Authors' response to referee comments for 'New calibration noise suppression techniques for the GLORIA limb imager'

T. Guggenmoser et al.

June 23, 2015

We thank both referees for their comments, which we will answer in this document. For better legibility, referee comments are marked up in orange, and proposed changes to the manuscript – where necessary – in red.

1 To anonymous referee no. 1

1.1 General comments

A principal component noise filter is applied to the calibration measurements of the GLORIA limb imager. This is a very interesting application, which is nicely introduced in the paper. However, the details of the proposed method are somewhat ad hoc and, in my opinion, not always properly justified. Furthermore, some information (for example the typical temperatures of the two black bodies or how many calibration measure- ments are obtained in each calibration sequence), which could be useful for a better understanding of the subject are omitted.

It is true that some details – e.g. the function used to represent the NCO shape – have been found empirically, which is also stated in the paper. Note that the techniques presented in the paper reflect our current understanding of instrument effects.

We will add a passage to the section on radiometric calibration with some contextual information (please see the reply to first specific comment).

1.2 Specific comments

Three different calibrations are achieved by using two of the three (DS, BBC and BBH) calibrations measurements to determine the calibration coefficients. It seems natural to use all three calibration measurements simultaneously to derive a single calibration using

a line fitting the three points best in a least squares sense. There might be reasons why this is not done, but as long as the reader is not told, he might wonder why not.

This approach would indeed seem natural and is a future goal, but currently not possible due to several issues. Most importantly, the temperature difference between the two blackbody sources during our measurement campaigns was smaller than originally expected, which led us to mostly use the BB-DS two-point method for calibration. This is discussed in more detail in another publication (Kleinert et al., 2014).

We will add the following passage to the section in question:

During the first GLORIA measurement campaigns, a typical calibration sequence consisted of 20 measurements of each blackbody, followed by 10 DS measurements. These sequences were performed every 30 to 45 minutes and took about 5 minutes each. Only measurements with matching interferometer direction are processed together. Typical temperatures for the blackbodies were about 240 vs. 256 K, but values varied throughout the flights, with the difference between the two ranging from 15 to 25 K. These temperatures are both higher and less far apart than expected, which presents problems with error magnification for the BB-BB calibration. For this and other reasons, including an imperfect nonlinearity correction, the three calibration sources can currently not all be used together in a three-point fit (Kleinert et al., 2014).

Figure 1: Shows that the parallelization speed up of a single measurement is far from perfect, so the reader asks himself why bother when, supposedly, different measure- ments can be processed completely independently of each other.

We use both thread- and process-based parallelisation for the bulk processing of campaign measurements. The best balance between the two depends on hardware constraints such as total memory of each compute node, memory bandwidth, and mass storage bandwidth. A typical bulk run is performed using 10-20 processes with 4 threads each.

We also use the same software library for diagnostic applications, where the fast processing of single measurements is advantageous.

To make this clearer to the reader, we will add the following paragraph to the end of the section:

In practice, the processing is usually performed on a cluster, with multiple measurements being processed simultaneously. This provides another level of parallelisation which is subject to different constraints. The optimal balance between the number of processes and number of threads depends on several factors such as mass storage bandwidth and memory available per compute node.

Eq.(7): Was found by trial and error and does not always seem to be appropriate (P.12659,L6), a very natural and more generic alternative would be to use a low rank approximation determined by an SVD. Has this been tried?

Recent results suggest that the deficiencies are likely caused by emissions of the outer window. We tried to approximate the negative offset by a low-rank SVD, but found

no improvement over the current method. The low-rank decomposition cannot generate different hyperbolas as efficient as the formula and has the same problem in reproducing the small scale and highly temperature dependent changes imposed, e.g., by the thermal emissions of the outer window. In addition, it is difficult to acquire a sufficient amount of training data so that the low-rank approximation retains no trace of the discontinuities imposed by instrument artefacts.

However, we work on providing more sophisticated models of the negative offset, e.g. by properly characterizing the emissions by the window dependent on the measured temperature to improve upon the described state. As we believe that the pseudo-hyperbola reproduces the self-emissions of the instrument rather accurately with the exception of the window, this would improve the accuracy of the method considerably.

We update the paper to read:

The pseudo-hyperbola function cannot reproduce the NCO for all spectral samples, as shown in Fig. 4 for the one at $840\,cm^{-1}$ *. Here, the difference between the BB-DS NCO with and without the correction shows ring-like structures, which are likely related to emissions of the outer interferometer germanium window that peak around this wavenumber. However, the improvements to the BB-BB calibration over the rest of the spectrum outweigh the cost of these artefacts, which only persist for a limited number of samples. Ongoing characterisation of the window emissions may eventually allow the inclusion of the window emissions as dedicated terms in the fit formula.*

P.12659,L17: I am not sure that the size of the NCO correction can be compared with the size of recorded spectra, which apparently has been done, to arrive at the quoted 8%. The NCO errors arise from extrapolation of noise in the BB measurements, but surely the effect of this noise is lower at the temperature of the recorded scenes.

Recorded interferograms are fourier-transformed, multiplied with the inverse gain matrix and finally the NCO is added. Thus the magnitude of the NCO directly impacts the calibrated spectra. To emphasize the importance of the correction, we compared the magnitude of the change in NCO due to the hyperbola smoothing to the magnitude of the measured atmospheric scene. And in fact, the effect of applying the hyperbola correction changes the value of the calibrated spectra by 8%. This was easily estimated by calibrating the spectra once with and once without the hyperbola correction and computing the relative differences of the result. The BB are comparatively hot and for the previous campaigns, the temperature difference was rather small so that errors are magnified. The NCO is rather large as only the detector is cooled to 50K, while the spectrometer itself is cooled down only to 200–220 K. Therefore, the NCO is often of similar magnitude to the atmospheric spectra (smaller for clouds, larger for upwards pointing measurements).

To anticipate any surprise for the reader, we will add the following paragraph:

The size of this correction is attributable to two reasons. As mentioned above, the BB-BB calibration currently suffers from error magnification due to unexpectedly large and

similar BB temperatures. A more fundamental reason is that the NCO itself is relatively large because the GLORIA spectrometer operates at 200–220 K (Friedl-Vallon et al., 2014), and the radiative background can be comparable in magnitude to atmospheric emissions, depending on the observed scene.

Figure 2: [bbbb_nh] As discussed in the paper the noise originates from noise in the BB measurements. It is a major weakness of the paper that this plot has not been redone after the PC noise filtering of the BB measurements.

We actually used a calibration dataset for this plot which had already been subjected to an earlier version of our PC filter. As we neglected to mention this in the text, and the structure of the paper might suggest otherwise to the reader, we have decided to re-generate the dataset without any PC filtering and to update the plots accordingly (see Figs. 1–4 in this document). Because the colour scale in the new plots is different, new plots using the original data from the manuscript have also been generated for comparison (see Figs. 5–7 in this document).

The inhomogeneities in the NCO function which we remove with the pseudo-hyperbola fit are due to both noise and pixel nonlinearities. Both errors are magnified due to the small BB temperature difference, but outside of spectral windows which are dominated by noise, the nonlinearities have a much higher impact. The visible artefacts (particularly the "crosses") in the bbbb_nh plot are due to these nonlinearities, rather than random measurement noise, which is why the PC filtering leads to little visible change in this plot.

P.12660,L13: Replace "Contributions to the original data which are fully uncorrelated, such as white noise, are distributed evenly among these principal components," with "If the original data is noise normalised, such that the random noise becomes uncorre- lated and uniform (white), any direction will carry the same amount of noise,"

The proposed change will be made in the revised manuscript.

Figure 3: There is a distinct shape of the IG, (a bit simplified) being highest in the bottom left corner and decreasing towards the upper right corner. What is the explanation for this? Even if no explanation is known, this should be mentioned in the paper.

This is due to small differences in the chemical composition of the detector, which become especially apparent in this plot because the wavenumber shown is close to the sensitivity cut-on. We will update the caption to include:

The variation from the lower left to the upper right corner is caused by a composition change in the detector material.

Figure 5: The green curve is almost always to the right of the blue curve. This does not seem right - when bbbb_nh is fitted with a pseudo-hyperbola we do not expect to introduce a bias, also not when averaged over rows. This is due to our exclusion of known more non-linear pixels from the fit. Certain pixels, which are readily identified

with simple threshold filters in the invgain are subject to especially large errors caused by their slightly larger non-linearity. These can be seen as dark blue crosses in Fig 2, lower row bbbb_nh. These non-linearities seem to be often biased in the same direction, i.e. generate a larger NCO. For example the large peak around row 62 is caused by the cluster of pixels on the r.h.s. of the detector.

Believing these pixels to be more incorrect, we exclude them from the fit of the hyperbola correction to remove a bias.

P12661,L1: True, that a training set based on previous measurements is not necessary. But it would likely be better - for instance by increasing the number of spectra, the noise distribution would improve - and faster.

We agree with this comment in principle. However, as we apply our method to calibration measurements, training with previously recorded spectra is a different proposition than it would be for radiometrically calibrated atmospheric measurements. It is also our intention in this paper to highlight methods that specifically exploit the instrument's imaging capabilities - and we believe nonreliance on a training set to be a distinct feature. For these reasons, PCA with a training set is beyond the scope of this work, although it may well be worth revisiting in the future, particularly for the treatment of atmospheric measurements.

P.12662,L1-2: I am puzzled by this sentence and the long-winded steps of "normalisation" it leads to. My approach would be to apply the PC noise filter directly to X (with noise normalisation of course) and would expect that the detector differences are easily accounted for by a few PCs. I do not have any evidence for this, but the authors do also not present any evidence of the benefits of their "normalisation".

It is indeed possible to achieve a smoothing effect with a direct SVD as suggested by the referee. However, we have found the mathematical interpretation of such a procedure, as well as of its results, rather questionable, and therefore developed the procedure presented in the paper.

The radiometric response of the instrument varies significantly over the detector pixels both in magnitude and phase, with each pixel having its own gain profile. This is illustrated by Fig. 3b in the paper; as explained above, the pattern observed there in the IG magnitude is specific to a certain spectral region.

Because of this wavenumber-dependent variation in response, it is unclear to us how spectra from individual pixels are supposed to measure the same set of statistical variables. One might even argue that their only a priori relationship was to be from the same vector space.

The procedure outlined in this paper is an approximate gain calibration which yields radiance spectra with matching physical units. In our view, this makes it more justified to treat them as measurements of the same statistical variables. Without this, we would not perform a PCA, but simply a low-rank approximation of Fourier-transformed raw data.

A secondary benefit is that we can further reduce the overall variance by applying the NCO pseudo-hyperbola fit before the decomposition, which would not be possible without the gain calibration.

It is our impression that the referee objects to our use of the term "normalisation" in this context. We agree that it is a bit too unspecific, especially considering that noise normalisation is also part of the procedure, and decided to change it to "gain calibration" and "homogenisation".

We will also change the passage in question to read:

When PCA was introduced in Sect. 4.1, it was assumed that all pixels had the same response function, so that each pixel measured the same statistical variables. More practically speaking, the set of pixel spectra resembled a series of measurements taken by the same instrument. This statistical interpretation provides the notion of a covariance matrix, whose eigendecomposition PCA is based on. In order to be able to apply PCA for noise suppression, the data set $\mathbf{X}$ *therefore needs to be brought into a form which closely resembles this.*

To stay terminologically consistent, we also update the following paragraph to read

We solve this problem by the following process. First, a smoothing of the blackbody and deep space measurements is performed based on singular value decomposition (SVD) without gain homogenisation, using 20 singular vectors for reconstruction.

and change *initial PCA* to *initial SVD* at the end of the section.

P.12663: Yes, the noise can be estimated by PC techniques, but to get the full noise, the part which is retained in the first 400 PC scores must be added. To estimate this part of the noise a first guess of the full noise is needed (with possible iterations). In parallel to this I was wondering if if wouldnt be possible to derive noise estimates directly from the BB measurements?

It is not our intention to estimate the full noise magnitude; the purpose here is only to estimate its spectral distribution for the normalisation matrix. Our chief interest is in suppressing calibration artefacts – if we were to apply our method to atmospheric spectra, a more sophisticated noise analysis would be required.

A noise estimate can be derived from the BB measurements. However, this is difficult at this stage even with gain homogenisation due to the presence of radiometric background. Also, it would give an estimate only for the BB measurements, while the reconstruction residual can easily be computed for the BB and DS measurements.

To make this distinction more clear, we will add the following statement:

. . . contains the SD of the calculated difference on the main diagonal. Note that this is not an approximation of the total noise, but only of the spectral distribution for the purpose of rescaling. To arrive at the full noise, a more thorough analysis is necessary (Antonelli et al., 2004).

Figure 10: What is meant by "Normalised covariance eigenvalues"? In any case this plot seems to confirm that the noise normalisation is not correct. Not even consistent between DS and BB. I would expect the noise for the two to be different, thus needing different N matrices, but this is not discussed in the paper.

The "normalised covariance eigenvalues" shown in Figs. 10-11 are the covariance eigenvalues, i.e. the squares of the singular values associated with each PC, divided by the total sum of all these values. Fig. 10 shows the individual values, while Fig. 11 shows the cumulative sum, i.e. the relative amount of variance restored by a number of principal components (see Eq. 14 in the paper). Because the eigenvalues are for the whole covariance matrix, not just the noise covariance, the differences between the three curves are expected. The difference between the DS and BB curves, in particular, is discussed in the text, albeit briefly.

The noise matrices used are always computed from the reconstruction residual of the current reconstruction, i.e. indeed different matrices are used for BB and DS. To make this explicit, we will add the following statement right after the previous addendum:

The **N** *matrix is computed separately for DS and BB measurements.*

Figure 14: This plot is showing a single spectrum. Yes? So how can you speak about a bias here?

We agree that the term "bias" is misleading, and will change the phrasing to:

. . . , a visible difference appears between the PCA methods and the pure LPF method. This difference . . . .

P12666,L3: Really it is not the homogeneity (a simple average would be enough to capture this), but the spectral correlation.

We concur that the choice of term "homogeneity" was unfortunate, and will change the wording accordingly:

is easily reconstructed due to its high spectral correlation.

Figure 15: It is misleading to refer to the reconstruction residual as reconstruction error. Reconstruction error normally refers only to the fraction of the signal which is lost in the residual.

We have seen the term used in this fashion in the literature, but we agree that "residual" is less ambiguous, and will modify the paper accordingly.

1.3 Technical Corrections

Eq.(9): Isnt NCO missing here?

In this context, the true radiance image includes the instrument background (i.e. the

NCO), which is why it does not appear as a separate term in the equations.

Eq.(12): Shouldnt X''' be X'''^* to get the dimensions right?

Indeed, this is a mistake. We will change the equation to read:

$$Y = X''' N^{-1} \tag{1}$$

Note that the noise normalisation matrix is always real-valued; we will also change the part after the equation to better reflect this.

2 To anonymous referee no. 2

2.1 General Comments

General question: It is not clear why the same technique is not applied to the radiometric gain which, in Fig 3, also seems smooth. Is it because some genuine pixel-pixel variability is expected in gain unlike radiometric offset?

Yes. The main difference is that the NCO consists of thermal emissions of the instrument, which are by their nature smooth. The gain, on the other hand, depends chiefly on electronic characteristics of the individual pixels and analog-digital converters. Because of this, there is significant and discrete variability between the pixels. In Fig. 3 of the paper, the large-scale change – caused by variations in the detector material (see revised caption) – may seem smooth, but the local change from pixel to pixel is certainly not smooth.

It is suggested that the above can be applied to the calibration of any imaging spectrometer, however I suspect that the individual characteristics and quirks of each such instrument will require tailored solutions with only marginal reference to this work.

Some details, such as the best function to fit the NCO, are likely particular to GLORIA. However, the instrument radiative background can generally be expected to be smooth. An exploitable symmetry around the optical axis is not an overly specific requirement, either.

The principal component technique makes even fewer assumptions about the instrument and should be readily adaptable to other imaging sounders, as long as the number of pixels is high enough and the calibration scenes offer enough spectral correlation.

We do agree, however, that the original statement may sound a little too strong. We will update the relevant sentence to read:

We expect that the techniques presented in this paper can also be adapted to other instruments with a focal plane array and similar absolute radiance calibration.

2.2 Specific Comments

The proposed technical changes (plots, notation) will be applied in the revised manuscript.

Fig 3: unclear what "AU-1" means. Arbitrary Units? (in which case why have any units at all?). AU also appears for y-axis radiance in other figures.

"AU" indeed stands for "arbitrary units". In this paper, it is used for uncalibrated radiance values, which is why it appears in the inverse gain matrix and the reconstructed calibration measurements.

3 Figures

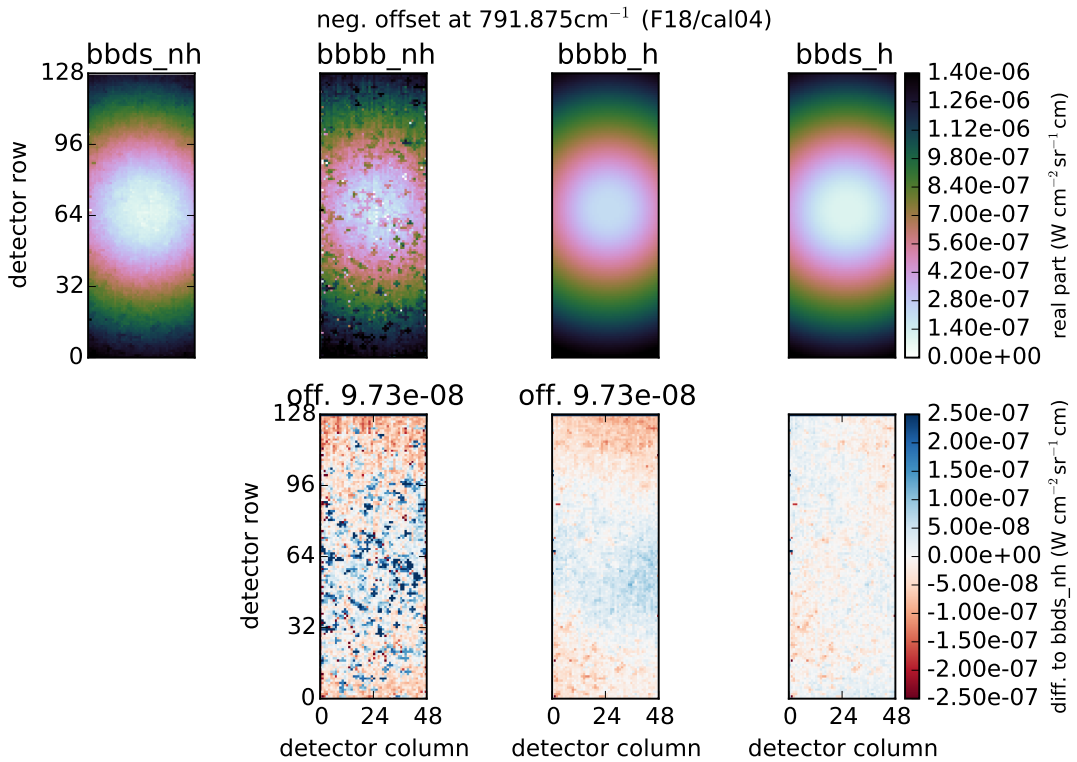

Figure 1: Proposed new figure 2.

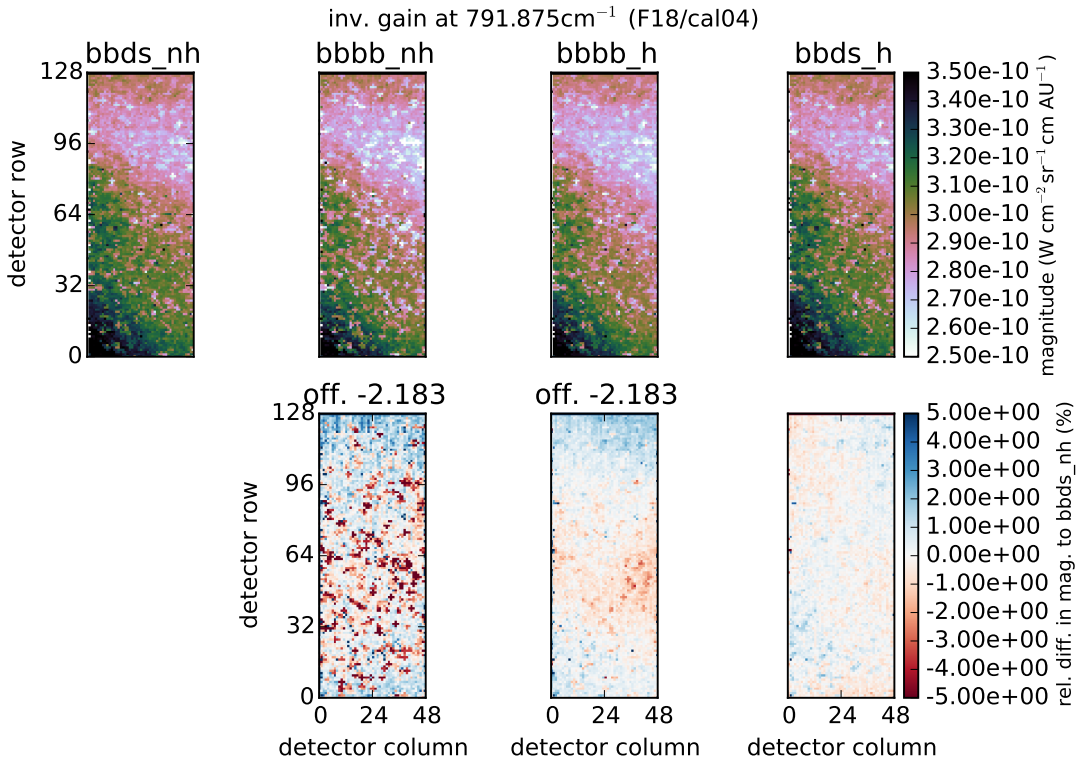

Figure 2: Proposed new figure 3.

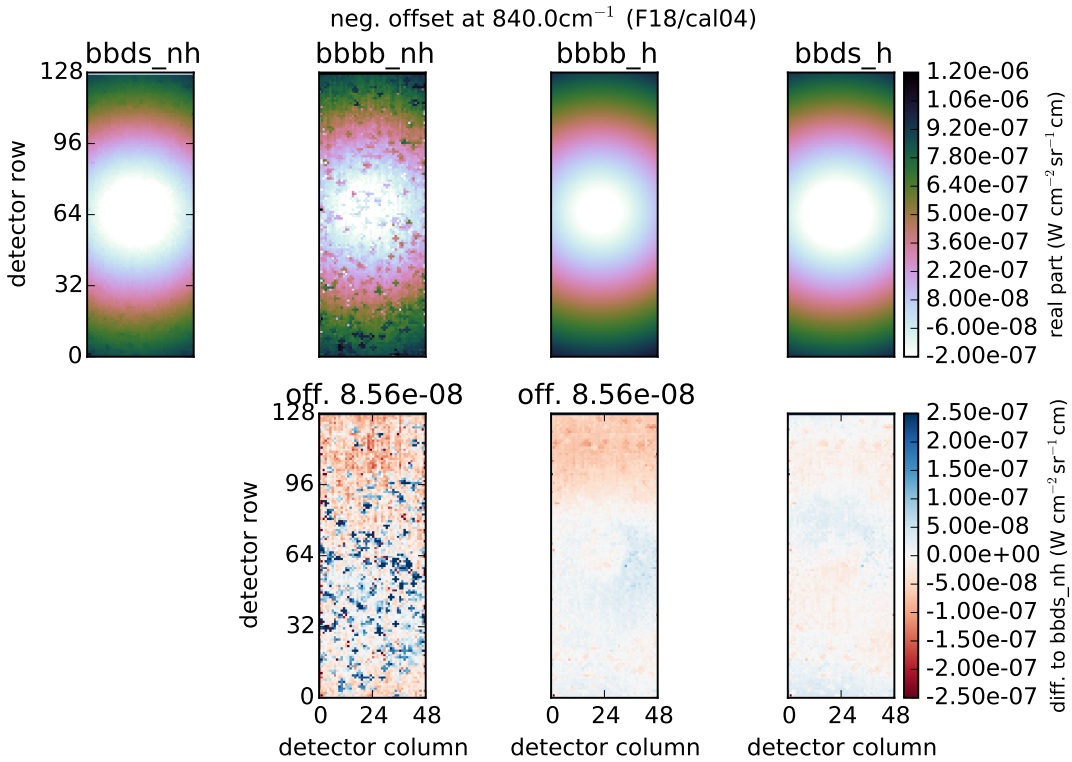

Figure 3: Proposed new figure 4.

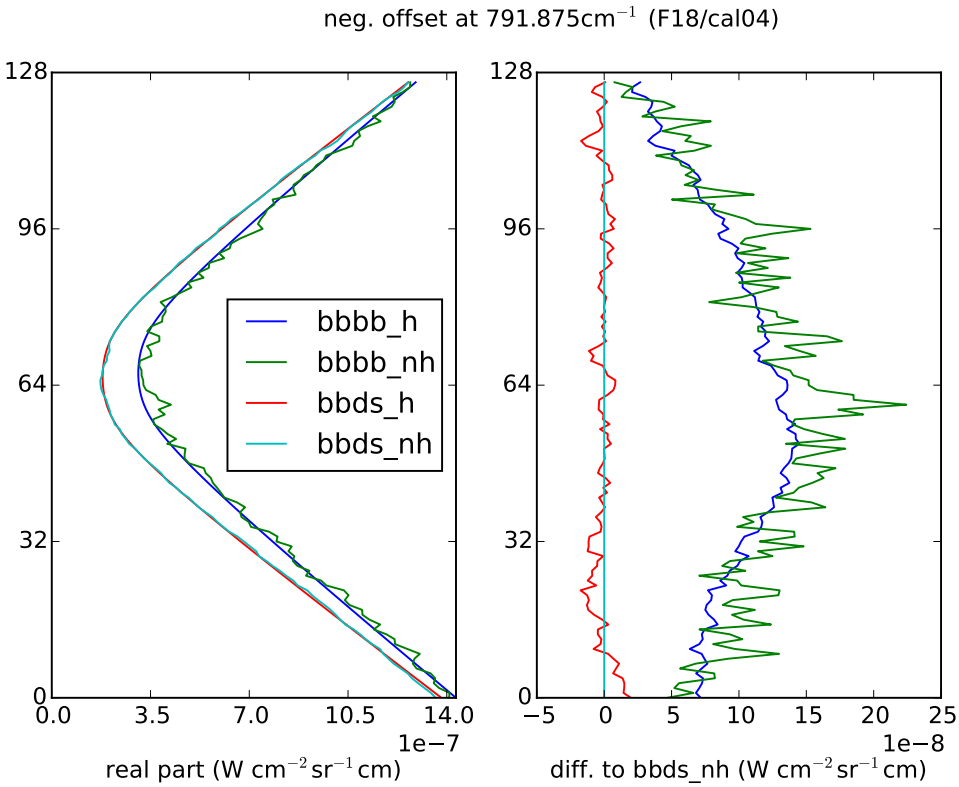

Figure 4: Proposed new figure 5.

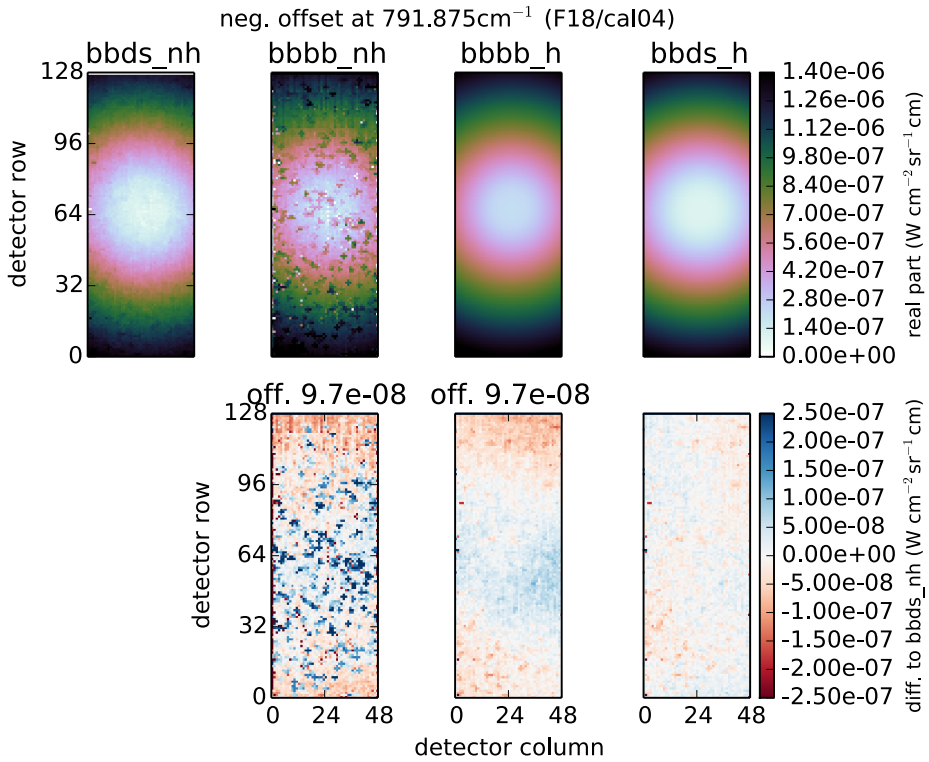

Figure 5: Figure 2 from the original manuscript in new colour scale.

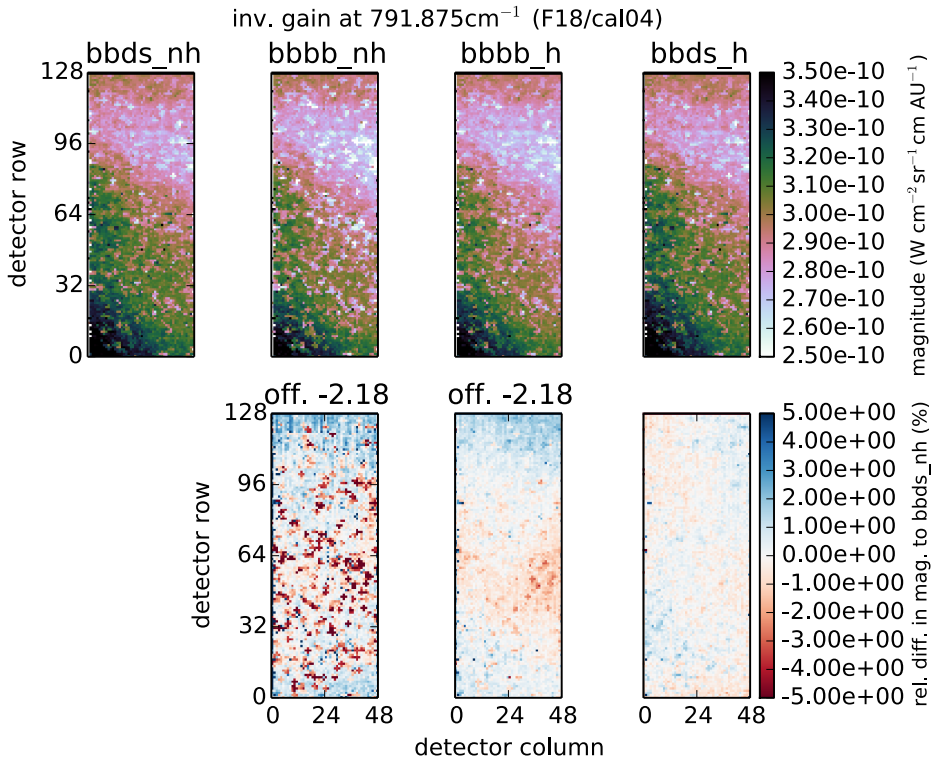

Figure 6: Figure 3 from the original manuscript in new colour scale.

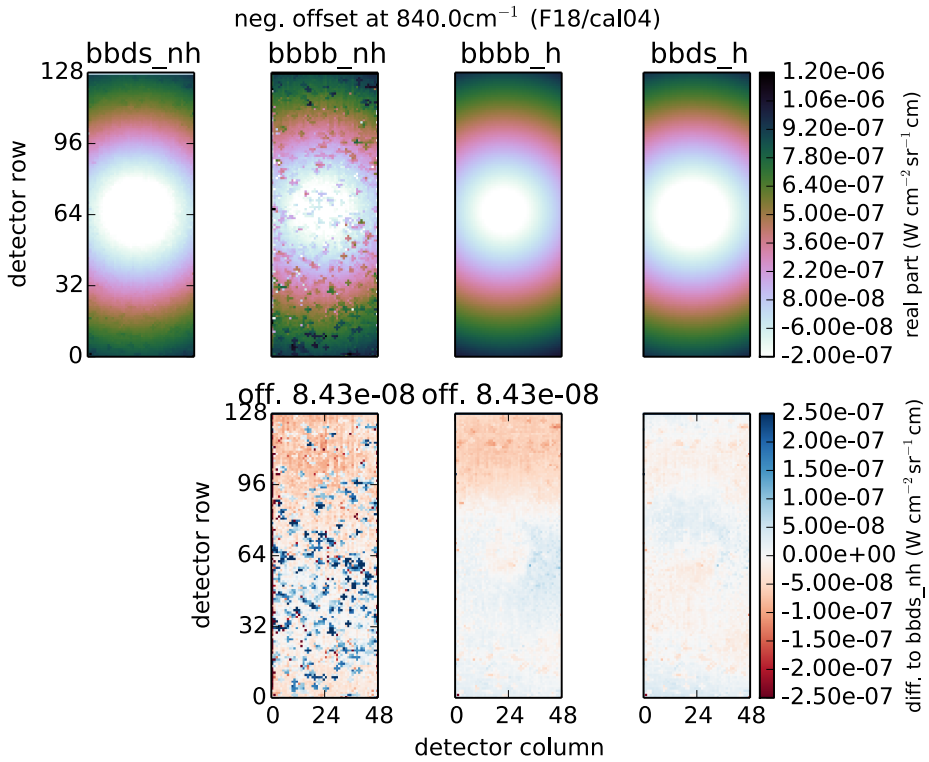

Figure 7: Figure 4 from the original manuscript in new colour scale.

References

Kleinert, A., Friedl-Vallon, F., Guggenmoser, T., Höpfner, M., Neubert, T., Ribalda, R., Sha, M. K., Ungermann, J., Blank, J., Ebersoldt, A., Kretschmer, E., Latzko, T., Oelhaf, H., Olschewski, F., and Preusse, P.: Level 0 to 1 processing of the imaging Fourier transform spectrometer GLORIA: generation of radiometrically and spectrally calibrated spectra, Atmos. Meas. Tech., 7, 4167–4184, doi:10.5194/amt-7-4167-2014, 2014.

Antonelli, P., Revercomb, H. E., Sromovsky, L. A., Smith, W. L., Knuteson, R. O., Tobin, D. C., Garcia, R. K., Howell, H. B., Huang, H.-L., and Best, F. A.: A principal component noise filter for high spectral resolution infrared measurements, J. Geophys. Res., 109, D23102, doi:10.1029/2004JD004862, 2004.

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.