We thank the reviewers for their time and constructive comments that have improved our manuscript. Below we include specific responses to the reviewer's comments. The reviewer's comments are in black. Authors' responses are in blue, quotes from the manuscript are in *italic*, and changes to the text are shown in red.

Bourgeois et al. presented comprehensive intercomparisons of airborne NO, NO2, HONO, NOy and CO in biomass burning plumes, each measured with differing techniques during FIREX-AQ in the summer of 2019. This study provides valuable dataset and the evaluation of accuracies of major techniques deployed in the challenging biomass burning plume conditions. Additional literature review on these species from major airborne field campaigns are helpful for understanding the accuracy of these measurements under different environmental conditions. The manuscript was written thoroughly, and the figures are made clear. Thus I recommend acceptance after revision. Below are my comments:

1. Line 204, hourly calibration of NO LIF was performed with [NO] 4-20 ppbv, did this concentration range apply for all the smoke conditions? How do you ensure the linear response beyond this range?

Response: We thank the reviewer for raising this point. As discussed in Rollins et al (2020), given the sensitivity typically observed during FIREX, nonlinearity associated with saturation of the LIF instrument is not problematic until mixing ratios well above 100 ppbv are encountered.
Additionally, linear response to mixing ratios up to 100 ppb has been tested in the laboratory.

We clarified this point in the manuscript lines 217-220:

"*As discussed in Rollins et al (2020), given the sensitivity typically observed during FIREX, nonlinearity associated with saturation of the LIF instrument is not problematic until mixing ratios well above 100 ppbv are encountered.*"

2. Lines 508-510, "Trajectories and ages that were grossly inconsistent with smoke transport patterns seen in geostationary satellite images were excluded from further analysis". Which group should these data categorized into.

Response: This question is not clear. As stated in the manuscript, these data were simply excluded.

3. Lines 648-649, what is the p-value of Figure S4 b and d, any explanation for the seemingly dependence of the difference on NO2 concentration?

Response: Few data points actually contributed to the pattern identified by the reviewer. We realized that this figure may be misleading and we replaced it with a new figure (now Figure S6) showing the same plots but color coded by data density.

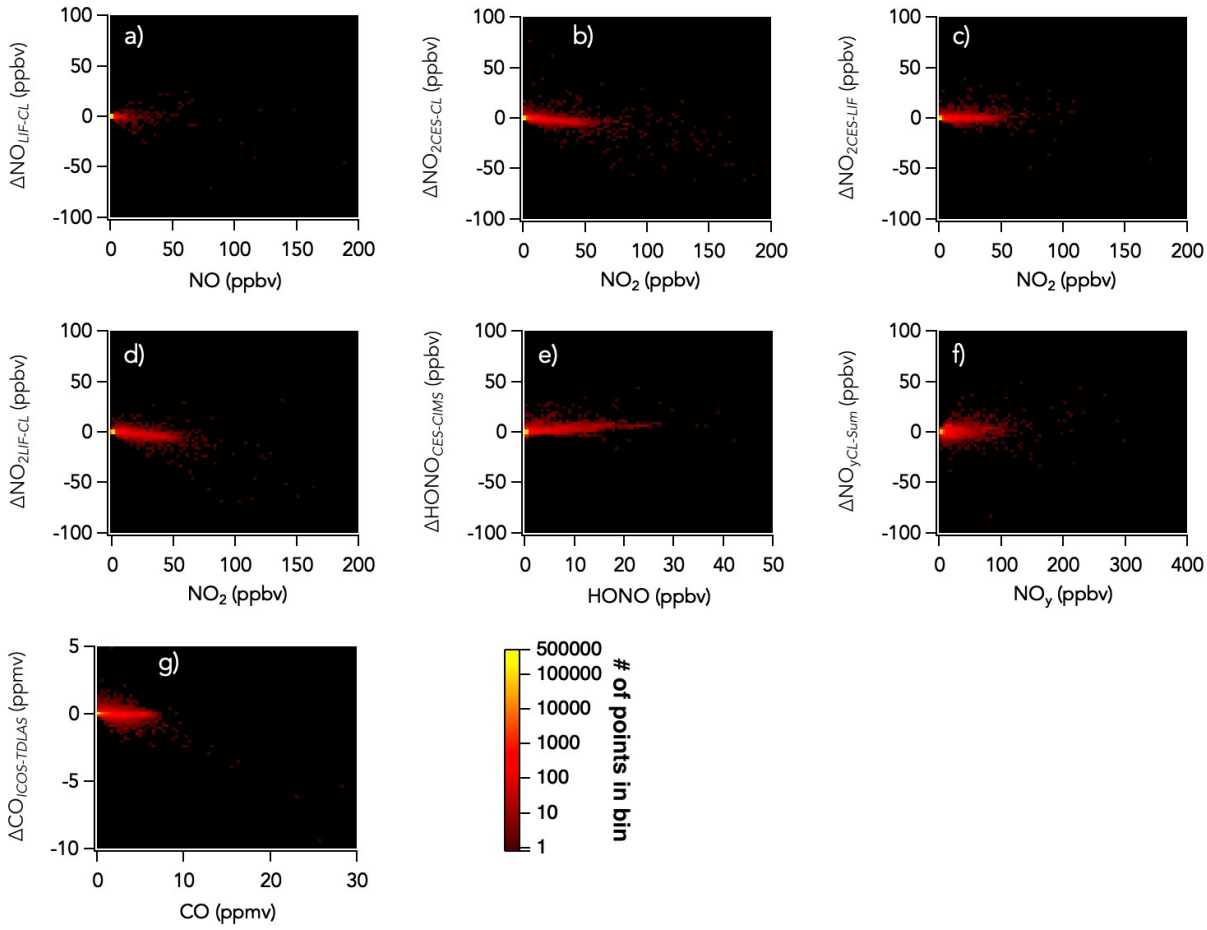

Figure S6 Measurement differences (1Hz data) of a) NO, b)–d) NO_2, e) HONO, f) NO_y, g) CO as a function of the species mixing ratios for the entire campaign. The color bar indicates the number of individual data points per bin of mixing ratios (bin size is 2.5×2.5 ppbv).

4. Figure 3 and Figure 4, no letter label (e.g., a to e) was assigned to any of the panel.

Response: Fixed.

5. In section 3.3.1, intercomparison between CES and CIMS measured HONO were presented. I have the following questions: 1) The slopes shown in Figure 8 suggests CES HONO was higher than CIMS HONO. However, it seems neither the flight averages of the absolute difference shown in Figure S9, nor the histograms of the absolute difference between the two methods suggest the CES-HONO > CIMS-HONO. Any explanation? 2) In Figure S9, why are there many missing points for intercepts (middle panel) and slopes (bottom panel), while the top panel (mean absolute difference) shows all the data on each sampling day? 3) it is interesting to see the measurement of HONO with CIMS are significantly affected by temperature, especially above 30°C, as is shown in Figure S10. Would the slope of CES-HONO vs CIMS-HONO be closer to 1 since it's not shown in this figure? 4) Could the inlets for the two methods be an issue that cause the discrepancy during FIREX-AQ?

Response: We thank the reviewer for raising these points.

1) As stated in the manuscript lines 581-585, the regression analysis (as presented in Figure 8) yields slightly different information than the calculation of the difference: while the former is weighted more by fire plumes, where mixing ratios were greatest, the latter is weighted more by background conditions, where most of the measurements took place. In background conditions, HONO mixing ratios were typically lower than the precision of the CES measurement, yielding a $\Delta HONO_{CES–CIMS}$ close to 0 on average (as reflected in Figures S9 (now Figure S11) and 5e).

2) We thank the reviewer for catching this. We fixed this issue in Figure S9 (now Figure S11) and in the other similar figures as well.

3) In Figure S10 (now Figure S12), the slope is closer to 1 at lower temperatures (slope of 1.3 at 33°C) than at higher temperature (slope of 3.9 at 38°C). A full description of the correction applied to the CIMS HONO data is provided in a follow-up paper that has been submitted to *AMTD* (Robinson et al., 2022)

4) Inlets are unlikely to be the issue. Please see responses above and note that the temperature dependence of the IMR as documented in a separate publication (Robinson et al. 2022) explains the difference without need to invoke inlet effects.

6. Do the measurements shown in Figure 10 (a) include both fresh smoke and aged smoke? If so, what if the fresh smoke and aged smoke were separately considered? Will the relative contribution of each NOy be significantly different? Are the large uncertainties associated with NO2, APNs and pNO3- driven by flight-to-flight difference, secondary processing, or environmental conditions (humidity and temperature)? What could be possible causes for the different contributions of major species (e.g. NO2, APNs and pNO3-) between western wildfires and eastern agriculture fires?

Response: Measurements in Figure 10 include both fresh and aged smoke. The separation of aged and fresh smoke and associated NO_y budget is presented in Figure S11 (now Figure S13). The large range of contribution of NO_2, APNs and pNO_3 is largely due to the wide range of photochemical conditions sampled during FIREX-AQ – as mentioned before, Figure 10 includes both aged and fresh smoke. The difference in NO_y budget between wildfires and eastern fires may be due to i) a difference in the photochemical aging of the smoke. Most eastern fires produced a thin and dilute smoke plume that was samples close to the fire whereas wildfires usually produced wide and thick plumes that were sampled both close to and further away from the fire; ii) a difference in the fuel. Eastern fires typically consisted of burned crops whereas wildfires fuel consisted of trees and grass.

7. In Figure S11(a), from the slopes determined for fresh versus aged smoke, can we say the sum of NOy outweigh CL-NOy for fresh smoke and the CL-NOy outweigh the sum of NOy, although the difference is within the combined instrumental uncertainties? If so what would the explanation be?

Response: The main difference in the NO_y budget between aged and fresh smoke is that pNO_3 becomes the main component of NO_y in aged smoke (Figure S13). Therefore, higher ΣNO_y than measured NO_y in aged smoke may be explained by the non-quantitative sampling of pNO_3 in the NO_y instrument, as detailed in the section S1 of the SI. In fresh smoke, pNO_3 is a smaller component of NO_y, and non-quantitative sampling of pNO_3 in the CL instrument may

have less impact on the comparison. We added a sentence reflecting this discussion in the text lines 820-827.

"*The variability in the ΣNO_y to NO_y correlation slope between aged and fresh smoke (Figure S13a) likely illustrates the non-quantitative sampling of pNO_3 in the NO_y instrument. Indeed, higher ΣNO_y than measured NO_y in aged smoke (slope of 1.05), where pNO_3 is one of the main components of ΣNO_y (Figure S13b), may be explained by the non-quantitative sampling of pNO_3 in the NO_y instrument. In fresh smoke, pNO_3 is a smaller component of NO_y, and non-quantitative sampling of pNO_3 in the CL instrument may have less impact on the comparison (slope of 0.98).*"

8. Lines 732-734 described what different NOy measurements were used to calculate total NOy. While I understand the choices are based on precision, I wonder why CIMS HONO instead of CES HONO was chosen, as CIMS HONO underestimated CES HONO and its accuracy seems to be significantly affected by temperature variation as is discussed in 3.3.1?

Response: The CIMS HONO measurements were used because they have much better precision than the CES HONO. Precision matters more than accuracy for plots like Figure 5, that are more weighted by background data than by smoke data. Using CES HONO rather than CIMS HONO only affects the slope of the correlation between measured NO_y and ΣNO_y by 6%. We clarified this aspect lines 370-372.

"*Using LIF NO, CES NO_2 and CES HONO as primary measurements changed the correlation slope between ΣNO_y and measured NO_y by -2%, -6% and 6%, respectively (Table S1).*"

9. Lines 747-779 are difficult to follow. Figure 12(a) should be well explained first followed by Figure 12 (b). The current order is reversed, and I don't quite get the idea of Figure 12 (a). For Figure 12 (b), it is unclear how the missing NOy fractions (bottom panel) were calculated. My understanding is that fraction of each individual NOy to total NOy was calculated from the individual measurements and sum of NOy, then particle sampling fraction was calculated from the model. Combining the two pieces will enable the quantification of missing NOy (0-24%) resulting from the CL-technique, but how? Thus, further clarification will be needed. Also, in section 3.4.1, it is interesting to see the possible reasons that cause the negative and positive mode of the discrepancy between CL-NOy and sum of NOy. The authors separated the two modes and interpreted them separately. However, if one reason is important (e.g. pNO3- loss through the CL inlet), it should be important throughout the entire campaign, instead of certain period. I might miss something, but a clarification would be helpful.

Response: The reference to Figure 12a (line 774 of the previous version of the manuscript – now line 840) was actually a typo and should have been Figure 12b. Now the discussion first discusses Figure 12a, then Figure 12b.

We calculate the missing NOy according to the following equation:

$$\text{Missing } NO_y = ((1 - \text{particle sampling fraction}) \times pNO_3) \,/\, NO_y$$

We added that equation to the text lines 828-835 to clarify the calculation of missing NO_y.

"*We calculated the fraction of measured NO_y in smoke initially attributed to pNO_3 that may result from other reactive nitrogen species than those included in the ΣNO_y according to equation 3:*

$$\textit{Missing } NO_y \textit{ fraction} = ((1 - \textit{particle sampling fraction}) \times pNO_3) \,/\, NO_y \quad \textit{(Eq. 3)}$$

Where particle sampling fraction corresponds to the modelled pNO_3 sampling fraction in the NO_y inlet. We found that missing NO_y accounted for 0–24% of the measured NO_y in smoke (assuming a sampled air speed 65% that of the aircraft; Figure 12b)."

We respectfully disagree with the reviewer that pNO_3 loss through the inlet should be important through the entire campaign. As stated in the manuscript lines 812-813, "*Particle sampling through the NO_y inlet is highly dependent on altitude, air speed (see section S1 and Figure SB) and pNO_3 mass size distribution (Figure 12a)*". Also, see our previous response to comment #7 on the effect of aged vs fresh smoke on NO_y closure and clarifications added to the text.

10. In section 3.5.1, it was noted the cause of the discrepancy between ICOS and TDLAS measured CO was unclear. I am curious whether temperature plays a role? Additionally, Figure 14(a) shows when CO goes above 10 ppmv, ICOS seems to outweigh TDLAS; as CO is higher the deviation from 1:1 line is larger. What are the possible explanations?

Response: We interpret this effect as the ICOS having a slower time response than the TDLAS instrument, which is most noticeable when the mixing ratio is high and the plume width is narrow.

References:

Robinson, M. A., Neuman, J. A., Huey, L. G., Roberts, J. M., Brown, S. S., and Veres, P. R.: Temperature dependent sensitivity of iodide chemical ionization mass spectrometers, EGUsphere, 1–17, https://doi.org/10.5194/amt-2022-295, 2022.