Interactive comment on "Retrieval algorithm for CO₂ and CH₄ column abundances from short-wavelength infrared spectral observations by the Greenhouse Gases Observing Satellite" by Y. Yoshida et al.

Anonymous Referee #3

Received and published: 28 December 2010

General comments:

The authors provide a detailed description of the current operational algorithm to retrieve CO2 and CH4 column abundances from GOSAT, which is the first satellite to specifically target greenhouse gases. The article is generally well written and organized. I recommend it to be published after addressing the following minor comments:

=> Thank you for your careful reading of our paper. The followings are our reply to your comments. The words with "double line (——)" were removed and those with "under bar (___)" were added.

Specific comments:

1. In line 9, page 4794, the description of SWIR is not accurate since SWIR observations is sensitive to the total abundance, I suggest changing it to "whereas SWIR observations are also sensitive to gas abundances near the surface" or "SWIR observations are sensitive to gas abundances down to the surface"

=> According to your suggestion, we revised as follows.
<p.4794, line 8>
"TIR observations are sensitive to CO_2 and CH_4 in the middle to upper troposphere, whereas SWIR observations are <u>also</u> sensitive to gas abundances near the surface."

2. Line 16, page 4794, the statement of precision of 1 percent order or better in column abundance is quite vague. Do you mean the precision for an individual satellite pixel? The references cited at the end of this sentence talked about monthly/weekly average on an 8 lat x 10 lon grid box. Please clarify this.

=> At the beginning of this sentence (p. 4794, line 15), we said "For spatially and temporally

averaged data".   However, as commented by other reviewers, spatial and temporal scales should be clarified.   Therefore, we revised this sentence as follows.

<p.4794, line 15>

"For spatially (few hundreds to thousands square kilometers) and temporally (weekly to monthly) averaged data, a precision of 1% order or better in CO_2 column abundances is required to improve our current knowledge of the surface CO_2 fluxes (Rayner and O'Brien, 2001; Houweling et al., 2004; Patra et al., 2003)."

3. It is confusing in lines 16-19, page 4794, to say precision of 1 percent or better is required and then says than random errors have less influence, because precision is normally referred to as random (-noise) errors. Please clarify this.

=> Revised as follows.

<p.4794, line 18>

"Random errors in the retrieved column abundance have ~~less influence~~ smaller impact on the estimated flux~~, but even a small~~ than bias in the retrieved value, which can lead to significant biases in the estimated flux (Patra et al., 2003; Houweling et al., 2005)."

4. Line 25, page 4794, I agree with the first reviewer. In addition, even retrievals with precision of 1-2 percent for individual pixels can be very useful for flux estimation according to the studies of Rayner and O'Brien, and Houwelling et al., 2004) because the precision can be much smaller when averaged weekly/monthly over a large area.

=> Revised as follows.

<p. 4794, line 25>

"~~However, the~~ The precision (1 to 2%) and accuracy (-1.5%) of CO_2 column abundances by current SCIAMACHY retrieval method (Schneising et al., 2008) are still insufficient for flux estimation, although the improvement of the retrieval algorithm has been continuously carried out."

5. I agree with reviewer 1 that the description of the DOAS method is incorrect.

=> Here, we want to mention that the optical path modification due to aerosol/cloud brings large

error.   We revised as follows.

<p. 4795, line 15>

"Many algorithms have been developed to retrieve column abundances of trace gases.   The differential optical absorption spectroscopy (DOAS) retrieval method has been widely used to retrieve the column abundances of trace gases (see Table 1 of Hönninger et al., 2004).   ~~The DOAS method works well when the measured signal is transmitted direct solar light.   However, if the measured signal is surface scattered light, the path radiance component modifies the equivalent optical path length, leading to large retrieval errors.~~   For the retrieval of SCIAMACHY data, several DOAS-based algorithms have been developed (Buchwitz et al., 2000; Barkley et al., 2006; Frankenberg et al., 2005a).   In those algorithms, however, aerosol scenarios and/or the surface albedo (two key parameters for the optical path length modification) are assumed.   <u>When the actual equivalent optical path length differs from the assumed one, it makes large errors in the retrieved results.</u>   Oshchepkov et al. (2008, 2009) proposed a new DOAS-based retrieval algorithm that simultaneously retrieves the photon path length probability density function parameters (Bril et al., 2007) to correct the optical path length modification effect."

6. About the descriptions of the spatial resolutions of TANSO-FTS and TANSO-CAI: for TANSO-FTS, the nadir footprint diameter is 10.5 km, so TANSO-FTS has circular pixel, right? How about TANSO-CAI, square or circular pixels? It was mentioned that the spatial resolutions of TANSO-CAI is 0.5, 0.5, 0.5 and 1.5 km, is it the same for both along-track and across-track directions? Please clarify.

=> The pixel shape for TANSO-FTS is circular and that for TANSO-CAI is rectangular.   For TANSO-CAI, its spatial resolution of 0.5/0.5/0.5/1.5 km is the nominal value for nadir pixels. We revised corresponding parts as follows.

<p,4796, line 24>

"The TANSO-FTS instantaneous field of view (IFOV) is 15.8 mrad, which corresponds to a nadir footprint diameter of about 10.5 km <u>circle</u> at sea level."

<p.4797, line 5>

"TANSO-CAI <u>is a push broom imager and</u> has four narrow bands in the near-ultraviolet to near-infrared regions at 0.38, 0.674, 0.87, and 1.6 µm (TANSO-CAI Band 1, 2, 3, and 4), with spatial resolutions of 0.5, 0.5, 0.5, and 1.5 km, respectively<u>, for nadir pixels.</u>"

7. Line 12-15, page 4798, it was mentioned that the high resolution solar irradiance database was used as the incident solar spectrum. Does the GOSAT measure its own solar irradiance at the GOSAT resolution? Because the absolute calibration of this high resolution reference might not be so good relative to the measured GOSAT radiance, large errors might occur. The high resolution solar irradiance is very good to be used to perform wavelength calibration, calculate ring effect, convolve radiances from high-resolution to instrument spectral resolution, but might not be good as incident solar spectrum due to systematic differences between it and measured GOSAT measurements.

=> This issue strongly relates with the radiometric calibration of the instrument. The most important point for the retrievals is the consistency between the radiometric calibration and the solar irradiance database. For the case of TANSO-FTS, it measures the diffuser-board-reflected solar irradiance in every orbit, and these data are used for the evaluation of the degradation of the TANSO-FTS (Kuze et al. 2009). The evaluated degradation as a function of time is used in the retrieval analysis. As for the degradation correction, we add following sentence at the end of section 2.1.
<p.4797, line 8>
"Details on TANSO-FTS and TANSO-CAI, along with the conversion method from raw interferograms to spectra (i.e., Level 1 processing) and pre-launch calibration results, are described by Kuze et al. (2009). Although the initial geometric and radiometric calibrations for TANSO-CAI have been performed, those for TANSO-FTS are still underway. The SWIR Level 2 V01.10, V01.20, and V01.30 only correct the radiometric degradation of the TANSO-FTS as a function of time."

8. Line 25, page 4798, please briefly explain why 2.1 um band is not used in the CO2 retrievals.

=> This is just because we did not finish the confirmation of the spectroscopic database for the 2.0-μm CO_2 absorption band. We plan to include this band in future (provably next) version.

9. Lines 1-6, page 4803, what are the initial values for lambda? Is the final lambda zero? If not, what are the typical values of lambda in the last iteration? What are the typical values for the diagonal scaling matrix? Is it fixed for all the retrievals? Please provide more detail so the reader can clearly understand the algorithm.

=> The initial value of lambda is zero. The final lambda vary scene by scene, and its value is within $10^{-6} \sim 10^{-3}$ in general. The value of lambda just determines the direction for next step (between the Gauss-Newton direction and the gradient-descent direction) and doesn't relate to the retrieved value, therefore we don't add the information of the final lambda. More detail formula for the scaling matrix is given as follows.

<p.4803, line 1>

"where the subscript i denotes the index of the i-th iteration; $\mathbf{K}$ is the Jacobian matrix, which is the derivative of the forward model as a function of the state vector $\mathbf{x}$, i.e., $\mathbf{K} = \partial \mathbf{F}(\mathbf{x}, \mathbf{b})/\partial \mathbf{x}$; λ is a non-negative factor and chosen at each step to minimize the cost function (initial value of λ is zero); and $\mathbf{D}^2 = diag(\mathbf{B}^T \mathbf{B})$ is a diagonal scaling matrix, where $\mathbf{B} = \left(\mathbf{T}_\varepsilon^{-T} \mathbf{K}_i \quad \mathbf{T}_{ainv} \right)^T$ is defined with the Cholesky decomposed matrices $\mathbf{S}_\varepsilon = \mathbf{T}_\varepsilon^T \mathbf{T}_\varepsilon$ and $\mathbf{S}_a^{-1} = \mathbf{T}_{ainv}^T \mathbf{T}_{ainv}$."

10. From Lines 20-25 on page 4803 to Lines 1-7 on page 4804: the use of weighting function might be confusing since it is normally means "Jacobians". Do you mean "profile of dry air partial columns (Wdry,i, i=1,n)". What are the units of the CO2/CH4 state vector Xx, volume mixing ratio? Some of the symbols are not clearly defined. I suggest to add "Xtarget" after "the column-averaged dry-air mole fraction" and add "SIGMAx" after "its error components"

=> We changed "weighting function" to "profile of dry air partial column". The unit of the CO_2 and CH_4 state vector is volume mixing ratio. We add the unit in Table 1. The sub-script X in eq. (12) has different meaning from the sub-script in eqs. (4-11, and 13), and it makes a confusion. We changed the sub-script X in eq. (12) to Y. The revised sentence becomes as follows.

<p.4803, line 20>

"Using the profile of dry air partial column $\mathbf{h}$, the column abundance V_{target}, the column-averaged dry-air mole fraction X_{target} and its error components σ_Y, and the column averaging kernel $\mathbf{a}_x$ are calculated as"

<p. 4804, line 2>

"$\sigma_Y = \dfrac{\sqrt{\mathbf{h}^T \mathbf{S}_Y \mathbf{h}}}{\mathbf{h}^T \mathbf{1}}$, $(Y = m, s, \text{ or } i)$,   (12)"

Table 1 State vector, its a priori, and the VCM for the TANSO-FTS SWIR retrieval. "SR" indicates data obtained by semi-real time processing, and "DB" indicates data prepared as monthly databases.

x (number of elements)		$\mathbf{x}_a$	$\mathbf{S}_a$
land	ocean		
CO_2 [ppmv] (15)		NIES TM (SR)	NIES TM (DB)
CH_4 [ppmv] (15)		NIES TM (SR)	NIES TM (DB)
H_2O [ppmv] (15)		GPV (SR)	GPV (DB)
AOD (1)		SPRINTARS (SR)	$(1.0)^2$ (fixed)
surface pressure [hPa] (1)		GPV (SR)	$(20)^2$ (fixed)
temperature profile bias [K] (1)		0.0 (fixed)	$(20)^2$ (fixed)
wavenumber dispersion (3)		0.0 (fixed)	$(10^{-5})^2$ (fixed)
surface albedo (22)		MODIS (DB)	$(1.0)^2$ (fixed)
	wind speed [m/s] (1)	GPV (SR)	GPV (DB)
	adjustment factor (2)	1.0 (fixed)	$(0.05)^2$ (fixed)

11. Also from lines 20-25, on page 4803, it is not clear about how partial dry air column abundance (e.g., h) is derived. In my understanding, it should be related to the retrieved surface pressure from the oxygen-A band, as well as the temperature profiles assumed and the retrieved temperature bias, right? If this is the case, when taking the ratio of x to h to calculate the Xtaget, the error in h should be propagated to calculate SIGMAx, but I did not see it in equation (12). Please provide more detail about this or to clarify this.

=> The partial dry air column abundance can be calculated from retrieved surface pressure and retrieved water vapor profile. In general, the retrieved temperature profile shift is enough small, therefore, we use the a priori temperature profile to calculate the partial dry air column abundance. The error propagation due to the error in **h** is not considered now. It should be corrected in future version, however, eq. (12) is correct as the description of the current algorithm.

<p. 4804, line 5>

"where the sub-script *target* indicates the target gaseous species (CO_2 or CH_4), n is the number of layers of the retrieval vertical grid, $w_{dry,l}$ is the partial dry air column abundance of the l-th layer, and **1** is the n-form column-vector with elements of unity. In each iteration step, $w_{dry,l}$ is re-calculated from the solved surface pressure and H_2O concentration profile."

12. Lines 9-16 on page 4804: maybe you could add "The state vector and it's a priori and it's a priori

covariance matrix is summarized in table 1 and will be described subsequently" at the end of this paragraph so the reader know where to find before reading the following several paragraphs.

=> According to your suggestion, the last sentence in section 4.2 is moved to at the end of the first paragraph of section 4.2.

<p.4804, line 9>

"The current retrieval algorithm uses the TANSO-FTS SWIR spectra within three selected wavenumber regions: 12950 to 13200 cm^{-1} (hereafter called O_2 sub-band), 6180 to 6380 cm^{-1} (CO_2 sub-band), and 5900 to 6150 cm^{-1} (CH_4 sub-band). In addition to the CO_2 and CH_4 concentration profiles, the H_2O concentration profile, the ground surface albedo for land, the surface wind speed and the radiance adjustment factor for ocean, aerosol optical depth (AOD), surface pressure, temperature profile bias, and the wavenumber dispersion are included in the state vector as auxiliary parameters and retrieved simultaneously. <u>The state vector and it's a priori and it's a priori variance-covariance matrix (VCM) are summarized in Table 1 and will be described subsequently.</u>"

<p.4807, line 21>

"~~The state vector, its a priori, and the VCM are summarized in Table 1.~~"

13. For aerosol optical depth, which wavelength is AOD defined?

=> AOD at 1.6 μm is used here. We add this information as follows.

<p.4810, line 19>

"Therefore, the retrieval results of VCO_2 and VCH_4 are dismissed when the retrieved AOD <u>defined at the wavelength of 1.6-μm</u> is larger than 0.5."

14. Line 17-18 on page 4805, it is not clear what do you mean "the target reflectance should be considered as retrieved?" Do you mean that "the target reflectance should be retrieved"

=> Yes. Revised as follows.

<p.4805, line 15>

"The radiance level of the observed spectrum was highly variable according to the target reflectance, the solar zenith angle, and the satellite viewing angle. ~~Because~~ <u>Although</u> the last two angles are easily determined, the target reflectance <u>varies with surface-type and time and</u> should be ~~considered as~~ retrieved."

15. In equations 17 and 18, what are the g and beta, please define.

=> The $g(k)$ and $\beta(k)$ are the interpolated parameters for the parameterization of the multiple-scattering components. They don't have physical meanings.

<p.4808, line 11>

"where k is the total optical depth due to the gaseous absorption, k' is the absorption optical depth from the TOA to a layer where substantial scattering occurs, ~~and~~ $\xi_0(k)$ is the average value of ξ for each fixed k, <u>and $g(k)$ and $\beta(k)$ are the interpolated parameters</u>. Values for $g(k)$, $\beta(k)$, and $\xi_0(k)$ are <u>calculated and</u> tabulated from accurately calculated multiple-scattering radiances for a small set of k values in k space."

16. Line 9 on page 4810, it is not clear about how you reject the retrievals? Do you mean that the sum of DFSs of CH4 and CO2 are less than 1? Or do you mean if the DFS for CO2 is less than 1, then you reject the retrieved CO2 value and if the DFS for CH4 is less than 1, then you reject the retrieved CH4 value?

=> We mean if the DFS for CO_2 is less than 1, then the retrieved CO_2 value is rejected, and if the DFS for CH_4 is less than 1, then the retrieved CH_4 value is rejected. Revise the sentence as follows.

<p.4810, line 8>

"Therefore, the retrieved VCO_2 or VCH_4 are dismissed when the <u>each</u> DFS~~s~~ of ~~those~~ <u>target</u> gas~~es~~ ~~are~~ <u>is</u> less than unity because the observed TANSO-FTS spectrum does not have enough information to retrieve them, and hence the retrieved values depend on the a priori values."

17. Lines 20-21 on page 4813, In "the retrieved Xco2 and Xch4 show appropriate patterns of global distribution. . .", the authors describes some latitudinal/seasonal variations but not how they vary with surface type (e.g., land, ocean, desert), I suggest to add some brief descriptions in this aspect. Otherwise, change "global distribution" to "latitudinal". Also pointed out by the first reviewer, the abstract says "agree well with the current state of knowledge", but it is only weakly supported in the text. I also think that it is good to add some comparison with model (i.e., also a priori) to support this in a more quantitative manner.

=> We add model comparison by using the NIES TM. Figures which show the monthly

averaged global distribution (corresponding to Fig. 6) and the latitudinal distributions of zonal mean (corresponding to Fig. 7) are added.   Followings are the revised sentences.

<p.4793, line 15>

"The interhemispherical differences and the temporal variation patterns of the retrieved column abundances ~~agree well with the current state of knowledge~~ show similar features with the atmospheric transport model."

<p. 4813, line 19>

"The global distributions ~~of retrieval results~~ and the latitudinal distributions of zonal averages of retrieval results and simulated results using the NIES TM are shown in Figs. A, B, and C ~~respectively~~.   For comparison, the matched NIES TM data with the GOSAT retrieval is used. The retrieved XCO_2 and XCH_4 show appropriate patterns of the latitudinal distributions and seasonal variations, although the retrieval results have biases and show ~~relatively~~ large variabilities as compared with the NIES TM.   The variation of XCO_2 and XCH_4 in the longitudinal direction over the ocean is smaller than those over the land mainly due to the distributions and strengths of sources and sinks of these gases.   Although the elements of the state vector are different for land cases and ocean cases, no clear gaps are found around the coastline."

<p.4815, line 7>

"The interhemispherical differences and the temporal variations of retrieved XCO_2 and XCH_4 ~~agree well with the current state of knowledge~~ show similar patterns with those simulated with the NIES TM, although there exist bias and amplitude difference."
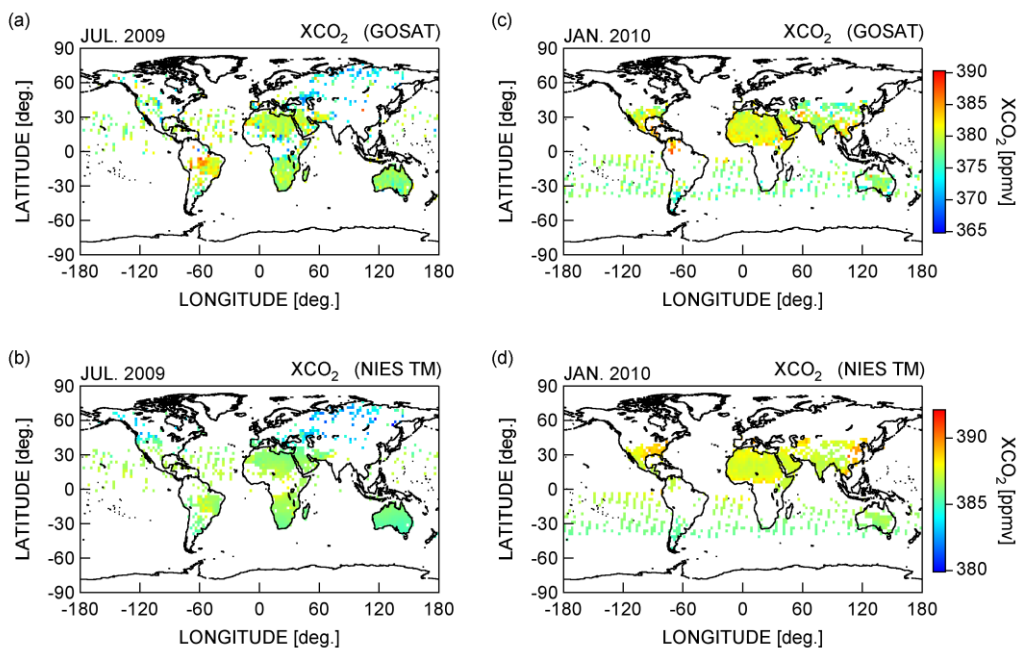
Fig. A   Monthly average of the XCO_2 [ppmv] retrieved by GOSAT (a, c) and simulated by the NIES TM (b, d) within a 2.5 x 2.5-degree grid box.   A blank indicates that no valid retrieval result was available within the grid box.   Different color-scales are used for GOSAT retrieval and the NIES TM simulation.

Fig. B   Monthly average of the XCH_4 [ppbv] retrieved by GOSAT (a, c) and simulated by the NIES TM (b, d) within a 2.5 x 2.5-degree grid box.   A blank indicates that no valid retrieval result was available within the grid box.   Different color-scales are used for GOSAT retrieval and the NIES TM simulation.
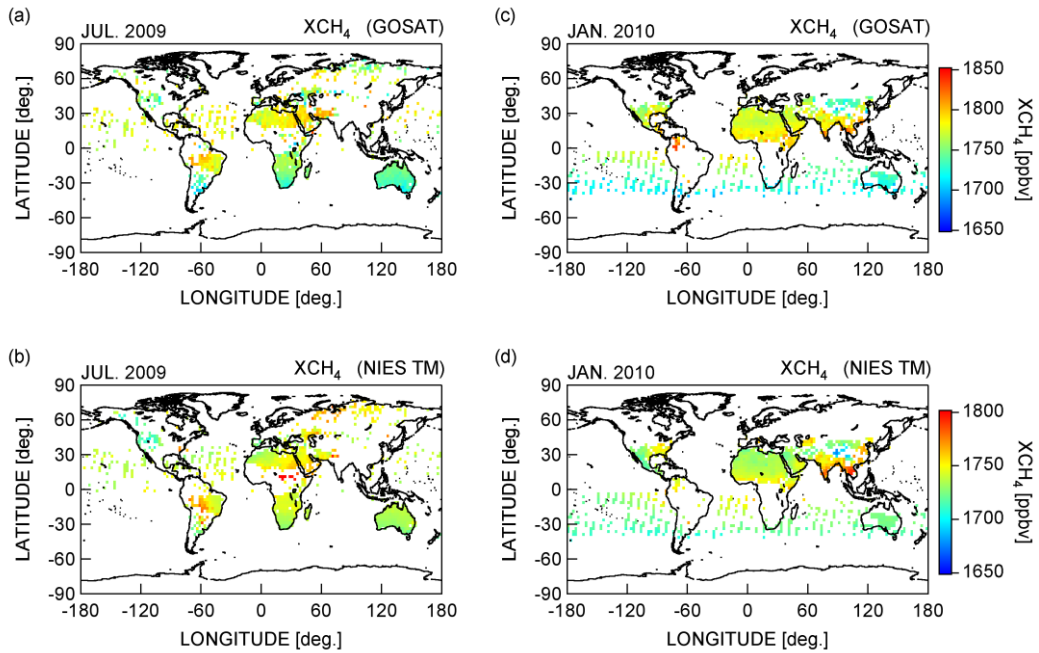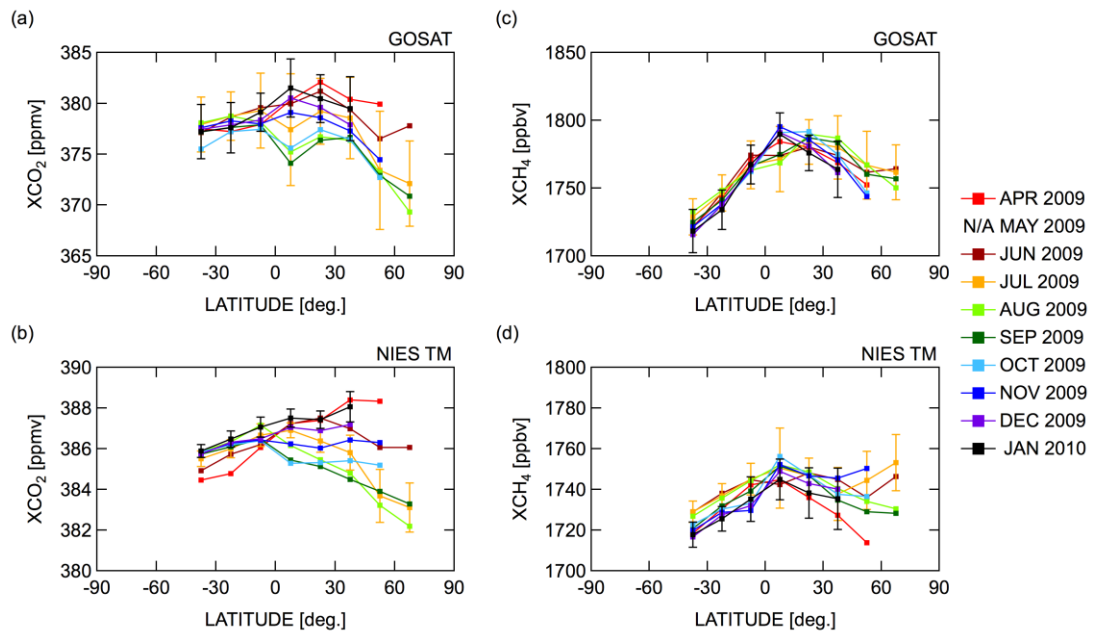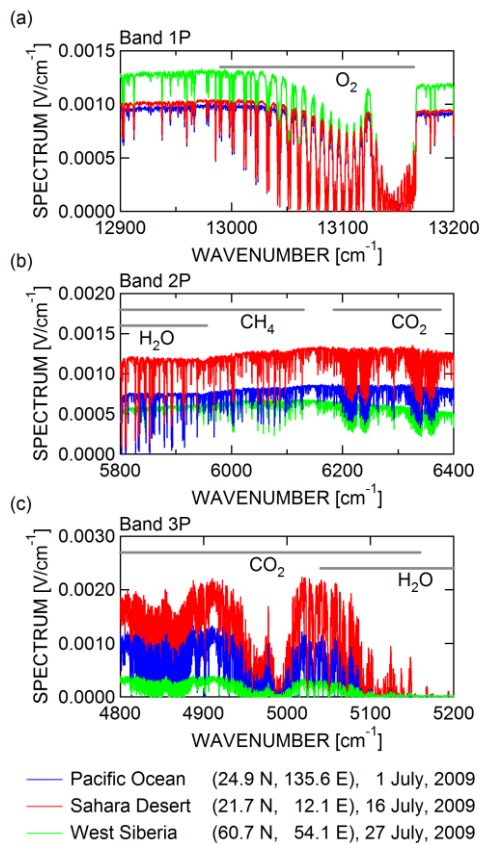
Fig. C   Latitudinal distributions of zonal mean of the retrieved and simulated XCO$_2$ [ppmv] (a, b) and XCH$_4$ [ppbv] (c, d).   The standard deviations of zonal variation for July 2009 and January 2010 are plotted as error bars.

18. For figure 1, it might be better to use a different color for each scene?

=> Done.

(a)

Band 1P

SPECTRUM [V/cm⁻¹]

0.0015

0.0010

0.0005

0.0000

O_2

12900   13000   13100   13200

WAVENUMBER [cm⁻¹]

(b)

Band 2P

SPECTRUM [V/cm⁻¹]

0.0020

0.0015

0.0010

0.0005

0.0000

H_2O CH_4 CO_2

5800   6000   6200   6400

WAVENUMBER [cm⁻¹]

(c)

Band 3P

SPECTRUM [V/cm⁻¹]

0.0030

0.0020

0.0010

0.0000

CO_2 H_2O

4800   4900   5000   5100   5200

WAVENUMBER [cm⁻¹]

— Pacific Ocean   (24.9 N, 135.6 E), 1 July, 2009
— Sahara Desert   (21.7 N, 12.1 E), 16 July, 2009
— West Siberia   (60.7 N, 54.1 E), 27 July, 2009

19. In Figure 5, label for color bar of the last panel should be "Uncertainty reduction (CH4)"

=> Done.

(g) A PRIORI UNCERTAINTY (CH$_4$) [ppbv]

(h) A POSTERIORI UNCERTAINTY (CH$_4$) [ppbv]

(i) UNCERTAINTY REDUCTION (CH$_4$)
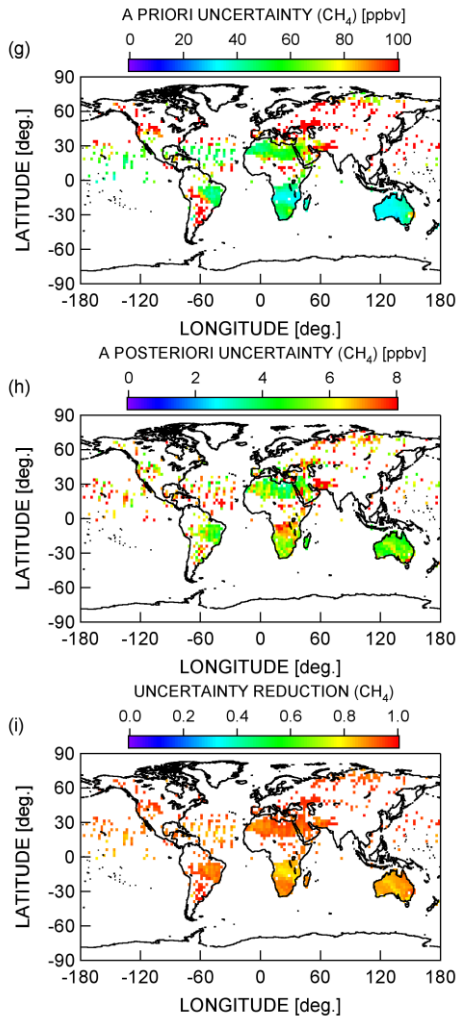
Technical comments:

1. In line 24, Page 4793, I suggest to rephrase the "high spatiotemporal-resolution monitoring of global greenhouse gas distributions" to "to monitoring the global distributions of greenhouse gases at high spatiotemporal resolution"

=> According to your comment, we revised as follows.

<p.4793, line 23>

"Satellite measurement is one of the most effective approaches to ~~high spatiotemporal-resolution monitoring of global greenhouse gas distributions~~ monitoring the global distributions of greenhouse gases at high spatiotemporal resolution."

2. Line 26, page 4796, change +/- to "±"

=> Done.

<p.4796, line 24>

"The TANSO-FTS has a pointing mechanism, which makes it possible to observe the off-nadir direction within the pointing mirror driving angles of ~~+/-~~ ± 35 degrees in the cross-track direction and ~~+/-~~ ± 20 degrees in the along-track direction."

3. Line 2 page 4802, "Consider . . . TANSO-FTS." is not a complete sentence, change "." to ","

=> Done.

<p.4802, line 2>

"Consider eight lines of sight (LOSs) that are shifted by 1.3 mrad around the center LOS of TANSO-FTS.~~We~~ , we calculate the mean surface elevation within TANSO-FTS IFOV for each LOS."