General Comments: Evaluating the overall quality of the preprint

Overall, the preprint is well-written and easy to read. A technical study assessing the long term, tower-based capabilities of an Aeris MIRA instrument for methane and ethane measurements is a valuable contribution to scientific literature. Except for several areas that need more clarity, instrument characterizations appear to be scientifically sound and of relevance for future field studies. My primary concerns in this preprint are the insufficient field assessment for ethane measurements and why some instruments did not undergo all or any of the laboratory tests. Additionally, there are several areas that need more clarity in the paper. In short, I think this paper is quite interesting and captures the potential for MIRA-based methane measurements in long-term monitoring, but more work needs to be done to assess the potential capability of long-term MIRA-based ethane measurements.

Specific Comments: Addressing individual scientific questions/issues

The introduction is easy to read with appropriate background information included. The introduction (line 61-62) and the title of the paper suggest that a large part of the novelty of this study is the laboratory and field assessment of MIRA Ultra for a long-term, tower-based network of both dry methane and ethane. The author understands that the collocated measurements of methane and ethane are important for distinguishing thermogenic methane emissions from total methane emissions. The preprint, however, does not adequately demonstrate confident tower-based ethane measurements. I recommend modifying language in the intro, discussion(line 514), and/or title, so they are more reflective of the results and analysis performed.

Authors perform technical and practical analysis, addressing field-based concerns such as time synchronization (which is problematic with the Aeris instrument) and cold/warm start delays. The preprint can benefit with some clarity on their methodology/results in the following areas:

- **Section 2.2: bias and precision goals:** The primary take away in this section is the reference precision goal, however, it is unclear how the authors derive 3ppb as the methane goal. < 10% of the typical enhancement should be 4ppm (based on line 488). Overall, this section can be written more clearly as well as concisely.
- **Why some instruments were chosen for tests/ studies:** An explanation should be provided why the authors did not perform all tests described in Table 1 for all instruments. The authors state they have 8 unique instruments used in the study, but Table 1 only describes 5 unique serial numbers (665, 792, 800, 886, and 778). If some instruments were upgraded by the manufacturer and returned with the same serial number, those instruments can be referenced as A792.v1 and A792.v2 or another shorthand to make it easier to follow.

It's worth noting that instruments A792 (used in 4 tests) and A665 (used in 3 tests and field deployment) were the most tested instruments and 792 was the sole instrument involved in determining *Uncertainty due to instrument noise, cylinder calibration, and ethane cylinder assignment uncertainty,* meanwhile A665 and A792 showed the greatest sensitivity to water vapor (lines 275-278) and the most unrealist deviations (lines 382-384). During the field study why would the authors choose to use A778 (which was involved in no other Table 1 tests) and A665 that had the most unrealistic deviations and the greatest sensitivity to water vapor? Given that your results often show instrument dependent characteristics, how do you justify using test results from other instruments to perform the field assessments.

- **Calibration cylinder usage and description:** Line 160-162 is confusing. It reads as though there are 17 cylinders and each calibration cycle includes 4 min for each cylinder, however, the paper says this is a 16-minute process. Does that mean only 4 cylinders are used? This needs to be clarified. Why was the test repeated for 8-16 hours? That's a large range.

  In test three, why did you include all calibration cycles (line 166), when the cylinders did not stabilize for the first few hours? Shouldn't the 'NOAA C2H6' column also include the cylinder assignment error?

  I noticed that your NOAA tertiary standards range from 1985.9-2284.7ppb methane and 1.3 and 22.9 ppb ethane. In test four, the maximum methane values exceeded 2350 ppb and the minimum ethane values were below 0.5 ppb (line 222). Do you generally trust the instrument's response outside the calibrated range, and have you checked how often your field measurements fall outside of the range?
- **Water Vapor corrections:** What does a perfect water correction mean (135-136)?
- **Field Deployment/Design:** Is a 20-minute lag time common for these sorts of measurements? I don't think it is correct to assume large variations in Aeris or Picarro measurements are from mismatched timing alone (lines 258-260). The author needs to add more support on why and how they chose to eliminate time series data points when either the Picarro or Aeris standard deviation exceeds the fiftieth percentiles. What is the local time period you are running these standard deviations? Can you site studies that employ a comparable methodology- it just seems a bit arbitrary?

  In Figure 9, I would recommend re-evaluating the Aeris data in early May and right after the version 2 switch that looks like lines - I suspect this is unreal data. The x-axis should likely be date (DD/YY) or something more description. It looks like Picarro data has more

missing datapoints, particularly during sharp peaks where the Picarro indicates a point at the maximum and the Aeris data is tracking multiple points along the enhancement peak.

- **Allan Deviation:** The allan test appears to be run for 40 minutes for V2 of A665, A800, and A886 and for 5 hours for A792. I am not convinced that A792 is the most representative instrument beyond 5 minutes, and I worry this study is putting too much emphasis on results from a single instrument, when there is a large variation between individual instrument performance and sensitivities.

Overall, I think the information, tables and figures are relevant for the main text, yet I am not sure it is worth including in such detail that standard deviation reduces as averaging time increases. Lines 232-234, 385-394, and Figure 8, are not particularly novel or necessary to include in the main text.

The discussion addressed some crucial concerns about field deployment of the MIRA instrument for long-term methane and ethane measurements. Using your logic in lines 486-496, the bias threshold (section 2.2) for each tower network should be dependent on the expected methane enhancement (e.g. Indianapolis should strive for an uncertainty threshold below 0.5ppb methane) and thus make this sort of analysis not practical outside of a large metro area or are with significant O&G operations. You addressed this issue similarly in lines 489-492. My main concern with your discussion is that you say "the [Aeris MIRA Ultra measurement] system shows promise for distinguishing among multiple methane sources by providing continuous ethane measurements, depending on the magnitudes of methane and ethane emissions". While showing promise is vague and the second part of your sentence creates a wide caveat, I think this is too strong of a statement for the lack of ethane results during the field deployment.

Technical Corrections: typos, etc.

- Line 19-31: Myhre et al., 2013 is an outdated source. I would recommend referencing the latest IPCC report.
- "also" used twice in the last sentence of section 2.2 (line 121-123)
- Subscripts for methane and ethane in table 2 display in the midline of text and uncertainty terms (like U_t) do not have subscripts ("Ut")
- Remove or replace "obviously" line 191
- Figure 2: "Latitude" y axis label is off center; caption and legend should clarify what the Oil & Gas data points are indicating. I would assume they are active sites during the time

of study but I'm not sure. On the tower illustrations it would be helpful if you added the approximate location of the picaro and the Aeris from the inlet line.

- Figure 5 and 6: subplots e and f should have the same x-label; Additionally, descriptions of subplots c and d can be worded clearer
- Figure 7: This figure would look more appealing if the zero horizontal line were aligned between subplot A and B. Thus, plot A y axis would range from 0+/-x and plot B y axis would range 0+/-z.
- Table 3: what does "typical" mean. In the top half of the table ()* means noise and in the bottom half of the table () means precisions. If that is not correct, please change symbols to be clearer.