Review Comments Referee #2

Comments made to manuscript-version5

The authors have answered most of the reviewers' questions adequately. I have a few more suggestions for this paper before I think it is ready for publication.

I think the Uncertainties section should come earlier in the paper, before the final estimates of emissions are presented. Are the uncertainties, say for the NO_x emissions, weighted by the daily uncertainties? In other words, was 182 ± 324 tonnes/d treated the same as 301 ± 166 tonnes/d? I also think the uncertainties should be clearly stated in this section, rather than scattered throughout. I found the boundary layer estimate of 650 ± 50 thrown in a section where it didn't really fit. A better place for it would be the Uncertainty section. Further, the authors state an uncertainty in the wind measurement of 0.1 m/s. I realize many of the coauthors wrote the original wind measurement paper, but my reading of the stated uncertainty in Conley et al. (2014) as < 0.2 m/s. Did the Mooney aircraft fly an L-pattern during this SJV project? Are the GPS measurements even accurate to 10 cm?

We have moved Section 4.2 Error Analysis up to the budgeting methods section, 3.4.2. to more directly address these questions in the context of the methodology.

The overall uncertainties expressed as the standard deviation of the mean, are not weighted by each flight's uncertainty estimate.

When reporting the average boundary layer height, 650 $\pm$50 m, it was not meant as a source of uncertainty, because it's not that significant of one, but simply to give a physical idea of the depths observed (and their day to day regularity, here $\pm$50 m is the standard deviation of each day's average. Nevertheless, the concern about these parameters being elided due to a companion paper is acknowledged and we have added a table of the boundary layer parameters into the supplementary materials (see below).

We agree that a better wind measurement error estimate would be 0.2 m/s in keeping with the reference and have changed that. We have experiential reason to believe that the wind errors are less than that, but there is no reason to exclude that from our error estimates here.

The Mooney aircraft did not routinely fly wind calibration patterns during this SJV project, but it did so in between the deployments due to other efforts to quantify small-scale site emissions. The calibration coefficients are most often found not to change significantly during these checks and are updated in the data analysis codes when any wind system configuration change occurs. Finally, the individual GPS measurements are not likely to be accurate to within 10 cm, but the utility in the differential GPS system lies in each antenna's offset approximately canceling. The error reported in Conley et al. (2014) of <0.2 m/s is an aggregate of empirical testing against other wind measurements, not bottom-up estimates of the system components.

If this paper is going to offload much of the boundary layer analysis to the companion paper, the companion paper should be provided along with this one.

As mentioned above, we have added a supplemental Table S1 to detail the observed boundary layer parameters from this study in order to decouple this work from the "future companion paper" as much as possible. Due to many modeling improvements the exact details (e.g. entrainment velocities) of the upcoming paper will be slightly different from the values used here (but are absolutely within our estimated errors) and we feel it is important to document such differences, albeit minor, in print.

p. 6, line 20, the authors should add the model number of the Picarro, like they did for the Eco Physics instrument.

Done.

In section 3.2, Aircraft Instrumentation, please add the estimated uncertainties to these measurements.

Done.

Section 4.1.1.2, how does the possible influence of the Soberanes fire compare with your estimate of the footprint? If the fire did affect the measurements, are your footprints too small? And if the fire didn't affect the measurements, why spend so much time discussing the possible interferences?

On some days the overlap appeared to be quite expansive (e.g., https://earthobservatory.nasa.gov/images/88440/soberanes-fire-california) however the fire effluent was mostly in the layer above the ABL until afternoon convective development brought some of it into the ABL.

*I do not understand the question, "are your footprints too small"? The footprints are determined by the aircraft paths from each flight, independent of the state of the fire effluent. We spend time discussing this possibility because it was an unanticipated environmental condition of the experiment that should be considered when trying to perform the budgets of these trace gases, which are all known to be influenced by wildfire emissions. We conclude that the **direct** emissions of fire NO_x likely did not make a significant impact on our budget, but other reactive nitrogen species may have contributed to the analytic artifact in our NO_x measurement that we also discuss. Our NO_x emission estimates have already been reported in the literature (Almaraz et al., 2018), but with scant justification of the methods. In the present work we are attempting to systematically address as many potential problems with the analysis as possible and then make a probabilistic argument that our estimates are still significantly larger than inventories because of a missing soil source.*

p. 21, line 10, what units are 3.6 and 2.4? Are these scaling factors?

They are ratios of top-down airborne estimates to emission inventory estimates as reported in Trousdell et al. (2016). We have adjusted the wording to make that clearer.

Maasakkers et al. (2016) should account for the seasonality of CH4 emissions. How do CH4 emissions in the SJV compare to that work?

We do not directly account for the seasonality of CH_4 emissions in this work, nor does Cui et al. (2017). We do, however, speculate that the seasonality of dairy emissions as reported in Arndt et al. (2018) could bring the estimates in much closer alignment.

*We have decided **not** to include a comparison to the different, national inventory of Maasakkers et al. (2016) for a few reasons. First, their inventory shows a very low correlation to that of CALGEM, which has been extensively studied and optimized for California, for their livestock category (r^2 ~ 0.21). Second, the comparison is particularly poor (their Figure 6) in the exact region of this analysis. They state, "The largest differences are from livestock emissions, as CALGEM uses more local data to distribute these emissions within the large California counties." Further, their accounting for annual variations are only applied to manure management and they admit, "Livestock emissions also vary subanually as a function of varying herd size, and management practices but those effects are not included in our inventory."*

Figure 2, you could also refer to Figure 1 for the locations of the cities.

OK.

Supplement, p. 3, is this a correction factor that is multiplied by the measured NOx? If so, please add the additional information, e.g., NOx(corrected) = (factor)*NOx(measured)

Corrected.

Grammar Suggestions
p. 2, line 3, change "seem" to "seems"
check.
p. 4, line 7, add a comma after "chemistry"
check.
p. 5, line 11-12, set the clause "coupled … floor" off with commas, and remove the other commas in this sentence
check.
p. 6, line 16 and throughout, "data" should be plural
check.
p. 7, line 21, I'm confused by this statement. The NOAA crew corresponded well with the Mooney crew? Or the ozone measurements agreed well?

Changed to, "who have shown excellent agreement between the ozone data collected by the aircraft and lidar (Langford et al., 2019)."

p. 7, line 23, change "diagnosis" to "diagnose"

check.

p. 12, line 4, "Results from …" is in boldface

check.

p. 16, line 24, change "in" to "by"

check.

p. 20, line 5, "Data are"

check.

p. 20, line 13, I think d(Cox) should be d(Ox)?

check.

Supplement, p. 2, line 1, subscript the 2 in NO2

check.