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 1:

… One concern is with the method applied to analyse the data from the coincident twinsoudings. Since these soundings are performed with two different radiosondes on the same balloon, this allows for a direct comparison of the profile data. Using a statistical method like the cumulative distribution function (CDF) seems to me like an unnecessary complication. To my understanding the CDF-based method as employed by Ciesielski et al., is applied when comparing radiosonde data taken under similar meteorological conditions albeit not coincidently on the same rig. The advantage of coincident twinsounding data is that it allows to directly determine the bias + associated uncertainty between both systems. Furthermore, the physical mechanism that is causing the bias between both radiosondes is warming by solar radiation that is counteracted by convective cooling by the air flowing over the sensor (ventilation). The efficiency of the convective cooling is directly linked to the altitude-dependent ambient air pressure. The CDF method that is applied in the manuscript does indeed derive corrections for pressure ranges, but the purpose of further analysing the differences in sense of ambient temperature is not clear to me.

For a description of the Storm Tracker radiosonde and its sensors, the reader is referred to another publication (Hwang et al. 2020) . However, for a good understanding and interpretation of the results presented in this manuscript a (brief) description of the radiosonde design, and specifications of the sensor are essential. Also a description of the payload configuration for the twinsoundings is necessary, currently the reader has to deduce the payload configuration from a rather poor-quality (underexposed) photograph. From the photograph in Figure 1 it becomes clear that the two radiosondes are connected back-to-back. The sensor boom of the RS41 is pointing upwards with the T + RH sensors probing undisturbed and uncontaminated air. However the integrated T & RH sensor of the Storm Tracker radiosonde is located about 20cm lower, directly in front of the radiosonde's housing. As a result the air flowing over the T & RH sensors is very likely contaminated by the housing/casing of the radiosonde, affecting the measurements. This payload configuration most likely explains the large temperature between the Storm Tracker and the RS41 shown in Figure 2. I was very surprised by the large temperature difference between both radiosondes (approx. 5 K at ground level), and my first thought was that this was caused by a calibration error of the temperature, but this bias is not present for the nighttime flight presented in Figure 10, so that it is indeed likely that this large bias results from solar radiation.

It can't be excluded that the large differences between Storm Tracker and RS41 are caused by the payload configuration, rather than by the performance of the individual sonde types. This raises the question whether the observed differences in this comparison study reflect differences that would be observed between both radiosondes when flying on separate balloons, or on payloads that are better configured for comparisons. Therefore, I am not sure whether the results found in this study are representative for the differences/bias between both radiosonde types and to my opinion it is doubtful that this study can be used to derive a generic correction for T & RH profiles of the Storm Tracker radiosonde.

If the temperature error due to solar radiation is that large, 5 K at ground level and up to 10 K at higher altitudes, it also inevitably limits the quality of the temperature profile after correction, i.e. there will be a considerable associated uncertainty remaining. In such a situation the preferred strategy is to improve the design of the temperature sensor so that it is less sensitive to solar radiation.

We acknowledge that the design of our co-launch experiment differs from some previous field campaigns. In this study, attaching the Storm Tracker (ST) to the Vaisala radiosonde (VS) allowed for direct inter-comparison at native temporal resolution (every second). Additionally, the Cumulative Distribution Function (CDF) is a well-established non-parametric statistical model for regression and classification within the community. The concept of correction processes for temperature (based on different pressure levels) and relative humidity (based on different temperature levels) between different instruments is similar to methods used in other field campaigns (e.g., Ciesielski et al. 2014). Furthermore, we compared CDF with a parametric alternative, Generalized Linear Models (GLM), for data correction purposes in this study. Our results indicate that both models perform well for this task.

In over 1,000 co-launches, we consistently bound ST and VS in the same configuration, as we believe this is essential for a controlled experiment. While it is challenging to definitively prove that the observed biases are unrelated to the binding method, we have conducted thorough checks, including in-lab measurements of both sensors and several co-launches with the sensors in separate balloons (although this data was not used for correction). We found that the bias patterns observed were consistent with those in the co-launch dataset.

Regarding the larger temperature biases near ground level, these issues are primarily due to the sun directly heating the ST while waiting to launch. In most cases, we prevent this, and the correction table excludes such instances as outliers.

We must also acknowledge that we cannot entirely rule out the possibility that the payload configuration (especially the casing) could contaminate the T and RH measurements of ST. The overall idea behind the ST hardware design is to leverage easily-accessible commercial sensors and serve as a supplement to regular sounding observations, specifically in the lower boundary layer. While further hardware upgrades are underway, the ST cannot replace the more mature design instruments. Therefore, we suggest co-launches for future applications to ensure data accuracy.

*Ciesielski, P. E., Yu, H., Johnson, R. H., Yoneyama, K., Katsumata, M., Long, C. N., ... & Van Hove, T. (2014). Quality-controlled upper-air sounding dataset for DYNAMO/CINDY/AMIE: Development and corrections. Journal of Atmospheric and Oceanic Technology, 31(4), 741-764.

The literature references provided in the introduction all are quite old, and recent work on the characterisation of solar radiation temperature error is missing. Important progress in this field is made by the efforts of the GRUAN community, which should be mentioned in the introduction.
Suggested publications are for example:
– von Rohden et al AMT2022 (doi 10.5194/amt-15-383-2022)
– Lee et al. AMT2022 (doi 10.5194/amt-15-1107-2022)
– Hoshino et al. AMT2022 (doi 10.5194/amt-15-5917-2022)
– GRUAN-TD-5 (https://www.gruan.org/documentation/gruan/td/gruan-td-5)

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

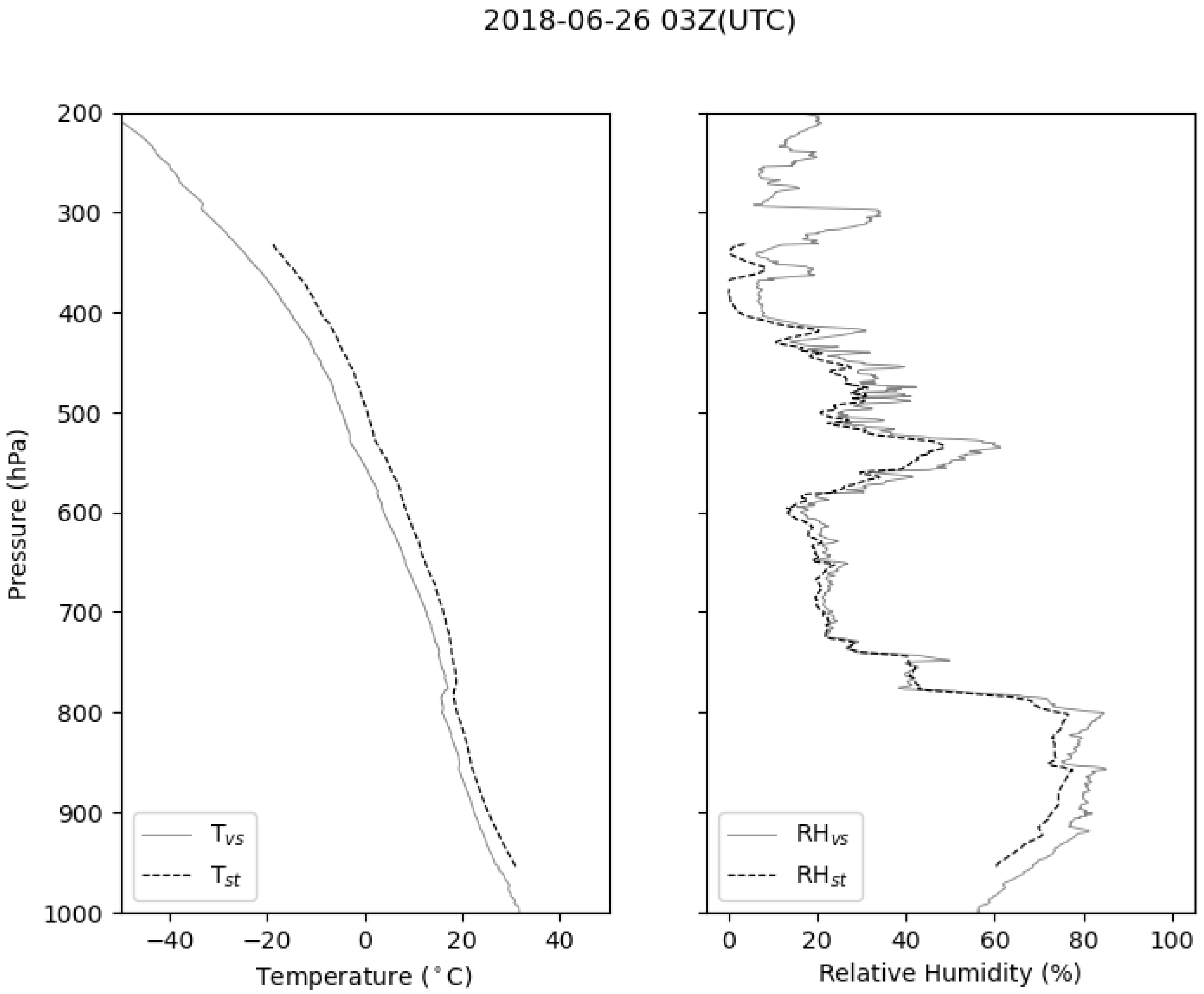
Briefly describe the characteristics of the radiosonde and its sensors. A more detailed description can indeed be provided in another paper.

The following paragraph was added to section 1 for more details on the ST sensors. Again, specific hardware details are described Hwang et al. 2020.

"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™). The sensors have an overall operation range from 1100 to 300 hPa in pressure and from -40℃ to 85℃ in temperature. The ST used a regular AAA battery for 2-4 hours of power; the total weight was 20g. More detailed hardware specifications can be found in Hwang et al., 2020. The design of ST aimed to leverage the low cost of sensors used in commercial electronics to enable high-frequency observations in the boundary layer. The receiver was designed to receive up to six STs simultaneously. It was ideal to use ST to gather supplemental data between regular sounding."

Use a regular plot for the comparison of radiosonde profiles, instead of a Skew-T diagram such as in Figure 2.

The new figure 2 as follows is updated as suggested.



In addition to Table 1, show a map with the location of the sites.

New Figure 3 is added as requested.

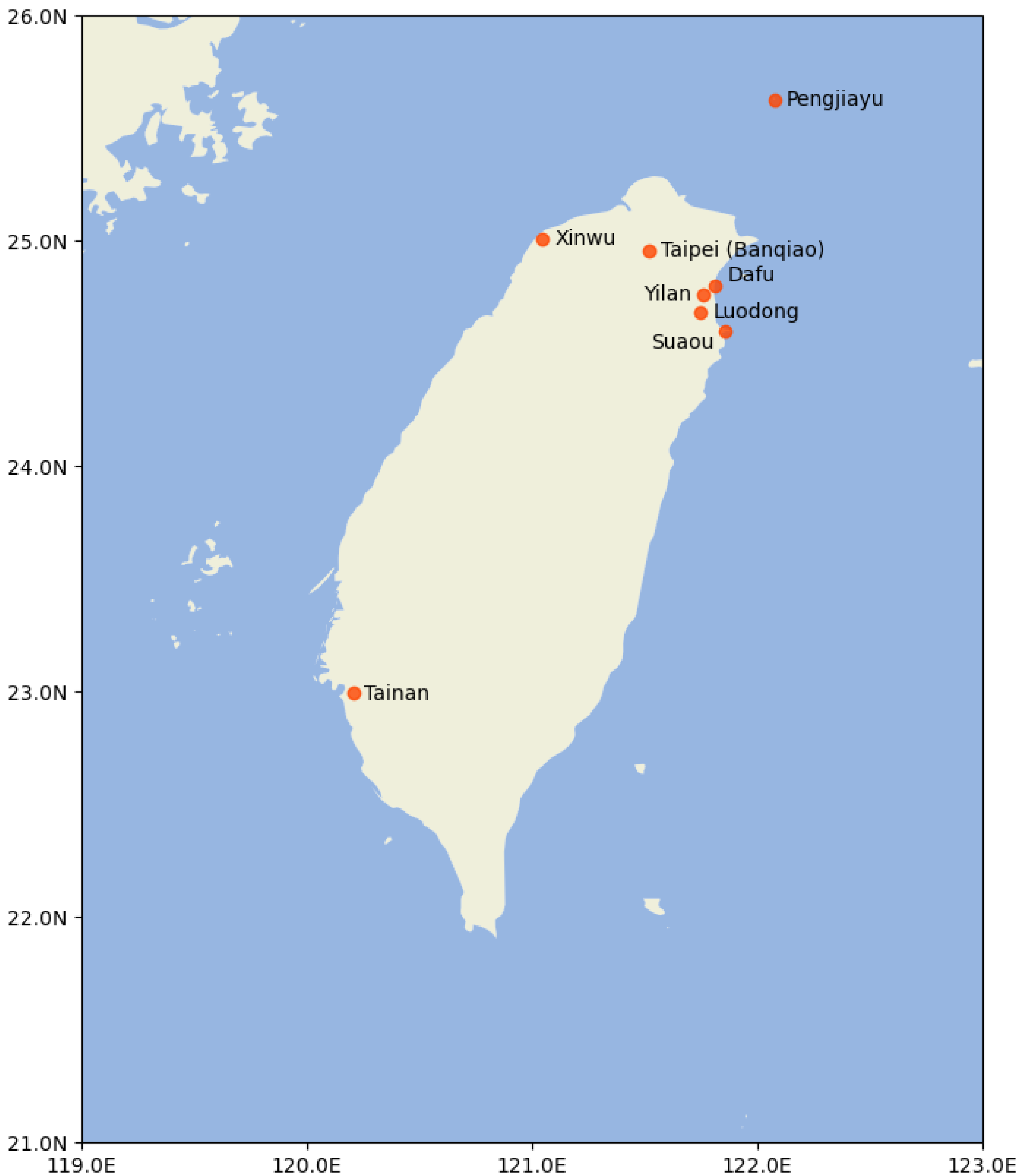
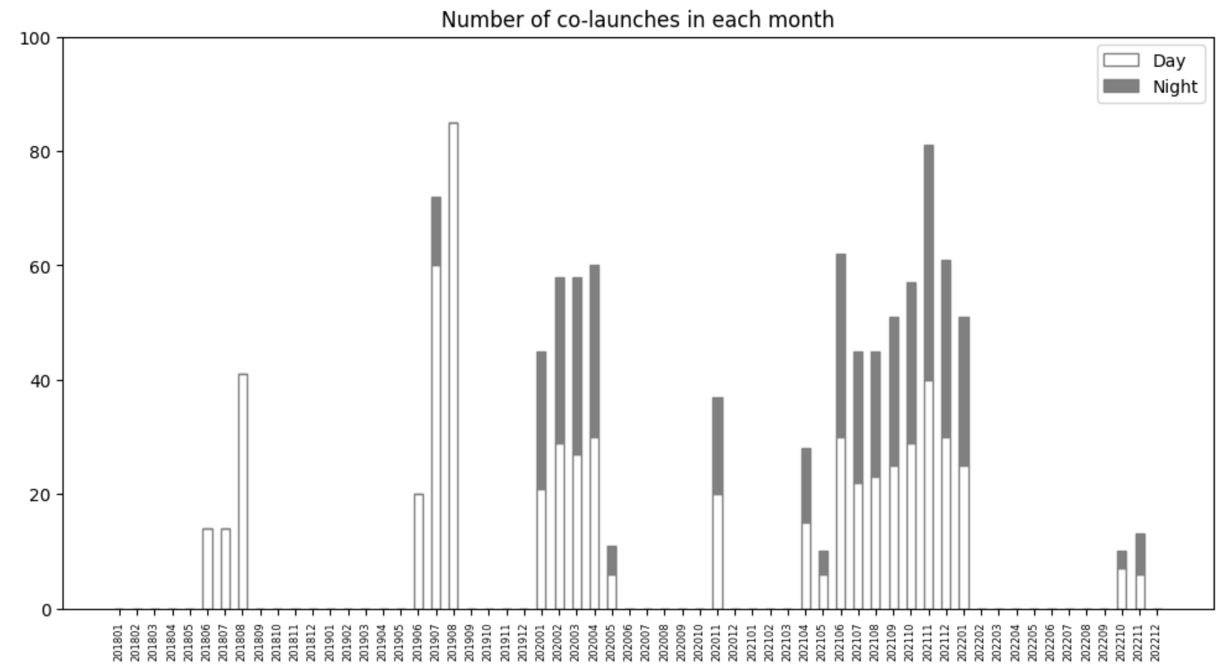


Table 2 appears very full. Leave out the zeroes, since they distract and don't provide any useful information. Alternatively use another way of presenting these data, such as a e.g. a bar chart.

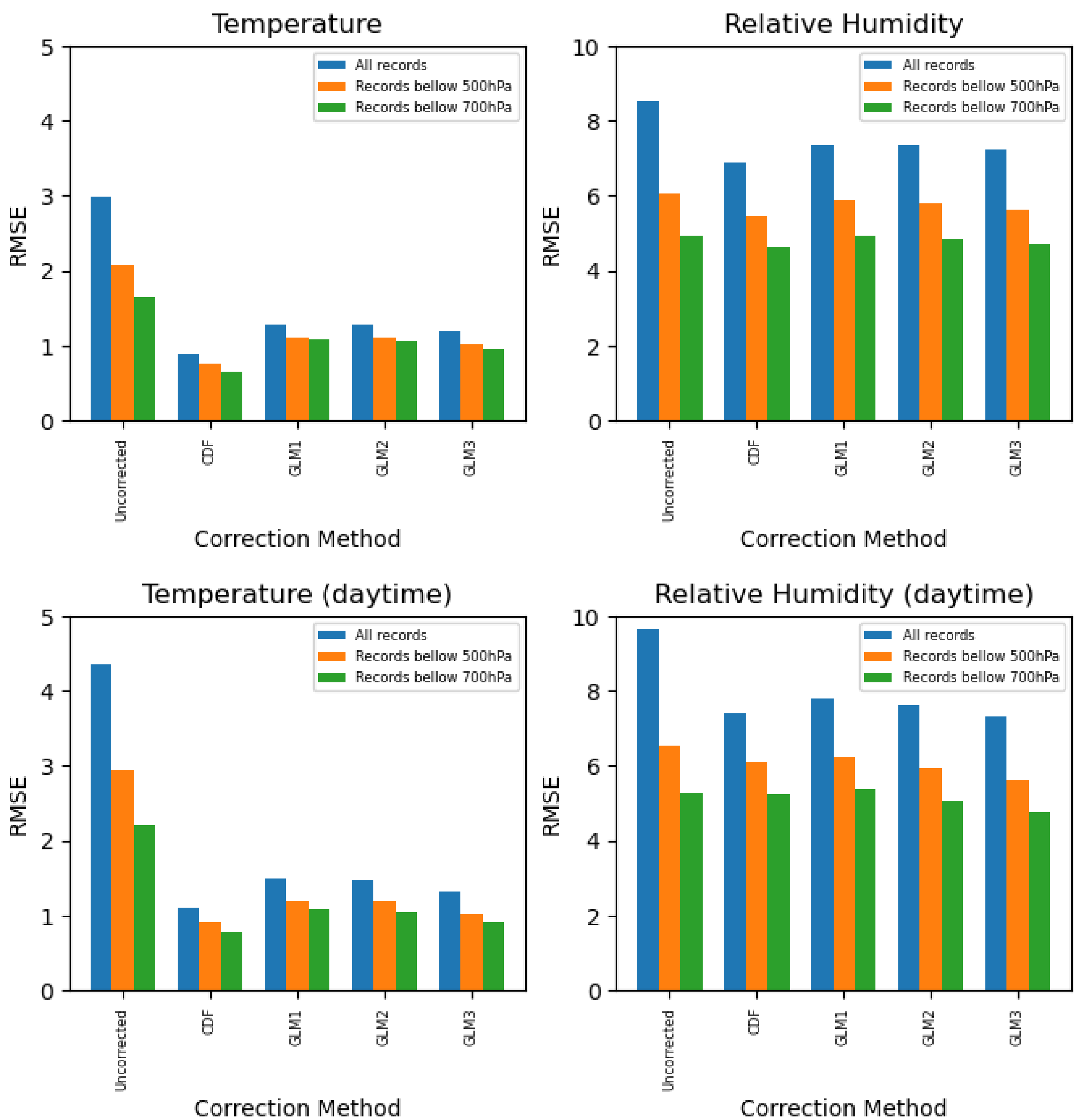
The bar chart as shown below is too dense for the information. Instead, we took the suggestion and left out the zeroes and updated the table as follows.

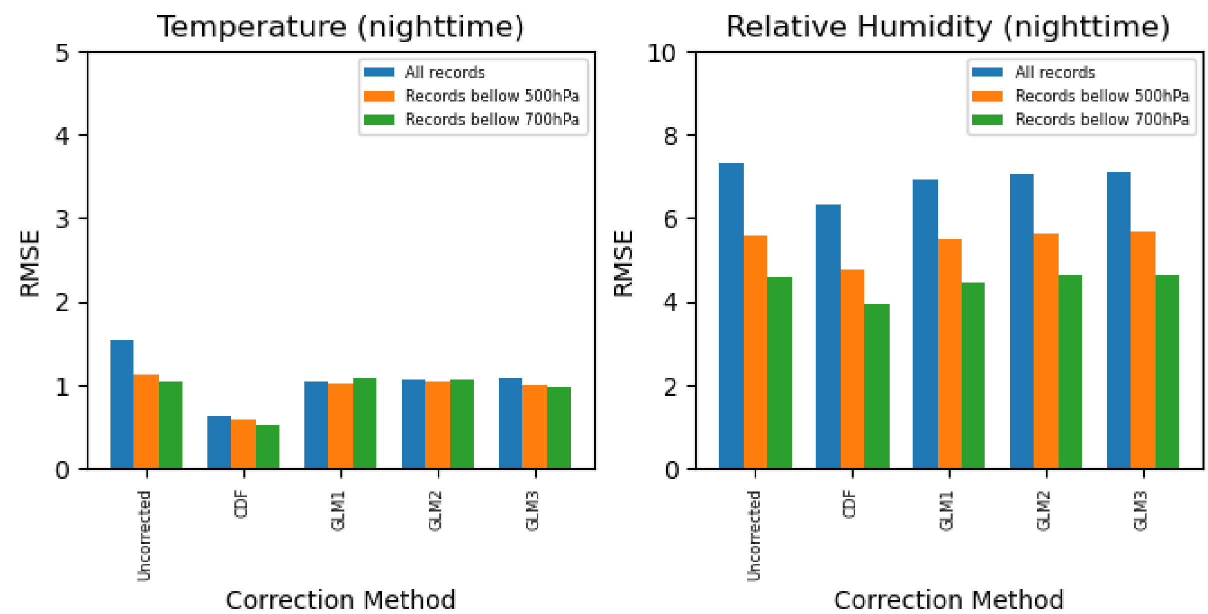


Month	2018		2019		2020		2021		2022		Total	
	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
1					21	24			25	26	46	50
2					29	29					29	29
3					27	31					27	31
4					30	30	15	13			45	43
5					6	5	6	4			12	9
6	14		20				30	32			64	32
7	14		60	12			22	23			96	35
8	41		85				23	22			149	22
9							25	26			25	26
10							29	28	7	3	36	31
11					20	17	40	41	6	7	66	65
12							30	31			30	31
Total	69	0	165	12	133	136	220	220	38	36	625	404

Figure 8: separate for daytime and nighttime data.

Day- and night-time analyses are separated as suggested.





The profile shown in Figure 11 exhibits an interesting discrepancy between RS41 and Storm Tracker around 800 hPa. There is a wiggle in the Storm Tracker's temperature profile that is not seen for the RS41. It coincides with a decrease in RH. Is this a radiation effect connected to a cloud top?

Though we cannot rule out the possibility of radiation effect, we tend to believe this belongs to a series of special cases when conducting co-launch during extreme weather. As described in section 5.1, a severe rainfall event was caused by the convergence of the tropical depression and the southwest monsoon from August 23 to August 30, 2018. The bias between ST and VS during this event was larger than usual. We presented this case to demonstrate the "worst case" scenario for using ST (well, it is not literally the worst case, but the 6th worst one if we didn't remove co-launches with too few records).