Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions

Quanyang Lu 1, 2, Yunliang Zhao 1, 2, 3, Allen L. Robinson 1, 2

1Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States
2Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States
3Now at: California Air Resources Board, Sacramento, California 95814, United States

Correspondence to: Allen L. Robinson (alr@andrew.cmu.edu)

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 \(C^* = 10^7 - 10^{11} \ \mu g/m^3\)), intermediate-volatile organic compounds (IVOCs, \(C^* = 10^3 - 10^6 \ \mu g/m^3\)), semi-volatile organic compounds (SVOCs, \(C^* = 1 - 10^2 \ \mu g/m^3\)), low-volatile organic compounds (LVOCs, \(C^* \leq 0.1 \ \mu g/m^3\)) and non-volatile organic compounds (NVOCs). 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 10th to 90th percentile) and 1.1% (0.4-3.6%) for gasoline engine emissions, respectively. This suggests that a single profile can be used to represent the emissions from sources operating on the same 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). The profiles are designed to be directly implemented into chemical transport models and inventories.

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 (PM_{2.5}) (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., 2013a, 2013b, 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 IVOCs and SVOC vapor. Laboratory experiments indicate that IVOC and SVOC 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). Omitting 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).

IVOCs and SVOCs 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 LVOC 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 data 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 study, we develop comprehensive organic emission profiles for mobile sources by integrating recently published data of organic emissions based on their volatility, including IVOCs and SVOCs. 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 Methods

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 develop comprehensive model-ready source profiles. All tests used same procedures to characterize IVOC and SVOC emissions to create a self-consistent dataset for low-volatile organics; however, slightly different procedures were used to characterize VOC emissions. In the results and discussion sections, we compare these data to other recently published measurements 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 a FID) minus CH$_4$ plus carbonyls. We define POA as organics collected by a bare quartz filter. 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. For gasoline and diesel sources, 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). For the gas-turbine engine, exhaust was sampled from a rake inlet installed 1-m downstream of the engine exit plane (Presto et al., 2011).

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 organics or unresolved complex mixture (UCM) was 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). IVOC lumped groups (C$_{12}$ – C$_{22}$) were 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 that has 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 Mapping organics into volatility basis sets (VBS)

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. In addition to comparing emissions using SAPRC speciation, we also compare them 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 low-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.

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,

$$C_i^* = \frac{M_i 10^6 \xi_i p_{L,i}^0}{760RT}$$

where $M_i$ is the molecular weight (g/mol), $\xi_i$ is the activity coefficient of compound i in the condensed phase (assumed to be 1), and $p_{L,i}^0$ is the vapor pressure (Torr) of compound i as a liquid, $R$ is the ideal gas constant (8.206x10^{-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 lumped groups of unspeciated organics, C* values are estimated using the retention time of n-alkanes as the 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, C* value of n-alkane in each bin is used to represents the group of compounds centered around n-alkanes. 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 lump all species into the VBS volatility distribution. Each volatility bin of C* = 10^n g/m^3 cover the volatility range from C* = 0.3 x 10^n g/m^3 to C* = 3 x 10^n g/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 x 10^6 g/m^3), n-dodecane (C* = 1.9 x 10^6 g/m^3) and naphthalene (C* = 1.1 x 10^6 g/m^3). Figure S1 (a-c) compares the quantification results of these species (Supporting Information).

The comparisons indicate that the filter/adsorbent tube sampling train quantitatively collects all organics less volatile than n-dodecane (C* = 1.9 x 10^6 g/m^3) while the bag/canister quantitatively collects all more volatile organics. n-dodecane falls within the 10^6 g/m^3 falls volatility bin. The upper bound of this bin is 3 x 10^6 g/m^3 falls in, which is close to the C* of n-dodecane. We therefore use 3 x 10^6 g/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 x 10^6 g/m^3 and discarded all species measured in the adsorbent tube more volatile than 3 x 10^6 g/m^3. Therefore, emissions in the C*=
10^7 to 10^{11} \mu g/m^3 bins are based on the bag/canister/cartridge data and that the emissions in the C^*= 10^{-1} to 10^6 \mu g/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 \times 10^6 (n-dodecane) and 3 \times 10^6 \mu g/m^3, however, they still measured, on average, 3.3 times organics in this range to the Tedlar bags (Fig. S1D).

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 detailed analysis of the bags/canister, cartridges, adsorbent tubes and filters.

Figure S2 indicates good mass closure for the on-road gasoline and diesel vehicle tests. The two estimates results agree within ±10% for more than 90% of non-DPF-equipped diesel engine tests (DPF = diesel particulate filter). For all LGDV 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 (light-duty gasoline vehicle) 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 (Fig. S3). 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 to make the two datasets consistent (see S1). For DPF-equipped diesel vehicles, the sum of speciated organics is up to 7 times as bulk measurement.

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 (frequently called unknown or UNK): (1) assume it is inert and therefore ignored in models, (2) renormalizing the 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, Z;Vukovich, JCarter, 2005). Therefore, it still lacks comprehensive IVOCs and SVOCs data.

In the following discussion, we normalize the residual/uncharacterized organics to the known composition, essentially by 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.
2.3 Box model for SOA yield calculation

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

\[ \gamma_{SOA} = \sum_i f_{gas,i} \times Y_i \quad (2) \]

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/m³ (a typical urban OA level).

SOA mass yields for each VOC groups were taken from CMAQ 5.1 (USEPA, 2016a). This required lumping the VOC emissions into the SAPRC species; the SAPRC07 versions of the new source profiles are listed in the SI. 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).

3. Results and discussion

3.1 Emission factors

Figure 1 shows the volatility distribution of the total characterized organic emissions for a typical gasoline (Fig. 1A) and diesel (Fig. 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 organics. 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 (Zhao et al., 2015, 2016). The adsorbent collects 5.2% and 54.8% of the total organic emissions from the gasoline and diesel engines, respectively. These organics are not characterized in traditional source tests and therefore not accounted for in most emission profiles (Pye and Poulion, 2012). The bag/canister collected roughly 30% of C_{12} IVOCs, which corresponds to 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 1 indicates there is also substantial breakthrough of SVOCs (on average, 37% for gasoline, 52% for non-DFP diesel and 89% for DPF-diesel) from the quartz filter during certification testing (e.g. 2007 CFR 86), which require maintaining a filter temperature of 47°C. Therefore, inclusion of gas-phase SVOC measured by the adsorbent is needed to improve the predictions of POA concentrations and SOA production.

Figure 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, there are distinct trends in total organic emissions. Gasoline small off road engines (SORE) have the highest emissions, with SORE-2S about one order of magnitude higher than SORE-4S (Gordon et al., 2013). This is due to less stringent regulations for off-road engine emissions (Cao et al., 2016), 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 emissions to the median LEV and 90% to the median ULEV. 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 dependences 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, Fig. 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 trends in total organic emission among the different source categories are clear even with this variability.

3.2 Volatility and chemical composition distributions

Figure 3 shows the median volatility distributions of the emissions for three different source categories: gasoline, gas-turbine and non-DPF diesel. For gas-turbine engine category, we plot the idle load (4% thrust) emission. Figure 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 \( \mu g \ m^{-3} \) for gasoline source, C*=10^6 \( \mu g \ m^{-3} \) for gas-turbine source, and C*=10^5 \( \mu g \ m^{-3} \) for diesel source. This mode contributes 72.6% (66.5-77.6% as 10th to 90th 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. For each source category, this mode has a similar volatility distribution and chemical composition as unburned fuel (Fig. S4). We therefore call it the ‘fuel’ mode.

Figure 3 highlights how the changes in fuel composition create systematic differences in emissions among the three source categories. Specifically, the exhaust ‘fuel’ mode shifts towards lower volatility from gasoline to diesel sources mirroring the trend in fuel volatility. Most compounds show close to 1:1 relations between fuel and exhaust composition (Fig. S4). However, there are some difference between the exhaust composition in the fuel mode and that of unburned fuel indicating that combustion and removal efficiencies vary by compound class. This is discussed further in section 3.4.

Emissions from each source has a low-volatility mode, comprised of SVOCs and even less volatile organics. This 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. Large variation range of SVOCs from gas-turbine and diesel sources reflects changes in
engine load/after-treatment: gas-turbine source shows 12.1% in low-volatility mode at 85% thrust load, and DPF-equipped vehicles show 14.8% (12.3-20.2%) in this mode on UDDS and high-speed cycles. For all three source categories, the peak in this low-volatility mode is at a C* of 10 µg 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, forming POA, while the rest exist as vapor. The volatility distribution of this mode is similar to that of lubricating oil (May et al., 2013a, 2013b); we therefore refer to the low-volatility mode as the oil mode. For diesel, the low-volatility and fuel modes blend together. NVOC (defined as the difference between the organics measured by thermal-optical and GC-MS analysis) contribute less than 15% of all organics less volatile than 10⁻¹ µg m⁻³.

The third mode is the most volatile one, peaking at a C* of 10¹⁰ or 10¹¹ µg m⁻³. It contributes about one-quarter the total organics in gasoline and diesel sources emissions, and 30-40% for gas-turbine engine exhaust. It is comprised of the smallest compounds, such as C₂-C₃ alkanes, alkenes and carbonyls, produced from the incomplete combustion and breakdown of fuel molecules (May et al., 2014). We therefore call it the ‘combustion by-product’ mode. The composition of this mode varies modestly by source class, but it is primarily comprised of alkenes and carbonyls.

Figure 2(b) indicates that IVOCs contribute 4.5% (2.4%-9.6%) of the total organic emissions from gasoline sources. This includes both heavily controlled and low emitting ULEV and uncontrolled and high emitting SOREs. IVOCs contribute 20-27% of gas-turbine engine emissions at idle and cruise 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%). 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; Gentry et al., 2012; May et al., 2014).

Figure 2(c) indicates that the contribution of SVOCs is also reasonably consistent across sources using same fuel type. For gasoline engines, SVOCs contribute 1.1% (0.4%-3.6%) of the total organic emission. For gas-turbines, SVOCs contribute 4-6% of total organic emissions. For diesel source, SVOCs contribute 4.6% (2.3%-16.1%), reflecting changes in engine load and after-treatment. 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).

Although the total organic emissions vary by more than five orders of magnitude, the volatility distribution and chemical composition are relatively consistent across sources using the same fuel type. Figure S5(a-c) shows scatter plots of the median volatility distribution of emissions of on- versus off-road gasoline engines, on- versus off-road non-DPF diesel engine, and on-road DPF versus non-DPF diesel engines. The R² is 0.91, 0.64 and 0.70, respectively, indicating very similar distributions between on- and off-road gasoline sources and generally consistent volatility distributions for diesel sources. This suggests that a median profile provides a reasonable representation of the emissions for gasoline sources. Since diesel source emissions would still be dominated by non-DFP vehicles till the vast majority of on-road diesel vehicles are DPF
vehicles, we are using the median profile of non-DPF diesel vehicle emissions as representative diesel source profile (Zhao et al., 2015).

3.3 New versus traditional source profiles

Figure 3 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). The VOCs, IVOCs, and SVOCs in traditional profiles are from the EPA SPECIATE database (USEPA, 2016b): profile #4674 for diesel, #8750a for gasoline, and #5565 for gas turbine sources. The NVOC (non-volatile POA) fraction is calculated in MOVES (USEPA, n.d.) at national average for 2014.

There is good agreement between our new and traditional profiles in the VOC range, with both having by-product and fuel modes (Fig. 3) and similar chemical composition. Figure S6 demonstrates the strong agreement for SARPC-lumped VOC groups between the new and traditional profiles for all three sources. For example, more than 90% of all SAPRC groups for the gasoline sources agree to within a factor of two.

However, the traditional profiles lack IVOCs and SVOCs, which are important classes of SOA precursors. As is illustrated in Fig. 1, this reflects the limitations of traditional source characterization techniques to quantitatively collect and analyze IVOCs (Kishan et al., 2008). For example, the traditional LDGV emission profile only attributes 0.2% of the total organics to IVOCs versus 4.5% in our new 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 more than 50% of organics in our new profiles. The traditional diesel source profile also contains about 20% unknown organics, part of which are likely IVOCs, as the collection and chemical analysis efficiency decrease towards lower volatility bins such as $10^3$ and $10^4$ µg/m$^3$ (Fig. 3). However, most UNK is not represented as IVOCs in models, as discussed in section 2.2.

3.4 Exhaust versus unburned fuel and IVOC enrichment factors

Figure 3 highlights the large contribution of unburned fuel to the exhaust. However, the combustion process and removal efficiency are compound dependent. For example, gasoline and gas-turbine emission are both enriched in IVOCs compared to fuel (C* = $10^6$ µg/m$^3$ for gasoline, and C* = $10^4$ µg/m$^3$ for gas-turbine). The difference between jet fuel and emissions in the C* = $10^7$ µg/m$^3$ bin suggests higher combustion efficiency of more volatile fuel components.

Figure S7 compares the chemical composition of the exhaust to unburned fuel. Overall, straight and branched alkanes (speciated and unspeciated) contribute a smaller fraction in the exhaust than in the fuel. For example, the median fractions decrease 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. 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 due to incomplete combustion 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, including IVOCs and single ring aromatics. To quantify the 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 Fig. S4 and S8, some exhaust components are enriched, and others are depleted relative to fuel. Benzene and 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). The enrichment factors of IVOCs in diesel exhaust are less than 1, which means relative depletion of IVOCs from diesel source emission, compared to the tracer compounds (C_8 to C_{10} n-alkanes).

Figure 5 shows box-whisker plots of the overall IVOC enrichment factors. Sources using more volatile fuel have higher IVOC enrichment factors. For example, gasoline engine exhaust has a median IVOC enrichment factor of 8.5 versus relative depletion (enrichment factor <1) to C_{8:10} n-alkanes in diesel source exhaust. Gas turbine exhaust falls in between. There are several possible explanations for this trend. Recent research shows that IVOCs are less efficiently combusted and/or removed by catalytic converters compared to VOCs (Pereira et al., 2017). 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).

4 Implications for OA formation

Simulation of ambient OA concentrations requires accurate representation of both emissions and effective SOA yields for SVOCs and IVOCs. Given the lack of IVOC data in traditional source profiles (Fig. 3), previous modelling studies have used different scaling approaches, most commonly POA (Koo et al., 2014; Murphy et al., 2017; Robinson et al., 2007; Woody et al., 2016) but also NMOG (Jathar et al., 2014, 2017) and naphthalene (Pye and Seinfeld, 2010), as well as using unburnt fuel surrogate to estimate IVOC emissions (Gentner et al., 2012). These estimates are then combined with SOA yield data.

In this section, we use our new data to evaluate these different approaches for estimating IVOC emissions to better understand their strengths and weaknesses for simulating ambient OA concentrations. Table 1 outlines several different approaches that have been used to estimate IVOC emissions: (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; (4) MUR: traditional profiles + 9.656 × POA as IVOCs; (5) PYE:
traditional profiles + $66 \times$ Naphthalene as IVOCs; (6) GEN: using unburnt fuel composition as surrogate; (7) JAT: 20\% of NMOG of gasoline emission and 25\% of diesel emissions are IVOCs. All but approach (2) provide an estimate of IVOCs.

Figure 5 compares our new data to six different previous estimates. Figure 5(a) shows the relative contribution of different classes of SOA precursors (VOC, IVOC and SVOC) to the NMOG emissions. Figure 5(b) shows the overall effective SOA yields of the total NMOG emissions for the different models.

As shown in Fig. 5(a), all estimates have similar VOC SOA precursor fractions, but widely divergent amounts of IVOCs. Our new profiles (1) and estimates (6) and (7) have lower VOC SOA precursors, due to the inclusion of IVOCs and gas-phase SVOC within NMOG emissions, while approaches (3) – (5) add additional IVOCs on top of the existing NMOG emissions. Given that NMOG is defined as the sum of non-methane organic gases, we think IVOCs and gas-phase SVOCs should be accounted within NMOG emissions. Traditional profiles do not account for IVOCs in all sources.

Approaches (3) ROB and (4) MUR: 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 these ratios vary by source category: the average IVOC-to-POA ratios for gasoline engines exhaust is $6.2 \pm 4.4$ (on-road) and $5.8 \pm 2.1$ (off-road), and for diesel engines exhaust is $12 \pm 7$ (non-DPF equipped) and 31 (DPF-equipped). Therefore, the widely used scaling factor of IVOC-to-POA of 1.5 (ROB in Fig. 5) (Robinson et al., 2007) 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 Fig. 5) overestimates IVOC emissions from gasoline and gas-turbine sources.

However, even if one uses source specific IVOC-to-POA scaling factors, we do not think that scaling POA provides a robust estimate 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 (Fig. 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) reported much stronger correlations between IVOC and total NMOG emissions than with POA over a range of operating conditions. 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 by scaling gas-phase organics not POA.

Approach (5) PYE: Pye and Seinfeld (2010) estimated IVOC emissions by scaling naphthalene using the same ratio for all sources. Our data indicate that naphthalene is not a good indicator of IVOCs, due to the large variation in fuel aromatics content. For example, naphthalene is 4 times as much in gasoline engine exhaust (0.4\%) and fuel (0.13\%) as compare to diesel engine exhaust (0.1\%) and fuel (0.04\%), resulting in a much higher estimates of IVOC emissions from gasoline than diesel sources, which is opposite of the actual emissions data (Fig. 2). This problem can be overcome with source specific
ratios, but, even if one uses 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.

Approach (6) JAT: 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. The ratios of Jathar et al. (2014) were derived to be used in combination with their empirically derived effective SOA yields. When used together they explain SOA yield production measured in smog chamber experiments with dilute exhaust. One cannot simply replace IVOC-to-NMOG of Jathar et al. (2014) with the ones reported here without also using different effective SOA yields.

Approach (7) GEN: 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 effective 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 (Fig. 4). For example, we predict 40% higher effective SOA yield for gasoline exhaust when compared to estimates based on unburned fuel.

In Fig. 5(b), we combine the emissions and SOA yield data for all different model approaches to compare predictions of overall effective SOA yields of NMOG emissions for all sources, assuming complete oxidation of all precursors. Our new profiles predict that IVOCs and SVOC vapor contribute substantially to SOA production, especially for sources using heavier fuels (e.g. diesel) where IVOCs and SVOC vapor dominate SOA production. 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 vapor combining contribute factors of 13 and 44 more SOA than VOCs, respectively. Figure 5(b) also compares the effective yields of NMOG emissions in all approaches (2) – (7). SOA yields of estimated IVOCs in all approaches are from the original literature (Table 1). The differences in effective yields are primarily due to differences in IVOC/SVOC emissions. Traditional profiles and ROB underpredict SOA production from all three source categories, due to the underestimation of IVOC emissions. As is discussed in Section 3.4, IVOCs are enriched in gasoline emissions compared to unburned fuel, therefore GEN underpredicts the effective SOA yield of gasoline emissions. In contrast, 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 in IVOC emissions by a factor of 4, and both underestimate the overall SOA yield for diesel source NMOG emissions.

To conclude, none of the previous modelling approaches provide a robust estimate of the IVOC fraction in the exhaust for all three sources. Figure 5(a) and (b) show that traditional profiles either completely omit IVOCs or incorrectly lumped them to VOC chemical mechanism groups, which greatly underestimate their overall SOA 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 \mu g/m) (May et al., 2013a, 2013b).

5 Recommendations and future research needs

Figure 5 highlights the importance of including IVOC and SVOC emissions in models and inventories. This paper facilities this by providing model-ready profiles that include direct measurements of IVOCs and SVOCs. These profiles, which are the median of the measurements reported here, are provided in the Table S3(a-c) (Supporting Information).

For other liquid-fuel internal engine sources, we recommend interpolating based on fuel composition and applying the IVOC enrichment factor estimated from fuel volatility (Fig. 4). For sources profiles that only contain speciated VOCs and unknown components, we recommend not normalizing to known species, which will likely misattribute low volatility organics to VOCs, but scaling to IVOCs or SVOCs based on emission data and fuel composition.

Future research needs:
1) IVOC and SVOC emissions data for other 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 and ready-to-use profiles are needed for these and other source categories (McDonald et al., 2018).

2) 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;

3) Detailed chemical composition of IVOCs and SVOCs. Although we have performed quantify 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, we have derived some composition information. However, 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);

4) 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.
Author contribution

Q.L., Y.Z. and A.L.R. designed the research. Q.L. and Y.Z. analyzed the data. Q.L., Y.Z. and A.L.R. wrote the paper.

Competing interests

The authors declare no conflict of interest.

5 Acknowledgments

This publication was developed 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 Dr. Albert Presto, Dr. Tim Gordon and Dr. Andrew May for their help in providing detailed datasets information. The authors would also like to thank Dr. Benjamin Murphy for helpful comments on this manuscript.
References


Deepwater Horizon Oil Spill, Science (80- ), 331(6022), 1295 LP-1299 [online] Available from: http://science.sciencemag.org/content/331/6022/1295.abstract, 2011.


Pye, H. O. T. and Pouliot, G. A.: Modeling the Role of Alkanes, Polycyclic Aromatic Hydrocarbons, and Their Oligomers in


### Tables

**Table 1 Comparison of different estimates of IVOC emissions and SOA yield for mobile sources**

<table>
<thead>
<tr>
<th>Estimated from</th>
<th>Reference</th>
<th>SOA yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>New</td>
<td>Direct measurements</td>
<td>This work</td>
</tr>
<tr>
<td>Trad</td>
<td>/</td>
<td>SPECIATE</td>
</tr>
<tr>
<td>ROB</td>
<td>(1.5*) POA</td>
<td>Robinson et al. (2007)</td>
</tr>
<tr>
<td>MUR</td>
<td>(9.656*) POA</td>
<td>Murphy et al. (2017)</td>
</tr>
<tr>
<td>GEN</td>
<td>Unburnt fuel</td>
<td>Gentner et al. (2012)</td>
</tr>
<tr>
<td>JAT</td>
<td>Inverting chamber experiment results</td>
<td>Jathar et al. (2014)</td>
</tr>
</tbody>
</table>
Fig. 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).
Fig. 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 all categories. Box-whisker plot represents range of emission for each category: 25th-75th percentiles and 10th-90th percentiles.)
Fig. 3 Volatility distribution of organic emissions (median, with chemical composition) for (a) on-road gasoline, (b) gas-turbine (idle) and (c) on-road non-DPF diesel engines. Bars indicate median data; dashed lines indicate distribution for unburned fuel; dots are SPECIATE profiles. The y-axis has a broken scale to amplify the least volatile emissions.
Fig. 4 IVOC mass enrichment factors as a function of IVOC content in fuel, $R_{\text{Enrichment},i} = \frac{m_i^{\text{exhaust}}/m_{C_{8-10}}^{\text{exhaust}}}{m_i^{\text{fuel}}/m_{C_{8-10}}^{\text{fuel}}}$.

The box-whisker plots indicate variability in ratio within a given source class: 25th-75th percentiles and 10th-90th percentiles.)
Fig. 5 Comparison of (a) SOA precursors as a fraction of total NMOG emissions (b) effective SOA yields as NMOG emissions from mobile sources via different estimation approaches. Asterisk (*) denotes no estimation available.