Dear the authors of the manuscript:

The study made an attempt to estimate traffic CO2 emissions from the city of Boulder using CO2 data collected from a ground-based remote sensing instrument. The team has developed and maintained their open-path dual-comb spectrometer. The instrument has been subjected to several comparisons prior to this study. The authors conducted a 7.5-week long observation (September-November) and found two time periods (10/22 (Sat) 11am-4pm and 10/25 (Tue) 10am-4pm) they think suitable for the emission estimation of this study. The authors employed a Gaussian plume model with a city-wide traffic emission distribution constructed using the traffic data collected in Boulder, in order to estimate the annual traffic emission from the city. The estimation yielded $6.9 \times 10^5$ MT CO2/yr for the year 2016, which is 153% (155% in the manuscript) of the scaled city emission estimate (the 2015 traffic emission was scaled up using the total vehicle miles traveled in 2016). While several sources of errors and uncertainties due to the estimation approach were acknowledged, the authors discussed them by citing previous work, but they remained fully unquantified. The significance of this study is the application of CO2 data collected from the unique ground-based remote sensing instrument. But due to the poor design of the estimation approach and the lack of the evaluations of the results, I feel the authors failed to fully conclude this study and thus I do not recommend this manuscript for publication. I listed my major concerns and some other comments below.

1. Poor experimental design

I don’t think the emission estimation was good enough to solely estimate traffic emissions. The authors had to make big assumptions to estimate annual traffic emissions from the city using two data period in 2016. Prior to the actual emission estimation, the authors needed to prove that they can get a reasonable estimation regardless of the assumptions made.

2. Traffic emissions

I checked the latest Boulder’s inventory (2016) and their traffic emission was reduced by 10% from the previous year, which contradicts with the conclusion of this study and suggests a larger discrepancy between the authors’ estimate and the inventory estimate.

Given the limited observation and the simple modeling, maybe it would have been a good idea to focus on the total city emission. I do not have an access to disaggregated CO2 emissions, but traffic emissions account for 28% of the city total emission, the residential is for 16% and commercial and industrial for 54%. Ignoring the contributions from other sectors does not seem to be a good idea, especially w/o doing any source attribution analysis.

3. Background problem?

This study used CO2 data from the reference path as a background. I understand that the air must be clean for the reference data, but my concern was the authors were comparing two different airmasses to calculate the CO2 enhancement. The only supporting information of background CO2 vs. city CO2 was the wind direction from a few observation points.

4. CO2-eq. I do not understand why the authors did not use emissions in the CO2 unit, rather than in the unit of CO2-eq. In collaboration with the city council, I would imagine it is not too difficult to obtain emission estimates solely for CO2.

5. Bottom-up vs. Top-down?

The discrepancy between bottom-up vs. top-down estimates are often large, as seen in previous studies. In many cases, the uncertainties associated with the inventory are assumed to be small. In this study, the authors have added a lot of potential errors when mapping the traffic emissions in space (approximated the spatial patterns) and time (scaled up to annual emission). Given that, less convincing than other studies if the authors did discussion just with citing papers.

Other comments:

P1, L31: “top-down measurement” sounds odd to me. How about “top-down approach using atmospheric measurements”?

P2, L58: A SoCAB CO2 top-down study has appeared on ACPD. Check out Hedelius et al. (2018) ACPD.

et al. (2018) ACPD.

P4, L162: This traffic emission modeling is based on huge assumptions. The errors and uncertainties associated with this modeling needs to be quantified at least to show if the emission estimation approach in this study has a good accuracy to show the utility of the CO2 data. Also, the authors might want to check the consistency/inconsistency between the traffic data the authors used and the new 2016 inventory used.

P6, L238: No correlations is good . . . but these are not shown anywhere.

P8, L321: 5% sounds small, but it is comparable to the potential emission changes we want to detect. I don’t think it is insignificant.

P8, L327: I think these are another set of big assumptions. These assumptions needed to be tested. The authors cited Gurney et al. (2017), but that is a case for Indianapolis. The authors could use the same logic for the large discrepancy between the top-down and bottom-up estimates. But the authors’ statements are not supported by any quantitative analyses.

P16, Figure 1: This figure needs to be improved. It does not even clearly show the traffic distributions.

P19, Figure 4: Given the small changes in CO2 we are discussing, I think the range of Y is too large. We can’t see the variability in CO2 data.