Response to Anonymous Referee 1 of wes-2025-198

The article entitled: "Temperature profiling at the American WAKE experimeNt (AWAKEN): methodology and uncertainty quantification" presents an assessment study of the accuracy in determining the atmospheric temperature vertical profile using a ground based atmospheric sounding spectrometer. For this purpose, the authors have conducted a series of validation tests before and during the AWAKEN experiment. I think that the manuscript is very well written and structured. The figures are clear and descriptive.

We appreciate the effort in reviewing a dense and comprehensive manuscript and the positive feedback. Please see below how we improved the paper based on the reviewer's suggestions.

I have only a few minor comments on the manuscript.

1. I find the manuscript a bit too long. I understand that this partly to the thorough description of all validation tests which is necessary. I suggest that the authors could consider moving parts in the Appendix (for example Figure 13) and consider if the presentation of the gravity wave case is necessary for this paper, or not. In the same context, Figure 1 is just mentioned in line 105, but it is not really discussed in the article.

   We agree with the reviewer that the manuscript is quite long. It is mostly because of the presence of graphic content, since the word count of about 8,000 words is in line with common practice. It is, however, a foundational paper on the subject, and we deemed it important to document carefully the methods to provide the many users of AWAKEN data with the necessary information. We tried to move Fig. 13 to the appendix, but it ended up being too disconnected and requiring additional text. The gravity wave case was also kept because it provides a quick insight into the physical knowledge enabled by this technology and adds some variety to the paper that otherwise feels overly abstract. We did increase the number of mentions of Fig. 1 as suggested.

2. On the contrary, I think that it would be constructive to include a bit more information regarding the ASSIST instrument.

   We included a concise but exhaustive description of the ASSIST at Sect. 2.1:
   The ASSIST is a hyperspectral infrared spectrometer that resolves the downwelling radiation in the wavenumber range 525-3300 cm^{-1} with a spectral resolution of about 0.5 cm^{-1}. It shares many similarities with the older Atmospheric Emitted Radiance Interferometer (AERI, Knuteson et al., 2004a). In fact, the core instrument is a Fourier transform interferometer of the Michelson type. The online radiometric calibration is

carried out by rotating a scene mirror that alternates views of two blackbodies and the sky. One blackbody is thermally regulated and the other is allowed to drift according to the ambient conditions, and both act as known sources of emission in the linear calibration process. The ASSIST is able to generate an independent spectral radiance estimation every 14 s during sky views. High spectral accuracy is obtained thanks to a stable laser source that triggers the sampling of the raw interferogram at precise intervals along the mirror path. The interested reader is referred to Michaud-Belleau et al. (2025); Letizia et al. (2025) for more details.

3. Line 145. It is not clear how often are radiosondes measurements required for the optimum performance of the atmospheric sounding spectrometer data analysis. Can you please comment on this.

   New Line 158 now reads:

   "Finally, prior radiosonde measurements used to constrain the calculation could be a source of bias if they are not representative of the observed climatology. This is relatively straightforward to diagnose (Maahn et al., 2020) and can be prevented by using a statistically converged dataset of radiosonde observations. The latter should capture both the seasonal and diurnal thermodynamic variability of the site (i.e., at least one year of launches in both stable and unstable conditions)."

4. Lines 180 – 185, Table 2. Add a brief description in the label of the table about the mean statistics.

   We changed the caption of table 2 as follows:

   "Data availability, mean uncertainty, σ_T, and vertical resolution, δz, of TROPoe retrievals for ASSIST 1 based on 12927 temperature profiles derived from data collected during the pre-campaign test."

5. Line 192. Is it daily or hourly averaged?

   Daily-averaged in this context means "averaged every day at the same hour". It is often synonym with hourly-averaged, but in our case, since we don't necessarily average every hour, we'd prefer to keep it as it is. We added at Line 184:

   " daily-averaged (i.e., averaged every day at the same hour)..."

6. Line 229, Figure 5. Add a description of the red line in the label of Figure 5.

   We added in the caption of Fig. 5:

“The red lines are the theoretical predictions.”

7. Lines 250 – 251. What is the minimum measuring height of the ASSIST? Is it 3 m?

   It measures down the optical inlet of the instrument, so less than 1 m a.g.l. We interpolated at 3 m just to compare to the met tower.

8. Lines 349 – 350. Isn’t it a redundancy to write that the “The Richardson number from TROPoe uses temperature and moisture content from TROPoe”?

   We think it is important to specify that what we label as Ri from TROPoe still uses winds from the met tower. Just a way to clarify that the thermodynamic profiler alone cannot measure the Ri.

9. Line 379. What was the type of the pulsed lidar of Shippert et al. 2023 and how is it used in this study?

   The citations at Line 349 the due references to the public data channels from which we extracted the cloud-base height. They contain details on the instrument models and the data products. We chose not to go into too many details on how the cloud-base height was calculated and instead referred the reader to the specialized literature.

Response to Anonymous Referee 2 of wes-2025-198

The manuscript provides an overview on temperature profiling from ground-based spectral infrared radiance observations using the rather novel ASSIST instrument and covers both, a comprehensive uncertainty analysis of the instrument itself and the TROPoe temperature profile retrieval algorithm, as well as the analysis of nearly one year of multi-site ASSIST temperature profiling during the AWAKEN campaign.

The manuscript is rather long, but the presented information justifies its length. It is in general well structured and well written and can be considered for publication in WES after addressing a few minor comments.

Thanks for the supportive feedback and the thoughtful review of our paper. The reviewer raised valuable points that we addressed as described below.

Line 51: opening "(" missing for the citations

Fixed.

First paragraph section 2 methods: The manuscript would highly benefit from a better and more detailed introduction of the ASSIST system

We included a concise but exhaustive description of the ASSIST in Sect. 2.1:
The ASSIST is a hyperspectral infrared spectrometer that resolves the downwelling radiation in the wavenumber range 525-3300 cm^{-1} with a spectral resolution of about 0.5 cm^{-1}. It shares many similarities with the older Atmospheric Emitted Radiance Interferometer (AERI, Knuteson et al., 2004a). In fact, the core instrument is a Fourier transform interferometer of the Michelson type. The online radiometric calibration is carried out by rotating a scene mirror that alternates views of two blackbodies and the sky. One blackbody is thermally regulated and the other is allowed to drift according to the ambient conditions, and both act as known sources of emission in the linear calibration process. The ASSIST is able to generate an independent spectral radiance estimation every 14 s during sky views. High spectral accuracy is obtained thanks to a stable laser source that triggers the sampling of the raw interferogram at precise intervals along the mirror path. The interested reader is referred to Michaud-Belleau et al. (2025); Letizia et al. (2025) for more details.

Lines 85/86: "TROPoe utilizes a spectrally-resolved and extensively validated radiative transfer model (Turner et al., 2004; Clough et al., 2005; Mlawer and Turner, 2016) to simulate the spectral radiance that is associated with a given temperature profile"; I think it is not only the temperature profile, but also the humidity profile that plays a role here, right?

Correct. We fixed that line by adding:

"…thermodynamic profile (mainly temperature and water vapor content)"

Line 224: "and exhibits a gradual monotonic trend with height"; what is the reason for this increase with height? Mainly the increase in range gate length/probe volume with altitude, or are there other effects in play?

The increase of the total error (noise+smoothing) with height is a well-established theoretical result in atmospheric sounding. However, the trend of the noise-only contribution has been shown very few times (e.g., Rodgers 2000, Fig. 3.6) and has a less consistent behavior with height. We hypothesize that one of the reasons why the noise error increases with height is the relatively lower signal-to-noise ratio in channels that are more transparent to infrared radiation. Transparent in the context of ground-based thermodynamic profiling also means carrying information from higher altitudes. Another reason could be the formulation of the numerical solution algorithm, which anchors the estimated profiles using first observations closer to the ground, thus reducing the noise-induced uncertainty at lower altitudes.

First par section 3.3., starting with line 244, description of the mast measurement setup; I miss here a bit more detailed description of the temperature measurement system; are the differences based on thermocouples? Which type is the absolute temperature sensor at 3 m?

The sensors for both absolute and differential temperature measurement are platinum RTD. We now say at Line 262:

"The total error of the temperature profiles is quantified through a comparison with the temperature readings of the Resistance Temperature Detector (RTD) platinum probes installed on the M5 tower. Specifically, the temperature profile is reconstructed by combining the absolute temperature measured by the RTD at 3 m and differential temperature measurements still from RTD pairs between 3 and 38, 38 and 87, and 87 and 122 m."

Line 293: "For an ergodic, horizontally homogeneous field, this reads…"; Is this assumption generally valid? I would expect this maybe in neutral and partially stable conditions, but not necessarily for daytime convective situations.

This is an interesting subtlety. During the daytime, there are persistent thermal plumes that may invalidate horizontal homogeneity for short periods of time. However, since we have no reason to believe that at the experimental site convective cells form consistently at specific locations, we can still assume that the expected value of the structure function is independent of the direction. Omni-directionality is the only simplification used in Eq. 4. This

is not too important, considering that later on we select only data within a 40-degree wind sector.

Fig. 16 f: Do you have any idea/hypothesis, why the distribution for the data at the site North is considerably broader compared to South and Middle? That site seems to have a different microclimate compared to the two others (see also some of the following comments)

Good catch, this is something we forgot to address. We now added at Line 453:

"The North site is practically unbiased, but shows slightly more scattering, possibly due to its proximity to the turbines, which may result in increased turbulence levels."

Figures 17 a-c indicate that it is station North causing the differences/deviations

We now clarify at Line 469:

"The previously observed bias at the South and Middle sites cancels out in Fig. 17d, but contributes to shifting the peak of the PDF in the other cases."

Fig 18: Again, North shows a distinct differing behavior

True, although it is merely a shift of the stability-dependent biases while the trend with stability (the main focus of this figure) is still consistent with the other sites. We chose to omit this last point for brevity.

Fig. 20: In all plots before the order was different: south-middle-north....., which feels a bit inconsistent; but I fully understand the point to plot the Northernmost station on top....

Correct observation, we keep it consistent with the "physical location" of the sites, like we did in the Fig.6.

Line 651 references: "XXXX-XXXXX" should there be another report identifier?

The reference has been updated (it was under review).

Final curiosity question: have there been deployed passive microwave radiometers for temperature profiling in parallel during AWAKEN, and if yes are there any plans to compare both methods?

We have not at AWAKEN, but other researchers in our team have made this comparison in the past (e.g., Blumberg et al., 2017, Turner and Lohnert, 2021, Bianco et al., 2024). Conclusions: the infrared part of the spectrum has more information content, which leads to generally smaller errors and better vertical resolution. However, MWR can see better through clouds.