

Replies to reviewers' comments

Data quality control and calibration for mini-radiosonde system “Storm Tracker” in Taiwan (egusphere-2024-661)

We really appreciate the comments to the manuscript from the two reviewers. We have carefully reviewed and adjusted accordingly, and prepared the replies to each comment as follows (the original comments are in blue, replies in black).

Reviewer 2:

... While making more than 1000 dual flights with VS to deduce correction tables/functions for the ST measurements is a very good effort and the results are worth being reported in AMT, I tend to feel that the current manuscript is “overselling” the ST. This is mainly because the temperature and RH sensors of the ST, as shown in Figures 2 and 3 of Hwang et al. (2020), are never optimal for balloon upper-air sounding, e.g. with possible large contaminations to the temperature measurements depending on solar radiative heating and thus cloud cover status as well as day-versus-night difference. I would also have concern about the production stability of the sensor, i.e. whether the characteristics and the quality are within the reasonable range of uncertainty in different production batches (e.g. in different years). This may mean that in the end, we may always need dual flights with a well characterized and reliable radiosonde like the VS. In addition, the radiowave used for ST is from 432 to 436.5 MHz (Hwang et al., 2020) which is for amateur radiolocation, not for meteorological aids (e.g. International Telecommunication Union (ITU), “Radio Regulations”, Volume 1, 2020, available from <https://www.itu.int/pub/R-REG-RR-2020>); thus, in some countries ST cannot be used as a meteorological radiosonde officially. This may mean that the ST cannot be “widely” used in the future. Note also that one-hourly sounding campaigns like the one shown in Section 5.2 are not impossible with modern radiosondes like the VS; thus it is not easy for me to imagine possible applications of the ST aiming at new scientific studies that are only possible with the ST.

Regarding the issue of overselling, it is important to clarify that while Soundings from Triggers (ST) are indeed suitable for Planetary Boundary Layer (PBL) studies and areas with complex terrain, they should be viewed as a supplement to Vaisala Soundings (VS) rather than a replacement. Atmospheric conditions at higher altitudes may not always require high-frequency observations due to geostrophic adjustment, but ST can significantly enhance our

understanding of PBL dynamics.

In regions such as Southeast Asia, where PBL conditions can vary within short distances, the high spatial frequency observations provided by ST are particularly beneficial. It is important to note, however, that the correction results presented are specific to Taiwan. To ensure broader applicability, we suggest conducting co-launches during field campaigns. This approach would allow users to derive in-situ correction formulas using the proposed methods.

Furthermore, we recommend utilizing ST between VS launches to optimize data collection and analysis. This combined approach will enhance the overall effectiveness of atmospheric observations and improve the accuracy of data interpretation. New statements are added in section 6 to address this.

Finally, regarding radio regulations, the hardware design for STs allows for adjustments to different radio bands if needed. While a hardware update is beyond the scope of this manuscript, it is important to note that radio regulation should not pose an issue for field campaigns using STs.

Line 16 (and line 369): I am afraid that “geostrophic adjustment dynamics” is never discussed in the manuscript.

We added a few sentences to elaborate our point.

“For synoptic weather, geostrophic adjustment dynamics suggest that spatial temperature variations in the free atmosphere may not be significant, reducing the need for high-frequency upper-air radiosonde observations. Consequently, most operational radiosonde observations worldwide are conducted daily at 00Z and 12Z, with intervals of 12–24 hours. However, atmospheric phenomena originating from the boundary layer are often smaller in scale and closely related to local terrain. For example, a single convective cell typically lasts for minutes, while thunderstorms persist for a few hours. To gain a better understanding of these types of weather, a low-cost device capable of deploying multiple sensors simultaneously or at intervals of less than an hour can enhance field experiments. This approach provides valuable insights into the lower atmosphere's significant variations in temperature and moisture, especially for convective systems that may lead to disastrous rainfall or flash flooding.”

Introduction and Figure 1: Please describe the main technological points of the ST in more detail, including the model and characteristics of the equipped temperature-RH sensor. Figure 1(b) should be much greater, and an enlarged photo for the sensor part may be added. Also, Figure 1(c) is not a good one, because the ST sensor boom (or “sensor box”) is not well shown.

We added more detailed information about the ST in the introduction section. We didn't modify Figure 1 as suggested because the Hwang et al., 2020 paper had the same figure, and we want to avoid copyright issue.

We added:

“The ST consists of a microcontroller (ATMEGA328p), a GPS sensor (U-blox MAX7-Q), a pressure sensor (Bosch BMP280), a temperature–humidity sensor (TE-Connectivity HTU21D), and a transmitter (LoRa™).”

Introduction, the review part of various radiosonde issues: The papers cited here tend to be too old. More recent papers for more recent radiosondes need to be cited. These include:

- Vaisala RS41 radiosonde (but with GRUAN data processing): Sommer et al., GRUAN characterisation and data processing of the Vaisala RS41 radiosonde. GRUAN Technical Document 8 (GRUAN-TD-8), v1.0.0 (2023-06-28), <https://www.gruan.org/documentation/gruan/td/gruan-td-8>.

- Modem M10 radiosonde: Dupont, J., M. Haeffelin, J. Badosa, G. Clain, C. Raux, and D. Vignelles, 2020: Characterization and Corrections of Relative Humidity Measurement from Meteomodem M10 Radiosondes at Midlatitude Stations. *J. Atmos. Oceanic Technol.*, 37, 857–871, <https://doi.org/10.1175/JTECH-D-18-0205.1>.

- Meisei RS-11G and iMS-100 radiosondes: Kizu et al., Technical characteristics and GRUAN data processing for the Meisei RS-11G and iMS-100 radiosondes, GRUAN-TD-5, v1.0 (2018-02-21), <https://www.gruan.org/documentation/gruan/td/gruan-td-5>.

These also include very useful information on the modern radiosonde sensor characteristics and on all necessary corrections to the radiosonde measurements.

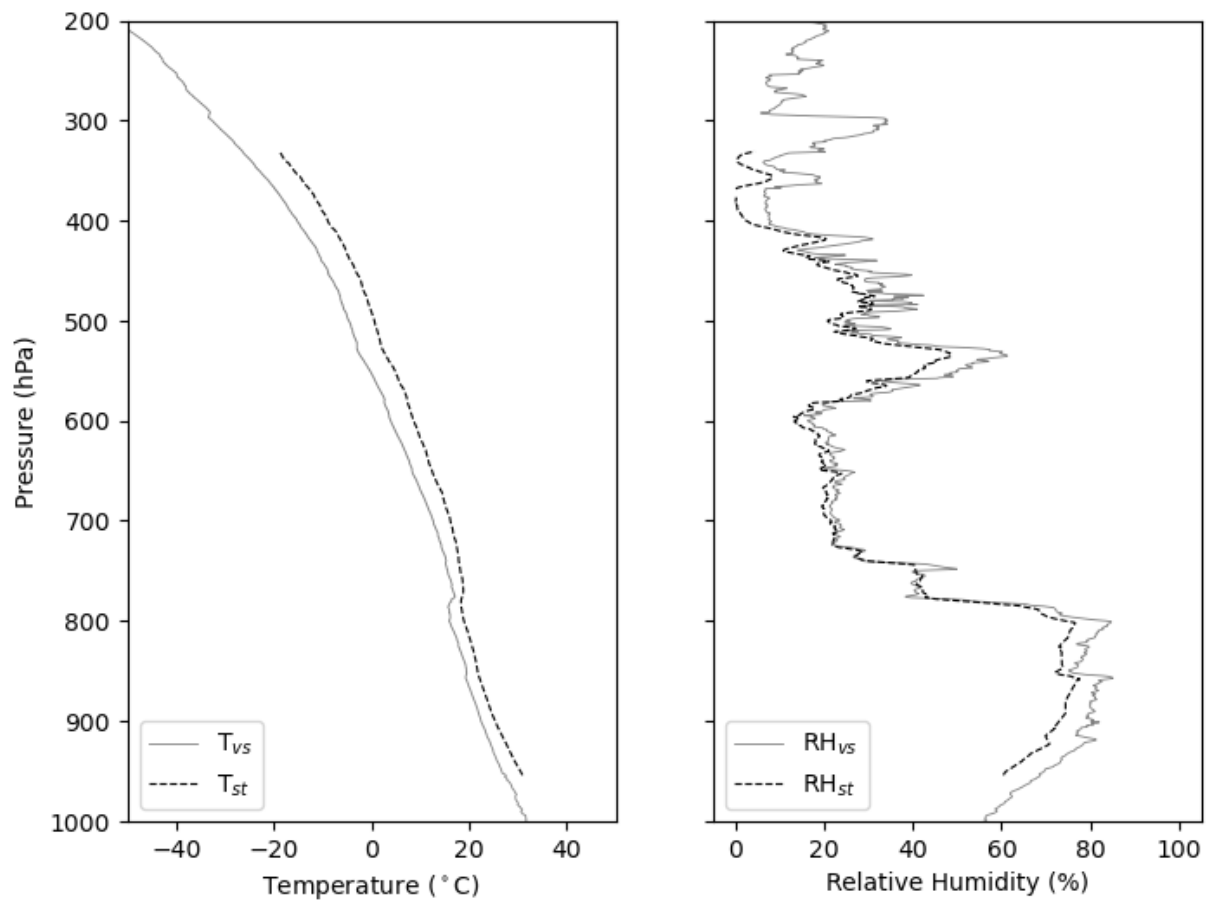
We have reviewed the suggested references and added them to the introduction.

Figure 2: Are the ST RH measurements really dry biased in comparison with the VS measurements below the 500 hPa level? They look wet biased in this case.

We followed the suggestions from other reviewers to replace the skew-T-log-p plot with a regular plot. The aspect ratio of our previous plot could be hard to read and misleading.

As illustrated in the new figure, the ST showed a dry bias in general except some mixed bias pattern above 800hPa.

2018-06-26 03Z(UTC)



Lines 98-105: Testing mathematically more sophisticated machine-learning-based methods is good and interesting, but if the sensor characteristics are not optimal for upper air sounding (e.g. ~5 deg.C temperature error as shown in Figure 2 is just too large!), I tend to think that we should improve the sensor itself before reconsidering correction methods.

We totally agree. Hence, a newly designed ST with different sensors is currently undergoing. However, the improvement of the hardware design is not within the scope of the presented work.

Table 1: Please also add the information on e.g. climate zone, season, etc., i.e. the information listed at Lines 92-93, to the place e.g. right after Location.

We added a map with co-launch sites marked on it. All sites are in the same climate zone. As shown on the map, those sites are of similar altitudes but with different terrain.

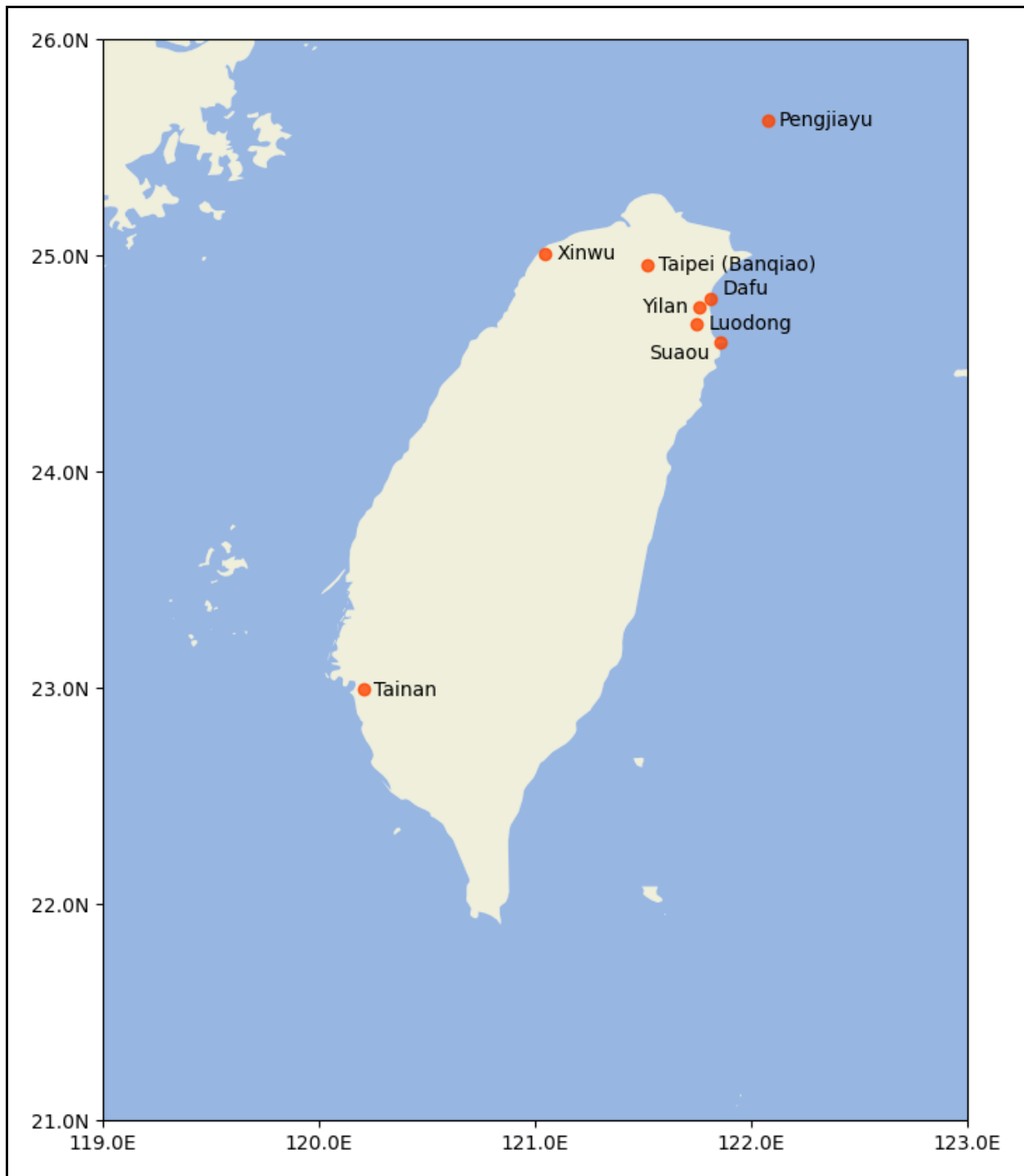
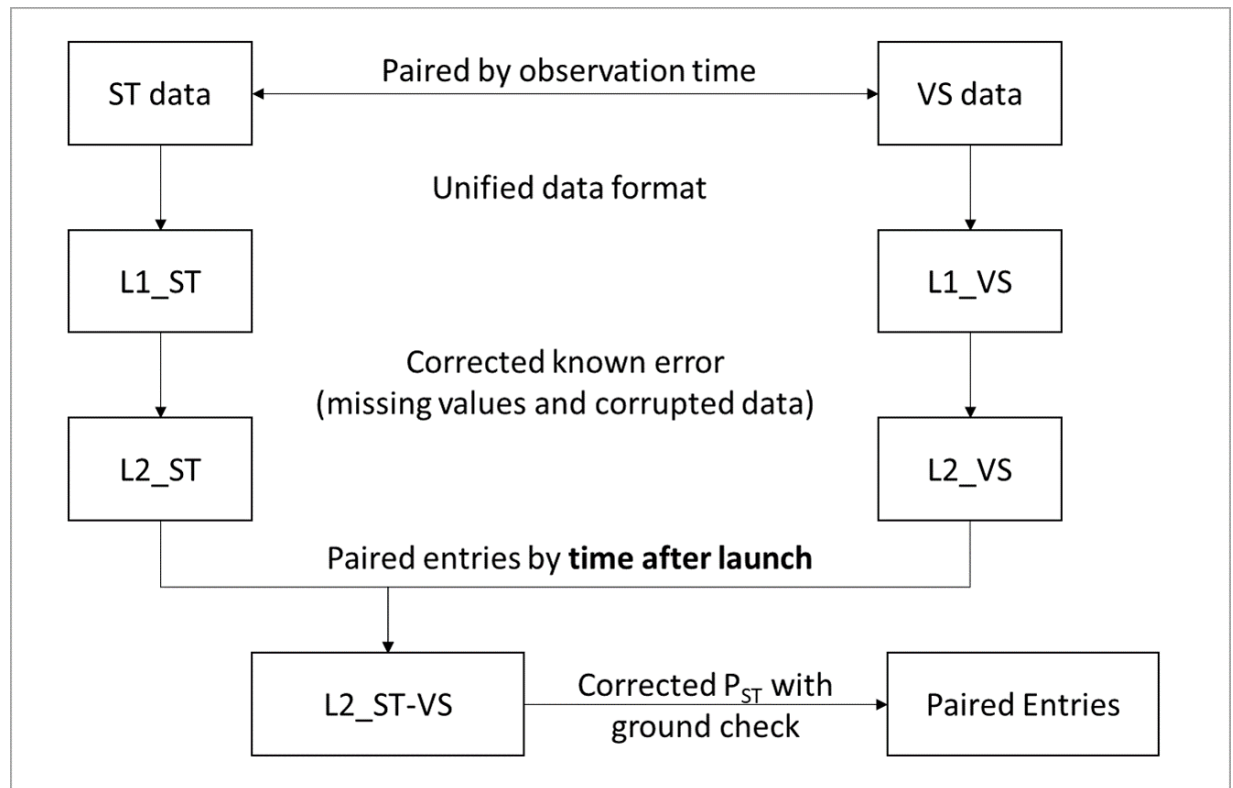


Figure 3: I confused with the three horizontal lines within the gray box. I thought that the data processing of the ST and the VS is independent before “L2_ST-VS”. In other words, I thought that the “Paired Entries” are used to establish correction tables/functions with various different methods. Or, in other words, I thought that for the VS, the authors simply use the manufacturer-processed data set, while I have impression from the current figure that the authors make their own (and perhaps common) corrections from L1_VS to L2_VS. The term Level 2 may be confusing if it is used for the ST at this stage, because the ST data will be

further corrected by using the VS data as the reference if my understanding is correct.

We have modified the figure to avoid confusion. The correction from level 1 to level 2 is done separately for ST and VS. These procedures aim to remove missing values and corrupted data. While VS data is exactly the manufacturer-processed data set, the procedure of L1_ST to L2_ST is done by implementing the specifications described in Hwang et al., 2020.

We want to position our work as a step between level 2 and level 3 data, in which we “identify biases in data and correct if possible” (Ciesielski et al., 2011) by aligning to the VS data using statistical methods.



Section 3.1 (and Figure 4 and its caption): Please add the explanation how to obtain delta T (and its uncertainty) from the CDF.

We added:

“The observed temperature records are sorted in ascending order, and then the proportion of observations is derived for every 0.1-degree interval from -80 to 40 degrees Celsius as the probability density.”

Line228: three-dimensional?

Yes, if we define the correction of RH as a function of temperature and pressure, the joint

probability distribution of $p(P, T, RH)$ is three-dimensional.

Lines 241-244: Please also add “(GLM1)”, “(GLM2)”, and “(GLM3)” here, not later.

We added the notes as suggested.

Lines 255-258: Solar elevation angle may be a better variable?

Yes, it is. And we even calculated the direct solar radiation as a predicting variable. However, we wanted to limit the scope of the presented work to the feasibility and recommended procedure of using ST for high-frequency boundary layer observations. Hence, the correction methods presented involved only basic statistics and data that are directly accessible.

We had another ongoing project for finding the best correction methods including both statistical models and predicting variables.

Figure 7: Results from which method, CDF, GLM*?

Figure 7 demonstrated the bias between VS and ST before correction. We modified the caption to avoid confusion.

Lines 283-285: I have a difficulty on this discussion. I am not an expert of metrology (i.e. science of measurement), but I think that that the “average errors” here are comparable to the uncertainties of the VS measurements does not mean that the uncertainties of the ST measurements (after the corrections) are comparable to the uncertainties of the VS measurements. Can we say roughly how large are the uncertainties of the corrected ST measurements in this case? Is it ~ 1.414 times?

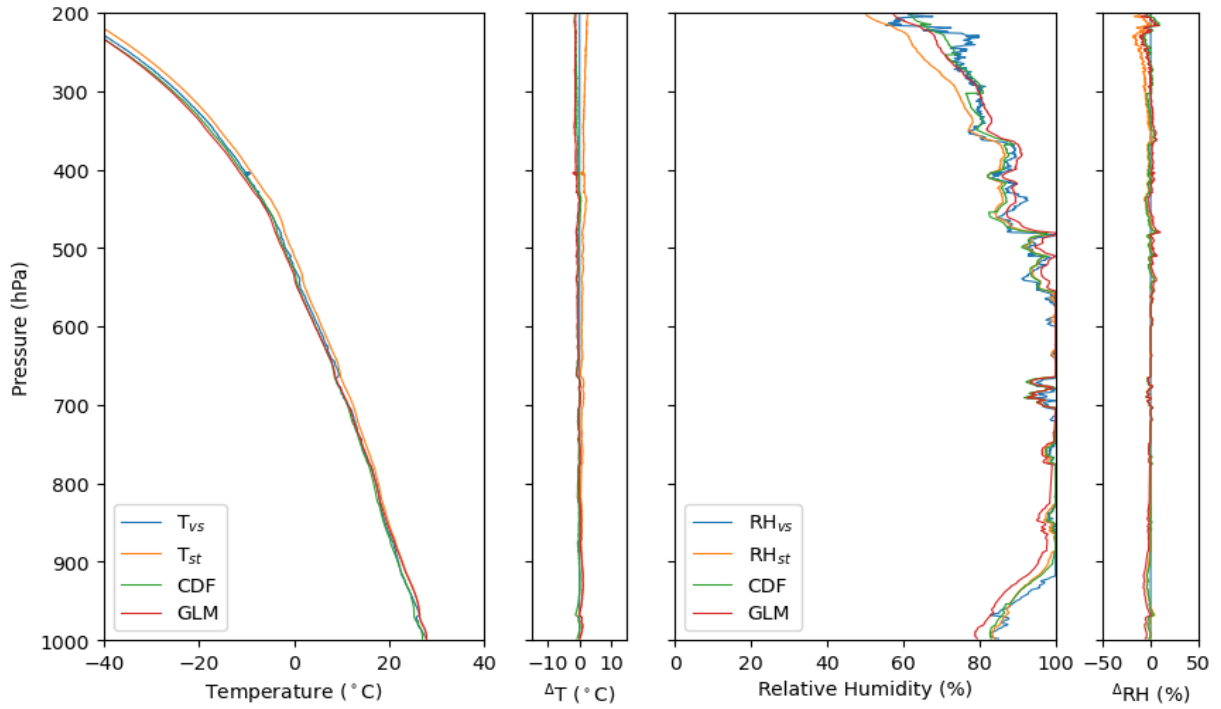
The “average error” (we modified it to “averaged RMSE”) means the bias of ST after correction. As for the uncertainties of “ST after correction” itself, it should be 1:1 in magnitude to the term “stdev” (in the right side of Table 3), which represents the “range of variation of the variable”.

The “errors” reported here are Root-Mean-Squared-Error (RMSE, the square root of the mean squared error), which is a common measurement for evaluating predicted values. The “squared” and “root” are used to ensure the positive and negative deviations will not cancel out each other, and hence this value should be 1:1 in magnitude to uncertainties.

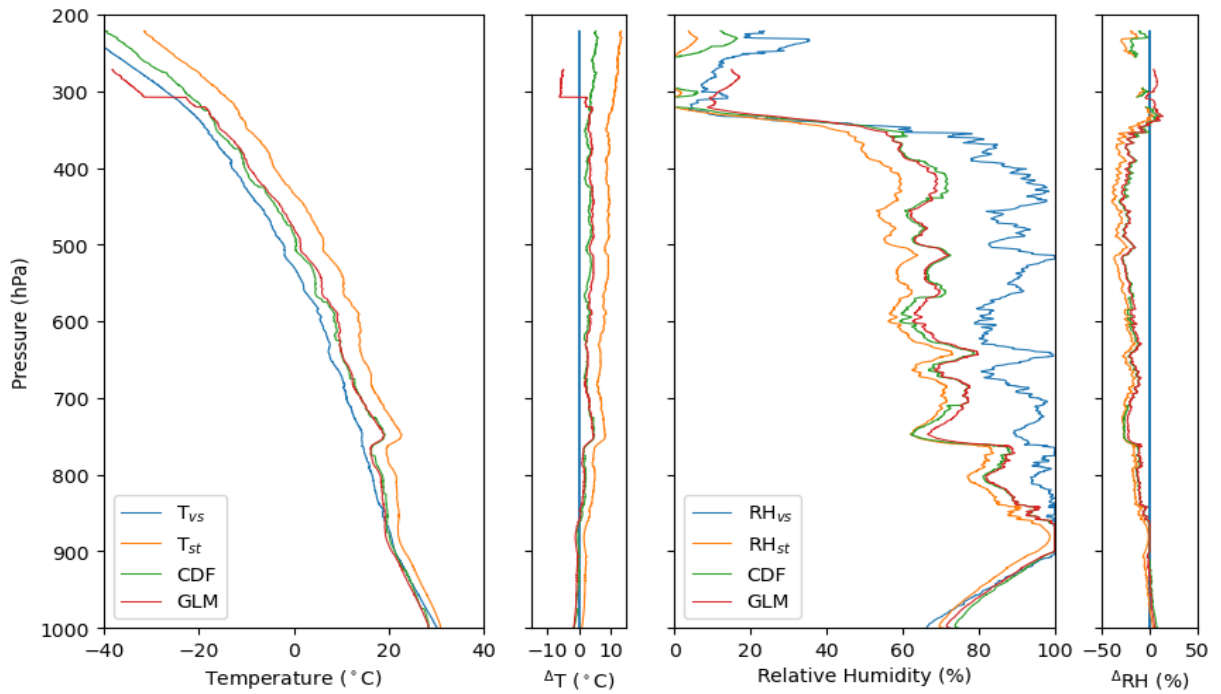
Figures 10 and 11: The authors should also show delta T and delta RH profiles as well. In particular, delta T profiles are very important because we need tropospheric temperature measurements with uncertainties less than e.g. 0.5 K or even 0.1 K usually, and this degree of differences is hard to see with the current panels.

We modified the figure as suggested. The bias in temperature of corrected ST can be as low as 0.66 degree below 700hPa in average with stdev of 0.34. This means the 95% of corrected T having bias between 0 to 1 degree.

2021080312



2018082706



Line 330: A wavy pattern. Which pattern are the authors referring to? With this figure, I am afraid that we cannot judge that GLM may be better.

We were not arguing the GLM is “better” here. We simply pointed out the characteristics of different statistical models. While the CDF approach used separate bins, the corrected profile will show segments corresponding to the bins. We admitted such segments were hardly noticeable if one doesn’t bear this concept in mind, and hence we removed the description to avoid confusion.

Section 5.2: One-hourly sounding campaign is possible with the VS as well.

Yes, it is possible, but with a much higher price. We didn’t argue that ST can do things that VS cannot do, but ST can provide data of reasonably good quality with a much lower cost. Furthermore, STs are also capable of simultaneous observations with just one system, which enhance the flexibility of field campaigns.

Line 369: Geostrophic adjustment dynamics has never been discussed, I am afraid.

We added a few sentences to elaborate our point.

“For synoptic weather, geostrophic adjustment dynamics suggest that spatial temperature variations in the free atmosphere may not be significant, reducing the need for high-frequency upper-air radiosonde observations. Consequently, most operational radiosonde observations worldwide are conducted daily at 00Z and 12Z, with intervals of 12–24 hours. However, atmospheric phenomena originating from the boundary layer are often smaller in scale and closely related to local terrain. For example, a single convective cell typically lasts for minutes, while thunderstorms persist for a few hours. To gain a better understanding of these types of weather, a low-cost device capable of deploying multiple sensors simultaneously or at intervals of less than an hour can enhance field experiments. This approach provides valuable insights into the lower atmosphere’s significant variations in temperature and moisture, especially for convective systems that may lead to disastrous rainfall or flash flooding.”

Lines 373-377: As discussed above, I personally think that the sensor characteristics (including the sensor covering and orientation) need to be improved and more optimized for upper-air sounding, before considering mathematically more sophisticated correction methods.

Yes, we agree. And as mentioned in the earlier response, the next generation of ST hardware design is under development. The represented work focused on the evaluation and recommended procedure of the current ST hardware.

Lines 378-386: As discussed in the beginning, I tend to think that the authors are overselling the ST here.

We modified the corresponding paragraphs to avoid the overselling. The full response to this point is in the beginning.