We thank the three reviewers for their thorough and thoughtful comments. Our responses to reviewers are in blue.

Anonymous Referee #1
Received and published: 5 April 2016

This is a long over due paper concerning the validation of satellite HCHO columns, retrieved by various sensors/groups data against in-situ aircraft measurements. It addresses an important topic and should be published when the following issues have been addressed. The paper is generally well written with nice graphics (apart from fig 7).

Although the analysis is decent, and it gives the reader an impression of the quality of the satellite HCHO column data, I do, however, feel that a somewhat opportunity has been missed since the aircraft data could have been used to really pick the retrievals apart. Maybe this is planned for the future, one can only hope!

Specific Comments:

Page 2, Line 25 & subsequent paragraph: Please note that Barkley et al (2013) compared OMI data to GABRIEL aircraft HCHO measurements and found OMI was about 40% too low.
We now cited Barkley et al [2013] in page 2, line 30–31.

Table 1 – I would include much more information about the details of the retrieval and AMF algorithms, e.g., fitting windows, absorbers fitted, use of radiance references, CTM vertical and horizontal resolutions, profile (monthly vs daily), resolution of scattering weight grids, treatment of clouds, terrain corrections, treatment of aerosols etc.
We now added Table S1 in the Supplementary Materials to summarize more retrieval details. We also referred to Table S1 in the text (page 3, line 15–16) and in the footnote of Table 1 (page 23, line 5).

Also discuss explcilty exactly how the reference sector correction is done in each case – yes they all follow the same approach - but there are subtle differences. This is a validation paper after all, so the reader needs to be well informed about the retrieval differences.
We now discussed the reference sector correction in the footnote of Table S1.

Page 3, Line 21: You have gridded the data yourself, so why quote the error on the monthly means from De Smedt's 2008 paper? Surely you can compute this error for each retrieval.
We cited De Smedt et al [2008] here to make the point that uncertainties in HCHO columns can be beaten down by temporal averaging. We have rewritten this part, please see page 3, line 28–29.

Equation (1) - please state if sigma_0 is a vertical (I assume it is) or a slant column background.
Sigma_0 (the reviewer meant to be omega_0) is a CTM-based vertical column. We have clarified this in the text (page 4, line 8).

Reference sector correction – since the retrieval groups are authors on the paper, I would urge the satellite data producers to provided their uncorrected slant columns (i.e. the raw retrieved slant

columns) so that a consistent a reference sector adjustment could be applied to all. This really would add substantial value to the paper. But I suspect that it will not be attempted, which is a shame.
Among all the 6 products, only BIRA retrievals provided the uncorrected slant columns. SAO and OMPS-PCA retrievals provided uncorrected vertical columns, which can be converted to uncorrected slant columns via AMF. We agree with the reviewer that retrieval groups should provide uncorrected columns in their products.

Page 4, line 27: "diurnal variability in the HCHO column is expected to be small..." Please quantify small – is this statement based on measurements or model studies?
We have rewritten this sentence, please see page 5, line 6–7.

Page 5: The normalisation process concerns me. I can sort of see why it was done but there are some opened questions.
We have rewritten this part, please see page 5, line 15–32.

For example, what are the spatial gradients in mixing layer depths, how does this impact the HCHO column.
We have normalized the impact of mixing layer depths on HCHO columns, please see page 5, line 16–18.

It is stated that day to day variability is "fitted well" by exp[0.11T] – please quantify well? What is the uncertainty in the comparison procedure due to fitting this function? How does the satellite comparisons change say if an arbitrary function of exp(cT) is used instead, where 'c' is a different constant?
We have added more details in the exponential fitting, see page 5, line 26.

Isoprene emissions, and thus HCHO columns, are very sensitive to temperature. What do you mean by 'local surface air temperature', where is this from (is it the GEOS-FP data) & how accurate is it?
"Local surface air temperature" refers to the "local surface air temperature at the time of the flight", please see page 5, line 29.

The conversion to a 'mean HCHO columns', using the mean vertical profile, again what is the error in this procedure and how does it impact the subsequent comparisons? What if you use a different assumed profile shape?
We assume here that the mean observed HCHO vertical profile (Figure 1, right panel) is the true profile.

Lastly, please discuss why it's really needed at all? Surely, you could compute HCHO columns for each flight, corresponding to a specific day and underlying surface footprint, and compare this to the satellite observations that lie in the same footprint (in which case you don't need to adjust for day-to-day variability)? I am, perhaps, being quite harsh but this step in the study is very important and yet it is given only a paragraph. Given the importance of the paper, I urge the authors to thoroughly discuss this key part of the analysis.

Single-scene HCHO retrieval has high uncertainty (~100%). So the direct comparison between coincide satellite pixel and aircraft point is not practical. The point here is to obtain the campaign-averaged satellite HCHO columns to beat down the retrieval noise. We have rewritten the text to clarify this, please see page 5, line 18–20.

Section 4: Why don't the retrievals apply the same GEOS-Chem model background correction, and probably better still, compute the AMFs using the GEOS-Chem model data? This would help isolate the columns differences due to the spectral fitting approaches. Yes, this would be some work but it would really add another dimension to the paper. You're already half way there, using the mean CAMS profile anyway.
This paper focuses on difference in HCHO columns among various products and potential drivers of the difference. So we just use the data provided by the retrieval groups without adding new information. Background corrections are almost the same (Table 3, column "Ω_o"), this won't cause too much difference.

Page 8, lines 1-2: the AMFs are also sensitive to the respective cloud parameters for each retrieval (which instrument specific and subject to their own uncertainties), hence it is not just the scattering weights and angles.
We have rewritten this sentence, please see page 8, line 15–17.

Page 8, line 5 onwards: Again using the same surface reflectance data set in the AMF for all retrievals, would give some idea of its relative importance in this validation exercise. Ok, each retrieval 'should' use a reflectance data set consistent with the cloud algorithm but it would be an interesting experiment to carry out.
The reviewer made a good point. But here we focus on differences in HCHO columns among various products and potential drivers of the difference. We have cited sensitivity of AMF to surface albedo from De Smedt et al. [2008], please see page 8, line 22–23.

Page 8, line 15: Am I missing something here, but you've corrected the GEOS-Chem model using the CAMS data, which is adjusted for day-to-day variability, and yet this shows 'good agreement' with temporal dependence of isoprene emissions. Isn't this a somewhat circular comparison/argument? Please elaborate, and if not, then make sure you adjust the text to avoid leaving the reader with this impression.
The reviewer seems mistaken. We applied a +10% correction to the mean GEOS-Chem column averaged over the Southeast US during the entire campaign. We now clarified what we meant, please see page 8, line 27–34.

Page 8, line 30: I would say spatial correlations are 'moderate to relatively high', rather than give the feeling that the retrievals are better than they really are.
Accepted, please see page 9, line 17.

Page 9, final paragraph: yes you are right in what you say! But couldn't the authors address some of these issues in the framework of this paper?
We do give our recommendations, there is no more can be done in the framework of this paper. We have rewritten this paragraph to clarify what we meant, please see page 9, line 28–31.

Figure 3: are the temperature and HCHO columns, both sampled according to OMI observations (i.e. under the same cloud conditions?)
Temperature was the midday (1200–1300, local time) surface air temperature. GEOS-Chem HCHO columns were sampled according to OMI observations. We now clarified this, please see page 18, line 7–8.

Figure 5, the GEOS-Chem data is sampled along the 'OMI schedule' – according to which OMI retrieval? I suspect that each retrieval will have different quality flags for their measurements, so a 'good' measurement in one might be flagged as 'bad' in another. Are the actual same OMI pixels gridded between the retrievals? This also holds for the OMPS data as well. Maybe discuss this in the main text to clarify.
We now clarified this, please see page 20, line 4–5.

Figure 6: why not show the AMFs alongside????
We have added a panel in Figure 6 to show AMFs ($w(\sigma)S(\sigma)$) as functions of pressures.

I'm sorry but Figure 7 looks like coloured spaghetti. Maybe make it multi-panelled, with GEOS-Chem plus OMI data, them GEOS-Chem plus the OMPS data etc.
We are now only showing OMI-SAO and OMI-BIRA in Figure 7. We mentioned other retrievals in the text (page 8, line 30–33).

Table 1 please explicitly state how the detection limits are calculated for each retrieval.
We now stated how the detection limits are determined for each retrieval in the text (page3, line 23–28).

Anonymous Referee #2
Received and published: 7 April 2016

General Comments This paper presents intercomparison results for HCHO vertical columns measurements from 4 satellite instruments and 6 different retrievals and comparison to an aircraft campaign over the Southeast US. The paper is easy to read and very interesting for the community in order to have a better view of the quality of the existing satellites HCHO retrievals. It is within the scope of ACP and should be published after taking into account the comments and improvements listed below.

Specific Comments

page 3, line 3: it's the only point in the paper where the in-situ data (reference and technique) are introduced, and as the whole paper is based on these measurements, it would be good to say a little more about them (maybe at the beginning of Section 3). What is their accuracy, sensitivity, ... see comment relative to page 4.
We have added two sentences to describe the two instruments, please see page 3, line 6–8.

page 3, line 10: to my view, Table 1 should contain more information on the different retrievals. It should highlight the main differences between the various retrievals made on the same satellite sensor. Other differences than only the CTM are important: the albedo, surface, cloud correction, etcc (that are only mentioned in page 8) that impact AMF calculation, but also differences in DOAS settings for the slant column retrieval.
We now added Table S1 in the Supplementary Materials to summarize more retrieval details. We also referred to Table S1 in the text (page 3, line 15–16) and in the footnote of Table 1 (page 23, line 5).

page 4, line 16: and applied Equation (1) "to" compute ... page 4, line 17: AMF based on the its reported: remove "the" or "its".
Accepted, please see page 4 line 25–26.

page 4, section 3: I would add a small paragraph on the 2 in situ techniques that are mentioned in the introduction, and that are considered the "truth" but that have 10% difference, as stated in line 31. Any hint on why?
We have added two sentences to describe the two instruments, please see page 3, line 6–8. The 10% difference is due to the fact that the two instruments are independently calibrated. This difference is generally within the mutual stated accuracy for both instruments. We now explained this difference in the text, please see page 5, line 9–12.

The organization of the figure sequence in this section is a little bit confusing for me (the different part of each figures are not discussed/mentioned together, but at different moments of section 3): first figure 1 (but only right and upper left panels), then Figure 2 (left panel), then bottom left panel of figure 1 (p.5, line 18), then figure 3 (p.5, line 20), then missing reference to figure 4 (p.5, line 29) and then only in page 6, the right panel of figure 2 is presented/mentioned. Maybe the bottom left panel of figure 1 could go with figure 3 and the right panel of figure 2 could go after figure 4?

We have considered rearrangement of those panels, but we would like to stick to the same layout.

page 4, line 24: it is stated that the typical extension of the mixed layer is to 1.5-2 Km in the afternoon, but in figure 1, the caption mention 1 to 3km. It would be good to add the height scale in km on the profile picture (and not only pressure) in order to easily make the link with the figure.
We have changed "1.5–2 km" to "1–3 km", please see page 5, line 3. We have also added a height scale (in km) in Figure 1.

page 4, line 26: most flight hours were in the afternoon. How much is "most"?
We have rewritten this sentence, please see page 5, line 5–6.

page 4, line 27: the diurnal variability is expected to be small. How much is "small"?
We have rewritten this sentence, please see page 5, line 6–7.

page 5, line 5: considering the 10% difference between the 2 in-situ technique, why is CAMS considered as the reference in the rest of the paper (GEOS-CHEM is increased by 10% to reach CAMS values) and not ISAF or the average of both?
We chose to use CAMS as reference because it is less biased against GEOS-Chem (Figure 1 and 2), but we also commented on the impact of using ISAF in the text. We have rewritten several sentences to clarify this, please see page 5, line 12–14.

page 5, line 29: add "as can be seen in figure 4, " before "Initial simulations ..."
Accepted, please see page 6, line 10.

page 5, line 31: add "previous"before "comparisons of GEOS-FP ..."
Accepted, please see page 6, line 12.

page 6, line 1: add "red line in" before "figure 4".
Accepted, please see page 6, line 15.

page 6, line 11: how do you "remove this dependence on altitude"? again, it would be helpful for an easy reading of the paper, to mention the limits of the mixed layer and of the free troposphere in pressure values.
We have rewritten this sentence, please see page 6, line 27. We have also added a height scale in Figure 1, and mentioned the limits of the mixed layer (page 5, line 3) and the free troposphere (page 6, line 30) in pressure in the text.

page 6, line 18: why only mention the spatial correlation with CAMS data and not with ISAF?
We have added the spatial correlation coefficient with ISAF data, see page 7, line 1–2.

page 6, line 26 and lower and Table 2: the spatial (and temporal) correlation coefficient calculation should be mentioned in the text here, and not only in the caption of Table 2
Accepted. We have rewritten this sentence, see page 7, line 9–10.

page 6, line 29: GOME-2A BIRA is noisier wrt GOME-2B, this is clear. How much is this due do degradation and how much to the reduced swath mode?

We found that noise of GOME2A columns is significantly the same before and after the swath mode reduction, so we attribute the noisier pattern seen by GOME2A to its instrument degradation rather than its reduced swath mode. We have rewritten the text to clarify this, please see page 7, line 12–15.

page 7, line 2: "GEOS-Chem columns are sampled on the same schedule and scenes as the individual retrievals" but as GOME2A and GOME2B does not have the same swath, the pixels geometry and cloud impact are not the same, so why the value of GEOS-Chem columns in table 3 is the same (1.59) for both instruments?
We have fixed this in Table 3.

page 7, line 8 to 14: This would be a good place to discuss the different retrieval choices and their impacts on the different retrieval steps (corrected slant column, AMF and background correction).
Actually, we did start to discuss impacts of retrieval choices on retrievals from this paragraph. Please see page 7 line 23–28 for corrected SCD, page 8 line 1–23 for AMF, page 8 line 24–26 for background correction.

page 7, line 22: after discussing the OMI-BIRA shape factors (ie the IMAGES input) wrt to CAMS profile, you could discuss the other model shapes too? and how are GEOS- CHEM and IFAS profile? is this difference between model expected/known? what is the possible difference explanation?
We have rewritten this, please see page 8, line 4–9. The discrepancies between GEOS-Chem and CAMS (and ISAF) are likely due to the insufficient deep convection in the model, we have mentioned this in the text (page 6, line 30).

page 7, line 29: the "much smaller" affirmation in the sentence "The differences with the original AMFs are less than 6.0%, a much smaller correction than for OMI-BIRA" would be more clear with "are less than 0.06, while for OMI-BIRA the change in AMF is around the double (-0.14)".
We have rewritten this sentence, please see page 8, line 10–11.

page 8, line 2: not only "driven by scattering weights and viewing angles". The clouds, albedo, aerosols also impacts the AMFs.
Accepted. We have rewritten this, please see page 8, line 15–17.

page 8, line 5: the difference in surface reflectivity between OMI-SAO and OMI-BIRA is not so small (~23 to 29%). this certinaly impact (and not only "would contribute in part) "the difference in scattering weights ar lower altitudes". Please refer to sensitivity tests and AMF error estimations from literature.
Accepted. We have rewritten this part, please see page 8, line 21–24.

page 8, line 8: the sentence "The scattering weights of OMI-BIRA are lower than those of GOME2B-BIRA (Figure 6) and GOME2A-BIRA (about the same as GOME2B-BIRA, not shown), even though all BIRA retrievals use the same surface reflectivity" is mis- leading. The retrieval is the same (same surface reflectivity), but the geometry of the 2 instruments is different, as well as the time of the day and clouds", so we expect different scattering weights.
We have removed this sentence from the text, please see page 8, line 21.

page 8, line 14: a discussion of table 3 wrt to differences in the slant columns values between the different retrieval is missing (discussion on VCD, AMF and background correction has been done before).
Actually, we did discuss the corrected slant column along with the AMF in page 7, line 22–28.

page 8, line 18: what is the difference between the temporal correlation reported in figure 7 (and here in the text) and temporal correlation values of table 2 ?.
Figure 7 shows the temporal correlation between satellite HCHO column and temperature, please see page 8, line 31–33. Table 2 summarizes the temporal correlation coefficients between HCHO columns for different pairs of satellite retrievals, computed from daily averages of each retrieval over the Southeast US domain. Please see page 7, line 8–9 for description of Table 2.

page 9, line 9: I find the sentence "Aside from OMI-BIRA, the shape factors used in the retrievals are not a significant source of error" a little bit odd. If I understand well table 3, the difference in VCD bias when considering the original profile and when using the CAMS profile is 8% (from -20% to -12% bias) for OMI-BIRA, which is indeed the highest change, but the other products ranges between 7% (for GOME2B-BIRA), to 6% (OMI-SAO), 5% (OMPS-SAO), 3% (GOME2A-BIRA) and finally 1% (OMPS-PCA), which are not so different than 8%. Please reformulate this sentence.
We now clarified this, please see page 9, line 26–27.

Anonymous Referee #3
Received and published: 20 April 2016

This manuscript presents an intercomparison exercise of tropospheric HCHO retrieved from satellites measurements using several products obtained from independent retrieval approaches. In order to validate the products, the authors use aircraft observations during the short term field deployment of SEAC4RS in the Southeast of the US. This is an important validation effort. In general, the paper is well written and within the scope of ACP. That said, I have important comments/suggestion that I recommend revisions before final publication.

Major comments:

- In the abstract and conclusion sections it is mentioned that the HCHO columns are biased low by 20-50%, which is a significant number. Then, it is concluded that HCHO from satellite provides a reliable proxy for isoprene emissions. How do the authors conclude that HCHO is reliable given the significant bias (plus any uncertainty associated to the assumptions of HCHO yields from isoprene conversion)?. The only section with results shown is section 4, I recommend to insert a new section(s) where specif details are given regarding errors and how they affect the isoprene emission.
The 20–50% bias is systematic among retrievals and we gave those correction factors based on in situ observations. We have added one paragraph to discuss effect on isoprene emission estimation, please see page 9, line 1–6.

- The emissions of natural biogenic volatile organic compound have a temperature dependence, hence a seasonal variation that it is not studied in the present study. How- ever, the statements along the manuscript are quite general regarding spatial and temporal distribution of HCHO. The focus of the study is only limited in the southeast of the US during less than two months. I suggest to state clearly that the results/conclusions shown are for this specific time/area.
We have specified this in multiple places in the text.

- The aerosol optical depth (AOD) observed in the southeast of the US is significantly higher than other parts in north America. In addition, the temporal distribution of AOD over the southeast of the US has a maximum peak in the summer months (similar months as this study). However, it is not well documented in section 2 how aerosols are treated in the satellite retrievals. Also, current investigation in the southeast of the US associate aerosol aloft (e.g., Goldstein et al., 2008). I encourage the authors to explain in better detail the effect of aerosols in the retrieval and final uncertainty. How do aerosol aloft impact the retrieval of HCHO columns?
"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." We have added this sentence in page 4, line 20–21.

- The in-situ sensors are taken as the ground truth in this study, however there is quite a bit of manipulation in the conversion of mixing ratios to columns. It is not clear to me why the HCHO need to be normalized if the columns are compared with satellite retrievals. Please give a thorough description of why this is needed and why the "Day-to-day variability in HCHO columns can be fitted well to exp[0.11T]".

We have rewritten the normalization part, please see page 5, line 15–32.

Is there any explanation of why CAMS and ISAF have differences of about 10% since they measure the same air mass?
The 10% difference is due to the fact that the two instruments are independently calibrated. This difference is generally within the mutual stated accuracy for both instruments. We now explained this difference in the text, please see page 5, line 9–12.

- Have the authors looked at the trace gas inhomogeneities captured by the in-situ observations and compare with results of satellite on a pixel-pixel resolution and ?. In other words, do the correlation improve if the air mass sampled is the same?
We didn't focus on individual pixel validation in this paper, because individual satellite has high uncertainty. To reduce the uncertainty, this validation work was done on monthly (Aug. 05 to Sep. 25) and regional (Southeast US) average basis.

- In the abstract it is mentioned that "The GEOS-Chem chemical transport model pro- vides a common intercomparison platform". However, it is not clear how GEOS-Chem is used. Consider expanding the sentence explaining how this is achieved.
We have rewritten this sentence, please see page 1, line 27–28.

- Page 3, line 23: "Here and elsewhere, we use only satellite pixels with solar zenith angle less than 60, cloud fraction less than 0.3, and row anomalies (for OMI) screened". In the same paragraph it is mentioned the detection limit of the satellites but it is not clear if you use only data above the detection limit for the analysis in this work.
We didn't filter satellite pixels by detection limits. Instead, we discarded pixels with too high (10×10^{16} molecules cm^{-2}) or too low (-0.5×10^{16} molecules cm^{-2}) column densities. We have clarified this in this text, please see page 3, line 31–32.

- Page 4, line 16: There are several abbreviations that need to be defined, e.g., SAO, OMPS-PCA, etc. A table/appendix with abbreviations would be useful.
We now summarized those abbreviations in the Appendix. Please see page 10.

- Page 5, line 5-17: The conversion of mixing ratios to columns is achieved by assuming that HCHO is co-located with aerosols (identified with the mixing height from DIAL- HSRL), how valid is this assumption?. I suggest to explain also why an exponential decay with a scale height of 1.9 is used and how do the background column is found.
We have rewritten this paragraph, please see page 5, line 15–32.

Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC^{4}RS 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 SEAC^{4}RS 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 to indirectly validate the satellite retrievals with in situ observations. All retrievals feature a HCHO maximum over Arkansas and Louisiana, consistent with the aircraft observations and GEOS-Chem, and reflecting high emissions of biogenic isoprene. The retrievals are also broadly consistent in their spatial variability over the Southeast US (r=0.4–0.8 on a 0.5°×0.5° grid) as well as their day-to-day variability (r=0.5–0.8). However, all satellite 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 the highest corrected slant columns and the lowest scattering weights in its air mass factor (*AMF*) calculation. Correcting the assumed HCHO vertical profiles (shape factors) would essentially eliminate the bias in the OMI-BIRA data. Based on the validation results during SEAC^{4}RS over the Southeast US, we conclude that current satellite HCHO

Formatted: English (UK)
Formatted: English (US)
Deleted: operational and research
Deleted: provides
Deleted: . We find that all
Deleted: capture the
Deleted: , and
Formatted: Font:Italic
Deleted: used in the AMF calculation
Deleted: further reduce
Deleted: We

data provide a reliable proxy for isoprene emission variability but with a low mean bias due both to the spectral fitting and the scattering weights used in the retrievals. 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].

Deleted: corrected slant columns

Deleted: on

Deleted: A

Deleted: , propagating through the atmosphere and

Deleted: , is retrieved

Deleted: satellite instruments

Formatted: Font:Italic

Formatted: Font:Italic

Deleted: Validation with

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 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, nor would these be helpful because of the large noise in individual retrievals [A. Richter et al., 2013]. 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 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 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 [De Smedt et al., 2012; 2015]. *AMF*s 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, down to 20–40% for monthly means [De Smedt et al., 2008], which corresponds to 0.1–0.3×10^{16} molecules cm^{-2} over the Southeast US. Here and elsewhere, we use only satellite pixels with solar zenith angle less than 60^o, 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.

Deleted: independent

Formatted: Font color: Text 1
Deleted: Except for OMPS-PCA,
Formatted: Font color: Text 1
Deleted: limits
Formatted: Font color: Text 1
Deleted: are
Formatted: Font color: Text 1
Deleted: uncertainties
Formatted: Font color: Text 1
Deleted: .
Formatted: Font color: Text 1
Formatted: Font:Italic, Font color: Text 1
Formatted: Font color: Text 1
Deleted: Total error on monthly means reduces to 20–40% (0.1–0.3×10^{16} molecules cm^{-2} for the Southeast US) [De Smedt et al., 2008].
Formatted: Font color: Text 1
Deleted: determined
Formatted: Font color: Text 1
Deleted: the
Formatted: Font color: Text 1
Deleted: of 1-sigma noise
Formatted: Font color: Text 1, English (UK)
Formatted: Font color: Text 1
Deleted: 2015].
Formatted: Font color: Text 1
Deleted: and
Formatted: Font color: Text 1
Formatted: Font color: Text 1

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.

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

$$\Omega = \frac{\Delta\Omega_S}{\mathrm{AMF}} + \Omega_o. \quad (1)$$

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 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*:

$$AMF_G = \frac{1}{\cos\theta_Z} + \frac{1}{\cos\theta_V}. \quad (2)$$

$$\mathrm{AMF} = AMF_G \int_{p_S}^{0} w(p)S(p)dp. \quad (3)$$

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

Formatted: Font:Not Italic, Not Superscript/ Subscript
Deleted: For
Deleted: , it
Formatted: Font:Italic
Formatted: Font:Italic
Deleted:
Deleted: $+\int_{p_S}^{0} w(p)S(p)dp$
Formatted: Font:Italic
Deleted: its
Deleted: shape factors
Deleted: by
Deleted: its

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 in the HCHO column is expected from models to be less than 10% 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 choose to use CAMS as reference because it is less biased against GEOS-Chem (Figures 1 and 2; see discussion below).

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 SEAC^{4}RS period (5 August–25 September) in order to compare to the satellite data averaged over the same period to reduce noise. 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 SEAC^{4}RS 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 SEAC^{4}RS period as inferred from the CAMS measurements. The spatial distribution is markedly different and smoother than for the original mixed layer data

Deleted: determined

Deleted: .5–2

Deleted: Most

Deleted: in the afternoon, and 95%

Deleted: .

Deleted: small

Deleted: between

Deleted: .

Deleted: good agreement

Deleted: .

Deleted: . ... [1]

Deleted: observed

Deleted: local

Deleted: tried

Deleted: remove these dependences in order to derive

Deleted: that could be compared to the mean spatial patterns in

Deleted: .

Deleted: to exp[

Deleted:]

Deleted: T is the surface air temperature in K and the exponential dependence

Deleted: that of isoprene emission

Deleted: HCHO

Deleted: concentrations in Figure 1

Deleted: We then converted these mean mixing ratios to HCHO columns by using the mean vertical profile information in Figure 1. This step assumed 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.

Deleted: ,

(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-

Deleted: to mean local temperatures for the SEAC^{4}RS period

Deleted: 2006;

Deleted: reveal

Deleted: (

Deleted:).

Deleted: Initial

Deleted: Comparisons

Deleted: is

Deleted: described

Deleted: very

Deleted: Removing

Deleted: ,

Deleted: within the mixed layer

Deleted: ,

Deleted: , which

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°×0.5° 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 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 on the 0.5°×0.5° 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° 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 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 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 $\Delta\Omega_S$ 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 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].

Formatted: Font:Not Bold

Deleted: and indicative of the maximum in isoprene emission

Deleted: They are also consistent in showing high values over the Southeast US due to isoprene emission.

Deleted: correlations of

Deleted: are given in Table 2.

Deleted: [De Smedt et al. 2015]. GOME2A has operated in

Deleted: , reducing its coverage by half and thus leading to greater

Deleted: in the time-averaged data

Formatted: Font:Italic

Formatted: Font:Italic

Deleted: Although

Deleted: was very fine for the

Deleted: 5

Deleted:], it is noisier than the others during the SEAC^{4}RS period, as pointed out above, reflecting instrument degradation and the lower number of observations.

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.

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

Formatted: Font:Italic

Formatted: Font:Italic

Deleted: factor

Deleted: OMI-BIRA retrieval

Deleted: underestimates

Deleted: overestimates

Deleted: , most likely due to the use of IMAGES profiles for the year 2012, as IMAGES profiles for the year 2013 were not available when the BIRA retrievals were processed

Formatted: Font:Italic

Formatted: Font:Italic

Formatted: Font:Italic

Deleted: .0%, a much smaller correction than

Deleted: .

Formatted: Font:Italic

Formatted: Font:Italic

Deleted: solely

Deleted: viewing angles

Formatted: Font:Italic

Formatted: Font:Italic

Formatted: Font:Italic

Deleted:

Deleted: This would contribute in part to the difference in scattering weights at lower altitudes. The scattering weights of OMI-BIRA are lower than those of GOME2B-BIRA (Figure 6) and GOME2A-BIRA (about the same as GOME2B-BIRA, not shown), even though all BIRA retrievals use the same surface reflectivity. Although reduced, part of the difference remains for cloud-free satellite pixels (cloud fraction <0.01).

Deleted: columns

Deleted: month-to-month

Deleted: six different

Deleted: ranges from 0.52 to 0.75 for the different retrievals.

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 consistence in capturing both spatial and daily variations in HCHO columns, as demonstrated by the indirect validation between SEAC^{4}RS observations and satellite retrievals. This suggests that HCHO columns observed from space can provide a reliable proxy for isoprene emission. 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

We have used SEAC^{4}RS 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 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-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 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*.

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 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.

Deleted: with

Deleted:

Deleted: all

Formatted: English (UK)

Deleted: success demonstrates that

Deleted: can provide

Deleted: reliable

Deleted: and reduced sampling

Formatted: Font:Italic

Deleted: (

Formatted: Font:Italic

Deleted:).

Deleted: Focus should be on slant column fitting, on corrected slant columns, and on the calculation of scattering weights. OMI-BIRA has the largest corrected slant columns and smallest scattering weights of all retrievals, yielding the best match to the SEAC^{4}RS observations. This can provide a comparison reference for other retrievals

Appendix

Abbreviations and acronyms

AMF	Air Mass Factor
BIRA	Belgian Institute for Space Aeronomy
CAMS	Compact Atmospheric Multispecies Spectrometer
CTM	Chemical Transport Model
GEOS-FP	Goddard Earth Observing System-Forward Processing
GMAO	Global Modeling and Assimilation Office
GMI	Global Modeling Initiative
GOME2	Global Ozone Monitoring Experiment-2
IMAGES	Intermediate Model of Global Evolution of Species
ISAF	In Situ Airborne Formaldehyde
MEGAN	Model of Emissions of Gases and Aerosols from Nature
OMI	Ozone Monitoring Instrument
OMPS	Ozone Mapping and Profiler Suite
PCA	Principal Component Analysis
RMA	Reduced Major Axis
SAO	(Harvard) Smithsonian Astrophysical Observatory
SCD	Slant Column Density
SCIAMACHY	Scanning Imaging Absorption spectrometer for Atmospheric Chartography
SEAC^{4}RS	Studies of Emissions, Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys
VCD	Vertical Column Density
VOCs	Volatile Organic Compounds

Acknowledgments

We acknowledge contributions from the NASA SEAC^{4}RS Science Team. We would also like to thank the SEAC^{4}RS flight crews and support staff for their outstanding efforts in the field. This work was funded by the US National Aeronautics and Space Administration. We thank Michel Van Roozendael for helpful discussions. Jenny A. Fisher acknowledges support from a University of Wollongong Vice Chancellor's Postdoctoral Fellowship. We thank three anonymous reviewers who provided thorough and thoughtful comments.

Deleted: , especially the CAMS, ISAF, and DIAL/HSRL lidar group

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 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.

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.

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.

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 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, 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.

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

Deleted: .

Moved down [1]: Discuss.,

Deleted: -52, in review

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.

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

Deleted: et al

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.

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., 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.

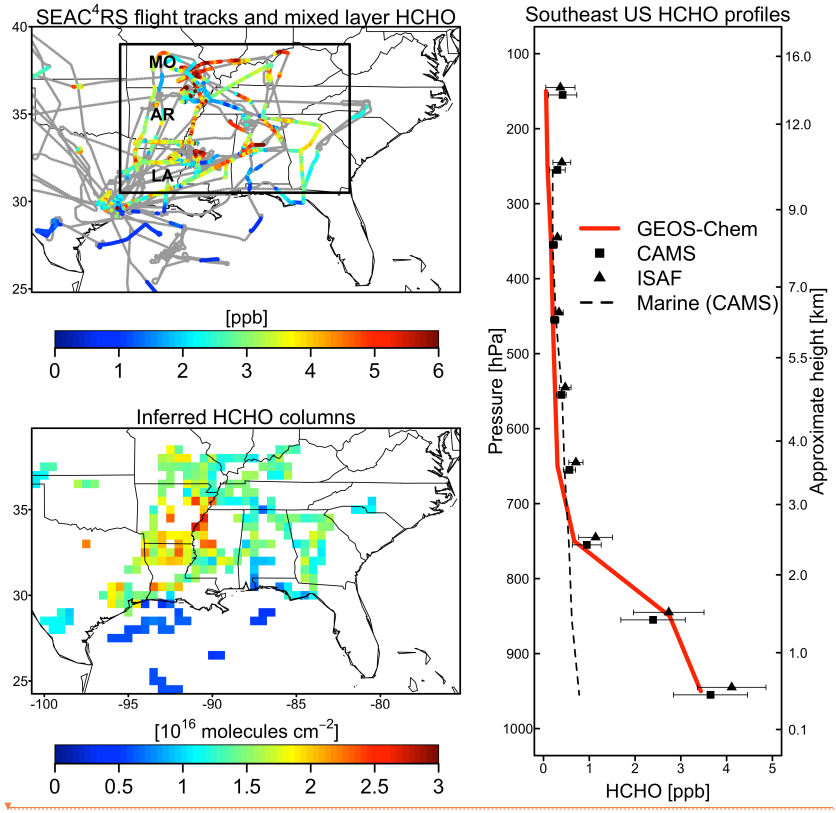
Zhu, L., D. J. Jacob, L. J. Mickley, E. A. Marais, D. S. Cohan, Y. Yoshida, B. N. Duncan, G. González Abad, and K. V. Chance, Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, Environ. Res.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

Deleted: submitted to

Moved (insertion) [1]

Deleted: Diss.,

Deleted: .



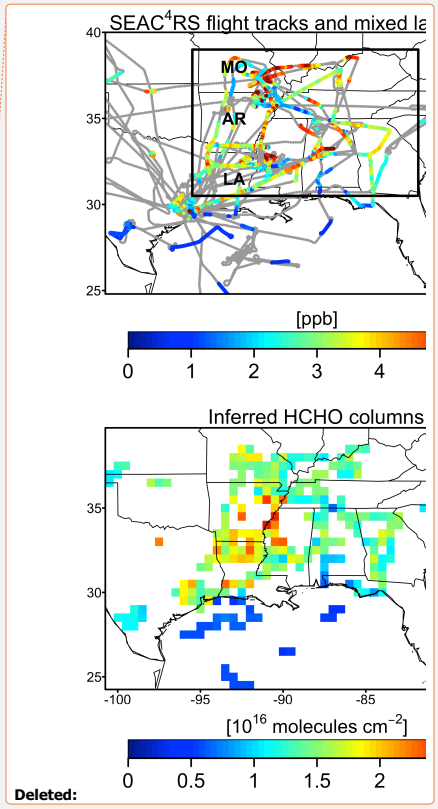


Figure 1: Formaldehyde (HCHO) concentrations along SEAC^{4}RS aircraft flight tracks (5 August–25 September 2013). The top left panel shows the DC-8 flight tracks (in grey) and the CAMS measurements aboard the aircraft in the mixed layer. The mixed layer is the convectively unstable region of the atmosphere 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°–39°N, 95°–81.5°W) 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°N, 96.5°–88.5°W), 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°×0.5° 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).

Deleted:

Deleted:).

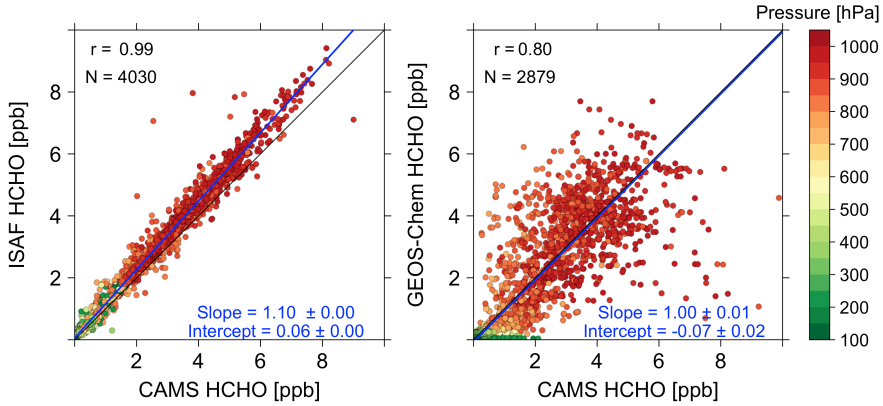


Figure 2. Comparisons between HCHO measurements from the CAMS and ISAF instruments aboard the SEAC^{4}RS aircraft, and simulated by GEOS-Chem, for the Southeast US flight tracks (box in Figure 1). The left panel compares 1-minute measurements from CAMS and ISAF. The right panel compares GEOS-Chem and CAMS HCHO. Here and elsewhere for model-observation comparisons, HCHO observations along the flight tracks are averaged onto the GEOS-Chem grids ($0.25^o \times 0.3125^o$, 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.

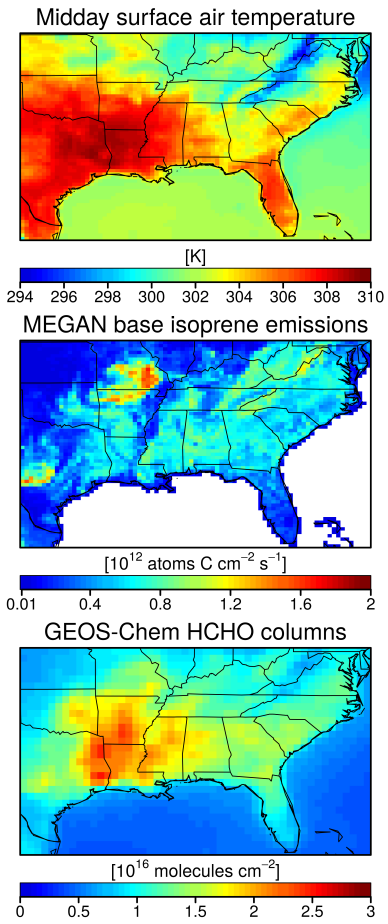


Figure 3. Mean temperature, base isoprene emissions, and HCHO columns in the GEOS-Chem model for the SEAC^{4}RS period (5 August–25 September 2013). The top panel shows the midday (1200–1300 local time) surface air temperature from the GEOS-FP assimilated meteorological data. The middle panel shows the 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.

Deleted: 2006;

Deleted: (

Deleted:).

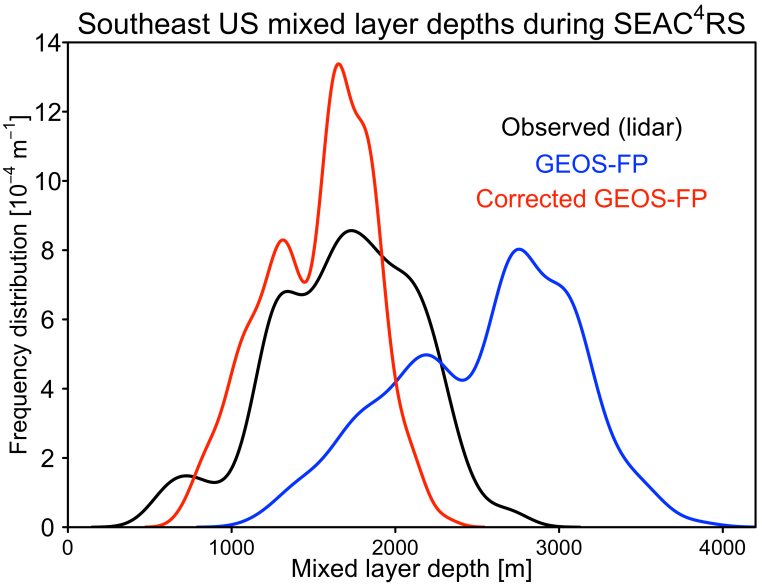


Figure 4. Frequency distribution of mixed layer depths over the Southeast US during the SEAC^{4}RS period (5 August–25 September 2013). Observations by aerosol lidar aboard the aircraft [DIAL-HSRL Mixed Layer Heights README, 2014] are compared to the local GEOS-FP data used to drive GEOS-Chem, before 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.

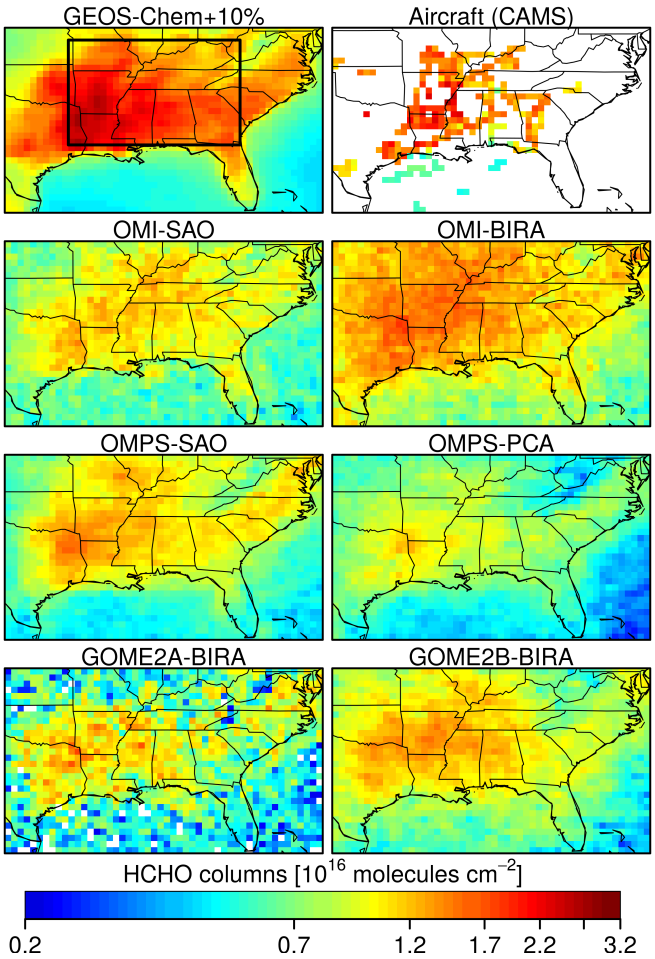


Figure 5. HCHO vertical column densities over the Southeast US averaged over the SEAC^{4}RS period (5 August–25 September 2013). The bottom panels show six retrievals from four satellites (OMI, GOME2A, GOME2B and OMPS) and three different groups (Table 1). The top panels show (1) GEOS-Chem model results 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.

Deleted: GOME-2A, GOME-2B

Deleted: Color bar

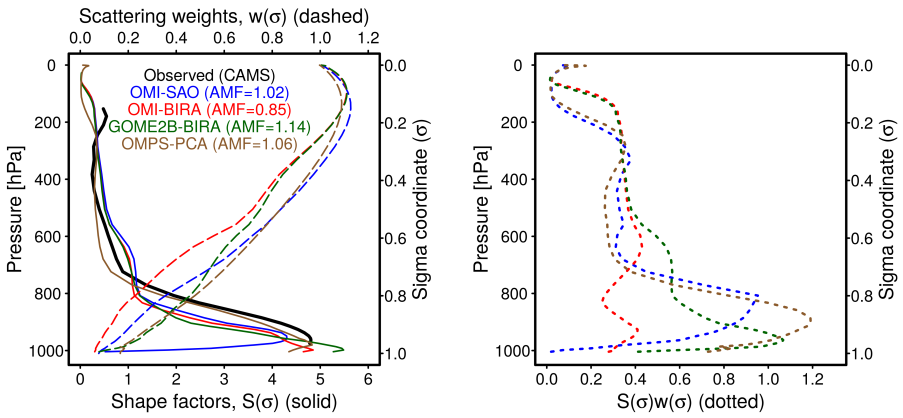
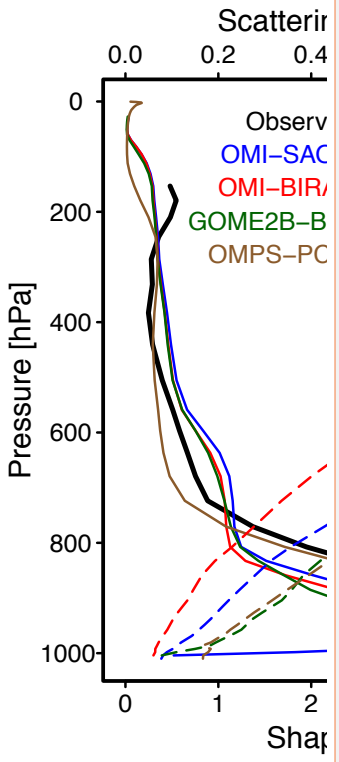


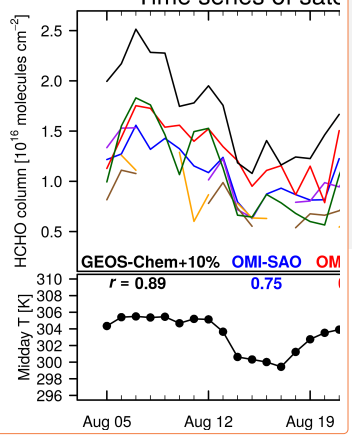
Figure 6. Air mass factor differences between retrievals. The left panel shows mean scattering weights (*w*) and shape factors (*S*) for HCHO retrievals over the Southeast US during the SEAC^{4}RS period, and the right panel shows the product of the two from which the *AMF* is derived by vertical integration 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.



Deleted: ... [2]

Deleted: resulting air mass factor (AMF) computed…he right panel shows the product of the two from equation (3)….hich 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 are…s the observed HCHO shape factors ... [3]

Formatted: Font color: Text 1



Deleted:

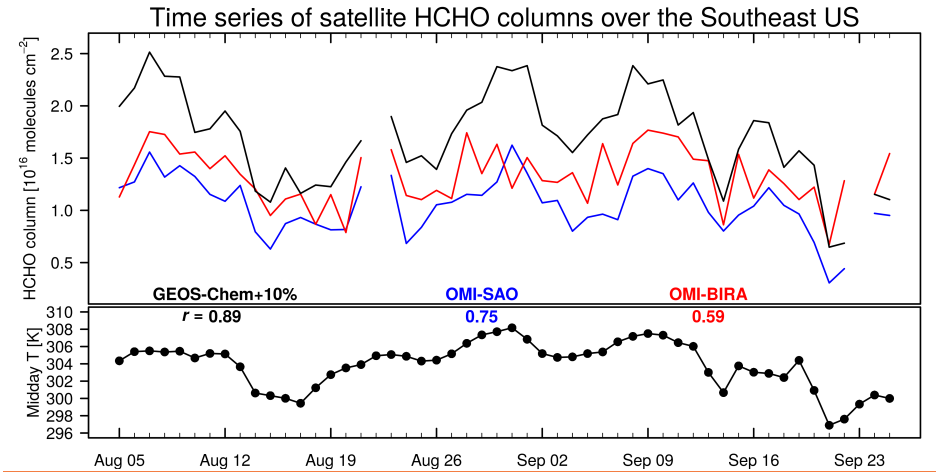


Figure 7. Daily variability of HCHO vertical column densities over the Southeast US during SEAC^{4}RS. The top panel shows daily HCHO columns averaged over the Southeast US (box in Figure 5) for the OMI-SAO and OMI-BIRA retrievals. GEOS-Chem columns (black) are sampled following the OMI viewing geometry 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.

Deleted: different satellite

Deleted: of Table 1

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

Retrieval	Nadir resolution [km^2]	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 SEAC^{4}RS 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.

Inserted Cells

Formatted: Position:Horizontal: 0.77", Relative to: Page, Vertical: 0.63", Relative to: Page, Horizontal: 0.13", Wrap Around

Deleted: limit[c]

Formatted Table

Deleted: Reference[d]

Deleted: (LT)

Deleted: retrievals from

Deleted: GOME-2A, GOME-2B

Deleted: Miller

Deleted: [c] Single-scene spectral fitting uncertainties (except for OMPS-PCA). Fitting windows for the different retrievals are in the 327.7–356.5 nm range. Detection limits of GOME2A and B were in 2013. Detection limit of OMPS-PCA is determined as the 4-times of 1-sigma noise over the Pacific Ocean [Li et al., 2015]. ... [4]

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 $0.5^o \times 0.5^o$ 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]

	Mean values[b]						with CAMS shape factors			GEOS-Chem+10%[c]
Retrieval	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.

Formatted Table
Deleted: 31
Deleted: 38
Deleted: 59
Deleted: operational and research
Deleted: AMF (AMF_G), air mass factors (AMF),
Deleted: equation
Deleted: [
Deleted:]
Formatted: Font:Italic
Deleted: section
Deleted: B
Deleted: Table
Formatted: Font:Italic
Formatted: Font:Italic

Page 5: [1] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

Part of the horizontal

Page 21: [2] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

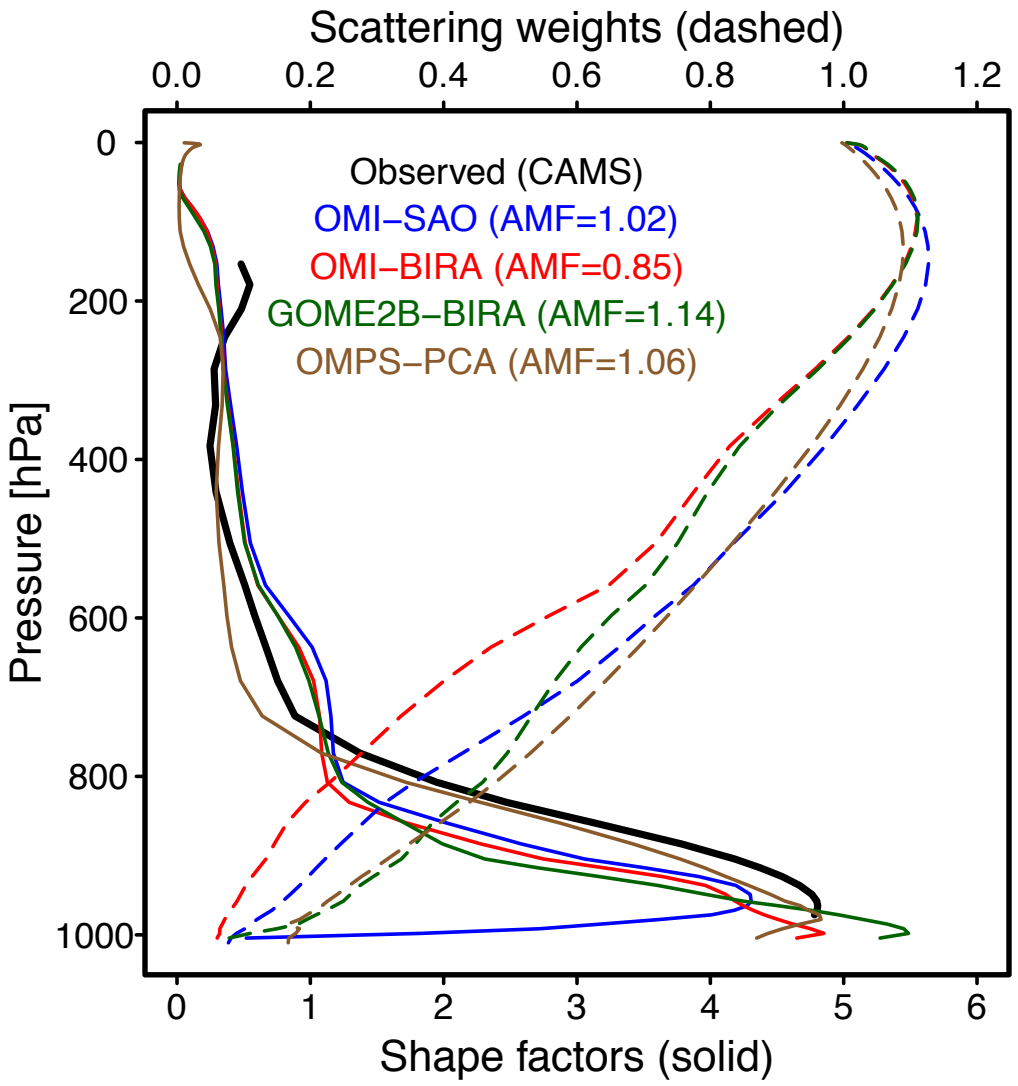


Figure 6. Mean

Page 21: [3] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

resulting air mass factor (AMF) computed

Page 21: [3] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

resulting air mass factor (AMF) computed

Page 21: [3] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

resulting air mass factor (AMF) computed

Page 21: [3] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

resulting air mass factor (AMF) computed

Page 23: [4] Deleted	Zhu, Lei	7/8/16 11:13:00 AM

[c] Single-scene spectral fitting uncertainties (except for OMPS-PCA). Fitting windows for the different retrievals are in the 327.7–356.5 nm range. Detection limits of GOME2A and B were in 2013. Detection limit of OMPS-PCA is determined as the 4-times of 1-sigma noise over the Pacific Ocean [Li et al., 2015].

[d]