Responses to Referee 2:

Thank you very much for your significant and useful comments on the paper "O_2:CO_2 exchange ratio for net turbulent flux observed in an Urban Area of Tokyo, Japan and its application to an evaluation of anthropogenic CO_2 emissions" by Ishidoya et al. The title of the paper has been changed from the ACPD paper. We have revised the manuscript, considering your comments and suggestions. Details of our revision are as follows;

Primary scientific concern

One of the major challenges of working with tower data is determining the region of influence (the "footprint") for the tower. Calculations of the OR fromO_2-CO_2covariation are particularly challenging, since the lower-frequency data from a paramagnetic analyzer lend themselves to aggregating data over extended periods. The authors acknowledge this in lines 133-135. However, the problem in this analysis is more profound than simply scaling footprints (inversely) by data-rate. This is because the OR slopes shown in Fig. 4 (lower panel) include data from the entire 18 month set of observations. Consequently, this is effectively a global average number with local influences superimposed.

To understand this, first consider a point in the plot with very low O_2 (and high CO_2). Maybe this parcel started with relatively high O_2 and was influenced by a great deal of local combustion. OR maybe it's part of an air mass that arrived from some dis- tant location (highly influenced by combustion) and was relatively unaffected by local fluxes. Compare this to a point with relatively high O_2 (and low CO_2). If this point was measured hours before the low-O_2 one, and the wind pattern was roughly constant, chances are good that O_2 fell due to local combustion. In contrast, if this high-O_2 point was measured days (or months) before, it might have come from a totally different re gion and the difference from the first point reflects local influences to a much smaller degree. One solution is to choose much shorter aggregation periods when determining OR_{atm}.

In short, all of the analysis of OR_{atm}, and the comparisons of OR_{atm} with OR_F need to be reconsidered.

For this reason, I will not comment further on the parts of the manuscript that involve

the interpretation of OR_{atm}.

Considering your comments, we have reconsidered the comparison of OR_{atm} with OR_F. We have added the sentence not only to state the problem to use OR_{atm} but also to clarify the purpose of the comparison (line 185-190). The comparisons by choosing 12-hour aggregation period have also been added (line 191-210 and Fig. 6). The discussion for the comparisons by choosing 1-week aggregation period has been modified, and we have concluded to use OR_F rather than OR_{atm} is more appropriate to validate inventory-based CO_2 emissions from gas, liquid and solid fuels in the flux footprint (line 211-252). Moreover, we have newly added discussion to estimate the average diurnal cycles of CO_2 fluxes from gas and liquid fuels consumption separately by using the OR_F, CO_2 flux, and inventory-based CO_2 emission from human respiration, in order to validate the inventory-based CO_2 emissions from gas consumption and traffic (line 290-344, Fig. 9 and Fig. 10). The inventory-based emission data have been updated from Hirano et al. (2015) for the present study.

Other scientific concerns:

1) Line 30: In addition to Mitchell et al., please cite Sargent et al., PNAS 2018.

Line 34: We have cited Sargent et al. (2018), as suggested.

2) Lines 59-60: Does the vegetated area actually change seasonally? Or is it that the vegetation is mostly dormant in the winter?

Line 66: The sentence has been modified as "The flux footprint includes vegetated area of 9% in the summer and 2% in the winter, reflecting seasonal changes in the wind direction." As seen in Fig. 1, the vegetated area included in the flux footprint actually change seasonally due to seasonal changes in the wind direction. It is noted that calculation of the flux footprint has been updated by using the model of Neftel et al. (2008) (line 63).

3) Lines 71-72: As I understand it, the samples are measured with a paramagnetic analyzer relative to secondary standards. It's the secondary standards that are measured against the primary standard with a mass spectrometer. This is not what this sentence says.

Lines 76-79: That is as you pointed out. We have changed the phrase as "In this study, $\delta(O_2/N_2)$ values of each air sample were measured with the paramagnetic analyzer using working standard air that was measured against our primary standard air (Cylinder No. CRC00045; AIST-scale) using a mass spectrometer (Thermo Scientific Delta-V) (Ishidoya and Murayama, 2014)."

4) Lines 73-75: Air is being drawn down at 10l/m and a very small subset of that airstream is being analyzed. There is no mention here of the possibility of fractionation at this sampling, of tests to detect fractionation, nor of measures to prevent it. This is something that Stephens et al. (DOI: 10.1175/JTECH1959.1) discusses extensively. Perhaps this is discussed in the original methods paper, but it should at least be mentioned here.

Lines 80-87: We have added the sentences to discuss the possible fractionation for the measurements in this study.

5) Lines 75-76: If air is measured first at one height, then the other, and air is measured for 10minutes at each height, isn't each measurement cycle 20minutes long (and thus, 9 cycles is 180 minutes)?

Lines 87-88: The phrase "After 9 measurement cycles (90 minutes)" has been changed to "After 9 cycles of measurements (5 and 4 cycles for 37 and 52 m, respectively)" to clarify the meaning.

6) Line 79: How is a correction made for Ar? The paramagnetic analyzer doesn't measure this species. Again, this might be presented in 2014 Tellus paper, but a few words of explanation here would be welcome.

Lines 90-93: We have modified the sentence as follows to explain the correction method briefly "The dilution effects on the O_2 mole fraction measured by the paramagnetic analyzer were corrected experimentally, not only for the changes in CO_2 of the sample air or standard gas measured by the NDIR, but also for the changes in Ar of the standard gas measured by the mass spectrometer as $\delta(Ar/N_2)$.".

7) Line 83: Why are uncertainties being quoted for 30minute averages when atmospheric measurements are only made for 10-minute intervals on each intake, and standards are measured for 5-minute intervals.

Lines 93-94: The sentence has been modified to show the analytical reproducibility for 2-minute average only.

8) Line 91: What does "span-difference" mean? Please clarify.

Lines 102-105: We avoid to use the word "span-difference" and changed the sentence as follows "Although the highest CO_2 concentration of the gravimetrically standard of the NIES-09 scale is similar to that of the TU-10 scale, a slope of 0.974 ppm ppm^{-1} is derived from a least-squares regression line fitted to the relationship between the CO_2 concentrations observed by NDIR on the TU-10 scale and those by CRDS on the NIES-09 scale with a correlation coefficient (r) of 0.978.".

9) Lines 114-115: Downward excursions in O_2 may be due to consumption within the canopy, or non-local influences being transported to the tower. If they coincide with positive excursions in deltaO_2, then I would be convinced that the cause is consumption within the local canopy, but until you show that the two excursions are coincident, you can't claim local consumption is the cause.

Lines 127-132 and Fig. 3: The sentences and figures have been added to show the two excursions are coincident.

10) Line 154: If errors in both species are non-negligible, a standard least-squares linear regression will give the wrong slope. Instead a Deming regression is required (which reduces to an orthogonal fit in the case of equal areas).

Lines 170-178: We have changed the regression method to Deming regression throughout the paper for calculating OR, as suggested.

11) Lines 183ff: A very basic back-of-the envelope calculation would be appropriate here to indicate whether human respiration really was utterly negligible or not. For example, the population density given for this area is 0.016 people m^{-2}. If each requires 2000 kcal/day, this could be supplied by metabolizing 3.34 moles of glucose, with a resulting consumption of 3.7μmol$m^{-2}s^{-1}$ of atmospheric O_2. This seems to be about

20% of the smallest values quoted on line 232: A modest, but non-negligible correction to the results presented here.

Lines 290-344, Fig. 9 and Fig. 10: We have added discussion to estimate the average diurnal cycles of CO_2 fluxes from gas and liquid fuels consumption separately by using the OR_F, CO_2 flux, and inventory-based CO_2 emission from human respiration, in order to validate the inventory-based CO_2 emissions from gas consumption and traffic. The inventory-based CO_2 emission from human respiration is close to the value in your comments.

Minor editorial comments:

1) Line 44: Change to "In this paper, we first present the. . ."

Line 48: The words "In this paper, we present firstly the…" have been changed to "In this paper, we first present the…", as suggested.

2) Line 74: should read "and 37m was introduced"

Lines 80-95: The sentences, including the words that were pointed out, have been rewritten.

3) Line 75: should read "100mL min^{-1} with the pressure stabilized to 0.1 Pa and measured"

Lines 83-84: The words have been changed, as suggested.

4) Line 85: should read "We used the gravimetrically prepared air-based"

Line 96: The words have been changed, as suggested.

5) Line 86: should read "1991) to determine"

Line 97: The words have been changed, as suggested.

6) Lines 87 and 90, "gravimetrically standard" should be replaced with "gravimetrically prepared standard"

Lines 98 and 101: The words "gravimetrically standard" have been replaced with "gravimetrically prepared standard".

7) Line 107: should read "activities. In contrast, the atmosphericO_2"

Line 119: The words "On the other hand" have been changed to "In contrast", as suggested.

8) Line 111: should read "Therefore, we attribute the opposite phase" and "in this study mainly to fossil"

Lines 123-124: The sentence has been modified as suggested.

9) Line 124: Remove "by"

Line 140: The word "by" has been removed.

10) Line 131: End the sentence with "troposphere" and simply remove "whereas. . ."

Line 147: The words "in the troposphere, whereas it is…" have been changed to "in the troposphere. It is…".

11) Line 134: should read "1994). We note that"

Line 150: The words "1994). It is noted" have been changed to "1994). We note that".

12) Line 204: should read "standard error ($\sigma/\sqrt{n}$)" (i.e. use symbols instead of writing it out).

Line 258: The words have been changed, as suggested.

13) Line 205: should read "negative values respectively, indicating"

Lines 258-259: The words have been changed, as suggested.

14) Line 206: end the sentence with "the year." and remove "respectively".

Line 260: The words have been changed, as suggested.

15) Figure 6: There is no legend explaining the filled and unfilled symbols in the upper panel.

Figure 8: The words to explain the filled and unfilled circles in the upper panel have been added to the figure caption. It is noted the number of the figure has been changed from that in the ACPD paper.