Reviewer #1:

We thank the review for their time in reading our paper and offering their thoughts and suggestions. The reviewer's comments/suggestions are in regular black text, and *our replies are in the blue italics.*

This study compares seven collocated infrared spectrometers over a month-long deployment to evaluate consistency in radiance measurements and retrieved thermodynamic profiles.

The instruments agree within 1% in radiance, display minimal spectral calibration differences, and produce temperature, humidity, and derived quantities (such as PWV and boundary-layer height) that match well within expected uncertainties.

General Comments

This is a clear, rigorous, and well-presented intercomparison study that convincingly demonstrates the reliability of modern IRS instruments for network applications.

The scientific approach is sound, the analysis thorough, and the presentation is clear and well structured. I find the work robust and valuable, and would recommend publication subject to minor technical corrections.

Technical corrections

r160 has -> have

Done

r162 provides -> provide

Done.

r182 missing Fig. number
We meant to reference Figure 6. That has been fixed

r218 TROPoe acronym is already defined in r71

We have removed that redundancy.

Reviewer #2:

We thank the review for their time in reading our paper and offering their thoughts and suggestions. The reviewer's comments/suggestions are in regular black text, and *our replies are in the blue italics.*

General Comments

This study provides an intercomparison between seven collocated ground-based infrared spectrometers (ASSIST systems) during a one-month period in the fall of 2023 in Boulder, CO. The purpose of this study is to show that identical instruments (i.e., ASSISTs) produce the same observations (within an instruments' uncertainty) and thus can be used/interpreted interchangeably when deployed as a network system. This is an important assumption to confirm in order to discern whether differences within an observational network are associated with atmospheric changes and not instrument characteristic differences. This work also provides further confidence in the use of infrared spectrometers for planetary boundary layer observational work, with real world potential applications particularly with estimating advection.

The instruments' measured infrared radiances agreed to within 1% of ambient radiance across both opaque and transparent spectral channels, and their spectral calibration differences were quite small (≤7.1 ppm). Using the TROPoe optimal estimation retrieval, temperature and water vapor profiles derived from each instrument showed near-zero mean biases and excellent agreement well within retrieval uncertainties under both clear sky and cloudy conditions. Derived quantities such as precipitable water vapor and planetary boundary layer height also closely matched across all systems. These results demonstrate high precision and repeatability of both the ASSIST instruments and the TROPoe retrieval framework, supporting their suitability for deployment in profiling networks where observed spatial differences can be confidently attributed to real atmospheric variability rather than instrumental or retrieval artifacts.

The presentation of this work, along with the methods used, are sound and well presented. I would recommend publication after minor revisions, which are mostly technical in nature.

Specific Comments

1. Line 72: Suggest rephrase to 'TROPoe uses the line by line radiative transfer model (LBLTRM; Clough et al. 2004)...'

We changed the phrase to "TROPoe uses the line-by-line radiative transfer model LBLRTM (Clough et al. 2004)…" because LBLRTM is actually its name, not an acronym (as per the LBLRTM lead Dr. Eli Mlawer, personal communication)

2. Lines 181-193: You point out the 'spikiness' in Fig. 6, associated with the sides of absorption lines and different residual instrument response functions. However, it isn't clear how this would effect the end user or how you address this (and the other instrument differences in this paragraph) in the following sections. You show clear agreement among retrievals, but don't close the loop of how the differences highlighted here correspond to the final product.
   Good point. We have added this sentence in the conclusion of the paper: "Importantly, the slight differences seen in Fig. 6 (i.e., the spikes in the spectral differences) associated with small differences in accounting for the variability in the response function from instrument to instrument, which was discussed in section 4, has no appreciable impact on the retrieved thermodynamic profiles."

3. This intercomparison study was completed in a relatively dry environment. Would you expect the instruments and/or retrievals to behave similarly in a humid region, such as in the deep south? I think this would be a beneficial piece of discussion for the conclusions section.
   Another good suggestion. We have added this short paragraph to the conclusion to address it. "This comparison was performed in the autumn at a relatively high elevation site in Boulder, Colorado, and the range of water vapor was relatively dry with the maximum PWV of approximately 1.9 cm. A natural question is: "Would these results be the same if the same comparison was performed in a much more humid environment with larger PWV values?" None of the calibration parameters (i.e., table 1) depend on the actual environment (i.e., they are determined in the laboratory before the instrument is deployed), and thus the question simplifies to how do the uncertainties in these parameters impact the calibrated radiance as the environment changes. This question is addressed using Fig 3: an environment with larger PWV will both be warmer (i.e., the ambient blackbody is warmer) and the amount of extrapolation to get to the scene's radiance will be smaller. Thus, the largest radiometric calibration uncertainty occurs when the day is warm but the PWV is small. Furthermore, the spectral calibration does not depend on the environment at all. Thus, it is expected that the agreement in the thermodynamic profiles would be essentially the same in a warmer, moister environment."

Technical Comments

1. The following references are missing from the reference list:
    - National Research Council 2009
    - Illingworth et al. 2018
    - Michaud-Belleau et al. 2025
    - Gero and Turner 2011
    - Shupe et al. 2013
    - Turner and Mlawer 2010
    - Adler et al. 2013 -- *this is Adler et al. 2023*
    - Mlawer et al. 2024
    - Minnett et al. 2001
    - Clough et al. 2004 – *this is Clough et al. 2005*
    - Hu et al. 2019

    We apologize for neglecting to add these to the list of references! They have been added.

2. The Sisterson et al. 2016 reference in the reference list is not referenced in the paper.
   This has been removed from the reference list
3. Line 182: Figure number is missing.
   It has been added (we were referring to Fig 6)
4. Line 218: TROPoe is already defined, no need to redefine
   This has been modified to no longer spell out the acronym here.
5. Figures: Consider using colorblind-friendly choices, particularly in Fig. 3 where red/green are used together in a way that would make it hard for an impaired individual to tell the difference between the HBB temperature and HBB emissivity apart.
   Good suggestion – we have updated the colors accordingly.
6. Figs. 4 and 5: Update with higher resolution, it appears fuzzy.
   We will be uploading EPS (encapsulated postscript) versions of each image, which will allow the printed images to be crisp.

Reviewer #3:

We thank the review for their time in reading our paper and offering their thoughts and suggestions. The reviewer's comments/suggestions are in regular black text, and *our replies are in the blue italics.*

This manuscript presents an intercomparison of seven co-located infrared spectrometer observations. The work is very valuable as the relative calibration and random uncertainties between different instruments proved to be very small, thus allowing a network use of these instruments.

General comments:

The manuscript is well written and structured. Also, concise and clear language is used.

Minor comments:

Abstract, line 27: you mean "quantities", not "qualities", I guess?

Yes, that is correct. We have fixed it.

Line 195: check phrase ("observed" is repeated)

We have removed the redundancy.

Line 253, and Fig. 10 (right): Do you have an idea why ASSIST-20 has a larger disagreement in the lower layers? If so, could you give a short explanation in the text?

We do not. We presume that the small difference in the water vapor in the A18 (which makes the differences with the other ASSISTs) is due to some small correlated error in A18, but we have not been able to identify it. Thus, we have not modified the text.

Line 266: PWV uncertainties are 0.1-0.3 mm (not 1-3 mm) according to Fig. 12

Thank you – this is an important catch! We have updated the text accordingly.

Figure 1 (and 3): ASSIST measures up to 3300 cm-1. Why do you show only the range until 1400 (1500) cm-1. If you have a good reason for this, please mention it in the paper why you restrict these plots to this selection of wavelengths. Otherwise, I would suggest to show the whole spectrum.

The original focus of this effort was to confirm that the thermodynamic profiling would be consistent among the units, which only uses observations between 550 to 1000 cm^{-1}. Thus, the careful attention to the laboratory-derived corrections was focused only on the MCT detector (i.e., 520 to 1800 cm^{-1}). (We didn't show the results between 1400-1800 cm^{-1} as that portion of the spectrum is largely opaque from the surface, and thus not very useful from a science perspective). Because the InSb band (1800-3200 cm^{-1} spectral region) is relatively unused, the laboratory calibration activities to refine the finite field-of-view and spectral calibration parameters for the InSb band was performed <u>after</u> this intercomparison activity in September 2023 (we wanted to confirm that the laboratory calibration approach worked for the MCT band first). What is important here is that this study confirmed that the procedures used in the laboratory to derive the needed instrument-specific coefficients are effective, and thus the company has adopted them for all future ASSISTs (including their use in both the MCT and InSb bands). We have added a new paragraph in the summary of the paper that captures this information.

Figure 4: Why is the upper axes range set to 150? It seems that there are also data above this value. Do you have a reason for that? I would suggest showing the whole range.

The axis range has been changed to 160 RU, which now encompasses all of the points. The interpretation of the results is unchanged.

Technical comments:

Please check the list of references. It seems that many references in the introduction are not listed.

The missing references have been fixed (thank you to reviewer #2 for their help).