Response to Referee #1 comments on "Photochemical Production of Ozone and Emissions of NO_x and CH_4 in the San Joaquin Valley" published 19-Feb-19

We thank the referee for their thorough reading of the manuscript, and address the individual comments below:

General Comments:

1) There are numerous typos and grammatical errors. Some of these are listed be- low, but I would suggest a thorough proof-reading before resubmitting the manuscript. Some of the language is also vague and inappropriate for a journal paper (e.g. “stuck out”, “more or less”, “more and more”, “a lot”).

Yes, thank you for pointing that out, we have double-checked the manuscript and removed the more imprecise colloquialisms.

2) No details of the flights are presented, with the exception of two sentences in Section 3.3. Figure 1 gives some idea of the horizontal extent of the two flight campaigns, although it is difficult to see the EPA flight lines and it is impossible to distinguish individual flights. We are given some windows of time (but no actual flight durations) and an altitude range “from the surface up to 4 km”. Presumably the surface measurements are at the start and end of each flight, unless there were multiple landings at different airports. Given the importance of the vertical coverage on the emission calculation, much more detail is required.

We have expanded on the paragraph in Section 3.3 detailing the flight strategies to read, "The main data set we use here comes from six flights sponsored by the US EPA (labelled EPA in Figure 1) during the California Baseline Ozone Transport Study (CABOTS) that were conducted on the afternoons of 26-28 July and 4-6 August, 2016 from 1100 to 1500 PST spanning an approximate altitude range from near surface up to ~3 km. The aircraft flights consisted of 6-7 back and forth level and profiling legs of approximately 15 minutes duration (~60 km) up and back along the mean wind direction (the valley axis) in order to capture the horizontal advection and vertical gradients of the measured scalars. The flight domain focused on the region of the SJV between Fresno and Visalia with approximately two-thirds of the data collected below ~1 km . The flight days were selected in coordination with a crew from NOAA operating a Tunable Optical Profiler for Aerosol and Ozone (TOPAZ) lidar in Visalia, California (Langford et al., 2019). Periodically the plane would make deep vertical profiles from ~10 m to 3 km in addition to two or three other profiling legs in order to diagnosis the ABL top, its growth, and vertical profiles of the measured scalars."

3) The description of the NOx processing (Section 4.1.1) lacks detail. Why one standard deviation? How much data are removed? If this all occurs in the late afternoon, why not just reject data from that time of day? And if fire smoke is entrained during that time period, why is this effect not discussed for O_3 and CH_4?

We have clarified and justified the explanation of the method of removing the NO_x spikes by stating, "NO_x ABL data was filtered by eliminating data greater than one standard deviation above the mean before being analyzed in order to remove the skewness from the distributions induced by numerous spikes encountered in the late afternoons. Variations of this threshold from 1-3 standard deviations did not change the mean flight concentration by more than 2-3 percent so the exact threshold was not considered critical for our analysis. The data filtering was done to eliminate the spikes which were consistently encountered throughout the latter part of the flights, each lasting no more than a few minutes and uncorrelated with any other species measured (CO_2, CH_4, and O_3.) We conjecture that their source may have been something in the fire smoke entrained in the late afternoon ABL that caused a transient interference in the NO_2 photolytic chamber (they were not observed in the NO measurements.) Furthermore, as we discuss later in conjunction with Figure 5, the influence of the fires on NO_x measured by the surface network (~1 ppbv) appears to be minimal relative to the clear signal enhancements in CO (~200 ppbv) and PM2.5 (~15 μgm^{-3})."

4) If emission rates are measured over the course of a few hours for 6 days, then presenting these values as month or yearly rates (i.e. tonnes/year) is not "averaging" or "converting" - It is extrapolating. For the most part this can be fixed by using the correct terminology. However, due to variations in emissions with time (some of which are discussed, such as the "weekend effect"), extrapolating these hourly time scale values to yearly values has an associated uncertainty that should be discussed.

Of course, all measurements of limited duration need to be extrapolated in time; however, there are conventional units used in inventories which make for more convenient comparison. Here we use ppb/h, tonnes/day, and Mg/h (and Gg/yr) for the O3, NOx, and CH4 source strengths, so there is not all that much extrapolation in these units. The conversion from the four hour flight to the entire daily emissions of NOx is discussed extensively in Section 4.1.1.3, and we have added a statement in the methane emissions section (4.1.3) to acknowledge their seasonal temperature dependence:

"Because the airborne measurements of our study take place during six summer days, comparisons with annually averaged inventories should be done with caution as methane emissions from the dominant sources (dairies) are likely to be seasonally temperature dependent. For example, a recent study of two dairies in the SJV (Arndt et al., 2019) reported facility-wide winter emissions to be only 40-50% of those during summer sampling."

5) The consideration of uncertainties is generally weak. For example, in Section 3.4 (page 8, line 16) it is simply stated that the approach is justified and that 20% in conservative without any reference to where that number comes from. In Section 4.4, two values (1 ppb, 50 ppb) are chosen "because the term is estimated by eye". For wind, 0.1 m/s is based on the "measurement capabilities" of an instrument which isn't named or referenced.

At the core of this issue, is our belief that the estimation of compound measurement errors is important yet ambiguous guesswork, and we generally refrain from the presentation of precise quantitative determinations of errors lest they give the impression that they are fully understood. When it comes to applying a mesoscale area over which to integrate and interpret our measurements, the exact uncertainty of the domain is, frankly, uncertain. Nonetheless, we have added the following discussion to justify our crude estimate as conservative in Section 3.4:

"The average area of this polygon was 5,200 km^2 (σ=940 km^2). To estimate an uncertainty in this area, we consider the average advection distance of the mean ABL wind (~3 ms^{-1}) over the course of a large eddy turnover time (boundary layer height divided by convective velocity scale ~ 8 minutes = 650 m/1.35 ms^{-1}) and multiply this on either end of the domain by an average cross-valley dimension (70 km) to generate a 'spread' in the sampled ABL volume influenced by the surface flux field. Although this additional area represents less than 4% of the overall domain, we include a conservative 20% error in the error analysis for it."

And in Section 4.4 Error Analysis we have added:

"The delta term error was assigned to be 1.0 ppb for NO_x and 50 ppb for methane based on variations in the data estimated by eye from inspection of the many vertical profiles."

"…and the horizontal wind at ABL height assigned an error of 0.1 ms^{-1} based on the measurement capabilities of the instrument (Conley et al. 2014)."

6) Correlation length is typically calculated at the 1/e value, not the "crossing of the zero-correlation line". This is primarily because small amount of noise in the correlation values can significantly change when the zero line is first crossed (Figure 12 demonstrates this effect). Smoothing the correlation or fitting an exponential decay to determine the value at 1/e gives a more accurate measure of the correlation length that isn't subject to the effects of noisy data.

The statement was in error because we did in fact use the e-folding depth for the calculation. We have changed the statement in the paper.

7) It is also not clear why correlation distances are important in the context of the study (expect to inform future satellite resolution values). What is the expected correlation length that would be associated with cities and traffic? Wouldn't this value depend on how far downwind from the source the measurement is (due to horizontal diffusion)? Why are potential temperature and water vapor compared? Do these values relate to the patchiness of land use and the separation of lakes and rivers? Why is this important?

We are not exactly sure of the answers to these questions, because we have not found this detail discussed directly in the literature. However, we wanted to present these results because we feel it may spark a useful discussion in the future literature on regional air quality. Because

the measurements represent a sampled area of order 75X75 km, the average decorrelation length should include some 'average' of a total urban plume downwind. The main city centers of this region are approximately 5-10 km in linear dimension, whereas California Highway 99 runs straight through the ~75 km flight domain, so it is not obvious what a line source spatial scale should be exactly.

We have amended the discussion to be more direct about our speculation: "Temperature and water represent ABL scalars dominated by surface fluxes, so in principle their correlation lengths are related to the scale of heterogeneity in their surface sources (irrigated or fallow fields, plots of differing albedos, urban heat islands, etc.) In the case of ozone, photochemical production dominates in the afternoon requiring the mixing together of NO_x and VOC emissions. The more spatially diffuse pattern of CH_4 and NO_x, comparable to that of ozone, may imply a preponderance of broad areal sources rather than localized emissions from cities (5-15 km) and/or highway traffic. This result for NO_x calls to mind the findings from Russell et al. (2010) and Pusede and Cohen (2012) previously mentioned which show broad scale homogeneity for NOx concentrations in the SJV unlike in other regions where urban hotspots are more localized."

Specific Comments:

Page 3, line 9-10: A claim like this is meaningless without defining “limited in duration”, “overextended in sampling”, and “altogether uncoordinated”. If those terms can be defined a citation will also be needed to a substantial review paper that backs up this claim.

This is an opinion of ours recalling one of the author's 20 years of experience with airborne atmospheric chemistry studies. As far as we can tell there is no author to date who has lamented the absence of such coordinated flight strategies, so there is nothing to cite.

Page 5, line 17: For the editor – Is a citation of a manuscript in preparation accepted?

Changed to: "The approximate residence time within this buffer layer was found to be about one week based on analysis of WRF model output which we plan to detail in a companion paper."

Page 7, line 11: The aircraft doesn’t measure “from the surface”.

In response to General Comment (2) above we have changed the wording to "near surface" and have explained about the occasional low-passes conducted at airports to sample within 5-10 m of the surface.

Page 9, line 14: Using WRF parameterizations isn’t a measurement.

Changed wording.

Page 20, line 12-13: This sentence is very confusing. (e.g. What is “a common time height stamp”?)

changed to, "All data was selected to be within the time dependent ABL and corrected to a common time and height within the ABL to remove biasing from temporal and vertical trends before the autocorrelation was run."

Minor comments/corrections:
Page 2, line 23. Sentence doesn’t make sense.

Split it into two sentences for clarity's sake: "In 1990, the San Joaquin Valley Air Quality Study (SJVAQS) was conducted. The largest study of its kind in the U.S. at the time, the SJVAQS targeted the complexities of the SJV at a time when it was considered the nation’s second worst overall air quality problem (Lagarias and Sylte, 1991)."

Page 3, line 14. Period should be outside bracket.

That is not my understanding of conventional grammar.

Page 5, line 3. What does “its” refer to?

changed to "…reaching a maximum…"

Page 5, line 11. Why is the air “unique”?

changed to, "The air above the ABL in the SJV is unique in that it does not purely consist of background air from the free troposphere (FT) as in most flat terrain."

Page 6, line 13. Should be “Vaisala”. “Ozone” should start new sentence.

Changed.

Page 9, line 22. Should be “50%”.

Changed.

Page 12, line 13. This is not a sentence.

Fixed.