Review of manuscript "Spatial resolution versus precision in CO_2 point-source emission retrievals from the DQ-1 spaceborne lidar" by Zhang et al.

Julia Marshall

This paper assesses the feasibility of estimating CO_2 emissions from point sources, using vertically integrated, remote-sensing measurements of XCO_2. The authors rightly point out that this capability is well-established for passive spaceborne sensors, but less so for lidar measurements. There are exceptions, however: both Zhang et al. (2025) and Han et al. (2024) have previously retrieved power plant emissions using spaceborne lidar data from DQ-1, and airborne lidar measurements have been used in various studies (e.g. Amediek et al., 2017; Wolff et al., 2021; Shi et al., 2023). The aspect that this paper aims to address is the appropriate spatial averaging length to improve the precision of the measurement without losing the ability to resolve the fine spatial structure. Unfortunately no conclusions can really be drawn – they show that changing the averaging length changes the retrieved emissions, and then apparently adjust this length in a somewhat arbitrary manor to best match inventory values. This is not how data-driven emission estimation is supposed to work.

The approach taken in the current study is quite different to that of Han et al. (2024), though they are using the same data. The approach is much closer to that of Zhang et al. (2025), though this paper (or rather the preprint, Zhang et al. (2024b)) is only cited in passing as a "validation study" to confirm the accuracy of the measurement. However, this is not at all the focus of Zhang et al. (2025), perhaps the authors meant instead to cite Zhang et al. (2024a)?

In any case: there are two published studies that have already looked at the retrieval of power plant emissions from DQ-1 data (Han et al., 2024; Zhang et al., 2025), but there is essentially no discussion of how this paper's approach differs. It is a pity that the authors did not engage a bit more with these studies, as they might have avoided some serious methodological errors, which essentially invalidate the results. These issues are outlined in more detail below. Given the impact of these errors on the results and conclusions, the paper cannot be considered for publication without major revisions.

Spatial sampling distance

The most serious concern relates to the assumptions about the spatial sampling distance of the instrument. The authors use 70 m throughout, which is the dimension of the footprint of a single shot measurement. Each shot pair is separated by 350 m, as can be seen in e.g. Fig. 1 of Han et al. (2024). (Zhang et al. (2025) give the footprint spacing as only 330 m.[1]) This is a fundamental mistake, and affects most of the results in the study. At very least, it changes all of the results in section 3.1 (Figs. 3–12).

The whole focus of this study is finding the optimal spatial averaging length to best retrieve small-scale features like plumes. This is a real limitation of current IPDA lidar technology. While the repetition rate is sufficient to resolve emission plumes from aircraft, faster-moving spaceborne platforms present additional challenges. This is made clear in several of the studies cited by the authors, such as this passage from Kiemle et al. (2017):

> Plume detection from quickly moving platforms such as satellites, however, requires high spatial resolution, which will only be achievable with a small laser footprint, a high laser pulse repetition frequency (PRF) and a single-measurement (non-averaged) retrieval.

Note that Kiemle et al. (2017) assume a PRF of 500 Hz for this feasability study, compared to the 20 Hz of DQ-1.

I cannot understand how the authors could have analyzed actual DQ-1 data (in Sects. 2.3 and 3.2, for example) without having realized the inconsistency between their 70 m assumption and the true sampling at some point. If the DQ-1 soundings were really averaged as explained in Fig. 13, there should no longer be a continuous-looking line of measurements for higher numbers of points included in the averaging. Compare to Figs. 4b or 8: The spacing of the plume cross-section changes as a function of the number of points that are averaged, as one would expect. If we consider the case of averaging 150 points for the overpasses shown in Fig. 13, given the actual spacing of 350 m between each measurement, there should only be one point per 52.5 km. This averaging would result in only two to three points for the Majuba plots, and perhaps three to four for the Lalitpur plots, but in Fig. 13 we see a continuous measurement track.

How can this be? The only thing I can think of is that a moving averaging window was used, essentially reusing all the measurements up to 150 times. This is statistically problematic, but also does not explain why the path length for the Majuba plots is the same for all averaging lengths: if 150 shot-pairs have to be averaged together, even allowing for the inappropriate use of a moving average, the first 74 shots ($\approx$ 26 km) should be truncated. This would be a visible difference in the scale shown in Fig. 13.

Emission estimation method

The description of the emission estimation method is insufficient, erroneous, and seems to have been copied verbatim from a previous study without citation.

Beginning with Eq. 1, the authors cite Bovensmann et al. (2010), but this formulation does not match the relevant equations in the source paper, some combination of their Eqs. A1, A2, and A4. To match Eqs. A1 and A2 in Bovensmann et al. (2010), $(\frac{x}{x_0})$ should be replaced by simply x. In this reference paper, x_0 is only introduced in Eqs. A4 and A5 to simulate emission sources with finite cross-sections, rather than point sources. In the manuscript under review, x_0 is never introduced.

Likewise, at first I struggled to decipher how Eq. 3 was derived from Eqs. 1 and 2, until I realized that the a in Eq. 3 was not the same a as in Eq. 1, nor is it ever defined. The variable μ was never introduced, and it seemed like an error to have b in both the first and second terms of Eq. 3. I consulted what seems to be the most relevant formulation in the cited paper (Eq. 1 in Reuter et al. (2019)), and it really did not relate to the derivation in the current paper.

Finally, I found the formulae (complete with some of the same errors) in Zhang et al. (2025). In fact, the majority of the text in Sect. 2.2 of the current manuscript was identical to that of Zhang et al. (2025). Given the "typo" in line 118, where the number 130 appears for no reason, I decided to check the preprint version of Zhang et al. (2025), Zhang et al. (2024b), which the authors did cite on line 58, but in a different (and incorrect) context. There I found the text in question, taken directly from Sect. 2.2, around line 130 of that paper. The preprint cited Bovensmann et al. (2010) in the introductory sentence of their Sect. 2.2 as well, but this was updated prior to final publication. The extra (first) b in Eq. 3 should have been a subscript, and the careless copy-paste explains some other formatting errors related to superscripts (e.g. in line 115). The manuscript under review kept the unit error for the variable w: it is given as $\mathrm{kg\,m^2}$ on line 135 of Zhang et al. (2024b) and on line 121 of the current paper, but was corrected to $\mathrm{kg\,m^{-2}}$ before publication in Zhang et al. (2025).

This is also flagged in the automated similarity report, though it appears less grievous as part of it shows up as copied from Zhang et al. (2024a) while the other part shows up as identical to Zhang et al. (2025). It is also relevant to point out that there is (only) one overlapping co-author between the manuscript under review and Zhang et al. (2025) – the second author of this paper is the seventh author of Zhang et al. (2025).

Besides these problems, the method section does not sufficiently explain how they solved for the flux rate, F. Which parameters had to be fit? Which values were compared to calculate the R^2 values that are quoted throughout? It is not possible to get a negative R^2 value, despite what is claimed in the text. (Could it be that the numbers given are rather the Pearson correlation coefficient?)

There is also no explanation of how the uncertainties on the emission estimates are calculated. What are the leading error terms? Looking at the comparisons to real data in Fig. 13, the wind direction seems quite far from perpendicular to the flight track. How was this dealt with?

Allan deviation

Given the larger methodological problem above this seems almost minor, but I was not really convinced by the application of the Allan deviation as it was described in the paper. Typically such plots are shown on a log-log scale, with the theoretical slope for white noise shown. This can be seen in e.g. Figs. 6 and 8 of (Amediek et al., 2017). (At very least, this $\tau^{-1/2}$ line should be shown.)

I was also a bit surprised by the formula given in Eq. 4, which is actually for the Allan variance, rather than its root, the Allan deviation. (There is also no explanation of what k and M are in the equation.) This formulation of the fixed-τ estimator of the Allan variance requires a new dataset for each value of τ. I think it might make more sense to use the overlapping Allan variance, which allows one to use all the data points in a given timeseries repeatedly. See e.g. Eq. 3 in Fix et al. (2018) for the both the formula and some discussion of its use.

Beyond these technical concerns, I was a bit confused about how this was applied. Is the assumption really that neither the albedo nor the true XCO_2 values change over single tracks for these regions? This seems very unrealistic. In the Arabian peninsula there is substantial oil and gas production, with associated emissions. Several of the tracks in the Boreal region go right over Lake Baikal, which would surely have an impact on the albedo, as would the presence of snow or the phenology of the larch forests in this region over the period from June through December. Even in the absence of local sources or sinks and albedo heterogeneity, the gradients of XCO_2 over zonal transects of over 1000 km can exceed 1 ppm simply due to synoptic-scale transport. The idea that these measurements should be uniform and represent only noise seems unreasonable.

Even if we assume for a moment that the approach is reasonable: how were the values in Fig. 2 applied? Is it expected that the albedo difference between the desert and the Taiga makes no difference to the measurement precision? Or was the precision for only one of the regions (fit to a curve, perhaps) applied for all the modelling studies that followed? The concrete use of these results was not sufficiently clear.

Minor comments

I will refrain from writing out most of the minor notes I made while reviewing the study, given the major issues make it unacceptable for publication in its current form. The one thing I will note is the poorly formatted references: they were not in alphabetical order, they were missing DOIs, they (and all paragraphs) lacked indentation, they were not linked in the PDF (minor), at least one was missing (Amediek et al., 2017), and they did not always contain information consistent with how they were cited (see e.g. Wang et al., 2022; Zhang et al., 2024b; Reuter et al., 2019).

References

Amediek, A., Ehret, G., Fix, A., Wirth, M., Büdenbender, C., Quatrevalet, M., Kiemle, C., and Gerbig, C.: CHARM-F-—a new airborne integrated-path differential-absorption lidar for carbon dioxide and methane observations: measurement performance and quantification of strong point source emissions, Applied Optics, 56, 5182, https://doi.org/10.1364/ao.56.005182, 2017.

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A., and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications, Atmospheric Measurement Techniques, 3, 781–811, https://doi.org/10.5194/amt-3-781-2010, 2010.

Fix, A., Quatrevalet, M., Amediek, A., and Wirth, M.: Energy calibration of integrated path differential absorption lidars, Applied Optics, 57, 7501, https://doi.org/10.1364/ao.57.007501, 2018.

Han, G., Huang, Y., Shi, T., Zhang, H., Li, S., Zhang, H., Chen, W., Liu, J., and Gong, W.: Quantifying CO_2 emissions of power plants with Aerosols and Carbon Dioxide Lidar onboard DQ-1, Remote Sensing of Environment, 313, 114 368, https://doi.org/10.1016/j.rse.2024.114368, 2024.

Kiemle, C., Ehret, G., Amediek, A., Fix, A., Quatrevalet, M., and Wirth, M.: Potential of Spaceborne Lidar Measurements of Carbon Dioxide and Methane Emissions from Strong Point Sources, Remote Sensing, 9, 1137, https://doi.org/10.3390/rs9111137, 2017.

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., Bovensmann, H., and Burrows, J. P.: Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites, Atmospheric Chemistry and Physics, 19, 9371–9383, https://doi.org/10.5194/acp-19-9371-2019, 2019.

Shi, T., Han, G., Ma, X., Pei, Z., Chen, W., Liu, J., Zhang, X., Li, S., and Gong, W.: Quantifying strong point sources emissions of CO_2 using spaceborne LiDAR: Method development and potential analysis, Energy Conversion and Management, 292, 117 346, https://doi.org/10.1016/j.enconman.2023.117346, 2023.

Wang, N., Zhang, K., Shen, X., Wang, Y., Li, J., Li, C., Mao, J., Malinka, A., Zhao, C., Russell, L. M., Guo, J., Gross, S., Liu, C., Yang, J., Chen, F., Wu, L., Chen, S., Ke, J., Xiao, D., Zhou, Y., Fang, J., and Liu, D.: Dual-field-of-view high-spectral-resolution lidar: Simultaneous profiling of aerosol and water cloud to study aerosol–cloud interaction, Proceedings of the National Academy of Sciences, 119, https://doi.org/10.1073/pnas.2110756119, 2022.

Wolff, S., Ehret, G., Kiemle, C., Amediek, A., Quatrevalet, M., Wirth, M., and Fix, A.: Determination of the emission rates of CO_2 point sources with airborne lidar, Atmospheric Measurement Techniques, 14, 2717–2736, https://doi.org/10.5194/amt-14-2717-2021, 2021.

Zhang, H., Han, G., Chen, W., Pei, Z., Liu, B., Liu, J., Zhang, T., Li, S., and Gong, W.: Validation Method for Spaceborne IPDA LIDAR X_{CO_2} Products via TCCON, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 16 984–16 992, https://doi.org/10.1109/jstars.2024.3418028, 2024a.

Zhang, X., Yang, H., Bu, L., Fan, Z., Xiao, W., Chen, B., Zhang, L., Liu, S., Wang, Z., Liu, J., Chen, W., and Lee, X.: Estimation of diurnal emissions of CO ₂ from thermal power plants using spaceborne IPDA lidar, https://doi.org/10.5194/egusphere-2024-3152, 2024b.

Zhang, X., Yang, H., Bu, L., Fan, Z., Xiao, W., Chen, B., Zhang, L., Liu, S., Wang, Z., Liu, J., Chen, W., and Lee, X.: Estimation of diurnal emissions of CO ₂ from thermal power plants using spaceborne integrated path differential absorption (IPDA) lidar, Atmospheric Chemistry and Physics, 25, 6725–6740, https://doi.org/10.5194/acp-25-6725-2025, 2025.

Notes

[1]Reviewing this paper led me to read the results in Zhang et al. (2025) as well, to see how they dealt with the averaging/spatial resolution question. Their results also seem to be fundamentally inconsistent with the spatial resolution of DQ-1. Consider Fig. 4 of that paper: the text states that there are 76 points that were designated as "in plume", shown in blue in Fig. 4d, along with the modelled emission plume (green and red dots). This modelled emission plume is apparently sown in Fig. 4c, and the position where the measurements cross the plume appears to be ≈ 10 km wide. However, 76 points should cover a distance of around 26 km. Is the model scale completely wrong? Are these even real measurements? The fact that the DQ-1 data have still not been publicly released makes it impossible to check.