Responses to reviewers' comments on "Validation of formaldehyde products from three satellite retrievals (OMI SAO, OMPS-NPP SAO, and OMI BIRA) in the marine atmosphere with four seasons of ATom aircraft observations"

We appreciate the valuable feedback and support from two reviewers and Jean-Francois Muller regarding the publication of this manuscript following revisions. In response to their suggestions, we have carefully revised the manuscript. To facilitate the review process, we have included the reviewers' comments in black text, with our responses in blue. All comments have been addressed, and the corresponding changes to the manuscript are tracked.

Referee #3:

Liao et al., utilize a four-season deployment of Atom aircraft observations to validate three HCHO retrieval products. They demonstrate that these HCHO products generally capture the spatial and seasonal distribution of HCHO in the remote ocean-atmosphere albeit with a low bias. An important result of this study is that the biases in slant column corrections have larger impacts on retrieval than AMFs. The paper is well-organized and includes technical details that fit well into the scope of AMT. I hope the authors can address the following comments before the paper is accepted for publication in AMT.

Major comments:

The conclusion of this paper could be further strengthened. While the study effectively validates these HCHO retrievals and addresses differences in HCHO columns across latitudes and seasons, it would be valuable for the authors to provide practical advice to users of these products. For example, do the authors have any recommendations on which retrieval product is preferable? Would averaging across multiple products yield more accurate results than using a single product? Alternatively, should the spread among the three products be treated as an indicator of uncertainty in HCHO retrieval?

According to the analysis of this study, we recommend OMI-SAO (v004) at least for the remote ocean atmosphere studies because this retrieval has the best agreement and smallest mean biases compared to ATom in situ data.

In the abstract, we have added "All retrievals are correlated with ATom integrated columns over remote oceans, with OMI SAO (v004) showing the best agreement." We now added "This is also reflected in the mean bias (MB) for OMI SAO (-0.73 ± 0.87), OMPS SAO (-0.76 ± 0.88), and OMI BIRA (-1.40 ± 1.11). We recommend the OMI-SAO (v004) retrieval for remote ocean atmosphere studies."

We would not advise averaging retrievals or using the spread as a measure of uncertainty, as the ATom profiles should serve as ground truth.

Other comments:

1. The OMI satellite overpass time is 1:30 pm local time while the Atom observations were conducted throughout the day. How do you account for the time difference when comparing HCHO retrievals to Atom observations?

   Section 3.2 added "Data on the diurnal variation of HCHO columns in the remote oceanic atmospheric are very limited (e.g., the Mauna Loa site in the supplementary information of Vigouroux et al. (2018)). Given the possible diurnal variation of HCHO, the difference between aircraft sampling time and satellite overpass time (1:30 pm) may account for some, but not the majority, of the discrepancies between satellite and ATom measurements at high latitudes (Fig. 4S and 5S). The differences across latitudes due to time variation may amount to approximately 0.2×10^{15} molecules cm^{-2}, based on the simulation results (Fig. 4S and 5S). Further research is needed to more accurately quantify the diurnal variation of HCHO over oceanic regions."

   Supplementary material added:

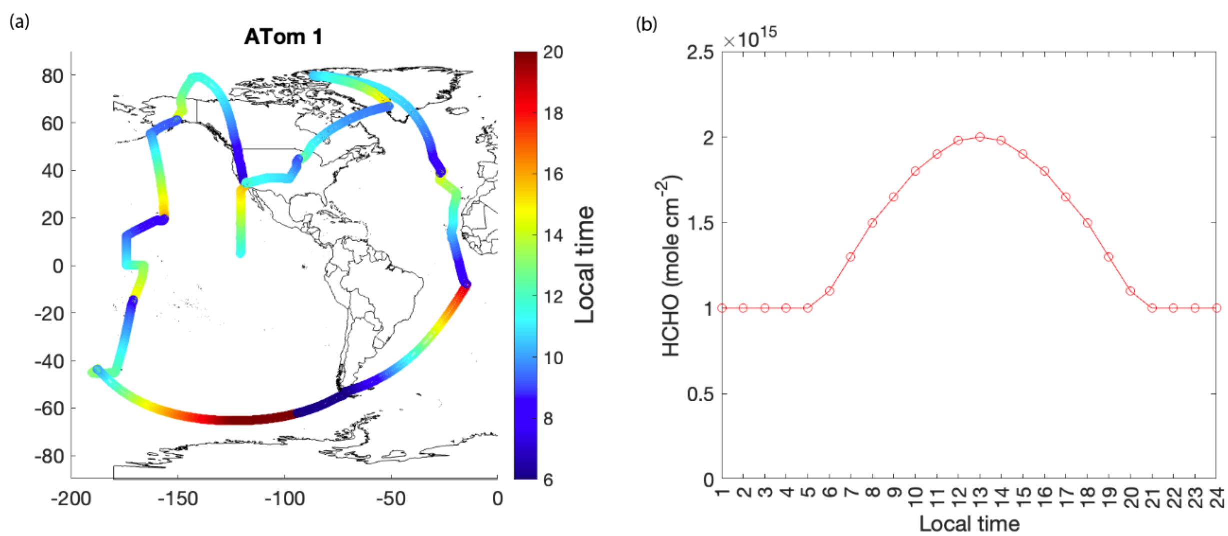

Figure S4. (a) Map of ATom1 flight track color-coded with local time. ATom 2, 3, and 4 maps are similar to ATom1 and not shown here. (b) Diurnal variation of HCHO columns with maximum value of 2.0×10^{15} molec cm^{-2} at 1: 00 pm is simulated, as an example. The diurnal variation is based on the profiles from Bruno Franco et al. (2016) and the maximum value is selected based on average satellite HCHO measurements at the northern high latitudes. It is important to note that the diurnal variation shown in (b) likely represents the upper limit of diurnal HCHO column fluctuations in the remote oceanic atmosphere, especially in high latitudes, as suggested by the measurements reported in Vigouroux et al. (2018).

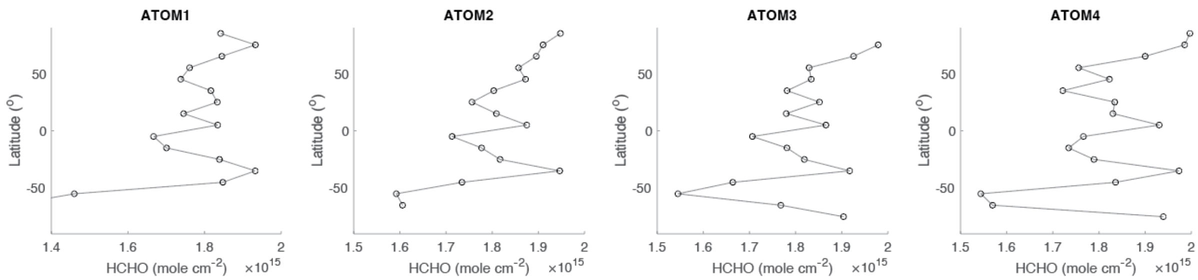

Figure 5S. The latitude-averaged distribution of simulated HCHO columns, using HCHO columns as a function of local time (Figure S4b) and the local time of the ATom flight tracks. This figure highlights the differences between satellite and ATom measurements across latitudes, which arise solely from the time discrepancies-- 1:30 pm local time for satellite measurements and varying local times for ATom measurements (Figure S4a). When comparing these measurements across latitudes, ATom measurements may appear higher than satellite measurements at higher latitudes (e.g., ATom1 70° N compared to 30° N in Figure 5S) due to local time differences. However, the local time effect contributing about 0.2×10^{15} molec cm^{-2}, is relatively minor compared to the overall differences between satellite and ATom measurements across latitudes (e.g., Figure 2 ATom1 70° N vs. 30° N). The relatively large variation in high southern latitudes may suggest that the simulated HCHO column variability is significantly overestimated for this region.

2. Line 121: the ascents and descents of aircraft measurement cover 200-450 km in horizontal distance, which is larger than the pixel size of satellite retrievals. Also, the aircraft provides *in-situ* measurements while the satellite measures pixel by pixel. How do you account for the differences in the spatial scales of these two observations?

   Line 121-122 change "In situ HCHO columns are calculated using the method described in Wolfe et al. (2019)" to "In situ HCHO columns are compared to the average of satellite grid cells intersected by the in situ profile area and calculated using the method described in Wolfe et al. (2019)."

3. It is unclear how you treat cloudy conditions when mapping satellite retrievals to ATOM observations. Do you only select satellite/ATOM observation under clear sky conditions?

   In section 2.2.5 Line 219- 221 we stated "SAO L2 data with solar zenith angle > 60°, cloud fraction > 40%, main data quality flag not equal to 0 are excluded. OMI BIRA L2 data with solar zenith angle > 60°, cloud fraction > 40%, and processing error flag ≠ 0 but ≤ 255 are excluded."

4. Figure 2: why is there an enhancement of the HCHO column at ~ -60 latitude bins in the OMI BIRA retrieval products?

   Section 3.2 added "The enhancement of HCHO columns around the -60° latitude bins may be attributed to noise in the OMI BIRA retrievals, specifically anomalous elevated values around filtering gaps when zoomed in, as observed over high southern latitudes in ATom 2 and ATom 3 (Figure 1)."

5. Line 313: since negative bias is more pronounced at higher latitudes, does it suggest that the latitude-dependent background correction is insufficient?

   In Line 314, we stated “This is probably indicative of issues with latitude-dependent background corrections in satellite retrievals and/or global model bias.”

6. Line 367-368: what does “variability” refer to here? If it refers to uncertainties, a factor of 10 seems too large. If it refers to the full range of corrected slant columns, I don’t understand why this implies that uncertainties in AMF are a minor contributor to overall retrieval error

   Changed “Partly this is because the range of variability in AMFs is small (factor of 2) compared to variability in corrected slant columns (factor of 10).” To “ This is primarily because the low OMI BIRA to OMI SAO AMF ratios correspond to the low HCHO column values and the data are spread.”