We would like to thank Referee#1 for his/her review. Below we provide our responses, with the referee's comments in black and our replies in blue. The lines mentioned are those of the revised manuscript with Simple Markup for better readability.

This paper compares radiosonde relative humidity measurements in the extratropical upper troposphere using Meteomodem M10 or M20 with collocated relative humidities from ERA5 over an extended period. Results show systematically higher relative humidities in the radiosondes compared to ERA5, particularly around 200 hPa. Two methods of raw data processing are compared, where GRUAN processing leads to systematically lower relative humidities than the Meteomodem processing at some altitudes. Little difference is found between M10 and M20 radiosondes.

Major issues: The paper does pay too little attention to the international radiosonde intercomparison in Lindenberg (Dirksen et al. 2024). This paper is cited and it is mentioned that M20 has a wet bias during nighttime compared to the standard defined in Dirksen's report. It seems highly likely that much of the presented discrepancy between ERA5 and M10/20 radiosondes is related to that wet bias. The wet bias can be partly reduced with GRUAN processing, which is the most important result of this study. The authors however attribute the discrepancy to potential errors in ERA5 (too little supersaturation over ice events due to deficient parameterization). To me this is highly doubtful, first because ERA5 relative humidity is reported relative to water, and it is unclear to me how the authors converted this relative humidity to relative humidity with respect to ice. There are uncertainties involved due to interpolation and averaging in ERA5 boxes at least.

In the present paper, as mentioned in section 2.1, ERA5 relative humidity is not systematically defined with respect to liquid water; instead it is defined with respect to liquid water or ice depending on temperature. If T is lower than 250.16 K, RH is relative to ice and if T is higher than 273.16 K, RH is relative to liquid water. Between 250.16K and 273.16K, a combination of both is used as described in equations 2 to 5. At the altitudes where the largest differences between radiosondes and ERA5 are observed (200-300 hPa), the temperature is always below 250.16 K and RH is defined with respect to ice.

In this paper, significant effort has been devoted to homogenize RH calculations (see section 2), allowing a consistent comparison between radiosondes and ERA5 RH without biases induced from different formulations of saturation water pressure.

Despite this homogenization, differences in RH between M10/M20 sondes and ERA5 remain. We agree that these difference may partly be attributed to wet bias previously reported by Dupont et al. (2020) for M10 and Dirksen et al. (2024) for M20 sondes. However, our results indicate that this cannot be the only explanation for the following reasons:

1) the wet bias observed for M10 in this study is larger than previously reported by Dupont et al. (2020).

Dupont et al. (2020) showed that differences relative to Vaisala RS92 RH values are within 2% (nighttime) and 5% (daytime) for the GRUAN processing, and within 6% and 9% for operational processing. In contrast, the bias shown in figure 1 of the present study increases with altitude and reach 10% at 300 hPa (~9-10 km) for both daytime and nighttime measurements and for both processing methods.

2) the bias cannot be explained by systematic errors as suggested by Dirksen et al. (2024).

By comparing the mean measurement error and standard deviation of the individual measurement errors (Figures L87 and L88 in Dirksen et al., 2024), these authors showed that the wet bias is mainly due to systematic errors, at least between 4 and 12 km for both daytime and nighttime measurements. However, in our analysis of the 200-300 hPa layer, the difference between radiosondes and ERA5 RH increases for ERA5 RH values above 90% with a bias between 20 and 25%. M10 sondes show more frequent supersaturation than ERA5, which shows a high occurrence of RH values close to 100 %. At these altitudes, the difference may result not only from the known wet bias in the M10 data, but also from a dry bias in ERA5, as reported in comparisons with lidar data (Alraddawi et al., 2025) and MOZAIC aircraft data under supersaturation conditions (Gierens et al., 2020). In addition, laboratory experiments and in flight comparisons conducted under cloudy conditions generally show a good agreement between M20 sondes and references instruments (Figures 10.24 and 10.25 of Dirksen et al., 2024). Only a few measurements indicate an underestimation of RH by M20 radiosondes, which would suggest a reduced wet bias in cloudy conditions, contrary to what is observed here.

Regarding uncertainties related to interpolation and averaging within ERA5 grid boxes, we minimized interpolation uncertainties by performing 3D interpolation using the radiosonde trajectory and by using the highest ERA5 resolution available (0.125° horizontal resolution and on 137 verticals model levels). Nevertheless, radiosonde measurements are point measurements while ERA5 provide grid boxes averages. Within each grid box, RH can be expressed as a cloud fraction weighted mean: $RH = C \times 100 + (1 - C) \times RH_{clear}$.

with C is the cloud fraction and $RH_{i,clear}$ is RH in clear sky condition. However, given the large number of profiles used in this study, there is no reason to expect a systematic sampling bias toward one or the other. Therefore, this effect is unlikely to explain the observed differences.

Instead, the difference is more probably related to the microphysical parametrization than spatial resolution. As mentioned by Gierens et al. (2020), “When a grid box contains a cloud, the humidity in the cloudy part of the box is immediately (within one time step) reduced to saturation” and “When the humidity in the clear part of the grid box, $RH_{i,clear}$, increases and surpasses a critical value, the cloud fraction C increases as well; this balancing effect inhibits the increase of grid mean humidity to observed maxima.” These mechanisms can lead to an underestimation of RH in cloudy conditions, consistent with our results (figures 3 and 4).

Additional sentences have been added in the revised manuscript to better discuss the observed bias and reference to the Dirksen et al. (2024) publication has also been added in this part.

These sentences are lines 173-184, 393-407, 441-443, 589-592, 602-604 and 614-615.

Parts of the conclusions are written rather vaguely, giving the impression of inconclusive results. This does not help the reader at all. It may help to give uncertainty estimates instead of formulations such as "shouldn't prevent climatological studies". At least some of these uncertainty values can be found in the main body of the paper. It is important to state these also in the conclusions since often only the conclusions are read.

To address this point, the conclusions has been improved to be more quantitative.

MInor revisons:

ERA5 (Hersbach et al. 2020) should be cited a lot earlier at the first occurence in the introduction.

Added line 131 page 5.

lines 262 and 270/271: This appears redundant, consider shortening this paragraph

The paragraph has been shortened on line 356.

387: Data processing evolution is a bit awkward in this context. It is more stepwise changes. Thus I would say "changes" or "upgrades"

The word upgrade is used in the revised version of the manuscript, line 234 for example.

line 213f: It is not entirely clear to me if this (taking RH from ERA5, and T from the sondes) could be a reason for the different supersaturation frequencies in ERA5 and the radiosondes. This should be elaborated with some sensitivity experiments/uncertainty estimates

This point has also been addressed in our response to the first major comment. As explained in section 2.2.3, RH is not taken from ERA5 and T from the sondes but RH is calculated using the dew point temperature T_e and the temperature T provided both by the sondes. The only adjustment is that we apply the same saturation vapor pressure formulations as those used in ECMWF-ERA5, in order to ensure a consistent and unbiased comparison between ERA5 and radiosonde-derived RH).

This is now clarified lines 272 to 274.

line 464: daytime! Radiosoundings

Sorry, it is now corrected.

line 506: larger (not greater)

Done.

lines 524, 548: IFS

Done

line 540: Is there a public document describing this campaign

At present, the only publication about this campaign is listed below. It has been added line 656.

Poignard, A., Farah, A., Sarkissian, A., Keckhut, P., Alraddawi, D., Khaykin, S., Dupont, J.-C., and Capo, J.: Comparison of TRACIS campaign data with data from radiosonde, IPRAL Lidar and ERA5, EGU General Assembly 2026, Vienna, Austria, 3–8 May 2026, EGU26-21164, https://doi.org/10.5194/egusphere-egu26-21164, 2026.

There is also a web page dedicated to this campaign but in French:

https://www.ipsl.fr/agenda/archives-evenements/tracis-tropospheric-research-campaign-on-air-humidity-content-by-ipral-at-sirta/

lines 529ff: as noted above: the conclusions are formulated very vaguely here ("shouldn't prevent", "appears reproducible"). It would be better to have uncertainty ranges

This paragraph has been removed in the revised version of the manuscript.