Response to the Editor's Comment (doi:10.5194/amt-2020-429-EC1) on "Retrieval of atmospheric CO_2 vertical profiles from ground-based near-infrared spectra" by Sébastien Roche et al.

We thank Frank Hase for his thoughtful comments. Below we address each comment and the resulting revisions to the text and figures will be added to the final revised manuscript which will also include revisions based on comments from the referees.

Comment 1: scaling retrieval and interlayer constraints

Page 2, line 66: "Scaling retrievals do not require inter-level constraints on a–priori concentration uncertainties" – I agree that a scaling retrieval does not require the explicit construction of inter-level constraints, but actually scaling equals the assumption of very strong inter-level constraints. I would suggest to rephrase the notion on how GFIT handles a scaling retrieval (I hope I understand the authors correctly), as, e.g. "Technically, GFIT handles the scaling retrieval by weakly constraining the fitted VSF factor. The approach is equivalent to performing an optimal estimation of the VSF, assigning a value of unity to the a-priori VSF and a value of 1e6 as its expected range of variability."

Response to comment 1

It was indeed meant that the scaling retrieval does not need explicit interlayer correlations.

The text will be updated as suggested.

Comment 2: line mixing

Page 3, line 96: the line mixing referred to by the authors is (I believe) Rosenkranz line mixing and should be referred to as such.

Response to comment 2

Yes, this is the Rosenkranz approximation for line mixing, the text will be updated to:

> This version of the code implements quadratic speed-dependent Voigt line shapes with line mixing (qSDV+LM) for CO_2 (Mendonca et al., 2016) and CH4 (Mendonca et al., 2017) bands, and qSDV line shapes for O2 in the band centered at 1.27 µm (Mendonca et al., 2019). The line mixing coefficients are derived with the first order Rosenkranz approximation (Rosenkranz et al., 1975).

With the added reference:

> Rosenkranz, P.: Shape of the 5 mm oxygen band in the atmosphere, IEEE Trans. Antennas Propag., 23(4), 498–506, doi:10.1109/TAP.1975.1141119, 1975.

Comment 3: sequential vs parallel retrievals

Page 6, line 147: "We see no advantage to fitting non-contiguous windows in parallel, rather than in series, and then averaging the results." In my opinion, this is a misjudgement. Especially for retrieving profile information, combining weak and strong bands in a simultaneous fit is known to be potentially very advantageous. While the line wings of saturated lines in a strong band carry information about the

lowermost atmospheric layers, weaker lines contribute information on higher atmospheric levels. We tried combination of bands in the context of the cited work by Dohe, which improved the uniformity of partial column sensitivities significantly over what is shown in Fig. 8.3 in the work of Dohe (using only the strong band). At that time, spectroscopic inconsistencies hindered a successful combined fit of weak and strong bands, given the progress on spectroscopic data this might look different today. In any case, the general statement that such a capability does not offer an advantage is in my opinion highly questionable. (The method of sequential estimation would be equivalent to a combined fit (if linearity can be assumed) and could be used as a makeshift solution if a fit of several windows is not supported by the code, but I assume this is not what the authors describe by "averaging the results".)

Response to comment 3

It is what we were describing, we should not have used the word "average". We were referring to the method of retrieving a profile from each window separately, and combining the resulting profiles in post processing by taking into account the jacobian of each window. We have not shown such combined profiles, they tend to be heavily weighted towards the profile retrieved from the Strong window. The combined averaging kernel indicates better sensitivity than any single window, but the combined profile is as affected by biases caused by temperature errors as the Strong window profiles. This should also be the case for parallel retrievals.

A simultaneous retrieval could be an advantage for gases for which the problem is more non-linear, where the change in the jacobian computed with the a priori state compared to the jacobian computed with the retrieved state is relatively large. But CO_2 is a special case as the a priori profiles compare well with the true profiles.

A sequential retrieval also gives the ability to diagnose potential issues between the different windows.

Comment 4: sensitivity study

Section 3.1: In my feeling, this section would better correspond to the following investigation using measured spectra if a similar (not much higher) SNR and a similar and more realistic a-priori covariance would be used. I miss a sensitivity study concerning ILS in section 3.1 (this should be possible to realize although the code does not yet explicitly support ILS parameters by using slightly different acceptance cones in the forward calculation and in the retrieval, thereby modelling a modulation loss via the selfapodisation). I also miss a sensitivity study with respect to the offset (as it cannot be fitted in the nonsaturated bands). By the way: as a DC-correction is implemented and as the non-linearity is well controlled (I assume), why is an offset observed in the saturated band? Does this offset behave like noise – varying from spectrum to spectrum – or is this a systematic offset? It would be interesting to provide a figure showing the behaviour of the fitted offset for a larger number of spectra (could be included in fig E2?). I find the fact that the fit residuals of the unperturbed spectrum shows these oscillatory high-resolution features in the residuals very irritating. Why is it not possible to achieve a self-consistent code configuration that applies the same approximations in the simulation of the spectrum as in the forward calculation performed as part of the retrieval? Note that for the disturbance of line parameters (panels e and f in figure 7), these artificial features seem to be of similar size as those invoked by the perturbation.

Response to comment 4

Constraints

The high SNR and loose prior constraint used in Sect. 3.1 highlights the source of variability in the retrieved profiles by forcing the retrieval to pull most of the information from the measurement. The section shows that including smoothness constraints on the CO_2 profile will mainly serve to reduce variability that is mainly caused by errors in the a priori temperature profile. Here we want to be able to identify the sources of variability to indicate which parts of the forward model should be improved in priority. The magnitude of scaling retrieval residuals caused by typical temperature errors is much larger that caused by incorrect CO_2 prior profile shape. Without a temperature retrieval scheme, the adjustment to the CO_2 profile in a profile retrieval will be mainly driven by variability in the temperature profile.

This is a critical point in the paper. The prior constraint should not be adjusted unless oscillations in the retrieved profiles caused by typical errors in sources other than CO_2 are smaller than typical CO_2 variability even when using a diagonal a priori covariance matrix.

ILS and Zero-level offset

Below a type of ILS error is considered (widening by perturbing the internal field of view diameter), and the effect of a zero-level offset is shown.

Figure R1(c) and R1(d) show the effect of perturbing the internal FOV diameter by +7%, which leads to a widening of the ILS. The deviations from the truth are within 1 ppm for $P > 0.5$ atm and within 3 ppm for $P < 0.5$ atm. The internal field of view diameter of the spectrometer is 2.4 mrad.

Figure R1(e) and R1(f) show the effect of a +0.002 perturbation in the zero-level offset, without retrieving it in the Strong window. This has a large effect in the profile retrieved from the Strong window, showing deviations from the a priori within 30 ppm, and a smaller but still significant effect in the other windows with deviations up to 10 ppm.

Figure R2 shows the zero-level offset (ZLO) retrieved from the Strong CO_2 window for 1 year of measurements at East Trout Lake, showing the data that pass TCCON quality flags (52884 spectra). ZLO is expressed as a fraction of the continuum level. The 500-point rolling mean ±1σ are shown as red and yellow lines. It stays within ±0.001 throughout the year, and there appears to be a decreasing trend. The median absolute value of the ZLO is 0.0002.

Figure R3 also shows ZLO but for all the Lamont spectra used in the study, with dashed lines marking the median for each date.

Figure R4 shows the daily median ZLO at various TCCON sites, since 2013 the maximum variation at a single site up to 2018 are ~0.002

This sensitivity test for a perturbed FOV and ZLO will be added as Appendix F. It shows the effect of typical zero-level offsets will not be a major source of variability in the retrieved profiles in the Weak and TCCON windows, at least not in the Lamont data used here. If the zero-level offset obtained from the Strong window with real spectra is added in the TCCON and Weak windows before the retrieval, the change in the retrieved profiles is less than 3 ppm at all altitudes for the Lamont spectra as shown in Fig. R5.

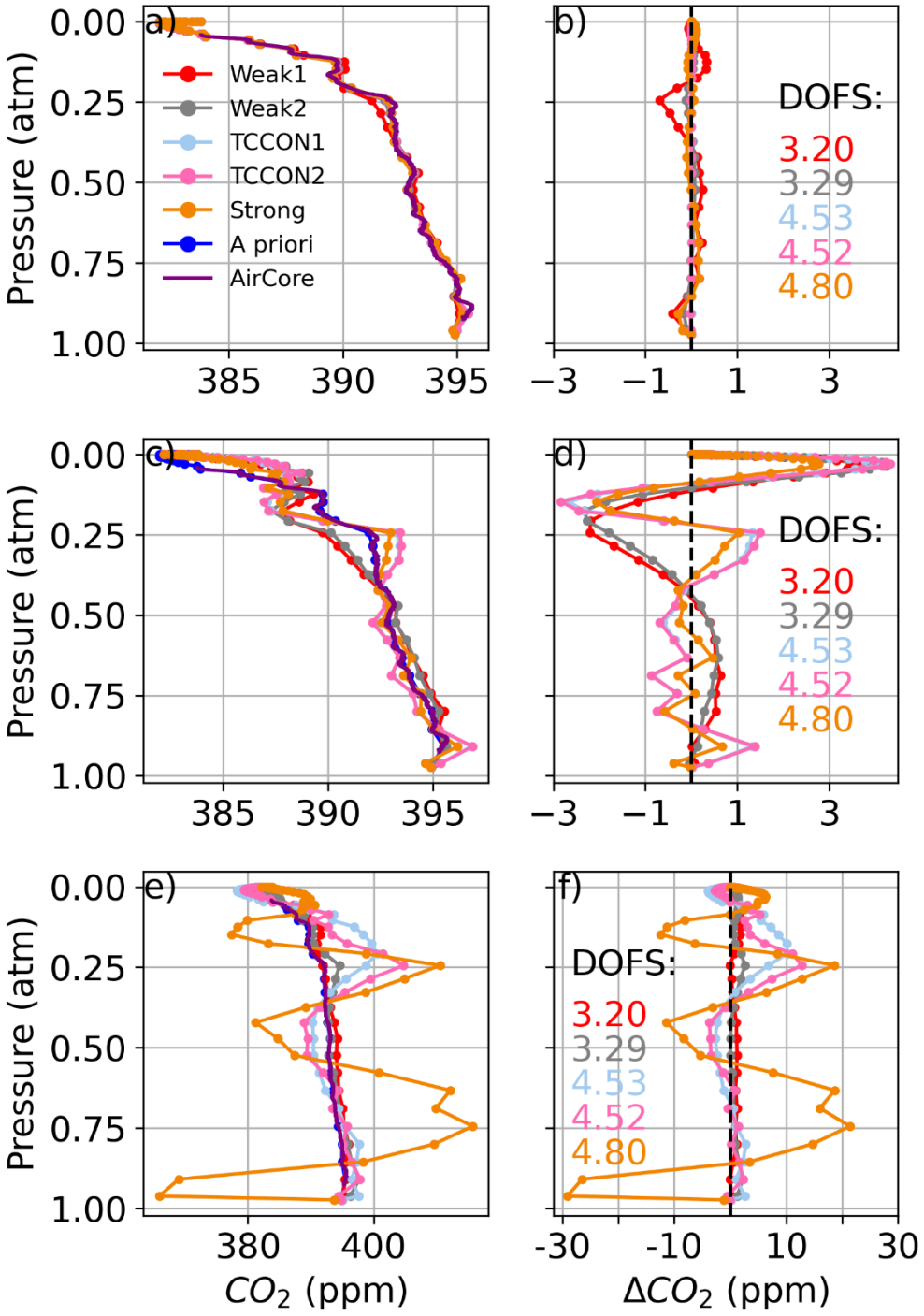

Figure R1: The left-hand panels show CO_2 profiles retrieved using synthetic spectra. In (a), we use the AirCore profile, which was used to generate the synthetic spectra, as the a priori. In (c), the field of view is perturbed by +7%, increasing the width

of the ILS. In (e), the zero level offset is perturbed by +0.002 and is not retrieved in the Strong window. The right-hand panels: (b), (d), and (f), show the difference between the retrieved profiles and AirCore, corresponding to (a), (c), and (e) respectively.

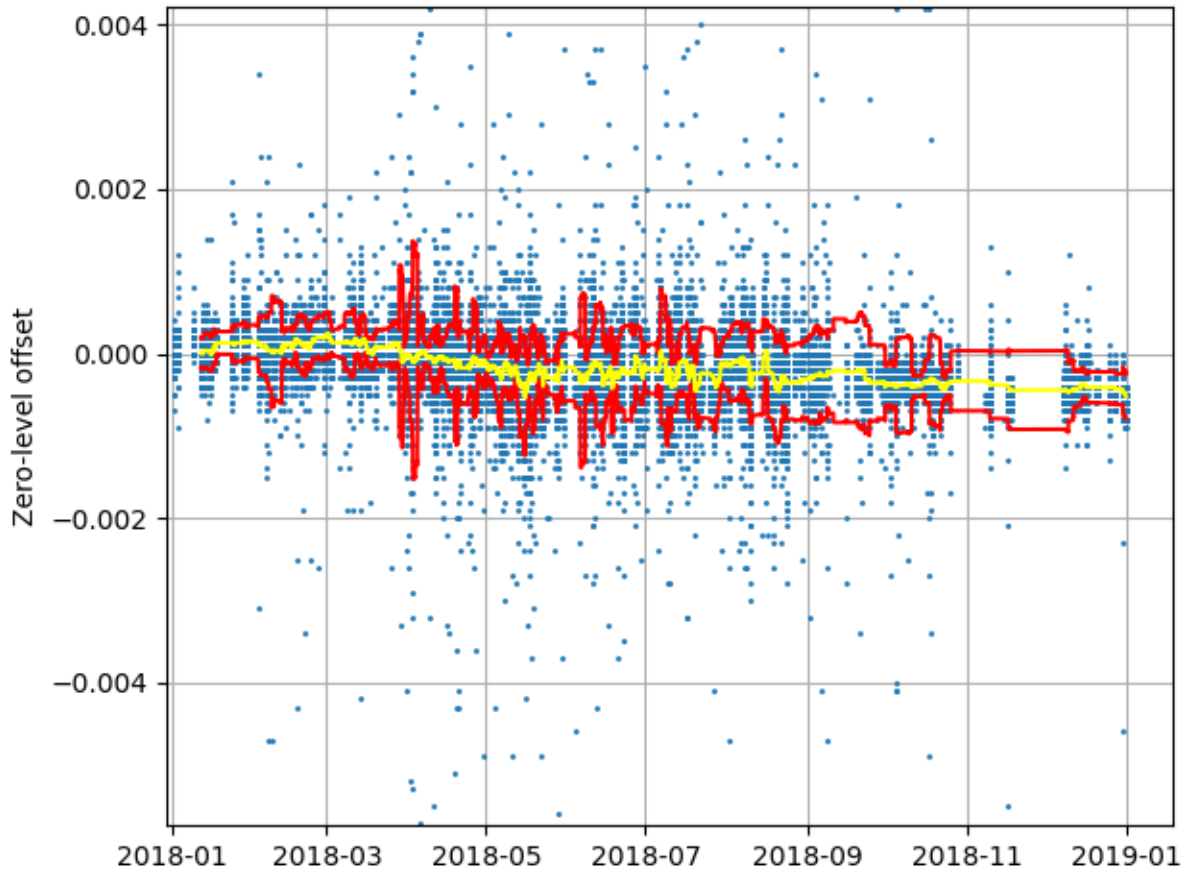

Figure R2: Zero-level offset retrieved from the Strong CO_2 window for 1 year of measurements at the East Trout Lake TCCON station, with 52884 total spectra. The yellow line shows the 500-point rolling mean and the red lines show the 500-point standard deviation.

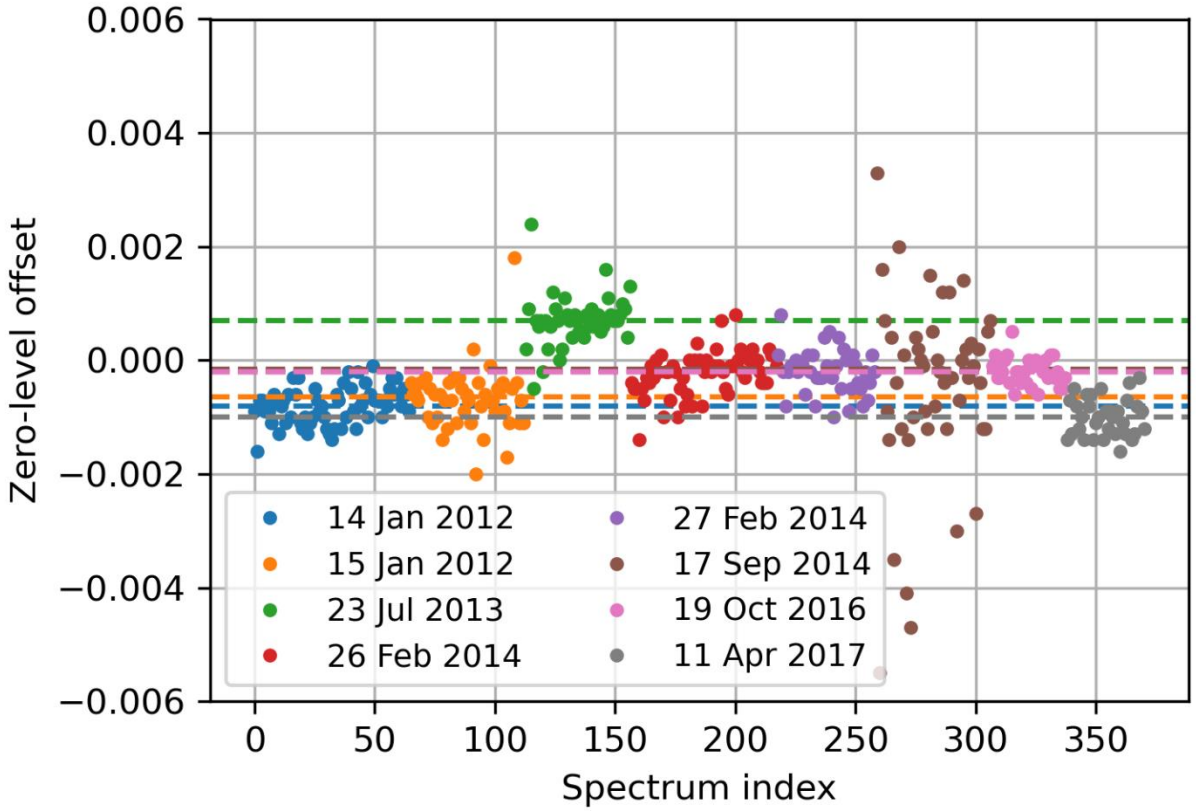

Figure R3: Zero-level offset retrieved from the Strong CO_2 window for the Lamont spectra coincicent within ±1 hour of the last AirCore sampling time and within ±1.5 hour of the closest a priori time on each day indicated by the legend (yyyymmdd). The dashed lines represent the median value for each date.

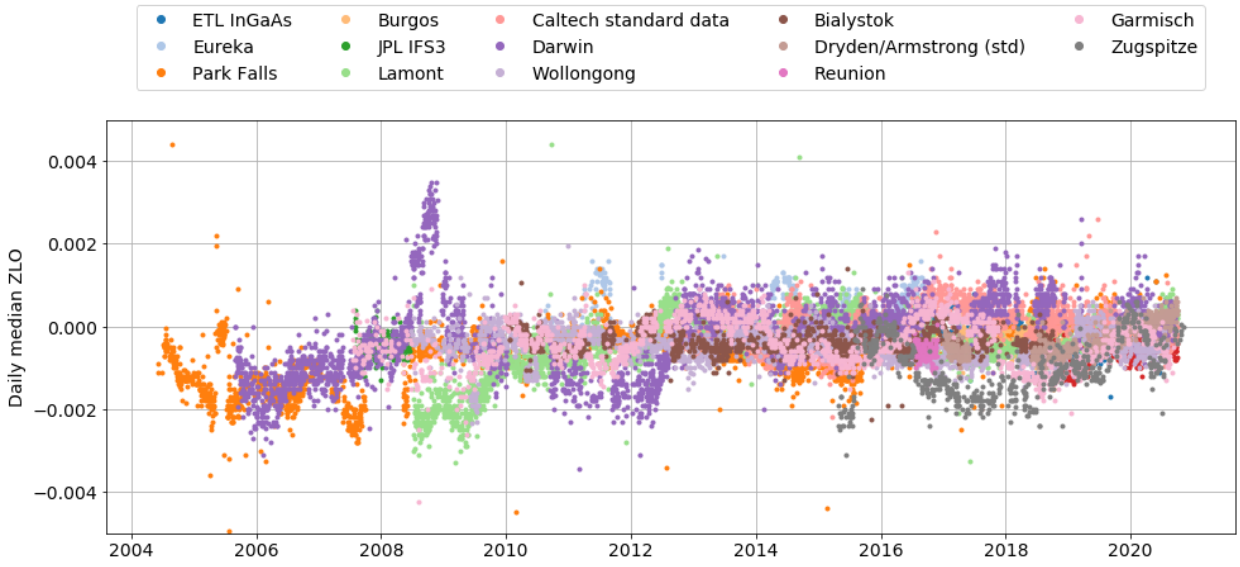

Figure R4: daily median zero-level offset at the TCCON sites indicated by the legend.

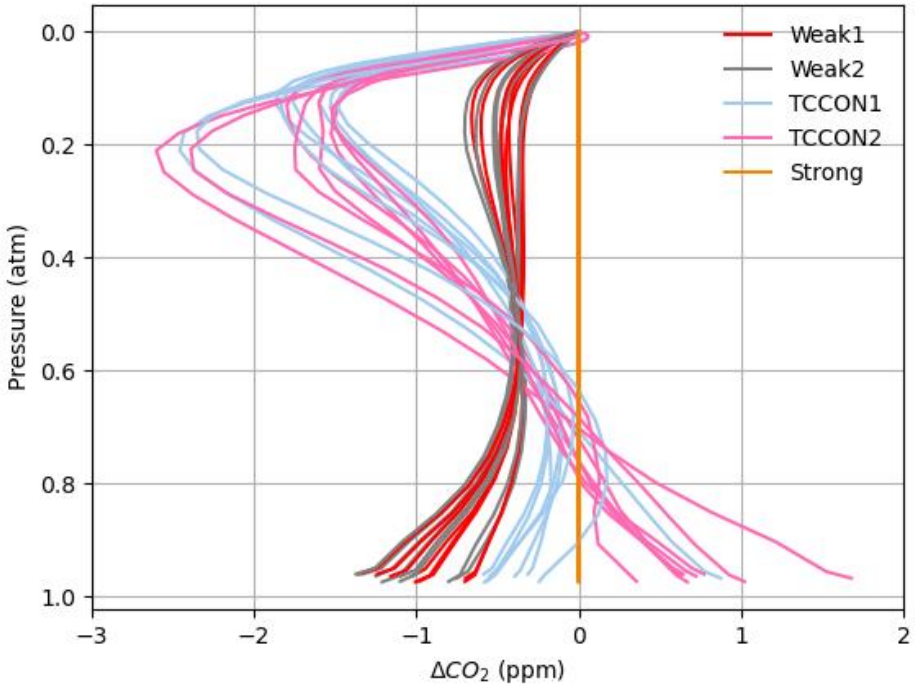

Figure R5: using real Lamont spectra and the AirCore profile as a priori, the zero-level offset was first retrieved from the Strong window and then added in the Weak and TCCON windows. The difference in the retrieved profiles with and without the added offset is shown for each window for all the days with AirCore profiles over Lamont. In the Strong window, where the offset is retrieved, the differences are less than 0.001 ppm.

Synthetic spectra, fit residuals in reference case

In GFIT, the measured spectrum is convolved with the truncated and windowed instrument function, but with the rectangular part of the ILS having 0 width. This improves the agreement between measured and calculated spectra. But this process was also applied when using synthetic spectra as observations and lead to the high frequency residuals observed in the reference case in Figure 7 of the paper, which correspond to differences in the synthetic spectrum before and after the convolution. Not applying this process for synthetic spectra removes most of the residuals in the reference case.

Figure R6 below shows the effect on profiles retrieved for the reference case when no perturbation is applied. Profiles become closer to the priori, within less than 0.3 ppm for the Weak and Strong windows, and within less than 0.01 ppm in the TCCON windows. The fits to synthetic spectra are still not perfect, but the residuals in the reference case are now an order of magnitude smaller than residuals obtained when perturbing the spectroscopic parameters.

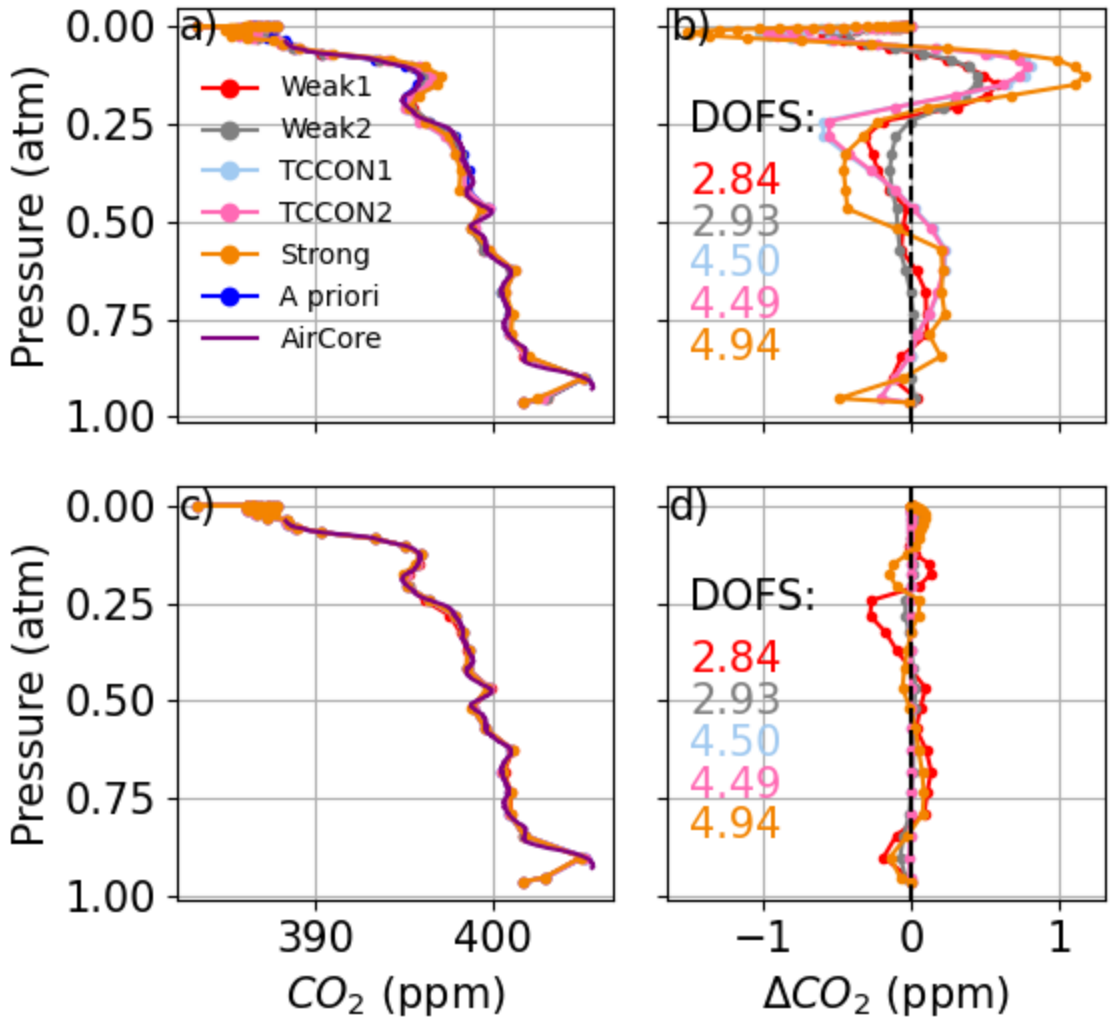

Figure R6: profiles retrieved from fits to synthetic spectra with no perturbations. In (a) with the extra convolution of the synthetic spectrum with the ILS, and in (c) without.

Figures of Section 3.1 were updated with this fix, it only has a visible effect on retrieved profiles in Fig. 3(a) and 3(b) (the reference case) and does not affect the results and conclusion of this section.

Comment 5: adding XCO_2 panel to figure 11

Figure 11: it might be instructive to add a further panel displaying XCO_2 for each retrieval setup.

Response to comment 5

The difference in XCO_2 between scaling and profile retrievals is less than 5 ppm while XCO_2 between the different days varies by ~20 pppm and putting XCO_2 from both profile and scaling retrieval on a same panel in Fig. 11 wouldn't be clear. Figure R7 shows an updated Fig. 11 with the profile retrieval XCO_2 and smoothed AirCore XCO_2.

And Fig. R8 shows an updated Fig. 11 but with the new panel showing the difference between profile and scaling retrieval XCO_2.

The number of spectra showed was also updated to correspond to the spectra used to produce the profile figures of Sect. 3, coincident within ±1.5 hours of the a priori closest to the AirCore time in addition to within ±1 h of the AirCore. Fig. 11 in the paper will be updated to Fig. R8 below.

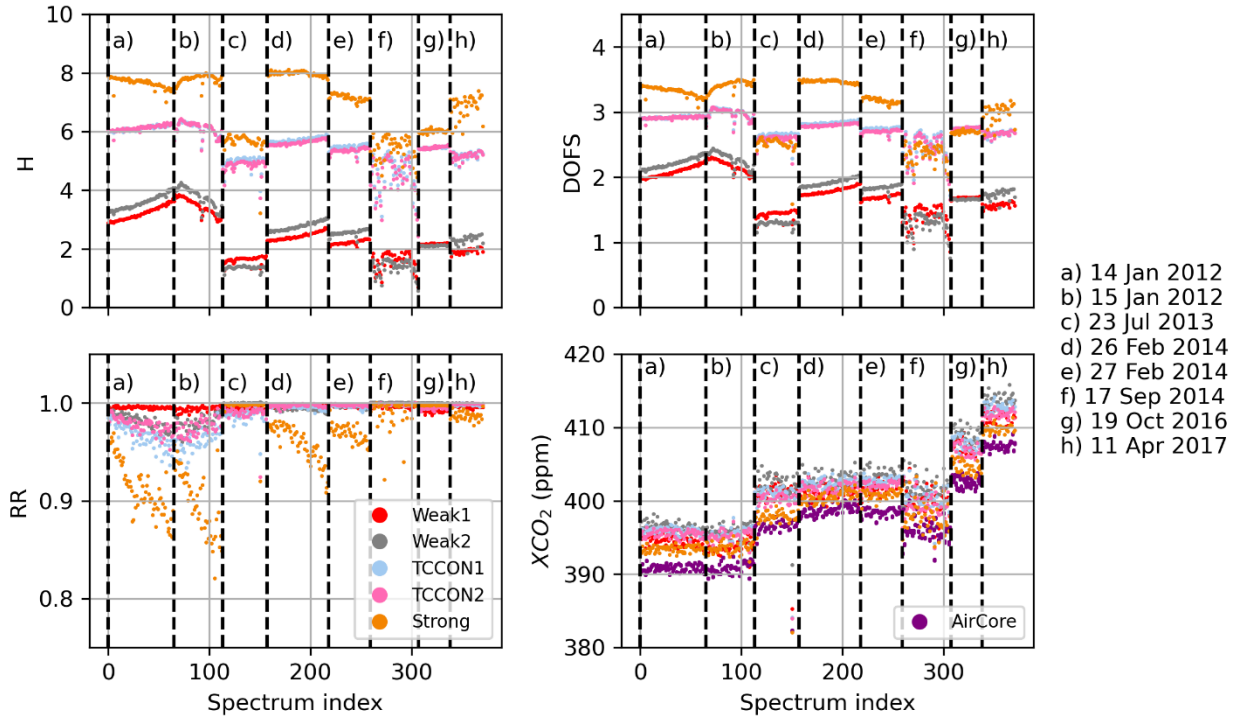

Figure R7: Shannon information content (top left), degrees of freedom for signal for the CO_2 profile (top right), ratio of residuals (bottom left), and profile retrieval XCO_2 as well as the smoothed AirCore XCO_2 (bottom right), for all Lamont spectra coincident within ±1 h of the AirCore last sampling time for AirCores launched on the dates indicated on the right, and within ±1.5 hours of the closest GGG2020 a priori in time. Each new date is marked by a vertical dashed line.

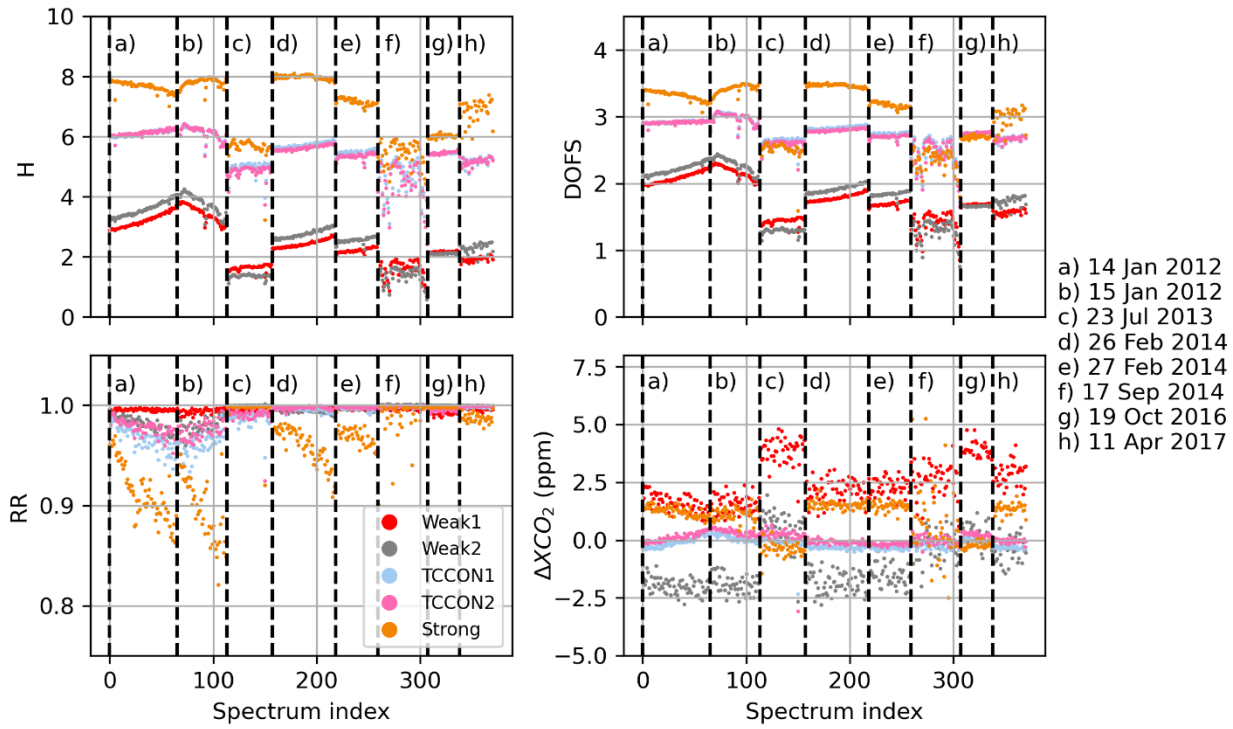

Figure R8: Shannon information content (top left), degrees of freedom for signal for the CO_2 profile (top right), ratio of residuals (bottom left), and profile retrieval minus scaling retrieval XCO_2 (bottom right), for all Lamont spectra coincident within ±1 h of the AirCore last sampling time for AirCores launched on the dates indicated on the right, and within ±1.5 hours of the closest GGG2020 a priori in time. Each new date is marked by a vertical dashed line.

Comment 6: figure 12, XCO_2 bias

Figure 13: Why is GGG XCO_2 so significantly off from the in situ value (scaling retrievals in the TCCON bands are off by about 4 ppm)? I assume that required calibration factors have not yet been applied?

Response to comment 6

There is no airmass dependent correction factors or in-situ correction factors applied. This ~1% offset might be due to uncertainties in line intensity of ~1% in the TCCON windows. The retrieved VSFs in the TCCON windows in GGG2020 are typically ~1.01. Recent line intensity measurements at the US National Institute of Standards and Technology (Long et al., 2020) in the region corresponding to the TCCON1 window can be made with ~0.1% uncertainty. A similar study is being conducted for the CO_2 band in the TCCON2 window. We recently tested including these line intensity measurements in the linelist for the TCCON2 window, with these the retrieved VSF become ~1.002 instead of ~1.01 when using the AirCore as a priori.

Long, D. A., Reed, Z. D., Fleisher, A. J., Mendonca, J., Roche, S. and Hodges, J. T.: High-Accuracy Near-Infrared Carbon Dioxide Intensity Measurements to Support Remote Sensing, Geophys. Res. Lett., 47(5), doi:10.1029/2019GL086344, 2020.

Comment 7: bias and smoothness constraint

Note that this problem of bias is connected to a significant problem in your retrieval setup: as far as I oversee, the CO_2 a-priori profiles used are intended to match with the actual CO_2 mixing ratios in the atmosphere. But the CO_2 becomes calibrated to WMO scale only further down the processing line (by rationing CO_2 column over the O_2 column and application of ADCFs + AICFs). Let's assume the calibration is off: this poses no problem for a scaling retrieval (same residuals, same solution, only a different scaling factor is reported). However, an optimal estimation setup will misbehave if the a-priori is biased: it will try to approach the a-priori values in altitudes where the sensitivity is low and in effect the retrieved profile will oscillate significantly. I think it would be superior to apply a pure smoothness constraint as used by Dohe instead of an optimal estimation with diagonal constraint and just some correlation on top. If a pure smoothness constraint is used, a bias of the a-priori remains without effect even for a profile retrieval (same residuals, same solution, only different scaling factors on identical profile shapes). Moreover, this setup is preferable because it allows a smooth transition between the profile retrieval and the simple scaling retrieval (just increase the smoothing constraint further and further).

Response to comment 7

With the a priori covariance used in the study, a bias in the a priori CO_2 profile has little effect on the retrieved profile shapes until the bias becomes larger than ~3%. This would be very uncommon for CO_2 priors at high altitudes. All the AirCore profiles over Lamont are within the GGG2020 a priori ±1% for P < 0.9 atm. Fig. R9 shows the effect of a +3% and +10% offset to the a priori CO_2 profile. It shows the retrieved profiles stay close to the a priori at the highest altitudes where there is little information. And to match the measured spectrum this is compensated with less CO_2 retrieved at P > 0.8 atm in the Weak and TCCON windows.

To avoid issues when the a priori bias is unusually high, a scaling retrieval could be performed first, and the resulting profile could be used as a priori for the profile retrieval. But with the data used in the study the effect of a bias in the a priori is small compared to the effect of temperature errors, or compared to the effect of other types of errors that produce the oscillations observed in Fig. R9(a).

Figure R10 shows the difference in profiles retrieved with the AirCore as a priori, and with the AirCore increased by 3% as a priori. This is the difference between retrieved profiles shown in Fig. R9(a) and R9(c), but for all 10 Lamont days with AirCore profiles. Here it is clearer that the deviation at P > 0.8 atm is larger in windows that have less sensitivity to the lowest altitudes while the Strong window shows differences within ±3 ppm at P > 0.1 atm. And the 3% offset considered here is unrealistically high.

Comment 7 is also addressed in Response to comment 4: constraints. It would be interesting to evaluate different a priori constraints after the implementation of a temperature retrieval or correction.

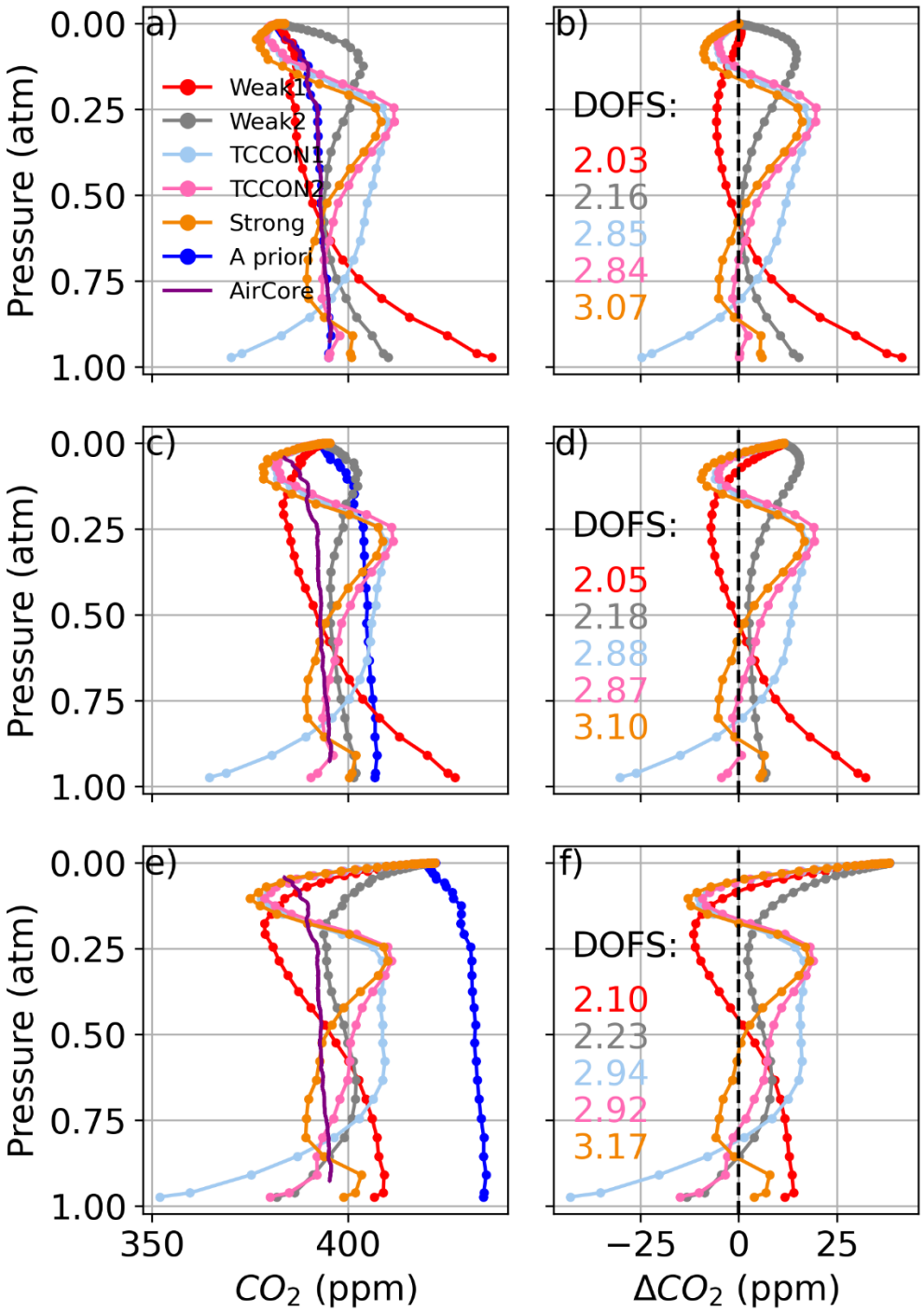

Figure R9: The left-hand panels show CO_2 profiles retrieved using real spectra on 14 January 2012. In (a), we use the AirCore profile as the a priori. In (c), the a priori is increased by 3%. In (e), the a priori is increased by +10%. The right-hand panels: (b), (d), and (f), show the difference between the retrieved profiles and AirCore, corresponding to (a), (c), and (e) respectively.

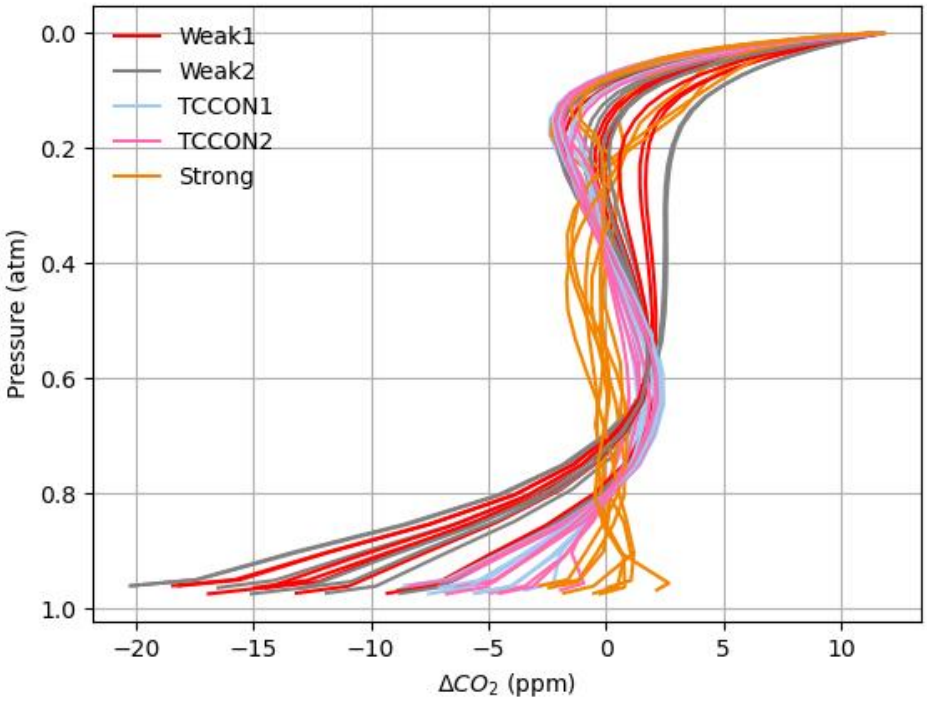

Figure R10: change in the profiles retrieved from real Lamont spectra when using AirCore as a priori and when using AirCore increased by 3% as a priori. Each window has 10 lines for each day with an AirCore profile over Lamont.