Dear Dietrich Feist, dear referees,

First of all thank you for your valuable comments. We think we could improve the manuscript following your suggestions.

Overall comments:

1. We changed the name of the ECMWF model from MACC (Monitoring of Atmospheric Composition and Climate) to CAMS (Copernicus Atmosphere Monitoring Service) to be consistent with the most recent naming conventions. The model data is unchanged.

2. Both referee #1 and referee #2 asked us to take a closer look at the model-measurement comparison regarding the smoothing effects of the averaging kernel matrix. In general, the retrieved column-average concentrations is given by

$$\hat{\boldsymbol{c}} = \boldsymbol{h}^{\boldsymbol{T}}\hat{\boldsymbol{x}} = \boldsymbol{h}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x}_{\boldsymbol{true}} + \boldsymbol{h}^{\boldsymbol{T}}(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_{\boldsymbol{apr}} + error\ terms \qquad (1)$$

where $\boldsymbol{h}^{\boldsymbol{T}}$ is the column-average operator that calculates the column-average mixing ratio $\hat{\boldsymbol{c}}$ from the retrieved profile $\hat{\boldsymbol{x}}$, $\boldsymbol{A}$ is the averaging kernel matrix, $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{x}_{\boldsymbol{true}}$ is the true vertical profile and $\boldsymbol{x}_{\boldsymbol{apr}}$ is the a priori profile [e.g., Butz et al., 2012, equations (15) –(17)]. In our case, the retrieval $\hat{\boldsymbol{x}}$ is simply the scaled prior $\boldsymbol{x}_{\boldsymbol{apr}}$. Thus, when comparing model to measurement, spurious differences originate from the second term on the right-hand-side of equation (1) if $\boldsymbol{x}_{\boldsymbol{apr}}$ is not equal to the model profile (assumed to be $\boldsymbol{x}_{\boldsymbol{true}}$ for comparison purposes).

To avoid this spurious difference pointed out by the referees, we set up a new retrieval run using the co-sampled CAMS (formerly MACC) profiles as a priori for our CO_2, CH_4, and H_2O retrievals. This renders a posteriori adjustments of the retrievals via the averaging kernel calculus of equation (1) obsolete. In the initial version of the manuscript, we used Carbontracker (CT) and Tracer Model 4 (TM4) monthly averaged concentration fields as a priori for CO_2 and CH_4, and a static a priori profile for water vapor. The revised version of the manuscript uses the new retrieval run based on CAMS priors. Therefore, all the datasets and plots saw mostly minor changes.
The impact of the use of different a priori profiles is shown in figure 3 of this document.

3. A significant change was caused by replacing the static H_2O prior by co-sampled CAMS profiles. The initial manuscript detected a discontinuous mismatch (on March 18) between retrieved ground-pressure (calculated from retrieved O_2 and H_2O) and in-situ measured ground-pressure. We speculated that the discontinuity was caused by a slight change in the instrument line shape after opening the instrument.

   With the new run using more realistic prior H_2O profiles, the discontinuity in ground-pressure mismatch is substantially reduced and our speculation on a discontinuous change in the ILS is most likely wrong. So we removed the entire part of section 3.3 that describes the ILS correction because it is not needed anymore. Figure 1 and 2 showing the differences in the ILS retrieval. Note the different scales.

4. Since the initial submission, the TCCON data from Karlsruhe saw a significant update which we decided to propagate into the revised manuscript. Our Polarstern data are affected through the calibration against TCCON standards (end of section 3.3). The new calibration yields XCO_2 and XCH_4

lower by approximately -0.92 ppm and -2.5 ppb, respectively. A global scaling factor added to the manuscript ensures the compatibility to the previous product.

Reply to Referee #1:

1. Section 3.1. In the description of the PROFFIT setup, it is stated that the software is configured to do a scaling of the a priori with a single factor. This is of course the manner in which the official TCCON analysis product, GFIT, operates its inverse method. The question is though is it completely the same? Is this the only regularization that is applied? Is PROFFIT effectively running in more or less the exact same way as GFIT.

Are the degrees of freedom for signal the same, more or less?

The aim of this campaign and the publication is not to perform a TCCON-equivalent measurement. Focus of this publication lies clearly on the mobile application and the interhemispheric gradient. It should be a demonstrator of the general performance of our remote sensing measurements under harsh ambient conditions. Therefore, the question whether PROFFIT96 does exactly the same retrieval as GFIT is outside the scope of this publication. There are other publications addressing this question to a certain extent (e.g. Hase et al. 2004 or Gisi et al. ,2012), but the question might deserve a dedicated study.

PROFFIT is set up to deliver degrees of freedom of exactly one (DOF=1). PROFFIT performs a scaling retrieval, meaning that the a priori profile is scaled by the algorithm such, that the total column fits the measurement best. So, no profile retrieval is being performed and the shape of the retrieved profile is equal to the shape of the prior profile.

2. Section 4, page 19. This leads onto another issue of the characteristics of the retrievals. The authors state that the effects of the measurement averaging kernels have been neglected. The EM27/SUN has a significantly different OPD to the HR125, the standard TCCON instrument. Since the averaging kernel (and dof) are largely driven by the OPD, apriori profile, solar zenith angle, and the apriori and measurement Covariances, I would have thought that it is important to first show that the averaging kernel of the EM27 (the early EM27 papers did not do this). The assumption the authors are making here is that the smoothing effects from the EM27 are relatively small. This should be demonstrated explicitly.

We do not make the assumption that the smoothing effects of the EM27 are relatively small. The instrument and the retrieval smooth the atmospheric profile such that we obtain 1 degree-of-freedom in the vertical, i.e. no vertical information is retrieved. New Figure 9 (discussed in new Appendix A) in the revised manuscript shows the column sensitivity of the EM27 retrievals as a function of pressure height and solar zenith angle. The EM27 sensitivity is very similar to the averaging kernel of an HR125 retrieval (e.g. Fig. A in Wunch et al., 2011). Thus, we expect averaging kernel effects to be small. However, in principle, we agree that averaging kernel effects should be considered when deriving the overall calibration factor between our ship records and TCCON. Since the focus of the manuscript is on relative differences along the latitudinal transect and not on the overall calibration, we did not change the manuscript and the data in that respect.

3. Section 4 discussion and Figure 9. Following on from point 2 above, all comparison data should be smoothed with the EM27/SUN averaging kernel, both model and satellite. That is, unless it can be demonstrated in point 2 above that the averaging kernel effects are very small.

As described by the introduction and equation (1), smoothing the model by the averaging kernel matrix is obsolete since we now use the model profiles as a priori. Smoothing the satellite retrievals (owing approximately 1 DOF) by the EM2/SUN averaging kernels (owing exactly 1 DOF) would not be adequate. Smoothing one dataset by the averaging kernel of the other is only reasonable if the first one has significantly better vertical resolution than the second. This is the case for the model but not for the satellite retrievals in comparison to the EM27/SUN retrievals.

However, we agree that, in principle, comparing the EM27/SUN retrievals to the satellite records would require adjusting the a priori term in equation (1) to a common a priori. We decided to perform that adjustment for the model-EM27/SUN comparison but not for the satellite–EM27/SUN comparison, since the model-EM27/SUN comparison turns out to be scientifically more interesting one.

4. Section 3.3: it would be interesting to know from an operational point of view, why the ILS of the EM27 changed by opening it. Other users of the EM27/SUN may want to know what was adjusted, if anything, and what this implies for the instrument stability while on a campaign. Clearly it is mandatory to measure the ILS before and after such a campaign. But what if, for some reason, it was required to open the instrument more than once? Does this mean the ILS must be remeasured each time?

In the paper we used an ILS-parameter (Instrumental Line Shape) retrieved from the oxygen column measurement, to indicate instrumental changes. The new retrieval run indicates that this approach is influenced by the choice of the water vapor a priori profile. Performing the same ILS retrieval with the water vapor a priori from CAMS, the retrieved ILS-parameter is more constant over time and the discontinuity detected in the initial manuscript is largely reduced. The ILS-correction procedure is not needed anymore and the corresponding discussion is removed. Essentially, our speculation that opening the instrument caused a slight change of the ILS was most likely wrong.

Minor comments

Thank you for mentioning the minor errors in the paper (e.g. typos or figure labels). We corrected them without further comment.

Reply to referee #2:

1) In the current state of the manuscript it is not clear how much information is coming from the measurement and how much from the a priori. This should be investigated and included in the final paper. I regard this as very important. Carbon Tracker / TM4 simulations are used as the apriori for the retrieval. Although these model runs are for a different year (I assume the months March/April), the latitudinal pattern is probably similar. Hence even in the extreme case that the measurements do not contribute any information, a reasonable agreement with the CAMS-model is expected.

The PROFFIT96 retrieval is set up such, that it performs a scaling retrieval and delivers XCO_2 and XCH_4 with degrees of freedom equal to one (DOF=1). The CarbonTracker/TM4 apriori profiles were taken from monthly mean profiles in March and April of the year 2009. The data were interpolated in space and time to avoid discontinuities. In the revised version, we replaced the CarbonTracker/TM4 a priori by the CAMS (formerly MACC) model data to avoid averaging kernel effects in the model-measurement comparison as explained in the general comments and the reply-to-reviewer #1.

Since our retrievals are constraint to deliver DOF=1, the total column of CH4 is largely unconstrained by the a priori. However, deviations between the shape of the true profile ($\boldsymbol{x}_{true}$ in equation (1)) and the shape of the a priori profile ($\boldsymbol{x}_{apr}$ in equation (1)) can have an impact on the retrieved columns according to the column sensitivity of the retrieval. New appendix A and new figure 9 discuss the column sensitivities as a function of level pressure and solar zenith angle. The column sensitivity for both, XCO_2 and XCH_4, is largely uniformly distributed throughout the atmosphere with somewhat higher sensitivity in the lower and somewhat lower sensitivity in the upper atmosphere as typical for ground-based direct sun measurements. Figure 3 in this document shows the retrieval difference for static or more realistic a priori profiles. Note that Carbontracker / TM4 a priori profiles in this study are temporal and/or zonal averages resulting in a less realistic a priori profile than the CAMS model run with high temporal and spatial resolution.

Finally, to illustrate the effect of the choice of the a priori, we conducted an (inferior) retrieval with constant CO_2, CH_4, and H_2O a priori and compared it to the new standard run with CAMS a priori profiles.

2) The averaging kernels are neglected in the model comparison. In my opinion this should be changed.

As outlined in the general comments and reply-to-reviewer 1, we reprocessed the entire campaign data using CAMS (formerly MACC) a priori profiles. This makes the use of the averaging kernels obsolete in the model-measurement comparison. We conducted the a priori adjustment for the model-EM27/SUN comparison but not for the satellite-EM27/SUN comparison since the former is the scientifically most interesting aspect. Further, the satellite-EM27/SUN comparison suffers from errors due to imperfect coincidence which we expect to be larger than the smoothing effects.

References:

Hase et al., 2004: F. Hase, J.W. Hannigan, M.T. Coffey, A. Goldman, M. Höpfner, N.B. Jones, C.P. Rinsland, S.W. Wood, Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements, Journal of Quantitative Spectroscopy and Radiative Transfer, Volume 87, Issue 1, 1 August 2004, Pages 25-52, ISSN 0022-4073, http://dx.doi.org/10.1016/j.jqsrt.2003.12.008.

(http://www.sciencedirect.com/science/article/pii/S0022407303003765)

Butz et al, 2012: Butz, A., A. Galli, O. Hasekamp, J. Landgraf, P. Tol, and I. Aben, TROPOMI aboard Sentinel-5 Precursor: Prospective performance of CH4 retrievals for aerosol and cirrus loaded atmospheres, Rem. Sens. Env., 120, 267, doi:10.1016/j.rse.2011.05.030, 2012

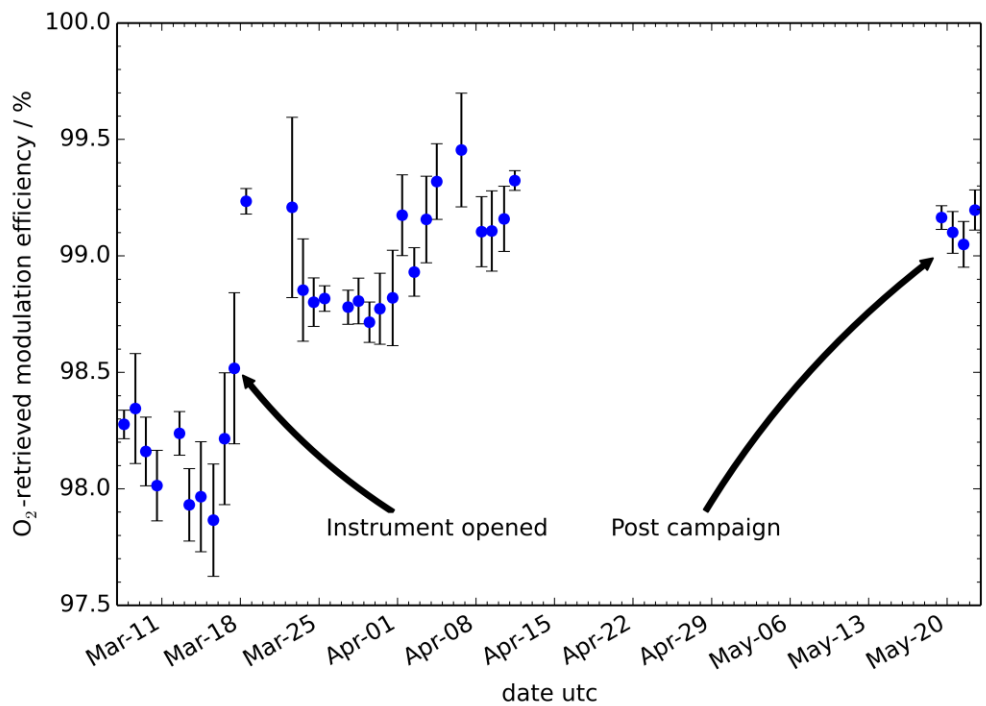

Figure 1: ILS parameter "Modulation efficiency" retrieved as a free fit parameter in a separate oxygen retrieval using a static water vapor apriori

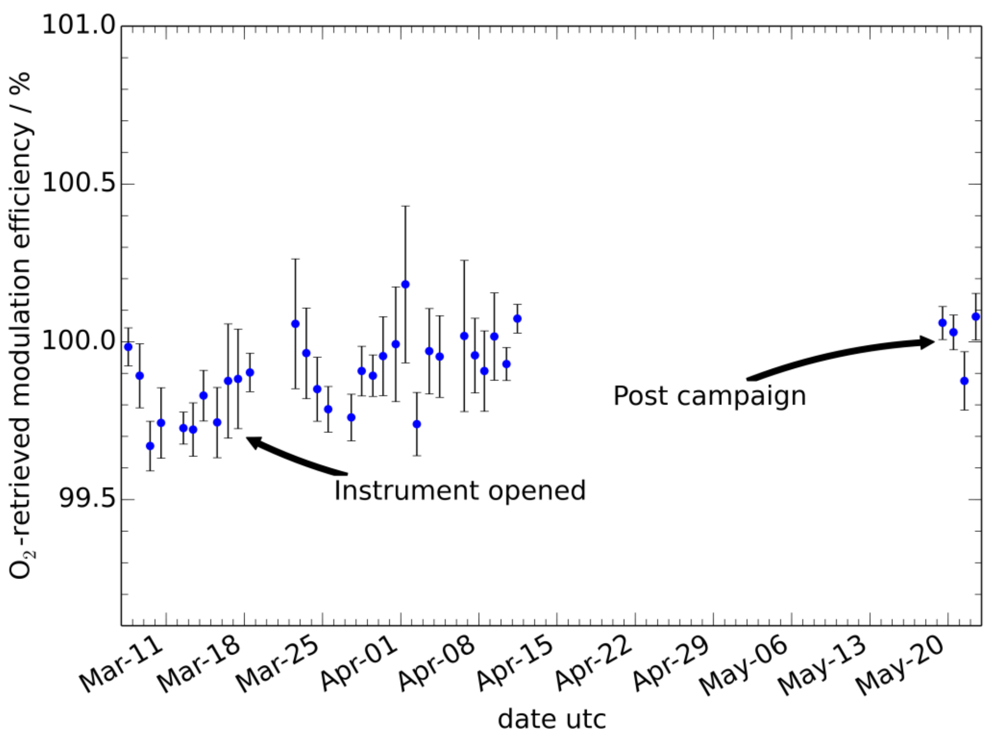

Figure 2: ILS parameter "Modulation efficiency" retrieved as a free fit parameter in a separate oxygen retrieval using a more realistic CAMS-apriori

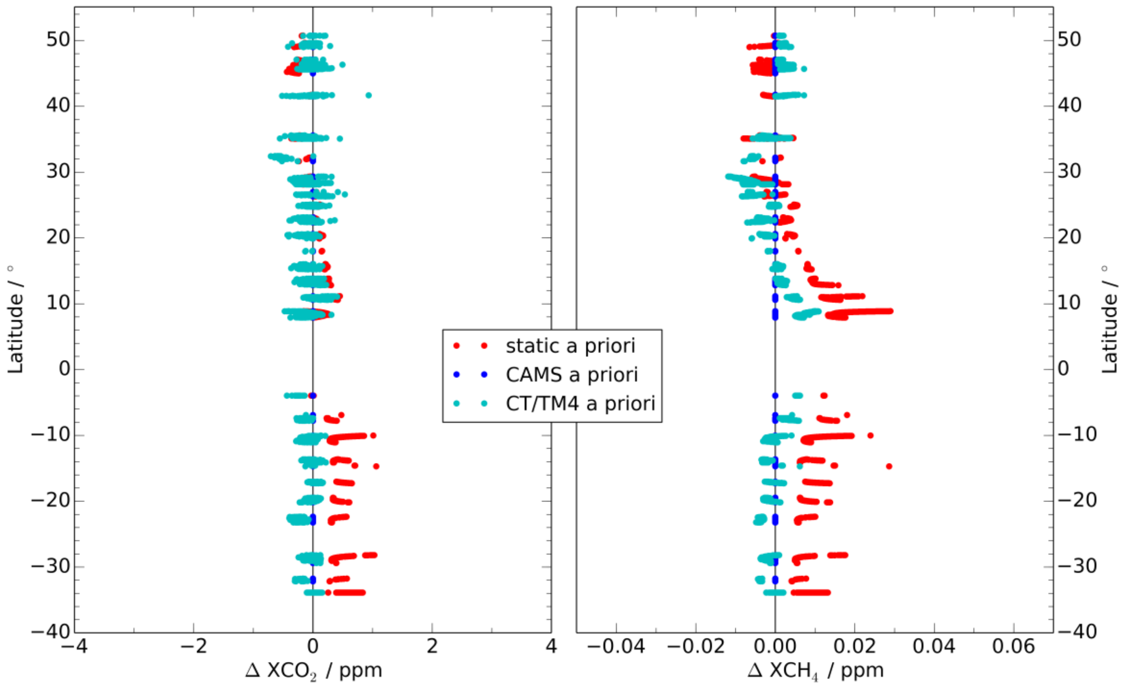

Figure 3: Dependence on retrieval products using various a priori profiles: A static, non temporal or spatial dependent a priori (red), the previous used Carbontracker (CT) (CO2,lat,lon,monthly mean) or Tracer Model 4 (TM4) (CH4,lon averages, monthly mean) a priori (light blue) compared to the updated retrieval using CAMS (lat,lon,8/day) dependent a priori. Note that the CT / TM4 a priori is a temporal and spatial average and for this reason we expect it to be less realistic than the high resolution CAMS model a priori.