Responses to Referee's Comments

We appreciate careful reading and lots of valuable comments.
We wrote referee's comments in black, our responses to comments in blue and italics, and the revised manuscript in red.

Referee #1:

In this paper, Kwon et al. described the HCHO retrieval algorithm to be implemented with the Geostationary Environment Monitoring Spectrometer (GEMS). The authors discussed the three main steps in the algorithm (namely preprocessing, spectral fitting, and postprocessing), carried out uncertainty analysis, and also compared GEMS results (using OMI radiance data) with existing OMI HCHO results and MAX-DOAS at a few stations. Once launched, the HCHO data from GEMS can potentially be used in studies on regional air quality, and biomass burning in large areas over East and Southeast Asia. A paper providing detailed documentation of the retrieval algorithm is certainly of great interest to data users and the satellite remote sensing community. Overall, the paper is well organized, and figures and tables are mostly clear. I would recommend publication in AMT after some clarifications (see below):

Specific comments: One would assume that there are some similarities and differences between the GEMS and OMHCHO algorithms. Some of these are discussed throughout the text, but it would be useful to have a table or a paragraph summarizing the different setups (and the resulting differences in HCHO) between the two instruments.

We used the same fitting options with OMHCHO products (González Abad et al., 2015), but auxiliary data such as model data for background corrections and AMF LUT are different. Also, we do not use undersampling correction and latitudinal bias correction for GEMS in default. For clarity, we only described GEMS fitting options in Section 2.2 and added Section 4.1 to explain differences of fitting options between GEMS and OMI as follows:

4.1 Retrieval of OMI HCHO

GEMS fitting options described in Table 1 are largely consistent with those of OMHCHO products (González Abad et al., 2015). However, we do not include spectral undersampling (Chance et al., 2005) in the fitting process for GEMS, and reference sectors for a radiance reference are 143-150°E (shaded areas in Fig. 1). For OMI products, spectral undersampling needs to be included, and radiance references are from the Pacific Ocean as described in González Abad et al. (2015). We use simulated HCHO vertical columns for the background correction, which are zonally and monthly averaged over the reference sector (140-160°W, 90°S-90°N) except for Hawaii (154-160°W, 19-22°N).

In addition, we need to correct latitudinal biases for OMI. Previous studies explained that the latitudinal biases result from spectral interferences of BrO and O_3, whose concentrations are a function of latitude and are high in high latitudes (De Smedt et al., 2008; De Smedt et al., 2015; González Abad et al., 2015). Therefore, the latitudinal biases were corrected when a radiance reference was used as the reference spectrum (De Smedt et al., 2008; González Abad et al., 2015; De Smedt et al., 2018). We correct the latitudinal biases, which are slant columns retrieved for a radiance reference and are averaged as a function of latitude, by subtracting the biases from the corrected slant columns in Eq. 11. Figure 6 shows OMI HCHO slant columns from OMHCHO products (Fig. 6a) and the GEMS algorithm without and with latitudinal bias corrections (Fig. 6b and 6c). HCHO slant columns without latitudinal bias corrections (Fig. 6b) are retrieved larger in 5°N-25°N than OMHCHO products, but HCHO slant columns with the bias corrections are in better agreement with OMHCHO products. Figure 6d shows the absolute differences between OMI HCHO slant columns with and without latitudinal bias corrections from the GEMS algorithm as latitudinal biases. Slant columns with bias corrections increase at latitudes lower than 5°N and higher than 25°N but decrease at latitudes from 5°N-25°N. However, latitudinal biases can be minimized when using a radiance reference as a function of each cross-track position in the south to north direction for GEMS. In default fitting options, therefore, we do not include latitudinal correction and do not analyze uncertainty of latitudinal corrections in Section 3. However, a further investigation for the latitudinal biases needs to be required after GEMS is launched.

Figure 7 shows an example of retrieved HCHO optical depths and fitting residuals as functions of wavelengths for a pixel in Indonesia (March 23 2005; orbit 3655). The retrieved HCHO slant column is 3.2×10^{16} molecules cm^{-2}, which is relatively high due to biomass burning in that region. Average slant column and random uncertainty for all pixels on the orbit are 7.6×10^{15} and 6.9×10^{15} molecules cm^{-2}, respectively, over the GEMS domain. The large random uncertainty of 100% or larger results from pixels with low concentrations, where averaged slant columns and random uncertainties are 2.2×10^{15} and 6.2×10^{15} molecules cm^{-2}.

It appears that the background correction is a main contributor to the differences between GEMS and OMHCHO. And the two used different versions of GEOS-Chem for background VCDs. Is it possible to compare the model HCHO VCDs from the same model over the GEMS "background" area and OMHCHO "background" area? The easternmost part of GEMS FOR is still relatively close to Asia (and biomass burning and CH4 sources). A comparison may help to determine if the GEMS background is "background" enough.

In operation for GEMS, we will use easternmost regions as GEMS reference sectors (143-150°E), which are relatively clean areas for GEMS. Figure S1 shows model HCHO VCDs in GEMS background area (dashed) and OMI background area (solid) and absolute differences between the two.
We discussed it as follows:

For GEMS, we plan to use simulated HCHO columns over easternmost regions (143-150°E) as GEMS reference sectors, which are shaded areas in Fig. 1. The GEMS reference sectors include part of islands near the equator and Japan but are relatively clean areas in south/north direction over the GEMS domain. In comparisons with background HCHO vertical columns over the Pacific Ocean for OMI (Fig. S1), annual mean of GEMS background columns over 4°S–45°N is 3.3×10^{15} molecules cm^{-2} slightly higher than that of OMI background columns (3.2×10^{15} molecules cm^{-2}), showing that we can use easternmost regions as background in the GEMS domain. Occasionally, local differences

between GEMS and OMI background columns can be as large as 3.8×10^{15} molecules cm^{-2} in the tropical region of the southern hemisphere due to biogenic activity and biomass burning, but the standard deviation of background values in that region is 5.1×10^{14} molecules cm^{-2} even lower than that of 1.2×10^{15} molecules cm^{-2} in the middle latitude (>30°N), indicating that the influences from biogenic activity and biomass burning can be corrected by model simulations.

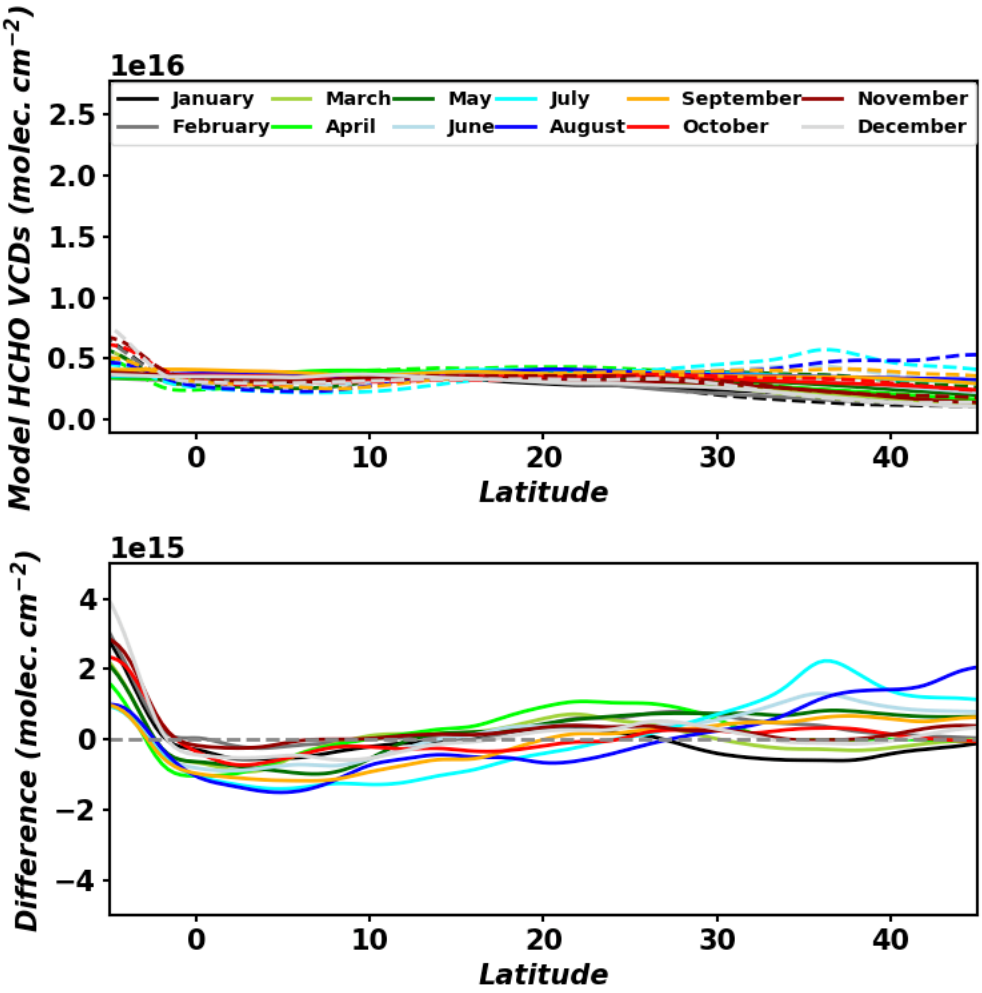

Figure S1. Simulated HCHO vertical column densities in GEMS background area (dashed lines) and OMI background area (solid lines) (top), and absolute differences between the two (bottom).

Some symbols used in equations are not defined (immediately before or after the

introduction of the equation), for example i and j in equation 11.

We explained symbols of i, j, and VCD_m after Eq. 11 as follows:

where i and j indicate pixel indices of cross and along tracks, respectively, and VCD_m denotes a background vertical column density from the model.

Page 7, Line 6: is 300 DU VCD of ozone for the pseudo cross section calculation?

The value of 300 DU is a slant column density of ozone for the pseudo cross section calculation and is scd_{ref} in Eq. 5.

Page 8, Line 7 and Figure 1: maybe you can define and plot the background areas?

I defined longitudinal ranges (143-150°E) for common modes and radiance references. We updated Fig. 1:

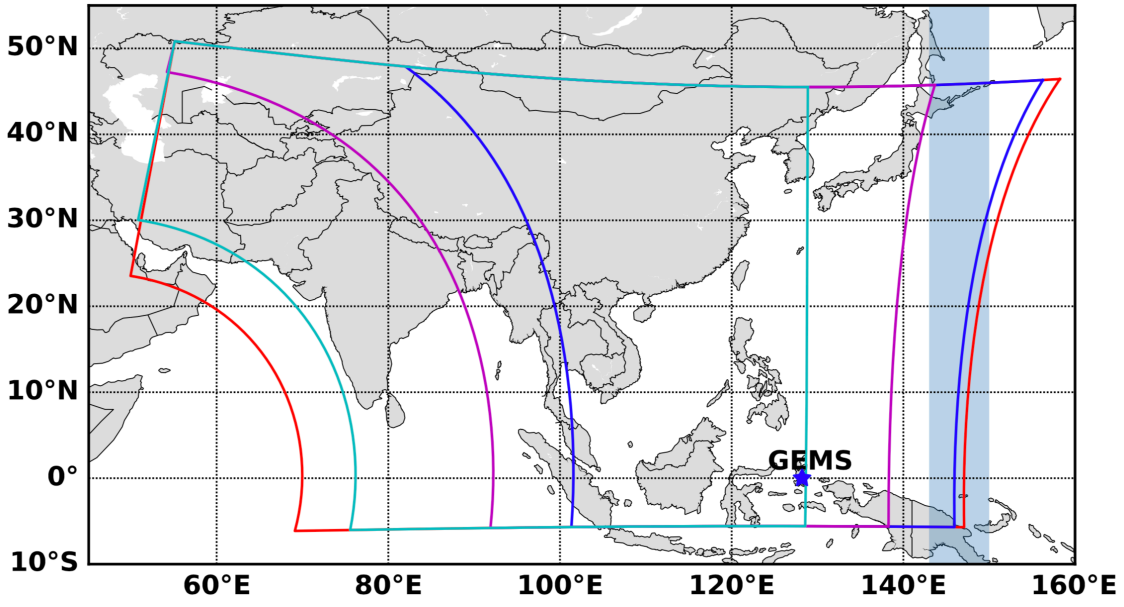

Figure 1. GEMS field of regard (red), nominal daily scan (blue), full central scan (magenta), full western scan (cyan), and GEMS location (blue star). Shaded areas (143-150°E) are regions for radiance references and common mode.

We also updated Table 1 and modified some sentences related with reference sectors as follows:

The common mode denotes fitting residuals caused by instrument properties which have not been determined from physical analysis. Accounting for the common mode can reduce fitting residuals and fitting uncertainties without affecting the retrieved slant columns (González Abad et al., 2015). The common mode for GEMS can be calculated by averaging fitting residuals at every cross-track over easternmost swaths (143-150°E) shown as shaded areas in Fig. 1, which are relatively clean regions.

Table 1 summarizes the detailed information used in the GEMS HCHO retrieval algorithm. We follow fitting options in González Abad et al. (2015). We use measured radiances as the reference spectrum, called a radiance reference, and measured radiances are averaged over the easternmost swaths (143-150°E; shaded areas in Fig. 1) for a day as a function of cross-track positions in the south to north direction. Background corrections are required when we use a radiance reference and are discussed in Section 2.2.5. Also, GEMS has cross-track swaths in the south to north directions while instruments such as OMI and TROPOMI have west to east swath. Therefore, latitudinal biases resulting from BrO and O_3 latitude-dependent interferences can be minimized for GEMS and are discussed in Section 4.1.

Table 1. Summary of GEMS system attributes, parameters for radiance fitting, and parameters for the AMF look-up table.

GEMS system attributes	
Spectral range	300–500 nm
Spectral resolution	< 0.6 nm
Wavelength sampling	< 0.2 nm
Signal-to-noise ratio	> 720 at 320 nm > 1500 at 430 nm
Field of regard	≥ 5000 (N/S) × 5000 (E/W) km^2 (5°S-45°N, 75°E-145°E)

Spatial resolution (at Seoul)	< 3.5 × 8 km^2 for aerosol < 7 × 8 km^2 for gas
Duty cycle	~ 8 times/day
Imaging time	≤ 30 minutes
Radiance fitting parameters[a]	
Fitting window (calibration window)	328.5–356.5 nm (325.5–358.5 nm)
Radiance reference	Measured radiances from far east swaths (143-150°E) for a day
Solar reference spectrum	Chance and Kurucz (2010)[b]
Absorption cross-sections	HCHO at 300 K (Chance and Orphal, 2011) O_3 at 228 K and 295 K (Malicet et al., 1995; Daumont et al., 1992) NO_2 at 220 K (Vandaele et al., 1998)[b] BrO at 228 K (Wilmouth et al., 1999) O_4 at 293 K (Thalman and Volkamer, 2013)[b]
Ring effect	Chance and Spurr (1997)[b]
Common mode	On-line common mode from easternmost swaths (143-150°E) for a day
Scaling and baseline polynomials	3rd order
AMF look-up table parameters	
Longitude (degree) (n=33)	70 to 150 with 2.5 grid
Latitude (degree) (n=30)	-4 to 54 with 2.0 grid
Solar Zenith Angle (degree) (n=9)	0, 10, 20, 30, 40, 50, 60, 70, 80
Viewing Zenith Angle (degree) (n=9)	0, 10, 20, 30, 40, 50, 60, 70, 80
Relative Azimuth Angle (degree) (n=3)	0, 90, 180
Cloud Top Pressure (hPa) (n=7)	900, 800, 700, 600, 500, 300, 100
Surface Albedo (n=7)	0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0

[a] GEMS fitting parameters follow González Abad et al. (2015). However, undersampling is not included in the fitting parameters for GEMS, and reference sectors for radiance reference and common mode are different.

[b] The datasets are used in QA4ECV retrievals. Please refer to De Smedt et al. (2018) for other datasets and fitting options.

Page 11, Line 15: would you expect that destriping would be necessary in the south/north direction?

GEMS has cross tracks (swaths) in the south to north directions. When we use radiance references, we make radiance references as a function of cross-track positions, already including effects of latitudinal dependent ozone absorption. Therefore, we think stripe patterns or biases would not be expected in the south/north direction.

We explained latitudinal biases in more detail as follows:

In addition, we need to correct latitudinal biases for OMI. Previous studies explained that the latitudinal biases result from spectral interferences of BrO and O_3, whose concentrations are a function of latitude and are high in high latitudes (De Smedt et al., 2008; De Smedt et al., 2015; González Abad et al., 2015). Therefore, the latitudinal biases were corrected when a radiance reference was used as the reference spectrum (De Smedt et al., 2008; González Abad et al., 2015; De Smedt et al., 2018). We correct the latitudinal biases, which are slant columns retrieved for a radiance reference and are averaged as a function of latitude, by subtracting the biases from the corrected slant columns in Eq. 11.

…

However, latitudinal biases can be minimized when using a radiance reference as a function of each cross-track position in the south to north direction for GEMS. In default fitting options, therefore, we do not include latitudinal correction and do not analyze uncertainty of latitudinal corrections in Section 3. However, a further investigation for the latitudinal biases needs to be required after GEMS is launched.

It appears that latitudinal correction is implemented for the GEMS prototype algorithm – can the authors discuss uncertainties associated with this?

As we mentioned above, latitudinal corrections are not included for GEMS. For clarity, we remained explanation only related with GEMS in Section 2.2 and explained fitting options for OMI in Section 4.1. Please refer to the first answer.

Figure 7: the figure is quite confusing – can the authors provide more detailed description and discussion?

AMF uncertainties are as functions of parameters in the AMF LUT and are sensitive to measurement conditions. Therefore, figure 7 is too confusing to explain the contributions of parameters so that we deleted Fig. 7. However, we added Table 2 to describe retrieval uncertainties of GEMS HCHO VCDs due to AMF uncertainties. We discussed it as follows:

Table 2 summarizes estimated retrieval uncertainties of GEMS HCHO VCDs due to AMF uncertainties as functions of surface albedos, cloud top pressures, and cloud fractions. Values are calculated assuming conditions with solar zenith angle of 30°, viewing zenith angle of 30°, relative azimuth angle of 0°, cloud fractions less than 0.3, and a profile height of 700 hPa. Uncertainties of HCHO VCDs due to AMF uncertainties can be as large as 20% and 24% of HCHO VCDs in clean and polluted areas, respectively. Maximum values occur for conditions with low surface albedo and clouds at high altitudes, and high cloud fractions, but they do not differ much between clean and polluted areas. However, AMF driven HCHO uncertainty with respect to the profile height in polluted areas is higher than that in clean areas, implying that accurate HCHO profile information in polluted areas is important for the GEMS HCHO retrieval. We can minimize the a priori HCHO profile uncertainties by using averaging kernels.

Table 2. Retrieval uncertainties of GEMS HCHO VCD due to AMF uncertainties as functions of surface albedos, cloud top pressures, cloud fractions, and HCHO profile heights for clean and polluted areas. Values are calculated for conditions with solar zenith angle of 30°, viewing zenith angle of 30°, relative azimuth angle of 0°, cloud fractions less than 0.3, and a profile height of 700 hPa.

AMF contribution to HCHO VCD uncertainty	Clean	Polluted
Surface albedo (α_s)	1-10%	1-12%
Cloud top pressure (p_c)	0-11%	0-11%
Cloud fraction (f_c)	0-19%	0-17%

HCHO height (p_h)	0-11%	0-17%
Total	2-20%	3-24%

Page 16, Line 29-34: can the authors briefly mention what kind of method/strategy/data will be used for aerosol correction, in the case of dust/smoke?

We briefly referred to methods for aerosol correction. We added sentences below the paragraph as follows:

We plan to update our AMF LUT as a function of aerosol optical depth, single scattering albedo, and aerosol height, which will be retrieved in GEMS, to account for the effect of absorbing aerosols. On-line AMF calculation can also be used for aerosol correction with cloud information and model simulation (Lin et al., 2014).

Figure 11 and related discussion: if ozone is an important contributor to the differences between GEMS and BIRA OMI, maybe the authors can also compare the results from tropics and mid-latitude areas separately? One may expect somewhat better agreement between the two in the tropics? Or maybe the authors can run some test GEMS retrievals using the ozone cross section as used in BIRA retrievals?

We found that differences between GEMS and QA4ECV (BIRA) OMI results are caused by polynomial orders. We tested effects of polynomial orders on slant columns. Also, we tested effects of different O_3 absorption cross-sections and non-linear O_3 absorption.

We found that using the 4th polynomial order improves both correlation coefficient and regression slopes although OMI GEMS HCHO slant columns are higher than those in QA4ECV (Fig. S2). Using different O_3 datasets and considering non-linear effects of O_3 improve statistics and relative differences between GEMS and QA4ECV products (Table S1 and S2).

We discussed them in the manuscript as follows:

The discrepancy between the two products could result from the radiance fitting. The OMI QA4ECV products use the DOAS method while the GEMS algorithm uses a non-linearized fitting method (BOAS) for radiance fitting. We also find that polynomial orders accounting for Rayleigh and Mie scatterings are important factors, causing differences between the two products. Retrieved slant columns using the 4th polynomial order are in better agreement with the QA4ECV products (Fig. S2). Both correlation coefficient and regression slope are improved although OMI GEMS HCHO values are higher than those of the QA4ECV. We use the 4th order polynomial instead of the 5th order used in the QA4ECV products because slant columns retrieved using the 5th order in the GEMS algorithms are much higher than the QA4ECV products.

Also, different O_3 absorption cross sections (Serdyuchenko et al., 2014) are used in the OMI QA4ECV at different temperatures (220 and 243 K), and a non-linear O_3 absorption effect (Puķīte et al., 2010) is included in the OMI QA4ECV. We examine the O_3 effects on retrieved slant columns in GEMS algorithm using O_3 datasets used in QA4ECV and considering a non-linear O_3 absorption effect. Correlation coefficient and regression slopes are slightly improved (Table S1), and relative differences in the four regions defined above are slightly reduced in most seasons and regions (Table S2).

Table S1. Spatial correlation coefficients and slopes between OMI GEMS and OMI QA4ECV. Left values are statistics in Fig. 10, and right values are statistics of OMI GEMS using O_3 datasets used in QA4ECV and considering non-linear O_3 absorption effects.

Statistics	OMI GEMS vs. OMI QA4ECV			
	Mar.	Jun.	Sep.	Dec.
R	0.76 / 0.75	0.66 / 0.70	0.64 / 0.67	0.52 / 0.54
Slope	0.92 / 1.02	0.76 / 0.82	0.85 / 0.91	0.79 / 0.84

Table S2. Relative differences between OMI GEMS HCHO slant columns and OMI QA4ECV

slant columns in four regions. Left values are relative differences in Table 3 and right values are relative differences of OMI GEMS using O_3 datasets used in QA4ECV and considering non-linear O_3 absorption effects.

Region	OMIGEMS vs. OMI QA4ECV			
	Mar.	Jun.	Sep.	Dec.
Sumatra/Malaysia (95°-110°E, 0°-7°N)	-0.5% / 3%	-18% / -17%	-6% / -4%	-15% / -13%
Indochina Peninsula (97°-110°E, 10°-20°N)	-7% / -3%	-20% / -18%	-20% / -15%	-17% / -12%
China (110°-120°E, 30°-40N)	-21% / -25 %	-25% / -20%	-20% / -14%	-23% / -23%
Borneo (110°-118°E, 5°S-0°)	-9% / -5%	-13% / -9%	-0.4% / 5%	-18% / -16%

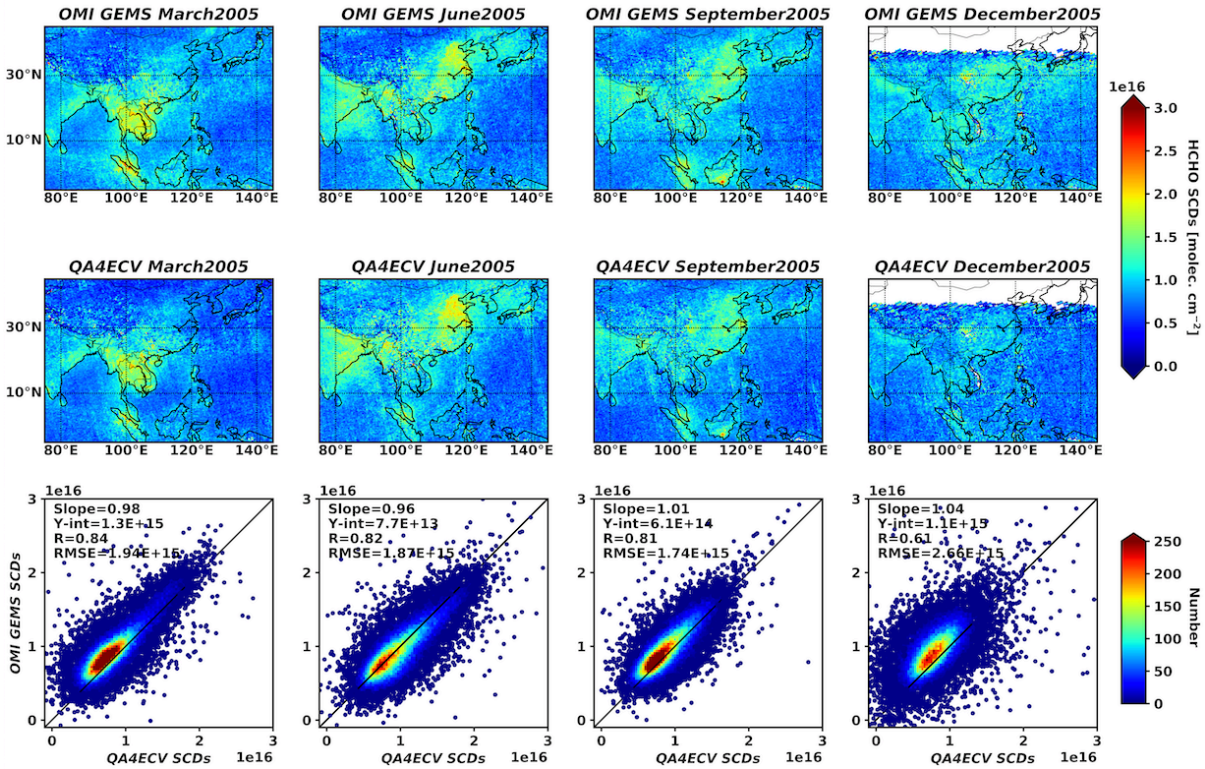

Figure S2. The same as Fig.10 but OMI GEMS products are retrieved using the 4th order polynomial instead of the 3rd order polynomial in default fitting options.