Response to reviewers for Sheese et al., "Validation of ACE-FTS version 3.5 NO_y species profiles using correlative satellite measurements" (doi:10.5194/amt-2016-69)

The authors wish to thank the anonymous reviewers for their thoughtful comments and valuable insight.

Reviewer #2

General comments:

The authors present a comprehensive validation study where five NOy species (NO, NO2, HNO3, N2O5, and ClONO2) of the most recent data version (v3.5) of the ACEFTS instrument on the Canadian SCISAT satellite are compared to data sets of up to 11 other satellite sensors. Differences between the previously validated v2.2 and the new v3.5 data versions are also discussed. The manuscript is clearly written and well structured and will be of interest to scientists working with the corresponding satellite data sets. I therefore recommend publishing this manuscript in AMT after addressing the comments below.

Major comment:

I would like to suggest that the authors should include the standard error of the mean (SEM) plotted as error bars on the mean difference profiles shown in Figures 3 to 18 (perhaps not in each altitude for better readability). This has the advantage that the reader may easily recognize whether a bias observed between instruments is significant or not.
All relative difference profiles now include SEM values for error bars every 5 km.

Specific comments:

Page 10, line 5: A validation study comparing MIPAS balloon data with the new MIPAS ESA v6 species including N2O5 has been published by Wetzel et al. (2013) (reference: see below in technical corrections). Observed differences are within ~20% in the middle stratosphere. You should include the citation here and change the corresponding sentence accordingly.
The Wetzel et al. reference has been added and is described as, "MIPAS ESA N_2O_5 and $ClONO_2$ data were compared with the balloon based MIPAS-B instrument by Wetzel et al. (2013). It was found that N_2O_5 concentrations typically agree within ±40% and $ClONO_2$ concentrations typically within ±30%."

Page 18, line 17: I cannot understand the meaning of the clause "...correlation with, and mean and standard deviation..." This sounds a bit confusing. Please rephrase for better understanding.
This has been rephrased to the weighted-average ACE-FTS correlation coefficients, mean differences, and standard deviations of the relative differences.

Page 20, line 21: From Fig. 1 I cannot see a difference of -22%. Shouldn't it read -12%?
This typo has been corrected.

Page 20, line 24: "Below 10 km, where HNO3..." Please include "v3.5" before "HNO3".
Text now reads "v3.5 HNO_3".

Page 25 line 7: Is there an idea why the evening comparisons are better than the morning ones? May this be caused by lower amounts of NO2 during local morning compared to local evening? Is this mainly a problem of ACE-FTS data? Please include one or two sentences on this issue.
The text now states that these results are "likely due to differences in the diurnal variation along the line of sight between sunrise and sunset observations. For sunrise (local morning) observations, ACE-FTS samples a region of the atmosphere that has yet to be sunlit long enough for NO_2 to be in equilibrium. For sunset (local evening), however, the entire sampled should be relatively stable." This is reiterated in the conclusions section.

Page 26, line 15: Please also comment on a possible reason of the local morning and evening differences in the N2O5 data.

Here we've added, "The poor agreement in the evening is mostly due to the low signal to noise in the ACE-FTS retrievals due to the lower N_2O_5 concentrations at sunset than at sunrise." At the end of the N_2O_5 section we have also added, "This indicates that the poor agreement seen in the evening data is most likely due to the high level of noise in the evening ACE-FTS N_2O_5 data." This is further addressed in the Discusion section.

Page 27, line 12: The expression "reasonably good agreement" is not very scientific. Are the observed differences within the combined systematic error limits of both instruments?
This line wasn't meant to give quantifiable results and has been reworded to "...profiles yield much better agreement." The specific values are given in the following lines.

Page 28, line 7: You state that the comparison results are typically better for the evening results but morning and evening differences have the same bias of -10%? I think this is a contradiction. Pleas change the text accordingly and write only "-10%" (not "-10±10%").
This sentence now reads, "In the 13-23 km region, where the comparison results are more consistent for the evening results, both the morning and evening results tend to exhibit a -10% bias."

Page 29, line 12: Please add a reference for the characterization of the ACE-FTS instrumental line shape here.
Boone et al. (2013) is now referenced here.

Page 29, line 22: You write that near 35 km ACE-FTS has a positive bias of about 20%. In the next sentence you say that this is an improvement to data version v2.2 which also shows an agreement of 20% with other satellite data. Hence, the improvement seems to refer to other altitude regions. Please rephrase this sentence to make this issue clearer.
These lines have been restructured to "HNO_3 comparisons near 35 km show that ACE-FTS has a positive bias that on average is ~20%. Within the 8-30 km range, ACE-FTS and correlative data sets on average are within ±7%, and around the HNO_3 peak (~20-26 km) on average ACE-FTS is within ±1% of the other measurements. These results represent an improvement from ACE-FTS v2.2 comparisons by Wolff et al. (2008), who found that ACE-FTS was typically within ±20% of correlative satellite data sets."

Page 30, line 4: From Table 1 I see that the vertical resolution of ACE-FTS and MIPAS is similar and with about 3-4 km high enough to be only slightly dependent on the a priori profile used. Hence, I don't believe that the systematic differences between ACE-FTS and MIPAS are largely due to differences in the a priori. Are the used a priori profiles really that different? You should check this to underpin your statement. Otherwise, please omit the sentence.
This line has been omitted.

Technical corrections:

Page 19, line 8: The "y" in "Bry" should be subscript.
This has been corrected.

Page 38, line 7: Please include the reference: Wetzel, G., Oelhaf, H., Friedl-Vallon, F., Kleinert, A., Maucher, G., Nordmeyer, H., and Orphal, J.: Long-term intercomparison of MIPAS additional species ClONO2, N2O5, CFC-11, and CFC-12 with MIPAS-B measurements, Annals of Geophysics, 56, Fast Track-1, doi:10.4401/ag-6329, 2013.
This reference is now included.

Page 40, Table 2: The uppermost reference should read Verronen et al. (2009) (not 2008).
This has been corrected.

Page 43, Fig. caption 1: Please write "...percent differences (v3.5 – v2.2 divided..." (not vice versa).
This has been corrected.

Page 45, Fig. caption 4: Please write "...ACE-FTS – INST..." (3rd line) because not only HALOE but also MIPAS is shown here.
This has been corrected.

Page 47, Fig. caption 6: Please omit the expression "legend shown in Figure 2" in the first line because it also occurs at the end of the Figure caption.
This has been corrected.

Page 55, Fig. caption 17: Please omit "top panel of" and "all" before the word ClONO2.
Caption now reads "Comparisons between ACE-FTS and MIPAS $ClONO_2$ measurements with coincidence criteria…"

Also note, that in the MIPAS Section we have added details on the different use of the a priori by the two algorithms.

Reviewer #3

This paper describes the validation of NOy species measured by ACE-FTS versus previous versions as well as several satellite datasets. The introduction needs work: the authors sometime show the reactions involved and some time they do not. They have to be more consistent. I strongly encourage them to show all the reactions. The methodology and datasets are presented in a verbose manner that could be more succinct. The figure shown for each species are inconsistent. For some the authors shown profiles, then differences, then seasonal biases, then hemispheric biases and for others a subset of these.

The description provided for the methodology and datasets was intentional. The authors would like to be as specific as possible such that there's no ambiguity in the methods that were used. With regards to the figures, plots of mean vmr profiles then comparison results are shown for all species. For some species, seasonal biases and/or hemispheric biases are shown; when they are not shown, the text explains why (e.g. no biases found, not enough coincident data).

Major comments:

A figure showing the diurnal variability for the NOy species used in the study needs to be shown. This would help the reader know when to expect a big impact due to this correction. See for example, Khosravi et al. 2013 Figure 2. (the figure can be normalized to be able to show all the species in the same figure)
A figure has been added that shows the diurnal variation (percent deviation from altitude mean) at each altitude for each species.

The authors needs to include the errors (a combination of the precision and accuracy) for the datasets. This would allow the reader to see if the differences between the datasets are significant or not.
Unfortunately, ACE-FTS does not have a known error budget for any of its retrieved species (determining this is an ongoing study). Instead, the relative difference profiles now include standard error of the mean (SEM) values as error bars, as requested by reviewer #2.

The author needs to add a figure showing the fake kernels constructed with the normalized Gaussian distributions against the kernels of the satellites that do provide kernels (at least MLS and MIPAS). Also the author needs to add a figure showing that the results of applying both kernels (the faked and the actual kernels) are similar. The impact of averaging kernels is crucial in this type of comparisons. Also, why not used the actual kernels when available and the fake kernels when not?
We now show, in an appendix, that average relative differences for the species we have validated are largely unaffected by any smoothing (typically < 1%). Therefore the method of smoothing is not as critical as it would be if we were comparing with, e.g., sondes or nadir viewing instruments. This method is done because it is much easier to store vertical resolution profiles than it is to store both averaging kernel matrices and a priori profiles (to use averaging kernel matrices properly, you need to multiply by the averaging kernel matrix and subtract the a priori). Also, this method is then at least consistent between instruments, as mentioned in the text.

Comments:

P1L31: NOy also includes [N] (See Brausser and Salomon)
N has been added here and in the introduction.

P1L36: The authors do not show the instrument average mean relative difference in figure 4, where do these estimates came from?
This now reads, "ACE-FTS typically agrees with correlative data to within -10%."

P2L1: for NO2, add "and up to 40% elsewhere." For HNO3 add "and up to 20% at 30-40km" Add a summary for the N2O5 evening.
We agree that these would be informative, unfortunately it is not possible to add these and keep within the abstract length limits.

P2L2: For ClONO2 add "and varies from -20 to 15% from 13 to 20km"
See above.

P2L6: NOy also includes [N] (See Brausser and Salomon)
"[N]" has been added.

P2L11: Not really a major role, that's ClO and BrO.
"major" has been replaced with "significant".

P2L14: This is not a detailed description of NOy chemistry. It is missing reactions for HNO4, NO3 and BRONO2. The authors should change this line to "A description of the chemistry of the molecules validated in this study is shown below [Brausser and Solomon, 2005]" or something similar. Or, show the complete NOy chemistry.
Text now reads "a summary for the species validated in this study is given below."

P2L20: Add N2O + hv (lambda <= 200nm) - > N2 + O1(D) R. 5.129 Brausser and Solomon, (2005) (from now on BS05 reaction numbers refer to the Third revised and enlarged edition)
This equation has been added.

P3 R4 is missing the (3P) that is O(^3 P)
The triplet P has been added.

P3L3 delete this line, not relevant to this paper.
This line has been deleted.

P3L5 HNO3 sources add reactions 5.151 (and keep 5.152 R5 in the paper) HNO3 sinks add reactions 5.153 and 5.101
These have been added to the text.

P3L9-11: Delete this paragraph not relevant to this study
Paragraph has been deleted.

P3L14: Add reactions 5.149a, 5.149b and 5.150 BS05
These have been added to the text.

P3L17: Add reactions 5.159a and 5.159b BS05
These have been added to the text.

P4 L15: Why HITRAN 2004? Why not HITRAN 2012?
HITRAN 2004 was the most recently validated version when ACE-FTS v3 started. The plan is to update to HITRAN 2012 for version 4. This is not discussed in the text because the focus of the paper is on the validation of the data products, not on the retrieval algorithm (which has been well documented). For more information on the specifics of the retrieval algorithm, readers are referred to Boone et al. [2005; 2013].

P4 L17-20: Could you explain briefly how the new microwindows were selected? Did you use propagation of random noise and select microwindows that maximize information content or degrees of freedom? Did you minimize total error (precision and accuracy)? Etc...
They were chosen to minimize the effects of interfering species. However, this is not discussed in the text because the focus of the paper is on the validation of the data products. For more on the specifics of the retrieval algorithm, readers are referred to Boone et al. [2005; 2013].

P4 L20: Why the exclusion of ClO?
It was not a reliable data product. This is not discussed further in the text because as noted above the focus of the paper is on the validation of the NOy data products. For more information on the specifics of the retrieval algorithm, readers are referred to Boone et al. [2005; 2013].

P5L6: What happen between 64 and 94km?
We added the sentence, "In that study, there were not enough coincident profiles that did not contain large errors in the 64-93 km region for statistically significant comparisons."

P5 L17: Why only one spectral window? Wolf et al. 2008 used two.
The number and range of the spectral windows are chosen to reduce the effects of interfering species. This is not discussed in the text because the focus of the paper is on the validation of the data products. For further details on the retrieval algorithm and microwindows, readers are referred to Boone et al. [2005; 2013].

P5 L24: The authors state "ACE-FTS N2O5 typically exhibited a low bias on the order of 30%..." This sentence only applies for daytime, during nighttime the biases were worse, in the order of 50% for uncorrected and around 35% for corrected.
The text now reads, "...on the order of 30-50%, whereas with diurnal scaling ACE-FTS typically exhibited a ~10-35% low bias."

P6 L1-4: Wolff et al 2008 states: That MIPAS and ACE-FTS agree within 1% (not 5% as the author states) between 16 and 27km and that ACE-FTS has a positive relative bias of 14% (not 20% as the author states) between 27 and 34 km.
The text states that Wolff et al. 2008 found that ACE-FTS differed by less than 5% below 25 km, which can be seen in Figure 8 of that paper. We have changed the text to read, "less than 1% between 16 and 24 km."

Wolff et al. 2008 found that ACE-FTS has a positive bias of "about" 14%, we state that they found bias of "up to 20% near 33 km," which can be seen in Figure 8 of that paper.

Section 2.1.1: The authors need to specify if the interfering species were retrieved or if their concentrations were set to climatological values. Did the retrieval used the O3 and H2O previously retrieved to constrain their impact, etc?
Interfering species are not set to a fixed profile, they are retrieved simultaneously with the target species. However, the "retrieved" interfering species profiles are not saved. This is not discussed in the text because the focus of the paper is on the validation of the data products. For more specific information on the retrieval algorithm, readers are referred to Boone et al. [2005; 2013].

P7 L1: the diurnal variations along the line-sight in the NO2 retrieval algorithm also affected ACE-FTS NO2 retrievals but the author forgot to mention it.
This was not "forgotten" as this section is reiterating the specific findings of Kerzenmacher et al. 2008. Furthermore, this issue is discussed for ACE-FTS in Sec. 4.2.1. We have added the sentence. "It should be noted that this issue would similarly affect ACE-FTS NO_2 retrievals."

P7L17: Add O2.
"air density" has been replaced with "O_2".

P7L18: the spectral range is 248-956nm See table 1 of Kyrola (2004).
Typo has been corrected to "954 nm," as per Kyrölä (2010).

P8 L12: five spectral bands (not channels)
Text has been changed to "bands".

P8 L16: an anomaly occurred in the interferometer mirror slide mechanism (not the instrument drive unit).
Text has been changed to "interferometer mirror slide mechanism"

P8 L25: Change Hanke, 1997 for Levenberg (1944) and Marquardt (1963)
The Hanke reference details the specific Levenberg-Marquardt regularization technique that is used and is therefore the appropriate reference.

P9 L2: the authors need a citation for the a-posteriori regularization.
References to Ceccherini 2005 and Ceccherini et al. 2007 have been added here. Also note, that in the MIPAS Section we have added details on the different use of the a priori by the two algorithms.

P9 L6: without pushing the results towards an a-priori profile (Neither the ESA one but the authors did not mention it)
In the MIPAS ESA section, we have added, "and does not drive the result towards an a priori profile."

P9 L10: As well, the forward model... → Further, the forward model can allow deviations from LTE which have an impact in the mesospheric retrievals.
Text now reads, "Further, the forward model can allow for deviations from local thermodynamic equilibrium (LTE), which mainly affects mesospheric retrievals…"

P13 L13 Waters citation should follow MLS not Aura.
This has been corrected.

P13 L18: the authors did not mention the ascending node time for the other satellites (MIPAS, GOMOS, SCIAMACHY, etc)
This ascending/descending node information has been provided for all sun-synchronous satellites (often in the satellite description section rather than the instrument section) so no change is needed.

P13 L20 temperature, GPH, concentrations of over 15 ... and cloud ice.
"and cloud ice" has been added.

P13 L23: radiometer not channel
This has been corrected.

P15 L25: radiometer not channel
This has been corrected.

Section 2.4: this is not the latest version for the MLS data. If the authors are not going to use it at least it should state that it exists. MLS version 3 was adversely impacted by clouds leading to noisy HNO3 in the UTLS. The adverse cloud impacts were substantially mitigated in version 4 (Livesey 2015).
After we mention that we did not use any profiles that were flagged as containing clouds, we have added, "However, the adverse effects on MLS v3.3 HNO_3 due to clouds were substantially mitigated in the most recent version, v4.2 (Livesey et al., 2015)."

P16L5: MLS also uses the JPL spectral catalog.
This is now mentioned in the MLS section.

P16L10: Why only from band C? The authors should include both or an average or indicate why HNO3 band A was not used. Also, the authors did not mention that SMILES covers the entire diurnal cycle in a period of about 2 months. A great plus of this dataset!

Text now reads, "Only level 2 SMILES data derived from band C measurements were used in the analysis, as the HNO_3 *retrievals from band A have been found to typically converge to a priori values (Dr. Makato Suzuki, personal communication)."*

We have also included that, "the local time coverage was such that it took two months to sample an entire diurnal cycle."

P17L19: Use sigma symbol (in latex \sigma for example)

This has been corrected.

P18L5: Please add the equation for the standard deviation of the relative differences instead of explaining with words.

This is now given as an equation in addition to the description.

P19L9-10 Why Sander et al (2003). That kinetics file have been superseded. The one the authors used is the evaluation number 14 while the newest one is 18. The author at least needs to check that no new rate constants have been recommend for the most important NOy and O3 reactions.

We are using a pre-compiled version of the Pratmo model. Coefficients for four NO_y *reactions, and zero* O_3 *reactions, are out of date in this version.* $NO_2 + OH \xrightarrow{M} HNO_3$ *is updated in evaluation 18, however* HNO_3 *is not expected to have a large diurnal variation and is therefore not crucial for this study.* $NO_2 + HO_2 \xrightarrow{M} HNO_4$ *and* $BrO + NO_2 \xrightarrow{M} BrONO_2$ *are updated in evaluation 18, however this study analyzes neither* HNO_4 *nor* $BrONO_2$*. The only update that is likely to have an effect on our results is the reaction* $NO_2 + NO_3 \xrightarrow{M} N_2O_5$*. We now state in the text that, "Updated reaction coefficients have more recently been suggested for Reactions R6 and R10 by Burkholder et al. (2015). Since* HNO_3 *does not have a significant diurnal variation, not including the updated coefficients for Reaction R6 is not likely to affect the results of this study; however, not including updates to the coefficients for Reaction R10 may add additional uncertainty to the comparisons of* N_2O_5 *that use diurnal scaling." Burkholder et al. 2015 has been added to the References section.*

P20L13: The difference are much worse because both versions retrieved closed to zero values.

"…and NO concentrations are over an order of magnitude smaller than above 22 km." has been added

P21L17: What is meant by statistically significant? 95% confidence more than 3, 10, 100 please explain.

This now reads, "…allow for a significant number of coincident profiles (minimum number of 10)…"

P21L23-24 Why do the number of coincidences drop at some altitudes? Please explain.

The end of the intro for section 4.2 now reads, "It should be noted that the profiles of number of coincidences are typically not constant in altitude due to screening of the data sets using metrics (e.g. retrieval response, quality flags) that are not always constant in altitude."

P22L23: Why did the comparison with MIPAS did not improve when using the diurnal scaling? ACE-FTS measures at sunset/sunrise while MIPAS measures at around 10, this should make a difference, unless the diurnal variation is flat.
As stated in the text, the correlation coefficients increased and standard deviations decreased with diurnal scaling, which is an improvement over comparisons without diurnal scaling. It is likely there wasn't a greater improvement because, above ~25 km, due to the relatively tight coincidence criteria used, the scaling factors were typically within 1±0.2.

P23L10-23: Please explain the top row of the figure without mention any seasonal bias. Then, mention that due to ACE coverage and MIPAS NO sensitivity there is only data during... Then explain the bottom row. Also, is there any reason to expect the comparison to work better over the summer months. Is the winter data wrong? If so, which data MIPAS or ACE.
The order of the explanation has been changed as suggested. The text states that "there is less NO variation in the polar summer regions than in the winter."

P26L15 Where is the figure showing the profiles? As Figure 3,6 and 10.
These are shown in figure 13, but we felt that it was important to first address the significant difference between morning and evening results. The intro to this section now reads "Before showing the N_2O_5 validation results, it should be noted that a significant difference…" in order to make this clearer.

Section 4.2.4: The authors should try to explain the much worse agreement found during the evening. Also, they need to explain a bit better the intention and the results of Figure 15.
Here we've added, "The poor agreement in the evening is mostly due to the low signal to noise in the ACE-FTS retrievals due to the lower N_2O_5 concentrations at sunset than at sunrise." At the end of the N_2O_5 section we have also added, "This indicates that the poor agreement seen in the evening data is most likely due to the high level of noise in the evening ACE-FTS N_2O_5 data." This is further addressed in the Discusion section.

The text now reads, "Although there was very poor agreement between local evening ACE-FTS and local evening MIPAS N_2O_5 profiles, comparisons between diurnally-scaled morning ACE-FTS and evening MIPAS N_2O_5 profiles yield much better agreement. This indicates that the poor agreement seen in the evening data is most likely due to the high level of noise in the evening ACE-FTS N_2O_5 data and is not likely an issue with the MIPAS data. Figure 15 shows…"

Section 4.2.5: There is no profile figure for morning and evening. Please be consistent with the figures shown among the species.
We are a bit confused by this comment. Figure 18 shows the $ClONO_2$ comparison profiles for morning and evening data. If this refers to morning and evening mean VMR profiles, showing those would be inconsistent, as these were not shown (for morning and evening data separately) for any of the other species.

P28L25: in line 15 (same page) the authors mention that the results for all species except for HNO3 were improved by diurnal scaling. Now the authors state that for NO the diurnal factors did not help improve the comparison results. Please clarify.
Text now states that "Results for NO_2, N_2O_5, and $ClONO_2$ were improved…"

Table 4 should also include NO no summer months and N2O5 evening
These have not been included, as we cannot say that these have been validated (especially evening N_2O_5). The caption to Table for now reads, "Summary of validated ACE-FTS NO_y…" and the text now reads, "…in the regions where the ACE-FTS data has been validated and where there is typically a strong…"

P29L2: In the comparison shown for NO, v3.5 still seem to have bias as big as 10% in this region. The -6% bias was versus MIPAS. So, it seems there was no improvement.
The HALOE average in this region is -6%, but "-6%" has been changed to "0 to -10%".

P29L4: The authors said that the bias was -6% and now they state approximately -5% (just a few lines apart). Please choose one. Also, add a statement: Above 40km, the bias is up to 40%.
The MIPAS average is -5% (the previous paragraph where the -6% was stated was discussing HALOE). The text now also states that "with values more negative than 50% below ~23 km and above ~50 km, and within…"

P29L10: Following figure 8 the statement that start "with diurnal scaling..." seems to be wrong. Something like: "With diurnal scaling, this negative bias varies from ~-10% to 3% for evening comparisons... and from -12% to 40% for morning comparisons between 20 and 40 km." will be better.
The bias being discussed here is for near 32 km. To be clearer, the text now states "…this negative bias near the peak is ~ -10% for…"

P29L13: please clarify where do this statement ("In the 14-25km...) came from. Figure 8 stops at 20km and figure 7 includes all the satellite datasets.
This comes from plots in both Fig 7 and 8. "In the 14-25 km…" has been changed to "Below 25 km…"

P29L12: Did the characterization of the lineshape changed from V2 to V3? If yes, how?
The characterization of the ILS did change between versions, and a reference to Boone et al., 2013 (which discusses the v3 ILS) has been added here.

P29L19: How can this be an improvement if v2 was within 20% and v3 can change up to 40%.
There is an improvement near the peak, and the v2 was "typically" within 20% and the max differences were near 40%. The text now states, "...ACE-FTS v2.2 NO_2 had a ~15% low bias near the peak, and between 20 and 40 km agreed with correlative data sets to within 40%."

P30L4: Did you check the a priori profiles? Is this just speculation? If it is, the author's need to change the language to reflect that.
This sentence has been deleted.

Figures: Titles should use subscripts. Also, the authors should use either 10ˆ-6, 10ˆ-9 or 10ˆ-12, that is to say, ppmv, ppbv or pptv, respectively.
All captions now use ppbv.

In the caption first caption, the author should specify the values of the dashed lines added for clarity.
Caption of Figure 1 now reads, "Dashed lines in the correlation (at 0.8) and relative difference plots (at -10, 0, and 10%) are provided for visual clarity."

Delete Figure 2 and add a caption on each of the following figures.
This has been deleted and legends have been added to the appropriate figures.

Figure 1: Presumably the authors shown v3.5 – v2.2 rather than v2.2 – v3.5
The caption has been corrected.

Figure 2: change MIPAS___IMK to MIPAS IMK/IAA Also SCIA to SCIAMACHY where needed.
These have been changed.

Figure 4: The two panels should be merged and a instrument average mean line (black thick line) should be added. Caption should state ACE-FTS – INST rather than ACE-FTS – HALOE.
These are split up because they have different coincidence criteria. "HALOE" has been changed to "INST".

Figure 5: Caption should state: mean relative difference profiles (ACE-FTS – MIPAS IMK/IAA)
This has been corrected.

Figure 8: Caption could say: Note that the occultation instruments have been excluded.
The text states, "Note that GOMOS and the solar occultation instruments have been excluded."

Figure 10: MIPAS__ESA and MIPAS___IMK need to be changed to MIPAS ESA and MIPAS IMK/IAA
These have been changed.