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:

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

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

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

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

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

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,

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

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