Reviewer #3

Review of “Errors induced by different approximations in handling atmospheric inhomogeneities in MIPAS/ENVISAT retrievals” by Castelli et al.

This paper provides simulated results that describe the need for a two-dimensional retrieval approach to infer vertical information from MIPAS limb emission measurements.
The authors simulate measurements and perform a variety of different retrievals in an attempt to quantify the errors induced by an assumption of horizontal homogeneity within the retrieval scheme. The simulations are by no means comprehensive but the reader is left with the impression that any retrieval scheme that uses two-dimensional information is less likely to produce artefacts related to atmospheric structure along the satellite track. This is of course by no means a surprising result. Although this paper provides no real information it is another work that supports the hypotheses that in order to accurately retrieve atmospheric composition information from limb measurements, a set of dense observations and a multi-dimensional retrieval are required.

The authors thank the reviewer for carefully reading the manuscript and for the useful comments. The replies to the reviewer's comments are reported in blue below each comment.

Major Comments:
I found the paper to be well written and the information that was presented was done so in an organized and easily understood manner.

It is very unclear to me what I am supposed to have learned from reading this paper.
The limited number of simulations performed do not allow me to quantify the typical error associated with MIPAS retrieved results. They may give me a feel for the seasonal dependence of certain errors and where in the vertical profiles these errors may occur, but I struggle to believe the errors have been “quantified” in any meaningful way. I think the main take home message of the paper is that to accurately retrieve information from the vertical profiles of limb emissions a two-dimensional scheme is required. This is well known. The paper is a report of some work but it’s use for the interpretation of MIPAS data is not clear. I would really like the authors to improve their discussion related to how their results guide the reader to better interpret artefacts within the standard MIPAS data products.

The reviewer's concerns are**:**
1) Limited number of simulations used to quantify the errors:
In the standard MIPAS systematic error analysis, the error due to to horizontal inhomogeneities is evaluated by using a temperature gradient of +/-1K/100 km (see Dudhia et al., 2002 and http://eodg.atm.ox.ac.uk/MIPAS/err/), while no error due to VMR horizontal inhomogeneity is considered. In Carlotti et al. 2013, the authors used a more sophisticated horizontal variability model with respect to the +/-1K/100 km to evaluate the so called “position” error on 1D retrievals. In this case the authors used a single atmosphere retrieved from real MIPAS measurements acquired during one orbit. Our analysis is based on simulations of 4 days of measurements in four different seasons to assess the errors due to neglecting the horizontal variability or to using a simplified approach to model it. This is a clear improvement with respect to the existing assessments.
2) **Relations between simulated retrieval results and MIPAS standard products not clear:**
As reported also in a reply to reviewer#2's comment, in the frame of the ESA-ESRIN Contract no. 21719/08/I-OL, we compared the 1D AX-DX differences obtained with our synthetic spectra with the ones calculated with ESA products when analysing real measurements for the same months in several years from 2005 to 2010. Apart from the differences owing to the day/night variations of the atmosphere, that are intentionally excluded from our simulations, the AX-DX differences derived form the ESA v6 1D products match pretty well both in shape and amplitude. These comparisons are reported in the Technical note ”Investigations on horizontal inhomogeneities issue: Outcome of WP 9000“. An example of the comparison for temperature is reported in Figure R1 below. Considering all the limitations of the simulations and the fact that in case of real measurements

part of the differences, especially at high altitudes are due to day-night variability of the atmosphere (see e.g. solar tides and changes due to photochemistry), the overall agreement is good.

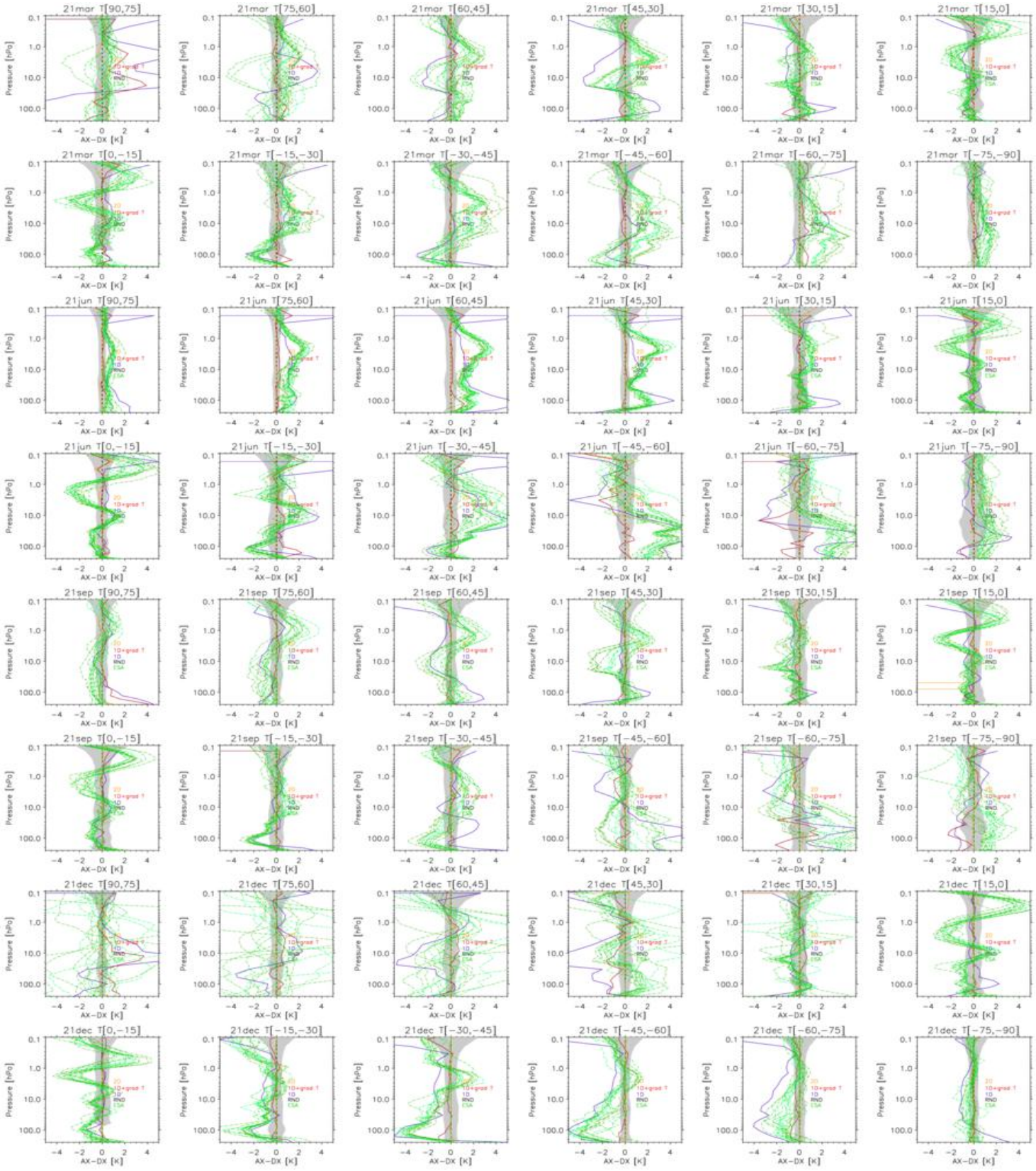

Figure R1: Temperature AX-DX differences from simulations for the four days considered in our study and from ESA V6 retrievals for months in corresponding season (e.g. December January February for 21 December) and years from 2005 to 2010.

In order to improve the paper, as suggested by the reviewer, this information is now included in the

revised version of the manuscript. We compare the AX-DX differences calculated with our simulated observations for the 21 December in the -45/-65 latitude band with those extracted from ESA Level 2 V6 monthly means for the corresponding month from 2005 to 2010, for each target parameter.
The detailed report of the changes we introduced in the manuscript due to these considerations are listed below, in the replies to comments labeled as “page 3-line20” and “Summary”.
These comparisons demonstrate that the synthetic measurements used in our tests and the analysis performed with those measurements can be considered representative to assess the error in MIPAS Level 2 products, due to different approximations in modeling the horizontal variability of the atmosphere. Thus we conclude that the results of our study can really be used for the interpretation of MIPAS retrieved profiles.

Following these considerations, we conclude that the results we show in the paper are 1) a clear improvement with respect to what was done so far to assess the errors due to neglecting horizontal variability, 2) representative when compared to real MIPAS data and thus can be used for the interpretation of real retrievals and for the evaluation of the different strategies used to cope with horizontal inhomogeneities.
For these reasons we think the work presented in the paper provides real information to the reader.

Comments and concerns

(page 2-line 7) it is stated that the MIPAS observations are exploited. This work is entirely simulation so I don’t see any exploitation of the MIPAS data.

We agree with the reviewer. Even if in the revised version of the paper we also add AX-DX differences calculated from ESA level 2 MIPAS data, the error estimation is based on synthetic observations. Thus in the revised version of the manuscript we changed the text accordingly: we replace “the measurements of” with “synthetic observations simulated for” in page 2 line 7 of the original manuscript.

(page 2-line 29) I think the authors should spend some more time justifying the statement that a 1.4 degree model is highly resolved. How does this resolution relate to MIPAS sampling resolution? The paper should do a better job of demonstrating the
forward model of the radiance measurements is sufficient to accurately simulate realistic MIPAS measurements. I think the paper is trying to tell me that MIPAS 1-D retrievals have errors so the forward model must be justified as an accurate representation of MIPAS measurements in order for me to interpret the two-dimensional results.

Regarding the horizontal resolution: the horizontal sampling of MIPAS nominal limb-scans is about 4 degrees, approximately equal to the horizontal resolution of the measurements. In this sense a model atmosphere with 1.4 degrees resolution is “highly resolved with respect to the MIPAS sampling”. Following the reviewer's suggestion we included this information in the revised text: “(1.4° in latitude, much finer than the MIPAS horizontal sampling of about 4°)”.
Regarding the forward model: We recall here that the 2D Forward Model (FM) used in the paper is the one internal to the 2D GMTR code. The capability of this FM to correctly model MIPAS measurements accounting for the horizontal variability was clearly demonstrated in Kiefer et al., 2010. In this paper the authors show that the AX-DX differences calculated from 2D GMTR retrieved profiles are very similar to those calculated from ECMWF profiles. The AX-DX differences owing to the 1D retrievals are always greater. This is a solid demonstration, based on real data, of the capability of the 2D FM to correctly reproduce MIPAS measurements and to model the horizontal variability. As suggested by the reviewer we included this information in the revised version of the manuscript (section 2.2 “Synthetic observations”, page 4 line 4 of the original manuscript) : “In Kiefer et al., 2010 the authors demonstrate the capability of the GMTR code to correctly model the features of MIPAS measurements due to the horizontal variability of the atmosphere. AX-DX differences calculated from profiles retrieved with the GMTR code are very similar to those calculated with the corresponding ECMWF data. In contrast, AX-DX differences calculated from profiles derived with 1D codes show features (not present in ECMWF data), hence

related to an incorrect modelling of atmospheric horizontal variability."

(page 3-line20) Is a four day and only eight orbit set of observations sufficient to quantify the errors in the one-dimensional approach? I believe that some specific errors have been quantified but I need more information to know that "the errors" have been quantified. Once again for this paper to be useful it must tell a complete MIPAS related story.

As above, the main reviewer's concern here is that the set of simulated observations used in the paper is not enough representative of the natural horizontal variability encountered by real MIPAS measurements to correctly quantify the errors of the 1D approach. In the revised paper we address this point by comparing the effect of horizontal variability on 1D retrievals performed with both real and our simulated data. As a quantifier for the error implied by the 1D approximation we use the AX-DX differences, that represent also the first experimental evidence of the effect of the horizontal homogeneity assumption on MIPAS-ESA 1D retrievals. As mentioned above, we compared the 1D AX-DX differences obtained with our synthetic spectra with those calculated from ESA products, for the same months of several years from 2005 to 2010. The good agreement between the observed and simulated AX-DX differences is both in shape and amplitude. This test proves that the simulated set of observations is suitable to pursue the objective of the paper, that is the assessment of the error implied by different approximations in modelling the horizontal variability of the atmosphere.

Owing to these considerations, we modified paper as follows:

1) In Fig.s 2, 3 and 4 we included also the curves related to the the ESA v6 AX-DX differences for the month of December, in the years 2005, 2006, 2007, 2008, 2009 and 2010.
2) We included a comment regarding the good agreement between observed and simulated AX-DX differences. The comment is in pag.6 line 6 of the original manuscript " For comparison purposes, we report in the same figures the values of AX-DX differences calculated from the 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 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 2005 to 2010". M. Kiefer who provided these differences has been added as an author and removed from acknowledgements.

If needed a supplement reporting analogous differences calculated for the other four days can be added (e.g. as Figure R1 in this document).

Summary
I found the paper to be well written but without much value in its current state. The results presented need to be linked more to the MIPAS measurements in order for them to be of value. It is very well known that two-dimensional retrievals do a better job of retrieving two-dimensional structure and without a more comprehensive link to the MIPAS data set this paper simply reiterates this well known fact. If the authors attempt to link their simulated results to the existing MIPAS data sets in a more realistic fashion I will be happy to look at the paper again.

We have already replied above to the reviewer why we think that the paper contains valuable information for the reader. In addition, as stated above, in the revised version of the paper we address the reviewer's concern and link our results, obtained with simulated spectra, with those obtained with real MIPAS measurements. This is done by including in plots and discussion the results from ESA V6 level 2 data. In particular in figures 2-3-4 we included the AX-DX differences calculated from ESA v6 data for December of the years from 2005 to 2010. The figures show that the real and simulated AX-DX differences obtained with the horizontal homogeneity assumption agree quite well both in amplitude and shape. Real AX-DX differences observed in ESA Level 2 v6

products in the days corresponding to the four seasons, and in all latitude bands can be given provided as a supplement, if required.