We would like to thank Referee #1 for his/her thoughtful comments and detailed suggestions to our manuscript. In the following, we answer to the reviewer's comments and indicate the changes in the manuscript that were implemented as a consequence of the recommendations. The comments are in black and italic. Our answers are in blue and plain text.

Anonymous Referee #1

Received and published: 5 August 2019

Review of Lian et al. (2019) Analysis of temporal and spatial variability of atmospheric CO_2 concentration within Paris from the GreenLITE™ laser imaging experiment.

Lian et al. describe the application of a long open-path spectroscopy technique for the measurement of CO_2 mixing ratios above a complex urban canopy, which could influence existing emissions estimates at the city scale. They compare the data measured using the GreenLITE™ system with fixed-site CO_2 measurements within the same urban environment, and contrast the results against two urban canopy schemes within the WRF-Chem model.

It is a well written paper which I would recommend for publication in ACP. The content of the paper, which covers greenhouse gas measurements with a possible climate change impact, is relevant to the journal and its readers.

We thank the reviewer for these very supportive comments.

General comments:

The authors acknowledge that calibration of long open-path spectroscopy techniques is difficult. A separate paper (Zaccheo et al., 2019), detailing a new calibration procedure applied to the GreenLITE™ data, is referenced by the authors here. This calibration procedure appears to use the fixed-site installations within the city to calibrate the open-path data. Whilst the authors state this "has no significant impact on chord-to-chord variations", they do not discuss the potential implications of using point-source measurements to adjust area/path averaged measurements. Zaccheo et al. (2019) does go into more detail but considering this is a key element of the calibration procedure, I believe it needs some more attention here.

We thank the reviewer for raising this question of data calibration. As he/she points out, the calibration procedure is described in some detail in (Zaccheo et al., 2019) so that we felt there is no need to go into the same level of detail. As it has been addressed in Zaccheo et al. (2019), while not desirable, it is often necessary to apply post-calibration corrections to such data to rectify residual differences between observation types. We acknowledge nevertheless that the reader may want to see more, and we therefore have provided more information about this calibration method as well as its limitation in the revised version of the manuscript:

"These slowly time-varying differences were most likely due to a slight systematic long-term drift in both the on- and off-line wavelengths as a function of continuous operations. Such drift may induce some non-linear impacts on the measured concentrations. It is therefore more appropriate to adjust the wavelengths rather than to apply a linear calibration to the retrieved concentrations. Unlike in-situ point measurement systems, there is no established method for calibration of long open-path systems to the WMO mole fraction scale used as an international standard for atmospheric CO_2 monitoring (Tans et al., 2011). Therefore, a bias correction method was developed by AER (Zaccheo et al., 2019) for addressing observed slowly drifting biases between the GreenLITE™ prototype system and the two in-situ sensors (CDS and JUS) that are near the GreenLITE™ chords. This method computed a time-varying adjustment to the offline

wavelength based on a non-linear optimization mechanism. This non-linear approach adjusts the GreenLITE™ offline wavelength considering not only the average values of hourly CO_2 concentrations at two in-situ stations, but also the corresponding average temperature, relative humidity, atmospheric pressure along the chord and an optimized online wavelength value during the measurement period. Finally, the median on- and off-line values over a 4-day window was used to recompute the GreenLITE™ data from all chords using a radiative transfer based iterative retrieval scheme based on the LBLRTM model (Clough et al., 2005). Even though this approach is not ideal as the two in-situ stations and the GreenLITE™ system do not sample the exact same area, it does provide a well-defined mechanism that reduces the systematic long-term biases with no significant impact on the chord-to-chord variations."

As most of our analyses focus on the spatial gradient on the concentrations, we feel that the important point of the calibration procedure is that it has no significant impact (and thus no significant uncertainty) on the spatial gradient between chords. (See below, our answers to the comment: Page 4, Line 23: What is meant by "no significant impact" – significant in what way?)

Font sizes in some figures could be larger. Some of the text is hard to read on a computer screen without zooming in.

This suggestion is well taken. We have increased the font size in all figures (see Figure 1 to 7 in the revised manuscript).

The authors should address the following points in a revised manuscript:

Page 3, Line 30: What is meant by 15/100 m above ground level? I assume there are two sampling inlets? This is not made clear.

Yes, the atmospheric CO_2 concentrations at SAC station are measured with two sampling inlets at 15 m and 100 m above ground level, on a tall tower at that location. This has been clarified in the text:

"OVS site is located about 26 km southwest of Paris center with the sampling height of 20 m above the ground level (AGL) on the top of a building. The SAC tall tower is located on the Plateau de Saclay (9.5 km southeast of OVS) with two air inlets placed at 15 m and 100 m AGL respectively."

In addition, we also changed "15/100" to "15 and 100" in Table 1.

Page 4, Line 23: What is meant by "no significant impact" – significant in what way?

We agree that the term "no significant impact" is an overly vague statement and should be clarified more precisely. We have added a figure in the revised supplement (Figure S2). It shows the distributions of the original and re-processed GreenLITE™ absolute CO_2 concentration differences between all pairs of chords for each transceiver. The differences between the medians of the inter-chord range of the re-processed and original data are within in the range of ±0.5 ppm for T1 and ±2 ppm for T2 with the respective yearly mean plus/minus one standard deviation of 0.04 ± 0.16 ppm for T1 and 0.48 ± 0.43 ppm for T2, which are relatively small. The modified text in the revised manuscript is as follows:

"Top panels in Figure S2 (a) and (b) show the distribution of the absolute values of the daily averaged CO_2 concentration difference between all pairs of chords for each transceiver before and after the calibration. The differences between the medians of the re-processed and original inter-chord range, shown in bottom panels, are within in the range of ±0.5 ppm for T1 and ±2 ppm for T2 with the respective yearly mean plus/minus one standard deviation of 0.04 ± 0.16 ppm for T1 and 0.48 ± 0.43 ppm for T2."

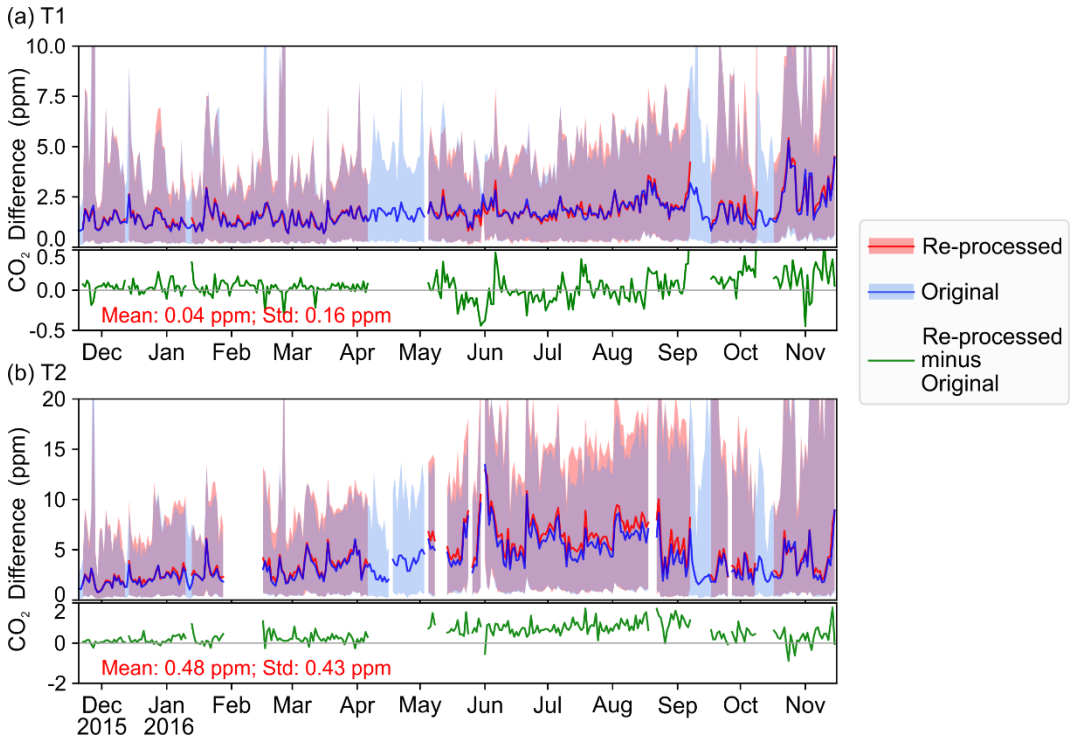

Figure S2: Distribution of the original and re-processed GreenLITE™ absolute CO_2 concentration differences between all pairs of chords for (a) T1 and (b) T2 from December 2015 to November 2016. The solid lines in top panels of (a) and (b) indicate the 0.5 quantile, and the shaded areas represent the 0.1 and 0.9 quantile intervals for original data in blue and re-processed data in red. The green line in bottom panels of (a) and (b) indicates the differences between the median values of the re-processed and original inter-chord range.

Page 4, Line 27: Why is a threshold of standard deviation < 10 ppm CO_2 applied to the hourly data? Is this because CO_2 is not expected to change by more than 10 ppm over the course of one hour? Is this justifiable?

The outlier detection for the 4-minute GreenLITE™ data is mainly based on the 3-sigma rule, which is used to remove the data outside three standard deviations from a mean in the positive direction. We have added a sentence in the main body of the revised manuscript and Figure S3 in the supplement to answer the reviewer's comment:

"The 10 ppm threshold was selected to be rough 3 times the typical standard deviation of the 4-minute measurements for any given chord within a one-hour period (Figure S3)."

For better clarity, we have also added the following statement in the supplement together with Figure S3.

"The outlier detection for the 4-minute GreenLITE™ data is mainly based on the 3-sigma rule, which is used to remove the data outside three standard deviations from a mean in the positive direction. Figure S3 (a) shows the frequency distribution of the standard deviations of the 4-minute CO_2 concentrations measured within one hour for one given chord (e.g. T2R08). Figure S3 (b) shows the three-sigma threshold (mean + 3σ) of the standard deviations of the 4-minute measurements within a one-hour period for each chord. In general, the threshold varies between 6.5 ppm and 11.9 ppm from chord to chord. We therefore choose to use a uniform threshold value of 10 ppm to remove the outliers for all chords."

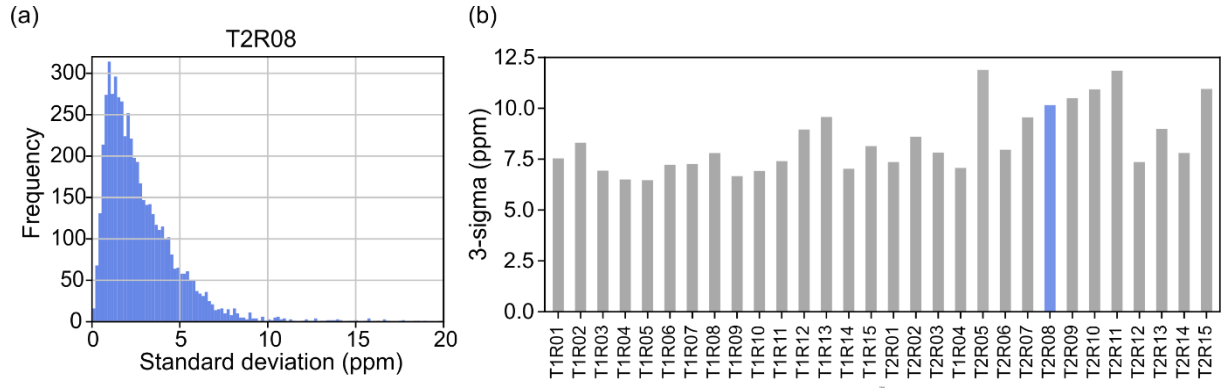

Figure S3: (a) Frequency distribution of the standard deviations of the 4-minute CO_2 concentrations measured within one hour for one chord (e.g. T2R08); (b) Three-sigma threshold (mean + 3σ) of the standard deviations of the 4-minute measurements within a one-hour period for each chord.

Page 5, Line 19: Can you quantify "much larger differences"?

For better clarity, we have added the following sentence in the revised manuscript:

"In order to select an adequate model physical configuration for Paris, we carried out some preliminary sensitivity experiments to test the impact of different physical schemes on the simulated CO_2 concentrations. These tests use up to five different PBL schemes and two urban canopy schemes. The simulations were carried out for two months, including one winter month (January 2016) and one summer month (July 2016). These preliminary sensitivity results indicate that different PBL schemes in the WRF-Chem model lead to monthly average differences of 2-3 ppm on the simulated CO_2 concentrations over Paris, whereas the two different urban canopy schemes lead to much larger differences of 8-10 ppm. Thus in this study, we carried out the 1-year simulation with two different urban canopy schemes as they are sufficient to address the paper main question regarding the ability of a configuration of the WRF-Chem model to simulate the CO_2 atmospheric transport in an urban environment, but also to provide an estimate of the modeling uncertainty. All of the other physics options remained the same for the two experiments (Table 2)."

Page 5, Line 29: Typo - "details" should be "detail".

Correction made.

Page 6, Line 5: Consider "accounting for" rather than "taking up"?

Text changed as suggested.

Page 11, Line 9: Rephrase "city surrounding" to "areas surrounding the city", or similar.

Text changed as suggested. The modified text is as follows:

"On the other hand, both models show similar performances in the areas surrounding the city."

Table 3: What are the colour scales showing? Better or worse values? This needs to be made more clear particularly because high correlation coefficient (red) is good but high RMSE (also red) is bad?

We agree with the reviewer that the color scales in Table 3 can be misleading. The color only represents the values from minimum (blue) to maximum (red) in the cells instead of indicating the goodness of fit between model and observation. We have added the following text in the caption of Table 3 in order to clarify this issue:

"The color highlights the value in the cell with the minimum in blue, the median in white and the maximum in red. All other cells are colored proportionally."

Figure 1: Some text is very small – a possible solution would be to refer the reader to the panel in Fig 2 in the caption and remove the chord labels. Also the caption refers to Figure S1 but this doesn't appear relevant to the text – the authors might mean Figure S2?

For better clarity, we have added a second panel in Figure 1 and noted the previous Figure 1 as Figure 1a. Now, Figure 1a shows the distribution of in-situ CO_2 stations and the GreenLITE™ laser system without the chord labels. Figure 1b is a high-resolution zoom of the inner Paris area and shows the GreenLITE™ laser system layout in detail.

Corrected, thanks. It should refer to Figure S5 in the revised manuscript.

Figure 2: Does the caption need to state that these emissions are taken from an emissions inventory i.e. not measured or modelled.

Yes. We have added the following sentence into the caption to stress this point:

"Figure 2: Total CO_2 emissions, according to the AirParif inventory (within IdF) and the IER inventory (outside IdF), for a weekday in March 2016."

Figure 3: See Figure 2.

We have modified the caption:

"Figure 3: Averaged anthropogenic CO_2 fluxes along each GreenLITE™ chord according to the AirParif inventory."

Figure 4b: Is there some way of better highlighting that this is not a continuous time series of data? Perhaps either thicker/bolder lines or a gap between each monthly diurnal cycle.

As suggested we have used thicker lines between each monthly diurnal cycle in the revised Figure 4b.

Figure 5: The blue (observation) line is quite difficult to see on these plots.

We have changed the line colors to a sharper contrast between the observation and the model results. (PS: following the recommendation from Referee #3, we have moved this figure to the supplement as Figure S7 in the revised manuscript)

Figure 7: y-axis titles should probably read "CO_2 difference (ppm)" as in Figure 6.

Changed as suggested. This figure is now Figure 6 in the revised manuscript.