Reply to interactive comment on Haszpra et al. "Variation of CO_2 mole fraction in the lower free troposphere, in the boundary layer and at the surface" by C. Gerbig (Referee#1)

The authors thank the Referee for his effort and constructive comments. Here are our responses to the comments.

General Comments:

1) It remains unclear why the authors used PBL heights from ECMWF, although temperature and moisture data were available for each flight. As the authors calculated PBL heights from these data for validation of the ECMWF PBL heights, I would recommend using these rather than model derived PBL heights. A bias of about 100m and 32% unexplained variance (rˆ2 of 0.68) seems significant and likely has an impact on the calculated mean CO2 mole fraction within the PBL, especially given that there are often strong vertical gradients in CO2 near the top of the PBL. This could also have an effect on the seasonal behaviour, as the bias might change throughout the year. It would also be interesting to know the standard deviation of the differences between aircraft data derived and ECMWF diagnosed PBL heights in addition to the bias.

On board temperature and humidity measurements were available only from November 2007, while the airborne measurements were started in 2001. For the consistency of the data evaluation it seemed reasonable to use the modeled PBL height data throughout the whole aircraft measurement period reported in the paper (2001-2008). Most of the transport models, which might profit from the present paper, also use model information for the stratification of the atmosphere. Comparison of the PBL heights obtained by the two different methods (ECMWF method, parcel method) is only a byproduct of the project and it does not form an essential part of the paper. We did not want to give special emphasis to the results as the method of comparison is arguable: the parcel method determines the PBL height for a well defined time (time of the measurements), while the ECMWF model gives the PBL height in 3-hour time steps and the linear interpolation in time between the model termini (page 11545, line 2-3) may not give the correct value for the time of the measurements because PBL does not evolve necessarily linearly. It may be one of the reasons for the bias between the PBL height data from the two different sources/methods.
We appreciate your suggestion to calculate and compare the mean CO_2 mole fraction for the PBL determined by the two different methods. For this exercise 15 flight days, a total of 25 ascending or descending profiles were available. The mean bias between the CO_2 mole fractions is 0.19 ± 0.72 μmol mol^{-1} (±1 σ). The correlation coefficient between the two data series is 0.993. These values indicate a statistically non-significant ($p>0.10$) deviation. It was not explicitly stated in the original manuscript that because of the potential large vertical concentration gradient around the top of the PBL the upper 10 % of the PBL was neglected in the calculation of the PBL-mean. This important information is inserted into the revised manuscript.

Taking into account that the evaluation of the ECMWF PBL model is not part of the main subject of the paper and the comparison presented in the original manuscript may be methodologically arguable it has been decided to completely remove this section from the revised version of the paper.

2) Traceability of profile data and tower data: when assessing the small difference between tall tower measurements and the mixed-layer mean derived from aircraft measurements, the accuracy of those measurements seem to be a crucial limitation. It needs to be made very clear, to what level each of the measurements are traceable to the WMO CCL scale, especially given the observed differences between co-located flask sample analysis and airborne profile data. In this regard, a plot comparing the lowest altitude of the profile measurements with those made at the top level of the tower should be presented.

The tower measurements are directly based on WMO CCL standards (page 11542, line 20-26). The working standards used for the on board in situ measurements were also decanted from WMO CCL standards (page 11544, line 5-9). Traceability of the standards used for the flask analyses was missed to be mentioned in the original manuscript but it is provided in the revised version (also WMO CCL). Following the suggestion of the Referee Section 2.3 (Validation of the in situ measurements) is completed with the comparison of the lowest aircraft measurements with the tower top ones. The additional text and a new figure show the results and discuss the similarities/differences including the analysis of three special cases. We think that this additional data evaluation has led to instructive results, and we thank again the Referee for his suggestion.

Detailed comments:

P 11544 L 6: When filling cylinders with calibration gas from a larger tank, there can be an impact on the CO2 mole fraction. Also, the mole fraction can change during time in the small tanks, for example through diffusion within pressure regulators. It should be described in detail how where the tanks, and if the mole fraction of the calibration gas in the small cylinders was compared to the original tank containing WMO CCL certified standard gases before and after use of the small cylinders for calibration of the airborne instrument. Note that any differences in the calibration of the aircraft instrument and the calibration of the instrument used for the tower measurements will result biased results for the assessment of differences of tower measurement and mean boundary layer mole fraction.

The CO_2 mole fractions in the small tanks ("field tanks") of the airborne CO_2 analyzer were compared with their mother cylinders (produced and certified by WMO CCL) before and after each refill. During this procedure the instrument's response function was determined using the certified mother cylinders temporarily replacing the low and the high field tanks. Then the high field tank of the instrument was replaced by the low mother cylinder having nominally the same mole fraction as the low field tank, while the *low field tank remained in its position in the instrument. In principle, the 'high-low' difference should have been zero in this case. This measurement was repeated with the high mother cylinder in the position of the high field tank. In this case the same signal was expected as during the normal operation. Finally, the low field tank was also replaced with the mother cylinders in the same way, while the high field tank was in its normal position. The results showed that the field tanks might deviate from the mother cylinders by a few tenth of* μ*mol mol*$^{-1}$ *(typically <0.3* μ*mol mol*$^{-1}$*) but it was not higher than the assumed field accuracy of the instrument. As the deviation was small and the temporal courses of the occasional drifts were uncertain no correction was applied on the raw data. The above intercomparison procedure is shortly described in the revised version of the paper.*

P11544 L 15: It is unclear how the comparison of the ascending and descending profiles can provide information on both, the changes in the atmosphere and on the performance of the instrument. The authors should explain how they distinguished the influence of these two on differences in the profiles.

Occasionally our instrument tended to record mole fraction decreasing in time without obvious signs of malfunctioning in the instrument parameters. In such a case the record showed decreasing mole fraction during the ascend and continuing decrease during the descent resulting up to 10 ppm difference at the bottom of the profile within short time without any atmospheric physical reason. These profiles were disqualified. Changing altitude of concentration jumps in the ascending and descending profiles might reflect actual atmospheric processes like increasing depth of the convective boundary layer. The sentences are rephrased in the revised version to be clearer for the reader.

P11544 L 22: It should be explained how the meteorological information was used to evaluate if the measured mole fraction profiles were realistic. In order to exclude erroneous data usually additional information on instrument parameters such as pressures and temperatures are used, otherwise a sampling bias might result when specifically assessing "unrealistic" data.

When no tower data were available, especially in wintertime when the PBL CO_2 *mole fraction varied in a wide range depending on the meteorological conditions (accumulation in the PBL), the meteorological conditions were also checked if the airborne measurements were within a realistic (expected on the basis of experience) range. As the present paper contains only those aircraft profiles when tower data were also available the sentence has no actual meaning here and has been omitted from the revised version.*

P111546 L25: Regarding the non-linearity: Was a multi-point calibration performed repeatedly at least on ground? Was the non-linear component of the calibration curve changing over time? When using a linear correction to the "raw in situ airborne measurements”, is there a problem with not capturing the non-linear component? As the linear correction is based on the comparison with flask data, does this mean the information from the in-flight calibrations was not used at all in the reported mole fraction data? This should be clarified. I would expect that e.g. the offset of the signal changes significantly during the time between flask samples, so that at least the information from “zero” checks contain valuable information. Also a scatter plot showing the comparison of in situ and flask CO2 mole fraction would be helpful.

Unfortunately, we had no technical possibility to check the non-linearity of the instrument and the manufacturer was not supportive in this business either. Therefore, the non-linearity, as the potential cause of the bias experienced, is only mentioned in the manuscript as a hypothesis. The instrument two-point calibration was used to set the scale and follow the drift of the instrument. The empirical linear correction was applied on the ‘calibrated’ data to be consistent with the flask and tower measurements. In the revised manuscript there is a scatter plot inset in Figure 2 showing the relation between the flask data and the uncorrected in situ data.

P11548 L23: the comparison of measurements made at the top level to mixed layer averaged mixing ratios reminds me a bit on the “virtual tall tower” concept by K.J. Davis (unfortunately I could not find the corrresponding publication detailing this). In that concept, CO2 measurements made at around 10 meters above ground are corrected for a vertical gradient to represent CO2 measurements from a tall tower. May be the authors can discuss this, and may be think about a similar correction that turns tall tower based CO2 into mixed layer averaged CO2.

Davis’ virtual tower concept was considered during the preparation of the original manuscript and it was decided that this topic deserves an individual, separate paper. Preparation of this paper is in progress.

P11549 L10: replace “largest” with “larger”

Corrected

P11551 L6: the authors probably meant “lay” instead of “lied”

Corrected

P11552 L10: From figure 5 it appears that the seasonal minimum is around day 230, which is more August 20 than July 20.

Unfortunately, day 230 was wrongly translated to calendar date. Corrected.

Fig 3 figure caption: please rewrite the first sentence

We have done our best to include all essential information for the interpretation of the figure still keeping the sentences readable.

Fig 7: please add “CO2” to the axis and the legend

Corrected