Response to review by Tim Hewison

Opening remarks

We warmly thank the four anonymous referees and Tim Hewison for taking the time to review our manuscript and to provide valuable feedback. As there are commonalities between several of the reviews, we start with some general remarks. To begin, we emphasise that our goal is not to encompass the entire AWS mission. First of all, this would be very challenging to cover within a standard manuscript length and would approximately double the number of co-authors. For example, the primary objective of AWS is numerical weather prediction (NWP), and addressing the aspects and applications of AWS within this area could be a manuscript in itself. The manuscript's aim is instead to provide the necessary information to understand the design of the AWS radiometer and to utilise the L1b data from this instrument. In the revision, we focus on improving the text around these aspects based on the provided feedback, as well as adding some new information.
A related question is how much in-orbit characterisation to include. Here, we hope to have an understanding of the difficulty of compiling the manuscript at the same time as the team is preoccupied with the satellite's commissioning phase. The initial aim was to submit the manuscript in 2024. In particular, the sudden deviating behaviour of the 174 GHz receiver (Sec. 6.3) caused significant concern and resulted in a substantial delay in the manuscript. Nevertheless, our approach is to include some initial basic results, primarily to indicate that the findings from the on-ground tests appear to be valid. We have added a sentence to exemplify this further and on the same time indicate the range of aspects that has to be considered. We avoid going further to leave room for one or several upcoming articles that are entirely focused on in-orbit testing. In addition, to fully cover the in-orbit testing would again require a considerable extension of the list of authors. This work is ongoing and far from complete. At least one update of the L1b processing algorithm is foreseen.
In summary, we find it reasonable to focus on the development of the instrument up to the launch. On this side, we think the manuscript is already more information-rich than usual. This brings us to an unstated objective. It is already difficult to find in the open literature the relevant background information about the satellite instruments we use for research. The trend towards new space and more substantial commercialisation risks making the situation worse; with this manuscript, we aim to demonstrate that this need not be the case.

The replies below refer to the revised version of the manuscript we have prepared.

Replies on referee's comments

- "L175: The scan rate is constant, giving an along-track distance between footprints of about 9.0 km"
  - We just wanted to give a rough number and it was a mistake to write 9.0 (instead of just 9). And we probably missed to include the scaling down to the ground altitude. Thanks for the correction, we have changed to your value.
- "I suggest to make it clear than the operational altitude of AWS has been a fairly constant 599km since 2024-12-01."
  - Changed from "altitude of about 610 km" to "a semi-major axis altitude of 599 km".
- "L315: Please highlight and justify the departure from the usual convention is to report bandwidths between 3dB points (not 6dB)."
  - As there is considerable variation inside the passbands, we selected -6dB (with respect to peak response) to be conservative. However, we agree that it is strange to not use the standard -3dB. Therefore, we have changed our method and motivation in the text. We now normalize each SRF with respect to the average response values between the peak-normalised -3dB points. Using this band average as reference, we take the usual -3dB bandwidth. This is described in Sec 4.1. The measured "Boxcar" difference values in Table 4 are updated.
- "L344: The minimum and maximum FWHM of the nadir response for some selected frequencies are reported in Table 4. - Add mean"
  - A good suggestion, adopted.
- "L349: Over what period were the standard deviations evaluated?"
- "L361 + L509: How was the In-Orbit NEDT evaluated? Is this based on Deep Space on OBCT views?"
  - On request from several of the referees, the way NEDTs have been derived is now described (in Sec 4.4). The same calculation approach was used for on-ground and in-orbit data, clarified in Sec 6.2.
- "L364: How is the short-term stability defined?"
- "L367: How were the inter-pixel error and orbital stability quantified?"
  - The choice of these words likely indicated a standard assessment of these aspects, that is not correct. Sec 4.5 has been rewritten and hopefully better clarifies the approach and ambition of the tests.
- "L419: What are the 5 atmospheric scenarios used to define the SRF sensitivity?"

    - We agree with this comment and a similar one by RC2 that this should be mentioned. We now explicitly mention the five scenarios and add a reference to the data source at the start of section 5.3.
- "L420: How is the SRF sensitivity actually quantified in Table 4? Is this the mean difference between the brightness temperatures simulated by ARTS with the actual SRF and with the boxcar approximation? The difference will be scene dependent – can you quantify its variance?"
    - Thanks for this comment. In our view, these questions are addressed in the manuscript. How SRF sensitivity in Table 4 is quantified is introduced at the beginning of Section 5.3, and O3 sensitivity in Section 5.4. Here, it is also described that the values are the maximum absolute difference between measured SRF and "Target" or "Boxcar" over five atmospheric scenarios (now also specified to FASCOD according to the previous comment).
    - The difference is indeed scene-dependent. However, only the atmospheric profile is changed according to the five Fascod atmospheres. These atmosphere profiles are intended to reflect a common state for tropical, midlatitude, or sub-arctic regions during summer or winter. In the interest of not cluttering Table 4 too much, and since there are no distinctive scenarios included, we think that a worst-case value across these simulations serves the analysis of SRF performance best.
    - However, to still address the question, we include the individual differences across cases for AWS15 and AWS42 (the two channels with the largest difference in the Boxcar column) in this reply:

channel	AbsSpeciesCase	FascodAtm	Ta abs(MEASURED-BOXCAR)	Ta abs(MEASURED-BOXCAR_FROM_MEASURED)
AWS15	RTTOV_v13x	midlatitude-summer	0.36	0.12
		midlatitude-winter	0.26	0.09
		subarctic-summer	0.27	0.10
		subarctic-winter	0.21	0.07
		tropical	0.43	0.14
AWS42	RTTOV_v13x	midlatitude-summer	0.50	0.22
		midlatitude-winter	0.45	0.17
		subarctic-summer	0.38	0.18
		subarctic-winter	0.39	0.15
		tropical	0.52	0.22