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.

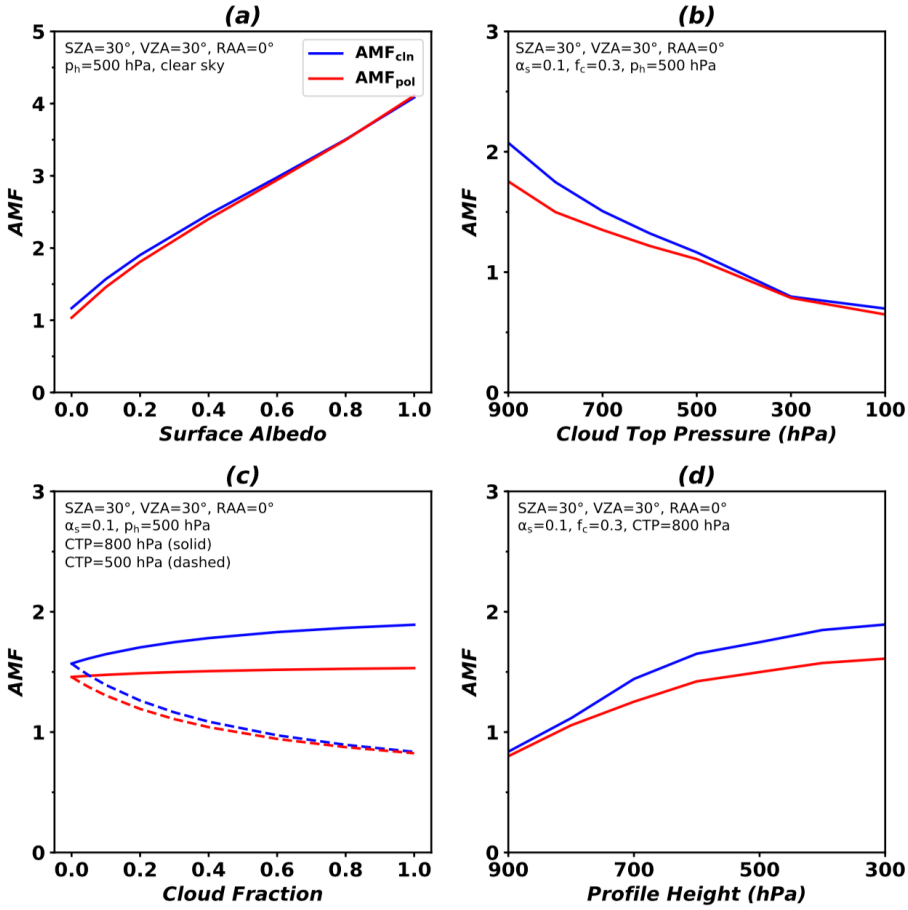

Figure 5. AMF variations as functions of (a) surface albedo, (b) cloud top pressure (CTP), (c) effective cloud fraction (f_c), and (d) profile height over clean (blue) and polluted (red) areas. Conditions of the AMF LUT are given in the figures. For sensitivity to surface albedo, cloud-free conditions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

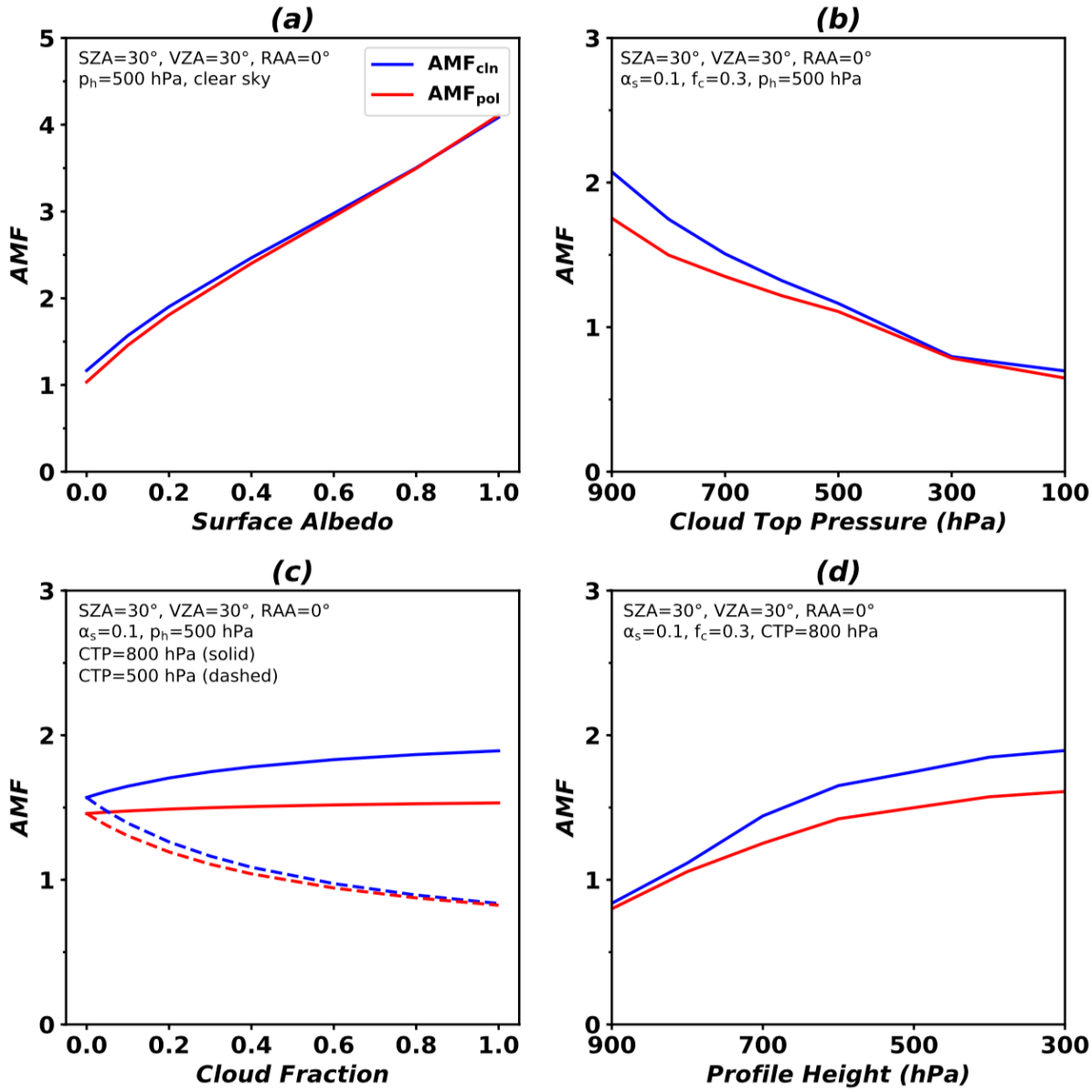

Figure 5. AMF variations as functions of (a) surface albedo, (b) cloud top pressure (CTP), (c) effective cloud fraction (f_c), and (d) profile height over clean (blue) and polluted (red) areas. Conditions of the AMF LUT are given in the figures. For sensitivity to surface albedo, cloud-free conditions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

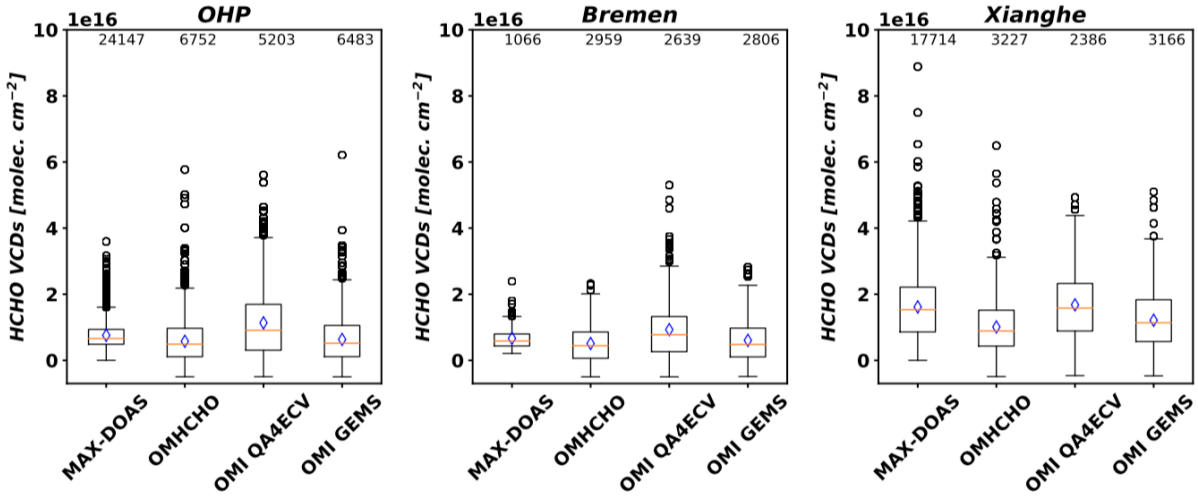

Figure 11. HCHO vertical columns from MAX-DOAS, OMHCHO, OMI QA4ECV, and OMI GEMS at OHP and Bremen in 2005, and at Xianghe in 2016. Orange lines are median values for each dataset, and blue diamonds are mean values. We computed mean values of each satellite product weighted by uncertainties and overlapped 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.