C* using *n*-alkanes can overestimate the volatility of PAHs and aromatic oxygenates (Presto et al., 2012), these compounds are expected to contribute only a small fraction of the total low-volatile organics. In addition, the VBS volatility bins are a factor of 10 apart, which reduces the chance of misclassification errors.

After assigning C* values, we compile all species into the VBS volatility distribution. Each volatility bin of C* = $10^n$ μg/$m^3$ cover the volatility range from C* = $0.3\times10^n$ μg/$m^3$ to C* = $3\times10^n$ μg/$m^3$ in a logarithmic space with n varying from -2 to 11.

One challenge is that the Tedlar bags/canister samples were collected in parallel to

the filter/adsorbent tubes, which creates concerns about double counting. We assessed this issue by comparing volatility of organics measured by both approaches. Three IVOC species were measured in both the Tedlar bags and adsorbent samples: *n*-pentyl-benzene ($C^*=2.8\times10^6$ μg/m$^3$), *n*-dodecane ($C^*=1.9\times10^6$ μg/m$^3$) and naphthalene ($C^*=1.1\times10^6$ μg/m$^3$). Figure S2 (a-c) compares the emissions of these species measured using the two approaches (Supporting Information). For the most volatile of these species, n-pentyl-benzene, the Tedlar bag measured, on average, 5.2 times more than the adsorbent tubes. Both approaches measured essentially the same amount of n-dodecane (ratio of 0.85 and $R^2$ of 0.9. For naphthalene (the least volatile of these species), the adsorbent tubes measured about 5 times more than the Tedlar bag, which we attribute to wall losses in the bag (Wang et al., 1996).

The comparisons indicate that the filter/adsorbent tube sampling train quantitatively collects all organics less volatile than *n*-dodecane ($C^* =1.9\times10^6$ μg/m$^3$) while the bag/canister quantitatively collects all more volatile organics. *N*-dodecane falls within the $10^6$ μg/m$^3$ volatility bin. The upper bound of this bin is 3 x $10^6$ μg/m$^3$, which is close to the C* of *n*-dodecane. We therefore use 3 x $10^6$ μg/m$^3$ as the boundary between the adsorbent tube and Tedlar bag samples. To avoid double counting, we discarded all organics measured using the bag/canister/cartridge that are less volatile than $3\times10^6$ μg/m$^3$ and discarded all species measured in the adsorbent tube more volatile than $3\times10^6$ μg/m$^3$. Therefore, emissions in the $C^*= 10^7$ to $10^{11}$ μg/m$^3$ bins are based on the bag/canister/cartridge data and that the emissions in the $C^*= 10^{-1}$ to $10^6$ μg/m$^3$ bins are based on the filter and adsorbent tube data. NVOCs are assigned to a non-volatile bin. The adsorbent tubes may underestimate the speciated emissions in C* between $1.9\times10^6$ (*n*-dodecane) and $3\times10^6$ μg/m$^3$; however, they still measured, on average, 3.3 times organics in this range to the Tedlar bags (Figure S2d).

A final issue is whether our sampling and analytical methods capture and recover all emitted organics. We evaluated this by comparing the sum of total characterized organics (integrated organics from bag, adsorbent tube and filter measurements) to our estimate of total organics by bulk measurements (NMOG+1.2*OC). The sum of the characterized organics includes the VOCs, IVOCs, SVOCs, LVOCs determined from the GC-based analysis of the bags/canister, cartridges, adsorbent tubes and filters. This includes both individual species and lumped groups of unspeciated material.

Figure S3 indicates good mass closure for the on-road gasoline and diesel vehicle tests. The two estimated results agree within ±10% for more than 90% of non-DPF-equipped diesel engine tests (DPF = diesel particulate filter). For all LDGV (light-duty gasoline vehicle) tests, total characterized organics are 82 ± 21% of the total organics by bulk measurements. We suspect that most of the missing organics from the LDGV tests could be VOCs since the VOC analysis only quantified a list of targeted compounds (Zhao et al., 2017). There was relatively poor mass closure for the off-road engine and DPF-equipped diesel tests. For the off-road engine emissions, the sum of total characterized organics was less than 50% of the bulk measurement. Comparisons with literature data (Gabele, 1997; Volckens et al., 2008) suggests that our speciated VOC groups to NMOG ratios are low (Figure S4). The cause of this bias is not known, but we attribute it to measurement error. We used a linear regression to the literature results to rescale our VOC data for off-road engines (see SI). For DPF-equipped diesel vehicles, the sum of speciated organics is up to 7 times the bulk measurement of total organics. The DPF-equipped diesel emission are quite low and this discrepancy is likely due to uncertainty in background corrections (Zhao et al., 2015).

Traditionally, there are three standard ways to treat these residual emissions -- the

difference between sum of characterized emissions and the total/bulk emissions (frequently called unknown or UNK): (1) assume it is inert and therefore ignored in models, (2) renormalizing the residual emissions to the known composition which assumes that the composition of the unspeciated material is the same as the speciated mass, or (3) by assigning a custom profile to the residual mass based on a representative list of compounds (Carter, 2015). The standard default profile for (3) was derived from the all-profile-average carbon number > 6, molecular weight > 120 compounds in SPECIATE database (Adelman et al. 2005). Therefore, it lacks comprehensive IVOCs and SVOCs data.

In the following discussion, we normalize the residual/uncharacterized organics to the known composition, assuming that the residual unknown organics have the same volatility and chemical characteristics as the total characterized organics. Since there was not an independent measurement of NMOG during the gas-turbine engine tests (Presto et al., 2011), we assume the supplemented speciated VOCs plus the sorbent and filter data is the total emitted organics.

### 2.2.3 Box model for SOA yield calculation

The overall SOA yield of gas-phase emissions (mass of SOA produced/mass of NMOG emissions) can be calculated as

$$y_{SOA} = \sum_i f_{gas,i} \times Y_i \quad (2)$$

where $f_{gas,\,i}$ is the mass fraction of SOA precursor *i* in NMOG; and $Y_i$ is the SOA mass yield of compound *i* at OA= 10 μg/m$^3$ (a typical urban OA level).

SOA mass yields for each VOCs are based on SAPRC groups and are taken from CMAQ 5.1 (USEPA, 2016a). The complete VOC composition for the new source profiles are listed in the Table S3. SOA mass yields for IVOCs are calculated using the mechanism

of Zhao et. al (2015). The gas-phase SVOCs are assumed to have a SOA mass yield of 1 (Presto et al., 2010). Equation (2) omits the OH reaction rates and therefore represents the ultimate SOA yield from NMOG emissions. The relative contribution of IVOCs and VOCs to SOA varies with time because IVOCs generally react faster with OH than VOCs (Zhao et al., 2016). Therefore, the ultimate yield approach (equation 2) provides a lower bound estimate of the contribution of IVOCs to SOA.

### 2.3 Results and discussion

Figure 2.1 shows the volatility distribution of the total characterized organic emissions for a typical gasoline (Figure 2.1a) and diesel (Figure 2.1b) test classified by collection media. It underscores the importance of using adsorbents (in addition to filters and Tedlar bags) to comprehensively characterize all of the organic emissions. The Tenax adsorbent tubes collect almost all of the IVOCs (> 90% for gasoline and > 97% for diesel), with the balance being collected by the quartz filter, presumably as adsorption artifact (Zhao et al., 2015, 2016). The Tenax adsorbent collects 5.2% and 54.8% of the total organic emissions from the gasoline and diesel engines, respectively. Since the vast majority of source testing does not employ adsorbents, IVOCs are not quantitatively accounted for in most emission profiles (Pye and Pouliot, 2012). In comparison to the adsorbent samples, the bag/canister collected only 12.9% and 4.0% of IVOCs for gasoline and diesel sources, respectively. We have discarded this component to avoid double counting, as discussed in the methods section.

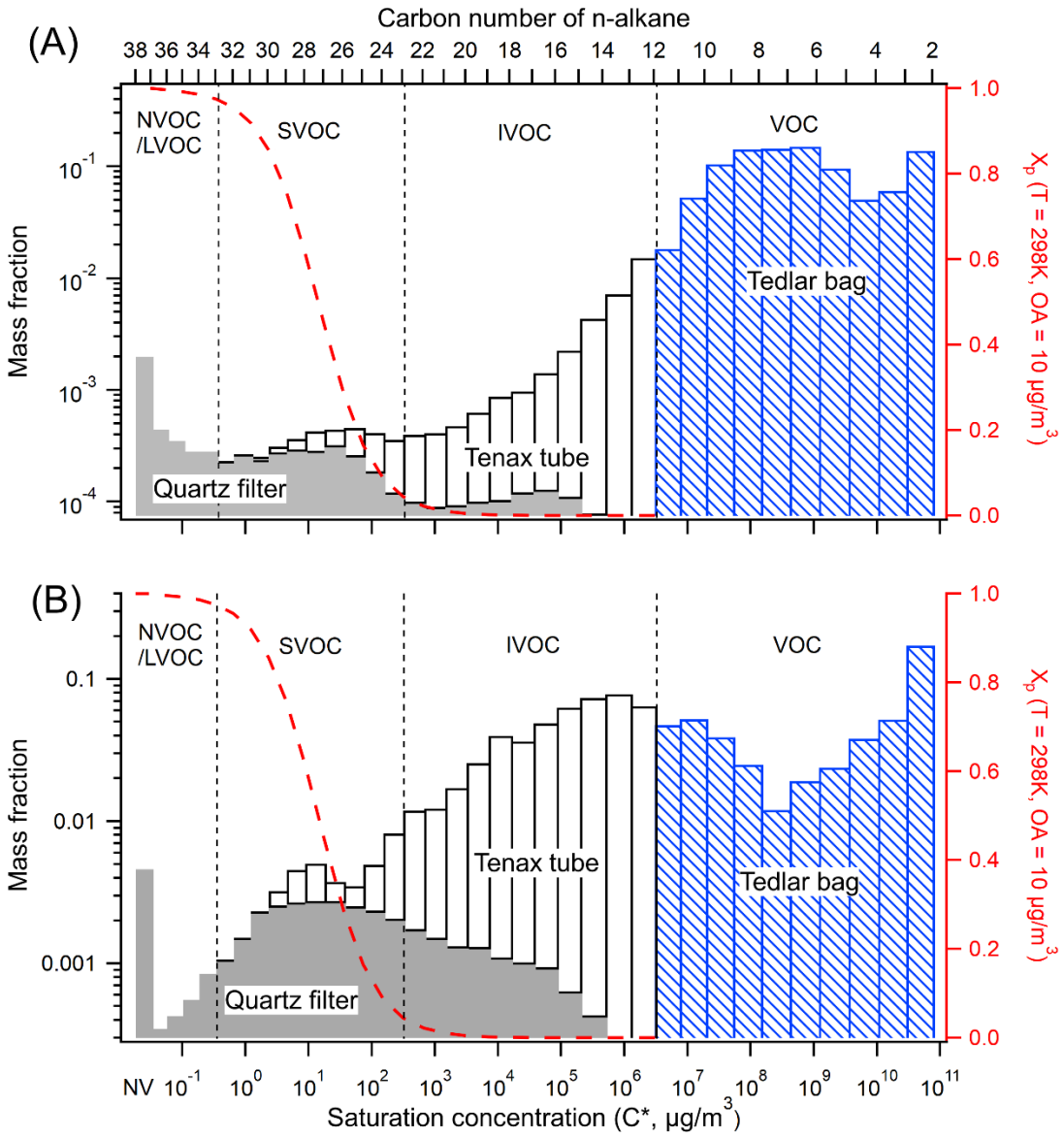



**Figure 2.1 Volatility distribution of organic emissions for a typical (a) gasoline (b) diesel vehicle. The emissions are classified by sampling media (line 1: Tedlar bag, line 2: bare quartz filter followed by two Tenax tubes). The red dashed line indicates the particle fraction assuming the emissions form a quasi-ideal solution at a $C_{OA}$ of 10 µg m$^{-3}$ and temperature of 298K.**

Figure 2.1 also shows the particle fraction ($X_p$) calculated assuming all organics form a quasi-ideal solution to illustrate gas-particle partitioning at typical atmospheric conditions (T=298K, OA=10 µg m$^{-3}$). At these conditions, IVOCs exist essentially exclusively in the gas-phase, while SVOCs exist in both phases. To illustrate the changes in gas-particle

partitioning of IVOCs and SVOCs across a wide range of atmospheric conditions, Figure S5 shows equilibrium particle fraction ($X_p$) for T between 273K – 320K and OA concentration from 1 – 10 μg $m^{-3}$. IVOCs are essentially exclusively in the gas-phase (>99%) except at very low temperature (T = 273K) and high OA loading (OA = 10 μg $m^{-3}$) conditions when about 6% of the lowest bin (C* = $10^3$ μg $m^{-3}$) partitions to particle-phase. In contrast, SVOCs are always present in both gas- and particle- phases, in both hot and dilute (T = 320K and OA = 1 μg/$m^3$) or cold and high OA loading (T = 273K and OA = 10 μg $m^{-3}$) conditions.

Figure 2.1 indicates there is also substantial breakthrough of SVOCs from the quartz filter during mobile source certification testing (e.g. 2007 CFR 86), which requires maintaining a filter temperature of 47°C. In our experiments, these SVOCs are collected by the downstream Tenax tubes. This breakthrough is denoted by the white bars in the SVOC range in Figure 2.1; the SVOC breakthrough corresponds to, on average, 37% of the total SVOC emissions from gasoline vehicles, 52% for non-DPF diesel and 89% for DPF-diesel (Zhao et al., 2015, 2016). Therefore, quantitatively accounting for all gas-phase SVOCs requires using adsorbents. This is needed to improve predictions of POA concentrations and SOA production.

We compared the sum of NVOCs, LVOCs and SVOCs to the quartz filter POA measurements. A linear regression of the on-road gasoline vehicle data yields a slope of 1.4 (Figure S6a), which indicates that the quartz-filter-based POA emission factors should be multiplied by 1.4 to account for missing gas-phase SVOC emissions. This factor would be larger if the quartz filter did not collect some IVOC vapors as adsorption artifact (Figure 2.1). For off-road gasoline sources, a linear regression yields a slope of 1.1 (Figure S6b), indicating a larger fraction of the SVOC are collected on quartz filters compared to on-road gasoline vehicles. This is presumably due to shifts in gas-particle partitioning towards the

particle-phase at the high OA concentrations in the off-road source tests. For diesel sources, a linear regression yields a slope of 0.9 (Figure S6c). This lower ratio is due to that filter measured POA also includes positive adsorption artifact from IVOCs, which more than offsets the gas-phase SVOC breakthrough (Figure 2.1b).

### 2.3.1 Emission factors

Figure 2.2 (a) compares the total organics emission factors (NMOG+1.2*OC) for on- and off-road gasoline vehicles, including LDGV, two-stroke small off-road engines (SORE-2S), and four-stroke small off-road engines (SORE-4S); gas-turbine engines; and on- and off-road diesel sources, including DPF-equipped engines. We subdivided the LDGV data based on emissions certification standard: pre-LEV (U.S. Tier0), LEV (California Low Emission Vehicle), and ULEV (California Ultra-Low Emission Vehicle).

Although there is source-to-source variability within a given source category (e.g. pre-LEV gasoline or DPF-equipped diesel), there are distinct trends in total organic emissions. Gasoline small off road engines (SORE) have the highest emissions, with SORE-2S having, on average, one order of magnitude higher emissions than SORE-4S (Gordon et al., 2013). This is due to less stringent regulations for off-road engine emissions (Cao et al., 2016a), and the unburnt fuel mixing in exhaust due to the two-stoke design in SORE-2S. The LDGV emissions decrease with tightening emission standards (Gentner et al., 2017; May et al., 2014). For example, relative to the median Pre-LEV, there is a 78% reduction in total organic emissions to the median LEV and 90% to the median ULEV. Although not shown here, total organic emission factors are dramatically higher during cold-start than during hot-stabilized operations after the catalytic converter has reached its operating temperature (Saliba et al., 2017).

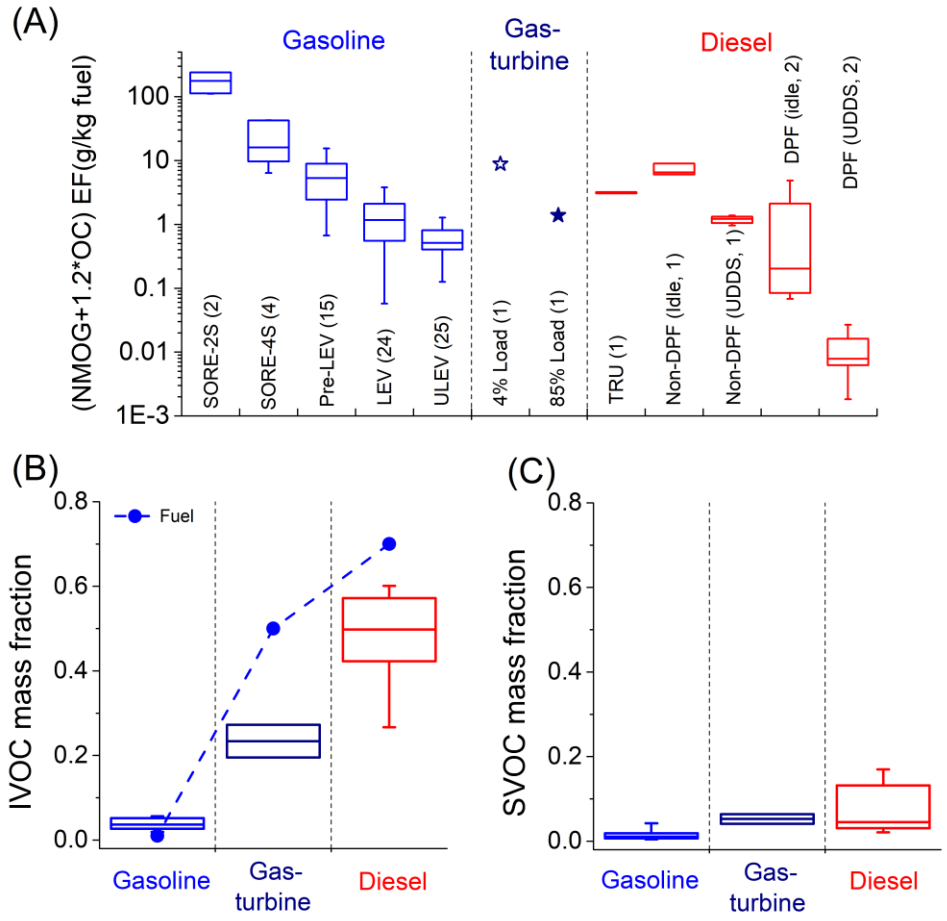



**Figure 2.2 (a) Emission factors for total organics (NMOG+1.2*OC) for different source categories. The number in parentheses indicates number of unique sources tested in each category. Mass fraction of (b) IVOCs and (c) SVOCs in total organics for each source category. This figure only shows cold-start and off-road gasoline engine emissions. Box-whisker plot represents range of emission for each category: 25$^{th}$ -75$^{th}$ percentiles and 10$^{th}$-90$^{th}$ percentiles.**

Gas-turbine engine emissions show strong load dependence; idle (4% thrust) emission is comparable to pre-LEV vehicles, and about an order of magnitude higher than high loads (85% thrust) emission. Diesel emissions show strong dependence on both after-treatment devices and test cycle. DPF-equipped diesel vehicles have the lowest emission factors among all tested engine types. Lower emission factors are measured for high speed transient operations (e.g. UDDS cycle) compared to idle/low speed operations. The trends in

gas-turbine and diesel emissions are qualitatively consistent with Cross et al. (2013, 2015) who showed similar load-dependent trend of decreasing THC or IVOC emission factors of gas-turbine and diesel engines with higher loads.

As expected, Figure 2.2 (a) indicates there is source-to-source variation in total organic emission for a given category (e.g. pre-LEV or ULEV). This variability reflects the effects of difference of engine design, engine calibration, after-treatment system, vehicle age, and maintenance history on emissions. However, the previously described trends in total organic emission among the different source categories are clear even with this variability.

### 2.3.2 Volatility and chemical composition distributions

Figure 2.3 shows the median volatility distributions of the emissions for three different source categories: gasoline (cold-start), gas-turbine and non-DPF diesel. For gas-turbine engine category, we plot the idle load (4% thrust) emission.

Figure 2.3 indicates that the organic emissions from all three source categories have tri-modal volatility distributions. The dominant mode is the middle one, with a peak at $C^*=10^8$ μg $m^{-3}$ for gasoline sources, $C^*=10^6$ μg $m^{-3}$ for gas-turbine sources, and $C^*=10^5$ μg $m^{-3}$ for diesel sources. For each source category, this mode has a similar volatility distribution and chemical composition as unburned fuel (Figure S7). We therefore call it the 'fuel mode'.

The fuel mode contributes 72.6% (66.5-77.6% as $10^{th}$ to $90^{th}$ percentile, same hereafter) of the total organic emissions in gasoline engine exhaust, 63.1% (48.9-84.4%) in diesel engine exhaust, and 37.5-38.5% in gas-turbine source emissions. The widely varying contribution of this mode to diesel emissions is due, in part, to after-treatment and duty-cycle effects. For example, low-speed operation (creep and idle) test results show higher mass

fractions in the fuel mode, 79.7% (62.2-83.0%), compared to high speed operations. The size of the fuel mode to DPF diesel vehicles is highly variable, 54.2% (29.3-79.9%), which is likely due in part to higher uncertainty associated with measuring very low emission rates.

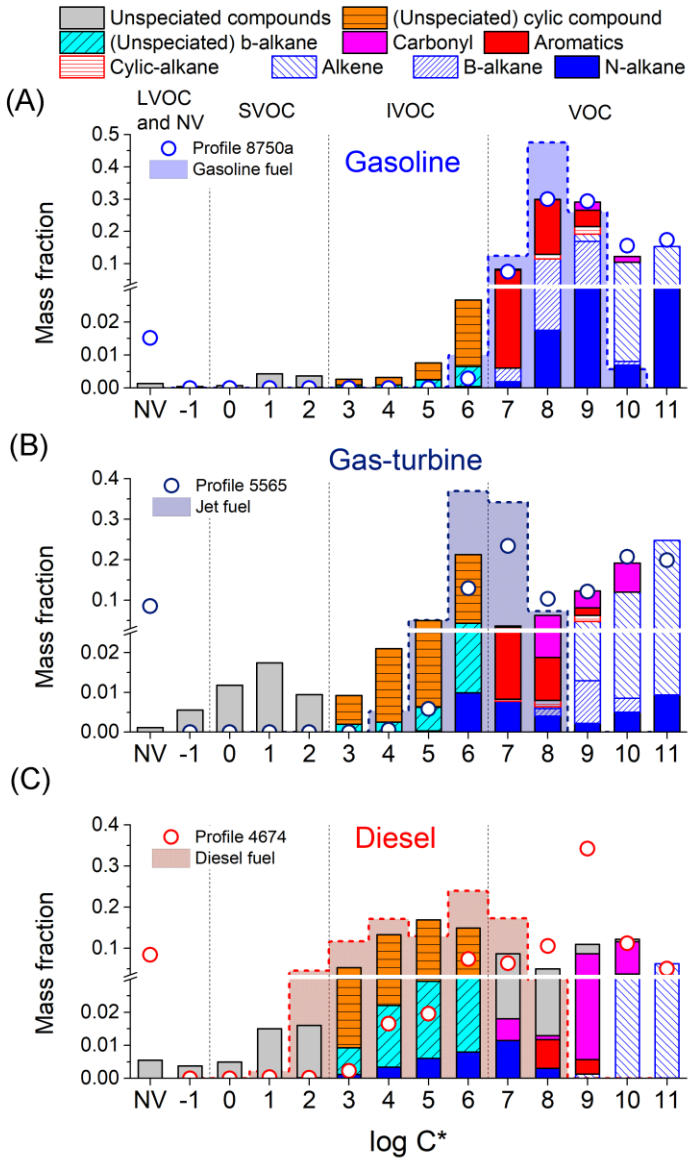



**Figure 2.3 Median volatility distribution of organic emissions for (a) cold-start on-road gasoline, (b) gas-turbine (idle) and (c) on-road non-DPF diesel engines. The color shading indicates composition. Shaded area indicated by dashed indicate distribution for unburned fuel; dots indicate traditional source profiles from EPA SPECIATE database. The y-axis has a broken scale to amplify the least volatile emissions.**

Figure 2.3 highlights how the changes in fuel composition create systematic

differences in volatility distribution of the emissions among the three source categories. Specifically, the 'fuel' mode of the exhaust shifts towards lower volatility from gasoline to diesel sources mirroring the trend in fuel volatility. Although the chemical composition of the fuel mode is also similar to that of unburned fuel (Figure S7), there are some important differences indicating that combustion and removal efficiencies vary by compound class, which are discussed in section 2.3.4.

Emissions from each source have a low-volatility mode, comprised of SVOCs and even less-volatile organics. For all three source categories, this low-volatility mode peaks at a $C^* = 10\ \mu g\ m^{-3}$, which is in the middle of the SVOC range. Therefore, some of the organics in the low-volatility mode partition into the particle phase in the atmosphere to form POA, while the rest exist as vapors. The volatility distribution of this mode is similar to that of lubricating oil (May et al., 2013a, 2013b; Worton et al., 2014); we therefore refer to the low-volatility mode as the 'oil mode'. For diesel, the low-volatility and fuel modes blend together. The oil mode contributes 1.4% (0.6-4.2%) of the total organic emissions for gasoline sources, 4.2-12.1% for the gas-turbine source, and 5.9% (3.1-17.7%) for diesel sources.

The size of the LDGV 'oil mode' varies with certification standard, with median values of 0.8% in Pre-LEV, 1.4% in LEV and 2.2% in ULEV. This trend indicates that improvements in after-treatment technology more effectively remove NMOG emission compared to POA emissions. The wide range of SVOC emissions from gas-turbine and diesel sources reflects the effects of changes in engine load/after-treatment: at 85% engine load, 12.1% of gas-turbine emissions are in the 'oil mode' versus only 4.2% at 4% load. DPF-equipped vehicles show 14.8% (10.1-30.6%) of the emissions for on high-speed cycles versus a much lower fraction 2.1% (1.7-5.7%) at low speed operations.

The third mode is the most volatile one, peaking at a C* = $10^{10}$ or $10^{11}$ μg $m^{-3}$. It contributes 25.9% (21.1–31.0%) of the total organics in gasoline emissions, 26.9% (9.4-40.6%) in diesel sources emissions, and 50.5-57.3% in gas-turbine engine emissions. It is comprised of the smallest compounds, such as $C_2$-$C_5$ alkanes, alkenes and carbonyls, produced from the incomplete combustion and breakdown of fuel molecules (May et al., 2014). It also contains other compounds such benzene in the C* = $10^9$ μg $m^{-3}$. We therefore call it the 'combustion by-product' mode. The composition of this mode varies modestly by source class.

The majority of the IVOC emissions are found in the lower volatility end of the fuel mode. For gasoline sources, IVOCs are in the lowest volatility tail of this mode. For the LDGV tested using cold-start unified cycle, IVOCs contribute 4.5% (2.4%-9.6%) of the total organic emissions. This includes both heavily controlled and low emitting ULEV and less controlled and higher emitting pre-LEVs. IVOCs contribute a similar fraction to the organic emissions from largely uncontrolled and high emitting SOREs (Figure 2.2). However, IVOCs contribute a larger fraction, 18.1% (5.8 – 31.1%) for organic emissions from LDGV operated over the hot-start unified cycle (Zhao et al., 2016). This suggests that catalytic converters may be less effective at removing lower volatility organics such as IVOCs, which is also consistent with the trends in SVOC data discussed above. However, only four vehicles were tested using the hot-start unified cycle and the IVOC faction varied widely. More research is needed to understand the effects of hot-operations and duty cycle in general on IVOC emissions.

For sources operating on less volatility fuels, IVOCs contribute a larger fraction of the emissions. For example, they contribute 20-27% of gas-turbine engine emissions at idle and 85% loads. This is somewhat larger than data from Cross et al. (2013) who reported 10-

20% of NMHC emissions are IVOCs at idle load. The difference could be due to multiple factors, including differences in collection techniques (cryogenic versus adsorbent) and/or differences in fuel composition (Corporan et al., 2009). Diesel sources emit the highest fraction of IVOCs, with median value of 51.3% (28.7-61.5%). Non-DPF diesel emissions have a more consistent IVOC fraction of 57.1% (46.3-66.4%) than DPF-diesel emissions (40.1%; 17.2-55.5%). Finally, the contribution of IVOCs qualitatively mirror the fuel composition: 1% of unburned gasoline is comprised of IVOCs, ~50% for JP-8, and ~70% for diesel (Corporan et al., 2009; Gentner et al., 2012; May et al., 2014).

Figure 2.2 (c) indicates that the contribution of SVOCs also differs by source type. For gasoline engines, SVOCs contribute 1.1% (0.4-3.6%) of the total organic emission. This variability is, in part, associated with the effects of tightening emissions certification standards as discussed above. For gas-turbines, SVOCs contribute 3.6-4.6% of total organic emissions. For diesel source, SVOCs contribute 4.6% (2.3-16.1%) of the total organic emission; the wide range reflects effects of duty cycle and after-treatment as discussed above. There are no SVOCs in unburned gasoline and jet fuel, and less than 2% for diesel fuel. The SVOCs in the emissions are likely predominantly from lubricating oil (Worton et al., 2014).

Given that the total organic emissions vary by more than five orders of magnitude (Figure 2.2a), the volatility distribution (and emissions profiles) are relatively consistent across sources using the same fuel type (Figure 2.3). As discussed previously, for a given fuel type, after-treatment technology (e.g. LDGV emission certification standard) and test cycle can also influence the volatility distribution, but their influence is much less than that of fuel. We therefore use the median distributions to represent the properties of the aggregate emissions from a large number of sources with a given source category. There is always

source-to-source variability, but for inventories we need to define representative profiles for distinct categories (we use medians as opposed to averages to reduce the influence of outliers).

An important question is the number of distinct source categories. To investigate this question, Figure 2.4 compares the volatility distributions of different sources in a 2-dimensional space of IVOC versus SVOC mass fractions. These are important SOA precursors so this framework highlights differences in SOA formation potential. There are three distinct clusters in Figure 2.4, one for each fuel type (gasoline, diesel and jet). Therefore, these source categories require different profiles. For example, the on-road (cold-start) and off-road gasoline sources emissions cluster, with a median mass IVOC and SVOC fractions of 4.5% and 1.1%, respectively, indicating similar volatility distributions between on- and off-road gasoline sources. Figure 2.4 also suggests two additional categories, but these distinctions are not as strong given the variability of the data. First, hot-start LDGV emissions have much higher IVOC and SVOC fractions than cold-start emissions (18.1% versus 4.5%, 4.7% versus 1.1%). This implies a roughly 4-time higher SOA yield for hot-start on-road gasoline emissions. Therefore, separate profiles should be used to represent cold-start and hot-operation emissions when constructing emission inventories for gasoline vehicles. Second, DPF and non-DPF equipped diesel sources also show significant different volatility distributions, especially in SVOC mass fraction (12.2% versus 3.8%). To account for these differences, we present five emission profiles in Table S3: gasoline (cold-start and off-road), gasoline (hot-start), diesel (non-DPF), diesel (DPF) and gas-turbine engines. Interestingly, the SVOC and IVOC mass fractions are strongly positive-correlated across all sources with an exponential fit between SVOC and IVOC mass fraction of $f_{SVOC} = 0.100\, f_{IVOC}^{\,0.700}$. One could certainly define additional source categories to, for example, account

for trends in SVOC fraction with emission certification of LDGVs, but it is not clear that those difference are large enough to improve model performance versus using an aggregate profile to represent all gasoline vehicles.

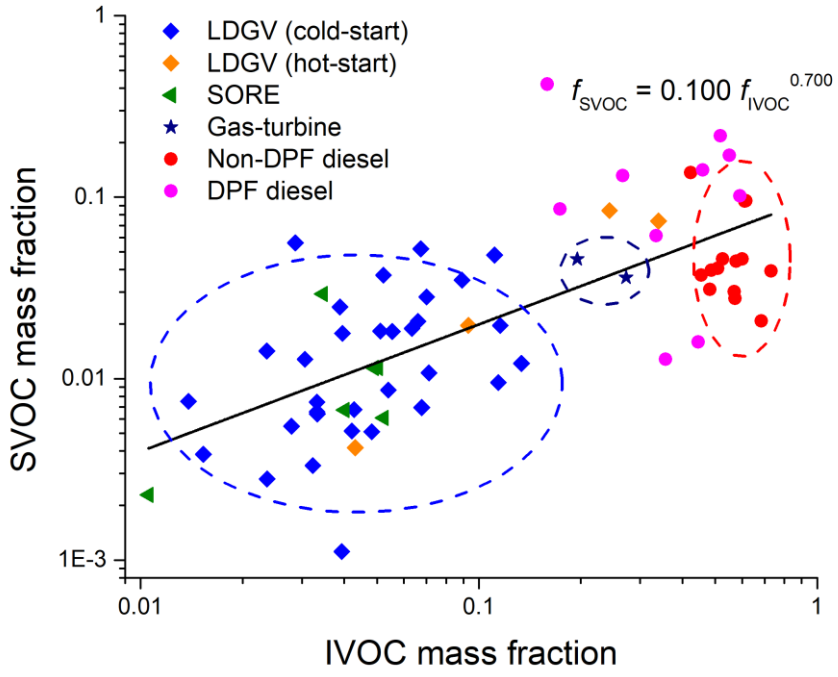



**Figure 2.4 Two-dimensional visualization of volatility distributions (x axis: IVOC mass fraction, y axis: SVOC mass fraction) of all tested sources. Dashed circles indicate clusters by fuel type: Blue cluster: gasoline (cold-start), navy: gas-turbine, red: non-DPF diesel source.**

### 2.3.3 New versus traditional source profiles

Figure 2.3 also compares our new comprehensive source profiles to traditional profiles used to construct the emission inventory to simulate air quality in the Los Angeles region during the 2010 CALNEX campaign (Baker et al., 2015). Our new profiles are the median value of the measured emission for gasoline (separate for cold-start and hot-operations), gas turbine, non-DPF and DPF-diesel sources; they are listed in Table S3 (Supporting Information). The traditional profiles are from the EPA SPECIATE database

(USEPA, 2016b) -- profile #4674 for diesel, #8750a for gasoline, and #5565 for gas turbine sources with the POA fraction (NVOC) calculated using MOVES (USEPA, 2014) .

There is good agreement between our new and traditional profiles in the VOC range, with both having similar chemical compositions and volatility distributions containing both by-product and fuel modes (Figure 2.3). For example, Figure 2.5 demonstrates the strong agreement for SARPC-lumped VOC groups between the new and traditional profiles for all three source categories with more than 90% of the SAPRC groups for the gasoline sources agree within a factor of two. We recommend using our new profiles for VOC composition because they have enhanced VOC speciation from combining the existing SPECIATE profiles with our new experimental data.

However, the traditional profiles dramatically underestimate IVOCs and SVOCs, which are important classes of SOA precursors. As illustrated in Figure 2.1, this is a consequence of the limitations of traditional source characterization techniques to quantitatively collect and analyze IVOCs. For example, the traditional LDGV emission profile only attributes 0.2% of the total organics to IVOCs versus 4.6% in our new cold-start profile. The traditional gas-turbine engine emission profile attributes 13% of the organics to IVOCs versus 27% IVOCs in our new profile. For diesels vehicle emissions, the traditional profile attributes 10% of total organic emission to IVOCs versus 54.2% of organics for non-DPF diesel in our new profile. The traditional diesel source profile does contain about 20% unknown organics (UNK), part of which are likely IVOCs, as the collection and chemical analysis efficiency decrease towards lower volatility bins such as $10^3$ and $10^4$ $\mu g/m^3$ (Figure 2.3). However, most UNK is not represented as IVOCs in models, as discussed in section 2.2.2.

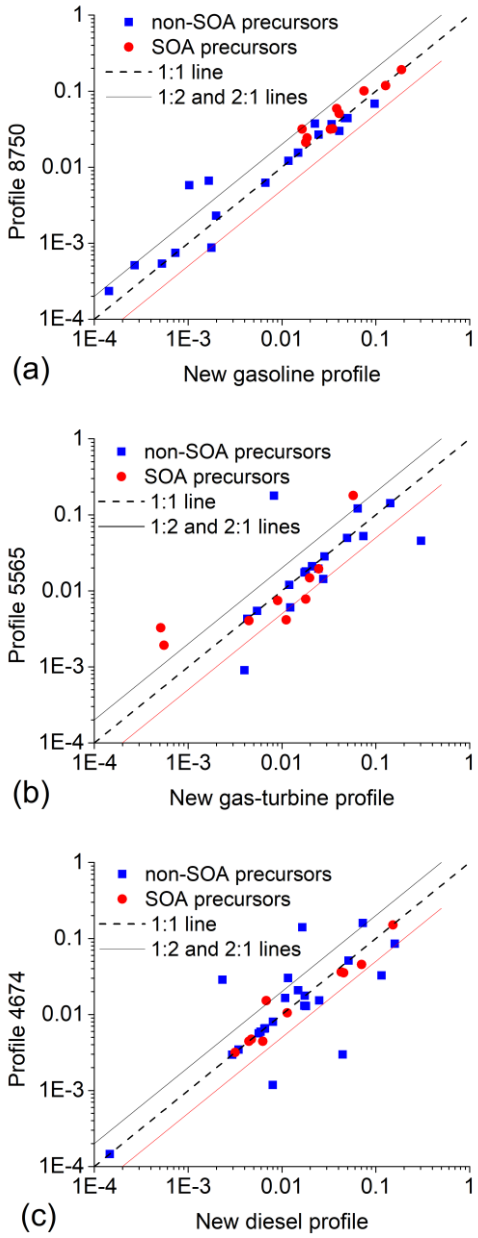



**Figure 2.5 Scatter plot of VOC groups in SAPRC mechanism in the new profiles and in SPECIATE database for (a) on-road gasoline, cold-start (b) gas-turbine (c) non-DPF diesel sources, demonstrating consistency between traditional and new profiles in VOC speciation**

### 2.3.4 Exhaust versus unburned fuel and IVOC enrichment factors

Figure 2.3 highlights the large contribution of unburned fuel to the exhaust, but careful examination of the data reveals that the combustion process and/or removal

efficiency by the after-treatment device are compound dependent. For example, gasoline and gas-turbine emission are both enriched in IVOCs compared to fuel (e.g. $C^*= 10^6$ μg/m$^3$ for gasoline, and $C^*= 10^4$ μg/m$^3$ for gas-turbine). The difference suggests higher combustion efficiency of more volatile fuel components.

Figure S8 compares the chemical composition of the exhaust to unburned fuel. Overall, straight and branched alkanes (speciated and unspeciated) contribute a smaller fraction to the exhaust than in the fuel with the median mass fractions decreasing from 46.6% (fuel) to 34.3% (exhaust) for gasoline sources, 50.0% to 9.8% for gas-turbine source, and 30.3% to 11.2% for diesel sources. In comparison the fraction of aromatic and cyclic compounds (speciated and unspeciated) are consistent between fuel and exhaust; for example, 37.2% (exhaust) versus 36.1% (exhaust) for gasoline source and 58.7% to 60.2% for diesel source. This comparison implies higher combustion efficiencies of *n*-/*b*- alkanes than cyclic/aromatic compounds in internal combustion engines, which could partly be explained by the flash points of different hydrocarbons. The mass fraction of alkenes, alkynes and carbonyls increase, indicating they are important product of incomplete combustion. For example, they increase from 3.5% (fuel) to 28.6% (exhaust) for gasoline sources, and 0% to 54.5% and 24% for gas-turbine and diesel sources, respectively. Gasoline emission have the highest single-ring aromatics fraction (~30%), compared to 5.5% in gas-turbine and 17% in diesel emissions. This mirrors fuel composition -- unburned gasoline fuel had the highest aromatic content (26.7%) of the fuels tested here.

We are especially interested in the enrichment or depletion of SOA precursors in the exhaust compared to fuel, including IVOCs and single ring aromatics. To quantify enrichment, we normalized SOA precursors in both the fuel and exhaust to $C_8$ to $C_{10}$ *n*-alkanes, a tracer for the unburned fuel. As shown in Figure S7 and S9, some SOA precursors

are enriched, and others depleted relative to fuel. Benzene and total IVOCs in gasoline and toluene and $C_8$ aromatics in diesel exhaust are enriched by more than a factor of two relative to unburned fuel. Enrichment of single-ring aromatics are likely due to pyrolysis of larger aromatic molecules (Akihama et al., 2002; Brezinsky, 1986). In contrast, total IVOCs (normalized to the $C_8$ to $C_{10}$ *n*-alkanes) are depleted in diesel exhaust compared to fuel (enrichment factors less than 1).

Figure 2.6 shows box-whisker plots of the total IVOC enrichment factors. Sources using more volatile fuel have higher IVOC enrichment factors. For example, relative to $C_{8-10}$ *n*-alkanes, gasoline engine exhaust has a median total IVOC enrichment factor of 8.5 versus modest depletion (enrichment factor <1) in diesel source exhaust with gas turbine exhaust in between. There are several possible explanations for this trend. IVOCs may be less efficiently combusted in the engines. Recent research also shows that less IVOCs are removed by catalytic converters compared to VOCs (Pereira et al., 2017). Figure S10 plot the IVOC enrichment factors of Pre-LEV, LEV and ULEV vehicles exhaust. Due to the different removal efficiency between IVOCs and VOCs, median ULEV vehicles show even higher (>10) IVOC enrichment factor. Lubricating oil decomposition products may also contribute to the IVOC emissions (May et al., 2013a; Worton et al., 2014). Finally, the IVOC fraction in fuel may be underestimated due to limitations in techniques used commonly to characterize fuel composition (Gentner et al., 2012).

### 2.4 Implications for OA formation

An important goal of this work is to develop emission profiles required to improve model predictions of SOA formation. Simulation of ambient OA concentrations requires accurate representation of both emissions and SOA yields for SVOCs and IVOCs. Given the

lack of IVOC data in traditional source profiles (Figure 2.3), previous modelling studies have used different scaling approaches, most commonly based on POA (Koo et al., 2014; Murphy et al., 2017; Robinson et al., 2007; Woody et al., 2016) but also using NMOG (Jathar et al., 2014, 2017) and naphthalene (Pye and Seinfeld, 2010). Finally, Gentner et al. (2012) used unburnt fuel surrogate to estimate IVOC emissions. These estimates are then combined with SOA yield data.

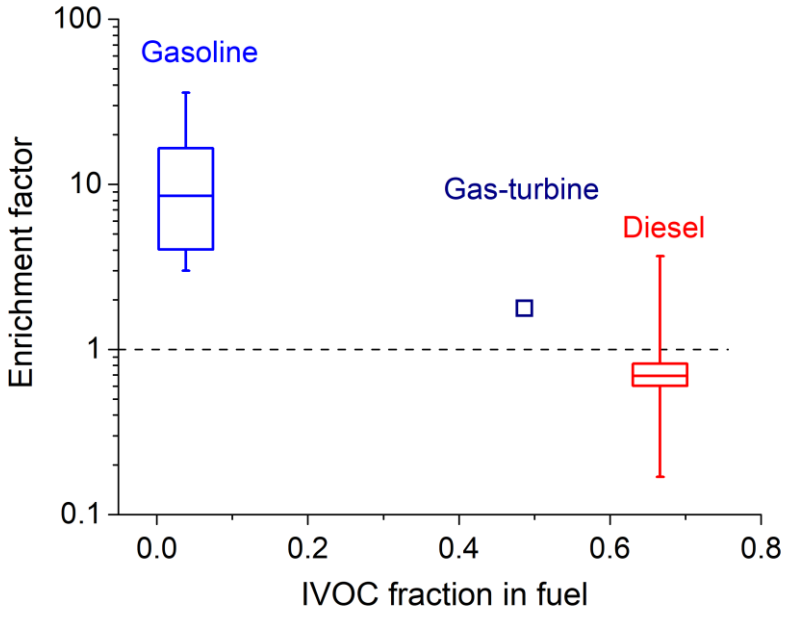



**Figure 2.6 IVOC mass enrichment factors as a function of IVOC content in fuel, $R_{Enrichment,i}$ = $(m_i^{exhaust}/m_{C_{8-10}}^{exhaust})/(m_i^{fuel}/m_{C_{8-10}}^{fuel})$. The box-whisker plots indicate variability in ratio within a given source class: 25$^{th}$ -75$^{th}$ percentiles and 10$^{th}$-90$^{th}$ percentiles.**

In this section, we use our new data to evaluate these different scaling approaches for estimating IVOC emissions to better understand their strengths and limitations for simulating ambient OA concentrations. Table 2.1 lists the different approaches we evaluated: (1) New – new profiles developed in this paper; (2) Trad – traditional profiles (SPECIATE #8750a for gasoline, #5565 for gas-turbine and #4674 for diesel sources emissions); (3) ROB:

traditional profiles + 1.5 × POA as IVOCs (Robinson et al., 2007); (4) MUR: traditional profiles + 9.656 × POA as IVOCs (Murphy et al., 2017); (5) PYE: traditional profiles + 66 × Naphthalene as IVOCs (Pye and Seinfeld, 2010); (6) GEN: using unburnt fuel composition as surrogate (Gentner et al., 2012); (7) JAT: 20% of NMOG of gasoline emission and 25% of diesel emissions are IVOCs (Jathar et al., 2014).

**Table 2.1 List of different estimates of IVOC emissions and SOA yield for mobile sources shown in Figure 2.7.**

| Label | IVOC Emissions | IVOC SOA yield (at OA = 10 μg/m$^3$) | Reference |
|---|---|---|---|
| **New** | Direct measurements | 0.22 - 0.30 | This work, Zhao et al. (2015, 2016) |
| **Trad** | N/A | N/A | EPA SPECIATE |
| **ROB** | 1.5 × POA | 0.15[a] | Robinson et al. (2007), Koo et al. (2014) |
| **MUR** | 9.656 × POA | N/A | Murphy et al. (2017) |
| **PYE** | 66 × Naphthalene | 0.22 | Pye and Seinfeld (2010) |
| **GEN** | Unburned fuel (1% of NMOG for gasoline, 62% for diesel) | 0.034 - 0.20 | Gentner et al. (2012) |
| **JAT** | Inverting chamber measurements (25% of NMOG for gasoline, 20% of NMOG for diesel) | 0.22 - 0.35 | Jathar et al. (2014) |

[a] From Koo et al. (2014)

Figure 2.7 compares our new data to the six previous estimates. Figure 2.7(a) shows the mass fraction of different classes of SOA precursors (VOC, IVOC and SVOC) in the NMOG emissions. Figure 2.7(b) shows the overall SOA yields of the total NMOG emissions for the different models (SOA mass/mass of NMOG emissions).

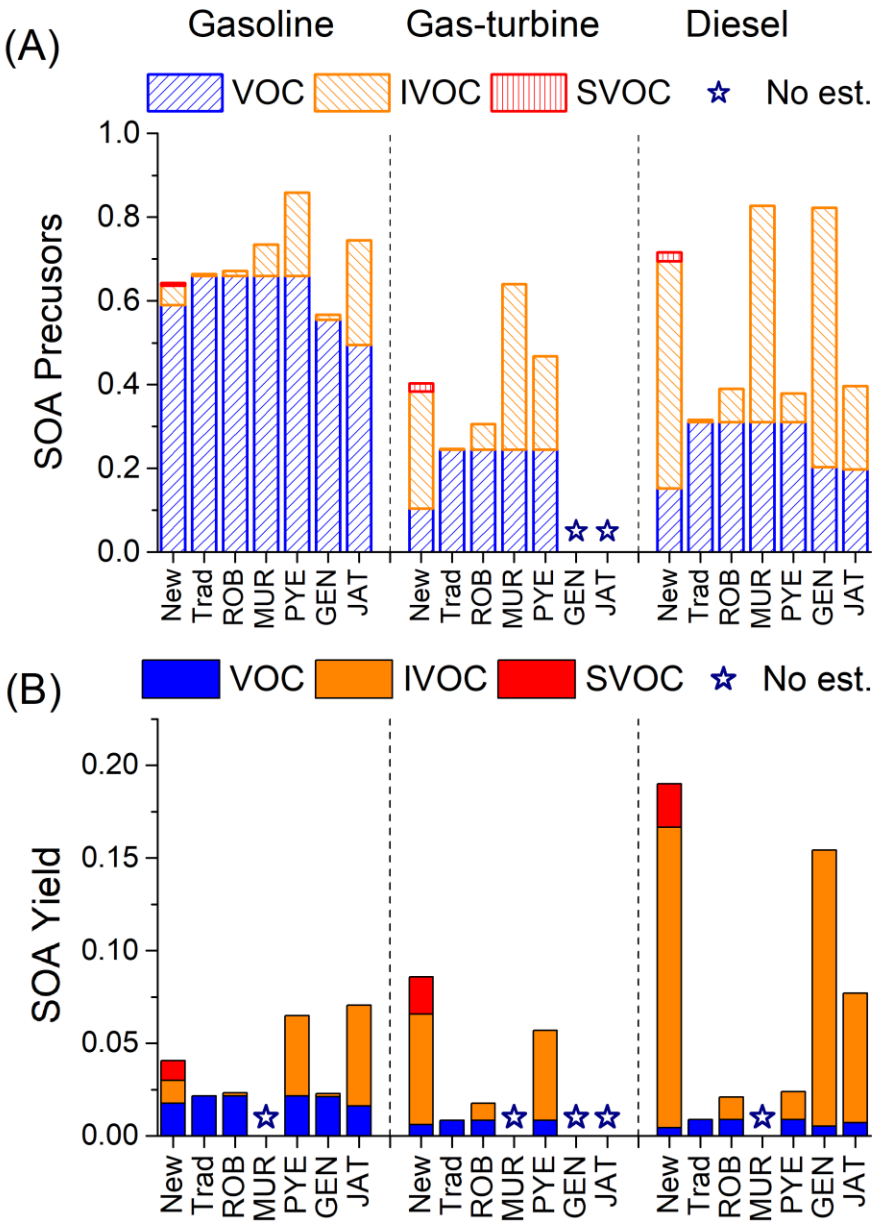



**Figure 2.7 Comparison of (a) mass fraction of SOA precursors in total NMOG emissions and (b) calculated total SOA yields of NMOG emissions from mobile sources based on the different estimation approaches listed in Table 2.1. Star denotes no estimate available.**

As shown in Figure 2.7(a), all estimates have similar VOC SOA precursor mass fractions, but widely divergent amounts of IVOCs. Our new profiles (1) and estimates (6) and (7) have modestly lower VOC SOA precursors, due to the inclusion of IVOCs and gas-

phase SVOC within NMOG emissions, while approaches (3) – (5) add additional IVOCs to the existing NMOG emissions. Since FID-based NMOG is a measure of all non-methane organic gases, we think IVOCs and gas-phase SVOCs are largely accounted for in existing NMOG emission factors for the types of sources measured here (see discussion of mass closure between bulk measurements and speciated measurements in section 2.2.2 and Figure S3). However, traditional source profiles do not correctly attribute these emissions to IVOCs/SVOCs in all sources.

The most common approach to incorporate IVOCs in models has been to scale POA emissions as defined by the organic mass collected on a quartz filter. The scaling ratios (e.g. IVOC-to-POA) were estimated from very limited data (a single or small number of sources) and the same ratio has typically been applied to all source categories. Our data indicate that scaling with POA is not a robust approach because IVOC-to-POA ratios vary by source category. For example, the average IVOC-to-POA ratios for gasoline engines exhaust are 6.2 ± 4.4 (cold-start) versus 12 ± 7 (non-DPF equipped) and 31 (DPF-equipped) for diesel exhaust (Zhao et al., 2015, 2016). In addition, these values are much larger than the widely used scaling factor of IVOC-to-POA of 1.5 (ROB in Figure 2.7) (Robinson et al., 2007), which grossly underestimates the IVOC emissions from the types of internal combustion engines considered here. While the IVOC-to-POA ratio of 9.6 by Murphy et al. (2017) (MUR in Figure 2.7) overestimates IVOC emissions from gasoline and gas-turbine sources, but underestimates it from diesel sources.

However, even if one uses source specific IVOC-to-POA scaling factors, we do not think that scaling POA provides a robust estimate of IVOC emissions from internal combustion engines. POA emissions are dominated by lubricant oil (Worton et al., 2014) while IVOC emissions appear to mainly arise from unburned fuel (Figure 2.2). In addition,

quartz filter measurements are subject to sampling artifacts and partitioning biases (May et al., 2013a, 2013b, 2013c). As a result, IVOC-to-POA ratios vary not only by source type (e.g. gasoline versus diesel) but also operating conditions (Zhao et al., 2015).

Zhao et al. (2015, 2016) reported stronger correlations between IVOC and total NMOG emissions than with POA over a range of operating conditions ($R^2$ = 0.96 vs 0.90 for gasoline and $R^2$ = 0.99 vs 0.91 for diesel sources, Figure S11). This is not surprising given that both NMOG and IVOC emissions arise from fuel and are controlled by similar processes. This suggests that IVOC emissions should be estimated using source specific scaling factors of NMOG not POA.

Jathar et al. (2014, 2017) estimated IVOC emissions by scaling NMOG. They also used different ratios for gasoline and diesel sources. However, they did not directly measure IVOCs. Instead they inferred IVOC-to-NMOG ratios using a combination of unspeciated emissions and inverse modelling of SOA production measured in a smog chamber. Using this approach, they attributed 25% of NMOG emission from gasoline engine and 20% from diesel engines to IVOCs. These values are very different than those reported here, which are based on direct measurements. A detail on the ratios of Jathar et al. (2014) is that they were derived to be used in combination with their empirically derived SOA yields. When used together they explain SOA yield production measured in smog chamber experiments with dilute exhaust. Therefore, one cannot simply replace IVOC-to-NMOG of Jathar et al. (2014) with the ones reported here without also using different SOA yields.

Pye and Seinfeld (2010) estimated IVOC emissions by scaling naphthalene using the same IVOC-to-naphthalene ratio for all sources. Our data indicate that naphthalene is not a good indicator of IVOCs, due to the large variation in fuel aromatic content. For example, there is four times more naphthalene in gasoline engine exhaust (0.4%) and fuel (0.13%)

compared to diesel engine exhaust (0.1%) and fuel (0.04%). Therefore, the approach of Pye and Seinfeld (2010) generates much higher estimates of IVOC emissions from gasoline than diesel sources, which is opposite of the actual emissions data (Figure 2.2). In principle, this problem can be overcome with source specific IVOC-to-naphthalene ratios, but even with source-specific ratios, individual organics are likely a less robust scaler for IVOCs than total NMOG because fuel composition (e.g. aromatic content) varies by location and season.

A final approach to estimate IVOC emissions is to use unburned fuel as a surrogate for the SOA production of exhaust. Gentner et al. (2012) used this approach to estimate the IVOC fraction, as well as the SOA yield of gasoline and diesel engine exhaust. This approach works for diesel, but not for gasoline given the enrichment of IVOCs in the exhaust (Figure 2.6).

In Figure 2.7(b), we combine the different emissions estimates with SOA yield data to calculate the SOA yield of the NMOG emissions for each source category, assuming complete oxidation of all precursors. Our new profiles predict that IVOCs and SVOC vapors contribute substantially to SOA production, especially for sources using lower-volatility fuels (e.g. diesel). For gasoline sources, we predict that IVOCs and SVOCs contribute as much SOA as traditional VOC precursors (mainly single-ring aromatics). Accounting for IVOCs in gasoline exhaust almost doubles the predicted SOA production compared to the traditional profile. For gas-turbine and diesel sources, IVOCs and SVOC vapors combining contribute factors of 13 and 44 more SOA than VOCs, respectively.

Figure 2.7(b) also compares the SOA yields of NMOG emissions for all the different approaches (2) – (7). The differences in effective yields are primarily due to differences in IVOC/SVOC emissions. Traditional profiles and ROB underpredicts SOA production from all three source categories because they underestimate IVOC emissions. As discussed in

section 2.3.4, IVOCs are enriched in gasoline emissions compared to unburned fuel, therefore GEN underpredicts the SOA yield of gasoline emissions. However, fuel composition provides a reasonable estimate for SOA production from diesel emissions, except for the lack of SVOCs potentially produced from the usage of lubricant oil. The approaches of PYE and JAT overpredict the overall SOA production from gasoline emissions, due to their overestimation of IVOC emissions, but both underestimate the overall SOA production for diesel emissions.

To conclude, none of the previous modelling approaches provide a robust estimate of the IVOC fraction in the exhaust for all three source categories. Figure 2.7(a) and (b) show that traditional profiles either completely omit IVOCs or incorrectly lumped them to VOC chemical mechanism groups, which greatly underestimate the overall SOA formation potential. Approaches (3) – (5) apply scaling factors to certain species, such as POA and naphthalene, but these factors vary by source and fuel composition, which may lead to significant bias for different sources. Using unburnt fuel composition as a surrogate in estimation (6) only works for sources that use lower volatility fuel, such as diesel.

In addition to better representing gas-phase SOA precursor emissions, the new profiles also account for the semi-volatile character of POA. Partitioning calculations predict that 40% to 50% of traditionally defined POA mass evaporates at typical atmospheric conditions (T=298K and OA=10 µg/m) (May et al., 2013b, 2013a).

## 2.5 Recommendations and future research needs

Figure 2.7 highlights the importance of including IVOC and SVOC emissions in models and inventories to improve predictions of SOA formation. This paper facilities this by providing model-ready profiles that include direct measurements of IVOCs and SVOCs.

They are designed to be applied to existing inventories of POA and NMOG emissions. These profiles (Table S3) are normalized to total organic emissions (VOC, IVOC, SVOC, LVOC plus NVOC), and therefore should be applied to the sum of gas- and particle-phase organic emissions. Since current emission inventories report gas- (NMOG or VOC) and particle-phase (PM or POA) emissions separately, the comprehensive profile can be separated into two parts: gas-phase (VOC and IVOC) and particle-phase (SVOC and less volatile components) profiles. These sub-profiles would be renormalized and then applied to the existing NMOG or POA emissions. With this approach, one needs to correct the POA data for missing SVOC vapors not collected during vehicle certification testing (the factor of 1.4 for LDGV discussed at the beginning of section 2.3).

Our new profiles intentionally do not define the phase state of the emissions. Phase state is not a property of the emissions, but determined by the combination of the volatility distribution of the emissions and atmospheric conditions because gas-particle partitioning depends on the concentration of organic aerosol and temperature. The profiles specify the volatility distribution of the emissions, which can then be used to calculate the gas-particle partitioning (phase state) for any atmospheric condition (Robinson et al., 2010). This approach is critical to correctly predict POA concentrations for sources that have substantial SVOC emissions, such as the sources tested here and biomass smoke (May et al., 2013c).

The three types of sources considered here account for 98.2% of the mobile source emissions in the 2014 US EPA National Emission Inventory. For other liquid-fuel internal combustion engine sources, we recommend interpolating based on fuel composition and applying the IVOC enrichment factor estimated from fuel volatility (Figure 2.6). For sources profiles that only contain speciated VOCs and unknown residual, we recommend not normalizing to known species, as this will likely misattribute low volatility organics to

VOCs.

The emission profiles described here (except for gas turbines) are based on experiments conducted using sources recruited from the California in-use fleet, at typical California summertime temperatures (10-25 °C) and using California commercial summertime fuels. Therefore, the data are most directly relevant to California summertime conditions. Ambient temperature can have a large influence on emissions. For example, George et al. (2015) measured about 10 times higher non-methane hydrocarbon (NMHC) emission rates during testing at -7 °C versus 24 °C. The VOC composition also changed with temperature with the fraction of $C_{9+}$ aromatics almost doubling at low temperature. These data suggest that winter emissions may have higher content of larger aromatics ($C_{9+}$ aromatics) and IVOCs, due to incomplete combustion or lower efficiency of catalytic converter. Our profiles therefore likely represent lower bounds to winter vehicle emission in terms of aromatics and IVOC contents. As discussed in section 2.3.2, unburned fuel is an important contributor to the emissions. Therefore, variations in fuel composition by, for example, season and/or location will influence the composition of the emissions. From an SOA formation perspective, we are most interested in changes in fuel IVOC and aromatic content. Figure S12 compares our new VOC profiles with data from China (Cao et al., 2016; Yao et al., 2015). There is good agreement for many compounds, but not all.

Future research needs:

1) IVOC and SVOC emissions data from vehicles operated over a wider range of conditions. Comprehensive emissions data are needed for a wider range of fuel compositions, test cycles (hot operations), and seasons (especially winter). However, given their major contribution to SOA formation, we recommend using our new profiles even for studies outside of California if the only other option is to use traditional profiles that don't

include IVOC and SVOC data.

2) IVOC and SVOC emissions data for non-mobile sources. Recent research has demonstrated that IVOCs and SVOCs are important contributors to biomass burning, oil sands, oil production, and volatile chemical product emissions (de Gouw et al., 2011; Hatch et al., 2017; Hunter et al., 2017; Liggio et al., 2016). More comprehensive, model-ready profiles that account for the full spectrum of organic emissions are needed for these and other source categories (McDonald et al., 2018).

3) Inclusion of IVOCs in air quality models and inventories. Our new profiles are designed to directly incorporate IVOCs into models and inventory. Since they are based on direct measurements, they do not have the large uncertainties associated with the previously developed scaling approaches.

4) Improved chemical composition of IVOCs and SVOCs. Although we have quantified the total IVOC emissions, the majority of these emissions were not resolved at the molecular level. Since the SOA yield of compounds depend on both molecular structure and volatility (Tkacik et al., 2012; Ziemann, 2011), future studies are needed to more fully speciate IVOCs and SVOCs in order to identify the class of compounds that needed for photo-oxidation experiments (Chan et al., 2013; Cross et al., 2015; Isaacman et al., 2012b).

5) Measurements and source apportionment of atmospheric IVOCs / SVOCs. Ambient measurements of IVOCs / SVOCs are needed to identify other important sources of atmospheric IVOCs / SVOCs. This will help future studies to prioritize which sources to characterize.

## 2.6 Acknowledgments

This publication was developed as part of the Center for Air, Climate, and Energy

Solutions (CACES), which is supported under Assistance Agreement No. RD83587301 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by EPA. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication. The authors would like to thank Albert Presto, Timothy Gordon and Andrew May for their help in providing detailed datasets information. The authors would also like to thank Benjamin Murphy and Havala Pye for helpful comments on this manuscript and valuable discussion on missing SVOC scaling approach.

## 2.7 References

Adelman, Z;Vukovich, JCarter, W. P. L.: Integration of the SAPRC Chemical Mechanism in the SMOKE Emissions Processor for the CMAQ/Models- 3 Airshed Model., 2005.

Akihama, K., Takatori, Y. and Nakakita, K.: Effect of hydrocarbon molecular structure on diesel exhaust emissions, R&D Rev. Toyota CSDL, 37, 46–52, 2002.

Baker, K. R., Carlton, A. G., Kleindienst, T. E., Offenberg, J. H., Beaver, M. R., Gentner, D. R., Goldstein, A. H., Hayes, P. L., Jimenez, J. L., Gilman, J. B., De Gouw, J. A., Woody, M. C., Pye, H. O. T. T., Kelly, J. T., Lewandowski, M., Jaoui, M., Stevens, P. S., Brune, W. H., Lin, Y. H., Rubitschun, C. L. and Surratt, J. D.: Gas and aerosol carbon in California: Comparison of measurements and model predictions in Pasadena and Bakersfield, Atmos. Chem. Phys., 15(9), 5243–5258, doi:10.5194/acp-15-5243-2015, 2015.

Brezinsky, K.: The high-temperature oxidation of aromatic hydrocarbons, Prog. Energy Combust. Sci., 12(1), 1–24, doi:https://doi.org/10.1016/0360-1285(86)90011-0, 1986.

Cao, T., Durbin, T. D., Russell, R. L., Cocker, D. R., Scora, G., Maldonado, H. and Johnson, K. C.: Evaluations of in-use emission factors from off-road construction equipment, Atmos. Environ., 147, 234–245, doi:https://doi.org/10.1016/j.atmosenv.2016.09.042, 2016.

Carter, W. P. L.: Saprc-07 Chemical Mechanism, , (05), 2010.

Carter, W. P. L. L.: Development of a database for chemical mechanism assignments for volatile organic emissions, J. Air Waste Manage. Assoc., 65(10), 1171–1184, doi:10.1080/10962247.2015.1013646, 2015.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crounse, J. D., Kürten, A., Wennberg, P. O., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: Implications for oxidation of intermediate volatility organic compounds (IVOCs), Atmos. Chem. Phys., 9(9), 3049–3060, doi:10.5194/acp-9-3049-2009, 2009.

Chan, A. W. H., Isaacman, G., Wilson, K. R., Worton, D. R., Ruehl, C. R., Nah, T., Gentner, D. R., Dallmann, T. R., Kirchstetter, T. W., Harley, R. A., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Offenberg, J. H., Kleindienst, T. E., Lin, Y. H., Rubitschun, C. L., Surratt, J. D., Hayes, P. L.,

Jimenez, J. L. and Goldstein, A. H.: Detailed chemical characterization of unresolved complex mixtures in atmospheric organics: Insights into emission sources, atmospheric processing, and secondary organic aerosol formation, J. Geophys. Res. Atmos., 118(12), 6783–6796, doi:10.1002/jgrd.50533, 2013.

Corporan, E., Edwards, T., Shafer, L., DeWitt, M. J., Klingshirn, C., Zabarnick, S., West, Z., Striebich, R., Graham, J. and Klein, J.: Chemical, thermal stability, seal swell and emissions characteristics of jet fuels from alternative sources, 11th Int. Conf. Stability, Handl. Use Liq. Fuels 2009, 2, 973–1014, 2009.

Cross, E. S., Hunter, J. F., Carrasquillo, A. J., Franklin, J. P., Herndon, S. C., Jayne, J. T., Worsnop, D. R., Miake-Lye, R. C. and Kroll, J. H.: Online measurements of the emissions of intermediate-volatility and semi-volatile organic compounds from aircraft, Atmos. Chem. Phys., 13(15), 7845–7858, doi:10.5194/acp-13-7845-2013, 2013.

Cross, E. S., Sappok, A. G., Wong, V. W. and Kroll, J. H.: Load-Dependent Emission Factors and Chemical Characteristics of IVOCs from a Medium-Duty Diesel Engine, Environ. Sci. Technol., 49(22), 13483–13491, doi:10.1021/acs.est.5b03954, 2015.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F. and Schwartz, J. D.: Air Pollution and Mortality in the Medicare Population, N. Engl. J. Med., 376(26), 2513–2522, doi:10.1056/NEJMoa1702747, 2017.

Donahue, N. M., Robinson, A. L., Stanier, C. O. and Pandis, S. N.: Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics, Environ. Sci. Technol., 40(8), 2635–2643, doi:10.1021/es052297c, 2006.

Drozd, G. T., Miracolo, M. A., Presto, A. A., Lipsky, E. M., Riemer, D. D., Corporan, E. and Robinson, A. L.: Particulate Matter and Organic Vapor Emissions from a Helicopter Engine Operating on Petroleum and Fischer–Tropsch Fuels, Energy & Fuels, 26(8), 4756–4766, doi:10.1021/ef300651t, 2012.

Ensberg, J. J., Hayes, P. L., Jimenez, J. L., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Holloway, J. S., Gordon, T. D., Jathar, S., Robinson, A. L. and Seinfeld, J. H.: Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation, Atmos. Chem. Phys., 14(5), 2383–2397, doi:10.5194/acp-14-2383-2014, 2014.

Fujitani, Y., Saitoh, K., Fushimi, A., Takahashi, K., Hasegawa, S., Tanabe, K., Kobayashi, S., Furuyama, A., Hirano, S. and Takami, A.: Effect of isothermal dilution on emission factors of organic carbon and n-alkanes in the particle and gas phases of diesel exhaust, Atmos. Environ., 59, 389–397, doi:10.1016/j.atmosenv.2012.06.010, 2012.

Gabele, P.: Exhaust Emissions from Four-Stroke Lawn Mower Engines, J. Air Waste Manage. Assoc., 47(9), 945–952, doi:10.1080/10473289.1997.10463951, 1997.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A. and Goldstein, A. H.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, Proc. Natl. Acad. Sci., 109(45), 18318–18323, doi:10.1073/pnas.1212272109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1212272109, 2012.

Gentner, D. R., Jathar, S. H., Gordon, T. D., Bahreini, R., Day, D. A., El Haddad, I., Hayes, P. L., Pieber, S. M., Platt, S. M., de Gouw, J., Goldstein, A. H., Harley, R. A., Jimenez, J. L., Prévôt, A. S. H. and Robinson, A. L.: Review of Urban Secondary Organic Aerosol Formation from Gasoline and Diesel Motor Vehicle Emissions, Environ. Sci. Technol., acs.est.6b04509, doi:10.1021/acs.est.6b04509, 2017.

Goldstein, A. H. and Galbally, I. E.: Known and Unexplored Organic Constituents in the Earth's Atmosphere, Environ. Sci. Technol., 41(5), 1514–1521, doi:10.1021/es072476p, 2007.

Gordon, T. D., Tkacik, D. S., Presto, A. A., Zhang, M. and Shantanu, H.: Primary Gas- and Particle-Phase Emissions and Secondary Organic Aerosol Production from Gasoline and Diesel Off-Road Engines, 2013.

de Gouw, J. A., Middlebrook, A. M., Warneke, C., Ahmadov, R., Atlas, E. L., Bahreini, R., Blake, D. R., Brock, C. A., Brioude, J., Fahey, D. W., Fehsenfeld, F. C., Holloway, J. S., Le Henaff, M., Lueb, R. A., McKeen, S. A., Meagher, J. F., Murphy, D. M., Paris, C., Parrish, D. D., Perring, A. E., Pollack, I. B., Ravishankara, A. R., Robinson, A. L., Ryerson, T. B., Schwarz, J. P., Spackman, J. R., Srinivasan, A. and Watts, L. A.: Organic Aerosol Formation Downwind from the Deepwater Horizon Oil Spill, Science (80-. )., 331(6022), 1295 LP-1299 [online] Available from: http://science.sciencemag.org/content/331/6022/1295.abstract, 2011.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R. and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9(14), 5155–5236, doi:10.5194/acp-9-5155-2009, 2009.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J. and Barsanti, K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors, Atmos. Chem. Phys., 17(2), 1471–1489, doi:10.5194/acp-17-1471-2017, 2017.

Hodzic, A., Jimenez, J. L., Madronich, S., Canagaratna, M. R., Decarlo, P. F., Kleinman, L. and Fast, J.: Modeling organic aerosols in a megacity: Potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation, Atmos. Chem. Phys., 10(12), 5491–5514, doi:10.5194/acp-10-5491-2010, 2010.

Hunter, J. F., Day, D. A., Palm, B. B., Yatavelli, R. L. N., Chan, A. W. H., Kaser, L., Cappellin, L., Hayes, P. L., Cross, E. S., Carrasquillo, A. J., Campuzano-Jost, P., Stark, H., Zhao, Y., Hohaus, T., Smith, J. N., Hansel, A., Karl, T., Goldstein, A. H., Guenther, A., Worsnop, D. R., Thornton, J. A., Heald, C. L., Jimenez, J. L. and Kroll, J. H.: Comprehensive characterization of atmospheric organic carbon at a forested site, Nat. Geosci., 10, 748 [online] Available from: http://dx.doi.org/10.1038/ngeo3018, 2017.

Isaacman, G., Chan, A. W. H., Nah, T., Worton, D. R., Ruehl, C. R., Wilson, K. R. and Goldstein, A. H.: Heterogeneous OH oxidation of motor oil particles causes selective depletion of branched and less cyclic hydrocarbons, Environ. Sci. Technol., 46(19), 10632–10640, doi:10.1021/es302768a, 2012a.

Isaacman, G., Wilson, K. R., Chan, A. W. H., Worton, D. R., Kimmel, J. R., Nah, T., Hohaus, T., Gonin, M., Kroll, J. H., Worsnop, D. R. and Goldstein, A. H.: Improved resolution of hydrocarbon structures and constitutional isomers in complex mixtures using gas chromatography-vacuum ultraviolet-mass spectrometry, Anal. Chem., 84(5), 2335–2342, doi:10.1021/ac2030464, 2012b.

Jathar, S. H., Gordon, T. D., Hennigan, C. J., Pye, H. O. T., Pouliot, G., Adams, P. J., Donahue, N. M. and Robinson, A. L.: Unspeciated organic emissions from combustion sources and their influence on the secondary organic aerosol budget in the United States, Proc. Natl. Acad. Sci., 111(29), 10473–10478, doi:10.1073/pnas.1323740111, 2014.

Jathar, S. H., Woody, M., Pye, H. O. T., Baker, K. R. and Robinson, A. L.: Chemical transport model simulations of organic aerosol in southern California: model evaluation and gasoline and diesel source contributions, Atmos. Chem. Phys., 17(6), 4305–4318, doi:10.5194/acp-17-4305-2017, 2017.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U. and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, Science (80-. )., 326(5959), 1525–1529 [online] Available from: http://science.sciencemag.org/content/326/5959/1525.abstract, 2009.

Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Van Dingenen, R., Ervens, B., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G. and Wilson, J.: Organic aerosol and global climate modelling: a review, Atmos. Chem. Phys., 5(35), 1053–1123, doi:10.5194/acp-5-1053-2005, 2005.

Kishan, S., Burnette, A. and Fincher, S.: Kansas City PM Characterization Study Final Report, , 1–462 [online] Available from: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Kansas+City+PM+Characterization+Study+Final+Report#1, 2008.

Komkoua Mbienda, A. J., Tchawoua, C., Vondou, D. A. and Mkankam Kamga, F.: Evaluation of vapor pressure estimation methods for use in simulating the dynamic of atmospheric organic aerosols, Int. J. Geophys., 2013, 2013.

Koo, B., Knipping, E. and Yarwood, G.: 1.5-Dimensional volatility basis set approach for modeling organic aerosol in CAMx and CMAQ, Atmos. Environ., 95, 158–164, doi:10.1016/j.atmosenv.2014.06.031, 2014.

Kroll, J. H. and Seinfeld, J. H.: Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere, Atmos. Environ., 42(16), 3593–3624, doi:10.1016/j.atmosenv.2008.01.003, 2008.

Kuwayama, T., Collier, S., Forestieri, S., Brady, J. M., Bertram, T. H., Cappa, C. D., Zhang, Q. and Kleeman, M. J.: Volatility of Primary Organic Aerosol Emitted from Light Duty Gasoline Vehicles, Environ. Sci. Technol., 49(3), 1569–1577, doi:10.1021/es504009w, 2015.

Li, X., Dallmann, T. R., May, A. A., Tkacik, D. S., Lambe, A. T., Jayne, J. T., Croteau, P. L. and Presto, A. A.: Gas-Particle Partitioning of Vehicle Emitted Primary Organic Aerosol Measured in a Traffic Tunnel, Environ. Sci. Technol., acs.est.6b01666, doi:10.1021/acs.est.6b01666, 2016.

Liggio, J., Li, S., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., Drollette, B. D., Gordon, M., Lee, P., Liu, P., Leithead, A., Moussa, S. G., Wang, D., Brien, J. O., Mittermeier, R. L., Osthoff, H. D., Makar, P. A., Zhang, J., Brook, J. R., Lu, G., Staebler, R. M., Han, Y., Travis, W., Plata, D. L. and Gentner, D. R.: Oil sands operations as a large source of secondary organic aerosols, Nature, 534(7605), 1–16, doi:10.1038/nature17646, 2016.

Lipsky, E. M. and Robinson, A. L.: Effects of Dilution on Fine Particle Mass and Partitioning of Semivolatile Organics in Diesel Exhaust and Wood Smoke, Environ. Sci. Technol., 40(1), 155–162, doi:10.1021/es050319p, 2006.

May, A. A., Presto, A. A., Hennigan, C. J., Nguyen, N. T., Gordon, T. D. and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: (1) Gasoline vehicle exhaust, Atmos. Environ., 77, 128–139, doi:10.1016/j.atmosenv.2013.04.060, 2013a.

May, A. A., Presto, A. A., Hennigan, C. J., Nguyen, N. T., Gordon, T. D. and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: (2) diesel vehicles, Environ. Sci. Technol., 47(15), 8288–8296, doi:10.1021/es400782j, 2013b.

May, A. A., Levin, E. J. T., Hennigan, C. J., Riipinen, I., Lee, T., Collett, J. L., Jimenez, J. L., Kreidenweis, S. M. and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: 3. Biomass burning, J. Geophys. Res. Atmos., 118(19), 11327–11338, doi:10.1002/jgrd.50828, 2013c.

May, A. A., Nguyen, N. T., Presto, A. A., Gordon, T. D., Lipsky, E. M., Karve, M., Gutierrez, A., Robertson, W. H., Zhang, M., Brandow, C., Chang, O., Chen, S., Cicero-Fernandez, P., Dinkins, L., Fuentes, M., Huang, S. M., Ling, R., Long, J., Maddox, C., Massetti, J., McCauley, E., Miguel, A., Na, K., Ong, R., Pang, Y., Rieger, P., Sax, T., Truong, T., Vo, T., Chattopadhyay, S., Maldonado, H., Maricq, M. M. and Robinson, A. L.: Gas- and particle-phase primary emissions from in-use, on-road gasoline and diesel vehicles, Atmos. Environ., 88, 247–260, doi:10.1016/j.atmosenv.2014.01.046, 2014.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B. and Trainer, M.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science (80-. )., 359(6377), 760–764, doi:10.1126/science.aaq0524, 2018.

Murphy, B. N., Woody, M. C., Jimenez, J. L., Carlton, A. M. G., Hayes, P. L., Liu, S., Ng, N. L., Russell, L. M., Setyan, A., Xu, L., Young, J., Zaveri, R. A., Zhang, Q. and Pye, H. O. T.: Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: Impacts on source strength and partitioning, Atmos. Chem. Phys., 17(18), 11107–11133, doi:10.5194/acp-17-11107-2017, 2017.

Pereira, K. L., Dunmore, R., Whitehead, J., Rami Alfarra, M., Allan, J. D., Alam, M. S., Harrison, R. M., Mcfiggans, G. and Hamilton, J. F.: The Effect of Varying Engine Conditions on Unregulated VOC Diesel Exhaust Emissions, Atmos. Chem. Phys. Discuss., (August), 1–25, doi:10.5194/acp-2017-603, 2017.

Presto, A. a., Hennigan, C. J., Nguyen, N. T. and Robinson, A. L.: Determination of Volatility Distributions of Primary Organic Aerosol Emissions from Internal Combustion Engines Using Thermal Desorption Gas Chromatography Mass Spectrometry, Aerosol Sci. Technol., 46(10), 1129–1139, doi:10.1080/02786826.2012.700430, 2012.

Presto, A. A., Nguyen, N. T., Ranjan, M., Reeder, A. J., Lipsky, E. M., Hennigan, C. J., Miracolo, M. A., Riemer, D. D. and Robinson, A. L.: Fine particle and organic vapor emissions from staged tests of an in-use aircraft engine, Atmos. Environ., 45(21), 3603–3612, doi:10.1016/j.atmosenv.2011.03.061, 2011.

Presto, A. a, Miracolo, M. a, Donahue, N. M. and Robinson, A. L.: Secondary Organic Aerosol Formation from High-NO x Photo-Oxidation of Low Volatility Precursors : n -Alkanes, Atmos. Chem. Phys., 44(6), 2029–2034, 2010.

Pye, H. O. T. and Pouliot, G. A.: Modeling the Role of Alkanes, Polycyclic Aromatic Hydrocarbons, and Their Oligomers in Secondary Organic Aerosol Formation, Environ. Sci. Technol., 46(11), 6041–6047, doi:10.1021/es300409w, 2012.

Pye, H. O. T. and Seinfeld, J. H.: A global perspective on aerosol from low-volatility organic compounds, Atmos. Chem. Phys., 10(9), 4377–4401, doi:10.5194/acp-10-4377-2010, 2010.

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. a, Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R. and Pandis, S. N.: Rethinking Organic Aerosols :, Science (80-. )., 315(March), 1259–1262, doi:10.1126/science.1133061, 2007.

Schauer, J. J., Kleeman, M. J., Cass, G. R. and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 1. C1 through C29 Organic Compounds from Meat Charbroiling, Environ. Sci. Technol., 33(10), 1566–1577, doi:10.1021/es980076j, 1999a.

Schauer, J. J., Kleeman, M. J., Cass, G. R. and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 2. C1through C30Organic Compounds from Medium Duty Diesel Trucks, Environ. Sci. Technol., 33(10), 1578–1587, doi:10.1021/es980081n, 1999b.

Schauer, J. J., Kleeman, M. J., Cass, G. R. and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 5. C1−C32 Organic Compounds from Gasoline-Powered Motor Vehicles, Environ. Sci. Technol., 36(6), 1169–1180, doi:10.1021/es0108077, 2002.

Shrivastava, M. K., Lane, T. E., Donahue, N. M., Pandis, S. N. and Robinson, A. L.: Effects of gas particle partitioning and aging of primary emissions on urban and regional organic aerosol concentrations, J. Geophys. Res. Atmos., 113(18), 1–16, doi:10.1029/2007JD009735, 2008.

Siegl, W. O., Hammerle, R. H., Herrmann, H. M., Wenclawiak, B. W. and Luers-Jongen, B.: Organic emissions profile for a light-duty diesel vehicle, Atmos. Environ., 33(5), 797–805, doi:10.1016/S1352-2310(98)00209-X, 1999.

Turpin, B. J. and Lim, H. J.: Species contributions to PM2.5 mass concentrations: Revisiting common assumptions for estimating organic mass, AEROSOL Sci. Technol., 35(1), 602–610, doi:10.1080/02786820119445, 2001.

USEPA-OAQPS: 2011 National Emissions Inventory Data & Documentation, , https://www.epa.gov/air-emissions-inventories/2014 [online] Available from: http://www3.epa.gov/ttnchie1/net/2011inventory.html, 2015.

USEPA: Estimation Programs Interface SuiteTM for Microsoft® Windows v 4.11, [online] Available from: https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface, 2012.

USEPA: CMAQv5.1 SOA Update, Community Model. Anal. Syst. Wiki [online] Available from: https://www.airqualitymodeling.org/index.php/CMAQv5.1_SOA_Update, 2016a.

USEPA: SPECIATE Version 4.5 Database Development Documentation, , (September) [online] Available from: https://www.epa.gov/sites/production/files/2016-09/documents/speciate_4.5.pdf, 2016b.

USEPA: MOVES 2014a, [online] Available from: https://www.epa.gov/moves/moves2014a-latest-version-motor-vehicle-emission-simulator-moves#manuals, n.d.

Volckens, J., Olson, D. A. and Hays, M. D.: Carbonaceous species emitted from handheld two-stroke engines, Atmos. Environ., 42(5), 1239–1248, doi:10.1016/j.atmosenv.2007.10.032, 2008.

Wang, Y., Raihala, T. S., Jackman, A. P. and St. John, R.: Use of Tedlar Bags in VOC Testing and Storage: Evidence of Significant VOC Losses, Environ. Sci. Technol., 30(10), 3115–3117, doi:10.1021/es950582y, 1996.

Woody, M. C., Baker, K. R., Hayes, P. L., Jimenez, J. L., Koo, B. and Pye, H. O. T. T.: Understanding sources of organic aerosol during CalNex-2010 using the CMAQ-VBS, Atmos. Chem. Phys., 16(6), 4081–4100, doi:10.5194/acp-16-4081-2016, 2016.

Worton, D. R., Isaacman, G., Gentner, D. R., Dallmann, T. R., Chan, A. W. H., Ruehl, C., Kirchstetter, T. W., Wilson, K. R., Harley, R. A. and Goldstein, A. H.: Lubricating Oil Dominates Primary Organic Aerosol Emissions from Motor Vehicles, Environ. Sci. Technol., 48(7), 3698–3706, doi:10.1021/es405375j, 2014.

Zhao, Y., Hennigan, C. J., May, A. A., Tkacik, D. S., De Gouw, J. A., Gilman, J. B., Kuster, W. C., Borbon, A. and Robinson, A. L.: Intermediate-volatility organic compounds: A large source of

secondary organic aerosol, Environ. Sci. Technol., 48(23), 13743–13750, doi:10.1021/es5035188, 2014.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environ. Sci. Technol., 49(19), 11516–11526, doi:10.1021/acs.est.5b02841, 2015.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines, Environ. Sci. Technol., 50(8), 4554–4563, doi:10.1021/acs.est.5b06247, 2016.

Zhao, Y., Saleh, R., Saliba, G., Presto, A. A., Gordon, T. D., Drozd, G. T., Goldstein, A. H., Donahue, N. M. and Robinson, A. L.: Reducing secondary organic aerosol formation from gasoline vehicle exhaust, Proc. Natl. Acad. Sci. , 114(27), 6984–6989, doi:10.1073/pnas.1620911114, 2017.

Zielinska, B., Sagebiel, J. C., Harshfield, G., Gertler, A. W. and Pierson, W. R.: F R O M M O T O R VEHICLES ; M E A S U R E M E N T M E T H O D S, , 30(12), 2269–2286, 1996.

# Chapter 3: Simulation of organic aerosol formation during the CalNex study: updated mobile emissions and secondary organic aerosol parameterization for intermediate volatility organic compounds *


## Abstract

We describe simulations using an updated version of the Community Multiscale Air Quality model version 5.3 (CMAQ v5.3) to investigate the contribution of intermediate volatile organic compounds (IVOCs) to secondary organic aerosol formation (SOA) in Southern California during the CalNex study. We first derive a model-ready parameterization for SOA formation from IVOC emissions from mobile sources. To account for SOA formation from both diesel and gasoline sources, the parameterization has six lumped precursor species that resolve both volatility and molecular structure (aromatic versus aliphatic). We also implement new mobile-source emission profiles that quantify all IVOCs based on direct measurements. The profiles have been released in SPECIATE 5.0. By incorporating both comprehensive mobile-source emissions profiles for SVOCs and IVOCs and experimentally constrained SOA yields, this CMAQ configuration best represents the contribution of mobile sources to urban and regional ambient OA. In the Los Angeles region, gasoline sources emit 4 times more non-methane organic gases (NMOG) than diesel sources, but diesel emits roughly 3 times more IVOCs on an absolute basis. The revised model predicts all mobile


---


* Originally published as Lu, Q., Murphy, B. N., Qin, M., Adams, P. J., Zhao, Y., Pye, H. O. T., Efstathiou, C., Allen, C., and Robinson, A. L.: Simulation of organic aerosol formation during the CalNex study: updated mobile emissions and secondary organic aerosol parameterization for intermediate-volatility organic compounds, Atmos. Chem. Phys., 20, 4313–4332, https://doi.org/10.5194/acp-20-4313-2020, 2020.

sources (including on- and off-road gasoline, aircraft and on- and off-road diesel) contribute ~1 μg $m^{-3}$ to the daily peak SOA concentration in Pasadena. This represents a ~70% increase in predicted daily peak SOA formation compared to the base version of CMAQ. Therefore, IVOCs in mobile-source emissions contribute almost as much SOA as traditional precursors such as single-ring aromatics. However, accounting for these emissions in CMAQ does not reproduce measurements of either ambient SOA or IVOCs. To investigate the potential contribution of other IVOC sources, we performed two exploratory simulations with varying amounts of IVOC emissions from non-mobile sources. To close the mass balance of primary hydrocarbon IVOCs, IVOCs would need to account for 12% of NMOG emissions from non-mobile sources (or equivalently 30.7 Ton $day^{-1}$ in the Los Angeles-Pasadena region), a value that is well within the reported range of IVOC content from volatile chemical products. To close the SOA mass balance and also explain the mildly oxygenated IVOCs in Pasadena, an additional 14.8% of non-mobile source NMOG emissions would need to be IVOCs (assuming SOA yields from the mobile IVOCs applies to non-mobile IVOCs). However, an IVOC-to-NMOG ratio of 26.8% (or equivalently 68.5 Ton $day^{-1}$ in Los Angeles-Pasadena region) for non-mobile sources is likely unrealistically high. Our results highlight the important contribution of IVOCs to SOA production in Los Angeles region, but underscore that other uncertainties must be addressed (multigenerational aging, aqueous chemistry, and vapor wall losses) to close the SOA mass balance. This research also highlights the effectiveness of regulations to reduce mobile-source emissions, which have in turn increased the relative importance of other sources, such as volatile chemical products.

### 3.1 Introduction

Exposure to fine particulate matter ($PM_{2.5}$ and $PM_1$) has been associated with

increased mortality, lung cancer and cardiovascular diseases (Apte et al., 2018; Di et al., 2017). Organic aerosol (OA) is a major component of ambient fine particulate matter (Jimenez et al., 2009; Zhang et al., 2015). The majority of OA, even in most urban areas, is secondary organic aerosol (SOA), formed from the atmospheric oxidation of gas-phase species. Over the past several decades, primary emissions have been greatly reduced in the United States, which has led to significant improvement in air quality, especially in the Los Angeles basin in California (Warneke et al., 2012; Zhang et al., 2018). However, SOA remains an important component of fine particulate matter, but its sources are uncertain (Ensberg et al., 2014; McDonald et al., 2018).

Intermediate volatility organic compounds (IVOCs) are an important class of SOA precursors (Chan et al., 2009; Liggio et al., 2016; Presto et al., 2009; Zhao et al., 2014). IVOCs, for example, $C_{12}$ to $C_{17}$ *n*-alkanes and polycyclic aromatic hydrocarbons, are efficient SOA precursors (Chan et al., 2009; Presto et al., 2010a). In addition, chamber experiments using unburnt fuel and diluted exhaust have demonstrated the importance of IVOCs to SOA production from mobile-source emissions (Gordon et al., 2014; Jathar et al., 2013; Miracolo et al., 2011; Platt et al., 2017).

Despite this evidence, IVOCs are not routinely or consistently accounted for in chemical transport models. A major challenge has been the lack of emissions data due to a combination of sampling challenges and the fact that the vast majority of IVOC emissions have not been speciated on a molecular basis. In addition, chemical mechanisms (e.g. SAPRC, Carbon Bond, etc.) often do not explicitly account for IVOCs, instead lumping them with VOCs or non-reactive gases (Lu et al. 2018). Several recent studies report total (speciated and unspeciated) IVOC emissions from a variety of mobile sources, including on- and off-road gasoline, diesel, aircraft and vessel engines (Cross et al., 2013; Huang et al.,

2018; Kroll et al., 2014; Pereira et al., 2018; Presto et al., 2011; Qi et al., 2019; Wang et al., 2012; Zhao et al., 2015, 2016). While these studies have not been able speciate all of the IVOCs emissions at the molecular level, some provide insight into the molecular structure of the unspeciated IVOCs (Drozd et al., 2019; Hatch et al., 2017; Hunter et al., 2017; Worton et al., 2014; Zhao et al., 2015, 2016). For example, IVOCs in diesel exhaust are primarily comprised of aliphatic compounds while IVOCs in gasoline exhaust are primarily aromatics with higher OH reaction rates and SOA yields. Zhao et al. (2015, 2017) used these new emissions data to explain the SOA formation in smog chamber experiments with diluted vehicle emissions. The SOA mechanism proposed by Zhao et al. (2015, 2017) accounts for all of the IVOC emissions. It represents them using 79 different "compounds", some of which are individual species and others are lumped groups assigned based on gas-chromatography and mass spectrometry data. However, this model is too computationally expensive for implementation in current operational CTM.

Because of the high levels of both ozone and PM exposure in the Los Angeles basin over the last several decades, extensive ambient measurement campaigns have explored the sources of poor air quality in the region, including the CalNex campaign in 2010 (Ryerson et al., 2013). During the CalNex campaign, average OA at the Pasadena supersite was 7 μg $m^{-3}$, of which SOA, defined as the sum of semi-volatile and low-volatile oxygenated OA (SV-OOA and LV-OOA) factors from AMS analysis, contributed 66% to total OA mass (Hayes et al., 2013). Zhao et al. (2014) measured the ambient IVOC concentration at the Pasadena site, and estimated that photo-oxidation of IVOCs contributed up to 57% of SV-OOA during CalNex.

A number of chemical transport model (CTM) studies have examined SOA formation in the LA basin during the CalNex campaign (Baker et al., 2015; Fast et al., 2014; Jathar et

al., 2017; Murphy et al., 2017; Woody et al., 2016). However, these studies used very different assumption for IVOC emissions and their SOA yields. IVOC emissions are commonly estimated by applying a scaling factor to some other species (generally POA). These scaling factors have been based on little experimental data and the typically the same factor is applied to all sources. For example, Fast et al. (2014) assumed additional SOA precursor (IVOC and/or SVOC; semivolatile organic compounds) mass of 6.5 × POA and Woody et al. (2016) assumed 7.5 × POA based on previous estimations (Hodzic et al., 2010; Koo et al., 2014), applied to all emission source categories. Jathar et al. (2017) assumed mobile IVOC emission as 25% of diesel NMOG emissions and 20% of gasoline NMOG emissions. Finally, Baker et al. (2015) did not explicitly account for IVOCs, but increased the SOA yields from VOCs by a factor of four compared to the base version of the Community Multiscale Air Quality (CMAQ) model. Despite these efforts, these studies still underpredicted the measured OA by a factor of 2 to 6 (Hayes et al., 2013). Murphy et al. (2017) largely closed the OA mass balance by defining a new lumped SOA precursor called potential combustion volatile organic compounds (pcVOC) with emissions equal to 9.6 × POA and an SOA yield of 1. However, all the above-mentioned models used scaling factors that are not based on actual emission data. They also only use a single IVOC surrogate, which does not account for differences in IVOC chemical composition. Lu et al. (2018) showed that a single scaling factor does not represent the magnitude of actual IVOC emissions across all mobile sources. Finally, none of these models account for the effects of differences in molecular structure in IVOC emissions on SOA yield.

Mobile sources are major sources of NMOG emissions, and therefore important sources of SOA precursors in urban environments (Gentner et al., 2017). Historically, mobile sources have been the dominant source of NMOG in many urban areas, but their contribution

has been reduced due to increasingly stringent emission regulations. The 2014 EPA National Emission Inventory (NEI) estimates that mobile sources contribute 32% of the anthropogenic VOC emissions nationally (and 43% in Los Angeles county). In Los Angeles county, on- and off-road gasoline and diesel sources account for more than 96% of mobile-source emissions.

Lu et al. (2018) recently compiled mobile-source emission data, including on- and off-road gasoline, aircraft and diesel engines, to create updated model-ready emission profiles that include explicit treatment of IVOCs. They found that mobile source NMOG emissions can be explained by trimodal distributions of by-product, fuel and oil modes. IVOC emissions originate from fuel components and similar distributions are observed across sources that use the same fuel (Cross et al., 2015; Lu et al., 2018; Presto et al., 2011). This applies to both low emitting heavily controlled sources (e.g. LEV-II certified gasoline vehicle) and uncontrolled high emitting sources (e.g. two stroke gasoline off-road sources) (Lu et al., 2018). Therefore, in this work, mobile IVOC emissions are modelled and grouped based on fuel type.

In this paper, we use an updated version of CMAQ v5.3 (US EPA Office of Research and Development, 2019) to investigate the sources and contribution of SVOCs and IVOCs to SOA formation in the Los Angeles region during the CalNex campaign. We updated CMAQ v5.3 with a new set of mobile-source NMOG and SVOC emission profiles that include 6 classes of IVOCs and a new parameterization of SOA formation from IVOC precursors designed for implementation into chemical transport models. The new emission profiles are based on direct measurement of IVOCs from on- and off-road mobile sources (Gordon et al., 2013; Lu et al., 2018; May et al., 2014; Presto et al., 2011; Zhao et al., 2015, 2016). These profiles (100VBS to 103VBS) are now available in SPECATE 5.0 (US EPA,

2019). The new SOA parameterization is derived from a comprehensive parameterization that explains the SOA formation from dilute mobile-source exhaust in smog chamber studies (Zhao et al., 2015, 2017). We evaluate the resulting model, now the most up-to-date representation of mobile-source organic compound emissions, using data collected during the CalNex campaign, including direct measurements of ambient IVOCs. Finally, we explore the potential contribution of non-mobile sources to IVOC and OA concentrations.

### 3.2 Parameterizing SOA formation from mobile-source IVOCs

Mobile sources are comprised of a complex mixture of on- and off-road sources, including gasoline, aircraft and diesel engines. However, they are predominantly gasoline- and diesel-powered, with a small fraction of aircraft emissions. In this work we apply the source profiles of Lu et al. (2018) to estimate the amount and composition of the IVOC emissions for different mobile sources. The IVOCs are normalized to total NMOG emissions, which only includes the organics in volatility range from $C^* = 10^3$ to $10^{11}$ μg $m^{-3}$. Table 3.1 summarizes the IVOC-to-NMOG ratios for different mobile sources. The ratios (and associated emission profiles) vary widely depending on the underlying fuel. For gasoline, aircraft to diesel sources, IVOCs comprise 4.6%, 28.5% and 55.5% of the NMOG emissions, respectively. IVOC emissions from gasoline source include high fractions of aromatics (Drozd et al., 2019; Zhao et al., 2016).

We developed a simplified parameterization to simulate first-generation SOA formation from IVOCs under high-$NO_x$ conditions. By first-generation we mean the amount of SOA that forms within a couple of hours in a smog chamber experiment with dilute exhaust at typical atmospheric oxidant levels. The parameterization is derived from the model of Zhao et al. (2015, 2016), which explicitly accounts for 79 different classes of

IVOCs. The chemistry and transport associated with 79 additional species in the gas and particle phases would be too computationally expensive in a CTM which normally has about 50 or less organic aerosol species. Our aim is to develop a model for IVOC SOA production that can be used in off-the-shelf regulatory and routine chemical transport modelling applications. For other applications, a more-explicit approach with multiple thousands of species may be more powerful for modelling reaction pathways (Ying and Li, 2011). From the IVOC measurement perspective, lumping similar IVOCs together based on their volatility and functionality is also more interpretable and compatible to data provide by most instruments.

**Table 3.1 Mass fractions (g per g-NMOG) of IVOCs in mobile NMOG emission profiles used in CMAQ simulations.**

| **Group** | | **Volatility** | **Source** | | |
|---|---|---|---|---|---|
| | | **(C* at 298K, µg $m^{-3}$)** | Gasoline | Aircraft | Diesel |
| **Non-aromatics** | IVOCP6-ALK | $10^6$ | 0.006 | 0.207 | 0.159 |
| | IVOCP5-ALK | $10^5$ | 0.002 | 0.048 | 0.187 |
| | IVOCP4-ALK | $10^4$ | 0.003 | 0.020 | 0.149 |
| | IVOCP3-ALK | $10^3$ | 0.003 | 0.009 | 0.054 |
| **Aromatics** | IVOCP6-ARO | $10^6$ | 0.025 | n/a | 0.002 |
| | IVOCP5-ARO | $10^5$ | 0.006 | n/a | 0.004 |
| **Total** | | | 0.046 | 0.285 | 0.555 |

The Zhao et al. (2015, 2016) model accounts for 57 individual IVOCs and 22 lumped IVOCs. The 22 lumped IVOCs are comprised of unspeciated IVOCs grouped based on gas chromatography (GC) retention time and an assigned chemical class based on its mass spectra. This model explains the SOA formation from dilute exhaust of gasoline and diesel vehicles measured in chamber experiments (Zhao et al., 2015, 2017). Our simplified SOA parameterization accounts for the key differences in chemical composition of the IVOC

emissions from different mobile sources. This is important because the composition of the IVOC emissions varies by source class (e.g. gasoline versus diesel) and SOA yield depends on both molecular weight (volatility) and chemical structure (aromatics versus alkanes) (Chan et al., 2009; Jathar et al., 2013; Lim and Ziemann, 2005, 2009; Presto et al., 2010a). For example, diesels emit more lower volatility IVOCs than gasoline engines, but diesel IVOC emissions are mainly comprised of aliphatic compounds versus aromatics for gasoline. These differences matter because, for a given chemical class, SOA yields generally increase with increasing molecular weight, which increases the effective SOA yield of diesel exhaust relative to gasoline exhaust. However, for a given carbon number, the SOA yield for hydrocarbon IVOCs generally follows aromatics > cyclic > linear > branched alkanes (Lim and Ziemann, 2009; Tkacik et al., 2012), thus gasoline IVOC yields increase when their structure is considered. Finally, aromatic IVOCs have higher OH reaction rates than alkanes (Chan et al., 2009; Zhao et al., 2017). In this study, we only account for IVOC-OH reactions because mobile-source IVOCs are mostly alkanes or aromatics, which will react slower with $O_3$. $NO_3$ oxidation can be important in night-time SOA formation (Fry et al., 2014; Hoyle et al., 2011), and these will be important to consider in the future, but experimental studies on SOA formation from anthropogenic IVOC reactions with $NO_3$ radical are limited at this time. To illustrate the complexity of the IVOC mechanisms of Zhao et al. (2015, 2016), Figure 3.1 plots the SOA yield (expressed as SOA mass divided by mass of precursor) as a function of volatility for the 79 different IVOCs in the model at a typical atmospheric OA concentration of 10 $\mu g\ m^{-3}$. This model likely provides a conservative estimate for SOA yields of lower volatility IVOCs, as $C_{18\text{-}22}$ *n*-alkanes are assumed to have the same SOA yields as $C_{17}$ *n*-alkanes. The scatter in the data highlights the complex relationship between molecular structure and SOA yield.

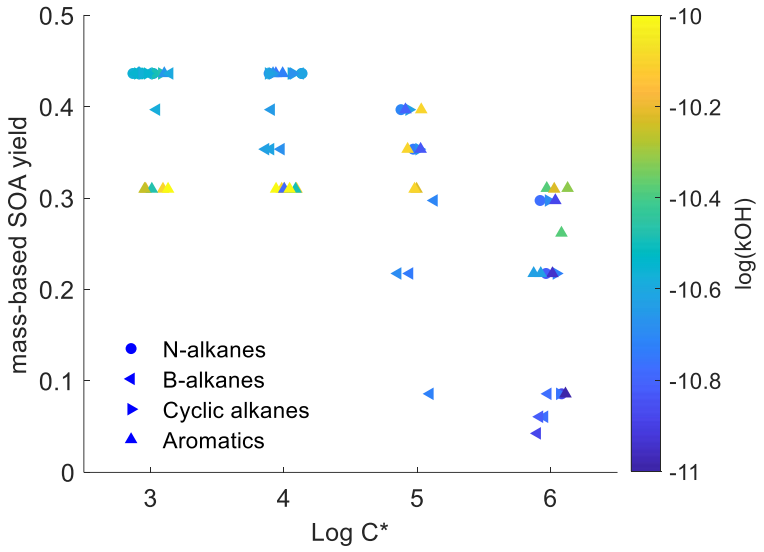



**Figure 3.1 Scatter plot of first-generation mass-based SOA yields versus volatility (log C*, μg / m$^3$) in the detailed parameterization (dots are colored by OH reaction rates)**

Our goal is to derive a semi-empirical SOA parameterization with the minimum number of surrogate species that reproduces the mechanism of Zhao et al. (2015, 2016). The simplified parameterization must account for the differences in SOA formation from IVOC emissions from different mobile source categories (gasoline, diesel and aircraft). We developed the simplified parameterization using the volatility basis set (VBS) framework of Donahue et al. (2006) following the approach of Presto et al. (2010b). The parameterization accounts for all IVOC emissions, which are lumped into surrogates based on gas-chromatography-retention-time (related to volatility) and mass spectral (composition information) data (Lu et al. 2018). Like the work of Zhao et al. (2015, 2016), the parameterization accounts for all IVOC mass, not just the mass that can be speciated at the molecular level (Lu et al. 2018). Briefly, to simulate SOA formation, each lumped IVOC group reacts with OH to form a set of semi-volatile products in Eq. (1):

$$IVOC_i + OH \rightarrow \alpha_{i,1}P_{C^*=0.1} + \alpha_{i,2}P_{C^*=1} + \alpha_{i,3}P_{C^*=10} + \alpha_{i,4}P_{C^*=100} \quad \text{for group } i = 1 \text{ to } 6 \quad (1)$$

where $\alpha_{i,1}$ to $\alpha_{i,4}$ are mass-based stoichiometric coefficients for $IVOC_i$ distributing the

reaction products across a second volatility basis set from 0.1 to 100 ug/$m^3$ (Presto et al., 2010b). For each lumped IVOC species there are 5 unknowns: four stoichiometric coefficients ($\alpha_{i,1}$ to $\alpha_{i,4}$) and the OH reaction rate $k_{OH,\,i}$. These coefficients and reaction rates are derived by fitting the mechanism of Zhao et al. (2015, 2016). All SOA parameters are set at fixed temperature of 298 K. Details of the fitting procedure are in the SI.

We initially tried using four lumped-IVOC-species distributed across the volatility basis set (C* = $10^3$ to $10^6$ ug/$m^3$) to account for the influence of precursor volatility based on gas-chromatography retention time but not molecular structure on SOA yield. However, that model poorly reproduced the SOA formation from gasoline vehicle emissions, especially at shorter time scales (Figure S1). The problem is that IVOCs in diesel exhaust are dominated by aliphatic compounds while IVOCs in gasoline exhaust are dominated by aromatics (Drozd et al., 2019; Zhao et al., 2016); as previously discussed, aromatics compounds have different OH reaction rates and SOA yields (Figure 3.1) (Lim and Ziemann, 2009; Tkacik et al., 2012).

We therefore defined two additional lumped IVOC species with C* = $10^5$ and $10^6$ µg $m^{-3}$ to account for the aromatic IVOCs in gasoline engine exhaust (Table 3.1). The IVOCs in these two bins were split based on mass spectral data (Zhao et al., 2015, 2016). Mobile-source IVOC emissions in the lower volatility bins of C* = $10^3$ and $10^4$ µg $m^{-3}$ are primarily alkanes from unburnt fuel or lubricant oil (Lu et al., 2018; Worton et al., 2014); therefore, the simplified mechanism only includes one lumped aliphatic IVOC species in each of those bins. IVOC emissions are assigned to these surrogate species using the source profiles listed in Table 3.1.

To illustrate the performance of the new parameterization, Figure 3.2 (a) compares the predicted SOA using our six-IVOC-group parameterization to the original mechanism of

Zhao et al. (2015, 2016). It shows that the two models agree with an absolute error for the mass-based SOA yield of less than 0.01 for all mobile sources at an OA concentration of 5 μg $m^{-3}$. Across a wide range of atmospherically relevant concentrations (OA of 1 to 50 μg $m^{-3}$), Figure 3.2 (b) shows that the relative error is less than 6% between our new parameterization and the original mechanisms of Zhao et al. (2015, 2016).

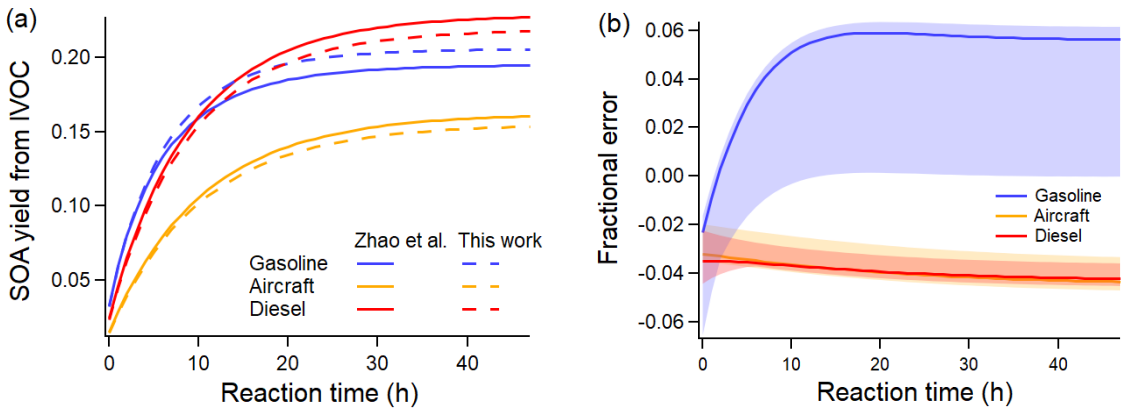



**Figure 3.2 (a) Comparison of predicted SOA formation per unit mass mobile IVOC emission of new parameterizations and model of Zhao et al. (2015, 2016) at OA = 5 μg $m^{-3}$, average [OH] = 3 × $10^6$ $cm^{-3}$. (b) Relative error in SOA formed between new and Zhao et al. (2015, 2016) parameterization (Solid line is the relative error at OA = 5 μg $m^{-3}$, shaded area corresponds to OA from 1 to 50 μg $m^{-3}$).**

The yields derived by the fitting make physical sense. The yields increase with decreasing volatility (Table 3.2). The fitting procedure assigns higher yields and faster reaction to the lumped aromatics compared to aliphatics in the same volatility bin (Drozd et al., 2019; Zhao et al., 2016). This explains the higher SOA production in the first 10 hours from gasoline exhaust compared to aircraft and diesel IVOC emissions. It also predicts that diesel IVOC emissions have the overall highest SOA yield due to their high fraction of lower volatility compounds compared to emission from gasoline engines and aircraft (Lu et al., 2018; Zhao et al., 2015).

**Table 3.2 Properties and stoichiometric mass-based product yields for 6-group IVOC-SOA parameterization**

| Group | C* (μg m$^{-3}$, at 298K) | MW (g/mol) | $k_{OH} \times 10^{11}$ (cm$^{-1}$ molec$^{-1}$ s$^{-1}$) | $\alpha_i$ (C* = 0.1 to 100 μg m$^{-3}$ at 298K) | | | | Yield at 10 μg m$^{-3}$ | $H_{vap}$ (kJ mol$^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | 0.1 | 1 | 10 | 100 | | |
| **IVOCP6-ALK** | $10^6$ | 184.4 | 1.55 | 0.009 | 0.045 | 0.118 | 0.470 | 0.15 | 19 |
| **IVOCP5-ALK** | $10^5$ | 219.4 | 1.89 | 0.051 | 0.061 | 0.394 | 0.494 | 0.35 | 30 |
| **IVOCP4-ALK** | $10^4$ | 254.9 | 2.25 | 0.068 | 0.083 | 0.523 | 0.239 | 0.43 | 41 |
| **IVOCP3-ALK** | $10^3$ | 296.6 | 2.65 | 0.067 | 0.086 | 0.544 | 0.198 | 0.43 | 52 |
| **IVOCP6-ARO** | $10^6$ | 162.3 | 3.05 | 0.022 | 0.109 | 0.251 | 0.005 | 0.25 | 19 |
| **IVOCP5-ARO** | $10^5$ | 197.3 | 7.56 | 0.143 | 0.021 | 0.329 | 0.358 | 0.36 | 30 |

Table 3.2 lists the set of $k_{OH}$ and $\alpha_i$ for the simplified six-IVOC-group parameterization for mobile-source emissions. Molecular weights (MW) are determined as the average MW of *n*-alkanes or speciated aromatics in each volatility bin. The IVOC MWs are used to convert mass-based SOA yields to molar units and calculate parameters needed to simulate dry deposition processes. Enthalpies of vaporization ($H_{vap}$) are determined using the fitted parameterization in Ranjan et al. (2012). In this work, we implement this six-lumped-IVOC-group parameterization to model the IVOC SOA formation in CMAQ v5.3. The first-generation products represented in equation (1) undergo multigenerational aging following the mechanism of Murphy et al. (2017) described in section 3.3.4.

## 3.3 CMAQ model

To evaluate the contribution of mobile source IVOC emissions to ambient SOA, we implemented our new six-lumped-IVOC-group SOA parameterization and emissions profiles into CMAQ v5.3. We used the model to simulate the air quality in California from 1 May to 30 June 2010, which includes the entire CalNex campaign (May and July 2010). Except as noted below, the simulations described here have essentially the same modelling domain and input parameters as previous modelling studies on CalNex (Baker et al., 2015; Murphy et al., 2017; Woody et al., 2016). We have extended this previous work by updating the emissions and SOA formation from IVOCs.

### 3.3.1 Model configuration

The model domain covered California and Nevada with a 4 km (325×225) grid resolution and 35 vertical layers. The input meteorology and NEI emission inventory are very similar to those used by Baker et al. (2015), Woody et al. (2016) and Murphy et al. (2017) and are identical to Qin et al. (2019). Meteorological inputs were generated using the Weather Research and Forecasting Model (WRF) Advanced Research WRF core version 3.8.1 (Skamarock et al., 2008) with one additional model layer at the surface compared to previous studies (i.e. the lowest layer of approximately 40 m depth has been split into two 20 m deep layers to better resolve surface gradients). The emissions inputs are based on the 2011 NEI version 2 with mobile, wildfire, and electric generating point source emissions calculated for 2010. Mobile on-road and non-road emissions are calculated by MOVES 2014a, except that on-road emissions for California are estimated by EMFAC and allocated using MOVES 2014a. Biogenic emissions are calculated online with BEIS v3.61 and improved land use cover from BELD4 (Bash et al., 2016). Sea-spray aerosols are calculated online and incorporate dynamic prediction of particle population size and standard deviation.

Wind-blown dust emissions are neglected and should not impact comparisons with the data collected by the AMS, which detects non-refractory particulate compounds. Moreover, previous studies (Cazorla et al., 2013) found little evidence of dust impacts during CalNex using both in-situ aircraft measurements and inference from AERONET retrievals. Gas-phase chemistry is simulated with the SAPRC07T chemical mechanism (Carter, 2010; Hutzell et al., 2012; Xie et al., 2013). Aerosols are simulated using the Aero-7 module (CMAQ-AE7) with monoterpene photo-oxidation updates (Xu et al., 2018) and organic water uptake (Pye et al., 2017). Boundary conditions were generated from a 12 km continental U.S. simulation of April to June 2010. We use the first 14 days of the simulation as a spin-up to minimize the influence of initial conditions.

Previous studies (Baker et al., 2015; Woody et al., 2016) have extensively evaluated different versions of CMAQ using CalNex data. These evaluations show good to excellent performance for many pollutants, with a notable exception of organic aerosols and SOA – the focus of this paper. We evaluated our model predictions with measurements of gas-phase pollutants such as CO, $O_3$ and $NO_x$, as they are typical indicators for model performance. Consistent with the previous applications of CMAQ to CalNex (Baker et al., 2015; Murphy et al., 2017; Woody et al., 2016), Figure S2(a) shows very good agreement between modelled and measured CO diurnal patterns in Pasadena, and the normalized mean bias (NMB) is 4.2%. Figure S3 compares the $O_3$, NO and $NO_2$ diurnal patterns with measurements in Pasadena, where the NMB is 10.7%, -6.7% and 5.4%, respectively. Figure S4 compares the CO, $O_3$ and NO diurnal patterns for three other sites: Bakersfield, Sacramento and Cool. The model NMB is within ±25% for all comparisons, except for $O_3$ and NO in Bakersfield. Thus, we can conclude that the CMAQ model perform reasonably well at all four sites for traditional gas-phase pollutants.

### 3.3.2 POA emissions

CMAQ v5.3 treats POA emissions as semi-volatile with variable gas-particle partitioning and multigenerational aging (Figure S5). The POA model, similar to the 1.5-VBS of Koo et al. (2014), contains five pairs of hydrocarbon-like vapor/particle species (1 LVOC, 3 SVOCs, and 1 IVOC) distributed across a volatility basis set with C* from $10^{-1}$ to $10^{3}$ μg $m^{-3}$, with O:C increasing slightly with decreasing volatility. POA emissions are then assigned to each of these species using the source-specific volatility profiles in Table 3.3 and CMAQ calculates gas-particle partitioning assuming equilibrium partitioning and treating the entire organic phase as a single, pseudo-ideal solution. For non-mobile sources, POA emissions are distributed into all five bins with C* from $10^{-1}$ to $10^{3}$ μg $m^{-3}$ while the mobile source POA profiles only map to the $10^{-1}$ to $10^{2}$ μg $m^{-3}$bins.

**Table 3.3 POA volatility distributions and filter artifact scaling factors**

| Source | Volatility, C* (μg $m^{-3}$, at 298K) | | | | Filter artifact scaling factor |
|---|---|---|---|---|---|
| | ≤ $10^{-1}$ | 1 | 10 | $10^{2}$ | |
| **Gasoline** | 0.16 | 0.08 | 0.37 | 0.39 | 1.4 |
| **Diesel** | 0.21 | 0.11 | 0.33 | 0.36 | 1 |
| **Gas-turbine** | 0.15 | 0.26 | 0.38 | 0.21 | 1 |

Comprehensive emissions profiles for semi-volatile POA include both SVOCs and lower volatility organics (Lu et al., 2018). In the base version of CMAQ v5.3, the volatility profile of Robinson et al. (2007) is used to represent all combustion sources. Here, we update the volatility distributions for mobile POA using the new mobile-source emission profiles in Lu et al. (2018). The profiles (8873VBS, 8992VBS to 8996VBS) are available in SPECIATE 5.0 (US EPA, 2019). For non-mobile combustion sources, we use the biomass burning POA

volatility distribution from May et al. (2013b) for wood-burning sources, the cooking POA volatility distribution from Woody et al. (2016) for cooking sources, and the diesel POA volatility distribution from May et al. (2013a) as a surrogate for all other combustion sources. According to our emission inventory, mobile, wood-burning and cooking sources combined emit more than 80% of total POA in LA region during the modelled period, where other combustion sources only emit 16.4% of the POA. We acknowledge that the diesel POA surrogate is modestly more volatile than biomass burning POA profiles. Thus, using diesel POA volatility as the surrogate for other combustion sources will possibly increase the regional SOA formation compared to if a different profile was used, but the potential bias is small. Table 3.3 summarizes the volatility distributions and scaling factors used in this work. The same POA emissions were used for all model runs.

A challenge is that most existing POA emission factors used to inform inventories such as NEI are based on filter measurements, which do not quantitatively collect all SVOCs. For example, filters collect only a portion of SVOC vapors. Estimating this error is complex because there are competing biases. First, source testing is often performed at low levels of dilution which creates high concentrations (relative to the more dilute atmosphere) that shifts gas-particle partitioning of SVOCs to the particle phase. In these situations, filters collect a larger fraction of SVOCs than more dilute conditions (of course, at high enough concentrations, filters will also collect some IVOC vapors). Second, during mobile source testing, filters are commonly collected at elevated temperatures (e.g. 47 ºC) to avoid water condensation, which shifts gas-particle partitioning towards the gas phase, reducing the fraction of SVOCs collected by a filter. Finally, filters collect some vapors as sampling artifacts, which depends on many factors, including filter material, filter face velocity, and filter pre-treatment (Subramanian et al., 2004). Therefore, the fraction of SVOCs collected

by filters depends on these competing effects, which are difficult to quantify. As expected, data from Zhao et al. (2015, 2016) and Lu et al. (2018) indicate that the fraction of SVOC collected depends on the OA concentration inside the sampling system.

To estimate potential biases in the amount of SVOC vapors in the filter-based POA emission factor measurements, we compared the mass of lower volatility organics (SVOC + LVOC + NV) collected on filters and Tenax tubes versus the mass collected on filters (regular POA measurement) (Lu et al., 2018). The two estimates for diesel and gas-turbine tests were within 10%, which is within experimental uncertainty. Therefore, we did not add any SVOC mass to these emissions. For gasoline sources, the data indicate an average bias of 40%, which means that lower volatility organics were only partially collected by the filter. This is consistent with the relatively low particle emissions of gasoline sources, which create lower concentration conditions inside of the dilution sampler and therefore gas-particle partitioning shifted more to the vapor phase. We therefore applied a filter artifact correction factor of 1.4 to gasoline POA emissions, as shown in Table 3.3. We add these SVOC vapors to address the bias in emissions measurements and to best estimate the potential local / regional SOA formation from mobile source SVOCs.

### 3.3.3 IVOC emissions

An important difference from previous implementations of CMAQ to simulate the CalNex campaign (Baker et al., 2015; Murphy et al., 2017; Woody et al., 2016) is the new mobile IVOC emission data and the application of the new six-lumped-IVOC-species SOA parameterization. Mobile sources contribute more than 40% of anthropogenic NMOG emissions in the South Coast Air Basin in the CalNex emission inventory (Baker et al., 2015). Given the consistency of the speciation and IVOC-to-NMOG ratio for sources using

same type of fuel (Lu et al., 2018), we assign mobile-source emissions profiles based on fuel type (gasoline, diesel, or jet fuel). NMOG emissions from all on- and off-road gasoline sources are represented using the same average gasoline exhaust profile (SPECIATE Profile #100VBS). NMOG emissions from all on-road, off-road diesel sources (including rail) are represented using the same average non-DPF diesel exhaust profile (SPECIATE Profile #103VBS). Studies have noted there can be significant differences in IVOC emissions between DPF-equipped and non-DPF vehicles (Dunmore et al., 2015; Lu et al., 2018; Platt et al., 2017). However, the total NMOG emission from diesel sources in southern California in 2010 were dominated (> 99%) by non-DPF vehicles (due to a combination of the fleet composition and the fact that non-DPF vehicles have much lower emission factors). Therefore, we use the IVOC emission profile for non-DPF vehicle for all diesel sources. Although only limited data are available for off-road diesel engine emissions (Qi et al., 2019), it suggests the emissions are similar to on-road diesel vehicles. NMOG emissions for all jet-fuelled sources are represented using the same gas-turbine exhaust profile (SPECIATE Profile #102VBS). The IVOC components of these profiles are summarized in Table 3.1 and complete profiles are given in SPECIATE 5.0 (US EPA, 2019). Total IVOC emissions are determined using the IVOC-to-NMOG ratios, which are more consistent across source types than IVOC-to-POA ratios (Lu et al., 2018).

For this work, IVOC emissions are added to existing NMOG emissions. This was done to keep the VOC emissions across the different models runs constant in order to better isolate the contribution of IVOCs to SOA. In addition, OH oxidation of IVOCs is assumed to regenerate OH radicals and thus have minimal impact on the oxidant budget and the production of $O_3$. However, Lu et al. (2018) argued that existing NMOG inventories largely include IVOCs, just that they are misattributed to VOCs. Therefore, future work should

proportionally reduce the VOC emissions to keep the overall NMOG emissions (VOC + IVOC) constant. This assumption minimally effects the OA model evaluation, because the base version of CMAQ predicts that traditional VOCs only contribute 7% of measured OA at Pasadena during the CalNex campaign (Baker et al., 2015).

SOA is produced from IVOC oxidation using the parameterization described in section 3.2. The SOA mass is determined by CMAQ based on the gas-particle partitioning of the SVOC products created from IVOC oxidation. CMAQ v5.3 calculates partitioning assuming thermodynamic equilibrium and that all organics form a single pseudo-ideal solution. The SVOC products also undergo multigenerational aging following the approach of Murphy et al. (2017) (see section 3.3.4).

### 3.3.4 Multi-generational aging and gas-particle partitioning

The semivolatile POA emissions and semivolatile products formed from oxidation of SOA precursors undergo multigenerational aging as described in Murphy et al. (2017). Figure S5 shows the schematic diagram for modelling OH oxidation first-generation and multigenerational aging. Briefly, the approach simulates the reaction of L/S/IVOC vapors with hydroxyl radical and distributes the product mass to a second set of five vapor-particle pairs of species at moderate O:C values. The stoichiometric ratios used to distribute the product mass were derived to match the SOA enhancement predicted by a full 2D-VBS simulation of the functionalization and fragmentation of SVOCs during three days of atmospheric oxidation. This model, unlike that of Koo et al. (2014), does transfer some of the aged products to higher volatility bins, and thus reduces SOA over multiple generations of OH reaction. The probability for fragmentation increases as a function of O:C in agreement with theory (Donahue et al., 2011). Although the competing effects of

fragmentation and functionalization at long timescales are represented in this model, the simplified framework is likely limited when trying to capture the full complexity of multigenerational aging. For this work, no changes were made to the chemical properties (e.g. carbon number, O:C, etc.) or reaction stoichiometry of the multigenerational aging mechanism of Murphy et al. (2017). Because IVOC products likely have lower carbon numbers than products of primary SVOC oxidation, our approach may represent an upper bound on the potential for IVOC SOA aging to further enhance particle mass downwind of sources.

### 3.3.5 Simulation cases

To systematically explore the effects of adding IVOC emissions from mobile and non-mobile sectors, we performed four simulation cases, summarized in Table 3.4. All cases use the same emission inputs as described earlier with differences in IVOC emissions. In the base case (Case 1), mobile SOA is only formed through the oxidation of traditional VOC emissions and SVOCs from evaporated semivolatile POA.

**Table 3.4 Total anthropogenic organic emissions (Ton day$^{-1}$) in Los Angeles Basin region in four CMAQ simulation cases**

| Case | Name | Inventory POA | POA after scaling | Inventory NMOG | Mobile IVOCs | Non-mobile IVOCs |
|---|---|---|---|---|---|---|
| **1** | Base | 26.4 | 28.9 | 450.2 | 0 | 0 |
| **2** | Mobile IVOC | 26.4 | 28.9 | 450.2 | 27.6 | 0 |
| **3** | Low non-mobile IVOC | 26.4 | 28.9 | 450.2 | 27.6 | 30.7 |
| **4** | High non-mobile IVOC | 26.4 | 28.9 | 450.2 | 27.6 | 68.5 |

Figure 3.3 (a) compares the anthropogenic NMOG emissions in the Los Angeles Basin region for the four simulation cases (geographical boundaries are defined by

simulation grid cells shown in Figure S6). In the base case (Case 1), mobile sources contribute 43% of anthropogenic NMOG emissions, of which gasoline sources contribute 35%, diesel sources 8% and aircraft less than 1%. Non-mobile sources contribute the remainder of the anthropogenic NMOG emissions (57%), of which VCP usage contributes 39%, followed by 17% from other sources. The emission inventory contains minimal cooking and biomass burning NMOG emissions during CalNex (1.5%).

Cases 2 to 4 incrementally add mobile IVOC emissions to the model. Table 3.4 shows that Case 2 adds on average 27.6 Ton $day^{-1}$ mobile source IVOC emissions, which is our best estimate of the mobile source IVOC emission based on the compilation of measurement data and source profiles in Lu et al. (2018) as described in section 3.3.3. The difference in SOA concentrations between Case 2 and Case 1 is the SOA contribution from mobile emitted IVOCs. In Case 3 and 4, we incrementally add IVOC emissions from non-mobile sources to the inventory to explore the contribution of non-mobile sources of IVOCs as discussed in section 3.4.2.

### 3.4 CMAQ simulation results

To evaluate model performance, we compared predictions to measured data from the CalNex campaign at Pasadena, CA, as well as the organic carbon (OC) measured at Chemical Speciation Network (CSN) sites in California. The CalNex campaign characterized atmospheric composition at two sites in southern California, Pasadena, and Bakersfield, from 15 May to 29 June 2010 (Ryerson et al., 2013). We focus on the Pasadena site, which is located 18 km northeast and generally downwind of downtown Los Angeles, because there were direct measurements of IVOCs (Zhao et al., 2014). We also evaluate model predictions at the Pasadena site for OA, BC, CO, select speciated VOCs and Planetary

Boundary Layer (PBL) height.

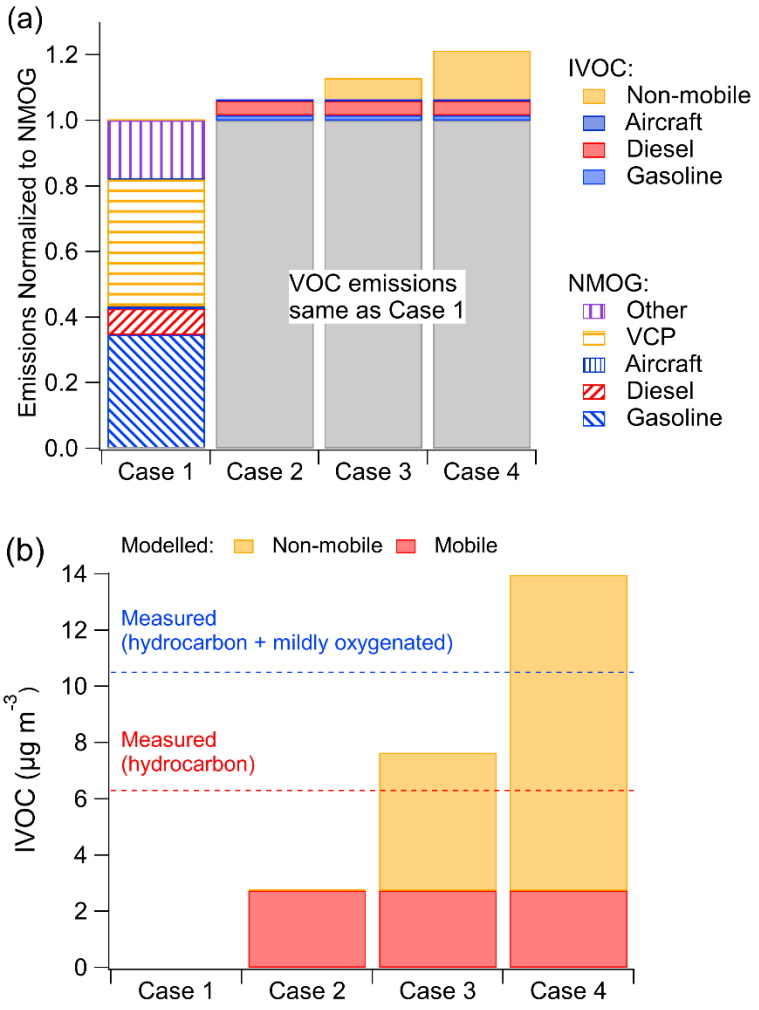



**Figure 3.3 (a) Modelled NMOG and IVOC emissions by source for the four simulation cases. (b) Measured and modelled IVOC mass concentrations at Pasadena, CA during CalNex for the four simulation cases. Measured data in (b) from Zhao et al. (2014).**

## 3.4.1 Base case and mobile IVOC case

### 3.4.1.1 IVOC mass concentrations

Figure 3.3 (b) compares the model-predicted and measured campaign-average IVOC mass concentration at the Pasadena site. Zhao et al. (2014) reported data for two classes of IVOCs differentiated based on mass spectral signature: hydrocarbon IVOCs and mildly oxygenated IVOCs. Zhao et al. (2014) attributes hydrocarbon IVOCs to primary emissions;

the mildly oxygenated IVOC could either be primary emissions or formed via atmospheric oxidation. The CalNex campaign-averaged measured hydrocarbon IVOCs at the Pasadena site was 6.3 μg $m^{-3}$; the measured mildly oxygenated IVOC concentration was 4.2 μg $m^{-3}$. The analytical techniques of Zhao et al. (2014) are not optimized for measuring oxygenated organics; therefore, their data provide a lower bound estimate of the total and oxygenated IVOCs.

The base case (Case 1) predicts essentially no IVOC concentrations as they are not explicitly included in the base inventory or model (though could be implicitly included as misclassified VOC species). Case 2 (mobile IVOC case) predicts 2.7 μg $m^{-3}$ of IVOCs at the Pasadena site, which corresponds to 43% of measured hydrocarbon IVOCs. This indicates that mobile sources are an important source of IVOCs in the LA region, but that more than half of the hydrocarbon IVOCs measured in Pasadena are likely emitted by non-mobile sources. In addition to hydrocarbon IVOCs, Zhao et al. (2014) measured 4.2 μg $m^{-3}$ of mildly oxygenated IVOCs, which are also not explained by mobile-source emissions.

While the comparison in Figure 3.3 (b) suggests that non-mobile sources may be important contributors to ambient IVOC concentrations, there are a number of potential uncertainties, including (1) uncertainty in mobile source activity, (2) uncertainty in mobile source NMOG emission factors, and (3) uncertainty in mobile source IVOC-to-NMOG emission ratios. The first potential uncertainty is mobile-source activities. BC and CO are commonly used as indicators of gasoline and diesel sources activity. The mobile-source CO emission inventory used here (EMFAC) agrees with another fuel-based CO inventory (Kim et al., 2016), both of which reproduce the observed weekly patterns. This suggests the mobile-source CO emission inventory in LA basin during CalNex is correctly modelled. While the model performs well for CO (Figure S2), it overestimates BC concentrations by a

factor of 2. These comparisons suggest that gasoline activity (the major of source of CO) is modelled correctly, but there may be a potential over-estimation of either diesel activity and/or the diesel BC emission factor (the major source of BC). If the diesel activity is overestimated, then diesel IVOC are likely overestimated, which only strengthens our conclusion that there are important non-mobile sources of IVOCs.

The second potential uncertainty is mobile-source NMOG emission factors. Comparisons in May et al. (2014) suggest that the EMFAC emission factors (which are used to create the mobile-source emission inventory for these simulations) are robust, except for LEV-2 vehicles. During the 2010 CalNex period, EMFAC estimates LEV-2 vehicles (considering model year after 2004) only emit 8.5% of total gasoline NMOG emissions in California and therefore are not major contributors in mobile emissions. Therefore this uncertainty also does not appear to alter our conclusion that there are important non-mobile sources of IVOCs.

The final potential uncertainty is the IVOC-to-NMOG ratios. Zhao et al. (2016) and Lu et al. (2018) show that IVOC-to-NMOG ratios of cold-start UC (unified cycle) emissions from gasoline sources are consistent across a large number of vehicles spanning a range of emission certification standards. Although IVOC emissions from hot-running gasoline vehicle exhaust are enriched by as much as a factor of 4 compared to the cold-start UC cycle (Lu et al., 2018; Zhao et al., 2016), EMFAC2017 estimates that running exhaust only contributes 34% of total gasoline summertime NMOG emissions in CA in 2010. A simple weighted average of 66% emission using cold-start UC emission profile and 34% of emission using hot-running emission profile increases the IVOC-to-NMOG fraction for gasoline vehicles by a factor of 2, from 4.5% to 9.1%. The IVOC-to-NMOG ratio for diesel sources is already high (55%) and thus it cannot be increased as much as the gasoline emissions (less

than a factor of 2). Therefore, the largest uncertainty in modelled mobile IVOCs is the gasoline source IVOC-to-NMOG ratio, which could be underestimated by as much as a factor of two. This means that the overall uncertainty in modelled mobile IVOC emissions is less than a factor of 2. Increasing the gasoline IVOC emissions to better account for hot-running operations would explain a larger fraction of the measured hydrocarbon IVOCs concentrations, but it seems unlikely that it would close the mass balance given that gasoline vehicles contribute less than half of the mobile IVOCs. Therefore, even acknowledging the existing uncertainty we still conclude that non-mobile sources are likely important contributors to ambient IVOC concentrations in Pasadena.

Jathar et al. (2017) also updated CMAQ with mobile-source IVOC emissions estimates. They assumed that IVOCs contribute 25% and 20% of the NMOG emissions from gasoline and diesel source, respectively. However, these ratios are not based on direct measurements, but instead inferred from SOA closure studies for chamber experiments. The model of Jathar et al. (2017) predicted mobile sources contribute 3.9 $\mu g\ m^{-3}$ of IVOCs, which is about factor of 1.5 higher than the IVOC concentrations predicted here (and about 65% of measured ambient hydrocarbon IVOC concentrations). The better closure is due the very high IVOC-to-NMOG ratio assumed for gasoline vehicles, which is not supported by direct measurements (Drozd et al., 2019; Zhao et al., 2016).

#### 3.4.1.2 Primary VOC/IVOC diurnal patterns

Figure 3.4 compares the measured and modelled campaign-average diurnal patterns of important anthropogenic VOCs (benzene, toluene, m-/p-/o- xylenes) and hydrocarbon IVOCs. Measured concentrations of benzene, toluene and hydrocarbon IVOCs are highest in the early afternoon (12pm - 2pm, in Figure 3.4 a, b and d). This has been attributed to the

transport of morning emissions from downtown Los Angeles to Pasadena (Borbon et al., 2013). Measured xylene concentrations show a slight decrease in daytime, which is attributed to their relatively high OH reaction rate and thus faster oxidation during the daytime (de Gouw et al., 2018).

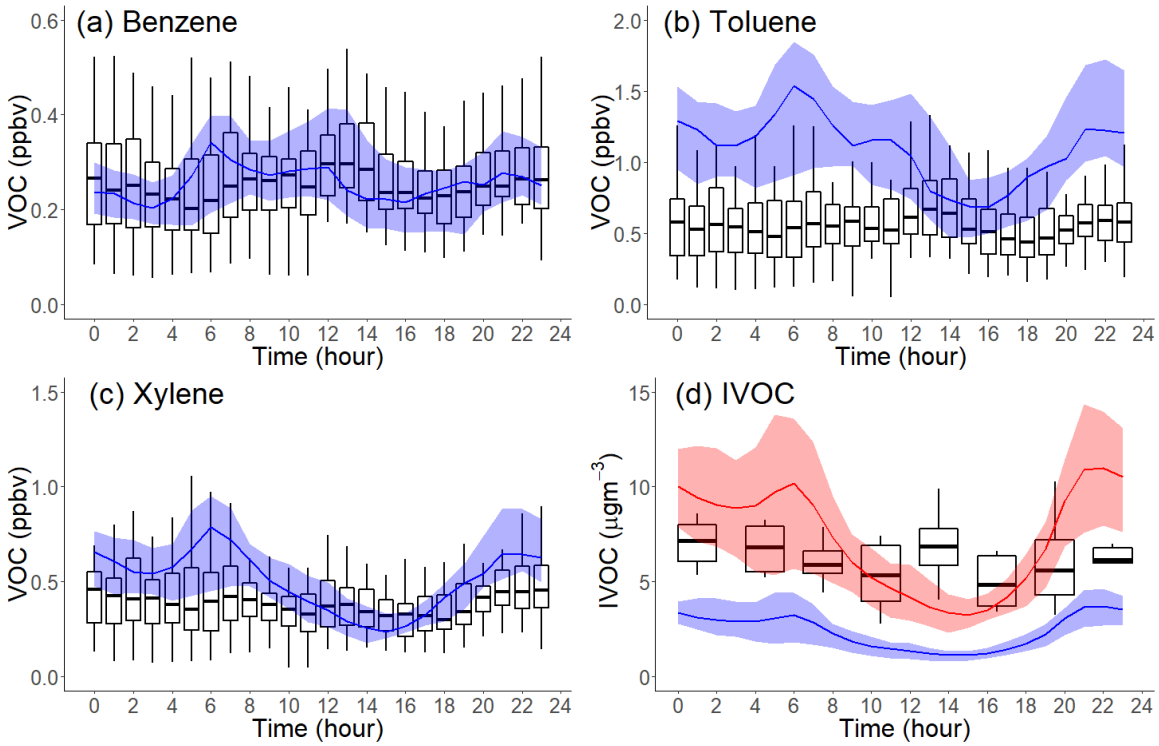



**Figure 3.4 Comparison of measured (boxplot, solid box denotes 25$^{th}$ to 75$^{th}$ percentiles and whiskers denote 10$^{th}$ to 90$^{th}$ percentiles) and modelled (line, shaded area denotes 25$^{th}$ to 75$^{th}$ percentiles) diurnal patterns in Pasadena, CA during CalNex for species: (a) benzene, $k_{OH}$ = 1.22 × 10 $^{-12}$ cm$^3$ molec$^{-1}$ s$^{-1}$ (b) toluene, $k_{OH}$ = 5.63 × 10 $^{-12}$ cm$^3$ molec$^{-1}$ s$^{-1}$ (c) xylene, $k_{OH}$ = 1.36 - 1.87 × 10 $^{-11}$ cm$^3$ molec$^{-1}$ s$^{-1}$ and (d) hydrocarbon IVOCs (blue: Case 2, red: Case 3), $k_{OH}$ = 1.55 – 7.56 × 10 $^{-11}$ cm$^3$ molec$^{-1}$ s$^{-1}$. Measured data from Borbon et al. (2013).**

Figure 3.4 indicates that the model reproduces the measured benzene diurnal pattern but not the toluene, xylene and hydrocarbon IVOC diurnal patterns. Figure 3.4 (b and c) shows that during night-time the model overpredicts toluene and xylene concentrations by a

factor of 2 and 1.4, respectively. Modelled hydrocarbon IVOCs mass concentration (Case 2) are underestimated throughout the day (Figure 3.4 d and Figure 3.3 b).

Figure 3.4 also shows modelled species concentrations peak around 6 AM and then steadily decrease from 6 AM to 4 PM, in contrast to the early-afternoon peaks (12 PM to 2 PM) in the measured data. A potential explanation for this difference is that the model is incorrectly simulating the PBL height. On average, the measured PBL height ranges from ~200 m at night to ~900 m at noon (Figure S7), while modelled PBL height ranges from ~60 m at night and up to 1500 m at noon. However, changing the predicted PBL height would degrade model performance for some species which are already predicted well (Figure S3 and S4). Another possible explanation is that additional unknown sources of IVOCs have large NMOG emissions that peaks at noon, for example some type of evaporative emissions. Additional research is needed to resolve the discrepancy between model and measured diurnal profiles shown in Figure 3.4.

#### 3.4.1.3 OA mass concentrations and diurnal patterns

Figure 3.5(a) plots the AMS-observed and CMAQ-modelled hourly-averaged $PM_1$-OA time series at the Pasadena site during CalNex. We consider the Pearson's correlation coefficient (r) and root-mean-square error (RMSE) as the evaluation metrics between measured and model OA time series. The definitions of r and RMSE are shown in Eq. (S1) and (S2):

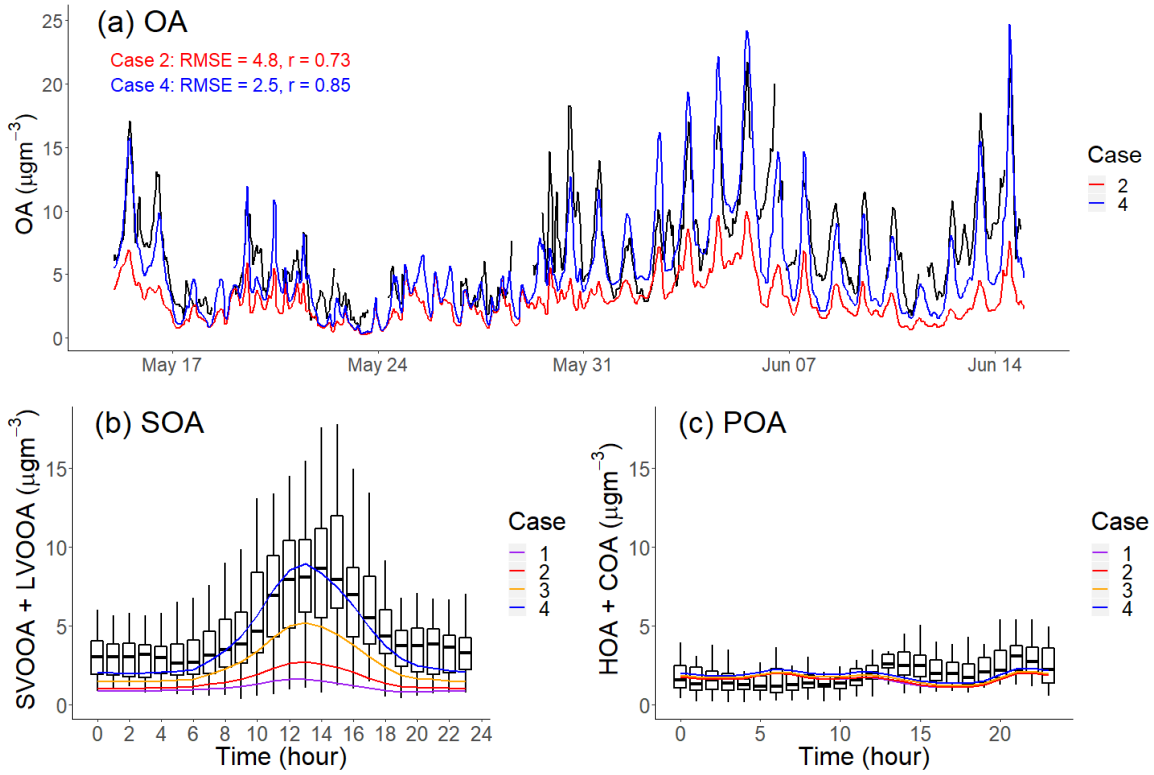



**Figure 3.5 (a) $PM_1$-OA component hourly-averaged time series of measured data and model output in Pasadena, CA during CalNex campaign. (b, c) Diurnal pattern of measured and modelled SOA and POA mass concentration in Pasadena, CA during CalNex. Measured data from Hayes et al. (2013).**

Our base model (Case 1) significantly underpredicts the OA concentration, often by more than a factor of 3, over the entire time period. Case 1 has a large RMSE = 5.3 μg m$^{-3}$, which is comparable to the average measured OA (6.9 μg m$^{3}$), and moderate positive correlation ($r$ = 0.69). To understand the source of this discrepancy, Figure 3.5 (b) and (c) compare the modelled average diurnal patterns for SOA and POA to PMF factors derived from Aerosol Mass Spectrometer data for OOA (SV-OOA plus LV-OOA) and POA (HOA plus COA) (Hayes et al., 2013). The observed OOA factor in Figure 3.5(b) has a strong peak in the early afternoon, similar to the OH radical concentration (de Gouw et al., 2018) and photo-chemical age (Hayes et al., 2015).

Figure 3.5(c) shows that the model correctly predicts average POA concentrations

(modelled: 1.73 μg $m^{-3}$ vs measured: 2.01 μg $m^{-3}$). It also reasonably reproduces the observed POA diurnal pattern. This applies to all four cases and suggests that our inventory (Table 3.3) has a reasonable representation for the POA emissions, volatility distributions, and correction for filter artifacts for gasoline sources. The mobile volatility profile predicts that a bit more than half of the semivolatile POA evaporates; therefore, if it treated POA as non-volatile then the model would have overpredicted the observed POA concentrations by about a factor of two.

Figure 3.5(b) shows that Case 1 produces very little SOA, similar to previous CMAQ simulations (Baker et al., 2015; Woody et al., 2016). In this study, we emphasize the peak in the diurnal SOA concentration because this enhancement is reflective of the strength of prompt SOA formation in both the observations and the model. In Case 1, the predicted peak SOA concentration is 1.65 μg $m^{-3}$ at the Pasadena site, which is 5 times lower than the AMS-observed value (8.63 μg $m^{-3}$). Both modelled LV-OOA and SV-OOA are much lower than AMS-observed factors.

Figure 3.2 indicates that mobile-source IVOC emissions contribute significantly to SOA formation, especially to the day-time SOA formation due to their high SOA yield and OH reaction rates. In Case 2, the addition of mobile IVOC emissions increases the peak SOA concentration by 60%, from 1.65 to 2.75 μg $m^{-3}$ and daytime SOA increase (peak SOA – night-time SOA) by 110% from 0.82 to 1.73 μg $m^{-3}$. The increase in night-time SOA from IVOC oxidation was about a factor of 4 smaller than the daytime increase. Adding mobile-source IVOC improves model performance, but Case 2 still only explained 32% of AMS-observed daytime peak SOA.

Our comparison demonstrates that mobile-source IVOC emissions need to be explicitly included in models and inventories. However, they do not close the mass balance

for hydrocarbon IVOCs or SOA in Pasadena. In the next section, we explore the potential contribution of IVOC emissions from non-mobile sources (McDonald et al., 2018).

### 3.4.2 Non-mobile IVOC emissions

#### 3.4.2.1 IVOC mass concentrations and diurnal pattern

Motivated by recent research on volatile chemical products (VCPs) (Khare and Gentner, 2018; McDonald et al., 2018), we also investigated potential IVOC emission from non-mobile source. For example, McDonald et al. (2018) estimated that 19.6% of total gas-phase VCP emissions are IVOCs. Khare and Gentner (2018) reported that the IVOC content in 12 commercially available VCPs range from 0 to 95%. However, many of these IVOCs in VCPs are heavily oxygenated.

Cases 3 and 4 explore different levels of IVOC emission for non-mobile sources. The IVOC-to-NMOG ratios are not based on independent laboratory data, but are set to close the gap between modelled and measured hydrocarbon IVOC concentration (Case 3) and SOA concentration (Case 4) at Pasadena, CA (Hayes et al., 2013; Zhao et al., 2014). Since there are limited data on non-mobile IVOC emissions, they are assumed to have the same properties as alkane-like IVOCs (IVOCP6-ALK to IVOCP3-ALK) with a uniform volatility distribution. Table 3.4 shows that Case 3 and 4 add an average 30.7 and 68.5 Ton $day^{-1}$ non-mobile IVOC emissions scaled from NMOG emissions as described in section 3.3.4.

For the low non-mobile-IVOC case (Case 3), we added IVOC emissions to the inventory equivalent to 12% non-mobile NMOG emission. The scaling coefficient was determined to roughly match the campaign-average hydrocarbon IVOC mass concentrations measured in Pasadena, CA (Zhao et al., 2014). The only difference between Cases 2 and 3 are the additional non-mobile hydrocarbon IVOC emissions.

For the high non-mobile-IVOC case (Case 4), we added non-mobile IVOC emissions equivalent to 26.8% of non-mobile NMOG emissions. This value was chosen to roughly close the mildly-oxygenated IVOC and SOA mass balance. It is obviously a very high estimate, but only somewhat higher than the 20% estimates of total VCP emissions in McDonald et al. (2018). The only difference between cases (4) and (3) is the additional non-mobile IVOC emissions equivalent to 14.8% of non-mobile NMOG.

Figure 3.3 (b) shows in Case 3, the model predicts 4.9 μg $m^{-3}$ of non-mobile hydrocarbon IVOC and 7.7 μg $m^{-3}$ of total hydrocarbon IVOC, which is only somewhat higher than the measured value (6.3 μg $m^{-3}$). Case 4 predicts additional 6.3 μg $m^{-3}$ of oxygenated IVOC from non-mobile sources and 14.0 μg $m^{-3}$ total IVOC (hydrocarbon + oxygenated), which exceeds the measured total IVOC (10.5 μg $m^{-3}$) by 30%. Given this overprediction and the fact that mildly-oxygenated IVOCs can also be formed through secondary chemistry, these results suggest that the IVOC-to-NMOG ratio for non-mobile sources is between Case 3 (12%) and Case 4 (26.8%). In addition, recent research suggests that up to a factor of 3 scale-up may be needed for VCP NMOG emissions (McDonald et al., 2018), which would drive down the IVOC-to-NMOG ratios to 4 - 9%.

#### 3.4.2.2 OA time series and diurnal patterns

Adding non-mobile IVOC emissions increased the predicted afternoon peak SOA concentration to 5.0 and 8.6 μg $m^{-3}$ for Cases 3 and 4, respectively. This highlights the potentially large contribution of non-mobile IVOC emissions to SOA formation. Figure 3.5(a) also shows that in Case 4, the modelled OA time series largely explains the observed SOA (RMSE = 2.5 μg $m^{-3}$, $r$ = 0.85), including explaining the observed peak values in the middle of the day. Since increasing OA concentrations also shifts the gas-particle

partitioning of SVOCs to the particle phase (Donahue et al., 2006), there are also minor shifts in POA partitioning from Case 1 to Case 4, but these changes are small and do not substantively alter the model-measurement POA comparison shown in Figure 3.5(c).

Adding non-mobile IVOC emissions also improves the model predictions of SOA contribution to OA in Pasadena. Hayes et al. (2013) apportioned 66% of the OA to OOA (SV-OOA plus LV-OOA) in Pasadena during CalNex campaign. Hersey et al. (2011) apportioned an even higher fraction of 77% OA to OOA in Pasadena in 2009. As a comparison, if no IVOCs are included in the model, Case 1 only predicts SOA only contributes 47% of the total OA. With additional mobile and non-mobile IVOC emissions, our model predicts 67% OA as SOA in Case 3, and 74% in Case 4.

Although Case 4 largely reproduces the measured OA, we do not think that missing IVOC emissions is the only contributor to the poor performance of the base model. The assumption of Case 4 that IVOC contribute 26.8% of non-mobile NMOG is likely too high, and it overpredicts the total measured IVOC concentrations. Other important uncertainties include (1) effect of vapor wall loss on SOA yield (Zhang et al., 2014), (2) PBL modelling, and (3) multigenerational SOA aging and (4) SVOC emission uncertainties. First, SOA yields for VOCs and IVOCs need to be corrected (typically increased) for vapor wall-losses (Akherati et al., 2018). Second, CMAQ likely overpredicts the afternoon PBL height in Pasadena, as discussed in section 3.4.1. Correcting this will likely increase SOA formation and concentrations, reducing the amount of IVOC emissions needed to reach SOA mass closure. Finally, the effects of multigenerational aging on secondary products of SOA precursor oxidation is uncertain. We have represented this phenomenon with model parameters designed for aging of SVOC emissions (Murphy et al., 2017), but the ratio of functionalization versus fragmentation could be different for products of IVOC oxidation

due to differences in carbon number and functionality. Figure 3.5(c) shows that simulated POA reproduces the measured concentrations, so we believe that the uncertainty in SVOC emissions are relatively small. We also acknowledge the model uncertainty in the oxidation and aging of SVOCs, and this can lead to the substantial changes in OA prediction.

Despite all of these potential uncertainties, the exploratory simulations (Cases 3 and 4) indicate non-mobile IVOC emissions are likely an important source of SOA precursors, but its contributions should be between Case 3 and 4 (12% and 26.8% of non-mobile NMOG). The lower value will close the hydrocarbon IVOC but not the SOA mass balance. Correcting the likely underestimate of VCP emissions (McDonald et al., 2018) in current inventories will drive down the needed non-mobile IVOC emissions to 4% to 9% of NMOG emissions.

### 3.4.3 Regional SOA formation

IVOCs also contribute to regional SOA formation. This is shown in Figure 3.6 (a, b), which presents maps of campaign-average NMOG emissions and modelled SOA concentrations. Primary NMOG emissions are concentrated in the densely populated urban areas such as Los Angeles, but due to the transport of SOA precursors, especially IVOCs, Figure 3.6(b) shows that SOA concentrations are spread over a much large spatial domain than the emissions. This is expected given the SOA production requires time for atmospheric oxidation.

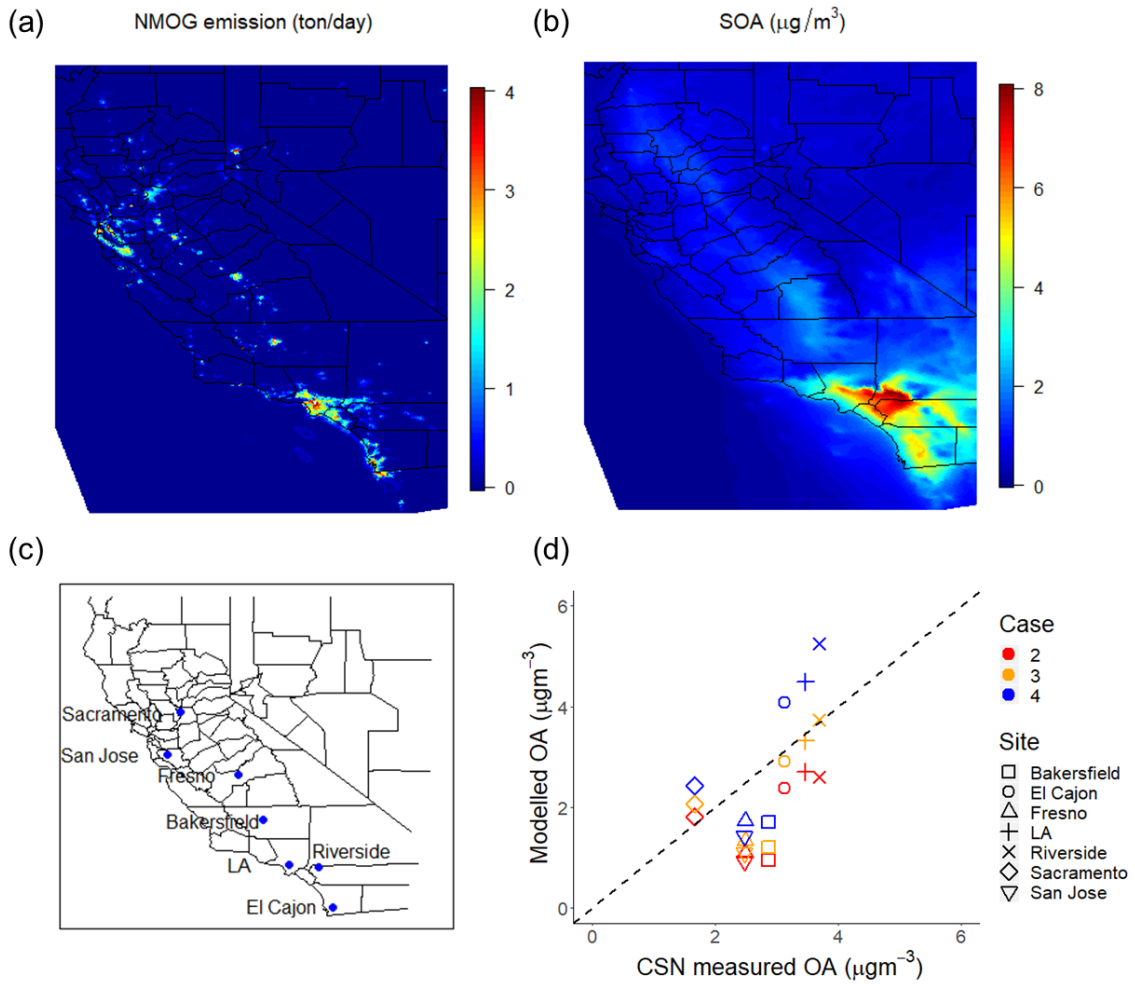



**Figure 3.6 (a) Campaign-average NMOG emissions (Ton day$^{-1}$) in emission inventory. (b) Modelled campaign-averaged SOA concentration in Case 4. (c) Location of CSN sites use for model evaluation. (d) Comparison of modelled OA to measured OA (OC*1.8) at CSN sites in California.**

To evaluate the spatial performance of the model, we compared model predictions of regional OA to CSN data at seven sites in California shown in Figure 3.6(c). Three of the sites are in southern California (LA, Riverside and El Cajon) while the others are central or northern California. Figure 3.6(d) shows the comparison between modelled OA and CSN data (OC*1.8 to account for non-carbonaceous components of the organic aerosol collected on the filters) for all seven sites from Case 1 to Case 4. Table 3.5 summarizes the evaluation metrics for all cases in site-aggregated comparisons.

**Table 3.5 Model OA performance metrics at all CSN sites (1.8*OC) for this study**

| Case | Fractional Bias | Fractional Error |
|---|---|---|
| **1 Baseline** | -0.59 | 0.67 |
| **2 Mobile IVOC** | -0.52 | 0.62 |
| **3 Low Non-mobile IVOC** | -0.33 | 0.49 |
| **4 High Non-mobile IVOC** | -0.10 | 0.42 |

Case 1 grossly underestimated the OA at all sites except for Sacramento, with an fractional bias (FB, definition in SI) of -0.59 and fractional error (FE, definition in SI) of 0.67, of which much of the measured OA are SOA (Docherty et al., 2008; Hayes et al., 2013). Case 2 and Case 3 reduce the fractional bias to -0.52 and -0.33, respectively, and the fractional error to 0.62 to 0.49. Of the four cases considered here, Figure 3.6(d) shows that Case 3 predicted the OA concentrations at three of the southern California CSN sites, but underpredicts at other sites such as Fresno, San Jose and Bakersfield. Case 4 overpredicts the OA concentrations at the southern California CSN sites (coincident with the highest average SOA concentrations), but still underpredicts in Bakersfield, San Jose and Fresno. However, this case has the best overall metrics (FB = -0.10 and FE = 0.42).

Figure 3.6(b) shows that the amount of SOA formed from additional IVOC emissions is much less in northern and central California compared to southern California. This could be due to the different meteorological conditions, or source variations, and/or inaccuracies in the multigenerational aging model. More research is needed to better understand the competition between functionalization and fragmentation of organic gases at long atmospheric timescales. Case 3 and Case 4 were estimated to roughly explain the measured hydrocarbon IVOC and SOA concentration in Pasadena, but measured data of source-specific IVOC-to-NMOG fractions are needed to correctly model the non-mobile emissions.

### 3.5 Conclusions

This paper presents new mobile-source emission profiles that explicitly account for IVOC emissions and a new SOA parameterization for mobile source IVOCs designed for implementation in chemical transport models. We implemented these new profiles and parameterization to investigate the contribution of mobile sources and IVOC emissions to SOA formation in California during the CalNex campaign. We have focused on mobile-source emissions because of the availability of data, but the same basic approach can be applied to other sectors of organic combustion in the future, such as wildfires, agricultural fires, and meat cooking, as additional data become available. The main findings are:

We developed a new parameterization to model SOA formation from mobile source IVOC emissions designed for implementation into CTMs. Explaining the SOA formation from both gasoline and diesel vehicles requires accounting for both the volatility and the chemical composition of the IVOC emissions. Our parameterization has six lumped IVOC species: two aromatic and four aliphatic.

We developed new source profiles for IVOC emissions from mobile sources that are available in SPECIATE 5.0 to facilitate their use in emissions inventory preparation and future CTM simulations. Applying these profiles to the existing EPA inventories predicts that mobile sources contribute 2.7 μg $m^{-3}$ of IVOCs at Pasadena site during CalNex, which is 43% of measured concentrations of hydrocarbon IVOCs.

Mobile source IVOC emissions are predicted to contribute ~1 μg $m^{-3}$ daily-peak SOA concentration, a 67% increase compared to the base case without IVOC emissions. Therefore, mobile-source IVOC emissions need to be included in CTM simulations. However, mobile-source emissions alone don't explain the measured IVOC or SOA concentrations. The growing importance of non-mobile sources underscores the

effectiveness of the decades-long regulatory effort to reduce mobile-source emissions. Results from exploratory model runs suggests that between 12% of 26.8% (or 30.7 to 68.5 Ton $day^{-1}$ in Los Angeles – Pasadena region) of non-mobile NMOG emissions are likely IVOCs.

Future research needs:

VCPs are likely a major source of IVOC and future research is needed to constrain their emissions using ambient observations, bottom-up emission inventory methods and computational models (McDonald et al., 2018; Qin et al., 2019). Measurements of both the volatility distribution and chemical composition of VCP emissions are needed. Modelling the SOA formation from these new IVOCs will likely require extension of existing chemical mechanisms to better represent more oxygenated IVOCs.

More measurements of ambient IVOC concentrations across a range of field sites are needed to better evaluate model performance. Given the lack of data, regional evaluations of ambient IVOC and OA predictions still have large uncertainty.

Improved understanding is needed on the effects of multigenerational aging on SOA formed from IVOC emissions (and other precursors). The impacts of polluted plumes on downwind receptors depends on the nature of aging processes and whether they result in the addition or reduction of particulate mass (e.g. fragmentation processes may enhance volatilization of OA downwind of sources).

### 3.6 Acknowledgments

This publication was developed as part of the Center for Air, Climate, and Energy Solutions (CACES), which is supported under Assistance Agreement No. RD83587301 awarded by the U.S. Environmental Protection Agency. This publication was also supported

by Oak Ridge Institute for Science and Education (ORISE) Research Participation Program for the U.S. Environmental Protection Agency (EPA). The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication. The authors thank Neil M. Donahue for helpful discussions.

## 3.7 References

Akherati, A., Cappa, C. D., Kleeman, M. J., Docherty, K. S., Jimenez, J. L., Griffith, S. M., Dusanter, S., Stevens, P. S. and Jathar, S. H.: Simulating secondary organic aerosol in a regional air quality model using the statistical oxidation model - Part 3: Assessing the influence of semi-volatile and intermediate volatility organic compounds and NOX, Atmos. Chem. Phys. Discuss., 2018, 1–44, doi:10.5194/acp-2018-616, 2018.

Apte, J. S., Brauer, M., Cohen, A. J., Ezzati, M. and Pope, C. A.: Ambient PM2.5 Reduces Global and Regional Life Expectancy, Environ. Sci. Technol. Lett., 5(9), 546–551, doi:10.1021/acs.estlett.8b00360, 2018.

Baker, K. R., Carlton, A. G., Kleindienst, T. E., Offenberg, J. H., Beaver, M. R., Gentner, D. R., Goldstein, A. H., Hayes, P. L., Jimenez, J. L., Gilman, J. B., De Gouw, J. A., Woody, M. C., Pye, H. O. T. T., Kelly, J. T., Lewandowski, M., Jaoui, M., Stevens, P. S., Brune, W. H., Lin, Y. H., Rubitschun, C. L. and Surratt, J. D.: Gas and aerosol carbon in California: Comparison of measurements and model predictions in Pasadena and Bakersfield, Atmos. Chem. Phys., 15(9), 5243–5258, doi:10.5194/acp-15-5243-2015, 2015.

Bash, J. O., Baker, K. R. and Beaver, M. R.: Evaluation of improved land use and canopy representation in BEIS v3.61 with biogenic VOC measurements in California, Geosci. Model Dev., 9(6), 2191–2207, doi:10.5194/gmd-9-2191-2016, 2016.

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, A., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D. and De Gouw, J. A.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris, J. Geophys. Res. Atmos., 118(4), 2041–2057, doi:10.1002/jgrd.50059, 2013.

Carter, W. P. L.: Saprc-07 Chemical Mechanism, , (05), 2010.

Cazorla, A., Bahadur, R., Suski, K. J., Cahill, J. F., Chand, D., Schmid, B., Ramanathan, V. and Prather, K. A.: Relating aerosol absorption due to soot, organic carbon, and dust to emission sources determined from in-situ chemical measurements, Atmos. Chem. Phys., 13(18), 9337–9350, doi:10.5194/acp-13-9337-2013, 2013.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crounse, J. D., Kürten, A., Wennberg, P. O., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: Implications for oxidation of intermediate volatility organic compounds (IVOCs), Atmos. Chem. Phys., 9(9), 3049–3060, doi:10.5194/acp-9-3049-2009, 2009.

Cross, E. S., Hunter, J. F., Carrasquillo, A. J., Franklin, J. P., Herndon, S. C., Jayne, J. T., Worsnop,

D. R., Miake-Lye, R. C. and Kroll, J. H.: Online measurements of the emissions of intermediate-volatility and semi-volatile organic compounds from aircraft, Atmos. Chem. Phys., 13(15), 7845–7858, doi:10.5194/acp-13-7845-2013, 2013.

Cross, E. S., Sappok, A. G., Wong, V. W. and Kroll, J. H.: Load-Dependent Emission Factors and Chemical Characteristics of IVOCs from a Medium-Duty Diesel Engine, Environ. Sci. Technol., 49(22), 13483–13491, doi:10.1021/acs.est.5b03954, 2015.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F. and Schwartz, J. D.: Air Pollution and Mortality in the Medicare Population, N. Engl. J. Med., 376(26), 2513–2522, doi:10.1056/NEJMoa1702747, 2017.

Docherty, K. S., Stone, E. A., Ulbrich, I. M., DeCarlo, P. F., Snyder, D. C., Schauer, J. J., Peltier, R. E., Weber, R. J., Murphy, S. M. and Seinfeld, J. H.: Apportionment of primary and secondary organic aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR-1), Environ. Sci. Technol., 42(20), 7655–7662, 2008.

Donahue, N. M., Robinson, A. L., Stanier, C. O. and Pandis, S. N.: Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics, Environ. Sci. Technol., 40(8), 2635–2643, doi:10.1021/es052297c, 2006.

Drozd, G. T., Zhao, Y., Saliba, G., Frodin, B., Maddox, C., Oliver Chang, M.-C., Maldonado, H., Sardar, S., Weber, R. J., Robinson, A. L. and Goldstein, A. H.: Detailed Speciation of Intermediate Volatility and Semivolatile Organic Compound Emissions from Gasoline Vehicles: Effects of Cold-Starts and Implications for Secondary Organic Aerosol Formation, Environ. Sci. Technol., 53(3), 1706–1714, doi:10.1021/acs.est.8b05600, 2019.

Dunmore, R. E., Hopkins, J. R., Lidster, R. T., Lee, J. D., Evans, M. J., Rickard, A. R., Lewis, A. C. and Hamilton, J. F.: Diesel-related hydrocarbons can dominate gas phase reactive carbon in megacities, Atmos. Chem. Phys., 15(17), 9983–9996, doi:10.5194/acp-15-9983-2015, 2015.

Ensberg, J. J., Hayes, P. L., Jimenez, J. L., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Holloway, J. S., Gordon, T. D., Jathar, S., Robinson, A. L. and Seinfeld, J. H.: Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation, Atmos. Chem. Phys., 14(5), 2383–2397, doi:10.5194/acp-14-2383-2014, 2014.

Fast, J. D., Allan, J., Bahreini, R., Craven, J., Emmons, L., Ferrare, R., Hayes, P. L., Hodzic, A., Holloway, J., Hostetler, C., Jimenez, J. L., Jonsson, H., Liu, S., Liu, Y., Metcalf, A., Middlebrook, A., Nowak, J., Pekour, M., Perring, A., Russell, L., Sedlacek, A., Seinfeld, J., Setyan, A., Shilling, J., Shrivastava, M., Springston, S., Song, C., Subramanian, R., Taylor, J. W., Vinoj, V., Yang, Q., Zaveri, R. A. and Zhang, Q.: Modeling regional aerosol and aerosol precursor variability over California and its sensitivity to emissions and long-range transport during the 2010 CalNex and CARES campaigns, Atmos. Chem. Phys., 14(18), 10013–10060, doi:10.5194/acp-14-10013-2014, 2014.

Fry, J. L., Draper, D. C., Barsanti, K. C., Smith, J. N., Ortega, J., Winkler, P. M., Lawler, M. J., Brown, S. S., Edwards, P. M., Cohen, R. C. and Lee, L.: Secondary Organic Aerosol Formation and Organic Nitrate Yield from NO3 Oxidation of Biogenic Hydrocarbons, Environ. Sci. Technol., 48(20), 11944–11953, doi:10.1021/es502204x, 2014.

Gentner, D. R., Jathar, S. H., Gordon, T. D., Bahreini, R., Day, D. A., El Haddad, I., Hayes, P. L., Pieber, S. M., Platt, S. M., de Gouw, J., Goldstein, A. H., Harley, R. A., Jimenez, J. L., Prévôt, A. S. H. and Robinson, A. L.: Review of Urban Secondary Organic Aerosol Formation from Gasoline and Diesel Motor Vehicle Emissions, Environ. Sci. Technol., acs.est.6b04509, doi:10.1021/acs.est.6b04509, 2017.

Gordon, T. D., Tkacik, D. S., Presto, A. A., Zhang, M. and Shantanu, H.: Primary Gas- and Particle-Phase Emissions and Secondary Organic Aerosol Production from Gasoline and Diesel Off-Road

Engines, 2013.

Gordon, T. D., Presto, A. A., May, A. A., Nguyen, N. T., Lipsky, E. M., Donahue, N. M., Gutierrez, A., Zhang, M., Maddox, C., Rieger, P., Chattopadhyay, S., Maldonado, H., Maricq, M. M. and Robinson, A. L.: Secondary organic aerosol formation exceeds primary particulate matter emissions for light-duty gasoline vehicles, Atmos. Chem. Phys., 14(9), 4661–4678, doi:10.5194/acp-14-4661-2014, 2014.

de Gouw, J. A., Gilman, J. B., Kim, S.-W., Alvarez, S. L., Dusanter, S., Graus, M., Griffith, S. M., Isaacman-VanWertz, G., Kuster, W. C., Lefer, B. L., Lerner, B. M., McDonald, B. C., Rappenglück, B., Roberts, J. M., Stevens, P. S., Stutz, J., Thalman, R., Veres, P. R., Volkamer, R., Warneke, C., Washenfelder, R. A. and Young, C. J.: Chemistry of Volatile Organic Compounds in the Los Angeles Basin: Formation of Oxygenated Compounds and Determination of Emission Ratios, J. Geophys. Res. Atmos., 123(4), 2298–2319, doi:10.1002/2017JD027976, 2018.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J. and Barsanti, K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors, Atmos. Chem. Phys., 17(2), 1471–1489, doi:10.5194/acp-17-1471-2017, 2017.

Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W., Toohey, D. W., Flynn, J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglück, B., Taylor, J. W., Allan, J. D., Holloway, J. S., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Massoli, P., Zhang, X., Liu, J., Weber, R. J., Corrigan, A. L., Russell, L. M., Isaacman, G., Worton, D. R., Kreisberg, N. M., Goldstein, A. H., Thalman, R., Waxman, E. M., Volkamer, R., Lin, Y. H., Surratt, J. D., Kleindienst, T. E., Offenberg, J. H., Dusanter, S., Griffith, S., Stevens, P. S., Brioude, J., Angevine, W. M. and Jimenez, J. L.: Organic aerosol composition and sources in Pasadena, California, during the 2010 CalNex campaign, J. Geophys. Res. Atmos., 118(16), 9233–9257, doi:10.1002/jgrd.50530, 2013.

Hayes, P. L., Carlton, a. G., Baker, K. R., Ahmadov, R., Washenfelder, R. A., Alvarez, S., Rappenglück, B., Gilman, J. B., Kuster, W. C., de Gouw, J. A., Zotter, P., Prévôt, a. S. H., Szidat, S., Kleindienst, T. E., Offenberg, J. H., Ma, P. K. and Jimenez, J. L.: Modeling the formation and aging of secondary organic aerosols in Los Angeles during CalNex 2010, Atmos. Chem. Phys., 15(10), 5773–5801, doi:10.5194/acp-15-5773-2015, 2015.

Hersey, S. P., Craven, J. S., Schilling, K. A., Metcalf, A. R., Sorooshian, A., Chan, M. N., Flagan, R. C. and Seinfeld, J. H.: The Pasadena Aerosol Characterization Observatory (PACO): chemical and physical analysis of the Western Los Angeles basin aerosol, Atmos. Chem. Phys., 11(15), 7417–7443, 2011.

Hodzic, A., Jimenez, J. L., Madronich, S., Canagaratna, M. R., Decarlo, P. F., Kleinman, L. and Fast, J.: Modeling organic aerosols in a megacity: Potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation, Atmos. Chem. Phys., 10(12), 5491–5514, doi:10.5194/acp-10-5491-2010, 2010.

Hoyle, C. R., Boy, M., Donahue, N. M., Fry, J. L., Glasius, M., Guenther, A., Hallar, A. G., Huff Hartz, K., Petters, M. D., Petäjä, T., Rosenoern, T. and Sullivan, A. P.: A review of the anthropogenic influence on biogenic secondary organic aerosol, Atmos. Chem. Phys., 11(1), 321–343, doi:10.5194/acp-11-321-2011, 2011.

Huang, C., Hu, Q., Li, Y., Tian, J., Ma, Y., Zhao, Y., Feng, J., An, J., Qiao, L., Wang, H., Jing, S., Huang, D., Lou, S., Zhou, M., Zhu, S., Tao, S. and Li, L.: Intermediate Volatility Organic Compound Emissions from a Large Cargo Vessel Operated under Real-World Conditions, Environ. Sci. Technol., 52(21), 12934–12942, doi:10.1021/acs.est.8b04418, 2018.

Hunter, J. F., Day, D. A., Palm, B. B., Yatavelli, R. L. N., Chan, A. W. H., Kaser, L., Cappellin, L., Hayes, P. L., Cross, E. S., Carrasquillo, A. J., Campuzano-Jost, P., Stark, H., Zhao, Y., Hohaus, T.,

Smith, J. N., Hansel, A., Karl, T., Goldstein, A. H., Guenther, A., Worsnop, D. R., Thornton, J. A., Heald, C. L., Jimenez, J. L. and Kroll, J. H.: Comprehensive characterization of atmospheric organic carbon at a forested site, Nat. Geosci., 10, 748 [online] Available from: http://dx.doi.org/10.1038/ngeo3018, 2017.

Hutzell, W. T., Luecken, D. J., Appel, K. W. and Carter, W. P. L.: Interpreting predictions from the SAPRC07 mechanism based on regional and continental simulations, Atmos. Environ., 46, 417–429, doi:https://doi.org/10.1016/j.atmosenv.2011.09.030, 2012.

Jathar, S. H., Miracolo, M. A., Tkacik, D. S., Donahue, N. M., Adams, P. J. and Robinson, A. L.: Secondary organic aerosol formation from photo-oxidation of unburned fuel: experimental results and implications for aerosol formation from combustion emissions., Environ. Sci. Technol., 47(22), 12886–93, doi:10.1021/es403445q, 2013.

Jathar, S. H., Woody, M., Pye, H. O. T., Baker, K. R. and Robinson, A. L.: Chemical transport model simulations of organic aerosol in southern California: model evaluation and gasoline and diesel source contributions, Atmos. Chem. Phys., 17(6), 4305–4318, doi:10.5194/acp-17-4305-2017, 2017.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U. and Worsnop, D. R.: Evolution of Organic Aerosols in the Atmosphere, Science (80-. )., 326(5959), 1525–1529 [online] Available from: http://science.sciencemag.org/content/326/5959/1525.abstract, 2009.

Khare, P. and Gentner, D. R.: Considering the future of anthropogenic gas-phase organic compound emissions and the increasing influence of non-combustion sources on urban air quality, Atmos. Chem. Phys., 18(8), 5391–5413, doi:10.5194/acp-18-5391-2018, 2018.

Kim, S.-W., McDonald, B. C., Baidar, S., Brown, S. S., Dube, B., Ferrare, R. A., Frost, G. J., Harley, R. A., Holloway, J. S., Lee, H.-J., McKeen, S. A., Neuman, J. A., Nowak, J. B., Oetjen, H., Ortega, I., Pollack, I. B., Roberts, J. M., Ryerson, T. B., Scarino, A. J., Senff, C. J., Thalman, R., Trainer, M., Volkamer, R., Wagner, N., Washenfelder, R. A., Waxman, E. and Young, C. J.: Modeling the weekly cycle of NOx and CO emissions and their impacts on O3 in the Los Angeles-South Coast Air Basin during the CalNex 2010 field campaign, J. Geophys. Res. Atmos., 121(3), 1340–1360, doi:10.1002/2015JD024292, 2016.

Koo, B., Knipping, E. and Yarwood, G.: 1.5-Dimensional volatility basis set approach for modeling organic aerosol in CAMx and CMAQ, Atmos. Environ., 95, 158–164, doi:10.1016/j.atmosenv.2014.06.031, 2014.

Kroll, J. H., Cross, E. S., Hunter, J. F., Carrasquillo, A. J., Franklin, J. P., Herndon, S. C., Jayne, J. T., Worsnop, D. R., Lye, R. C. M. and Onasch, T. B.: Emissions of Gas - Phase Low - Volatility Organic Compounds from Mobile Sources, , (March), 2014.

Liggio, J., Li, S., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., Drollette, B. D., Gordon, M., Lee, P., Liu, P., Leithead, A., Moussa, S. G., Wang, D., Brien, J. O., Mittermeier, R. L., Osthoff, H. D., Makar, P. A., Zhang, J., Brook, J. R., Lu, G., Staebler, R. M., Han, Y., Travis, W., Plata, D. L. and Gentner, D. R.: Oil sands operations as a large source of secondary organic aerosols, Nature, 534(7605), 1–16, doi:10.1038/nature17646, 2016.

Lim, Y. Bin and Ziemann, P. J.: Products and mechanism of secondary organic aerosol formation from reactions of n-alkanes with OH radicals in the presence of NO x, Environ. Sci. Technol., 39(23), 9229–9236, 2005.

Lim, Y. Bin and Ziemann, P. J.: Chemistry of Secondary Organic Aerosol Formation from OH Radical-Initiated Reactions of Linear, Branched, and Cyclic Alkanes in the Presence of NO x , Aerosol Sci. Technol., 43(6), 604–619, doi:10.1080/02786820902802567, 2009.

Lu, Q., Zhao, Y. and Robinson, A. L.: Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions, Atmos. Chem. Phys. Discuss., 2018, 1–28, doi:10.5194/acp-2018-752, 2018.

May, A. A., Presto, A. A., Hennigan, C. J., Nguyen, N. T., Gordon, T. D. and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: (2) diesel vehicles, Environ. Sci. Technol., 47(15), 8288–8296, doi:10.1021/es400782j, 2013a.

May, A. A., Levin, E. J. T., Hennigan, C. J., Riipinen, I., Lee, T., Collett, J. L., Jimenez, J. L., Kreidenweis, S. M. and Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions: 3. Biomass burning, J. Geophys. Res. Atmos., 118(19), 11327–11338, doi:10.1002/jgrd.50828, 2013b.

May, A. A., Nguyen, N. T., Presto, A. A., Gordon, T. D., Lipsky, E. M., Karve, M., Gutierrez, A., Robertson, W. H., Zhang, M., Brandow, C., Chang, O., Chen, S., Cicero-Fernandez, P., Dinkins, L., Fuentes, M., Huang, S. M., Ling, R., Long, J., Maddox, C., Massetti, J., McCauley, E., Miguel, A., Na, K., Ong, R., Pang, Y., Rieger, P., Sax, T., Truong, T., Vo, T., Chattopadhyay, S., Maldonado, H., Maricq, M. M. and Robinson, A. L.: Gas- and particle-phase primary emissions from in-use, on-road gasoline and diesel vehicles, Atmos. Environ., 88, 247–260, doi:10.1016/j.atmosenv.2014.01.046, 2014.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B. and Trainer, M.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science (80-. )., 359(6377), 760–764, doi:10.1126/science.aaq0524, 2018.

Miracolo, M. A., Hennigan, C. J., Ranjan, M., Nguyen, N. T., Gordon, T. D., Lipsky, E. M., Presto, A. A., Donahue, N. M. and Robinson, A. L.: Secondary aerosol formation from photochemical aging of aircraft exhaust in a smog chamber, Atmos. Chem. Phys., 11(9), 4135–4147, 2011.

Murphy, B. N., Woody, M. C., Jimenez, J. L., Carlton, A. M. G., Hayes, P. L., Liu, S., Ng, N. L., Russell, L. M., Setyan, A., Xu, L., Young, J., Zaveri, R. A., Zhang, Q. and Pye, H. O. T.: Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: Impacts on source strength and partitioning, Atmos. Chem. Phys., 17(18), 11107–11133, doi:10.5194/acp-17-11107-2017, 2017.

Pereira, K. L., Dunmore, R., Whitehead, J., Alfarra, M. R., Allan, J. D., Alam, M. S., Harrison, R. M., McFiggans, G. and Hamilton, J. F.: Technical note: Use of an atmospheric simulation chamber to investigate the effect of different engine conditions on unregulated VOC-IVOC diesel exhaust emissions, Atmos. Chem. Phys., 18(15), 11073–11096, doi:10.5194/acp-18-11073-2018, 2018.

Platt, S. M., El Haddad, I., Pieber, S. M., Zardini, A. A., Suarez-Bertoa, R., Clairotte, M., Daellenbach, K. R., Huang, R.-J., Slowik, J. G., Hellebust, S., Temime-Roussel, B., Marchand, N., de Gouw, J., Jimenez, J. L., Hayes, P. L., Robinson, A. L., Baltensperger, U., Astorga, C. and Prévôt, A. S. H.: Gasoline cars produce more carbonaceous particulate matter than modern filter-equipped diesel cars, Sci. Rep., 7(1), 4926, doi:10.1038/s41598-017-03714-9, 2017.

Presto, A. A., Miracolo, M. A., Kroll, J. H., Worsnop, D. R., Robinson, A. L. and Donahue, N. M.: Intermediate-volatility organic compounds: A potential source of ambient oxidized organic aerosol, Environ. Sci. Technol., 43(13), 4744–4749, doi:10.1021/es803219q, 2009.

Presto, A. A., Miracolo, M. A., Donahue, N. M. and Robinson, A. L.: Secondary Organic Aerosol Formation from High-NOx Photo-Oxidation of Low Volatility Precursors: n-Alkanes, Environ. Sci. Technol., 44(6), 2029–2034, doi:10.1021/es903712r, 2010a.

Presto, A. A., Nguyen, N. T., Ranjan, M., Reeder, A. J., Lipsky, E. M., Hennigan, C. J., Miracolo, M. A., Riemer, D. D. and Robinson, A. L.: Fine particle and organic vapor emissions from staged tests of an in-use aircraft engine, Atmos. Environ., 45(21), 3603–3612, doi:10.1016/j.atmosenv.2011.03.061, 2011.

Presto, A. a, Miracolo, M. a, Donahue, N. M. and Robinson, A. L.: Secondary Organic Aerosol Formation from High-NO x Photo-Oxidation of Low Volatility Precursors : n -Alkanes, Atmos. Chem. Phys., 44(6), 2029–2034, 2010b.

Pye, H. O. T., Murphy, B. N., Xu, L., Ng, N. L., Carlton, A. G., Guo, H., Weber, R., Vasilakos, P., Appel, K. W., Budisulistiorini, S. H., Surratt, J. D., Nenes, A., Hu, W., Jimenez, J. L., Isaacman-VanWertz, G., Misztal, P. K. and Goldstein, A. H.: On the implications of aerosol liquid water and phase separation for organic aerosol mass, Atmos. Chem. Phys., 17(1), 343–369, doi:10.5194/acp-17-343-2017, 2017.

Qi, L., Liu, H., Shen, X., Fu, M., Huang, F., Man, H., Deng, F., Shaikh, A. A., Wang, X., Dong, R., Song, C. and He, K.: Intermediate-Volatility Organic Compound Emissions from Nonroad Construction Machinery under Different Operation Modes, Environ. Sci. Technol., 53(23), 13832–13840, doi:10.1021/acs.est.9b01316, 2019.

Qin, M., Murphy, B. N., McDonald, B. C. B., McKeen, S. A., Koval, L. and al, et: Impacts of volatile chemical products on criteria pollutants in an urban atmosphere (in preparation), 2019.

Ranjan, M., Presto, A. a., May, A. a. and Robinson, A. L.: Temperature Dependence of Gas–Particle Partitioning of Primary Organic Aerosol Emissions from a Small Diesel Engine, Aerosol Sci. Technol., 46(1), 13–21, doi:10.1080/02786826.2011.602761, 2012.

Ryerson, T. B., Andrews, A. E., Angevine, W. M., Bates, T. S., Brock, C. A., Cairns, B., Cohen, R. C., Cooper, O. R., de Gouw, J. A., Fehsenfeld, F. C., Ferrare, R. A., Fischer, M. L., Flagan, R. C., Goldstein, A. H., Hair, J. W., Hardesty, R. M., Hostetler, C. A., Jimenez, J. L., Langford, A. O., McCauley, E., McKeen, S. A., Molina, L. T., Nenes, A., Oltmans, S. J., Parrish, D. D., Pederson, J. R., Pierce, R. B., Prather, K., Quinn, P. K., Seinfeld, J. H., Senff, C. J., Sorooshian, A., Stutz, J., Surratt, J. D., Trainer, M., Volkamer, R., Williams, E. J. and Wofsy, S. C.: The 2010 California Research at the Nexus of Air Quality and Climate Change (CalNex) field study, J. Geophys. Res. Atmos., 118(11), 5830–5866, doi:10.1002/jgrd.50331, 2013.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W. and Powers, J. G.: A Description of the Advanced Research WRF Version 3, 2008.

Subramanian, R., Khlystov, A. Y., Cabada, J. C. and Robinson, A. L.: Positive and Negative Artifacts in Particulate Organic Carbon Measurements with Denuded and Undenuded Sampler Configurations Special Issue of Aerosol Science and Technology on Findings from the Fine Particulate Matter Supersites Program, Aerosol Sci. Technol., 38(sup1), 27–48, doi:10.1080/02786820390229354, 2004.

Tkacik, D. S., Presto, A. A., Donahue, N. M. and Robinson, A. L.: Secondary organic aerosol formation from intermediate-volatility organic compounds: Cyclic, linear, and branched alkanes, Environ. Sci. Technol., 46(16), 8773–8781, doi:10.1021/es301112c, 2012.

US EPA: Final Report, SPECIATE Version 5.0, Database Development Documentation. [online] Available from: https://www.epa.gov/sites/production/files/2019-07/documents/speciate_5.0.pdf, 2019.

US EPA Office of Research and Development: CMAQ (Version 5.3), , doi:10.5281/zenodo.3379043, 2019.

Wang, X., Cheung, C. S., Di, Y. and Huang, Z.: Diesel engine gaseous and particle emissions fueled with diesel-oxygenate blends, Fuel, 94, 317–323, doi:10.1016/j.fuel.2011.09.016, 2012.

Warneke, C., de Gouw, J. A., Holloway, J. S., Peischl, J., Ryerson, T. B., Atlas, E., Blake, D., Trainer, M. and Parrish, D. D.: Multiyear trends in volatile organic compounds in Los Angeles, California: Five decades of decreasing emissions, J. Geophys. Res. Atmos., 117(D21), doi:10.1029/2012JD017899, 2012.

Woody, M. C., Baker, K. R., Hayes, P. L., Jimenez, J. L., Koo, B. and Pye, H. O. T. T.: Understanding sources of organic aerosol during CalNex-2010 using the CMAQ-VBS, Atmos. Chem. Phys., 16(6), 4081–4100, doi:10.5194/acp-16-4081-2016, 2016.

Worton, D. R., Isaacman, G., Gentner, D. R., Dallmann, T. R., Chan, A. W. H., Ruehl, C., Kirchstetter, T. W., Wilson, K. R., Harley, R. A. and Goldstein, A. H.: Lubricating Oil Dominates Primary Organic Aerosol Emissions from Motor Vehicles, Environ. Sci. Technol., 48(7), 3698–3706, doi:10.1021/es405375j, 2014.

Xie, Y., Paulot, F., Carter, W. P. L., Nolte, C. G., Luecken, D. J., Hutzell, W. T., Wennberg, P. O., Cohen, R. C. and Pinder, R. W.: Understanding the impact of recent advances in isoprene photooxidation on simulations of regional air quality, Atmos. Chem. Phys., 13(16), 8439–8455, doi:10.5194/acp-13-8439-2013, 2013.

Xu, L., Pye, H. O. T., He, J., Chen, Y., Murphy, B. N. and Ng, N. L.: Experimental and model estimates of the contributions from biogenic monoterpenes and sesquiterpenes to secondary organic aerosol in the southeastern United States, Atmos. Chem. Phys., 18(17), 12613–12637, doi:10.5194/acp-18-12613-2018, 2018.

Ying, Q. and Li, J.: Implementation and initial application of the near-explicit Master Chemical Mechanism in the 3D Community Multiscale Air Quality (CMAQ) model, Atmos. Environ., 45(19), 3244–3256, doi:https://doi.org/10.1016/j.atmosenv.2011.03.043, 2011.

Zhang, R., Wang, G., Guo, S., Zamora, M. L., Ying, Q., Lin, Y., Wang, W., Hu, M. and Wang, Y.: Formation of Urban Fine Particulate Matter, Chem. Rev., 115(10), 3803–3855, doi:10.1021/acs.chemrev.5b00067, 2015.

Zhang, X., Cappa, C. D., Jathar, S. H., McVay, R. C., Ensberg, J. J., Kleeman, M. J., Seinfeld, J. H. and Christopher D. Cappa: Influence of vapor wall loss in laboratory chambers on yields of secondary organic aerosol., Proc. Natl. Acad. Sci. U. S. A., 111(16), 1–6, doi:10.1073/pnas.1404727111, 2014.

Zhang, Y., West, J. J., Mathur, R., Xing, J., Hogrefe, C., Roselle, S. J., Bash, J. O., Pleim, J. E., Gan, C.-M. and Wong, D. C.: Long-term trends in the ambient PM2.5- and O3-related mortality burdens in the United States under emission reductions from 1990 to 2010, Atmos. Chem. Phys., 18(20), 15003–15016, doi:10.5194/acp-18-15003-2018, 2018.

Zhao, Y., Hennigan, C. J., May, A. A., Tkacik, D. S., De Gouw, J. A., Gilman, J. B., Kuster, W. C., Borbon, A. and Robinson, A. L.: Intermediate-volatility organic compounds: A large source of secondary organic aerosol, Environ. Sci. Technol., 48(23), 13743–13750, doi:10.1021/es5035188, 2014.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environ. Sci. Technol., 49(19), 11516–11526, doi:10.1021/acs.est.5b02841, 2015.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines, Environ. Sci. Technol., 50(8), 4554–4563, doi:10.1021/acs.est.5b06247, 2016.

Zhao, Y., Saleh, R., Saliba, G., Presto, A. A., Gordon, T. D., Drozd, G. T., Goldstein, A. H., Donahue, N. M. and Robinson, A. L.: Reducing secondary organic aerosol formation from gasoline vehicle exhaust, Proc. Natl. Acad. Sci. , 114(27), 6984–6989, doi:10.1073/pnas.1620911114, 2017.

# Chapter 4: Analysis of the Unresolved Complex Mixture of Intermediate Volatile Organic Compounds in Gas Chromatograph-Mass Spectrum Data using Positive Matrix Factorization


## Abstract

We present a new approach based on Positive Matrix Factorization (PMF) to better characterize the unresolved complex mixture (UCMs) in gas chromatograph-mass spectrum (GC-MS) data. The UCM is a complex mixture of typically higher molecular weight compounds that co-elute during traditional one-dimensional (1-D) GC analysis of sorbent tube and filter samples and therefore cannot be speciated at the molecular level. The goal of this work is to recover information on the chemical composition of the intermediate volatility organic compounds (IVOC) UCM for use in source apportionment and for simulation of secondary organic aerosol formation. To demonstrate the technique, we applied it to a range of emissions and ambient samples. We identified the chemical character of the PMF factors by comparison with reference compounds in the NIST mass spectral database. The PMF technique can recover more detailed information measured by more sophisticated instrumentation. PMF-decomposed mobile source samples (gasoline, diesel and aircraft) show the effects of emissions control technologies, fuel composition and engine load on IVOC composition. For example, gasoline vehicle sources show increasing fraction of alkanes and oxygenates and decreasing fraction of single-ring aromatics (SRAs) from small off-road engine (SORE) to newer on-road vehicles. IVOC emission from biomass burning show very different composition than mobile sources, with a large fraction (>50%) of IVOC oxygenates. We also performed source apportionment analysis combining PMF technique

with chemical mass balance model (CMB) on samples collected in a highway tunnel; even though diesel vehicles only used 20% of the fuel consumption in the tunnel they contributed around 70% of mobile source IVOCs. PMF analysis on ambient samples in Pasadena, CA show very high fractions (>70%) of oxygenates, where minor hydrocarbon IVOC peak in early afternoon suggesting contributions from unknown local evaporative IVOC emission sources. SOA modelling on mobile sources emissions show an 80% increase of SOA yield of gasoline vehicle emissions under low-$NO_x$ conditions, highlighting the importance of IVOC chemical information in chemical transport models. This technique can be applied to both archival and future GC-MS data analysis.

## 4.1 Introduction

Gas chromatography-mass spectrometry (GC-MS) analysis is commonly used to characterize atmospheric organics, including volatile organic compounds (VOCs) (Chow et al., 2007; May et al., 2014; Yao et al., 2015), intermediate-volatile organic compounds (IVOCs) (Presto et al., 2011; Zhao et al., 2014, 2015, 2016) and organic aerosol (OA) (Kreisberg et al., 2014; Williams et al., 2006; Zhao et al., 2013). A challenge is the exponential increase in structural isomers as the carbon number increases (Goldstein and Galbally, 2007). This means that majority of lower volatility organics cannot be characterized at the molecular level by traditional one-dimensional GC-MS techniques. Instead these organics frequently appear as a large hump of co-eluting compounds in the chromatogram, commonly referred to as unresolved complex mixture (UCM) (Presto et al., 2011).

The UCM is a prominent feature in 1D GC-MS chromatograms of sorbent (VOCs, IVOCs and SVOCs) and filter (SVOCs and lower volatility organics) samples. It frequently

contributes more than 80% of the mass on these types of samples (Presto et al., 2011; Zhao et al., 2014, 2015, 2016). The lack of chemical information complicates including lower volatility organics such as SVOCs and IVOCs in atmospheric chemistry models used for source apportionment analysis and the simulation of secondary organic aerosol formation (Lu et al., 2019; Pye and Pouliot, 2012).

This work focuses on IVOCs, which are ubiquitous in emissions from mobile sources (Drozd et al., 2019; Lu et al., 2018; Qi et al., 2019; Schauer et al., 1999; Zhao et al., 2015), coal combustion (Cai et al., 2019), biomass burning (Hatch et al., 2018), volatile chemical products (Khare and Gentner, 2018; McDonald et al., 2018), biogenics (Hunter et al., 2017). They are also an important component of the organic budget in the ambient urban atmosphere (Zhao et al., 2014). IVOCs are organic compounds that have an effective saturation concentration (C*) at 298K between $10^3$ and $10^6$ ug/$m^3$ (roughly equivalent to $C_{12}$ to $C_{22}$ *n*-alkanes) (Presto et al., 2012), and have been identified as important precursors to secondary organic aerosol (SOA) formation (Hayes et al., 2015; Hodzic et al., 2010; Jathar et al., 2013; Presto et al., 2009; Pye and Seinfeld, 2010; Shah et al., 2020; Tkacik et al., 2014). SOA is a major component of fine particulate matter, which adversely affects human health (Apte et al., 2018; Di et al., 2017).

SOA formation depends on the volatility and molecular structure of the precursor compounds (Lim and Ziemann, 2009a; Loza et al., 2014; Presto et al., 2010; Pye and Pouliot, 2012). Therefore, accurate simulation of SOA in atmospheric chemistry models likely requires additional chemical information on the IVOC UCM (Lu et al., 2019). Most existing parameterization for SOA formation from IVOCs either only has single specie surrogate or only rely on volatility (Akherati et al., 2018; Hayes et al., 2015; Hodzic et al., 2010; Murphy and Pandis, 2009; Robinson et al., 2007).

Researchers have used different approaches to characterize the chemical composition of IVOC UCM. The simplest approach is to use some basic mass spectral information to crudely classify the bulk IVOC UCM into one or two different chemical classes (Zhao et al., 2015, 2016). Researchers have also used more advanced analytic techniques to characterize the UCM composition (Drozd et al., 2019; Hatch et al., 2018), such as soft-ionization (GC-VUV-MS) or 2D-GC. For example, Drozd et al. (2019) used 2D-GC for gasoline vehicle IVOC emissions measurements, where IVOCs are classified into 3 categories (aliphatic, single ring aromatics and PAHs). Hatch et al. (2018) used 2D-GC to quantify the IVOC/SVOC emissions on biomass burning smokes. Sheu et al. (2018) and Khare et al. (2019) describe advances in chemically speciated measurements of trace gas-phase organics, but these approaches have been applied on relatively few samples. There is a need to provide more information from traditional GC-MS results.

Positive matrix factorization (PMF) has been widely used in aerosol chemistry community for source apportionment of $PM_{2.5}$ mass and organic aerosols (Lee and Seung, 2001; Pedregosa et al., 2011). PMF decomposes a data matrix to two lower-rank non-negative matrixes, whose product approximate the original matrix, and minimize the residual (Paatero and Tapper, 1994).

There have been a few applications of PMF technique to GC-MS data of particle phase organics (Fortenberry et al., 2018; Gao et al., 2018; Zhang et al., 2014, 2016). Zhang et al. (2014) used PMF to resolve the chemical composition of ambient OA and compared to aerosol mass spectrometer (AMS) source apportionment. Others have applied to GC-MS results from biomass burning aerosols and catalytic lignin depolymerization measurements (Fortenberry et al., 2018; Gao et al., 2018).

In this work, we apply a PMF technique to analyze the GC-MS data of gas-phase

sorbent samples of emissions and ambient IVOCs. Our main goal is to better resolve the chemical composition of the IVOC UCM by decomposing its mass spectra signal to different chemical classes. To interpret the chemical character of each PMF factor, we compare them to reference mass spectra in the NIST database, rather than attributing factors to sources categories from AMS analysis (Zhang et al., 2016). To illustrate application of the technique, we use it in conjunction with a chemical mass balance (CMB) regression model to estimate IVOC sources in a highway tunnel. We also analyze ambient samples collected in Pasadena CA during the CalNex campaign. We conclude with a discussion of the importance of having improved chemical resolution of IVOC UCM on SOA modeling.

## 4.2 Methods

### 4.2.1 Datasets

The GC-MS datasets analyzed using PMF were compiled from published works on mobile source emissions (Presto et al., 2011; Zhao et al., 2015, 2016), biomass burning emissions (Hennigan et al., 2011), highway tunnel measurements (Tkacik et al., 2014), and ambient samples (Zhao et al., 2014). Detailed information on sample collection can be found in the original papers, here we only briefly describe the experimental design and analytic procedures to provide context for the PMF decomposition of GC-MS data.

All samples are collected on Tenax TA adsorbent tubes, which were then analyzed by the same thermal desorption system (Gerstel, Baltimore, MD) coupled with gas chromatography/mass spectrometry (Agilent 6890 GC/5975 MS) using a capillary GC column (Agilent HP- 5MS, 30 m × 0.25 mm). There were minor differences in the temperature protocol used for some samples. We selected the data between $C_{12}$ to $C_{22}$ n-alkanes retention times as the IVOC volatility range (Zhao et al., 2015).

#### 4.2.1.1 Mobile sources emissions

The mobile source dataset included 55 samples from gasoline vehicles / engines, including small off-road engines (SORE), not LEV-certificated vehicles (pre-LEV), low emission vehicles (LEV) and ultra-low emission vehicles (ULEV), 15 from diesel vehicles, including diesel particulate filter (DPF) equipped and non-DPF vehicles and 2 from aircraft. All samples were from dilute exhaust. For vehicle and engine tests, samples were collected from a constant volume sampler (CVS). The aircraft samples were collected using a sampling rake install approximately 1 meter downstream of a CFM-56 turbofan engine.

Information of vehicle type, certification and test cycle is listed in Table 4.1. We examined the data to determine the effects of certification standard, aftertreatment devices and / or test cycle on mobile sources IVOC composition. The IVOC emissions data from the mobile source experiments have been previously reported (Presto et al., 2011; Zhao et al., 2015, 2016). Diesel vehicles were tested using the hot-start Urban Dynamometer Driving Schedule (UDDS) and select modes of low-speed (creep and idle) and high-speed (high-cruise) operations of the Heavy Heavy-Duty Diesel Truck (HHDDT) driving schedule. We grouped the results from UDDS and high-cruise cycles together as higher-speed (HS) cycles (Zhao et al., 2015).

We use 1:1 gas and diesel split to make sure that the model is resolving the factors from different sources more equally. Another reason is to create a more balanced mixture of single-ring aromatics (SRA), polycyclic aromatic hydrocarbon (PAH) and aliphatic IVOCs as training data. We trained the PMF model with a fraction of samples (15 gasoline, 12 non-DPF diesel, 3 DPF-diesel and 2 aircraft) emission samples and 1 dynamic blank. The training samples are only selected based on fuel type. The model is then evaluated and applied to all samples.

**Table 4.1 Mobile sources emissions**

| Source type | Certification | Test cycle | No. of samples in dataset |
|---|---|---|---|
| Gasoline | SORE | Unified cycle (Cold UC) | 3 |
| | Pre-LEV | | 10 |
| | LEV | | 13 |
| | ULEV | | 23 |
| | All | Hot-running | 6 |
| Diesel | Non-DPF | Idle | 4 |
| | | Higher-speed | 7 |
| | DPF | Idle | 1 |
| | | Higher-speed | 3 |
| Aircraft | | 4% Load | 1 |
| | | 85% Load | 1 |
| Total | | | 82 |

#### 4.2.1.2 Biomass burning smoke

The biomass burning emission samples are from the experiments conducted during the FLAME 2 campaign at the Fire Science Lab (FSL) in Missoula MT (Hennigan et al., 2011). The experiments were designed to simulate open burning of 11 different important wildland fuels (14 total experiments). During each experiment burn, a small mass (0.3–1.0 kg) of fuel was burned to completion, filling the FSL 3000$m^3$ combustion chamber with smoke from the entire burn event. After the smoke was well mixed, sorbent samples were collected for analysis. Table S1 lists the fuel for all samples. The IVOC emissions from the biomass burning experiments has not been previously reported.

#### 4.2.1.3 Highway tunnel

To examine the sources of highway tunnel IVOCs, we analyzed the samples collected in Fort Pitt Tunnel locate on Interstate-376 in Pittsburgh PA (Tkacik et al., 2014) during a two-week period in May 2013. Separate samples of tunnel air were collected during morning rush hour, middle of the day and evening rush hours. Limited data on the IVOC samples collected in the tunnel have been published (Tkacik et al., 2014).

#### 4.2.1.4 Pasadena CA during CalNex campaign

We analyzed ambient samples collected in Pasadena CA during the CalNex campaign in May and June 2010 (Zhao et al., 2014). The goal is to understand the relative contribution of hydrocarbon and oxygenated compounds as a function of weekday/weekend pattern or time of the day. Six sets of composite samples were collected to characterize average weekday and weekend profiles. In each set, 8 samples were collected at different times of the day to characterize average diurnal profiles, totaling 48 samples in the dataset. Data from the Pasadena IVOC samples have been published (Zhao et al., 2014).

### 4.2.2 Positive Matrix Factorization (PMF)

We used the PMF algorithm implemented in the open source software package scikit-learn (Pedregosa et al., 2011). The PMF algorithm performs decomposition of a data matrix X to two lower-rank non-negative matrixes W and H, and to minimize the Frobenius norm of the residual matrix E (Paatero and Tapper, 1994; Pedregosa et al., 2011). For this analysis the original matrix (dimension: K × N × M) is the GC-MS data, where K = number of samples, N = number of scans (retention time) in IVOC range per sample, and M = the *m/z* values in a scan, usually from *m/z* = 36 to *m/z* = 360. This 3-d matrix is then transformed to a 2-d matrix X (dimension: KN × M). Figure S1 shows the schematic plot of overall process of GC-MS data dimension reduction using PMF. The matrix W is the 'scan matrix' or 'abundance matrix' because it has the dimension of scan number (or equivalently retention time). More specifically, every row of W denotes a scan, and the values in this row denotes the abundance of each factor. And H now become the 'factor matrix', each row of H denotes a factor, and the values in this row denotes the abundance of the mass fragments (*m/z*) in this factor.

This paper focus on PMF solutions with 10 to 20 factors, examining the solution for

both factor resolution and interpretability. There are no 'gold standard' to determine the best number of PMF factors, and often more detailed analysis and interpretation are needed (Ulbrich et al., 2009). Too few factors (P < 10) leads to lower resolution and contributions from co-eluted species, and too many factor (P > 20) often lead to duplicate factors or split factors (factors that split the compound mass spectra to two or more parts). To address this issue, we used two criteria: (1) reconstruction error and (2) max correlation efficient ($r$) between factors. Generally, reconstruction error decrease as no. of factors increase, but the max correlation increase. Our goal is to lower the error and keep the max $r$ below certain threshold (for example, $r < 0.7$).

### 4.3 Results and Discussion

In this section, we report results from the PMF analysis on four emission datasets (gasoline, diesel, aircraft, and biomass burning smoke) and two ambient datasets (highway tunnel and Pasadena CA). We also compare the PMF-resolved IVOC composition to published literature results using more sophisticated instruments whenever available.

#### 4.3.1 PMF decomposition of IVOC emissions

Figure 4.1 illustrates the PMF analysis using data from a non-DPF vehicle. Figure 4.1(a) shows the raw GC-MS data (similar figures gasoline, biomass burning and ambient GC-MS data are shown in Figure S2). The raw data are a 2-D matrix of scan number (retention time) and *m/z* ratios.

The feature patterns of the data matrix vary by source type. For example, in the mobile source dataset, the detected mass fragments shift to longer retention time and higher *m/z* from gasoline to aircraft and diesel. The gasoline vehicle sample (Figure S2a) shows

more mass fragments between *m/z* = 110 and 135 at shorter retention times, which are likely alkylated single ring aromatics. Whereas, both aircraft (Figure S2b) and diesel vehicle (Figure 4.1a) samples have clusters of fragments with *m/z* difference of 14 ($CH_2$), linearly increasing with retention time. They are likely alkanes, where the largest fragment *m/z* equals to the MW of the compound itself (molecular ions) (Isaacman et al., 2012; Worton et al., 2014). Biomass burning (Figure S2c) has high signal in the smaller retention time range with different *m/z* patterns than gasoline vehicle, indicating potential different IVOC chemical composition. Ambient data (Figure S2d) shows signal over a wide span of retention time, but with less obvious patterns.

Figure 4.1(b) shows the raw TIC signal of a non-DPF diesel vehicle exhaust (solid line), which is the sum of all *m/z* fragments at each scan. Scan number is approximately linearly correlated with retention time (Figure S4). The large hump UCM is unclear in this representation of the data. Figure S5 shows the chromatogrram for a gasoline vehicle sample.

Figure 4.1(b) shows the PMF decomposed data matrix as a function of retention time (stacked area plot, plotting each column of the abundance matrix W and stack over the previous column). The inset in Figure 4.1(b) shows that reconstructed TIC reproduces the raw data with a correlation coefficient ($r$) = 0.99.

Each factor was linked to specific class of compounds based on comparisons with reference spectra in the NIST webbook. We select the reference compounds semi-empirically by locating the peaks in the original chromatogram and search in the NIST database using ChemStation, or proposing possible surrogate based on existing knowledge. We refer to the PMF factors by the most abundant ion. Figure 4.1(c) shows an example of this process for the PMF resolved factor (*m/z* = 57). The inset of Figure 4.1(c) shows very good agreement ($r$ = 0.997) between *m/z* = 57 factor and NIST reference mass spectra of *n*-

Dodecane. Therefore we attribute this factor to *n*-alkanes. Detailed comparison of the mass spectra between all of the PMF factors in the mobile source solution are shown in Figure S6.

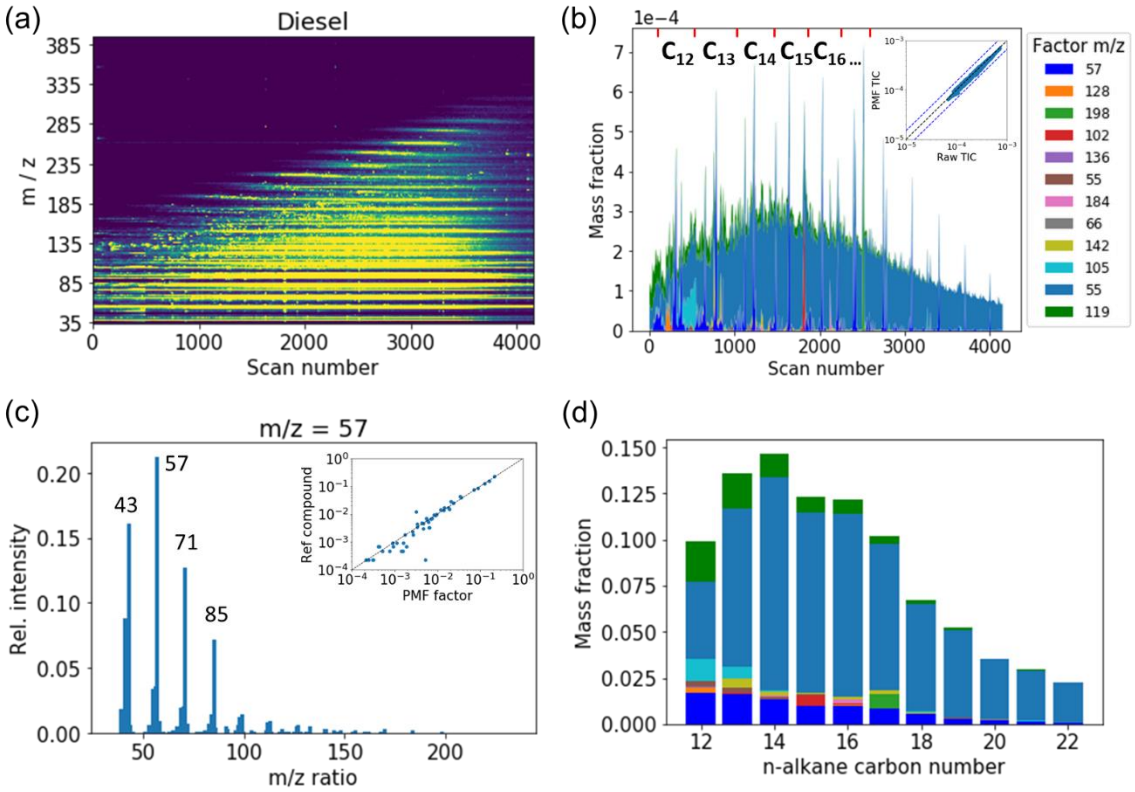



**Figure 4.1 (a) Raw GC-MS data and (b) TIC signal from non-DPF diesel vehicle exhaust (inset: scatter plot of raw and PMF reconstructed TIC signal, *r* = 0.990) (c) Example of resolved factor for *n*- alkanes (inset: scatter plot of PMF resolved factor versus reference compound mass spectra, *r* = 0.999), other factors shown in Figure S3 (d) PMF resolved non-DPF diesel vehicle exhaust decomposed into 12 PMF factors**

Integrating over the retention time bins, we get the resolved chemical composition profile in 10 PMF factors (excluding internal standards) and 11 carbon number bins. Figure 4.1(d) shows the PMF decomposed chemical composition as a function of volatility bins.

Figure S7 show the reconstruction error and max *r* as a function of number of factors in mobile source emission dataset, so we choose P = 12 to achieve best overall performance and interpretability. Table 4.2 lists the 12 PMF factors derived from the best solution for the

mobile sources identified based on their peak *m/z*.

**Table 4.2 PMF resolved factors in mobile IVOC emissions dataset**

| No. | Peak *m/z* | Group | Reference Compound | Formula |
|---|---|---|---|---|
| 1 | 57 | *n*- and *b*-alkanes | *n*-Dodecane | $C_{12}H_{26}$ |
| 2 | 55 | Cyclic alkanes | Cyclopentane, 1-methyl-2-(4-methylpentyl)-, trans- | $C_{12}H_{24}$ |
| 3 | 55 | Cyclic alkanes | Cyclopentane, 1-methyl-2-(4-methylpentyl)-, trans- | $C_{12}H_{24}$ |
| 4 | 119 | Single-ring aromatics | Benzene, 1,2,4,5-tetramethyl- | $C_{10}H_{14}$ |
| 5 | 128 | PAH | Naphthalene | $C_{10}H_{8}$ |
| 6 | 142 | PAH | Naphthalene, 1-methyl- | $C_{11}H_{10}$ |
| 7 | 102 | Oxygenates | Phenyl maleic anhydride | $C_{10}H_{6}O_{3}$ |
| 8 | 105 | Oxygenates | Benzoic acid | $C_{7}H_{6}O_{2}$ |
| 9 | 184 | Unclassified | 2,5-Cyclohexadiene-1,4-dione, 2-phenyl- | $C_{12}H_{8}O_{2}$ |
| 10 | 198 | Unclassified | 2-Methyl-5-phenyl-1,4,cyclohediene-1,4-dione | $C_{13}H_{10}O_{2}$ |
| 11 | 66 | Internal standards | *n*-Dodecane-D26 | $C_{12}D_{26}$ |
| 12 | 136 | Internal standards | Naphthalene-D8 | $C_{10}D_{8}$ |

Figure S6 shows the scatter plots of the mass spectra of all PMF resolved factors and their reference compounds in NIST database. All factors except *m/z* = 55 and *m/z* = 119 show strong agreement ($r > 0.99$) with the proposed reference compounds. The *m/z* = 55 and 119 factors appear associated with a more complicated mix of compounds. We believe that *m/z* = 55 factor is primarily cyclic alkanes while the *m/z* = 119 factor is primarily alkylated SRAs. These two classes of species have many isomers, and the factors are mixture of these compounds. Figure S8 show a few examples of IVOC cyclic alkanes and SRAs. These compounds show repeating mass fragments that are 14 apart. For example, SRAs show peak *m/z* = 91, 105, 119, 133 and cyclic alkanes show peak *m/z* = 41, 55, 69, which are also the peaks in our PMF resolved factors. Cyclic alkanes (*m/z* = 55) and SRAs (*m/z* = 119) contributes the majority of the emissions from diesel and gasoline emissions, respectively, consistent with analysis of Zhao et al. (2015, 2016) and Gentner et al. (2012).

Table 4.2 also includes two unclassified factors (*m/z* = 184 and 198). They are only found prominent in DPF-vehicle emission samples. DPF-vehicles are very low-emitting vehicles, so they are possibly from experiment artifact (Tenax reaction by-product). In this work, we listed the associated reference compounds, but group them as 'unclassified' to address the uncertainty here.

We used the same PMF technique to analyze the biomass smoke samples (which were fit seperately from mobile source). For biomass smoke, the best-fit PMF solution had 13 factors (Table 4.3). Figure S9 plots the two criteria: (1) reconstruction error and (2) max *r* between factors versus the number of factors; the 13-factor solution show lowest error and max $r < 0.6$. PMF factors that are internal standard, siloxane and solvent are not from actual emission, they are either spiked onto the samples as internal standards or possible artifacts (Coggon et al., 2018; McDonald et al., 2018), and are ignored in the subsequent anaysis.

**Table 4.3 PMF resolved factors in biomass burning IVOC emissions**

| No. | Peak *m/z* | Group | Reference compound | Formula |
|---|---|---|---|---|
| 1 | 57 | Alkanes | *n*-Dodecane | $C_{12}H_{26}$ |
| 2 | 128 | PAH | Naphthalene | $C_{10}H_8$ |
| 3 | 141 | PAH | 1H-Indene, 1-ethylidene- | $C_{11}H_{10}$ |
| 4 | 152 | PAH | Acenaphthylene | $C_{12}H_8$ |
| 5 | 155 | PAH | Naphthalene, 1,6,7-trimethyl- | $C_{13}H_{14}$ |
| 6 | 95 | Oxygenates (aliphatic) | Bicyclo[2.2.1]heptan-2-one, 1,7,7-trimethyl-, (1S)- | $C_{10}H_{16}O$ |
| 7 | 137 | Oxygenates (aromatic) | Phenol, 4-ethyl-2-methoxy- | $C_9H_{12}O_2$ |
| 8 | 138 | Oxygenates (aromatic) | Phenol, 2-methoxy-4-methyl- | $C_8H_{10}O_2$ |
| 9 | 150 | Oxygenates (aromatic) | 2-Methoxy-4-vinylphenol | $C_9H_{10}O_2$ |
| 10 | 120 | Furan | Benzofuran, 2,3-dihydro- | $C_8H_8O$ |
| 11 | 122 | N-containing | Hydrazine, 1-methyl-1-phenyl- | $C_7H_{10}N_2$ |
| 12 | 66 | Internal standard | *n*-Dodecane-D26 | $C_{12}D_{26}$ |
| 13 | 40 | Solvent | n/a | n/a |
| 14 | 91 | Solvent | Toluene | $C_7H_8$ |
| 15 | 91 | Solvent | Toluene | $C_7H_8$ |
| 16 | 73 | Siloxane | n/a | $C_{10}H_{30}O_5Si_5$ |
| 17 | 73 | Siloxane | n/a | $C_{10}H_{30}O_5Si_5$ |

Figure S10 shows the mass spectra scatter plots of resolved factor versus reference compounds from the NIST webbook. All factors show strong agreement ($r$ > 0.9) with reference mass spectra, except $m/z$ = 95 ($r$ = 0.60) and $m/z$ = 122 ($r$ = 0.85). Each of these could be from a mixture of similar compounds, and not represented by a single reference compound.

### 4.3.2 Mobile Source IVOC Emissions

Figure 4.2 summarizes the PMF analysis results of mobile source IVOC emissions. Figure 4.2(a) shows the previously reported emissions factors of total IVOCs (Lu et al., 2018; Presto et al., 2011; Tkacik et al., 2014; Zhao et al., 2015, 2016). We mulitplied the previously quantified IVOC emission profiles with PMF resolved chemical composition by each *n*-alkane bin. Figure 4.2(b) shows the overall PMF resolved chemical composition.

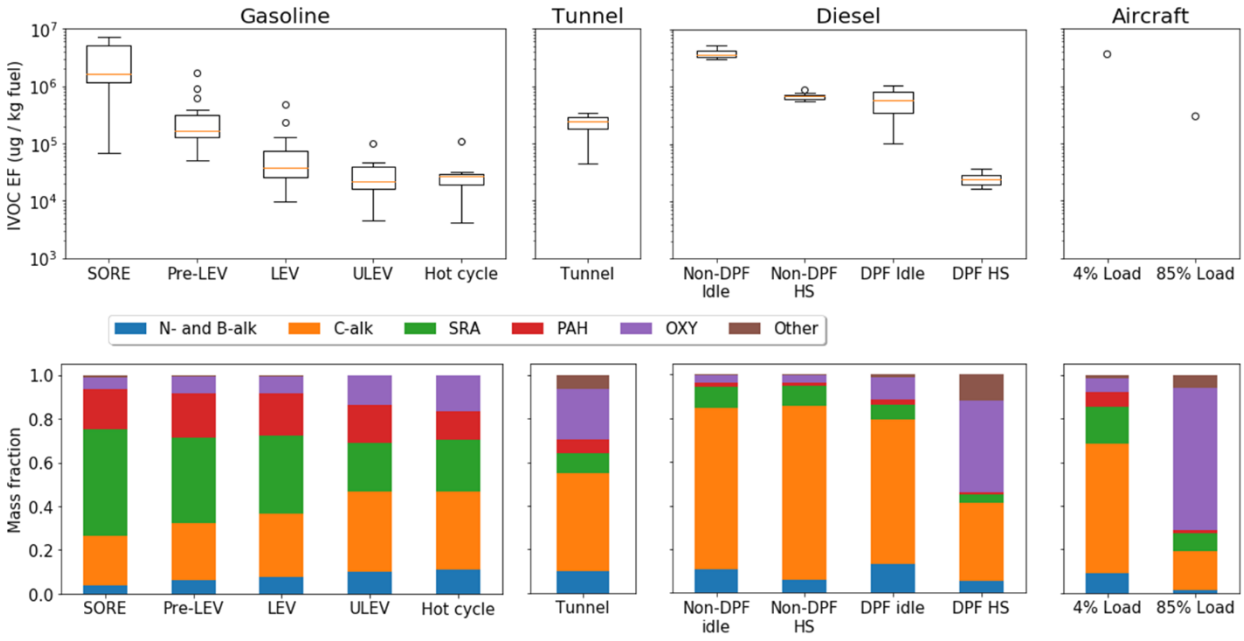



**Figure 4.2 (a) IVOC emission factors (b) PMF-resolved IVOC chemical composition in tested vehicles / engines / aircraft by certification, test cycle and load**

The PMF technique enables more detailed analysis on IVOC chemical composition,

as more than 80% of mass were only attributed to two bulk UCM groups in previous work (Presto et al., 2011; Zhao et al., 2015, 2016). With the PMF technique, we can attribute measured IVOC mass in every retention time bin to 10+ chemically resolved factors along with their reference compounds.

In Figure 4.2, we categorized the resolved PMF factors into 6 lumped groups listed in Table 4.2: (1) *n*- and *b*-alkanes (2) cyclic alkanes (3) single-ring aromatics (SRAs), (4) PAHs, (5) oxygenates and (6) internal deuterated standards. These groups are based on molecular structure of the reference compounds and to account for the effects of molecular structure on SOA yield.

Figure 4.2 shows striking differences in PMF resolved chemical composition between diesel- and gasoline-fueled sources. Gasoline-fueled source emissions show largest contribution (40% to 70%) from aromatics (SRAs and PAHs), whereas diesel-fueled source emissions are dominated (>80%) by alkanes (*n*-/*b*- and cyclic alkanes). Figure 4.3 (a to d) show PMF SRAs and PAHs dominates the smaller $C_{12}$ to $C_{14}$ bins in emission from gasoline-fueled sources, but diesel-fueled source emissions show large contribution from PMF cyclic alkanes similar to the unburnt fuel chemical composition (Gentner et al., 2012). Previously, large fraction of UCMs are classified as unspeciated cyclic compounds and its molecular structure are informed from overall mass spectra (Zhao et al., 2015, 2016). Therefore, this technique provides more insights into the IVOC chemical information, and enables more detailed PMF compositional analysis on certification standard and test cycle effects.

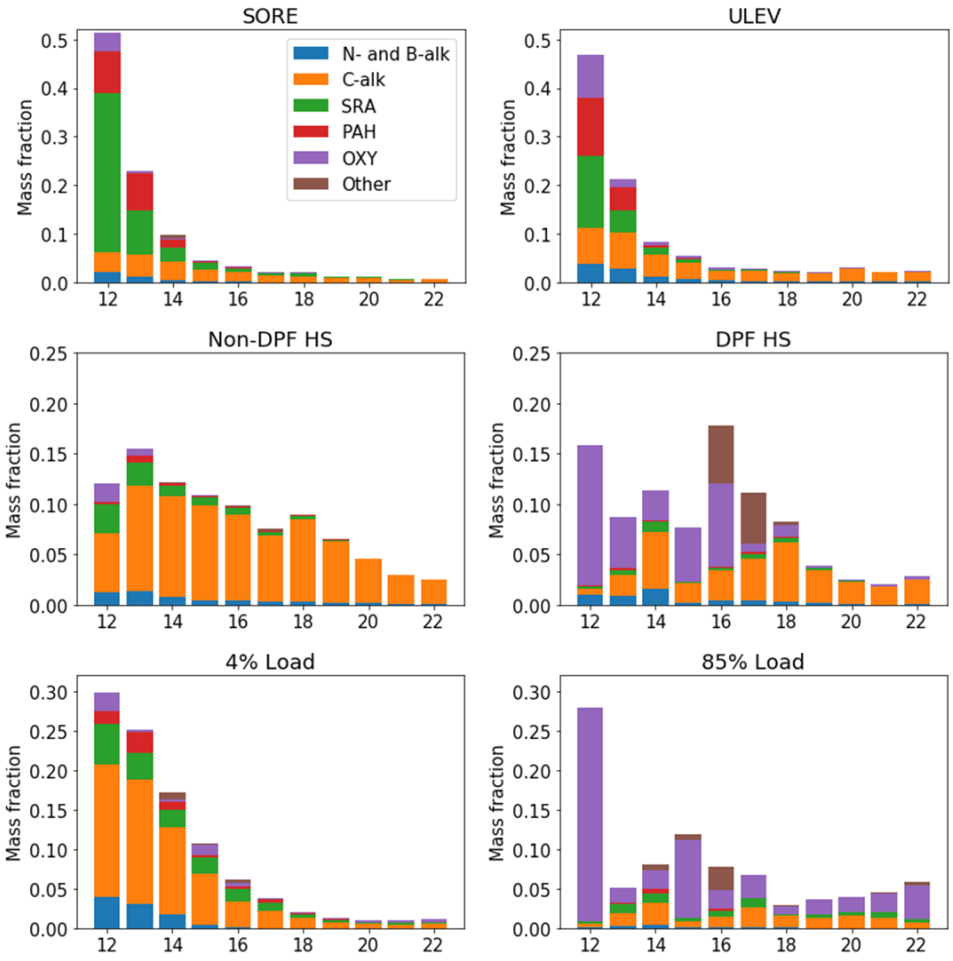



**Figure 4.3 PMF-resolved IVOC composition as a function of volatility (a, b) gasoline sources: SORE, and ULEV, (c, d) diesel mobile sources: Non-DPF diesel vehicles and DPF diesel vehicles tested at high speed cycle, (e, f) aircraft sources: emissions at 4% and 85% load**

Figure 4.2 also shows the test cycle effect on the PMF resolved chemical composition. For example, IVOC EFs tested on creep and idle cycle are about a factor of 5 to 20 higher than high speed cycle for both DPF-equipped and non-DPF vehicles. No significant PMF compositional difference is observed for non-DPF vehicle emissions, but large increase (>50%) in oxygenates are found for DPF vehicles tested under high-speed cycle. Emissions from hot cycles are similar in compsoition as ULEV tested on cold unified cycle (UC). This indicates that cycle effect is not as important for newer gasoline vehicles. Test cycles also

matter for aircraft. At idle, the exhaust chemical composition is similar to jet fuel as dominated by alkanes (70%) and aromatics (17%)(Drozd et al., 2012), except for about 10% emissions assigned to an oxygenated factor. At high load, more than 60% of measured IVOCs are oxygenated compounds. Figure 4.3 (e, f) suggest the idle exhaust volatility distribution is very similar to that of jet fuel (Presto et al., 2011), but 85% load results show more spiky volatility distribution. These trends are consistent with data from Cross et al. (2013)

Figure 4.2 shows that there are systematic trends in the PMF resolved compostion within gasoline vehicle emissions. Starting from SORE to pre-LEV, LEV and ULEV, there's a decreasing trend of the fraction of PMF SRAs (49% down to 23%) matched by an increase in the contribution of other factors (n- and b-alkanes (4% up to 10%), cyclic alkanes (23% up to 37%) and oxygenates (6% up to 14%)). The fraction of PAH remains at about 20%. We suspect that IVOC SRAs are relatively smaller compounds, they are easier to be depleted by higer combustion efficiency or catalytically removed in the engine exhaust aftertreatment.

The data for gasoline vehicle emissions are consistent with results from 2D-GC analysis (Drozd et al., 2019). Drozd et al. (2019) reported an average fraction of 30% aliphatic compounds, 30% of SRA-IVOC and 35% of PAH for IVOC emission from ULEV. Our PMF resolved profile for ULEV shows 47% alkanes, 26% SRAs, 17% PAHs, which has more alkanes (+17%) and less PAHs (-18%) comparing to the 2D-GC results. We only included napthalene-related compounds (*m/z* = 128 and 142) in the PMF PAH group, therefore the fraction is actually comparable (20% versus 17%). Drozd et al. (2019) also identified clusters of larger PAHs in 2D-GC plots, where our PMF solution might misclassify them into cyclic alkanes (*m/z* = 55). Figure 4.2 also shows the decreasing trend of SRAs with newer vehicles, and as the additional cyclic alkanes might be Gen-IVOCs identified in 2D-

GC results. Despite this modest difference, we conclude that this demonstrate the overall effectiveness of this PMF-based analysis technique.

Figure S11 shows IVOC volatiltity distributions of the mobile source emissions. As the engine and emission control technology advances with stricter regulations (from Pre-LEV to ULEV), the fraction of lower volatility IVOCs gradually increases. Hot-cycle on-road vehicle test results show the lowest volatility distribution. Similar trend in volatility distribution is found in off-road diesel emission tests under different operation modes as well (Qi et al., 2019). But as the IVOC emission factor is decreasing dramatically, so the total SOA burden is certainly falling with newer vehicles.

Figure 4.2a shows that IVOC emission factors from diesel vehicles also vary by nearly three orders of magnitude ($10^4$ to $10^7$ ug/kg fuel) with diesel particulate filter (DPF) equipped vehicles emitting a factor of 5 to 20 times lower than non-DPF vehicles.

Figure 4.2b indicates distinct PMF resolved chemical composition between DPF-equipped and non-DPF diesel vehicles. DPF-equipped vehicle emissions show 10% to 50% oxygenated IVOCs versus 3.3% for non-DPF. This trend is consistent with data from Alam et al. (2019). Figure 4.2 (b) and 4.3 show that IVOC oxygenates contribute more than 50% from $C_{12}$ to $C_{16}$ retention time bins in DPF-equipped vehicle emission, whereas more than 80% of mass are alkanes in non-DPF vehicle emission, consistent with results from Zhao et al. (2015). For non-DPF diesel vehicles, the emission volatility distribution and PMF resolved chemical composition is similar to diesel fuel (Zhao et al., 2015), but DPF vehicle emissions feature high fraction of oxygenates which could be from combustion or aftertreatment by-production.

The PMF analysis also shows how the composition of the IVOC emission varies systematically with fuel composition. The non-DPF diesel vehicle was tested with three

different fuels (low, medium and high aromatics). The PMF results in Figure S12 shows the SRA factors changing from 22%, 10% and 3% in IVOC emissions, for these fuel (May et al., 2014). This is expected and in turn validate the chemical composition information extracted by PMF technique.

### 4.3.3 IVOC emissions from biomass burning

We categorized the biomass burning PMF resolved factors into 5 groups: alkanes, PAHs, oxygenates (aliphatic and aromatic) and N-containing compounds, in a similar fashion as Hatch et al. (2017). Previously, we were unable to analyze the biomass burning IVOC data measured by 1D GC-MS due to highly complex UCM mass spectra, including artifacts such as siloxane, solvent and internal standards .

The PMF factors cover most of the groups in Hatch et al. (2017), only except terpenes, which is usually less than 5% of the IVOC mass fraction.

Figure 4.4 and Figure S13 compares the PMF results for the 11 different wildland fuels. There is substantial variation, on average about 26% alkanes, 27% PAHs, 7% furans, 19% oxygenates and 21% N-containing compounds. This shows only roughly half of the detected compounds (53%) are hydrocarbons, in contrast to more than 80% in most mobile sources. Figure S14 shows that biomass burning smoke is dominated by $C_{12}$ IVOCs. PAHs and N-containing aromatics are the dominant chemical groups (> 80%) in the $C_{12}$ to $C_{14}$ retention time bins, whereas alkanes contribute more than 80% in lower volatility bins ($C_{15}$ to $C_{22}$). Due to the partial elution of oxygenated compounds in the non-polar GC column we used for this analysis, we likely underestimate the mass of oxygenates in our samples (Zhao et al., 2014).

There are no clear trends in IVOC composition with either emission rate or fuel type,

even multiple experiments with same fuel give different compositions. This likely reflects the poorly controlled nature of biomass combustion. The IVOC EFs are in the range of 0.1 to 1 µg / kg fuel, which is about 2 times lower than the 2D-GC measured EF (0.27 to 1.8 µg / kg fuel) reported in Hatch et al. (2017). It is likely due to the partial elution of oxygenates (Zhao et al., 2014). Compared to the results of Hatch et al. (2017) , our results show somewhat lower oxygenates (19% vs 50% to 75%) but higher alkanes (26% vs 10% to 25%), PAHs (27% vs 10% to 20%) and N-containing compounds (21% vs less than 10%). The fractions of furans (5% to 15%) are similar. One Possible explanation is that Hatch et al. (2017) compiled organics measured by four instruments, of which some oxygenates were not measured by GC-MS.

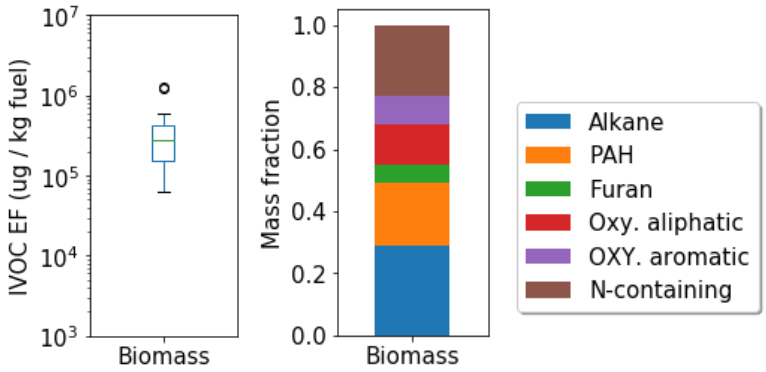



**Figure 4.4 IVOC EFs and PMF decomposed composition in biomass burning smoke samples (n=14)**

Figure S15 shows the correlation coefficient (*r*) between IVOC EFs and other pollutants. The *r* of IVOC-to-OC and IVOC-to-CO is 0.56 and 0.64, and *r* of IVOC to $CO_2$ and modified combustion efficiency (MCE) is -0.37 and -0.55, respectively. This analysis indicates that OC and CO are reasonable estimators for biomass burning IVOCs. This is expected, as in biomass burning smoke, the IVOCs are from incomplete combustion, whereas in vehicle emissions, they could be from unburnt fuel, incomplete combustion or oil components (Gentner et al., 2012; Lu et al., 2018; Worton et al., 2014).

### 4.3.4 Source apportionment IVOCs measured in a highway tunnel

We also applied PMF analysis to highway tunnel samples to estimate the relative contribution of gasoline and diesel vehicles to IVOC emissions. This was done using a two-step process. First, we ran PMF analysis on tunnel samples concatenated with vehicle emission samples (similar to Figure S2) to resolve chemical composition into a consistent set of factors. Second, we perform source apportionment analysis by chemical mass balance (CMB) regression method to apportion the PMF factors to sources,

$$T_i = \beta_1 \times G + \beta_2 \times D$$

where $T_i$ is the PMF factor profile from $i_{\text{th}}$ tunnel sample, G and D are the PMF resolved profiles for gasoline and diesel vehicle emissions and $\beta_1$ and $\beta_2$ are the relative contributions of gasoline and diesel vehicles, respectively. All factor profiles are 44-dimension vectors (11 carbon number bins times 4 PMF factors), and the sum of all components in each vector are normalized to 1. We only include hydrocarbon PMF factors (alkanes, SRAs and PAHs) in the CMB calculation. This assumes that the PMF factor with oxygenates is from the background air. We want to calculate the ratio of $\beta_2 : \beta_1$, which is the diesel-to-gasoline ratio in IVOC emissions. Because each tunnel sample was taken at different times with varying fleet composition ($\beta_1$ and $\beta_2$). We ran the CMB model separately for each sample using the sample PMF factor source profiles.

We choose a 15-factor PMF solution, because it best approximates the solution in mobile sources dataset with max $r < 0.5$. Table S2 lists all the resolved factors, peak *m/z*, and their reference compounds. Similarly to mobile sources, PMF factors are grouped into 6 groups: n- and b-alkanes, cyclic alkanes, SRAs, PAHs and oxygenates and 2 other groups for internal standard and unclassified compounds. Previously, we were unable to resolve the

tunnel IVOC chemical composition measured by 1D GC-MS due to the UCMs of both hydrocarbon and oxygenated compounds.

Figure 4.2 shows that both IVOC EFs and PMF resolved composition from tunnel data fall between the data from gasoline and diesel vehicle source profiles. This is expected as tunnel IVOC concentrations are dominated by emissions from these two sources. The tunnel sample has a higher fraction of oxygenated compared to the vehicle samples (except for DPF-diesel), indicating more contribution from oxidation products in ambient background.

Figure 4.5(a) shows the PMF factors from a typical tunnel sample. There is a strong signal of the SRA PMF factor in $C_{12}$ to $C_{14}$ *n*-alkane bins and high fraction of PMF cyclic alkanes factor in $C_{17}$ to $C_{22}$ *n*-alkane bins, indicating substantial contributions from both gasoline and diesel vehicular IVOCs.

Figure 4.5(b) shows the scatter plot between CMB fitted mixture PMF profile (gasoline + diesel) and GC-MS measured tunnel PMF profile. The $R^2$ of the linear regression ranges from 0.73 to 0.88 for all tunnel samples, indicating that the CMB model explain more than 80% of the variation in the primary IVOC PM factors in the tunnel samples. The average ratio of diesel-to-gasoline IVOCs ($\beta_2 : \beta_1$) for all samples is 2.1 (ranging from 1.4 to 2.6). Therefore, although diesel vehicles only used 20% of fuel (Tkacik et al., 2014), they are estimated to contribute around 70% of the primary IVOCs in tunnel samples.

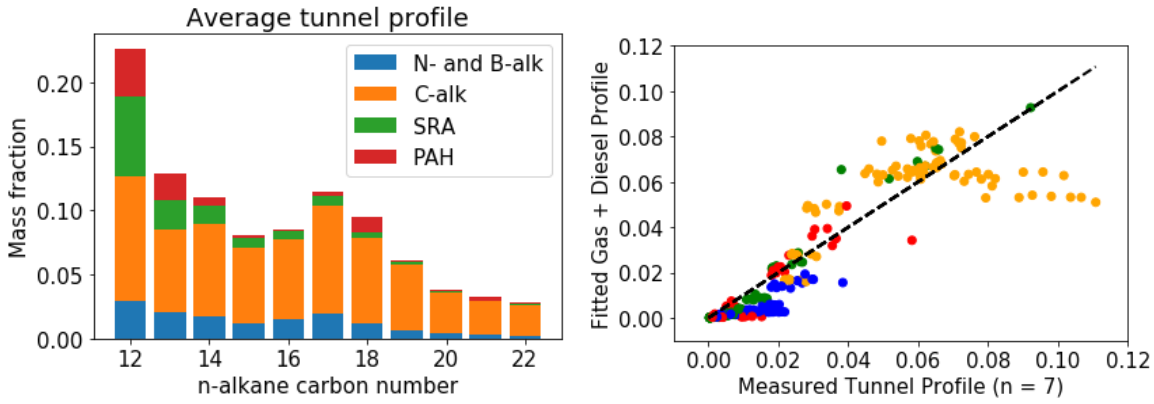



**Figure 4.5 (a) Volatility distribution and primary PMF factors in a typical tunnel emission sample (b) Scatter plot of mixture of gasoline and diesel profiles versus tunnel profile**

### 4.3.5 IVOC composition in ambient samples in Pasadena, CA

To investigate the sources of ambient IVOCs, we performed PMF analysis on ambient samples collected in Pasadena, CA during the CalNex campaign. The PMF model is trained with ambient samples concatenated with vehicle emission samples (similar to Figure S2). We used 16 mobile source samples (4151 scans per sample) and 48 ambient samples (1336 scans per sample), which is to keep similar number of mass spectra scans in both datasets. The concatenation of sources emissions and ambient samples is to simultaneously resolve hydrocarbon and oxygenated IVOC factors for source apportionment analysis. A 17-factor PMF solution was most interpretable, including similar hydrocarbon factors as in mobile-source, and extra oxygenate factors. Table S3 lists the resolved factors, peak *m/z*, and reference compounds. Previously, IVOC UCMs were only attributed to two subgroups (Zhao et al., 2014). We grouped the PMF factors into 6 subgroups: alkanes, SRAs, PAHs, oxygenates, acids, internal standards, and unidentified. Ten out of 17 resolved factors are oxygenates. Table S3 shows all oxygenated factors have more than 2 oxygen atoms.

Figure 4.6(a) shows the PMF resolved chemical composition of a typical ambient

sample. Measured IVOCs mass is dominated by oxygenated factors in the $C_{13}$ and $C_{15}$ bins, 3 times more than hydrocarbon factors. The hydrocarbon PMF factors decreases from $C_{12}$ to $C_{22}$ (Zhao et al., 2014), which suggests that diesel emissions may not be the dominant hydrocarbon IVOC sources in Pasadena, CA (Ensberg et al., 2014; Lu et al., 2019; McDonald et al., 2018; Zhao et al., 2014). We observed much higher fraction of oxygenate factor from $C_{17}$ to $C_{22}$ bins than in mobile-source emissions samples. The reference compounds of resolved oxygenated factors are all oxygenates with one or more rings (Table S3), indicating complex atmospheric oxidation products originated from primary SRA or PAH emissions.

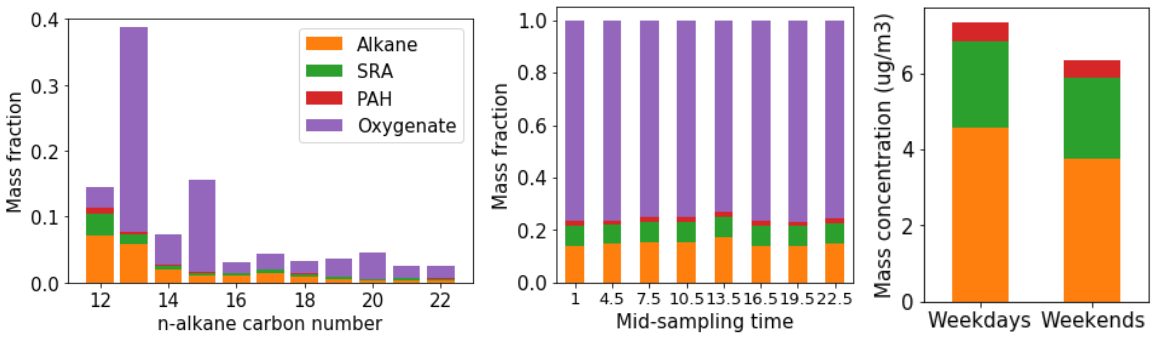



**Figure 4.6 (a) Volatility distribution of PMF factors of typical ambient sample (b) Average chemical composition diurnal pattern (c) Average weekday-weekend hydrocarbon IVOC concentrations. We removed the benzoic acid factors and unidentified group, as it could be the Tenax reaction by-product (Zhao et al., 2014).**

On average, oxygenated PMF factor contributes 75% measured IVOC mass, with only around 25% are hydrocarbon factors (alkanes, SRAs and PAHs). This likely overestimate the IVOC oxygenates, because some identified oxygenate factor could also be reaction byproduct with Tenax. Detailed bottom-up analysis is needed to examine the factors to determine if they are artifacts or actual sampled pollutants. Since PMF analysis on mobile-

source emissions and highway tunnel samples are dominated (more than 80% and 70%, respectively) by hydrocarbon factors, we expect some primary emissions of oxygenated species in ambient (Khare and Gentner, 2018; McDonald et al., 2018; Qin et al., 2019; Shah et al., 2020) (for example, volatile chemical products, VCPs) and secondary species formed from atmospheric oxidation (Lu et al., 2019). Figure S16 shows the diurnal pattern of IVOC mass concentration. Despite the dilution effect by planetary boundary layer (PBL) (Lu et al., 2019), the hydrocarbon PMF factors slightly increased at 1:30pm (Zhao et al., 2014), but the oxygenated factors decreased. Therefore, oxygenated factors are likely the combination of primary emission and secondary oxidation product, but dominated by the latter.

Figure 4.6(b) shows hydrocarbon IVOCs has minor peak (+3%) in composition only in the middle of the day at 1:30 pm (Zhao et al., 2014). This leads us to two hypotheses, which are: (1) regional transport of hydrocarbon IVOCs emitted in the morning rush hour or (2) fresh local evaporative IVOC emissions around noon. But transported hydrocarbon IVOCs would have been most oxidized if they reach Pasadena around 1:30 pm. Also, due to the strong photo-oxidation around noon, we would expect more oxygenated than hydrocarbon factors (Hayes et al., 2013). Therefore, the most likely explanation is that there should be some unknown evaporative sources peaked at noon with more (>50%) hydrocarbon than oxygenates emissions (Lu et al., 2019). This does not contradict with the assertion that most IVOCs are oxygenated and secondary, but only to propose unknown sources to explain the early-afternoon peak of hydrocarbon IVOCs.

Figure 4.6(c) shows that the weekday-to-weekend pattern of hydrocarbon IVOCs. There have been a series of studies on the mobile-source contribution to SOA formation in Los Angeles area (Ensberg et al., 2014; Gentner et al., 2012; Jathar et al., 2017; Lu et al., 2019). Diesel vehicle emissions were once hypothesized as important sources, our PMF

results show higher alkane factors (+0.8 μg/$m^3$) on weekdays compared to weekends. Considering the less than half diesel vehicle activities on weekends (Bahreini et al., 2012), we can estimate the weekday diesel-related IVOCs are 1.6 to 2 μg/$m^3$. Recent simulation study also shows that mobile sources only explain about 40% of measured hydrocarbon IVOCs (Lu et al., 2019), of which diesel-related sources contribution is around 2 μg / $m^3$ (Zhao et al., 2014). Therefore, our PMF-based analysis shows similar diesel-related contribution as the bottom-up chemical transport model simulation. We conclude there should be important non-mobile sources of hydrocarbon IVOC emissions, without weekday-weekend pattern. Figure S17 shows that there is no significant weekday-to-weekend pattern in PMF resolved total IVOC chemical composition(Zhao et al., 2014).

### 4.4 Implications and future work

Knowing the chemical composition and molecular structure of IVOCs are cruicial to model their SOA formation in simulations, because the SOA yield depends on volatility and molecular structure (Chacon-Madrid et al., 2010; Loza et al., 2014; Presto et al., 2010). SOA yields increase with lower volatility IVOCs (Presto et al., 2010), but for a given carbon number, the SOA yield for hydrocarbon IVOCs generally follows aromatics > cyclic > linear > branched alkanes (Lim and Ziemann, 2009b; Loza et al., 2014; Tkacik et al., 2012). Aromatics have higher SOA yield under low-$NO_x$ conditions (Chan et al., 2009; Ng et al., 2007). To date, most modelling studies have not incorporated IVOC chemical composition in simulations.

We will explore the importance of these factors with the PMF resolved composition. The SOA yield for each PMF group is assigned using the surrogate listed in Table 4.2. Namely, *n*-alkanes and PAHs have corresponding SOA yield data for their reference

compounds (Chan et al., 2009; Presto et al., 2010). For cyclic alkane factor, we use alkyl-cycloalkanes as the reference compound, and chamber results on $C_{12}$ alkanes show their SOA yield is similar to *n*-alkanes under high-$NO_x$ conditions (Loza et al., 2014; Yee et al., 2013). For SRA factor, we use the SOA yield for tri-methyl benzene for $C_{12}$ bin, because Drozd et al. (2019) shows most SRAs in gasoline vehicle emissions have two or more branches. Oxygenate factors are modelled using aldehydes as surrogate, and their SOA yield are assumed as 1/3 of *n*-alkanes (Chacon-Madrid et al., 2010; Chacon-Madrid and Donahue, 2011). Chamber experiment results of alkanes, SRAs, PAHs under low-$NO_x$ conditions are also used to inform the SOA parameterization (Chen et al., 2016; Li et al., 2016; Yee et al., 2013). Detailed information can be found in the Table S4.

Figure 4.7 indicates that the predicted the mass-based SOA yield increase significantly (+80%) in gasoline vehicle emissions from 0.17 to 0.3 between high- and low-$NO_x$ cases. The SOA precursors in gasoline vehicle exhaust are dominated by SRAs and PAHs (Section 3.2), whose SOA yields are strongly $NO_x$ dependent. Therefore, the difference in predicted SOA yield between high- and low-$NO_x$ conditions due to the SRAs and PAHs highlights the need to include IVOC-aromatic groups in chemical transport model studies (Akherati et al., 2018; Lu et al., 2019).

A second question is whether the compositional differences between different classes of gasoline vehicles shown in Figure 4.2 affect SOA formation. Figure S18 show less than ± 20% changes on SOA yield within subclasses in gasoline vehicle sources. This is much less than the three orders of magnitude variation in IVOC EFs.. This supports that a single profile can be used to represent SOA formation from gasoline vehicle IVOC emissions (Lu et al., 2018).

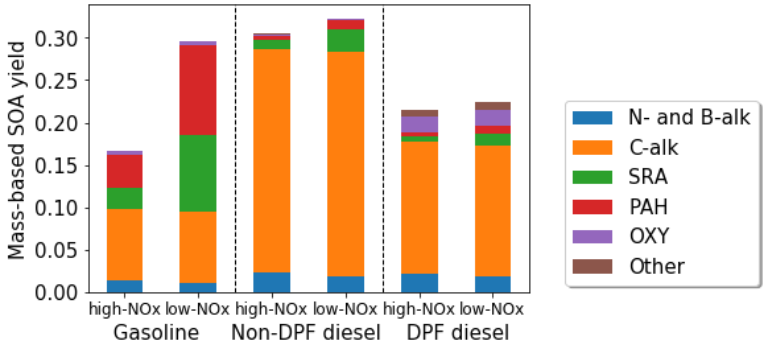



**Figure 4.7 Comparison of modelled SOA yield for moibile sources under high- and low-$NO_x$ conditions**

Figure 4.7 also shows the mass-based SOA yield from diesel vehicle IVOC emissions. Under both high- and low- $NO_x$ conditions, the SOA yield of non-DPF diesel vehicle emissions range is around 0.3, because of its lower volatility distribution that is dominated by alkanes. The SOA yield of alkanes is less sensitive to $NO_x$ conditions than aromatics (Loza et al., 2014). Due to much higher oxygenates fraction (lower SOA yield), we predict that the for DPF diesel is lower than for non-DPF diesel.

To summarize, we proposed a PMF-based technique to resolve the chemical composition of the IVOC UCM measured in traditional 1-D GC-MS. Prominent PMF factors can be automatically found and mass spectra comparison to reference compounds in NIST show high confidence ($r$ > 0.9) for most factors. The results can provide similar level of speciation information that is in more sophisticated 2D-GC results. Therefore, it can be applied to both archival and future GC-MS data analysis.

The key of PMF technique is the underlying structure in UCM mass spectra. In nature, the mass spectra detected by GC-MS are linear separable, and the good thing is the individual components are often distinct enough (thinking of alkanes, PAHs and oxygenates). Because of the strong fragmentation induced by electron ionization, the mass fragments provide unique signal (including peak *m/z* and -$CH_2$ patterns) for recovery of molecular information.

We also admit that PMF works the best for mixture of intrinsically different components. Koss et al. (2020) showed that PMF is not suitable for chamber experiment data analysis, which is a dynamic and evolving system, of which the mass spectra are slowing changing and not distinct enough for separation.

The PMF-based technique that show promising results. But some resolved factors, usually 2 to 3, are less correlated ($r < 0.9$) to the reference compounds in NIST database, for example, $m/z = 55$ (cyclic alkane) and $m/z = 119$ (SRAs) factors. These factors are likely mixture of multiple compounds and have less clean mass spectra. Therefore, we recommend future research on how to speciate such factors into more detailed factors. There are two possible directions to address this issue: (1) split the entire data matrix into multiple slides according to retention time, and run PMF separately on these sub-matrixes. But this may lead to different number of factors for each sub-matrix; or (2) apply regularization term or sparsity constrain on the PMF algorithm, such that mixture factors are penalized more than regular factors, so as to get cleaner mass spectra solutions.

### 4.5 Acknowledgments

This publication was developed as part of the Center for Air, Climate, and Energy Solutions (CACES), which is supported under Assistance Agreement No. RD83587301 awarded by the U.S. Environmental Protection Agency. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication. The authors thank Benjamin Murphy and Jian Zhao for helpful discussions.

## 4.6 References


Akherati, A., Cappa, C. D., Kleeman, M. J., Docherty, K. S., Jimenez, J. L., Griffith, S. M., Dusanter, S., Stevens, P. S. and Jathar, S. H.: Simulating secondary organic aerosol in a regional air quality model using the statistical oxidation model - Part 3: Assessing the influence of semi-volatile and intermediate volatility organic compounds and NOX, Atmos. Chem. Phys. Discuss., 2018, 1–44, doi:10.5194/acp-2018-616, 2018.

Alam, M. S., Zeraati-Rezaei, S., Xu, H. and Harrison, R. M.: Characterization of Gas and Particulate Phase Organic Emissions (C9–C37) from a Diesel Engine and the Effect of Abatement Devices, Environ. Sci. Technol., 53(19), 11345–11352, doi:10.1021/acs.est.9b03053, 2019.

Apte, J. S., Brauer, M., Cohen, A. J., Ezzati, M. and Pope, C. A.: Ambient PM2.5 Reduces Global and Regional Life Expectancy, Environ. Sci. Technol. Lett., 5(9), 546–551, doi:10.1021/acs.estlett.8b00360, 2018.

Bahreini, R., Middlebrook, A. M., De Gouw, J. A., Warneke, C., Trainer, M., Brock, C. A., Stark, H., Brown, S. S., Dube, W. P., Gilman, J. B., Hall, K., Holloway, J. S., Kuster, W. C., Perring, A. E., Prevot, A. S. H., Schwarz, J. P., Spackman, J. R., Szidat, S., Wagner, N. L., Weber, R. J., Zotter, P. and Parrish, D. D.: Gasoline emissions dominate over diesel in formation of secondary organic aerosol mass, Geophys. Res. Lett., 39(6), 2–7, doi:10.1029/2011GL050718, 2012.

Cai, S., Zhu, L., Wang, S., Wisthaler, A., Li, Q., Jiang, J. and Hao, J.: Time-Resolved Intermediate-Volatility and Semivolatile Organic Compound Emissions from Household Coal Combustion in Northern China, Environ. Sci. Technol., 53(15), 9269–9278, doi:10.1021/acs.est.9b00734, 2019.

Chacon-Madrid, H. J. and Donahue, N. M.: Fragmentation vs. functionalization: chemical aging and organic aerosol formation, Atmos. Chem. Phys., 11(20), 10553–10563, doi:10.5194/acp-11-10553-2011, 2011.

Chacon-Madrid, H. J., Presto, A. A. and Donahue, N. M.: Functionalization vs. fragmentation: n-aldehyde oxidation mechanisms and secondary organic aerosol formation, Phys. Chem. Chem. Phys., 12(42), 13975–13982, doi:10.1039/C0CP00200C, 2010.

Chan, A. W. H., Kautzman, K. E., Chhabra, P. S., Surratt, J. D., Chan, M. N., Crounse, J. D., Kürten, A., Wennberg, P. O., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol formation from photooxidation of naphthalene and alkylnaphthalenes: Implications for oxidation of intermediate volatility organic compounds (IVOCs), Atmos. Chem. Phys., 9(9), 3049–3060, doi:10.5194/acp-9-3049-2009, 2009.

Chen, C.-L., Kacarab, M., Tang, P. and Cocker, D. R.: SOA formation from naphthalene, 1-methylnaphthalene, and 2-methylnaphthalene photooxidation, Atmos. Environ., 131, 424–433, doi:https://doi.org/10.1016/j.atmosenv.2016.02.007, 2016.

Chow, J., Bachmann, J., Hsu, Y.-C., Chen, S.-K., Tsai, J.-H., Chiang, H.-L., Fujita, E., Zielinska, B., Campbell, D., Arnott, W., Sagebiel,Lynn, J., Gabele, P., Crews, W., Snow, R., Clark, N., Wayne, W., Lawson, D., Arnott, W., Brown, S., Frankel, A., Raffuse, S., Roberts, P., Hafner, H., Anderson, D., Eklund, B., Simon, M., Regmi, S., Ongwandee, M., Morrison, G., Fitch, M. and Surampalli, R.: Variations in Speciated Emissions from Spark-Ignition and Compression-Ignition Motor Vehicles in California's South Coast Air Basin, J. Air Waste Manage. Assoc., 57(6), 705–720, doi:10.3155/1047-3289.57.6.705, 2007.

Coggon, M. M., McDonald, B. C., Vlasenko, A., Veres, P. R., Bernard, F., Koss, A. R., Yuan, B., Gilman, J. B., Peischl, J., Aikin, K. C., DuRant, J., Warneke, C., Li, S.-M. and de Gouw, J. A.: Diurnal Variability and Emission Pattern of Decamethylcyclopentasiloxane (D5) from the Application of Personal Care Products in Two North American Cities, Environ. Sci. Technol., 52(10), 5610–5618, doi:10.1021/acs.est.8b00506, 2018.

Cross, E. S., Hunter, J. F., Carrasquillo, A. J., Franklin, J. P., Herndon, S. C., Jayne, J. T., Worsnop, D. R., Miake-Lye, R. C. and Kroll, J. H.: Online measurements of the emissions of intermediate-volatility and semi-volatile organic compounds from aircraft, Atmos. Chem. Phys., 13(15), 7845–7858, doi:10.5194/acp-13-7845-2013, 2013.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F. and Schwartz, J. D.: Air Pollution and Mortality in the Medicare Population, N. Engl. J. Med., 376(26), 2513–2522, doi:10.1056/NEJMoa1702747, 2017.

Drozd, G. T., Miracolo, M. A., Presto, A. A., Lipsky, E. M., Riemer, D. D., Corporan, E. and Robinson, A. L.: Particulate Matter and Organic Vapor Emissions from a Helicopter Engine Operating on Petroleum and Fischer–Tropsch Fuels, Energy & Fuels, 26(8), 4756–4766, doi:10.1021/ef300651t, 2012.

Drozd, G. T., Zhao, Y., Saliba, G., Frodin, B., Maddox, C., Oliver Chang, M.-C., Maldonado, H., Sardar, S., Weber, R. J., Robinson, A. L. and Goldstein, A. H.: Detailed Speciation of Intermediate Volatility and Semivolatile Organic Compound Emissions from Gasoline Vehicles: Effects of Cold-Starts and Implications for Secondary Organic Aerosol Formation, Environ. Sci. Technol., 53(3), 1706–1714, doi:10.1021/acs.est.8b05600, 2019.

Ensberg, J. J., Hayes, P. L., Jimenez, J. L., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Holloway, J. S., Gordon, T. D., Jathar, S., Robinson, A. L. and Seinfeld, J. H.: Emission factor ratios, SOA mass yields, and the impact of vehicular emissions on SOA formation, Atmos. Chem. Phys., 14(5), 2383–2397, doi:10.5194/acp-14-2383-2014, 2014.

Fortenberry, C. F., Walker, M. J., Zhang, Y., Mitroo, D., Brune, W. H. and Williams, B. J.: Bulk and molecular-level characterization of laboratory-aged biomass burning organic aerosol from oak leaf and heartwood fuels, Atmos. Chem. Phys., 18(3), 2199–2224, 2018.

Gao, Y., Walker, M. J., Barrett, J. A., Hosseinaei, O., Harper, D. P., Ford, P. C., Williams, B. J. and Foston, M. B.: Analysis of gas chromatography/mass spectrometry data for catalytic lignin depolymerization using positive matrix factorization, Green Chem., 20(18), 4366–4377, doi:10.1039/C8GC01474D, 2018.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A. and Goldstein, A. H.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, Proc. Natl. Acad. Sci., 109(45), 18318–18323, doi:10.1073/pnas.1212272109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1212272109, 2012.

Goldstein, A. H. and Galbally, I. E.: Known and Unexplored Organic Constituents in the Earth's Atmosphere, Environ. Sci. Technol., 41(5), 1514–1521, doi:10.1021/es072476p, 2007.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J. and Barsanti, K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors, Atmos. Chem. Phys., 17(2), 1471–1489, doi:10.5194/acp-17-1471-2017, 2017.

Hatch, L. E., Rivas-Ubach, A., Jen, C. N., Lipton, M., Goldstein, A. H. and Barsanti, K. C.: Measurements of I/SVOCs in biomass-burning smoke using solid-phase extraction disks and two-dimensional gas chromatography, Atmos. Chem. Phys., 18(24), 17801–17817, doi:10.5194/acp-18-17801-2018, 2018.

Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W., Toohey, D. W., Flynn, J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglück, B., Taylor, J. W., Allan, J. D., Holloway, J. S., Gilman, J. B., Kuster, W. C., De Gouw, J. A., Massoli, P., Zhang, X., Liu, J., Weber, R. J., Corrigan, A. L., Russell, L. M., Isaacman, G., Worton, D. R., Kreisberg, N. M.,

Goldstein, A. H., Thalman, R., Waxman, E. M., Volkamer, R., Lin, Y. H., Surratt, J. D., Kleindienst, T. E., Offenberg, J. H., Dusanter, S., Griffith, S., Stevens, P. S., Brioude, J., Angevine, W. M. and Jimenez, J. L.: Organic aerosol composition and sources in Pasadena, California, during the 2010 CalNex campaign, J. Geophys. Res. Atmos., 118(16), 9233–9257, doi:10.1002/jgrd.50530, 2013.

Hayes, P. L., Carlton, a. G., Baker, K. R., Ahmadov, R., Washenfelder, R. A., Alvarez, S., Rappenglück, B., Gilman, J. B., Kuster, W. C., de Gouw, J. A., Zotter, P., Prévôt, a. S. H., Szidat, S., Kleindienst, T. E., Offenberg, J. H., Ma, P. K. and Jimenez, J. L.: Modeling the formation and aging of secondary organic aerosols in Los Angeles during CalNex 2010, Atmos. Chem. Phys., 15(10), 5773–5801, doi:10.5194/acp-15-5773-2015, 2015.

Hennigan, C. J., Miracolo, M. A., Engelhart, G. J., May, A. A., Presto, A. A., Lee, T., Sullivan, A. P., McMeeking, G. R., Coe, H., Wold, C. E., Hao, W.-M., Gilman, J. B., Kuster, W. C., de Gouw, J., Schichtel, B. A., Collett Jr., J. L., Kreidenweis, S. M. and Robinson, A. L.: Chemical and physical transformations of organic aerosol from the photo-oxidation of open biomass burning emissions in an environmental chamber, Atmos. Chem. Phys., 11(15), 7669–7686, doi:10.5194/acp-11-7669-2011, 2011.

Hodzic, A., Jimenez, J. L., Madronich, S., Canagaratna, M. R., Decarlo, P. F., Kleinman, L. and Fast, J.: Modeling organic aerosols in a megacity: Potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation, Atmos. Chem. Phys., 10(12), 5491–5514, doi:10.5194/acp-10-5491-2010, 2010.

Hunter, J. F., Day, D. A., Palm, B. B., Yatavelli, R. L. N., Chan, A. W. H., Kaser, L., Cappellin, L., Hayes, P. L., Cross, E. S., Carrasquillo, A. J., Campuzano-Jost, P., Stark, H., Zhao, Y., Hohaus, T., Smith, J. N., Hansel, A., Karl, T., Goldstein, A. H., Guenther, A., Worsnop, D. R., Thornton, J. A., Heald, C. L., Jimenez, J. L. and Kroll, J. H.: Comprehensive characterization of atmospheric organic carbon at a forested site, Nat. Geosci., 10, 748 [online] Available from: http://dx.doi.org/10.1038/ngeo3018, 2017.

Isaacman, G., Wilson, K. R., Chan, A. W. H., Worton, D. R., Kimmel, J. R., Nah, T., Hohaus, T., Gonin, M., Kroll, J. H., Worsnop, D. R. and Goldstein, A. H.: Improved resolution of hydrocarbon structures and constitutional isomers in complex mixtures using gas chromatography-vacuum ultraviolet-mass spectrometry, Anal. Chem., 84(5), 2335–2342, doi:10.1021/ac2030464, 2012.

Jathar, S. H., Miracolo, M. A., Tkacik, D. S., Donahue, N. M., Adams, P. J. and Robinson, A. L.: Secondary organic aerosol formation from photo-oxidation of unburned fuel: experimental results and implications for aerosol formation from combustion emissions., Environ. Sci. Technol., 47(22), 12886–93, doi:10.1021/es403445q, 2013.

Jathar, S. H., Woody, M., Pye, H. O. T., Baker, K. R. and Robinson, A. L.: Chemical transport model simulations of organic aerosol in southern California: model evaluation and gasoline and diesel source contributions, Atmos. Chem. Phys., 17(6), 4305–4318, doi:10.5194/acp-17-4305-2017, 2017.

Khare, P. and Gentner, D. R.: Considering the future of anthropogenic gas-phase organic compound emissions and the increasing influence of non-combustion sources on urban air quality, Atmos. Chem. Phys., 18(8), 5391–5413, doi:10.5194/acp-18-5391-2018, 2018.

Khare, P., Marcotte, A., Sheu, R., Walsh, A. N., Ditto, J. C. and Gentner, D. R.: Advances in offline approaches for trace measurements of complex organic compound mixtures via soft ionization and high-resolution tandem mass spectrometry, J. Chromatogr. A, 1598, 163–174, doi:https://doi.org/10.1016/j.chroma.2019.03.037, 2019.

Koss, A. R., Canagaratna, M. R., Zaytsev, A., Krechmer, J. E., Breitenlechner, M., Nihill, K. J., Lim, C. Y., Rowe, J. C., Roscioli, J. R., Keutsch, F. N. and Kroll, J. H.: Dimensionality-reduction techniques for complex mass spectrometric datasets: application to laboratory atmospheric organic oxidation experiments, Atmos. Chem. Phys., 20(2), 1021–1041, doi:10.5194/acp-20-1021-2020,

2020.

Kreisberg, N. M., Worton, D. R., Zhao, Y., Isaacman, G., Goldstein, A. H. and Hering, S. V.: Development of an automated high-temperature valveless injection system for online gas chromatography, Atmos. Meas. Tech., 7(12), 4431–4444, doi:10.5194/amt-7-4431-2014, 2014.

Lee, D. D. and Seung, H. S.: Algorithms for non-negative matrix factorization, in Advances in neural information processing systems, pp. 556–562., 2001.

Li, L., Tang, P., Nakao, S. and Cocker III, D. R.: Impact of molecular structure on secondary organic aerosol formation from aromatic hydrocarbon photooxidation under low-NOx conditions, Atmos. Chem. Phys., 16(17), 10793–10808, doi:10.5194/acp-16-10793-2016, 2016.

Lim, Y. B. and Ziemann, P. J.: Effects of Molecular Structure on Aerosol Yields from OH Radical-Initiated Reactions of Linear, Branched, and Cyclic Alkanes in the Presence of NOx, Environ. Sci. Technol., 43(7), 2328–2334, doi:10.1021/es803389s, 2009a.

Lim, Y. Bin and Ziemann, P. J.: Chemistry of Secondary Organic Aerosol Formation from OH Radical-Initiated Reactions of Linear, Branched, and Cyclic Alkanes in the Presence of NO x , Aerosol Sci. Technol., 43(6), 604–619, doi:10.1080/02786820902802567, 2009b.

Loza, C. L., Craven, J. S., Yee, L. D., Coggon, M. M., Schwantes, R. H., Shiraiwa, M., Zhang, X., Schilling, K. A., Ng, N. L., Canagaratna, M. R., Ziemann, P. J., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol yields of 12-carbon alkanes, Atmos. Chem. Phys., 14(3), 1423–1439, doi:10.5194/acp-14-1423-2014, 2014.

Lu, Q., Zhao, Y. and Robinson, A. L.: Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions, Atmos. Chem. Phys. Discuss., 2018, 1–28, doi:10.5194/acp-2018-752, 2018.

Lu, Q., Murphy, B. N., Qin, M., Adams, P. J., Zhao, Y., Pye, H. O. T., Efstathiou, C., Allen, C. and Robinson, A. L.: Simulation of organic aerosol formation during the CalNex study: updated mobile emissions and simplified secondary organic aerosol parameterization for intermediate volatility organic compounds, Atmos. Chem. Phys. Discuss., 2019, 1–36, doi:10.5194/acp-2019-986, 2019.

May, A. A., Nguyen, N. T., Presto, A. A., Gordon, T. D., Lipsky, E. M., Karve, M., Gutierrez, A., Robertson, W. H., Zhang, M., Brandow, C., Chang, O., Chen, S., Cicero-Fernandez, P., Dinkins, L., Fuentes, M., Huang, S. M., Ling, R., Long, J., Maddox, C., Massetti, J., McCauley, E., Miguel, A., Na, K., Ong, R., Pang, Y., Rieger, P., Sax, T., Truong, T., Vo, T., Chattopadhyay, S., Maldonado, H., Maricq, M. M. and Robinson, A. L.: Gas- and particle-phase primary emissions from in-use, on-road gasoline and diesel vehicles, Atmos. Environ., 88, 247–260, doi:10.1016/j.atmosenv.2014.01.046, 2014.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B. and Trainer, M.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science (80-. )., 359(6377), 760–764, doi:10.1126/science.aaq0524, 2018.

Murphy, B. N. and Pandis, S. N.: Simulating the formation of semivolatile primary and secondary organic aerosol in a regional chemical transport model, Environ. Sci. Technol., 43(13), 4722–4728, doi:10.1021/es803168a, 2009.

Ng, N. L., Kroll, J. H., Chan, a W. H., Chhabra, P. S., Flagan, R. C. and Seinfeld, J. H.: Secondary organic aerosol formation from m-xylene, toluene, and benzene, Atmos. Chem. Phys., 7(14), 3909–3922, doi:10.5194/acp-7-3909-2007, 2007.

Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5(2), 111–126,

doi:10.1002/env.3170050203, 1994.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R. and Dubourg, V.: Scikit-learn: Machine learning in Python, J. Mach. Learn. Res., 12(Oct), 2825–2830, 2011.

Presto, A. a., Hennigan, C. J., Nguyen, N. T. and Robinson, A. L.: Determination of Volatility Distributions of Primary Organic Aerosol Emissions from Internal Combustion Engines Using Thermal Desorption Gas Chromatography Mass Spectrometry, Aerosol Sci. Technol., 46(10), 1129–1139, doi:10.1080/02786826.2012.700430, 2012.

Presto, A. A., Miracolo, M. A., Kroll, J. H., Worsnop, D. R., Robinson, A. L. and Donahue, N. M.: Intermediate-volatility organic compounds: A potential source of ambient oxidized organic aerosol, Environ. Sci. Technol., 43(13), 4744–4749, doi:10.1021/es803219q, 2009.

Presto, A. A., Miracolo, M. A., Donahue, N. M. and Robinson, A. L.: Secondary Organic Aerosol Formation from High-NOx Photo-Oxidation of Low Volatility Precursors: n-Alkanes, Environ. Sci. Technol., 44(6), 2029–2034, doi:10.1021/es903712r, 2010.

Presto, A. A., Nguyen, N. T., Ranjan, M., Reeder, A. J., Lipsky, E. M., Hennigan, C. J., Miracolo, M. A., Riemer, D. D. and Robinson, A. L.: Fine particle and organic vapor emissions from staged tests of an in-use aircraft engine, Atmos. Environ., 45(21), 3603–3612, doi:10.1016/j.atmosenv.2011.03.061, 2011.

Pye, H. O. T. and Pouliot, G. A.: Modeling the Role of Alkanes, Polycyclic Aromatic Hydrocarbons, and Their Oligomers in Secondary Organic Aerosol Formation, Environ. Sci. Technol., 46(11), 6041–6047, doi:10.1021/es300409w, 2012.

Pye, H. O. T. and Seinfeld, J. H.: A global perspective on aerosol from low-volatility organic compounds, Atmos. Chem. Phys., 10(9), 4377–4401, doi:10.5194/acp-10-4377-2010, 2010.

Qi, L., Liu, H., Shen, X., Fu, M., Huang, F., Man, H., Deng, F., Shaikh, A. A., Wang, X., Dong, R., Song, C. and He, K.: Intermediate-Volatility Organic Compound Emissions from Nonroad Construction Machinery under Different Operation Modes, Environ. Sci. Technol., 53(23), 13832–13840, doi:10.1021/acs.est.9b01316, 2019.

Qin, M., Murphy, B. N., McDonald, B. C. B., McKeen, S. A., Koval, L. and al, et: Impacts of volatile chemical products on criteria pollutants in an urban atmosphere (in preparation), 2019.

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. a, Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R. and Pandis, S. N.: Rethinking Organic Aerosols :, Science (80-. )., 315(March), 1259–1262, doi:10.1126/science.1133061, 2007.

Schauer, J. J., Kleeman, M. J., Cass, G. R. and Simoneit, B. R. T.: Measurement of Emissions from Air Pollution Sources. 2. C1through C30Organic Compounds from Medium Duty Diesel Trucks, Environ. Sci. Technol., 33(10), 1578–1587, doi:10.1021/es980081n, 1999.

Shah, R. U., Coggon, M. M., Gkatzelis, G. I., McDonald, B. C., Tasoglou, A., Huber, H., Gilman, J., Warneke, C., Robinson, A. L. and Presto, A. A.: Urban Oxidation Flow Reactor Measurements Reveal Significant Secondary Organic Aerosol Contributions from Volatile Emissions of Emerging Importance, Environ. Sci. Technol., 54(2), 714–725, doi:10.1021/acs.est.9b06531, 2020.

Sheu, R., Marcotte, A., Khare, P., Charan, S., Ditto, J. C. and Gentner, D. R.: Advances in offline approaches for chemically speciated measurements of trace gas-phase organic compounds via adsorbent tubes in an integrated sampling-to-analysis system, J. Chromatogr. A, 1575, 80–90, doi:https://doi.org/10.1016/j.chroma.2018.09.014, 2018.

Tkacik, D. S., Presto, A. A., Donahue, N. M. and Robinson, A. L.: Secondary organic aerosol formation from intermediate-volatility organic compounds: Cyclic, linear, and branched alkanes, Environ. Sci. Technol., 46(16), 8773–8781, doi:10.1021/es301112c, 2012.

Tkacik, D. S., Lambe, A. T., Jathar, S., Li, X., Presto, A. a., Zhao, Y., Blake, D., Meinardi, S., Jayne, J. T., Croteau, P. L. and Robinson, A. L.: Secondary organic aerosol formation from in-use motor vehicle emissions using a potential aerosol mass reactor., Environ. Sci. Technol., 48(19), 11235–42, doi:10.1021/es502239v, 2014.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R. and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9(9), 2891–2918, doi:10.5194/acp-9-2891-2009, 2009.

Williams, B. J., Goldstein, A. H., Kreisberg, N. M., Hering, S. V and Abbatt, J. P. D.: An In-Situ Instrument for Speciated Organic Composition of Atmospheric Aerosols: Thermal Desorption Aerosol GC/MS-FID (TAG), Aerosol Sci. Technol., 40, 627–638, doi:10.1080/02786820600754631, 2006.

Worton, D. R., Isaacman, G., Gentner, D. R., Dallmann, T. R., Chan, A. W. H., Ruehl, C., Kirchstetter, T. W., Wilson, K. R., Harley, R. A. and Goldstein, A. H.: Lubricating Oil Dominates Primary Organic Aerosol Emissions from Motor Vehicles, Environ. Sci. Technol., 48(7), 3698–3706, doi:10.1021/es405375j, 2014.

Yao, Z., Shen, X., Ye, Y., Cao, X., Jiang, X., Zhang, Y. and He, K.: On-road emission characteristics of VOCs from diesel trucks in Beijing, China, Atmos. Environ., 103, 87–93, doi:10.1016/j.atmosenv.2014.12.028, 2015.

Yee, L. D., Craven, J. S., Loza, C. L., Schilling, K. A., Ng, N. L., Canagaratna, M. R., Ziemann, P. J., Flagan, R. C. and Seinfeld, J. H.: Effect of chemical structure on secondary organic aerosol formation from $C_{12}$ alkanes, Atmos. Chem. Phys., 13(21), 11121–11140, doi:10.5194/acp-13-11121-2013, 2013.

Zhang, Y., Williams, B. J., Goldstein, A. H., Docherty, K., Ulbrich, I. M. and Jimenez, J. L.: A Technique for Rapid Gas Chromatography Analysis Applied to Ambient Organic Aerosol Measurements from the Thermal Desorption Aerosol Gas Chromatograph (TAG), Aerosol Sci. Technol., 48(11), 1166–1182, doi:10.1080/02786826.2014.967832, 2014.

Zhang, Y., Williams, B. J., Goldstein, A. H., Docherty, K. S. and Jimenez, J. L.: A technique for rapid source apportionment applied to ambient organic aerosol measurements from a thermal desorption aerosol gas chromatograph (TAG), Atmos. Meas. Tech., 9(11), 2016.

Zhao, Y., Kreisberg, N. M., Worton, D. R., Teng, A. P., Hering, S. V. and Goldstein, A. H.: Development of an In Situ Thermal Desorption Gas Chromatography Instrument for Quantifying Atmospheric Semi-Volatile Organic Compounds, Aerosol Sci. Technol., 47(3), 258–266, doi:10.1080/02786826.2012.747673, 2013.

Zhao, Y., Hennigan, C. J., May, A. A., Tkacik, D. S., De Gouw, J. A., Gilman, J. B., Kuster, W. C., Borbon, A. and Robinson, A. L.: Intermediate-volatility organic compounds: A large source of secondary organic aerosol, Environ. Sci. Technol., 48(23), 13743–13750, doi:10.1021/es5035188, 2014.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environ. Sci. Technol., 49(19), 11516–11526, doi:10.1021/acs.est.5b02841, 2015.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines, Environ. Sci. Technol., 50(8), 4554–4563, doi:10.1021/acs.est.5b06247, 2016.

# Chapter 5: Conclusions

## 5.1 Summary of scientific findings

The research in this thesis presents three key improvements to IVOC related studies: (1) It provides comprehensive model-ready mobile-source profiles for others to use; (2) It provides an SOA parameterization to represent hydrocarbon IVOCs in CTMs; and (3) It provides a technique to analyze GC-MS data to get necessary chemical information. More key results and conclusions are summarized below.

IVOC is an important component in organic emissions for mobile sources, ranging from around 4% for gasoline vehicles to more than 50% for diesel vehicles. Volatility distributions of all mobile-source organic emissions show tri-modal volatility distributions, which are defined as 'by-product', 'unburnt fuel' and 'lubricant oil' modes. Despite the large variation in total organic emission factors, their volatility distributions are largely consistent across sources using the same fuel type. The main difference in gasoline IVOC emissions is due to cycle effect (cold UC vs hot-running), whereas the top factor is the after-treatment device (DPF vs non-DPF) for diesel vehicle emissions and load dependency for gas-turbine source.

IVOC is the dominant contributor to predicted SOA formation from mobile source emissions. We underscored the importance of including IVOCs in models and inventories due to their high SOA formation potential using the box model. Large gaps in IVOCs and SVOCs are found between traditional and new profiles based on direct measurements. Using new comprehensive emission profiles, the box model almost doubles the predicted SOA compared to the traditional profile for gasoline vehicle exhaust. For gas-turbine and diesel sources, IVOCs and SVOC vapors combining contribute factors of 13 and 44 more SOA than VOCs, respectively. Our new profiles are compiled into model-ready formats and

available in SPECIATE 5.0.

To simulate the SOA formation from IVOCs in CTMs, we developed a new parameterization from mobile source IVOC emissions designed for implementation in CTMs. The parameterization has six lumped IVOC species: two aromatics and four aliphatics, to account for both the volatility and chemical composition of the IVOC emissions. IVOC emissions can be assigned to these classes using GC-MS data. The relative error between the simplified mechanisms is less than 6% compared to the complete mechanism (Zhao et al., 2015, 2016).

Simulations show that mobile source IVOCs contribute considerably to ambient IVOCs and SOA formation. We performed CTM simulations with new mobile-source organic emission profiles. Results show mobile sources contribute 2.7 μg $m^{-3}$ of IVOCs at Pasadena site during CalNex, which is around 40% of measured concentrations of hydrocarbon IVOCs. They are also predicted to contribute around 1 μg $m^{-3}$ daily peak SOA concentration, a 67% increase compared to the base case. Therefore, it is crucial to include mobile-source IVOC emissions in simulations. To reach the IVOC or SOA mass closure with measurement data, results from exploratory model runs suggest that additional 12% to 26.8% of non-mobile NMOG emissions are likely IVOCs.

Chemical composition information is crucial to predicting the atmospheric properties of IVOCs, including SOA yields. We developed a PMF-based technique to resolve the chemical composition of IVOC UCMs in traditional GC-MS data. This technique can extract the important groups of compounds in IVOC UCMs, and most resolved factors show very similar spectra ($r > 0.9$) to the associated reference compounds in the NIST database. This technique can be applied to a wide range of emission samples, including mobile sources and biomass burning samples. It can also be combined with the CMB model to perform source

apportionment analysis. Analysis of tunnel samples shows, although diesel vehicles only used 20% of the fuel, they contributed around 70% of mobile-source IVOCs in the tunnel. PMF analysis on ambient samples in Pasadena, CA shows up to 70% measured IVOC are oxygenated indicating complex atmospheric oxidation.

This thesis presents systematic efforts to better incorporate IVOCs from direct measurements into CTM simulations. We compiled model-ready emission profiles for mobile sources that include IVOCs and SVOCs from direct measurements (chapter 2). We proposed an SOA parameterization to model the SOA formation from IVOCs, and perform CTM simulations which highlight the contribution of mobile-source and potentially non-mobile source IVOCs (chapter 3). We developed a PMF-based technique to resolve the chemical composition of IVOC UCMs in traditional GC-MS data, and can be applied to a wide range of emission sources (chapter 4).

## 5.2 Recommendations for data integration into CTMs

To incorporate new emission data from other sources into CTMs, we recommend the following procedure as shown in Fig 5.1.

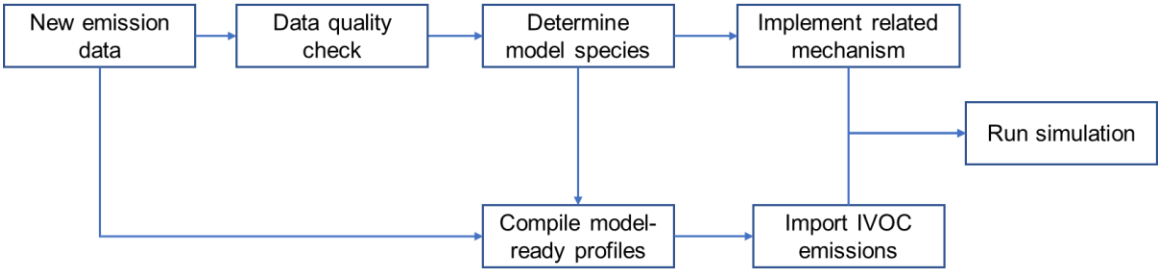



**Figure 5.1 Recommended procedure for new emission data integration into models**

First, understand the emission sources and sampling media of the measurements. To build a comprehensive emission profile, we need to collect all phases (gas, semi-volatile and

particle) of organic emissions. In this work, organic emissions were sampled onto Tedlar bag, adsorbent tubes and quartz filters. This set of sampling media collects the majority of organic emissions compared to traditional approaches (chapter 2). Because there is a potential overlap between data from multiple sampling media, for example, Tedlar bag and adsorbent tubes both collect IVOCs, but our results show that IVOCs less volatile than n-dodecane are better captured by adsorbent tubes. Therefore, a line must be drawn to combine the quality data on both media and avoid double counting. Also, a good quality check is to make sure that the data is of similar quality with previous measurements, for example, measured VOC data agrees with previous profiles, and the sum of measured organics from different media approximately reach the mass closure with bulk measurements, for example, total hydrocarbon (THC) and OC measurements.

Second, quantify the chemical composition of measured IVOCs. For some more advanced instruments, high-resolution chemical information is readily available. But for traditional 1-D GC-MS, the PMF-based technique in chapter 4 may be used to determine the important factors from GC-MS data. Meanwhile, the retention time of n-alkanes is used as the approximation of volatility. This is important because chemical information could be vastly different in a lot of non-mobile sources compared to mobile sources. SOA production yields are dependent on the molecular structure of the primary organic vapor precursor.

Third, determine the chemical transport model species for IVOCs and compile the data into model-ready profiles. Following the chemical information analysis, for example, if measured IVOCs are dominated by hydrocarbons, then the six lumped IVOC species described in chapter 3 can be reused. If not, new IVOC model species and SOA parameterization may be needed. This may be especially important for IVOCs found in volatile chemical products (VCPs) which are often highly oxygenated. Because there is often

computation cost associated with new model species, classifying IVOCs into a few groups that propagate important chemical information (e.g. volatility and functionality) shows lower cost compared to molecular level speciation. Next, implement the IVOC SOA mechanism with new model species (if needed) in your air quality CTM and RCM. An existing chemical mechanism will need to be modified to include the SOA yields corresponding to the IVOC model species. For emissions inputs, employing one of the two options in the following paragraph will generate files ready for use by a CTM. A short-term alternative, for testing and impact study purposes only, would be to skip reprocessing emissions and instead scale new IVOC emissions to the emissions of, for example, NMOG when reading into your air quality CTM. The Community Multiscale Air Quality (CMAQ) model is equipped with online chemical scaling for just this purpose. Tracking the IVOC and product model species could give you the full information on their effects on certain processes and better motivate the inclusion of these species in emissions databases like SPECIATE.

For users of the SPECIATE emissions profile database and the SMOKE emissions processing software, there is a choice at this point for where to include the new IVOC speciation information medium- and long-term. In order to disseminate the new profile to a large community, it is recommended to submit the new profile to the SPECIATE team for inclusion in the database. Once the profile is approved and included, the SPECIATE Team will then be in the position to update the Speciation Tool software to propagate the new model species from the SPECIATE Database to SMOKE. Alternatively, you may use the Speciation Tool to directly create a mapping from NMOG to the new IVOC species and apply it to the Source Classification Codes (SCCs) that correspond to the source you have investigated. This approach would be quicker to implement (no iteration with SPECIATE Team) but would not be available for the wider community to use. For either choice, you

would rerun your emissions processing software (e.g. SMOKE) to generate new emissions inputs for your air quality CTM.

### 5.3 Recommendations for reduced complexity models (RCMs)

Reduced complexity models (RCMs) (Heo et al., 2016b, 2016a; Muller et al., 2011; Tessum et al., 2017) are a group of models that use primary emission data, for example, $SO_2$, $NH_3$ and VOCs, to infer their contribution to ambient PM concentration and related social and health costs. These models usually have much simpler representations of the physical and chemical mechanisms compared to state-of-the-art CTMs. For example, AP2 (Muller et al., 2011) utilized a source–receptor matrix framework to map VOC emissions to ambient $PM_{2.5}$ concentrations. But currently, all RCMs do not account for IVOCs in their model representation, therefore no social cost is assigned to $PM_{2.5}$ contribution from IVOC emissions.

Given the major contribution of IVOCs from mobile sources, we recommend adding a model species 'IVOC' into RCMs and its related social cost is associated with emission quantity and SOA yield. The emission quantity of IVOCs can be calculated from inventory NMOG or VOC using model-ready emission profiles. For example, on-road diesel vehicles show a higher than 50% IVOC-to-NMOG ratio, where gasoline vehicles have around 4% IVOC-to-NMOG ratio, which is more than a factor of 10 difference. Meanwhile, the SOA yield of IVOCs from all mobile sources range from 0.15 to 0.3, a good approximation could be an SOA yield between 0.2 and 0.25, so the related uncertainty on SOA yield is much smaller than emission quantity. Since current RCMs only apply one SOA yield for all VOCs (Gilmore et al., 2019), the number of IVOC species in RCMs should reduce from 6 to 1 compared to our CTM implementation.

## 5.4 Future work

IVOCs account for around 4% to more than 50% of the NMOG emissions for mobile sources. Simulations show that IVOC precursors can contribute up to 70% SOA formation in the southern California region. The PMF-based technique is able to recover most of the IVOC chemical composition in traditional GC-MS, which is largely consistent with results from more sophisticated instruments. Based on our findings in this thesis, we propose a few future work directions.

More vehicle IVOC emissions data are needed across the country, and from different regions of the world with a wide range of fuel compositions, test cycles (hot operations), and seasons (especially winter). Also, IVOCs are important components of NMOG emissions in biomass burning, oil sands, and volatile chemical product emissions (de Gouw et al., 2011; Hatch et al., 2017; Hunter et al., 2017; Liggio et al., 2016; McDonald et al., 2018). Model-ready comprehensive profiles that account for all organic emissions are needed for these non-mobile source categories. We incorporated mobile-source IVOCs into a CTM model (CMAQ v5.3) based on the new profiles documented in SPECIATE 5.0 and we expect this to become regular practices in CTM simulations. But protocols and test methods are still needed so that data can be gathered to fully characterize each sector in inventories.

Non-mobile sources are likely a major source of hydrocarbon and oxygenated IVOC and future research is needed to constrain their emissions on the volatility distribution and chemical composition. Modeling the SOA formation from these IVOCs will likely require an extension of existing chemical mechanisms, for example, new model species such as ‘oxygenated IVOCs’. Also, measurements of ambient IVOC across a range of field sites are needed to better constrain the model parameters. Multigenerational aging on SOA formed

from IVOC emissions still has a large uncertainty, which also has a large effect on downwind receptors.

Our PMF-based technique has shown promising results in a wide range of datasets. But some resolved factors, are less correlated ($r < 0.9$) to the reference compounds in the NIST database, for example, cyclic alkane and SRA factors. These factors are likely mixtures of multiple compounds and have less clean mass spectra. Therefore, we recommend future research on how to speciate such factors into more detailed factors. There are two possible ways to address this issue: (1) split the entire data matrix into multiple slides according to retention time, and run PMF separately on these sub-matrixes. But this may lead to different no. of factors for each sub-matrix; or (2) apply regularization term or sparsity constrain on the PMF algorithm, such that mixture factors are penalized more than regular factors, so as to get better mass spectra solutions.

## 5.5 References


Gilmore, E. A., Heo, J., Muller, N. Z., Tessum, C. W., Hill, J. D., Marshall, J. D. and Adams, P. J.: An inter-comparison of the social costs of air quality from reduced-complexity models, Environ. Res. Lett., 14(7), 74016, 2019.

de Gouw, J. A., Middlebrook, A. M., Warneke, C., Ahmadov, R., Atlas, E. L., Bahreini, R., Blake, D. R., Brock, C. A., Brioude, J., Fahey, D. W., Fehsenfeld, F. C., Holloway, J. S., Le Henaff, M., Lueb, R. A., McKeen, S. A., Meagher, J. F., Murphy, D. M., Paris, C., Parrish, D. D., Perring, A. E., Pollack, I. B., Ravishankara, A. R., Robinson, A. L., Ryerson, T. B., Schwarz, J. P., Spackman, J. R., Srinivasan, A. and Watts, L. A.: Organic Aerosol Formation Downwind from the Deepwater Horizon Oil Spill, Science (80-. )., 331(6022), 1295 LP – 1299 [online] Available from: http://science.sciencemag.org/content/331/6022/1295.abstract, 2011.

Hatch, L. E., Yokelson, R. J., Stockwell, C. E., Veres, P. R., Simpson, I. J., Blake, D. R., Orlando, J. J. and Barsanti, K. C.: Multi-instrument comparison and compilation of non-methane organic gas emissions from biomass burning and implications for smoke-derived secondary organic aerosol precursors, Atmos. Chem. Phys., 17(2), 1471–1489, doi:10.5194/acp-17-1471-2017, 2017.

Heo, J., Adams, P. J. and Gao, H. O.: Public Health Costs of Primary PM2.5 and Inorganic PM2.5 Precursor Emissions in the United States, Environ. Sci. Technol., 50(11), 6061–6070, doi:10.1021/acs.est.5b06125, 2016a.

Heo, J., Adams, P. J. and Gao, H. O.: Reduced-form modeling of public health impacts of inorganic PM2.5 and precursor emissions, Atmos. Environ., 137, 80–89, doi:https://doi.org/10.1016/j.atmosenv.2016.04.026, 2016b.

Hunter, J. F., Day, D. A., Palm, B. B., Yatavelli, R. L. N., Chan, A. W. H., Kaser, L., Cappellin, L., Hayes, P. L., Cross, E. S., Carrasquillo, A. J., Campuzano-Jost, P., Stark, H., Zhao, Y., Hohaus, T., Smith, J. N., Hansel, A., Karl, T., Goldstein, A. H., Guenther, A., Worsnop, D. R., Thornton, J. A., Heald, C. L., Jimenez, J. L. and Kroll, J. H.: Comprehensive characterization of atmospheric organic carbon at a forested site, Nat. Geosci., 10, 748 [online] Available from: http://dx.doi.org/10.1038/ngeo3018, 2017.

Liggio, J., Li, S., Hayden, K., Taha, Y. M., Stroud, C., Darlington, A., Drollette, B. D., Gordon, M., Lee, P., Liu, P., Leithead, A., Moussa, S. G., Wang, D., Brien, J. O., Mittermeier, R. L., Osthoff, H. D., Makar, P. A., Zhang, J., Brook, J. R., Lu, G., Staebler, R. M., Han, Y., Travis, W., Plata, D. L. and Gentner, D. R.: Oil sands operations as a large source of secondary organic aerosols, Nature, 534(7605), 1–16, doi:10.1038/nature17646, 2016.

McDonald, B. C., de Gouw, J. A., Gilman, J. B., Jathar, S. H., Akherati, A., Cappa, C. D., Jimenez, J. L., Lee-Taylor, J., Hayes, P. L., McKeen, S. A., Cui, Y. Y., Kim, S.-W., Gentner, D. R., Isaacman-VanWertz, G., Goldstein, A. H., Harley, R. A., Frost, G. J., Roberts, J. M., Ryerson, T. B. and Trainer, M.: Volatile chemical products emerging as largest petrochemical source of urban organic emissions, Science (80-. )., 359(6377), 760–764, doi:10.1126/science.aaq0524, 2018.

Muller, N. Z., Mendelsohn, R. and Nordhaus, W.: Environmental Accounting for Pollution in the United States Economy, Am. Econ. Rev., 101(5), 1649–1675, doi:10.1257/aer.101.5.1649, 2011.

Tessum, C. W., Hill, J. D. and Marshall, J. D.: InMAP: A model for air pollution interventions, PLoS One, 12(4), e0176131 [online] Available from: https://doi.org/10.1371/journal.pone.0176131, 2017.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production, Environ. Sci. Technol., 49(19), 11516–11526, doi:10.1021/acs.est.5b02841, 2015.

Zhao, Y., Nguyen, N. T., Presto, A. A., Hennigan, C. J., May, A. A. and Robinson, A. L.: Intermediate Volatility Organic Compound Emissions from On-Road Gasoline Vehicles and Small Off-Road Gasoline Engines, Environ. Sci. Technol., 50(8), 4554–4563, doi:10.1021/acs.est.5b06247, 2016.

# Appendix A: Supporting information for Chapter 2

## A.1. Comparison of three individual compounds measured using GC/MS analysis of Tedlar bag versus Tenax adsorbent samples

Figure S2(a) shows that, the most volatile of these species, *n*-pentyl-benzene, the Tedlar bags measurement averaged 5.2 times the adsorbent tubes. We attribute this difference to incomplete collection of this relatively volatile species by the adsorbent tubes or incomplete recovery of thermal desorption method. Figure S2(b) show essentially the same amount of *n*-dodecane was measured using both approaches, a linear regression yields a slope of 0.85 and $R^2$ of 0.9. As for naphthalene (the least volatile of these species), Figure S2(c) show the adsorbent tubes measured about 5 times more than the Tedlar bag, which we attribute to wall losses in the bag. [54]

## A.2. Supplementing gas-turbine and diesel VOC speciation with traditional emission profiles

Given the different levels of VOC characterization, we supplemented our gas-turbine and diesel VOC data with existing speciation profiles (SPECIATE profiles 4674 and 5565). For gas turbine exhaust, 86 individual VOCs were identified, which can be classified as 21 SAPRC groups. Meanwhile, SPECIATE profile 5565 includes 81 individual species, which can be lumped into 27 groups (not including IVOCs in the profile). Of the two grouping results, 17 groups are identified in both profiles, and 10 groups are unique only in profile 5565.

We then complement our VOC results with 10 unique groups as additional 31.9% of VOC mass (all carbonyls, 31.9% of total VOC mass in profile 5565).

For diesel exhaust, 57 individual VOCs and 11 Kovats lumped groups were identified, which can be classified as 25 SAPRC groups. SPECIATE profile 4674 includes 144 individual species, which can be lumped into 34 groups (not including IVOCs in the profile). Of the two grouping results, 22 groups are identified in both profiles, and 12 groups are unique only in profile 4674.

We then complement our VOC results with 12 unique groups as additional 10.8% of VOC mass (10.8% of total VOC mass in profile 4674).

### A.3. Preparing NMOG and POA emission profiles from VBS version

The VBS version in Table S3(a) is designed to be applied to total organic emissions (NMOG + 1.2×OC). The VOC composition profiles in Table S3(b-f) is designed to be applied to VOC emissions. In most emission inventories, where NMOG and POA are provided separately, we recommend partition the VBS profiles into gas- and particle-phase sub-profiles for NMOG and POA emissions respectively, and then apply the VOC profiles in Table S3(b-f) to speciate VOC emissions.

## A.4. Table S1 to S7

**Table S1 Summary of all tested engines, fuels and test cycles with complete characterization data**

| Source type | Category | Class / Model | Number of engines | Number of tests | Model year | Fuel | Test cycle / Thrust |
|---|---|---|---|---|---|---|---|
| Gasoline | LDGVs (on-road) | Pre-LEV | 10 | 10 | 1987-1999 | Commercial summertime California gasoline | Cold-start unified cycle (UC) / Hot-start |
| | | LEV | 9 | 10 | 1991-2009 | | |
| | | ULEV | 10 | 12 | 2003-2012 | | |
| | SOREs (off-road) | SORE-2S | 2 | 3 | 2002, 2005 | | CARB SORE certification cycle |
| | | SORE-4S | 2 | 3 | 2004, 2005 | | |
| Gas-turbine | KC-135 Stratotanker | CFM56-2B1 | 1 | 2 | / | Commercial JP-8 | 4% and 85% |
| Diesel | HDDVs (on-road) | DPF-equipped | 2 | 8 | 2007, 2010 | Three different ULSD fuels with varying aromatic content (9-28%) | Urban Dynamometer Driving Schedule (UDDS) / Creep and idle / High-speed cruise |
| | | Non-DPF | 1 | 10 | 2006 | | |
| | MDDV (on-road) | Non-DPF | 1 | 1 | 2005 | Commercial California ultra-low sulfur diesel (ULSD) | Cold-start UC |
| | TRU (off-road) | Non-DPF | 1 | 2 | 1998 | | CARB procedures for engine certification |

**Table S2 Effect of temperature on gas/particle partitioning at equilibrium**

| Log (C*, T = 298K) | C* (ug/m$^3$) | | | Particle-phase fraction ($X_p$) | | | $\Delta H_{vap}$ (kJ / mol) |
|---|---|---|---|---|---|---|---|
| | T = 298K | T = 273K | T = 320K | T = 298K | T = 273K | T=320 K | |
| Nonvolatile | n/a | n/a | n/a | 1.00 | 1.00 | 1.00 | n/a |
| -1 | 1.00E-01 | 3.14E-03 | 1.34E+00 | 0.99 | 1.00 | 0.88 | 96 |
| 0 | 1.00E+00 | 4.72E-02 | 9.85E+00 | 0.91 | 1.00 | 0.50 | 85 |
| 1 | 1.00E+01 | 7.08E-01 | 7.26E+01 | 0.50 | 0.93 | 0.12 | 74 |
| 2 | 1.00E+02 | 1.06E+01 | 5.35E+02 | 0.09 | 0.48 | 0.02 | 63 |
| 3 | 1.00E+03 | 1.60E+02 | 3.94E+03 | 0.01 | 0.06 | 0.00 | 52 |
| 4 | 1.00E+04 | 2.40E+03 | 2.91E+04 | 0.00 | 0.00 | 0.00 | 41 |
| 5 | 1.00E+05 | 3.60E+04 | 2.14E+05 | 0.00 | 0.00 | 0.00 | 30 |
| 6 | 1.00E+06 | 5.41E+05 | 1.58E+06 | 0.00 | 0.00 | 0.00 | 19 |

**Table S3a VBS distribution of complete profiles**

| | Gasoline (cold-start) | Gasoline (hot operation) | Gas-turbine | Diesel (non-DPF) | Diesel (DPF) |
|---|---|---|---|---|---|
| **NV** | 0.002 | 0.006 | 0.001 | 0.006 | 0.006 |
| **log C* = -1** | 0.000 | 0.012 | 0.005 | 0.004 | 0.001 |
| **log C* = 0** | 0.001 | 0.004 | 0.011 | 0.005 | 0.001 |
| **log C* = 1** | 0.005 | 0.024 | 0.016 | 0.016 | 0.083 |
| **log C* = 2** | 0.005 | 0.019 | 0.009 | 0.018 | 0.038 |
| **log C* = 3** | 0.003 | 0.019 | 0.009 | 0.052 | 0.032 |
| **log C* = 4** | 0.003 | 0.022 | 0.020 | 0.134 | 0.080 |
| **log C* = 5** | 0.008 | 0.035 | 0.046 | 0.178 | 0.176 |
| **log C* = 6** | 0.031 | 0.105 | 0.198 | 0.177 | 0.168 |
| **log C* = 7** | 0.084 | 0.072 | 0.035 | 0.088 | 0.073 |
| **log C* = 8** | 0.293 | 0.242 | 0.078 | 0.050 | 0.024 |
| **log C* = 9** | 0.302 | 0.273 | 0.133 | 0.098 | 0.267 |
| **log C* = 10** | 0.117 | 0.077 | 0.209 | 0.116 | 0.038 |
| **log C* = 11** | 0.145 | 0.089 | 0.230 | 0.057 | 0.014 |

**Table S3b VOC emission composition for gasoline source (cold-start)**

| CAS | Name | Mass fraction |
|---|---|---|
| **50-00-0** | formaldehyde | 0.098448714 |
| **75-07-0** | acetaldehyde | 0.079453169 |
| **74-85-1** | ethene | 0.077569019 |
| **108-88-3** | toluene | 0.06013721 |
| **115-07-1** | propene | 0.039523976 |
| **78-78-4** | 2-methyl-butane | 0.038157017 |
| **71-43-2** | benzene | 0.036696589 |
| **540-84-1** | 2,2,4-trimethyl-pentane | 0.03316184 |
| **74-84-0** | ethane | 0.028506629 |
| **108-38-3** | m-xylene | 0.027485361 |
| **107-83-5** | 2-methyl-pentane | 0.021537976 |
| **74-86-2** | ethyne | 0.018596122 |
| **95-63-6** | 1,2,4-trimethyl-benzene | 0.017161219 |
| **67-64-1** | acetone | 0.015787522 |
| **95-47-6** | o-xylene | 0.0155095 |
| **110-54-3** | n-hexane | 0.014871602 |
| **565-59-3** | 2,3-dimethyl-pentane | 0.014521667 |
| **106-42-3** | p-xylene | 0.014017881 |
| **109-66-0** | n-pentane | 0.01400488 |
| **96-14-0** | 3-methyl-pentane | 0.013934972 |
| **115-11-7** | 2-methyl-propene | 0.01357629 |
| **100-52-7** | benzaldehyde | 0.013233636 |
| **96-37-7** | methylcyclo-pentane | 0.011870355 |
| **100-41-4** | ethyl-benzene | 0.011868345 |
| **589-34-4** | 3-methyl-hexane | 0.010887573 |
| **620-14-4** | 1-methyl-3-ethyl-benzene | 0.010858917 |
| **591-76-4** | 2-methyl-hexane | 0.010740291 |
| **74-98-6** | propane | 0.009575236 |
| **108-08-7** | 2,4-dimethyl-pentane | 0.009478111 |
| **107-02-8** | acrolein | 0.008345552 |
| **142-82-5** | n-heptane | 0.007876648 |
| **565-75-3** | 2,3,4-trimethyl-pentane | 0.007770152 |
| **79-29-8** | 2,3-dimethyl-butane | 0.007424302 |
| **106-98-9** | 1-butene | 0.006543969 |
| **106-97-8** | n-butane | 0.006108686 |
| **108-67-8** | 1,3,5-trimethyl-benzene | 0.005792158 |
| **78-85-3** | methacrolein | 0.005491469 |
| **106-99-0** | 1,3-butadiene | 0.005412884 |
| **589-43-5** | 2,4-dimethyl-hexane | 0.005381599 |
| **622-96-8** | 1-methyl-4-ethyl-benzene | 0.005206562 |
| **123-38-6** | propanal | 0.005191645 |

| | | |
|---|---|---|
| **589-81-1** | 3-methyl-heptane | 0.00513135 |
| **3522-94-9** | 2,2,5-trimethyl-hexane | 0.005041825 |
| **592-27-8** | 2-methyl-heptane | 0.004587965 |
| **611-14-3** | 1-methyl-2-ethyl-benzene | 0.004358071 |
| **592-13-2** | 2,5-dimethyl-hexane | 0.004256891 |
| **108-87-2** | methylcyclo-hexane | 0.004101934 |
| **111-65-9** | n-octane | 0.00406272 |
| **110-82-7** | cyclo-hexane | 0.003913606 |
| **526-73-8** | 1,2,3-trimethyl-benzene | 0.003840964 |
| **624-64-6** | trans-2-butene | 0.003525753 |
| **871-83-0** | 2-methyl-nonane | 0.00324077 |
| **584-94-1** | 2,3-dimethyl-hexane | 0.003084652 |
| **75-83-2** | 2,2-dimethyl-butane | 0.003026964 |
| **590-18-1** | cis-2-butene | 0.002915496 |
| **2216-34-4** | 4-methyl-octane | 0.002711622 |
| **513-35-9** | 2-methyl-2-butene | 0.002711362 |
| **78-93-3** | methylethyl-ketone | 0.002552226 |
| **1759-58-6** | trans-1,3-dimethyl-cyclopentane | 0.00245383 |
| **934-74-7** | 1,3-dimethyl-5-ethyl-benzene | 0.00241612 |
| **2216-33-3** | 3-methyl-octane | 0.002381883 |
| **563-46-2** | 2-methyl-1-butene | 0.002269309 |
| **103-65-1** | n-propyl-benzene | 0.002160127 |
| **4170-30-3** | crotonaldehyde | 0.002111139 |
| **123-72-8** | butanal | 0.00199658 |
| **287-92-3** | cyclo-pentane | 0.001916573 |
| **463-49-0** | 1,2-propadiene | 0.001861323 |
| **589-53-7** | 4-methyl-heptane | 0.001768663 |
| **496-11-7** | 2,3-dihydro-indene | 0.001724763 |
| **100-42-5** | styrene | 0.00169368 |
| **646-04-8** | trans-2-pentene | 0.001657373 |
| **111-84-2** | n-nonane | 0.001633027 |
| **617-78-7** | 3-ethylpentane | 0.00159796 |
| **638-04-0** | cis-1,3-dimethyl-cyclohexane | 0.001556882 |
| **74-99-7** | 1-propyne | 0.001490404 |
| **2532-58-3** | cis-1,3-dimethyl-cyclopentane | 0.001481396 |
| **1074-43-7** | 1-methyl-3-n-propyl-benzene | 0.00144506 |
| **66-25-1** | hexanal | 0.001348987 |
| **110-83-8** | cyclo-hexene | 0.001348008 |
| **926-82-9** | 3,5-dimethyl-heptane | 0.001306079 |
| **124-18-5** | n-decane | 0.001261442 |
| **75-28-5** | methyl-propane | 0.001100569 |
| **109-67-1** | 1-pentene | 0.001087023 |
| **2815-58-9** | 1,2,4-trimethyl-cyclopentane | 0.001064466 |
| **922-62-3** | 3-methyl-cis-2-pentene | 0.00105455 |

| | | |
|---|---|---|
| **1758-88-9** | 1,4-dimethyl-2-ethyl-benzene | 0.001052893 |
| **527-53-7** | 1,2,3,5-tetramethyl-benzene | 0.001039611 |
| **822-50-4** | trans-1,2-dimethyl-cyclopentane | 0.001029538 |
| **2613-66-3** | cis-1-methyl-3-ethyl-cyclopentane | 0.001013644 |
| **874-41-9** | 1,3-dimethyl-4-ethyl-benzene | 0.00096792 |
| **627-20-3** | cis-2-pentene | 0.00090946 |
| **934-80-5** | 1,2-dimethyl-4-ethyl-benzene | 0.000905253 |
| **933-98-2** | 1,2-dimethyl-3-ethyl-benzene | 0.000897434 |
| **625-65-0** | 2,4-dimethyl-2-pentene | 0.000872819 |
| **2207-03-6** | trans-1,3-dimethyl-cyclohexane | 0.000839409 |
| **583-48-2** | 3,4-dimethyl-hexane | 0.000825466 |
| **142-29-0** | cyclo-pentene | 0.000816028 |
| **95-93-2** | 1,2,4,5-tetramethyl-benzene | 0.000715707 |
| **563-45-1** | 3-methyl-1-butene | 0.000696766 |
| **590-35-2** | 2,2-dimethyl-pentane | 0.000651296 |
| **1074-55-1** | 1-methyl-4-n-propyl-benzene | 0.000619922 |
| **141-93-5** | 1,3-diethyl-benzene | 0.000607649 |
| **562-49-2** | 3,3-dimethyl-pentane | 0.000602954 |
| **1069-53-0** | 2,3,5-trimethyl-hexane | 0.000578116 |
| **20291-95-6** | 2,2,5-trimethyl-heptane | 0.000577085 |
| **4032-94-4** | 2,4-dimethyl-octane | 0.000566084 |
| **2207-04-7** | trans-1,4-dimethyl-cyclohexane | 0.000556876 |
| **527-84-4** | 1-methyl-2-(1-methylethyl) benzene | 0.000526823 |
| **98-82-8** | (1-methylethyl) benzene | 0.000516953 |
| **1074-17-5** | 1-methyl-2-n-propyl-benzene | 0.000516353 |
| **2815-57-8** | (1a,2a,3b)-1,2,3-trimethyl-cyclopentane | 0.000504966 |
| **2613-65-2** | trans-1-methyl-3-ethyl-cyclopentane | 0.000501094 |
| **78-79-5** | 2-methyl-1,3-butadiene | 0.000499289 |
| **590-19-2** | 1,2-butadiene | 0.000490403 |
| **2870-04-4** | 1,3-dimethyl-2-ethyl-benzene | 0.000481461 |
| **1072-05-5** | 2,6-dimethyl-heptane | 0.000476217 |
| **14720-74-2** | 2,2,4-trimethyl-heptane | 0.000467016 |
| **824-63-5** | 2-methyl-indan | 0.000439344 |
| **1074-92-6** | 1-(1,1-dimethylethyl)-2-methyl-benzene | 0.000426956 |
| **592-41-6** | 1-hexene | 0.000423546 |
| **13269-52-8** | trans-3-hexene | 0.000422622 |
| **2213-23-2** | 2,4-dimethyl-heptane | 0.000422147 |
| **105-05-5** | 1,4-diethyl-benzene | 0.000390783 |
| **1120-21-4** | n-undecane | 0.000378441 |
| **110-62-3** | pentanal | 0.000376131 |
| **4050-45-7** | trans-2-hexene | 0.000355995 |
| **14686-14-7** | trans-3-heptene | 0.00035565 |
| **7642-04-8** | cis-2-octene | 0.000344011 |
| **1678-91-7** | ethyl-cyclohexane | 0.000318917 |

| | | |
|---|---|---|
| **2207-01-4** | cis-1,2-dimethyl-cyclohexane | 0.000271176 |
| **1839-63-0** | 1,3,5-trimethyl-cyclohexane | 0.000268792 |
| **15869-87-1** | 2,2-dimethyl-octane | 0.000237057 |
| **7146-60-3** | 2,3-dimethyl-octane | 0.000234718 |
| **563-78-0** | 2,3-dimethyl-1-butene | 0.000231376 |
| **535-77-3** | 1-methyl-3-(1-methylethyl) benzene | 0.000228013 |
| **4110-44-5** | 3,3-dimethyl-octane | 0.000224977 |
| **538-93-2** | (2-methyl-propyl) benzene | 0.00022236 |
| **625-27-4** | 2-methyl-2-pentene | 0.000221064 |
| **463-82-1** | 2,2-dimethyl-propane | 0.000208707 |
| **760-20-3** | 3-methyl-1-pentene | 0.000207391 |
| **16747-30-1** | 2,4,4-trimethyl-hexane | 0.000207128 |
| **15869-89-3** | 2,5-dimethyl-octane | 0.000200048 |
| **692-24-0** | 2-methyl-trans-3-hexene | 0.000199602 |
| **3074-71-3** | 2,3-dimethyl-heptane | 0.000180652 |
| **124-11-8** | 1-nonene | 0.000168512 |
| **1640-89-7** | ethylcyclo-pentane | 0.000161199 |
| **2738-19-4** | 2-methyl-2-hexene | 0.000156125 |
| **111-66-0** | 1-octene | 0.000154377 |
| **7688-21-3** | cis-2-hexene | 0.000152575 |
| **674-76-0** | 4-methyl-trans-2-pentene | 0.000147646 |
| **689-97-4** | 1-buten-3-yne | 0.000137369 |
| **590-73-8** | 2,2-dimethyl-hexane | 0.000125526 |
| **763-29-1** | 2-methyl-1-pentene | 0.000124347 |
| **6443-92-1** | cis-2-heptene | 0.000120364 |
| **691-37-2** | 4-methyl-1-pentene | 0.000116634 |
| **16747-26-5** | 2,2,4-trimethyl-hexane | 0.000114231 |
| **135-98-8** | (1-methyl-propyl) benzene | 0.000111651 |
| **14850-23-8** | trans-4-octene | 0.000110369 |
| **1120-62-3** | 3-methyl-cyclopentene | 0.000104472 |
| **6236-88-0** | 1-methyl-4-ethyl-cyclohexane | 0.000103236 |
| **107-00-6** | 1-butyne | 9.25532E-05 |
| **135-01-3** | 1,2-diethyl-benzene | 8.68528E-05 |
| **592-76-7** | 1-heptene | 8.66527E-05 |
| **560-21-4** | 2,3,3-trimethyl-pentane | 8.65911E-05 |
| **464-06-2** | 2,2,3-trimethyl-butane | 8.47673E-05 |
| **558-37-2** | 3,3-dimethyl-1-butene | 8.42842E-05 |
| **616-12-6** | 3-methyl-trans-2-pentene | 8.00602E-05 |
| **693-89-0** | 1-methyl-cyclopentene | 7.4531E-05 |
| **99-87-6** | 1-methyl-4-(1-methylethyl) benzene | 7.35496E-05 |
| **2051-30-1** | 2,6-dimethyl-octane | 7.15307E-05 |
| **107-40-4** | 2,4,4-trimethyl-2-pentene | 6.86095E-05 |
| **16021-20-8** | 1-ethyl-2-n-propyl-benzene | 6.08539E-05 |
| **2004-70-8** | trans-1,3-pentadiene | 5.97573E-05 |

| | | |
|---|---|---|
| **13389-42-9** | trans-2-octene | 5.96843E-05 |
| **14686-13-6** | trans-2-heptene | 5.89585E-05 |
| **10574-37-5** | 2,3-dimethyl-2-pentene | 5.72033E-05 |
| **563-16-6** | 3,3-dimethyl-hexane | 5.69252E-05 |
| **107-39-1** | 2,4,4-trimethyl-1-pentene | 5.63071E-05 |
| **2213-32-3** | 2,4-dimethyl-1-pentene | 4.90584E-05 |
| **3683-22-5** | 4-methyl-trans-2-hexene | 4.52079E-05 |
| **7385-78-6** | 3,4-dimethyl-1-pentene | 4.14052E-05 |
| **460-12-8** | 1,3-butadiyne | 4.12311E-05 |
| **10574-36-4** | 3-methyl-cis-2-hexene | 3.21055E-05 |
| **7642-09-3** | cis-3-hexene | 2.42751E-05 |
| **3221-61-2** | 2-methyl-octane | 2.02113E-05 |
| **3899-36-3** | 3-methyl-trans-3-hexene | 1.90901E-05 |
| **816-79-5** | 3-ethyl-2-pentene | 1.61224E-05 |
| **3404-61-3** | 3-methyl-1-hexene | 8.47651E-06 |
| **503-17-3** | 2-butyne | 5.77414E-06 |
| **1634-04-4** | methyl-tert-butyl-ether | 0 |
| **637-92-3** | 1-ethyl-tert-butyl-ether | 0 |
| **691-38-3** | 4-methyl-cis-2-pentene | 0 |

**Table S3c VOC emission composition for gasoline source (hot operation)**

| CAS | Name | Mass fraction |
|---|---|---|
| **71-43-2** | benzene | 0.08835029 |
| **108-88-3** | toluene | 0.073388053 |
| **107-83-5** | 2-methyl-pentane | 0.064165863 |
| **74-85-1** | ethene | 0.054665319 |
| **75-28-5** | methyl-propane | 0.051135102 |
| **108-38-3** | m-xylene | 0.042019507 |
| **74-84-0** | ethane | 0.039221797 |
| **50-00-0** | formaldehyde | 0.035221052 |
| **115-07-1** | propene | 0.033526361 |
| **96-14-0** | 3-methyl-pentane | 0.032662285 |
| **95-63-6** | 1,2,4-trimethyl-benzene | 0.030130967 |
| **95-47-6** | o-xylene | 0.026465306 |
| **1839-63-0** | 1,3,5-trimethyl-cyclohexane | 0.02437512 |
| **111-65-9** | n-octane | 0.023831474 |
| **106-42-3** | p-xylene | 0.021570907 |
| **16747-30-1** | 2,4,4-trimethyl-hexane | 0.020658974 |
| **2532-58-3** | cis-1,3-dimethyl-cyclopentane | 0.020123322 |
| **592-41-6** | 1-hexene | 0.020014748 |
| **589-81-1** | 3-methyl-heptane | 0.018470098 |
| **527-53-7** | 1,2,3,5-tetramethyl-benzene | 0.018378861 |
| **824-63-5** | 2-methyl-indan | 0.0181645 |
| **287-92-3** | cyclo-pentane | 0.016316706 |
| **565-75-3** | 2,3,4-trimethyl-pentane | 0.015032351 |
| **100-52-7** | benzaldehyde | 0.013300528 |
| **75-07-0** | acetaldehyde | 0.012849327 |
| **142-82-5** | n-heptane | 0.012457509 |
| **108-08-7** | 2,4-dimethyl-pentane | 0.011178691 |
| **464-06-2** | 2,2,3-trimethyl-butane | 0.01066526 |
| **7642-04-8** | cis-2-octene | 0.010612309 |
| **590-73-8** | 2,2-dimethyl-hexane | 0.009991342 |
| **2213-23-2** | 2,4-dimethyl-heptane | 0.009658211 |
| **2613-66-3** | cis-1-methyl-3-ethyl-cyclopentane | 0.008147075 |
| **589-53-7** | 4-methyl-heptane | 0.007966429 |
| **2613-65-2** | trans-1-methyl-3-ethyl-cyclopentane | 0.007891499 |
| **96-37-7** | methylcyclo-pentane | 0.0078344 |
| **3404-61-3** | 3-methyl-1-hexene | 0.007194667 |
| **74-86-2** | ethyne | 0.006985866 |
| **1758-88-9** | 1,4-dimethyl-2-ethyl-benzene | 0.006709835 |
| **496-11-7** | 2,3-dihydro-indene | 0.006578582 |
| **99-87-6** | 1-methyl-4-(1-methylethyl) benzene | 0.006186673 |
| **926-82-9** | 3,5-dimethyl-heptane | 0.005925649 |

| | | |
|---|---|---|
| **535-77-3** | 1-methyl-3-(1-methylethyl) benzene | 0.005883676 |
| **540-84-1** | 2,2,4-trimethyl-pentane | 0.005783134 |
| **4032-94-4** | 2,4-dimethyl-octane | 0.004813196 |
| **1074-17-5** | 1-methyl-2-n-propyl-benzene | 0.004213845 |
| **141-93-5** | 1,3-diethyl-benzene | 0.004128507 |
| **1678-91-7** | ethyl-cyclohexane | 0.003027222 |
| **67-64-1** | acetone | 0.002802591 |
| **2815-57-8** | (1a,2a,3b)-1,2,3-trimethyl-cyclopentane | 0.002568004 |
| **463-49-0** | 1,2-propadiene | 0.002566361 |
| **15869-89-3** | 2,5-dimethyl-octane | 0.002522685 |
| **100-42-5** | styrene | 0.001948572 |
| **616-12-6** | 3-methyl-trans-2-pentene | 0.001735617 |
| **106-99-0** | 1,3-butadiene | 0.001282624 |
| **6236-88-0** | 1-methyl-4-ethyl-cyclohexane | 0.001263392 |
| **16021-20-8** | 1-ethyl-2-n-propyl-benzene | 0.001216907 |
| **1120-21-4** | n-undecane | 0.000986545 |
| **933-98-2** | 1,2-dimethyl-3-ethyl-benzene | 0.000663197 |
| **1074-43-7** | 1-methyl-3-n-propyl-benzene | 0.000452997 |
| **135-98-8** | (1-methyl-propyl) benzene | 0.000395385 |
| **620-14-4** | 1-methyl-3-ethyl-benzene | 0.000374894 |
| **110-83-8** | cyclo-hexene | 0.000327168 |
| **591-76-4** | 2-methyl-hexane | 0.000309917 |
| **590-35-2** | 2,2-dimethyl-pentane | 0.000257643 |
| **66-25-1** | hexanal | 0.000166723 |
| **98-82-8** | (1-methylethyl) benzene | 0.000140577 |
| **2738-19-4** | 2-methyl-2-hexene | 0.000107625 |
| **624-64-6** | trans-2-butene | 2.15507E-05 |
| **123-72-8** | butanal | 1.66303E-05 |
| **538-93-2** | (2-methyl-propyl) benzene | 0 |
| **1074-92-6** | 1-(1,1-dimethylethyl)-2-methyl-benzene | 0 |
| **526-73-8** | 1,2,3-trimethyl-benzene | 0 |
| **95-93-2** | 1,2,4,5-tetramethyl-benzene | 0 |
| **2815-58-9** | 1,2,4-trimethyl-cyclopentane | 0 |
| **590-19-2** | 1,2-butadiene | 0 |
| **135-01-3** | 1,2-diethyl-benzene | 0 |
| **934-80-5** | 1,2-dimethyl-4-ethyl-benzene | 0 |
| **108-67-8** | 1,3,5-trimethyl-benzene | 0 |
| **460-12-8** | 1,3-butadiyne | 0 |
| **2870-04-4** | 1,3-dimethyl-2-ethyl-benzene | 0 |
| **874-41-9** | 1,3-dimethyl-4-ethyl-benzene | 0 |
| **934-74-7** | 1,3-dimethyl-5-ethyl-benzene | 0 |
| **105-05-5** | 1,4-diethyl-benzene | 0 |
| **689-97-4** | 1-buten-3-yne | 0 |
| **106-98-9** | 1-butene | 0 |

| | | |
|---|---|---|
| **107-00-6** | 1-butyne | 0 |
| **637-92-3** | 1-ethyl-tert-butyl-ether | 0 |
| **592-76-7** | 1-heptene | 0 |
| **527-84-4** | 1-methyl-2-(1-methylethyl) benzene | 0 |
| **611-14-3** | 1-methyl-2-ethyl-benzene | 0 |
| **622-96-8** | 1-methyl-4-ethyl-benzene | 0 |
| **1074-55-1** | 1-methyl-4-n-propyl-benzene | 0 |
| **693-89-0** | 1-methyl-cyclopentene | 0 |
| **124-11-8** | 1-nonene | 0 |
| **111-66-0** | 1-octene | 0 |
| **109-67-1** | 1-pentene | 0 |
| **74-99-7** | 1-propyne | 0 |
| **14720-74-2** | 2,2,4-trimethyl-heptane | 0 |
| **16747-26-5** | 2,2,4-trimethyl-hexane | 0 |
| **20291-95-6** | 2,2,5-trimethyl-heptane | 0 |
| **3522-94-9** | 2,2,5-trimethyl-hexane | 0 |
| **75-83-2** | 2,2-dimethyl-butane | 0 |
| **15869-87-1** | 2,2-dimethyl-octane | 0 |
| **463-82-1** | 2,2-dimethyl-propane | 0 |
| **560-21-4** | 2,3,3-trimethyl-pentane | 0 |
| **1069-53-0** | 2,3,5-trimethyl-hexane | 0 |
| **563-78-0** | 2,3-dimethyl-1-butene | 0 |
| **10574-37-5** | 2,3-dimethyl-2-pentene | 0 |
| **79-29-8** | 2,3-dimethyl-butane | 0 |
| **3074-71-3** | 2,3-dimethyl-heptane | 0 |
| **584-94-1** | 2,3-dimethyl-hexane | 0 |
| **7146-60-3** | 2,3-dimethyl-octane | 0 |
| **565-59-3** | 2,3-dimethyl-pentane | 0 |
| **107-39-1** | 2,4,4-trimethyl-1-pentene | 0 |
| **107-40-4** | 2,4,4-trimethyl-2-pentene | 0 |
| **2213-32-3** | 2,4-dimethyl-1-pentene | 0 |
| **625-65-0** | 2,4-dimethyl-2-pentene | 0 |
| **589-43-5** | 2,4-dimethyl-hexane | 0 |
| **592-13-2** | 2,5-dimethyl-hexane | 0 |
| **1072-05-5** | 2,6-dimethyl-heptane | 0 |
| **2051-30-1** | 2,6-dimethyl-octane | 0 |
| **503-17-3** | 2-butyne | 0 |
| **78-79-5** | 2-methyl-1,3-butadiene | 0 |
| **563-46-2** | 2-methyl-1-butene | 0 |
| **763-29-1** | 2-methyl-1-pentene | 0 |
| **513-35-9** | 2-methyl-2-butene | 0 |
| **625-27-4** | 2-methyl-2-pentene | 0 |
| **78-78-4** | 2-methyl-butane | 0 |
| **592-27-8** | 2-methyl-heptane | 0 |

| | | |
|---|---|---|
| **871-83-0** | 2-methyl-nonane | 0 |
| **3221-61-2** | 2-methyl-octane | 0 |
| **115-11-7** | 2-methyl-propene | 0 |
| **692-24-0** | 2-methyl-trans-3-hexene | 0 |
| **558-37-2** | 3,3-dimethyl-1-butene | 0 |
| **563-16-6** | 3,3-dimethyl-hexane | 0 |
| **4110-44-5** | 3,3-dimethyl-octane | 0 |
| **562-49-2** | 3,3-dimethyl-pentane | 0 |
| **7385-78-6** | 3,4-dimethyl-1-pentene | 0 |
| **583-48-2** | 3,4-dimethyl-hexane | 0 |
| **816-79-5** | 3-ethyl-2-pentene | 0 |
| **617-78-7** | 3-ethylpentane | 0 |
| **563-45-1** | 3-methyl-1-butene | 0 |
| **760-20-3** | 3-methyl-1-pentene | 0 |
| **10574-36-4** | 3-methyl-cis-2-hexene | 0 |
| **922-62-3** | 3-methyl-cis-2-pentene | 0 |
| **1120-62-3** | 3-methyl-cyclopentene | 0 |
| **589-34-4** | 3-methyl-hexane | 0 |
| **2216-33-3** | 3-methyl-octane | 0 |
| **3899-36-3** | 3-methyl-trans-3-hexene | 0 |
| **691-37-2** | 4-methyl-1-pentene | 0 |
| **691-38-3** | 4-methyl-cis-2-pentene | 0 |
| **2216-34-4** | 4-methyl-octane | 0 |
| **3683-22-5** | 4-methyl-trans-2-hexene | 0 |
| **674-76-0** | 4-methyl-trans-2-pentene | 0 |
| **107-02-8** | acrolein | 0 |
| **2207-01-4** | cis-1,2-dimethyl-cyclohexane | 0 |
| **638-04-0** | cis-1,3-dimethyl-cyclohexane | 0 |
| **590-18-1** | cis-2-butene | 0 |
| **6443-92-1** | cis-2-heptene | 0 |
| **7688-21-3** | cis-2-hexene | 0 |
| **627-20-3** | cis-2-pentene | 0 |
| **7642-09-3** | cis-3-hexene | 0 |
| **4170-30-3** | crotonaldehyde | 0 |
| **110-82-7** | cyclo-hexane | 0 |
| **142-29-0** | cyclo-pentene | 0 |
| **100-41-4** | ethyl-benzene | 0 |
| **1640-89-7** | ethylcyclo-pentane | 0 |
| **78-85-3** | methacrolein | 0 |
| **108-87-2** | methylcyclo-hexane | 0 |
| **78-93-3** | methylethyl-ketone | 0 |
| **1634-04-4** | methyl-tert-butyl-ether | 0 |
| **106-97-8** | n-butane | 0 |
| **124-18-5** | n-decane | 0 |

| | | |
|---|---|---|
| **110-54-3** | n-hexane | 0 |
| **111-84-2** | n-nonane | 0 |
| **109-66-0** | n-pentane | 0 |
| **103-65-1** | n-propyl-benzene | 0 |
| **110-62-3** | pentanal | 0 |
| **123-38-6** | propanal | 0 |
| **74-98-6** | propane | 0 |
| **822-50-4** | trans-1,2-dimethyl-cyclopentane | 0 |
| **2207-03-6** | trans-1,3-dimethyl-cyclohexane | 0 |
| **1759-58-6** | trans-1,3-dimethyl-cyclopentane | 0 |
| **2004-70-8** | trans-1,3-pentadiene | 0 |
| **2207-04-7** | trans-1,4-dimethyl-cyclohexane | 0 |
| **14686-13-6** | trans-2-heptene | 0 |
| **4050-45-7** | trans-2-hexene | 0 |
| **13389-42-9** | trans-2-octene | 0 |
| **646-04-8** | trans-2-pentene | 0 |
| **14686-14-7** | trans-3-heptene | 0 |
| **13269-52-8** | trans-3-hexene | 0 |
| **14850-23-8** | trans-4-octene | 0 |

**Table S3d VOC emission composition for gas-turbine engine (idle)**

| CAS | Name | Mass fraction |
|---|---|---|
| **74-86-2** | ethyne | 0.339280582 |
| **50-00-0** | Formaldehyde | 0.093587659 |
| **115-07-1** | Propene | 0.082621687 |
| **112-31-2** | Decanal | 0.044421827 |
| **75-07-0** | Acetaldehyde | 0.032478187 |
| **71-43-2** | benzene | 0.027532651 |
| **106-99-0** | 1,3-butadiene | 0.027330903 |
| **106-98-9** | 1-butene | 0.023094197 |
| **107-02-8** | Acrolein (2-propenal) | 0.018618698 |
| **107-22-2** | Glyoxal | 0.01380627 |
| **67-56-1** | Methyl alcohol (methanol) | 0.013722642 |
| **74-84-0** | Ethane | 0.013706988 |
| **78-98-8** | Methylglyoxal | 0.011426665 |
| **142-29-0** | Cyclo-pentene | 0.011333483 |
| **1120-21-4** | n-undecane | 0.011119868 |
| **109-67-1** | 1-pentene | 0.01082318 |
| **108-88-3** | toluene | 0.010051791 |
| **592-41-6** | 1-hexene | 0.00962456 |
| **74-85-1** | Ethene | 0.009173594 |
| **74-99-7** | 1-Propyne | 0.008580218 |
| **115-11-7** | Iso-butene | 0.008509013 |
| **4170-30-3** | Crotonaldehyde | 0.007853457 |
| **592-76-7** | 1-heptene | 0.007298526 |
| **624-64-6** | Trans-2-butene | 0.007239188 |
| **78-79-5** | 2-methyl-1,3-butadiene | 0.006645812 |
| **110-82-7** | cyclo-hexane | 0.006159244 |
| **107-83-5** | 2-methyl-pentane | 0.005957496 |
| **526-73-8** | 1,2,3-trimethyl-benzene | 0.005577735 |
| **105-05-5** | 1,4-diethyl-benzene | 0.005542133 |
| **123-38-6** | Propionaldehyde | 0.00552707 |
| **108-95-2** | Phenol (carbolic acid) | 0.005519467 |
| **75-28-5** | Iso-butane | 0.005067432 |
| **95-63-6** | 1,2,4-trimethyl-benzene | 0.004972492 |
| **135-98-8** | Sec-butylbenzene | 0.004675804 |
| **74-98-6** | Propane | 0.004438453 |
| **111-84-2** | n-nonane | 0.004284175 |
| **78-78-4** | Iso-pentane | 0.004034957 |
| **563-46-2** | 2-methyl-1-butene | 0.003595859 |
| **100-52-7** | Benzaldehyde | 0.003573209 |
| **563-45-1** | 3-methyl-1-butene | 0.003500919 |
| **25339-56-4** | Heptene | 0.003329926 |

| | | |
|---|---|---|
| **78-85-3** | 2-methyl-2-propenal | 0.003261503 |
| **691-37-2** | 4-methyl-1-pentene | 0.003227966 |
| **95-93-2** | 1,2,4,5-tetramethyl-benzene | 0.003227966 |
| **108-38-3** | m-xylene | 0.003133026 |
| **106-97-8** | n-Butane | 0.002943145 |
| **589-34-4** | 3-methyl-hexane | 0.002907543 |
| **67-64-1** | Acetone | 0.002805349 |
| **103-65-1** | n-propyl-benzene | 0.001970009 |
| **620-14-4** | 3-ethyl-toluene | 0.001875068 |
| **124-11-8** | 1-nonene | 0.001870233 |
| **646-04-8** | Trans-2-pentene | 0.001863201 |
| **110-62-3** | Valeraldehyde | 0.00186263 |
| **110-54-3** | n-hexane | 0.001827598 |
| **110-83-8** | cyclo-hexene | 0.001720791 |
| **108-67-8** | 1,3,5-trimethyl-benzene | 0.001708923 |
| **108-87-2** | 1-methyl-cyclohexane | 0.001708923 |
| **611-14-3** | 2-ethyl-toluene | 0.001495308 |
| **287-92-3** | Cyclo-pentane | 0.001495308 |
| **96-14-0** | 3-methyl-pentane | 0.00148344 |
| **109-66-0** | n-Pentane | 0.001424103 |
| **872-05-9** | 1-decene | 0.001406476 |
| **590-18-1** | Cis-2-butene | 0.0013885 |
| **96-37-7** | 1-methyl-cyclopentane | 0.001329162 |
| **135-01-3** | 1,2-diethyl-benzene | 0.00129356 |
| **763-29-1** | 2-methyl-1-pentene | 0.001257957 |
| **141-93-5** | 1,3-diethyl-benzene | 0.001210487 |
| **4050-45-7** | Trans-2-hexene | 0.001127415 |
| **104-51-8** | 1-butyl-benzene | 0.001008739 |
| **627-20-3** | Cis-2-pentene | 0.000996872 |
| **100-42-5** | styrene | 0.000973137 |
| **496-11-7** | indane | 0.000937534 |
| **622-96-8** | 4-ethyl-toluene | 0.000913799 |
| **123-72-8** | Butyraldehyde or butanal | 0.000904706 |
| **111-65-9** | n-octane | 0.000890064 |
| **10574-37-5** | 2,3-dimethyl-2-pentene | 0.000890064 |
| **7642-09-3** | Cis-3-hexene | 0.000854462 |
| **592-27-8** | 2-methyl-heptane | 0.000842594 |
| **591-76-4** | 2-methyl-hexane | 0.000795124 |
| **590-19-2** | 1,2-butadiene | 0.000759521 |
| **592-48-3** | Trans-1,3-hexadiene | 0.000747654 |
| **80-56-8** | a-pinene | 0.000735786 |
| **7688-21-3** | Cis-2-hexene | 0.000723919 |
| **513-35-9** | 2-methyl-2-butene | 0.000712051 |
| **111-66-0** | 1-octene | 0.000700184 |

| | | |
|---|---|---|
| **142-82-5** | n-heptane | 0.000700184 |
| **589-81-1** | 3-methyl-heptane | 0.000676449 |
| **589-53-7** | 4-methyl-heptane | 0.000664581 |
| **565-75-3** | 2,3,4-trimethyl-pentane | 0.000628979 |
| **95-47-6** | o-xylene | 0.000617111 |
| **591-49-1** | 1-methyl-cyclohexene | 0.000617111 |
| **98-82-8** | iso-propyl-benzene | 0.000569641 |
| **106-42-3** | p-xylene | 0.000569641 |
| **100-41-4** | ethyl-benzene | 0.000462833 |
| **75-19-4** | Cyclo-propane | 0.000344158 |
| **79-29-8** | 2,3-dimethyl-butane | 0.000332291 |
| **124-18-5** | n-decane | 0.000296688 |
| **625-27-4** | 2-methyl-2-pentene | 0.000249218 |
| **590-86-3** | Isovaleraldehyde | 0.000243282 |
| **75-83-2** | 2,2-dimethyl-butane | 0.000178013 |

**Table S3e VOC emission composition for diesel source (non-DPF)**

| CAS | Name | Mass fraction |
|---|---|---|
| **50-00-0** | formaldehyde | 0.190515683 |
| **74-85-1** | ethene | 0.131286669 |
| | UNK | 0.088762269 |
| **75-07-0** | acetaldehyde | 0.085135263 |
| **763-29-1** | 2-methyl-1-pentene | 0.070828265 |
| **115-07-1** | propene | 0.048637884 |
| **96-37-7** | Methylcyclopentane | 0.044640885 |
| **78-85-3** | methacrolein | 0.029117448 |
| **67-64-1** | acetone | 0.025364537 |
| **74-86-2** | ethyne | 0.022825338 |
| **107-02-8** | acrolein | 0.021514473 |
| **100-52-7** | benzaldehyde | 0.012106658 |
| **106-98-9** | 1-butene | 0.011809058 |
| **1120-21-4** | n-undecane | 0.011300307 |
| **71-43-2** | benzene | 0.011096203 |
| **123-38-6** | propanal | 0.010664552 |
| **124-18-5** | n-decane | 0.010074798 |
| **106-99-0** | 1,3-butadiene | 0.008732024 |
| **109-66-0** | N-pentane | 0.007825628 |
| **110-54-3** | N-hexane | 0.007666122 |
| **107-22-2** | Glyoxal | 0.007393502 |
| **123-72-8** | butanal | 0.006838604 |
| **108-88-3** | toluene | 0.006799629 |
| **540-84-1** | 2,2,4-trimethylpentane | 0.006541924 |
| **108-38-3** | m-xylene | 0.00628432 |
| **108-08-7** | 2,4-dimethylpentane | 0.006077602 |
| **115-11-7** | 2-methyl-propene | 0.005897392 |
| **111-84-2** | n-nonane | 0.005078584 |
| **109-67-1** | 1-pentene | 0.005069993 |
| **592-76-7** | 1-heptene | 0.003802851 |
| **696-29-7** | Isopropylcyclohexane | 0.003760477 |
| **2216-33-3** | 3-methyloctane | 0.00353982 |
| **100-41-4** | ethyl-benzene | 0.003265949 |
| **589-34-4** | 3-methylhexane | 0.003252282 |
| **142-82-5** | N-heptane | 0.00312251 |
| **110-82-7** | Cyclohexane | 0.003000046 |
| **513-35-9** | 2-methyl-2-butene | 0.002915484 |
| **108-87-2** | Methylcyclohexane | 0.002868535 |
| **78-93-3** | methylethyl-ketone | 0.002835303 |
| **616-12-6** | 3-methyl-trans-2-pentene | 0.002762539 |
| **95-47-6** | o-xylene | 0.002742249 |

| | | |
|---|---|---|
| **111-65-9** | n-octane | 0.002302202 |
| **105-05-5** | 1,4-diethylbenzene (para) | 0.002264525 |
| **624-64-6** | trans-2-butene | 0.002072279 |
| **611-14-3** | 1-Methyl-2-ethylbenzene | 0.001955146 |
| **565-59-3** | 2,3-dimethylpentane | 0.001948398 |
| **591-76-4** | 2-methylhexane | 0.00176243 |
| **527-53-7** | 1,2,3,5-tetramethylbenzene | 0.001720505 |
| **141-93-5** | 1,3-diethylbenzene (meta) | 0.001705933 |
| **79-29-8** | 2,3-dimethylbutane | 0.001685282 |
| **110-62-3** | pentanal | 0.001603763 |
| **463-49-0** | 1,2-propadiene | 0.001572021 |
| **25155-15-1** | Isopropyltoluene | 0.001557047 |
| **592-27-8** | 2-methylheptane | 0.001546028 |
| **78-79-5** | Isoprene (2-methyl-1,3-butadiene) | 0.001529748 |
| **142-29-0** | Cyclopentene | 0.001511457 |
| **3522-94-9** | 2,2,5-trimethylhexane | 0.001421593 |
| **563-46-2** | 2-methyl-1-butene | 0.001415324 |
| **590-18-1** | cis-2-butene | 0.001401075 |
| **589-81-1** | 3-methylheptane | 0.001353489 |
| **110-83-8** | Cyclohexene | 0.001273925 |
| **28729-54-6** | Propyltoluene | 0.001252591 |
| **526-73-8** | 1,2,3-trimethylbenzene | 0.001240402 |
| **563-45-1** | 3-methyl-1-butene | 0.001190237 |
| **620-14-4** | 1-Methyl-3-ethylbenzene (3-Ethyltoluene) | 0.001164947 |
| **75-83-2** | 2,2-dimethylbutane | 0.001065483 |
| **108-67-8** | 1,3,5-trimethylbenzene | 0.001045319 |
| **74-84-0** | ethane | 0.000990347 |
| **622-96-8** | 1-Methyl-4-ethylbenzene (4-ethyltoluene) | 0.000984128 |
| **107-83-5** | 2-methylpentane (isohexane) | 0.000980939 |
| **287-92-3** | Cyclopentane | 0.00096242 |
| **4050-45-7** | Trans-2-hexene | 0.000908185 |
| **80-56-8** | Alpha-pinene | 0.000874869 |
| **95-93-2** | 1,2,4,5-tetramethylbenzene | 0.000825046 |
| **565-75-3** | 2,3,4-trimethylpentane | 0.000805802 |
| **78-78-4** | 2-methyl-butane | 0.000791653 |
| **767-58-8** | 1-Methylindan | 0.000705877 |
| **496-11-7** | Indan | 0.000700212 |
| **66-25-1** | hexanal | 0.000697103 |
| **103-65-1** | N-propylbenzene | 0.000639453 |
| **100-42-5** | styrene | 0.000552381 |
| **95-63-6** | 1,2,4-trimethylbenzene & t-butylbenzene | 0.000523508 |
| **104-51-8** | N-butylbenzene | 0.000523049 |
| **7688-21-3** | Cis-2-hexene | 0.000518564 |
| **589-53-7** | 4-methylheptane | 0.000445284 |

| | | |
|---|---|---|
| **824-63-5** | 2-methylindan | 0.000428208 |
| **592-48-3** | 1,3-hexadiene (trans) | 0.000371216 |
| **98-82-8** | Isopropylbenzene (cumene) | 0.000313981 |
| **625-27-4** | 2-methyl-2-pentene | 0.000309001 |
| **138-86-3** | Dl-limonene (dipentene) | 0.000297207 |
| **922-62-3** | 3-methyl-cis-2-pentene | 0.000227492 |
| **584-94-1** | 2,3-dimethylhexane | 0.00016783 |
| **10574-37-5** | 2,3-dimethyl-2-pentene | 5.8815E-05 |
| **10061-02-6** | Trans-1,3-dichloropropene | 5.39753E-05 |
| **127-91-3** | Beta-pinene | 1.62063E-05 |
| **1072-05-5; 2040-96-2** | 2,6-dimethylheptane, propylcyclopentane | 1.24439E-05 |
| **106-97-8** | n-butane | 0 |
| **4170-30-3** | crotonaldehyde | 0 |
| **590-19-2** | 1,2-butadiene | 0 |
| **627-20-3** | cis-2-pentene | 0 |
| **646-04-8** | trans-2-pentene | 0 |
| **74-98-6** | propane | 0 |
| **74-99-7** | 1-propyne | 0 |
| **75-28-5** | methyl-propane | 0 |
| **107-00-6** | 1-butyne | 0 |
| **463-82-1** | 2,2-dimethyl-propane | 0 |
| **503-17-3** | 2-butyne | 0 |
| **689-97-4** | 1-buten-3-yne | 0 |

**Table S3f VOC emission composition for diesel source (DPF)**

| CAS | Name | Mass fraction |
|---|---|---|
| | UNK | 0.244298674 |
| **75-07-0** | acetaldehyde | 0.225707262 |
| **78-85-3** | methacrolein | 0.118274705 |
| **74-85-1** | ethene | 0.086046842 |
| **78-93-3** | methylethyl-ketone | 0.056155762 |
| **74-98-6** | propane | 0.052065601 |
| **67-64-1** | acetone | 0.039655278 |
| **50-00-0** | formaldehyde | 0.032694644 |
| **74-84-0** | ethane | 0.030199727 |
| **115-11-7** | 2-methyl-propene | 0.01927645 |
| **123-38-6** | propanal | 0.017448155 |
| **124-18-5** | n-decane | 0.012309522 |
| **71-43-2** | benzene | 0.01148514 |
| **115-07-1** | propene | 0.008937853 |
| **4170-30-3** | crotonaldehyde | 0.007600155 |
| **1120-21-4** | n-undecane | 0.006662249 |
| **123-72-8** | butanal | 0.005667583 |
| **66-25-1** | hexanal | 0.005301555 |
| **108-88-3** | toluene | 0.003167642 |
| **106-99-0** | 1,3-butadiene | 0.002427048 |
| **100-52-7** | benzaldehyde | 0.002201009 |
| **78-78-4** | 2-methyl-butane | 0.002120476 |
| **75-28-5** | methyl-propane | 0.002033569 |
| **106-97-8** | n-butane | 0.001762 |
| **111-84-2** | n-nonane | 0.001524398 |
| **108-38-3** | m-xylene | 0.000779215 |
| **106-98-9** | 1-butene | 0.000759658 |
| **111-65-9** | n-octane | 0.000754035 |
| **100-41-4** | ethyl-benzene | 0.000678447 |
| **74-86-2** | ethyne | 0.000677368 |
| **95-47-6** | o-xylene | 0.000512759 |
| **463-49-0** | 1,2-propadiene | 0.000334792 |
| **110-62-3** | pentanal | 0.00014345 |
| **109-67-1** | 1-pentene | 0.000127798 |
| **624-64-6** | trans-2-butene | 0.000106502 |
| **107-02-8** | acrolein | 0.000102675 |
| **590-19-2** | 1,2-butadiene | 0 |
| **689-97-4** | 1-buten-3-yne | 0 |
| **107-00-6** | 1-butyne | 0 |
| **74-99-7** | 1-propyne | 0 |
| **463-82-1** | 2,2-dimethyl-propane | 0 |

| | | |
|---|---|---|
| **503-17-3** | 2-butyne | 0 |
| **563-46-2** | 2-methyl-1-butene | 0 |
| **563-45-1** | 3-methyl-1-butene | 0 |
| **590-18-1** | cis-2-butene | 0 |
| **627-20-3** | cis-2-pentene | 0 |
| **100-42-5** | styrene | 0 |
| **646-04-8** | trans-2-pentene | 0 |

**Table S4 Speciated IVOCs and their OH reaction rate constants (cm3 molec-1 s-1) and SOA yields at the OA concentration of 9 μg/m3 from Zhao et al. (2016)**

| Compound code | Compound name | OH reaction rate | SOA yield |
|---|---|---|---|
| **1** | Dodecane | 1.32E-11 | 0.08 |
| **2** | Tridecane | 1.51E-11 | 0.21 |
| **3** | Tetradecane | 1.68E-11 | 0.28 |
| **4** | Pentadecane | 1.82E-11 | 0.34 |
| **5** | Hexadecaen | 1.96E-11 | 0.38 |
| **6** | Heptadecane | 2.10E-11 | 0.42 |
| **7** | Octadecane | 2.24E-11 | 0.42 |
| **8** | Nonadecane | 2.38E-11 | 0.42 |
| **9** | Eicosane | 2.52E-11 | 0.42 |
| **10** | Heneicosane | 2.67E-11 | 0.42 |
| **11** | Docosane | 2.81E-11 | 0.42 |
| **12** | 2,6,10-Trimethylundecane | 1.70E-11 | 0.04 |
| **13** | 2,6,10-Trimethyldodecane | 1.87E-11 | 0.08 |
| **14** | 2,6,10-Trimethyltridecane | 2.01E-11 | 0.21 |
| **15** | 2,6,10-Trimethylpentadecane | 2.30E-11 | 0.34 |
| **16** | Pristane | 2.44E-11 | 0.34 |
| **17** | Phytane | 2.61E-11 | 0.38 |
| **18** | Hexylclohexane | 1.76E-11 | 0.08 |
| **19** | Heptylcyclohexane | 1.91E-11 | 0.21 |
| **20** | Octylcyclohexane | 2.05E-11 | 0.28 |
| **21** | Nonylcyclohexane | 2.19E-11 | 0.34 |
| **22** | Decylcyclohexane | 2.33E-11 | 0.38 |
| **23** | Undecylcyclohexane | 2.47E-11 | 0.42 |
| **24** | Dodecylcyclohexane | 2.61E-11 | 0.42 |
| **25** | Tridecylcyclohexane | 2.75E-11 | 0.42 |
| **26** | Tetradecylcyclohexane | 2.89E-11 | 0.42 |
| **27** | Pentadecylcyclohexane | 3.04E-11 | 0.42 |
| **28** | Hexadecylcyclohexane | 3.18E-11 | 0.42 |
| **29** | Heptadecylcyclohexane | 3.32E-11 | 0.42 |
| **30** | Naphthalene | 2.30E-11 | 0.21 |
| **31** | 2-methylnaphthalene | 4.86E-11 | 0.30 |
| **32** | 1-methylnaphthalene | 4.09E-11 | 0.25 |
| **33** | C2-naphthalene | 6.00E-11 | 0.31 |
| **34** | C3-naphthalene | 8.00E-11 | 0.31 |
| **35** | C4-naphthalene | 8.00E-11 | 0.31 |
| **36** | Acenaphthylene | 1.24E-10 | 0.31 |
| **37** | Acenaphthene | 8.00E-11 | 0.31 |

| | | | |
|---|---|---|---|
| **38** | Fluorene | 1.60E-11 | 0.31 |
| **39** | C1-Fluorene | 8.00E-11 | 0.31 |
| **40** | Phenanthrene | 3.20E-11 | 0.31 |
| **41** | Anthracene | 1.78E-10 | 0.31 |
| **42** | C1-Phenanthrene/anthracene | 5.89E-11 | 0.31 |
| **43** | C2-Phenanthrene/anthracene | 8.00E-11 | 0.31 |
| **44** | Fluoranthene | 3.30E-11 | 0.31 |
| **45** | Pyrene | 5.60E-11 | 0.31 |
| **46** | C1-Fluoranthene/pyrene | 1.31E-10 | 0.31 |
| **47** | Pentylbenzene | 1.01E-11 | 0.04 |
| **48** | Hexylbenzene | 1.15E-11 | 0.08 |
| **49** | Heptylbenzene | 1.30E-11 | 0.21 |
| **50** | Octylbenzene | 1.44E-11 | 0.28 |
| **51** | Nonylbenzene | 1.58E-11 | 0.34 |
| **52** | Decylbenzene | 1.72E-11 | 0.38 |
| **53** | Undecylbenzene | 1.86E-11 | 0.42 |
| **54** | Dodecylbenzene | 2.00E-11 | 0.42 |
| **55** | Tridecylbenzene | 2.14E-11 | 0.42 |
| **56** | Tetradecylbenzene | 2.29E-11 | 0.42 |
| **57** | Pentadecylbenzene | 2.43E-11 | 0.42 |

**Table S5a Surrogate compounds (*n*-alkanes) for OH reaction rate constants ($cm^3$ $molec^{-1}$ $s^{-1}$) and SOA yields of unspeciated IVOC bins under the IVOC-cyclic case from Zhao et al. (2016)**

| Bin# | OH rate constant | Surrogate compounds for SOA yields | |
|---|---|---|---|
| | | Unspeciated *b*-alkanes | Unspeciated cyclic compounds (IVOC-cyclic) |
| B12 | C12 | C10 | C12 |
| B13 | C13 | C11 | C13 |
| B14 | C14 | C12 | C14 |
| B15 | C15 | C13 | C15 |
| B16 | C16 | C14 | C16 |
| B17 | C17 | C15 | C17 |
| B18 | C18 | C16 | C18 |
| B19 | C19 | C17 | C19 |
| B20 | C20 | C18 | C20 |
| B21 | C21 | C19 | C21 |
| B22 | C22 | C20 | C22 |

**Table S5b Surrogate compounds (*n*-alkanes and naphthalenes) for OH reaction rate constants ($cm^3$ $molec^{-1}$ $s^{-1}$) and SOA yields of unspeciated cyclic compounds in each IVOC bin under the IVOC-aromatics case from Zhao et al. (2016)**

| Bin# | OH rate constant | Unspeciated cyclic compounds (IVOC-aromatic) |
|---|---|---|
| B12 | Naphthalene | Naphthalene |
| B13 | C1-naphthalene | 2-Methylnaphthalene |
| B14 | C2-naphthalene | 1,2-Dimethylnaphthalene |
| B15 | C3-naphthalene | C15 |
| B16 | C4-naphthalene | C16 |
| B17 | C17 | C17 |
| B18 | C18 | C18 |
| B19 | C19 | C19 |
| B20 | C20 | C20 |
| B21 | C21 | C21 |
| B22 | C22 | C22 |

**Table S6 Comparison of different estimates of IVOC fraction and overall SOA yield of NMOG emissions**

| | IVOC mass fraction | | | SOA yield | | |
|---|---|---|---|---|---|---|
| | **Gasoline** | **Gas-turbine** | **Diesel** | **Gasoline** | **Gas-turbine** | **Diesel** |
| **This work** | 4.6% | 27.9% | 54.3% | 0.041 | 0.086 | 0.190 |
| **Traditional** | / | / | / | 0.022 | 0.008 | 0.009 |
| **ROB** | 1.2% | 6.1% | 8.0% | 0.023 | 0.018 | 0.021 |
| **MUR** | 7.5% | 39.5% | 51.7% | N/A | N/A | N/A |
| **PYE** | 19.9% | 22.3% | 6.9% | 0.065 | 0.057 | 0.024 |
| **GEN** | 1% | N/A | 62% | 0.023 | N/A | 0.154 |
| **JAT** | 25% | N/A | 20% | 0.071 | N/A | 0.077 |

**Table S7a Volatility distributions for gasoline fuelled sources**

| Test ID | Vehicle name | Certification | Cycle | Non-volatile | -1 |
|---|---|---|---|---|---|
| **1027837** | LEV1-1 | Tier I | Cold UC | 0.001628419 | 0 |
| **1027852** | LEV2-4 | ULEV; Tier II | Cold UC | 0.002325422 | 0 |
| **1027863** | LEV2-13 | LEV II, ULEV; Tier II, Bin 5 | Cold UC | 0.004138927 | 0 |
| **1027867** | LEV2-9 | LEV II, ULEV; Tier II, Bin 5 | Cold UC | 0.001656056 | 0 |
| **1027872** | PreLEV-14 | Pre-LEV | Cold UC | 0.000708387 | 0.000787 |
| **1027905** | LEV2-19 | LEV II, ULEV; Tier II, Bin 5 | Cold UC | 0.001915585 | 0 |
| **1027907** | LEV2-20 | LEV II, ULEV | Cold UC | 0.001813142 | 0.000245 |
| **1027918** | LEV1-6 | LEV I, NLEV | Cold UC | 0.002281633 | 0.005443 |
| **1027921** | PreLEV-11 | Pre-LEV | Cold UC | 0.002133104 | 0.000104 |
| **1027970** | LEV1-26 | LEV I, ULEV | Cold UC | 0.003518602 | 4.58E-05 |
| **1027971** | LEV2-4 | ULEV; Tier II | Cold UC | 0.003787772 | 0 |
| **1027973** | LEV2-11 | LEV II, ULEV; Tier II, Bin 5 | Cold UC | 0.001053403 | 0.00069 |
| **1028021** | LEV2-18 | Tier II | Cold UC | 0.002258575 | 4.29E-05 |
| **1028023** | LEV1-25 | LEV I; Tier I | Cold UC | 0.00277297 | 0.008253 |
| **1028027** | LEV1-16 | LEV I, TLEV | Cold UC | 0.003218517 | 0 |
| **1028029** | PreLEV-10 | Pre-LEV | Cold UC | 0.001051418 | 4.44E-06 |
| **1028075** | LEV1-19 | LEV I | Cold UC | 0.001941466 | 2.17E-05 |
| **1032302** | LEV1-2 | LEV | Cold UC | 0.000993942 | 0.000414 |
| **1032304** | LEV1-2 | LEV | Cold UC | 0.000500407 | 9E-05 |
| **1032309** | LEV2-6 | LEV II, ULEV | Cold UC | 0.001493257 | 0.000152 |
| **1032320** | PreLEV-5 | LEV | Cold UC | 0.000885352 | 0.00067 |
| **1032321** | LEV2-6 | LEV II, ULEV | Cold UC | 0.001249077 | 0.000747 |
| **1032342** | LEV2-5 | ULEV | Cold UC | 0.003335295 | 0.004184 |
| **1032347** | LEV1-3 | LEV | Arterial | 0.003122842 | 0.007244 |
| **1032348** | LEV1-3 | LEV | Freeway | 0.011303132 | 0.020941 |
| **1032360** | LEV2-3 | LEV2O | Hot UC | 0.008903379 | 0.014291 |
| **1032383** | LEV2-24 | ULEV | Cold UC | 0.003249821 | 0.005709 |
| **1032388** | LEV1-21 | LEV | Cold UC | 0.001186722 | 0.002211 |
| **1032392** | PreLEV-4 | Tier I | Cold UC | 0.001315303 | 0.001075 |
| **1032393** | LEV1-4 | TLEV | Cold UC | 0.000567566 | 0.000498 |
| **1032426** | PreLEV-3 | Tier I | Hot UC | 0.001230972 | 0.000103 |
| **1032442** | PreLEV-1 | Tier I | Cold UC | 0.002268793 | 0.002457 |
| **1032443** | PreLEV-9 | Tier I | Cold UC | 0.001052782 | 0.000825 |
| **1032444** | PreLEV-2 | Tier I | Cold UC | 0.000959409 | 0.001283 |
| **1032445** | PreLEV-15 | Tier I | Cold UC | 0.001434841 | 0.001888 |
| **1032472** | LEV1-24 | LEV | Cold UC | 0.001130255 | 0.002817 |
| **1032431** | SORE4S-1.2 | | | 0.001218536 | 0.001301 |
| **1032439** | SORE2S-2.2 | | | 0.001762609 | 0.007281 |
| **1032463** | SORE4S-2.1 | | | 0.001698994 | 0.002875 |
| **1032464** | SORE4S-2.2 | | | 0.001675373 | 0.002932 |

**Table S7a (continued)**

| Test ID | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| **1027837** | 0.000323 | 0.002944 | 0.005377 | 0.001712 | 0.003248 | 0.011497 | 0.037531 | 0.104545 |
| **1027852** | 0.002154 | 0.018577 | 0.016471 | 0.004856 | 0.004855 | 0.007981 | 0.034545 | 0.026406 |
| **1027863** | 0.002059 | 0.020102 | 0.025805 | 0.014092 | 0.009346 | 0.015124 | 0.072661 | 0.136798 |
| **1027867** | 0.001144 | 0.009333 | 0.014395 | 0.004733 | 0.003162 | 0.006208 | 0.024719 | 0.055051 |
| **1027872** | 0.000547 | 0.001262 | 0.000992 | 0.001052 | 0.002026 | 0.00508 | 0.015499 | 0.060464 |
| **1027905** | 0.001928 | 0.009183 | 0.007127 | 0.003566 | 0.006513 | 0.011937 | 0.033422 | 0.065096 |
| **1027907** | 0.002202 | 0.008442 | 0.007634 | 0.005056 | 0.004139 | 0.00876 | 0.033258 | 0.10387 |
| **1027918** | 0.005037 | 0.008673 | 0.005217 | 0.004624 | 0.007128 | 0.017738 | 0.033897 | 0.074425 |
| **1027921** | 0.001009 | 0.004854 | 0.004903 | 0.004712 | 0.005758 | 0.015503 | 0.045199 | 0.072506 |
| **1027970** | 0.001379 | 0.009273 | 0.009004 | 0.009164 | 0.007557 | 0.019935 | 0.078973 | 0.096416 |
| **1027971** | 0.000683 | 0.004511 | 0.006934 | 0.00984 | 0.00639 | 0.016603 | 0.100723 | 0.056252 |
| **1027973** | 0.000973 | 0.003034 | 0.002568 | 0.002241 | 0.001569 | 0.004328 | 0.025113 | 0.060636 |
| **1028021** | 0.001129 | 0.010055 | 0.009557 | 0.004414 | 0.004681 | 0.013209 | 0.043779 | 0.111393 |
| **1028023** | 0.010196 | 0.012405 | 0.005649 | 0.010066 | 0.00732 | 0.013078 | 0.039685 | 0.062458 |
| **1028027** | 0.000683 | 0.004618 | 0.004213 | 0.003812 | 0.005072 | 0.020003 | 0.085388 | 0.093764 |
| **1028029** | 9E-05 | 0.000484 | 0.000538 | 0.001074 | 0.001543 | 0.006798 | 0.029907 | 0.089178 |
| **1028075** | 0.000525 | 0.003122 | 0.003288 | 0.003549 | 0.003659 | 0.013296 | 0.047212 | 0.077817 |
| **1032302** | 0.000542 | 0.006032 | 0.007643 | 0.00239 | 0.002171 | 0.003539 | 0.015498 | 0.078918 |
| **1032304** | 0.000264 | 0.002146 | 0.001426 | 0.001312 | 0.001164 | 0.002141 | 0.010703 | 0.075779 |
| **1032309** | 0.000647 | 0.007958 | 0.009179 | 0.003252 | 0.003505 | 0.007166 | 0.025574 | 0.073904 |
| **1032320** | 0.000437 | 0.002815 | 0.00221 | 0.001464 | 0.002012 | 0.004501 | 0.019943 | 0.081247 |
| **1032321** | 0.000433 | 0.002297 | 0.002416 | 0.00231 | 0.002779 | 0.006516 | 0.030544 | 0.089589 |
| **1032342** | 0.002252 | 0.013368 | 0.019293 | 0.009241 | 0.009666 | 0.018333 | 0.051944 | 0.08621 |
| **1032347** | 0.002524 | 0.009238 | 0.007937 | 0.00857 | 0.011585 | 0.020323 | 0.052688 | 0.199558 |
| **1032348** | 0.011669 | 0.035094 | 0.02725 | 0.030449 | 0.044967 | 0.083473 | 0.180894 | 0.069113 |
| **1032360** | 0.005492 | 0.037132 | 0.041905 | 0.027046 | 0.029663 | 0.044119 | 0.14279 | 0.052725 |
| **1032383** | 0.004186 | 0.025524 | 0.022233 | 0.019081 | 0.011316 | 0.015063 | 0.021881 | 0.096476 |
| **1032388** | 0.002505 | 0.006881 | 0.003423 | 0.002716 | 0.003314 | 0.005267 | 0.019326 | 0.08805 |
| **1032392** | 0.000967 | 0.003451 | 0.00234 | 0.001734 | 0.002699 | 0.006559 | 0.031763 | 0.092602 |
| **1032393** | 0.000491 | 0.003322 | 0.003686 | 0.002318 | 0.003307 | 0.003538 | 0.004669 | 0.065855 |
| **1032426** | 0.000339 | 0.001797 | 0.002026 | 0.002473 | 0.002595 | 0.006758 | 0.031254 | 0.065145 |
| **1032442** | 0.003369 | 0.029259 | 0.023495 | 0.001884 | 0.002318 | 0.005132 | 0.019348 | 0.075035 |
| **1032443** | 0.000608 | 0.003395 | 0.002389 | 0.001635 | 0.002562 | 0.006163 | 0.022914 | 0.146022 |
| **1032444** | 0.000691 | 0.001462 | 0.001165 | 0.001215 | 0.002343 | 0.007587 | 0.021155 | 0.089774 |
| **1032445** | 0.00102 | 0.001871 | 0.002208 | 0.002224 | 0.003293 | 0.00979 | 0.032893 | 0.092111 |
| **1032472** | 0.001419 | 0.003065 | 0.002944 | 0.00298 | 0.00258 | 0.005224 | 0.022443 | 0.095154 |
| **1032431** | 0.001211 | 0.001783 | 0.004071 | 0.00537 | 0.004547 | 0.005433 | 0.023152 | 0.05715 |
| **1032439** | 0.007017 | 0.01152 | 0.005256 | 0.003402 | 0.002973 | 0.003877 | 0.026467 | 0.091827 |
| **1032463** | 0.002501 | 0.003993 | 0.002745 | 0.003359 | 0.004706 | 0.00634 | 0.038827 | 0.078075 |
| **1032464** | 0.002619 | 0.004044 | 0.002603 | 0.00337 | 0.004228 | 0.005964 | 0.038678 | 0.071393 |

**Table S7a (continued)**

| Test ID | 8 | 9 | 10 | 11 | Total emission factor (VOC+IVOC+SVOC+LVOC+NV, mg/kg fuel) |
|---|---|---|---|---|---|
| **1027837** | 0.34316 | 0.28323 | 0.08974 | 0.115066 | 1197.045677 |
| **1027852** | 0.266282 | 0.30345 | 0.124656 | 0.187441 | 249.7791584 |
| **1027863** | 0.250371 | 0.206057 | 0.101977 | 0.141469 | 270.9979288 |
| **1027867** | 0.180129 | 0.389159 | 0.211024 | 0.099286 | 273.52397 |
| **1027872** | 0.388894 | 0.291441 | 0.079671 | 0.151577 | 6711.277796 |
| **1027905** | 0.265269 | 0.303474 | 0.128182 | 0.162386 | 257.9684118 |
| **1027907** | 0.333345 | 0.324761 | 0.130977 | 0.035496 | 391.1795365 |
| **1027918** | 0.258073 | 0.312597 | 0.118381 | 0.146488 | 7924.924112 |
| **1027921** | 0.231865 | 0.311065 | 0.137789 | 0.162599 | 2869.336247 |
| **1027970** | 0.276639 | 0.27757 | 0.110068 | 0.100458 | 422.2442435 |
| **1027971** | 0.244361 | 0.264108 | 0.117562 | 0.168246 | 374.2027651 |
| **1027973** | 0.294357 | 0.396121 | 0.059604 | 0.147712 | 3214.576957 |
| **1028021** | 0.294816 | 0.281955 | 0.150127 | 0.072583 | 482.9393712 |
| **1028023** | 0.306084 | 0.278146 | 0.122877 | 0.12101 | 1496.489328 |
| **1028027** | 0.269157 | 0.271891 | 0.124836 | 0.113344 | 1204.859468 |
| **1028029** | 0.299053 | 0.299552 | 0.10369 | 0.167036 | 4388.747026 |
| **1028075** | 0.293955 | 0.337219 | 0.106244 | 0.108151 | 869.3988182 |
| **1032302** | 0.293231 | 0.333571 | 0.112677 | 0.142381 | 1174.716757 |
| **1032304** | 0.303732 | 0.336229 | 0.116018 | 0.148495 | 1406.205838 |
| **1032309** | 0.277483 | 0.287782 | 0.121697 | 0.180207 | 571.2164043 |
| **1032320** | 0.281398 | 0.308193 | 0.141464 | 0.152761 | 2358.84952 |
| **1032321** | 0.299682 | 0.248253 | 0.134226 | 0.17896 | 709.7481613 |
| **1032342** | 0.291062 | 0.267085 | 0.0993 | 0.124727 | 438.4555372 |
| **1032347** | 0.234783 | 0.283334 | 0.087205 | 0.071888 | 54.68838036 |
| **1032348** | 0.190808 | 0.18471 | 0.054251 | 0.055079 | 90.35410772 |
| **1032360** | 0.214585 | 0.224929 | 0.063389 | 0.093031 | 100.9442248 |
| **1032383** | 0.243624 | 0.228635 | 0.159952 | 0.14307 | 429.8286328 |
| **1032388** | 0.300516 | 0.305276 | 0.106157 | 0.153171 | 931.1555508 |
| **1032392** | 0.304724 | 0.311941 | 0.103541 | 0.135288 | 6008.179735 |
| **1032393** | 0.246726 | 0.283774 | 0.270103 | 0.111146 | 875.3871449 |
| **1032426** | 0.294838 | 0.397397 | 0.080388 | 0.113656 | 2547.087082 |
| **1032442** | 0.271856 | 0.304769 | 0.0988 | 0.16001 | 9133.157736 |
| **1032443** | 0.335941 | 0.252885 | 0.106407 | 0.117201 | 4691.830657 |
| **1032444** | 0.291986 | 0.317343 | 0.132332 | 0.130705 | 28711.63898 |
| **1032445** | 0.283502 | 0.296956 | 0.111203 | 0.159607 | 13220.45075 |
| **1032472** | 0.300784 | 0.302598 | 0.101893 | 0.15497 | 1207.253523 |
| **1032431** | 0.208683 | 0.206668 | 0.159461 | 0.319952 | 1406.457345 |
| **1032439** | 0.316433 | 0.388595 | 0.089783 | 0.043807 | 119452.2234 |
| **1032463** | 0.22666 | 0.244437 | 0.141158 | 0.242625 | 20740.43167 |
| **1032464** | 0.213187 | 0.233717 | 0.162212 | 0.253379 | 20569.97937 |

**Table S7b Volatility distributions for gas-turbine sources**

| Test | Non-volatile | -1 | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|
| **4% load** | 0.001095331 | 0.005212 | 0.01101 | 0.016255 | 0.008797 | 0.008586 | 0.019535 |
| **85% load** | 0.056450185 | 0.018476 | 0.010983 | 0.015643 | 0.018957 | 0.029844 | 0.033446 |

| Test | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|
| **4% load** | 0.046212 | 0.198147 | 0.034958 | 0.077688 | 0.132911 | 0.209386 | 0.230207 |
| **85% load** | 0.041386 | 0.090459 | 0.075209 | 0.104446 | 0.192397 | 0.233367 | 0.078934 |

| Test | Total emission factor (VOC+IVOC+SVOC+LVOC+NV, mg/kg fuel) |
|---|---|
| **4% load** | 13256.32 |
| **85% load** | 1527.866 |

**Table S7c Volatility distributions for diesel fuelled sources**

| Test ID | Vehicle ID | After-treatment | Cycle | Non-volatile | -1 | 0 |
|---|---|---|---|---|---|---|
| **1406** | 2 | DOC + DPF + SCR | UDDS2 | 0.005801 | 0.000613 | 0.000634 |
| **1411** | 2 | DOC + DPF + SCR | UDDS2 | 0.006944 | 0.001028 | 0.001649 |
| **1413** | 2 | DOC + DPF + SCR | UDDS2 | 0.007211 | 0.000938 | 0.002588 |
| **1417** | 2 | DOC + DPF + SCR | UDDS2 | 0.006018 | 0.000782 | 0.00216 |
| **1418** | 2 | DOC + DPF + SCR | UDDS2 | 0.007364 | 0.000517 | 0.001258 |
| **1419** | 2 | DOC + DPF + SCR | CREEP3IDLE30 | 0.004615 | 4.11E-05 | 0.00011 |
| **1420** | 2 | DOC + DPF + SCR | CREEP3IDLE30 | 0.003694 | 3.29E-05 | 8.77E-05 |
| **1433** | 3 | none | UDDS2 | 0.004937 | 0.002134 | 0.004639 |
| **1434** | 3 | none | CREEP3IDLE30 | 0.00656 | 0.009365 | 0.008272 |
| **1435** | 3 | none | HHDDT50MPH3 | 0.005528 | 0.00345 | 0.004304 |
| **1437** | 3 | none | HHDDT50MPH3 | 0.005755 | 0.00437 | 0.005641 |
| **1439** | 3 | none | CREEP3IDLE30 | 0.007763 | 0.006893 | 0.005651 |
| **1440** | 3 | none | HHDDT50MPH3 | 0.006238 | 0.003458 | 0.004088 |
| **1441** | 3 | none | UDDS2 | 0.00602 | 0.001939 | 0.003127 |
| **1443** | 3 | none | CREEP3IDLE30 | 0.006069 | 0.006864 | 0.005242 |
| **1444** | 3 | none | HHDDT50MPH3 | 0.005908 | 0.031092 | 0.023742 |
| **1444** | 3 | none | HHDDT50MPH3 | 0.005304 | 0.003951 | 0.005458 |
| **1445** | 3 | none | UDDS2 | 0.005167 | 0.003739 | 0.004674 |
| **1450** | 1 | DOC + DPF | CREEP3IDLE30 | 0.003968 | 0.0009 | 0.002451 |
| **1458** | 1 | DOC + DPF | UDDS2 | 0.002597 | 4.57E-05 | 0.000292 |
| **1461** | 1 | DOC + DPF | UDDS2 | 0.003984 | 7.01E-05 | 0.000449 |
| **1028019** | 4 | DOC | Cold UC | 0.007056 | 0.001106 | 0.001385 |
| **1032421** | TRU | | | 0.028815 | 0.014355 | 0.012146 |
| **1032422** | TRU | | | 0.029073 | 0.014483 | 0.012255 |

**Table S7c (continued)**

| Test ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| **1406** | 0.366219 | 0.053298 | 0.007734 | 0.034229 | 0.067908 | 0.049448 | 0.000992 |
| **1411** | 0.060297 | 0.039716 | 0.037491 | 0.145914 | 0.23054 | 0.177791 | 0 |
| **1413** | 0.083815 | 0.08338 | 0.060841 | 0.151389 | 0.20374 | 0.134456 | 0 |
| **1417** | 0.069943 | 0.06958 | 0.050771 | 0.126332 | 0.170018 | 0.112202 | 0.122283 |
| **1418** | 0.133927 | 0.083349 | 0.030334 | 0.129965 | 0.203995 | 0.153017 | 0.003642 |
| **1419** | 0.008404 | 0.007449 | 0.00582 | 0.010152 | 0.166264 | 0.263216 | 0.218311 |
| **1420** | 0.006727 | 0.005963 | 0.004659 | 0.008126 | 0.13309 | 0.210697 | 0.333359 |
| **1433** | 0.016277 | 0.016332 | 0.036272 | 0.097038 | 0.180442 | 0.140537 | 0.095736 |
| **1434** | 0.020784 | 0.016623 | 0.03459 | 0.118383 | 0.222157 | 0.225791 | 0.141066 |
| **1435** | 0.014379 | 0.021784 | 0.045987 | 0.109787 | 0.164098 | 0.189053 | 0.091453 |
| **1437** | 0.015757 | 0.024377 | 0.054381 | 0.125271 | 0.177799 | 0.167892 | 0.076834 |
| **1439** | 0.015639 | 0.017937 | 0.043336 | 0.146241 | 0.294659 | 0.245904 | 0.080887 |
| **1440** | 0.012701 | 0.027651 | 0.065049 | 0.145708 | 0.215587 | 0.149584 | 0.072404 |
| **1441** | 0.010992 | 0.013605 | 0.036833 | 0.109185 | 0.204153 | 0.222186 | 0.085503 |
| **1443** | 0.014108 | 0.010857 | 0.055441 | 0.177694 | 0.184439 | 0.152206 | 0.136111 |
| **1444** | 0.063899 | 0.049175 | 0.063885 | 0.140315 | 0.121249 | 0.097444 | 0.070029 |
| **1444** | 0.016801 | 0.017265 | 0.073553 | 0.161551 | 0.1396 | 0.112191 | 0.080628 |
| **1445** | 0.013915 | 0.012387 | 0.050601 | 0.14333 | 0.148054 | 0.14 | 0.075854 |
| **1450** | 0.034124 | 0.024783 | 0.024084 | 0.06434 | 0.095491 | 0.150673 | 0.117575 |
| **1458** | 0.071106 | 0.014493 | 0.021468 | 0.046728 | 0.046789 | 0.058802 | 0.300269 |
| **1461** | 0.109078 | 0.022232 | 0.032932 | 0.071681 | 0.071774 | 0.090203 | 0.006995 |
| **1028019** | 0.008945 | 0.010469 | 0.037017 | 0.153741 | 0.253882 | 0.239092 | 0.127579 |
| **1032421** | 0.046923 | 0.035994 | 0.058055 | 0.127522 | 0.125312 | 0.300075 | 0.104096 |
| **1032422** | 0.047342 | 0.036315 | 0.058573 | 0.128662 | 0.126431 | 0.302755 | 0.099736 |

**Table S7c (continued)**

| Test ID | 8 | 9 | 10 | 11 | Total emission factor (VOC+IVOC+SVOC+LVOC+NV, mg/kg fuel) |
|---|---|---|---|---|---|
| **1406** | 0 | 0.390368 | 0.021057 | 0.001698 | 237.1973 |
| **1411** | 0.01198 | 0.255576 | 0.01905 | 0.012024 | 52.68983 |
| **1413** | 0.012035 | 0.23001 | 0.021805 | 0.007791 | 34.73976 |
| **1417** | 0.002007 | 0.21249 | 0.024177 | 0.031235 | 41.48411 |
| **1418** | 0 | 0.224097 | 0.017742 | 0.010793 | 49.61973 |
| **1419** | 0.134201 | 0.043589 | 0.045877 | 0.091952 | 2428.335 |
| **1420** | 0.117621 | 0.044799 | 0.051355 | 0.079791 | 3025.363 |
| **1433** | 0.047577 | 0.121348 | 0.164126 | 0.072607 | 1322.425 |
| **1434** | 0.053682 | 0.028171 | 0.046438 | 0.068119 | 5212.944 |
| **1435** | 0.051711 | 0.121276 | 0.122168 | 0.055021 | 1373.43 |
| **1437** | 0.048252 | 0.115709 | 0.121795 | 0.056166 | 1326.31 |
| **1439** | 0.041579 | 0.020232 | 0.02386 | 0.04942 | 7502.684 |
| **1440** | 0.029003 | 0.098878 | 0.116968 | 0.052682 | 1409.18 |
| **1441** | 0.024244 | 0.098442 | 0.118912 | 0.06486 | 1598.323 |
| **1443** | 0.055645 | 0.042362 | 0.053132 | 0.09983 | 6161.867 |
| **1444** | 0.053739 | 0.104065 | 0.116308 | 0.05915 | 1359.754 |
| **1444** | 0.061872 | 0.119814 | 0.133911 | 0.068102 | 1178.912 |
| **1445** | 0.07357 | 0.12173 | 0.131934 | 0.075044 | 1320.863 |
| **1450** | 0.029513 | 0.113557 | 0.040251 | 0.298289 | 317.1129 |
| **1458** | 0.04175 | 0.337552 | 0.046607 | 0.0115 | 95.6635 |
| **1461** | 0.033801 | 0.466079 | 0.088769 | 0.001954 | 62.61919 |
| **1028019** | 0.093705 | 0.061635 | 0.003747 | 0.00064 | 1161.057 |
| **1032421** | 0.048841 | 0.025755 | 0.030854 | 0.041258 | 1236.639 |
| **1032422** | 0.049695 | 0.025582 | 0.026655 | 0.042442 | 1225.69 |

## A.5. Figure S1 to S12

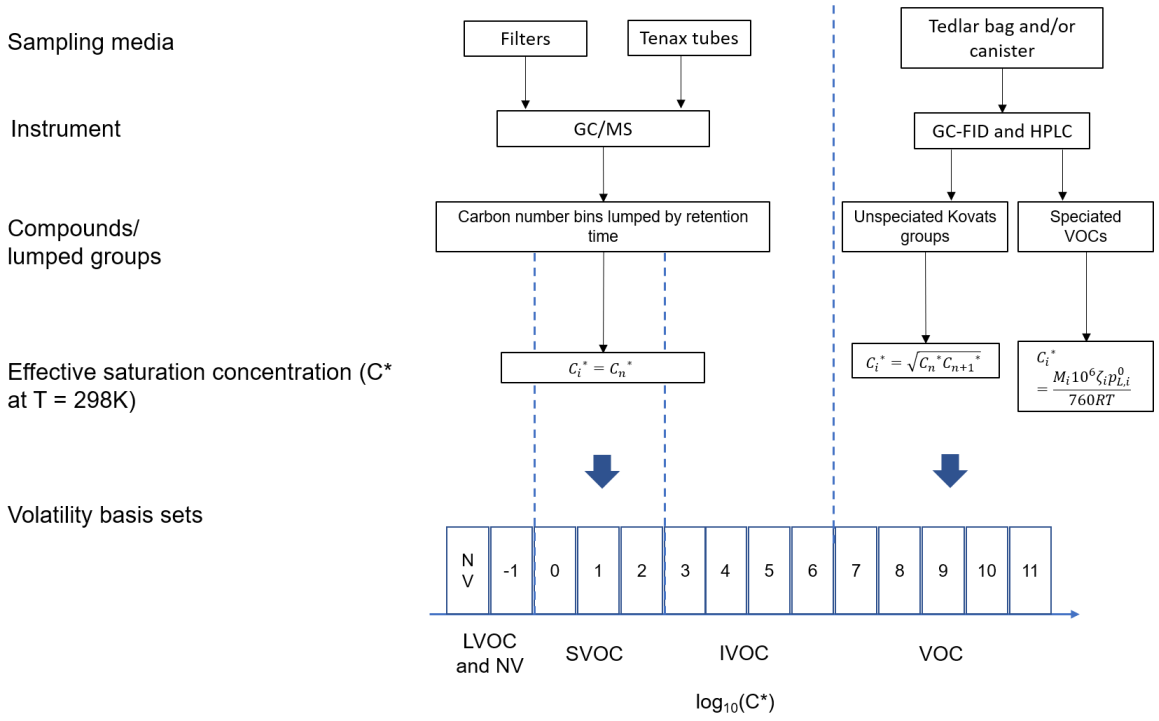
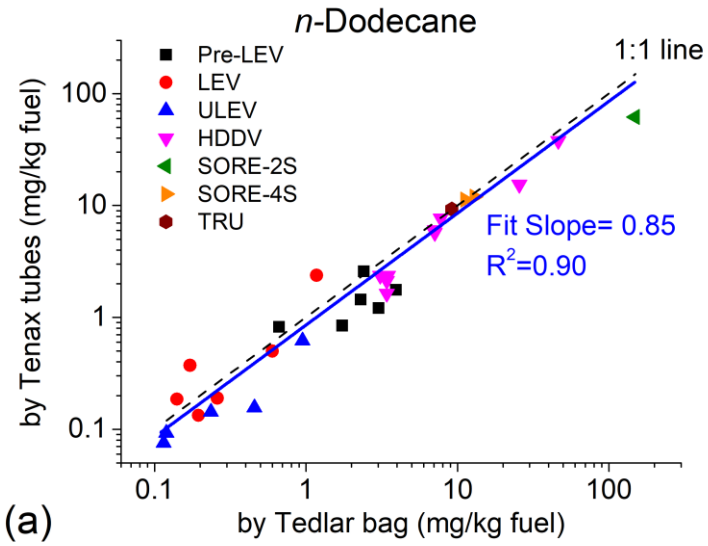
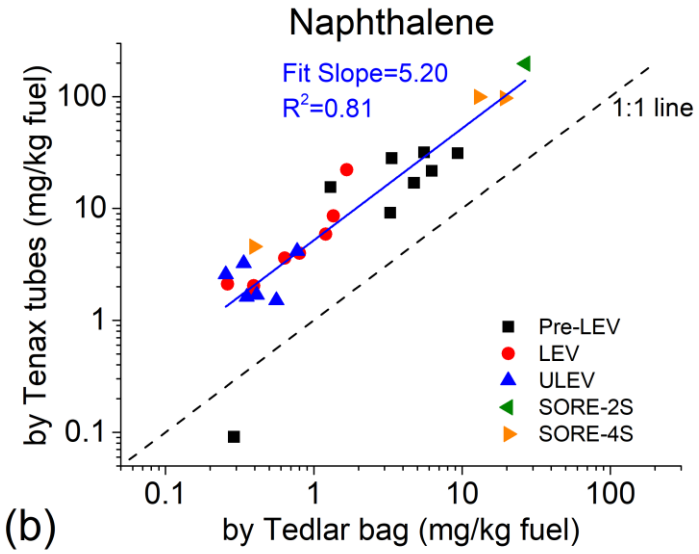



**Figure S1 Schematic diagram of mapping speciated and unspeciated compounds data to volatility basis set (VBS), $C_n^*$ denotes the C* value for n-alkanes as surrogate.**


n-Dodecane
Pre-LEV
LEV
ULEV
HDDV
SORE-2S
SORE-4S
TRU
1:1 line
Fit Slope= 0.85
R2=0.90
by Tenax tubes (mg/kg fuel)
by Tedlar bag (mg/kg fuel)
0.1
1
10
100
(a)


Naphthalene
Fit Slope=5.20
R2=0.81
1:1 line
Pre-LEV
LEV
ULEV
SORE-2S
SORE-4S
by Tenax tubes (mg/kg fuel)
by Tedlar bag (mg/kg fuel)
0.1
1
10
100
(b)

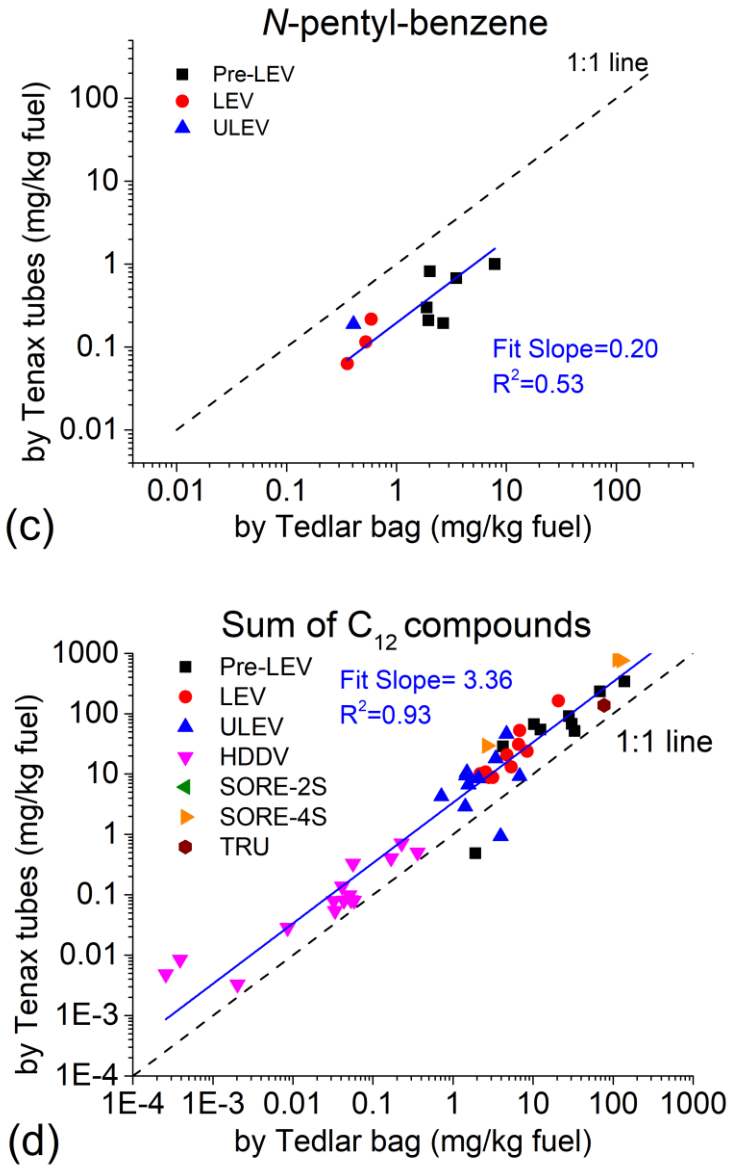



**Figure S2 Comparison of (a) *n*-dodecane (b) naphthalene (c) n-pentyl-benzene (d) sum of all $C_{12}$ compounds results measured using GC/MS analysis of Tedlar bag versus Tenax adsorbent samples (everything elutes in the $C_{12}$ carbon number bin), demonstrating the consistency of two technique in *n*-dodecane measurement**

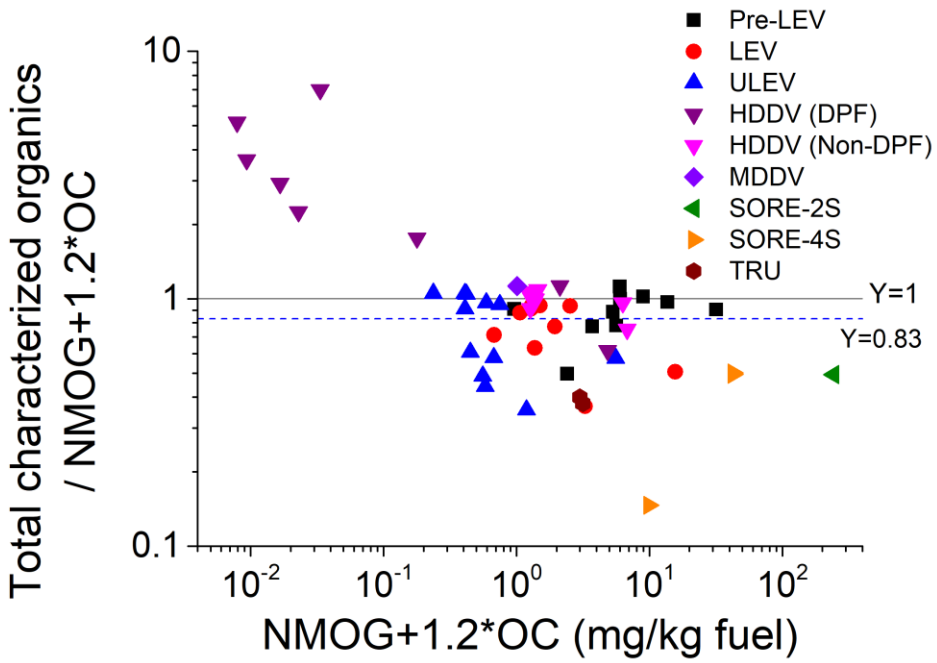



**Figure S3 Ratio of total characterized organics integrated from all techniques to total organics by bulk measurement (NMOG+1.2*OC), indicating mass closure for on-road non-DPF diesel source and partial mass closure (0.83) for on-road gasoline sources**

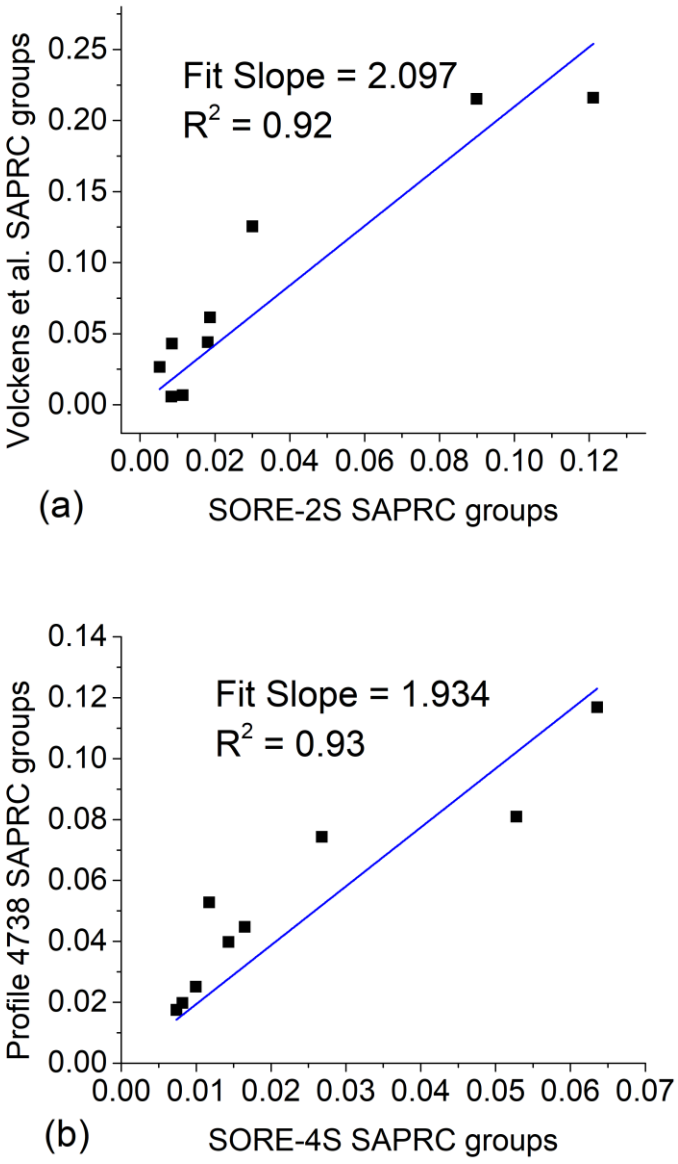



**Figure S4 Scatter plot of SOA-forming SAPRC groups (ALK4, ALK5, ARO1, ARO2, BENZ, TOLU, PXYL, MXYL, OXYL, B124) mass fraction in NMOG emission versus literature value for (a) SORE-2S (b) SORE-4S, indicating the need to factorize VOC mass in off-road measurement results by 2**

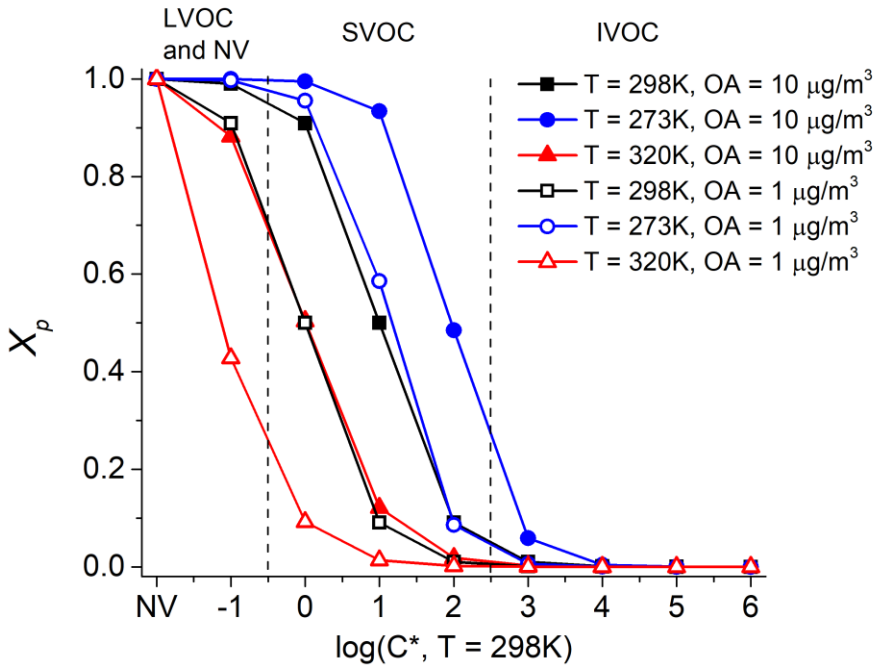
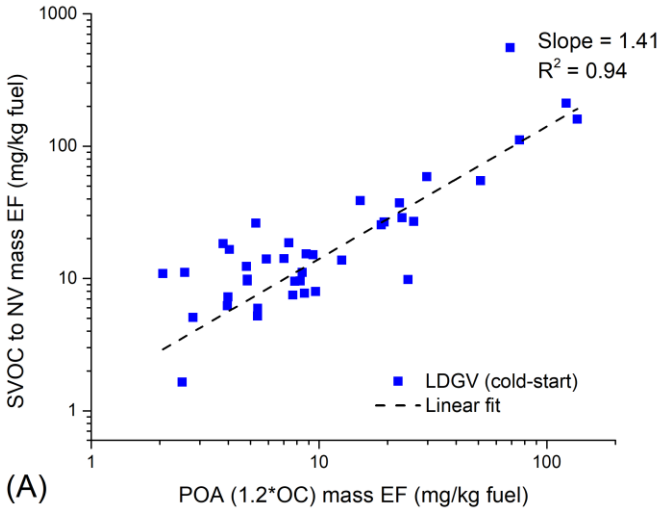
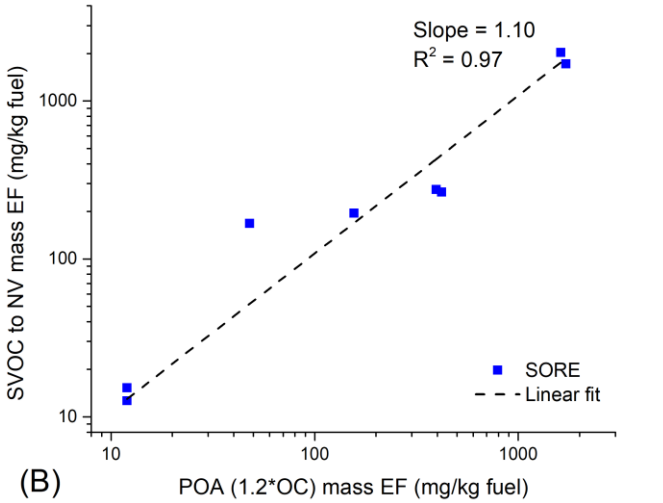



**Figure S5 Effect of temperature and OA loading on gas/particle partitioning at equilibrium**


Slope = 1.41
R2 = 0.94
LDGV (cold-start)
Linear fit
SVOC to NV mass EF (mg/kg fuel)
POA (1.2*OC) mass EF (mg/kg fuel)
1000
100
10
1
1
10
100

(A)


Slope = 1.10
R2 = 0.97
SORE
Linear fit
SVOC to NV mass EF (mg/kg fuel)
POA (1.2*OC) mass EF (mg/kg fuel)
1000
100
10
10
100
1000

(B)

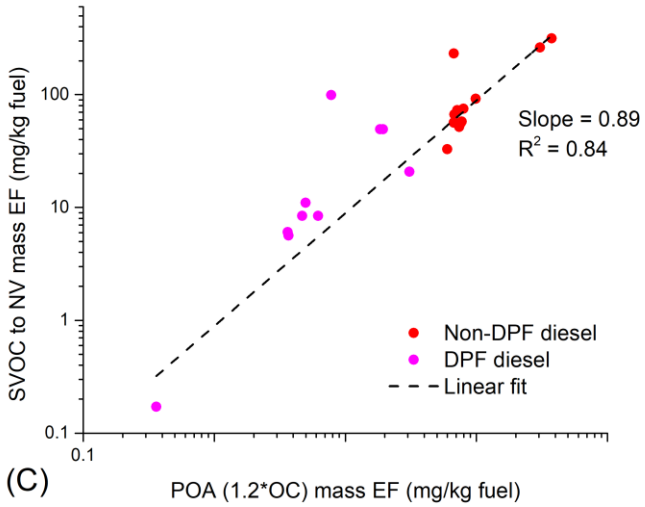



**Figure S6 Scatter plot of sum of SVOC to NV mass versus filter-based POA measurements (A: On-road gasoline, cold-start B: off-road gasoline, C: non-DPF and DPF diesel)**

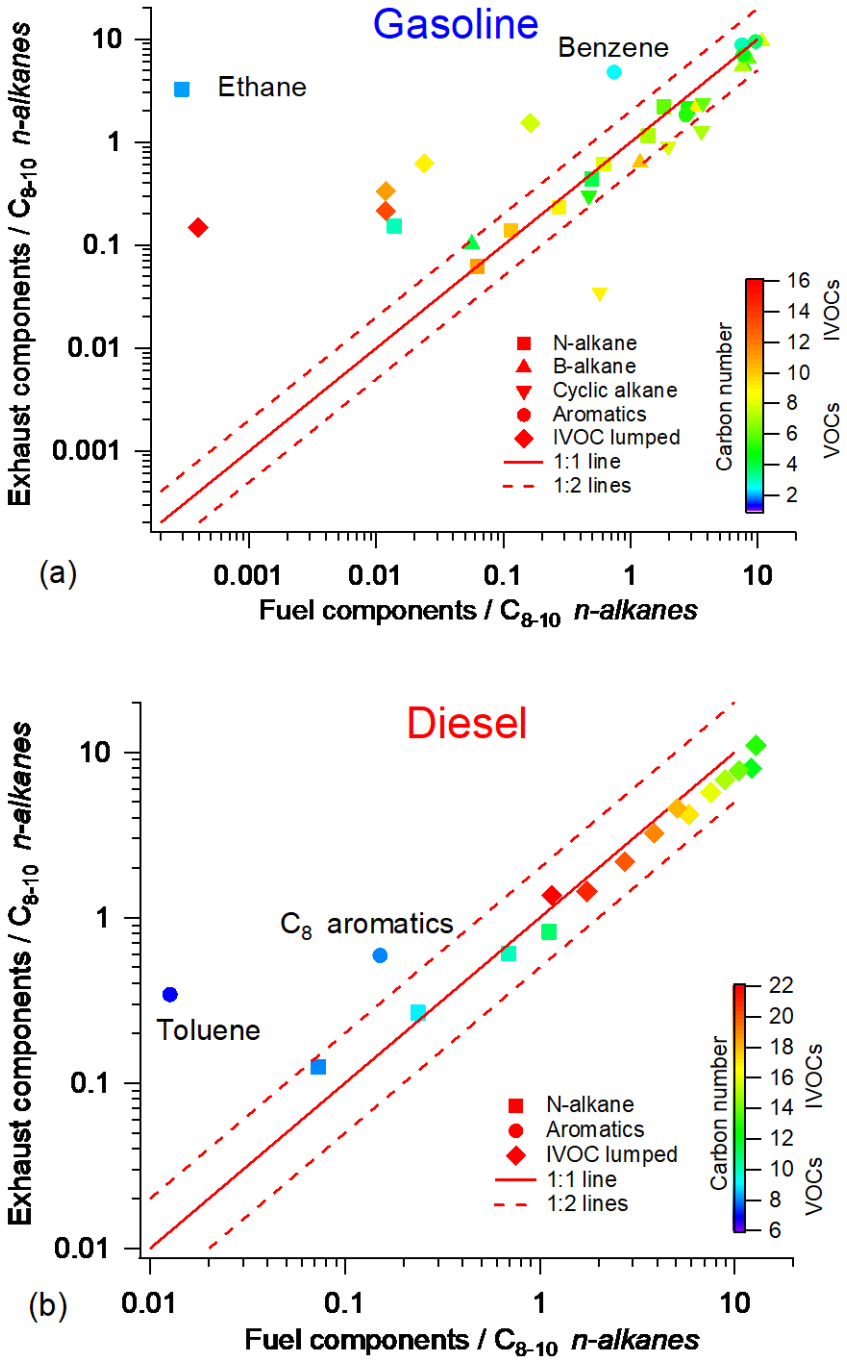



**Figure S7 Scatter plot of exhaust components versus fuel (VOCs and IVOCs) normalized by $C_{8\text{-}10}$ *n*-alkanes for (a) Gasoline (b) Diesel sources, demonstrating the overall consistency of chemical composition between exhaust and fuel**

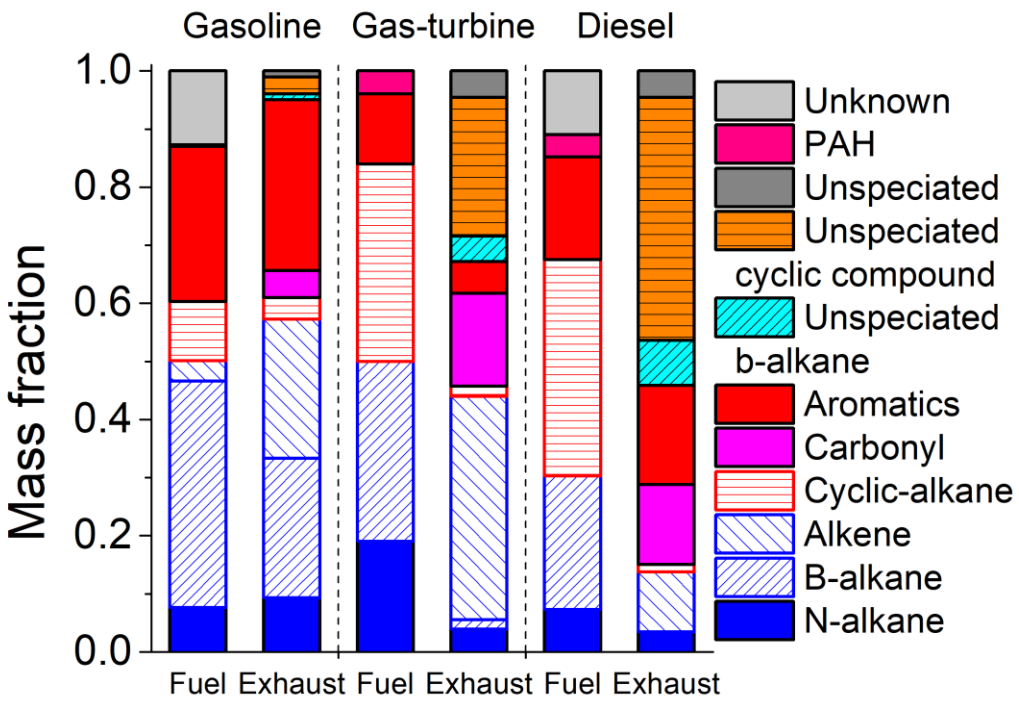



**Figure S8 Comparison of median chemical composition between fuel and exhaust for gasoline, gas-turbine and diesel sources, indicating the compositional changes after combustion**

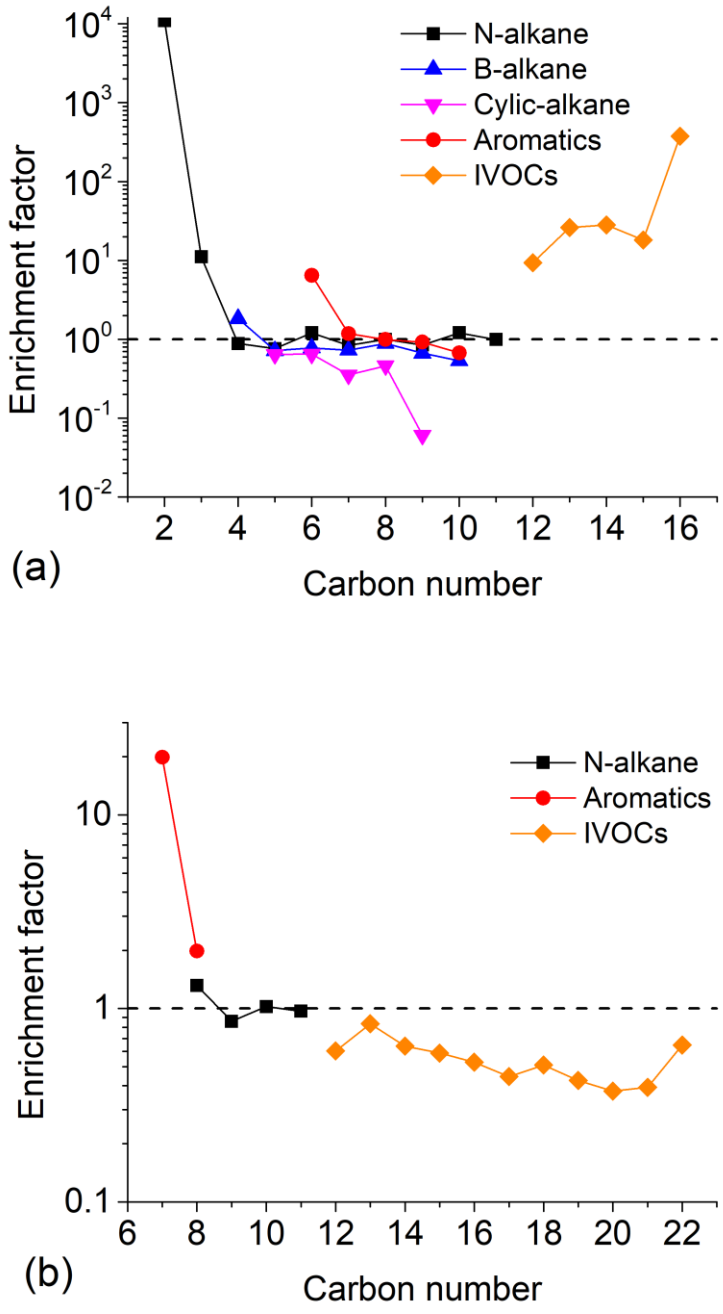



**Figure S9 Enrichment factors of exhaust and fuel components (VOCs and IVOCs) normalized by $C_{8\text{-}10}$ *n*-alkanes (a) Gasoline (cold-start) (b) non-DPF diesel, demonstrating the enrichment for certain compounds (VOCs and/or IVOCs) in difference sources**

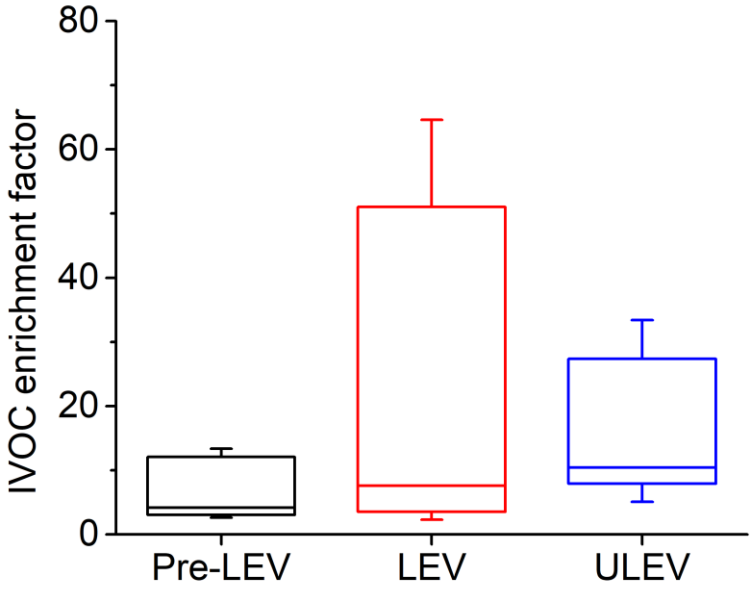



**Figure S10 IVOC enrichment factors of Pre-LEV, LEV and ULEV exhaust**

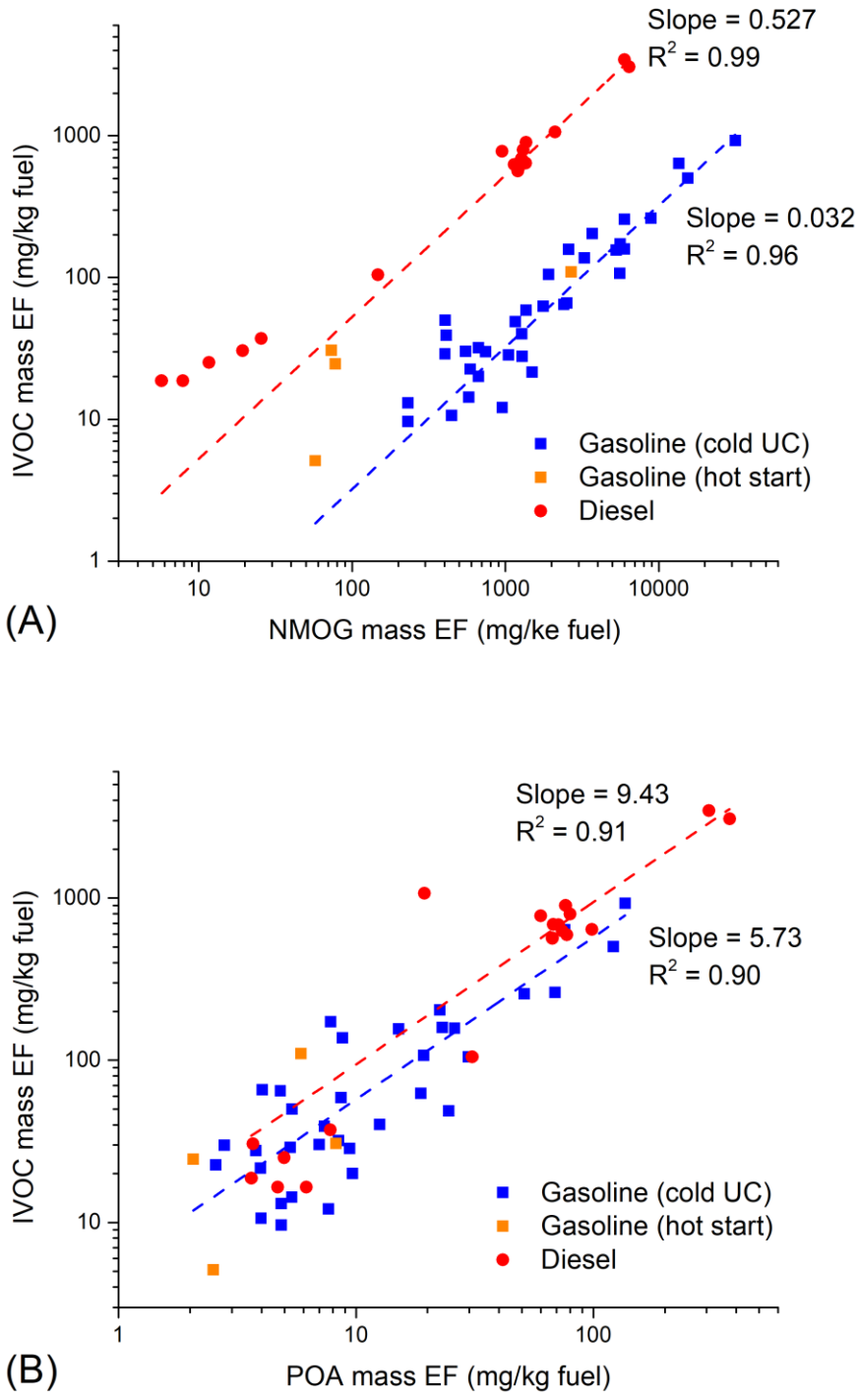



**Figure S11 Comparison of the ratios and coefficients of determination, $R^2$ between (a) IVOC and NMOG and (b) IVOC and POA for tested gasoline and diesel sources**

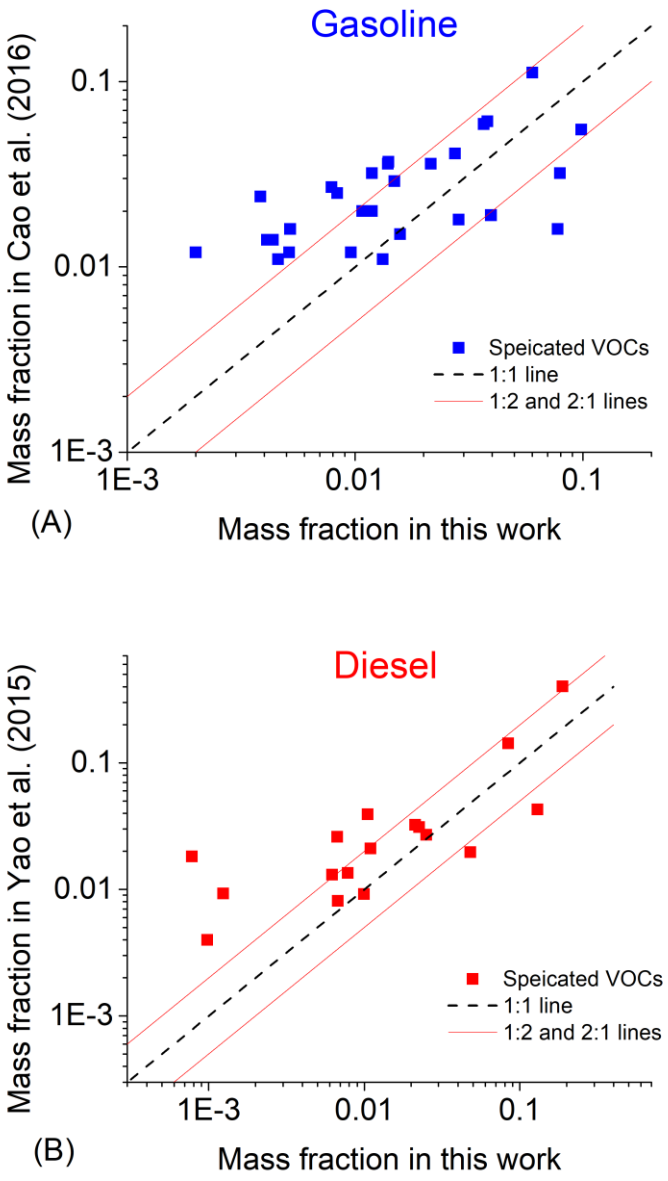



**Figure S12 Scatter plots of mass fractions of speciated VOCs in new VOC emission profiles versus: on-road gasoline (left) and on-road diesel (right) vehicle VOC emissions in China**

# Appendix B: Supporting information for Chapter 3

## B.1 Parameter fitting for SOA formation from lumped IVOC species

The loss term is defined as squared error between two surfaces: $m_{SOA,simp}(OA,t)$ and $m_{SOA,79}(OA,t)$:

$$Loss = \Sigma_{OA=1}^{10}\Sigma_{t=1}^{48}(m_{SOA,simp}(OA,t) - m_{SOA,79}(OA,t))^2 \quad (1)$$

which minimizes the squared distances between two surfaces in (OA concentration, time) space. Due to very high non-linearity in Eq. (1), the optimization is decoupled into step 1: '$k_{OH}$ fitting' and step 2: 'SOA yield fitting'.

Step 1: Relax the constrain on SOA yield to fit $k_{OH}$, Eq. (2) can be rewritten as,

$$m_{SOA,simp}(t) = \sum_j m_j\,\gamma_j f(k_{OH,j},t) = \sum_j m_j\,\gamma_j(1-\mathrm{e}^{-k_{OH,j}[OH]\Delta t}) \quad (2)$$

where $\gamma_j$ is the free variable representing SOA yield of surrogate j at given OA concentration, [OH] is assuming to be $3\times10^6$ $cm^{-3}$. Solving Eq. (2) with 2 unknowns: $k_{OH,j}$ and $\gamma_j$, $k_{OH,j}$ is the fitted OH reaction rate for the new lumped IVOC group.

Step 2: After solving for $k_{OH,j}$, we now eliminate the non-linearity in the time term of Eq. (2) by replacing unknown $f(k_{OH,j},t)$ with calculated reacted fraction $r_{j,t} = 1-\mathrm{e}^{-k_{OH,j}[OH]\Delta t}$ from fitted $k_{OH,j}$. Therefore, we can minimize the loss in Eq. (1) for each reduced IVOC groups,

$$Loss = \Sigma_{OA=1}^{10}\Sigma_{t=1}^{48}(\sum_{i\in j} m_{SOA,i}\,(OA,t) - \sum_{i\in j} m_{IVOC,i}\left[\alpha_{j,1}\xi_{OA,C^*=0.1} + \alpha_{j,2}\xi_{OA,C^*=1} + \alpha_{j,3}\xi_{OA,C^*=10} + \alpha_{j,4}\xi_{OA,C^*=100}\right]r_{j,t})^2 \quad (3)$$

where $\alpha_{j,1}$ to $\alpha_{j,4}$ are the fitted SOA parameterization for reduced IVOC group j. Minimization of the loss between $m_{SOA,simp,j}(OA,t)$ to $\sum_{i\in j} m_{SOA,i}\,(OA,t)$ is performed with the surface fitting toolbox in MATLAB.

## B.2 Equations

$$\mathrm{r} = \frac{\mathrm{Cov}\,(OA_{measured}, OA_{model})}{\mathrm{Var}(OA_{measured})\,\mathrm{Var}(OA_{model})} \quad (S1)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(OA_{measured}-OA_{model})^2}{N}} \quad (S2)$$

,where OA is the series of hourly-average value from measurements and model, and S1 and S2 are taking the statistics over hourly values.

$$\text{Fractional bias} = \frac{1}{N}\sum_{i=1}^{N}\frac{P-M}{\frac{P+M}{2}} \quad (S3)$$

$$\text{Fractional error} = \frac{1}{N}\sum_{i=1}^{N}\frac{|P-M|}{\frac{P+M}{2}} \quad (S4)$$

,where P is the predicted value, M is the measured value, and N is the sample size.

## B.3 Figure S1 to S7

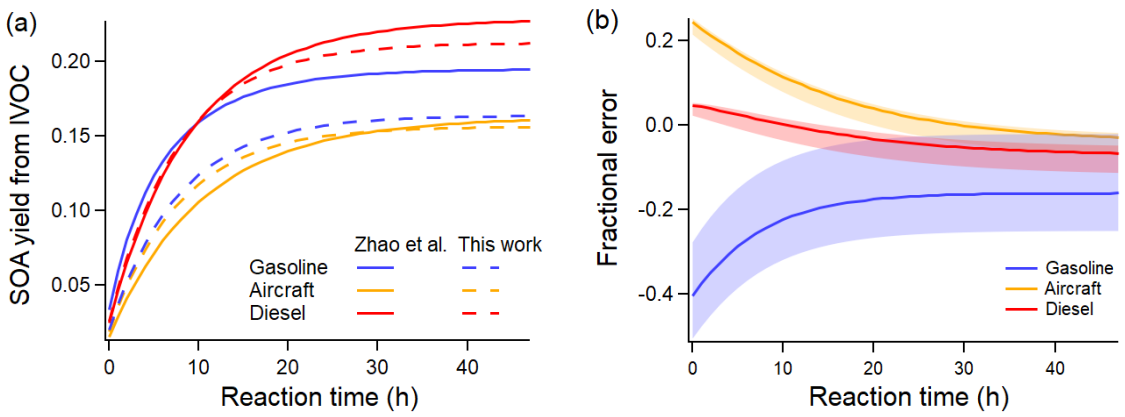



**Figure S1 (a) Comparison of predicted SOA formation per unit mass mobile IVOC emission using original and four-lumped-species parameterizations at OA = 5 μg $m^{-3}$, average [OH] = 3 × $10^{6}$ $cm^{-3}$ (b) Relative error in SOA formed between original and four-lumped-species parameterizations (Solid line is the relative error at OA = 5 μg $m^{-3}$, shaded area corresponds to OA = 1 to 50 μg $m^{-3}$)**

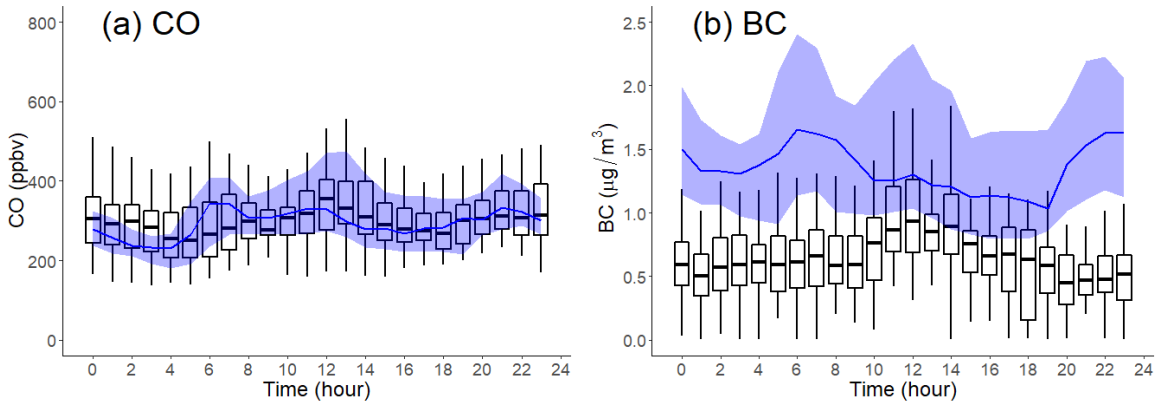



**Figure S2 Comparison of measured (boxplot, solid box denotes 25th to 75th percentiles and whiskers denote 10th to 90th percentiles) and modelled (line, shaded area denotes 25th to 75th percentiles) diurnal patterns in Pasadena, CA during CalNex for species: (a) CO (b) BC**

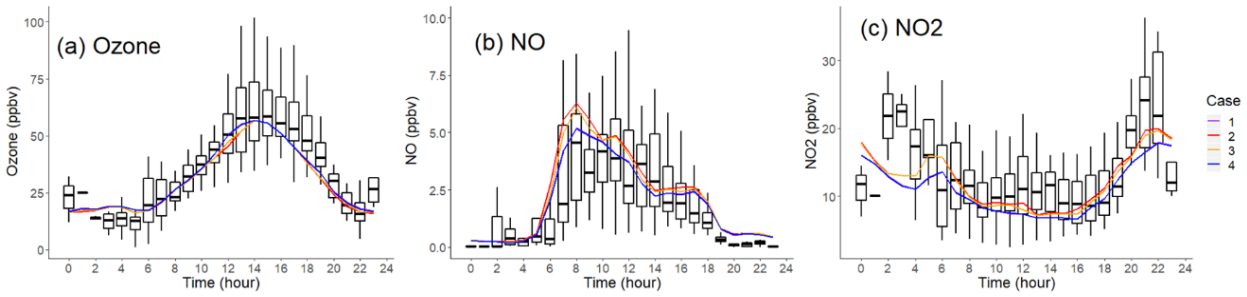



**Figure S3 Comparison of measured (boxplot, solid box denotes 25$^{th}$ to 75$^{th}$ percentiles and whiskers denote 10$^{th}$ to 90$^{th}$ percentiles) and modelled (line, from Case 1 to Case 4) diurnal patterns in Pasadena, CA during CalNex for species: (a) Ozone (b) NO and (c) $NO_2$**

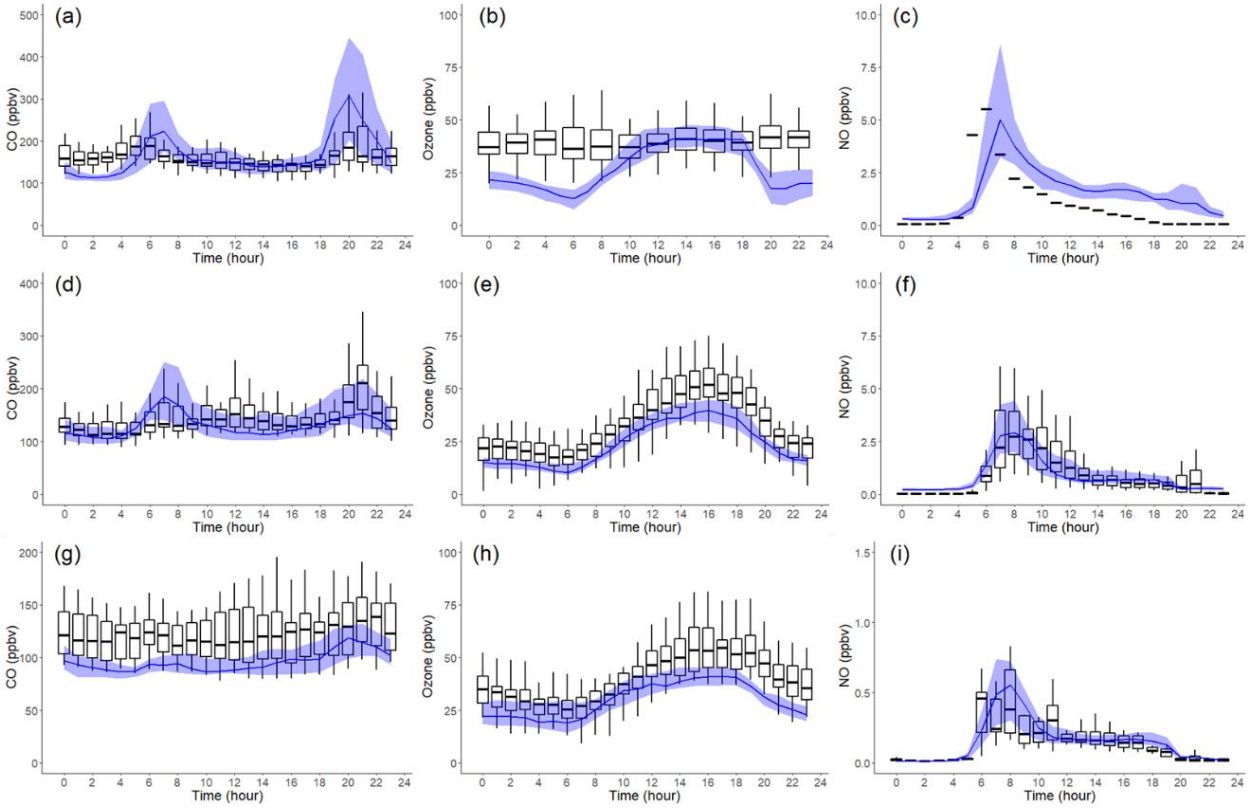



**Figure S4 Comparison of measured (boxplot, solid box denotes 25$^{th}$ to 75$^{th}$ percentiles and whiskers denote 10$^{th}$ to 90$^{th}$ percentiles) and modelled (line, shaded area denotes 25$^{th}$ to 75$^{th}$ percentiles) diurnal patterns during CalNex and CARES for species: CO, $O_3$ and NO in (a-c) Bakersfield, (d-f) Sacramento and (g-i) Cool**

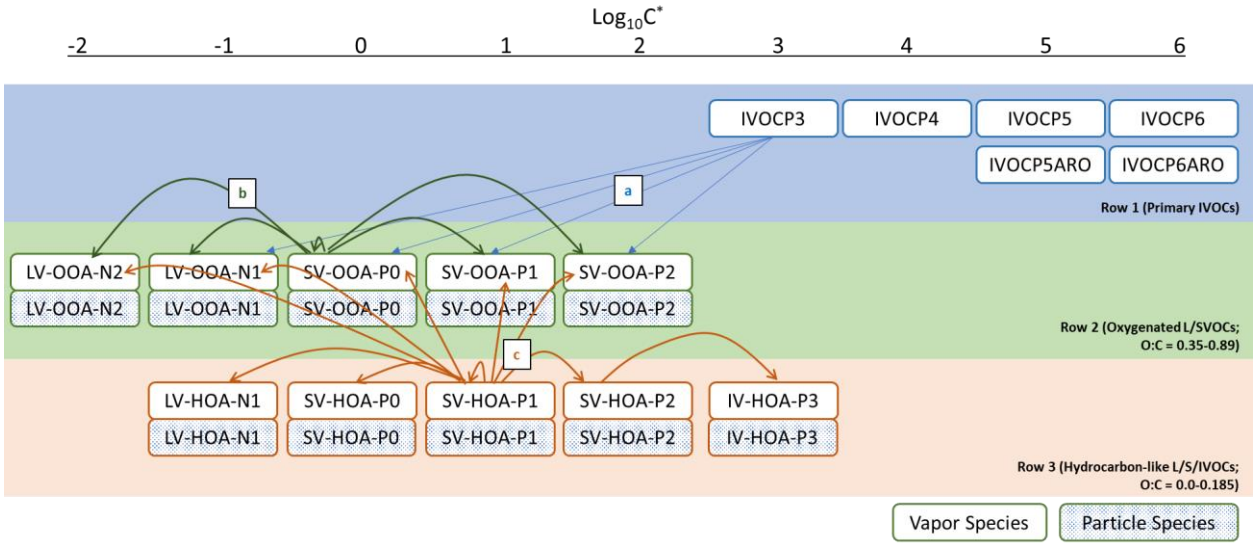



**Figure S5 Schematic of stoichiometry for OH oxidation first-generation and multigenerational aging. Species are segregated into primary IVOC species (row 1, blue), substantially oxygenated LVOC and SVOC species (row 2, green) and hydrocarbon-like or mildly oxygenated species (row 3, orange). Species in row 2 are closely aligned with SOA while species in row 3 are aligned with POA. Particle species are in equilibrium with associated vapor-phase species. Oxidation only occurs in the gas-phase. a) First-generation oxidation of primary IVOCs. All six species in Row 1 form products across the four OOA species in row 2 from LV-OOA-N1 to SV-OOA-P2. b) Multigenerational oxidation of oxygenated LVOCs and SVOCs. These reactions do not produce hydrocarbon-like species. Oxidation of all five vapor-phase species in Row 2 can produce mass in all five bins; thus functionalization and fragmentation pathways are represented. c) Multigenerational oxidation of hydrocarbon-like LVOCs, SVOCs, and IVOCs. These reactions may produce oxygenated or hydrocarbon-like species. Oxidation of all five vapor-phase species in Row 3 may produce mass in all ten vapor-phase species in Rows 2 and 3; thus functionalization and fragmentation are possible. Oxidation of species in row 2 is more likely to lead to fragmentation than is oxidation of species in row 3. Gas and particle emissions are applied to species in Rows 1 and 3.**

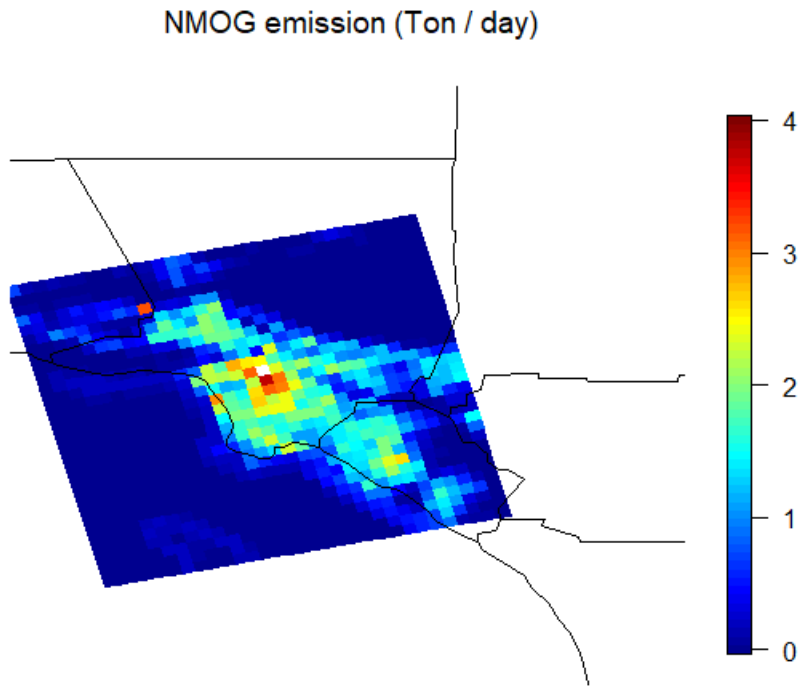



**Figure S6 (a) Los Angeles region in this study as defined by simulation grid cells (30 × 30 grid cell with 4 km resolution, equivalent to 120 km × 120 km)**

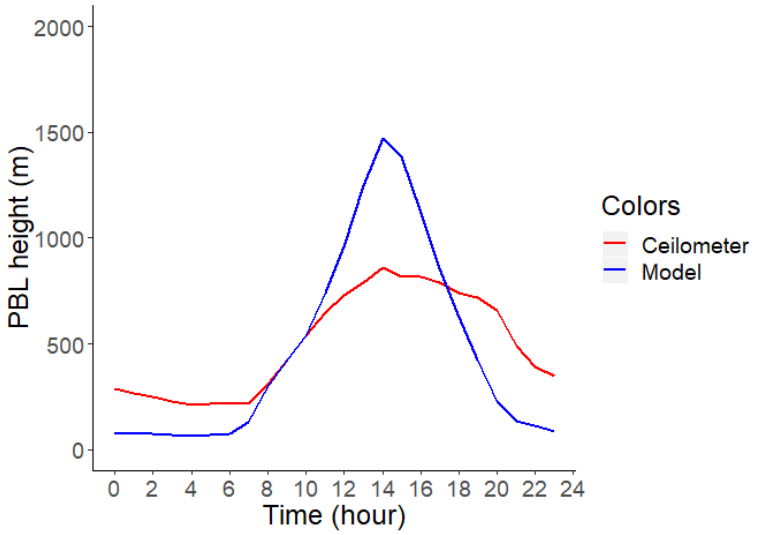



**Figure S7 Comparison of ceilometer measured (h1) and modelled PBL height diurnal patterns at Pasadena during CalNex (line denotes median value)**

## B.4 Table S1

**Table S1 Nomenclature of species in Figure S5 and CMAQ v5.3**

| Species Name in Figure S5 | Species Name in CMAQv5.3 (Gas/Particle) |
|---|---|
| LV-OOA-N2 | VLVOO1/ALVOO1 |
| LV-OOA-N1 | VLVOO2/ALVOO2 |
| SV-OOA-P0 | VSVOO1/ASVOO1 |
| SV-OOA-P1 | VSVOO2/ASVOO2 |
| SV-OOA-P2 | VSVOO3/ASVOO3 |
| LV-HOA-N1 | VLVPO1/ALVPO1 |
| SV-HOA-P0 | VSVPO1/ASVPO1 |
| SV-HOA-P1 | VSVPO2/ASVPO2 |
| SV-HOA-P2 | VSVPO3/ASVOO3 |
| SV-HOA-P3 | VIVPO1/AIVPO1 |

# Appendix C: Supporting information for Chapter 4

## C.1 Detailed description of GC-MS data pre-processing for PMF

(1) Read raw TIC, mass spectrum and retention times from .AIA files.

Using gcmstools, create a *data* object from raw .AIA file, export three matrixes (*data.TIC, data.intensity* and *data.times*) from each object. *data.TIC* is an 1D Numpy array of the total ion chromatogram (TIC) intensity values. *data.intensity* is an 2D Numpy array of raw MS intensity data. And *data.times* is a 1D Numpy array of the elution time points.

(2) Select IVOC range data from all (*data.TIC, data.intensity* and *data.times*) matrixes.

All samples were analyzed using the same TD-GC-MS setup, so the retention times for IVOC compounds are very close for all datasets. IVOCs are classified into 11 retention time bins ranging from $C_{12}$ n-alkane to $C_{22}$ n-alkanes. The starting retention time is the mean between $C_{11}$ and $C_{12}$ n-alkane, and the cut-off retention time is the mean between $C_{22}$ and $C_{23}$ n-alkane. Therefore, we plot the *data.times* to identify the IVOC starting and end scan numbers. We subset the original *data.intensity* matrixes based on scan numbers to get *data_ivoc* matrixes.

(3) Normalize *data_ivoc* matrixes to get *data_ivoc_norm* matrixes.

The absolute emission of IVOCs may vary several orders of magnitude, therefore, we focus on the relative contributions from all chemical species. To make every sample equally important in the following PMF analysis, *data_ivoc_norm* is computed by dividing *data_ivoc* by the sum of all elements in *data_ivoc*.

(4) Concatenate all *data_ivoc_norm* matrixes into *dataset_ivoc_norm* matrixes.

For all selected samples in a dataset, we concatenate the *data_ivoc_norm* matrixes along the rows to get a very long 2D matrixes *dataset_ivoc_norm.* The sum of all elements in *dataset_ivoc_norm* equals to the number of selected samples in the dataset.

## C.2 Table S1 to S3

**Table S1 List of biomass fuel used in biomass burning dataset**

| Sample no. | Fuel |
|---|---|
| 1 | Lodgepole Pine |
| 2 | Lodgepole Pine |
| 3 | Wire Grass |
| 4 | Saw Grass |
| 5 | Turkey Oak |
| 6 | Gallberry |
| 7 | Sage |
| 8 | AK Duff |
| 9 | Sage |
| 10 | White Spruce |
| 11 | Ponderosa Pine |
| 12 | Chamise |
| 13 | Lodgepole Pine |
| 14 | Pocosin |

**Table S2 PMF resolved factors in tunnel samples dataset**

| no. | Peak *m/z* | Group | Reference compound | Formula |
|---|---|---|---|---|
| 1 | 57 | n- and b-alkane | *n*-Dodecane | $C_{12}H_{26}$ |
| 2 | 55 | Cyclic alkane | Cyclopentane, 1-methyl-2-(4-methylpentyl)-, trans- | $C_{12}H_{24}$ |
| 3 | 55 | Cyclic alkane | Cyclopentane, 1-methyl-2-(4-methylpentyl)-, trans- | $C_{12}H_{24}$ |
| 4 | 83 | Cyclic alkane | Cyclohexane, undecyl- | $C_{17}H_{34}$ |
| 5 | 119 | Single-ring aromatics | Benzene, 1,2,4,5-tetramethyl- | $C_{10}H_{14}$ |
| 6 | 128 | PAH | Naphthalene | $C_{10}H_{8}$ |
| 7 | 142 | PAH | Naphthalene, 1-methyl- | $C_{11}H_{10}$ |
| 8 | 178 | PAH | Phenanthrene | $C_{14}H_{10}$ |
| 9 | 202 | PAH | Pyrene | $C_{16}H_{10}$ |
| 10 | 102 | Oxygenate | Phenyl maleic anhydride | $C_{10}H_{6}O_{3}$ |
| 11 | 105 | Oxygenate | Benzoic acid | $C_{7}H_{6}O_{2}$ |
| 12 | 122 | Oxygenate | Benzoic acid | |
| 13 | 66 | Internal standard | *n*-Dodecane-D26 | $C_{12}D_{26}$ |
| 14 | 136 | Internal standard | Naphthalene-D8 | $C_{10}D_{8}$ |
| 15 | 40 | Other | n/a | n/a |

**Table S3 PMF resolved factors in ambient samples dataset**

| no. | Peak *m/z* | Group | Reference compound | Formula |
|---|---|---|---|---|
| 1 | 55 | Alkane | Cyclopentane, 1-methyl-2-(4-methylpentyl)-, trans- | $C_{12}H_{24}$ |
| 2 | 57 | Alkane | *n*-Dodecane | $C_{12}H_{26}$ |
| 3 | 119 | Single-ring aromatics | Benzene, 1,2,4,5-tetramethyl- | $C_{10}H_{14}$ |
| 4 | 128 | PAH | Naphthalene | $C_{10}H_8$ |
| 5 | 142 | PAH | Naphthalene, 1-methyl- | $C_{11}H_{10}$ |
| 6 | 91 | Oxygenate | Benzeneacetic acid, 4-tetradecyl ester | $C_{22}H_{36}O_2$ |
| 7 | 102 | Oxygenate | Phenylmaleic anhydride | $C_{10}H_6O_3$ |
| 8 | 102 | Oxygenate | Phenylmaleic anhydride | $C_{10}H_6O_3$ |
| 9 | 104 | Oxygenate | 1,2-Benzenedicarboxylic acid | $C_8H_6O_4$ |
| 11 | 105 | Oxygenate | Benzoic acid | $C_7H_6O_2$ |
| 12 | 120 | Oxygenate | Salicylic acid | $C_7H_6O_3$ |
| 13 | 121 | Oxygenate | Benzoic acid, 3-hydroxy- | $C_7H_6O_3$ |
| 14 | 122 | Oxygenate | Benzoic acid | $C_7H_6O_2$ |
| 10 | 222 | Oxygenate | Methanone, 2-benzofuranylphenyl- | $C_{15}H_{10}O_2$ |
| 15 | 223 | Oxygenate | 1,3-Propanedione, 1,3-diphenyl- | $C_{15}H_{12}O_2$ |
| 16 | 66 | Internal standard | *n*-Dodecane-D26 | $C_{12}D_{26}$ |
| 17 | 191 | Other | n/a | n/a |

## C.3 Figure S1 to S18

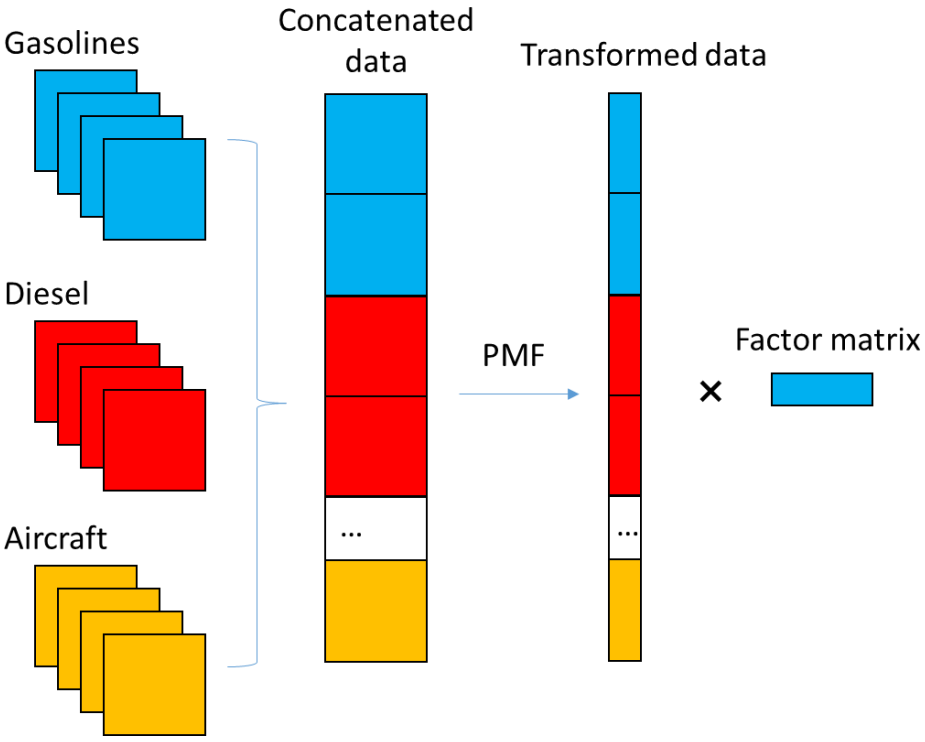



**Figure S1 Schematics plot of the overall process of PMF dimension reduction**

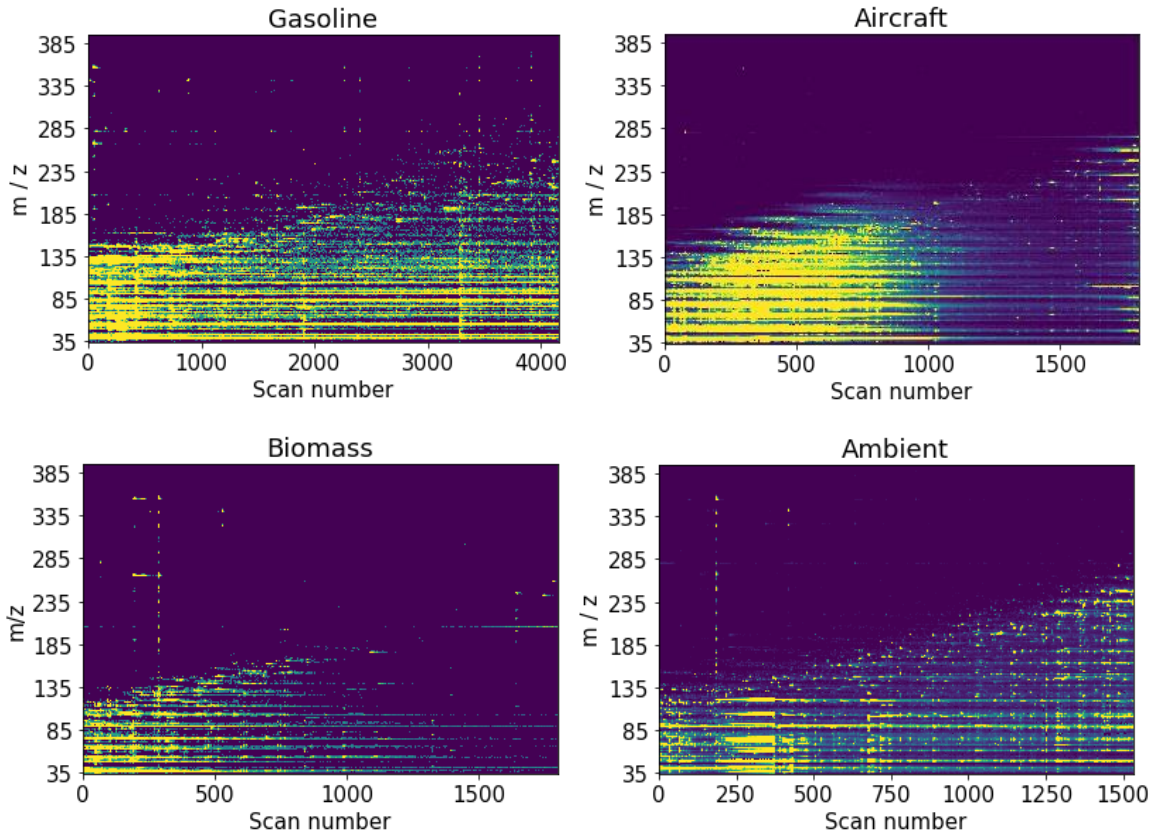
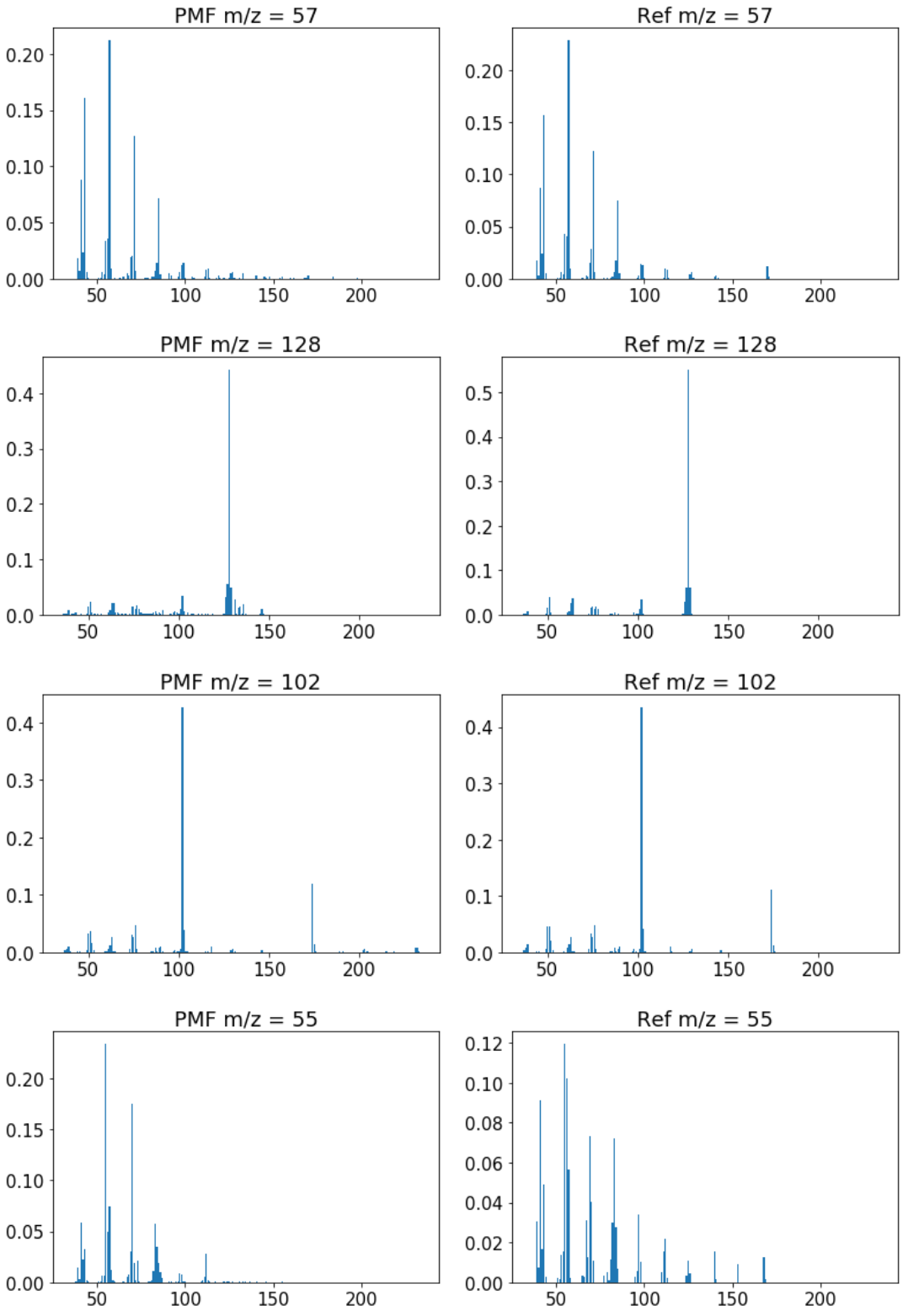



**Figure S2 Visualization of example GC-MS data matrix in the IVOC range**


PMF m/z = 57
Ref m/z = 57
PMF m/z = 128
Ref m/z = 128
PMF m/z = 102
Ref m/z = 102
PMF m/z = 55
Ref m/z = 55

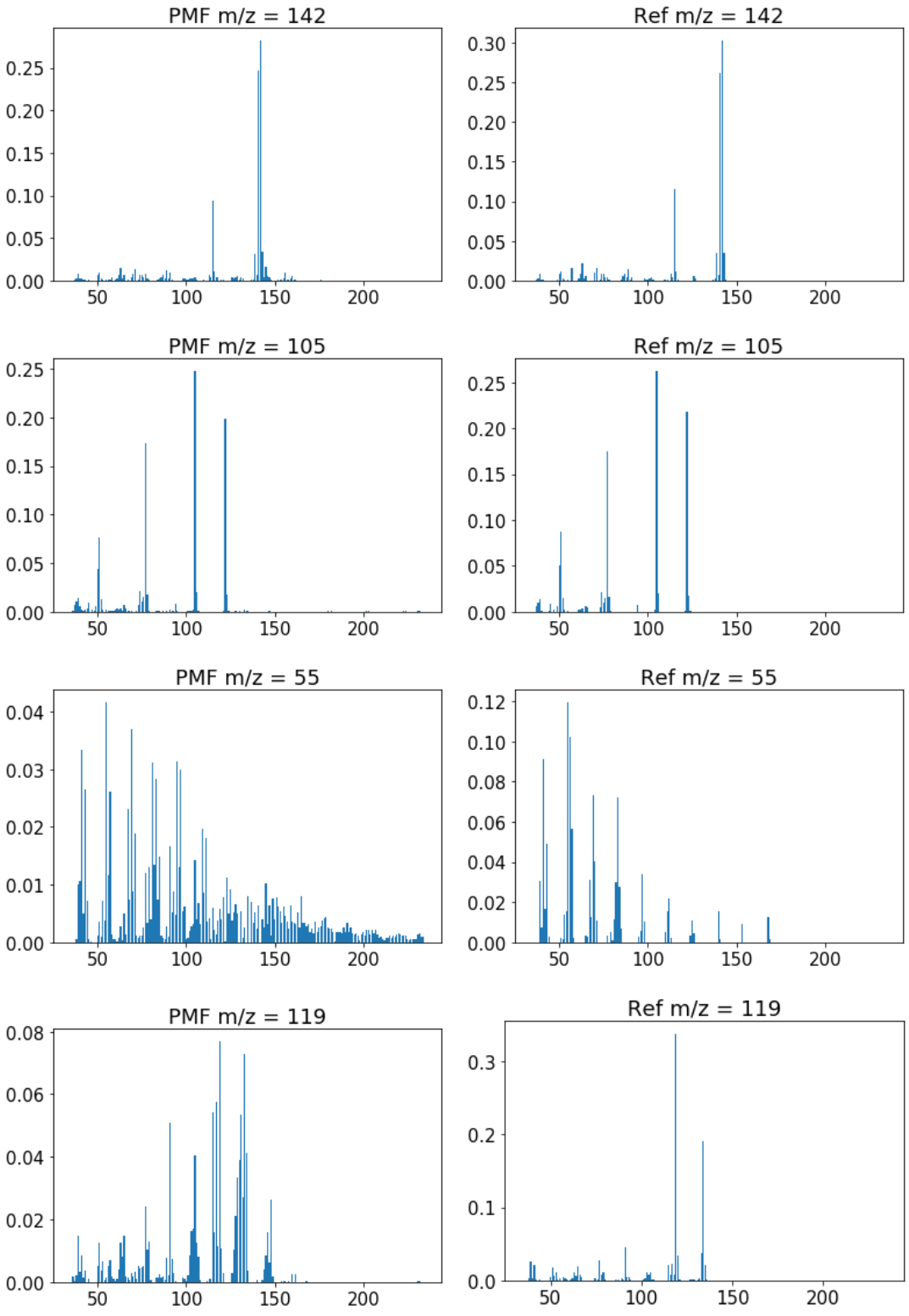



**Figure S3 Mass spectra of PMF resolved factors and reference compounds in mobile sources dataset**

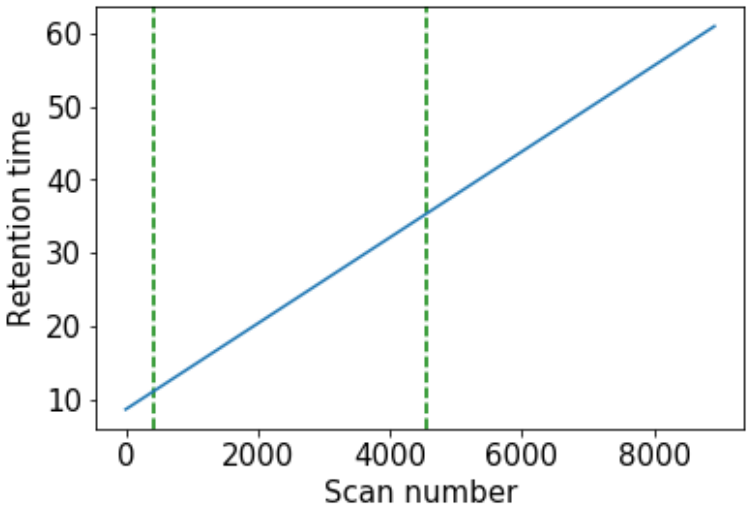



**Figure S4 Scan number and retention time relation for mobile-source dataset (dashed line: selected IVOC scan number range based on retention time)**

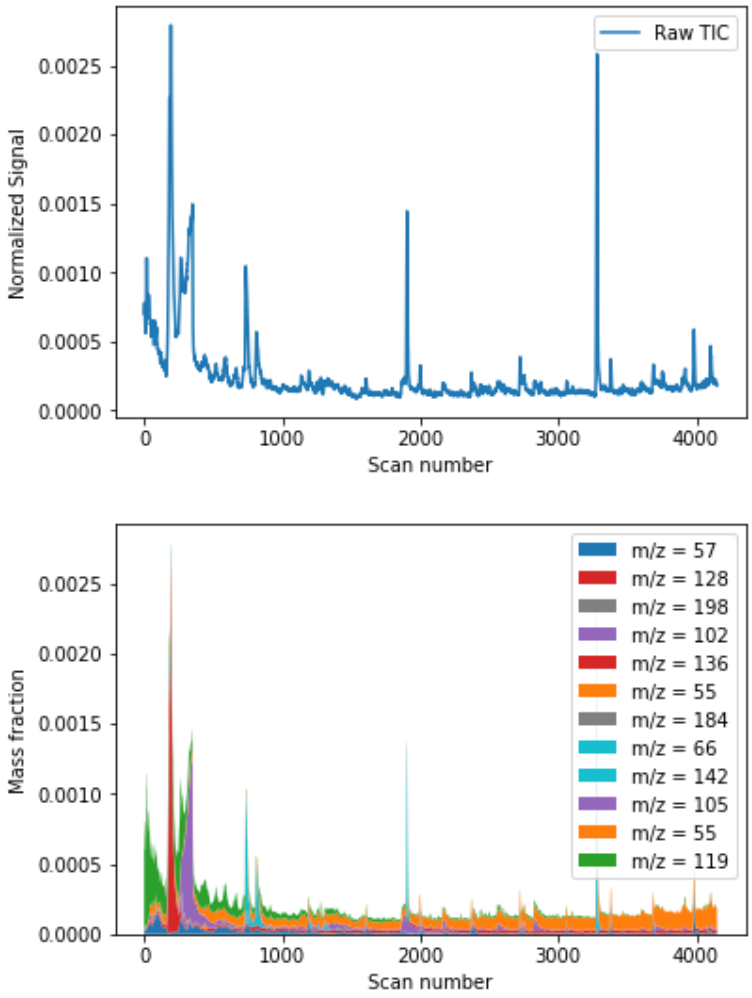



**Figure S5 (a) Example of gasoline vehicle exhaust raw TIC signal and (b) PMF resolved gasoline vehicle exhaust decomposed into 12 PMF factors**

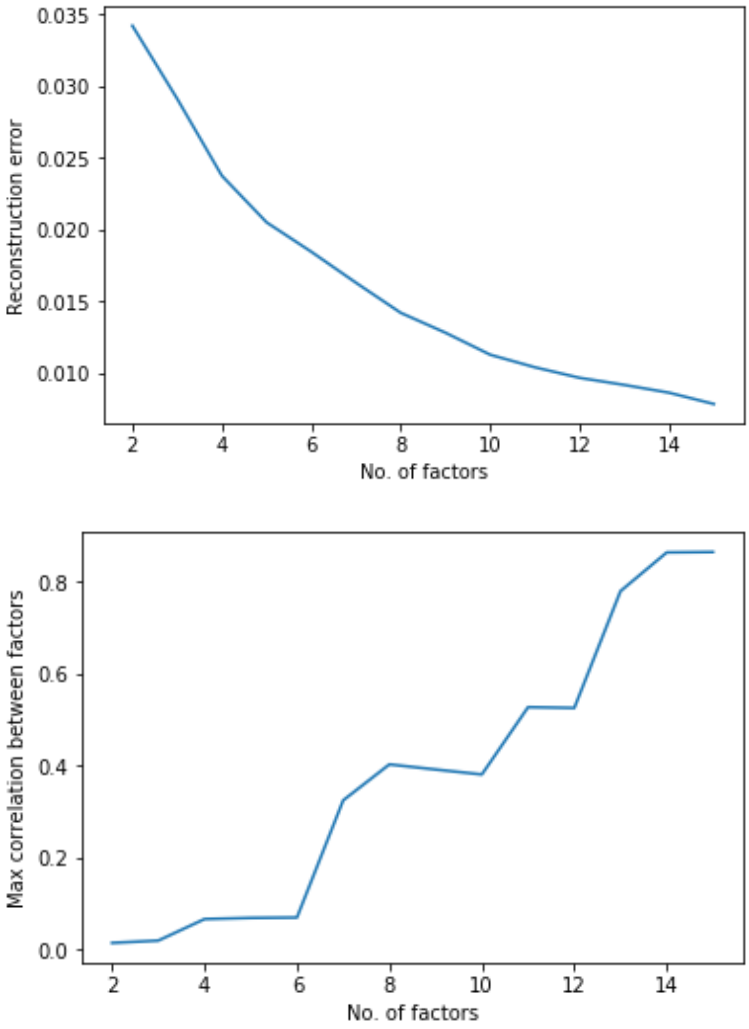



**Figure S6 Reconstruction error and max *r* as a function of PMF factors for mobile source emission dataset**

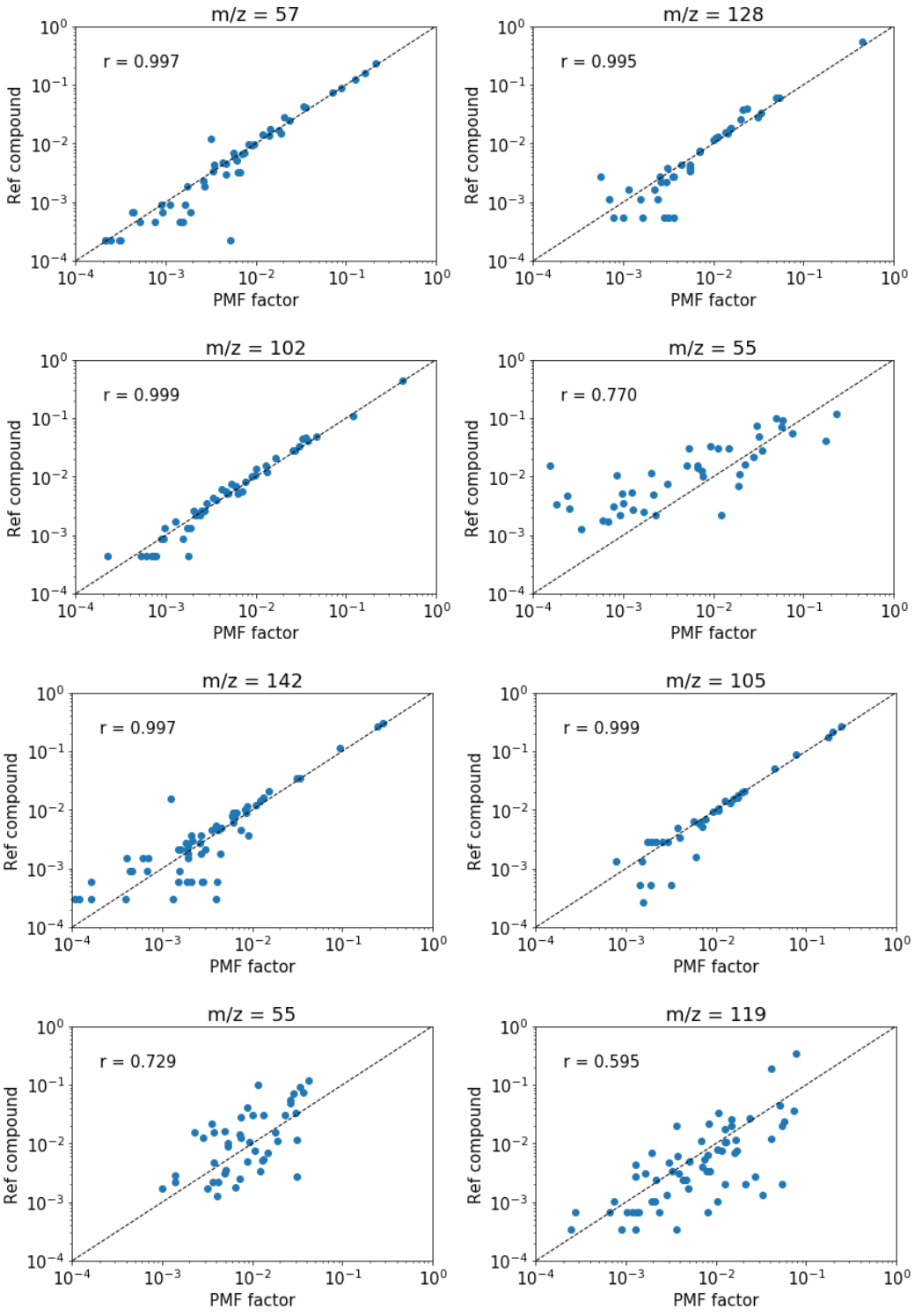



**Figure S7 Scatter plots of PMF resolved factors and reference compounds mass spectra in mobile sources dataset**

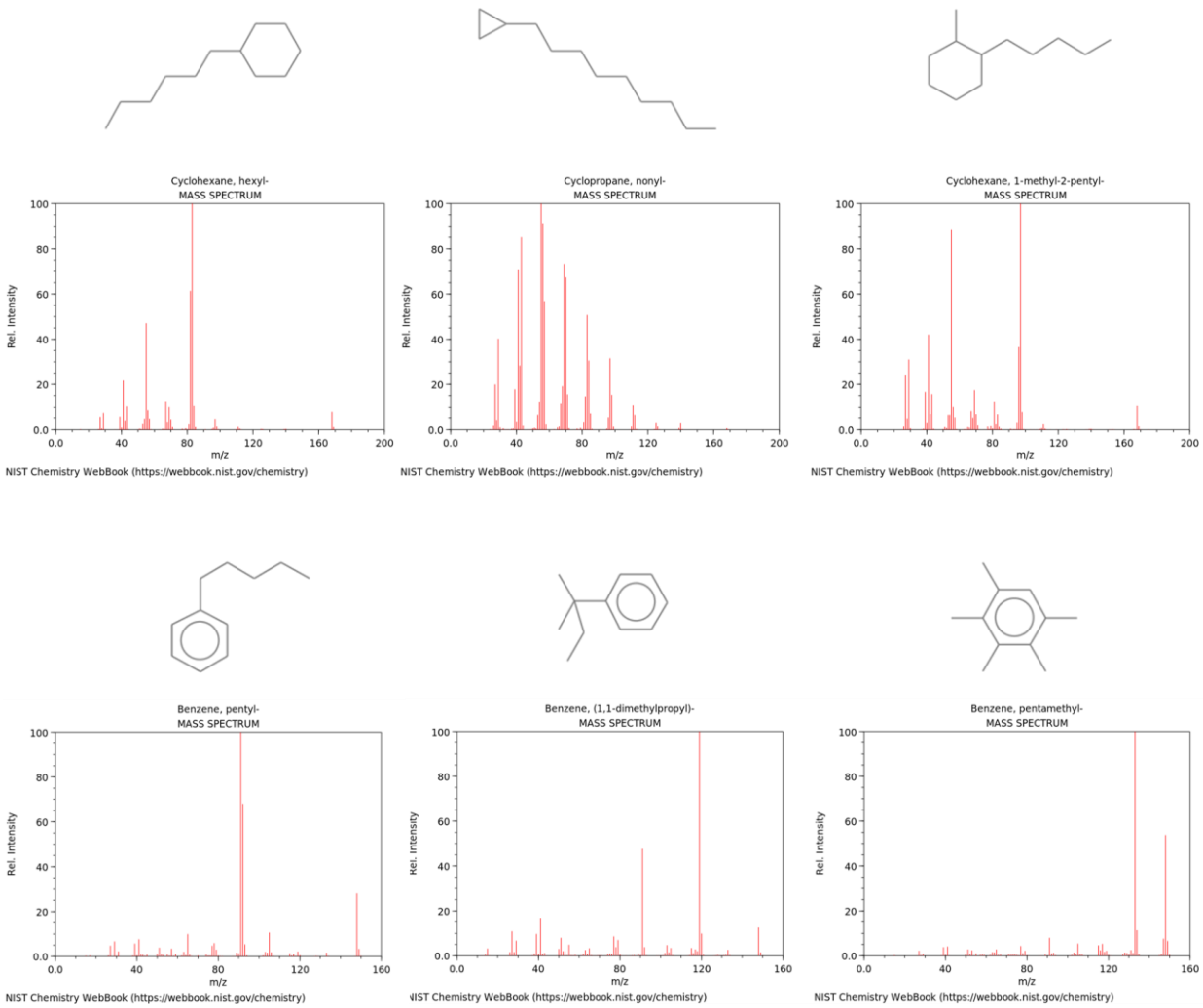



**Figure S8 Examples of mass spectra for cyclic alkanes and single-ring aromatics from NIST database**

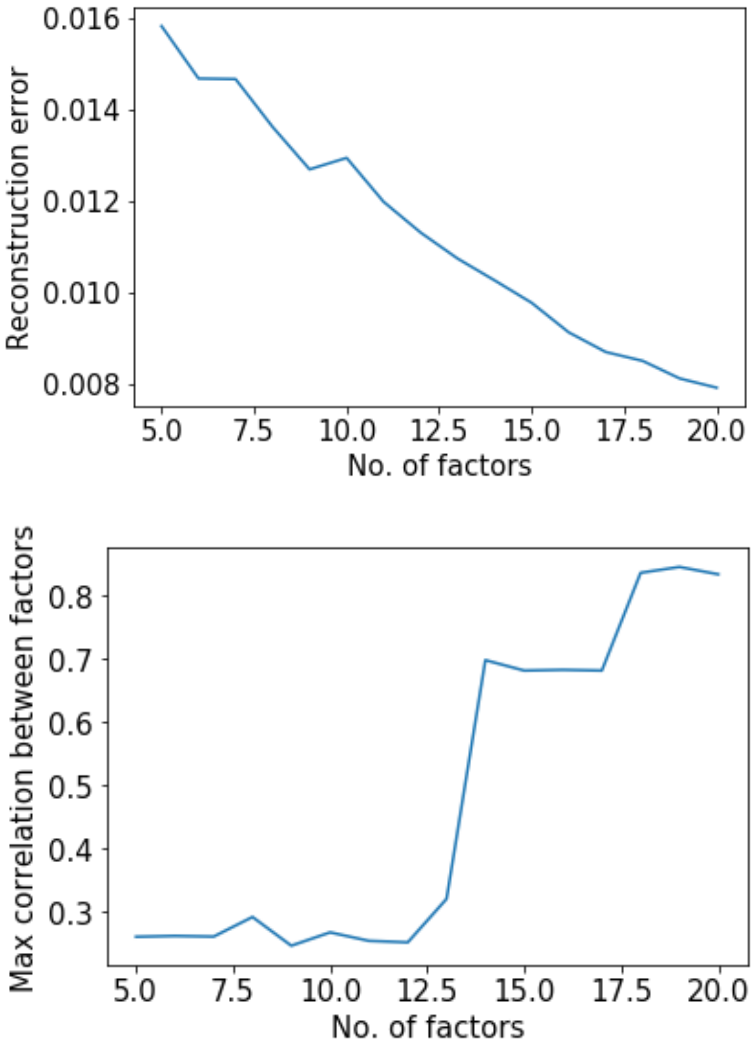
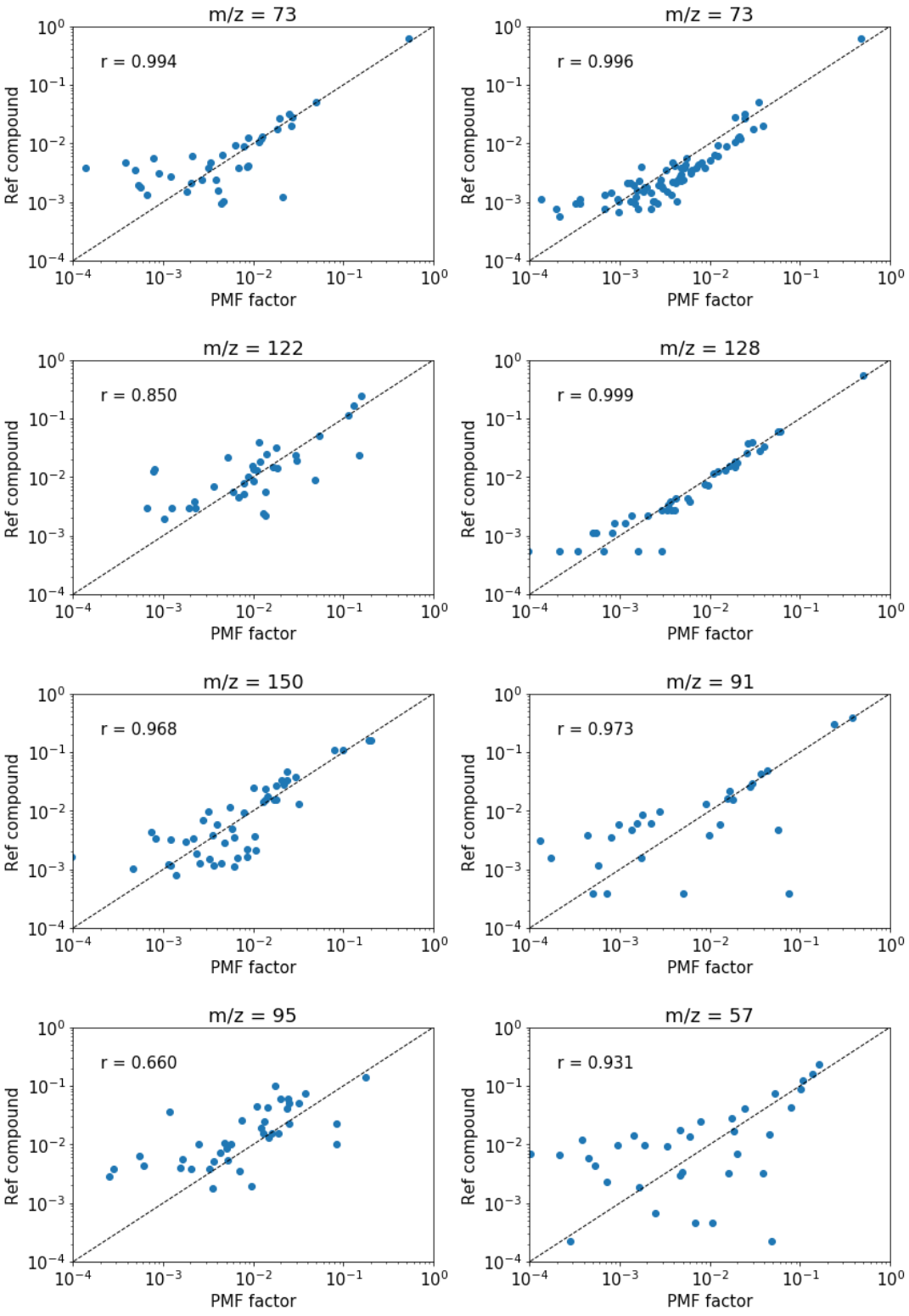



**Figure S9 Reconstruction error and max *r* as a function of PMF factors for biomass burning dataset**


m/z = 73
r = 0.994
m/z = 73
r = 0.996
m/z = 122
r = 0.850
m/z = 128
r = 0.999
m/z = 150
r = 0.968
m/z = 91
r = 0.973
m/z = 95
r = 0.660
m/z = 57
r = 0.931
Ref compound
PMF factor

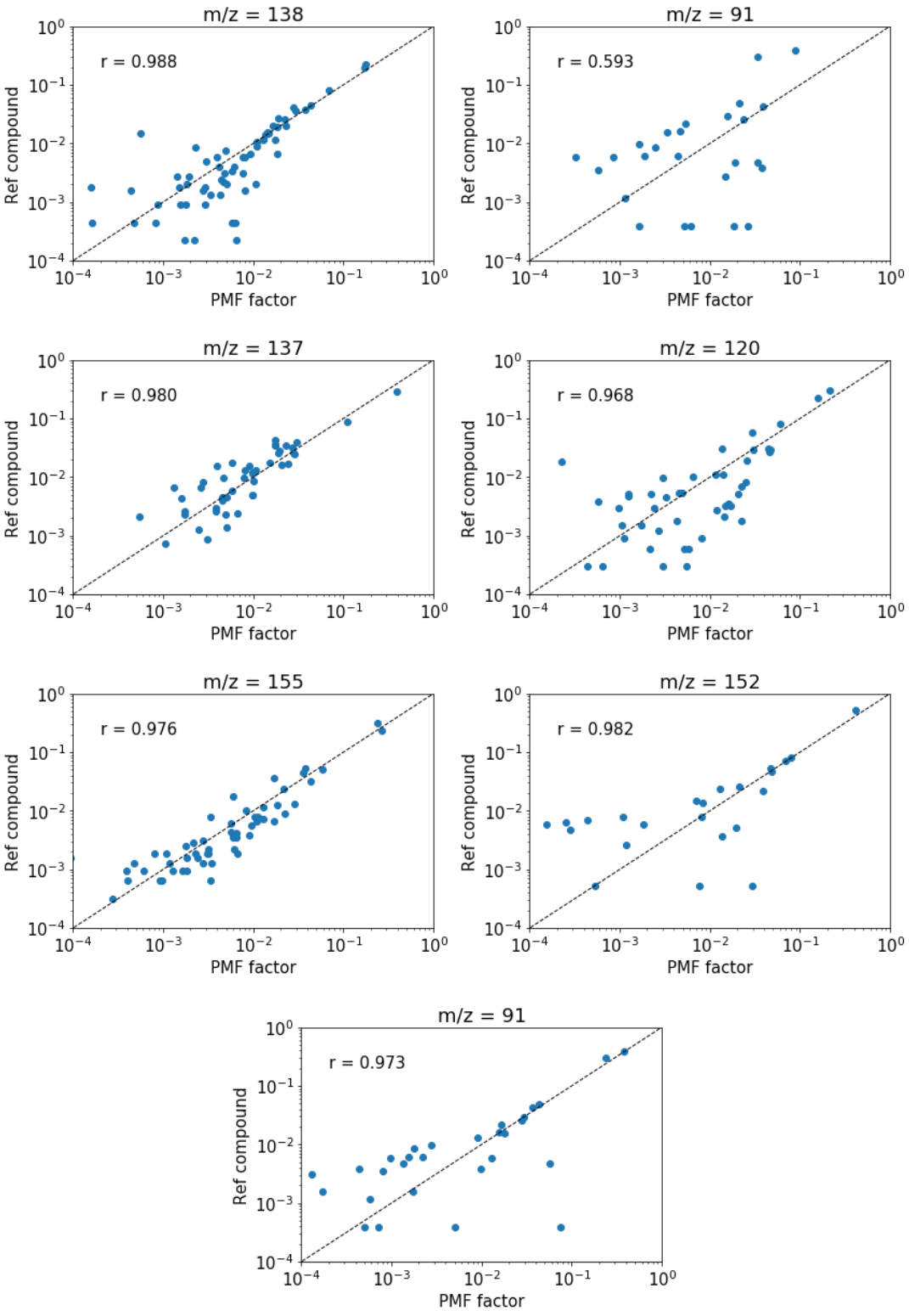



**Figure S10 Scatter plots of PMF resolved factors and reference compounds mass spectra in biomass burning dataset**

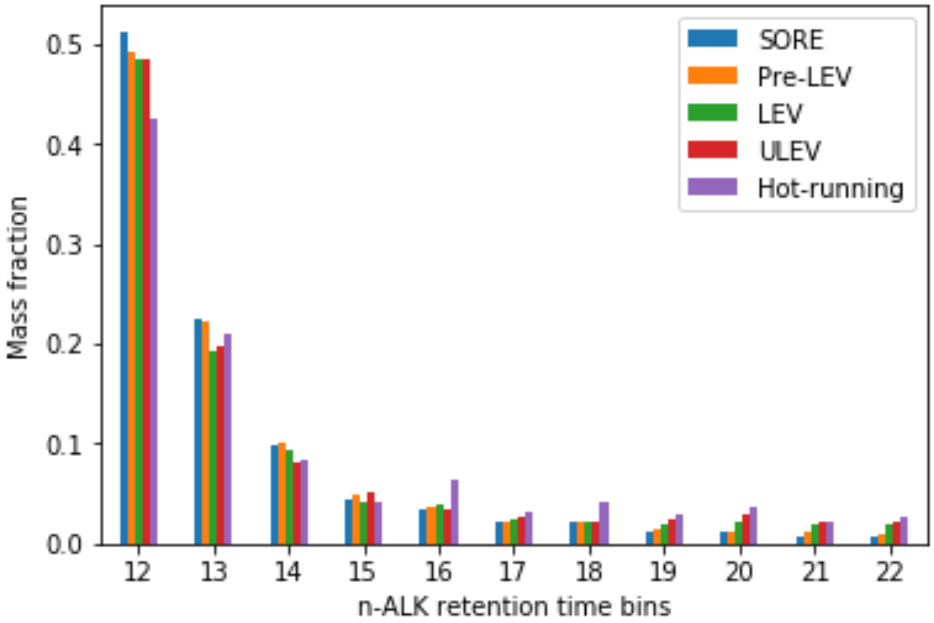



**Figure S11 Volatility distribution of IVOCs for all certification classes / test cycle of gasoline vehicles**

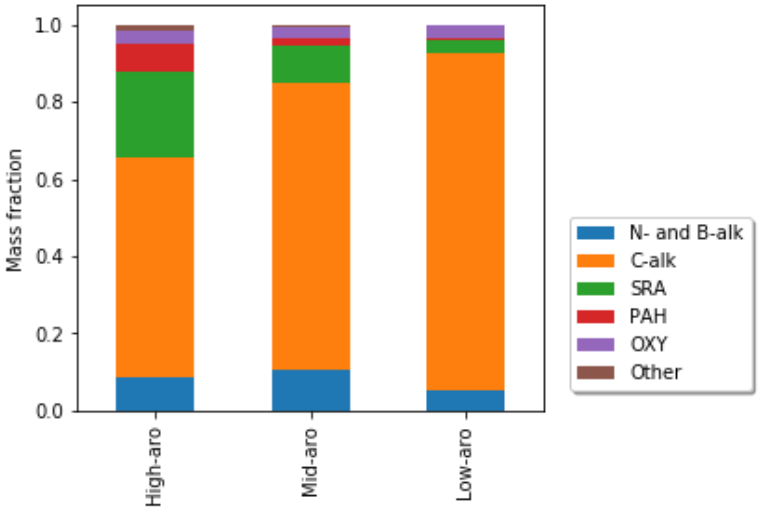



**Figure S12 PMF decomposed composition of non-DPF diesel vehicle emissions operated on different fuel**

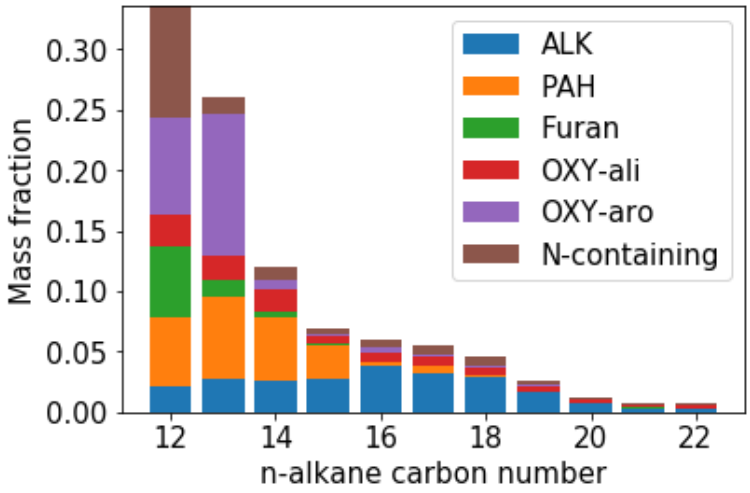



**Figure S13 Example of PMF decomposed chemical composition of biomass burning emissions**

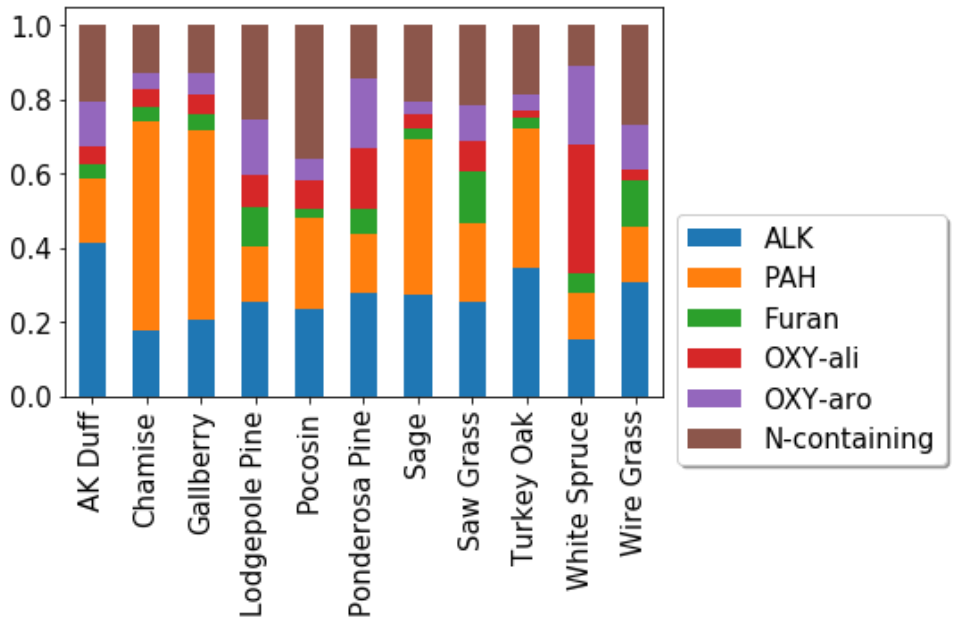



**Figure S14 PMF resolved chemical composition of all samples in biomass burning dataset**

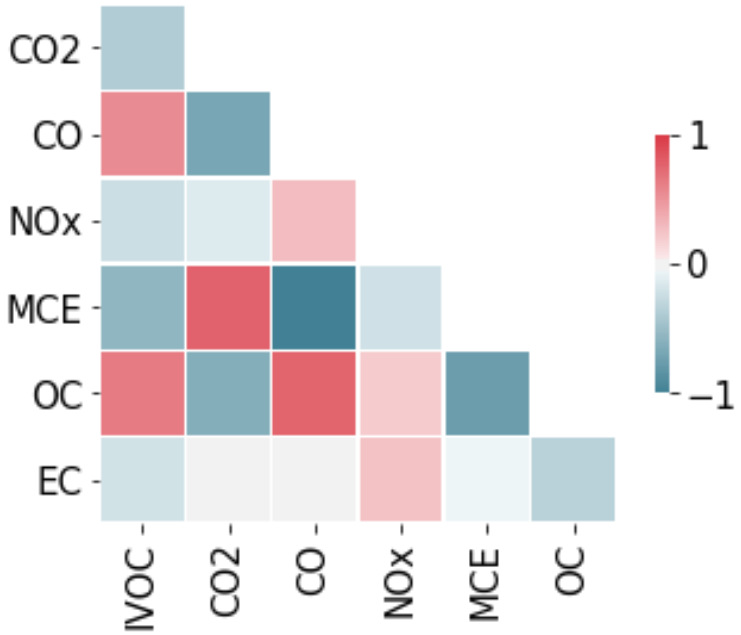



**Figure S15 Correlation coefficient (*r*) between quantified IVOC EFs and all other measured pollutants**

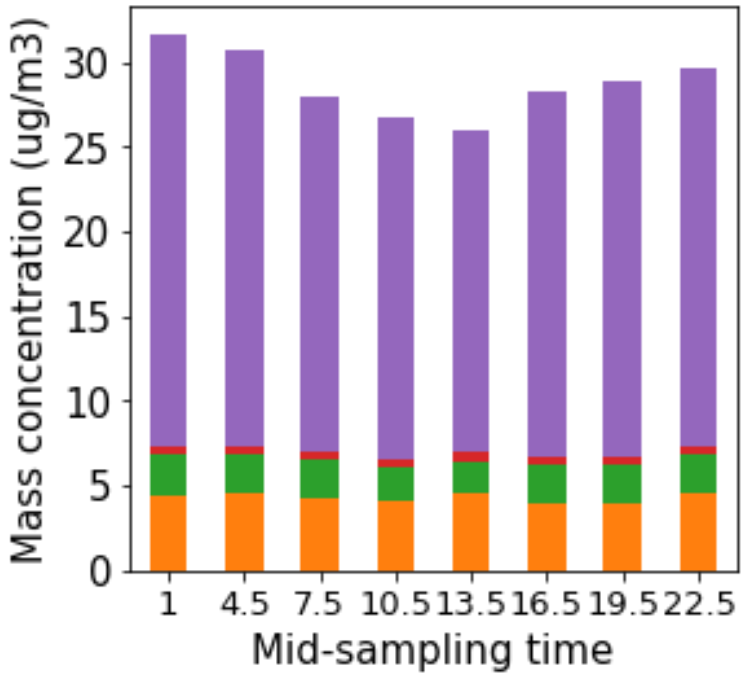



**Figure S16 Diurnal pattern of PMF decomposed total IVOC mass concentration**

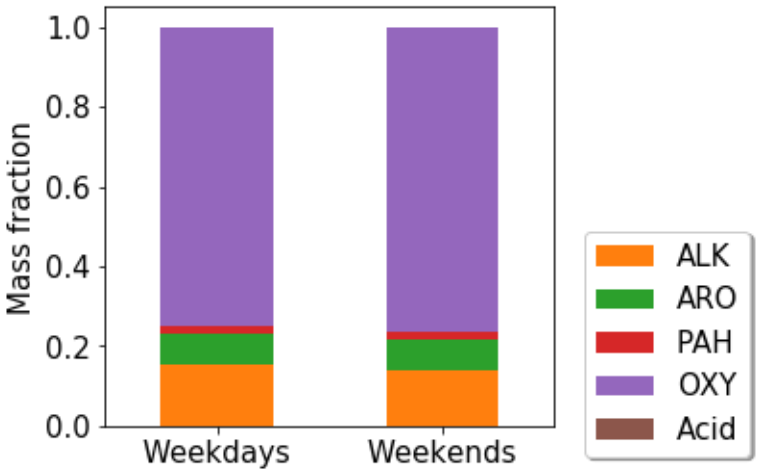



**Figure S17 Weekdays vs weekend ambient IVOCs chemical composition for all samples from CalNex**

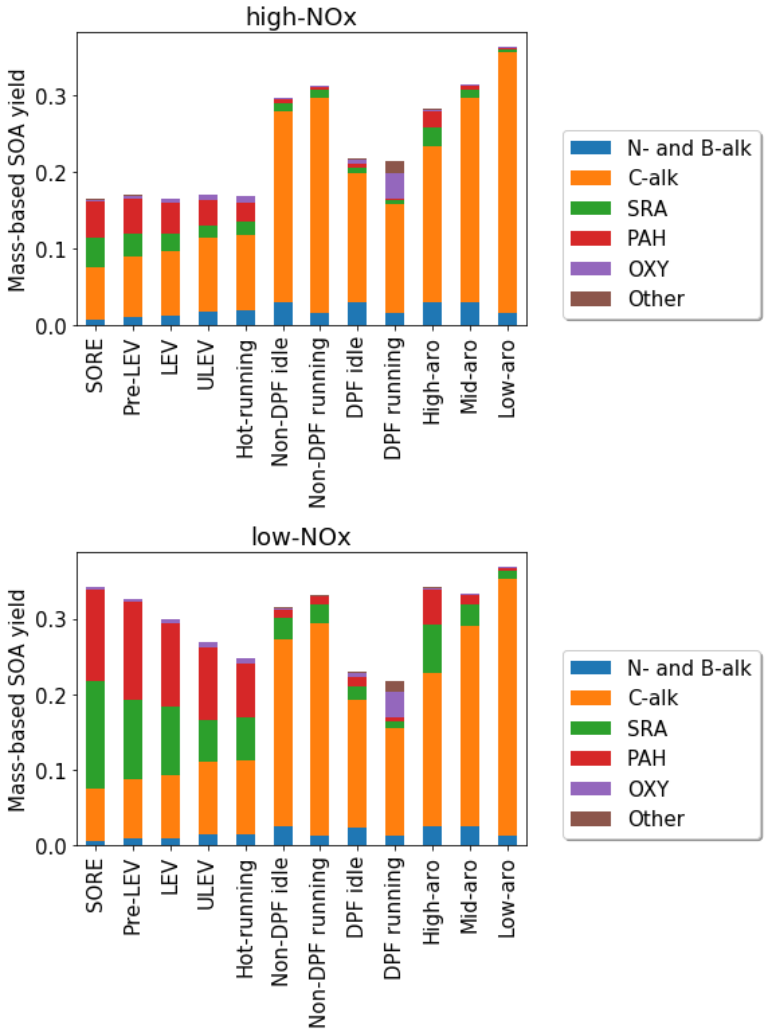



**Figure S18 Mass-based SOA yield for gasoline and diesel vehicles IVOC emissions under (a) high- and (b) low-$NO_x$ conditions**