

Please check carefully all the Durre citations. You have Durre 2016 which describes IGRA 2 as a NOAA internal document, Durre et al. 2016 which references the actual IGRA 2 dataset, and Durre et al. 2018 which provides a published description of data coverage improvements in IGRA 2. I appreciate that at least one of these references appeared during evaluation of this manuscript, but you (and only you) can help future readers by ensuring correct references in appropriate locations. I see the qualifications on page 9 but I suspect that Durre et al. 2018 might replace Durre 2016 in many places?

Please pay close attention to tense. Technically, all references to processes and features of IGRA should occur in past tense, while all references to this work should occur in present tense. E.g. this statement from page 9 (and dozens others like it) should change: “is assured in IGRA since its first version” should instead become ‘was assured in IGRA since its first version’. Otherwise, readers get confused about what happened before and what you have added here.

Page 4 line 25 “should address that issue too.” I think you mean that rigorous assessment of completeness of radiosonde humidity data should include vertical and temporal coverage as well as geographic coverage? Perhaps you should write ‘should address these issues simultaneously’?

Page 6 line 9: reader first encounters RAOB term. Define it as an acronym or as a coding term here? Not done in this manuscript until page 8 line 15.

Page 6 line 18: “That was no always been so” Need correction to proper English here.

Page 6 lines 28,29: I think you mean that, as opposed to expense and challenge of chilled-mirror hygrometers, weather services instead need to rely on lower-cost lighter instrumentation packages for their regular radiosonde operations? Perhaps some small changes in wording here?

Page 6 line 30: “since the time when balloon sondes were abandoned by national weather services”. Very confusing, please revise. I think you mean the change from ground-tracked balloons to radio-tracked balloons, with associated changes in sensors? “electric hygrometers” not quite the correct word, I think you mean electronic sensors?

Page 7 line 3: were, not where

Page 7 line 7: “measurements in that region” What? I think you mean measurements at those altitudes and temperatures?

Page 7 line 13: “RH varied in the range 10–20 %; the lowest temperature “ Make this two separate sentences: ... RH varied in the range 10–20 %. The lowest temperature ...

Page 7 line 15: ‘new’ instead of “newly”.

Page 7 line 16: a Wang et al. paper that you do not cite, published jointly with Vaisala, showed that most contamination came from outgassing of radiosonde packaging materials. Protective cap solved that problem, would not have solved a ‘rain’ problem. Two sensors alternately heated, used initially more often in dropsondes than in up-sondes, did address the rain / cloud saturation problem. If you need to recount the contamination work, you should do it accurately. (I found the DOI for a paper in JTech, listed below.)

Page 7 line 21: again, if you intend to discuss humidity measurements in the low temperature conditions over Antarctica, the French-US Concordiasi data set seems quite relevant. Notably, it includes dropsondes, upsondes and GPS occultation intercomparisons. Another Wang et al. paper, in GRL I think (I found the citation, included below). Again, if you choose to go into this detail, you should at least have the details correct! (Note: I do not have the Dirksen paper, which may include some of these references.)

Page 8 line 6: “on the observing practices intricated with sensor limitations” use a word other than ‘intricated’? I think you mean ‘combined with’ or ‘complicated by’?

Page 8 line 10: Leslie Hartten and her team did a very nice job of outlining present-day daunting logistic, communication and scientific challenges of radiosonde operations in remote environments, e.g. Earth Syst. Sci. Data, 10, 1165-1183, <https://doi.org/10.5194/essd-10-1165-2018>. One expects their data to appear in IGRA v3?

Page 8 line 21: confusing section. I think you miss a ‘)’ after the Durre 2016 reference? The reference to WMO stations also introduces confusion because all or most of the 2700 IGRA stations already mentioned carry WMO station ID numbers. What distinction do you draw between “derived data for a selection of WMO stations” and standard IGRA station data?

Page 8 line 23: again, confusing. “and the former version”. ‘Former version’ refers to IGRA? If so, state it clearly.

Page 8 line 25: ‘prior’ not “priority”.

Page 9 line 2: ‘also’ not “too”.

Page 9 lines 12, 13 “In wind data from pilot-balloons, the vertical coordinate is geopotential height (presumably adjusted from geometrical height measurements and the gravitational field).” Because this discussion refers to wind data, why do we care? Delete this sentence?

Page 10 line 11: “virtually” Not sure what you mean here, but perhaps you do not need this word?

Page 10 line 20: “the second world war” usually capitalized as ‘second World War’ or more often designated as World War II.

Page 10 line 20: ‘tripled’ not “triplicated”.

Page 11 lines 1,2: I think you intend that readers should compare total numbers here? E.g. for 800 to 900 stations at each twice per day one would expect 1600 to 1800 soundings. Instead one observes consistently fewer than 1500 soundings per day, perhaps more like 1200 soundings. If so, be explicit. Do not assume readers will imagine your intentions. Tell us what you want us to see.

Page 11 line 5: but you have just told readers that IGRA avoids soundings without valid temperature data. So, now, by RAOB, you mean valid temperature, RH, pressure, etc? But technically a RAOB might include temperature but not RH or vice versa? So in fact you mean ‘valid’ or ‘complete’ RAOB messages, instead of generic RAOB? If I get confused, readers will get more confused.

Page 11 line 7: “relatively few RAOB data”. Here, clearly, you define RAOB as consisting of full valid all-parameter data (T, RH, pressure, etc.). But earlier you have defined RAOB as a generic

radiosonde observation, quality unknown. To understand your selections and corrections, you need to adhere to, and give the reader, a clearer definition of what you mean by RAOB. Any sounding data, or only a valid full-parameter sounding? Which definition you intend makes a very large difference. Making the definition clearer will greatly improve the text of this manuscript.

Page 11 line 12 “the IGRA-RS subset retains practically all the RAOB soundings” Because of confusing definitions, how could any IGRA-RS subset not automatically consist entirely of valid RAOB data? I think the point you mean to make here is that taking only the radiosonde fraction of IGRA, e.g. IGRA-RS, still retains most of the original IGRA RAOB data. If so, then your earlier definition of a RAOB as a radiosonde observation (page 8 line 15) seems again confusing. Some RAOBs are not valid RAOBs? Or, some RAOBs are not valid radiosonde data, even though you define a RAOB as a radiosonde observation? I understand confusing often inconsistent meteorological terms, but here you have amplified the confusion? Readers need your best guidance but do not get it.

Page 11 line 16: “missing years are considered”. I think you mean ‘included’ or ‘included and identified’. ‘Considered’ does not tell us how your selection process treated missing years.

Page 11 line 17: here a reader learns that 1300 of 1700 stations (75%) carry WMO ID numbers. This does not explain nor accord with the statements on page 8 (noted above) about a “selection of WMO stations”.

Page 11 line 21: “integrating” I think you mean ‘integrated into’ or ‘coordinated through’?

Page 11 line 23: “together with the surface stations of the GCOS Surface Network” Why do we need this? Is this somehow relevant to the upper air humidity data? If not, omit?

P12 section 2.3: a very good description of the core motivation of this work! These questions should move to the top, even to the abstract. Readers should not need to wait until this point to understand the motivation!

Page 14 line 25: here a reader again finds reference to ‘RAOB’ reports when in fact the discussion pertains to IGRA-RS? More confusion?

Page 14 line 29: here the authors include moving stations but earlier, under global coverage, they only included fixed IGRA stations. Readers need better information about which subsets used when? If moving stations only a small fraction of total IGRA stations, why include them in this analysis? What value, if any, do they add?

Page 16 line 2: ‘shown’ rather than “show”.

Page 16 line 7: increases in a step-wise manner

Page 16 line 20: “readiness”? I do not understand this word in this context. Change, please.

Page 16 line 23: now we have “RS” stations. So, IGRA, IGRA RS, RAOB, RS. Do the authors follow a deliberate plan to confuse readers?

Page 16 line 29: “repeated, by restricting” Remove this comma.

Page 17 lines 3 to 5: finally, here, a reader learns about fixed versus mobile and which analysis used which subset. We should have had this information much earlier, at the start of section 2 or even as part of the introduction?

Page 17, line 11: "IGRA-RS excludes the IGRA stations without any RAOB at all" this statement is **NOT** consistent with earlier use of or definitions of terms RAOB and IGRA-RS. I suspect the authors know what they intend, but they have only confused their readers.

Page 17 line 11: "comparison between Fig. 3 and Fig. 1a" Following (correctly, I hope) the authors' intent, extrapolating from Fig 1a, in Fig 3 I should see, between 1955 and 1975, a 10-fold increase in non-humidity (e.g. PIBAL) stations - which my eyes do NOT see - and, between 1975 and 2015, a steady number of humidity stations accompanied by a decreased number of PIBAL stations. Why do I not see the 1955 to 1975 differences? Bad eyes? Bad figure? Can I actually confirm the drop in number of PIBAL stations in 2015 relative to 1975? Authors should provide guidance to readers about what the authors wish readers to see, and ensure that Figures support that evidence? Not clear in this instance?

Page 17 section 3.1 At the low given resolution of Figure 3 (it does zoom in nicely on my screen), the reader doubts whether we can confirm the temporal changes in geographic patterns described by the authors. A reader almost certainly has zero ability to detect "four fixed weather-ships". We either need descriptions better scaled to the maps or better maps.

Page 17 line 29: observations reported by Driemel actually included a large fraction gathered during transit, e.g. north and south along the Atlantic oceans, with perhaps the largest fraction between 60N and 60S? Polar yes, and very valuable, but not exclusively polar.

Page 18 line 6: use arctic or Arctic, but at least use it consistently. Copernicus style sheet suggests 'Arctic'.

Page 18 and Figure 4: Whatever the authors may have intended here, they have largely failed. Figure 4 remains almost impossible to understand, readers need to spend way too much time trying to understand it. What the authors' claim as climate zones actually represent latitudinal bands instead, and not evenly distributed in any case. Properly speaking, climate zones include elevations, distance from coastlines, location with respect to monsoonal circulations, etc. We get (combined) 46 degrees of equatorial, 23 degrees of sub-tropical, 60 degrees of temperate and 46 degrees of polar (using my own guesses at names for the zones). Figure 4 exacerbates this confusion, with quantities and lines in no particular order or calibration. How do we compare a northern sub-tropical range of 12 (23 to 35) degrees with an equatorial region of 46 degrees or a south polar region of 23 degrees? I have no doubt the authors understand the data to the resolution of "two ships reporting radiosonde observations in waters around the Arctic Circle" (page 18 line 9) but readers will not find anything like that detail in these figures. Going back to questions on page 12 (and notice there that the authors used the word "latitudes" rather than climate zones) do we really need any of this detailed location by location discussion? Instead, authors could help readers by defining latitude zones appropriately (30, 60, 90, etc.) or at least of equal latitudinal extent and then draw our attention to temporal patterns within those zones. Describe the data sufficiently so that subsequent users can explore specific latitudinal or zonal features based on their own criteria?

Page 18 lines 11,12: "relative weight has been decreasing" "weight"? I think you mean number or frequency?

Page 18 throughout: "Tropics", "extratropics", "climate zone", "latitude band" - terminology and punctuation very inconsistent throughout this section. Needs careful attention and

correction to achieve consistency as well as accuracy. Really too many to note them all, needs thorough scrubbing and appropriate revisions. Authors make appropriate notice of land to ocean differences by hemisphere but then compare Arctic (ocean) with Antarctic (land) without any such qualifications.

Page 18 line 30: "While the same is impracticable in many other parts of the world and over the oceans, distances up to two or three times larger than ideal are accepted, in view of the relatively mild climatic conditions on oceans and the fulfillment from surface and satellite observations." I believe the authors intend this as a description based on practical realities but many researchers would not agree that we should find the situation acceptable? We certainly need sea surface temperatures, surface roughness, cloudiness and rainfall, and interior ocean temperatures (e.g. by Argo) at much higher temporal and spatial resolution. Last sentence in this paragraph (e.g. top of page 19) also contradicts this statement? Statement represents a lightning rod, authors might do themselves a favor to omit it?

Page 19 line 9: "Besides, the corresponding ...". Delete the first word.

Page 19 line 14: "we only care with sub-year missing days" I think you mean 'we focus only on sub-year'?

Page 19 line 15: "Fig. 7 gives a glint of the typical continuity ...". I think you mean 'Fig 7 offers a summary of the typical ....' Or 'offers an indication'.

Page 19 line 16: "between 1945 and 1960" In fact, Fig 7 shows that number of missing days dropped much faster than your phrases suggests, over not more than 5 years. The pattern looks like an initiation or spin-up problem, which you have hinted at elsewhere.

Page 20 line 12:  $3/4$  (also in line 13). Please use percentages as you do elsewhere, not fractions.

Page 20 line 13: "data reach 22 km". In the previous line you gave us pressure then altitude, e.g. 100 hPa roughly 10 km. Here you should do the same, e.g. something like 50 hPa roughly 22 km.

Page 20 line 14: "This difference " refers to differences in maximum height or to differences in height in temperature records versus humidity records? Need clarity here.

Page 20 line 23: "This last feature and is coincident " remove the word 'and'

Page 20 line 30 - use percentage not fraction.

Page 20 line 31, 32 "fairly recent measurements in the upper-troposphere, and certainly more above too, was considered inadequate for climate". I think you mean 'and certainly into the lower stratosphere'? Recent as used here means before the Durre 2005 reference, so not the most recent. No widespread globally-useful solution, certainly, but other people have worked on this problem?

Page 21 line 14: Most mobile soundings come from ships and for those the baseline elevation is always sea level plus/minus 10 m at most?

Page 22 lines 5,6: Confusing. I think you mean that, for a fairly high standard such as expecting 95% of stations to have valid lower-troposphere humidity data, the number of stations meeting that standard remains very low from start of the records in 1945 to as recently as 1990. You use P in percentage to indicate fraction of stations while many readers familiar with statistical

analysis will understand P as probability. You need to give explicit explanation of what you describe and how you measure it.

Page 22 line 19: "two noteworthy change points: a sudden increase around 1970 and 2000". Singular plural problem: two change points result in sudden increases around 1970 and 2000.

Page 23 line 25: "Evidently, Fig. 11 is only ..." Remove the first word.

Page 23 line 25: "the question of since when we have enough" remove the word 'since'

Page 25 lines 10, 11: "selecting the stations with a minimal amount of radiosonde" I think you mean by selecting those stations with a sufficient amount? Minimal as you use it in this case technically indicates few or fewest, not what you intend. You mean 'exceeded minimal standards'?

Note - the first five sentences of section 6 make a very good abstract, much better and more concise than the one you have.

Again, arctic vs Arctic. Please check and correct!

Page 26 line 9 - use percentage not fraction

Page 26 line 22 "was not standard and it was rarely used" delete word 'it'

Page 26 line 23 "consecutive years of data until a given year" I think you mean number of consecutive years greater than some specified value? Later in that sentence, depends on the value, the time span and the completeness criteria.

Page 26 line 24 "E.g.: the station-based" Do not start a sentence with an abbreviation. Instead, write it out: For example.

Page 26 line 27 "Evidently, the equivalent time-series ..." delete the first word.

Page 38 - why do we get Figure 2 in black and while while we get other figures in useful helpful colors?

Page 39 - Figure 3 still very hard to read at page resolution, but works okay with page zoom.

Page 40 - Figure 4, latitude bands not climate bands, latitude bands inconsistent in extent, figure hard to read and harder to understand. Change to standard latitude bands as used in Figures 6 and 7.

Page 41 - potentially useful information but limited by use of the variable latitude bands. Use standard latitude bands as in Figures 6 and 7?

Page 42, Figure 6 - at a std deviation of 20%, evidently no significant differences among day fraction by latitude. If not, combine the lines into a composite, both for the absolute value and for the standard deviations. E.g. no valid distinctions between these lines so why show them separately?

Page 42 Figure 7 - see comment above about data that I made in reference to Page 19, above; needs color.

Page 43 Figure 8a - no differences statistically or visible between T and RH mean and quartile values in this panel, nothing gained by showing them super-imposed? Combine them? Or delete this panel and focus instead on panel b?

Page 44, figures 9 and 10. Why not color instead of grey-scale? Do we really need both 9 and 10, as they basically tell the same story? Panel a in Figure 10 could group to two lines, one before 1990 and one after? Comparison with Panel b would still hold?

Page 45, Figure 11, useful, but why not use the same color scheme both panels?

Page 46, Table 1: relevant to RAOB, IGRA, IGRA RS confusion above, here RAOB IGRA and RAOB IGRA RS have almost exactly identical numbers of soundings with humidity. Differences in total soundings only roughly 300 out of nearly 30 million, e.g. roughly 0.001% difference? What exactly drives the distinction in terms?

Two additional references mentioned above, authors to include if considered useful and relevant:

Wang et al., [https://doi.org/10.1175/1520-0426\(2002\)019<0981:COHMEF>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0981:COHMEF>2.0.CO;2)

Wang, J., et al., *Geophys. Res. Lett.*, 40, 1231–1236, doi:10.1002/grl.50246