

Referee report on "The first decadal-scale ground-based microwave radiometer dataset in China: Brightness temperature and thermodynamic profiles from Xianghe (2013–2022)" by Yueyuan Gong et al.

The data set presented contains measured microwave and infrared brightness temperatures, corresponding quality flags and retrieved integrated products on a minutely time resolution. In addition, retrieved temperature and relative humidity profiles are given on a 10 min time resolution. Both types cover the time period of Apr 2013 - Dec 2022.

In the presented article, the authors describe a multi-year data set of quality controlled microwave radiometer observations, including comparisons to simulated brightness temperatures using atmospheric profiles from reanalysis. Two different retrieval approaches are presented, applied to cloudy and clear sky conditions, respectively, and compared to the retrieval approach provided by the manufacturer with respect to radiosonde data. Furthermore, a statistical analysis is carried out including a temperature trend and inversion analysis connected to air pollution.

The study presents a valuable data set with a high temporal resolution and the performed statistical analysis provides insights in long-term developments of the thermodynamic structure of the lower troposphere as observed from a ground-based remote sensing instrument. However, the argumentation for choosing different retrieval approaches for clear sky and cloudy conditions remains unclear. In addition, more information on the provided retrieval by the manufacturer is needed to ensure that a fair comparison is made in terms of underlying training data and radiative transfer methods. Instead the newly developed approaches could be compared for cloudy and clear sky cases. Also, the retrieval seems to be trained on observations rather than simulated data.

Therefore the article / data set could be accepted after revision. All major and minor points, corresponding to the data set and manuscript, supporting this decision, are listed in detail below.

Manuscript

Major points:

- How do the DNN and OE retrievals compare to radiosondes in all sky or clear sky / cloudy conditions? This comparison would add more insights rather than comparing it to the retrieval from the manufacturer (where additional information is missing). Also, the argumentation of using different approaches is unclear.
- The analysis and data set should be separate for the different approaches to ensure a meaningful interpretation, and it should be made clear which retrieval approach and conditions (cloudy / clear sky) were used for the statistical analysis in Chapter 4 and in the final data set.
- l. 283: the retrieval training should be conducted with multi-year simulated TB (from ERA5), but observed TB from 2019 are mentioned. This needs to be checked and clarified.
- The authors should add more information on the retrieval methods (e.g. utilized auxiliary data, noise) and on the radiative transfer method (e.g. absorption models). Was it tested against line-by-line models?

Minor points:

- More information on the site could be presented (e.g. map)

- Were elevation scans available for improved temperature profiles?
- The phrasing "weather type classification" is misleading, as it only contains a rain and cloudiness check and should be changed (e.g. to weather related information / flags) throughout the manuscript (e.g. l. 20).
- l.28 (and others): Change "MWR's self-developed products" to "retrieval approach provided by the manufacturer". This also needs a more consistent naming, as "LV2" is too general.
- l. 70: Statistical retrievals can also rely on e.g. model data (as in this study; not only radiosondes)
- l. 77: cited paper doesn't support statement and needs to be replaced to justify the decision of using multiple retrieval approaches for different atmospheric conditions
- l. 81: add reference "Rose, T., Crewell, S., and Löhnert, U.: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, Atmos. Res., 75, 183–200, 2005." for the RPG-HATPRO system
- l. 106: reference doesn't describe the specific retrieval developed for the presented study
- l. 108: change "microwave window" to "radome". Also, how often was the radome replaced?
- l. 113: what is the distance to the radiosonde station?
- chapter 3.1.1: this approach flags valid "extreme" events, whereas erroneous data with small deviations remain in the data set, which could cause a bias. This needs to be discussed.
- l. 143: 4 hours seem to be too long; was a blower/heater system deployed with the instrument?
- l. 160: change "stability" to "variability"
- chapter 3.1.3 b): how is CBH derived/retrieved from IRT?
- l. 181: is 1 min the chosen output resolution, or are temporally higher resolved values averaged?
- l. 189: add "model uncertainties" to the argumentation
- l. 193: add "or bias from input data"
- l. 201: He et al. (2021) don't show a comparison of radiative transfer models
- l. 212: more details on the data matching should be presented
- l. 238: how are LWP & IWV retrieved? Retrieval approach from the manufacturer? Was an often applied LWP offset correction carried out?
- l. 250-257: this part is repeated from the introduction and could be removed
- Table 4: change to "Available samples after QC"
- chapter 4.1: could the different occurrence of clouds and rain be linked to atmospheric conditions in the different years? Maybe using findings from other studies to provide more context.
- Figure 8: the temperature from the AWS could be added
- chapter 4.3: the generally "coarse vertical resolution" (limited information content) of MWR derived temperature profiles should be mentioned, especially regarding the EI layers. And what is the reason for increased STD in winter (Inversion Intensity)
- l. 456: how many cases are included in each bin?

Data set

Major points:

- All variables should have names and units
- It should be made clear which retrieval method was used
- Uncertainty estimates (e.g. from the optimal estimation method) should be included

Minor points:

- The data repository site lacks details about file contents (such as variable names, units, metadata), exact instrument location and information on how the products were derived
- Time is encoded in two ways in the *cflag.csv files, which could be confusing
- Is there a difference between "Rainflag" and "n2"? Could one of them be omitted?
- Are there two IR channels? IRT2 is not mentioned in the manuscript (and has a different unit than IRT1)
- "cflag" is not defined in the manuscript
- Why do some quality flags show NaN?