Satellite-based global monitoring of urban-scale methane emissions

Reviewer #1

Long et al. present urban methane emission estimates for London, New York, and Los Angeles using an updated source pixel method and TROPOMI methane retrievals. The improved method uses a revised wind analysis to incorporate pressure-weighted PBL winds into the emission calculation, rather than the 10m wind data provided with the TROPOMI data. The manuscript is well written and thoroughly presented. There are, however, a few areas that I believe require attention prior to publication.

1. In the abstract, the work is motivated by the following statement: "Previous studies of urban methane emissions have employed a range of emission quantification methods, leading to inconsistent estimates and different associated uncertainties." However, the conclusion does not address whether this work has helped improve these inconsistencies.

Response: Thank you for this insightful comment. The inconsistencies between different top-down emission estimates and bottom-up inventories have been well documented in previous studies, and they do remain despite this study. But we hope we can offer an improved method with greater possibility for repeatability going forward. Discrepancies can be partially attributed to the thermodynamic variability (i.e. wind fields) within the planetary boundary layer and the methods used to define background concentrations. In addition, inconsistencies among previous top-down studies are largely driven by differences in the definition of what constitutes the urban area of interest (i.e. city boundary) and the reporting periods (i.e. timescales), which makes direct comparison between studies challenging. By integrating pressure-weighted winds under the planetary boundary layer with an upwind background concentration approach, the method we describe could help bridge methodological differences in flux quantification by providing a more repeatable methodology and approach to urban boundary definitions. This is now better clarified in the conclusion section of the revised manuscript (Line 458-470).

2. Section 2.2: Oversampling method. How are missing pixels handled with this approach? Is the oversample done on the column mixing ratios or on the calculated methane enhancements?

Response: The oversampling method used in our study is applied to the methane column mixing ratios retrieved from TROPOMI and not the calculated methane enhancements, and is only applicable to pixels that first passed data filtering criteria described in Section 2.1.1. This has now been more clearly specified in the revised manuscript (Lines 180-181). In addition, we have changed the terminology from "oversampling" to "remapping" throughout the revised manuscript.

3. A major component of this work is the use of pressure-weighted PBL wind instead of 10m wind in the calculation of the emission rate. Unless I missed it, I do not see a comparison of how these two approaches compare in emission space. In the comparison with other studies for these cities, how do the emission rates calculated using 10 m winds compare? Put another way, how do you know the pressure-weighted PBL results result in a more accurate emission estimate?

Response: We didn't present specific methane emission estimates obtained using pressure-weighted PBL wind fields and 10 m winds in the paper. However, conceptually, Equation (2), shows that the calculated emission is linearly proportional to wind speed. Therefore, the comparison between the wind fields included in the TROPOMI retrieval products (10 m wind) and those from ERA5 reanalysis (hourly resolved, pressure-weighted PBL winds) in Figure 2 serve as a proxy for comparing emissions using different wind inputs. The significant differences between the two wind fields shown in Figure 2 indicate that considerable discrepancies in the estimated emission magnitudes can be expected. For example, for orbit 17480, which passed over the London area on 26 February 2021, the emission estimates derived using PBL pressure-weighted wind (3.18 m s^{-1}) and 10 m wind (1.86 m s^{-1}) are 25.74 $\pm$ 14.13 t h^{-1} and 15.44 $\pm$ 7.65 t h^{-1}, respectively. Such a difference would be expected and conceptually, we would expect that a wind that is more representative of whole boundary layer advection is more realistic, compared with typically much lower surface wind speeds which would negatively bias emissions estimates. The advance in the new method lies in the incorporation of hourly resolved PBL-pressure weighted wind fields and an explicitly defined upwind background concentration.

4. Line 212: Is the methane enhancement calculation done for each overpass, or over the averaged column mixing ratio data? If it is the former, this should be clarified. If it is the latter, would this not skew the result towards overpasses with larger overall column retrievals?

Response: Thank you for pointing out the potential confusion here. It is indeed the former. This study uses TROPOMI data to estimate orbit-level methane emissions over urban areas based on the methane enhancement calculated for each overpass. Specifically, this has been explicitly clarified in the revised manuscript (Line 232).

5. Line 276: "While most previous studies using the source pixel method have analyzed an aggregated set of observations through equally weighting each day in the average, we use the spatial coverage (in %) of valid observations in the source region to weight each overpass in an effort to reduce bias." It is not clear to me which bias this will reduce. There may be more valid pixels in a part of the domain than another, but there will still be transport within the domain such that a column retrieval for a given pixel will include surface emissions from outside the pixel geometry.

Response: In our study, the spatial coverage of valid observations in the source region is considered when deriving urban emission estimates at monthly or yearly timescales from orbit-level estimates. For each overpass, the number of valid pixels over the source region may vary, even under cloud-free conditions, and potential heterogeneity of methane column mixing ratios may exist between different pixels, which need to be carefully treated when averaging multiple overpasses. If a conventional mean of orbit-level estimates were used, overpasses with limited valid data coverage (for example 25%) would be assigned the same weight as those with full coverage (100%), which could introduce bias into the final estimate. By weighting valid observation coverage, we are seeking to reduce spatial bias arising from potential heterogeneity in methane mixing ratios in the source region. Although emission transfer among pixels may occur, it can be expected to be strongly influenced by meteorological and topographical conditions, which may vary over the source region and therefore impossible to accurately account for in any systematic way. Conceptually, if there was complete sampling of the source region, transport within the domain would not affect the flux calculated. But it is true that this could cause issues when the source region is only sparsely sampled, which is actually further justification for weighting emissions for spatial coverage. Thus, in this study, we don't consider surface emission transfer among different pixels, but it is correct to note its existence. This clarification has been added in the revised manuscript (Line 292-297).

6. Figures 4,5,6: What is driving the differences in the size of the error bars between different months? Some months have very tight error bars.

Response: This is indeed highly variable. The size of the error bars is influenced by the uncertainty associated with each valid overpass in a given month, which is further derived from uncertainties in methane enhancement and wind variability based on Equation (5). When aggregating orbit-level estimates to the monthly timescale, the uncertainty of each valid overpass is propagated into the final estimate uncertainty through Equation (6). For months with several overpasses that have relatively small emission estimates but large uncertainties, or for months in which the emission estimates vary considerably among overpasses, the final error bar will be large. Otherwise, the error bar will be smaller. In our detailed examination of a great many individual overpasses when evaluating the method, we noted that it was the variability in the wind direction (i.e. wind angle variance relative to the mean direction used to define the background box) that resulted in the largest uncertainty, and that this is highly variable from overpass to overpass (especially for low wind speed conditions where boundary layer flow over a city may be expected to much less uniform).

7. In the several months, there have been a few new studies examining urban methane using TROPOMI (https://doi.org/10.1073/pnas.2504211123, https://www.science.org/doi/full/10.1126/sciadv.adz9308). They are not included here, so I am left wondering whether or how this work fits within the context of these more recent works. Or perhaps, they are not strictly relevant to the work here?

Response: We agree with the reviewer that these recent studies on urban-scale methane emission estimates should be included for comparison. We note that the total annual methane emissions for urban areas reported by Wang et al. (2026) are not accompanied by detailed uncertainty information in the Supplementary Materials, making comparison here impossible. But we have now included the results reported by Whiting et al. (2026) and added this comparison to Figure 7 in the revised manuscript. Their study derives emission estimates broadly comparable to ours for the three case study cities from 2021 to 2023. However, given the differences in the definitions of urban domains and background concentrations, direct comparisons remain complicated. This discussion (and complexity in comparing such studies because of urban boundaries) has been clarified in the revised manuscript (Line 414-423).