Reply to the comments of reviewer # 1 and reviewer # 2

We would like to thank both reviewers for their careful review of the manuscript, and their helpful and constructive comments.

Reviewer comments are in **bold face**, replies in *italic.*

Reviewer # 1:

This paper presents a comprehensive validation study that will be valuable to anyone using MIPAS data or planning a similar comparison of satellite data with surface-based measurements and other satellite instruments.

It gives a thorough explanation of how bias and significance are determined. It also discusses the validation of measurement precision estimates. In many applications, such as data assimilation, it is critical to have realistic measurement precisions as well as known biases. The paper is well written and as concise as could be expected for such a large set of instruments in the comparison. I only have a few minor comments.

ABSTRACT:

1) I would include a sentence that these results confirm that MIPAS is biased with respect to ECMWF for temperature around the tropopause.
Reply: This will be done.

I would include a sentence about also validating measurement precisions.
Reply: This will be done.

Sec. 4 Validation method
p4418 Line 6: delete the phrase: “which in many cases was small anyhow”
p4418 Line 8: “do not differ in any case” to “do not differ significantly”?
Reply: Both changes will be done.

Reviewer # 2:

General Comments
This paper describes the validation of MIPAS IMK-IAA temperature, water vapour, and ozone profiles using measurements obtained during the MOHAVE-2009 campaign. The validation data sets are obtained from radiosondes, frost point hygrometers, lidars, microwave radiometers, and FTIR spectrometers, as well as coincident satellite measurements by Aura-MLS, ACE-FTS, and AIRS. In most cases, the number of coincident measurements is relatively small (less than 100 for all but four pairs) given the single location and limited duration of the campaign. However, the comparisons are done thoroughly, with careful attention to assessing the observed differences in the context of the instrument errors and precision. The validation methodology is clearly explained, including the formalism for accounting for different altitude grids and vertical resolutions. The paper is logically organized and the results are described as concisely as possible given the large number of comparisons. This paper will be a useful reference for anyone using the MIPAS IMK-IAA temperature, water vapour, and ozone versions that are validated in this study. I recommend publication in AMT after the minor comments below are addressed.

Specific Comments

Section 3.1: There is some inconsistency in the level of detail provided for each instrument. In some cases, instrument performance is assessed, including estimates of systematic errors and precision, as well as vertical resolution. However, for other instruments, no such details are provided. Since this information must be available for all instruments, as it is used in the validation comparisons, it would be helpful to provide it for all instruments for greater consistency. One possibility would be to add a table summarizing vertical resolution, systematic error, and precision for all validation instruments.
Reply: We will provide the requested table.

Page 4410, line 20: Elaborate briefly on the "few typical signal corrections".
Reply: We will include a description of the signal corrections applied. An in-depth description of the TWF can be found in the following two papers:
McDermid, I. S., Leblanc, T., and Walsh, T. D.: Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere – Part 1: Instrument development, optimization, and validation, Atmos. Meas. Tech. Discuss., 4, 5079-5109, doi:10.5194/amtd-4-5079-2011, 2011.

Leblanc, T., McDermid, I. S., and Walsh, T. D.: Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere – Part 2: Data analysis and calibration for long-term monitoring, Atmos. Meas. Tech. Discuss., 4, 5111-5145, doi:10.5194/amtd-4-5111-2011, 2011.

Page 4411, line 22-29: This paragraph should be revised for clarity. e.g., line 23: The first was the ozone mode. . . line 24-25: ozone, temperature, aerosol, and water vapor were retrieved from these data. Same for line 26-27. Line 27: Explain the wavelengths included in third mode - the sentence first says only 355 nm was transmitted, but then says the filter blocked 355 nm, transmitting 387 and 407 nm.
Reply: We will reword this paragraph. Some more description can be found in Leblanc et al., 2011a. This reference will also be given here.

Page 4412, line 10-13: Is this paragraph needed? This section 3.2.1 has described the three lidars used in the study: TMF/TMW Raman lidar, ALVICE, and STROZ. What are the "Two other lidars permanently deployed at TMF"? If their data sets are being used in the paper (apparently not), why mention them? If they are, include full descriptions. Why mention STROZ and ALVICE again? List the measurements made by each lidar in the paragraphs describing them.
Reply: We have not added a specific description of the ozone lidars which were employed in addition to the TMF/TMW, ALVICE and STROZ lidars, because they were not part of the specific instrumentation of the MOHAVE-2009 campaign. We shall include some description of the ozone lidars in the revised version. We agree that STROZ does not need to be mentioned again, since its ozone measurement capability has been described earlier.

Page 4413, line 9-14 and 15-19: Add references for the frost point hygrometer biases, and for the ozonesondes.
Reply: We will add the following references: For frost point hygrometer biases (lines 9-14):
Hurst, D. F., Hall, E. G., Jordan, A. F., Miloshevich, L. M., Whiteman, D. N., Leblanc, T., Walsh, D., Vömel, H., and Oltmans, S. J.: Comparisons of temperature, pressure and humidity measurements by balloon-borne radiosondes and frost point hygrometers during MOHAVE 2009, Atmos. Meas. Tech. Discuss., 4, 4357–4401, doi:10.5194/amtd-4-4357-2011, 2011.
For the ozonesondes (lines 15-19):
Komhyr, W.D., "Electrochemical Concentration Cells for Gas Analysis," Ann. Geophys., 25: 203–210, 1969.
Komhyr, W.D., J.A. Lathrop, D.P. Opperman, R.A. Barnes, and G.B. Brothers, "ECC Ozonesonde Performance Evaluation during STOIC 1989," J. Geophys. Res., 100(D5): 9231-9244, 1994.

Sections 3.1.2 and 3.1.3: Perhaps the order of these two sections should be swapped, as the RS92 radiosondes are introduced in 3.1.3, but are discussed in the context of comparisons the frost point hygrometers in 3.1.2.
Reply: agreed.

Page 4416, line 16: This sentence leaves the reader hanging. State briefly the results of the MIPAS-MLS v2.2 temperature, water vapor, and ozone comparisons by Chauhan et al.
Reply: We will include a few sentences on the results of this comparison in quantitative terms.

Page 4417, line 6: Are there really oscillations in pressure profiles, or just in the temperature profiles?
Reply: We have reworded this sentence as follows: "This newest data version has reduced the occurrence of oscillations in the temperature profiles and the microwindows for all trace gases have been updated."

Page 4421: It would be helpful to add a table summarizing the transformations that were made on each MIPAS-validation data set.
Reply: This kind of table will be provided.

Page 4424, line 18: Qualify this statement as "just above". Not true for all altitudes above the tropopause.
Reply: This is correct, and the change will be done.

Page 4425, para 2: Comment on the STROZ results being opposite to those from the TMF lidar and ALVICE - negative differences below the tropopause, while the other two comparisons are positive in this region.
Reply: We will point out this behavior more explicitly.

Page 4434, Section 5.2.6: State how many profiles are included in Figure 15. Is there a consistent difference between the LTE and non-LTE profiles? If so, could a systematic correction be applied to the MIPAS data?
Reply: This figure is not the result of real measurements, but of test retrievals with synthetic data simulating nominal mode MIPAS measurements (with tangent altitudes between 6 and 70 km). Spectra have been simulated including non-LTE effects, and retrievals have been done from these spectra with and without including non-LTE. The figure shows the difference between the two retrieval results. In principle, a correction could be applied to the MIPAS data, but since this correction would depend on illumination (non-LTE excitation), the actual kinetic temperature profiles, and the actual amounts of other atmospheric constituents, and thus not be constant for different atmospheric states, we have refrained from applying such a correction. We will add a respective explanation to the text.

Page 4443, Acknowledgements: This section seems short given the large number of validation instruments; acknowledgements are missing for many, including MLS, ACE, and AIRS.
Reply: Acknowledgments for ACE, AIRS, and MLS will be added.

Page 4454, Figure 1: The caption for the right panel refers to a dotted line, which turned out as a solid line when printed. However, it is a dashed line (really two lines) in the electronic version.
Reply: We will improve the representation in the figures and correct the figure caption.

Page 44, Figure 2: This figure is very small. It would also be preferable to replace the dotted line in the right panels with a solid line for this figure and all similar ones, 4, 6, 7, 9, etc.
Reply: This will be done. Moreover, in the final layout of AMT the figures will be larger since it will take

a whole page.

Technical Corrections

Page 4405, line 11: define FTIR
Page 4406, line 1: altitude-resolved
Page 4406, line 12: particularly
Page 4407, line 6: measurements at the original
Page 4407, line 8: The text below seems to more consistently use the phrase "optimized-resolution nominal observation mode". Use that here for consistency.
Page 4407, line 18: allows coverage of the globe
Page 4407, line 21: field-of-view
Page 4409, line 6: Network for the Detection of Atmospheric Composition Change
Page 4409, line 9: Hygrometer
Page 4409, line 12: campaigns
Page 4409, line 22: The MOHAVE-2009 campaign not only. . . or MOHAVE-2009 not only. . . . Hosted used three times in this sentence.
Page 4409, line 26: centered (US) or centred (UK)
Page 4410, line 3: define PV; provide a reference for MIMOSA or elaborate on (Hauchecorne/CNRS)
Page 4410, line 8: and instruments operated
Page 4410, line 15: elsewhere referred to as Raman lidar
Page 4410, line 20: radiosondes
Page 4411, lines 8, 12, 13: add commas after measurements (8), measurements (12), and vapor (13)
Page 4411, line 17: signal-dependent
Page 4413, line 5: delete the
Page 4413, line 17: were the same model
Page 4414, line 3: launched during MOHAVE-2009
Page 4414, line 16: from the University of Bern
Page 4414, line 23: the altitude covered depends on the signal-to-noise ratios
Page 4414, line 26: provides an estimate
Page 4415, line 9: high-resolution
Page 4415, line 17: MLS has been used several times, but has not yet been defined
Page 4416, line 3 and 9: geo-potential or geopotential - choose one
Page 4416, line 5: in the case of
Page 4416, line 21: high-resolution
Page 4417, line 1: field-of-view . . . details of the ACE-FTS
Page 4417, line 9: define AIRS . . . into Earth orbit
Page 4417, line 10: medium-resolution
Page 4417, line 18: used CO2 above (line 3)
Page 4418, line 3 and 4: coincident or co-incident? Choose one and use consistently. Both are used throughout the paper.
Page 4418, line 13: better-resolved profiles
Page 4418, line 14: lower-resolved
Page 4418, line 22: (Rodgers, 2000; Chapter 10.3.1)
Page 4419, line 1-3: Here and elsewhere, inconsistent use of hyphens in compound adjectives: better-resolved, low-resolved - change to lower-resolution?
Page 4419, line 13: what does set union mean?
Reply: We thought this is a standard mathematical term describing the ensemble of data points (grid points in this case) which belongs either to the first or to the second or to both ensembles. We will reword this sentence:
"In these cases the profiles were interpolated to a grid which was the combination of both original

grids, containing all grid points from the one and the other original grid, and ….".
Page 4420, line 13, 16: add periods after equations (also Eqns. 13 and 15 on page 4422)
Page 4420, line 19: define vmr (why not VMR throughout?)
Page 4421, line 7: use water vapor for consistency
Page 4421, line 19: remove paragraph indent
Page 4421, line 20: on the coarser grid
Page 4421, line 26: is the mean difference
Page 4422, line 8: Why is root mean squares used throughout? Usually rms/RMS means root mean square.
Page 4422, line 22: into day
Page 4423, line 1: Replace "in special sections" with the actual section numbers for those following sections that discuss non-LTE, e.g., 5.1.5, 5.2.6, 5.3.4
Page 4423, line 4: nightly means
Page 4424, line 18: Qualify this statement as "just above". Not true for all altitudes above the tropopause.
Page 4424, line 19: differences relative to ECMWF
Page 4425, line 7: temperatures throughout the stratosphere
Page 4426, line 28: with MIPAS
Page 4426, line 27: hinting at overly optimistic
Page 4427, line 25-16: two mean temperature profiles . . . in the case of MIPAS, which
Page 4429, line 22: but did not
Page 4431, line 18: Here and elsewhere in the paper, replace "in the order of" with "on the order of".
Page 4431, line 19: consistent with
Page 4431, line 23: This is inconsistent with . . . or This disagrees with . . .
Page 4433, line 20: which causes
Page 4433, line 24: provides a picture consistent with
Page 4434, Section 5.2.6: State how many profiles are included in Figure 15. Is there a consistent difference between the LTE and non-LTE profiles? If so, could a systematic correction be applied to the MIPAS data?
Reply: see our reply above
Page 4435, line 5: thus the latter are
Page 4436, line 1: consistent with
Page 4437, line 13-14: below 25 km and above 35 km . . . differences are larger.
Page 4438, line 4 and 7: 60-65 and 65-70 use single and double hyphens – use AMT format
Page 4439, line 4: well captured . . . with the TMF lidar
Page 4439, line 5: throughout the stratosphere
Page 4439, line 10: temperatures
Page 4442, line 25: explicitly
Page 4443, line 1: as a consequence
Page 4443, line 5: into the instrument's optics by internal parts.
Page 4443, line 7: are accounted for
Page 4443, line 12 and 14: change "not further identified" to unidentified
Page 4453, Table 2 caption: Number of observations coincident with MIPAS for
Page 4455, Figure 2 caption, line 3: coincident with the MIPAS
Page 4468, Figure 15: State the number profiles included in this comparison - one or many?
Reply: Thank you for reading the paper so carefully. All suggested corrections will be included.

In addition to the changes suggested by the reviewers, a number of improvements became possible during the submission and discussion phase of the paper:

1) The MLS team made more correlative measurements available after the submission of this paper, and extended the altitude range of the provided profiles to higher altitudes. We have

used these data to improve the comparisons to MLS. Instead of only 3 co-incidences, 47 coincidences have become available now. The improved inter-comparisons between MLS and MIPAS on temperature, water vapor and ozone will be presented in the revised version of the manuscript, and related text will be adjusted accordingly. In particular, the comparison on temperature and water vapor have become much more favorable now (compare Fig. 1, new comparisons, to Fig. 2, old comparisons).

2) The unresolved issue with the application of WVMS averaging kernels on MIPAS profiles has been resolved meanwhile; further, Gerald Nedoluha has informed us that WVMS profiles should not been used for comparisons below 26 km. We will replace the comparison of MIPAS water vapor profiles to WVMS by a new figure including both the application of WVMS averaging kernels to MIPAS profiles and the reduction of the altitude range (see Fig. 3 for an example).
3) Within discussions among the co-author team, some criticism on the evaluation of the systematic error of MIPAS water vapor data with respect to uncertainties in spectroscopic data has been applied. In the original version of the manuscript, we have assessed the impact of uncertainties of spectroscopic data by perturbing both the line intensities and the line widths of spectroscopic lines used within the water vapor retrieval simultaneously within their uncertainties and calculating the impact on the retrieved volume mixing ratios. Gerald Nedoluha pointed out that by this approach, some error compensation could occur which is probably the reason that the systematic error of MIPAS water vapor profiles is virtually zero around 50 km (see, e.g. Fig. 12, middle panel, of the original manuscript). Within an improved approach to assess the systematic error of MIPAS water vapor profiles, we now perturb line intensities and line widths of spectroscopic lines independently from each other and add the derived errors linearly (in absolute terms). This leads to slightly increased systematic errors around 50 km (see Fig. 3 of this document). The improved assessment of systematic errors of MIPAS water vapor profiles will be used in all Figures of the revised manuscript where systematic errors of MIPAS water vapor have been shown before (Figs. 9 and 11-14 of the original manuscript), and the related text will be adjusted accordingly.
4) In order to account for the changes listed above, the summary figure (Fig. 19 of the original manuscript) will be adapted accordingly.

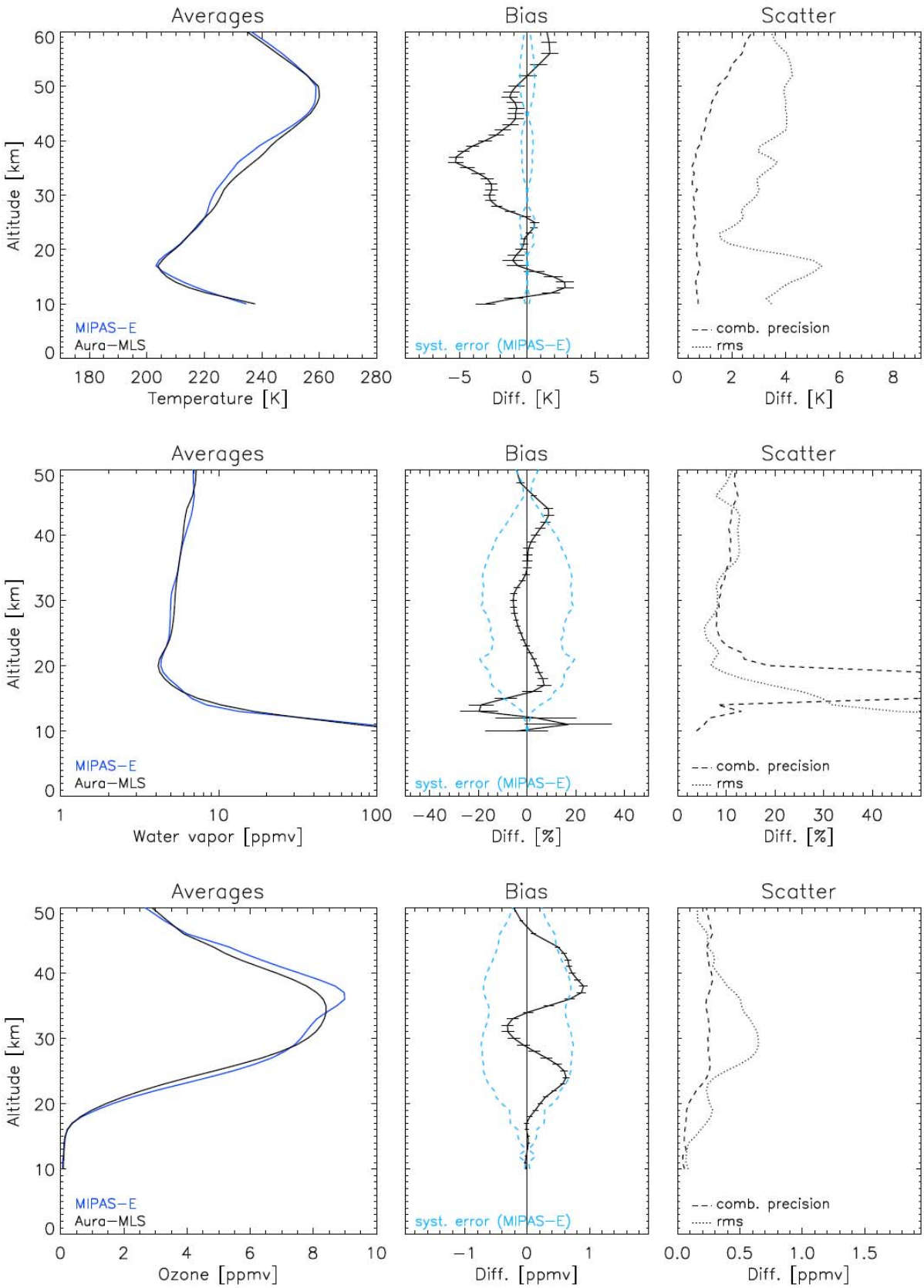

Fig. 1: Comparison of MIPAS and MLS temperatures (top), water vapor (middle) and ozone (bottom), based on 47 coincidences (new figure).

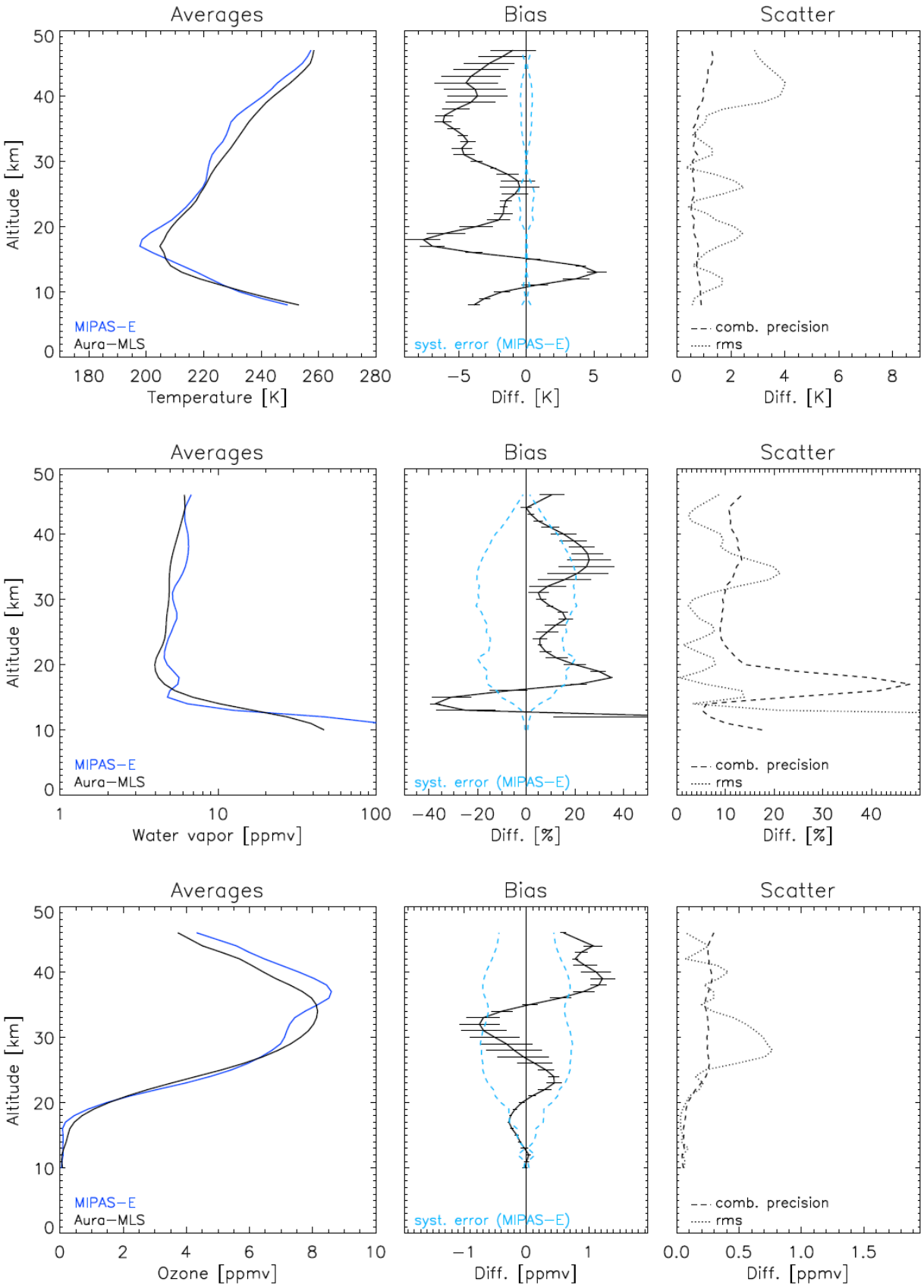

Fig. 2: Comparison of MIPAS and MLS temperatures (top), water vapor (middle) and ozone (bottom), based on 3 coincidences (old figure).

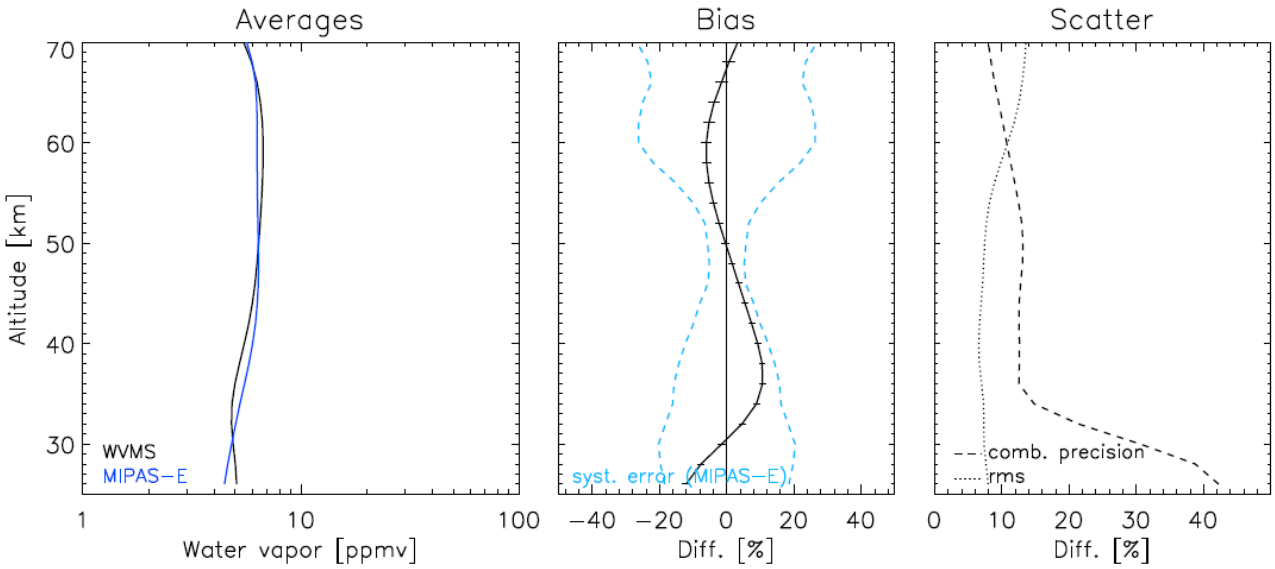

Fig.3: New Figure on the comparison of MIPAS and WVMS, with averaging kernel of WVMS applied to MIPAS.