Reviewer #2

This is a nice paper that clearly describes a solid study of the impact of horizontal gradients on various approaches to retrievals from the MIPAS instrument. In principal I am happy to see this manuscript proceed to AMT, however, I only have one concern that I'd like to understand beforehand (plus some minor suggestions/comments for the authors to consider). The standard of English is excellent and the figures are very clear.

We gratefully acknowledge the reviewer for the constructive comments and suggestions that help in improving the quality of the paper.

The replies to the reviewer's comments are reported in blue below each comment.

My main concern surrounds the discussion on page 4, lines 10-15. As I understand it, the authors have taken model temperature and pressure profiles on a fixed height grid, and forced the pressures to be horizontally homogeneous while retaining the model horizontal temperature variability. If the altitude grid indeed remains fixed, then surely the model atmosphere that results is not in hydrostatic balance. As such, this becomes an unphysical simulation from which it is arguably hard to draw meaningful "real world" insights. Furthermore, are the lines being observed not subject to significant pressure broadening, making pressure the dominant contributor to the radiance signals? As such, I would have thought that horizontal gradients in pressure are arguably the most important thing to study the impact of (though one would probably ultimately quantify it in terms of impact to the temperature/composition profiles as interpolated to a fixed pressure grid, being the product most widely used in the community). While I understand the authors desire to "focus the analysis on the error caused solely by the approximations in modeling the horizontal variability of temperature and target gas", surely, if the pressure gradients are indeed the dominant term, they should have been included in the analysis. I'd like the authors to address this point, and consider revising their approach, or at least undertaking a separate quantification of the impact of pressure gradients (but again, the unphysical nature of their atmosphere would limit the usefulness of the result).
Perhaps I have misunderstood the description in the manuscript, in which case, greater clarity is required.

The reference atmosphere used for the analysis was made using temperature and pressure profiles from the model on a fixed altitude grid, and forcing the pressures to be horizontally homogeneous while retaining the model horizontal temperature variability (as correctly stated by the reviewer). Then we retrieve both temperature and pressure (plus VMR when required) on a fixed altitude grid, retaining the horizontal variability of either T or VMR as in the model if 1D + gradient retrievals are performed. Since the retrievals are all performed with the GMTR code and into the GMTR code the pressure is left free to vary during the p,T retrieval procedure without imposing hydrostatic equilibrium, we obtain T and P on a fixed vertical grid. The retrieved pressure is then used as vertical coordinate for the retrieved Temperature when comparing the profiles with the reference ones. Then, as correctly stated by the reviewer, the final impact is evaluated on temperature/composition profiles on a fixed pressure grid.
The reviewer's main concern is that such an approach, where the pressure is considered as horizontally homogeneous and no horizontal pressure gradient is accounted for, is not enough realistic to reproduce real measurement conditions. At the beginning of this work, to test the capability of our scheme to correctly reproduce real MIPAS L2 behavior as part of ESA-ESRIN Contract no. 21719/08/I-OL, we compared the 1D AX-DX differences obtained with our approach with the ones calculated with ESA level 2 products (real data) for the same months in several years. Apart from differences due to day/night conditions (not modeled in our study) in general the ESA 1D AX-DX differences and our 1D results matches well in both shapes and amplitude. These comparisons are reported in the Technical note "Investigations on horizontal inhomogeneities issue: Outcome of WP 9000".
Further tests were performed comparing the results we get with the 1D and 1D+T gradient approach with the ones reported in Kiefer et al. 2010. As also stated in the manuscript, we compared our results for T, HNO_3 and CFC-11 in terms of AX-DX differences with the one in

Kiefer et al., 2010 and find similar results. All these findings suggest that 1) the scenario used in our test is realistic enough to draw meaningful conclusions and that 2) the pressure gradients have a second order effect with respect to temperature and VMR gradients.
In addition, following the reviewer's suggestion, we performed a test using the reference atmosphere of 21th March in which both pressure and temperature are not forced to be horizontally homogeneous. Then we performed a 1D and 1D+T gradient retrieval and calculated the differences with respect to the reference (not homogeneous in pressure) Temperature field. These results are finally compared to the ones used for the analysis reported into the manuscript (obtained with the horizontal homogeneous pressure field). As we can see from Figure R1 reported below, we can hardly discriminate between the case where the pressure is horizontally homogeneous (left column maps) or not (right column maps). In both cases the temperature horizontal gradient is by far the major contributor to the 1D error. This finding is in agreement with what stated in Kiefer et al., 2010: "Furthermore, the facts that in the 1-D temperature retrievals there is already a clear effect, and that species retrievals from mid-IR emission spectra strongly depend on temperature, suggest to refine this hypothesis: the major part of the ascending/descending differences of 1-D retrieval results is caused by the neglect of the horizontal temperature inhomogeneities in the retrieval algorithm."

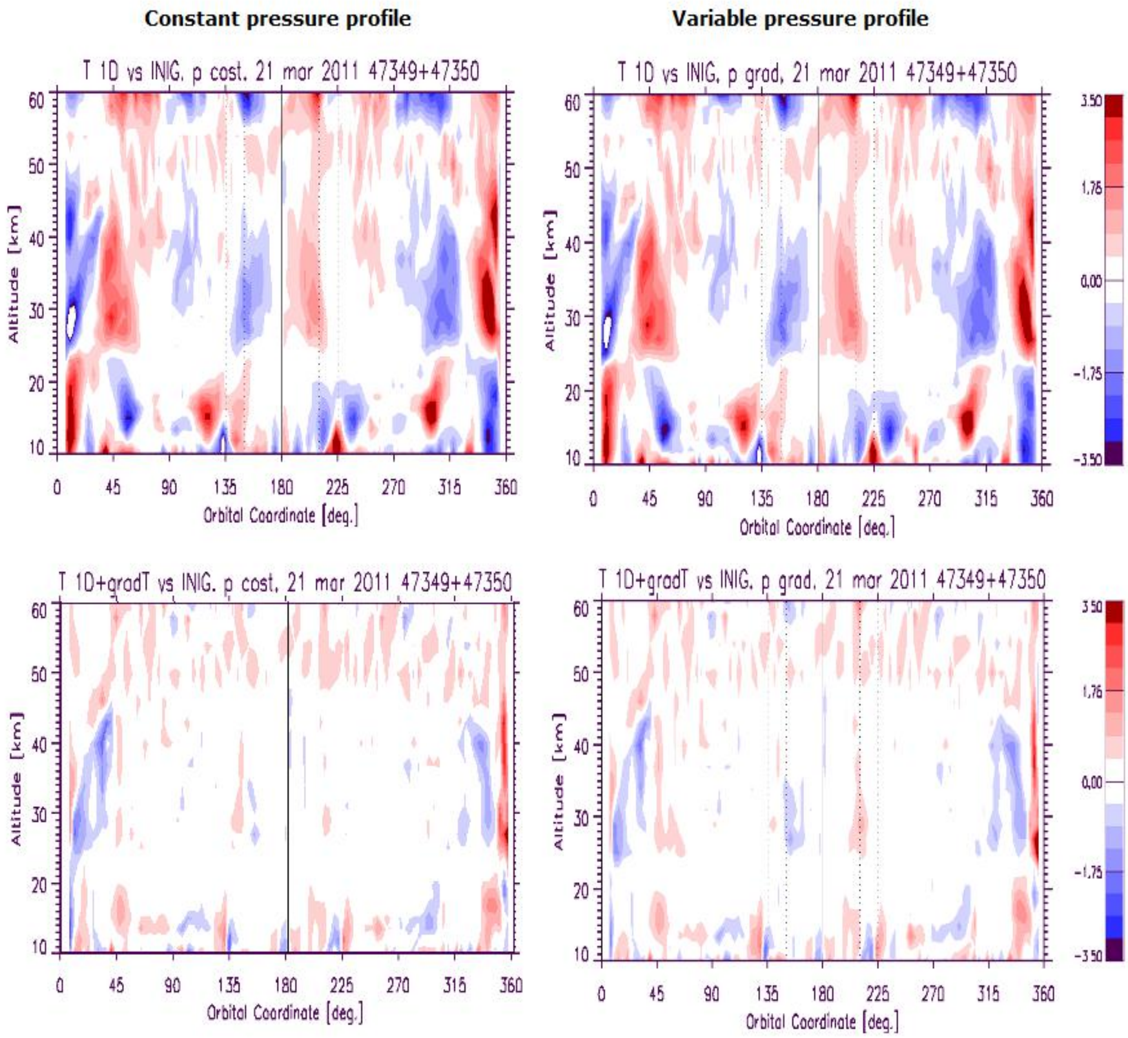

Figure R1: Left column: Temperature field retrieved with 1D code minus reference (top) and 1D+gradT minus reference (bottom), using in the reference atmosphere a horizontally homogeneous pressure field. Right column: Temperature field retrieved with 1D code minus reference (top) and 1D+gradT minus reference (bottom) using a reference atmosphere with horizontally inhomogeneous pressure field.

In order to quantify the impact of pressure gradients, as suggested by the reviewer, we built a reference atmosphere for the 21 of Mach where both temperature and composition are horizontally homogeneous while only the pressure profiles varies with latitude. Then we performed a 1D retrieval. The impact on temperature field is evaluated for this analysis comparing the results with the ones obtained using the 2D retrieval approach. The results are reported in Figure R2. As we can see from this figure, the effect of pressure gradients on 1D retrieval is negligible and it is of the same order of the smoothing error component.

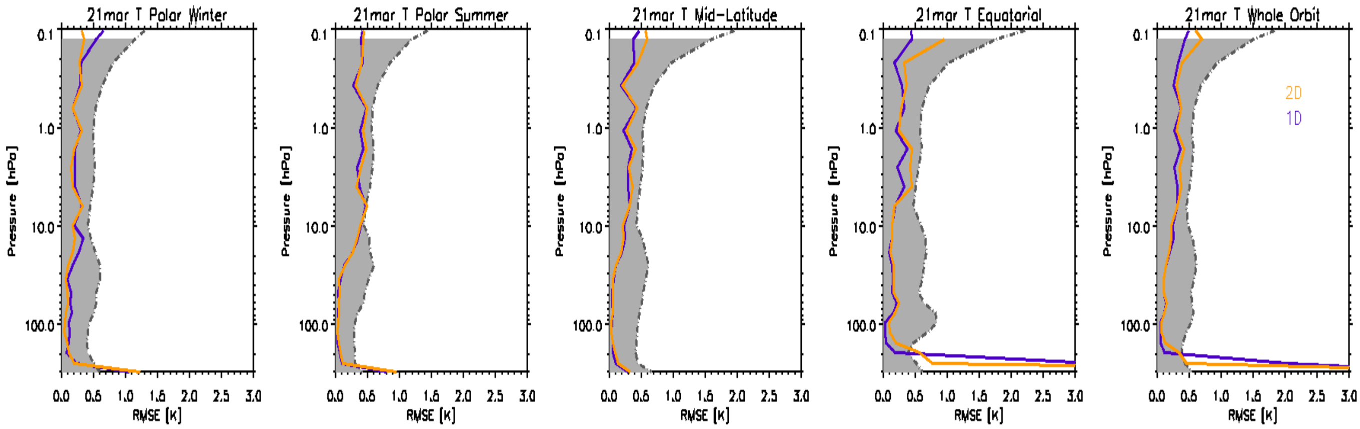

Figure R2: RMSE error on temperature due to pressure gradients on 1D retrievals (blue line) and on 2D (yellow, smoothing error). The grey areas represent the average ORM random error.

Finally, in the frame of MIPAS QWG (Quality Working Group) activities, we performed some tests with real data using a 1D+gradient approach. In these tests we also accounted for the horizontal pressure gradients. The results show that accounting or not for these gradients produced negligible differences on AX-DX differences. (see slide 15/49 of https://earth.esa.int/documents/700255/2551278/2.9_Ridolfi_and_Sgheri_testing_orm_v8.pdf).

According to reviewer's suggestions we revised our paper as follows:

1) We now assess the validity of our assumptions in reproducing real measurements behaviour, by comparing the AX-DX differences from simulated data with the ones from ESA IPF V6.0 Level2 MIPAS data computed by M. Kiefer in the frame of the MIPAS QWG activities. We compare those differences for the corresponding month in the years from 2005 to 2010 to our results by overplotting them in figures 2-3-4. In the text we comment about the good agreement between those data by adding at pag.6 line 3 of the original manuscript "For comparison purposes, in the same figures we show the values of AX-DX differences calculated from ESA IPF V6.0 level2 MIPAS data of December 2005 to 2010. Simulated 1D retrievals and real measurements show a very similar behaviour for most of the target species in the altitude range where ESA AX-DX differences are available, despite of the fact that different years are used. The amplitude of the 1D AX-DX differences is comparable to that of real data, confirming the fact that the simulated observations used in our tests are suitable for reproducing the behaviour of real MIPAS measurements." Accordingly in the caption of Figure 2 we added: "in green the differences from ESA IPF V6.0 Level2 MIPAS data in December form 2005 to 2010". M.Kiefer how provided the ESA V6.0 AX-DX differences has been added to the list of authors.
2) At page 4 line 12 of the original manuscript (see also the reply to reviewer#1's comment) we included a sentence about the impact of the missing horizontal pressure variability mod-el in the 1D retrievals. The sentence is: "Further tests, on simulated and real data (private communication), demonstrate that the pressure horizontal gradient has an almost negligible effect on 1D retrievals. For this reason in this study we set the pressure profiles to be latitu-dinally constant to better isolate the error due to the 1D assumption on the other targets."

==== Minor comments
—- Page 3
Lines 27/28: I don't understand this sentence. By "average atmosphere in a given latitude band" do you mean a zonal mean, i.e., an average over all longitudes? If so, hasn't longitude, by definition, been ruled out. Do you mean averages over a longitude range should not depend on the choice of

range (clearly not an appropriate assumption geophysically). Please clarify.

This sentence refers to the work performed by Kiefer et al., 2010 using real measurements. In their paper the authors use a set of MIPAS measurements covering 10days/one month to calculate AX and DX averages in each latitude band. Given the large size of the considered dataset and considering the MIPAS sampling rate, it turns-out that in a given latitude belt the different longitudes are evenly sampled in the AX and the DX parts of the satellite orbits. For this reason the deviations from zero of the AX-DX differences cannot be ascribed to the different sampling longitudinal intervals in the AX and DX parts of the orbits. In this sense "the average atmosphere in a given latitude band should not depend on longitude therefore, with a perfect retrieval scheme, the zonal averages of retrieved parameters should be equal when computed from measurements in the ascending or descending parts of the orbit".
Since in the model the AX and DX atmosphere could be different, in our one-orbit tests we used a AX-DX symmetric atmosphere. If we would have used a non-symmetric atmosphere, differences due to different longitudinal positions of the profiles in AX and DX part of the orbit could influence the AX-DX calculation.
In order to clarify this point, in the revised version of the paper, we modified the sentence "The results of Kiefer et al (2010) are based on the assumption that the average atmosphere in a given latitude band should not depend on longitude, therefore with a perfect retrieval scheme the zonal averages of retrieved parameters should be equal when computed from measurements in the ascending or descending parts of the orbit." in this way:
"The results of Kiefer et al. (2010) are based on the assumption that the average atmosphere in a given latitude band should not depend on longitude. Actually those averages are computed using several days of measurements allowing for an almost equal longitudinal distribution of ascending and descending profiles. Therefore with a perfect retrieval scheme (and a constant in time atmosphere) the zonal averages of retrieved parameters should be equal when computed from measurements in the ascending or descending parts of the orbit."

—- Page 4
Line 5: Please be more quantitative; what is the spatial resolution used in the FM?

The spatial resolution of the FM is the one of the EMAC atmospheric model: 1.4 degrees. We included this information in the revised version of the manuscript: we modified page 4 line 5 "the discretization of the atmosphere is sufficiently fine .." into "the discretization of the atmosphere is sufficiently fine (1.4 °)".

Lines 10-15: Please see discussion above.

Please see answer above.

Lines 17-19: First, please explain why you needed to add any noise at all? Why not simply do a noise free simulation. Also, please state whether, in addition to adding 1/40th of the expected noise, you also quote that 1/40th value as the radiance precision in the retrieval calculation (S_y in the Rodgers formulation), or does the retrieval still believe that the noise is at the 100% level?

We use 1/40th of the nominal noise specification both to add a synthetic pseudo-random error to the simulated measurements and to define their error covariance matrix Sy in the retrieval. Due to the different discretizations of the atmosphere in the FM that simulates synthetic measurements and in the FM internal to the inversion algorithm, we find that convergence of the retrieval is much more difficult if no noise is used. On the other hand the noise error must be significantly reduced with respect to the nominal case in order to make the horizontal model approximation the main source of error as already performed in in Carlotti et al., 2013 for the evaluation of the position error on MIPAS 1D retrievals. This sentence has been added into the revised manuscript page 4 line 27

—- Page 5
Line 20 (and 22). Your "diff" has been typeset in math mode, you'll want to typeset it in text mode

(e.g., \text{diff}, using the amsmath.sty package).

Ok, done thanks.

—- Page 6
Line 26: I think this would be better "Polar and mid-latitude winter conditions" if that's what you mean. As it is it could mean "Polar winter" and "mid-latitude all seasons".

It is “Polar winter” and “mid-latitudes” all seasons as described in section 2.4 page 5 lines 28-29.

—- Page 7
Line 2: As for page 6, line 26. Possibly elsewhere that I didn't catch also.

See answer to comment above.

—- Figures
Figures 2 and on: I would much prefer to see temperature errors quantified in K than in %. K are much more accessible to the general reader.

Percentage values are used only in Figure 2, in Figure 5 and onward we used K. According to the reviewer's suggestion, in the revised version of the paper we changed Figure 2 using absolute values in K instead of percentage values for AX-DX differences.