

Reviewer 2

General Comments

Katz et al. combine a very unique long-term eddy covariance (EC) VOC flux dataset with mobile observations to investigate the sources and spatial variability of ubiquitous OVOCs in Berkeley and Oakland. This is a valuable dataset, and the combination of tower and mobile measurements provides rare and much needed insights into urban VOC emissions and the representativeness of the tower footprint for those that are locally derived with the EC method. The large contribution of ethanol to the measured VOC flux is consistent with observations from other urban areas, and the flux-based source apportionment provides an important complement to the many published concentration-based apportionment studies.

My main concern is that some of the quantitative source apportionment results are presented with more confidence than the current analysis supports. In particular, many estimates rely on inventory- or observation-derived scaling factors and source ratios without fully evaluating the associated uncertainty. The comparison with GRA2PES is also useful, but some of the apparent agreement is not fully independent, and several conclusions about the causes of model-measurement differences seem stronger than the available evidence. Overall, I think the manuscript is suitable for publication after the authors provide additional sensitivity and uncertainty analyses and soften several of the source-specific and inventory-comparison conclusions. I hope the comments below are helpful in strengthening the analysis and interpretation.

We thank the reviewer and appreciate these thoughtful suggestions, which we agree will strengthen the manuscript.

Specific Comments

Lines 98: Some more experimental methods could be helpful here. Can the authors please provide some additional information about the mobile platform and inlet, including the inlet length and location relative to the vehicle, whether the reported tubing diameter is the inner diameter, and the flow rate through the tubing.

We thank the reviewer for the suggestion. We added the following details so Sect. 2.1.1.

“The Vocus sampled ambient air through filtered (47 mm PTFE membrane filters with 2-micron hole size) 0.32 cm inner diameter Teflon tubing at ~2 liters per minute. The inlet tubing was mounted approximately half a meter above the van roof with the inlet facing away from the exhaust. Inlet tubing was ~3.8 meters long.”

Lines 117-118: Some experimental details should also be included here. How were zeros produced (cylinder, heat platinum catalyst, etc.), and were calibrations performed using serial dilutions with multiple concentration steps?

The text says the calibration factors and zeros were stable, but it would be helpful to provide a quantitative value for what that stability threshold was such as typical variability or maximum deviation within a sampling day.

Thank you for the suggestions. We have revised the text to address these comments in Sect 2.1.2.

“Calibrations were performed approximately every 2-4 hours using two gravimetrically prepared cylinders (Apel Riemer). Zero air was generated using a platinum catalyst and was used to dilute air from the cylinders in multipoint calibrations (3-4 points). Calibration factors and zeros were stable (within 10 %) throughout the measurements each day, so one set of calibration factors and instrument zeros were used for each day.”

Lines 120-121: How was the default kPTR value determined, and what compounds were used to derive it? Please also provide an uncertainty for the default sensitivity, like a range of potential values based on observed or representative species.

Thank you for the comment. The sensitivity of well-behaved (i.e., those with minimal to no fragmentation) VOCs in the PTR-MS is linearly related to the proton transfer rate constant (kPTR). Holzinger et al., (2019) (<https://doi.org/10.5194/amt-12-6193-2019>) report a plausible kPTR range of 1.85 to 3.39 cm³ s⁻¹ molecule⁻¹. Cappellin et al., (2012) (<https://doi.org/10.1021/es203985t>) calculated kPTR for an extensive list (n=180) of compounds under variable operating conditions. Depending on the operating conditions, the median kPTR ranges from 2.2 to 2.5 cm³ s⁻¹ molecule⁻¹, and we believe 2.25 cm³ s⁻¹ molecule⁻¹ is the most representative choice for our instrument. Similar approaches using a representative kPTR to estimate sensitivities of uncalibrated VOCs have been applied in PTR-MS measurements with uncertainties of ~ 50 % (e.g., Arata et al. 2021, <https://doi.org/10.1111/ina.12906>).

We agree that an uncertainty should be reported with the default sensitivity. Since the actual kPTR of an uncalibrated VOC can differ from this representative value, we are considering the possible range from Holzinger et al. (2019). We added the following text to Sect. 2.1.2.

“On average, the default sensitivity was 5500 (4600-8400) counts per second (cps) per ppb. The uncertainty range is based on the possible kPTR values for ambient VOCs, 1.85 to 3.39 cm³ s⁻¹ molecule⁻¹, as reported by Holzinger et al. (2019).”

Lines 133-134: How was the 200 x 200 m grid cell size selected, and was the sensitivity of the spatial patterns to this choice evaluated? In other words does this resolution average over important sub-grid variability? The grid size may not strongly affect comparisons using the overall climatological footprint, but it could matter if individual averaging period footprints are considered in this analysis.

On a related note, reporting the 50% relative standard deviation in daily means is good but I also think that there should be a more quantitative presentation of the relationship between high RSD and VOC emission hotspots.

We thank the reviewer for the question and interesting ideas. We're focusing on similar ideas (mobile laboratory methodology for pinpointing direct VOC sources in urban areas) in a separate manuscript. We chose 200 x 200 m grid cells because they are roughly the size of a city block and that level of resolution worked well for comparing the mobile laboratory data to the flux tower data and contextualizing the flux data within the broader urban region. Assessing VOC variability on smaller spatial scales is beyond the scope of this current manuscript, and we did not consider

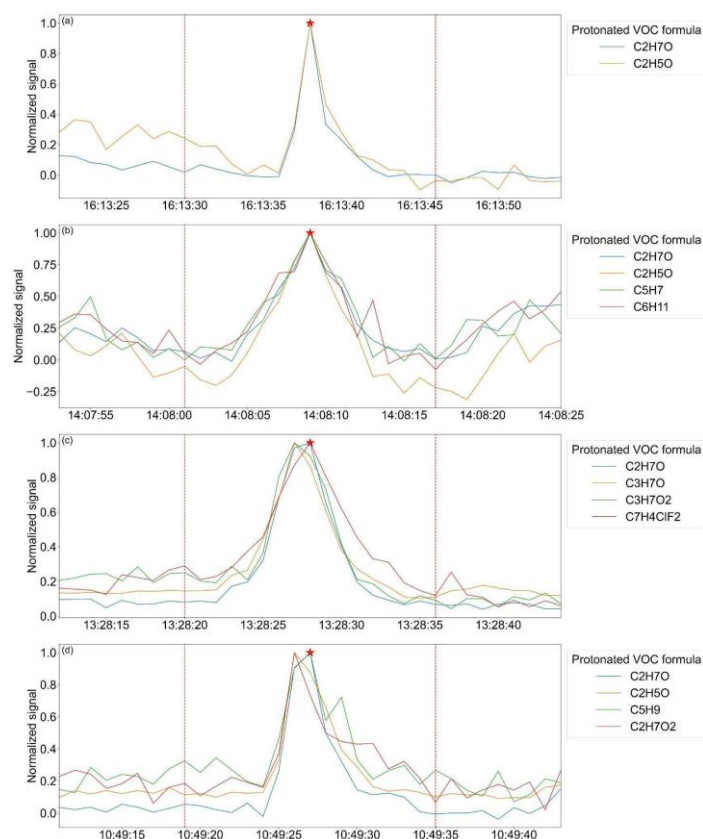
footprints on short time scales, and added the following sentence to indicate why we focused on the average footprint. “Flux footprint modeling, especially in urban environments, can be uncertain, so the intention with this analysis was to show the average geographic area of influence instead of to pinpoint exact emission sources on short time scales.”

Lines 142-144: The spike-correlation analysis is interesting, but correlated signals within a short time window may represent transport of a mixed source plume rather than necessarily direct co-emission, so I suggest using “correlated” or “coincident” spikes unless co-emission can be more directly established. Is there any example figure like Fig. S4 where there are some spikes for a group of compounds with a common source and no spike for a group or VOC from a different source?

Please also clarify whether ions known to be fragments of the same parent compound were excluded from the correlation analysis, since shared fragmentation could otherwise produce high correlations that do not represent separate co-emitted compounds. It seems like correlation analyses were only used for calibrated species known not to be fragments?

We changed the wording to refer to spikes as ‘co-occurrences’ or ‘coincident.’ We also added a few additional examples of spike correlations to Fig. S4 (see below). Known fragments and water clusters were excluded from the spike analysis (e.g., acetone water cluster and small OVOC fragments). Modified text in Section 2.1.3 is below.

“For a given VOC, spike co-occurrences were found by calculating VOC correlations within an 8 second window around each spike. VOCs with correlations (R^2) above 0.85 in this window were considered coincident. Figure S4 demonstrates how VOC correlations were found around each aggregated spike data point. Known fragments and water clusters were omitted from the correlation analysis.”



Lines 152-153: While it is mentioned that additional information is available in the Katz et al. (2025) reference, some experimental details relevant to the EC measurements should be mentioned here. For example, the inlet length, inner diameter of the tubing, flow rate, and approximate inlet residence time should be provided here. Further, what was the calculated high frequency attenuation (Horst, 1997; DOI: 10.1023/A:1000229130034) and random uncertainty (Langford et al., 2015; <https://doi.org/10.5194/amt-8-4197-2015>) for select VOCs?

Thank you for the suggestion; we added more experimental details to Section 2.2.1.

“Instruments were housed in a small room on the roof, and ~36 meters of heated, filtered, Teflon inlet tubing (0.32 cm inner diameter) ran from the top of the tower to the instruments. The tubing flow rate was set to 10 liters per minute, which resulted in residence times under 3 seconds.”

We calculated the flux noise root mean square error (RMSE), which is the RMSE of the covariance spectrum ($w'c'$) at unreasonably high lag times. This is one of the outputs of InnFLUX (Striednig et al. 2020, <https://doi.org/10.5194/amt-13-1447-2020>) and is described as the most conservative estimate of flux LOD. To obtain the RMSE values listed below, we divided the flux RMSE by the measured flux for each 30-minute averaging interval. Median and interquartile range (IQR) values of the resulting random error are Ethanol: 16 % (IQR: 10-26 %), Methanol: 15 % (IQR: 11-24 %), Acetone: 14 % (IQR: 9-23 %), Acetaldehyde: 14 % (IQR: 9-22 %). We added the following text to Sect. 2.2.3.

“The flux noise root mean square error (RMSE), the most conservative estimate of flux noise, for the four OVOs of interest was between 14 and 16 % of the measured flux (median value) (Langford et al., 2015; Striednig et al., 2020).”

To calculate high frequency attenuation, we followed the approach outlined by Striednig et al., (2020) in the InnFLUX manuscript (<https://doi.org/10.5194/amt-13-1447-2020>, see SI section S4). Briefly, in urban environments, the sources of sensible heat flux are distributed differently than VOC sources, which can cause dissimilar co-spectra. We have a sufficiently long dataset to filter co-spectra using strict criteria, and the geometry of our sampling set-up is defined, which allows for transfer functions of the main loss processes to be calculated. We created a model w’C’ co-spectrum using strictly filtered toluene flux data and applied the transfer functions outlined in Section S4 of Striednig et al., (2020). The key transfer functions we used were tubing attenuation (based on inlet geometry) and sensor separation. When applied to our model cospectra, the resulting integral was less than 5 % different than the original, indicating most of the flux was attained. Similar results were described by Striednig et al., (2020).

Line 168: The manuscript assigns 50% uncertainty to compounds using the default sensitivity, but Katz et al. (2025) states that this value was assumed. How was the 50% value determined?

Holzinger et al., (2019) state that uncalibrated compounds can be quantified with 30 % (and up to 50 %) uncertainty. We have chosen the more conservative uncertainty estimate based on the range of possible kPTR values.

Line 171: How large was the isoprene fragmentation correction, either avg. magnitude or % correction during day and night?

The isoprene correction resulted in a 68% reduction in mixing ratio during the day and a 83% reduction at night. The correction resulted in a 64% reduction in flux during the day and a 89% reduction at night. This is also reported in our previous manuscript. We added the following text to Sect. 2.2.2.

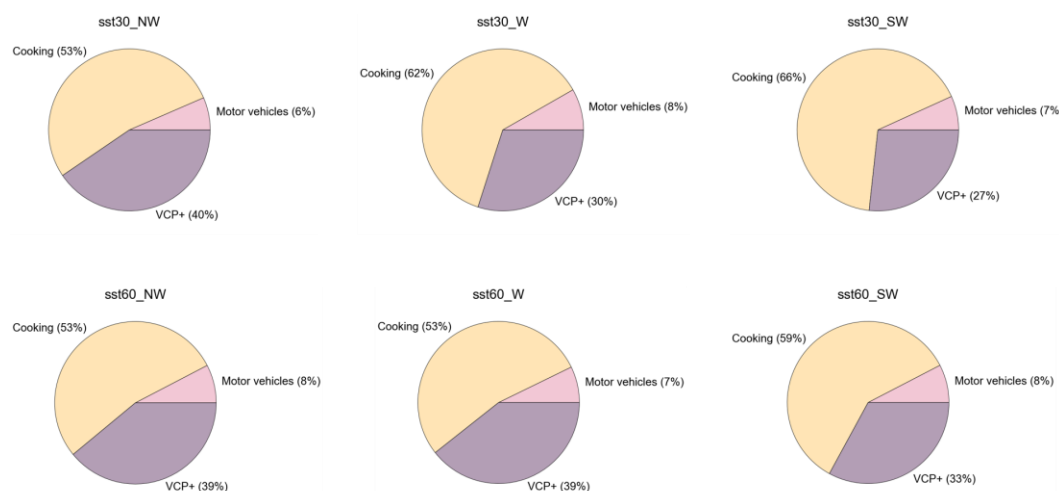
“On average, the isoprene correction resulted in a 68% reduction in mixing ratio during the day and a 83% reduction at night. For fluxes, the correction resulted in a 64% reduction during the day and a 89% reduction at night.”

Also, was interference of ethanol fragments in the acetaldehyde signal quantified? The Coggon et al. (2024b) reference goes through this correction.

We looked for but did not observe the impact of ethanol fragmentation on acetaldehyde in our dataset.

Lines 189-190: Why was a 60% stationarity threshold selected rather than generally recommended 30%? If the covariance changes substantially within an averaging period, the calculated flux may combine different turbulent, footprint, or source conditions and may not represent a single steady-state period. This seems particularly relevant in a heterogeneous urban footprint with intermittent emissions. Are the main source-apportionment results sensitive to a stricter threshold?

The stationarity test resulted in removal of a large fraction of the dataset. The typical 30% stationarity threshold was relaxed to 60% to allow for more data to be included in the analysis. We relaxed this threshold following methods published by other urban eddy covariance studies, e.g., Velasco et al., 2009 (<https://doi.org/10.5194/acp-9-7325-2009>) and Acton et al., 2020 (<https://doi.org/10.5194/acp-20-15101-2020>). The main source apportionment results were not sensitive to the stricter threshold. Below are results from 2022 in different wind sectors with different stationarity tests applied (sst30 indicates the stationarity test had to be below 30%, and sst60 indicates below 60%).



Lines 192-193: What were the typical lag times and did this align with what was expected for a calculated inlet residence time? Were these calculated lag times still robust in the evening and early morning when covariances are typically low and approach the noise?

Lag times agreed with the inlet residence time (~3 s). The lag time analysis in Innflux allows for lag times determined during periods of high flux to be interpolated so they can be applied during times with noisy covariance spectra. Updated text is below.

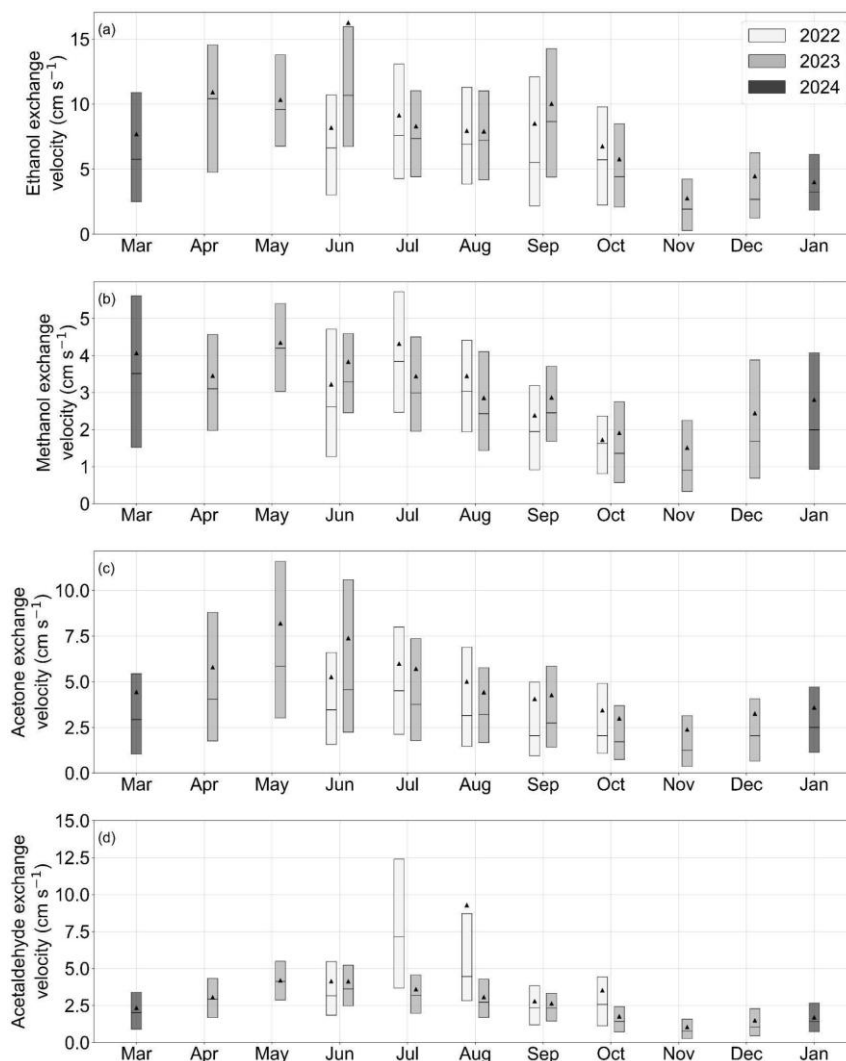
“Lag times were determined for each VOC separately using the maximum covariance between vertical windspeed and mixing ratio for averaging intervals with high signal-to-noise ratios and interpolated across the dataset following a procedure in the InnFLUX code (described by Striednig et al., (2020)).”

Line 266: The authors suggest that lower winter ethanol fluxes may partly result from enhanced wet deposition. Do the authors mean enhanced dry deposition and/or uptake to wet surfaces rather than wet deposition? Could this proposed mechanism be evaluated using exchange velocity, dew point departure, surface wetness, or another available metric?

We thank the reviewer for the suggestion. We calculated exchange velocity for the OVOCs and found that it was generally lower in October through January. This aligns with time periods with the highest mixing ratios (and lowest fluxes). During those months, the source strength is weaker

relative to the atmospheric concentration, implying local emissions are lower. This contrast also implies that the concentrations during those months are decoupled from processes in the flux footprint, and are most likely impacted by processes like boundary layer height and atmospheric transport. We removed the suggestion that lower winter ethanol fluxes may partly result from enhanced wet deposition.

“Lower fluxes during these months were most likely caused by lower local emissions. The exchange velocity (V_{ex}), the ratio of flux to concentration, was lowest in November through January (Figure S10), while mixing ratios were highest during these months. This implies that the source strength in the flux footprint was relatively weaker, and that high mixing ratios were largely decoupled from processes within the flux footprint.”



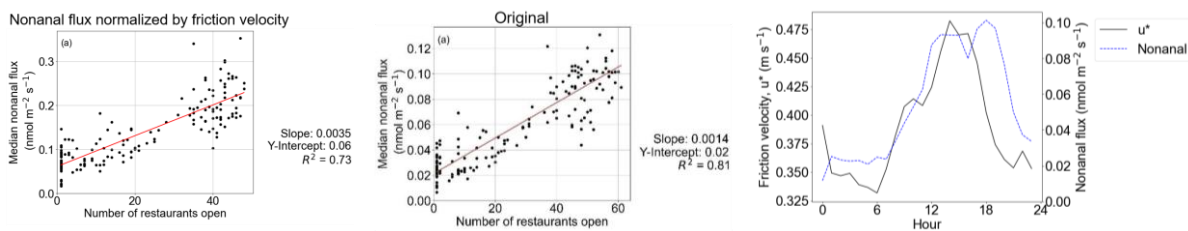
Lines 300-305: The relationship between nonanal flux and the number of open restaurants provides useful supporting evidence, but both quantities have strong diel patterns. I think the diel profile needs to be

removed or accounted for before this correlation is interpreted. For example, the authors could compare anomalies at the same hour on different days or normalize the flux using sensible heat, friction velocity, or another measure of turbulent transport. Further, was the number of open restaurants determined dynamically within each half-hour footprint or within the overall climatological footprint? Those might provide different results and interpretations.

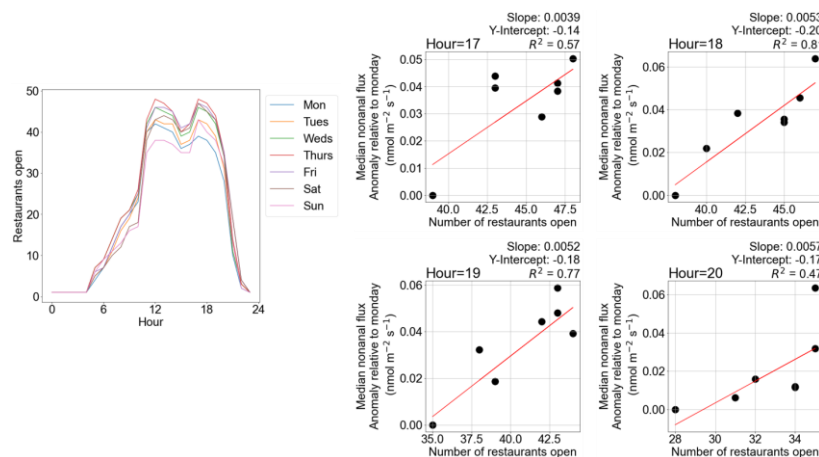
We thank the reviewer for the thoughtful comment. The goal of this plot is to show that there is a relationship between the number of open restaurants and nonanal flux, thus providing more evidence that nonanal is a robust tracer for cooking emissions in our flux footprint. The quantitative results of the fit are not used to determine any of the results reported in the manuscript, so a different fitting approach would not change the results and interpretations. We believe the normalization tests below show that the correlation between nonanal and restaurants is robust and can remain in its current format in the manuscript.

The number of open restaurants was determined for the overall average climatological footprint. We felt this provided a more representative estimate than half-hour averages given that flux footprint modeling can be uncertain, especially in urban areas. Further, most of the results in the manuscript consider averages or medians over longer periods of time which includes the entire footprint. Further, this approach also ensured that we were averaging over all of the different cooking styles present in the flux footprint, providing a more generalizable result.

We did not find evidence that the nonanal flux diel profile is only driven by turbulent transport. The median diurnal profiles of friction velocity and nonanal flux are shown below. Notably, the nonanal flux remains elevated after the friction velocity decreases, which corresponds with dinner hours. Also, the nonanal flux diurnal profile directly matches the trend of the number of open restaurants. We also normalized the nonanal flux by the friction velocity and did not see a change in the trend (see below). There was still a robust correlation between the normalized nonanal flux and the number of restaurants open.



We also examined nonanal flux anomalies at the same hour on different days per the reviewers request. We determined the nonanal flux anomaly relative to the Monday flux, which was the day with the lowest number of restaurants open for dinner hours. There is a clear correlation between the nonanal flux anomaly and the number of restaurants open (see plot below, each dot represents a different day of the week).



Line 326: Did the authors mean that cooking intervals are when ethanol/benzene are greater than 0.01? I would expect cooking periods to be when ethanol is high, but benzene is low, and even then I would think it's much more than 0.01. Please clarify this threshold.

Thank you for catching this. We clarified the text. It was a typo and should have been the benzene/ethanol ratio instead of ethanol/benzene.

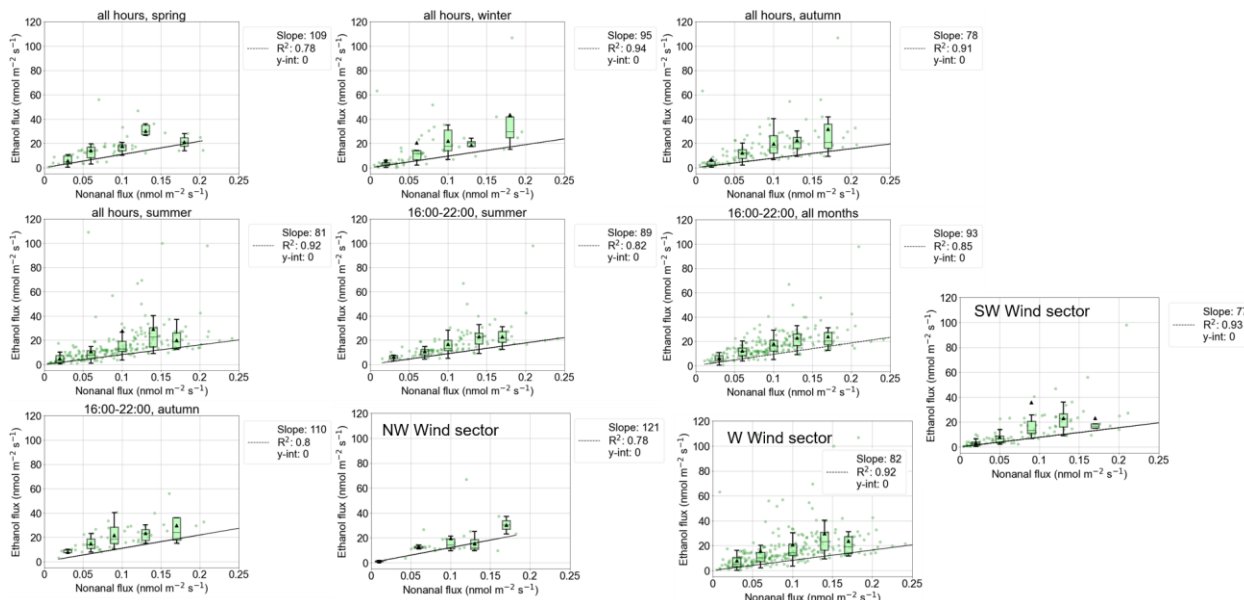
“...when benzene/ethanol was less than 0.01 to avoid influence from on-road sources.”

Line 327-328: The rationale for using the lower portion of the ethanol-nonanal relationship seems reasonable, since additional ethanol sources should only move points above the cooking relationship, and the similar slopes from the 5th to 20th percentiles provide a useful sensitivity test. However, I am still not fully convinced that the lower edge necessarily gives a representative ethanol/nonanal ratio for cooking. Even if the measurement uncertainty, cooking period selection, and binning are all reasonable, couldn't the lower edge still be influenced by differences among cooking sources or by non-cooking sources of nonanal? It would be helpful to show whether the slope is similar across different wind sectors, times of day, seasons, or reasonable cooking selection criteria, or by using quantile regression on the unbinned data. That range could then be used to show how sensitive the reported cooking contribution is to this assumption.

We expect non-cooking sources of nonanal to be extremely minor in the flux footprint relative to the cooking source. These could include skin oil oxidation, indoor surface oxidation, (<https://doi.org/10.1111/ina.12381>) and ocean surface film oxidation (<https://doi.org/10.5194/acp-24-3729-2024>) (the ocean was not in our footprint). By taking the footprint average ethanol vs. nonanal relationship, we include all restaurants and residential cooking and thus provide a representative description of all types of cooking styles and types in the footprint.

We ran a few sensitivity tests according to your suggestions. The original slope in the manuscript was 80, and the binned slope range was 71 to 92. The slope is similar across different wind sectors (W slope=82, SW slope=77, NW slope=121). However, note that NW has much fewer data points available (see figure below), which reduces points that we can use to define the lower

edge, which increases the uncertainty. NW data also therefore includes fewer cooking styles than W and SW. We also performed the fit (binned data fit through 10th percentile) at different times of day and different seasons (see figure below). Under these conditions, the slope ranged from 78 to 110 with an average of 94 and standard deviation of 12. These results show that the fit of the lower edge is stable across different conditions.



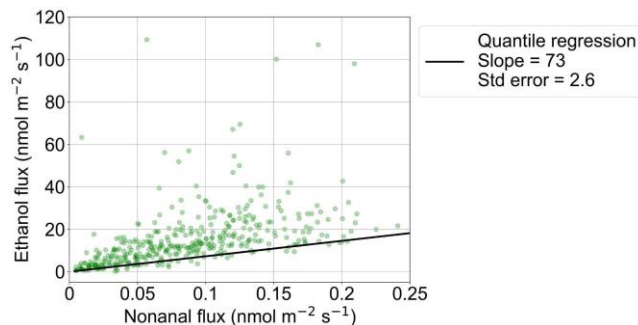
We appreciate the suggestion to use quantile regression on the unbinned data and found this to be a more robust fit. We decided to use this method in the revised manuscript instead of the original approach which involved using binned data. We performed quantile regression on the whole dataset, fitting the 5th percentile through the 20th percentile. With this fitting method, the slopes ranged from 64 to 90 with an average of 77 and standard deviation of 9. Ultimately, we believe that fitting the 10th percentile leads to the most accurate fit of the lower edge of the slope of ethanol versus nonanal, and that including all data (not limiting to wind sectors or certain times of day or season) allows for the most accurate definition of the lower edge and includes the greatest diversity of cooking styles.

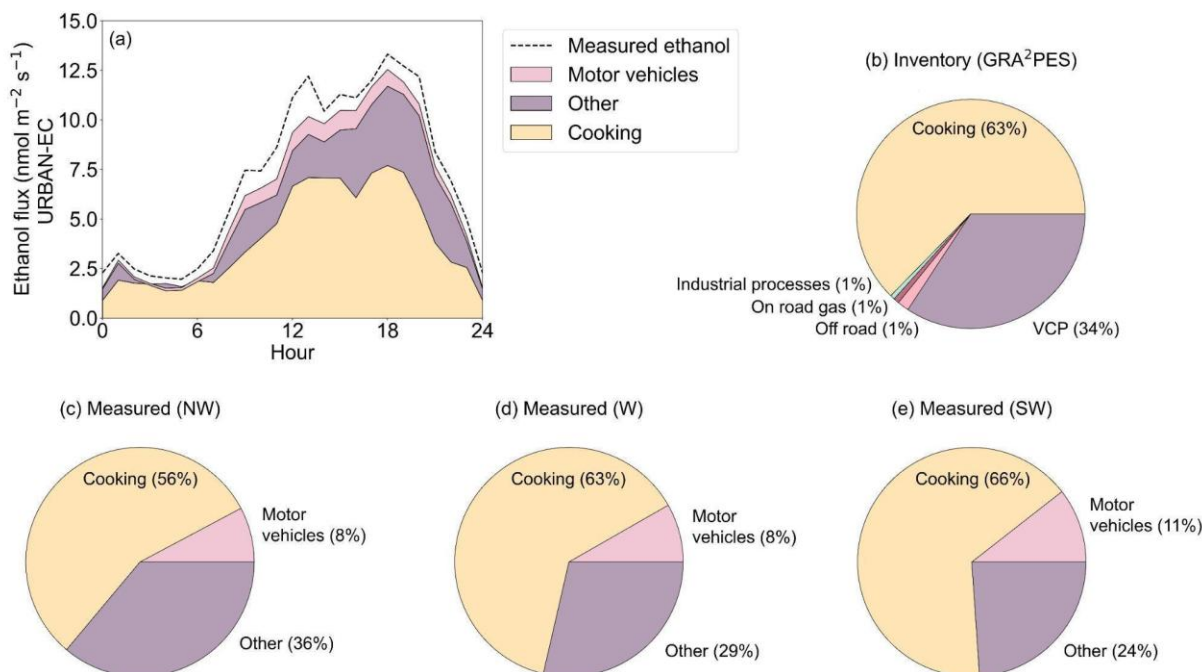
We estimated the uncertainty in the ethanol/nonanal ratio from cooking using bootstrapping. We sampled the dataset with replacement 10,000 times. For each iteration, we fit the ethanol vs. nonanal data using the quantile regression approach (through the 10th percentile). The median slope was 76 (95% confidence interval: 70 to 82). This result is slightly lower than the ratio we reported in the first draft, but it includes a more robust uncertainty estimate and it is closer to the estimate (ethanol/nonanal = 75) from Arata et al., (2021). The updated text, SI figure, and main results figure are below.

“The nonanal to ethanol ratios reported in the literature vary over two orders of magnitude. Ditto et al., (2023), who measured VOC emissions from cooking and cleaning in a restaurant, reported ethanol to nonanal ratios of 479 to 588. The ethanol to nonanal ratio for a cooking profile derived using positive matrix factorization of ambient VOC data was 11 (Coggon et al., 2024a). Also, Arata et al., (2021) measured indoor air at a test house during cooking experiments and reported

an ethanol to nonanal ratio of 75. 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 during active cooking. We estimated $S_{\text{(EtOH_cooking)}}$ by examining the observed relationship between ethanol and nonanal when cooking emissions were sampled. Averaging intervals were assumed to be primarily influenced by cooking when nonanal/octanal was within 20 % of the observed nonanal/octanal ratio during dinner time (17:00-22:00) and when benzene/ethanol was less than 0.01 to avoid influences from on-road sources. The scatter plot of ethanol versus nonanal during these times had a clear lower edge with positive ethanol scatter (Fig. S13). The lower edge without scatter indicates that nonanal is co-emitted with ethanol. Non-cooking sources of nonanal (e.g., surface ozone interactions) were expected to be negligible relative to cooking and likely did not influence the lower edge. $S_{\text{(EtOH_cooking)}}$ was the slope of this lower edge and was found by fitting the 10th percentile of ethanol flux using the quantile regression approach (median $S_{\text{(EtOH_cooking)}}$ = 76, 95 % confidence interval = 70 to 82). The median value and the 95 % confidence interval (CI) were determined by bootstrapping the data. We sampled the data with replacement 10,000 times and calculated the slope of the lower edge of the ethanol versus nonanal scatter plot using quantile regression. Figure S13 shows quantile regression results using the entire dataset.

We tested the sensitivity of $S_{\text{(EtOH_cooking)}}$ to a variety of factors. First, we performed the quantile regression using the 5th to 20th percentiles, which resulted in slopes ranging from 64 to 90 (average = 77, standard deviation = 9). Lower percentiles (20th and below) were chosen to capture the lower edge and because ethanol fluxes were influenced by sources in addition to cooking. We also considered binned linear fits through the 10th percentile of ethanol flux for certain subsets of the data, including selected wind sectors, dinner hours, and seasons. These conditions resulted in slopes ranging from 77 to 121. However, some subsets of the data (e.g., NW wind sector) had fewer data points available, which reduced the points available to define the lower edge, increasing the uncertainty. Considering each iteration, the average slope was 83 and the standard deviation was 12. The ethanol/nonanal ratio we determined through quantile regression and bootstrapping ($S_{\text{(EtOH_cooking)}}$ = 76, 95 % CI = 70 to 82) is consistent with observations from cooking experiments at a test house (ethanol/nonanal = 75, Arata et al., 2021), the GRA²PES inventory (ethanol/nonanal = 110), and average of possible ethanol/nonanal ratios from the additional cases tested from our dataset (ethanol/nonanal = 83). Ethanol/nonanal likely varies with food and cooking methods, so this process was chosen to represent all cooking emissions within this flux footprint. With this scaler-based source apportionment approach, 56 % of the median daytime ethanol flux was attributed to cooking.”





Line 349: The VCP+ category is calculated as the residual after subtracting cooking and vehicle emissions. As described in the manuscript, it may include alcohol consumption, biogenic emissions, unaccounted sources, and errors in the other estimates. I think there should be some presentation of an independent tracer showing that VCPs dominate this term or it should be called “other” or “residual” to avoid being misleading. The statement that these sources are difficult to quantify also seems somewhat inconsistent with presenting this residual as a quantitative VCP contribution later on in the text.

We thank the reviewer for the suggestion. We changed the category to be “other” instead of “VCP+”.

Lines 368-375: The vehicle contribution is estimated using an ethanol/benzene ratio from GRA2PES, and the resulting source apportionment is then compared with the GRA2PES source fractions. The agreement in source fractions is therefore not fully independent, at least for the vehicle contribution, and I think this should be acknowledged.

We thank the reviewer for acknowledging this. The following sentence was added.

“We note that since $F_{\text{(EtOH, mobile)}}$ is derived using the measured benzene flux and the ethanol/benzene ratio from GRA2PES, the comparison is not entirely independent.”

Lines 384-387: The GRA2PES ethanol emission in the lower population northern grid cell is closer to the tower observations, but I do not think this comparison alone demonstrates that the population-density scaler is overestimated. Even if population scaling is the main reason the inventory differs between these

two cells, the tower does not measure the northern cell, so the comparison only shows that a lower modeled emission magnitude agrees better with the observations. I think the language should be softened unless there is other supporting evidence.

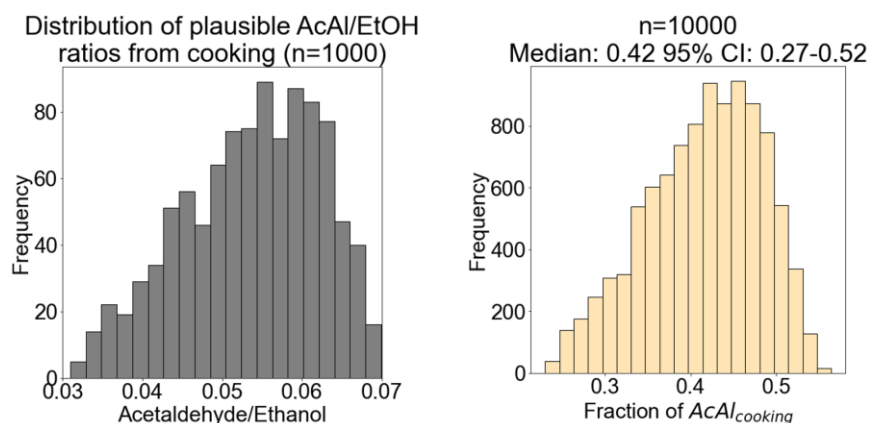
We thank the reviewer for the suggestion, we modified the text.

“This grid cell comparison indicates that population density scalars used to compute ethanol emissions in this region may be currently overestimated. Direct comparisons in neighboring grid cells with variable population densities would confirm this assumption.”

Lines 581-585: The acetaldehyde cooking estimate uses an acetaldehyde/ethanol ratio of 0.06, while the literature values span 0.03-1.33. Can the authors provide more context for this range? For example, how are the reported values distributed, and do they vary systematically with the food or cooking method? The finding that ratios above 0.1 produce more cooking acetaldehyde than the total measured flux provides a useful upper constraint, but it does not by itself establish that 0.06 is the most appropriate value. Since the reported 57% contribution depends directly on this ratio, please show how the estimated cooking fraction changes across a reasonable range of ratios and reflect this uncertainty in the reported contribution.

We thank the reviewer for the suggestion and acknowledge that this uncertainty analysis is necessary. First, we tested acetaldehyde/ethanol ratios ($R_{AcAl/EtOH}$) between 0.03 and 0.1 (lower bound from the literature and upper bound for what makes sense physically with our dataset). Ratios of 0.08 and above resulted in some seasonal total acetaldehyde medians above the total measured flux, so we narrowed the range to 0.03 to 0.07. We performed a bootstrapping + Monte Carlo analysis to estimate the uncertainty in the estimated cooking fraction. Updated text from Sect. 3.4.1 is below.

“To evaluate the uncertainty in this estimate, we performed a bootstrap analysis where observations were sampled with replacement 10,000 times. For each sample, we performed Monte Carlo sampling of $R_{AcAl/EtOH}$ by randomly varying $R_{AcAl/EtOH}$ within its plausible range (Figure S19). The plausible values were selected from a triangular distribution with a mode of 0.06 between 0.03 (lowest ratio from the literature) and 0.07. The range of for cooking in the literature spans three orders of magnitude (0.03 to 1.33) (Arata et al., 2021; Coggon et al., 2024a; Kumar et al., 2025). However, using above 0.08 in the tracer-based source apportionment approach caused the total calculated acetaldehyde flux to be higher than the measured acetaldehyde flux, indicating a lower ratio is more representative of the integrated cooking emissions sampled during this study. Through this method, 42 % (median value, 95 % CI = 27 – 52 %) of the acetaldehyde flux was attributed to cooking.”



Lines 588-589: The combustion contribution also depends on inventory-derived acetaldehyde/benzene ratios. Please provide a reasonable uncertainty range for these ratios and show how it affects the reported 23% combustion contribution.

Unfortunately, the uncertainty is not provided in the GRA²PES inventory.

Lines 593-603: The correlation between the residual acetaldehyde flux and PPFD is interesting, and I appreciate the calculation and consideration of below-sensor gas-phase chemistry, but I do not think it is sufficient to identify photochemical degradation of surface materials as the source. Why was PPFD used rather than a surface UV measurement that is more directly related to the proposed process? Could thermal degradation also contribute to this process, which would be consistent with cooler surfaces in the winter? Is there evidence that ambient aerosols or urban surfaces can produce acetaldehyde under atmospherically relevant light and temperature conditions through heterogeneous chemistry?

Our tower setup only included a PAR sensor, so we are using that as a proxy for overall UV exposure. We acknowledge that other sources are possible, but it is also speculative. We updated the text.

“Instead, photochemical or thermal degradation of materials at the surface or heterogeneous aerosol chemistry could be sources of acetaldehyde.”

Fig. 11: The interannual variability in both the cooking fraction and the magnitude of the acetaldehyde flux is notable. Can the authors provide some potential causes these differences? Are they related to differences in restaurant activity, wind direction and footprint, the estimated ethanol/nonanal ratio, or meteorology? This variability should be discussed because it affects how confidently the campaign-average source fractions can be interpreted.

The average wind direction and footprint were not significantly different between years. The most likely cause of the difference is changes in restaurant activity between 2022 and 2023. However, we do not have the data to prove this. We will acknowledge this difference in the text and speculate the change is from restaurants.

“The fraction of acetaldehyde attributed to cooking was higher in 2023 than 2022, which may be caused by higher restaurant emissions, as meteorological drivers were not significantly different.”

Minor Comments/Technical Corrections

Line 186: This is a really minor comment but the text says meteorology inputs were 20 Hz but the sonic data was actually downsampled to 10 Hz before calculating the covariance, correct?

The InnFLUX code upsamples the 10 Hz data.. Reported that detail in Sect. 2.2.3

“The inputs included 20-Hz sonic anemometer data (3D wind, temperature, pressure, etc.) and upsampled 10-Hz calibrated Vocus data.”

Line 191: There is an extra “all” in “All data is all on Pacific Daylight Time.”

Corrected.

Line 302: Please correct “they was highest” to “were highest”.

Corrected.

Line 355: $F(\text{EtOH}, \text{cooking})$ should be subtracted, not added if this is a residual.

We thank the reviewer for catching this. Here is the corrected formula:

$$F_{\text{EtOH}, \text{VCP}+} = F_{\text{EtOH}} - (F_{\text{EtOH}, \text{FF}} + F_{\text{EtOH}, \text{cooking}})$$

Line 417: I know what the authors mean but it may be more straightforward to say a 200 x 200 m grid rather than “200 square meter grid” since it’s technically not 200 m².

Corrected.

Lines 421-422: The sentence stating that the tower could be “missing key sources of ethanol” could be made more clear. I think the intended point is that the footprint does not include some important regional hotspots, rather than that the local flux measurement is somehow incomplete.

Thanks, the corrected sentence is below.

“Importantly, the highest average ethanol concentrations observed via mobile monitoring were not within the flux footprint, indicating the flux observations could have missed key regional hotspots of ethanol emission (described below in Sect. 3.1.3).”