# Answer to Review by Editor

17.01.2022

We thank the Editor Mathias Palm very much for his time, thorough revision, and helpful comments. In the following, we provide our answers to each of the comments and corrections. The original Editor comment is repeated in bold, and our answers are provided in *italic*. Changes in the manuscript are indicated in *blue italic.*

**Comments to the author**:
**Dear authors, many thanks for your work and the detailed reply to the reviews. However there is one point which I do not find convincing:**

**In your paragraph about the AVK use (page 10 first paragraph), you write, that you do not use AVK's because you do not expect that the situation improves significantly. I find this is an odd statement in the context of your work.**

**You cite Ungermann (2011) as a justification, that the AVK's would not have an effect. But when I understand Ungermann (2011) correctly, this is the case when tomographic retrievals are used. Are you using the tomographic retrieval scheme? I did not find this information in your paper.**

**A major part of this work is to compare the GLORIA measurements with the models ICON and EMAC. You write several time, that even the fine structures are remarkably well resolved.**

*We thank the Editor for pointing out this weakness of our discussion and agree that clarification is required. A tomographic retrieval scheme was not used in our study. We applied conventional 1-D retrievals of single vertical profiles. Profiles are combined to 2-D vertical cross sections along flight track (see Johansson et al., 2018a). For clarification, we modified P5/L9-10 as follows:*

*"Details on the applied 1-D trace gas retrieval and the data products used here are provided by Johansson et al. (2018a). The retrieved individual trace gas profiles of GLORIA are combined to 2-D vertical cross sections of the respective species along the flight track."*

*The retrievals can be characterised with respect to (i) vertical smoothing by a conventional 1-D averaging kernel and (ii) horizontal smoothing along the viewing direction (i.e. to the right hand side of the aircraft) by a more complex 2-D averaging kernel. The latter is discussed by Ungermann et al. (2011) in their Sect 3.2 and Eqn (9) (although these authors focus at tomographic retrievals). The application of the 2-D averaging kernel specific to GLORIA chemistry mode observations (such as used here) is discussed by Woiwode et al. (2018). We added this reference in the manuscript.*

*In the vertical domain, conventional 1-D averaging kernels from the trace gas retrievals describe the amount of smoothing, which results from the constraint (in our case a Tikhonov approach), which reduces the vertical resolution. However, the achieved resolution in the altitude domain used for the respective trace gases is here in the order of 500 m and therefore well comparable to the vertical resolution of the models. Therefore, the application of conventional 1-D averaging kernels to the model data in the vertical domain is not expected to affect the comparison notably and the direct interpolation of model data to the GLORIA tangent points, such as done in many other studies (see e.g. Johansson et al., 2019; Khosrawi et al., 2017; Braun et al., 2019), is applied here. The quality of the GLORIA data used here is furthermore confirmed by in situ comparisons. As discussed by Johansson et al. (2018a), median* *differences and median absolute deviations between the GLORIA chemistry mode data and in situ observations during the entire PGS campaign were as low as -0.13 ± 0.63 ppmv (water vapour), -3.5 ± 116.8 ppbv (ozone), and -0.03 ± 0.85 ppbv (nitric acid). Since the in situ data measured during PGS covered regions characterised by strong vertical gradients, this gives us confidence that the GLORIA data are not affected by overall systematic biases that are relevant here.*

*In the horizontal domain along the viewing direction (i.e. to the right side of the flight track), model comparisons can be affected by limited horizontal resolution of a limb sounder along the line of sight. This effect can be taken into account by applying a 2-D averaging kernel for the interpolation of the model data, which takes into account also the dimension along viewing direction (Ungermann et al., 2011). However, this approach is computationally demanding, and leads, if strong gradients along the line of sight are avoided, only to a moderate improvement of direct comparisons of fine structures in vertical cross sections. This aspect has been investigated by Woiwode et al. (2018) (see their Appendix A) for the case of water vapour, when detailed observations of a tropopause fold were compared with high resolution ECMWF IFS data. In contrast, in case of strong local variations and gradients along the line of sight, as in the case of temperature variations due to gravity waves (which are not analysed here), the application of 2-D averaging kernels can strongly improve direct comparisons.*

*The flight analysed here was planned with the help of forecasts so that strong gradients in the meteorological fields along the line of sight were mostly avoided. The GLORIA observations were mostly aligned in a way that the viewing direction was aligned into relatively homogeneous air masses along viewing direction (compare Fig. 1b with 4b). This enabled us to resolve the fine filaments discussed in Section 4.3 that agree remarkably well with the model data. Naturally, an optimal alignment is not possible in all cases, and local differences between the GLORIA and model cross sections of trace gases due to remaining effects by horizontal gradients cannot be excluded. However, when the GLORIA and model data are analysed as ensemble (e.g. in mean values of correlations), these remaining smoothing effects are expected to cancel out on average in the large amount of data used here.*

*For clarification, we modified P10/L1-8 as follows:*
*"The vertical resolution of the GLORIA data used here is in the order of 500 m, depending on altitude and parameter (see Johansson et al., 2018a), and therefore comparable with the vertical resolution of the simulations by both models in the tropopause region. Therefore, the use of 1-D averaging kernels in the vertical domain, such as often used in context of vertical profiles retrieved from satellite limb observations (e.g. Microwave Limb Sounder (MLS)) that are characterized by notably coarser vertical resolution is not expected to improve the comparison significantly. The absence of relevant overall systematic biases in the GLORIA data used here is furthermore confirmed by in situ comparisons (see Johansson et al., 2018a).*

*Due to the limb viewing geometry, strong horizontal gradients along the line of sight of GLORIA (i.e. towards the right hand side of the flight track) can affect direct comparisons of vertical cross sections of atmospheric parameters derived from the GLORIA observations and interpolated from the models at the tangent points. This effect can be taken into account by interpolating the model data with the help of 2-D averaging kernels (Ungermann et al., 2011, their Sect. 3.2). As discussed by Woiwode et al. (2018) in a case study where the mesoscale fine structure of a tropopause fold was investigated, the application of 2-D averaging kernels improves the model comparison only moderately if the observations are aligned such that horizontal gradients in the trace gas fields along the line of sight are small (see their Appendix A).*

*Aided by meteorological forecasts, the flight analysed here was planned so that the GLORIA observations were mostly aligned in such a way. This can be seen by comparing Fig. 1b with Fig. 4b, for example during the backward leg to Kiruna, when the GLORIA limb views were aligned along the* *direction of moist filaments above Greenland. Therefore, the viewing geometry allowed us to resolve the fine structures of the narrow filaments discussed in Sect. 4.3 remarkably well. Due to the suitable alignment of the GLORIA observations during the discussed flight and since the application of 2-D averaging kernels is computationally demanding (particularly in case of the GLORIA high spectral resolution chemistry mode observations that employ a large number of spectral sampling points), 2-D averaging kernels are not applied here. Therefore, local discrepancies between the GLORIA and model cross sections due to remaining effects by horizontal gradients along the line of sight cannot be excluded.*

*However, when the complete ensemble of GLORIA and model data points is analysed, such remaining effects by horizontal gradients are expected to cancel out on average due to the large amount of data points. Therefore, we consider the estimation of model biases in Sect. 4.4 to be robust."*

**On page 16 lines 27 to 30 you find a mismatch and argue, that this can be explained by 'line of sight effects in the GLORIA observations'. But at page 10, line 1 - 8 you write line of sight effects are expected to cancel out due to the large amount of data (is this because you use a tomographic retrieval?).**

*We thank the Editor for pointing out the unclear discussion. On P10/L1–8 we meant to refer to the trace gas retrievals and the discussed correlations with the respective trace gases in the models. As mentioned above, improvements of local structures in the vertical cross section comparisons are possible, if 2-D averaging kernels were applied to the model results. However, when looking at the whole ensemble of data from the flight here, involving all kinds of viewing directions, we expect remaining line of sight effects to cancel out, when we analyse the model biases.*

*However, when looking at the distributions of clouds in the 2-D cross sections along the flight track, as described on page 16, line of sight effects could explain discrepancies between models and GLORIA observations. Note that, different from the trace gases, the cloud signal represented by the cloud index (CI) is directly inferred from the GLORIA spectral measurements, not by a retrieval procedure. While GLORIA measurements show an integrated cloud signal along the line of sight, the model data has been interpolated to the tangent point of the GLORIA measurements, which represents one single point in space. If, for instance, clouds were situated in front or behind the tangent point along the line of sight, this comparison would lead to a discrepancy between the model results and the measurements. In addition, the properties of clouds with their strong optical gradients (transparent/opaque) are considerably different from the trace gas distributions. Especially, complex small-scale cloud structures can differ in coverage and orientation from the trace gas fields. Therefore, we consider the comparison of GLORIA cloud detection to the simulated clouds to be more difficult, in particular for small clouds or edges of clouds. Despite these limitations of the comparison, we mostly found good agreement between GLORIA and the models.*

*For clarification, we added after P16/L27:*

*"Note that line of sight related effects are capable of particularly strong influences on the comparison with respect to clouds. If, for instance, clouds were situated in front or behind the tangent point along the line of sight, this comparison would lead to a discrepancy between the model results and the measurements. Especially, complex small-scale cloud structures with strong optical gradients (transparent/opaque) can differ in coverage and orientation when compared to the trace gas fields. Therefore, we consider the comparison of GLORIA cloud detection to the simulated clouds to be more*

*difficult, in particular for small clouds or edges of clouds. Despite these limitations of the comparison, we mostly found good agreement between GLORIA and the models."*

**On page 19 line 23 - 32 you describe biases around the tropopause. Especially in regions with large gradients AVK effects may cause under- and over-estimations similar to the describes effects.**

**In my view the AVK issue leaves a gap in your paper, because it is not clear, if discrepancies in some cases are due to the neglect of resolution effects, which would disappear if AVK's would be used.**

*See discussion above: In the vertical domain, the vertical resolution of the GLORIA retrieval results is comparable with the vertical model resolution and the absence of relevant overall systematic biases in the GLORIA data is confirmed by in situ comparisons. In the horizontal domain, the use of 2-D averaging kernels is capable to improve the comparison in the presence of strong horizontal gradients along the line of sight. However, since such gradients were mostly avoided during the discussed flight and since the application is demanding, 2-D averaging kernels were not applied here. However, when the complete ensemble of GLORIA and model data points is analysed, such remaining effects by horizontal gradients are expected to cancel out on average due to the large amount of data points. Therefore, we consider the estimation of model biases in Sect. 4.4 to be robust.*

**Some language issues. I am not an expert in English, so I am not entirely sure, if my corrections are entirely correct. I also may have missed other errors.**

**Page 11 line 9**
**accompanied by with a notably positive NAO -> accompanied by a notably positive NAO**

*Done*

**Page 26 line 16**
**than in the case of _THE_ ICON-ART forecast**

*Done*

**Page 28 line 5**
**in significant low biases -> in significantly lower biases**

*For clarification, we rephrased to "in significantly lower mixing ratios"*

**Page 29 line 6**
**by tuning of this scheme -> by tuning this scheme**

*Done*

***References***

*Woiwode, W., Dörnbrack, A., Bramberger, M., Friedl-Vallon, F., Haenel, F., Höpfner, M., Johansson, S., Kretschmer, E., Krisch, I., Latzko, T., Oelhaf, H., Orphal, J., Preusse, P., Sinnhuber, B.-M., and Ungermann, J.: Mesoscale fine structure of a tropopause fold over mountains, Atmos. Chem. Phys., 18, 15643–15667, https://doi.org/10.5194/acp-18-15643-2018, 2018.*