Towards a satellite – in situ hybrid estimate for organic aerosol abundance

Jin Liao¹,², Thomas F. Hanisco¹, Glenn M. Wolfe¹,³, Jason St. Clair¹,³, Jose L. Jimenez⁴,⁵, Pedro Campuzano-Jost⁴,⁵, Benjamin A. Nault⁴,⁵, Alan Fried⁶, Eloise A. Marais⁷, Gonzalo Gonzalez Abad⁸, Kelly Chance⁸, Hiren T. Jethva¹,², Thomas B. Ryerson⁹, Carsten Warneke⁵,⁶, Armin Wisthaler¹⁰,¹¹

¹Atmospheric Chemistry and Dynamic Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
²Universities Space Research Association, GESTAR, Columbia, MD, USA
³University of Maryland Baltimore County, Joint Center for Earth Systems Technology, Baltimore, MD, USA
⁴Department of Chemistry, University of Colorado, Boulder, Colorado, USA
⁵Cooperative Institute for Research in the Environmental Sciences, University of Colorado, Boulder, Colorado, USA
⁶Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado, USA
⁷School of Geography, Earth and Environmental Sciences, University of Birmingham, UK
⁸Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
⁹NOAA Earth System Research Laboratory (ESRL), Chemical Sciences Division, Boulder, CO, USA
¹⁰Department of Chemistry, University of Oslo, Oslo, Norway
¹¹Institute for Ion Physics and Applied Physics, University of Innsbruck, Innsbruck, Austria

Correspondence email: jin.liao@nasa.gov
Abstract

Organic aerosol (OA) is one of the main components of the global particulate burden and intimately links natural and anthropogenic emissions with air quality and climate. It is challenging to accurately represent OA in global models. Direct quantification of global OA abundance is not possible with current remote sensing technology; however, it may be possible to exploit correlations of OA with remotely observable quantities to infer OA spatiotemporal variability. In particular, formaldehyde (HCHO) and OA share common sources via both primary emissions and secondary production from oxidation of volatile organic compounds (VOCs). We examine OA-HCHO correlations using data from summer-time airborne campaigns investigating biogenic (NASA SEAC^4RS and DC3), biomass burning (NASA SEAC^4RS) and anthropogenic conditions (NOAA CalNex and NASA KORUS-AQ). In situ OA correlates well with HCHO ($r = 0.59 - 0.97$) but the slope and intercept of this relationship vary with chemical regime. For biogenic and anthropogenic regions, the OA-vs-HCHO slope is higher in low NOx conditions, where HCHO yields are lower and aerosol yields are likely higher. The OA-vs-HCHO slope of wild fires is more than 9 times higher than that associated with biogenic and anthropogenic sources. An estimate of near-surface OA is derived by combining observed in situ relationships with HCHO column retrievals from NASA’s Ozone Monitoring Instrument (OMI). We evaluate this OA estimate against OA observations from the US EPA IMPROVE network and simulated OA from the GEOS-Chem global chemical transport model. The OA estimate compares well with IMPROVE data obtained over summer months (e.g. slope = 0.62, $r = 0.56$ for August 2013), comparable to intensively validated GEOS-Chem performance (e.g. slope = 0.57, $r = 0.56$) and superior
to the correlation with satellite-derived total aerosol extinction ($r = 0.41$). Improving the detection limit of satellite HCHO and expanding in situ airborne HCHO and OA coverage in future missions will improve the quality and spatiotemporal coverage of this OA estimate, potentially enabling constraints on the global OA distribution.
1. Introduction

Aerosols are the largest source of uncertainty in climate radiative forcing (IPCC 2013; Carslaw et al., 2013) and also decrease atmospheric visibility and impact human health (Pope 2002). Organic aerosols (OA) comprise a large portion (~50%) of submicron aerosols (Jimenez et al., 2009; Murphy et al., 2006; Shrivastava et al., 2017), and this fraction will grow with continued decline in SO$_2$ emissions (Attwood et al., 2014; Marais et al., 2017; Ridley et al., 2018). In addition, OA affect cloud formation and climate radiative forcing such as the inclusion of phase-separation effects from organic material into cloud droplet activation thermodynamics in climate models can lead to important changes in cloud condensation nuclei (CCN) concentrations (Ovadnevaite et al., 2017).

OA components also have adverse health effects (e.g. Walgraeve et al., 2010) and contribute significantly to regional severe haze events (e.g. Hayes et al., 2013). Finally, because the response of temperature to changes in climate forcing is non-linear (Taylor and Penner, 1994) and the forcing by aerosols has strong regional character (Kiehl and Briegleb, 1993), it is necessary to separate out different climate forcing components to accurately forecast the climate response to changes in forcing.

It is challenging to accurately represent OA in global models. Chemical transport models (CTMs) often under-predict OA (e.g., more than 2 times less near the ground) compared to observations, and model-to-model variability can exceed a factor of 100 in the free troposphere (Tsигаридис et al., 2014; Heald et al., 2008; Heald et al., 2011). Fully explicit mechanisms (e.g. Lee-Taylor et al., 2015) attempt to capture the full OA chemical formation mechanisms, but many reaction rate constants and yields lack laboratory
validation. In addition, it is too computationally expensive to apply these mechanisms to OA formation in global CTMs at a useful resolution. For computational efficiency, 3-D models such as GEOS-Chem include direct emissions of primary OA (POA) and either represents secondary OA (SOA) formation by lumping SOA products according to similar hydrocarbon classes (Kim et al., 2015) or based on the volatility of the oxidation products (Pye et al., 2010). Marais et al. (2016) applied an aqueous phase mechanism for SOA formation from isoprene in GEOS-Chem to reasonably simulate isoprene SOA in the southeast US, but this is only one location and it is likely that isoprene SOA is not the dominant source of SOA in summer (Hu et al., 2015; Zhang et al., 2018). Accurate emission inventories are also needed to correctly represent volatile organic compounds (VOCs) and NOx (NOx = NO + NO2) inputs, and these often have biases compared to observational constraints (Kaiser et al., 2018, Travis et al., 2016, Anderson et al., 2014; McDonald et al., 2018).

A quantitative measure of OA from space would be very helpful for verifying emissions and aerosol processes in models. However, direct measurements of OA from space are currently unavailable. Aerosol optical depth (AOD) measured by satellite sensors provides a coarse but global picture of total aerosol distributions. MISR provides aerosol property information such as size, shape and absorbing properties, which allows retrieving AOD of a subset of aerosols when AOD is above a certain threshold (Kahn and Gaitley, 2015). Classification algorithms have been developed to speciate different aerosol types such as OA based on AOD, extinction Angstrom exponent, UV Aerosol Index, and trace gas columns from satellite instruments (de Vries et al., 2015).
Formaldehyde (HCHO) is one of the few VOCs that can be directly observed from space. Sources emitting POA (e.g., biomass burning (BB)) often simultaneously release VOCs. HCHO and SOA are also both produced from emitted VOCs. VOCs, as well as semi- and intermediate-volatility organic compounds (S/IVOCs), are oxidized by hydroxyl radicals (OH) to form peroxy radicals (RO$_2$), which then react with NO, RO$_2$, hydroperoxy radicals (HO$_2$) or isomerize. These oxidation processes produce HCHO and oxidized organic compounds with low volatility that condense to form SOA (Robinson et al., 2007; Ziemann and Atkinson, 2012). The yield of HCHO and SOA from hydrocarbon oxidation varies, depending on the VOC precursors, oxidants (OH, O$_3$ and NO$_3$), RO$_2$ reaction pathway (e.g. NO levels), and pre-existing aerosol abundance and properties (Wolfe et al., 2016; Pye et al., 2010; Marais et al., 2016 and 2017; Xu et al., 2016).

Although the lifetime of HCHO (1-3 hrs) is shorter than OA (1 week), HCHO continues to form from slower reacting VOCs, as well as from the oxidation of later generation products. Observations across megacities around the world show that OA formation in polluted/urban area happens over about 1 day (e.g. DeCarlo et al., 2010; Hodzic and Jimenez, 2010; Hayes et al., 2013; 2015), and HCHO is also significantly formed over this timescale (Nault et al., in preparation). Veefkind et al. (2011) found that satellite AOD correlated with HCHO over the summertime Southeast US, BB regions, and Southeast Asian industrialized regions. This also suggests that OA, a major contributor to AOD in the above cases (Jimenez et al., 2009), and HCHO share common emission sources and photochemical processes. Marais et al. (2016) further used the relationship
between aircraft OA and satellite HCHO to evaluate GEOS-Chem representation of SOA mass yields from biogenic isoprene in the Southeast US.

We present an OA surface mass concentration estimate (OA estimate) derived from a combination of satellite HCHO column observations and in situ OA and HCHO relationships. The detection limit of satellite HCHO column observations limit the quality of OA estimate, thus we are focusing on summer time when HCHO levels are high. The OA estimate is evaluated against OA measurements at ground sites. A 3-D model GEOS-Chem OA simulation is also shown for comparison.

2. Methods

2.1 In situ airborne observations

Figure 1 shows flight tracks of field campaigns used in the current study. The Studies of Emissions, Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) mission (Toon et al., 2016) covered the continental US with a focus on the southeast US in August-September 2013. The Deep Convective Clouds & Chemistry Experiment (DC3) (Barth et al., 2015) surveyed the central and southeast US in May-June 2012 while targeting isolated deep convective thunderstorms and mesoscale convective systems. The California Research at the Nexus of Air Quality and Climate Change (CalNex) (Ryerson et al., 2013) investigated the California region in May-June 2010, targeting the Los Angeles (LA) Basin and Central Valley. The Korea-United States Air Quality Study (KORUS-AQ) studied South Korean air quality, including sampling many large urban areas in South Korea and continental Asian outflow over the West Sea,
in May-June 2016 (https://www-air.larc.nasa.gov/missions/korus-aq/). KORUS-AQ only includes data with longitude < 133° E to exclude the transit from US because it targeted South Korea and the nearby region. These field campaigns were selected as they had recent high-quality in situ HCHO and OA data measured with state-of-the-art instruments and studied summertime regional tropospheric chemical composition.

In situ airborne HCHO observations were acquired by multiple instruments. The DC3 NASA DC-8 payloads featured two HCHO measurements: the NASA In Situ Airborne Formaldehyde (ISAF) (Cazorla et al., 2015) and the Difference Frequency Generation Absorption Spectrometer (DFGAS) (Weibring et al., 2006). The SEAC4RS NASA DC-8 payloads also featured two HCHO measurements: the NASA ISAF and the Compact Atmospheric Multispecies Spectrometer (CAMS) (Richter et al., 2015). HCHO measurements from ISAF were found to be in good agreement with CAMS, with a correlation coefficient of 0.99 and a slope of 1.10 (Zhu et al., 2016). Because ISAF has higher data density, we used ISAF HCHO data for DC3 and SEAC4RS. During KORUS-AQ, CAMS was the only HCHO instrument onboard the DC-8. In CalNex a proton transfer reaction mass spectrometer (PTRMS) (Warneke et al., 2011) was used to measure HCHO on board the NOAA P3 aircraft.

In situ airborne OA from SEAC4RS, DC3, and KORUS-AQ was measured by the University of Colorado High-Resolution Time-of-Flight Aerosol Mass Spectrometer (AMS, DeCarlo et al., 2006; Dunlea et al., 2009; Canagaratna et al., 2007; Jimenez et al., 2016) and in situ airborne OA from CalNex was measured by the NOAA Compact Time-
of-Flight Aerosol Mass Spectrometer (Drewnick et al., 2005; Canagaratna et al., 2007; Bahreini et al., 2012). The OA measurements are from 1 min merge data and converted from $\mu$g sm$^{-3}$ to $\mu$g m$^{-3}$ under local T & P for each data point.

Although NO modulates the RO$_2$ lifetime, and thus, the production of HCHO and SOA, it cannot be directly observed via remote sensing. Instead, NO$_2$ can be directly observed in space by satellites, and since NO$_2$ represents typically ~80% (e.g. SEAC$^4$RS and KORUS-AQ) of the boundary layer NO$_x$ concentrations during the daytime, it can be used as a surrogate for NO concentrations and oxidative conditions around the globe. In situ airborne NO$_2$ was measured by the NOAA Chemiluminescence NO$_2$O$_3$ instrument (Ryerson et al., 2001) during SEAC$^4$RS and by University of Berkeley laser induced fluorescence NO$_2$ instrument (Day et al., 2002) during KORUS-AQ.

2.2 Ground-based OA measurements

Ground-based OA measurements over the US are from the EPA Interagency Monitoring of Protected Visual Environments (IMPROVE) (Malm et al., 1994; Solomon et al., 2014; Hand et al., 2014; Hand et al., 2013; Malm et al., 2017) and Southeastern Aerosol Research and Characterization (SEARCH) (Edgerton et al., 2006) networks. In the IMPROVE network, aerosols were collected on quartz fiber filters and analyzed in the lab by thermal optical reflectance for organic and elemental carbon. The data were reported every three days from 1988 to 2014. Monthly averages were used for comparison in this study. IMPROVE OA data over the southeast US (east of 70ºW) in summertime were multiplied by a factor of 1.37 to correct for partial evaporation during
filter transport, following the recommendation of a comparison study with SEARCH organic carbon (OC) measurements (Kim et al., 2015; Hand et al., 2013). Although IMPROVE OA corrected for evaporation has potential uncertainties with the constant scaling factor, the IMPROVE measurements have high spatial coverage. SEARCH network (Edgerton et al., 2006; Hidy et al., 2014) OC is determined by the difference between total carbon (TC) detected by a tapered element oscillating microbalance (TEOM) and black carbon (BC) measured by an in situ Thermal Optical instrument. This allows real-time measurement of OC and prevents evaporation during filter transport. Although the SEARCH network only has 5 sites available, we also use observations from this network due to their high accuracy. The IMPROVE and SEARCH network OC measurements were converted to OA by multiplying by a factor of 2.1 based on ground and aircraft observations (Pye et al., 2017; Schroder et al., 2018).

2.3 Satellite measurements

Satellite HCHO column observations are derived from the NASA’s Ozone Monitoring Instrument (OMI), a UV/Vis nadir solar backscatter spectrometer on the Aura satellite (Levelt et al., 2006). Aura overpasses the equator at 1:30 pm local time, daily. Here we use the OMI HCHO version 2.0 (collection 3) retrieval (Gonzalez Abad et al., 2015) from the Smithsonian Astrophysical Observatory (SAO). Satellite data for HCHO columns were subjected to data quality filters: 1) solar zenith angle lower than 70°, 2) cloud fraction less than 40%, and 3) main quality flag and the xtrackquality flag both equal to zero (https://www.cfa.harvard.edu/atmosphere/Instruments/OMI/PGEReleases/READMEs/O
The monthly average HCHO columns were also weighted by the column uncertainties of the pixels. The HCHO retrieval used a priori profiles without aerosol information from the GEOS-Chem model (Gonzalez Abad et al., 2015). Satellite NO₂ column observations are also derived from NASA’s OMI level 3 data, archived at https://disc.sci.gsfc.nasa.gov as “OMI-Aura_L3-OMNO2d” (Lamsal et al., 2014).

Satellite AOD observations were acquired from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Aqua satellite, using overpasses at about 1:30 pm local time. Here, we use collection 06 (MYD04_L2, https://ladsweb.nascom.nasa.gov/), retrieved using the Dark Target (DT) and Deep Blue (DB) algorithms (Levy et al., 2015).

2.4 GEOS-Chem

We used GEOS-Chem (v9-02) at 2° × 2.5° with 47 vertical layers to simulate HCHO and OA globally, the same as that in Marais et al. (2016). GEOS-Chem is driven with meteorological fields from the NASA Global Modeling and Assimilation Office (GMAO). The OA simulation includes POA from fires and anthropogenic activity and SOA from the volatility-based reversible partitioning scheme (VBS) of Pye et al. (2010) for anthropogenic, fire, and monoterpene sources, and an irreversible aqueous-phase reactive uptake mechanism for isoprene. The aqueous-phase mechanism is coupled to gas-phase isoprene chemistry and has been extensively validated using surface and aircraft observations of isoprene SOA components in the Southeast US (Marais et al.,
The model is driven with Goddard Earth Observing System – Forward Processing (GEOS-FP) meteorology for 2013 and sampled along the SEAC4RS (2013) and KORUS-AQ (2016) flight tracks. The model is also run with 10% decrease in biomass burning, biogenic, or anthropogenic emissions as sensitivity test to evaluate the contributions of different sources to OA and HCHO budget. Model monthly mean surface layer OA and total column formaldehyde are obtained around the OMI overpass time (12:00-15:00 local time) for 2008-2013 using MERRA Modern-Era Retrospective analysis for Research and Applications (MERRA) meteorology, as GEOS-FP is only available from 2012. This is compared to the OA estimate derived from satellite HCHO.

Global isoprene emissions are used to calculate an isoprene and NO2 dependent OA estimate. Global isoprene emissions are from the Model of Emissions of Gases and Aerosols from Nature version 2.1 (Guenther et al., 2006) as implemented in GEOS-Chem and driven with MERRA (MEGAN-MERRA).

2.5 Estimation of surface organic aerosol mass concentrations

An estimate for surface OA mass concentration is calculated based on a simple linear transformation.

\[ \varepsilon(i) = \Omega_{HCHO}(i)\eta(i)\alpha(i) + \beta(i) \]  

Eq. (1)

Here, \( \varepsilon(i) \) is the OA estimate for grid cell i (µg m\(^{-3} \)), \( \Omega_{HCHO}(i) \) is the OMI HCHO column density (molec cm\(^{-2} \)) in each 0.25°×0.25° grid cell (similar resolution to OMI HCHO nadir pixel data), \( \eta(i) \) is the ratio of midday surface layer (~140 m) HCHO concentrations (molec cm\(^{3} \)) to column concentrations (molec cm\(^{-2} \)) from GEOS-Chem, and \( \alpha(i) \) and
\( \beta(i) \) are the slope and intercept of a linear regression between OA and HCHO from low altitude (<1 km) airborne in situ measurements. The in situ to column conversion factor \( \eta(i) \) is similar to that used by Zhu et al. (2017) to convert HCHO columns into surface concentrations. \( \eta(i) \) is derived from GEOS-Chem (v9-02), which includes updated isoprene scheme for OA and is the next version of the model (v9-01-03) for a priori profiles used in SAO satellite HCHO retrievals. The detailed information about \( \eta \) is provided in Sect. 5.

### 2.6 Aerosol extinction from satellite measurements

Currently remote sensing observes aerosols by quantifying AOD. The MISR satellite instrument can estimate the fraction of OA relative to total AOD, due to constraints on size range, shape and absorbing properties, but cannot distinguish OA from other submicron aerosol compounds such as sulfate and nitrate and also requires AOD above 0.1. Moreover, OA account for a large and relatively constant fraction of submicron aerosols in the Southeast US (Kim et al., 2015; Wagner et al., 2015). Therefore, AOD was converted to extinction to represent OA for comparison.

\[
A_{\text{ext}} = \text{AOD}(i) \delta(i)
\]  

Eq. (2)

where AOD(i) is aerosol optical depth from MODIS (see Sect. 2.3) in each 0.25°×0.25° grid cell, and \( \delta(i) (\text{m}^{-1}) \) is the ratio of surface layer OA concentrations (µg m\(^{-3}\)) to column OA concentrations (µg m\(^{-2}\)) from GEOS-Chem multiplied by \( 10^6 \text{Mm}^{-1}/\text{m}^{-1} \). The shape of average vertical profile of OA was very close to that of total aerosol mass (Wagner et al., 2015).
3. In situ OA and HCHO relationship

Although OA and HCHO share common VOC emission sources and photochemical processes, their production rates from different emission sources and photochemical conditions vary, as do their loss rates. The following section discusses the main factors that modulate OA-HCHO relationships.

3.1 Regional and Source-Driven Variability

For all regions and/or sources investigated, in situ OA and HCHO are well correlated. A scatter plot of in situ OA vs. HCHO at low altitudes (<1 km) from a number of field campaigns (SEAC4RS, DC3, CalNex, and KORUS-AQ) is displayed in Fig. 2. The slope, intercept and correlation coefficient are provided in Table 1. SEAC4RS, DC3, and CalNex excluded BB data when acetonitrile > 200 pptv (Hudson et al., 2004). KORUS-AQ used a BB filter with higher acetonitrile (>500 pptv) because the air masses with moderate acetonitrile enhancement (200-500 pptv) were actually from anthropogenic emissions. This attribution is based on high levels of acetonitrile detected downwind of Seoul and west coastal petrochemical facilities, the slope between acetonitrile and CO being to urban emissions (Warneke et al., 2006), and the concentrations of anthropogenic tracer CHCl3 being high (Warneke et al., 2006). Similar to acetonitrile, another common BB tracer hydrogen cyanide (HCN) was also enhanced in these air masses. BB data (acetonitrile > 200 pptv) for SEAC4RS were analyzed separately and are inset in Fig. 2. Although all CalNex data had a tight correlation, we only included the flight data near LA basin to target the area strongly influenced by anthropogenic emissions. In general,
the correlation coefficients between in situ OA and HCHO were strong ($r = 0.59 - 0.97$) (Table 1). The variety in OA-HCHO regression coefficients among different campaigns reflects the regional and/or source-driven OA-HCHO variability. Considering only the non-biomass burning (non-BB) air masses sampled, OA and HCHO had the tightest correlation for CalNex, because CalNex focused on the LA area (shown in Fig. 2) and Central Valley while SEAC4RS and DC3 covered a larger area with a potentially larger variety of sources and chemical conditions. Although SEAC4RS and DC3 both sampled the continental US, SEAC4RS had more spatial coverage and sampled more air masses at low altitudes, while DC3 was designed to sample convective outflow air masses and had more data at high altitudes. Although KORUS-AQ covered a much smaller area compared to SEAC4RS, KORUS-AQ data also had a large spread, which may be due to the complicated South Korean anthropogenic sources mixed with transported air masses (mainly from China) and possibly biogenic sources. OA exhibits a tight correlation with HCHO for both wildfires and agricultural fires. This is because the production of HCHO and OA is much higher in BB air masses compared to background. This may also suggest that the emissions of OA and HCHO in these air masses are relatively constant. More intensive fire sampling is needed to probe the correlation between OA and HCHO across fuel types and environmental conditions.

The different slopes of OA-HCHO among different campaigns also reflect the regional or source-driven OA-HCHO variability. The slopes of OA vs. HCHO were more similar between DC3 and SEAC4RS. The slope of OA vs. HCHO was higher for South Korea
(KORUS-AQ), which is dominated by anthropogenic emissions compared to the biogenic-dominated emission in the Southeast US (SEAC4RS and DC3). The high OA air masses also had high acetonitrile during KORUS-AQ. The slope of OA-HCHO for California LA basin, dominated by anthropogenic emissions, was also not as high as South Korea. The potential difference in the anthropogenic emissions mix could contribute to the different OA-HCHO slopes from US LA region and South Korea anthropogenic sources (Baker et al., 2007; Na et al., 2005; Na et al., 2002). The slopes of OA vs. HCHO for BB air masses were higher than for anthropogenic and biogenic sources. This is consistent with high POA emission in BB conditions (Heald et al., 2008; Lamarque et al., 2010; Cubison et al., 2011), despite low addition of mass due to SOA formation (Cubison et al., 2011; Shrivastava et al., 2017). The slope of OA to HCHO was higher for wildfires than agricultural fires. This may indicate that more OA is emitted in wildfires which often have higher intensity than agricultural fires (Liu et al., 2017; Forrister et al., 2015). As SEAC4RS had the largest geographic coverage for low altitude data over US, the campaign average slope of OA vs. HCHO was used to represent the US region in summer, except large cities where CalNex LA Basin data were used.

3.2 Dependence on NOx and VOCs speciation

Biogenic and anthropogenic VOCs are oxidized by atmospheric oxidants (e.g. OH as the dominant oxidant) to form RO2. HCHO is produced from the reactions of RO2 with HO2 or NO, with RO2+NO typically producing more HCHO than RO2 + HO2 (e.g. Wolfe et al., 2016). RO2 can react with HO2, NO, or isomerize to form oxidized organic
compounds with high molecular weight and low volatility, which condense on existing particles to form SOA. The products of RO₂ + NO tend to fragment instead of functionalize and often lead to higher volatility compounds (e.g. HCHO) and thus less SOA formation compared to the products of RO₂ + HO₂ (Kroll et al., 2006; Worton et al., 2013). Therefore, with the same VOC, we expect more HCHO and less OA formed at high NO conditions and vice versa. As mentioned before, NO₂ instead of NO is easily measured from space and NO₂ typically is ~80% of NOₓ in the boundary layer during the day. Therefore, NO₂ is used as a surrogate for the NO levels influencing OA and HCHO production. The yields of HCHO and SOA also depend on VOC speciation (e.g. Lee et al., 2006). Specifically, isoprene has a higher yield of HCHO than most non-alkene VOCs (Dufour et al., 2009).

A scatter plot of OA vs. HCHO for SEAC₄RS low altitude data is shown in Fig. 3(a). The data are color-coded by the product of in-situ isoprene and NO₂, attempting to capture time periods strongly influenced by oxidation products of isoprene at high NO conditions. No obvious trends are evident when the data are instead color coded by NO₂ or isoprene only. This may be because isoprene (biogenic source) and NO₂ (anthropogenic sources) are generally not co-located in the US (Yu et al., 2016) and isoprene is the dominant source of HCHO compared to anthropogenic VOCs in the US (e.g. Millet et al., 2008). This plot shows that, at high NO₂ and high isoprene conditions, less OA was formed for each HCHO produced generally. This is consistent with high NO and isoprene conditions promote HCHO formation over SOA formation. We also looked at the dependence on peroxy acetyl nitrate (PAN), as PAN is a product of the photo
oxidation of VOCs, including isoprene, in the presence of NO$_2$. The dependence on PAN was not as clear as on the product of NO$_2$ and isoprene.

KORUS-AQ OA vs. HCHO, color-coded with NO$_2$, is plotted in Fig. 3(b). The OA / HCHO ratio clearly decreased as NO$_2$ levels increased during KORUS-AQ, suggesting that high NO conditions accelerated HCHO formation more than they did SOA production. Compared to SEAC4RS, the KORUS-AQ OA/HCHO ratio does not depend on VOCs. This may be consistent with the dominant VOC being anthropogenic VOCs that are co-located with NO sources. This may also suggest the anthropogenic VOCs generally have a lower HCHO yield than does isoprene.

4. Comparison of OA and HCHO relationship: in-situ vs. GEOS-Chem

In situ OA vs. HCHO relationships from SEAC$^4$RS low altitude non-BB (Fig. 4a), KORUS-AQ low altitude (Fig. 4b), and SEAC$^4$RS BB (Fig. 4c) air masses were compared to GEOS-Chem model simulations (Fig. 4d-4f) sampling along the corresponding flight tracks. Similar to the in situ data, GEOS-Chem model simulations also found correlations between OA and HCHO for these three regions, especially for SEAC$^4$RS non-BB. GEOS-Chem was intensively validated with in situ measurements for Southeast US (e.g. Marais et al., 2016; Kim et al., 2015). The ratios of the slopes between OA and HCHO for the US (SEAC$^4$RS), South Korea (KORUS-AQ), and wildfire cases (SEAC$^4$RS) from GEOS-Chem were 1:1:1:0.4, which was different from the in situ measurements ratios of 1:1:4:13 (Table 1). GEOS-Chem could not capture any wild fires
in US during SEAC4RS, which is probably due to poor representation of BB emission inventory for US wildfire and also the coarse grid in GEOS-Chem. GEOS-Chem also significantly underpredicted the slope of OA to HCHO for South Korea. We attribute this to a likely underprediction of anthropogenic SOA, which was dominant in South Korea, in GEOS-Chem (Schroder et al., 2018), as well as a different mix of OA and HCHO sources in the US compared to South Korea and representation of these in GEOS-Chem. Although GEOS-Chem contains isoprene chemistry with a focus on the Southeast US (Marais et al. 2016), there is still room to improve the model especially for anthropogenic and BB sources, as well as anthropogenic OA formation mechanisms. For example, in the model biogenic sources are more important than anthropogenic sources for the OA and HCHO budgets in South Korea, which is not the case from KORUS-AQ in situ measurements. A 10% decrease of emissions from biogenic, anthropogenic and BB sources results in a 6%, 3%, and 1% decrease in OA and 2%, 1%, and 0% decrease in HCHO over South Korea in May 2016. However, the in situ airborne field campaign KORUS-AQ found that OA and HCHO were higher near anthropogenic emission sources compared to rural regions. The larger impact of biogenic sources compared to anthropogenic sources on OA and HCHO in the model can be due to both low-biased anthropogenic emission inventories and low-biased anthropogenic SOA. Improving anthropogenic emissions inventories in the models can also bring model results closer to observations. Improving anthropogenic SOA, such as implementation of the SIMPLE model, in GEOS-Chem (Hodzic and Jimenez, 2011) can also improve the model results compared to observations. Measurements or measurement-constrained estimation with sufficient spatial and temporal coverage can help to narrow down the key factors (e.g.
emission inventories or chemical schemes) in GEOS-Chem to better represent VOCs and OA globally. Furthermore, we did also find that GEOS-Chem could not capture the observed higher slope of OA to HCHO at high altitudes (not shown), which could be due to issues such as transport, OA lifetime, and OA production.

5. Relating satellite HCHO column to surface HCHO concentrations

To utilize the derived in-situ OA and HCHO relationship, the satellite HCHO column needs to be converted to surface HCHO concentrations. We used a vertical distribution factor $\eta$ (cm$^{-1}$) (Sect. 2.5), which is defined as the ratio of surface HCHO concentrations (molec cm$^{-3}$) to HCHO column (molec cm$^{-2}$), to estimate surface HCHO concentrations from satellite column measurements. Zhu et al. (2017) used the same vertical distribution factor for their study. The use of this factor is justified by the fact that the derived surface HCHO retained the spatial pattern of the satellite HCHO column and agreed with local surface measurements of HCHO for a multi-year average (Zhu et al., 2017).

We also investigated the main factors affecting the variation of the vertical distribution factor $\eta$. Because the factor is determined by HCHO vertical distributions, we examined three typical normalized HCHO vertical distribution profiles with the highest, median and lowest $\eta$ values for the SEAC$^4$RS field campaign (Fig. 5). GEOS-Chem can generally capture the vertical profiles of measured HCHO. Boundary layer mixing height and surface emission strength are the dominant factors in determining the fraction of HCHO.
near the surface. We can see that higher boundary layer mixing height results in lower $\eta$ for southeast US profiles, where there are biogenic sources of HCHO from the surface and HCHO has distinct concentration difference below and above the boundary layer. However, there are exceptions, such as for the profiles over the ocean and the coastal regions. Although the boundary layer is shallow in these regions, a large portion of HCHO resides above the boundary layer, resulting in low $\eta$. In these cases, surface emissions of HCHO or precursors are very small and therefore methane oxidation makes a large contribution to the total HCHO column. High concentrations of HCHO (e.g., in BB plumes) lofted by convection can also impact the vertical profile (Barth et al., 2015). Overall, the source intensities and boundary layer mixing height mostly determined the HCHO vertical profiles.

6. Construction of the OA estimate

6.1 Variables to construct OA estimate

As mentioned in Sect. 2.5, the OA estimate value in each grid cell is estimated from satellite HCHO column observation by the linear Eq. (1). Satellite HCHO columns, $\Omega_{\text{HCHO}}$, are converted to surface HCHO concentrations by multiplying by the $\eta(i)$ factor. Surface OA is then estimated by the derived surface HCHO concentrations and applying the linear regression equation (slope $\alpha(i)$ and intercept $\beta(i)$) between in situ OA and HCHO determined from in-situ aircraft field campaign data. The relationship between OA and HCHO varies but previous sections demonstrated that we can quantify the surface OA-HCHO relationship by their regions, sources and chemical conditions (e.g., NOx levels). To test the impact of the chosen OA-HCHO relationship on the calculated
OA estimate, the OA estimate in the US was calculated using four different methods (see Table 2).

6.2 OA estimate over US

The monthly average surface OA estimate over the US in August 2013 for case 1 (Sect. 6.1) is shown in Fig. 6(a). Because BB regions in the US are not covered by smoke continuously during a period of time and it is challenging for satellite retrieval to separate thick BB plumes and clouds without information on the time and location of the burning, thick BB events (OMI UV Aerosol Index (UVAI) > 1.6) (Torres et al., 2007) were excluded and shown as the blank (white) grid cells in Fig. 6(a). The same filter was also applied to aerosol extinction and GEOS-Chem OA abundance. To evaluate the representative quality of the OA estimate, OA estimate data were compared to the EPA IMPROVE ground sites corrected-OA measurements over the US and SEARCH ground sites OA measurements in the Southeast US (Sect. 2.2). The locations of IMPROVE and SEARCH sites are displayed in Fig. 6(g) as small and large dots, respectively. The dot color represents the average OA mass concentrations for August 2013.

The good correlation between the OA estimate and corrected IMPROVE network measurements (Fig. 6(e)) indicates that the OA estimate generally captured the variation of OA loading over the US. The correlation between the OA estimate and IMPROVE OA measurements yielded a slope of 0.62, indicating that the OA estimate slightly underestimated OA. Satellite HCHO data were measured in mid-day, in situ airborne OA and HCHO were measured during the daytime and IMPROVE network organic carbon
was collected day and night. Because ground OA in the Southeast US were observed to have little diurnal variation (Xu et al., 2015; Hu et al., 2015), the different sampling time of ground and airborne OA probably does not have a significant impact on the comparison of OA estimate and IMPROVE OA. The difference in mid-day and daytime HCHO concentrations is not prominent, depends on the location and may contribute to a small bias to the mid-day OA estimate (DiGangi et al., 2012). This is probably due to increased boundary layer height diluting the photochemical formation of HCHO in the mid-day. Instead, the potential underestimation of HCHO from satellite retrieval (by 37%) (Zhu et al., 2016) compared to SEAC4RS may cause the low slope between the OA estimate and IMPROVE OA according to Eq. (1). Also, the uncertainty in η estimated from GEOS-Chem could have also contributed to the low slope between the OA estimate and EPA ground site OA. Furthermore, the uncertainties in IMPROVE OA measurements, such as using a correction factor to correct the partial evaporation across all southeast US sites, and real variations on the OA/OC ratio, may also have contributed to the discrepancies between the OA estimate and EPA IMPROVE sites OA.

SEARCH OA data were also used to compare to the OA estimate. The correlation was good for August 2013. Although the SEARCH network OA measurements have better accuracy, the number of SEARCH sites is limited (5 sites). The correlation of OA estimate and SEARCH OA varied dramatically 2008-2013 (Fig. S1). GEOS-Chem OA does not correlate with SEARCH OA except for the year 2013 (Fig. S1). As the IMPROVE network has more sites and spatial coverage, we use IMPROVE network data as ground OA measurements for comparison in the remainder of the discussion.
6.3 Comparison to aerosol extinction from AOD

To further evaluate the method of using satellite HCHO to derive an OA surface estimate, satellite measurements of AOD were converted to extinction for comparison. Studies showed that OA were a dominant component of aerosol mass and extinction during SEAC4RS (Kim et al., 2015; Wagner et al., 2015) and the fractions of OA were relatively constant (interdecile 62-74%) (Wagner et al., 2015). Therefore AOD variation is expected to generally reflect the OA variation during SEAC4RS. Satellite measurements from MISR can provide more aerosol property information to apportion total AOD to AOD of a subset of aerosols with small to medium size and round shape, which can better capture OA, when AOD is above 0.15 to 0.2 (Kahn and Gaitley, 2015; personal communication with R. Kahn, 2018). Because MISR cannot distinguish OA and other submicron aerosol components (e.g. sulfate and nitrate), the fractions of OA were relatively constant, and AOD below 0.15 or 0.2 accounts for near half of the data over US in August, 2013, we use total AOD to derive extinction for our comparison. The AOD-derived extinction map is shown in Fig. 6(b), and the scatter plot of AOD-derived extinction and EPA corrected OA is displayed in Fig. 6(f). The same filter of high AI was also applied to AOD-derived extinction to remove BB plumes. Generally, the derived aerosol extinction had a correlation with IMPROVE OA, but the correlation is not as good as for the OA estimate with IMPROVE OA. This indicates that the OA estimate derived from HCHO may be better than AOD at representing the concentrations of OA, even for the regions where AOD is dominated by OA (Xu et al., 2015).
6.4 Comparison to GEOS-Chem OA

Surface OA over the US from a GEOS-Chem simulation for August 2013 is shown in Fig. 6(c), and the scatter plot of GEOS-Chem OA with IMPROVE OA is in Fig. 6(g). The GEOS-Chem simulation had a coarser resolution than satellite HCHO data. To be comparable to the OA estimate, the scatter plot Fig. 6(g) used GEOS-Chem results for the grid squares that overlap with individual IMPROVE sites. Compared to the OA estimate, GEOS-Chem OA had a similar correlation coefficient with IMPROVE OA. Although the GEOS-Chem OA plot appears more scattered, there are many GEOS-Chem data points close to zero when IMPROVE OA was low, making the overall correlation coefficient similar to that for the OA estimate. GEOS-Chem under predicted IMPROVE OA more with a slope of 0.57 compared to the OA estimate. This is consistent with underprediction of anthropogenic OA in Marais et al. (2016).

6.5 OA estimate with different OA-HCHO relationships

In general, OA estimate results from the four cases were similar. The OA proxies from the four cases (Table 2) were compared to IMPROVE OA and the correlation coefficients are shown in Fig. 7. Applying CalNex LA Basin in situ data (case 3) to bring down the OA concentrations of LA resulted in a similar correlation. The uncertainties in IMPROVE OA measurements may contribute to this. Including the NO₂–isoprene-dependent OA and HCHO relationship (case 2) showed a similar (or slightly worse) correlation between the OA estimate and IMPROVE OA. As the in situ data showed a NO₂–isoprene-dependent OA and HCHO relationship, we attributed this to the uncertainty of isoprene emissions from MEGAN or IMPROVE network measurements.
Because separating large urban areas and other regions and applying a chemical region dependent in situ OA and HCHO relationship did not improve the agreement between the OA estimate and IMPROVE OA, we used the base case OA and HCHO relationship (case 1) to derive the OA estimate (shown in Fig. 6).

### 6.6 Temporal variation of the agreement between OA estimate and IMPROVE OA

Besides August 2013 (see Fig. 6), the correlations between the OA estimate and IMPROVE OA for the summer months June-July-August 2008-2013 were also examined and shown in Fig. 7. Generally, the correlation coefficients between the OA estimate and IMPROVE OA were >0.5 for summer months of the years investigated. The correlation coefficients were generally higher in June compared to July and August. The lower average temperature in June might be related to the higher correlation coefficients. IMPROVE network aerosol samples were transported at ambient temperature in a truck and more organic vapors likely evaporated at higher temperature. The different temperatures and distances from IMPROVE sites to the laboratory may lead to inhomogeneous evaporation among the samples and result in lower correlation coefficients. Although higher temperatures in July and August may also lead to more BB, average aerosol index over the US was not higher in July (mean: 0.35) and August (mean: 0.36) compared to June (mean: 0.39) for these years. The underlying cause for the lowest correlation coefficients in July and August 2012 is not clear and may be related the severe drought in 2012 (Seco et al., 2015). The correlation coefficients were also low for the linear regressions (not shown) of IMPROVE OA with both GEOS-Chem OA and
AOD-derived extinction. Because the lowest correlation coefficients were consistently observed for multiple OA-related products and not just the OA estimate, we attributed this to uncertainties in the IMPROVE OA measurements or some unknown bias shared by the satellite HCHO, GEOS-Chem OA, and satellite AOD.

6.7 South Korea OA estimate

We attempted to estimate an OA estimate for South Korea, using airborne in situ measurements of OA and HCHO from the KORUS-AQ field campaign (https://www-air.larc.nasa.gov/missions/korus-aq/) and SAO OMI HCHO measurements. The National Institute of Environmental Research (NIER) ground sites OC measurements during KORUS-AQ over South Korea could be used to validate the OA estimate. However, OMI HCHO measurements were below the detection limit (Zhu et al., 2016) in May 2016. Also, there were no OMI data available in June 2016 when airborne measurements and ground sites OC measurements were available during KORUS-AQ. Because an OA estimate for South Korea could not be well retrieved and validated, it was not presented in this study.

7 Limitations of the OA estimate and future work

Because the OA estimate is based on satellite HCHO data, the detection limit of satellite HCHO data affects the quality of the OA estimate. Currently, due to the limited sensitivity of OMI for HCHO, the OA estimate is valid only when high levels of HCHO are present, such as during summer time and near large HCHO sources. With the new TROPOMI satellite instrument and future missions TEMPO and GEMS, satellite HCHO
measurements will have higher spatial and temporal resolutions and lower detection limits. These higher quality satellite HCHO measurements will improve our OA estimate quality and also its spatial and temporal coverage.

Because the OA estimate uses the relationship of in situ HCHO and OA measurements, the coverage of in situ aircraft field campaigns will impact the OA estimate quality. Currently, in situ airborne measurements of OA and HCHO focus on the continental US. Extending measurements to regions such as Africa BB, South America, and East Asia, where HCHO and OA have high concentrations, will increase the spatial coverage of the OA estimate product. Ground site measurements of OA with consistent quality control in those regions will also be important for validating the OA estimate.

Improvement of satellite HCHO retrieval during the BB cases will also improve OA estimate quality. BB cases with high UV aerosol index over the US were excluded in the current OA estimate analysis. With improvement in the satellite retrieval of HCHO, we may be able to estimate OA during BB cases over the US. Upcoming field campaigns such as the Fire Influence on Regional and Global Environments Experiment – Air Quality (FIREX-AQ) will provide opportunities to improve the analysis of OA estimate in BB cases in the US.

This OA estimate method has limitations in remote regions far away from HCHO sources. Because the lifetimes of HCHO (1-3 hours) and OA (1 week) are different, the slopes and intercepts between HCHO and OA are expected to change when air masses
are aged (e.g. in remote regions). OA vs. HCHO from SEAC4RS and KORUS-AQ field campaigns, color-coded with altitude, are plotted in Fig. S2 (a) and (b), respectively. A relative depletion of HCHO at high altitudes was observed due to its shorter lifetime. This also suggests that, at remote regions far away from the sources, the ratios of OA and HCHO could be much higher and the relationship between OA and HCHO derived near the sources may no longer apply. On the other hand, the lifetime of 1-3 hrs for HCHO does not imply that the OA estimate only work within this timescale. HCHO is formed from oxidation of transported gas phase VOCs, including the oxidation products of the primary emitted VOCs, as well as of the slower reacting VOCs (e.g. Ethane and Benzene). Most gas-to-particle oxidation processes that might produce HCHO can last up to 1-2 days (Palm et al., 2018). Fig. S3 shows the ratios of OA and HCHO did not change significantly downwind for the Rim Fire plume for about 1 day of aging, which was determined by the distance from the source and the wind speed. A lower photolysis rate of HCHO in the plume can also contribute to this. However, we do not expect that the relationship of OA and HCHO remains past 1-2 boundary layer ventilation cycles (Palm et al., 2018).

8 Summary

We have developed a satellite-based estimate of the surface OA concentration (“OA estimate”) based on in situ observations. This estimate is based on the empirical relationships of in-situ OA and HCHO for several regions. OA and HCHO share VOC sources with different yields and lifetimes. Using surface OA and HCHO linear
regression slopes and intercepts can relate surface HCHO to OA. To estimate the surface
HCHO concentration from satellite HCHO column, we used a vertical distribution factor
\( \eta \), which is largely determined by boundary layer height and surface emissions and found
to reasonably retrieve surface HCHO from column HCHO.

The OA estimate over the continental US generally correlated well with EPA IMPROVE
network OA measurements corrected for partial evaporation. The good correlations are
not only for the time during SEAC4RS but also for most summer months over several
years (2008-2013) investigated. Compared to aerosol extinction derived from AOD, the
OA estimate had slightly higher correlation coefficients with IMPROVE OA. GEOS-
Chem can predict OA with a similar correlation coefficient with IMPROVE OA
compared to the OA estimate when GEOS-Chem was intensively validated with in situ
measurements for Southeast US. Better satellite HCHO data from TROPOMI and future
TEMPO and GEMS and extending spatiotemporal coverage of in situ measurements will
improve the quality and coverage of the OA estimate.
Acknowledgements:

JL, TFH, GMW, and JSC were supported by NASA grant NNH15ZDA001N and NNH10ZDA001N. BAN, PCJ, and JLJ were supported by NASA grant NNX15AT96G and 80NSSC18K0630. AW and PTR-MS measurements during DC3, SEAC4RS and KORUS-AQ were supported by the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit) through the Austrian Space Applications Programme (ASAP) of the Austrian Research Promotion Agency (FFG). The PTR-MS instrument team (P. Eichler, L. Kaser, T. Mikoviny, M. Müller) is acknowledged for their field support.
References:


conditions, Atmos Chem Phys, 16, 4369-4378, 10.5194/acp-16-4369-2016, 2016.


Table 1. Linear regression parameters for OA vs. HCHO at low altitudes (<1 km)

<table>
<thead>
<tr>
<th>Method</th>
<th>US (SEAC^RS)</th>
<th>US (DC3)</th>
<th>US (CalNex)</th>
<th>South Korea (KORUS-AQ)</th>
<th>Wild Fires (SEAC^RS)</th>
<th>Agricultural Fires (SEAC^RS)</th>
<th>SEAC^RS Low and Isoprene</th>
<th>SEAC^RS High and Isoprene</th>
</tr>
</thead>
<tbody>
<tr>
<td>In situ measurements OA v.s. HCHO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope (× 10^{-11})</td>
<td>1.93 ± 0.07</td>
<td>1.30 ± 0.10</td>
<td>1.34 ± 0.02</td>
<td>2.75 ± 0.05</td>
<td>25.09 ± 0.04</td>
<td>13.32 ± 0.27</td>
<td>1.99 ± 0.07</td>
<td>1.19 ± 0.16</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation coefficient r</td>
<td>0.76</td>
<td>0.88</td>
<td>0.70</td>
<td>0.97</td>
<td>0.88</td>
<td>0.83</td>
<td>0.64</td>
<td>0.44</td>
</tr>
</tbody>
</table>

GEOS-Chem model sampled along the flight track OA v.s. HCHO

<table>
<thead>
<tr>
<th>Method</th>
<th>US (SEAC^RS)</th>
<th>US (DC3)</th>
<th>US (CalNex)</th>
<th>South Korea (KORUS-AQ)</th>
<th>Wild Fires (SEAC^RS)</th>
<th>Agricultural Fires (SEAC^RS)</th>
<th>SEAC^RS Low and Isoprene</th>
<th>SEAC^RS High and Isoprene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope (× 10^{-11})</td>
<td>1.25 ± 0.03</td>
<td>1.39 ± 0.05</td>
<td>0.48 ± 0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.32 ± 0.11</td>
<td>1.88 ± 0.07</td>
<td>0.12 ± 0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Coefficient r</td>
<td>0.76</td>
<td>0.43</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The unit of the slope is g g^{-1}.

The uncertainties are one standard deviation.

Table 2. Methods to estimate OA surface concentrations, based on the choice of slope and intercept from a linear regression relationship between OA and HCHO data found in Table 1.

1* Using non-BB SEAC^RS relationship to represent all continental US

2* Using NO_2 and isoprene dependent non-BB SEAC^RS relationship for all continental US

3 Using the CalNex LA Basin relationship for large urban cites and the non-biomass burning SEAC^RS relationship for other US regions

4* Using the CalNex LA Basin relationship for large urban cites and the NO_2 and isoprene dependent non-BB SEAC^RS relationship for other US regions

*SEAC^RS was chosen to represent all continental US because it had the largest horizontal and vertical coverage.

b In methods 2 and 4, when the product of NO_2 column (Sect. 2.3) and surface isoprene emission rate (Sect. 2.4) was above threshold of 5×10^{12} molec cm^{-2} atom C cm^{-2} s^{-1}, the slope and intercept from SEAC^RS high isoprene and NO_2 conditions were used. When the NO_2 column–isoprene emission product was below that threshold, the slope and intercept from SEAC^RS low isoprene and NO_2 conditions were used. Threshold of “isoprene × NO_2” is determined by its mean value over southeast US (83° to 96° W and 32° to 35°N). Large urban cities are categorized with high NO_2 vertical columns (>4 ×10^{15} molec cm^{-2}) (Tong et al., 2015) based on the satellite NO_2 levels over LA.
emissions instead of concentrations are used because global models use isoprene emission inventory to simulate isoprene concentrations and isoprene emission inventory is easier to access. Since isoprene has a short-lifetime of up to a few hours (Guenther et al., 2006), the emissions have a similar spatiotemporal distribution as the concentrations.
Figure 1. Flight tracks of airborne field campaigns SEAC4RS (blue), DC3 (black), CalNex (red) and KORUS-AQ (green), of which in situ OA and HCHO measurements were used.
Figure 2

Scatter plots of in situ OA (µg m$^{-3}$) vs. HCHO (µg m$^{-3}$ or molec cm$^{-3}$) from SEAC4RS (excluding biomass burning) (blue), DC3 (dark grey), CalNex (pink), and KORUS-AQ (green) low altitude (< 1 km) data. Inset shows wildfire (brown), and agricultural fire (grey) SEAC4RS data. SEAC4RS biomass burning cases are defined as acetonitrile > 200 pptv. The linear regression fits are shown as the darker lines and correlation coefficients are provided.
Figure 3. (a) A scatter plot of OA vs. HCHO for SEAC4RS non-biomass burning low altitude data color-coded with the product of NO$_2$ and isoprene in log scale. The red and blue lines are the linear regression fits of high (> 0.5) and low (<0.5) product of NO$_2$ (ppbv) and isoprene (ppbv), respectively. (b) A scatter plot of OA vs. HCHO for KORUS-AQ data color-coded by log(NO$_2$).

Figure 4 Scatter plots of OA vs. HCHO for US (SEAC4RS altitude < 1 km non-biomass burning), South Korea (KORUS-AQ altitude < 1 km) and wildfire (SEAC4RS) from in situ measurements (a, b, c) and GEOS-Chem outputs sampled along the flight tracks (d,e,f).
Figure 5. Three typical vertical profiles of the ratio of in situ HCHO concentrations (molec cm$^{-3}$) to integrated HCHO column from SEAC$^4$RS flight track. These three profiles were located at Kansas-Oklahoma border (red), Arkansas-Tennessee border (black), and Gulf of Mexico (blue). Solid curves were from GEOS-Chem results and the dashed were from ISAF measurements. HCHO columns were integrated HCHO concentrations of these vertical profiles extrapolated from 0 to 10 km, assuming the HCHO below and above the measured HCHO vertical profiles were the same as the HCHO at the lowest and highest altitudes sampled, respectively. The boundary layer heights (BLH) of these three profiles were plotted by the shaded areas.
Figure 6. (a) The maps of (a) surface OA estimate (Case 1), (b) surface aerosol extinction derived from MODIS AOD, (c) GEOS-Chem simulated surface OA, and (d) EPA IMPROVE (small dots) and SEARCH (large dots) network ground sites color coded with OA concentrations for August 2013. The scatter plots of (e) surface OA estimate, (f) surface aerosol extinction derived from MODIS AOD, and (g) surface GEOS-Chem OA vs. EPA IMPROVE network ground sites OA. IMPROVE sites OA were corrected for evaporation. (h) The scatter plots of surface OA estimate and GEOS-Chem OA vs. SEARCH network ground sites OA for August 2013. GEOS-Chem OA and OA estimate did not have good correlations with SEARCH OA for other years (SI). For the scatter plots, linear regressions are shown (blue and green lines) and regression equations and correlation coefficients for the scatter plots are listed. The dashed lines in the scatter plots indicate the 1:1 line. Biomass burning data (UV aerosol index > 1.6) were excluded in all panels.
Figure 7. The correlation coefficients of the linear regression between the OA estimate from 4 cases (red, blue, gray, and yellow) vs. EPA-corrected OA from 2008-2013 for June, July, and August. The monthly average ambient temperature is in black.