The authors thank the reviewers for the constructive comments which have helped to improve the paper.

Answer to Reviewer #1.

1. Figure 4 should give the definition of the color scales.
The color scale is defined in the caption of the Figure : "Monthly latitudinal distributions of collocated measurements of MIPAS with reference instruments, *in percents*."

2. In section 5, the author proposed to evaluate natural variability by eq. 6, and point out that it might not be good for higher latitudes. Have you ever evaluate that natural variability using an independent way, such as the method proposed by Von Clarmann (2006).
Equations 57-59 in von Clarmann 2006 provide the natural variability within a latitude bin which include multiple measurements of the same instrument. Thus this bin has to be quite large. Contrary to that, the natural variability in Eq. 6 of present paper is the RESIDUAL variability related to matched pairs, which is caused by the finite width of the collocation criteria. Since natural variability is a function of the scales involved, the one does not help to nail down the other.
Further, it would be a logical circle to calculate natural variability from MIPAS measurements (assuming known MIPAS measurements errors) and to use these to verify the MIPAS error estimates via Eq. 6.

An estimation of natural variability comes as a side-result of MIPAS precision validation, which is shown in Laeng, A., Hubert, D., Verhoelst, T. et al, "The Ozone Climate Change Initiative: Comparison of four Level-2 Processors for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS)", under revision in Remote Sensing of Environment, 2014. The method used is the same as in recently published analysis by Sofieva et al. "Validation of GOMOS precision estimates in the stratosphere", AMT, 2014, http://www.atmos-meas-tech.net/7/2147/2014/amt-7-2147-2014.html. See in particular Figure 6 in the last paper.

In order to make the implications of natural variability clearer, the text after Eq (6) has been rewritten.

3. Follow the previous comments, in page 3965, line 1-6, the author claims the reason is large geophysical variability. But you also claim that in section 5, the estimates might not be good for higher latitudes, please address this conflict issue.
We claim that the reasons for big RV values are two-fold : large geophysical variability and underestimation of its uncertainties by POAM. We are not making any estimates of RV at higher latitudes. In order to be more clear, we change the text
"where geophysical variability tends to be higher relative to low latitude coincidences. "
to :
"In the northern hemisphere, POAM coincidences occur in the region impacted the springtime breakdown of the polar vortex. In the southern hemisphere, many coincidences are the near the edge of the fall/winter vortex. Both of these regions can be expected to have large geophysical variability."

4. Page 3966, Figure 7 compares the slopes with 1:1 relationship, but I do not see any slope in the figure. The text is not easy to follow. A linear regression should be plotted, and it is better to also provide some statics: y=ax+b, R value, for example. My suggestion would be to pick up several typical altitudes instead of all levels.
The 1:1 slope curves on the panels were made thicker. No regression line has been fitted, because the actual distribution of data points is not expected to be a linear function. We also provide figure 8 comparisons.

5. Also in Figure 7, the bottom right panel (with OSIRIS), MIPAS varies from minor values to 1 ppmv while OSIRIS keep zero, how does that happen?

The scatter plot with OSIRIS looks like if it was cut by a zero or close to zero lines on OSIRIS side : this reflects that in OSIRIS processors negative vmrs are cut off, filtered or replaced by a fixed value close to 0. In contrast, MIPAS IMK/IAA processors retrievals of negative vmrs, although unphysical, are allowed, hence avoiding biasing the statistics.

And the top right panel (with SAGE) seems using a different symbols, if so, you may need to re-plot it.

It was done intentionally: as reported in Table 2, there are only 189 collocations with SAGE, against 5 000 with ACE-FTS, 360 000 with MLS and 140 000 with OSIRIS. The 189 points of SAGE-MIPAS are barely visible on the plot. In contrast, the number of collocations with ACE, MLS and OSIRIS allow to do the plots with points.

6. Figure 8 shows that the bias with respect to MLS depends on time, how about similar comparison with other measurements.

The spatial sampling is not sufficient in some measurements, and thus does not allow capturing seasonal cycle well enough for meaningful comparisons with MIPAS. This analysis was performed for reference instruments that have known small bias, sufficient sampling, and sufficient time overlap: GOMOS, MLS, and OSIRIS. Similar patterns in bias were found in comparisons in all latitudes with all three instruments, GOMOS, MLS, and OSIRIS. We mention this in the revised version.

Similar comparisons are shown in figure 9, it's better to combine these two figures together.

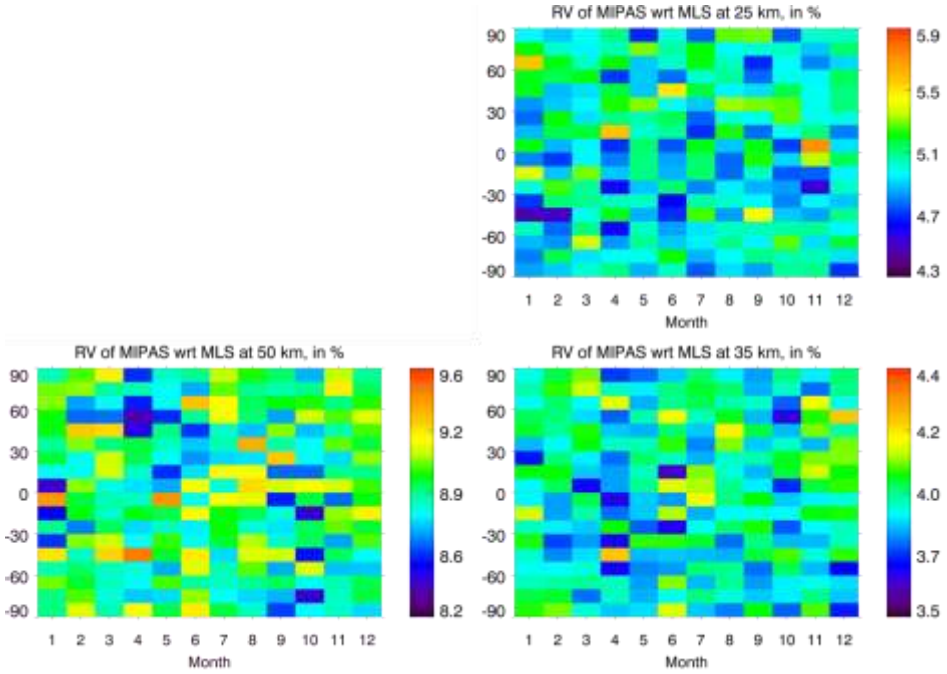
Figure 8 shows the evolution of ozone with time as seen by MLS (top panel) and MIPAS (middle panel). The bottom panel of Figure 8 is similar to bottom panels of Figure 9 by showing the evolution of relative bias with time.

Figure 9 shows the time evolution of the absolute and relative biases for two small-biased instruments with similar to MIPAS coverage, MLS and OSIRIS, with respect to MIPAS. Data for comparisons are selected in the same latitude bands, 30-60 N. New version now has the same latitude bands as in the original version the OSIRIS/MIPAS comparisons were shown in the 0-30N band.

These two figures illustrate different analyses and we prefer to keep them separated as in the original manuscript .

What about the pattern of the residual variability? Could you give some statements about this?

Following your suggestion, RV distributions have been analyzed (see plots below).



No significant patterns have been detected. This is attributed to the fact that the mean distance between the MIPAS measurement and the comparison measurement can in itself depend on latitude and thus mask any pattern of latitudinal variability. Since with respect to this no result has been found and in order not to overload the paper, we have decided not to include this analysis in the paper, but to mention the results in the text.

In Figure 8 and 9, only some particular latitude bins were chosen, could you give reasons?
The same patterns are observed on all latitudes bins, high-southern and mid-northern latitudes were chosen only as representatives.

7. In section 8, page 3971; the final statement is "over all, this MIPAS dataset has a small bias with respect to standard small-biased data : : :", this may not be correct and does not consist with the statements in page 3965.
Page 3965 deals with instruments that have a known bias (HALOE, HIRDLS, IASI, SBUV, SCIAMACHY, SMILES, POAM_III, SMR). This phrase refers to small-biased instruments: ACE-FTS, GOMOS, MLS, OSIRIS, SAGE_II, corresponding discussion is on pages 3963-3964.

And in the conclusion part you should also give some statement about the time dependent bias.
Drifts, i.e. long term deviations of the bias, as a function of latitude and altitude, have been analyzed by Eckert at al, Drift-corrected trends and periodic variations in MIPAS IMK/IAA ozone measurements, Atmos. Chem. Phys., 14, 2571–2589, doi:10.5194/acp-14-2571-2014, 2014. Short resume of their findings is added in the text of the paper. This discussion is also mentions now that new version 7 of Level 1 Spectra of MIPAS were released by ESA. Its production uses a new set of time dependent correction coefficients for the detector non-linearity (which changes in time due to ageing of the detectors for channels A, AB, B). The effect can be in the order of 3% in ozone change per decade (if AB info is used in the retrieval), i.e. trends will be determined more accurately with new data. Dedicated baseline for ozone retrieval is under development.

8. Figure 12 is poorly plotted. Lines with different colors are not fully explained, neither in the caption nor in the text.
The Figure was changed, the legend was added; the caption and explaining text were changed.