We thank the reviewer for the thorough and thoughtful comments. Our responses to the reviewer are in blue.

Second Review of Zhu et al.

Specific issues:

Page 3 Lines 9-10. You imply that direct satellite validation profiles are not useful, as the noise in the satellite data is high. This is not really true is it? One can still average validation profiles coincident with satellite observations, to mitigate for noise in the retrievals.
We have rewritten this sentence to avoid ambiguity. Please see page 3, line 9–11.

Page 3, lines 28-29: The gridded HCHO errors on the de Smedt et al (2008) correspond to GOME & SCIAMACHY averages, which have sample less frequently than the sensors in this study. Please give the appropriate errors for OMI, GOME-2 (A/B) and OMPS – as asked for originally.
We have listed error associated with each retrieval. Please see page 3, line 28.

Page 3, lines 31-32. Why was data only in the range -0.5 to 10 x1016 molecules/cm2 used?
We have clarified this. Please see page 4, line 1–2.

Page 5, lines 6-7. It would be more accurate to say "Diurnal variability in the HCHO columns over this regions is expected from models to be less than 10%, assuming a correctly simulated diurnal photochemical cycle…" (or similar).
Accepted. We have rewritten this sentence, please see page 5, line 6–7.

Page 5, lines 18-20. Okay, you construct mean HCHO columns to reduce "noise" (here I assume you mean in the retrieval). However, I still think computing the 'aircraft' HCHO columns for each flight and their spatial extent, and comparing that to the satellite data that lie within the aircraft footprint, should be a first step. OMI has good coverage, so you should still get a sufficient number of observations to average. Even if the comparison is 'bad' as you indicate, it would still be useful to show it maybe in the supplementary materials (& if the approach is poor then it of course adds credibility to your chosen method).
This would be too noisy. Averaging data prior to comparison is standard practice in satellite validation studies, in particular for HCHO which is so noisy.

Page 5, lines 20+. The uncertainty in the conversion using the mean HCHO profile has not been quantified. Assuming the mean observed HCHO profile, is a good approach, but it still needs some error associated with it (yes, the error maybe small but it needs to be in the paper). This was asked for in the previous review & should have been included in the revised manuscript, but for some reason the authors have not done it.
Error in the mean HCHO profile is negligible: CAMS has a precision of 40 ppt, then the error in the mean mixing ratios at each pressure bin is tiny (1–5 ppt) by averaging. "We estimate the error in this mean HCHO columns is ~15%, which is mainly from the mixing depths, assumed background, scale height and temperature dependence." We have added this sentence in page 5, line 34–page 6, line1.

Section 4. Again I still think computing consistent AMFs using GEOS-Chem would be a useful exercise and would help to identify and explain the retrieval differences, but it shouldn't limit publication.

Actually, we did compute AMFs using GEOS-Chem profiles, but we realized that working with a single CAMS profile has the advantage of being more transparent. Therefore, we just apply the same observed CAMS mean profile (assumed to be true) to all the six retrievals to identify other possible factors of the retrieval differences. We have included a figure in the Supplemental Materials to show GEOS-Chem based AMFs. We have also added a sentence in the text to that effect, please see page 8, line 14–15.

One final point: There are several places were you have written "HCHO columns serve as a useful proxy isoprene emissions" (or similar). It should be something like "our analysis/results show that HCHO columns may serve as a useful proxy isoprene emissions over this region" – although we already knew this based on the works of Palmer et al and Millet et al.!

We understand the reviewer's point and have modified the text in the several places. We don't want to say "over this region" because there is no reason to presume that our results would not have generality.

Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC4RS aircraft observations over the Southeast US

Lei Zhu[1], Daniel J. Jacob[1,2], Patrick S. Kim[2], Jenny A. Fisher[3,4], Karen Yu[1], Katherine R. Travis[1], Loretta J. Mickley[1], Robert M. Yantosca[1], Melissa P. Sulprizio[1], Isabelle De Smedt[5], Gonzalo Gonzalez Abad[6], Kelly Chance[6], Can Li[7,8], Richard Ferrare[9], Alan Fried[10], Johnathan W. Hair[9], Thomas F. Hanisco[8], Dirk Richter[10], Amy Jo Scarino[11], James Walega[10], Petter Weibring[10], Glenn M. Wolfe[8,12]

[1]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[2]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA
[3]Centre for Atmospheric Chemistry, School of Chemistry, University of Wollongong, Wollongong, NSW, Australia
[4]School of Earth and Environmental Sciences, University of Wollongong, Wollongong, NSW, Australia
[5]Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[6]Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
[7]Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA
[8]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[9]NASA Langley Research Center, Hampton, VA 23681, USA
[10]Institute of Arctic and Alpine Research, Univ. of Colorado, Boulder, CO, USA
[11]Science Systems and Applications, Inc., Hampton, VA, USA
[12]Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, Maryland, USA

Correspondence to: Lei Zhu (leizhu@fas.harvard.edu)

Abstract. Formaldehyde (HCHO) column data from satellites are widely used as a proxy for emissions of volatile organic compounds (VOCs) but validation of the data has been extremely limited. Here we use highly accurate HCHO aircraft observations from the NASA SEAC4RS campaign over the Southeast US in August–September 2013 to validate and intercompare six retrievals of HCHO columns from four different satellite instruments (OMI, GOME2A, GOME2B and OMPS) and three different research groups. The GEOS-Chem chemical transport model is used as a common intercomparison platform. All retrievals feature a HCHO maximum over Arkansas and Louisiana, consistent with the aircraft observations and reflecting high emissions of biogenic isoprene. The retrievals are also interconsistent in their spatial variability over the Southeast US (r=0.4–0.8 on a $0.5^\circ \times 0.5^\circ$ grid) and in their day-to-day variability (r=0.5–0.8). However, all retrievals are biased low in the mean by 20–51%, which would lead to corresponding bias in estimates of isoprene emissions from the satellite data. The smallest bias is for OMI-BIRA, which has high corrected slant columns relative to the other retrievals and low scattering weights in its air mass factor (*AMF*) calculation. OMI-BIRA has systematic error in its assumed vertical HCHO shape profiles for the AMF calculation and correcting this would eliminate its bias relative to the SEAC4RS data. Our results support the use

1

of satellite HCHO data as a quantitative proxy for isoprene emission after correction of the low mean bias. There is no evident pattern in the bias, suggesting that a uniform correction factor may be applied to the data until better understanding is achieved.

1 Introduction

Formaldehyde (HCHO) is a high-yield product from the atmospheric oxidation of volatile organic compounds (VOCs). Methane oxidation largely defines the tropospheric HCHO background. Higher HCHO concentrations over continents are due to short-lived non-methane VOCs (NMVOCs). Loss of HCHO is mainly by photolysis and oxidation by OH, resulting in an atmospheric lifetime of the order of a few hours. HCHO is detectable from space by solar UV backscatter between 325 and 360 nm [Chance et al, 2000]. HCHO column data from satellites have been used in a number of studies as top-down constraints on NMVOC emissions from biogenic, anthropogenic, and open fire sources [Palmer et al 2003; Shim et al 2005; Stavrakou et al 2009; Marais et al 2012; Barkley et al., 2013; Zhu et al., 2014]. However, the satellite data have received little validation so far. Here we validate and intercompare six different HCHO retrievals from four satellites instruments (OMI, GOME2A, GOME2B, OMPS) and three different groups with aircraft observations from the NASA SEAC[4]RS (Studies of Emissions, Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys) campaign over the Southeast US in summer 2013 [Toon et al., 2015].

HCHO columns (molecules cm^{-2}) have been continuously observed from space since GOME (1996–2003; Chance et al. [2000]) and SCIAMACHY (2003–2012; Wittrock et al. [2006]). Observations are presently available from OMI (2004–), GOME2A (2006–), OMPS (2011–), and GOME2B (2012–). The satellite detects a slant column density of HCHO along the atmospheric path of the solar radiation back-scattered to the satellite from the surface and the atmosphere. Conversion to a vertical column is done with an air mass factor (*AMF*) that depends on the satellite viewing geometry, the surface albedo, the vertical HCHO profile, and the vertical distributions of clouds and aerosols [Palmer et al., 2001]. Scattering by air molecules causes the *AMF* to be highly sensitive to the HCHO vertical distribution, which has to be independently specified [Hewson et al., 2015]. The resulting HCHO vertical column retrieved from the satellite includes errors from the slant column fitting and from the *AMF* estimate [Marais et al., 2012].

Validation of HCHO satellite data sets has been extremely limited due to (1) the large noise in individual satellite retrievals, requiring extensive data averaging to enhance detection; and (2) the limited number of HCHO column measurements acquired from aircraft or from the ground. Martin et al. [2004] validated GOME HCHO columns with aircraft observations in eastern Texas averaged over two campaigns (June–July 1999 and August–September 2000), and found GOME to be too high by 16% on average. Comparison of SCIAMACHY data to ground-based measurements of HCHO columns found no significant mean bias [Wittrock et al., 2006; Vigouroux et al., 2009]. Barkley et al. [2013] found that OMI was 37% lower than aircraft measurements made in October 2005 over Guyana. Validation with ground-based remotely sensed vertical profiles indicates a 20–40% underestimate in OMI and GOME2 data [De Smedt et al. 2015].

2

The SEAC[4]RS campaign offers an exceptional opportunity for validating satellite HCHO data. HCHO columns over the Southeast US in summer are among the highest in the world [Kurosu et al., 2004], due to large emissions of biogenic isoprene from vegetation [Guenther et al., 2006]. Several studies have used HCHO data from space as constraints on isoprene emission in the Southeast US [Palmer et al., 2006; Millet et al., 2008; Valin et al., 2016]. The SEAC[4]RS aircraft payload

5 included two independently-calibrated in situ HCHO measurements: the Compact Atmospheric Multispecies Spectrometer (CAMS) [D. Richter et al., 2015] and the NASA GSFC In Situ Airborne Formaldehyde (ISAF) [Cazorla et al., 2015]. CAMS is a mid-IR laser-based spectrometer, which has 1 Hz detection sensitivity of ~40 ppt HCHO [D. Richter et al., 2015]. ISAF uses rotational-state specific laser for detection of HCHO with a 1 Hz detection limit of 36 ppt [Cazorla et al., 2015].

The SEAC[4]RS aircraft did not conduct direct satellite validation profiles, hence we did not apply the direct validation

10 method. Instead we use here an indirect validation method involving joint comparisons of satellite and in situ HCHO observations with the GEOS-Chem chemical transport model (CTM; Bey et al., [2001]). Satellite and in situ observations do not need to be concurrent, thus increasing considerably the range of data and conditions that can be used for validation.

2 Satellite data sets

Table 1 lists the six different satellite retrievals of HCHO produced during the SEAC[4]RS campaign. Additional details

15 on the retrievals are in the Supplementary Material. These are from four satellite instruments (OMI, GOME2A, GOME2B, OMPS) on different platforms, with retrievals produced by independent groups for OMI and OMPS. OMI, flown on the NASA Aura research satellite, has much higher spatial resolution than the other instruments. GOME2A and GOME2B are the first successive instruments of a long-term operational commitment by the EUMETSAT European agency for observing atmospheric composition from space [Callies et al., 2000]. OMPS is the first instrument of a similar long-term operational

20 commitment by NOAA in the US [Dittman, et al., 2002].

All instruments in Table 1 provide dense data sets, with full coverage of the Earth's surface in 1 day for OMI and OMPS, 3 days for GOME2A (since July 2013), and 1.5 days for GOME2B. The single-scene detection limit (0.5–1.0×10^{16} molecules cm^{-2}) is determined by uncertainty in fitting the backscattered solar spectra for SAO retrievals [Gonzalez Abad et al., 2015a; b] and is estimated as the standard deviation of HCHO column amounts over the remote Pacific for BIRA retrievals

25 [De Smedt et al., 2012; 2015]. *AMFs* add another error of 30–100% for single-scene retrievals [Gonzalez Abad et al., 2015a]. OMPS-PCA has a single-scene detection limit of 1.2×10^{16} molecules cm^{-2} estimated as 4-times the standard deviation of HCHO column amounts over the Pacific Ocean [Li et al., 2015]. Uncertainties in HCHO columns can be reduced for monthly means, down to 20–40% for GOME-2A [De Smedt et al., 2008], 38% for OMI, 46% for GOME-2B and ~30% for OMPS, corresponding to 0.1–0.4×10^{16} molecules cm^{-2} over the Southeast US. Here and elsewhere, we use only satellite pixels with

30 solar zenith angle less than 60°, cloud fraction less than 0.3, row anomalies (for OMI) screened, quality check passed (for SAO retrievals), and vertical column density within the -0.5 to 10×10^{16} molecules cm^{-2} range. The latter criterion excludes data that have passed quality tests but are nevertheless suspect as outliers. It excludes 5.8% of the data.

3

All retrievals (except OMPS-PCA) fit the slant column density (SCD) of HCHO from the backscattered solar radiance spectra and then subtract the SCD over the remote Pacific (known as reference sector correction) for the same latitude and observing time to remove offsets [Khokhar et al., 2005]. The resulting corrected SCD ($\Delta\Omega_S$) thus represents a HCHO enhancement over the Pacific background. Additional details on reference sector correction are in the Supplementary Material.

5 $\Delta\Omega_S$ is converted to the HCHO vertical column density (VCD, Ω) by applying an air mass factor (AMF) and a background correction (Ω_o):

$$\Omega = \frac{\Delta\Omega_S}{AMF} + \Omega_o. \tag{1}$$

The background correction, Ω_o, is the HCHO vertical column simulated by a CTM (Table 1) for the remote Pacific at the corresponding latitude and observing time. OMPS-PCA derives the VCD in one step using spectrally varying Jacobians [Li et

10 al., 2015].

The AMF depends on the solar zenith angle (θ_Z) and satellite viewing angle (θ_V), on the scattering properties of the atmosphere and the surface, and on the vertical profile of HCHO concentration. It is computed following Palmer et al. [2001], as the product of a geometrical AMF (AMF_G) describing the viewing geometry in a non-scattering atmosphere, and a correction with scattering weights w applied to the vertical shape factors S:

15 $$AMF_G = \frac{1}{\cos\theta_Z} + \frac{1}{\cos\theta_V}. \tag{2}$$

$$AMF = AMF_G \int_{P_S}^{0} w(p)S(p)dp. \tag{3}$$

Here the integration is over the pressure (p) coordinate from the surface (p_S) to the top of the atmosphere. The shape factor is the normalized vertical profile of mixing ratio: $S(p)=C(p)\Omega_A/\Omega$ where C is the HCHO mixing ratio and Ω_A is the total air column [Palmer et al., 2001]. The scattering weight measures the sensitivity of the backscattered radiation to the presence of

20 HCHO at a given pressure. Impact of aerosols is not explicitly addressed in HCHO retrievals because it is considered to be implicitly included in the cloud correction scheme to the scattering weights [De Smedt et al., 2012; 2015].

All satellite data products (except OMPS-PCA) in Table 1 report for each retrieval Ω, AMF_G, and AMF, and the scattering weights $w(p)$ or equivalent averaging kernels $A(p)=w(p)/AMF$. The BIRA retrievals report in addition the corrected SCD $\Delta\Omega_S$, and background correction Ω_o. To be able to interpret differences between retrievals, we obtained the Ω_o values

25 used by the SAO retrievals and applied Equation (1) to compute their values of $\Delta\Omega_S$. For OMPS-PCA, we computed the AMF based on the reported $S(p)$, $w(p)$ and AMF_G using Equation (3), computed Ω_o based on the reported uncorrected and corrected VCDs, and then obtained $\Delta\Omega_S$ by Equation (1).

3 Aircraft observations and GEOS-Chem model simulation

The SEAC[4]RS DC-8 aircraft flew 21 flights over the Southeast US between 5 August and 25 September 2013,

30 providing extensive mapping of the mixed layer and vertical profiling from the mixed layer to the upper troposphere (Figure

4

1). The mixed layer is defined here as the convectively unstable region of the atmosphere in contact with the ground, as measured from the aircraft by aerosol lidar [Browell et al., 1989; Hair et al., 2008; DIAL-HSRL Mixed Layer Heights README, 2014; Scarino et al., 2014]. It typically extended to 1–3 km altitude (~700–900 hPa) during the afternoon. The mixed layer was often capped by a convective cloud layer of fair-weather cumuli extending to about 3 km, with the free troposphere above [Kim et al., 2015]. 95% of flight hours were between 0930 and 1800 local time for the data in Figure 1, and 78% in the afternoon. Diurnal variability of the HCHO columns is expected from models to be less than 10%, assuming a correctly simulated diurnal photochemical cycle, since photochemistry is both a source and a sink [Millet et al., 2008; Valin et al., 2016].

Figure 2 (left panel) shows a point-to-point comparison of 1-minute averaged ISAF and CAMS HCHO observations (R3 version) aboard the aircraft. There is excellent correlation in the mixed layer (r=0.96) and above (r=0.99). Reduced major axis (RMA) regression of the two data sets yields a slope of 1.10±0.00, with ISAF 10% higher than CAMS due to the fact that the two instruments are independently calibrated. This difference is generally within the mutual stated accuracy for both instruments. The strong correlation between CAMS and ISAF provides confidence that they can be used for satellite validation purposes, and suggests that they can be used as equivalent data sets after 10% bias correction. We use CAMS in what follows.

The aircraft data show high concentrations in the mixed layer due to biogenic isoprene emission, and a sharp drop above the mixed layer because of the short lifetimes of isoprene (~1 h) and of HCHO itself (~2 h). Horizontal variability in the mixed layer reflects the density of isoprene-emitting vegetation but also surface air temperature (affecting isoprene emission) and mixing depth (affecting vertical mixing) at the time of the flights. We wish to convert the data to mean HCHO columns for the SEAC⁴RS period (5 August–25 September) in order to compare to the satellite data averaged over the same period. This requires time-averaging of the local surface air temperature and mixing depth, and conversion of the mixed layer concentration to a total column. We convert the aircraft HCHO mixing ratios in Figure 1 to HCHO columns by assuming uniform HCHO mixing ratios from the surface up through the local mixing depth measured from the aircraft [DIAL-HSRL Mixed Layer Heights README, 2014], an exponential decay from the top of the mixed layer to 650 hPa with a scale height of 1.9 km, and a fixed background of 0.40×10^{16} molecules cm⁻² above, based on the mean vertical profiles information in Figure 1. Day-to-day variability in HCHO columns in the Southeast US is mainly driven by the temperature dependence of isoprene emission and can be fitted well by $\ln\Omega = 0.11T + 2.62$ (r^2=0.64), where Ω is in units of 10^{15} molecules cm⁻² and T [K] is the surface air temperature [Palmer et al., 2006; Zhu et al., 2014]. We applied this temperature dependence to the local HCHO columns inferred from the aircraft mixed layer data converted HCHO columns in order to correct for the difference between the local surface air temperature at the time of the flight and the local mean midday (1200–1300 local time) surface air temperature for the SEAC⁴RS period. Temperatures were taken from the Goddard Earth Observing System-Forward Processing (version 5.11.0, GEOS-FP hereafter) assimilated meteorological data product of the NASA Global Modeling and Assimilation Office (GMAO) [Molod et al., 2012].

The bottom left panel of Figure 1 shows the resulting mean HCHO columns for the SEAC⁴RS period as inferred from the CAMS measurements. We estimate the error in this mean HCHO columns is ~15%, which is mainly from the mixing

5

Deleted: in

Deleted: column

Deleted: %

Deleted: We choose to use CAMS as reference because it is less biased against GEOS-Chem (Figures 1 and 2; see discussion below).

Deleted: to reduce noise.

depths, assumed background, scale height and temperature dependence. The spatial distribution is markedly different and smoother than for the original mixed layer data (top left panel), reflecting in large part the temperature normalization. Figure 3 shows the spatial distribution of midday temperatures for the SEAC[4]RS period, along with base isoprene emissions at 303 K from the MEGAN 2.1 model [Guenther et al., 2012]. The base isoprene emissions feature a hotspot in the Ozarks region of Southeast Missouri, where there is dense oak cover. This region was repeatedly sampled by the aircraft on hot days. The HCHO aircraft observations are particularly high there but this feature is muted after correction for the mean August–September temperatures, which are much cooler in Missouri than further south. Inferred HCHO columns in Figure 1 are instead highest over Arkansas and Louisiana, where August–September temperatures are high.

We simulated the SEAC[4]RS period using the GEOS-Chem v9-02 CTM (http://geos-chem.org) with 0.25°×0.3125° horizontal resolution over North America driven by NASA GEOS-FP assimilated meteorological fields. The model has 47 vertical levels including 18 below 3 km. As can be seen in Figure 4, initial simulations of the SEAC[4]RS data with GEOS-Chem pointed to a positive bias in the daytime GEOS-FP diagnostic for the height of the mixed layer (mixing depth), used in GEOS-Chem for surface-driven vertical mixing. Previous comparisons of GEOS-FP mixing depths to lidar and ceilometer data for other field studies in the Southeast US found a 30–50% high bias [Scarino et al., 2014; Millet et al., 2015]. For the SEAC[4]RS simulation we decreased the GEOS-FP mixing depths by 40%, and comparison to the aircraft lidar measurements along the DC-8 flight tracks shows that this corrects the bias (red line in Figure 4). Corrected afternoon (1200–1700 local time) GEOS-FP mixing depths along the flight tracks in the Southeast US average 1530±330 m, compared to 1690±440 m in the lidar data.

Formaldehyde production in GEOS-Chem over the Southeast US in summer is mainly from isoprene. Companion papers by Fisher et al. [2016], Marais et al. [2016], and Travis et al. [2016] describe the GEOS-Chem simulation of isoprene chemistry in SEAC[4]RS and comparisons to aircraft and surface observations. Biogenic VOC emissions are from the MEGAN 2.1 model as implemented in GEOS-Chem by Hu et al. [2015] and with a 15% decrease applied to isoprene [Wolfe et al., 2015]. Surface-driven vertical mixing up to the mixing depth uses the non-local mixing scheme of [Holtslag and Boville, 1993], as implemented in GEOS-Chem by Lin and McElroy [2010].

Figure 2 (right panel) compares simulated and observed HCHO mixing ratios along the SEAC[4]RS flight tracks, averaged over the GEOS-Chem grid and time step. Comparison for the ensemble of data shows high correlation (r=0.80) and no significant bias. Part of the correlation reflects the dependence on altitude, which is well captured by GEOS-Chem (Figure 1, right panel). After removing this dependence on altitude (by only examining observations within the mixed layer), the correlation between model and observations remains high (r=0.64) with only a small bias (-3±2%) indicated by the RMA linear regression. GEOS-Chem is less successful in reproducing the HCHO concentrations in the free troposphere (3–12 km, ~700–200 hPa), with a -41% normalized mean bias. This may be due to insufficient deep convection in the model.

Integration of the mean vertical profiles in Figure 1 indicates a mean GEOS-Chem HCHO column of 1.46×10^{16} molecules cm^{-2} over the Southeast US during the SEAC[4]RS period, which is 10% lower than observed by CAMS (1.63×10^{16} molecules cm^{-2}), and 23% lower than observed by ISAF (1.90×10^{16} molecules cm^{-2}). The spatial correlation between GEOS-

6

Chem mean HCHO columns (Figure 3, bottom panel) and the HCHO columns inferred from the CAMS data is 0.44 (0.47 for ISAF) on the $0.5^o \times 0.5^o$ grid, with GEOS-Chem capturing the region of maximum HCHO in Arkansas and Louisiana.

4 Intercomparison and validation of satellite data sets over the Southeast US

Figure 5 shows the spatial distribution of mean HCHO columns over the SEAC[4]RS period taken from the six satellite
5   retrievals of Table 1, along with values from GEOS-Chem and columns inferred from the CAMS aircraft observations. All retrievals feature high values over the Southeast US due to isoprene emission and maximum values over and around Arkansas and Louisiana, consistent with GEOS-Chem and CAMS (Figure 3).

Spatial correlation coefficients between HCHO columns for different pairs of satellite retrieval data in Figure 5 are given in Table 2. The correlation coefficients are computed for the temporally averaged (5 August–25 September 2013) data
10   on the $0.5^o \times 0.5^o$ grid of Figure 5 for the Southeast US domain (box in Figure 1 and 5). Correlation coefficients for the different satellite retrievals are only 0.24–0.44 with CAMS but 0.38–0.85 with GEOS-Chem and typically 0.4–0.8 between pairs of retrievals. We conclude that there is consistency between retrievals in the spatial information even at the 0.5^o scale. The GOME2A-BIRA retrieval is noisier than the others and we attribute this to degradation of the instrument after 7 years of operations rather than its reduced swath mode operated since July 2013 [De Smedt et al. 2015], because the noise of GOME2A
15   columns is significantly the same before and after the swath mode reduction.

We see from Figure 5 that all retrievals are biased low relative to CAMS and GEOS-Chem. Table 3 gives statistics for these biases as spatial averages for the Southeast US. GEOS-Chem columns are sampled on the same schedule and scenes as the individual retrievals, and are increased by 10% to correct for the bias with CAMS. Satellite retrieval biases relative to the corrected GEOS-Chem values range from -20% (OMI-BIRA) to -51% (OMPS-PCA). The GOME2A and GOME2B
20   observations are made at 0930 local time, while the OMI and OMPS observations are made at 1330 local time. GEOS-Chem columns increase by 6% from 0930 to 1330 and this is accounted for in the GEOS-Chem comparisons of Table 3.

Retrieval biases in the vertical column Ω could be contributed by the corrected slant column ($\Delta\Omega_S$), the *AMF*, and the background correction Ω_o (Equation (1)). Table 3 gives mean values for these different terms. We see that the OMI-BIRA column is the highest because it has the highest $\Delta\Omega_S$ and lowest *AMF*, while the OMPS-PCA column is the lowest because its
25    $\Delta\Omega_S$ is the lowest. OMPS-SAO and OMPS-PCA use the same OMPS spectra but the OMPS-SAO $\Delta\Omega_S$ are much higher and more consistent with the other retrievals. One caveat is that the derived $\Delta\Omega_S$ of OMPS-PCA may not be the best measure for its algorithm sensitivity, since OMPS-PCA doesn't retrieve a slant column nor does it subtract the Pacific SCD to remove offsets, as described in section 2.

GOME2A-BIRA columns average 18% lower than GOME2B-BIRA despite sharing the same retrieval algorithm and
30   overpass time. This reflects instrument degradation as pointed out above. GOME2A performed much better during its first five years of operation (2007–2011) [De Smedt et al. 2012; 2015].

7

The OMI-BIRA retrieval has the smallest bias relative to the GEOS-Chem and CAMS HCHO columns, and this is due in part to its low *AMF* (0.88). Figure 6 shows the mean reported scattering weights and shape factors for that retrieval (Equation (3)), in comparison to other retrievals and to the CAMS aircraft observations. OMI-BIRA has lower scattering weights than the other retrievals, contributing to the lower *AMF*, and we discuss that below. The shape factors in the SAO (from GEOS-Chem CTM with horizontal resolution of $2^o \times 2.5^o$) and BIRA retrievals (from the IMAGES CTM with horizontal resolution of $2^o \times 2.5^o$) underestimate HCHO in the boundary layer and overestimate it in the free troposphere. With the correct shape factor from CAMS the OMI-BIRA retrieval has an even lower *AMF* (0.74), as shown in Table 3, making it even better in comparison to GEOS-Chem and to the aircraft data. The shape factor from ISAF is consistent with that from CAMS (Figure 1).

Table 3 also gives the *AMFs* for the other retrievals re-computed using CAMS shape factors. The differences with the original *AMFs* are less than 6% except for OMI-BIRA (14%). Although the results for OMI-BIRA illustrate how sensitive the *AMF* calculation is to the specification of shape factor, we find that this is not a significant source of bias in the other retrievals. This may reflect compensating errors in the vertical profile, as illustrated in Figure 6 with the OMI-SAO shape factors in comparison with CAMS. We also compute *AMFs* using GEOS-Chem shape factors for each retrieval (Figure S1 in the Supplemental Materials), but we find this makes no difference to the results.

When the *AMFs* for all retrievals are re-computed with common CAMS shape factors, as shown in Table 3, the remaining differences in *AMFs* are driven by viewing angles (as described by AMF_G in Table 3), scattering weights and cloud parameters. Figure 6 shows that scattering weights are 10–30% higher in the OMI-SAO retrieval (*AMF*=1.02) than in the OMI-BIRA retrieval (*AMF*=0.85). The difference remains for cloud-free satellite pixels (cloud fraction<0.01) and so is not due to different treatments of cloud effects. Surface reflectivity averages 0.048 in OMI-SAO and 0.037 in OMI-BIRA. Although both use the OMI surface reflectance climatology of Kleipool et al. [2008], OMI-SAO applies monthly mean reflectivities while OMI-BIRA applies monthly minimum reflectivities. This can explain some though not all of the difference in scattering weights. De Smedt et al. [2008] found that the HCHO *AMF* increases from 0.4 to 4.0 when the surface albedo changes from 0 to 1.

The background corrections (Ω_o=0.30–0.38×10^{16} molecules cm^{-2}) in the different retrievals are all consistent and amount to about 30% of the mean Ω over the Southeast US. They agree with background HCHO columns measured by aircraft over the remote North Pacific (0.37±0.09×10^{16} molecules cm^{-2}, Table 8 in Singh et al. [2009]).

Previous studies have shown that variability in HCHO columns seen from space over the Southeast US in summer is mainly driven by the temperature dependence of isoprene emission [Palmer et al., 2006; Millet et al., 2008; Duncan et al., 2009; Zhu et al., 2014]. Figure 7 shows time series of daily HCHO columns averaged spatially over the Southeast US for the OMI-SAO and OMI-BIRA retrievals. All retrievals have day-to-day temporal coherence consistent with the temperature dependence of isoprene emission. Temporal correlation between the daily HCHO column and midday temperature is 0.52 for GOME2A-BIRA, 0.59 for OMPS-PCA, 0.59 for OMI-BIRA, 0.69 for GOME2A-BIRA, 0.71 for OMPS-SAO and 0.75 for

8

OMI-SAO. GOME2A-BIRA shows the lowest correlation with temperature, again likely due to noise from instrument degradation.

HCHO over the Southeast US in summer is mainly from oxidation of isoprene [Millet et al., 2006; 2008]. Satellite retrievals validated in this study show consistency in capturing both spatial and daily variations in HCHO columns, as
5   demonstrated by the indirect validation between SEAC⁴RS observations and satellite retrievals. This supports their use as a quantitative proxy for isoprene emissions However, the systematic low bias (20%–51%) in the HCHO retrievals needs to be corrected. Our results show no indication of a pattern in the biases, suggesting that these could be removed as a uniform correction until better understanding is achieved.

5 Conclusions

10   We have used SEAC⁴RS aircraft observations of formaldehyde (HCHO) from two redundant in situ instruments over the Southeast US for 5 August–25 September 2013, together with a GEOS-Chem chemical transport model simulation at $0.25^{o} \times 0.3125^{o}$ horizontal resolution, to validate and intercompare six HCHO retrievals from four different satellite instruments operational during that period. The combination of aircraft data and GEOS-Chem model fields provides strong constraints on the mean HCHO columns and their variability over the Southeast US, where high column amounts are driven by biogenic
15   isoprene emission.

We find that the different retrievals show a large degree of consistency in their simulation of spatial and temporal variability. All retrievals capture the HCHO maximum over Arkansas and Louisiana seen in the aircraft data and in GEOS-Chem, and corresponding to the region of highest isoprene emission. Spatial correlation coefficients between retrievals are moderate to relatively high (0.4–0.8) even on a $0.5^{o} \times 0.5^{o}$ grid. All retrievals are also consistent in their simulation of day-to-
20   day variability correlated with temperature. This supports the use of HCHO columns observed from space as a proxy for isoprene emission. GOME2A-BIRA (launched in 2006) is noisier than other retrievals. We attribute this to instrument degradation.

Despite this success and consistency in observing HCHO variability from space, we find that all satellite retrievals are biased low in the mean, by 20% to 51% depending on the retrieval. This would cause a corresponding bias in estimates of
25   isoprene emission made from the satellite data. The bias is smallest for OMI-BIRA and could be further reduced by correcting the assumed HCHO vertical profiles (shape factors) assumed in the *AMF* calculation. Other retrievals have larger biases that appear to reflect a combination of (1) spectral fitting affecting the corrected slant columns, and (2) scattering weights in the radiative transfer model affecting the *AMF*. Aside from OMI-BIRA, the shape factors used in the retrievals are not a significant source of error in determining the *AMF*.
30   Our work points to the need for improvement in satellite HCHO retrievals to correct the mean low bias. We find no evident spatial or temporal pattern in the bias, at least for the Southeast US in summer, that would compromise the

9

interpretation of the satellite data to estimate patterns of isoprene emission. The biases may be removed by applying uniform correction factors until better understanding is achieved.

Appendix

Abbreviations and acronyms

5	**AMF**	**A**ir **M**ass **F**actor
	BIRA	**B**elgian **I**nstitute for **S**pace **A**eronomy
	CAMS	**C**ompact **A**tmospheric **M**ultispecies **S**pectrometer
	CTM	**C**hemical **T**ransport **M**odel
	GEOS-FP	**G**oddard **E**arth **O**bserving **S**ystem-**F**orward **P**rocessing
10	**GMAO**	**G**lobal **M**odeling and **A**ssimilation **O**ffice
	GMI	**G**lobal **M**odeling **I**nitiative
	GOME2	**G**lobal **O**zone **M**onitoring **E**xperiment-**2**
	IMAGES	**I**ntermediate **M**odel of **G**lobal **E**volution of **S**pecies
	ISAF	**I**n **S**itu **A**irborne **F**ormaldehyde
15	**MEGAN**	**M**odel of **E**missions of **G**ases and **A**erosols from **N**ature
	OMI	**O**zone **M**onitoring **I**nstrument
	OMPS	**O**zone **M**apping and **P**rofiler **S**uite
	PCA	**P**rincipal **C**omponent **A**nalysis
	RMA	**R**educed **M**ajor **A**xis
20	**SAO**	(Harvard) **S**mithsonian **A**strophysical **O**bservatory
	SCD	**S**lant **C**olumn **D**ensity
	SCIAMACHY	**Sc**anning **I**maging **A**bsorption spectro**m**eter for **A**tmospheric **Ch**artography
	SEAC⁴RS	**S**tudies of **E**missions, **A**tmospheric **C**omposition, **C**louds and **C**limate **C**oupling by **R**egional **S**urveys
	VCD	**V**ertical **C**olumn **D**ensity
25	**VOCs**	**V**olatile **O**rganic **C**ompounds

10

11

References

Barkley, M. P., et al.: Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns, J. Geophys. Res. Atmos., 118, 6849–6868, doi:10.1002/jgrd.50552, 2013.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. Field, A. M. Fiore, Q. Li, H. Liu, L. J. Mickley, and M. Schultz, Global
5   modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, J. Geophys. Res., 106, 23,073-23,096, 2001.

Browell, E. V., Differential absorption lidar sensing of ozone, Proc. of the IEEE, 77, 419-432, 1989.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2- Metop's second-generation sensor for operational ozone monitoring, ESA Bull., 102, 28–36, 2000.

10   Cazorla, M., Wolfe, G. M., Bailey, S. A., Swanson, A. K., Arkinson, H. L., and Hanisco, T. F.: A new airborne laser-induced fluorescence instrument for in situ detection of formaldehyde throughout the troposphere and lower stratosphere, Atmos. Meas. Tech., 8, 541-552, doi:10.5194/amt-8-541-2015, 2015.

Chan Miller, C., Gonzalez Abad, G., Wang, H., Liu, X., Kurosu, T., Jacob, D. J., and Chance, K.: Glyoxal retrieval from the Ozone Monitoring Instrument, Atmos. Meas. Tech., 7, 3891-3907, doi:10.5194/amt-7-3891-2014, 2014.

15   Chance, K., Palmer, P. I., Spurr, R. J. D., Martin, R. V., Kurosu, T. P., and Jacob, D. J.: Satellite observations of formaldehyde over North America from GOME, Geophys. Res. Lett., 27, 3461–3464, doi:10.1029/2000GL011857, 2000.

De Smedt, I., Müller, J.-F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors, Atmos. Chem. Phys., 8, 4947-4963, doi:10.5194/acp-8-4947-2008, 2008.

20   De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., Valks, P., Hao, N., and van der A, R.: Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmos. Meas. Tech., 5, 2933-2949, doi:10.5194/amt-5-2933-2012, 2012.

De Smedt, I., Stavrakou, T., Hendrick, F., Danckaert, T., Vlemmix, T., Pinardi, G., Theys, N., Lerot, C., Gielen, C., Vigouroux, C., Hermans, C., Fayt, C., Veefkind, P., Müller, J.-F., and Van Roozendael, M.: Diurnal, seasonal and long-term
25   variations of global formaldehyde columns inferred from combined OMI and GOME-2 observations, Atmos. Chem. Phys., 15, 12519-12545, doi:10.5194/acp-15-12519-2015, 2015.

DIAL-HSRL Mixed Layer Heights README File SEAC4RS 2013: http://www-air.larc.nasa.gov/cgi-bin/ArcView/seac4rs?MLH=1, last access: 2015, 2014.

Dittman, M. G., Ramberg, E., Chrisp, M., Rodriguez, J. V., Sparks, A. L., Zaun, N. H., Hendershot, P., Dixon, T., Philbrick,
30   R. H., and Wasinger, D.: Nadir ultraviolet imaging spectrometer for the NPOESS Ozone Mapping and Profiler Suite (OMPS), in: Proceedings of SPIE, vol. 4814, pp. 111–119, 2002.

Duncan B., Y. Yoshida, M. Damon, A. Douglass, J. Witte: Temperature dependence of factors controlling isoprene emissions, Geophysical Research Letters, 36, p. L05813 http://dx.doi.org/10.1029/2008GL037090, 2009.

12

Fisher, J. A., Jacob, D. J., Travis, K. R., Kim, P. S., Marais, E. A., Chan Miller, C., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Mao, J., Wennberg, P. O., Crounse, J. D., Teng, A. P., Nguyen, T. B., St. Clair, J. M., Cohen, R. C., Romer, P., Nault, B. A., Wooldridge, P. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., Shepson, P. B., Xiong, F., Blake, D. R., Goldstein, A. H., Misztal, P. K., Hanisco, T. F., Wolfe, G. M., Ryerson, T. B., Wisthaler, A., and Mikoviny, T.: Organic nitrate chemistry and its implications for nitrogen budgets in an isoprene- and monoterpene-rich atmosphere: constraints from aircraft (SEAC4RS) and ground-based (SOAS) observations in the Southeast US, Atmos. Chem. Phys., 16, 5969-5991, doi:10.5194/acp-16-5969-2016, 2016.

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., and Suleiman, R.: Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, Atmos. Meas. Tech., 8, 19-32, doi:10.5194/amt-8-19-2015, 2015a.

González Abad, G., Vasilkov, A., Seftor, C., Liu, X., and Chance, K.: Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval, Atmos. Meas. Tech. Discuss., 8, 9209-9240, doi:10.5194/amtd-8-9209-2015, 2015b.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181–3210, 2006.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci. Model Dev., 5, 1471–1492, doi:10.5194/gmd-5-1471-2012, 2012.

Hair, J. W., C. A. Hostetler, A. L. Cook, D. B. Harper, R. A. Ferrare, T. L. Mack, W. Welch, L. R., Izquierdo, F. E. Hovis: Airborne High Spectral Resolution Lidar for Profiling Aerosol Optical Properties, Applied Optics, 47, doi: 10.1364/AO.47.006734, 2008.

Hewson, W., Barkley, M. P., Gonzalez Abad, G., Bösch, H., Kurosu, T., and Spurr, R.: Development and characterisation of a state-of-the-art GOME-2 formaldehyde air-mass factor algorithm, Atmos. Meas. Tech. Discuss., 8, 1109-1150, doi:10.5194/amtd-8-1109-2015, 2015.

Holtslag, A., Boville, B., Local versus nonlocal boundary-layer diffusion in a global climate model, Journal of Climate, 6, 1825, 1993.

Hu, L., Millet, D. B., Baasandorj, M., Griffis, T. J., Turner, P., Helmig, D., Curtis, A. J., and Hueber, J.: Isoprene emissions and impacts over an ecological transition region in the US Upper Midwest inferred from tall tower measurements, J. Geophys. Res.-Atmos., 120, 3553– 3571, doi:10.1002/2014JD022732, 2015.

Kim, P. S., Jacob, D. J., Fisher, J. A., Travis, K., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M. P., Jimenez, J. L., Campuzano-Jost, P., Froyd, K. D., Liao, J., Hair, J. W., Fenn, M. A., Butler, C. F., Wagner, N. L., Gordon, T. D., Welti, A., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Teng, A. P., Millet, D. B., Schwarz, J. P., Markovic, M. Z., and Perring, A. E.: Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite

13

observations with the GEOS-Chem chemical transport model, Atmos. Chem. Phys., 15, 10411-10433, doi:10.5194/acp-15-10411-2015, 2015.

Khokhar, M., Frankenberg, C., Roozendael, M. V., Beirle, S., Kuhl, S., Richter, A., Platt, U., and Wagner, T.: Satellite observations of atmospheric SO2 from volcanic eruptions during the time-period of 1996–2002, Adv. Space Res., 36, 879–887, 2005.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F.m and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res., 113, D18308, doi:10.1029/2008JD010290, 2008.

Kurosu, T. P., Chance, K., and Sioris, C. E.: Preliminary results for HCHO and BrO from the EOS-Aura Ozone Monitoring Instrument, Proc. SPIE, 5652, 116–123, doi:10.1117/12.578606, 2004.

Li, C., J. Joiner, N. A. Krotkov, and L. Dunlap: A newmethod for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, Geophys. Res. Lett., 42, doi:10.1002/2015GL063204, 2015.

Lin, J.-T. and McElroy, M.: Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing, Atmos. Environ., 44, 1726–1739, doi:10.1016/j.atmosenv.2010.02.009, 2010.

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F., and Mao, J.: Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns, Atmos. Chem. Phys., 12, 6219-6235, doi:10.5194/acp-12-6219-2012, 2012.

Marais, E. A., Jacob, D. J., Jimenez, J. L., Campuzano-Jost, P., Day, D. A., Hu, W., Krechmer, J., Zhu, L., Kim, P. S., Miller, C. C., Fisher, J. A., Travis, K., Yu, K., Hanisco, T. F., Wolfe, G. M., Arkinson, H. L., Pye, H. O. T., Froyd, K. D., Liao, J., and McNeill, V. F.: Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: application to the southeast United States and co-benefit of SO2 emission controls, Atmos. Chem. Phys., 16, 1603-1618, doi:10.5194/acp-16-1603-2016, 2016.

Martin, R. V., D. D. Parrish, T. B. Ryerson, D. K. Nicks Jr., K. Chance, T. P. Kurosu, D. J. Jacob, E. D. Sturges, A. Fried, and B. P. Wert: Evaluation of GOME satellite measurements of tropospheric NO2 and HCHO using regional data from aircraft campaigns in the southeastern United States, J. Geophys. Res., 109, D24307, doi:10.1029/2004JD004869, 2004.

Millet, D. B., D. J. Jacob, K. F. Boersma, T. M. Fu, T. P. Kurosu, K. Chance, C. L. Heald, and A. Guenther: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, J. Geophys. Res., 113, D02307, doi:10.1029/2007JD008950, 2008.

Millet, D. B., Baasandorj, M., Farmer, D. K., Thornton, J. A., Baumann, K., Brophy, P., Chaliyakunnel, S., de Gouw, J. A., Graus, M., Hu, L., Koss, A., Lee, B. H., Lopez- Hilfiker, F. D., Neuman, J. A., Paulot, F., Peischl, J., Pollack, I. B., Ryerson, T. B., Warneke, C., Williams, B. J., and Xu, J.: A large and ubiquitous source of atmospheric formic acid, Atmos. Chem. Phys., 15, 6283–6304, doi:10.5194/acp-15-6283-2015, 2015.

Molod A., Takacs L., Suarez M., Bacmeister J., Song I-S, Eichmann A.: The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna, NASA/TM–2012, 104606, vol. 28, 2012.

14

Palmer, P. I., et al.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106, 14,539–14,550, 2001.

Palmer, P. I., D. J. Jacob, A. M. Fiore, R. V. Martin, K. Chance, and T. P. Kurosu: Mapping isoprene emissions over North America using formaldehyde column observations from space, J. Geophys. Res., 108(D6), 4180, doi:10.1029/2002JD002153, 2003.

Palmer, P. I., et al.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, J. Geophys. Res., 111, D12315, doi:10.1029/2005JD006689, 2006.

Richter, A., Weber, M., Burrows, J., Lambert, J.-C., and van Gi- jsel, A.: Validation strategy for satellite observations of tropospheric reactive gases, Ann. Geophys.-Italy, 56, doi:10.4401/ag- 6335, 2013.

Richter D., P. Weibring, J. G. Walega, A. Fried, S. M. Spuler, M. S. Taubman: Compact highly sensitive multi-species airborne mid-IR spectrometer, Appl. Phys. B, doi: 10.1007/s00340-015-6038-8, 2015.

Rodriguez, J. M.: Chapter 6 in atmospheric effects of aviation: First report of the subsonic assessment project, in Global Modeling Initiative, NASA Reference Publication, vol. 1385, edited by A. M. Thompson, R. R. Friedl, and H. L. Wesoky, National Aeronautics and Space Administration, Washington, D. C., 1996.

Scarino, A. J., Obland, M. D., Fast, J. D., Burton, S. P., Ferrare, R. A., Hostetler, C. A., Berg, L. K., Lefer, B., Haman, C., Hair, J. W., Rogers, R. R., Butler, C., Cook, A. L., and Harper, D. B.: Comparison of mixed layer heights from airborne high spectral resolution lidar, ground-based measurements, and the WRFChem model during CalNex and CARES, Atmos. Chem. Phys., 14, 5547–5560, doi:10.5194/acp-14-5547-2014, 2014.

Shim, C., Wang, Y., Choi, Y., Palmer, P. I., Abbot, D. S., and Chance, K.: Constraining global isoprene emissions with Global Ozone Monitoring Experiment (GOME) formaldehyde column measurements, J. Geophys. Res., 110, D24301, doi:10.1029/2004JD005629, 2005.

Singh, H. B., Brune, W. H., Crawford, J. H., Flocke, F., and Jacob, D. J.: Chemistry and transport of pollution over the Gulf of Mexico and the Pacific: spring 2006 INTEX-B campaign overview and first results, Atmos. Chem. Phys., 9, 2301-2318, doi:10.5194/acp-9-2301-2009, 2009.

Stavrakou, T., M¨uller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., and Guenther, A.: Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006, Atmos. Chem. Phys., 9, 3663–3679, doi:10.5194/acp-9-3663-2009, 2009.

Toon, O. B., H. Maring, J. Dibb, R. Ferrare, D. J. Jacob, E. J. Jensen, Z. J. Luo, G. G. Mace, L. L. Pan, L. Pfister, K. H. Rosenlof, J.s Redemann, J. S. Reid, H. B. Singh, R. Yokelson, G. Chen, K. W. Jucks, and A. Pszenny: Planning, implementation and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) field mission, J. Geophys. Res., under review, 2016.

Travis, K. R., Jacob, D. J., Fisher, J. A., Kim, P. S., Marais, E. A., Zhu, L., Yu, K., Miller, C. C., Yantosca, R. M., Sulprizio, M. P., Thompson, A. M., Wennberg, P. O., Crounse, J. D., St. Clair, J. M., Cohen, R. C., Laugher, J. L., Dibb, J. E., Hall, S. R., Ullmann, K., Wolfe, G. M., Pollack, I. B., Peischl, J., Neuman, J. A., and Zhou, X.: NOx emissions, isoprene

15

oxidation pathways, vertical mixing, and implications for surface ozone in the Southeast United States, Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-110, in review, 2016.

Valin, L. C., A. M. Fiore, K. Chance, and G. González Abad: The role of OH production in interpreting the variability of CH2O columns in the southeast U.S., J. Geophys. Res. Atmos., 121, doi:10.1002/2015JD024012, 2016.

5 Vigouroux, C., Hendrick, F., Stavrakou, T., Dils, B., De Smedt, I., Hermans, C., Merlaud, A., Scolas, F., Senten, C., Vanhaelewyn, G., Fally, S., Carleer, M., Metzger, J.-M., Müller, J.-F., Van Roozendael, M., and De Mazière, M.: Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data, Atmos. Chem. Phys., 9, 9523-9544, doi:10.5194/acp-9-9523-2009, 2009.

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U.,
10 and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space, Geophys. Res. Lett., 33, L16804, doi:10.1029/2006GL026310, 2006.

Wolfe, G. M., et al.: Quantifying sources and sinks of reactive gases in the lower atmosphere using airborne flux observations, Geophys. Res. Lett., 42, 8231–8240, doi:10.1002/2015GL065839, 2015.

Zhu, L., D. J. Jacob, L. J. Mickley, E. A. Marais, D. S. Cohan, Y. Yoshida, B. N. Duncan, G. González Abad, and K. V.
15 Chance, Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, Environ. Res. Lett., 9, 114004, doi:10.1088/1748-9326/9/11/114004, 2014.
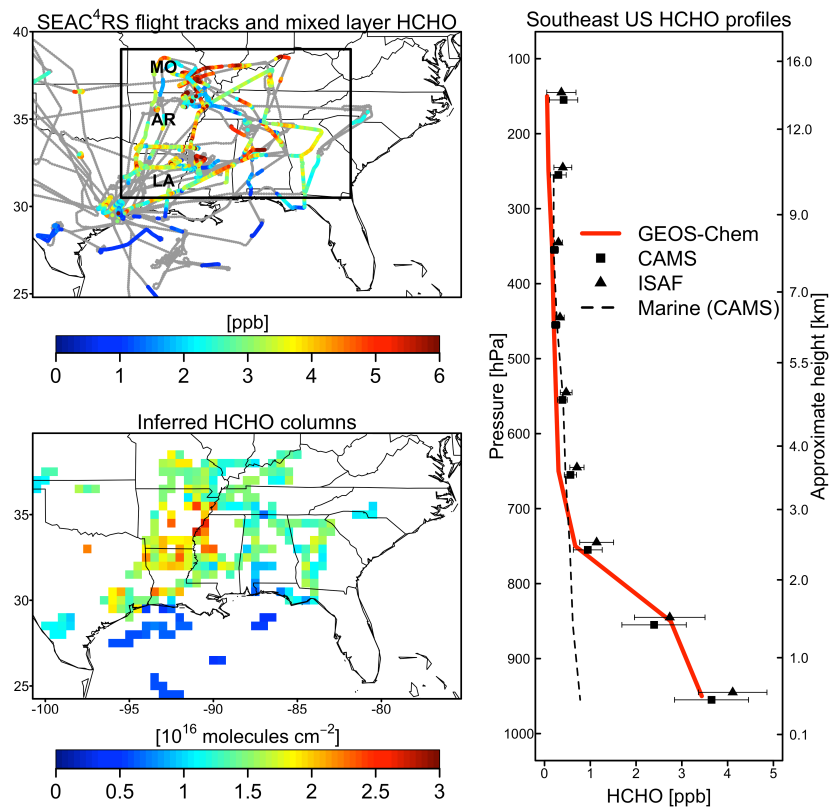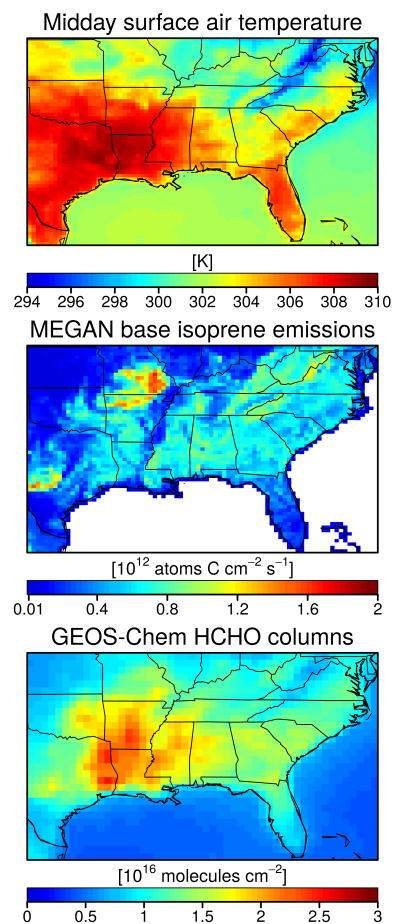
16

Figure 1: Formaldehyde (HCHO) concentrations along SEAC[4]RS aircraft flight tracks (5 August–25 September 2013). The top left panel shows the DC-8 flight tracks (in grey) and the CAMS measurements aboard the aircraft in the mixed layer. The mixed layer is the convectively unstable region of the atmosphere in contact with the surface, diagnosed locally from aerosol lidar observations aboard the aircraft [DIAL-HSRL Mixed Layer Heights README, 2014] and typically extending to 1–3 km altitude. The states of Missouri (MO), Arkansas (AR), and Louisiana (LA) are indicated. The right panel shows the mean vertical profiles observed by the CAMS and ISAF instruments, and simulated by GEOS-Chem, for the Southeast US domain (30.5^o–39^oN, 95°–81.5^oW) defined by the black rectangle in the top left panel. Horizontal bars represent observed standard deviations. GEOS-Chem is sampled along the flight tracks at the time of the measurements. The dashed black line shows the mean vertical CAMS profile in marine air over the Gulf of Mexico (22°–28^oN, 96.5^o–88.5^oW), which is used in determining background HCHO column (0.40×10^{16} molecules cm^{-2}, see Section 3). The bottom left panel shows the mean HCHO columns on a $0.5^o\times0.5^o$ grid derived from the CAMS measurements after normalizing for temperature, for mixed layer depth, and for the contribution from HCHO aloft (see text in Section 3).

17
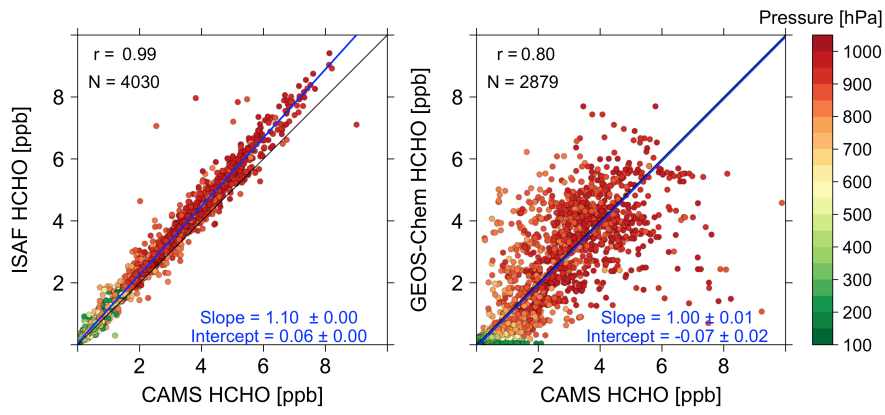
Figure 2. Comparisons between HCHO measurements from the CAMS and ISAF instruments aboard the SEAC[4]RS aircraft, and simulated by GEOS-Chem, for the Southeast US flight tracks (box in Figure 1). The left panel compares 1-minute measurements from CAMS and ISAF. The right panel compares GEOS-Chem and CAMS HCHO. Here and elsewhere for model-observation comparisons, HCHO observations along the flight tracks are averaged onto the GEOS-Chem grids (0.25°×0.3125°, 47 vertical layers) and time steps (10 minutes). HCHO data points are colored by atmospheric pressure. Slopes and intercepts of reduced major axis (RMA) regressions are shown along with the correlation coefficient (*r*), sample size (*N*), RMA regression line (in blue), and 1:1 line.

Midday surface air temperature
[K]
294 296 298 300 302 304 306 308 310

MEGAN base isoprene emissions
[10^{12} atoms C cm^{-2} s^{-1}]
0.01 0.4 0.8 1.2 1.6 2

GEOS-Chem HCHO columns
[10^{16} molecules cm^{-2}]
0 0.5 1 1.5 2 2.5 3

Figure 3. Mean temperature, base isoprene emissions, and HCHO columns in the GEOS-Chem model for the SEAC[4]RS period (5 August–25 September 2013). The top panel shows the midday (1200–1300 local time) surface air temperature from the GEOS-FP assimilated meteorological data. The middle panel shows the MEGAN 2.1 base isoprene emissions from Guenther et al. [2012] for standard conditions (air temperature=303 K, photosynthetic photon flux density=200 µmol m^{-2} s^{-1} for sunlit leaves and 50 µmol m^{-2} s^{-1} for shaded leaves.) The bottom panel shows the GEOS-Chem HCHO columns computed with MEGAN 2.1 isoprene emissions and sampled at 1330 local time, under OMI-SAO schedule, and filtered by OMI-SAO quality flags and cloud conditions.
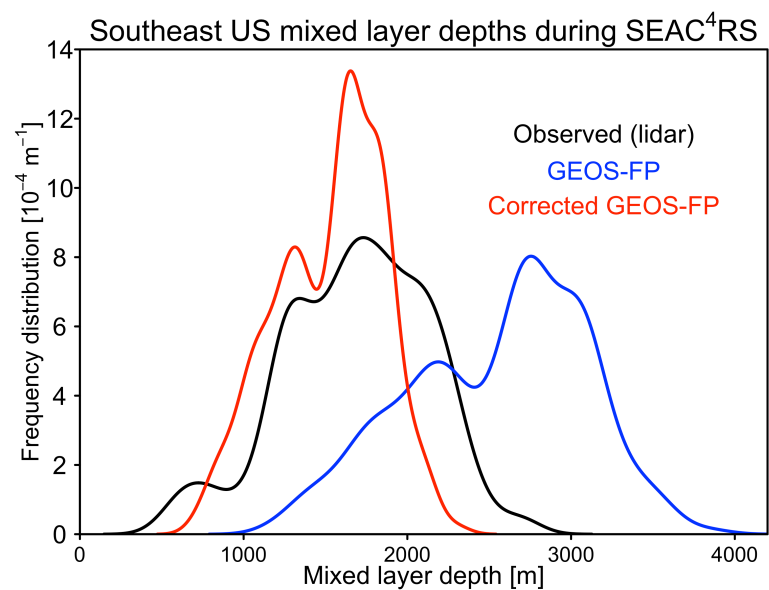
19

Figure 4. Frequency distribution of mixed layer depths over the Southeast US during the SEAC[4]RS period (5 August–25 September 2013). Observations by aerosol lidar aboard the aircraft [DIAL-HSRL Mixed Layer Heights README, 2014] are compared to the local GEOS-FP data used to drive GEOS-Chem, before and after the 40% downward correction. The frequency distributions are constructed from 1-minute average data along the aircraft flight tracks over the Southeast US (box in Figure 1) for the 1200–1700 local time window.

20
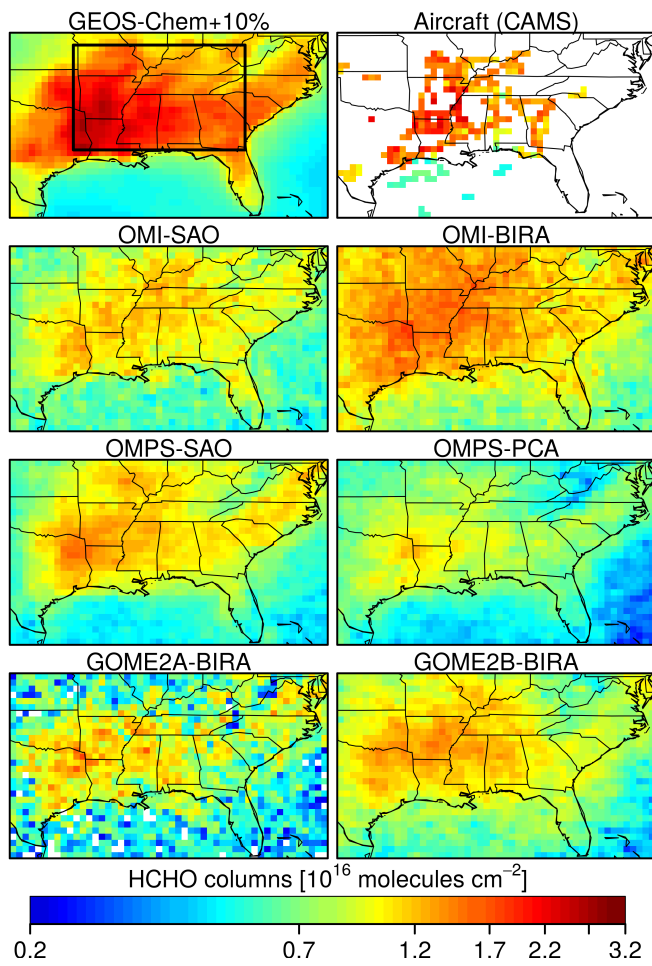
Figure 5. HCHO vertical column densities over the Southeast US averaged over the SEAC⁴RS period (5 August–25 September 2013). The bottom panels show six retrievals from four satellites (OMI, GOME2A, GOME2B and OMPS) and three different groups (Table 1). The top panels show (1) GEOS-Chem model results sampled on the OMI schedule with filtering by OMI-SAO quality flags and cloud conditions, and increased by 10% to correct for the bias relative to CAMS aircraft measurements; and (2) columns derived from the CAMS aircraft measurements (same as bottom left panel of Figure 1 but on a different color scale). The black rectangle represents the Southeast US domain (same as in Figure 1). Colorbar is a logarithmic scale.

21
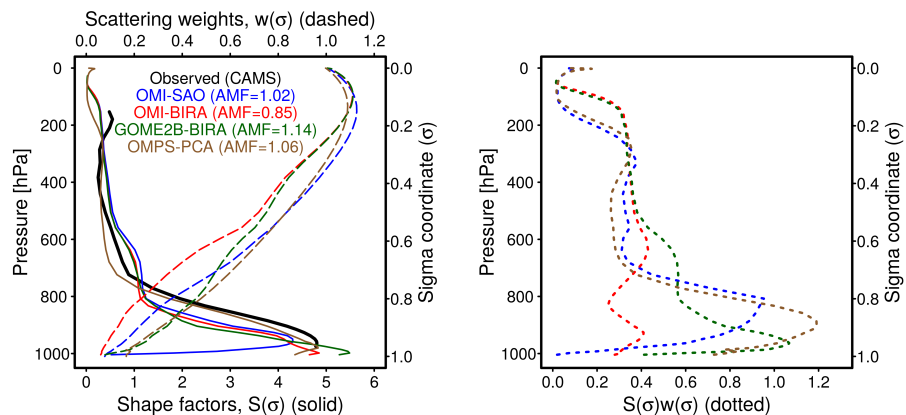
Figure 6. Air mass factor differences between retrievals. The left panel shows mean scattering weights (w) and shape factors (S) for HCHO retrievals over the Southeast US during the SEAC[4]RS period, and the right panel shows the product of the two from which the *AMF* is derived by vertical integration using Equation (3). Values are shown for the OMI-SAO, OMI-BIRA, GOME2B-BIRA, and OMPS-PCA retrievals. Mean *AMF* values are given in legend. Also shown is the observed HCHO shape factor (black) from the mean CAMS profile in Figure 1.

22
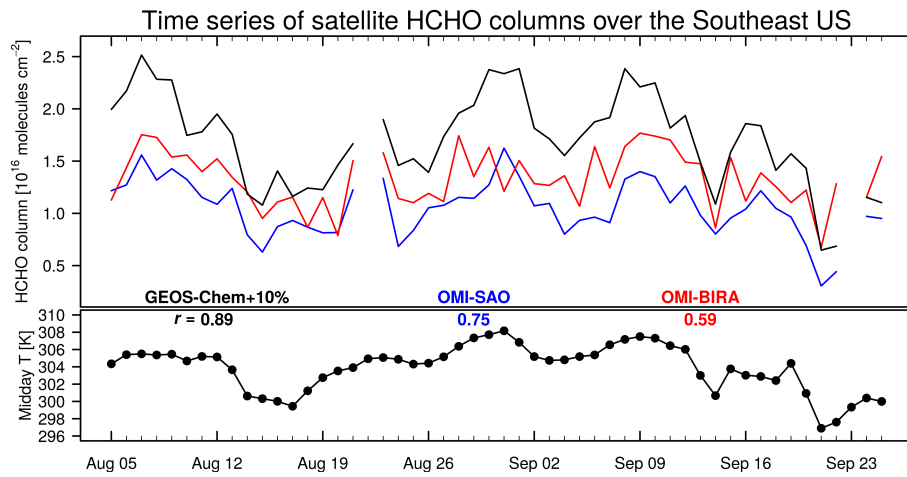
Figure 7. Daily variability of HCHO vertical column densities over the Southeast US during SEAC[4]RS. The top panel shows daily HCHO columns averaged over the Southeast US (box in Figure 5) for the OMI-SAO and OMI-BIRA retrievals. GEOS-Chem columns (black) are sampled following the OMI viewing geometry with filtering by OMI-SAO quality flags and cloud conditions, and scaled up by 10% on the basis of comparison with CAMS aircraft columns. The bottom panel shows the local midday (1200–1300 local time) surface air temperature over the Southeast US domain from the GEOS-FP assimilated meteorological data. Also shown for each data set is the temporal correlation coefficient (*r*) with temperature.

Table 1. Satellite HCHO products validated and intercompared in this work[a]

Retrieval	Nadir resolution [km²]	Local viewing time	Fitting windows [nm]	Chemical transport model[b]	Detection limit [10^{16} molecules cm^{-2}]	Reference[c]
OMI-SAO (V003)	24×13	1330	328.5–356.5	GEOS-Chem v09-01-03	1.0	[1]
OMI-BIRA (V14)	24×13	1330	328.5–346.0	IMAGES v2	0.7	[2]
GOME2A-BIRA (V14)	40×40	0930	328.5–346.0	IMAGES v2	0.8	[3]
GOME2B-BIRA (V14)	80×40	0930	328.5–346.0	IMAGES v2	0.5	[2]
OMPS-SAO	50×50	1330	327.7–356.0	GEOS-Chem v09-01-03	0.75	[4]
OMPS-PCA	50×50	1330	328.5–356.5	GMI	1.2	[5]

[a] Retrievals operational during the SEAC⁴RS aircraft campaign (5 August–25 September 2013). These include four different sensors (OMI, GOME2A, GOME2B and OMPS), flown on different platforms, with different retrievals for OMI and OMPS produced by the Harvard Smithsonian Astrophysical Observatory (SAO), the Belgian Institute for Space Aeronomy (BIRA), and the NASA Goddard Space Flight Center by Principal Component Analysis (PCA). Further retrieval details are in the Supplementary Material.

[b] Chemical transport model (CTM) supplying the normalized mixing ratio vertical profiles (shape factors) and background correction (Ω_o see section 2) used in the retrieval. References are ChanMiller et al. [2014] for GEOS-Chem v09-01-03, Stavrakou et al. [2009] for IMAGES v2, and Rodriguez [1996] for GMI.

[c] [1] Gonzalez Abad et al. [2015a]; [2] De Smedt et al. [2015]; [3] De Smedt et al. [2012]; [4] Gonzalez Abad et al. [2015b]; [5] Li et al. [2015]. OMI-SAO data were downloaded from: http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/OMI/omhcho_v003.shtml. GOME2A-BIRA and GOME2B-BIRA data were downloaded from: http://h2co.aeronomie.be. Other data were courtesy of the retrieval group.

24

Table 2. Spatial/temporal correlation coefficients (r) between pairs of HCHO column products[a]

HCHO product	OMI-SAO (V003)	OMI-BIRA	GOME2A-BIRA (V14)	GOME2B-BIRA (V14)	OMPS-SAO	OMPS-PCA
OMI-SAO (V003)	1/1					
OMI-BIRA	0.55/0.67	1/1				
GOME2A-BIRA (V14)	0.28/0.48	0.38/0.50	1/1			
GOME2B-BIRA (V14)	0.50/0.76	0.65/0.60	0.49/0.26	1/1		
OMPS-SAO	0.48/0.77	0.70/0.50	0.45/0.55	0.72/0.76	1/1	
OMPS-PCA	0.40/0.70	0.60/0.51	0.53/0.63	0.71/0.68	0.85/0.84	1/1
GEOS-Chem[b]	0.38/0.88	0.50/0.65	0.68/0.82	0.85/0.88	0.74/0.86	0.82/0.75
Aircraft (CAMS)[c]	0.24/-	0.44/-	0.26/-	0.35/-	0.43/-	0.37/-

[a] Correlation coefficients between HCHO columns for different pairs of satellite retrievals, GEOS-Chem, and CAMS aircraft observations. Values are for the Southeast US domain (box in Figure 1 and 5) during SEAC[4]RS (5 August–25 September 2013). Spatial correlation coefficients are computed for the temporally averaged data on the 0.5°×0.5° grid of Figure 5. Temporal correlation coefficients are computed from daily averages of each retrieval over the Southeast US domain.

[b] GEOS-Chem CTM columns sampled for the same scenes as the individual retrievals.

[c] Aircraft column data are temporal averages for the SEAC[4]RS period as shown in Figure 1 (bottom left panel) and Figure 5 (top right panel).

Table 3. Satellite retrievals of HCHO columns over the Southeast US[a]

Retrieval	Mean values[b]						with CAMS shape factors			GEOS-Chem+10%[c]
	Ω	$\Delta\Omega_S$	AMF_G	AMF	Ω_o	Bias[d]	AMF^e	Ω^f	Bias[d]	Ω
OMI-SAO (V003)	1.06	0.65	2.66	0.95	0.38	-37%	1.01	0.96	-43%	1.69
OMI-BIRA	1.33	0.87	2.62	0.88	0.31	-20%	0.74	1.47	-12%	1.67
GOME2A-BIRA (V14)	0.89	0.62	2.37	1.12	0.30	-44%	1.14	0.84	-47%	1.59
GOME2B-BIRA (V14)	1.09	0.86	2.56	1.22	0.30	-34%	1.27	0.98	-41%	1.65
OMPS-SAO	1.09	0.72	2.54	1.01	0.38	-34%	1.02	1.01	-39%	1.66
OMPS-PCA	0.80	0.49	2.53	1.11	0.35	-51%	1.15	0.78	-52%	1.63

[a] Mean values over the Southeast US domain (box in Figure 1 and 5) for the data in Figure 5 collected during the SEAC[4]RS period (5 August–25 September, 2013).

[b] Mean values provided as part of the retrieval product including vertical HCHO columns (Ω), corrected slant columns ($\Delta\Omega_S$), geometrical and scatter-corrected AMFs, and background correction (Ω_o), following Equation (1). Columns are in units of 10^{16} molecules cm^{-2} and AMFs are dimensionless. The corrected slant columns and background correction are not reported in the SAO and OMPS-PCA retrievals and are reconstructed here to enable comparison with the other retrievals (see Section 2).

[c] GEOS-Chem columns sampled for the same scenes as the individual retrievals and increased by 10% to correct for the bias relative to the SEAC[4]RS CAMS aircraft measurements (Figure 1). Mean GEOS-Chem columns increase with time of day by 6.0% from 0930 local time (GOME2A and GOME2B) to 1330 (OMI and OMPS).

[d] Normalized mean bias relative to the corrected GEOS-Chem values (last column in the table).

[e] AMFs recalculated using the mean HCHO vertical shape factor from the CAMS aircraft instrument (Figure 1 and 6) and the scattering weights or averaging kernels provided as part of the satellite product (Figure 6).

[f] Columns recomputed using AMFs constrained by the CAMS aircraft measurements.
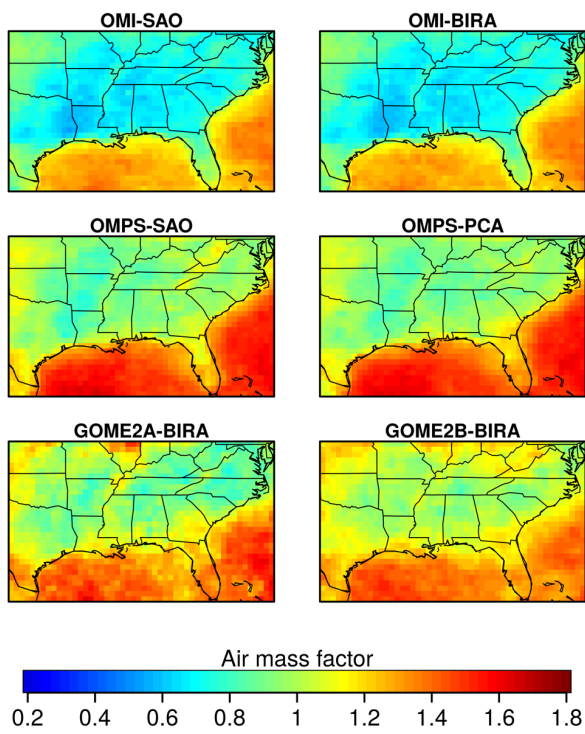
26

Table S1 Summary of different HCHO retrievals validated and intercompared in this work

	OMI-SAO (V003)	OMI-BIRA	GOME2A-BIRA (V14)	GOME2B-BIRA (V14)	OMPS-SAO	OMPS-PCA
Fitting window (nm)	328.5–356.5	328.5–346.0	328.5–346.0	328.5–346.0	327.7–356.0	328.5–356.5
Spectral resolution (nm, at 340 nm)	0.42	0.42	0.27	0.28	~1	~1
Fitted species	O_3, NO_2, BrO, O_2-O_2, Ring	O_3, NO_2, BrO, O_2-O_2, Ring	O_3, NO_2, BrO, O_2-O_2, Ring	O_3, NO_2, BrO, O_2-O_2, Ring	O_3, NO_2, BrO, O_2-O_2, Ring	BrO (O_3, Ring, NO_2 features are from principal components)
HCHO cross-section	Chance and Orphal [2011], 300K	Meller and Moortgat [2000], 298K	Meller and Moortgat [2000], 298K	Meller and Moortgat [2000], 298K	Chance and Orphal [2011], 300K	-
Closure polynomial	3rd order	5th order	5th order	5th order	3rd order	-
Underdamping correction	Yes	No	No	No	Yes	-
Additional fit parameters	Residual common mode spectrum	Non-linear O_3 absorption effect	Non-linear O_3 absorption effect; Polarisation cross-section	Non-linear O_3 absorption effect; Polarisation cross-section	Residual common mode spectrum	Principal components derived from radiances
Cloud product used	OMI O4 cloud product, Acarreta et al. [2004]	OMI O4 cloud product, Stammes et al. [2008].	GOME-2 O2 A-band Frescov6 product, Wang et al. [2008]	GOME-2 O2 A-band Frescov6 product, Wang et al. [2008]	OMPS cloud product, Vasilkov et al. [2014]	OMPS cloud product, Vasilkov et al. [2014]
Vertical layers of the CTM	47	40	40	40	47	72
Horizontal resolutions of the CTM	2°×2.5°	2°×2.5°	2°×2.5°	2°×2.5°	2°×2.5°	2°×2.5°
A prior HCHO profiles	1300 to 1400 Monthly average in 2007	Sampled at the satellite overpass times	Sampled at the satellite overpass times	Sampled at the satellite overpass times	1300 to 1400 Monthly average in 2007	Monthly average, 2005–2007

27

Resolution of scattering weight grids	2°×2.5°	Interpolated for each satellite pixel (surface albedo, surface altitude, observation angles)	Interpolated for each satellite pixel (surface albedo, surface altitude, observation angles)	Interpolated for each satellite pixel (surface albedo, surface altitude, observation angles)	2°×2.5°	2°×2.5°
Terrain corrections	Yes	Yes, based on Zhou et al. [2009]	Yes	Yes	Yes	Yes
Treatment of aerosols	No	No	No	No	No	No
Reference sector correction[a]	Pixel-by-pixel correction	Modelled by a polynomial	Modelled by a polynomial	Modelled by a polynomial	Pixel-by-pixel correction	Model climatology

[a] SAO retrievals applied the "pixel-by-pixel correction" method, which accounts for the difference between retrieved and GEOS-Chem modelled slant columns over the remote Pacific Ocean. For BIRA retrievals, the reference sector correction refers to the latitudinal dependency (modelled by a polynomial) of the HCHO slant columns in the reference sector. For OMPS PCA retrievals, the latitude-dependent GMI monthly climatology HCHO column amount over the remote Pacific is added to all pixels, without pixel-specific correction.

5

OMI-SAO **OMI-BIRA**

OMPS-SAO **OMPS-PCA**

GOME2A-BIRA **GOME2B-BIRA**

Air mass factor

0.2   0.4   0.6   0.8   1   1.2   1.4   1.6   1.8

Figure S1. GEOS-Chem based air mass factors (AMF) over the Southeast US averaged over the SEAC[4]RS period (5 August–25 September 2013). Each panel shows the AMF computed using GEOS-Chem HCHO shape factors with scattering weights from the six retrievals (Table 1). For each retrieval, GEOS-Chem shape factors are sampled under its schedule, and filtered by its quality flags and cloud conditions.

5

29

References

Acarreta, J. R., De Haan, J. F., and Stammes, P.: Cloud pressure retrieval using the O2–O2 absorption band at 477 nm, J. Geophys. Res.-Atmos., 109, D05204, doi:10.1029/2003JD003915, 2004.

Chance, K. and Orphal, J.: Revised ultraviolet absorption cross sections of H2CO for the HITRAN database, J. Quant. Spectrosc. Ra., 112, 1509–1510, doi:10.1016/j.jqsrt.2011.02.002, 2011.

Meller, R., and Moortgat, G. K.: Temperature dependence of the absorption cross section of HCHO between 223 and 323K in the wavelength range 225–375 nm, J. Geophys. Res., 105, 7089–7102, doi:10.1029/1999JD901074, 2000.

Stammes, P., Sneep, M., de Haan, J. F., Veefkind, J. P., Wang, P., and Levelt, P. F.: Effective cloud fractions from the Ozone Monitoring Instrument: Theoretical framework and validation, J. Geophys. Res., 113, D16S38, doi:10.1029/2007JD008820, 2008.

Vasilkov, A., Joiner, J., and Seftor, C.: First results from a rotational Raman scattering cloud algorithm applied to the Suomi National Polar-orbiting Partnership (NPP) Ozone Mapping and Profiler Suite (OMPS) Nadir Mapper, Atmos. Meas. Tech., 7, 2897–2906, doi:10.5194/amt-7-2897-2014, 2014. 9217, 9220.

Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M.: FRESCO+: an improved O2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals, Atmos. Chem. Phys., 8, 6565–6576, doi:10.5194/acp-8-6565-2008, 2008.

Zhou, Y., Brunner, D., Boersma, K. F., Dirksen, R., and Wang, P.: An improved tropospheric NO2 retrieval for OMI observations in the vicinity of mountainous terrain, Atmos. Meas. Tech., 2, 401-416, doi:10.5194/amt-2-401-2009, 2009.