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.