Reply to Reviewer 1

We sincerely thank the reviewer for carefully reading our manuscript and for providing valuable and constructive comments and suggestions, all of which we have accepted. Below, we present our point-by-point responses to each comment.

Estimation of CO2 and other greenhouse gases from megacities is an important research topics, and requires set of observations at strategically located sites using suitable instruments or measurement systems. This study discusses measurements of CO2 using two types of CO2 measurement systems and comes up with useful conclusions about the usefulness of tier-2 instrument (often referred to as low-cost sensor or LCS) for tracking emissions of CO2. The results are supplemented using Max-DOAS NO2 and BC measurements. The article is very well written and a little more efforts in improving the Abstract and Introduction are needed to properly reflect the outcomes of this study and also to put this study in perspective of growing number of literature for this area of study. I recommend a major revision before accepting for publication in AMT, but I feel the works involved are not difficult to conduct. Details below.

Reply: We appreciate your positive assessment.

Line 14: Not sure if this is the right terminology - may be Greater Tokyo area ? and please explain a bit what you mean by this.

Reply: In response to the comments, we revised the manuscript to use the term 'Greater Tokyo Area' as a more appropriate terminology.

Line 28-29 : reference needed, I feel

Reply: In response to the comments, we added the reference of Li et al. (2026).

Line 30 : Definition needed, e.g., Greater Tokyo includes Tokyo Metropolis (includes 23 special wards) Yokohama Kawasaki Saitama Kawaguchi Chiba Sagamihara

Reply: The revised manuscript now uses the term 'Greater Tokyo Area' with its clear definition.

Only at line 61 you say that the site "is located east of Tokyo", which is correct in my view.

Reply: Corrected as suggested.

Introduction : Efforts should be made to improve. It will help the readers to provide a review of existing literatures and how the results from this study improves our knowledge. There are several more recent papers (a few listed below, in addition to Yamada et al.) who have tried to discuss CO2 measurements and models from the Tokyo and surrounding areas like this paper, e.g.,

Pisso et al., https://doi.org/10.1186/s13021-019-0118-8
Ballav et al., https://doi.org/10.1007/s12040‐015‐0653‐y
Sugawara et al., https://doi.org/10.1029/2021GL092600
Bisht et al. https://doi.org/10.1029/2025JD043589

Reply: Thank you very much for the comment. Accordingly, the Introduction section has been improved largely to cover more recent papers and provide how the results from this study improve our knowledge.

Line 71ff : Why different treatment to the intake air samples ?

Reply: In response to the comment, the revised manuscript now states that "On the other hand, the observations using the LI-7810 were conducted without performing regular calibrations and without dehumidifying the sampled air. This approach was adopted to explore whether useful automated long-term measurements could be achieved with lower effort.

Line 120ff : referring to Fig. 1d ? I think this is not a reasonable metric to discuss, without water vapour correction. Additionally, I would also like the authors to analyse % error for the "localised" spikes, not the absolute differences, e.g., if a peak height is frequently seen at 60 ppm, the 12 ppm is a 20% uncertainty - not terrible. For baseline observation, 12 ppm is a serious error.

Reply: We understand your comment to indicate that a water vapor correction is necessary, and we fully agree. Based on your comment and by reading Pathakoti et al. (2024) at https://www.nature.com/articles/s41597-024-03243-x, we attempted to clarify our understanding of the discrepancy. Then, we realized that there are two different water vapor effects; 1) the effect due to the omission of the water vapor correction to derive the dry mole fraction and 2) the effect due to H2O-related interference, while the LI-7810 detects inherently weak CO2 absorption signals that overlap with those of H2O. Considering this, we realized also that we had not mentioned that the LI-7810 outputs CO_2 concentration data for dry air, in which the water vapor correction (the water vapor effect 1 described above) has already been applied, and that our analysis in this study was conducted using those corrected data. Accordingly, we now clearly state this point in the revised manuscript. In addition, a value of % error for the localized spikes is now mentioned in the revised manuscript.

Figure 2 and associated discussions: Why not apply a water vapour correction - otherwise the two instruments are not comparable. See for example - https://www.nature.com/articles/s41597-024-03243-x

Reply: As mentioned above, LI-7810 outputs CO_2 concentration data for dry air, in which water vapor correction has already been applied, and that our analysis in this study was conducted using those corrected data.

Line 150ff : Why this selection ? Why not apply mathematically a well-known correction term for H2O.

Reply: This selection was chosen based on Fig. 3, while it was difficult to determine the threshold. This is now clearly stated in the revised manuscript. As mentioned above, LI-7810 outputs CO_2 concentration data for dry air, in which water vapor correction has already been applied, and that our analysis in this study was conducted using those corrected data.

Figure 5 and associated discussions: Can this analysis be done by month or so? There seems to be more than one factor in the scatter plot.

Reply: In response to the comment, Figure 5 was revised to indicate temporal variations on a time scale of month. Relevant discussion was added in the text.

Figure 6 and associated discussions: Why the same analysis is not performed using the data from G4301 instrument ?

Reply: In accordance with your suggestion, we have revised the manuscript to include the results from the G4301 as well.

This decision here and also many earlier decisions to select data and analysis methods are not well argued - why one method/data is chosen over other. A general checking by the authors to improve the contents would be much appreciated by the readers.

Reply: Thank you for pointing this out. In response to the comment, we have revised the manuscript to include the results from the G4301 as well.

Line 208 ff: I guess this is something similar to Tohjima et al., 2020; but there is another way that GHG community use for deriving seasonal and synoptic variability in time series data (Nakazawa et al., environometrics, 1997). Please take a look and if possible a comparison of your method and fitting-filtering method would be helpful for the readers.

Reply: Based on the comment, we revisited papers such as Nakazawa et al. (1997) and Tohjima et al. (2020). Nakazawa et al. (1997) aimed to extract seasonal variations and long-term trends by deriving a best-fit curve from discretely obtained flask data, and our understanding is that their method is not suitable for estimating baselines under an urban atmosphere. Tohjima et al. (2020) discussed temporal variations in CO2 emissions by focusing on the ratio of concentration variabilities (standard deviations) of CO2 and CH4. In contrast, the baseline derivation in this study is intended to remove contributions other than those from fossil-fuel combustion sources near the observation site under an urban atmosphere. Although we considered including these points in the main text, we ultimately decided not to do so in order to avoid making the description overly complex.

Line 245 : Or is the high BC point causing this trouble ? e.g., if you did a fitting without the highest BC value the intercept would come down? Then it could mean that the high BC emission activities do not emit CO2 as much (say the BBQ restaurants nearby!), because the del-CO2 level remained flat for the two high BC values.

Reply: Thank you for the comment. Accordingly, we have added a discussion of the results obtained after removing high BC data in the revised manuscript.

Line 257ff: I strongly urge the authors to prepare another scatter plot after applying water vapour correction to the LI-7810 data because the G4301 is making dry air mole fraction measurement while the LI isn't. This study is making significant conclusions and can go further by a little additional calculation.

Reply: As mentioned above, LI-7810 outputs CO_2 concentration data for dry air, in which water vapor correction has already been applied, and that our analysis in this study was conducted using those corrected data.

Looks like repitation in Acknowledgements and Financial support - "...by the Environment Research and Technology Development Fund (JPMEERF21S20810 and JPMEERF24S12202) of the Environmental Restoration and Conservation Agency, provided by the Ministry of the Environment of Japan."

Reply: Corrected, as suggested.