We thank the referee for the time spent in evaluating our manuscript. In the following we address the referee's comments, which are copied in the following as text in blue fonts. There are referee comments which we share, but there are also important referee comments with which we disagree. In case of disagreement we will define our position and provide a detailed argumentation.

This paper presents for the first time FTIR ozone measurements over Latin America, at 2 close but different sites (Altzomoni, a mountain observatory; and Mexico city, a polluted site). Note that ozone FTIR data are available in the NDACC database at Paramaribo (Suriname, 5_48'N, 55_12'W), but to my knowledge these results have not been published. Furthermore, the authors combine the measurements at these 2 sites to obtain a new product: an O3 boundary layer. This could provide interesting new information and would be promising for future investigation of the Mexico city pollution. Therefore, I recommend the publication of this paper in AMT. However, I have some few questions to be clarified, and some specific comments that need to be taken into account before publication. The 2 main points being the fact that the authors forgot about the smoothing error in their uncertainty budget, and that their "new product" is not convincing to me at present. So it should be at the very least clarified, or omitted in the new manuscript.

We find it important to clarify that we have not forgotten about any important component of the error analysis. We present a detailed analysis of the many parameter and measurement noise errors and clearly document the smoothing characteristic of the remote sensing data by providing the averaging kernels.

In the comments about the "smoothing error" the referee mainly refers to different works of Rodgers. Our error analysis method is fully in line with these works and we would like to draw the attention to Section 3.2 in Rodgers (2000), where the theory for reporting the errors of remote sensing systems is developed and the so called "smoothing error" is introduced: It says: "For the purpose of carrying out an error analysis, the retrieval can either be regarded as an estimate of a state smoothed by the averaging kernel rather than an estimate of the true state, or as an estimate of the true state, but with an error contribution due to smoothing. The error analysis will be different in the two cases, because in the second case there is an extra term, the smoothing error."

The referee recommends an analysis according to case two and asks us to document the smoothing characteristics of the remote sensing data by a "smoothing error" term, i.e. to do an error analysis taking the infinitesimal fine structured atmosphere as the reference. We provide a comprehensive error analysis according to case one, i.e. for the smoothed state, whereby the kind of smoothing is fully documented by the averaging kernels. This case one error analysis concept is actually similar to the referee's idea of working with partial columns, because partial columns mean that boxcar-like coarse vertical structures are used as reference for the error analysis. However, while the referee suggests introducing an additional smoothing function (partial columns, i.e. smoothing functions having boxcar-like vertical structures) we and Rodgers suggest working directly with the smoothing functions as given by the averaging kernels.

There is full agreement that when using remote sensing data for scientific studies, the smoothing characteristics have to be considered. In this context, we think it is important to reflect about the needs in practice. In practice the averaging kernels are the decisive measure for considering the different and application specific issues of the smoothing effect:

- Evaluation of models:
  Every model has limitations in its vertical resolution and the “smoothing error” of the remote sensing data would be too large for comparisons to the model atmosphere. In practice, model evaluations are made by applying the averaging kernels to the model data (e.g. Langerock et al., 2015, http://www.geosci-model-dev.net/8/911/2015/). This is a reasonable approach whenever the model’s vertical resolution is much finer than the vertical resolution of the remote sensing data.
- Inter-comparisons between two remote sensing data products:
  Most remote sensing data product have their own limitations in vertical resolution and summing up the “smoothing errors” from both remote sensing products will strongly overestimate the smoothing effect on the inter-comparison. In practice this smoothing effect is estimated by applying the averaging kernels in the context of the diagnostic metric as suggested by Rodgers and Connor (2003).
- Comparison with vertically very fine resolved profile data:
  O3 ECC sonde data have a very fine vertical resolution. In practice the averaging kernels are applied to the ECC profile, thereby calculating an ECC profile that has practically the same smoothing characteristics as the remote sensing data (e.g., Schneider et al. 2008). Then the smoothed sonde data can be adequately compared to the remote sensing data. For validation purposes it might be also interesting not smoothing the ECC data thereby leaving it completely independent from the remote sensing data, and compare to partial columns (e.g. Schneider et al. 2008; García et al., 2012). Then the averaging kernels are needed for calculating the smoothing charactistics (“smoothing errors”) specifically for the covariances as observed in the ECC ensemble to which the remote sensing data are compared to. A more general “smoothing error” calculated by assuming less specific covariances is of limited value only.
- Sensitivity analyses:
  An atmospheric signal is only detectable if it is larger than the errors and the interferences due to other atmospheric variabilities that are not related to the atmospheric signal of interest. Therefore, one has to simulate how the signal of interest on the one hand and the other atmospheric variabilities on the other hand, propagate through the averaging kernels (e.g. see example for tropospheric CH4 as in Sepúlveda et al, 2014, http://www.atmos-meas-tech.net/7/2337/2014/). These analyses can be made easily by the data user for any specific application whenever the averaging kernels are available. The “smoothing error” does not help to investigate the detectability of specific atmospheric signals.

Finally, we think it would be important to consider von Clarmann (2014, http://www.atmos-meas-tech.net/7/3023/2014/) in addition to the publications of Rodgers. Von Clarmann (2014) outlines and resumes in detail the important restrictions of the “smoothing error” concept. He concludes that a safe communication of the smoothing characteristics is best ensured by providing the averaging kernels and strongly disadvises from providing the “smoothing error” as a part of the error budget.

In summary, we would like to state that we present a detailed and comprehensive error analysis. We opt for documenting the smoothing characteristics by presenting the averaging kernels, thereby avoiding the misleading use of the “smoothing error”. We are aware that in different previous remote sensing data studies error bars are provided that include the “smoothing error” component, however, there are very good arguments for deviating from this practice.

Concerning the new boundary layer product (the combined product), we think that it will be very useful for air quality research in Mexico City by providing continuous boundary layer O3 concentrations, in addition to the surface data from the in-situ network. In the manuscript we present the theory for a posteriori combining two remote sensing observations, develop a method for analyzing the errors of the combined product (including the comprehensive consideration of correlated uncertainties) and document its performance by comparison to coincident in-situ reference data. We think that the manuscript theoretically and empirically demonstrates the good performance of the combined product. Nevertheless, since it is something new, additional intuitive explanations might be helpful for better communicating the working principle of the combined product.

The working principle of the combined product:
For our combined product we are able to optimize the boundary layer data quality, because we add to the measurement in the boundary layer (the UNAM spectra) a second measurement above the boundary layer (the Altzomoni spectra). The best way to do that would be to fit both measured spectra within a single inversion process (similar to what is done for limb sounding retrievals, e.g., Fischer et al. 2008, http://www.atmos-chem-phys.net/8/2151/2008/), which would, however, mean the setup of a new inversion algorithm software. Our method is a workaround of such multispectral inversion algorithm, because it works with the two individual retrieval results and not with the individual measurements. The method consists in an a posteriori combination the two individual retrieval results with the objective of optimally exploiting the synergies of the two individual measurements.

Specific comments/questions

1) Section 2.2 (Retrieval method) and 2.3 (Error analysis)
a) Are the baseline (channeling and offset), the ILS, the temperature, the line of sight, and the solar lines fixed model parameters or do you retrieve them ? In the case of UNAM, the Linefit results (Hase, 2012) are given as input in the forward model or Linefit is just used to check the alignment (you specify only for Altzomoni) ?

The specific retrieval setups are described in Section 2.4 (for Altzomoni) and 2.5 (for UNAM). We perform no fit of temperature, no fit of the line of site, and no fit of the ILS. For UNAM we additionally fit a baselin channeling (frequency of 0.39cm^{-1}).
As ILS for the Altzomoni retrieval we use the Linefit results. At UNAM we assume nominal ILS but regularly control the ILS by the Linefit method.
For the Altzomoni retrieval we determine the shift of solar lines with respect to the terrestrial lines by a previous analysis of a spectral window that contains well isolated contains solar and terrestrial lines and is from the spectra that we use for the retrieval.
For UNAM we use solar lines in form of fixed model parameters.

b) You wrote that all interfering species are from WACCM. Does that mean that this is also the case for H2O ? Maybe the 100% uncertainty taken for H2O interference could be reduced by using e.g. either preliminary retrieval of H2O or 6-hourly H2O profiles from NCEP. However, you obtain very small errors due to H2O in Table 2, so this seems not so important at your sites anyway.

Yes, all interfering species are from WACCM.
Yes, the H2O interference error is not too important.

At Altzomoni we estimate H2O profiles by a previous retrieval that uses a spectral window with a well isolated H2O line and use this as the H2O first guess for the O3 retrieval. At UNAM we perform a profile scaling retrieval of WACCM H2O profile during the O3 retrieval process.

c) Spectroscopic errors: p.6, l.4-5, you say that your spectroscopic errors are coming from the HITRAN line list and in the discussion p. 16, l.16-17, you say that the uncertainty on pressure broadening parameters could be as large as 20% (instead of 5%). Could you explain ? Does this mean that the 5% should not be trusted ?

Many HITRAN parameter data are available together with uncertainty estimates. The considered O3 parameter values are typically within 5% and this value is confirmed by studies that compare the FTIR data with ECC sondes (e.g. Schneider et al., 2008; Garcia et al., 2012) or Brewer total column amounts (e.g., Schneider et al., 2008). However, not all O3 parameters have the same uncertainty values and so far comparisons of FTIR O3 data to in-situ reference data have been made for the retrieval setting that we apply at Altzomoni, not for the retrieval setting that uses medium resolution spectra and fit a very broad spectral window.

There are also HITRAN parameters that are provided without any uncertainty estimate. Even for parameters where there is an uncertainty estimate, measured spectra occasionally indicate large residuals that cannot be explained by the uncertainty values, i.e. there are inconsistencies between the parameters or the provided uncertainty estimates are unrealistic. Like for any scientific data a critical reflection on the HITRAN parameter data does certainly not harm (HITRAN provides very good and carefully elaborated data, but uncertainties might be always there…).

For explaining the large difference we observe between the remote sensing data and the in-situ data in the boundary layer we discuss three possible causes (not only HITRAN pressure broadening parameters). This is described in the manuscript on page 16, line 15 -30:

- A pressure broadening parameter uncertainty of 20% (which we think is rather unrealistic, but maybe cannot be counted out completely).
- The neglection of O3 continuum absorption (might be important when fitting a broad spectral window, like for the UNAM retrieval setup)
- There is no error at all and the observed difference is due to the actual atmospheric O3 concentrations being higher 1 km above ground than at ground.

d) Table 1: I don't understand how the numbers in the column "Statistical/Systematic" are obtained, and how they are applied in the error analysis. Can you explain in the paper ?

The profile error covariances are calculated according to Eq. (6) and (12). The uncertainty of parameter p (sampled in the covariance matrix $\mathbf{S_p}$) can be a random uncertainty or a systematic uncertainty. For instance, Table 1 lists a baseline uncertainty of 0.2%. We assume 50% being random and 50% being systematic. This means that the parameters p for the baseline uncertainty have a random uncertainty of 0.1% and a systematic uncertainty of also 0.1%.

e) My main concern about this paper is that one of the dominating random error source is missing in your budget (Tables 2 3; Figs. 3 and 5): the smoothing error. You need to replace your Eq. 2 by, e.g., Eq. 1 of Schneider et al. ACP 2008 ("Quality assessment of O3 profiles. . ."). And calculate the smoothing error, which is clearly dominant on ozone profiles and partial columns.

In accordance to von Clarmann et al (2014) the provision of the averaging kernels is the safest option for communicating the smoothing characteristics of the remote sensing data product. Since we document the error for the smoothed profiles there is no additional error term ("smoothing error" term) that has to be considered (see Rodgers, 2000).

What does the referee mean by replacing Eq. 2 (the cost function) by Eq. 1 from Schneider et al., 2008 (a function that describes error contributions and the smoothing effect)? Is there a confusion with equation numbers? We document the smoothing characteristics with the averaging kernels (see Figs 3, 5, 10 and Eqs. 4 and 11) and the errors (see Figs. 3 and 5 and Table 3) by error covariances calculated according to Eqs. 6 and 12 and thereby comprehensively document the characteristic of the data products.

2) Section 3: Free tropospheric and stratospheric O3
a) Seasonal cycle at UNAM: you could add the seasonal cycle of the total columns at UNAM in Fig.6 (top panel), for comparison. Same for the different altitudes, if you define common altitudes where both instruments have information. (as you did for Fig.7).
b) Fig. 6 (lower panel): I would prefer 4 different subplots instead of all curves in the same plot. You could also add UNAM seasonal cycle (see comment a)).

Ok, we can make an additional plot showing the UNAM data: total columns and vmr at 17 and 32 km, preferably in an appendix.

c) Seasonal cycle discussion: since you have a model available (waccm), I wonder if the model reproduces well your seasonal cycles. It could be interesting to check this, and maybe add the model seasonal cycle on the plots.

Yes, this could be certainly interesting, but we chose AMT as the journal because we want to focus on the technical aspects. And the presentation and evaluation of two data sets from different instruments and the introduction into theory and practice for combining two remote sensing data retrievals merits a publication focused on the technical aspects.

d) For me, since FTIR has low vertical resolution, it would be better to describe your data in partial columns (PC) where you have about one DOFS, rather than at a given altitude point. If you take into account the smoothing error, you will find a larger error on a single altitude point (the dominant random error on the profiles is the smoothing error), than on a PC. This division in PC based on DOFS is commonly used in FTIR studies (e.g. García et al., 2012; Vigouroux et al., 2015; Duchatelet et al., 2010;. . .). I would use 4 PC in Fig. 6 (lower panel) instead of altitude points, and also in Fig.7. Your agreement between Altzomoni and UNAM PC would be probably better using common PC rather than with single altitude points (Sect. 3.2).

We can measure the vertical structures as given by the averaging kernels. Calculating partial columns will further smooth the profiles and remove profile information, however, we are interested in the profile information.

The referee's statement about partial columns having lower smoothing error is actually a good opportunity for bridging our different point of views. What we are actually discussing are the two cases/options suggested by Rodgers (2000) for the error analysis: option one means an analysis with respect to smoothed profiles and option two with respect to infinitesimal fine gridded profiles, which means that a "smoothing error" term has to be provided. In their argumentation the referee so far has clearly preferred option two whereas we do the analysis according to option one.

Now, by suggesting the use of partial columns, the referee proposes the use of boxcar-like coarse vertical structures of the atmosphere as the reference in contrast to an infinitesimal fine gridded atmosphere. With this suggestion the referee comes close to option one. The main difference to option one is that the referee introduces a new smoothing function (partial columns, i.e. smoothing functions having boxcar-like vertical structures), instead of directly working with the averaging kernels.

e) Scmp formula (Eq.8): I think you should give the reference for this formula (Rodgers and Connor, 2003). In Rodgers and Connor (2003), Scmp is not only your Eq. 8, you also have the terms SUNAM + AUNAM SALTZ AUNAM in addition (Eq. 30 of Rodgers and Connor, 2003). Since you do not use Eq. 8 in this Sect. 3.2, I am not sure that you are indeed talking about the complete error covariance matrix of the comparison between the 2 instruments ("error on the differences", are you ? If yes, then you should use the complete Eq. 30 of Rodgers and Connor (2003).

We agree with the referee that there will also be a scatter between the two data due to the parameter errors (but only if they are not correlated). However, the purpose of this calculation is not to get a theoretical estimate for the scatter that can be expected when comparing the two remote sensing data. The purpose of this calculation is to be able to filter out coinciding Altzomoni and UNAM remote sensing data, for which a comparison makes no sense because of their significantly different smoothing characteristics. We don't want an additional filter for low quality data (data with large errors), only data for which a comparison makes no sense due to different smoothing characteristics.

We agree with the referee that this should be explained better in the manuscript and we suggest modifying page 11, line 10/11 as: "In order to assure that we perform a reasonable comparison we calculate the covariances $\mathbf{S_{cmp}}$ that estimates the averaging kernel induced uncertainty for the comparison between the UNAM remote sensing data and the Altzomoni remote sensing data after smoothing according to Eq. (7):".

f) In principle, you should use Scmp (complete formula) to conclude if the 2 instruments are in agreement: standard deviation of the differences (1.2%; 5.2%, and 2.1% for your 3 comparisons) < "random error on the differences". The bias of 8.7% at 17 km is probably not explained by the systematic uncertainty budget on the differences (same spectroscopy, same temperature profile), but this bias might improve if you use PC instead of a point at 17 km.

We can mention in the discussion of the differences that the scatter between the two remote sensing data is explained by the sum of the errors and the different smoothing characteristics and refer to Rodgers and Conner (2003). For estimating the sum of errors it is important to consider correlated errors, for instance, those due to temperature profile uncertainties above 12.5 km (please see also discussion at the end of Section 4.3.1).

g) p.11, l. 15: Scov= 100% O3 variability: this is a large assumed variability, especially in the stratosphere. I think you are overestimating then your error on comparison. Why not taking Scov from the model WACCM ? (or any other climatology, e.g. from satellite?)

Please be aware that with Eq. 8 we define a simple metric for identifying situations where the two coinciding remote sensing data cannot be compared due to the significantly different smoothing characteristics. To do so we define $100^2\%^2$ for the diagonal of $\mathbf{S_{cov}}$ and then filter for $\mathbf{S_{cmp}}$ with diagonal values smaller than $10^2\%^2$.

The referee suggests using apriori covariances from WACCM data for constructing $\mathbf{S_{cov}}$. This can of course be done, but then we would also need to adapt our threshold (diagonal element of the resulting $\mathbf{S_{cmp}}$, for which we still can do a reliable comparison). We can then define as threshold that the diagonal elements of $\mathbf{S_{cov}}$ and $\mathbf{S_{cmp}}$ must have a ratio of at least $10^2\%^2$.

h) p.11, l. 16-17: I don't understand how you can "require that the square root of. . . is smaller than 10%". If you fix Scov (100% O3 variability + 5 km correlation), and you have AUNAM and AALTZ fixed from your retrievals, how can you control Scmp ? Sorry, I am missing what you mean here.

We use this metric for filtering out coinciding Altzomoni and UNAM remote sensing data, for which a comparison makes no sense because of their significantly different smoothing characteristics. We calculate $\mathbf{S_{cmp}}$ for all coincidences individually and define as data pairs that can be compared only those for which the diagonal of $\mathbf{S_{cmp}}$ is below $10^2\%^2$, i.e. for which $\mathbf{A_{UNAM}}$ and $\mathbf{A_{UNAM}}*\mathbf{A_{ALTZ}}$ have very similar smoothing characteristics.

i) p.11, l. 26-27: you said that the bias of 2% is due to the different altitudes between the 2 instruments. But if I get right how you construct x*ALTZ, you are not using zero values from 2.3 and 4km, but the values of xa. Then you should not have some bias (at least not so large). But maybe you calculate from x*ALTZ a total column that starts only at 4 km ? If yes, please specify in the text.

Yes, we calculate the column for Altzomoni using x^*_{ALTZ} above 4km and will specify this in the text.

For Eqs. 7 and 8 (as well as for Eqs. 9-12) it is necessary to expand the Altzomoni state vectors and averaging kernels to altitudes below 4 km. We use the same x_a at Altzomoni and UNAM and since the averaging kernel for Altzomoni has cero entries for altitudes below 4 km (no sensitivity) it is $x_{ALTZ} = x_{a,ALTZ} = x_{a,UNAM}$.

3) Section 4: Boundary layer
a) Sect. 4.2: Again here, you could reduce the error on the FTIR products by using PC instead of a single value at 2.3 or 4 km (as in, e.g., Sepúlveda et al., 2014) ? These values are correlated to the whole tropospheric column anyway. I would be curious to see Fig. 9 with PC instead of single altitude values.

For partial columns the smoothing error is mainly reduced because the reference for the "smoothing error" calculation changes from an infinitesimal fine gridded atmosphere to a boxcar-like coarse gridded atmosphere. However, calculating partial columns will actually mean a reduction of the vertical resolution of the remote sensing data.

The idea of accepting that remote sensing data are only representative for coarse vertical structures like partial columns comes close to what we do throughout the paper by accepting the coarse structures as given by the averaging kernels. However, the calculation of partial columns means the application of an additional boxcar-like smoothing function on top of the natural smoothing function of the remote sensing data given by the averaging kernels.

In Sepúlveda et al., (2014) we calculate the mean mixing ratios for a certain altitude range above ground, i.e. we reduce the vertically resolution. There our motivation was to remove the oscillation in the lowermost profile levels, which we think occur due to errors in modelling the CH4 line shapes, we focused on a broad layer (ground to about 8-12 km above ground), and on

removing interferences from the atmosphere above 12 km. To the contrary, here we focus on a rather fine layer 2.3 – 4km and on interferences from the atmosphere above 4 km. In order to maintain this focus on a fine layer we have to avoid any unnecessary smoothing of the profile information.

b) Sect. 4.3: p.13, l.24-30: Note that if you use Eq.30 of Rodgers and Connor (2003), you directly have the combined error on the differences, i.e., your 4-5% value.

Yes, correct. Here we are discussing the sum of the scatter due to different smoothing characteristics and the scatter due to uncorrelated errors of the two products.

c) I don't understand why having a boundary layer product coming from the total columns differences with an error of about 4-5%, would be worse than the combined product that you propose for which you reach 21% ! Therefore, I am really surprised that the correlation with in situ data using your combined product is better than using the total column differences.

The 4-5% is the error in the difference of the total columns, i.e. it is an error relative to the total column amounts and we have to consider that the boundary layer is only a very small portion of the total column.

The boundary layer partial column is about 2.5% of the total column. Furthermore, the remote sensing system is only sensitive to about to about 35% of the boundary layer variations. In summary, by calculating the difference of the total columns we can expect a boundary layer signal that is about 1% of the total column. At the same time we estimate a 5% error in the difference of the total columns. We cannot detect a 1% signal in a value that has an error of 5%. We think that this is well explained on page 13, line 29-34, but we can try to further improve the explanation.

The 21% error is the error relative to the boundary layer concentrations.

d) I am not convinced by the construction of the combined product. So I strongly suggest that you document better what you are doing. Is there any reference that could be added to your calculation ?

Equations. 10, 11, and 12 are written in analogy to Rodgers (2000) and many other works. We expand this "Rodger formalism" by means of an operator **C**. The procedure is similar to the aposteriori processing as presented in Sepúlveda et al. (2014) or Schneider et al. (2012), with the difference that the aposteriori processing in those works is not supported by a second measurement.

An additional intuitive explanation might be helpful for better communicating the working principle of the combined product. In the following we copy what we have written as reply to the general comment:
"For our combined product we are able to optimize the boundary layer data quality, because we add to the measurement in the boundary layer (the UNAM spectra) a second measurement above the boundary layer (the Altzomoni spectra). The best way to do that would be to fit both measured spectra within a single inversion process (similar to what is done for limb sounding retrievals, e.g., Fischer et al. 2008, http://www.atmos-chem-phys.net/8/2151/2008/), which would, however, mean the setup of a new inversion algorithm software. Our method is a workaround of such multispectral inversion algorithm, because it works with the two individual retrieval results and not with the individual measurements. The method a posteriori combines

the two individual retrieval results with the objective of optimally exploiting the synergies of the two individual measurements."

At present these iterations between Eq. 9 and 11 (why only 2 ?) seem arbitrary.

Maybe it is better to replace "iterations" by "steps" and to make clear that the procedure has nothing to do with an iterative retrieval. The two steps are simply needed in order to normalise the kernel for the combined product.

I have the feeling that the DOFS are increased "arbitrarily" in the boundary layer (p.15, l. 29 – p.26, l.1; and p.26, l.9-10). There is one instrument at 2.3 km with "little sensitivity up to 4 km (from 0.05 to 0.12 Fig.10), and one instrument that is not measuring below 4 km,

Actually there is a certain kind of arbitrariness in any product retrieved by an ill-posed remote sensing inversion procedure, because the product depends not only on the measurement, it also depends on the a priori information and the strength of constraint with respect to the a priori information. Reducing the constraint will make the product less dependent on the a priori data, but at the same time more sensitive to the different error sources. Increasing the constraint will make to product more dependent on the a priori data, but less sensitive to the different error sources. Essential is that the constraint dependent errors and the constraint dependent averaging kernels are provided together with the data product.

For the combined product we reduce the constraint in the boundary layer by normalizing the operator **C** (Eq. 9). Since we add the information from the Altzomoni data we can thereby increase the boundary layer sensitivity without being affected by increasing interferences from altitudes above the boundary layer. So we can indeed say that we tune the kernels for being optimally sensitive in the boundary layer (the result can be observed in Fig. 10). It is important to state that there is nothing artificial and that the characteristics of this combined product are fully traceable back to the Jacobians matrices, gain matrices and averaging kernel matrices of the individual retrieval products (see Eqs. 10 - 12). Reducing the boundary layer constraint has the negative side effect that the errors increase. So it is important to make the error estimation according to Eq. (12) and analyse the error in the combined product without forgetting about the possibility of correlated uncertainty sources (see Table 3).

so to me what you are doing by using the information above 4 km from the 2 instruments, removing it (Acom come to about zero above 4 km; Fig 10), and transferring the information to the 2.3-4km is equivalent to use the differences between the 2 instruments.

Yes, in principle we calculate the differences of the two instruments, but also account for the different sensitivities and assure that the kernel of the combined product is normalized for the boundary layer (we require that the sum of its boundary layer diagonal elements are 1.0). Maybe it is helpful to execute the matrix calculation of Eq. (10) by inserting **C**. We get:

$x_{comb} = (1/norm) * [x_{UNAM} - \mathbf{A_{UNAM}} * (x_{ALTZ} - x_{a,ALTZ})]$.

Here *norm* is the normalization factor from Eq. (9).

And if you want to do so, the more precise way should be to use the total columns since the errors of the instruments are the lowest for total columns, because the smoothing error is much reduced for total columns. It might be that I am missing something, but then more information and references are needed in the paper to explain why this combined product is not only

"artificial", and that real information is appearing in this boundary layer, with an added value compared to differences of total columns.

The problem with the total columns is that only a very small part of the total column is the boundary layer partial column. Although the total column has lower errors than the boundary layer profile data, the error in the difference of the two total column data is by far too large for detecting the small signals of the boundary layer partial columns. See also our reply to a similar comment above and the explanation on page 13, line 29-34.

Technical / minor comments:

- Abstract, l. 6: "three" and "four" should be inverted.
Yes, thank you!

- p.2, l. 21: " of Paris, Viatte et., 2011)" : add the comma.
Yes, thank you!

- p.2, l. 22: Missing "." And "However" instead of "however".
Yes, maybe it is better to separate the long sentence into two shorter sentences.

- p.2, l. 26-27: A new FTIR station is measuring now in Brazil (Porto Velho), since July 2016. Although it is not yet an "NDACC" station, I think it's worth to mention it.
Yes, we can add it. Does the referee know when O3 data will be retrieved and made available? What spectra have been measured so far? Should we refer to it as "Vigouroux, private communication" or "De Mazière, private communication"?

- p.5, l. 22: "listed in in Table 1": remove one "in".
Yes, thank you!

- p.6, l. 8: "Observations at Altzomoni" (not Atzomoni)
Yes, thank you!

- p.13, l.30: "expected", not expect.
Yes, thank you!