

Reviewer 1

This manuscript discusses oxygenated volatile organic compound emissions and sources in the Berkeley/Oakland area of California. The authors tied together both flux tower measurements and mobile monitoring in this study to investigate emissions and sources of ethanol, methanol, acetone, and acetaldehyde with significant depth. They assessed the role of cooking in contributing to ethanol and acetaldehyde emissions, and the contributions of VCPs for acetone and methanol. They concluded that the total oxygenated volatile organic compound contribution to total measured VOCs was 70% of molar flux, 69% of measured mixing ratios, and 27% of calculated OH reactivity, with ethanol standing out as having the highest mixing ratio and flux of the focus chemicals. They also performed some comparisons of their flux measurements with the GRA2PES inventory and showed good agreement for ethanol sources but not with emission factor estimates for ethanol. Emission factor estimates for acetaldehyde and methanol similarly had poor agreement with GRA2PES, with the inventory generally overestimating emissions. This was a strong paper and is a good fit for the journal. I have a few minor comments and suggestions, which are listed in detail below.

We thank the reviewer for the compliments and thoughtful suggestions.

Methods: Minor clarifying questions.

Line 104-105: When you say roughly the same roads were sampled for each domain, do you mean roughly the same roads daily, or across some other unit of time?

Roughly the same roads daily. Updated text is below.

“Roughly, the same roads were sampled on each day for each domain, and measurements were randomized in terms of street sampling order and sampling time.

Line 218: It would help readers to provide slightly more detail about how the flux footprints were computed rather than relying too much on the cited literature, and also to potentially move this discussion closer to the occurrence of Figure 1 where the flux footprint first appears.

Thank you for the suggestion.

“Flux footprints, which are the geographic area impacting tower measurements, were calculated using the methodology described by Kljun et al., (2015). Briefly, the footprint depends on the tower height, wind direction, wind speed, turbulence, stability, and surface roughness length, which was determined in 30° wind quadrants using Lidar scans. These parameters are described in more detail in the preceding publication (Katz et al., 2025). Flux footprint modeling, especially in urban environments, can be uncertain, so the intention with this analysis was to show the average geographic area of influence instead of to pinpoint exact emission sources on short time scales.”

Line 133: Can you clarify how many times per day was the grid cell sampled (as you mention that you report the median of daily means)?

“Grid cells were visited one or two distinct times per day, and there were multiple data points per visit since data was collected at 1 Hz.”

Line 139-140: Can you provide some additional justification, validation, or references for this definition of background concentration used?

Thank you for requesting clarification. The background-subtracted concentrations were used solely to identify transient spikes in concentration above the local background and were not used for any quantitative analyses in the manuscript. Therefore, the choice in percentile may impact the visualization of individual spikes but does not impact other results in the manuscript.

Our approach follows the methodology of Apte et al. (2017) (<https://doi.org/10.1021/acs.est.7b00891>) who estimated the local background using the 5th percentile concentration within a rolling 120 second window. We also used a rolling window to estimate the local background but used the 10th percentile based on inspection of the resulting background-subtracted time series. We found that the 10th percentile provided a stable representation of the local background while producing spike identification consistent with the 5th-percentile approach. Changing the percentile which defined the background produced negligible differences in the identified spikes. We clarified this methodology in Sect 2.1.3 and noted that background subtraction was only used for spike identification.

“Spikes in VOC concentration were identified using background-subtracted data. The background concentration for each VOC was designated as the 10th percentile in 10-minute rolling windows. Background subtracted data were only used for spike identification and were not used to present quantitative metrics throughout the manuscript. We considered designating the background as the 5th percentile, similar to Apte et al., (2017), but the differences were negligible.”

Generally: How were isobaric interferences accounted for in the PTR-MS measurements (e.g., for acetone measurements, did you make any assumptions about the contribution of propanal?)?

We thank the reviewer for asking this important question. We could not quantify the contribution from propanal and assumed the signal at $C_3H_7O^+$ was predominantly from acetone. We updated the text in Sect. 2.2.2.

“The influence of isobaric interferences for the compounds of interest could not be quantified. Fluxes of $C_3H_7O^+$ were assumed to be predominantly from acetone, though some influence from propanal was possible. Reports of propanal fluxes in urban areas are sparse. In Beijing, the emission ratio of propanal (relative to acetylene) was ~4 % of the emission ratio of acetone, while background concentrations of propanal were ~16 % of acetone concentrations (Yuan et al., 2012). Together, these comparisons indicate direct emissions of propanal in urban areas are likely minimal relative to acetone, while regional transport or photochemical production could influence the observed $C_3H_7O^+$ mixing ratios by up to 15 to 20 %.”

Results and Discussion:

Line 234: Any speculation as to why there was no significant correlation between OVOC fluxes and temperature? Can you show this explicitly in a figure (perhaps in the SI)?

We think that the main driver of the OVOC fluxes is human activity. Episodic activities (e.g., cooking, VCP use) cause spikes in the flux. There is probably an underlying dependence on temperature and other meteorological factors, but the main driver of the variability we observe is activities. Correlation plots are below for each OVOC of interest. We added this to the SI. Each datapoint represents a one-day average (only daytime data is included). Updated text below.

“We looked for but did not observe a significant correlation between OVOC fluxes and temperature (Fig. SX), which may indicate that the trends in OVOC fluxes are predominantly driven by human activity.”

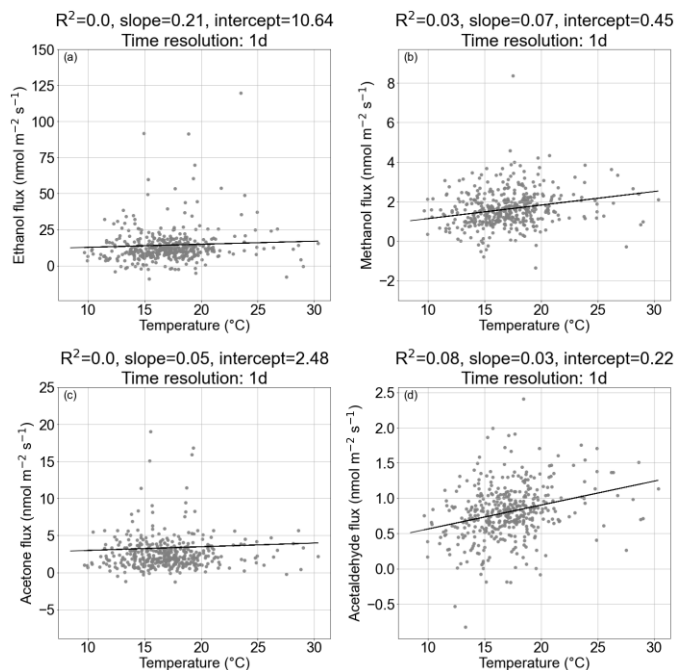
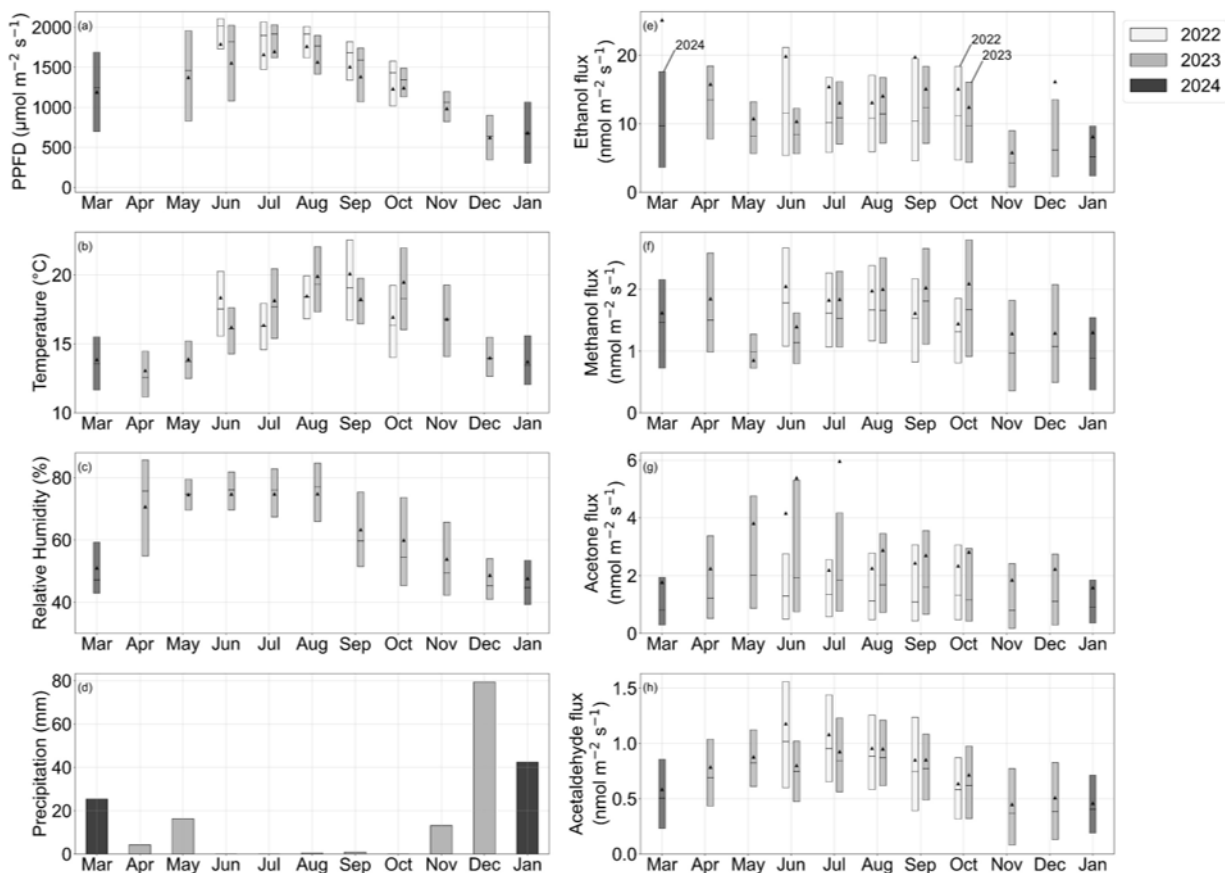


Figure 2: It would help to note the year on top of each bar for readability.

Thank you for the suggestion. This figure is already quite busy, so we decided not to add text to each bar. Instead, we added the annotation to one panel. (See comment below)

Figure 3: I wonder if this could be merged with Figure 4 so that the reader can directly compare the chemical fluxes in Figure 4 with the environmental parameters shown in Figure 3. This would also address my comment above with respect to line 234 (of showing temperature and fluxes explicitly together).

Thank you for the suggestion. The merged figure is shown below. We also adjusted PPFD from the original figure to only show the data during peak sunlight hours (10:00 to 14:00).



For the literature cited for ethanol to nonanal ratios in lines 219 to 323, two of the papers measure indoor mixing ratios and one measures outdoor. While each of these observations is valuable in constraining the possible range of ratios, what caveats need to be considered when interpreting the indoor values when considering outdoor measurements? I.e., other confounding sources of these chemicals either indoors or outdoors?

“Indoors, other sources of nonanal could be present such as surface ozone interactions and skin oil oxidation (Kruza et al., 2017), though we expect their influence to be minor in these studies because the ratios were taken while cooking was actively occurring.”

Generally: can the authors address, in a paragraph, any impacts of the flux tower and mobile measurements not being temporally aligned, and also any considerations regarding the only 12 days of mobile measurements included (and thus some commentary on how representative these are)? If possible, I wonder if some kind of figure similar to Figure S6 with the range of mobile measurements overlaid on the flux tower data would help establish the relationship between the two data sets (in terms of their relative mixing ratio ranges and relative levels between observed chemicals).

We thank the reviewer for the suggestion. We commented on the representativeness of the mobile dataset in Section 2.1.1 and 2.1.2. We added a paragraph to Section 3.1.2 which discusses the range of mixing ratios observed in the two datasets and added the suggested figure to the SI.

“Mobile and tower measurements were not contemporaneous; mobile data collection occurred before the tower was established. Comparing the two datasets assumes that the major OVOC sources sampled were roughly the same between years. Figure S14 compares the range of mixing ratios observed with the mobile laboratory to the flux tower for overlapping months (November, December, January). Generally, there was good agreement between the two datasets for November and December, while tower data from January was lower. On average, mobile laboratory data was 1.1 to 3.6 times higher than tower observations. Since mobile measurements were closer to VOC sources and the sampled airmasses had undergone less dilution, we expected mobile laboratory mixing ratios would be higher than those measured at the flux tower 40 meters above the ground.”

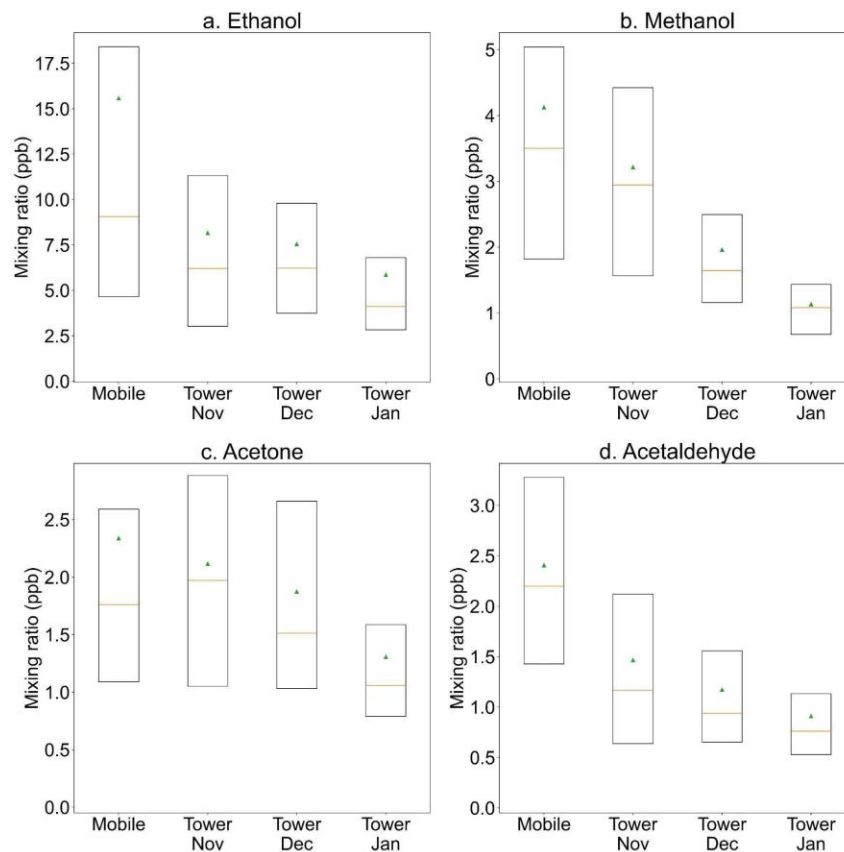
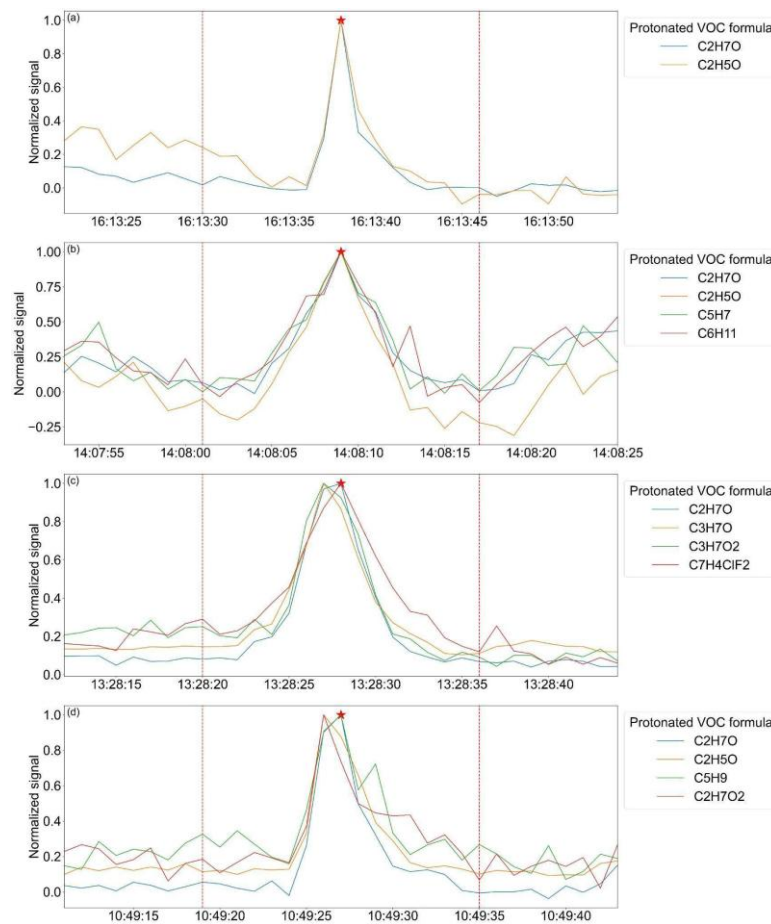


Figure 9: Can the authors provide a few more example time series to show spike correlations? Over what timescale is the co-occurrence quantified (i.e., do the peaks need to co-occur over the course of a few seconds like the SI example shows (Fig. S4)? Few minutes? And does the length of time remain consistent across spikes defined as “correlated”)?

Co-occurrences were found by calculating VOC correlations within a ± 8 second window around the spike peak (Sect. 2.1.3). The same length window was used for all spikes. So, peaks needed to co-occur over the course of a few seconds. This short window was chosen after visually examining concentration spikes in the dataset. We assumed these short spikes corresponded to

driving by a point source or group of point sources. We expanded Figure S4 to display a few more examples of spike correlations.



Line 679-680: This is a really important and big point that feels glossed over at the end (about the inventory/measurement comparison being more accurate for lower population density grid cells and potential associated issues with population density scaling). Can the authors elaborate slightly more here? This sounds like an excellent future work direction.

We thank the reviewer for this suggestion. The second reviewer suggested that this claim needed to be softened, since we did not make measurements in lower population grid cells. We softened the claim and added this sentence to the conclusion section.

“Future work evaluating emissions inventories should consider multiple regions with diverse population and restaurant densities to understand how population scaling impacts the comparisons.”