Line numbers are references to file with edits in suggestion mode.

REVIEWER 1

The current version is improved over the original, but I have a good number of questions about the study that should be addressed:

1. I assume that the data shown in Fig. 1 are from calculations based on this work, but that isn't clear.

We have changed the figure legend such that it is clear that the dots shown in upper panes are calculated from the IGRA data used in this study.

504

2. Line 109: Does this temperature limit vary with height? 373K is possible near the surface, but not near the tropopause, for example.

We agree this is very crude and would remove only extreme outliers mostly by coding errors. These do not occur in practice because such outliers are already set to missing within the IGRA2 quality control system https://www.ncei.noaa.gov/data/integrated-global-radiosonde-archive/doc/igra2-product-description.pdf.

We therefore now write: For temperature, needed for geopotential calculations, we relied on the IGRA2 quality control (Durre et al. 2018) that already removes gross errors. An additional very crude check (temperature between 173 and 373K) was applied just to verify that the data were read correctly.

113

3. Line 192: How is "improvement" measured?

Calculated displacements are verified against GNSS measured displacements, using the difference as error metric. The differences found when using a polynomial fit for ascent speeds were only negligible smaller than when using a constant ascent speed. We modified the text accordingly.

205

4. Lines 232-234: I believe that the older data are available in TAC format, which is in kt, and in degrees and speed. Is there code to do the conversion from TAC format to what is needed here, or a way to directly input the older data without conversion?

This is correct. Since there must be code to read from different formats, like IGRA or directly from TAC or other archives anyway, the software relies on the users to do these unit conversions.

5. Figure 4: It is interesting that meridional displacement differences are much smaller than zonal ones. Is this because the zonal wind is usually faster than the meridional wind?

Yes, that is definitely the reason, particularly in the jet stream regions. The majority of stations can also be found in the mid-latitudes.

6. Figure 8: It would be useful to know how many ascents are used at each level.

The figure now contains the number of ascents as in the other figures.

541

7. Section 5: I believe that the ERA5 assimilates the rawinsonde data as vertical profiles, which should be mentioned.

“This is to be expected, since radiosondes were assimilated as vertical profiles in ERA5.” was present in line 329.

8. Lines 309-310: I think you mean comparing the radiosonde observations as slanted profiles, not the forecasts.

We meant that the better fit to ERA5 can be found only with background forecasts, not the analyses, since those were drawn to the measurements at the wrong place. To make that a little clearer, we added (in contrast to analyses) to the sentence.

9: Line 522: I'm not sure what "vertical temperature" is. Is it the ERA5 temperature?

This was indeed not well formulated. We changed this passage now to “RMSE (obs - ERA5) of temperature assuming vertical ascents (orange, lower x-axis) and RMSE (obs - ERA5) of temperature from slanted ascents, taking balloon drift into account (red, lower x-axis).”

555

10. Figure 13: The caption says u (zonal wind), but the description in the text just talks about speed. Is this total wind speed, u-speed, or zonal velocity?

This has now changed to “zonal (u)” wind, to make it clear that only the u-wind component is meant.

576

Minor and pedantic comments:

1. When upper air is used as an adjective, sometimes it is hyphenated, sometimes not. It should always be hyphenated when it is an adjective (upper-air measurements), but not when it is a noun (which is never the case in this version of the manuscript).

It is now upper-air throughout the text.

2. Distinguish between wind (the motion of air) and quantification of that motion (wind speed, wind velocity, etc.). For example, line 77, the wind velocity is calculated.

Done.

3. Line 114: What is "ERA5 feedback?"

We removed feedback, which is actually just a technical term for some statistics calculated during assimilation. This has now been changed to “ERA5 background forecast differences”

4. Line 121: What are the "input data," and what is it put into?

“input data” is now replaced by radiosonde data for the displacement calculation

5. Base coordinates is defined twice (lines 125 and 133)

Second definition is now removed.

145

6. Table 3 is mentioned before Tables 1 and 2.

Changed.

7. Line 182: The ascents are the same, whether the data are reported in high-resolution or not. Suggest changing to something like high-resolution reports or high-resolution data (note that high-resolution should be hyphenated when it is an adjective.)

We changed to high-resolution data

8. Line 536: Figure 12 is labeled Figure 13.

This is now corrected.

576

REVIEWER 2

General: Quality control.

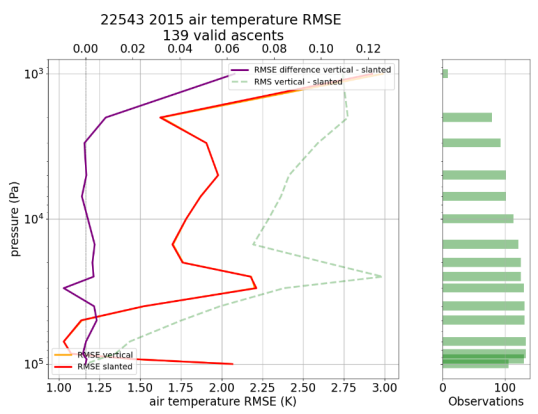
I am pleased to see that the results in section 6 use the standard ECMWF QC. It slightly reduces the independence of the background, but it does 'clean up' the data. One point that I forgot to mention originally is that non-GNSS Russian radiosondes can have large height errors at larger displacements and low radar-elevation angles (this would affect the comparison of all the variables with model fields) this is a consequence of relying on radar heights without a pressure sensor.

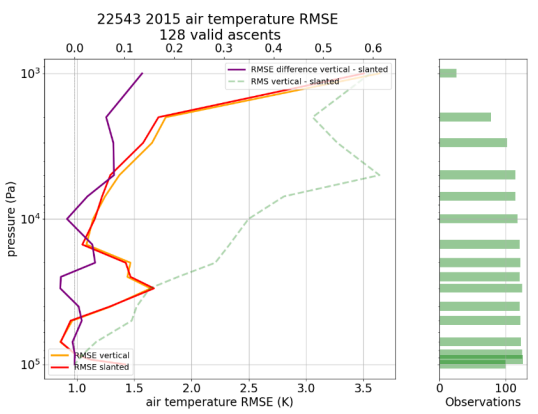
See Kats A., Balagourov A & Grinchenko V. (2005) The impact of new RF95 radiosonde, introduction on upper-air data quality in the North-west region of Russia. Poster Pw(07), TECO-2005, WMO/TD- No. 1265; IOM Report- No. 82.

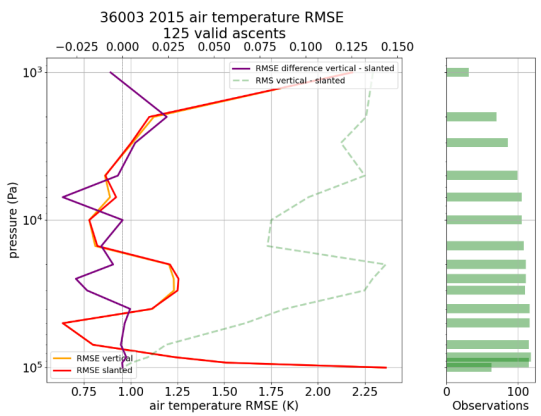
Thanks for this comment. We added the following text to the paragraph on quality control:

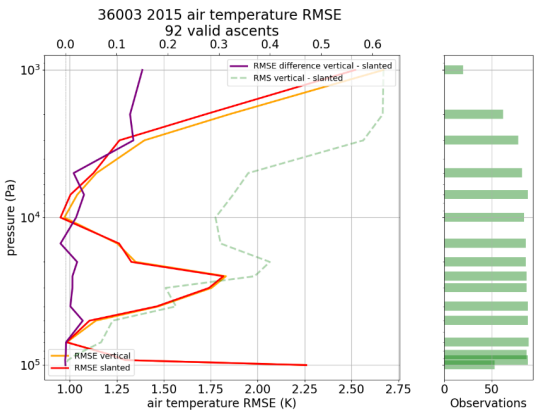
“The reason was not always the displacements themselves but also the fact that large lateral displacements can lead to large height errors in profiles from non-GNSS Russian radiosondes, since those have no pressure sensor but rely on radar heights (Kats et al. , 2005). However, even for these sondes, we found that taking into account the balloon drift reduces the differences to the ERA5 background forecasts.”

The two figures below show temperature RMSE profiles at two radiosonde stations using Russian sensors (22543 used a MRZ-3AK1 with MARL-A radar, 36003 used a PAZA-22M with MARL-A radar). The left panels show the differences for ascents with maximum displacements less than 0.5 degrees, the right panels for ascents with maximum displacements above 0.5 degrees. For the larger displacements we get better agreement by taking into account the drift. However this does not remove the height error problem, which appears to cause the peak in rms errors around 300 hPa, which can be seen at many Russian radiosondes.









Winds near the poles

The topic is mentioned now, which is a step in the right direction, but I think the issue of adjusting the wind direction should be discussed directly. My suggestion: "The WMO Manual on Codes states that for stations within 1° of either pole wind direction shall be reported in such a way that the azimuth ring shall be aligned with its zero coinciding with the Greenwich 0° meridian. There is currently an attempt to update this advice for BUFR reports, such that wind direction should be reported relative to the current reported longitude - to help in NWP use of such winds. Before comparing winds from the South Pole station with NWP fields they should have their direction adjusted when the drift positions are calculated (not currently done)."

The comparison of 'South Pole' winds with the ECMWF background winds could be looked at before and after drift calculation for the experiment in section 6. I am fairly sure that no NWP system uses that 1° convention when outputting model winds near the pole so _someone_ needs to adjust the wind direction.

It occurs to me that there are possible issues at other stations too, eg 71082 launches at 82.5°N, a 220 km drift could be about 15° longitude - should a direction adjustment be made?

Thanks for this comment. We implemented your suggested text. You are right that also for the northernmost stations this effect may lead to noticeable wind direction errors and should also be taken into account, but we did not mention that in the text.

Detailed comments

31. 'steep horizontal gradients' - 'sharp horizontal gradients' slightly better

Done

32

40. '(Ingleby, 2018)' - '(Ingleby et al., 2018)'

Done

42

68. 'procedural errors' Optional: could also mention the work of Gandin and Collins eg. Collins, WG, 2001: The operational complex quality control of radiosonde heights and temperatures at the National Centers for Environmental Prediction. Part II: Examples of error diagnosis and correction from operational use.

JOURNAL OF APPLIED METEOROLOGY 40, pp 152-168

Done, but we did not this reference, since it is a bit redundant to the literature already cited.

82. 'in the used input databases' - delete 'used'

Done

85

154. 'between the layers of the profile' - 'between the levels in the profile'

Done

168