Source and variability of formaldehyde (HCHO) at northern high latitude: an integrated satellite, aircraft, and model study

Tianlang Zhao1, Jingqiu Mao1, William R. Simpson1, Isabelle De Smedt2, Lei Zhu3, Thomas F. Hanisco4, Glenn M. Wolfe4, Jason M. St. Clair4,5, Gonzalo González Abad6, Caroline R. Nowlan6, Barbara Barletta7, Simone Meinardi7, Donald R. Blake7, Eric C. Apel8 and Rebecca S. Hombrook8

1 University of Alaska Fairbanks, Department of Chemistry and Biochemistry & Geophysical Institute, Fairbanks, AK, United States

2 Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

3 Southern University of Science and Technology, School of Environmental Science and Engineering, Shenzhen, China

4 NASA Goddard Space Flight Center, Atmospheric Chemistry and Dynamics Lab, Greenbelt, MD, United States

5 University of Maryland Baltimore County, Baltimore, MD, United States

6 Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, United States

7 University of California Irvine, Irvine, CA, United States
Here we use satellite observations of formaldehyde (HCHO) vertical column densities (VCD) from the TROPOspheric Monitoring Instrument (TROPOMI), aircraft measurements, combined with a nested regional chemical transport model (GEOS-Chem at 0.5°×0.625° resolution), to understand the variability and sources of summertime HCHO better in Alaska. We first evaluate GEOS-Chem with in-situ airborne measurements during Atmospheric Tomography Mission 1 (ATom-1) aircraft campaign. We show reasonable agreement between observed and modeled HCHO, isoprene, monoterpenes, and the sum of methyl vinyl ketone and methacrolein (MVK+MACR) in continental boundary layer. In particular, HCHO profiles show spatial homogeneity in Alaska, suggesting a minor contribution of biogenic emissions to HCHO VCD. We further examine the TROPOMI HCHO product in Alaska summer, reprocessed by GEOS-Chem model output for a priori profiles and shape factors.

For the year with low wildfire activity (e.g., 2018), we find that HCHO VCDs are largely dominated by background HCHO (66-71%), with minor contributions from wildfires (20-32%) and biogenic VOC emissions (8-10%). For the year with intense wildfires (e.g., 2019), summertime HCHO VCD is dominated by wildfire emissions (50-72%), with minor contributions from background (22-41%) and...
biogenic VOCs (6-10%). In particular, the model indicates a major contribution of wildfires from direct emissions of HCHO, instead of secondary production of HCHO from oxidation of larger VOCs. We find that the column contributed by biogenic VOC is often small and below the TROPOMI detection limit, in part due to the slow HCHO production from isoprene oxidation under low NOx conditions. This work highlights challenges for quantifying HCHO and its precursors in remote pristine regions.

1. Introduction

The Arctic (north of 66.5°N) and boreal region (between 45°N and 65°N) have undergone dramatic temperature and ecological changes over the past century and the rate of this change has accelerated in recent decades (Cohen et al., 2014). Satellite-based observations of leaf area index (LAI) and normalized difference vegetation index (NDVI) suggest that northern high latitudes show a significant trend of greening in the past three decades as a result of vegetation growth (Bhatt et al., 2017; Keeling et al., 1996; Myers-Smith et al., 2011; Myneni et al., 1997; Xu et al., 2013; Zhou et al., 2001; Zhu et al., 2016), in part because the temperature is the limiting factor for vegetation growth in this region (Nemani et al., 2003). In the meantime, boreal forest fires have shown an increasing trend over the past few decades, which is likely to continue (Abatzoglou and Williams, 2016).

Volatile organic compounds (VOCs) emitted from terrestrial vegetation play a major role in air quality and chemistry-climate interactions (Guenther et al., 1995). These biogenic VOCs (BVOCs) undergo
photochemical degradation, leading to the formation of ozone and aerosol particles that play major roles in climate and air quality (Mao et al., 2018). Biogenic VOCs account for more than 80% of global VOC emissions and represent a major source of reactive carbon to the atmosphere (Guenther et al., 1995, 2006). Primary biogenic VOC emissions include both isoprene (2-methyl-1,3-butadiene, C$_5$H$_8$) and monoterpenes (a class of terpenes that consist of two isoprene units, C$_{10}$H$_{16}$), rapidly producing HCHO through oxidation after emitted to the atmosphere (Millet et al., 2006; Palmer et al., 2006). The emissions of these biogenic VOCs are dependent on the air temperature, light intensity, plant functional type (PFT), leaf area index (LAI), leaf age, soil moisture, ambient carbon dioxide (CO$_2$) concentrations and a number of other factors (Guenther et al., 2006). It has been suggested that at least some ecosystems in the northern high latitudes are highly sensitive to temperature, leading to a strong increase in BVOC emissions in recent years (Kramshøj et al., 2016; Lindwall et al., 2016). BVOC emissions are further complicated by land cover and LAI changes in this region (Tang et al., 2016).

Biogenic VOC emissions in the Arctic and boreal region are poorly characterized, due to lack of measurements. Previous measurements have been generally focused on European boreal forests with a major focus on monoterpenes (Bäck et al., 2012; Juráň et al., 2017; Rantala et al., 2015; Rinne et al., 2000; Spirig et al., 2004; Zhou et al., 2017). Biogenic VOC emissions in other boreal forests outside of Europe have been rarely quantified. Some early aircraft-based measurements show abundant isoprene in Alaskan boreal forests (Blake et al., 1992), suggesting a major gap in current understanding of BVOC
emissions in this region. Isoprene fluxes in tundra systems have been measured in Greenland (Kramshøj et al., 2016; Lindwall et al., 2016; Vedel-Petersen et al., 2015), northern Sweden (Faubert et al., 2010; Tang et al., 2016) and the Alaskan North Slope (Angot et al., 2020; Potosnak et al., 2013). All these tundra measurements show a very strong positive temperature dependence for isoprene fluxes, likely due to higher emission potentials for isoprenoids than temperate species (Rinnan et al., 2014). The high temperature sensitivity suggests an important role of climate warming on BVOC emissions.

HCHO serves as an important indicator of BVOC emissions on regional and global scales (Millet et al., 2006). The HCHO column density has been observed from space by several satellite sensors including the Global Ozone Monitoring Experiment (GOME) (Palmer et al., 2001), Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) (De Smedt et al., 2008), and Ozone Monitoring Instrument (OMI) (González Abad et al., 2015). A number of studies use satellite-based observations of the HCHO column density to quantify regional and global isoprene emissions in vegetated regions (Guenther et al., 2006; Millet et al., 2008; Palmer et al., 2003, 2006; Stavrakou et al., 2009, 2014), and their interannual variability (De Smedt et al., 2010, 2015; Stavrakou et al., 2018, 2015, 2014; Zhu et al., 2017; Bauwens et al., 2016).

Biomass burning represents another major source of HCHO from both primary emissions and secondary production from VOC precursors. Biomass burning is the second largest source of global non-methane...
volatile organic compounds (NMVOCs) after biogenic emissions (Yokelson et al., 2008). The GFED4s burned area dataset including small fires shows that boreal forests are responsible for 2.5% of global burned area but 9% of fire carbon emission and 15% of fire methane (CH$_4$) emission (van der Werf et al., 2017). Several studies have reported a high level of HCHO emitted from wildfire plumes. Liu et al (2017) found formaldehyde as the second most abundant NMVOC from wildfires in western US, with an emission factor of 2.3 (±0.3) g/kg dry matter for temperate forests and a similar emission factor for boreal forest fires. We-CAN aircraft measurement reports the HCHO emission factor in near-fire smoke plume to be 1.9 (±0.43) g/kg (Permar et al., 2021). As boreal fires have become more intense in the past few decades (Macias Fauria and Johnson, 2008), HCHO from boreal fires are likely to play an important role in the temporal and spatial variability of HCHO in this region.

While satellite-based observations of HCHO appear promising, their application in air quality and regional photochemical modeling remains challenging. There are large uncertainties and inconsistencies among different satellite-based sensors and retrieval methods for HCHO, due to instrumental sensitivity, retrieval algorithms, timing of observation with respect to the diurnal cycle, as well as several other factors (De Smedt et al., 2015; Zhu et al., 2016). Zhu et al (2016) show that differences among these satellite sensors can be as much as a factor of two, posing a challenge for comparing different satellite based HCHO observations. Another uncertainty lies in the reference sector correction, which is usually done by subtracting the retrieved SCD measured over the remote Pacific from the
retrieved terrestrial SCD observed at the same latitude (Khokhar et al., 2005). The corrected differential SCD, which we call the dSCD, represents a HCHO enhancement relative to the Pacific background (Zhu et al., 2016). Several studies have shown systematic biases in satellite HCHO products. Wolfe et al. (2019) finds a small bias in OMI HCHO when comparing to ATom-1 and ATom2 datasets. Using FTIR ground-based measurements, Vigouroux et al. (2020) finds a positive bias of 25% in TROPOMI HCHO vertical column density in regions with low HCHO (<2.5×10^{15} molecules cm^{-2}) and a negative bias of 31% in regions with high HCHO (>8.0×10^{15} molecules cm^{-2}), consistent with a recent comparison between MAX-DOAS and TROPOMI (De Smedt et al., 2021). Zhu et al. (2020) finds a similar bias for OMI HCHO product, with in-situ measurements from aircraft campaigns.

Here we use satellite-based observations of HCHO VCDs from TROPOMI, aircraft measurements, combined with a high-resolution chemical transport model (GEOS-Chem at 0.5° × 0.625° resolution), to better understand the sources and variability of summertime HCHO in Alaska.

2. Observations and Model

2.1. TROPOMI

In this study, we use the TROPOMI operational level 2 (L2) HCHO vertical column density (VCD) product, version 1.1.5-7. The TROPOMI sensor, on board the Sentinel-5 Precursor (S5P) satellite, provides a horizontal resolution of 3.5 km × 7 km from May 2018 to August 2019, 3.5 km × 5.5 km
since August 2019. This product provides a continuous record of reprocessed + offline data
(RPRO+OFFL) since 2018 May. More details can be found in the S5P TROPOMI HCHO L2 product
user manual (Veefkind et al., 2012).

The retrieval algorithm for the S5P TROPOMI HCHO product is based on DOAS technique, following
the OMI QA4ECV product retrieval algorithm (http://www.qa4ecv.eu/ecv/hcho-p/data) detailed in De
Smedt et al. (2018). The HCHO slant column density (SCD_{SAT}) is retrieved in the fitting window of
328.5-359 nm (TROPOMI channel 3). The DOAS reference spectrum is based on the spectra averaged
over tropical Pacific region from previous day (Vigouroux et al., 2020). Therefore, since dSCD_{SAT} is
derived from the difference between local spectra and reference spectrum, it quantifies the slant column
exceeding the average Pacific background. The L2 product provides an air mass factor (AMF_{SAT}) to
convert slant column absorbances of trace gases to vertical column absorbances. AMF_{SAT} is computed
from a radiative transfer model (RTM) VLIDORT v2.6 (Spurr, 2008) and is dependent on observation
geometry, surface albedo, cloud properties, and the vertical distribution of relevant species. The
retrieval uses the 1°×1° monthly averaged surface albedo measured by OMI (Kleipool et al., 2008). A
priori vertical profiles of relevant species are provided by the daily forecast of a chemical transport
model, TM5-MP, at 1° × 1° spatial resolution (Williams et al., 2017).
To correct for possible systematic time- and latitude-dependent offsets, a reference sector correction is applied to calculate the differential slant column, \( \text{dSCD}_{\text{SAT}} \). This correction is based on the background HCHO column over remote oceanic regions, provided by a chemical transport model (TM5-MP). The TROPOMI-measured HCHO differential slant column, \( \text{dSCD}_{\text{SAT}} \) equals the \( \text{SCD}_{1,\text{SAT}} \) minus the reference sector \( \text{SCD}_{\text{Ref, SAT}} \). The reference sector \( \text{SCD}_{\text{Ref, SAT}} \) consists of two parts, an across-track correction (the mean \( \text{SCD}_{1,\text{SAT}} \) in the equatorial reference sector \([-5^\circ, 5^\circ], [180^\circ, 240^\circ])\)) and the zonal along-track correction (a polynomial of all-rows-combined mean \( \text{SCD}_{1,\text{SAT}} \) in 5° latitude bins (only selecting \( \text{SCD}_{1,\text{SAT}} \) that is lower than \( 5 \times 10^{16} \) molecules cm\(^{-2} \)) in the reference sector \([-90^\circ, 90^\circ], [180^\circ, 240^\circ])\).

The resulting differential column, \( \text{dSCD}_{\text{SAT}} \), is then added to the background slant column calculated by the TM5-MP CTM, for the tropospheric vertical column (VCD\(_{\text{SAT}}\)):

\[
\text{VCD}_{\text{SAT}} = \frac{\text{dSCD}_{\text{SAT}}}{\text{AMF}_{\text{SAT}}} + \text{VCD}_{0,\text{SAT}} = \frac{\text{SCD}_{1,\text{SAT}} - \text{SCD}_{\text{Ref, SAT}}}{\text{AMF}_{\text{SAT}}} + \frac{\text{AMF}_{0,\text{SAT}} \text{VCD}_{0,\text{CTM}}}{\text{AMF}_{\text{SAT}}}
\]

Here \( \text{SCD}_{1,\text{SAT}} \) is the measured slant column density, \( \text{SCD}_{\text{Ref, SAT}} \) is the background slant column correction in reference sector. \( \text{AMF}_{\text{SAT}} \) is the air mass factor provided by the TROPOMI HCHO product. \( \text{AMF}_{0,\text{SAT}} \) is the air mass factor for the background column in the reference sector. \( \text{VCD}_{0,\text{CTM}} \) is the vertical column in reference sector calculated by a CTM model (TM5-MP CTM), in the TROPOMI HCHO product. Further details can be referred to De Smedt et al. (2018).
Following S5P TROPOMI HCHO L2 user manual (Veefkind et al., 2012), we applied several criteria to ensure the data quality in this work. This includes: (1) quality assurance values (QA) greater than 0.5; (2) cloud fraction at 340 nm less than 0.5; (3) Solar Zenith Angle (SZA) less than 60°; (4) surface albedo less than 0.1, and (5) derived AMF greater than 0.1. In particular, northern Alaska can be covered by snow and ice even in summer with the criteria of surface albedo. We do not use the data over snow/ice surface as the retrieval algorithm may not work well on these surfaces (De Smedt et al., 2018). We use the overpass data in the local time window 12:00–15:00 AKDT (20:00–23:00 UTC).

To compare the HCHO column density from TROPOMI with our model, we recalculate the AMF based on vertical shapes derived from GEOS-Chem simulations and scattering weight from TROPOMI HCHO product. GEOS-Chem vertical profiles are updated hourly with collocated TROPOMI HCHO pixels. This method has been applied in a number of previous studies (Palmer et al., 2001; Boersma et al., 2004; González Abad et al., 2015; Zhu et al., 2016).

\[
AMF_{GC} = \int_{p_{A,GC}}^{p_{A,GC}} \left( \frac{ \Omega_{GC}(p) }{ \Omega_{A,GC} } \right) w(p) dp
\]

Here \( \Omega_{GC}(p) \) is the column density of the air parcel at vertical air pressure \( p \), for a specific air column. \( \Omega_{A,GC} \) is the total column of the specific air column. \( w(p) \) is scattering weight of TROPOMI HCHO.
product at each altitude, calculated by the product of TROPOMI averaging kernel and air mass factor

\[ \text{AMF}_{\text{SAT}}. P_s \text{ is surface layer pressure.} \]

We further replace the original background HCHO (VCD\textsubscript{0,SAT} in Equation (1)), taken from the TM5-MP model, with VCD\textsubscript{0,GC} from nested GEOS-Chem background simulation (González Abad et al., 2015; Kaiser et al., 2018). The GEOS-Chem background simulation is performed over reference sector, with both biogenic and biomass burning emissions switched off (Table 1). Following De Smedt et al. (2018), we do not consider the variability of the AMF\textsubscript{0,GC}/AMF\textsubscript{GC} (AMF\textsubscript{0,GC} is AMF\textsubscript{GC} in reference sector averaged in 5° latitude bins), and the reprocessed TROPOMI HCHO VCD is expressed as:

\[ VCD_{\text{SAT,GC}} = \frac{\text{dSCD}_{\text{SAT}}}{\text{AMF}_{\text{GC}}} + VCD_{0,\text{GC}} \] (3)

We estimate the total uncertainty of reprocessed TROPOMI HCHO vertical column to be ≥ 90% for fire-free region (TROPOMI L2 HCHO Algorithm Theoretical Basis Document, https://sentinels.copernicus.eu/documents/247904/2476257/Sentinel-5P-ATBD-HCHO-TROPOMI.pdf/7e8af8ac69c24a68bed8891b28261616?_=1646910030856, and references therein).

This includes 75% of uncertainties from the AMF\textsubscript{SAT}, 25% from dSCD\textsubscript{SAT} and 40% from VCD\textsubscript{0,SAT}.

The uncertainties in regions with strong fire are estimated to be ≥ 35%, including 30% of uncertainties.
from AMF\textsubscript{SAT}, 15% from dSCD\textsubscript{SAT} and 10% from VCD\textsubscript{SAT}. The relative lower uncertainties reflect much stronger VCDs in these wildfire regions.

2.2. ATom-1 aircraft campaign

The NASA Atmospheric Tomography (ATom) studied atmospheric composition in remote regions (Wofsy et al., 2018). ATom had four phases over a 2-year period, with each phase sampling the global atmosphere in one of four seasons. ATom deployed a comprehensive gas and aerosol particle measurement payload on the NASA DC-8 aircraft. During ATom-1, two flights performed vertical profiling over Alaska during August 1–3 in 2016. We make use of 1-minute averaged measurements of HCHO, isoprene, monoterpenes (\(\alpha\)-pinene and \(\beta\)-pinene) and the sum of methyl vinyl ketone and methacrolein (MVK+MACR). HCHO measurements sampled in 1-Hz frequency were made by laser-induced fluorescence by the NASA In Situ Airborne Formaldehyde (ISAF) instrument (Cazorla et al., 2015). Isoprene and monoterpenes were measured by two instruments: the University of Irvine Whole Air Sampler (WAS) followed by laboratory Gas Chromatography (GC) analysis, sampled every 3-5 minutes (Simpson et al., 2020), and the National Center for Atmospheric Research (NCAR) Trace Organic Gas Analyzer (TOGA), sampled every 2 minutes with a 35-second integrated sampling time (Apel et al., 2021). MVK and MACR were also measured by TOGA. These measurements are interpolated to 1-minute time resolution for model comparison. Within our study domain, there are 341 1-minute averaged mixing ratio values for HCHO, 101 and 231 for isoprene and \(\alpha\)-pinene/\(\beta\)-pinene.
from WAS, 337 for isoprene, \(\alpha\)-pinene/\(\beta\)-pinene and MVK/MACR from TOGA. The reported measurement uncertainties are \(\pm 10\%\) for HCHO, \(\pm 10\%\) for WAS isoprene and monoterpenes, \(\pm 15\%\) for TOGA isoprene and \(\pm 30\%\) for TOGA monoterpenes, \(\pm 30\%\) for MVK and \(\pm 20\%\) for MACR.

2.3. Nested GEOS-Chem simulation

GEOS-Chem is a 3-D global chemical transport model driven by Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) by the Global Modeling and Assimilation Office (GMAO) at NASA’s Goddard Space Flight Center (Rienecker et al., 2011), at a horizontal resolution of 0.5\(^\circ\) \times 0.625\(^\circ\) and 72 vertical layers from surface to 0.01 hPa. Here we use GEOS-Chem v12.7.2 (http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-Chem_12#12.7.2), with an update on cloud chemistry (https://github.com/geoschem/geos-chem/issues/906). GEOS-Chem v12.7.2 provides a new nested capability, FlexGrid, allowing users to define the model grid at run time (http://wiki.seas.harvard.edu/geos-chem/index.php/FlexGrid). We take advantage of this nested capability to investigate the spatial variability of HCHO and VOCs over Alaska domain (170\(^\circ\)W–130\(^\circ\)W, 50\(^{\circ}\)N–75\(^{\circ}\)N), at a horizontal resolution of 0.5\(^\circ\) \times 0.625\(^\circ\). The boundary conditions for the nested...
Biomass burning emissions follow the Global Fire Emission Database, GFED4.1s biomass burning emissions processed for GEOS-Chem (Giglio et al., 2013). We use 3-hourly emissions calculated in GFED4.1s based on fire detection and burning area from MODIS satellite (van der Werf et al., 2017). The biomass burning emissions in 2018 and 2019 has been updated to reflect the year-specific emissions. The GFED4s inventory reports the HCHO emission factor to be 1.86 g/kg dry matter for boreal forest fires and 2.09 g/kg dry matter for temperate forest fires, consistent with recent field measurements (Liu et al., 2017; Permar et al., 2021).

BVOC emissions in the model are calculated using the Model of Emissions of Gases and Aerosols from Nature (MEGAN, v2.1) (Guenther et al., 2006, 2012). In this work, BVOC emission activity factors are calculated online, expressed as:

$$\gamma = C_e \cdot LAI \cdot \gamma_p \cdot \gamma_T \cdot \gamma_A \cdot \gamma_{SM} \cdot \gamma_{CO_2}$$

Here $C_e$ is a standard environment coefficient normalizing $\gamma$ to 1 under standard environmental condition. LAI is the leaf area index (m$^2$ m$^{-2}$), $\gamma_p$ and $\gamma_T$ are emission activity factors accounting for light and temperature effects, respectively. $\gamma_p$ is calculated based on the photosynthetic photon flux
density (PPFD) (µmol of photons in 400–700 nm range m$^{-2}$ s$^{-1}$). Terrestrial vegetation for BVOC emissions is based on the plant functional type (PFT) distribution derived from Community Land Model (CLM4) (Lawrence et al., 2011; Oleson et al., 2013). CLM4 output suggests two dominating PFTs in the state of Alaska: needle leaf evergreen boreal tree (mainly in the interior boreal forest region) and broadleaf deciduous boreal shrub (mainly over north slope and southwest Alaska), both with high emission factors in isoprene (3000 µg m$^{-2}$ h$^{-1}$ and 4000 µg m$^{-2}$ h$^{-1}$, respectively) and low EFs in monoterpenes ($\alpha$-pinene + $\beta$-pinene, 800 µg m$^{-2}$ h$^{-1}$ and 300 µg m$^{-2}$ h$^{-1}$, respectively). Thus, we expect a major contribution from isoprene to BVOC emissions in Alaska in model results. Despite the shrub having a higher emission factor of isoprene, we expect a larger isoprene emission flux from the central Alaska boreal forest region mainly due to higher temperatures and LAI.

In this work we use the detailed O$_3$-NO$_x$-HO$_x$-VOC chemistry (“tropchem” mechanism) (Park et al., 2004; Mao et al., 2010, 2013), with updates on isoprene chemistry (Fisher et al., 2016). This version of isoprene chemistry in GEOS-Chem have been extensively evaluated by recent field campaigns and satellite observations over southeast US (Fisher et al., 2016; Travis et al., 2016), including HCHO production from isoprene oxidation (Zhu et al., 2016, 2020, Kaiser et al. 2018). To our knowledge, this chemistry has not been evaluated at northern high latitude. In general, under high-NO$_x$ condition (1 ppbv), the HCHO production is prompt, reaching 70-80% of its maximum yield within a few hours. While under low-NOx condition (0.1 ppbv or lower), it takes several days to reach the maximum yield.
and the cumulative yield is still lower than the high-NOx condition by a factor of 2–3 (Marais et al., 2012).

For the comparisons between observations and model shown below, we sample the model output along
the flight track at the flight time with 1-min time resolution. To examine the influence of different
sources on HCHO columns in Alaska, we conducted a series of nested GEOS-Chem simulations, as
described in Table 1. The background HCHO column (VCD_{0,GC}) is calculated from a GEOS-Chem
simulation where both biogenic emissions and biomass burning emissions are turned off. The HCHO
differential column induced by wildfire or biogenic emission is derived from the difference between the
control run and the run with wildfire or biogenic emission turned off. The boundary conditions are
provided by a global simulation with both wildfire and biogenic emission turned on.

**Table 1. Configurations of GEOS-Chem nested simulations in this study.**

<table>
<thead>
<tr>
<th>Simulations</th>
<th>Biogenic emission</th>
<th>Wildfire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (Ctrl)</td>
<td>On</td>
<td>On</td>
</tr>
<tr>
<td>Background (BG)</td>
<td>Off</td>
<td>Off</td>
</tr>
<tr>
<td>No Fire (NF)</td>
<td>On</td>
<td>Off</td>
</tr>
<tr>
<td>No biogenic emission (NB)</td>
<td>Off</td>
<td>On</td>
</tr>
</tbody>
</table>
3. Model evaluation by ATom-1

Figure 1 shows a comparison of measured and modeled vertical profiles of formaldehyde, isoprene, monoterpenes and MVK+MACR across the Alaska domain during ATom-1. We show that the measured HCHO mixing ratio decreases exponentially from <2 km near surface (405 pptv) to the ~10 km upper troposphere (100 pptv). The HCHO surface mixing ratio in Alaska is an order of magnitude lower than other high-BVOC regions such as Southeast US (Li et al., 2016). Observed isoprene and monoterpenes show highest mixing ratios in the lowest 2-km layer. The mean observed isoprene mixing ratio from WAS is ~82 pptv in the boundary layer, a factor of 10 higher than that of monoterpenes from WAS (7 pptv). Measurements from TOGA also shows the predominant role of isoprene but in relatively lower levels near surface (mean value, isoprene: ~28 pptv, monoterpenes: ~2 pptv). We also find that observed sum of MVK and MACR shows enhancement in the boundary layer with a mean mixing ratio of 38 pptv, in line with observed isoprene. As isoprene has a shorter lifetime (~1.1 hours) than monoterpenes (~2.1 hours), this indicates a stronger isoprene emission flux than monoterpane emission flux over Alaskan boreal forests. The predominance of isoprene emissions in Alaskan boreal forests is different from some European boreal forests, where monoterpenes are often the predominant BVOC species (Juráň et al., 2017; Bäck et al., 2012).

To evaluate model performance with ATom-1 measurements, a nested GEOS-Chem simulation is conducted during ATom-1 mission period over Alaska. We sampled the model along the flight track at the flight time with 1-hour model time resolution for comparisons between model and observations. As shown in Figure 1(d) – (f), our nested GEOS-Chem model reproduces the ATom-1 vertical and spatial variability of HCHO, isoprene and monoterpenes mixing ratios. Modeled isoprene and monoterpenes mixing ratios concentrate in the surface layer (0 - 2 km) and show a median value of around 100 ppv and 10 ppv respectively. Modeled isoprene mixing ratios are comparable with ATom-1 observations, while monoterpenes mixing ratio is lower than ATom-1 averaged value in the lowest 2 km (around 40 pptv).
Our model shows reasonable agreement with measurements in the boundary layer (<2 km). Modeled HCHO has a mean mixing ratio of 431 pptv, slightly higher than the observed value (405 pptv).

Modeled isoprene has a mean mixing ratio of 227 pptv in the boundary layer, in agreement with observed values given the large variability of observations. Both observations and model show significantly less monoterpenes compared to isoprene, on the order of tens of pptv. In addition, modeled MVK+MACR shows average mixing ratio of 78 pptv while observations show 38 pptv, providing additional constraints on isoprene oxidation.

Our model tends to underestimate HCHO above boundary layer (>2 km). We show in Figure 1 that mean modeled HCHO is 98 pptv at 3-6 km, and ~46 pptv at 6-10 km, compared to observed values of 212 pptv and 104 pptv respectively. The reason is unknown, but could be related to the large underestimate of CH$_3$OH in the same region (Bates et al., 2021). As a result, the model-derived HCHO VCD is likely lower than that calculated from ATom measurements, by $2.5 \times 10^{15}$ molecules cm$^{-2}$. Such bias may lead to a systematic bias on our estimate of background HCHO VCD$_0$ in this region.

One remarkable feature in Figure 1 is the spatial homogeneity in HCHO vertical profiles, as shown in both observations and model. We find that all sampled HCHO vertical profiles in Alaska show similar magnitude and vertical distribution, despite different land types and locations of these sampled profiles. The homogeneity is not observed in isoprene and monoterpene mixing ratios, which show maximums in
central and south Alaska, where boreal forests are located (Figure S1). Such spatial discrepancies between HCHO and isoprene/monoterpenes suggest a minor contribution of biogenic VOC emissions to HCHO column density over Alaska during summertime.

We further examine the abundance of isoprene and monoterpenes in Alaska with available surface VOC measurements from field campaigns at Toolik Field Station (TFS; 68°38'N, 149°36'W). Angot et al. (2020) reported surface-level ambient mixing ratios of isoprene (0–505 pptv, mean of 36.1 pptv) and monoterpenes (3–537 pptv ± 14±18 pptv, median ± standard deviation) in 2018 and 2019 summers. GEOS-Chem is in reasonable agreement with measurement at TFS, with mean isoprene and monoterpene mixing ratios of 151 pptv and 7 pptv respectively, during corresponding measurement periods. Both field measurements and model suggest that isoprene is the predominant BVOC in this region.
**Figure 1.** Vertical profiles of HCHO, isoprene, monoterpenes and MVK+MACR mixing ratios from ATom-1 and GEOS-Chem, along the ATom-1 flight track from 1 August 2016 20:11 to 3 August 2016 21:20 UTC. GEOS-Chem data are resampled along ATom-1 flight track. (a) to (d) are from ATom-1, (e) to (h) are from GEOS-Chem simulation. The subpanel in (e) shows GEOS-Chem HCHO mixing ratio along the ATom-1 flight track crossing Alaska. For (b)-(d), circles represent values measured by WAS and crosses represent measurements from TOGA. Box plots represent data distribution in < 2 km layers. Orange dashes show the median values, black dashes show the mean values.
4. Evaluating TROPOMI HCHO product

In this section we evaluate the TROPOMI HCHO product over Alaska during the summer of 2018 and 2019. As noted above, these two years differ substantially on local wildfire emissions, providing useful information on satellite capability of detecting biogenic and wildfire HCHO in remote regions.

4.1. Background HCHO VCD in Alaska summer

Figure 2 (a) shows the reprocessed monthly TROPOMI HCHO vertical column density (VCD$_{SAT,GC}$) along with background (VCD$_{0,GC}$) in Alaska during May-August of 2018. Over Alaska domain, HCHO VCD$_{SAT,GC}$ peaks around the interior Alaska boreal forest region (Figure S1), with VCD$_{SAT,GC}$ as $3.7 \times 10^{15}$ molecules cm$^{-2}$ in July; near the north slope and Gulf of Alaska, VCD$_{SAT,GC}$ is around $2 \times 10^{15}$ molecules cm$^{-2}$ in July.

To understand the drivers of HCHO variability, we first examine the background HCHO VCD provided by GEOS-Chem (VCD$_{0,GC}$). Figure 2(b) shows that from May to August 2018, VCD$_{0,GC}$ in central Alaska increases from $1.7 \times 10^{15}$ molecules cm$^{-2}$ to $2.6 \times 10^{15}$ molecules cm$^{-2}$, then decreases to $2.1 \times 10^{15}$ molecules cm$^{-2}$, accounting for 66%–74% of VCD$_{SAT,GC}$. This indicates that VCD$_{SAT,GC}$ is largely dominated by background signals VCD$_{0,GC}$ in 2018. The spatial pattern of VCD$_{0,GC}$, most noticeable in July, is driven by the geography in Alaska, instead of surface vegetation or snow. As the majority of
HCHO VCD resides in lowest atmospheric layers (Figure 1), the high elevation in the Alaska Range in southern Alaska (63°N, 151°W, peaks at Denali, elevation 6190 m) and the Brooks Range in northern Alaska (68°N, 152°W, peaks at Mount Isto, elevation 2736 m) are responsible for the significantly lower HCHO VCD in these regions. We also find that VCD$_{0,GC}$ (2.4–2.8 x 10$^{15}$ molecules cm$^{-2}$) over northern Pacific in July and August is higher than that in central Alaska. Enhanced methane oxidation likely results from the increase of water vapor and therefore OH production, leading to a higher HCHO production via CH$_3$O + NO reactions near surface and CH$_2$O + CH$_3$O at higher altitudes. This enhanced methane oxidation also leads to temperature dependence of VCD$_{0,GC}$ (Figure S2).

![Reprocessed TROPOMI HCHO VCD$_{SAT,GC}$ and background in 2018 summer.](image)

Figure 2. Reprocessed TROPOMI HCHO VCD$_{SAT,GC}$ and background in 2018 summer. (a) reprocessed TROPOMI HCHO VCD$_{SAT,GC}$, (b) HCHO background VCD$_{0,GC}$ used in the reprocessed TROPOMI product, provided by GEOS-Chem. GEOS-Chem results are applied the same local noon.
time window (12:00-15:00) AKDT. TROPOMI data are regridded to GEOS-Chem output spatial resolution (0.5° × 0.625°).

### 4.2. Evaluating TROPOMI HCHO dVCD

Now we further examine measured HCHO signals other than modeled background. Figure 3(a) shows a monthly spatial pattern of TROPOMI differential HCHO vertical column (dVCD_{SAT,GC} = VCD_{SAT,GC}− VCD_{0,GC}), persistent throughout summer 2018. In July 2018, monthly dVCD_{SAT,GC} is positive over central Alaska (1.1×10^{15} molecules cm^{-2}) and north slope (2.0×10^{14} molecules cm^{-2}) and is negative over southwest Alaska and Gulf of Alaska (-5.0×10^{14} molecules cm^{-2}). Negative values reflect the fact that averaged HCHO dVCD_{SAT} is close to zero as a result of reference sector correction (TROPOMI L2 HCHO ATBD). This pattern is also seen in summer 2019 outside wildfire region.

To quantify the sources of HCHO dVCD, we derive two variables: dVCD induced by wildfire emission (dVCD_{GC,Fire}) and biogenic emission (dVCD_{GC,Bio}), computed by the differences between model control run and sensitivity runs with wildfire or biogenic emissions turned off (Table 1).
We show in Figure 3(c) that \( dVCD_{GC,Bio} \) presents a similar spatial pattern and monthly cycle as modeled isoprene emission (Figure S6), with high values over central boreal forest region (\( 4.5 \times 10^{14} \) molecules cm\(^{-2} \)) and low values in other parts (\( 3.5 - 11.0 \times 10^{13} \) molecules cm\(^{-2} \)), in July 2018. The widespread biogenic HCHO enhancement can be in part explained by the slow photooxidation in Alaska and low HCHO yield under low NO\(_x\) conditions (~25-35 pptv near surface in GEOS-Chem) (Marais et al., 2012). Indeed, the HCHO production from isoprene and monoterpenes emissions is lower under low NO\(_x\) conditions than high NO\(_x\) conditions (~1 ppbv) by a factor of 10 after 24-h of oxidation, and it only reaches 20% of its 5-day cumulative yield, leading to a suppressed but prolonged HCHO production (Marais et al., 2012). As a result, \( dVCD_{GC,Bio} \) in Alaska is lower than that in mid-latitude by more than a factor of 10 for the same amount of isoprene emissions.

Despite the relatively weak Alaskan fire in summer 2018, we find a higher fraction of \( dVCD_{GC,Fire} \) than \( dVCD_{GC,Bio} \) in total \( dVCD_{GC} \). Figure 3(b) shows a \( dVCD_{GC,Fire} \) enhancement of \( \sim 1.0 \times 10^{15} \) molecules cm\(^{-2} \) co-located around fire hot spots. The GFED4s burning area measured by MODIS is shown in Figure S5. A model sensitivity test in 2018 suggests that over 90% of \( dVCD_{GC,Fire} \) is from wildfire direct emission, instead of secondary production of HCHO from oxidation of other VOCs. The HCHO emission factor in our model is consistent to recent in-situ measurements in wildfire smoke, providing indirect evaluation on the model HCHO under strong wildfires (Liu et al., 2017; Permar et al., 2021). It can also partly due to the missing of wildfire VOC emissions (Akagi et al., 2011) and the...
underestimation of secondary wildfire VOC oxidation (Liao et al., 2021; Alvarado et al., 2020). The predominance of combustion HCHO in $dVCD_{GC,Fire}$ is consistent with the strong localization of $dVCD_{GC,Fire}$ enhancement, as the HCHO lifetime is on the order of hours in the presence of sunlight. This also explains why weak wildfire emission (46 GgC) can lead to a stronger HCHO dVCD than biogenic emission (281 GgC) does. As a result, $dVCD_{GC,Fire}$ contributes to 20-32% of $dVCD_{GC}$, while $dVCD_{GC,Bio}$ contributes to 8-10% of $dVCD_{GC}$. Wildfire and biogenic emission are both important for $dVCD_{GC}$ and most active in central boreal forest region, posing a challenge to attribute TROPOMI $dVCD_{SAT,GC}$ to individual sources.
Figure 3. dVCD and emission in Alaska in summer 2018. The first-row panels are reprocessed TROPOMI monthly HCHO dVCD_{SAT,GC} in May, June, July and August (unit: molecules cm^{-2}), The second-row panels are GEOS-Chem wildfire emission induced monthly dVCD_{GC,Fire}. The third-row panels are GEOS-Chem biogenic emission induced monthly dVCD_{GC,Bio}. The fourth row is total NMVOC carbon emission from terrestrial vegetation and biomass burning in Alaska in each month (unit: GgC). In (d), blue bars are biogenic isoprene emissions, cyan bars are biogenic monoterpenes.
emissions, purple bars represent wildfire direct HCHO emissions, orange bars represent other NMVOCs emitted by wildfires.

4.3. Impacts of wildfire on HCHO column in Alaska

Figure 4(a) shows monthly VCD$_{\text{SAT,GC}}$ in the 2019 Alaskan summer. In contrast to 2018, TROPOMI observations show an extensive HCHO VCD enhancement over central Alaska in July 2019. The monthly average value reaches $9.4 \times 10^{15}$ molecules cm$^{-2}$. In Figure 5, GEOS-Chem VCD$_{\text{GC}}$ reproduces the spatial and temporal variation of VCD$_{\text{SAT,GC}}$ for the summer of 2019. VCD$_{\text{GC}}$ shows a monthly HCHO VCD value of $1.2 \times 10^{16}$ molecules cm$^{-2}$ in central Alaska for July of 2019, similar to VCD$_{\text{SAT,GC}}$. We expect VCD$_{\text{SAT,GC}}$ to be higher by 20-30% due to a smaller AMF, when strongly absorbing aerosols from wildfire are treated explicitly (Jung et al., 2019). The spatial pattern of VCD$_{\text{GC}}$ enhancements agree well with burned area (Figure S5(b)), indicating that the enhancements of VCD$_{\text{SAT,GC}}$ and VCD$_{\text{GC}}$ in July 2019 are both strongly induced by wildfire sources. In contrast, HCHO VCD outside of the central Alaska are close to the background level, with little enhancement on background HCHO.
Further detailed examination shows the predominance of wildfire emissions on central Alaskan HCHO VCD in summer 2019. We find little change on VCD\(_{\text{GC}}\) and dVCD\(_{\text{GC,Bio}}\) (\(~2 \times 10^{14} \text{ molecules cm}^{-2}\)) from summer 2018 to 2019 in model sensitivity tests, while dVCD\(_{\text{GC,Fire}}\) appears to be solely responsible for the 2018-2019 HCHO VCD difference, especially in July. As a result, dVCD\(_{\text{GC,Fire}}\) accounts for 72% of VCD\(_{\text{GC}}\) in central Alaska, while dVCD\(_{\text{GC,Bio}}\) only takes 6% of VCD\(_{\text{GC}}\) and background oxidation accounts for 22% of VCD\(_{\text{GC}}\). We emphasize that modeled direct emissions of HCHO from wildfires contributes to \(~60%\) of dVCD\(_{\text{GC,Fire}}\). Consequently, dVCD\(_{\text{GC,Fire}}\) is higher than...
dVCD_{GC,Bio} by a factor of 10 in 2019 Alaska summer, despite that NMVOC from wildfires (498 GgC) are only higher than biogenic emissions (389 GgC) by 30%. Due to model sensitivity tests, intercontinental transport of wildfire emissions contributes a minor part of dVCD_{GC,Fire} (~1% in interior Alaska, ~10% in southwest Alaska for 2019 July).

Satellite retrievals of HCHO in wildfire region remains as a major challenge. One source of uncertainty stems from \textit{a priori} profiles used in AMF calculation (Kwon et al., 2017). We find that for regions with heavy smoke, our calculated GEOS-Chem AMF_{GC} is 50% lower than the AMF_{SAT} in the operational product, due to the difference in HCHO vertical profiles (Figure S3). As a result, our reprocessed HCHO VCD product, VCD_{SAT,GC}, is higher than the operational product by 3-5×10^{15} molecules cm^{-2} in heavy smoke regions in July 2019 (Figure S10). Another uncertainty lies in the aerosol optical properties. Wildfire smoke is a major source of brown carbon (June et al., 2020). As the current retrieval algorithm for HCHO does not account for absorbing aerosols, smoke can reduce the sensitivity of satellite measurements to atmospheric layers below and above the aerosol layer, leading to a smaller AMF by 20-30% (Jung et al., 2019; Martin et al., 2003).
Figure 5. HCHO dVCD and emission in Alaska in summer 2019. Similar to Figure 3 but for 2019. The y-axis range of (d) is larger than that in Figure 3.

5. Discussion and Conclusions

The Arctic boreal terrestrial ecosystem is undergoing rapid changes in recent decades, but VOC emissions from Arctic and boreal vegetation and wildfires remains poorly quantified, limiting our
capability for understanding biosphere-atmosphere exchange in this region and its feedback on Arctic climate and air quality. In this work, we use satellite-based observations of HCHO VCD from the TROPOMI instrument on-board S5P satellite, combined with a nested grid chemical transport model, to examine the source and variability of HCHO VCD in Alaska for the summers with low fire activity (2018) and high fire activity (2019).

We first evaluate the GEOS-Chem nested simulation (0.5° × 0.625°) with in-situ airborne measurements in Alaska from the ATom-1 mission. We show reasonable agreement between observed and modeled HCHO, isoprene, monoterpenes and the sum of MVK+MACR in the continental boundary layer. In particular, HCHO profiles show spatial homogeneity in Alaska, suggesting a minor contribution of biogenic emissions to HCHO VCD.

We further compared GEOS-Chem results with TROPOMI HCHO L2 product, reprocessed with background HCHO VCD and AMF using GEOS-Chem model output. The reprocessed product may benefit from the finer horizontal and vertical resolution of GEOS-Chem than TM5-MP model, as well as the year-specific wildfire emissions. We find that reprocessed TROPOMI HCHO VCD\textsubscript{NAT,GC} is dominated by background HCHO VCD\textsubscript{0,GC} from methane oxidation in a mild wildfire summer.
Wildfires have a larger contribution to HCHO total column than biogenic emissions, even in a year with mild wildfires. This result is in part due to the direct emission of HCHO from wildfires, and in part due to the slow and small production of HCHO from isoprene and monoterpenes oxidation under low NOx conditions.

For the year with large wildfires in Alaska (2019), we find that TROPOMI and model show good agreement on magnitude and spatial pattern of HCHO VCD, and wildfire becomes the largest contributor to HCHO VCD. Model sensitivity suggests the direct emission of HCHO from wildfires accounts for the majority of HCHO VCD. While the emission factor of HCHO from wildfires (1.86 g/kg dry matter for boreal forest) applied in our model largely agree with field measurements, the role of secondary production of HCHO is likely underestimated due to unaccounted VOCs and underrepresented plume chemistry. We show that wildfire signals can be detected by TROPOMI HCHO product, making TROPOMI a semi-quantitative tool to constrain wildfire emissions in Alaska given the large uncertainties associated with HCHO retrieval in wildfire plumes. As the Arctic and boreal region continue to warm, we expect HCHO VCD in Alaska continues to be driven by wildfires and background methane oxidation.

Quantifying HCHO at northern high latitudes can be further improved in several aspects. First, we show that background signal, often taken from model output, can be dominant in final product of HCHO.
VCD. However, model results differ significantly on HCHO even over the Pacific Ocean (Figure S8), leading to a large uncertainty in the final satellite product in this region. Second, reference sector correction represents another major uncertainty (Zhu et al., 2020). This is particularly a problem for Alaska, as it lies in the reference sector defined by most retrieval algorithms (González Abad et al., 2015; De Smedt et al., 2018). Any systematic bias in Alaska can propagate to retrievals in other regions. Third, pristine regions can also be influenced by wildfire plumes, which can largely impact HCHO retrieval. Future work is warranted to improve HCHO retrieval and therefore our understanding of HCHO at northern high latitudes.

Acknowledgements. T.Z., J. M. and W.R.S. acknowledge funding from NASA 80NSSC19M0154. This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. We thank Dylan Millet, Xiaoyi Zhao for helpful discussions.

References


