We thank the reviewers for their helpful comments. We expanded the description of the instrument and the regularisation scheme and clarified several misleading sentences. We applied all minor technical corrections unless specifically noted and do not list them in the referee comment recapture below. Major textual changes have been marked in green in both the revised manuscript as in excerpts below.

1 Reply to Referee #1

1.1 General Comments

> General Comments: This paper describes the retrieval of temperature and H2O, O3 and HNO3 from limb radiances measured by the GLORIA instrument on the German HALO aircraft during 2 campaigns in 2012. The paper is generally well written, and de- scribes the GLORIA instrument (briefly), and the retrieval process. The error analysis is very illuminating. Comparisons of the results from flights during 2 campaigns with the accompanying in situ measurements serve as validation of the results. Some of the descriptions and explanations should be expanded, so that the paper can be read as a stand-alone contribution. GLORIA and its 1-D and 3-D data analysis show promise of providing very useful data for UTLS studies.

1.2 Specific Comments

1. Abstract, l. 12- From capitalization, shouldnt acronym be BAHAMASS?
   l. 14- FAIRO written as acronym, without explanation

We changed the capitalization of one S to match the acronym. FAIRO is according to its users no acronym but a name. We thus changed FAIRO to Fairo.

2. Page 4: line 2- add references to Nakamura (1996) and Haynes and Shuckburgh (2000).
   Line 9- add reference to (Gille et al. JGR, 2014), Line 14- add reference (Peevey et al., JGR 2014)

We followed the suggestions.

3. Page 5, ll5-6- Someplace the paper should show or discuss how different the 1-D and 3-D results are.

The focus of this paper is the description and validation of 1-D retrievals. 3-D retrievals are more complex as they pose higher demands on the consistency of the calibrated spectra.

As such, the 3-D retrievals are currently being improved upon and a direct comparison of results for the discussed flight of 2014-09-26 is in preparation. The paper at hand mentions the robustness of 3-D retrievals against horizontal gradients in the

Conclusion section, which compensates the sole major weakness of the limb sounding technique.

4. GLORIA Instrument- More detail is needed, even in this short overview. What are the number of pixels in the horizontal and vertical dimensions, What is the composition of the detector array, and its temperature. How precisely can the movements of the A/C be compensated by the gimbal mounting? Does this define the pointing angle precision, or are there additional components from the angle measurements? How closely spaced along track are the measured profiles? Is this the spacing of the profiles shown later in the paper? What is the range of GLORIA azimuth angles?

The textual instrument description was enhanced to "The GLORIA instrument is a Fourier-Transform-Spectrometer (FTS) with a HgCdTe infrared image detector array (cooled to an operating temperature of 50 K) allowing to take up to 16 384 spectra simultaneously. To reduce the read out time, only 6144 of these are currently used. The usable spectral coverage ranges from approximately 780 to 1400 cm^{-1} while the spectral sampling can be adjusted quite freely (see Tab. 2)."

In addition a table was added that summarises the key instrument characteristics relevant for this paper (observation geometry, spectral sampling, etc.).

To answer the question also here: It is a HgCdTe detector array cooled down to 50 K. The gimbal mount stabilises the vertical view within 0.012° (1σ precision). The accuracy is currently not fully characterised, but seems to be better than 0.1°. We work additionally on a way to further increase the precision in L0 processing for the chemistry mode, which is more susceptible due to its longer measurement time (Latzko and Graf, 2015). The spacing along track would be 0.5/3.2 km in dynamics/chemistry mode, respectively. However, in dynamics mode, the instrument swivels so that the same azimuth angle is measured less frequently. The results shown later are reduced to use only measurements pointing at about 89° and employing the forward direction of the interferogram sled. So far we have not identified any problems with the spectra using the other sled direction, but their processing requires a significant amount of additional calibration effort, which was not yet spent. Depending on the measurement mode, this results in a spatial resolution of about ≈15 km on average. The azimuth can be tuned from 45° to 135°, whereby the full 90° cannot be exploited on HALO due to an obstruction by the wing at the last couple of degrees. In dynamics mode the yaw angles are swiped through in 4° steps and it is intended to tomographically process these sections. In this paper only the dynamics mode measurements pointing to 89°are employed, which are far fewer.

5. Page 7; line 9- please explain what a genetic algorithm is. Also artifacts.

The paragraph was extended to

,,The listed integrated spectral windows (ISW) used for the retrievals were selected by a genetic algorithm, which identifies the location and width of ISWs that maximises the information gain. The algorithm recombines initially randomly selected sets of ISWs, preferring ”good” sets and thus identifies a (nearly) optimal set much faster than a simple brute force search. Details are given by (Blank, 2013). The resulting windows were then modified to mitigate discovered instrument artefacts such as imperfectly compensated emissions of the outer window due to fast temperature changes.”

6. Tables 1 and 3 mentioned, no mention of Table 2.

An appropriate reference was added to the section on regularisation.

7. Why necessary to extend O3 and HNO3 to 60 km, so high above A/C.

It is not strictly necessary to extend that high. However, *some* extension is needed to allow to compensate for errors in the a priori background profiles. Picking an altitude well above the ozone and nitric acid maxima is fully sufficient for this, even if a lower altitude might have sufficed, too. With CRISTA-NF measurements, we have even made good experience with deriving the shape of the profiles above the plane with ≈1 DOF from upwards pointing measurements. While the state of GLORIA calibration does not allow for this currently, we plan on doing this in the future.

8. Line 23 ff: The explanation of the regularization is not nearly as clear in the authors earlier papers. It should be made clearer with an equation. Why not follow Rodgers (2000), as is much more usual?

We expanded the section to fully cover the details: “The precision matrix $\mathbf{S}_\mathrm{a}^{-1}$ is defined as

$$\mathbf{S}_\mathrm{a}^{-1} = (\alpha_0)^2\mathbf{L}_0^\mathrm{T}\mathbf{L}_0 + (\alpha_1)^2\mathbf{L}_1^\mathrm{T}\mathbf{L}_1 + (\alpha_2)^2\mathbf{L}_2^\mathrm{T}\mathbf{L}_2, \tag{1}$$

with $\alpha_0, \alpha_1, \alpha_2 \in \mathbb{R}$ and $\mathbf{L}_0, \mathbf{L}_1, \mathbf{L}_2 \in \mathbb{R}^{n\times n}$. The constraint can be separated into one constraint on the absolute value of retrieved target compared to a (climatological) mean weighted with its standard deviation and two smoothness criteria. The matrix $\mathbf{L}_0$ thus consists of a diagonal containing the reciprocal values of the standard deviations of the retrieved entities. The matrix $\mathbf{L}_1$ is a matrix to compute the first derivative of the vector $\boldsymbol{x_i}$ by finite differences (it has -1 is on the main diagonal and 1 is on the first upper side diagonal, except for some rows that would take the difference of different quantities or non-neighbouring values). In addition each row of $\mathbf{L}_1$ is scaled with the reciprocal of the standard deviation and $\sqrt{c_\mathrm{q}/(2h_i)}$, with c_q being a quantity q specific correlation length and h_i being the vertical distance between the elements of the vector that are being subtracted from each other (Steck and von Clarmann, 2001). Tab. 3 lists the empirically derived correlations lengths. $\mathbf{L}_2$ is set up similarly to $\mathbf{L}_1$ but with finite differences approximating the second derivative instead of the first. The sources for a priori values, background values and standard deviations are listed in Tab. 4.”

In effect, we do follow Rodgers with some exceptions in notation to make the non-linear nature of the retrieval more apparent (e.g. we refrain from using K_i for the Jacobi-matrix of the forward model for the i-th iteration and use the more unambiguous $\mathbf{F}'(\mathbf{x}_i)$ instead.).

However, due to our focus on large-scale retrievals, we focus heavily on directly constructing the precision matrix instead of the covariance matrix — otherwise we would need to start our setup from scratch when progressing towards tomographic retrievals. It thus resembles the regularisation employed by MLS, which is also the sum of a statistical and a smoothness term (e.g. Livesey et al., 2006). This requirement prevents our use of some popular choices of setting up the covariance in an empirical way. Secondly, the optimal estimation approach is more useful if appropriate true covariance matrices derived from in situ measurements are available or if the remote sensing measurements can contribute only few degrees of measurements — neither is true for our retrievals.

9. Page 8; l. 7; vertical correlation lengths mentioned here, but no mention of Table 2.

We added a reference to the table.

10. Page 12; l.10: Define F'

We added the following text: ,,where $\mathbf{F}'(\boldsymbol{x}_f)$ is the first derivative (Jacobian matrix) of the forward model $\mathbf{F}$ evaluated at the retrieval result $\boldsymbol{x}_f$."

11. l. 24: Is noise reduced as square root of number of spectral samples, or linearly?

The text has been clarified to note that the square root of the number of involved spectral samples is used. The relative noise components is independent of the amount of samples used.

12. Page 16; l. 24: Can you say more about the problem with the 792 Q branch?
page 17; l. 15: Is it clear that the problem is with the measurement of the Q branch, as opposed to modeling its contribution?

As stated in the paper, the artefact is still undergoing investigation and time-consuming laboratory characterisation. It expresses itself most strongly in the vicinity of strong spectral features and may be caused by the read-out circuitry of the detector chip. A second spectral region, where the effect can be observed is the vicinity of the methane peak close to $1300\,\mathrm{cm}^{-1}$.

Obviously, the CO2 Q-branch is notouriously difficult to model due to, e.g., the imperfect modelling of line-mixing. However, simply plotting the measured radiances around $792\,\mathrm{cm}^{-1}$ spectrally and spatially shows that there is coloured ,,noise" much stronger than expected present. Summarising, we expect a larger disagreement between model and measurements at the Q-branch than in other spectral regions (e.g. from experience with CRISTA-NF), but not to the extent observed in GLORIA observations.

13. l. 20: temperature As biases towards instrument location, but arent along track data being used to correct for this?

In principle, yes. We added: ,,However, we expect that the effect is mitigated by the application of ECMWF temperature gradients."

14. l. 21: In Table 4 the here is a consistent negative bias for the temperatures for the three flights, which appears to contradict the sentence that ...the mean difference is..but of opposite sign, indicating no consistent systematic problem. Please clarify.

The statement was indeed misplaced and belonged into the water vapour section. temperature seems to have a low bias as spelled out in the relevant section.

15. Page 19; l. 18: FAIRO is written as an acronym- please explain what it means.

According to its maintainers FAIRO is no acronym but a name. We thus changed FAIRO to Fairo.

16. Page 23; l. 23: The explanation of different air masses should be supported by ECMWF or other data- it doesnt have to be shown, but to dismiss all systematic differences as due to different air masses is not very convincing.

We did not state that all differences stem from different airmasses. The conclusions states that the discrepancies can be partially explained by the different geometries and a time-lag.

However, when looking closely at those discrepancies, we so far always found a plausible explanation, if the two stratospheric tracers (O3 and HNO3) both disagree consistently. For example, between 8:30 and 9:00 on 2012-09-26, GLORIA measures lower stratospheric tracers than the in situ instruments. In this case, according to retrieval results, ECMWF PV and CLaMS O3 modelling, there is a large airmass of low-PV, low-O3, low-HNO3 below the flightpath. In this case, the vertical FOV and averaging due to regularization (and partly probably horizontal gradients - this is difficult to quantify) lower the retrieved values compared to the in situ instruments. Other differences, like the drop in HNO3 at 11:30 on the same flight could meanwhile be attributed to a shift in instrument offset due to a rapid warming of the window by direct sunlight.

17. Figure 1: It is hard to distinguish the yellow and green. I could not find the purple tangent points. Are these the same as the blue lines?

The flightpath lines cannot be made thicker without generating too much overlap over Europe. Purple may be a misnomer, the colour seems to be actually called something like “medium slate blue”. The textual description will be changed to “blue”. The blue points are rather large to be visible, so they often run together and look line lines. Zooming into the PDF allows to distinguish them at least at higher latitudes.

18. Figure 2: Again it is hard to distinguish the green and yellow segments-can the lines be made thicker?

Here, the available space allows for a thickening. See revised version.

19. Figure 3: Please explain what the red lines are.

We added ,,The vertical red lines separate regions that employ different aerosol/extinction profiles.”.

2 Reply to Referee #2

2.1 General Comments

This paper describes a 1-D retrieval method for measurements taken by the GLORIA infrared limb-imager during two validation campaigns in 2012. GLORIA is a novel instrument; the data processing for it is in its infancy and not much has previously been published about it. The authors are well known in their specialist field of work. They are affiliated to the institutes that have developed, built and deployed the GLORIA instrument, the data processing of which is the subject of this work. They also have a proven track record of publishing atmospheric measurement from remote sensing instruments, as well as theoretical work on retrieval methods. The authors consortium is therefore well placed to have conducted this work.

I expect a publication of a retrieval scheme to contain a theoretical description of the algorithm used, as well as an encyclopaedia of input parameters (measurement data and errors/problems therewith, prior information, approximations used and their impact on the result, correlations, etc.). There must also be a comprehensive section on the validation of the results, wherever possible. The manuscript as such addresses all these topics, with the exception of the issues raised in the section Specific Comments. If these are addressed satisfactorily, I recommend the manuscript to be published in AMT.

2.2 Specific Comments

1. Page 12040, line 27: The difference between dynamics and chemistry mode could be explained in more detail: I.e. what is the extent of changes to the spectral and spatial resolutions from one mode to the other? Also, how are

the different modes implemented at instrument level? The details of this presumably affect the data processing.

The section on the GLORIA instrument has been expanded with a lot of technical details including also a table with instrument specifics.

There is not much difference between dynamics mode and chemistry mode from the processing side up to L1. The optical path employed for the acquisition of the interferogram is ten times longer, which implies also a roughly ten times as long measurement time and as much more data to process.

The major differences come up only in level 2 processing, where currently slightly different philosophies are employed (retrieve many targets at once with ISWs covering the whole spectrum in contrast to zooming in on a single or a small set of lines for one target).

2. Page 12041, line 19: the used configuration for the GLORIA data processing. Is there a version number to help identify this used configuration in future references? If not I think there should be one.

We added a V1.00 designation to the described configuration.

3. Page 12043, line 17: It is said that Table 1 describes what the optical properties of the aerosol extinction coefficients are, i.e. it would be interesting to know what the prior information for aerosol retrieval was. However the table just lists an aerosol index, which I presume is a placeholder for an unspecified set of aerosol parameters?

Table 1 simply specifies, which ISWs share a common aerosol/extinction profile. ISWs with a common index share this aerosol/extinction profile in the forward simulation. Each of the five aerosol/extinction profiles has an identical a priori distribution that decreases exponentially with increasing altitude and no spectral shape (aside the one implied by the Planck curve).

4. Page 12068, Table 2: The vertical correlations lengths seem quite large. What provision has been taken to ensure that the 5km correlation length for water vapour doesnt affect the retrieved Tropopause altitude?

The correlation lengths can be related to correlation length in an auto-regressive model (see Rodgers (2000)), but they are, in addition, fudge factors to regulate the regularisation strength as deemed necessary from retrieval results. Further, changing them by a factor of 2 changes the result only in a very small way. For example, ozone has such a high correlation factor as there are so many ISWs with a high signal content that otherwise the smoothing had no effect and the retrieval would overfit the data (which in light of some known systematic instrument errors and a still ongoing instrument characterisation would not be too good).

With respect to the correlation length for water vapour one needs to keep in mind that the log of water vapour is being retrieved so that very strong gradients are feasible even with high correlation lengths. Also, water vapour is not scaled with a climatologically derived standard deviation (or effectively a SD of 1 is assumed), so interpreting the configured length is a bit up for discussion.

So far, we have not found that changes in the parametrisation of water vapour affect temperature (more the other way round). Ozone has an influence on temperature due to the use of optically thin and thick ISWs around 990 cm^{-1}.

With respect to the chemical tropopause, vertical smoothing obviously has an effect and this was the major motivation for the change from linear-H2O to log-H2O retrieval. As indicated by Fig. 5, the vertical resolution is typically below 500 m for trace gases, which gives an indication of how the employed correlation length translates into resolvability.

Currently, our thermal and chemical tropopause seem to fit very well to each other and the ExTL seems to have the expected thickness so we are not aware of any adverse side effects of the chosen regularisation.

5. Page 12048, lines 8ff: An error estimate for the fast forward models is given by comparing the band model with the more accurate monochromatic model, and then both of the fast models with the more detailed RFM. However, the RFM explicitly uses the same ray tracing, so how are errors from the ray tracing estimated? This is exuberated by the fact that the band model is used as a priori for the monochromatic forward model in the retrieval.

The L2 result of the band model is employed as initial guess for the retrieval with the monochromatic forward model, not as a priori.

The influence of raytracing, which includes both refraction and mapping of the discrete representation of the atmosphere onto a number of homogeneous cells has not been examined in this study. The raytracing employed by JURASSIC samples the atmosphere along the line-of-sight in regular intervals and assumes that the cell surrounding the sample point is homogeneous. In the limit of an infinite amount of samples, this introduces no representation error. Studies by Hoffmann (2006) showed that a sampling distance of 2 km introduces an error of 0.025% with a standard deviation of less than 0.1% compared to a calculation with a sampling distance of 100 m (smaller step lengths did not result in meaningful differences) Thus we assume that the uncertainties introduced by the band approximation dominate the uncertainties introduced by the raytracing. The refraction scheme employed by JURASSIC has been described by (Hase and Höpfner, 1999) and seems to be state-of-the-art.

The RFM follows a different approach by defining horizontal layers (which are finely samples and averaged using Curtis-Godson means), for which the monochromatic optical path values are calculated in a first step. In a second one, the length of the rays intersecting each layer is determined and the optical path values are then mapped. This approach does not lead inherently to a fine representation close to the tangent point as the method JURASSIC employs, but is more efficient as multiple rays can use the optical path values precalculated for one layer. Especially for optically dense spectral samples, a fine spatial sampling is important. It is thus expected that differences occur. Quantifying this is beyond the scope of this paper.

However, the largest discrepancies in studies comparing JURASSIC to RFM in the past were caused by the different vertical interpolation of aerosol extinction (log-linear compared to linear), which had an effect in the mean of up to 0.5% close to atmospheric windows.

6. Page 12049, line 20: The characterisation of actual noise figures is still in progress. The measurement noise figure is of central importance to the retrieval algorithm. Its reasonable to expect than uncertainties in its knowledge will have a major impact on the results. The authors claim that they have evidence that the estimates they are using are accurate enough. It would strengthen their case if they could quantify this statement.

We have quantified the thermal/white noise to an extent that gives us confidence in the quality of our results but not to be already fully publishable as all differences between different way to deduce the noise are not yet accounted for. For example, Friedl-Vallon et al. (2014) gives a noise figure derived from black bodies, which is very similar to the one employed. We currently use a noise estimate derived from the variation of the imaginary part of atmospheric spectra to also partly account for coloured noise such as imperfections of the calibration. This overestimate causes our chi^2 to be in the order of 0.8 on average (as some effects of an imperfect calibration may be corrected for by the extinction retrieval).

Numerical experiments involving small changes of the employed noise figures (+-50%) did not show meaningfully different results except for a notable effect on the vertical resolution, which could be compensated by corresponding changes to the assumed correlation lengths.

7. Page 12055, line 12: The correlation of O3 with HNO3 and anti-correlation with H2O reflects the distinction between stratospheric air (dry, O3 rich) vs. tropospheric air. This is worth pointing out in the text since its an important self-validation of the retrieval! In fact, the actual discussion of the scientific findings is quite marginal - a mere couple of lines. Surely this could be extended.

This may indeed not be obvious to all interested readers and deserves spelling out. We added This is expected due to the typical chemical composition of stratospheric air (dry, O_3 and HNO_3 rich air) and tropospheric air (wet and deprived of O_3 and HNO_3) and makes the observed filamentary structure plausible. From the given figures, one can directly identify air masses, which were recently mixed from the troposphere into the stratosphere like the filament of comparatively wet air at 12 km around 10:00 UTC.

A detailed discussion of the observed situation certainly deserves more room than can be found in this paper that focuses on the processing side of things.

8. Page 12058, line 9: It is expected that the vertical resolution of temperature can be further improved when the instrument artefacts around the CO2 Q-branch have been resolved. This, together with the statements that not all of the campaign spectra have been processed and that the Level1 processing hasnt reached a final version, is the main issue I have with the current manuscript. For a work that aims to become the canonical reference for future scientific publications of GLORIA campaign data, I would have expected it to be based on the comprehensive set of measurement data. Conclusions generally stand on wobbly ground if the input data lacks the seal of approval. At the very least I would like to see a solid case being made to corroborate that whatever instrumental effects are possibly to be identified from i.e. the CO2 Q-branch or from any of the missing scans for that matter will not require significant alterations to the retrieval processor as it is described in this work.

We added to the paper that the instrument problems at the Q-branch introduce spatially and spectrally correlated coloured noise. This requires an increase of the vertical smoothing to compensate, which causes the comparatively bad vertical resolution of temperature of 1 km (as mentioned in Sect.) instead of also ≈0.5 km that we get from synthetic measurements and a reduced vertical smoothing.

Please note that we did not need to exclude the affected spectral regions from the L2-processing for the presented flights. The only foreseen change for the processing of future campaigns (or the current ones, in case we identify a way to correct for the artefact in a preprocessing step) is a reduction of the c_{temp} value for the vertical correlation length.

Thus, we expect the current L2 setup to be stable as much as such things may be foreseen.

9. Page 12074, Figure 4 (and similarly Figure 5): There is a periodic structure (oscillations) at the top altitudes of the error profiles for offset and spectroscopic parameters for CO2. I wonder what the reason for this is? Its striking for O3 and HNO3, not so for Temperature and H2O (probably masked by the different scales).

The effect of some of the systematic errors in individual profiles increases rapidly above the flight altitude. As the airplane flew on effectively three distinct levels, these average together to form two visible maxima (the third was above 15 km). Previously, we included data up to 250 m above the flight altitude in the averaging. We prepared new figures that discard all information above the flight altitude and the peaks in the error plots all but disappeared (see Fig. 1 in this document).

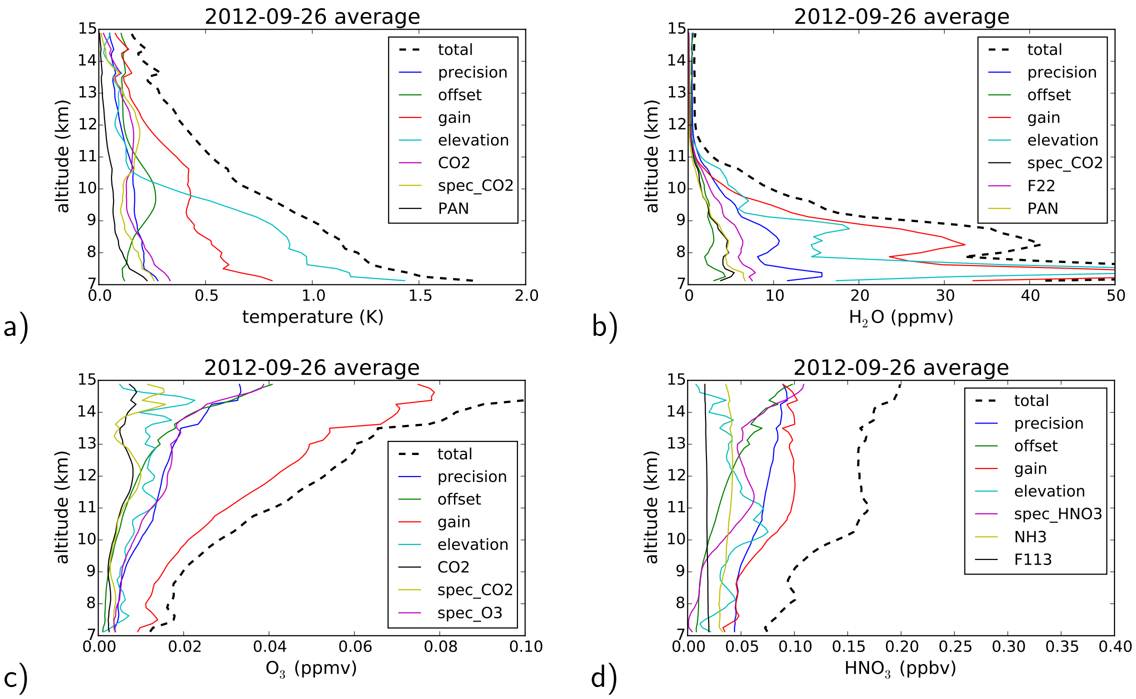

Figure 1: Total error and major error sources for the four primary targets temperature (panel **a**), H_2O (panel **b**), O_3 (panel **c**), and HNO_3 (panel **d**) averaged over the profiles of the flight of 26 September 2012. A "spec" prefix notes the error induced by spectral uncertainty of line intensities.

For the resolution, some peaks remain, which is expected as the resolution increases rather rapidly with decreasing altitude and becomes best at flight altitude. Thus, there are minima in the resolution at the three typical flight levels. We thinks that the remaining discontinuities in the error plots have the same underlying cause (minimal values for errors at flight altitudes due to the high measurement density).

2.3 Technical Corrections

10. Header: The secondary affiliation mark for author T. Guggenmoser for his current po- sition at ESETC (*) is barely legible. Given that an affiliation entry for ESTEC already exists (3) why not just re-use the latter.

The "3" was in error as the co-author was at IEK during the work on this paper. It has been corrected. Further, we followed the suggestion to replace the asterisk with a regular number.

11. Page 12039, line 1: Could explain what the word gimballed means. Its instrumental to the instrument concept, yet its quite an exotic term and as such not universally understood.

We added “To operate on an aircraft, it is placed in a gimbal (a cardanic frame), which is used to stabilise the pointing against movements of the carrier and to adjust the viewing direction.” to the instrument section.

12. Page 12039, line 9: What are the different modes of operation, and how is dynamics mode different from chemistry mode?

We added the spectral sampling of dynamics mode spectra to the abstract. The distinction between the two major operating modes can be found in the enhanced Section 2 including a table listing key instrument characteristics.

13. Page 12039, line 15: What is FAIRO?

Fairo is a name and not an acronym. Correspondingly, it is written now as ”Fairo”, not ”FAIRO”.

14. Page 12059, line 25: q_{log} is later called q_{H2O} in the next formula. This is slightly confusing. How about the following notation q_{log}^{H2O} and q_{vmr}^{H2O}, or using italic variables for log-space and roman variables for vmr-space?

The H_2O suffix was given in error, it should have been log (it obviously also represents water vapour, but the paragraph deals with the conversion between log space and VMR space, which is generally applicable).

15. Page 12073, Figure 3. The caption to Figure 3 could be improved. What are the red vertical lines? Half of the time these seem to correlate with singularities in the residuals; what is the reason for the latter? Also, the text mentions a instrumental effect at the CO2 line at 792cm-1. I cant recognise this in the figure, and the residuals are within their boundaries, which would imply that the effect is present in the simulations too?

The vertical lines indicate the transition from one aerosol/extinction profile to another, causing a discontinuity in the simulated spectra. This has been noted in the caption in the revised version.

The instrumental effect at 792 wavenumbers causes coloured noise beyond the original instrument specification. Please note that most crosses in this spectral region are outside the target noise region and some even outside the threshold noise region. A strong constraint on smoothness in temperature is currently used for mitigation.

16. Page 12075, Figure 5 and Page 12076, Figure 6: Sub-panels (a) (c) not attributed in caption.

We added the description.

17. Page 12078, Figure 8; Page 12079, Figure 9; Page 12080, Figure 10: Sub-panels not (a), (b) not attributed in caption.

Figures 8, 9, and 10 refer back to Figure 7, which gives the full explanation on the displayed data.

References

Blank, J.: Tomographic Retrieval of Atmospheric Trace Gases Observed by GLORIA, Forschungszentrum Jülich, Jülich, Ph.D. thesis, Wuppertal University, 2013.

Fahey, D. W., Gao, R.-S., Möhler, O., Saathoff, H., Schiller, C., Ebert, V., Krämer, M., Peter, T., Amarouche, N., Avallone, L. M., Bauer, R., Bozóki, Z., Christensen, L. E., Davis, S. M., Durry, G., Dyroff, C., Herman, R. L., Hunsmann, S., Khaykin, S. M., Mackrodt, P., Meyer, J., Smith, J. B., Spelten, N., Troy, R. F., Vömel, H., Wagner, S., and Wienhold, F. G.: The AquaVIT-1 intercomparison of atmospheric water vapor measurement techniques, Atmos. Meas. Tech., 7, 3177–3213, doi:10.5194/amt-7-3177-2014, 2014.

Friedl-Vallon, F., Gulde, T., Hase, F., Kleinert, A., Kulessa, T., Maucher, G., Neubert, T., Olschewski, F., Piesch, C., Preusse, P., Rongen, H., Sartorius, C., Schneider, H., Schönfeld, A., Tan, V., Bayer, N., Blank, J., Dapp, R., Ebersoldt, A., Fischer, H., Graf, F., Guggenmoser, T., Höpfner, M., Kaufmann, M., Kretschmer, E., Latzko, T., Nordmeyer, H., Oelhaf, H., Orphal, J., Riese, M., Schardt, G., Schillings, J., Sha, M. K., Suminska-Ebersoldt, O., and Ungermann, J.: Instrument concept of the imaging Fourier transform spectrometer GLORIA, Atmos. Meas. Tech., 7, 3565–3577, doi:10.5194/amt-7-3565-2014, 2014.

Gille, J., Karol, S., Kinnison, D., Lamarque, J.-F., and Yudin, V.: The role of midlatitude mixing barriers in creating the annual variation of total ozone in high northern latitudes, J. Geophys. Res., 119, 9578–9595, doi:10.1002/2013JD021416, 2014.

Hase, F. and Höpfner, M.: Atmospheric ray path modeling for radiative transfer algorithms, Appl. Optics, 38, 3129–3133, doi:10.1364/AO.38.003129, 1999.

Haynes, P. and Shuckburgh, E.: Effective diffusivity as a diagnostic of atmospheric transport: 1. Stratosphere, J. Geophys. Res., 105, 22 777–22 794, doi:10.1029/2000JD900093, 2000.

Hoffmann, L.: Schnelle Spurengasretrieval für das Satellitenexperiment Envisat MIPAS, Tech. Rep. JUEL-4207, Forschungszentrum Jülich, Jülich, Germany, ISSN 0944-2952, 2006.

Latzko, T. and Graf, F.: Evaluation of the pointing accuracy of a limb sounder using cloud edge detection, in: Fourier Transform Spectroscopy and Hyperspectral Imaging and Sounding of the Environment, p. FW3A.5, Optical Society of America, doi:10.1364/FTS.2015.FW3A.5, 2015.

Livesey, N., Van Snyder, W., Read, W., and Wagner, P.: Retrieval algorithms for the EOS Microwave limb sounder (MLS), IEEE T. Geosci. Remote. Sens., 44, 1144 – 1155, doi:10.1109/TGRS.2006.872327, 2006.

Nakamura, N.: Two-dimensional mixing, edge formation, and permeability diagnosed in area coordinates, J. Atmos. Sci., 53, 1524–1537, 1996.

Peevey, T. R., Gille, J. C., Homeyer, C. R., and Manney, G. L.: The double tropopause and its dynamical relationship to the tropopause inversion layer in storm track regions, J. Geophys. Res., 119, doi:10.1002/2014JD021808, 2014.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, vol. 2 of *Series on Atmospheric, Oceanic and Planetary Physics*, World Scientific, Singapore, 2000.

Steck, T. and von Clarmann, T.: Constrained profile retrieval applied to the observation mode of the Michelson Interferometer for Passive Atmospheric Sounding, Appl. Optics, 40, 3559–3571, doi:10.1364/AO.40.003559, 2001.

Voigt, C., Schumann, U., Jurkat, T., Schäuble, D., Schlager, H., Petzold, A., Gayet, J.-F., Krämer, M., Schneider, J., Borrmann, S., Schmale, J., Jessberger, P., Hamburger, T., Lichtenstern, M., Scheibe, M., Gourbeyre, C., Meyer, J., Kübbeler, M., Frey, W., Kalesse, H., Butler, T., Lawrence, M. G., Holzäpfel, F., Arnold, F., Wendisch, M., Döpelheuer, A., Gottschaldt, K., Baumann, R., Zöger, M., Sölch, I., Rautenhaus, M., and Dörnbrack, A.: In-situ observations of young contrails - overview and selected results from the CONCERT campaign, Atmos. Chem. Phys., 10, 9039–9056, doi:10.5194/acp-10-9039-2010, 2010.