Response to reviewer 2

We thank reviewer 2 for their effort. Many of the points raised by the reviewer are points which we had listed as weaknesses or deficiencies in our manuscript. Comments by the reviewer are in italics.

1. Could the authors provide a first figure with a typical observation of ACE, maybe two spectra corresponding to tangent heights would be valuable to see the changes in the absorption features and clearly understand the choice of microwindows. These microwindows and some color code could also easily indicate which line belongs to which species;

This needs to be a four-panel figure with spectra at two tangent heights and two spectral intervals. It will be Figure 3 in the revised manuscript. The spectra in the $^{18}O^{12}C^{16}O$ ν_1 fundamental region will not be shown. The idea is to show:

1) the quality of the spectral fit in the $^{18}O^{12}C^{16}O$ ν_1 first overtone region where most of the CO_2 microwindows are located, and
2) the quality of the spectral fit in the N_2CIA region.

In the $^{18}O^{12}C^{16}O$ ν_1 first overtone region, all of the strong lines are due to CO_2, so there is no need to plot interferers in different colours. In the N_2CIA region, we followed the reviewer's suggestion to use a colour scheme to illustrate the contribution of different interferers in each microwindow for two different tangent heights. At higher tangent heights, only $^{14}N_2{}^{16}O$ is a significant interferer in the chosen microwindows.

an indication of the microwindows used for the aerosol retrieval would also enlighten the reader.

A portion of a single microwindow (~2636.46-2636.52 cm^{-1}) is used for aerosol retrieval. This is illustrated in Figure 3 of the original manuscript. Rather than writing "can provide" on p1702, L3, we now write "provides". We also refer to this figure in the text now (p1702, L1) as follows:

> …are used to determine the observed total transmittance (Fig. 4),…

This is a general remark for the complete paper: the author obviously describe knowingly their data, but it is very hard for a reader who is not involved in the mission and instrument to follow. Some introductory words to start each section would make the reading more easy.

We have re-structured the paper, moving the section on post-processing filters and the paragraphs describing the in-situ measurements to the method section (now sub-sections 2.4 and 2.5). Following the suggestion by reviewer 1, we have omitted some details in Sect. 2 to improve readability and we have created a Sect 3.1 which contains measured tangent heights, compared with ACE v3.x THs and a discussion of sources of error relating to ACE v3.x THs.
We have added the following segues:

(Sect. 2.1) Since the first guess of altitudes comes from ACE v3.x data, it is important to consider the finding of Foucher (2009) that there was a discontinuity…

(Sect. 2.2) We begin the error budget description with a review of the available literature on error analysis for ACE-FTS tropospheric CO_2 retrieval.

2. The inclusion of the CASS-FTS is not well justified nor described: one sentence at the end of the introduction and a very short (15 lines) description on the error budget. No description of the instrument or mission, of the differences relative to ACE, or of any improvements, no simulated data, analysis , etc. If this stays at this level, I would recommend to remove the reference to this instrument completely from the paper.

We now write:

The proposed CASS mission (Melo et al., 2013) is under consideration by the Canadian Space Agency. The objectives of the mission are:

1) "climate and ozone balance monitoring", and
2) improved "knowledge of atmospheric processes driving climate and its changes".

The CASS-FTS is very similar to the ACE-FTS, both consisting of Michelson interferometers, which use a pair of moving cube corner mirrors on a V-shaped scan arm. The optical path is folded in this design to give a double pass of the beams in each arm of the interferometer resulting in high spectral resolution for a compact size. The CASS-FTS has the same spectral sampling (0.02 cm^{-1}) as the ACE-FTS and also uses the solar occultation technique to measure transmittance spectra in the 750-4400 cm^{-1} range. A tangent height range of 5 to 100 km is expected. The CASS orbit will likely have a lower inclination than ACE, offering better coverage of tropical and mid-latitudes. (…)

In the original manuscript, we mentioned that:

'accompanied by solar imagers with the potential to independently provide improved pointing knowledge (Melo et al., 2013). Thus the error budget for CASS is different than the error budget for ACE…'

This would be an improvement over ACE.

Regarding simulated CASS-FTS data, we interpret this to mean Level 1 data (i.e. spectra). CASS-FTS spectra are expected to be very similar to ACE-FTS spectra, with one minor difference relating to the field-of-view being narrower in the vertical direction for CASS-FTS to improve vertical resolution. Figure 7 of the original manuscript is related to the error budget for CASS-FTS and involved some CASS-related 'analysis'. For example, we propagated the CASS-FTS temperature uncertainty profile through the CO_2 retrieval algorithm.

3. Could the authors add page number (and line number) it facilitates the commenting of the paper

The revised version we submitted has page and line numbers. Also, the version on the web has page and line numbers. I believe the reviewer was sent the first submitted manuscript.

4. section 1, 2§: the statement 'well-known atmospheric pathlengths' : is this not in contradiction to the point demonstrated in the paper ? Error on the tangent heights would impact highly on the pathlength so if THs are calculated only on the s/c information it will reflect in wrong pathlengths.

We were essentially trying to contrast solar occultation with nadir backscatter and limb scatter measurements. We have replaced “well-known atmospheric pathlengths” with “a single, geometric pathlength at each tangent height”.

5.section 1, §starting with “The Lafferty et al : : : N2 CIA ” - determination of β_0 : β_0 is determined using spectra taken from HITRAN : what are those spectra? It is very confusing and not easy to follow which data sets are used and in which way.

We now write:

The Lafferty et al. (1996) N_2 CIA cross-section spectra measured in the laboratory at five temperatures are considered to be the best available for temperatures below 300 K according to Richard et al. (2012).

Figure 1 contains data which are not cited or discussed in the text. The figure is just mentioned but not presented nor discussed.

The data in Figure 1 is cited in the text (p1697, L28). The reviewer is correct that a sentence was needed to discuss the derived temperature-dependence spectra illustrated in Figure 1. We now write:

The temperature-dependence spectra that we obtained independently using the Menoux et al. (1993) and Lafferty et al. (1996) N_2CIA spectra are more consistent, particularly near 2500 cm^{-1}, than the temperature-dependence spectra obtained previously (Lafferty et al., 1996; Foucher, 2009) as shown in Fig. 1.

It shows differences and is not a direct illustration of the method to obtain the β_0, which is what is discussed where the reference to the figure is inserted in the text.

We slightly modify the text at p1697, L28 as follows:

… to reduce noise. This results in the pink curve in Fig. 1.

By the way in the text, it is mentioned that the spectral interval goes from 2130 to 2600 cm-1, and the figure does not cover this interval.

Figure 1 has been extended to 2130 cm^{-1}. The Menoux et al. absorption coefficient spectra are not precise for wavenumbers <2150 cm^{-1}. We only attempted to determine the temperature-dependence down to 2200 cm^{-1}.

Figure 2 does not bring any information, I would suppress it.

Even though we don’t agree, Figure 2 is not critical and has been removed.

Why not replace it with the spectra that were used for this analysis. This will moreover show to the reader the shape of the N2 CIA spectra, and have more feeling about the observed data.

The derived N_2CIA temperature dependence is shown in Figure 1 but the shape of the B_0 spectrum is already in the literature (Lafferty et al., 1996) and should not be duplicated. Figure 3a-b in the revised manuscript give the reader a sense of the shape and strength of the observed and simulated N_2CIA at different tangent heights and wavenumbers.

Maybe this discussion about CIA spectra should be introduced by a sub-section title ?

We have added a sub-section entitled "2.1 N_2 CIA modelling and uncertainties".

6. Similarly, sub-sections could be set up in section 2.1: cloud detection, Tangent heights, CO2, and aerosols. It would render the text more readable.

We thank the reviewer for this suggestion, which was implemented.
As a result, we moved the paragraph about hydrostatic equilibrium to the tangent height section. This should improve readability and makes the sequence of paragraphs more logical.

7. Description of clouds detection : Here again a plot of a spectrum would help following the discussion.

We have included a plot of a series of observed spectra near 970.00 cm^{-1} (Fig. 2) from a single occultation. The 2505.5 cm^{-1} region is shown in Fig. 3a. Figure 2 shows that the spectra are flat (i.e. no strong, discrete absorption) with high transmittance (0.9) until the instrument is looking through the cloud top. Figure 3a also shows that 2505.5 cm^{-1} is in the continuum between some $^{18}O^{12}C^{16}O$ lines.

Some numbers are specified (0.0689, 0.076) without justification: either they are absolutely needed and then a more detailed explanation for the values is necessary; or they are not necessary for the understanding of the method, and maybe reference to a published paper is which more details can be found should be indicated.

The numbers specified are empirically-derived from a small subset of occultations. Increasing these thresholds leads to inclusion of occultations with optically-thin cloud, which significantly reduces CO_2 data quality. On p1699, L19-20, we state the purpose is to reduce "cloud-related error in the determination of THs". This is the justification, but instead, we now write:

These empirical settings are very stringent for the purpose of reducing cloud-related error in the determination of THs to a level where it is not a dominant source of CO_2 retrieval error (see Sect. 2.3).

These numbers are indicating changes between successive tangent height spectra: is the vertical sampling always about the same that such a rough law can be applied on all occultations.

We agree that the approach is rough. The vertical sampling varies with beta angle as mentioned in the original manuscript. For occultations in the upper troposphere using all clear-sky CO_2 profiles, the vertical sampling in the 5-17 km range is 1.4±0.4 km (1σ, N=963). Retrievals do not extend below 5 km and clouds generally do not extend above 17 km, although some clouds occur in the polar stratosphere in winter. Foucher (2009) used 0.1 for both cloud microwindows. In the next version of the algorithm, we intend to define the threshold as the change in transmittance per change in TH. This will likely require much trial-and-error and cannot be done in this work. We now write:

At 970 cm^{-1}, when the transmittance falls below 0.8 or when the change in transmittance between adjacent tangent heights exceeds an empirically-determined value of 0.0689, there is considered to be a cloud. Similarly at 2505.5 cm^{-1}, if the change in transmittance between adjacent tangent heights exceeds 0.076, a cloud is assumed to be present. Foucher (2009) used 0.1 as a threshold for both cloud microwindows.

The authors state that only 8% of the spectra pass the selection (+15% half pass): what happens to the rest, are these data completely rejected and never analysed; I hope not, then in that case how are the tangent heights (and CO2 profiles) determined.

In the original manuscript, we write "77% are deemed to be cloudy and not processed." We also state (p1699,L18): "The CO_2 retrieval is not applied to cloudy occultations."

We believe these statements are clear enough already. Rejecting observations impacted by clouds is unfortunately the reality for satellite measurements of CO_2. For nadir-viewing CO_2 missions, like GOSAT, fewer than 10% of observations pass the cloud, aerosol and other filters used in the various retrieval algorithms. In future, we hope to limit the retrieval range to extend down only to the first tangent height immediately above the cloud top. This would greatly increase the number of CO_2 profiles. Our current approach to reject cloudy occultations was also used by Foucher et al. and Rinsland et al.

8. section 2.1 – description of tangent height The authors mention that some lines have been chosen to 'increase the temperature-insensitive CO2 signal' : this might be difficult to understand for non spectroscopists and the distinction between sensitive and insensitive lines has not been introduced before (it is somewhat in the next §). Maybe add a column in the Table to clearly indicate which lines are or are not sensitive to temperature.

This paragraph is describing microwindow selection for N_2 CIA; however, there happens to be $^{18}O^{12}C^{16}O$ lines in the microwindows targeting the N_2 CIA. We confirm that these CO_2 lines are not very temperature-sensitive, similar to the ones chosen in the ν_1 fundamental region and the 2600 cm^{-1} wavenumber region. All CO_2 lines in Table 2 are also temperature-insensitive. We have moved this sentence to the CO_2 microwindow sub-section, following the sentence on temperature-insensitive lines. Temperature sensitivity can be distinguished on the basis of the lower state energy. We cannot add a column to Table 1 since there are many $^{18}O^{12}C^{16}O$ lines in some of the N_2CIA microwindows as shown in the new Figure 3a-b.

9. section 2.1 last sentence of the first page : Is the comment inside brackets '(which does not have a Q branch)' pertinent to the discussion ?

It is absolutely pertinent to the discussion. If there was a *Q branch*, this would not be a good spectral interval for a retrieval microwindow.

10. section 2.1 – discussion on CO2 retrieval: There is again a discussion of a comparison of data, this time spectroscopic parameters, which is not easy to follow because not well introduced. The authors should stress that the discussion concern spectroscopic parameters, and laboratory measurements. The misleading sentence is 'we *have compared the line intensities to measured ones'. Even knowing that Toth and the other cited authors are laboratory related, the sentence hints about comparing with atmospheric measurements.*

There is nothing misleading and no such hint about measuring line intensities using the atmosphere. This is an inference on the part of the reviewer. HITRAN is a spectroscopic database and a reference to HITRAN 2012 is provided before this paragraph in the manuscript. Nevertheless, we now write:

The line intensities in the HITRAN 2012 spectroscopic database are from a model for both the fundamental and the 20002←00001 band. Because the line intensity uncertainties in HITRAN 2012 are conservatively set to >20% for all lines mentioned above, we have compared the line intensities to ones measured in the laboratory …

11 section 2.2 – The paragraph starting with 'One source of error : : :' is incomprehensible and in contradiction with the list of all error sources just stated above.

The purpose of this paragraph is to mention sources of error which were not included in the error budget. To improve readability, we add an introductory sentence and modify the second sentence of this paragraph as follows:

Before delving into each error source, we list sources which are not included in the current error budget. The uncertainty on the N_2 VMR was not considered because it was expected to be trivial in the 5-25 km retrieval range.

For example ' The uncertainty due to spectroscopic parameters of the interferers : : : should be considered in future work' tells exactly the contrary to uncertainty nb 8 in the list. I suppose that this paragraph has been copied from a previous work, and not adapted consequently with the new approach.

Uncertainty #8 in the list (p1704, L3) is wavelength calibration. The uncertainty due to spectroscopic parameters of the interferers is certainly not the same error source as wavelength calibration. Spectroscopic parameters include the line positions, but also the line widths, their temperature dependence, and pressure-induced shifts.

12 section 2.2 is very heavy, not always well structured and should be rewritten. One typical example is that the different uncertainty types addressed are not following the order presented in the list.

Once again, the referee appears to be reviewing the manuscript that we initially submitted to AMTD, not the one that was revised and included suggestions to this comment from the editor (i.e. the online version). Note that the list (p1703, L19) only includes sources of error not included by previous reviewers. The reviewer is correct that the error source #9, namely CO_2 first guess profile is discussed before the error sources 7-8, namely pressure profile and wavelength calibration. This stems from the suggestion by the editor to list the error sources in order of importance. However, the discussion of errors relating to assumed profiles of the relevant gases were grouped and thus the paragraph on the error due to the assumed CO_2 first guess profile was moved ahead of error sources #7-8. We change the list as follows:

7) CO_2 first guess profile (above and within the retrieval range)

8) pressure profile

9) wavelength calibration

Maybe a clear separation (sub-sections) between the different topics will help.

We don't feel that a sub-section is needed for each error source since some error sources are discussed in as little as two sentences. There is a clear separation of different topics using paragraphs and the outline for the section is presented to guide the reader (p1704, L22-27).

Maybe also clearly separate the ones that affect CO2 via biases in TH and the others.

The error sources that affect CO_2 via biases to TH have been grouped (i.e. separated) in the published AMTD version.

13. section 2.2 §starting with 'Biases in TH: : :' : Fig 4 and 5 are not discussed in enough details. Fig 5 contains 2 curves derived from 2 methods which are not described nor discussed in the text.

For each retrieval altitude, we have a CO_2 concentration and a TH error (relative to the ACE v3.x corresponding TH estimate). We perform a simple linear regression between these quantities and use the slope term to quantify the sensitivity. The modified text below should help the reader understand how we obtained Figure 4 and the purple curve in Figure 5 (using Figure numbering in the original manuscript). We modify the text at p1705, L3:

…from a linear regression of TH offset versus CO_2 VMR using all altitudes. Each TH offset is determined by the difference between our retrieved TH and the corresponding ACE v3.x TH. In doing so over a large number of occultations, we obtain a CO_2 sensitivity to tangent height offsets of 0.09 ppm/m from the slope term of the linear regression.

A long discussion of Figure 5 (original manuscript numbering) begins at p1708L12 after all of the individual sources of TH-related error have been discussed. This relates mostly to the orange curve and discusses the comparison. We have added detail to the text at p1708L20:

"…there would be a major difference in the magnitude of errors from the empirical method across the hygropause…"

14 section 2.2 §starting with 'Perturbing the N2': How was the value of 0.9% chosen?

The value of 0.9% was chosen based on the upper limit of the quadrature-sum of uncertainties stated by Lafferty et al. (1996). We now write the following, with an introductory sentence to help guide the reader:

Next, we analyze the individual sources of TH-related error, which are mostly related to errors in model inputs. Perturbing the N_2 absorption coefficient spectrum by a constant value of 0.9%, which is the upper limit of the quadrature-sum of the uncertainties stated by Lafferty et al. (1996) as discussed in Sect. 2.1, we find a bias that grows with decreasing TH.

15 section 2.2 §starting with 'We also considered the impact' : the authors introduce *a 'scaling factor' to take into account the spectral dependence of the aerosols, is this factor fitted on spectra, if not could it be; is this factor the same for all spectra from one occultation series, whatever the tangent height values ?*

The reviewer makes a great suggestion here, which one of us (C. Boone) has also suggested: the spectral dependence of the aerosols could be determined from the measurements. Currently, it is not. Currently, a single value of 1 is used for all tangent heights. We would need again to test such an algorithm development on a large number of occultations, probably involving more than one iteration of processing a thousand occultations, so it cannot be included in this work (i.e. left to future work). We now write:

The slope is currently assumed to be 0 (i.e. scaling factor of 1 between 2637 cm^{-1} and the N_2 CIA microwindows) at all THs.

16 section 2.2 §starting with 'In summary': the authors list the 7 sources of error which MOSTLY affect CO2. The 7th one, sub Lorentzian line shape, was just described in the previous paragraph and its effect was deemed 'trivial'. Moreover this topic is not even further discussed in the following paragraph. Probably it should be removed from the list.

The reviewer has misunderstood here. These are the "seven theoretical sources of error which affect CO_2 via biases in retrieved TH". They are not the sources "which mostly affect CO_2". The last paragraph of the section entitled "Error budget – ACE-FTS" discusses the largest sources of error as a function of height. It is not discussed in the following paragraph because it is trivial. It will not be removed from the list of error sources which affect CO_2 via biases in retrieved TH.

17 Maybe add a Table summarizing the numbers?

Such a table would be full of footnotes. We prefer to leave the numbers in the text. The text discusses the random and systematic component of some errors. Not all error sources are systematic (i.e. purely random) and some error sources are essentially entirely systematic (i.e. no random component). As a result, a table would have a number of 'N/A' (not applicable) entries under columns of 'systematic' and 'random' error. Reviewer 1 has stated that the method is too long. Adding a table (and a second table for CASS-FTS) would not shorten the paper, because most of the 'Error budget' text consists mostly of a description of the perturbation done to assess each error. This would remain in the paper even if a table or two were added.

18 section 2.2 §starting with 'With respect to the overall ' : the enrichment factor is given with far too many decimal, restrict to the first decimal

Thank you for this correction. Done….

19 section 2.3 : either remove or complete the description From figure 7 the total uncertainty seems to be of the same order of magnitude, I would not state that CASS error are lower.

The reviewer makes another good point. The magnitude is about the same (insignificantly less). This is because TH errors are assumed to be 50 m for CASS, which is not much better than what is retrieved from ACE-FTS from the N_2 CIA (i.e. ~60 m). We now write:

The CASS-FTS total CO_2 uncertainty (~7-8 ppm) is not significantly lower than for ACE at most altitudes, but the magnitude depends strongly on the assumed tangent height uncertainty (Fig. 8).

20 section 3.2: the authors apply different versions of the software on two parts of the datasets. Would it not be more consistent to apply the same one on all data ? Since it seems that the older version cannot do it, why not apply the V3.5 on all?

It would be more consistent but the differences between v3.0 and v3.5 are trivial for the period before October 2010, however, at the time of this work, the period January 2009 to September 2010 was only partially processed using the version 3.5 software. After October 2010, v3.0 data have biases due to an error in the processing which is corrected in v3.5. The differences stem from difficulties ingesting the assimilated temperature and pressure fields from the 'regional model' used as the first guess for the v3.0 retrieval. The v3.5 retrieval instead uses the 'global model' fields. That is why we have a mix of data versions. We compared the retrieved CO_2 using v3.0 and v3.5 data. We find that the differences in TH are typically 20 metres and the differences in CO_2 are 0.5 ppm. The CO_2 VMR retrieval uncertainties are essentially unchanged. The differences in TH and CO_2 VMR are smaller than their uncertainties. We intend to use only v3.5 from now on when processing other time periods until it is replaced with version 4 which is under development.

Other comments

We have made these corrections. Thanks!