Reply to the comments by Referee #2

To Referee #2,

Overall, this is a good paper dealing with difficult but necessary bias corrections to TANSO-FTS observations of mid-troposphere CO_2. It's a tricky subject, but the methodology is generally sound. However, the paper is difficult to follow in some sections, and in many cases, the figures need some improvement and clarification. I would recommend publication after some revisions in the text, and if the authors could better address the issue of the number of layers in the forward model (see comment for page 10, line 32 below.)

We appreciate you reading our paper carefully and giving valuable comments and suggestions. We have considered your recommendations for revisions and made the necessary changes. The major points that we deal with in the revised manuscript are as follows:

1. Following your advice, we have added Table 1 to show representative pressure levels of each of the retrieval grid layers of GOSAT/TANSO-FTS thermal infrared (TIR) version 1 (V1) Level 2 (L2) CO_2 product.
2. Relating to the above, we have referred to the retrieval grid layers by the representative pressure levels throughout the text.

Individual responses to the Referee's comments are listed below.

General comment: Throughout the paper, the authors refer to the retrieval layers by number (layer 3, layer 4, etc.), rather than, say, its log mean pressure. These layer numbers are specific to their algorithm, and referencing the layers by number is a little burdensome to the reader, even where the pressures are provided. For example, Page 6, line 23 reads "Saitoh et al. (2016) showed that TIR V1 CO_2 data agreed well with CME level flight CO_2 data in the UT region corresponding to retrieval layers 9 and 10." This would read better if the pressures were given instead of the layer numbers. I suggest they prepare a table listing the retrieval layer numbers, layer boundary pressures, and the log-mean pressures of the layers (similar to Table 1 of Saitoh et al., 2016), and then just refer to a layer by its mean pressure rather than its number.

Reply:

We greatly appreciate your comments. As described above, we have added Table 1 to show representative pressure levels of each of the retrieval grid layers used in the GOSAT/TANSO-FTS TIR V1 L2 CO_2 retrieval processing and referred to the retrieval grid layers by the representative pressure levels instead of retrieval grid numbers. In Table 1, we have

kept the retrieval grid numbers for the convenience of TIR CO_2 data users. In the TIR V1 L2 CO_2 retrieval algorithm, we have calculated representative pressure level P_{rlay}, which is thermodynamically mean pressure level, by the following expression [Gallery et al., 1983]:

$$P_{rlay} = \{\frac{H_P H_\rho}{H_P + H_\rho}(P_{rlev\_j}\,\rho_{rlev\_j} - P_{rlev\_j+1}\,\rho_{rlev\_j+1})\}/\{H_\rho(\rho_{rlev\_j} - \rho_{rlev\_j+1})\}$$

$$H_P = \frac{-\Delta z}{\log_e(P_{rlev\_j+1}/P_{rlev\_j})}$$

$$H_\rho = \frac{-\Delta z}{\log_e(\rho_{rlev\_j+1}/\rho_{rlev\_j})}$$

$$\Delta z = \log_e \frac{P_{rlev\_j+1}}{P_{rlev\_j}} \times -\frac{Rd}{g} \times \frac{\left|T_{rlev\_j+1} - T_{rlev\_j}\right|}{2}$$

,

where P_{rlev_j} and P_{rlev_j+1} are lower and upper pressure levels of each retrieval grid layer, respectively, T_{rlev_j} and T_{rlev_j+1} are temperatures at the two pressure levels, ρ_{rlev_j} and ρ_{rlev_j+1} are air densities at the two pressure levels, Rd is the gas constant, and g is the acceleration of gravity. Representative pressure levels change depending on temperature, which are stored in each of the TIR V1 L2 CO_2 data files, but their variabilities are quite small. In Table 1, we have presented the averages of representative pressure levels of each retrieval grid layer calculated by using all GOSAT/TANSO-FTS measurements in 2010.

Page 1, line 14: "…good spatial representability." It's not obvious what 'representability' means here. Would "resolution and precision" be a better phrase to use?

Reply:

CO_2 concentrations in the free troposphere are well mixed compared to the concentrations near the surface and less affected by local point sources of CO_2; in that context, observations in the free troposphere can obtain CO_2 concentrations representative of regions, which can be dealt with in a global model estimating CO_2 surface fluxes. In the revised manuscript, we have modified the sentence to clarify this point as follows:

"CO_2 observations in the free troposphere can be useful for constraining CO_2 source and sink estimates at the surface due to their representativeness being away from local point sources of CO_2."

Page 1, line 24: "(retrieval layers 5−6), …" It's not necessary to get into the details of their retrieval method in the abstract.

Reply:

We have deleted the phrase in the abstract of the revised manuscript following your advice.

Page 2, line 3: Suggest changing "(e.g., Gurney et al., 2002 Gurney et al., 2004)" to "(e.g., Gurney et al., 2002; 2004)".

Reply:

Following your suggestion, we have modified the text in how to cite the references.

Page 2, line 24: "spatial representability." Again, not obvious what it means here.

Reply:

XCO_2 data obtained by measurements utilizing short-wave infrared (SWIR) band contained information on CO_2 concentrations near the surface compared to free tropospheric CO_2 measurements utilizing TIR band. However, satellite-borne sensors have relatively large field-of-views, and therefore their XCO_2 data are averaged concentrations in their field of views of several kilometers that are not too much affected by strong local point sources of CO_2. In the revised manuscript, we have modified the sentence as follows:

"Global XCO_2 data based on satellite observations are averaged concentrations in their field of views of several kilometers that are not too much affected by strong local point sources of CO_2, and have therefore been used to estimate surface CO_2 fluxes (Maksyutov et al., 2013; Saeki et al., 2013a; Chevallier et al., 2014; Basu et al., 2013, 2014; Takagi et al., 2014)."

Page 3, line 16: Suggest changing "…and has continued CO_2 and CH_4 operational measurements for approximately eight years." to "and has continued operational measurements of CO_2 and CH_4 for approximately eight years.

Reply:

Following your suggestion, we have modified the sentence.

Page 3, line 23: Suggest shortening "These studies showed the following: 1) TIR UT CO_2 data agreed…" to "These studies showed: 1) TIR UT CO_2 data agreed…"

Reply:

Following your suggestion, we have modified the sentence.

Page 5, line 14: Suggest more explanation of why the averaging kernels are applied to the CME data and then comparison made. This would be useful to the reader not well versed in averaging kernels etc.

Reply:

Following your advice, we have added more explanation of why we should apply TIR CO_2

averaging kernel functions to CME aircraft profiles as follows:

"Observations by satellite-borne nadir-viewing sensors like TANSO-FTS have much lower vertical resolution than aircraft observations. Therefore, we smoothed the CME_obs. profile to fit its vertical resolution to the vertical resolution of corresponding TIR CO_2 profile by applying TIR CO_2 averaging kernel functions (AK) to the CME_obs. profile, as follows (Rodgers and Connor, 2003):"

Page 6, Section 4.2: It's not obvious why an "average" averaging kernel can be applied and not sometimes be misleading. In addition to the effect of instrument parameters (SNR, spectral resolution, view angle etc.) and assuming clear scenes only, the averaging kernel could vary by temperature gradient and thermal contrast with the surface. How much does an averaging kernel vary within a grid box? It would help if the authors briefly explain why they're using an averaged AK here and discuss the limitations of doing so.

Reply:

We agree with you. TIR CO_2 averaging kernel functions depend on TIR measurement spectral noise, a priori CO_2 profile variability, and CO_2 Jacobians. In the TIR V1 L2 CO_2 retrieval algorithm, we set covariance matrices of the TIR measurement noise and a priori CO_2 profile in the same manner for all TIR CO_2 measurements, as described in Saitoh et al. (2016). The CO_2 Jacobians depend on temperature and CO_2 profiles, and therefore change with location and time. For a validation purpose based on one-by-one comparisons like TIR versus CME CO_2 profiles, we should apply corresponding TIR CO_2 averaging kernel functions, not averaged one. On the other hand, the purpose of comparisons between TIR and NICAM-TM CO_2 data is to evaluate the bias-correction values determined for each vertical layer, latitude band, and season. In addition, TIR CO_2 averaging kernel functions showed nearly identical structures with each other when collected for each 2.5° grid in one month, which means that applying the monthly averaged TIR CO_2 averaging kernel functions did not affect the conclusions of this study. From this standpoint, using monthly averaged TIR CO_2 averaging kernel functions instead of individual one is enough for our purpose. In the revised manuscript, we have added one paragraph in Section 4.2 and discussed the effect of using monthly averaged TIR CO_2 averaging kernel functions on our analysis. We appreciate your comments.

Page 7, line 14 "In addition, negative biases of TIR CO_2 data against NICAM-TM CO_2 data increased by 1 ppm or less per year in all seasons, judging from the mode values, although the increase in negative biases was not evident in the comparisons over airports shown in Figure 6." I did not quite understand what is meant by this. Do they mean the bias varied by 1ppm or less?

Reply:

We intended to say the following: negative biases of TIR CO_2 data against NICAM-TM CO_2 data

seemed to increase over time, judging from each of the mode values for the three years and the rate of the increase was around and less than 1 ppm; however, the increase in the negative biases against NICAM-TM CO_2 data was not evident as was the case with the negative biases against CME CO_2 data discussed in Section 5.1. In the revised manuscript, we have modified the sentence as follows:

"In addition, negative biases of TIR CO_2 data against NICAM-TM CO_2 data in all seasons slightly increased over time, judging from the mode values, although the increase in negative biases was not also evident as in the comparisons over airports shown in Figure 6."

Page 8, line 27: Typo: "… in the LT and ML regions." Did they mean "MT" regions?

Reply:

We have modified the sentence. We appreciate you pointing out our mistake.

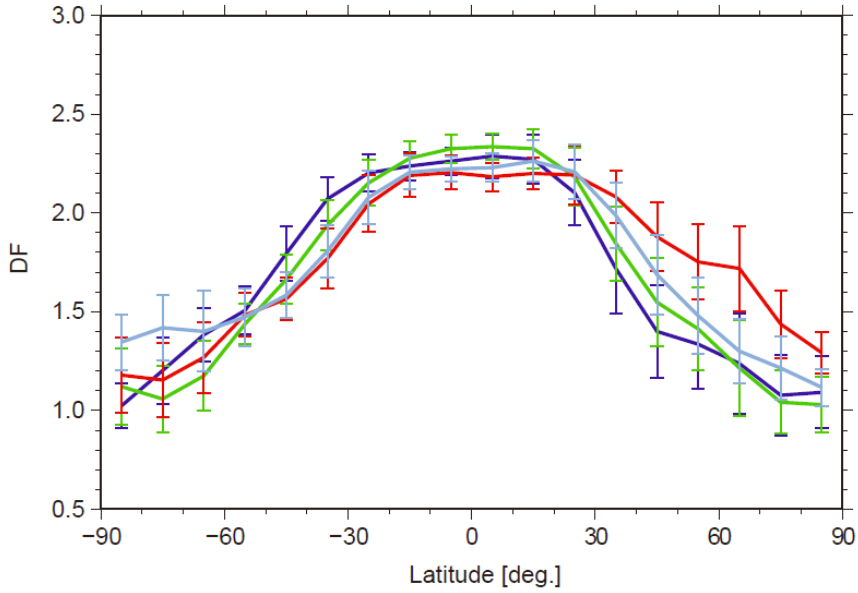
Page 9, line 13: "As shown in Figure 6, the largest negative biases in TIR V1 CO_2 data existed in the MT region in middle and low latitudes during spring and summer, where TANSO-FTS TIR measurements have relatively large sensitivity to CO_2 concentrations and thus the retrievals are less constrained to a priori concentrations." Some kind of comparison is in order to quantify the difference in CO_2 sensitivity here – say average row-sum of averaging kernels, or total DOFS as a function of latitude.

Reply:

We totally agree with you. We have modified the related sentences for consistency with the sentences in the second paragraph of Section 5.1, and then provided information on degrees of freedom of TIR V1 CO_2 data in low latitudes where the largest negative biases existed:

"As shown in Figure 6, the largest negative biases in TIR V1 CO_2 data existed in the MT region in low latitudes (20°S−20°N) during the JJA season. Degrees of freedom (DF) of TIR V1 CO_2 data were highest in low latitudes, exceeding 2.2 in all seasons, which means retrieved CO_2 concentrations there contained more information coming from TANSO-FTS TIR L1B spectra and thus were relatively less constrained to a priori concentrations."

The DF values have been referred from the below figure that shows monthly averaged DF values for each 10° latitude in January (blue), April (green), July (red), and October (light blue) in 2010.



Reference figure. Monthly averaged DF values of TIR V1 CO_2 data for each 10° latitude in January, April, July, and October 2010, shown by blue, green, red, and light blue lines, respectively. Here, GOSAT/TANSO-FTS observations with high elevated areas (surface pressure less than 736 hPa) were excluded.

Page 9, line 15: "This implies that biases in L1B spectra are a major cause of the negative biases in retrieved CO_2 concentrations, as Saitoh et al. (2016) noted in the UT region." The wording is confusing. Does this mean there are biases in the L1b radiances related to latitude and season, or are there fitting biases from the retrieval algorithm? Judging from the rest of the paragraph where the authors write about retrieval of surface parameters, I think they're referring to fitting bias, but whatever the bias is, it should be explicitly described.

Reply:

According to comparisons between TANSO-FTS TIR and S-HIS radiance spectra (Kataoka et al., 2014) and theoretical radiance error estimations (Kuze et al., 2016), TANSO-FTS TIR L1B radiance spectra had considerable biases. In low latitudes, retrieved CO_2 data contained more information coming from TANSO-FTS TIR L1B spectra judging from their highest DF values. This means that the effect of the L1B radiance biases should be also largest in TIR CO_2 data in low latitudes. The magnitude of the TIR L1B radiance biases may change by scene, but we have not yet drawn any conclusion on the dependence of the radiance biases on time, location, viewing angle, thermal condition of TANSO-FTS instrument, and so on. As the related three paragraphs in Discussion were less organized, we have reorganized the discussion on the relation between L1B radiance biases and L2 CO_2 negative biases against CME CO_2 data in the revised manuscript.

Page 10, line 4: "From these results, we conclude that using the 10-μm band in conjunction with the 15-μm

and 9-μm bands in the V1 retrieval algorithm is a probable cause of the negative biases in retrieved CO_2 concentrations in the LT and MT regions." While I don't disagree with this, this would be more convincing if the authors compared their results using the different mixes of CO_2 bands directly against the aircraft measurements.

Reply:

We totally agree with you. We have also showed nearby CME CO_2 profiles by gray lines in Figure 10 of the revised manuscript other than TIR CO_2 retrieval results. We appreciate your suggestion.

Page 10, Line 13: "According to Figure 13 in Kuze et al. (2016), there was no distinct uncertainty in the 10-μm band in the latest version of the TANSO-FTS TIR spectra." The wording of this leaves me uncertain of what they're claiming. Uncertainty of linestrengths or low fitting residual? Are they saying that using the 10 micron band of CO_2 does not add significant bias? This should be clarified.

Reply:

Kuze et al. (2016) performed theoretical estimation of radiance biases of TANSO-FTS TIR L1B V161 and newer version V201 spectra. The radiance biases inherent in the TANSO-FTS TIR L1B spectra were attributable to several calibration issues, mainly due to polarization correction. According to theoretical calculations shown in Figure 13 in Kuze et al. (2016), there were no distinct radiance biases in the 10-μm band (930−990 cm^{-1}) in the latest version of the TANSO-FTS TIR spectra. If it is true for observed TIR radiances, our test retrievals imply that simultaneous retrieval of surface parameters for TIR spectra at the 10-μm band with less radiance bias worsened CO_2 retrieval results. We have clearly stated this in the revised manuscript.

Page 10, paragraph beginning line 17: As noted earlier, it would really help the reader if the authors referred to the retrieval layers by pressure and not layer number.

Reply:

Following your advice, we have referred to the lower and upper pressure levels of the two retrieval grid layers that we focused on.

Page 10, line 32: "In retrieval from TIR spectra, the more atmospheric layers in which we retrieve CO_2 concentrations, the lower the information content of the retrieval result in each layer becomes; as a result, the retrieved concentrations are constrained by a priori model data. Thus, there is a high possibility of large biases in retrieved TIR CO_2 concentrations in low latitudes." This assertion needs to be tested. It is true that with more layers, the information is spread out more, but the overall information content, as measured by the degrees-of-freedom-of-signal (trace of the averaging kernel) can be the same or very similar, as can the

retrieved profiles (depending on what the off-diagonals are for the a priori background covariance.) It's quite possible that if the background a priori is biased, then a TIR retrieval can also be biased not because of the number of retrieval layers, but, particularly at low latitudes, because of water vapor interference, undetected boundary- layer clouds changing the thermal contrast with the surface, or biases in the temperature. Again, this needs to be tested, or the statement removed or at least reworded as a hypothesizing.

Reply:

We totally agree with you. Our wording in the original manuscript leads to misunderstanding. We here intended to say that TIR CO_2 retrieval were somewhat constrained by a priori concentrations. In the MT region in low latitudes, a priori CO_2 concentrations taken from the NIES-TM05 model probably have larger uncertainties due to the parameterization of vertical transport. Therefore, there is a possibility of more biases attributed to the a priori uncertainties in retrieved TIR CO_2 data there. Following your suggestion, we have removed the related statement and modified the sentences in the revised manuscript as follows:

"In low latitudes, there are relatively strong updrafts, and thus there are larger uncertainties among models than in other areas due to differences in the parameterization of vertical transport. Therefore, a priori CO_2 concentrations taken from the NIES-TM05 model (Saeki et al., 2013b) probably have larger uncertainties in the MT region in low latitudes. As retrieved TIR CO_2 concentrations were to some extent constrained by a priori concentrations, they possibly had more biases attributed to the a priori uncertainties in the MT region in low latitudes."

We greatly appreciate your comment.

Figure 5: It would be much clearer to the reader if they provided guidance to the different panels and lines in a legend box on the figure, rather than only in the caption. It would also help, for a reader skimming the paper, to describe what "CME_AK CO_2" means in the caption as well as the text of the paper.

Reply:

Following your advice, we have provided information on seasons in each panel and described each line in both left and right sides of the panel (a). In the caption of the revised manuscript, we have described what CME_AK CO_2 means as follows:

"The CME_AK CO_2 data are CME CO_2 data to which TIR CO_2 averaging kernel functions are applied."

Figure 6: Use pressures and not layer numbers on vertical axis. It would also be better if latitude information and season (line color) were provided as a legend on the figure. It would help if the lines in the top panels had slight vertical offsets to clarify how different the error bars are from each other.

Reply:

Following your advice, we have presented the representative pressure levels of the six retrieval grid layers shown in Table 1 instead of their layer numbers. We have provided information on latitude bands and colors for seasons as a legend and slightly shifted horizontal bars for 1-σ standard deviations in Figure 6 of the revised manuscript. We appreciate your comments.

Figure 7: It's not clear here (or in the text) at what pressures they are comparing avg CO_2 with NICAM. The contrast between the mid-gray and light-gray lines is not enough to easily distinguish between them.

Reply:

Figure 7 includes all comparison results between TIR and NICAM-TM CO_2 data in the six retrieval grid layers from 736 to 287 hPa (retrieval layers 3−8). In the revised manuscript, we have stated this clearly in the revised manuscript as follows:

"Figure 7 shows the frequency distributions of differences in monthly averaged CO_2 concentrations between TIR and NICAM-TM CO_2 data in all retrieval layers from 736 to 287 hPa in all 2.5° grids over the latitude range of 40°S to 60°N.".

Following your advice, we have presented the lower and upper pressure levels of the six retrieval layers that we focused on and used red and blue colors instead of light-gray and mid-gray colors in Figure 7 of the revised manuscript. We appreciate your comments.

Figure 8: Please use pressures instead of layer numbers. Again, the contrast between the mid-gray and light-gray lines is not enough to easily distinguish between them.

Reply:

Following your advice, we have presented the lower and upper pressure levels of each set of the six retrieval grid layers that we focused on and used red and blue colors instead of light-gray and mid-gray colors in Figure 8 of the revised manuscript.

Figure 9: Again, please state the pressures instead of "layer 7-8."

Reply:

Following your advice, we have modified Figure 9 to present the lower and upper pressure levels of the two retrieval grid layers that we focused on.

Figure 10: Please also describe the lines and the location/times the different panels represent as a legend rather than just in the caption.

Reply:

Following your advice, we have modified Figure 10: we have separated the two results of Figure 10(b) and discarded the result of Figure 10(a) of the original manuscript to simplify the figure, provided information on the locations (both over Narita airport) and dates ((a) April 1, 2010 and (b) April 30, 2010) of the two results in the caption and each of the panels, and described each of the five lines in the panel (b).