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 A. Andrews (Referee#2)

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

1. Units. The units micromole per mol are used throughout, instead of the more common and compact parts per million (ppm). If allowed by this journal, consider stating that units are reported as a dry air mole fraction, and then using "ppm".

According to the ACP guideline SI units have to be used in the papers. Although, ppm is widely accepted and used by atmospheric scientists including us, it is not an SI unit. That is why we use micromole per mole throughout the paper. If the editor does not oppose it we are willing to change for ppm.

2. Pg 11541 line 2. Authors should contact Dr. Colm Sweeney from NOAA ESRL regarding the preferred citation for the North American aircraft network.

Reference suggested by Colm Sweeney has been inserted into the revised manuscript.

3. Pg 11544 first paragraph. I agree with review by Christoph Gerbig that decanting standard gases from large cylinders into small flight tanks can cause concentration changes in both tanks of order a few tenths of a ppm. This can be detected by calibrating both fill tanks and flight tanks against an independent set of standards. Line 9. It seems that the word "completed" should be replaced by "complimented". Also, consider using "baseline" in place of "zero"and "analyzer-drift" in place of "scale-drift".

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. For the details of the procedure, please, see our response to Christoph Gerbig. The results showed that the field tanks may deviate from the mother cylinders by a few tenth of ppm (typically <0.3 ppm) 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 intercomparison procedure is shortly described in the revised version of the paper. The comparison of the mother tanks with an independent set of standards is a reasonable suggestion. Unfortunately, we did not think about it during the project. However, the mother tanks were returned to the WMO CCL at NOAA after the project for recalibration. The recalibration did not indicate any significant change therefore missing of the regular intercomparison did not distort the measurements. "Zero" and "zero check" have been replaced by "baseline" and "baseline check" in the revised *manuscript. "Completed" has been replaced by "complemented" in the revised manuscript.*

4. Pg 11544 2nd paragraph. Flights were performed in late morning-early afternoon. I expect that the PBL might still be developing through early afternoon and that the best time to fly would be mid to late afternoon. Perhaps add a few sentences about how this time was selected? Perhaps discuss magnitude of bias e.g., if profiles were obtained earlier than the time that the PBL reached its maximum height on a particular day.

Mostly for logistic reasons occasionally the measurements were performed before the formation of the fully developed PBL. These occasions occurred rather randomly, therefore they might not influence the calculated seasonal variations systematically. The occasional early flights were accepted supposing that sampling in different times might give additional insight to the atmospheric processes. In the present paper we handled the PBL and the free troposphere rather separately, therefore, the actual height of the PBL, the phase of its development within the day, does not seem crucial. Obviously, the absolute concentration in the PBL depends on the height of the PBL because the surface flux is diluted in smaller or larger volume. We have added a few sentences on the flight time in the revised manuscript.

5. Pg 11545. It would be useful to state the mean PBL height so that the bias could be interpreted as a percentage. Even better would be to compute the percentage bias for each profile and then report the mean of that in addition to the mean absolute bias.

Comparison of the PBL heights obtained by the two different methods (ECMWF method, parcel method) was only a byproduct of the project and it did 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. Following the suggestion of the other Referee, Christoph Gerbig, we calculated and compared 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 ppm ($\pm1\ \sigma$). The correlation coefficient between the two data series is 0.993. These values indicate a statistically non-significant ($p>0.10$) deviation, which can be considered as random error.
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.

6. Pg 11546-47. When interpreting flask vs in situ differences, it would be helpful to consider the standard deviations corresponding to the 2-min averages. In that case, instead of a histogram, it would be useful to present the differences as a scatter-plot with error bars corresponding to the std dev of the in situ data. Line 22: 0.34 ppm bias: which is higher flasks or in situ? Information about the standard deviation of insitu measurements during calibrations versus atmospheric samples could inform about what portion of the variability is geophysical vs instrument-related. Consider replacing "The most likely cause could be : : :" with "A possible cause is non-linearity: : :"

Flask vs. in situ difference can be analyzed and presented in several different ways but the size of the paper is limited, as well as we tried to keep the balance between the methodological section and the results, the presentation of which the main aim of the paper was. Nevertheless, we have added a new figure to the manuscript presenting the relation of the flask data and the uncorrected in situ data supporting the empirical linear correction applied. The standard deviation of the 2-min averages is <0.1 ppm. 5-min averages and standard deviations were also calculated for test because flushing and pressurizing of the flasks took so long time. The typical standard deviation was <0.4 ppm, but in the case of the outlier discussed in the paper it was as high as 3.08 ppm. It calls the attention to the problem of the synchronization between the flask sampling and the in situ measurements. As an overall average, the in situ measurements exceeded the flask samples by 0.34 ppm. Now, it is clearly stated in the revised manuscript. "The most likely cause could be" is replaced in the revised manuscript as it is suggested.

7. Pg 11549. For those flights where on-board data are used to estimate PBL height, what is the difference in mean PBL CO2 concentration when the observed PBL is used compared to ECMWF PBL height?

As it was mentioned above at item 5 we calculated and compared the mean CO_2 mole fraction for the PBL determined by the two different methods (ECMWF vs. parcel method). 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 ppm ($\pm1\ \sigma$). The correlation coefficient between the two data series is 0.993. These values indicate a statistically non-significant ($p>0.10$) deviation.

8. Section 3.1. It seems that the CO2 gradient observed across all heights on the tower might provide information about which days the 115m level represents a reasonable estimate of the PBL mean values.

It might be assumed that the lower the gradient along the tower the more accurately the PBL mean is estimated by the top of the tower, but it is a rather uncertain assumption because of possible stratifications above the tower within the PBL.

9. Pg 11551/Figure 5. Consider adding data from a high altitude site like Jungfraujoch to compare with the >2500m curve in Figure 5.

Jungfraujoch has had continuous CO_2 measurements only since 2005, while our measurements started in 2001 and ended in January 2009. The 3 years overlap is not enough for the reliable comparison of the mean seasonal cycles but as a qualitative comparison the mean seasonal cycle at Jungfraujoch based on the period of 2005-2011 is also presented on the revised Fig. 5. (Fig. 6. in the revised manuscript).

10. Pg 11552-53: The analysis of fair-weather bias using the 115m tower data (Figure 7) is interesting and robust, but I don't understand the basis for assuming that fairweather bias at higher altitudes is insignificant. Consider a case where cloudy/foggy conditions prevent flying. Such conditions may be associated with particular synoptic conditions (such as Northerly flow bringing cold air). Given the large-scale latitude gradients that exist in the free-troposphere, it seems rather likely that there could be a fair-weather bias aloft. The likelihood and magnitude of fair weather bias at high altitudes could be studied using continuous output from a model with reasonable fluxes and realistic weather (e.g. CarbonTracker or similar).

Usually persistent foggy/hazy weather in the Carpathian Basin in winter is associated with anticyclonic conditions. It might bias the composition of the lower free troposphere indeed as the Referee has called our attention. This phenomenon, its magnitude must be studied indeed in cooperation with a research group having access to and experience with transport models. For the present manuscript we have reformulated the critical, speculative sentences removing the statement about the upper layers. We also mention the need for 3D transport model results as a possible future research direction.

11. These papers are highly relevant, and form some of the basis for the Virtual Tall Tower concept that was also mentioned by reviewer Christoph Gerbig:

Yi, C. X., K. J. Davis, B. W. Berger, and P. S. Bakwin. "Long-Term Observations of the Dynamics of the Continental Planetary Boundary Layer." Journal of the Atmospheric Sciences 58, no. 10 (2001): 1288-99.

Yi, C., K. J. Davis, P. S. Bakwin, A. S. Denning, N. Zhang, A. Desai, J. C. Lin, and C. Gerbig. "Observed Covariance between Ecosystem Carbon Exchange and Atmospheric Boundary Layer Dynamics at a Site in Northern Wisconsin." Journal of Geophysical Research|Journal of Geophysical Research 109, no. D8 (2004): 9 pp.

Thank the Referee very much for calling our attention to the papers of Yi and his coworkers. Inclusion of Ken Davis' virtual tall tower concept was considered during the preparation of the original manuscript, but it was agreed that this topic deserves an individual, separate paper. Preparation of this paper is in progress.