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.

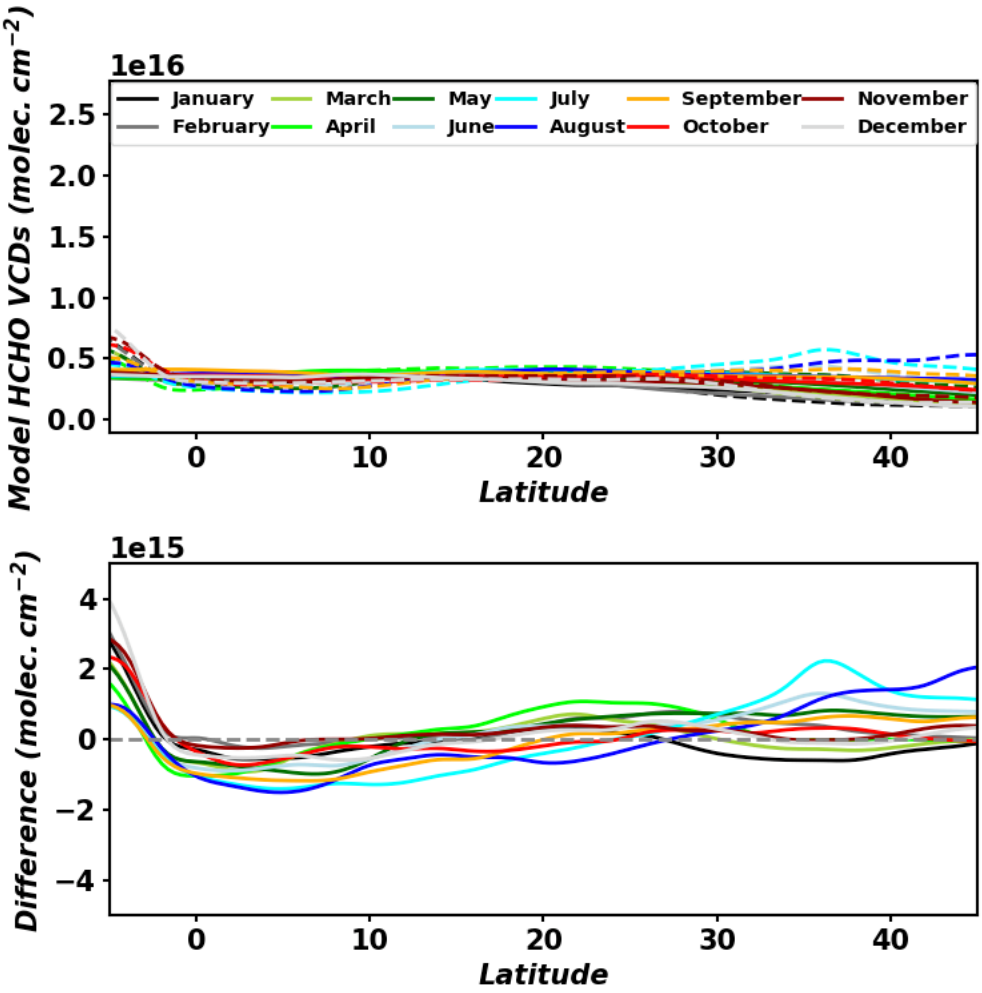


Figure S1. Simulated HCHO vertical column densities in GEMS background area (dashed lines) and OMI background area (solid lines) (top), and absolute differences between the two (bottom).

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:

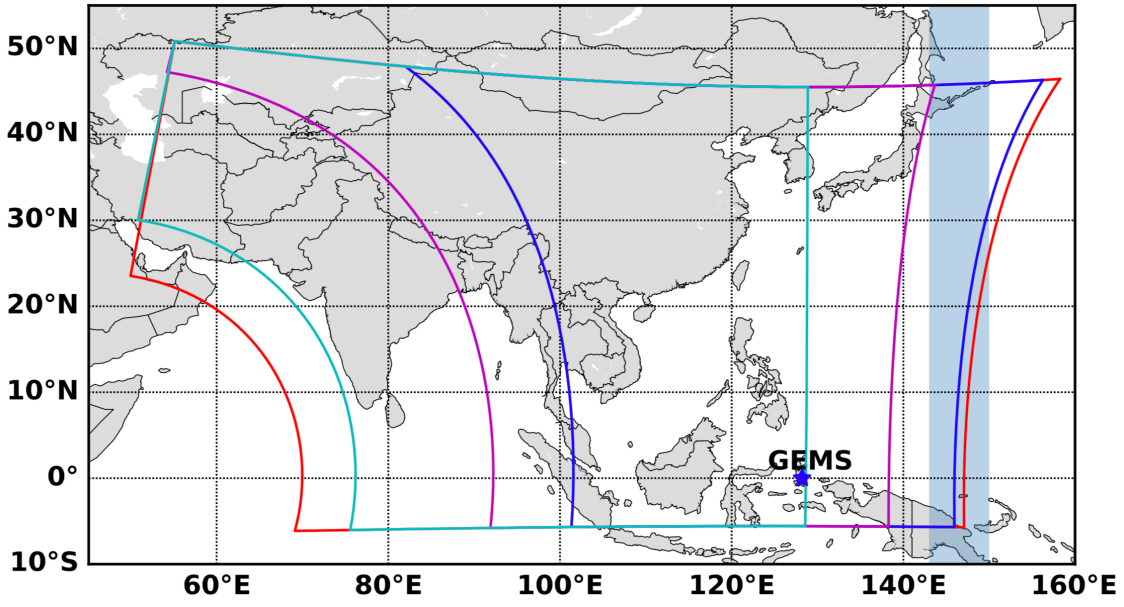


Figure 1. GEMS field of regard (red), nominal daily scan (blue), full central scan (magenta), full western scan (cyan), and GEMS location (blue star). Shaded areas (143-150°E) are regions for radiance references and common mode.

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%

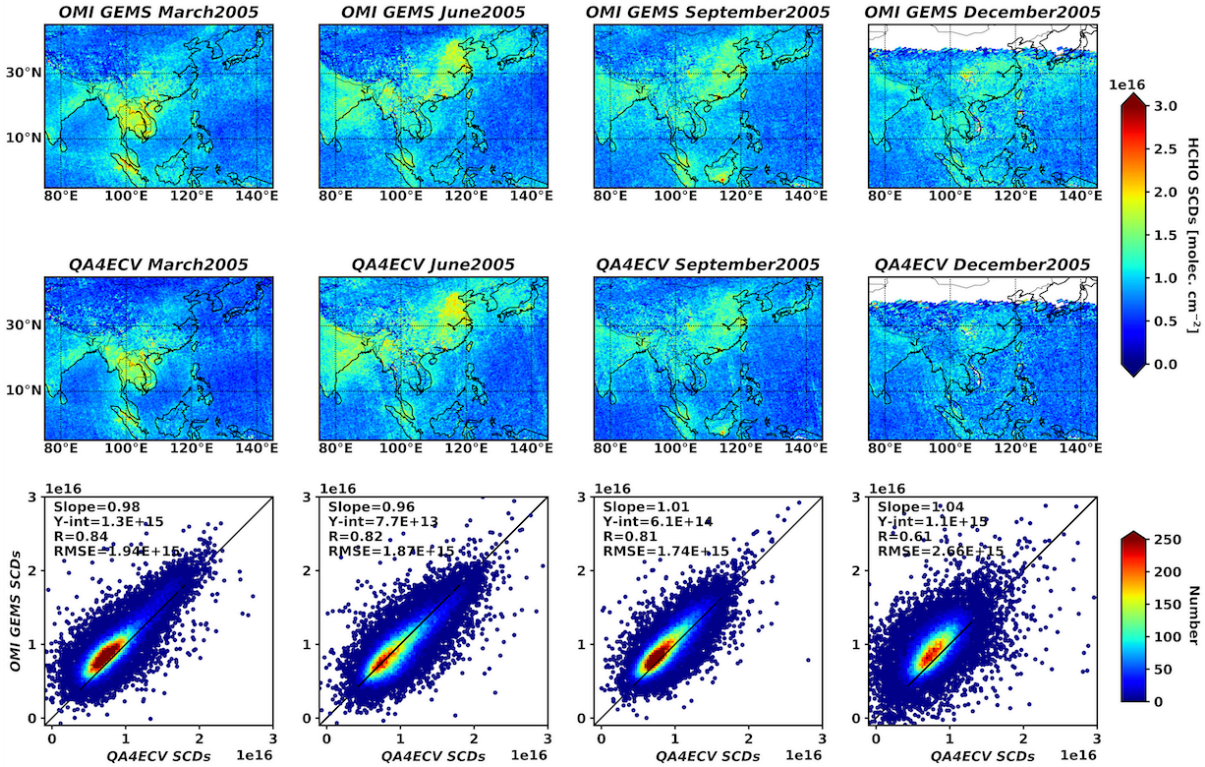


Figure S2. The same as Fig.10 but OMI GEMS products are retrieved using the 4th order polynomial instead of the 3rd order polynomial in default fitting options.

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 #2:

This is a useful and timely manuscript on the algorithm for HCHO retrievals with the GEMS geostationary sensor that will observe the atmosphere in the near future over eastern Asia. It is useful because the algorithm is discussed in a step-by-step manner, and a thorough uncertainty assessment is included, and a comparison to independent data is provided. The discussion of the systematic component of the uncertainty is very strong. It is timely because the launch of GEMS is imminent, and the community would like to learn how retrievals are different or better than what we know from OMI and TROPOMI. I recommend publication of the paper after the following issues are accounted or considered for.

Major issues

1. The paper focuses on testing the retrieval algorithm for OMI-type viewing conditions. It therefore remains unclear how the GEMS HCHO retrieval approach will account for diurnally varying measurement conditions. Surface reflectivity, HCHO profile shape, clouds will all change throughout the day, and it is unclear how these changes will affect the retrieval and their uncertainties. This is a major hiatus in this paper should be addressed.

Thanks for your suggestions about our weakness. As you mentioned, HCHO products can be sensitive to diurnally varying parameters such as surface reflectivity, HCHO profile shape, and clouds. Even though these parameters can affect the radiance fitting, we can estimate effects of variations of the parameters on AMF. Uncertainty of AMF as a function of the parameters was discussed in Section 3.2.

To examine sensitivity of AMF to HCHO profile height, we added Fig. 5d.

In addition, we define a profile height parameter (p_h) as an altitude below which 75% of HCHO VCDs exist from the surface, to estimate AMF uncertainty with respect to a HCHO profile shape.

…

Figure 5d shows increasing AMF values with an increase in the profile height, resulting from increased HCHO absorptions at high altitudes. The AMF sensitivity to profile heights in clean areas is higher than that in polluted areas because HCHO distributions are more uniform in clean areas than polluted areas.

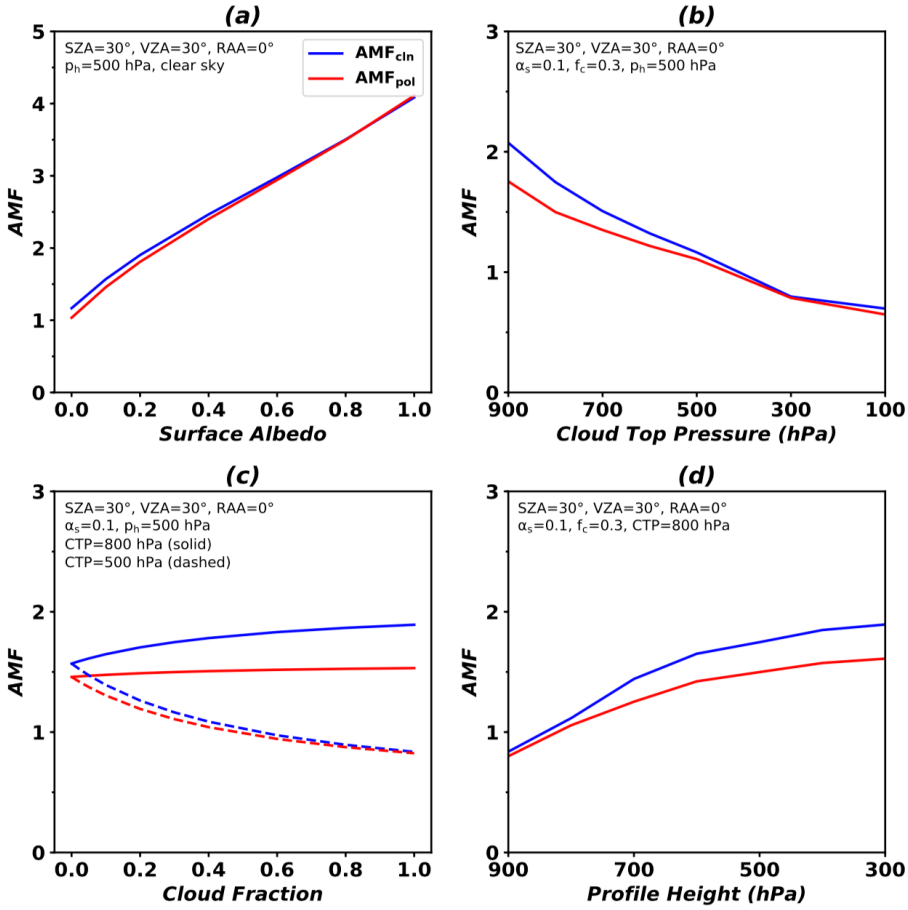


Figure 5. AMF variations as functions of (a) surface albedo, (b) cloud top pressure (CTP), (c) effective cloud fraction (f_c), and (d) profile height over clean (blue) and polluted (red) areas. Conditions of the AMF LUT are given in the figures. For sensitivity to surface albedo, cloud-free conditions are assumed. For sensitivity to cloud fraction, cloud top pressures are 800 hPa (solid line) and 500 hPa (dashed line).

We deleted Fig. 7 because it is too confusing to explain the contributions of parameters. Instead, 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 can be as large as 20% and 24% 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.

HCHO VCD uncertainty due to AMF 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%

We additionally discussed how to consider diurnally varying parameters in the GEMS in detail.

Surface albedo, effective cloud fraction, and cloud top pressure are retrieved from GEMS and are used in the AMF calculations. GEMS Level 2 surface properties include Lambertian equivalent reflectivity (LER) and the daily bidirectional reflectance distribution function (BRDF) (Lee and Yoo, 2018). GEMS LER products are retrieved as composites of minimum LER values for 15 days every hour with fixed viewing geometry so that geometry dependent LER are yielded. The effective cloud fraction and cloud top pressure (effective cloud pressure) are retrieved from GEMS with the assumption of a Lambertian cloud surface (cloud surface albedo = 0.8) (Veefkind et al., 2016). GEMS surface reflectivity products are also used for cloud retrievals. In addition, the radiative cloud fraction (f_{rc}) will be provided from GEMS Level 2 cloud products, and is defined by Eq. 9, where I_{cld} and I_{clr} are radiances over cloud and cloud-free surfaces, respectively.

…

However, the horizontal resolution of 2° × 2.5° for HCHO profiles in AMF LUT is much coarser than the GEMS horizontal resolution of 7 × 8 km^2 to discern spatial variations by local source emissions. HCHO profiles in AMF LUT are monthly averaged so that hourly variations are not accounted for. In order to resolve these rough conditions, we can use HCHO profiles with a finer resolution as a function of time. For example, Kwon et al. (2017) showed that HCHO retrievals using monthly mean hourly AMF values were in better agreement with the model simulations in observation system simulation experiments (OSSE) than those using monthly mean AMF values. Also, air quality forecasting data can be used to consider hourly varying HCHO profiles. Further studies are required to examine the dependency of AMF calculations on spatial resolutions and temporal variations of HCHO profiles and its effect on GEMS retrieval.

2. Even though the instrument is still to be launched, the paper should give more information on the GEMS instrument and how its data will be explored. What is the anticipated signal-to-noise for the HCHO spectral window, or how would it compare to OMI and TROPOMI?

Requirements of the signal-to-noise ratio for GEMS are greater than 720 at 320 nm and 1500 at 430 nm for natural spatial resolutions of 3.5 × 8 km^2. Required signal-to noise-ratios of OMI are 1450 in 335-365 nm, 700 in 365-420 nm, and 2600 in 420-450 nm for spatial resolution of 13 × 24 km^2 (OMI L1B ATBD). Signal-to-noise ratios of TROPOMI are 800-1000 in 310-405 nm and 405-500 nm (Veefkind et al., 2012). GEMS signal-to-noise ratios are comparable with those of OMI and TROPOMI. We added sentences and updated Table 1 as follows:

Requirements of signal-to-noise ratio for GEMS are 720 and 1500 at 320 and 430 nm, respectively, for natural spatial resolutions (3.5 × 8 km^2 over Seoul). However, pixels are co-added in order to increase signal-to-noise ratio, and GEMS will provide spatial resolutions of 7 × 8 km^2 or less over Seoul, South Korea for trace gases.

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.

How will the cloud retrieval from GEMS work? What surface reflectivity data will be used for the cloud and HCHO retrievals? How does the GEMS team address the issue of viewing geometry dependent surface reflectivity? These issues are not discussed, and thus the paper runs the risk of being read as just another OMI HCHO approach, i.e. of little specificity to GEMS.

Thanks for your comments. Effective cloud fraction and cloud top pressure (effective

cloud pressure) from GEMS will be retrieved by using O_4 absorption band with the assumption of Lambertian surface reflectors. Surface reflectance is provided as Lambertian equivalent reflectivity (LER) from GEMS Level 2 surface properties and is used for cloud and HCHO retrievals. GEMS LER products are retrieved as composites of minimum LER values for 15 days every hour with fixed viewing geometry so that geometry dependent LER is yielded.

We briefly referred to input parameters provided from GEMS Level 2 for AMF calculation and added references as follows:

Surface albedo, effective cloud fraction, and cloud top pressure are retrieved from GEMS and are used in the AMF calculations. GEMS Level 2 surface properties include Lambertian equivalent reflectivity (LER) and the daily bidirectional reflectance distribution function (BRDF) (Lee and Yoo, 2018). GEMS LER products are retrieved as composites of minimum LER values for 15 days every hour with fixed viewing geometry so that geometry dependent LER are yielded. The effective cloud fraction and cloud top pressure (effective cloud pressure) are retrieved from GEMS with the assumption of a Lambertian cloud surface (cloud surface albedo = 0.8) (Veefkind et al., 2016). GEMS surface reflectivity products are also used for cloud retrievals. In addition, the radiative cloud fraction (f_{rc}) will be provided from GEMS Level 2 cloud products, and is defined by Eq. 9, where I_{cld} and I_{clr} are radiances over cloud and cloud-free surfaces, respectively.

3. The paper would be strengthened if the authors would provide a breakup of the uncertainty budget for typical polluted and clean conditions, e.g. in the form of a table.

We summarized the uncertainties of HCHO VCD due to AMF uncertainties for polluted and clean regions in Table 2 and 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%

4. I'm missing a discussion of the GEOS-Chem 2x2.5 a priori profile shapes. These are much coarser than the 7x8 km2 viewing scenes, and this will result in a substantial AMF uncertainty. It is true that this can be accounted for via application of the averaging kernels, or by recomputing the AMFs with high-resolution profiles from a regional CTM or zoom-version of GEOS-Chem. In any case this issue should be discussed in more detail, and also included in the uncertainty budget.

We defined a profile height parameter (p_h) as an altitude below which 75% of HCHO VCDs exist from the surface to estimate AMF uncertainty with respect to a HCHO

profile shape. We included AMF uncertainty with respect to a profile height. We discussed as follows:

The AMF uncertainty can be estimated by each parameter in Eq. 16. We examine AMF uncertainties for surface albedo (α_s), cloud top pressure (p_c), and effective cloud fraction (f_c) with a solar zenith angle of 30°, a viewing zenith angle of 30°, and a relative azimuth angle of 0°. In addition, we define a profile height parameter (p_h) as an altitude below which 75% of HCHO VCDs exist from the surface, to estimate AMF uncertainty with respect to a HCHO profile shape. The uncertainties of parameters ($\sigma_{\alpha_s} = 0.02$, $\sigma_{p_c} = 50\ hPa$, and $\sigma_{f_c} = 0.05$) are based on De Smedt et al. (2018) and will be replaced to those from GEMS Level 2 products. The uncertainty of profile height (σ_{p_h}) is defined as a standard deviation of profile heights in AMF LUT, and σ_{p_h} in polluted and clean areas are 84 and 55 hPa, respectively.

$$\sigma_{AMF}^2 = \left(\frac{\partial AMF}{\partial \alpha_s}\sigma_{\alpha_s}\right)^2 + \left(\frac{\partial AMF}{\partial p_c}\sigma_{p_c}\right)^2 + \left(\frac{\partial AMF}{\partial f_c}\sigma_{f_c}\right)^2 + \left(\frac{\partial AMF}{\partial p_h}\sigma_{p_h}\right)^2 \quad (16)$$

…

Figure 5d shows increasing AMF values with an increase in the profile height, resulting from increased HCHO absorptions at high altitudes. The AMF sensitivity to profile heights in clean areas is higher than that in polluted areas because HCHO distributions are more uniform in clean areas than polluted areas.

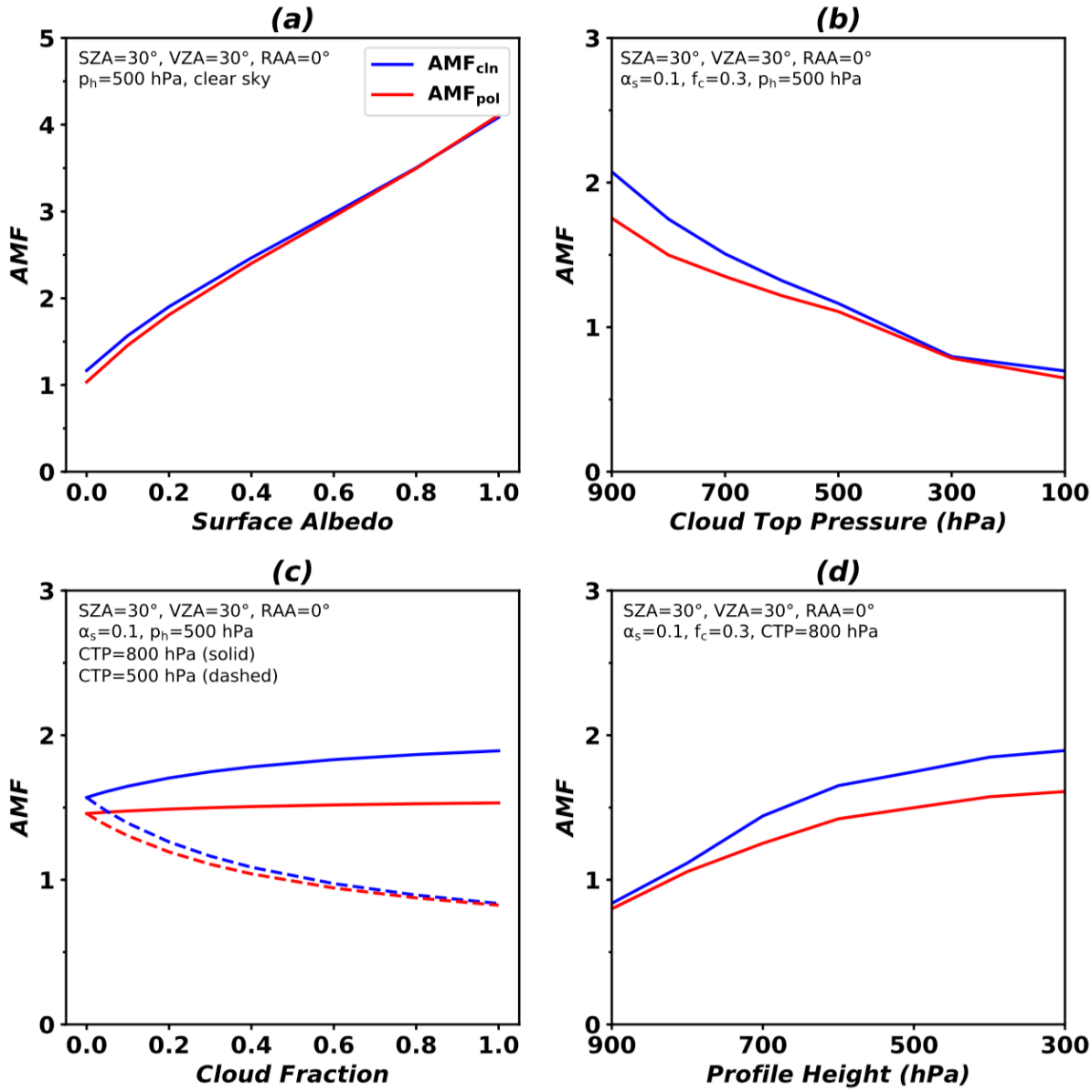


Figure 5. AMF variations as functions of (a) surface albedo, (b) cloud top pressure (CTP), (c) effective cloud fraction (f_c), and (d) profile height over clean (blue) and polluted (red) areas. Conditions of the AMF LUT are given in the figures. For sensitivity to surface albedo, cloud-free conditions are assumed. For sensitivity to cloud fraction, cloud top pressures are 800 hPa (solid line) and 500 hPa (dashed line).

5. It remains very much unclear how the latitude-bias is being determined. The text on page 11 (lines 9-11) is not clear, and the patterns shown in Figure 5(d) need explanation.

Latitudinal bias is determined by retrieved slant columns for radiance references.

Figure 5d (Figure 6d in the new manuscript) shows latitudinal bias, which is equal to averaged slant columns for radiance references as a function of latitude.

We modified sentences 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.

…

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.

Minor comments

P2, L12: suggest to remove 'instrument' after SCIAMACHY.

We removed it.

P3, L15: suggest to use air quality in the singular

We changed the in the singular.

P3, L16-17: compared to TROPOMI's 7x7 km2 pixels, the 7x8 km2 resolution from GEMS is not that superior, so I suggest to nuance that statement.

We modified the sentence as follows:

Instruments on-board these geostationary satellites have good spatial resolutions corresponding to those of TROPOMI and high signal-to-noise ratios, …

P6: eq. (3) and (4) – suggest to use that mathematical e rather than exp which reads as computer code.

We changed "exp" into "e" in other equations (Eq. (5) and (6)) as well as Eq. (3) and (4).

Eq. (15) appears wrong.

We corrected Eq. 15 as follows:

$$\sigma_{s,j}^2 = RMS^2 \frac{m}{m-n} C_{j,j} C_{j,j}, \quad (15)$$

Figure 7: what are the relative uncertainties in the AMF?

We deleted Figure 7 but add Table 2 to describe retrieval uncertainties of GEMS HCHO VCDs due to AMF uncertainties. We discussed Table 2 above.

Validation: which spatio-temporal selection criteria were used?

For comparison with OMI other products, we used monthly averages weighted by fitting uncertainties and overlapped areas between pixels and grid boxes with a horizontal resolution of 0.25° × 0.25°.

For MAX-DOAS comparison, we also used the weighted monthly averages for OMI in a grid box of 0.25° at the center of site locations, and MAX-DOAS data were arithmetically averaged within OMI overpass time (12:00-15:00 local time). We updated comparisons between MAX-DOAS and OMI products for a year at OHP and Bremen in 2005 and at Xianghe in 2016.

We modified paragraphs as follows:

We also compare satellite results with MAX-DOAS ground observations at Haute-Provence Observatory (OHP) in France, Bremen in Germany, and Xianghe in China (Table 4). MAX-DOAS data are collected within the OMI overpass time (12:00–15:00

local time) at OHP and Bremen in 2005, and at Xianghe in 2016, respectively. We collect OMI data pixels that are overlapped by a grid box of 0.25° at the center of the site location, and average values of OMI data are weighted by uncertainties and overlapped areas between pixels and grid boxes.

Comparisons of HCHO VCDs between MAX-DOAS and satellite products are shown in Fig. 11 and Table 4. Averaged MAX-DOAS HCHO VCDs for a year are 7.6×10^{15}, 6.7×10^{15}, and 1.6×10^{16} molecules cm^{-2} at OHP, Bremen, and Xianghe, respectively. HCHO VCDs show a seasonal variation with the maximum concentrations in summer at all sites (Fig. S3). The largest monthly change is shown at Xianghe, likely driven by abundant VOC precursors for HCHO productions compared to OHP and Bremen.

Averaged HCHO VCDs from OMI GEMS are by 16%, 9%, 25% lower than those from MAX-DOAS at OHP, Bremen, and Xianghe. At Bremen, HCHO VCDs from the GEMS algorithm are in the best agreement with those of MAX-DOAS and show similar monthly variations with MAX-DOAS. OMI GEMS results at Xianghe show a monthly variation but at OHP do not show monthly variation despite of a bit increment in summer. In particular, the GEMS algorithm yields lower HCHO VCDs in summer. These lower values may be caused by the a priori HCHO profiles used in AMF calculation. In summer, HCHO is produced and concentrated near the surface, which results in lower AMFs (higher VCDs). S. W. Kim et al. (2018) showed the anti-correlation between AMF values and the HCHO mixing ratios at 200 m above ground level. OMHCHO products show similar tendencies as OMI GEMS, but they are much lower than those of OMI GEMS. OMI QA4ECV products are higher than MAX-DOAS at OHP and Bremen but are in the best agreement with MAX-DOAS at Xianghe compared to other satellite products.

Table 4. Averaged HCHO VCDs (molecules cm^{-2}) from MAX-DOAS ground observations and OMI satellite data at OHP in France, Bremen in Germany, and Xianghe in China. For satellites, mean values are weighted by uncertainties and overlapped areas between satellite pixels and 0.25° grid cells for each site. Relative differences between OMI and MAX-DOAS are given in parentheses.

Site[a]	Class[b]	MAX-DOAS[c]	OMHCHO	OMI QA4ECV	OMI GEMS

OHP (44°N, 5.5°E)	Rural	7.5×10^{15}	5.8×10^{15} (-24%)	1.1×10^{16} (50%)	6.3×10^{15} (-16%)
Bremen (53°N, 9°E)	Urban	6.7×10^{15}	5.1×10^{15} (-23%)	9.3×10^{15} (40%)	6.1×10^{15} (-9%)
Xianghe (39°N, 117°E)	Sub-urban	1.6×10^{16}	1.0×10^{16} (-37%)	1.7×10^{16} (4%)	1.2×10^{16} (-25%)

[a] HCHO VCDs are averaged at OHP and Bremen in 2005 and at Xianghe in 2016.

[b] Class is assigned in a QA4ECV MAXDOAS website (http://uv-vis.aeronomie.be/groundbased/QA4ECV_MAXDOAS)

[c] Fitting windows of 336–359 nm and 324–359 nm are used at OHP and Bremen, and at Xianghe, respectively.

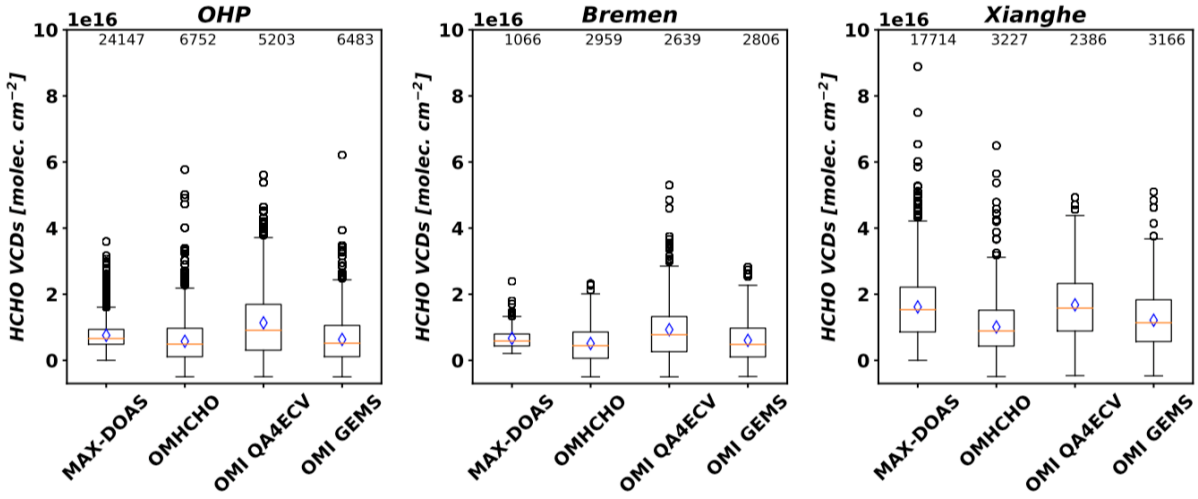


Figure 11. HCHO vertical columns from MAX-DOAS, OMHCHO, OMI QA4ECV, and OMI GEMS at OHP and Bremen in 2005, and at Xianghe in 2016. Orange lines are median values for each dataset, and blue diamonds are mean values. We computed mean values of each satellite product weighted by uncertainties and overlapped areas between satellite pixels and 0.25° grid cells for each site. Boundaries of boxes indicate first and last quantiles of datasets.

Reference

OMI L1B ATBD, https://projects.knmi.nl/omi/documents/data/OMI_ATBD_Volume_1_V1d1.pdf

Veefkind, J. P., et al. (2012). "TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications." Remote Sensing of Environment **120**: 70-83.

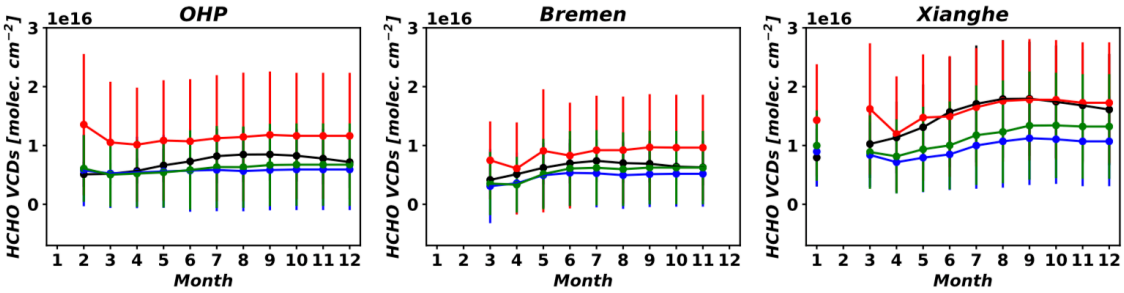


Figure S3. Monthly variations of HCHO VCDs from MAX-DOAS (black) and OMI at OHP, Bremen, and Xianghe. Blue indicates OMHCHO, red indicates QA4ECV, and green indicates GEMS.

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 #3:

General Description:

The authors describe the retrieval algorithm of formaldehyde (HCHO) for the future GEMS instrument and estimate the likely uncertainties and biases relative to OMI and ground-based MAX-DOAS measurements. The content is appropriate for AMT. Suggested changes, comments and concerns are included below.

General Comments:

It's not clear what's unique about the retrieval to GEMS. Seems more like a recapitulation of the OMI retrieval description paper of González Abad et al. (2015). A way to address this would be to assess the implication of the unique temporal component of GEMS (i.e. observations throughout the day) on uncertainties in the retrieval.

Thanks for suggestions. We analyzed expected random uncertainty for GEMS by using simulated radiances, which are convoluted with GEMS bandpass functions at 330 nm and include noises based on signal-to-noise ratio for co-added pixels with spatial resolutions of 7 × 8 km^2. We updated related paragraphs as follows:

We analyze expected uncertainties for the GEMS algorithm by using simulated radiances from Kwon et al. (2017) and OMI Level 1B data. In order to estimate the expected random uncertainty for GEMS (Section 3.1.1), we use simulated radiances, which are convoluted with GEMS bandpass functions at 330 nm as a function of cross-track positions in the south to north direction. Simulated radiances include noises based on the expected signal-to-noise ratio for co-added pixels with spatial resolutions of 7 × 8 km^2. We use absorption cross-sections of Ring effect, O_3, NO_2, HCHO, and additionally SO_2 (Hermans et al.,

2009; Vandaele et al., 2009) in radiance fitting because O_3, NO_2, and HCHO, and SO_2 were considered in radiance calculation (Kwon et al. 2017).

For other uncertainty analyses, we use OMI Level 1B data with OMI slit function data (Dirksen et al., 2006) in order to examine algorithm sensitivities to individual parameters. Fitting options such as absorption cross-section data and the fitting window are summarized in Table 1. It will be necessary to conduct an additional uncertainty analysis for GEMS HCHO retrievals after GEMS is launched.

…

Random uncertainties from the GEMS algorithm are estimated using simulated radiances. RMS of fitting residuals and random uncertainty for the GEMS domain range from 2.9×10^{-4} to 2.1×10^{-3} and 2.1×10^{15} to 1.6×10^{16} molecules cm^{-2}, respectively, which are comparable with those (RMS: 4×10^{-4} to 2.0×10^{-3}; random uncertainty: 3.3×10^{15} to 1.8×10^{16} molecules cm^{-2}) obtained from the GEMS algorithm using OMI Level 1B data. GEMS measures target species every hour in daytime so that changes of solar location for a day can affect the accuracy of radiance fitting. An averaged fitting RMS value and a random uncertainty are 6.9×10^{-4} and 5.0×10^{15} molecules cm^{-2} for conditions with both solar and viewing zenith angles less than 70, which happen at 8:00–18:00 and 9:00–16:00 local time of Seoul in summer and winter, respectively. However, the fitting RMS value and the random uncertainty increase to 1.1×10^{-3} and 8.2×10^{15} molecules cm^{-2}, respectively, when solar and viewing zenith angles are higher than 70.

To clarify, we remained descriptions related with GEMS in the Section 2. Descriptions related with OMI to validate the GEMS algorithm were moved to new Section 4.1.

We described a radiance reference for GEMS in Section 2.2.3 as follows:

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.

We described GEMS surface reflectivity and cloud information used for AMF calculation in Section 2.2.4.

Surface albedo, effective cloud fraction, and cloud top pressure are retrieved from GEMS and are used in the AMF calculations. GEMS Level 2 surface properties include Lambertian equivalent reflectivity (LER) and the daily bidirectional reflectance distribution function (BRDF) (Lee and Yoo, 2018). GEMS LER products are retrieved as composites of minimum LER values for 15 days every hour with fixed viewing geometry so that geometry dependent LER are yielded. The effective cloud fraction and cloud top pressure (effective cloud pressure) are retrieved from GEMS with the assumption of a Lambertian cloud surface (cloud surface albedo = 0.8) (Veefkind et al., 2016). GEMS surface reflectivity products are also used for cloud retrievals. In addition, the radiative cloud fraction (f_{rc}) will be provided from GEMS Level 2 cloud products, and is defined by Eq. 9, where I_{cld} and I_{clr} are radiances over cloud and cloud-free surfaces, respectively.

Also, we wrote a plan to consider temporal variations of a priori HCHO profiles as follows:

However, the horizontal resolution of 2° × 2.5° for HCHO profiles in AMF LUT is much coarser than the GEMS horizontal resolution of 7 × 8 km^2 to discern spatial variations by

local source emissions. HCHO profiles in AMF LUT are monthly averaged so that hourly variations are not accounted for. In order to resolve these rough conditions, we can use HCHO profiles with a finer resolution as a function of time. For example, Kwon et al. (2017) showed that HCHO retrievals using monthly mean hourly AMF values were in better agreement with the model simulations in observation system simulation experiments (OSSE) than those using monthly mean AMF values. Also, air quality forecasting data can be used to consider hourly varying HCHO profiles. Further studies are required to examine the dependency of AMF calculations on spatial resolutions and temporal variations of HCHO profiles and its effect on GEMS retrieval.

Throughout, use the standard symbol ⊗ for convolution. This will help clarify terms in equations that are confusing, as brackets are used to denote dependence, but also operators, e.g. $f \otimes g(\lambda)$ to replace $(f * g)(\lambda)$ in Equation (2) is clearer. Please correct these issues throughout.

Thanks for suggestion. We replaced the symbol * to the symbol ⊗ in Eq. (1)-(5) as follows:

$$I_R(\lambda) = I_0^h \otimes g(\lambda + \Delta\lambda) P_{sc}(\lambda) + P_{bl}(\lambda), \quad (1)$$

$$f \otimes g(\lambda) = \int_{-\infty}^{\infty} f(\Lambda) g(\lambda - \Lambda) d\Lambda \quad (2)$$

$$\text{attenuated radiance in radiance fitting} = I_0^h \otimes g(\lambda) e^{-\tau^h \otimes g(\lambda)}, \quad (3)$$

$$\text{attenuated radiance in reality} = (I_0^h(\lambda) e^{-\tau^h(\lambda)}) \otimes g(\lambda). \quad (4)$$

$$\sigma_{ps}(\lambda) = \frac{1}{scd_{ref}} \ln\left(\frac{I_0^h \otimes g(\lambda)}{\left(I_0^h(\lambda) e^{-scd_{ref}\sigma^h(\lambda)}\right) \otimes g(\lambda)}\right), \quad (5)$$

Inconsistent use of wavelength dependence in equations. For example, why do I_R and I_0^h not depend on wavelength in Equations (1)-(4), but do in Equation (5)?

We changed those equations above, and we also modified variables related with Eq. (6) as a function of wavelength.

$$I(\lambda) = \left[\left(aI_0(\lambda) + c_r\sigma_r(\lambda)\right)e^{-\sum_i SCD_i\sigma_i(\lambda)} + c_{cm}\sigma_{cm}(\lambda)\right]P_{sc}(\lambda) + P_{bl}(\lambda), \quad (6)$$

Many sub-sections in Section 2.2. are the same as in González Abad et al. (2015). Why not just refer the reader to that paper and only state aspects specific to GEMS and that are different between the two approaches?

As we answered to first comments, we remained descriptions related to GEMS. We also added new sub-section 4.1 to describe fitting options in the GEMS algorithm for OMI HCHO retrievals.

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's not clear why Section 2.2.5 is relevant, as it describes bias corrections specific to OMI. Is it anticipated that the same bias corrections will be needed for GEMS? If this section is relevant, the readers could just be referred to González Abad et al. (2015) and this section be kept brief.

Thanks for your comments. For GEMS, background corrections are only used when we use a radiance reference. To clarify, therefore, we explained background corrections for GEMS in Section 2.2.5, and corrections and discussions for OMI were moved to Section 4.1.

We modified paragraphs in Section 2.2.5 as follows:

An alternative method to avoid the above-mentioned biases in the fitting procedure is to

use measured radiances over a clean background region (referred to as radiance references) as the reference spectrum in radiance fitting. As measured radiance includes instrument noise and attenuation by interfering gases in the background atmosphere, the interfering effects can be minimized in radiance fitting, which results in negligible cross-track biases. 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.

The retrieved slant columns using a radiance reference are differential slant columns ($\Delta SCD = SCD - SCD_0$) and do not include background HCHO columns (SCD_0) that are mainly from the oxidation of methane. To account for the background columns, we use HCHO vertical columns simulated in 2014 from a chemical transport model, GEOS-Chem (Bey et al., 2001) with a spatial resolution of 2° × 2.5°. Simulated HCHO vertical columns are zonally and monthly averaged over the reference sectors and are interpolated to 720 latitudinal grid points with a resolution of 0.25° from 90°S to 90°N.

In order to account for dependency of measured radiances on geometric angles, we then convert simulated background vertical columns into slant columns by applying AMF values over the reference sector (AMF_0), which are calculated with cloud information and geometric angles on the reference sectors. Corrected GEMS HCHO slant columns are formulated as the sum of the retrieved differential slant columns and the simulated background slant columns as shown in Eq. 11,

$$\Omega_s(i,j) = SCD_{corr}(i,j) = \Delta SCD(i,j) + AMF_0(lat)VCD_m(lat), \quad (11)$$

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. We finally apply AMF values from the LUT to the corrected slant columns to obtain GEMS HCHO vertical column densities.

It's also not clear why data quality flags are provided for a future product. This would only be important for the user when the data is ready for release.

The data quality flag is provided for basic information of data quality in radiance fitting, and we followed the flag definition from González Abad et al. (2015). We are planning to provide flags including much information such as geometry angles, clouds, surface information.

Section 3 appears to just be testing uncertainties inherent in fitting parameters and retrieval terms that would be an issue for all space-based instruments measuring HCHO, rather than being specific to GEMS. Is there anything unique to GEMS (instrument configuration, viewing domain, repeat time etc.) that would increase or decrease sensitivity to these uncertainties relative to other instruments?

Uncertainty related to GEMS instrument is considered in random uncertainty. Random uncertainty, called fitting uncertainty, is calculated from fitting residuals caused by instrument noise, radiance measurement uncertainty from dark current and stray lights, and polarization. We estimated expected random uncertainty by using simulated radiances with GEMS bandpass functions and signal-to-noise ratios. We discussed it in the first answer.

GEMS does not include a polarization scrambler while OMI and TROPOMI include a polarization scrambler. In the operation, polarization correction will be conducted when L1B data are produced. The correction could minimize polarization, but it would

not be perfect. The effects could increase random uncertainty. We need to have a process to minimize polarization.

We discussed it in Section 5 as follows:

We currently use a broad fitting window (328.5–356.0 nm). However, we may need to use a different fitting window to reduce interference from polarization effects because GEMS does not include a polarization scrambler. A polarization correction is planned to minimize its interference during GEMS Level 1B production, but we need to examine the retrieval sensitivity to polarization.

Specific Comments:

P2, Line 18: the spatial resolution of TROPOMI is finer than 7 x 7 km2 for HCHO (De Smedt et al., 2018).

Veefkind et al. (2012) showed the spatial resolution of TROPOMI UVIS band 3 (310-405 nm) is 7 x 7 km^2. However, we found TROPOMI HCHO products are provided with 7 x 3.5 km^2. We corrected it from 7 x 7 km^2 to 7 x 3.5 km^2.

Equation (1): Why are *Psc* and *Pbl* not dependent on wavelength?

P_{sc} and P_{bl} are functions of wavelength. Therefore, we changed it as follows:

$$I_R(\lambda) = I_0^h \otimes g(\lambda + \Delta\lambda) P_{sc}(\lambda) + P_{bl}(\lambda), \quad (1)$$

P3, Line 22: Can aerosol optical properties be retrieved across this wavelength and for this type of instrument? Do the authors mean AOD and aerosol index (AI)?

We meant AOD and SSA. We clarified it as follows:

Geostationary Environment Monitoring Spectrometer (GEMS) will be launched by South Korea, and it will measure radiances ranging from 300 to 500 nm every hour with fine spatial resolutions of 3.5 × 8 km^2 for aerosols or 7 × 8 km^2 for gases over Seoul in South Korea to monitor column concentrations of air pollutants including O_3, NO_2, SO_2, and HCHO, and aerosol optical properties (aerosol optical depth and single scattering albedo).

Table 1: add references for these parameters as footnotes to point to consistency with existing retrievals.

We marked 'a' and 'b' on datasets used in OMHCHO and QA4ECV, respectively, and explanation was written in footnotes as follows:

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

[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.

P8, Lines 14-17: What about clouds (Millet et al., 2006)?

We corrected the sentence as follows:

AMF uncertainties contribute to retrieval uncertainties by multiple factors including cloud, HCHO vertical distribution, aerosol vertical distribution, and aerosol optical properties (Millet et al., 2006; Chimot et al., 2016; Kwon et al., 2017; Hewson et al., 2015).

P18, Lines 8-14: Comment too on the implications of more observations over the same scene per day on uncertainty compared to OMI.

Thank your comments. We wanted to show an example of OMI HCHO results retrieved from the GEMS algorithm. Therefore, we showed HCHO optical depths and fitting residuals and explained averaged HCHO slant column density and random uncertainty. In addition, we explained slant columns and random uncertainties in pixels with low concentrations as follows:

Averaged 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}.

P18, Lines 26-28: Provide an appropriate reference for this statement.

We added a reference as follow:

Zhong, L., Louie, P. K. K., Zheng, J., Yuan, Z., Yue, D., Ho, J. W. K., and Lau, A. K. H.: Science–policy interplay: Air quality management in the Pearl River Delta region and Hong Kong, Atmospheric Environment, 76, 3-10,

https://doi.org/10.1016/j.atmosenv.2013.03.012, 2013

Referencing: some references are missing the doi number (e.g., González Abad et al., 2015).

DOI numbers are added as follows:

González Abad et al. (2015): 10.5194/amt-8-19-2015
Barkley et al. (2013): 118, 6849-6868, 10.1002/jgrd.50552, 2013
Bey et al. (2001): 10.1029/2001JD000807
Cantrell et al. (1990): 10.1021/j100373a008
Chance et al. (1997): 10.1364/AO.36.005224
Chance et al. (2000): 10.1029/2000GL011857
Daumont et al. (1992): 10.1007/BF00053756
De Smedt et al. (2008): 10.5194/acp-8-4947-2008
Hewson et al. (2013): 10.5194/amt-6-371-2013
Malicet et al. (1995): 10.1007/BF00696758
Marais et al. (2012): 10.5194/acp-12-6219-2012
Palmer et al. (2001): 10.1029/2000JD900772
Spurr (2006): 10.1016/j.jqsrt.2006.05.005
Zhu et al. (2014): 10.1088/1748-9326/9/11/114004

References:

González Abad et al., Atmos. Meas. Tech., 8, 19-32, 2015, doi:10.5194/amt-8-19-2015. De Smedt et al., Atmos. Meas. Tech., 11, 2395–2426, 2018, doi:10.5194/amt-11-2395-2018. Millet et al., J. Geophys. Res., doi:10.1029/2005JD006853, 2006

Description of a formaldehyde retrieval algorithm for the Geostationary Environment Monitoring Spectrometer (GEMS)

Hyeong-Ahn Kwon[1], Rokjin J. Park[1], Gonzalo González Abad[2], Kelly Chance[2], Thomas P. Kurosu[3], Jhoon Kim[4], Isabelle De Smedt[5], Michel Van Roozendael[5], Enno Peters[6,*], and John Burrows[6]

[1]School of Earth and Environmental Science, Seoul National University, Seoul, Republic of Korea
[2]Atomic and Molecular Physics Division, Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
[3]Earth Science, Jet Propulsion Laboratory, Pasadena, California, USA
[4]Department of Atmospheric Sciences, Yonsei University, Seoul, Republic of Korea
[5]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[6]Institute of Environmental Physics, University of Bremen, Bremen, Germany
*now at: DLR - Institute for Protection of Maritime Infrastructures, German Aerospace Center, Bremerhaven, Germany

Correspondence to: Rokjin J. Park (rjpark@snu.ac.kr)

Abstract. We describe a formaldehyde (HCHO) retrieval algorithm for the Geostationary Environment Monitoring Spectrometer (GEMS) that will be launched by the Korean Ministry of Environment in 2019. The algorithm comprises three steps: pre-processes, radiance fitting, and post-processes. The pre-processes include a wavelength calibration, and interpolation and convolution of absorption cross-sections; radiance fitting is conducted using a non-linear fitting method referred to as basic optical absorption spectroscopy (BOAS); and post-processes include air mass factor calculations and bias corrections. In this study, several sensitivity tests are conducted to examine the retrieval uncertainties using the GEMS HCHO algorithm. We evaluate the algorithm with the OMI Level 1B irradiance/radiance data by comparing our retrieved HCHO column densities with OMI HCHO products of the Smithsonian Astrophysical Observatory (OMHCHO) and of the Quality Assurance for Essential Climate Variables project (OMI QA4ECV). Results show that OMI HCHO slant columns retrieved using the GEMS algorithm are in good agreement with OMHCHO, with correlation coefficients of 0.77–0.91 and regression slopes of 0.94–1.04 for March, June, September, and December 2005. Spatial

Deleted: protection of maritime infrastructures

Deleted: for stripe patterns, background HCHO, and latitudinal biases.…

Deleted: Belgian Institute

Deleted: Space Aeronomy

Deleted: BIRA

Deleted: 92

distributions of HCHO slant columns from the GEMS algorithm are consistent with the OMI QA4ECV products, but relatively poorer correlation coefficients of 0.52 to 0.76 are found compared to those against the OMHCHO products. Also, we compare the satellite results with ground-based MAX-DOAS observations. OMI GEMS HCHO vertical columns are by 9–25% lower than those of MAX-DOAS at Haute-Provence Observatory (OHP) in France, Bremen in Germany, and Xianghe in China. We find that the OMI GEMS retrievals have less bias than the OMHCHO and OMI QA4ECV products at OHP and Bremen in comparison with MAX-DOAS.

Deleted: BIRA

Deleted: have

Deleted: 0.8–30

Deleted: BIRA

Deleted: for

1 Introduction

Formaldehyde (HCHO) is mainly produced by the oxidation of non-methane volatile organic compounds (NMVOCs), and it has been observed from space since the GOME instrument on ERS-2 satellite first began conducting column measurements in 1995 (Chance et al., 2000). The subsequent instrument, SCIAMACHY on ENVISAT, collected continuous HCHO column data for 2002–2012 (Wittrock et al., 2006), and GOME-2A and -2B instruments have been conducting measurements from 2007 to the present day (De Smedt et al., 2012; De Smedt et al., 2015). The OMI instrument was launched in 2004 and has provided global HCHO vertical column data with a higher spatial resolution of 13 × 24 km^2 at the nadir than that of former instruments. Furthermore, the TROPOMI equivalent to OMI has been offering consecutive data with an even finer spatial resolution of 7 × 3.5 km^2 in UVIS bands at the nadir since 2017. There are thus more than 20 years of HCHO column data available from these various instruments, which enable analyses to be conducted on the global changes of HCHO columns throughout this time period.

Deleted: instrument

Deleted: 7

All of the low-orbiting sun-synchronous satellite measurements of HCHO columns have played an important role in filling gaps for regions where limited (or no) in-situ measurements of HCHO have been made, and these measurements have been used to constrain top-down estimates of biogenic and anthropogenic emissions of NMVOCs (Marais et al., 2012; Barkley et al., 2013; Stavrakou et al., 2014; Zhu et al., 2014). Together with satellite glyoxal measurements, HCHO satellite measurements have been used to distinguish dominant VOC sources (e.g., biogenic vs. anthropogenic) (DiGangi et al., 2012; Vrekoussis et al., 2010). In addition, the ratio of HCHO to nitrogen dioxide (NO_2) columns has been used to determine NOx-limited or VOC-limited ozone production regimes (Martin et al., 2004; Duncan et al.,

Deleted: mentioned above

2010; Choi et al., 2012). Continuous HCHO column measurements from sun-synchronous satellites are thus invaluable for evaluating and monitoring NMVOC emission trends over long-term periods.

However, as sun-synchronous satellites have measurement frequencies of once or twice a day, they provide limited explorations of diurnal cycles and transboundary transport of air pollutants. Moreover, their coarse spatial resolutions make discerning local source emissions difficult. With the aim of overcoming the issues, Zhu et al. (2014) used the oversampling method and temporally averaged out pixels of OMI HCHO vertical columns over high-resolution grids of ~2 km, and Kim et al. (2016) developed a downscaling method for OMI NO_2 measurements by adopting the spatial distribution information from a regional air quality model. However, both these methods have inherent limitations: the former method involves a trade-off between spatial and temporal information and the latter method includes the uncertainties of emission distributions.

To tackle the limitations inherent in low-orbiting satellites measurements, environmental geostationary satellites will be launched in 2019 (or later) by South Korea and the United States, and in 2021 by European Union, to monitor air quality over East Asia, North America, and Europe, respectively. Instruments onboard these geostationary satellites have spatial resolutions corresponding well with those of TROPOMI and high signal-to-noise ratios, and they will conduct column measurements of air pollutants every hour during the daytime. The Geostationary Environment Monitoring Spectrometer (GEMS) will be launched by South Korea, and it will measure radiances ranging from 300 to 500 nm every hour with fine spatial resolutions of 3.5 × 8 km^2 for aerosols or 7 × 8 km^2 for gases over Seoul in South Korea to monitor column concentrations of air pollutants including O_3, NO_2, SO_2, and HCHO, and aerosol optical properties (aerosol optical depth and single scattering albedo). The measurements from GEMS will then be used in applications such as data assimilation of air quality forecasts and top-down constraints of air pollutant emissions.

This paper describes a GEMS retrieval algorithm for HCHO. It also presents an uncertainty analysis and an evaluation of the algorithm, which involves comparing OMI GEMS HCHO results with OMI HCHO products from the Smithsonian Astrophysical Observatory (OMHCHO) and those from the Quality Assurance for Essential Climate Variables (QA4ECV) project. In addition, OMI HCHO results are compared with those of MAX-DOAS ground observations. In Section 2, we describe the GEMS instrument and provide the theoretical basis for HCHO retrievals. In Section 3, sensitivity tests are conducted to examine the retrieval

Deleted: observations periods
Deleted: transports
Deleted: qualities
Deleted: on-board
Deleted: superior
Deleted: to
Deleted: sun-synchronous satellites
Deleted: higher
Deleted: a
Deleted: resolution
Deleted: .
Deleted: Belgian Institute for Space Aeronomy (OMI BIRA).

uncertainties; and in Section 4, we discuss an evaluation of HCHO results retrieved from the GEMS algorithm.

2 GEMS HCHO algorithm

2.1 GEMS instrument

GEMS is a scanning ultraviolet-visible spectrometer which will be launched by the Korean Ministry of Environment in 2019 onboard a geostationary satellite (GEO-KOMPSAT 2B), which also carries a Geostationary Ocean Color Imager-2 (GOCI-2). GEMS will be located at ~128.2°E near the equator and will cover East and Southeast Asia (5°S-45°N, 75°-145°E). The instrument will conduct hourly measurements during the day (8 times) over the whole domain. It will measure one swath from south to north and then turn a scan mirror from east to west using an imaging time of 30 minutes and a transmission time of 30 minutes to enable GOCI-2 measurements for a 30 min period.

Solar backscattered radiances will be measured in the 300–500 nm wavelength range with a spectral resolution of 0.6 nm and a wavelength interval of 0.2 nm. Signal-to-noise ratio requirements for GEMS are 720 and 1500 at 320 and 430 nm, respectively, for natural spatial resolutions (3.5 × 8 km^2 over Seoul). However, pixels are co-added in order to increase signal-to-noise ratio, and GEMS will provide spatial resolutions of 7 × 8 km^2 or less over Seoul, South Korea for trace gases. The field of regards (FOR) and information about GEMS are shown in Fig. 1 and Table 1, respectively. Detailed information about the GEMS instrument and algorithms for species other than HCHO are found elsewhere (Kim, J. et al., 2018; Kim, M. et al., 2018; Go et al., 2019).

2.2 HCHO algorithm description

Figure 2 is a flow chart of the HCHO retrieval algorithm for GEMS. The retrieval procedure consists of three steps: pre-processes, radiance fitting, and post-processes. The pre-processes begin with a wavelength calibration of Level 1B data (irradiance and radiance), and interpolation and convolution of absorption cross-sections at calibrated wavelength grid points. Radiance fitting is then conducted to derive the HCHO slant columns using a non-linear least square method. Finally, the post-processes include an air mass factor (AMF) calculation that employs a look-up table to convert HCHO slant columns to vertical columns, an assignment

Deleted: on-board

Deleted: scan

Deleted: GEMS will provide finer

Deleted: , compared to those of previous low-orbiting satellites (for example, OMI has a spatial resolution of 13 × 24 km^2 at nadir).

Deleted: ; detailed

Deleted: 2018a

of data quality flags for each pixel, the removal of a possible stripe pattern along each cross-track position, and corrections for background values. Each retrieval step is described in more detail in the following sections.

Deleted: making

Deleted: and latitudinal biases

2.2.1 Wavelength calibration and GEMS bandpass function

Wavelength grid points of measured irradiances and radiances in a charge-coupled device (CCD) sensor are often shifted or squeezed, and such systematic biases due to wavelength shifts or squeezes need to be corrected when producing Level 1B data. However, as precise wavelength alignments between irradiances/radiances and absorption cross-sections are required to achieve accurate radiance fitting, it is also necessary to conduct wavelength calibration during retrieval.

In wavelength calibration, the solar reference spectrum (Chance and Kurucz, 2010) is firstly convolved with a GEMS bandpass function and is then interpolated to the wavelength grids of the measured spectrum. A convolved solar reference spectrum with wavelength shift parameters and polynomial parameters (Eq. 1) is then fitted to the measured irradiances and radiances in a broader fitting window (325.5–358.5 nm) than that of the radiance fitting for HCHO retrievals as follows:

Deleted: ,

$$I_R(\lambda) = I_0^h \otimes g(\lambda + \Delta\lambda) P_{sc}(\lambda) + P_{bl}(\lambda), \quad (1)$$

Deleted: = (I_0^h *

Deleted:)(

Deleted: $+P_{bl}$,

where $I_R(\lambda)$ is the modeled irradiance and radiance, $I_0^h(\lambda)$ is the solar reference spectrum with a high spectral resolution of 0.01 nm wavelength interval, $\Delta\lambda$ is the wavelength shift, $g(\lambda)$ is a bandpass function, and $P_{sc}(\lambda)$ and $P_{bl}(\lambda)$ are scaling and baseline polynomials, respectively. The symbol $\otimes$ denotes the convolution procedure, as shown in Eq. 2.

Deleted: asterisk

$$f \otimes g(\lambda) = \int_{-\infty}^{\infty} f(\Lambda) g(\lambda - \Lambda) d\Lambda \quad (2)$$

Deleted: $(f * g)(\lambda)$

We use the GEMS bandpass functions for the convolution in wavelength calibration to ensure consistency with the spectral resolutions of measured irradiances and radiances. GEMS bandpass functions are provided at seven center wavelengths ranging from -1.8 nm to 1.8 nm

at center wavelengths, with wavelength sampling intervals of 0.06 nm (Fig. 3). The right panel in Fig. 3 shows bandpass functions averaged for spatial indices at 330 and 365 nm, and it also shows the relative differences between bandpass functions at 365 nm and 330 nm. The relative differences are smaller near the wavelength center, but they increase over each wing of the function. For GEMS HCHO retrieval, we will conduct calibration for every spatial pixel of the sensor using the bandpass functions at 330 nm. However, as bandpass functions are not linear with wavelengths, it will be necessary to estimate uncertainties for the wavelength dependency of the bandpass functions after GEMS is launched.

Deleted: wavelength

2.2.2 Convolution and reference spectra sampling

Table 1 shows a summary of absorption cross-section datasets used in GEMS HCHO retrieval. In the retrieval algorithm, absorption cross-section data with fine spectral resolutions (for example, HCHO absorption cross-section data with a spectral resolution of 0.011 nm) are first convoluted with the bandpass functions described in Section 2.2.1, and they are then interpolated to the calibrated wavelength grids of measured radiances. Finally, radiance fitting accounts for attenuation of a reference spectrum (measured irradiance or radiance) by gas absorption using Eq. 3 with convoluted cross-section data as follows:

Deleted: ,

$$\text{attenuated radiance in radiance fitting} = I_0^h \otimes g(\lambda) e^{-\tau^h \otimes g(\lambda)}, \quad (3)$$

Deleted: = (I_0^h *

Deleted:)(λ)exp (−(τ^h * g)(λ)),

where τ^h is the optical depth of interfering gases with fine spectral resolutions.

Deleted: where the asterisks denote convolution (Eq. 2) and

However, radiative transfer in the atmosphere occurs in a slightly different manner. Solar irradiance is firstly reduced by the absorption of interfering gases, and radiances are subsequently measured on discrete wavelength grids of an instrument with predetermined spectral resolutions, as shown in Eq. 4,

$$\text{attenuated radiance in reality} = (I_0^h(\lambda) e^{-\tau^h(\lambda)}) \otimes g(\lambda). \quad (4)$$

Deleted: = ((I_0^h exp(−τ^h)) * g)(λ)

Therefore, the difference between measured and calculated radiances on discrete grids may provide biases in radiance fitting when sensors have limited spectral resolutions, which is referred to as the solar I_0 effect (Aliwell et al., 2002; Chan Miller et al., 2014).

We calculate pseudo absorption cross-sections to account for the differences in gas absorptions between reality and radiance fitting. We assume that the absorption process in radiance fitting is the same with that in reality, and the pseudo absorption cross-sections are computed using the following equation (Aliwell et al., 2002; Chan Miller et al., 2014),

$$\sigma_{ps}(\lambda) = \frac{1}{scd_{ref}} \ln\left(\frac{I_0^h \otimes g(\lambda)}{\left(I_0^h(\lambda) e^{-scd_{ref}\sigma^h(\lambda)}\right) \otimes g(\lambda)}\right), \quad (5)$$

Deleted: $\ln\left(\frac{I_0^h(\lambda)*g(\lambda)}{[I_0^h(\lambda)\exp(-scd_{ref}\sigma^h(\lambda))]*g(\lambda)}\right)$

where $\sigma_{ps}(\lambda)$ is a pseudo absorption cross-section, scd_{ref} is a reference slant column density, and $\sigma^h(\lambda)$ is an absorption cross-section with a fine spectral resolution.

Although the corrected absorption cross-section can be applied to all the species, we only apply the correction to the O_3 absorption cross-section, which is the most important interfering species in the fitting window of HCHO retrievals, and we use an O_3 reference slant column density of 300 DU for the correction.

2.2.3 Radiance fitting

Three different methods have been used with sun-synchronous satellite measurements in previous HCHO retrievals: differential optical absorption spectroscopy (DOAS) method (De Smedt et al., 2008; De Smedt et al., 2012), a non-linearized fitting method, which is known as basic optical absorption spectroscopy (BOAS) (Chance et al., 2000; González Abad et al., 2015; González Abad et al., 2016), and principal component analysis (Li et al., 2015). Zhu et al. (2016) conducted an inter-comparison of HCHO vertical column densities retrieved using the three retrieval methods for four instruments such as OMI, GOME-2A, GOME-2B, and OMPS; they found that the different retrieval results were consistent in terms of temporal and spatial variations of HCHO columns in the southeastern United States.

To yield HCHO slant columns in this study, we use the BOAS method, which is based on a non-linearized form of the Lambert-Beer law, as shown in Eq. 6 (González Abad et al., 2015). One advantage of the BOAS method is that it uses unprocessed radiance data, and it is thus

more intuitive than the widely used DOAS method, which uses a linearized form by taking the logarithm of radiance to irradiance and high-pass filtering the result. The modeled radiative equation is given as follows:

$$I(\lambda) = [(aI_0(\lambda) + c_r\sigma_r(\lambda))e^{-\sum_i SCD_i\sigma_i(\lambda)} + c_{cm}\sigma_{cm}(\lambda)]P_{sc}(\lambda) + P_{bl}(\lambda), \quad (6)$$

where a is an amplitude factor; $I_0(\lambda)$ is a reference spectrum (solar irradiance or radiance reference); $c_r\sigma_r(\lambda)$ is the contribution of the Ring effect; $e^{-\sum_i SCD_i\sigma_i(\lambda)}$ is the contributions of all gas absorptions; SCD_i and $\sigma_i(\lambda)$ are slant column densities and absorption cross-sections for species i, respectively; $c_{cm}\sigma_{cm}(\lambda)$ is the contribution of the common mode; and $P_{sc}(\lambda)$ and $P_{bl}(\lambda)$ are scaling and baseline polynomials, respectively, considering low frequency variations due to Rayleigh and Mie scattering. Furthermore, the modeled spectrum is fitted to measured radiances using a non-linear least square method (Wedin and Lindström, 1987) to yield HCHO slant columns.

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. Fitting options in González Abad et al. (2015) are followed. 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 direction 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.

Deleted: ,

Deleted: $[(aI_0(\lambda) + c_r\sigma_r(\lambda))\exp(-\sum_i SCD_i\sigma_i) + c_{cm}\sigma_{cm}]$

Deleted: ,→

Deleted: $\exp(-\sum_i SCD_i\sigma_i)$

Moved down [1]: Table 1 summarizes the detailed information used in the GEMS HCHO retrieval algorithm.

Deleted:).

Deleted: far east swaths

Moved (insertion) [1]

2.2.4 Air mass factor

HCHO slant column densities (Ω_s) from the radiance fitting are then converted to vertical columns (Ω_v) with an AMF (Eq. 7), which is a correction factor for the light slant path to the vertical path. Previous studies have shown that AMF uncertainty is one of the crucial factors causing retrieval uncertainties (De Smedt et al., 2012 and 2018), and AMF uncertainties contribute to retrieval uncertainties by multiple factors including cloud top pressure, cloud fraction, HCHO vertical distribution, aerosol vertical distribution, and aerosol optical properties (Millet et al., 2006; Chimot et al., 2016; Kwon et al., 2017; Hewson et al., 2015).

Deleted: mainly

Deleted: over polluted areas because AMFs can be influenced

Deleted: (

Deleted:) (

$$\Omega_v = \frac{\Omega_s}{AMF} \quad (7)$$

An AMF can be decoupled with a scattering weight (w_l) and a vertical shape factor (S_l) of the target species (Eq. 8), which represent radiative sensitivity to the optical depth of the absorber and a partial column density profile normalized by total vertical column density, respectively, at each layer l ($l = 1, 2, \dots, n$) (Palmer et al., 2001). Scattering weights are dependent on the solar zenith angle (θ_s), viewing zenith angle (θ_v), relative azimuth angle (θ_r), surface albedo (α_s), cloud top pressure (p_{cld}), and effective cloud fraction (f_c).

$$AMF = \sum_{l=1}^{n} w_l(lat, lon, month, \theta_s, \theta_v, \theta_r, \alpha_s, p_{cld}, f_c) S_l(lat, lon, month) \quad (8)$$

Surface albedo, effective cloud fraction, and cloud top pressure are retrieved from GEMS and used in the AMF calculations. GEMS Level 2 surface properties include Lambertian equivalent reflectivity (LER) and the daily bidirectional reflectance distribution function (BRDF) (Lee and Yoo, 2018). GEMS LER products are retrieved as composites of minimum LER values for 15 days every hour with fixed viewing geometry so that geometry dependent LER are yielded. The effective cloud fraction and cloud top pressure (effective cloud pressure) are retrieved from GEMS with the assumption of a Lambertian cloud surface (cloud surface albedo = 0.8) (Veefkind et al., 2016). GEMS surface reflectivity products are also used for cloud retrievals. In addition, the radiative cloud fraction (f_{rc}) will be provided from GEMS

Deleted: can be

Moved (insertion) [2]

Deleted:).

Level 2 cloud products, and is defined by Eq. 9, where I_{cld} and I_{clr} are radiances over cloud and cloud-free surfaces, respectively.

$$f_{rc} = \frac{f_c I_{cld}}{(1-f_c) I_{clr} + f_c I_{cld}} \quad (9)$$

In order to consider the presence of clouds in the AMF calculations, scattering weights in the partial cloudy scenes are linearly interpolated as a function of radiative cloud fractions with scattering weights for clear sky ($w_{l,nc}$) and fully covered cloudy sky ($w_{l,wc}$) (Eq. 10) (Martin et al., 2002; González Abad et al., 2015). The latter ($w_{l,wc}$) is calculated as a function of cloud top pressures using the assumption of Lambertian cloud surface with a cloud surface albedo of 0.8.

$$w_l = (1 - f_{rc}) w_{l,nc}(lat, lon, month, \theta_s, \theta_v, \theta_r, \alpha_s) + f_{rc} w_{l,wc}(lat, lon, month, \theta_s, \theta_v, \theta_r, p_{cld}) \quad (10)$$

For AMF calculations, we compile an AMF look-up table (LUT) at 340 nm as a function of the variables described in Eqs. 8–10 and Table 1. González Abad et al. (2015) showed that the wavelength dependence of scattering weights in the HCHO fitting window is small. Therefore, we ignore the wavelength dependence and use AMF values for one wavelength. The AMF LUT is calculated using VLIDORT v2.6 (Spurr, 2006) with a priori data including temperature, pressure, and gas profiles (O_3, NO_2, SO_2, and HCHO), which were simulated from a 3-D chemical transport model (GEOS-Chem v9-01-02; Bey et al., 2001) driven by Modern-Era Retrospective Analysis for Research and Applications (MERRA) with 47 vertical levels and a 2° × 2.5° horizontal resolution, for 2014.

However, the horizontal resolution of 2° × 2.5° for HCHO profiles in AMF LUT is much coarser than the GEMS horizontal resolution of 7 × 8 km^2 to discern spatial variations by local source emissions. HCHO profiles in AMF LUT are monthly averaged so that hourly variations are not taken into account. In order to resolve these coarser conditions, we can use HCHO profiles with a finer resolution as a function of time. For example, Kwon et al. (2017) showed that HCHO retrievals using monthly mean hourly AMF values were in better agreement with the model simulations in observation system simulation experiments (OSSE) than those using

Deleted: We also
Deleted: :
Deleted: $w_{l,yc}$)
Deleted: $w_{l,yc}$
Deleted: in
Deleted: (see Table 1)
Deleted: $w_{l,yc}$
Deleted: are

monthly mean AMF values. Additionally, air quality forecasting data can be used to consider hourly varying HCHO profiles. Further studies are required to examine the dependency of AMF calculations on spatial resolutions and temporal variations of HCHO profiles and its effect on GEMS retrieval.

Figure 4 shows examples of scattering weights and vertical profile shapes from the AMF LUT in June with conditions of a solar zenith angle of 30°, a viewing zenith angle of 0°, a relative azimuth angle of 90°, a surface albedo of 0.1, a cloud top pressure of 800 hPa, and an effective cloud fraction of 0.3 over clean and polluted grids. Clean grids are classified as having a HCHO column density less than 3.0×10^{15} molecules cm^{-2} and a surface pressure higher than 990 hPa, and polluted grids have a HCHO column higher than 1.0×10^{16} molecules cm^{-2} and a surface pressure higher than 990 hPa. In Fig. 4, although scattering weights are not significantly changed, the normalized vertical profile (a vertical shape factor) over a polluted area is larger near the surface compared to a clean area, which results in AMF values of 1.55 and 1.28 over the clean and polluted areas, respectively.

Deleted:

2.2.5 Post processes

Post processes include systematic bias corrections and a statistic data quality flag calculation for each pixel. Systematic bias corrections include cross-track bias correction and background HCHO column correction. Cross-track biases can appear in two-dimensional CCD sensors such as GEMS as functions of each cross-track position when a solar irradiance is used as a reference spectrum (Chan Miller et al., 2014; Nowlan et al., 2016). The cross-track biases are attributed to cross-track variability of the measured irradiance. For example, the biases for OMI are constant at different latitudes; therefore, the biases are shown as stripes in the along-track direction. The cross-track biases are estimated by a polynomial fit through medians of HCHO slant columns for each cross-track position in clean sectors, which are the easternmost swaths for GEMS. The biases are removed from all data measured on the same day for each cross-track position.

Deleted: appear

Deleted: and

Deleted: The

Deleted: show

Deleted: Pacific Ocean for OMI and

An alternative method to avoid the above-mentioned biases in the fitting procedure is to use measured radiances over a clean background region (referred to as radiance references) as the reference spectrum in radiance fitting. As measured radiance includes instrument noise and attenuation by interfering gases in the background atmosphere, the interfering effects can be minimized in radiance fitting, which results in negligible cross-track biases.

Deleted: that can be used

For GEMS, we plan to use simulated HCHO columns over the easternmost regions (143-150°E) as GEMS reference sectors, shown as 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 comparison with background HCHO vertical columns over the Pacific Ocean for OMI (Fig. S1), the 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 the 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}, which is 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.

The retrieved slant columns using a radiance reference are differential slant columns ($\Delta SCD = SCD - SCD_0$) and do not include background HCHO columns (SCD_0) that are mainly from the oxidation of methane. To account for the background columns, we use HCHO vertical columns simulated in 2014 from a chemical transport model, GEOS-Chem (Bey et al., 2001) with a spatial resolution of 2° × 2.5°. Simulated HCHO vertical columns are averaged zonally and monthly over the reference sectors and are interpolated to 720 latitudinal grid points with a resolution of 0.25° from 90°S to 90°N.

In order to account for dependency of measured radiances on geometric angles, we then convert simulated background vertical columns into slant columns by applying AMF values over the reference sector (AMF_0), which are calculated with cloud information and geometric angles on the reference sectors. Corrected GEMS HCHO slant columns are formulated as the sum of the retrieved differential slant columns and the simulated background slant columns as shown in Eq. 11,

$$\Omega_s(i,j) = SCD_{corr}(i,j) = \Delta SCD(i,j) + AMF_0(lat)VCD_m(lat), \quad (11)$$

Deleted: In this case, the

Deleted: related to

Deleted: For OMI products, simulated

Deleted: zonally

Deleted: clean background regions for

Deleted: sector (140-160°W, 90°S-90°N), except for Hawaii (154-160°W, 19-22°N),

Deleted: ,

Deleted: . For GEMS, we plan to use simulated HCHO columns over easternmost regions as GEMS reference sectors

Deleted: To derive

Deleted: from the model,

Deleted: same orbit. Pixels for AMF_0 are selected with cloud fractions of less than 0.4, data quality flags of 0 (good pixels), and retrieved slant columns of ± 1.0 × 10^{16} molecules cm^{-2}.

Deleted:).

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. We finally apply AMF values from the LUT to the corrected slant columns to obtain GEMS HCHO vertical column densities.

After the correction of systematic biases and conversion to vertical column densities with AMFs, all pixels are flagged with vertical columns and fitting uncertainties (González Abad et al., 2015). We assign a data quality flag of 0 for good pixels, where retrieved vertical columns plus two-times fitting uncertainties are positive. The pixels in which retrieved vertical columns are negative within two-times fitting uncertainties, but positive within three-times fitting uncertainties, are assigned with a data quality flag of 1, which represents suspected quality pixels. Pixels with negative vertical columns within three-times fitting uncertainties are designated as bad quality pixels and given a data quality flag of 2, and missing values are flagged by -1. It is of note that these conditions are generous and that tighter conditions for good data may be required for analysis.

3 Uncertainty analysis

We use error and uncertainty terminology from the Guide to the Expression of Uncertainties in Measurements (GUM) (https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf). Following GUM, "error" means the difference between a measurement result and a true value, and "uncertainty" means the dispersion of measurement values, such as a standard deviation and a full width at the half maximum. Because there is a lack of true values for HCHO vertical columns, we consider that the word "uncertainty" is more appropriate for use in our analysis.

Uncertainties in the retrieval steps mentioned in Section 2.2 are assumed to be uncorrelated, because the steps are independently performed. Total uncertainty in HCHO vertical column density (VCD) using a radiance reference can be formulated as follows (Boersma et al., 2004; De Smedt et al., 2008),

$$\sigma_v^2 = \left(\frac{\partial \Omega_v}{\partial SCD}\sigma_s\right)^2 + \left(\frac{\partial \Omega_v}{\partial \mathrm{AMF}}\sigma_{AMF}\right)^2 + \left(\frac{\partial \Omega_v}{\partial VCD_m}\sigma_m\right)^2 + \left(\frac{\partial \Omega_v}{\partial AMF_0}\sigma_{AMF_0}\right)^2, \quad (12)$$

where σ is the uncertainty in each part, Ω_v is the vertical column density, and subscripts v, s, and m represent vertical, slant, and model, respectively. The total uncertainty equation is transformed using Eqs. 7 and 11 into Eq. 13,

Deleted: However, we may also need to correct latitudinal biases resulting from O_3 interference dependent on latitudes. Previous studies conducted latitudinal bias corrections when using a radiance reference as the reference spectrum (González Abad et al., 2015; De Smedt et al., 2018). Latitudinal biases result from stratospheric O_3 distributions dependent on latitude; however, they can be corrected by subtracting slant columns retrieved for a radiance reference from the slant columns analyzed. ¶
Figure 5 shows OMI HCHO slant columns from OMHCHO products (Fig. 5a) and the GEMS algorithm without and with latitudinal bias corrections (Fig. 5b and 5c).

Moved down [3]: HCHO slant columns without latitudinal bias corrections (Fig. 5b)

Deleted: 5b) 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 5d shows the absolute differences between OMI HCHO slant columns with and without latitudinal bias corrections from the GEMS algorithm. 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, GEMS has south to north swaths, while sun-synchronous satellites have west to east swaths; therefore, latitudinal biases can be minimized when using a radiance reference for GEMS. However, the latitudinal biases need to be further investigated after GEMS is launched.¶

Deleted: .

Deleted: where

Deleted: ↵

Deleted: As

Deleted: exact

$$\sigma_v^2 = \frac{1}{AMF^2}\left[\sigma_s^2 + \frac{(\Delta SCD + AMF_0 VCD_m)^2}{\mathrm{AMF}^2}\sigma_{AMF}^2 + AMF_0^2\sigma_m^2 + VCD_m^2\sigma_{AMF_0}^2\right]. \quad (13)$$

We analyze expected uncertainties for the GEMS algorithm by using simulated radiances from Kwon et al. (2017) and OMI Level 1B data. In order to estimate the expected random uncertainty for GEMS (Section 3.1.1), we use simulated radiances which are convoluted with GEMS bandpass functions at 330 nm as a function of cross-track positions in the south to north direction. Simulated radiances include noises based on the expected signal-to-noise ratio for co-added pixels with spatial resolutions of 7×8 km^2. We use absorption cross-sections of Ring effect, O_3, NO_2, HCHO, and additionally SO_2 (Hermans et al., 2009; Vandaele et al., 2009) in radiance fitting because O_3, NO_2, and HCHO, and SO_2 were considered in the radiance calculation (Kwon et al. 2017).

For other uncertainty analyses, we use OMI Level 1B data with OMI slit function data (Dirksen et al., 2006) in order to examine algorithm sensitivities to different parameters. Fitting options such as absorption cross-section data and the fitting window are summarized in Table 1. It will be necessary to conduct an additional uncertainty analysis for GEMS HCHO retrievals after GEMS is launched.

Deleted: this
Deleted: retrieving HCHO VCDs with
Deleted: level
Deleted: For the
Deleted: analysis, we use
Deleted: , and use OMI slit function data (Dirksen et al., 2006). However, when GEMS is launched, it
Deleted: another

3.1 Uncertainties in slant columns

3.1.1 Random uncertainty

Uncertainties in slant columns result from random uncertainties (σ_{rand}) and systematic uncertainties (σ_{sys}) (Eq. 14) (De Smedt et al., 2018),

$$\sigma_s^2 = \sigma_{rand}^2 + \sigma_{sys}^2. \quad (14)$$

Random uncertainties are fitting uncertainties when yielding slant columns, and they mainly result from instrument noises. We can reduce random uncertainties by using measured radiances over clean areas as reference spectra instead of irradiances, as the use of measured radiances can minimize instrument noises and interference of O_3 and BrO in the stratosphere. In addition, averaging the resulting slant columns for individual pixels can reduce random

Deleted: noise
Deleted: noise

uncertainties, but this is achieved at the expense of a loss of temporal and spatial resolution (De Smedt et al., 2018).

Random uncertainties can be calculated from root mean square (RMS) values of fitting residuals, degrees of freedom ($m - n$), and diagonal components of a covariance matrix ($C_{j,j}$) for fitting parameters,

$$\sigma_{s,j}^2 = RMS^2 \frac{m}{m-n} C_{j,j} C_{j,j}, \quad (15)$$

where m and n are the number of spectral grids and fitting parameters, respectively, and j denotes specific species in fitting parameters.

Random uncertainties from the GEMS algorithm are estimated using simulated radiances. RMS of fitting residuals and random uncertainty for the GEMS domain range from 2.9×10^{-4} to 2.1×10^{-3} and 2.1×10^{15} to 1.6×10^{16} molecules cm^{-2}, respectively, which are comparable with those (RMS: 4×10^{-4} to 2.0×10^{-3}; random uncertainty: 3.3×10^{15} to 1.8×10^{16} molecules cm^{-2}) obtained from the GEMS algorithm using OMI Level 1B data. GEMS measures target species every hour in daytime , and changes of solar location during the day can affect the accuracy of radiance fitting. An averaged fitting RMS value and a random uncertainty are 6.9×10^{-4} and 5.0×10^{15} molecules cm^{-2} for conditions with both solar and viewing zenith angles less than 70, which happen at 8:00–18:00 and 9:00–16:00 local time of Seoul in summer and winter, respectively. However, the fitting RMS value and the random uncertainty increase to 1.1×10^{-3} and 8.2×10^{15} molecules cm^{-2}, respectively, when solar and viewing zenith angles are higher than 70.

3.1.2 Systematic uncertainty

Systematic uncertainties result from uncertainties of wavelength calibration, the bandpass function for convolution, and absorption cross-sections. We estimate systematic uncertainties from sensitivity tests to parameters using OMI Level 1B data. First, systematic uncertainties associated with absorption cross-sections are estimated using alternative absorption cross-sections: we compare resulting slant columns to the baseline calculation with conditions in Table 1 for a one-month period (March 2005). In the analysis, we define an uncertainty as a standard deviation of differences between the sensitivity and baseline calculations. Absorption

Deleted: component

Deleted: $\frac{RMS^2}{m-n}$

Deleted: →

Deleted: To examine random uncertainties from the GEMS algorithm for HCHO, we select pixels with the following conditions: a main data quality flag of 0, an effective cloud fraction of less than 0.3, and a solar zenith angle of less than 60° in the GEMS field of view (5°S-45°N, 75°-145°E) for four months (March, June, September, and December) in 2005. Random uncertainties vary between 3.3×10^{15} and 1.8×10^{16} molecules cm^{-2}, corresponding to pixels with fitting RMS from 4.4×10^{-4} to 2.0×10^{-3}, and accounting for 97% of all pixels. In comparison with OMHCHO retrievals, González Abad et al. (2015) showed fitting uncertainties ranging from 6×10^{15} molecules cm^{-2} to 100% or larger of slant columns and RMS values from 0.4×10^{-3} to 2.0×10^{-3}, which are similar to those from the GEMS algorithm.¶

Deleted: .

Deleted: run

Deleted: run

Deleted: run

cross sections are convoluted and interpolated using the same spectral resolution and wavelength sampling to enable comparisons between datasets.

To test the retrieval sensitivity to HCHO absorption cross-sections, we use HCHO absorption cross-section datasets from Cantrell et al. (1990) instead of those of Chance and Orphal (2011), which provide a rescaling of the datasets in Cantrell et al. (1990) to those of Meller and Moortgat (2000). Absorption cross-sections of Cantrell et al. (1990) are ~10% lower than those of Chance and Orphal (2011), and the differences are directly linked to slant column retrieval (Pinardi et al., 2013). Therefore, slant columns using Cantrell et al. (1990) are by a factor of 1.1 higher than those of the baseline calculation. The slant column changes are similar to values from previous studies (Pinardi et al., 2013; De Smedt et al., 2018).

We conduct a sensitivity test to O_3 absorption cross-sections at 223 K and 293 K from Chehade et al. (2013). Uncertainties of these datasets at both temperatures are ~4% in the fitting window of the GEMS algorithm. Compared to the baseline calculation, use of the O_3 datasets of Chehade et al. (2013) at 223 K and 293 K changes the slant columns by ~20% and ~8% on average, respectively, which provides uncertainties in slant columns of 1.4×10^{15} and 0.57×10^{15} molecules cm^{-2}. These uncertainties are larger than those of 13% and 5% from Pinardi et al. (2013) and De Smedt et al. (2018), respectively. It thus appears that the GEMS HCHO retrieval algorithm is the most sensitive to O_3 absorption, especially at low temperatures in the stratosphere, due to strong absorption in the ultraviolet.

A sensitivity test is conducted against the BrO dataset of Fleischmann et al. (2004), which is by ~9% lower than the baseline BrO dataset in the GEMS HCHO fitting window, and results in ~4% slant column changes compared to the baseline calculation (with an uncertainty of 0.28×10^{15} molecules cm^{-2}).

We then examine slant column uncertainties for O_4 and NO_2 absorption cross-sections. We use alternative O_4 absorption cross-sections from http://spectrolab.aeronomie.be/o2.htm that have differences of 28% compared to the dataset used in the baseline calculation. Due to the large uncertainties of the data compared to other absorption cross-section data, the resulting slant column changes are significant, ~24%, with an uncertainty of 1.6×10^{15} molecules cm^{-2}. This uncertainty could thus be decreased by reducing the uncertainties of the O_4 datasets. Also, O_4 should be included in the large fitting interval for HCHO because it has strong peaks near 344.0 and 361 nm (De Smedt et al., 2015; Thalman and Volkamer, 2013). The NO_2 datasets

Deleted:)

Deleted: run

Deleted: run

Deleted: run

Deleted: run

from Burrows et al. (1998) are 2% larger than those in the baseline calculation; switching to them leads to ~5% slant column changes with uncertainty of 0.37×10^{15} molecules cm^{-2}.

We also estimate the systematic uncertainties of slant columns for wavelength calibration and solar I_0 effects by using an alternative solar irradiance reference. As described in Section 2.2.1 and 2.2.2, a solar irradiance reference spectrum was used in the wavelength calibration and calculation of pseudo absorption cross-sections related with solar I_0 effect. An alternative reference spectrum from Kitt Peak National Observatory (Kurucz et al., 1984) is almost identical to that of the baseline calculation, but the resulting slant column changes are up to ~14% with uncertainty of 0.92×10^{15} molecules cm^{-2}. We thus consider that the solar I_0 effect associated with the strongest interfering gas (O_3) is very sensitive to the reference spectrum.

The total systematic uncertainty of slant columns for the parameters discussed is 38% of the slant column densities on average for one month. This uncertainty is larger than that of De Smedt et al. (2018), which is 20% of the slant column densities. However, there are remaining slant column uncertainties resulting from uncertainties relating to other parameters, such as polynomial orders in Eq. 6, instrument bandpass functions, and temperature dependency of cross-sections. Therefore, we estimate the systematic uncertainty as being 38% of the slant columns, prior to conducting uncertainty analyses on other parameters.

3.2 Uncertainty in AMF

The AMF uncertainty can be estimated by each parameter in Eq. 16. We examine AMF uncertainties for surface albedo (α_s), cloud top pressure (p_c), and effective cloud fraction (f_c) with a solar zenith angle of 30°, a viewing zenith angle of 30°, and a relative azimuth angle of 0°. In addition, we define a profile height parameter (p_h) as an altitude below which 75% of HCHO VCDs exist from the surface to estimate AMF uncertainty with respect to a HCHO profile shape. The uncertainties of parameters ($\sigma_{\alpha_s} = 0.02$, $\sigma_{p_c} = 50\ hPa$, and $\sigma_{f_c} = 0.05$) are based on De Smedt et al. (2018) and will be replaced to those from GEMS Level 2 products. The uncertainty of profile height (σ_{p_h}) is defined as a standard deviation of profile heights in AMF LUT, and σ_{p_h} in polluted and clean areas is 84 and 55 hPa, respectively.

$$\sigma_{AMF}^2 = \left(\frac{\partial AMF}{\partial \alpha_s}\sigma_{\alpha_s}\right)^2 + \left(\frac{\partial AMF}{\partial p_c}\sigma_{p_c}\right)^2 + \left(\frac{\partial AMF}{\partial f_c}\sigma_{f_c}\right)^2 + \left(\frac{\partial AMF}{\partial p_h}\sigma_{p_h}\right)^2 \quad (16)$$

Deleted: run

Deleted: run

Deleted: Derivatives

Deleted: AMF to parameters (slope) are calculated

Deleted: AMF LUT, while the

Deleted: ,

Deleted: ,

Deleted:)

Deleted:),

Deleted: .

Figure 5a shows AMF sensitivities to surface albedos with clear sky conditions ($f_c = 0$). AMF values increase linearly with an increase in surface albedos and are slightly higher for clean regions than polluted regions in which HCHO is concentrated near the surface.

Clouds below or within HCHO layers increase photon path lengths due to multiple scattering, while clouds at altitudes above HCHO layers shield photons from reaching the surface. Therefore, the AMF values decrease with a decrease in cloud top pressures (with an increase in cloud heights) due to fewer photons reaching to the surface (Fig. 5b). AMFs rapidly change with increasing cloud top pressures near the surface, implying a high AMF sensitivity to the changes in photon path lengths with respect to multiple scattering.

Cloud fractions also play an important role in determining AMF values and their effects are shown in Fig. 5c. For a cloud with its top pressure at 800 hPa, AMF values increase with an increase in the cloud fraction, which implies that more photons arrive at the near surface due to multiple scattering by clouds with increasing cloud fractions. However, for clouds with its top pressure at high altitudes (500 hPa), AMF values decrease with an increase in cloud fractions due to the shielding effect of clouds.

Figure 5d shows increasing AMF values with an increase in the profile height, resulting from increased HCHO absorptions at high altitudes. The AMF sensitivity to profile heights in clean areas is higher than that in polluted areas because HCHO distributions are more uniform in clean areas than polluted areas.

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, clouds at high altitudes, and high cloud fractions, but they do not differ significantly 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.

Aerosol vertical distributions and aerosol optical properties are also important sources of AMF uncertainty (Chimot et al., 2016; Kwon et al., 2017; Hewson et al., 2013). Non-absorbing

Deleted: 6a

Deleted: from AMF LUT

Deleted: albedo,

Deleted: sensitivities of AMF to surface albedo

Deleted: low surface albedos of less

Deleted: 0.2. The

Deleted: mixing ratio

Deleted: higher

Deleted: in polluted areas compared to clean areas, which relates to lower AMF values in polluted areas with low surface albedos.

Deleted: high

Deleted: near

Deleted: 6b

Deleted: are more sensitive to

Deleted: pressure

Deleted: ; this implies

Deleted: is more sensitive

Deleted: than blocking by clouds near the surface

Deleted: roles

Deleted: 6c

Deleted: clouds

Deleted: the

Deleted: from the

Deleted: Figure 7 shows AMF uncertainty as functions of each parameter. AMF uncertainties with respect to surface albedo are estimated using a surface albedo (σ_{α_s}) uncertainty of 0.02 (Kleipool et al., 2008) and are uniform compared to those with respect to cloud properties. In cloud free conditions ($f_c = 0$ in Fig. 7c), uncertainties of surface albedo almost account for total AMF uncertainty. For cloudy conditions, we use the uncertainties of cloud top pressure ($\sigma_{p_c} = 50$ hPa) and cloud fraction ($\sigma_{f_c} = 0.05$

Moved up [2]:) (Veefkind et al., 2016).

Deleted: In conditions where the surface albedo is less than 0.6 and the cloud top pressure is higher than 550 hPa, AMF uncertainties for cloud top pressures account for most of the total AMF uncertainties (Fig. 7a and 7b). In the opposite conditions, the uncertainties of cloud fractions contribute to most of the total AMF uncertainties. In addition, AMF uncertainties for cloud top pressures account for almost all the AMF uncertainties with a cloud fraction higher than 0.4 (Fig. 7c). Total AMF uncertainties result in 4% to 16% of HCHO vertical column uncertainties from Eqs. 13 and 16.¶
HCHO profile shapes, aerosol vertical distributions,

Deleted: AMF uncertainties from HCHO profiles are estimated to range from 0.27–0.44; they are thus the largest uncertainties of all other parameters in Eq. 16. The AMF uncertainty in HCHO profiles leads to a 15%–63% uncertainty in HCHO vertical columns. However, we can minimize the a priori HCHO profile uncertainties by using averaging kernels.¶

aerosols play a similar role as that of clouds in radiative transfer and are implicitly considered when cloud information is used. However, absorbing aerosols, such as mineral dust and black carbon, counteract the effects of non-absorbing aerosols and clouds. In the GEMS domain, dust storms and biomass burning occur seasonally, and we may therefore need to consider the effect of absorbing aerosols on the retrieval. We plan to update our AMF LUT as a function of aerosol optical depth, single scattering albedo, and aerosol height – all of which will be retrieved by 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).

Deleted: have
Deleted: roles to
Deleted: ,
Deleted: non-absorbing aerosols
Deleted: using aerosol information and investigating the interaction between aerosols and clouds.

3.3 Uncertainty in background correction

The uncertainties of model results for background concentrations (σ_m) and AMFs over reference sectors (σ_{AMF_0}) are important errors for background corrections (σ_{bg}) in Eq. 17, which represents the 3rd and 4th terms on the right-hand side of Eq. 13,

Deleted: contribute to uncertainties in

$$\sigma_{bg}^2 = \frac{1}{AMF^2}\left[AMF_0^2\sigma_m^2 + VCD_m^2\sigma_{AMF_0}^2\right]. \quad (17)$$

We estimate model uncertainties by using standard deviations of model results as a function of latitude. Standard deviations range from 2.3 × 10^{14} to 1.4 × 10^{15} molecules cm^{-2}. Uncertainties related to AMF in the reference sectors are identical to those discussed in Section 3.2.

Deleted: the differences in
Deleted: between the GEMS algorithm and the OMHCHO products, and find that model HCHO vertical columns used in the GEMS algorithm are lower than those used in OMHCHO, with absolute differences ranging…
Deleted: -1
Deleted: to 6.1 × 10^{14}. In Section 4, we discuss the impacts of these differences on retrieved HCHO columns

4 Results and validation

Deleted: Validation

In this section, we use OMI Level 1B data to validate our retrieval algorithm. Resulting HCHO products are compared with OMI products of other institutes for one month of each season (March, June, September, and December) in 2005 to provide seasonal variation in the GEMS domain. We also compared OMI products with ground-based MAX-DOAS at two sites in Europe for 2005 and one site in China for 2016.

Deleted: is validated by a comparison
Deleted: . First, we yield HCHO column densities using the GEMS HCHO algorithm with OMI Level 1B data in the fitting options described in Table 1, and also include spectral undersampling in the fitting process (Chance et al., 2005). We conduct our calculations

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. Averaged 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

Moved (insertion) [3]

Deleted: Figure 8

Deleted: Average

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}.

4.2 Comparison with other OMI products

Figure 8 compares monthly mean slant columns retrieved using the GEMS algorithm and those of OMHCHO products (González Abad et al., 2015). We select pixels with (1) vertical columns ranging from -5.0×10^{15} to 10×10^{16} molecules cm^{-2}, (2) a main data quality flag of 0, (3) an effective cloud fraction of less than 0.3, and (4) a solar zenith angle of less than 60. Monthly mean slant column densities are weighted by uncertainties and overlapped areas between pixels and grid boxes with a horizontal resolution of 0.25° × 0.25°.

We find similar spatial patterns of HCHO slant columns in both products; this shows that high HCHO columns occur over Indonesia and the Indochina Peninsula in March and over Indonesia in September, owing to biomass burning and biogenic activities. In summer, HCHO enhancements over China are caused by biomass burning and the oxidation of biogenic and anthropogenic VOCs due to photochemical reactions. In addition, high HCHO slant column densities occur over the Pearl River Delta, where anthropogenic VOCs are emitted from petrochemical industries, cargo ports, paint production, and many other activities (Zhong et al., 2013). The scatter-plot comparisons with OMHCHO products show that GEMS HCHO slant columns are in good agreement with OMHCHO products, with correlation coefficients of 0.77–0.91 and regression slopes of 0.94–1.04. The relative differences between GEMS slant columns and those of OMHCHO products are -3% to 0.1% on average over the domain.

However, some discrepancies are found despite overall good agreement between GEMS and OMHCHO products, and these are mainly related to the background correction from the different model results. Figure 9 shows the simulated HCHO vertical columns used for background corrections in OMHCHO (solid lines) and the GEMS algorithm (dashed lines), respectively. Similar latitudinal variations are shown with a peak in the tropics and gradual decreases in high latitudes, which reflects the photochemical production of HCHO; however, the magnitudes slightly differ. The model results used in GEMS are smaller than those used in the OMHCHO products (especially within ± 20° latitudes) by -1.3×10^{15} molecules cm^{-2} near 6°N in September. Both results were obtained from the same 3-D global chemical transport model (GEOS-Chem), but different assimilated meteorological products were employed.

Deleted: results from random uncertainties
Deleted: in
Deleted: 9
Deleted: (1)
Deleted: Latitudinal biases are also corrected using the method discussed in Section 2.2.5.
Deleted: .
Deleted: 92
Deleted: 10
Deleted: Values show similar
Deleted:)

HCHO vertical columns used in the OMHCHO products were from GEOS-Chem with GEOS-4 meteorological data (Millet et al., 2006), which have lower cloud optical depths near the equator compared to GEOS-3 and GEOS-5. Therefore, low cloud optical depths in the tropics resulted in faster methane oxidation and the production of more HCHO in relation to high hydroxyl radical concentrations (http://wiki.seas.harvard.edu/geos-chem/index.php/GMAO_GEOS-4). González Abad et al. (2016) also showed that the OMHCHO products are larger than OMI BIRA-IASB products near the equator, such as in Southeastern Asia, Tropical Africa, and the Amazon Basin.

We select four regions (Sumatra/Malaysia, the Indochina Peninsula, China, and Borneo), where HCHO is abundant from biomass burning and biogenic and anthropogenic sources. Table 3 provides the relative differences between OMI GEMS HCHO slant columns and OMHCHO slant columns in these four regions. GEMS HCHO slant columns are 1% to 13% lower than those of OMHCHO in Sumatra/Malaysia and Borneo, because the differences in simulated HCHO column densities for background corrections are relatively large near the equator compared to the mid-latitudes, as previously mentioned. In the Indochina Peninsula and China, however, the GEMS HCHO slant columns are 6% to 25% higher than the OMHCHO slant columns, although the simulated HCHO concentrations used in the GEMS algorithm for background corrections are lower than those used in OMHCHO.

We also compare OMI GEMS HCHO slant columns with OMI QA4ECV products (OMI QA4ECV) (De Smedt et al., 2018). The QA4ECV project was proposed to provide reliable satellite and ground-based measurements of climate and air quality variables with detailed uncertainty information (http://www.qa4ecv.eu). Figure 10 shows that the spatial distributions of the GEMS HCHO slant columns are consistent with those of the OMI QA4ECV products, but relatively poorer correlation coefficients of 0.52 to 0.76 are found compared to those with the OMHCHO products. The relative differences between GEMS and QA4ECV slant columns range from -11% to -22% on average over the GEMS domain.

Magnitudes of relative differences between OMI GEMS and OMI QA4ECV slant columns vary regionally and seasonally (Table 3). The differences are relatively small near the equator (Sumatra/Malaysia and Borneo) compared to those in subtropics and mid-latitudes (Indochina Peninsula and China) because biomass burning often occurs and biogenic sources are more abundant near the equator. The HCHO differences near the equator are lower in spring and fall than those in summer and winter due to biomass burning. The effects of biomass burning on HCHO slant columns are also found over Indochina Peninsular in spring. As a result,

Deleted: 2

Deleted: from BIRA-IASB

Deleted: BIRA) in Fig. 11

Deleted: The

Deleted: BIRA

Deleted: and provide

Deleted: of

Deleted: BIRA

Deleted: This

magnitudes of relative differences in regions and seasons with high HCHO concentrations are small, implying that HCHO can be well-retrieved because of the abundant HCHO concentrations, regardless of the fitting method used.

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).

4.3 Comparison with ground- based MAX-DOAS

We also compare satellite results with MAX-DOAS ground observations at Haute-Provence Observatory (OHP) in France, Bremen in Germany, and Xianghe in China (Table 4). MAX-DOAS data are collected within the OMI overpass time (12:00–15:00 local time) at OHP and Bremen in 2005, and at Xianghe in 2016, respectively. We collect OMI data pixels that are overlapped by a grid box of 0.25° at the center of the site location, and average values of OMI data are weighted by uncertainties and overlapped areas between pixels and grid boxes.

Comparisons of HCHO VCDs between MAX-DOAS and satellite products are shown in Fig. 11 and Table 4. Averaged MAX-DOAS HCHO VCDs for a year are 7.6×10^{15}, 6.7×10^{15}, and 1.6×10^{16} molecules cm^{-2} at OHP, Bremen, and Xianghe, respectively. HCHO VCDs show a seasonal variation with the maximum concentrations in summer at all sites (Fig. S3).

Deleted: BIRA
Deleted: use
Deleted: OMI BIRA uses a two-step method: radiance fitting is conducted in a broad fitting window (328.5–359.0 nm) to remove interferences from BrO and O_4, and HCHO slant columns are then derived in a narrow fitting window (328.5–346.0 nm) with fitted BrO slant columns. …
Deleted: BIRA
Deleted: . Differences caused by
Deleted: absorption datasets is the largest term in the
Deleted: (as mentioned in Section 3.1.2); therefore, relatively lower slant columns in the GEMS algorithm compared to OMI BIRA may result from the different
Deleted: employed
Deleted: the
Deleted: The regional and seasonal magnitudes of relative differences between OMI GEMS and OMI BIRA slant columns are larger over the Indochina Peninsula and China and in summer and winter, respectively, compared to those over Sumatra/Malaysia and Borneo, and compared to in spring and fall (Table 2). Biomass burning often occurs in spring and fall within Sumatra/Malaysia and Borneo, and during the spring in the Indochina Peninsula, which results in strong HCHO enhancements (see Fig. 11). That implies that HCHO can be well-retrieved because of the abundant HCHO concentrations, regardless of the fitting method used.¶
Deleted: QA4ECV
Deleted: 3
Deleted: daily averaged
Deleted: and are collected
Deleted: in March, June, and September 2005, at
Deleted: June and September
Deleted: May
Deleted: 12
Deleted: 3.
Deleted: averaged
Deleted: each period
Deleted: 9
Deleted: .2
Deleted: 8

The largest monthly change is shown at Xianghe, likely driven by abundant VOC precursors for HCHO production compared to OHP and Bremen.

Averaged HCHO VCDs from OMI GEMS are by 16%, 9%, 25% lower than those from MAX-DOAS at OHP, Bremen, and Xianghe. At Bremen, HCHO VCDs from the GEMS algorithm are in the best agreement with those of MAX-DOAS and show similar monthly variations with MAX-DOAS. OMI GEMS results at Xianghe show a monthly variation but, at OHP, do not show monthly variation despite of a bit increment in summer. In particular, the GEMS algorithm yields lower HCHO VCDs in summer. These lower values may be caused by the a priori HCHO profiles used in AMF calculation. In summer, HCHO is produced and concentrated near the surface, which results in lower AMFs (higher VCDs). S. W. Kim et al. (2018) showed the anti-correlation between AMF values and the HCHO mixing ratios at 200 m above ground level. OMHCHO products show similar tendencies as OMI GEMS, but they are much lower than those of OMI GEMS. OMI QA4ECV products are higher than MAX-DOAS at OHP and Bremen but are in the best agreement with MAX-DOAS at Xianghe compared to other satellite products.

5 Conclusions and discussions

We have developed a GEMS HCHO algorithm based on a non-linearized fitting method and described the algorithm in detail. The GEMS HCHO algorithm consists of three steps: pre-processes, radiance fitting, and post-processes. Pre-processes include wavelength calibration, and interpolation and convolution of absorption cross-sections. In the radiance fitting, HCHO slant column densities are retrieved by minimizing the difference between calculated radiances with initial guesses of absorbing gases and measured radiances in HCHO fitting windows. Finally, AMF values are calculated from an AMF LUT, and bias corrections are conducted if necessary.

We estimated the uncertainties of slant columns, AMF, and background corrections using simulated radiances and OMI Level 1B data. The random uncertainties of slant columns are estimated using simulated radiances and are comparable with those of OMI GEMS products. The systematic uncertainty is 38% of the slant columns, which is higher than that of De Smedt et al. (2018). However, the systematic uncertainty can be reduced by using up-to-date absorption cross-sections. AMF uncertainty amounts to by 20% and 24% of the HCHO vertical columns in clean and polluted areas, respectively, and mainly results from uncertainties

Deleted: results
Deleted: %, 0.8%, and 30
Deleted: , respectively,
Deleted: by 45% and 27% in June at OHP and Bremen, respectively, compared to those from MAX-DOAS
Deleted: 2018b
Deleted: OMHCHO products show similar tendencies as OMI GEMS, but they are much lower than those of OMI GEMS. OMI BIRA products are higher by 19%–66% than ground observations at all three sites. Of the three satellite products, the OMI GEMS products show the smallest relative differences at OHP and Bremen against MAX-DOAS, where HCHO VCDs are relatively low. However, compared to the other two sites, there are smaller differences for OMI BIRA products in Xianghe where there are high HCHO VCDs.¶
Deleted: for stripe patterns, background concentrations, and latitudinal biases …
Deleted: .
Deleted: uncertainty
Deleted: corresponds
Deleted: OMHCHO
Deleted: , but the
Deleted: with respect
Deleted: surface albedo, cloud top pressure,
Deleted: cloud fraction contributes to 4–16
Deleted: . AMF uncertainty with respect to

associated with HCHO profile heights and cloud information (cloud top pressure and cloud fractions).

OMI HCHO columns from the GEMS algorithm were compared to the OMHCHO products with consistent fitting conditions applied. OMI GEMS slant columns show good agreement with OMHCHO products, with correlation coefficients of 0.77–0.91 and regression slopes of 0.94–1.04. However, some differences between two products were found because of background corrections. Both products used the model results simulated by GEOS-Chem, but driven with different assimilated meteorological datasets. The simulated HCHO vertical columns used in OMHCHO products were from GEOS-Chem with GEOS-4 meteorological data, which have lower cloud optical depths near the equator compared to GEOS-3 and GEOS-5. Low cloud optical depths in the tropics result in faster methane oxidation and greater HCHO production caused by high hydroxyl radical concentrations. Therefore, OMHCHO slant columns are 1% to 13% higher than OMI GEMS slant columns in Sumatra/Malaysia and Borneo near the equator.

The spatial distributions of GEMS HCHO slant columns were consistent with OMI QA4ECV products, but relatively poorer correlation coefficients of 0.52 to 0.76 are found compared to those with OMHCHO products. Relative differences between GEMS and QA4ECV slant columns range from -11% to -22% on average over the GEMS domain. We found that the discrepancy between the GEMS and QA4ECV products were mainly caused by the different radiance fitting methods, polynomial orders in the fitting window, the different O_3 absorption datasets, and consideration of the non-linear O_3 absorption effect.

We also compared satellite results with MAX-DOAS ground observations at OHP in France, Bremen in Germany, and Xianghe in China. HCHO VCDs from the GEMS algorithm were by 16%, 9%, and 25% lower than those of MAX-DOAS, but the GEMS discrepancies at OHP and Bremen were the smallest compared to the other satellites against the in-situ data.

After GEMS is launched, several options need to be tested. As described in Section 2.2.1, it will be necessary to estimate uncertainties resulting from the wavelength dependence of bandpass functions. We may also need to conduct additional sensitivity tests to optimize the fitting window and fitting options. We currently use a broad fitting window (328.5–356.0 nm). However, we may need to use a different fitting window to reduce interference from polarization effects because GEMS does not include a polarization scrambler. A polarization correction is planned to minimize its interference during GEMS Level 1B production, but we

Deleted: height leads to 15%–63% uncertainty in the HCHO vertical columns, but this can be reduced by using averaging kernels.

Deleted: results

Deleted: 92

Deleted: were found despite good agreement

Deleted: GEMS and OMHCHO

Deleted: , and these are caused by the use

Deleted: different model results for

Deleted: Although both

Deleted: are

Deleted: products are employed.

Deleted: are

Deleted: are

Deleted: BIRA

Deleted: they have

Deleted: the

Deleted: BIRA

Deleted: consider

Deleted: BIRA

Deleted: may be

Deleted: employed

Deleted: compare

Deleted: are 25%, 0.8

Deleted: 30

Deleted: are

Deleted: among

Deleted: three

Deleted: as GEMS bandpass functions depend on wavelength,

Deleted: .

Deleted: narrow

Deleted: (De Smedt et al., 2018). Also, latitudinal biases for a radiance reference can also be minimized in GEMS employing its south to north swaths, as we plan to use easternmost swaths (except for islands such as Japan and Indonesia) for a radiance reference. Therefore, we will further investigate latitudinal biases after GEMS is launched.…

need to examine the retrieval sensitivity to polarization. Additionally, an update to the optimized AMF LUT with finer spatial and temporal resolutions is required.

Acknowledgements

This subject is supported by Korea Ministry of Environment (MOE) as "Public Technology Program based on Environmental Policy (2017000160001)". Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

References

Aliwell, S. R., Van Roozendael, M., Johnston, P. V., Richter, A., Wagner, T., Arlander, D. W., Burrows, J. P., Fish, D. J., Jones, R. L., Tørnkvist, K. K., Lambert, J. C., Pfeilsticker, K., and Pundt, I.: Analysis for BrO in zenith-sky spectra: An intercomparison exercise for analysis improvement, Journal of Geophysical Research: Atmospheres, 107, ACH 10-11-ACH 10-20, 10.1029/2001JD000329, 2002.

Barkley, M. P., De Smedt, I., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., and Marais, E.: Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns, Journal of Geophysical Research: Atmospheres, J18, 6849-6868, 10.1002/jgrd.50552, 2013.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research Atmospheres, 106, 23073-23095, 10.1029/2001JD000807, 2001.

Boersma, K. F., Eskes, H. J., and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, Journal of Geophysical Research: Atmospheres, 109, n/a-n/a, 10.1029/2003JD003962, 2004.

Burrows, J. P., Dehn, A., Deters, B., Himmelmann, S., Richter, A., Voigt, S., and Orphal, J.: ATMOSPHERIC REMOTE-SENSING REFERENCE DATA FROM GOME: PART 1. TEMPERATURE-DEPENDENT ABSORPTION CROSS-SECTIONS OF NO2 IN THE 231–794 nm RANGE, Journal of Quantitative Spectroscopy and Radiative Transfer, 60, 1025-1031, https://doi.org/10.1016/S0022-4073(97)00197-0, 1998.

Cantrell, C. A., Davidson, J. A., McDaniel, A. H., Shetter, R. E., and Calvert, J. G.: Temperature-dependent formaldehyde cross sections in the near-ultraviolet spectral region, Journal of Physical Chemistry, 94, 3902-3908, 10.1021/j100373a008, 1990.

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, 10.5194/amt-7-3891-2014, 2014.

Chance, K., and Spurr, R. J. D.: Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, Applied Optics, 36, 5224-5230, 10.1364/AO.36.005224, 1997.

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, Geophysical Research Letters, 27, 3461-3464, 10.1029/2000GL011857, 2000.

Chance, K., Kurosu, T. P., and Sioris, C. E.: Undersampling correction for array detector-based satellite spectrometers, Applied Optics, 44, 1296-1304, 10.1364/AO.44.001296, 2005.

Chance, K., and Kurucz, R. L.: An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible, and near infrared, Journal of Quantitative Spectroscopy and Radiative Transfer, 111, 1289-1295, https://doi.org/10.1016/j.jqsrt.2010.01.036, 2010.

Chance, K., and Orphal, J.: Revised ultraviolet absorption cross sections of H2CO for the HITRAN database, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 1509-1510, http://dx.doi.org/10.1016/j.jqsrt.2011.02.002, 2011.

Chehade, W., Gür, B., Spietz, P., Gorshelev, V., Serdyuchenko, A., Burrows, J. P., and Weber, M.: Temperature dependent ozone absorption cross section spectra measured with the GOME-2 FM3 spectrometer and first application in satellite retrievals, Atmos. Meas. Tech., 6, 1623-1632, 10.5194/amt-6-1623-2013, 2013.

Chimot, J., Vlemmix, T., Veefkind, J. P., de Haan, J. F., and Levelt, P. F.: Impact of aerosols on the OMI tropospheric NO2 retrievals over industrialized regions: how accurate is the aerosol correction of cloud-free scenes via a simple cloud model?, Atmos. Meas. Tech., 9, 359-382, 10.5194/amt-9-359-2016, 2016.

Choi, Y., Kim, H., Tong, D., and Lee, P.: Summertime weekly cycles of observed and modeled NO_x and O₃ concentrations as a function of satellite-derived ozone production sensitivity and land use types over the Continental United States, Atmos. Chem. Phys., 12, 6291-6307, 10.5194/acp-12-6291-2012, 2012.

Deleted: ¶

Deleted: 2013.

Daumont, D., Brion, J., Charbonnier, J., and Malicet, J.: Ozone UV spectroscopy I: Absorption cross-sections at room temperature, Journal of atmospheric chemistry, 15, 145-155, 10.1007/BF00053756, 1992.

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, Atmospheric Chemistry & Physics, 8, 4947-4963, 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, 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, 10.5194/acp-15-12519-2015, 2015.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmos. Meas. Tech., 11, 2395-2426, 10.5194/amt-11-2395-2018, 2018.

DiGangi, J. P., Henry, S. B., Kammrath, A., Boyle, E. S., Kaser, L., Schnitzhofer, R., Graus, M., Turnipseed, A., Park, J. H., Weber, R. J., Hornbrook, R. S., Cantrell, C. A., Maudlin Iii, R. L., Kim, S., Nakashima, Y., Wolfe, G. M., Kajii, Y., Apel, E. C., Goldstein, A. H., Guenther, A., Karl, T., Hansel, A., and Keutsch, F. N.: Observations of glyoxal and formaldehyde as metrics for the anthropogenic impact on rural photochemistry, Atmospheric Chemistry and Physics, 12, 9529-9543, 10.5194/acp-12-9529-2012, 2012.

Dirksen, R., Dobber, M., Voors, R., and Levelt, P.: Prelaunch characterization of the Ozone Monitoring Instrument transfer function in the spectral domain, Applied Optics, 45, 3972-3981, 10.1364/AO.45.003972, 2006.

Duncan, B. N., Yoshida, Y., Olson, J. R., Sillman, S., Martin, R. V., Lamsal, L., Hu, Y., Pickering, K. E., Retscher, C., Allen, D. J., and Crawford, J. H.: Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation, Atmospheric Environment, 44, 2213-2223, http://dx.doi.org/10.1016/j.atmosenv.2010.03.010, 2010.

Fleischmann, O. C., Hartmann, M., Burrows, J. P., and Orphal, J.: New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, Journal of Photochemistry and Photobiology A: Chemistry, 168, 117-132, https://doi.org/10.1016/j.jphotochem.2004.03.026, 2004.

Go, S., Kim, J., Mok, J., Irie, H., Yoon, J. M., Torres, O., Krotkov, N., Labow, G., Kim, M., Koo, J.-H., Choi, M., and Lim, H.: Column Effective Imaginary part of refractive index derived from UV-MFRSR and SKYNET in Seoul, and implications for retrieving UV Aerosol Optical Properties from GEMS measurements, Remote Sensing of Environment, summited, 2019.

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

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., 9, 2797-2812, 10.5194/amt-9-2797-2016, 2016.

Hermans, C., Vandaele, A. C., and Fally, S.: Fourier transform measurements of SO2 absorption cross sections:: I. Temperature dependence in the 24 000–29 000 cm−1 (345–420 nm) region, Journal of Quantitative Spectroscopy and Radiative Transfer, 110, 756-765, http://dx.doi.org/10.1016/j.jqsrt.2009.01.031, 2009.

Hewson, W., Bösch, H., Barkley, M., and De Smedt, I.: Characterisation of GOME-2 formaldehyde retrieval sensitivity, Atmospheric Measurement Techniques, 6, 371-386, 10.5194/amt-6-371-2013, 2013.

Hewson, W., Barkley, M. P., Gonzalez Abad, G., Bösch, H., Kurosu, T., Spurr, R., and Tilstra, L. G.: Development and characterisation of a state-of-the-art GOME-2 formaldehyde air-mass factor algorithm, Atmos. Meas. Tech., 8, 4055-4074, 10.5194/amt-8-4055-2015, 2015.

Kim, H. C., Lee, P., Judd, L., Pan, L., and Lefer, B.: OMI NO2 column densities over North American urban cities: the effect of satellite footprint resolution, Geosci. Model Dev., 9, 1111-1123, 10.5194/gmd-9-1111-2016, 2016.

Kim, J., Ukkyo Jeong, Myoung-Hwan Ahn, Jae H. Kim, Rokjin J. Park, Hanlim Lee, Chul Han Song, Yong-Sang Choi, Kwon-Ho Lee, Jung-Moon Yoo, Myeong-Jae Jeong, Seon Ki Park, Kwang-Mog Lee, Chang-Keun Song, Sang-Woo Kim, Young-Joon Kim, Si-Wan Kim, Xiong Liu, Kelly Chance, Christopher Chan Miller, Jay Al-Saadi, Ben Veihelmann, Pawan K. Bhartia, Omar Torres, Gonzalo González Abad, David P. Haffner, Dai Ho Ko, Seung Hoon Lee, Jung-Hun Woo, Mijin Kim, Sujung Go, Heesung Chong, Sang Seo Park, Dennis Nicks, Won Jun Choi, Kyung-Jung Moon, Ara Cho, Jong-Min Yoon, Sang-Kyun Kim, Hyunkee Hong, Kyunghwa Lee, Hana Lee, Seoyoung Lee, Myungje Choi, Pepijn Veefkind, Pieternel Levelt, David Edwards, Mina Kang, Mijin Eo, Juseon Bak, Kanghyun Baek, Hyeong-Ahn Kwon, Jiwon Yang, Kyung Man Han, Boram Kim, Hee-Woo Shin, Haklim Choi, Ebony Lee, Jihyo Chong, Yesol Cha, Hitoshi Irie, Sachiko Hayashida, Yasko Kasai, Yugo Kanaya, C. L., Jintai Lin, James H. Crawford, Gregory R. Carmichael, Michael J. Newchurch, Barry L. Lefer, Robert J. Swap, Alexis K H Lau, Thomas P. Kurosu, Glen Jaross, Berit Ahlers, and Dobber, M.: New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS) Bulletin of the American Meteorological Society, submitted, 2018.

Kim, M., Kim, J., Torres, O., Ahn, C., Kim, W., Jeong, U., Go, S., Liu, X., Moon, J. K., and Kim, D.-R.: Optimal Estimation-Based Algorithm to Retrieve Aerosol Optical Properties for GEMS Measurements over Asia, Remote Sensing, 10, 10.3390/rs10020162, 2018.

Deleted: 2018a

Kim, S. W., Natraj, V., Lee, S., Kwon, H. A., Park, R., de Gouw, J., Frost, G., Kim, J., Stutz, J., Trainer, M., Tsai, C., and Warneke, C.: Impact of high-resolution a priori profiles on satellite-based formaldehyde retrievals, Atmos. Chem. Phys., 18, 7639-7655, 10.5194/acp-18-7639-2018, 2018.

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

Kurucz, R. L., Furenlid, I., Brault, J., and Testerman, L.: Solar flux atlas from 296 to 1300nm. National Solar Observatory, Sunspot, New Mexico, 1984.

Kwon, H. A., Park, R. J., Jeong, J. I., Lee, S., González Abad, G., Kurosu, T. P., Palmer, P. I., and Chance, K.: Sensitivity of formaldehyde (HCHO) column measurements from a geostationary satellite to temporal variation of the air mass factor in East Asia, Atmos. Chem. Phys., 17, 4673-4686, 10.5194/acp-17-4673-2017, 2017.

Lee, K., and Yoo, J.-M.: Determination of directional surface reflectance from geostationary satellite observations, SPIE Asia-Pacific Remote Sensing, 10777-52, 2018.

Li, C., Joiner, J., Krotkov, N. A., and Dunlap, L.: A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, Geophysical Research Letters, 42, 2515-2522, 10.1002/2015GL063204, 2015.

Lin, J. T., Martin, R. V., Boersma, K. F., Sneep, M., Stammes, P., Spurr, R., Wang, P., Van Roozendael, M., Clémer, K., and Irie, H.: Retrieving tropospheric nitrogen dioxide from the Ozone Monitoring Instrument: effects of aerosols, surface reflectance anisotropy, and vertical profile of nitrogen dioxide, Atmos. Chem. Phys., 14, 1441-1461, 10.5194/acp-14-1441-2014, 2014.

Malicet, J., Daumont, D., Charbonnier, J., Parisse, C., Chakir, A., and Brion, J.: Ozone UV spectroscopy. II. Absorption cross-sections and temperature dependence, Journal of atmospheric chemistry, 21, 263-273, 10.1007/BF00696758, 1995.

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

Martin, R. V., Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I., Fiore, A. M., Li, Q., Yantosca, R. M., and Koelemeijer, R. B. A.: An improved retrieval of tropospheric nitrogen dioxide from GOME, Journal of Geophysical Research: Atmospheres, 107, ACH 9-1-ACH 9-21, 10.1029/2001JD001027, 2002.

Martin, R. V., Fiore, A. M., and Van Donkelaar, A.: Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions, Geophysical Research Letters, 31, n/a-n/a, 10.1029/2004GL019416, 2004.

Meller, R., and Moortgat, G. K.: Temperature dependence of the absorption cross sections of formaldehyde between 223 and 323 K in the wavelength range 225–375 nm, Journal of Geophysical Research: Atmospheres, 105, 7089-7101, 10.1029/1999JD901074, 2000.

Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B., Anderson, B. E., and Clarke, A. D.: Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006853, 2006.

Nowlan, C. R., Liu, X., Leitch, J. W., Chance, K., González Abad, G., Liu, C., Zoogman, P., Cole, J., Delker, T., Good, W., Murcray, F., Ruppert, L., Soo, D., Follette-Cook, M. B., Janz, S. J., Kowalewski, M. G., Loughner, C. P., Pickering, K. E., Herman, J. R., Beaver, M. R., Long, R. W., Szykman, J. J., Judd, L. M., Kelley, P., Luke, W. T., Ren, X., and Al-Saadi, J. A.: Nitrogen dioxide observations from the Geostationary Trace gas and Aerosol Sensor Optimization (GeoTASO) airborne instrument: Retrieval algorithm and measurements during DISCOVER-AQ Texas 2013, Atmos. Meas. Tech., 9, 2647-2668, 10.5194/amt-9-2647-2016, 2016.

Palmer, P. I., Jacob, D. J., Fiore, A. M., and Martin, R. V.: Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, Journal of geophysical research, 106, 14,539-514,550, 10.1029/2000JD900772, 2001.

Pinardi, G., Van Roozendael, M., Abuhassan, N., Adams, C., Cede, A., Clémer, K., Fayt, C., Frieß, U., Gil, M., Herman, J., Hermans, C., Hendrick, F., Irie, H., Merlaud, A., Navarro Comas, M., Peters, E., Piters, A. J. M., Puentedura, O., Richter, A., Schönhardt, A., Shaiganfar, R., Spinei, E., Strong, K., Takashima, H., Vrekoussis, M., Wagner, T., Wittrock, F., and Yilmaz, S.: MAX-DOAS formaldehyde slant column measurements during CINDI: intercomparison and analysis improvement, Atmos. Meas. Tech., 6, 167-185, 10.5194/amt-6-167-2013, 2013.

Puķīte, J., Kühl, S., Deutschmann, T., Platt, U., and Wagner, T.: Extending differential optical absorption spectroscopy for limb measurements in the UV, Atmos. Meas. Tech., 3, 631-653, 10.5194/amt-3-631-2010, 2010.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections – Part 2: Temperature dependence, Atmos. Meas. Tech., 7, 625-636, 10.5194/amt-7-625-2014, 2014.

Spurr, R. J.: VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media, Journal of Quantitative Spectroscopy and Radiative Transfer, 102, 316-342, 10.1016/j.jqsrt.2006.05.005, 2006.

Stavrakou, T., Müller, J. F., Bauwens, M., De Smedt, I., Van Roozendael, M., Guenther, A., Wild, M., and Xia, X.: Isoprene emissions over Asia 1979-2012: impact of climate and land-use changes, Atmos. Chem. Phys., 14, 4587-4605, 10.5194/acp-14-4587-2014, 2014.

Thalman, R., and Volkamer, R.: Temperature dependent absorption cross-sections of O2–O2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Physical Chemistry Chemical Physics, 15, 15371-15381, 10.1039/C3CP50968K, 2013.

Vandaele, A. C., Hermans, C., Simon, P. C., Carleer, M., Colin, R., Fally, S., Mérienne, M. F., Jenouvrier, A., and Coquart, B.: Measurements of the NO2 absorption cross-section from 42 000 cm−1 to 10 000 cm−1 (238–1000 nm) at 220 K and 294

Deleted: 2018b

K, Journal of Quantitative Spectroscopy and Radiative Transfer, 59, 171-184, http://dx.doi.org/10.1016/S0022-4073(97)00168-4, 1998.

Vandaele, A. C., Hermans, C., and Fally, S.: Fourier transform measurements of SO2 absorption cross sections: II.: Temperature dependence in the 29 000–44 000 cm−1 (227–345 nm) region, Journal of Quantitative Spectroscopy and Radiative Transfer, 110, 2115-2126, http://dx.doi.org/10.1016/j.jqsrt.2009.05.006, 2009.

Veefkind, J. P., de Haan, J. F., Sneep, M., and Levelt, P. F.: Improvements to the OMI O2–O2 operational cloud algorithm and comparisons with ground-based radar–lidar observations, Atmos. Meas. Tech., 9, 6035-6049, 10.5194/amt-9-6035-2016, 2016.

Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, Atmos. Chem. Phys., 10, 10145-10160, 10.5194/acp-10-10145-2010, 2010.

Wedin, P. Å., and Lindström, P.: Methods and Software for Nonlinear Least Squares Problems, Umeå University. Institute of Information Processing, 1987.

Wilmouth, D. M., Hanisco, T. F., Donahue, N. M., and Anderson, J. G.: Fourier Transform Ultraviolet Spectroscopy of the A 2Π3/2 ← X 2Π3/2 Transition of BrO, The Journal of Physical Chemistry A, 103, 8935-8945, 10.1021/jp991651o, 1999.

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, Geophysical Research Letters, 33, n/a-n/a, 10.1029/2006GL026310, 2006.

Zhong, L., Louie, P. K. K., Zheng, J., Yuan, Z., Yue, D., Ho, J. W. K., and Lau, A. K. H.: Science–policy interplay: Air quality management in the Pearl River Delta region and Hong Kong, Atmospheric Environment, 76, 3-10, https://doi.org/10.1016/j.atmosenv.2013.03.012, 2013.

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

Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., Mickley, L. J., Yantosca, R. M., Sulprizio, M. P., De Smedt, I., González Abad, G., Chance, K., Li, C., Ferrare, R., Fried, A., Hair, J. W., Hanisco, T. F., Richter, D., Jo Scarino, A., Walega, J., Weibring, P., and Wolfe, G. M.: Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC4RS aircraft observations over the southeast US, Atmos. Chem. Phys., 16, 13477-13490, 10.5194/acp-16-13477-2016, 2016.

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

Deleted: < 56 km^2

Deleted: Seoul

Deleted: Spatial resolution

Deleted: Observation coverage

Deleted: ≥

Deleted: parameters

Deleted: 294

Deleted: far east

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] Datasets used in QA4ECV retrievals. Please refer to De Smedt et al. (2018) for details on other datasets and fitting options.

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.**

HCHO VCD uncertainty due to AMF uncertainty	**Clean area**	**Polluted area**
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%

Deleted: . Relative differences between OMI GEMS HCHO slant columns and OMHCHO and OMI BIRA slant columns in four regions.¶

Deleted: Region

Deleted: GEMS vs. OMHCHO

Deleted: GEMS vs. OMI BIRA

Table 3. Relative differences between OMI GEMS HCHO slant columns and OMHCHO and OMI QA4ECV slant columns in four regions.

Region	GEMS vs. OMHCHO				GEMS vs. OMI QA4ECV			
	Mar.	**Jun.**	**Sep.**	**Dec.**	**Mar.**	**Jun.**	**Sep.**	**Dec.**
Sumatra/Malaysia (95°-110°E, 0°-7°N)	-7%	-12%	-8%	-4%	-0.5%	-18%	-6%	-15%
Indochina Peninsula (97°-110°E, 10°-20°N)	8%	11%	9%	25%	-7%	-20%	-20%	-17%
China (110°-120°E, 30°-40°N)	12%	6%	10%	6%	-21%	-25%	-20%	-23%
Borneo (110°-118°E, 5°S-0°)	-7%	-13%	-7%	-1%	-9%	-13%	0.4%	-18%

Deleted: 40N

Table 4. Averaged HCHO VCDs (molecules cm^{-2}) from MAX-DOAS ground observations and OMI satellite data at OHP in France, Bremen in Germany, and Xianghe in China. For satellites, mean values are weighted by uncertainties and overlapped areas between satellite pixels and 0.25° grid cells for each site. Relative differences between OMI and MAX-DOAS are given in parentheses.

Site[a]	Class[b]	MAX-DOAS[c]	OMHCHO	OMI QA4ECV	OMI GEMS
OHP (44°N, 5.5°E)	Rural	7.5×10^{15}	5.8×10^{15} (-24%)	1.1×10^{16} (50%)	6.3×10^{15} (-16%)
Bremen (53°N, 9°E)	Urban	6.7×10^{15}	5.1×10^{15} (-23%)	9.3×10^{15} (40%)	6.1×10^{15} (-9%)
Xianghe (39°N, 117°E)	Sub-urban	1.6×10^{16}	1.0×10^{16} (-37%)	1.7×10^{16} (4%)	1.2×10^{16} (-25%)

[a] HCHO VCDs are averaged at OHP and Bremen in 2005 and at Xianghe in 2016.

[b] Class is assigned in a QA4ECV MAXDOAS website (http://uv-vis.aeronomie.be/groundbased/QA4ECV_MAXDOAS)

[c] Fitting windows of 336–359 nm and 324–359 nm are used at OHP and Bremen, and at Xianghe, respectively.

Deleted: ¶
Deleted: area
Deleted: of 0.25° at locations
Deleted: Site*
Deleted: Class
Deleted: DOAS**
Deleted: BIRA
Deleted: 4
Deleted: 3
Deleted: 5.9
Deleted: 9
Deleted: 32
Deleted: 61
Deleted: 25
Deleted: 6.0
Deleted: 8.8
Deleted: 7.2
Deleted: .2
Deleted: 17
Deleted: 21
Deleted: 0.8
Deleted: 1
Deleted: 2.2
Deleted: 3
Deleted: 8
Deleted: 41
Deleted: 19
Deleted: 30
Deleted: *
Deleted: in March, June, September 2005; at
Deleted: June and September
Deleted: ;
Deleted: May
Deleted: **

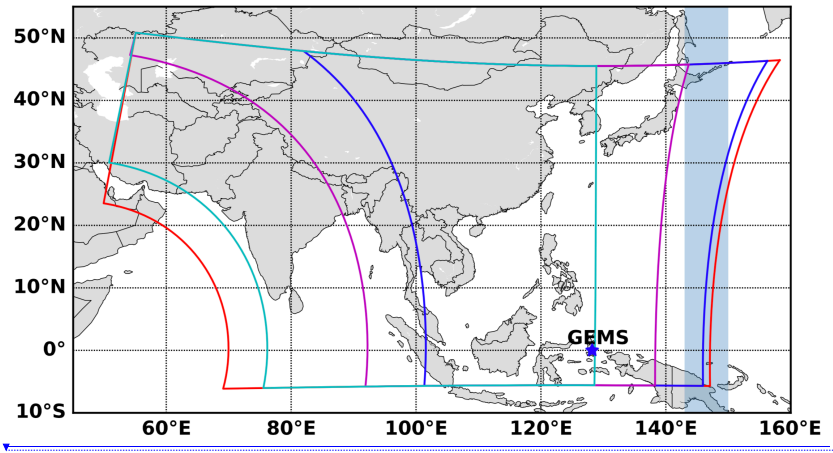
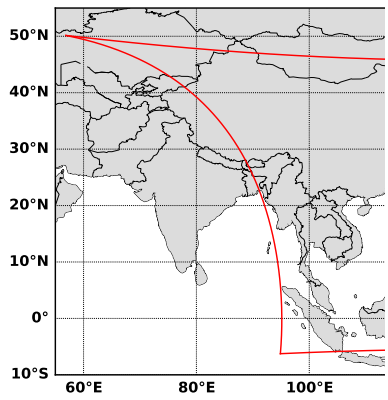


Figure 1. GEMS field of regard (red), nominal daily scan (blue), full central scan (magenta), full western scan (cyan), and GEMS location (blue star). Shaded areas (143-150°E) are regions for radiance references and common mode.

Deleted: ¶
¶



Deleted: FOR)

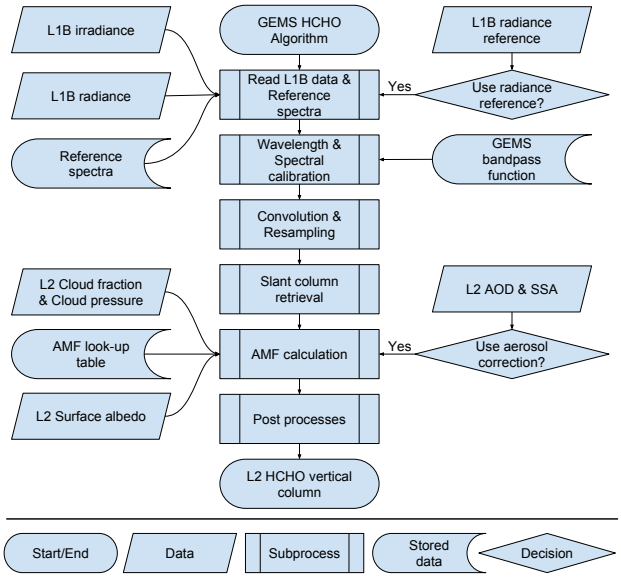


Figure 2. Flow chart for GEMS HCHO algorithm.

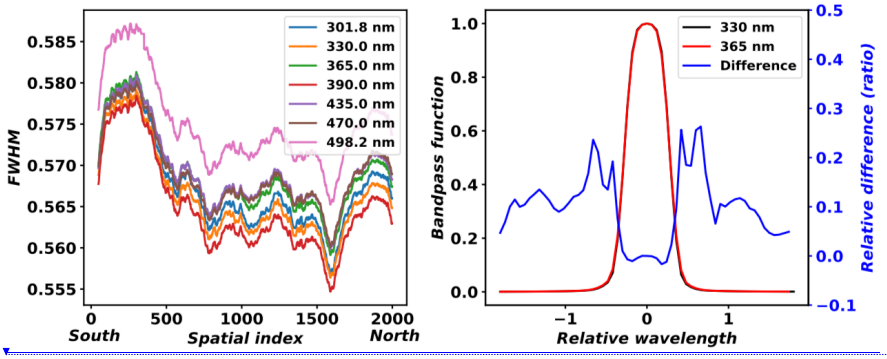
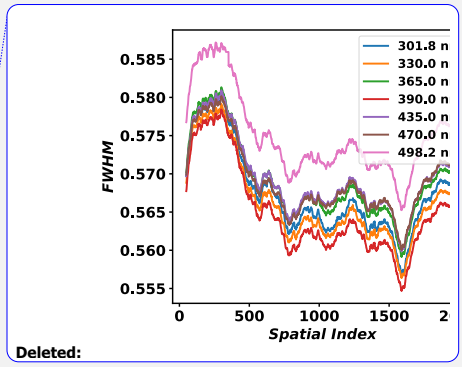


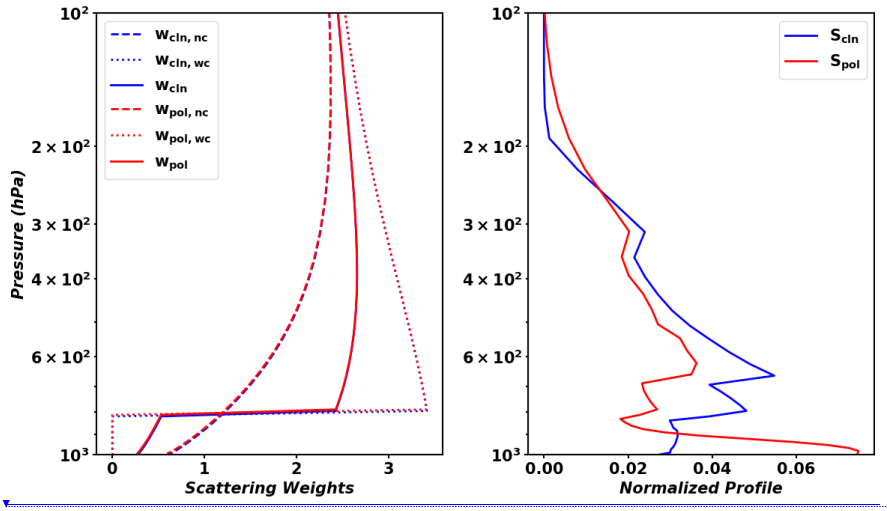
Figure 3. Full width at half maximum (FWHM) of GEMS bandpass as a function of spatial index of GEMS detection rows (left) and averaged bandpass functions for spatial indices at 330 and 365 nm and relative differences (right).



Deleted:

Deleted: to

Deleted: index



Deleted:

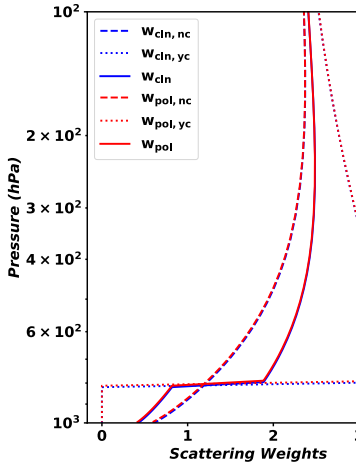


Figure 4. Scattering weights (left) and normalized profiles (right) from the AMF LUT over clean (blue) and polluted (red) grids. Dashed and dotted lines in the left figure indicate scattering weights with no cloud (nc) and with cloud (wc), respectively. The solid line in the left figure indicates scattering weights for a partial cloudy scene calculated using Eq. 10.

Deleted: without

Deleted: ,

Deleted: by

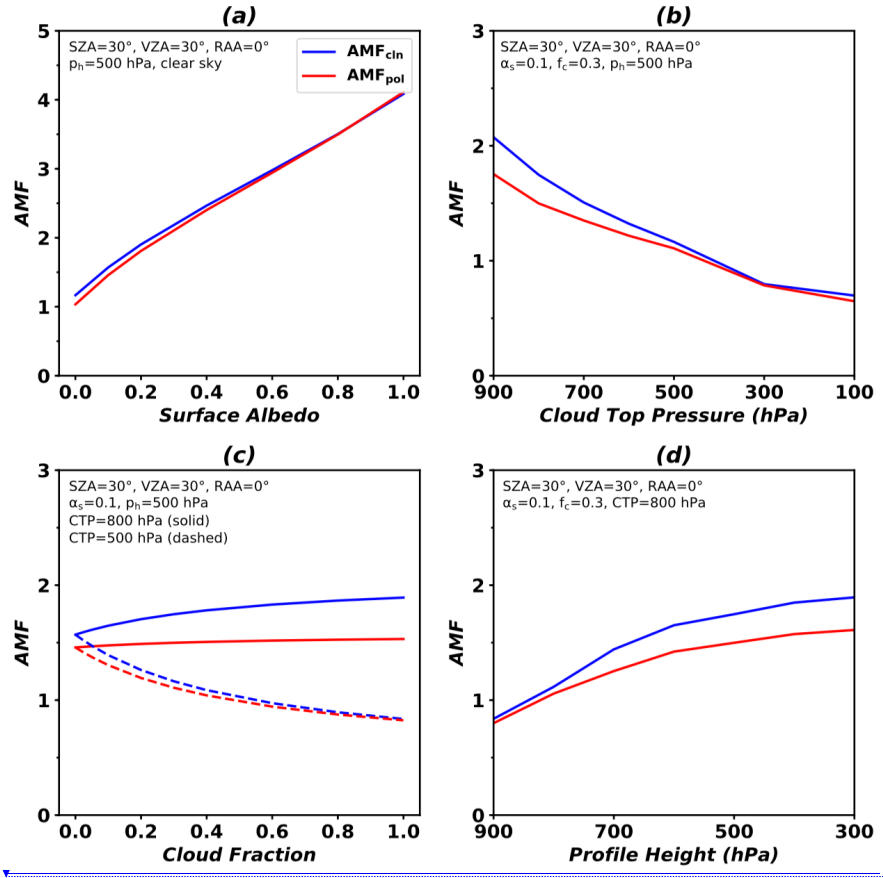
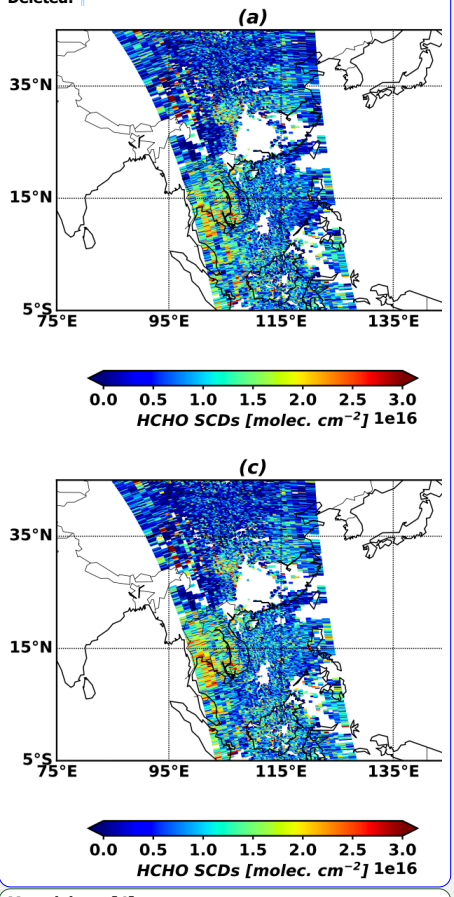


Figure 5. AMF variations as functions of (a) surface albedo, (b) cloud top pressure (CTP), (c) effective cloud fraction (f_c), and (d) profile height over clean (blue) and polluted (red) areas. Conditions of the AMF LUT are given in the figures. For sensitivity to surface albedo, cloud-free conditions are assumed. For sensitivity to cloud fraction, cloud top pressures are 800 hPa (solid line) and 500 hPa (dashed line).

Deleted: ¶



Moved down [4]: HCHO slant column densities (March 8, 2005, orbit 3436): (a) OMHCHO products, (b) GEMS algorithm without latitudinal bias corrections and (c) GEMS with latitudinal bias corrections, and (d) differences ((c) – (b)).¶

Deleted: ¶
¶
¶
¶
¶
¶
¶
... [1]

Deleted: and

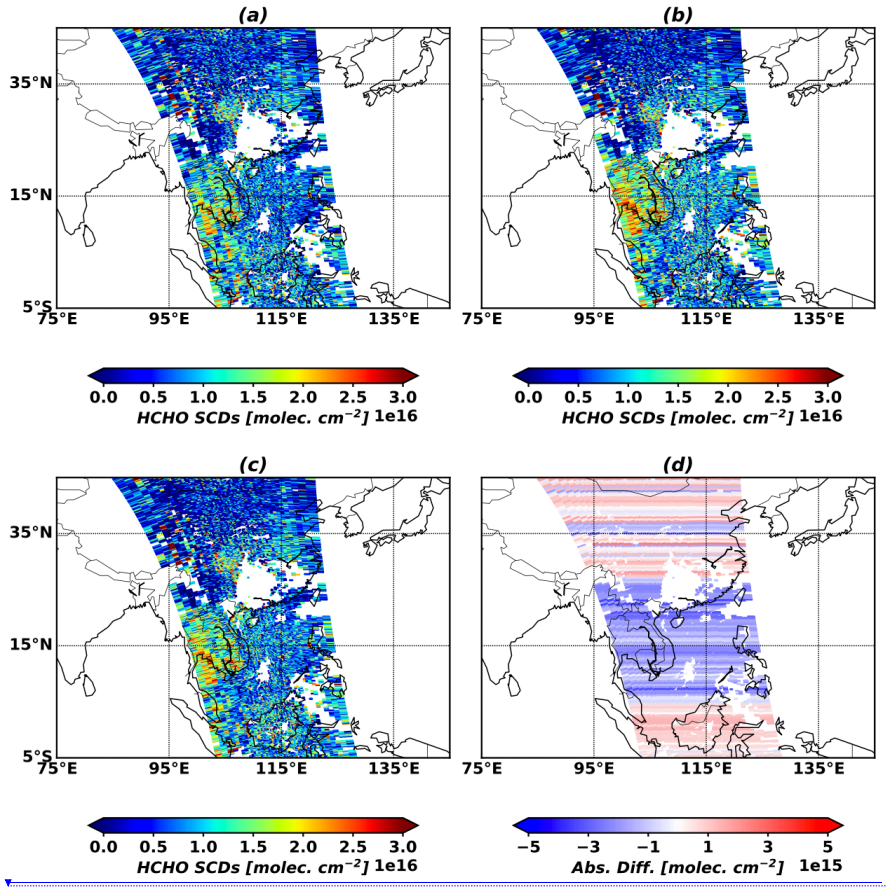
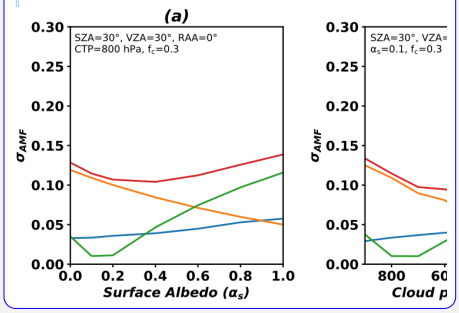


Figure 6. HCHO slant column densities (March 8, 2005, orbit 3436): (a) OMHCHO products, (b) GEMS algorithm without latitudinal bias corrections and (c) GEMS with latitudinal bias corrections, and (d) differences ((c) – (b)).

Deleted: ¶



Moved (insertion) [4]

Deleted: . AMF uncertainties from (a) surface albedo (α_s; blue), (b) cloud top pressure (CTP; orange), and (c) effective cloud fraction (f_c; green) and total AMF uncertainty (red) to variations of each parameter. Conditions of the AMF LUT are given in the figures.¶

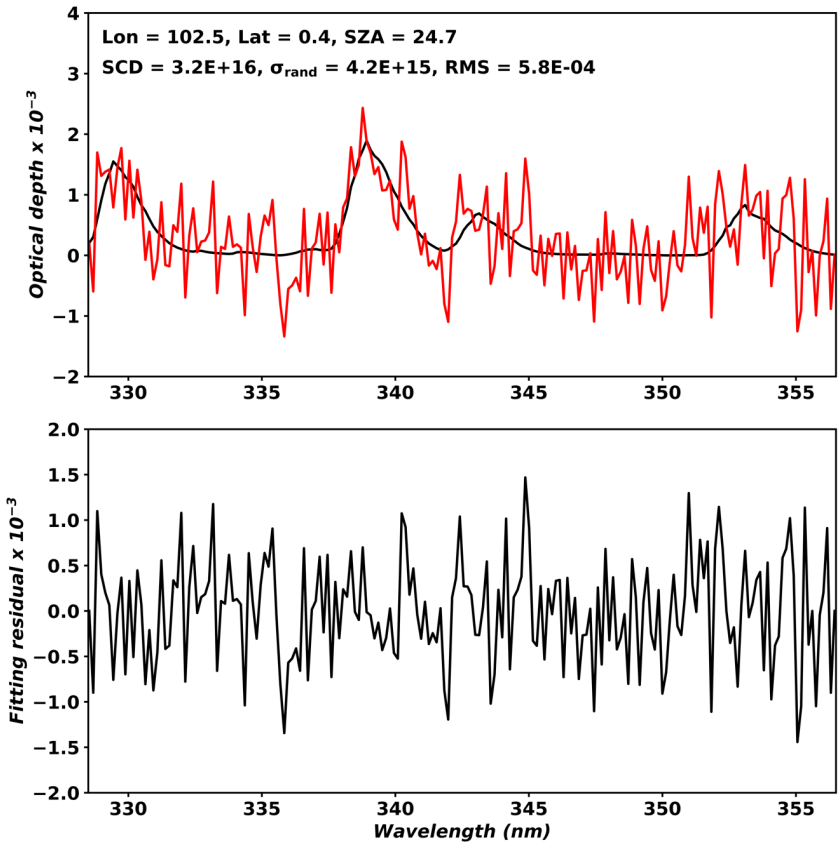


Figure 7. Fitted HCHO optical depth (top) and fitting residuals (bottom) on a pixel (March 23, 2005; orbit 3655) with main data quality flag of 0 and effective cloud fraction less than 0.3. In the top panel, the black solid line indicates optical depth, and the red solid line indicates HCHO optical depth plus fitting residuals.

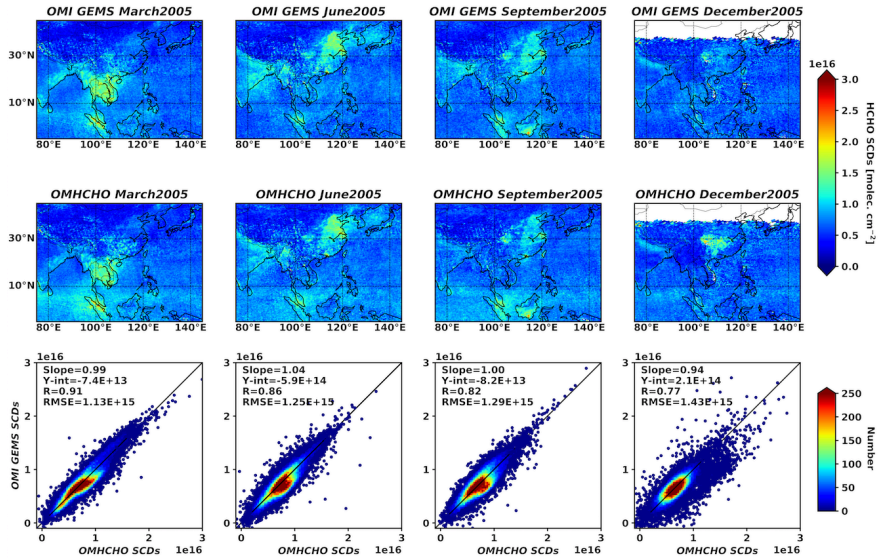


Figure 8. Monthly mean slant column densities (SCDs) from GEMS algorithm (1st row) and OMHCHO products (2nd row) for four months in 2005, and scatter plots between GEMS and OMHCHO. Statistics are given in the figures.

Deleted: columns

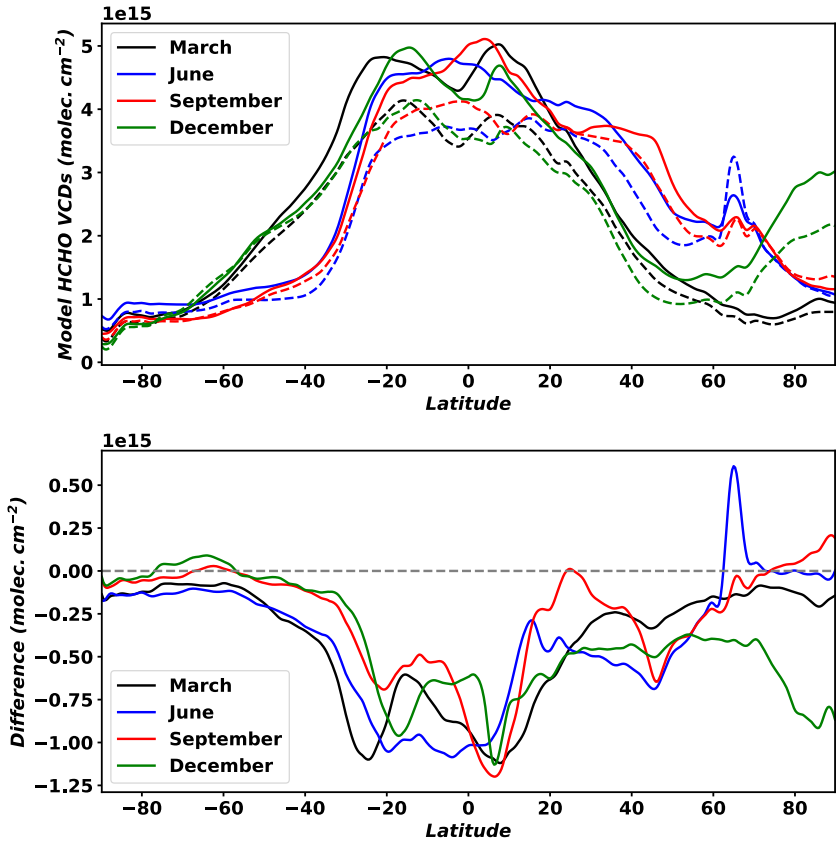


Figure 9. Simulated HCHO vertical column densities used in OMHCHO (solid lines) and GEMS algorithm (dashed lines) for background correction (top), and absolute differences of model results between GEMS algorithm and OMHCHO.

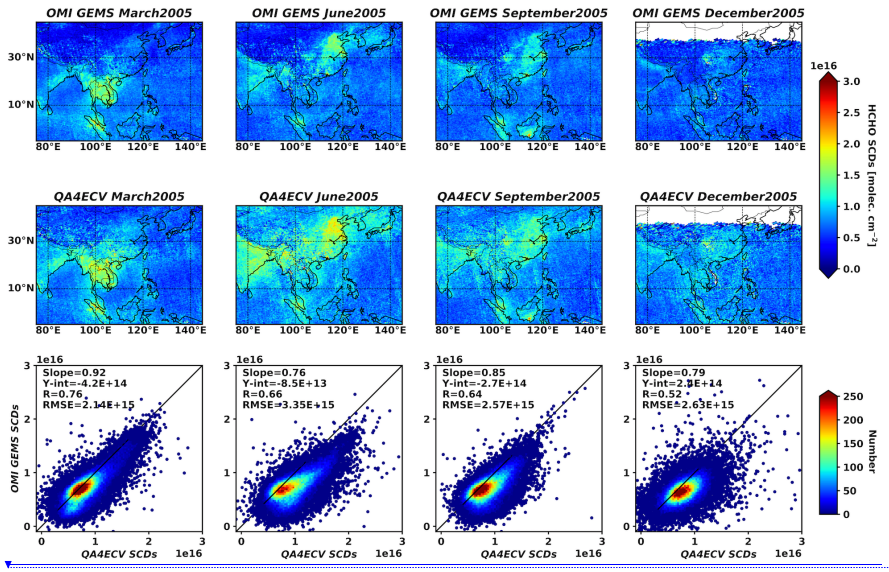
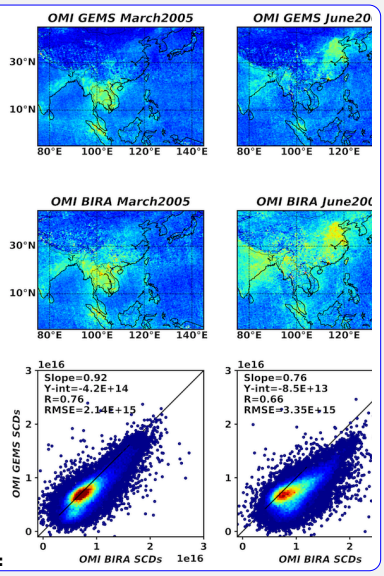


Figure 10. Monthly mean slant column densities (SCDs) from GEMS algorithm (1st row) and OMI QA4ECV products (2nd row) for four months in 2005, and scatter plots between GEMS and OMI QA4ECV. Statistics are given in the figures.

Deleted:



Moved down [5]: Figure 11.

Deleted: columns

Deleted: BIRA

Deleted: BIRA

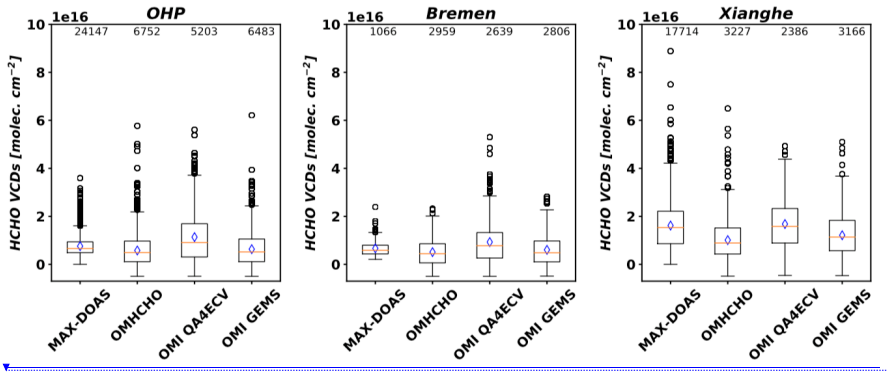
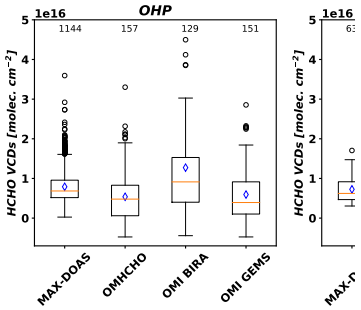


Figure 11. HCHO vertical columns from MAX-DOAS, OMHCHO, OMI QA4ECV, and OMI GEMS at OHP and Bremen in 2005, and at Xianghe in 2016. Orange lines are median values for each dataset, and blue diamonds are mean values. We computed mean values of each satellite product weighted by uncertainties and overlapped areas between satellite pixels and 0.25° grid cells for each site. Boundaries of boxes indicate first and last quantiles of datasets.

Deleted: Figure 12.



Moved (insertion) [5]

Deleted: BIRA

Deleted: (March, June,

Deleted: September

Deleted:), Bremen (June, September in 2005),

Deleted: (May

Deleted:).

Deleted: diamond markers

Deleted: For satellites,

Deleted: are

Deleted: area

Deleted: of 0.25° at locations.

Page 39: [1] Deleted	Author	5/16/19 2:00:00 PM