06 April 2019

Dear David Carlson,

I and my co-authors would like to thank you for all your corrections and suggestions. We hope our revision is sufficient for the manuscript to be accepted for publication in ESSD. The authors´ response is provided below, followed by the marked-up manuscript.

Sincerely,
António P. Ferreira

---

## Authors response to Topical Editor report for the manuscript

'Completeness of radiosonde humidity observations based on the IGRA'
by António P. Ferreira et al. (MS No.: essd-2018-95)

*Topical Editor's comments are highlighted in blue. Page and line numbers in the authors responses refer to the previously revised manuscript, unless stated otherwise*

---

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?

Since the (meta)dataset and the analysis presented in our work derive from the current IGRA, it would be impossible to neglect the proper citations.

The first version of the manuscript included an obvious reference to the paper by Durre et al. (2006), which describes IGRA v1; another obvious reference to Durre et al. (2016) indicating the IGRA v2 dataset used by us; and two personal communication citations, since Imke Durre had clarified some questions on IGRA 2 during the preparation of the first draft of the manuscript and gently provided us a first-author technical description of IGRA 2 (the NOAA internal document). Two other references to Imke Durre and co-authors (2005, 2009) were also included.

During the revision process, we added two major references for common-sense reasons. 1) Durre et al. (2018), the recent paper describing IGRA 2; when preparing the discussing manuscript, we were unaware of the publication of that paper a few weeks before; that explains why the paper was only cited in our first revision (on page 3 in the Introduction on page 9 in the section about IGRA sounding data). 2) Durre (2016), referring the NOAA internal document, replaced a personal communication citation since we noticed that

that document became available on the NOAA website for IGRA practically at the same time of the communication in which Imke Durre gently provided it to us.

Concerning the present revision:

**->** Durre et al. (2016) is added to Durre (2018) on page 3, line 10, where IGRA is first mentioned;

**->** Durre et al. (2018) is added to Durre (2016) on page 9, line 10, since both works give the same information in different ways;

**->** Durre et al. (2016) and Durre et al. (2018) are both cited again in the opening lines of the concluding section 6.

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.

We made a number of significant corrections. Note - sometimes we keep past tense to refer what we did before performing calculations.

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'?

Changed:

"issue of geographical coverage" → 'geographical coverage of stations'
"should address that issue too" → 'should address both 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.

The mistake resulted from the addition of the related paragraph in the last revision. We prefer not to us the term RAOB here since it is only needed from section 2 to 4. The less technical parts of the manuscript (abstract, introduction and concluding section) avoid unnecessary use of acronyms. "RAOB reports" → 'radiosonde reports'

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

'That has not always been so'.

Page 6 lines 28,29: I think you mean that, as opposed to expense and challenge of chilledmirror 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?

"meteorological radiosondes are" was replaced by 'whether services need to rely on meteorological radiosondes consisting of'.

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?

"Balloon sondes" is one of the names for registering balloons – very confusing indeed since it can also refer (uncommonly) to radiosondes. Replaced by 'registering balloons'.

The terminology "electric hygrometers" follows from DuBois (2002) but it can be found in other places, e.g., Meteorological Measurement Systems by Fred V. Brock and Scott J. Richardson (Oxford University Press). Although radiosonde measurements involve a lot of electronics besides radio transmission of data, the above terminology derives from the fact that most humidity sensors for remote applications are based on the change of an electrical parameter caused by moisture: capacitance, in polymer capacitive sensors; resistance, in the carbon hygristor; ionic conductivity, in electrolytic humidity sensors like the LiCl electrolyte sensor invented in 1937.

Page 7 line 3: were, not where

Typo corrected.

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

We mean low HR region in the domain 0-100%. Replaced by 'measurements in that low RH region', where "that" refers to the low values mentioned in the previous sentence (less than 15–20%).

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 …

The two sentences form part of a list with many items. This is now separated by subject.

P7, L12-15 – Rephrasing (changes are underlined):

'Note that changes in instrument and reporting practices in different countries took place at different times: the threshold value of RH varied in the range 10–20 %; the lowest temperature of −40$^{o}$C for reporting humidity, and the shift to lower temperatures, was applied in different periods depending on country; humidity could be reported up to a specified pressure level. Moreover, mechanical sensors are not exclusive to pre-1940s radiosondes: hair hygrometers were only abandoned in the mid-1950s and rolled hair hygrometers were used in a few places until about 1980; the goldbeater's skin sensors introduced in the 1950s became particularly important in the Soviet Union. [For historical details on these changes, see Gaffen (1993).]'

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

Corrected.

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.)

Thank you for enlightening us. We were referring to the protective rain cap, not to the protective shield to avoid chemical contamination. Both were used in RS80 radiosondes, but absent in RS90 and RS92. Now we understand the outgassing issue affecting RS80 sondes until 2000. While for RS90 and RS92 the major problem was solar radiation.

The text was extended for the sake of accuracy. Telling the main features of humidity sensors in the main Vaisala radiosondes seems a solution. We have cited the suggested paper (Wang et al., 2002) and also a book chapter (Smit et al. , 2013). Please see the changes on p. 8 of the marked-up manuscript.

Page 7 line 21: again, if you intend to discuss humidity measurements in the low temperature conditions over Antartica, 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.)

We have mentioned the low temperatures over Antarctica by way of example and see no need for further discussing the subject. We appreciate the suggestion of Wang et al. (2013), which we found interesting; however, it only deals with temperature sensing.

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’?

Yes, we mean “complicated by’ in the sense that it may be difficult to tell if missing humidity data in a sounding report results from limitations of a specific humidity sensor (nominal working range) or rather to reporting practices related to sensor limitations; ‘combined with’ is suitable.

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?

That paper is quite illustrating. It is cited now, with few words in sense indicated by you.

It would be wonderful if IGRA could include observations from such field campaigns. But for now, "data sources containing either field campaign observations or fewer than two years of data were set aside because of both their limited contribution to the dataset as a whole and the often considerable effort required for reformatting any one data source" (Durre et al. 2018)

P8, L9 – Added text:

'Radiosonde operations in remote environments, particularly performed from ships, present their own challenges; Hartten et al. (2018) give a vivid illustration.'

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?

Apart from the fact that we have missed the parenthesis after the Durre el al. (2016) reference, the sentence is accurate. There are hundreds of stations in IGRA that do not have a WMO ID number. The distinction between sounding data (raob and pibal) and derived data for a selection of WMO stations (derived from raob at surface level and standard levels) is inherent to IGRA, with separate datasets. However, since the latter is irrelevant to the paper, we have deleted the first part of the sentence, which now reads:

'This paper concerns with sounding data (Durre et al. 2016), comprising over 45 million soundings from 2761 (2662 fixed and 99 mobile) stations [based on data accessed in September 2017].'

Page 8 line 23: again, confusing. "and the former version". 'Former version' refers to IGRA? If so, state it clearly.

Yes, it refers to IGRA 1, as stated now.

Page 8 line 25: 'prior' not "priory".

Corrected.

Page 9 line 2: 'also' not "too".

Replaced by 'mostly'.

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?

We find this information useful to distinguish PIBAL from RAOB data in IGRA. In IGRA files (one file per station) the sounding data are arranged in the same way no matter if they contain PIBAL or RAOB data. RAOB uses pressure as default vertical coordinate, although geopotential height may be given. Pressure is only present in 28 PIBAL stations (0.03% of all PIBAL stations), as noted on footnote 5. Note: The word "presumably" means that we cannot ascertain if geopotential height was always properly calculated; since nothing proves the contrary either, that word was removed.

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

We mean almost or practically. Since we give the end year (2016), it can be deleted.

Page 10 line 20: "the second world war" usually capitalized as 'second World War' or more often designated as World War II.

Corrected to capitalized and shortest form.

Page 10 line 20: 'tripled' not "triplicated".

Corrected.

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.

Your numbers refer to the years after around 1970, when the numbers become stable. Numbers of observing stations and of raob per day for other years are quite different, while the conclusion is the same since around 1955.

In the table below, NO-T (NO-H) denotes missing obs. for temperature (humidity) in %

| YEAR: | 1945 | 1955 | 1965 | 1975 | 1985 | 1995 | 2005 | 2011 |
|---|---|---|---|---|---|---|---|---|
| NO-T: | 76.8 | 25.5 | 24.7 | 13.2 | 14.8 | 26.9 | 26.0 | 23.6 |
| NO-H: | 77.4 | 30.5 | 24.6 | 13.6 | 14.9 | 27.0 | 26.0 | 23.6 |

(Note: for 1950, NO-T and NO-H exceed 60%)

Averaging the above values for the years 1955 through 2015, we get NO-T = 18.4% and NO-H = 22.9% – i.e., roughly 20% for both parameters (1 in 5 soundings). Note that this estimate agrees with Figure 6 for the selected stations 'IGRA-RS'.

We prefer to present a crude estimate rather than overloading the reader with numbers, and suggest the following adjustment:

"roughly 1 in 5 days during the year, on average, as it can be concluded by comparing Fig. 1a (RAOB) with Fig. 1b (TEMP, HUM)."

→

‘roughly 1 in 5 days during the year, on average for the years after the mid-1950s [as concluded by comparing the yearly number of observing stations (TEMP, HUM in Fig. 1a) with half the global number of daily observations (TEMP, HUM in Fig. 1b)].’

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.

The identification of RAOB reports in IGRA, as opposed to PILOT reports, is essential for selecting the IGRA stations of interest to study humidity data and to derive our dataset. Both acronyms are defined in the beginning of section 2, on PP 8–9.

Technically, all RAOB messages include pressure, temperature, humidity and wind at several pressure levels (wind measurements require an external tracking device; only the vertical position is derived from radiosonde data). But humidity may be absent, as explained in Sections 1.1 and 1.3. This happens more often in the highest sounding levels, as shown in Section 3.3 (Fig. 8b), but it can happen also in the lower and middle troposphere. In addition, most radiosonde measurements prior to 1945 did not included humidity. This is shown in Fig. 1. Apparently, some stations begun to measure humidity much later than the advent of radiosondes. For example, the Lindenberg station was pioneering in measuring air temperature at fixed pressure levels (using meteographs before at least 1930). But humidity data for this station begins in 1958 in IGRA; it is very unlikely that this is has to do with IGRA’s quality assurance procedures.

RH and DPD cannot be obtained without temperature measurements, of course. IGRA takes this into account in its quality checks by removing humidity data which is not accompanied by valid temperature data.

In short, a RAOB message must have at least temperature data, while humidity and/or wind may be missing. So, temperature data, together with pressure, clearly identifies a RAOB. While many IGRA stations have only PIBAL data, many other stations have both PILOT and RAOB data during their period of record – with PILOT (i.e., non-RAOB) data denoting the earlier years in many cases.

On page 9 we explain how the ‘IGRA-RS’ stations are selected based on the percentage of RAOB out of the yearly soundings of any kind (RAOB + PIBAL).

However, we understand that the manuscript might be unclear in some parts and hope that the following changes bring clarity. (Please see also response to comment on page 17, line 11.)

<u>P10, L4-6 – Rewording</u>:

“Figure 1a shows the yearly number of the RAOB stations – i.e., the stations with radiosonde observations at any time of the year, at least of temperature, regardless of humidity/wind – and of the PIBAL stations – i.e., the stations measuring only wind throughout the year (assuming data are not lost).”

→

‘Figure 1a shows the yearly number of stations reporting RAOB any time of the year – meaning they have at least observations of temperature regardless of the simultaneous humidity/wind observations –, and of the stations reporting PIBAL observations alone – i.e., reporting only wind throughout the year.’

P9, L10-13 – Rephrasing, additions and shift to a separate paragraph:

“In radiosonde data, consistence between pressure and geopotential height (whenever the latter is reported in source data) is assured in IGRA since its first version. In wind data from pilot-balloons, the vertical coordinate is geopotential height (presumably adjusted from geometrical height measurements and the gravitational field).”

→

‘Note that a RAOB message must have at least temperature data at several pressure levels, while humidity or wind data may be missing, and geopotential height is not always given. The recording of pressure levels, and consistence between pressure and geopotential height whenever the latter is reported in source data, has been assured in IGRA since its first version. IGRA uses a consistent data format, irrespective of the provenience of the data (PIBAL or RAOB). Therefore, RAOBs in IGRA can be simply identified by the presence of temperature data. Wind observations from pilot-balloons (PIBAL) have only wind data at several geopotential heights (adjusted from geometrical height measurements and the gravitational field).’

P10, L4 – Word addition:

“because temperature is required to measure RH.”

→

‘because temperature is required to measure RH or DPD and so all humidity data in IGRA are accompanied by temperature data.’

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.

The phrase “some of them contribute with relatively few RAOB data” means that some IGRA-RS stations have an amount of RAOBs below average, because of the relatively high percentage of PIBAL observations in their period of record. The term RAOB is always used as an abbreviation for radiosonde observation.

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.

By "IGRA-RS subset" we mean the subset of IGRA stations which we refer to as "IGRA-RS" and is previously defined in Section 2.2 (as well as in the Abstract with less detail). However, despite discharging pilot-balloon stations (without any RAOBS), IGRA-RS does not automatically consist entirely of RAOB data. Some IGRA stations have performed a certain number of PIBAL observations during its period of activity, as explained in the beginning of Section 2.2. IGRA-RS includes those stations because the primary goal of our work is to derive a dataset based on data content for each station (not to extract humidity data from IGRA, which is up to users based on our metadata). Nevertheless, as shown in Table 1, "the IGRA-RS subset retains practically all the RAOB soundings" of IGRA.

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.

"Since missing years are considered" was replaced by "Since the humidity time series may be interrupted for long periods of time". Missing years are included in the statistics of record in Table 1, as explained in the remainder of the sentence: "the full period of record of one station may be segmented into two or more periods for humidity". We added '(both are rounded to years)' at the end of the sentence.

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".

The number 1300 clearly refers to WMO stations among the IGRA-RS stations listed in Table 1. To avoid confusion, the clause about "a selection WMO stations" was deleted on page 8 (although referring explicitly to the derived data set belonging to IGRA 2, such data are not used in our study).

Page 11 line 21: "integrating" I think you mean 'integrated into' or 'coordinated through'?

Changed to 'integrated into'.

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?

It is only complementary, so it can be omitted.

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!

Actually, the abstract includes all the five points, although in a much formal manner. More importantly, our work has two purposes: 1) to elucidate the completeness of the radiosonde humidity observations in global terms, using IGRA since this a wide-ranging archive; 2) to provide metadata describing the completeness of humidity observations for each station selected from IGRA. This two-fold purpose is stated in the abstract and clarified in the Introduction (P3, L13-28). Lastly, we prefer to have these five questions here. They are intended to motivate the reader to go through Section 2.3 [before looking the results of Section 3] which is central to the paper and is structured in five related sub-sections.

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

Not at all. The sentence refers to finding the vertical extent of humidity soundings from RAOB reports in general. IGRA retains source data if they pass certain quality checks. So, the discussion applies equally to the RAOB data found in IGRA – particularly in IGRA-RS, since this is the subset *of IGRA stations* that contains practically all RAOB data.

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?

Moving stations (about one hundred) form part of IGRA and so they are included in the dataset introduced in our work. Although individual time series are relatively short-lived in many cases, they have made an important contribution to upper-air observations.

The analysis provided by paper serves the purpose of illustrating the completeness of radiosonde humidity of observations on average terms. The dataset contains detailed information for each station.

Among the several aspects regarding completeness of observations, global coverage and temporal completeness are the ones which are expected to depend more on latitude. Including moving stations in a statistical analysis by latitude bands is impracticable. That is why Sections 3.1 and 3.2 excludes moving stations.

Concerning methodology, the use of only fixed station or all stations in the analysis of sounding data was stated in each of the five subsections of section 2.3 [see P12, L28; P14, L9; P14, L29; P16, L13; P16, L26]. However, for clarity we do now the following word additions:

P13, L27: "Eq. (1) was applied to the IGRA-RS stations" → 'We have applied Eq. (1) to the IGRA-RS fixed stations'

P15, L4: "humidity soundings" → 'humidity soundings from all IGRA-RS stations (including mobile)'

P16, L13: "percentage of stations" → 'percentage of stations (fixed and mobile)'

Page 16 line 2: 'shown' rather than "show".

Corrected.

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

We take the chance to be accurate by rephrasing:

'is almost step-wised, with a discontinuity around 1992" → 'increased almost as a step-function around 1992'

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

Corrected as 'To simplify'

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?

We can only agree that the acronym "RS" is superfluous and was used in an ambiguous way. It is now replaced by one of two different words (either 'radiosonde' or 'IGRA-RS') throughout the manuscript, depending on context:

P16, L23: "RS" → 'radiosonde'

P16, L24: "RS" → 'IGRA-RS'

P22, L5: "RS" → 'radiosonde'

P22, L10: "RS" → 'IGRA-RS'

P22,L24: "RS" → 'IGRA-RS'

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

Removed.

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?

The lines in question specify the number of mobile and fixed stations involved in the results shown in Section 3. Section 2.3 (describing the method of analysis of IGRA data) provides this information beforehand. We hope this is now clearer in our revision: please see the changes in that section, detailed in our response to the comment on P14, L29. But we agree this may not be enough.

The same information is now made clear in the summary of contents given at the end of the introductory part of the Introduction. We take the chance to outline the selection of IGRA stations in that same part, introducing the term IGRA-RS (after the Abstract) even though this is formally defined in Sect. 2.2.

P3. L33 – P4, L4:

"Section 2 indicates the IGRA data set used in the study and explains the data analysis. Section 3 presents a global picture of the completeness of humidity observations over the years, as derived from the IGRA stations reporting a minimum of radiosonde data, i.e., discharging stations with practically wind-only data in their period of record. Section 4 provides the definition of the metadata parameters describing the completeness of humidity observations from each relevant IGRA station – both in terms of annual statistics and in individual soundings – and the format description of the corresponding data sets. (…)"

→

'Section 2 indicates the IGRA data set used in the study; selects the IGRA stations reporting a minimum of radiosonde data (coined as 'IGRA-RS') by discharging stations with practically wind-only data in their period of record; and explains the data analysis. Section 3 presents a global picture of the completeness of humidity observations over the years, as derived from the IGRA-RS stations. The geographical coverage and temporal completeness of annual observations are detailed by latitude bands, thus restricting to fixed stations, while other aspects on the global scale include both fixed and mobile stations indistinctively. Section 4 provides the definition of metadata parameters describing the completeness of humidity observations from each IGRA-RS station – either as annual statistics or for individual soundings – and the format description of the corresponding data sets. (…)'

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.

RAOB stands for radiosonde observation(s), as defined on P8. IGRA-RS refer to the IGRA *stations* selected as described in Section 2.2. By excluding the IGRA with less than 5% RAOBs out of the annual soundings in every year [equivalent to retaining the stations with at least 5% RAOBs in any year], PIBAL *stations* are immediately removed (meaning the IGRA stations with only pilot-balloon observations during their period of record, i.e., without ant RAOB).

Please note that a single station can perform PIBAL observations during part of its period of activity and RAOB during other part. Thus, some IGRA-RS stations do have a significant fraction of PIBAL observations besides RAOB. Our analysis is based on all data from IGRA-RS stations. It is impossible in many cases to classify a station as a RAOB or PIBAL station – except for the sites which performed only PIBAL launches during their existence. Of course, some stations have performed consistently radiosonde launchings since they were open. Usually the term "radiosonde station" refers to sites that have taken RAOBs, even if not always (particularly at stations with a long history). The term "RAOB station" was used by us exceptionally in relation to Figure 1a, to denote a station reporting any RAOB during a specific year. Conversely, the stations not reporting RAOB during a given year, but only PIBAL observations, were called “PIBAL stations”. This terminology was convenient to study the yearly number of stations reporting either RAOB or PIBAL data. Now we understand that this terminology is quite confusing, which demands minor but significant changes in Figure 1 – by detailing the figure caption, removing unnecessary information and correcting the figure legend in panel-a – and in the related text on P8.

P8, L3-4

“RAOB stations” → ‘stations reporting RAOB’

“PIBAL stations” → ‘stations reporting PIBAL observations alone’

Figure legend of Figure 1a:

“RAOB” → ‘TEMP’

“Higrom.“ → ‘HUM’

“Higrom. > 95%’ → ‘HUM > 95%’

“PIBAL” → ‘WIND-only’

For the related changes in figure caption, see p. 40 in the marked-up manuscript.

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?

As stated both in the text and figure captions, Figure 1 refers to IGRA whereas Figure 3 refers to IGRA-RS. IGRA has almost 1000 PIBAL stations (no RAOB in period of record), while IGRA-RS excludes them all. So, to answer to the objection, the evolution of PIBAL stations is not expected to be seen in Figure 3.

We wish readers can understand the difference between the total number of IGRA-RS stations (red crosses + blue dots) and the number of IGRA-RS stations reporting humidity (blue dotes), as both numbers are given above each map of Figure 3. But we admit that drawing a conclusion by comparing Figures 1 and 3 is not straightforward. The argument we had in mind is the following: Figure 1 shows that the relative number of IGRA stations with temperature data any time of the year but without humidity data at all (relative difference between the black and solid blue lines in Fig. 1a) is very small. Thus, most of the stations not reporting humidity in a given year must report PIBAL data rather than temperature data.

This is now clarified on P17, L11:

"Although IGRA-RS excludes the IGRA stations without any RAOB at all, a comparison between Fig. 3 and Fig. 1a indicates that most of the IGRA-RS station-years without humidity data correspond to periods of wind-only observations."

→

'The IGRA-RS retains practically all RAOB data of IGRA; however, some stations have years with only PIBAL observations. Since almost all of the IGRA stations measuring temperature in a given year do also measure humidity at least part of the time (as seen by comparing black and solid blue lines in Fig. 1a), it is clear that most of the IGRA-RS station-years without humidity data (red crosses in Fig. 3) correspond to years of PIBAL observations alone (no RAOB).'

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.

The size of dots and crosses in Figure 3 was optimized to resolve the location of stations, without merging the ones very close to each other. Since ocean weather-ships are not marked on the maps (being very few in the years represented), we think it is better to remove details about their varying number when describing specifically each map. Also, the lines about the importance of ships of any kind is now in a separate paragraph and was slightly revised. Please see changes on pp. 18-19 of the marked-up manuscript.

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.

We agree that Atlantic Ocean regions are worth mentioning. Driemel et al. (2016) also reported a large proportion of radiosonde launches at latitudes above the polar circles. Note: 29 Dec 1982 was reported as the date of the first radiosonde launching. However, data in IGRA begin in 1985, with an interruption of 6 years between 1994 and 1999 without any data. Maybe missing years have insufficient radiosonde data to be in IGRA.

P17, L28:

"The polar missions between 1985 and 2014 of the ice-breaker and research vessel Polarstern (Driemel et. al, 2017) (…)"

→

'The missions to the Arctic and Antarctica performed by the ice-breaker and research vessel *Polarstern* (Driemel et al., 2016), covering also Atlantic Ocean regions during transit, provided substantial radiosonde humidity data in the periods 1985–1993 and 2000–2014 (…)'

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

The adjective arctic was replaced by the noun Arctic as part of the changes related to changes in Figure 4 (see next response)

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?

Concerning Figure 4

Figure 4 was intended to show how the number of humidity-reporting stations evolved with time in different latitudes, and to highlight hemispheric differences. We agree that is was not easy to read because of the large number of latitude bands plotted; and that the curves could not be compared between each other, except for homologous curves in opposite hemispheres.

So, Fig. 4 was modified by restricting the plot to four latitude bands of equal area, two per hemisphere: 0–30$^{o}$ and 30$^{o}$–90$^{o}$, representing tropical and extratropical latitudes on each hemisphere. In addition, the vertical scale changed from linear to logarithmic to better detail the rapid grow in the early years.

Concerning terminology

By "climate zones" we mean the Earth´s major climatic zones, which can be schematized as follows: tropical zone → region between the tropic circles (0–23.5$^{o}$N/S); subtropics → 23.5$^{o}$– 35$^{o}$; temperate zone → 23.5$^{o}$–66.5$^{o}$; polar zone → latitudes above the polar circle (66.5$^{o}$–90.5$^{o}$). We know that it's possible to refine the definition of these major zones based on climatic parameters and surface features, so that they are not exactly belt-shaped. Although the average boundary (mean latitude) between each other is arguable, the values given above are often used, among others certainly. The subdivision 30, 60, 90 is a modern simplification that rounds latitudes and neglects the subtropics; that is precisely why we used it later in Section 3.2, Figures 6-7 (further combining hemispheres since hemispheric differences are irrelevant in that context).

However, to avoid possible confusion with climate zones of specific areas of the globe (closely related to climate types), in the revised manuscript we replace the ambiguous term "climate zone" by 'latitude band'.

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.

Maybe the reason why this page of Section 3.1 seems inconsistent in punctuation is that the paragraph break appearing in line 10 should be at the end of line 15. However, we must recognize that page 18 has deficiencies regarding clearness.

The comparison between the Arctic and Antarctica refer to Figure 4, appearing before the discussion of Figure 5 which is when land to ocean differences are considered.

We keep the detailed latitude bands only in Figure 5 for the sake of detailing the stations density by latitude. But we do not use the term "climate zone" anymore.

Apart from correcting the paragraph break, the changes in the manuscript consists in adapting the discussion of Figure 4 (modified as explained in the previous response), defining more precisely terminology, and caring about clarity and accuracy.

We also adjust the first bullet paragraph of Section 6, due to corrections to Section 3.1.

Please see the changes on pp. 18-20 and p. 28 of the marked-up manuscript.

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?

We understand the objection, although we have written "accepted", not "acceptable". Yes, the description intends to represent reality. To circumvent misinterpretation, we rephrased the last part of the quoted sentence:

P18, L30: "are accepted, in view of the relatively mild climatic conditions on oceans and the fulfilment from surface and satellite observations." → 'need to be filled by satellite-based data and supplemented by surface observations, which are generally much denser than radiosonde stations.'

The last sentences in the same paragraph (end of page 18, top of page 19) are intended to quantify and stress the poor coverage over oceans even for climatic purposes. This is in line with the concern with spatial resolution. To highlight this point, we made slight but significant changes:

P18, L32: "On a scale suitable for climate monitoring, the WMO recommends" → 'Nevertheless, on a scale suitable for climate monitoring, the WMO recommends'

P19, L2: "indicates larger distances in most oceans" → 'indicates that fixed stations are too far apart in most oceans'

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

Deleted.

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

Absolutely. Corrected.

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'.

Changed to 'summarizes'

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.

Indeed, the fastest decrease seems to be in period 1945–50, although the curves of Fig. 7 are very noisy in subsequent years until 1960.

P19, L16 - Clause changed as follows:

"the average size of missing days decreased from 4–6 months to about 1 month between 1945 and 1960"

→ 'the average size of missing days dropped from 4–6 months to about 1 month between 1945 and 1960; much of this change occurred before 1950, indicating that radiosonde measurements became rapidly regular in the early years'

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

Converted to percentages. See next response for other changes in the same paragraph.

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.

IGRA et al. (2006) used observed pressure, while our results (P20) uses computed height. (Incidentally, we have wrongly given the approx. height of 250 hPa instead of 100 hPa.) So that readers can clearly understand the terms being compared, we have rephrased and shortened lines 10-15 as follows:

‘According to Fig. 8b, by 2003, 75% of the temperature soundings with surface data reached an altitude of 22 km, i.e. ~ 50 hPa. Note that Durre et al. (2006) reported that, by the same year, 74% of all IGRA 1 soundings reached at least the 100-hPa level, i.e. ~ 16 km; this lower height is due to the inclusion of PIBAL data in their analysis of IGRA 1, whereas our analysis is restricted to RAOB in IGRA 2.’

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.

The text was clarified as explained in the previous response.

Page 20 line 23: “This last feature and is coincident “ remove the word ‘and’

Typo corrected.

Page 20 line 30 - use percentage not fraction.

In this context, we think that “¾” (opposed to “half” in the same sentence) is a nice way to mean about 75% without needing to give an exact percentage. To be consistent, we prefer ‘three quarters’

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?

The words “fairly recent” refer to 2003, when the international Workshop to Improve the Usefulness of Operational Radiosonde Data (involving 30 data users and providers) took place, with conclusions reported in Durre et al. (2005). To our knowledge, things have not changed significantly in the meantime, nor the problem has been specifically addressed in a manner as to clarify the real value of humidity measurements close and above the tropopause – despite more evidences from observational studies using parallel soundings with standard sensors (like the frost point hygrometer) as reference.

The words “and certainly more above“ were added by us (as an inference) and can be deleted. To be a bit more accurate in the citation, we have rephrased the sentence:

“; however, the accuracy of fairly recent measurements in the upper-troposphere, and certainly more above too, was considered inadequate for climate studies and a challenge for future operational radiosondes (Durre et al., 2005).”

→ ‘. However, only 15 years ago, international experts pointed out that the accuracy of current operational measurements of humidity in the upper-troposphere was inadequate for addressing climate variability and change (despite the usefulness of some sensors) and a challenge for future operational radiosondes (Durre et al., 2005)’

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?

A very good suggestion to improve the dataset in the next updated version, by replacing missing values for TOPZ (max. height of hum. measuring level, rounded to decametres) where appropriate. This will only affect a tiny percentage of the metadata for individual observations, not the metadata by year which uses pressure instead of height. However, assuming that baseline for our calculations require caution when a ship-based vertical sounding has no other upper-air humidity except at the level of the balloon release or very close it.

We made the following corrections (please see different contexts):

P14, L29:

“For moving stations, the elevation of station is taken equal to the geopotential height at the surface level, if given (otherwise, the vertical extent was not calculated)”

→

‘For mobile stations (ships and buoys), the elevation of the stations can be approximated to zero, unless the vertical extent of the sounding is too small, requiring data for the balloon release height.’

P20, L9:

“Also, the soundings from mobile stations with missing geopotential height at the surface level had to be excluded.”

→

‘In addition, the soundings from mobile stations with missing geopotential height data at the surface level are excluded, for consistency with the hypsometric calculations used in the dataset presented in this paper.’

P21, L14:

“The same applies to soundings from mobile stations, since the station’s elevation is variable, and the geopotential height of the surface level is missing in 30 % of the corresponding RAOB”

→

‘The same applies to soundings from mobile stations if geopotential height is not given at the surface level (it is missing in 30 % of the corresponding RAOB). Although the local sea level is normally within ± 10 m from mean sea level, taking zero as the baseline height can lead to large relative errors if humidity is only measured very close to the radiosonde station elevation (balloon release height)’

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.

P represents the percentage of Sfc-to-500hPa humidity soundings, out of all soundings with humidity in each year from an arbitrary station (P is defined on the top labels of both Figure 9 and Figure 10). Note that Sfc-to-500hPa humidity soundings are defined in two alternate ways just a few lines above (see end of page 21). The lines in question refer to the strict definition 'HUM-A'.

What the text means is that for a very high standard such as expecting to have more than 95% HUM-A soundings, the number of stations meeting that standard (shown in Fig. 9a) remains much smaller than the total number of radiosonde stations performing observations (shown in Fig. 1a) from 1945 to as recently as 1990.

<u>To improve clarity, we made the following amendment</u>:

P22, L3-6

"Figure 9a shows the evolution of the number of stations with the percentage of Hum-A soundings exceeding given values, out of all soundings with humidity in each year – since nearly the time radiosonde humidity data at the surface level are first available. Comparing Fig. 1a with Fig. 9a, we can see that between 1945 and 1991 only a small fraction of the RS stations carried out Hum-A observations in most of the soundings, say, within a percentage range P > 95 %."

→

'Let P be the percentage of Sfc-to-500hPa soundings (HUM-A or HUM-B, at our choice), out of all humidity soundings from an arbitrary station in a given year. Figure 9a shows the evolution of the number of stations with P for Hum-A soundings exceeding given values, since nearly the time radiosonde humidity data at the surface level are first available. Comparing Fig. 9a with Fig. 1a, we can see that between 1945 and 1991 only a small fraction of the radiosonde stations carried out Hum-A observations in most of the soundings, say, with P within the range P > 95 %.'

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.

<u>Changed to</u> 'two noteworthy change points: a sudden increase around 1970 and another one around 2000'

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

<u>Removed</u>.

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

Correct, curves in Fig. 11 are not all monotonic. Word removed.

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’?

Given the changes related to the next comment, the word is no longer needed.

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

We agree that the style of the first lines of the concluding section make a good abstract. The Abstract cannot be abridged too much since it includes other important elements: what do we  mean by completeness of humidity observations; what is the content of the dataset presented by the paper; and how it may be useful to the scientific community.

Upon careful consideration, we hope to have reached this time a compromise between concision and comprehensiveness. The abstract is much shorter now.

Please see the changes to the Abstract in the marked-up manuscript, as well as the related changes in section 6 (appearing there on p. 28).

Again, arctic vs Arctic. Please check and correct!

Referring to page 26, we now have only ‘Arctic’. This first bullet paragraph was revised to fit the revision of section 3.1

Page 26 line 9 - use percentage not fraction

Maybe ‘three quarters’ is better, since we have “at present” while 13 km refers to 2016.

Page 26 line 22 “was not standard and it was rarely used” delete word ‘it’

Delected.

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.

We mean what we wrote, perhaps not well written. In other words, and giving more emphasis to the number of years of past data:

P26,L23 – Changed to:

‘The amount of humidity time-series with a given number of consecutive years of data until a given year depends not only on the year and the length of record, but also on 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.

Corrected.

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

Deleted.

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

Changed to colour-blind friendly solid lines.

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

See previous response to comment on page 17, section 3.1.

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.

The figure was reshaped two show number of stations is equal-area latitude bands in each hemisphere. "climate bands" changed to 'latitude bands'. Note that the more usual latitude bands used in Figures 6 and 7 do not show hemispheric differences since they are irrelevant for the case. In Figure 4, hemispheric differences are very important.

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

For a figure representing the station density in different latitudes, we think that dividing the globe in major latitudes with climatic significance – chiefly defined by the Tropic and Polar Circles – will be more informative to many readers. The interval 23-5$^{o}$–35$^{o}$ is a very common approximation for the subtropical regions.

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?

Considering the average values in Figure 6, we think that the differences between the tropics and extratropics are large enough to be highlighted. Making a 'composite' of the AVG and STD curves (AVG ± STD) would make the graph confusing; besides, AVG + STD sometimes exceeds a little 100%. This is not unusual in a descriptive statistic. In the case, a radiosonde station cannot have data in more than 100% of the days of the year; but the STD represents both the dispersion of the values above and below the mean.

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

Color added. See text correction in our response above.

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?

Panel a is important to show how vertical resolution of RAOB has evolved over time. The almost coincidence between the curves for T and RH demonstrates that temperature measurements alone, with missing humidity at any levels up to the maximum height of humidity measurements, are exceptional. We prefer showing it rather than just stating it.

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?

Figure 9 uses absolute number of stations and gives a detailed temporal description. Figure 10 uses percentages of stations and a course temporal description. Although we give more emphasis to Figure 9, readers familiar with (inverse) cumulative distribution functions will probably prefer to see Figure 10. The evolution with time is an important feature also in Figure 10. Grouping lines as suggest would be less informative, and, yes, panels a and b will still be different.

Figures 9 and 10 changed to color.

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

Changed as suggested.

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?

Please see above responses to comments referring to p. 11, line 12 and p. 17, line 11.

The negligible difference noted in the last comment shows that our selection of IGRA stations (i.e., IGRA-RS) retains practically all humidity data; in other words, the IGRA stations left out are irrelevant to our study. This point is elucidated in Sect. 2.2,namely on P11, L3-13. The percentage (99.999%) of RAOB present in IGRA-RS is now reported.

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

# Completeness of radiosonde humidity observations based on the IGRA

António P. Ferreira, Raquel Nieto, Luis Gimeno

Environmental Physics Laboratory, Faculty of Science, University of Vigo, Ourense, 32004, Spain

*Correspondence to*: António P. Ferreira (ap.ferreira@uvigo.es)

**Abstract.** Radiosonde measurements from the 1930s to present give unique information on the distribution and variability of water vapor in the troposphere. The data ~~compiled in~~of the Integrated Global Radiosonde Archive (IGRA) Version 2 ~~(released by the NOAA's National Centers for Environmental Information)~~are examined here until the end of 2016, aiming to describe the completeness of ~~radiosondes~~ humidity observations ~~– i.e., of~~ (simultaneous measurements of pressure, temperature and ~~either relative~~ humidity ~~or dewpoint depression –~~) in different times and locations. ~~The IGRA stations reporting radiosonde data in at least 5 % of the annual soundings for at least one year are evaluated according to specified completeness parameters for every year in their period of record. The selection of source data essentially removes pilot-balloon sites, retaining a set of 1723 stations (designated 'IGRA-RS')~~ Upon finding the stations with a non-negligible amount of radiosonde observations in their period of record, thus removing pilot-balloon stations from IGRA, the selected set (designated IGRA-RS) comprises 1723 stations, including 1300 WMO ~~upper-air~~ stations, of which 178 belong to the current GCOS Upper-Air Network (GUAN) and 16 to the GCOS Reference Upper-Air Network (GRUAN). Completeness of humidity observations for a radiosonde station and a full year is herein defined by five basic parameters: number of humidity soundings; fraction of days ~~having~~ with humidity data; ~~mean~~ average vertical resolution ~~of humidity data~~; ~~mean~~ average atmospheric pressure ~~or~~ and altitude at the highest measuring level; and maximum number of consecutive days without ~~humidity~~ data. The ~~completeness of the~~ observations eligible for calculating precipitable water vapor – i.e., having adequate vertical sampling between the surface and 500 hPa – ~~is~~ are particularly studied. ~~Individual soundings are described by the (vertically averaged) vertical resolution and the pressure level and altitude of the top of humidity measurements. For illustration, t~~The present study presents the global coverage of humidity data and an ~~global~~ overall picture of the temporal and vertical completeness parameters ~~of radiosonde humidity observations~~ over ~~the years~~time~~, including their latitudinal coverage~~. This overview indicates that the number of radiosonde stations ~~having a record length long enough to be~~ potentially useful for climate studies involving humidity~~-related quantities~~ depends not only on ~~the temporal range (multidecadal)~~their record length, but also on the continuity, regularity and vertical sampling of the humidity time-series. ~~More generally, it is hoped that the derived metadata will help climate and environmental scientists to find the most appropriate radiosonde data for humidity studies by selecting upper-air stations, observing years or individual soundings~~ Furthermore, a dataset is provided with the purpose of helping climate and environmental scientists to select radiosonde data according to various completeness criteria – even if differences in instrumentation and observing practices require extra attention. ~~A~~ This dataset ~~is presented for that purpose, consisting~~ consists of two main sub-sets: 1)

~~humidity~~statistical metadata for each ~~of the~~ IGRA-RS station~~s~~ and year within ~~the~~ period of record ~~(yearly metadata)~~; and 2) ~~humidity~~ metadata for individual observations from ~~the same~~each station~~s (ascent metadata)~~. These are complemented by 3) a list of the stations represented in the whole dataset, along with the observing periods for humidity (relative humidity or dew-point depression) and the corresponding counts of observations. The dataset is to be updated on a two-year basis, starting in 2019, and is available at https://doi.org/10.5281/zenodo.1332686.

## 1 Introduction

For more than three-quarters of a century, the global radiosonde network designed and developed for weather forecasting has provided in situ observations of humidity from the surface up the middle troposphere, eventually reaching the stratosphere. Satellite-based remote sensing of atmospheric water vapor is part of modern weather forecasting and climate monitoring (Kley et al., 2000; Andersson et al., 2007). In the present state-of-the-art, some satellite retrievals of moisture-related quantities are used as a reference to compare humidity measurements from different radiosonde types, aiming to monitor radiosonde stations and improve satellite calibration (Kuo et al., 2005; John and Buehler, 2005; Sun et al., 2010; Moradi et al. 2013). However, limb-sounding satellite techniques with high vertical resolution (albeit very course in the horizontal), using GPS radio-occultation, are a recent acquisition, of main interest to access water vapor in the upper troposphere and lower stratosphere so far (Kishore et al., 2010; Shangguan et al., 2016; Rieckh et al., 2018; Vergados et al., 2018). Reanalysis outputs based on past radiosonde data, assimilating also satellite data when available, offer multiple-level, global gridded, synoptic-scale moisture fields up to 4 times daily from a beginning year (e.g.: 1948 in NCEP/NCAR Reanalysis 1; 1979 in NCEP/NCAR Reanalysis 2 and ECWMF's ERA Interim) to present time – even though radiosonde observations are scarce over the ocean, unevenly spaced over land, and taken normally twice a day, with significant differences in vertical coverage. Naturally, since air moisture is highly variable in time and space, humidity data from different reanalysis models show discrepancies and can differ significantly from the collocated radiosonde data (e.g., Noh et al., 2016). Therefore, the radiosonde archives represent the primary source of information on the short and long-term distribution of moisture in the troposphere, despite various data inhomogeneities. Namely, geographical-temporal sampling differences (Wallis, 1998), uncertainties related to observation time and balloon drift (Kitchen, 1989; McGrath et al., 2006; Seidel et al., 2011; Laroche and Sarrazin, 2013), differences in vertical coverage and data gaps related to reporting practices of humidity (Dai et al. 2011 and references therein) and differences in humidity data accuracy – which depend on humidity sensors and vary with measured conditions (WMO, 1995; Nash, 2002; Sappuci et al., 2005; Moradi et al., 2013; Dirksen et al., 2014).

The growth interest in climate change motivated a renewed attention to radiosonde data since the 1990s. Soon it was realized that the usefulness of radiosonde data archives to investigate climate trends relies on homogenization procedures to overcome biases and sudden shifts arising from instrument changes, reporting practices and sampling differences (Elliott and Gaffen, 1991; Schwartz and Doswell, 1991; Parker and Cox, 1995; Luers and Eskridge, 1998; Lanzante et al. 2003).

Subsequent climate studies based on radiosondes have mostly focused on the detection of climate change in temperature trends (Free and Seidel, 2005; Thorne et al., 2005; Haimberger et al. 2008). Concerning humidity, radiosonde-based climatic studies are for now confined to the lower and middle troposphere, because of the large uncertainty of measurements and biases in the upper troposphere and lower stratosphere (Elliot and Gaffen, 1991; Soden and Lanzante, 1996; Wang et al., 2003) and the extremely large relative biases and insufficient data in the lower stratosphere (Miloshevich et al., 2006; Nash et al. 2011). Radiosonde data have been used for studying the long-term trends and the annual cycle of several humidity parameters (precipitable water vapor, specific humidity and relative humidity), at least in well-sampled regions of the globe and if data inhomogeneities are removed (Elliot et al., 1991; Gaffen et al., 1992; Ross and Elliott, 1996; Ross and Elliot, 2001; McCarthy el al., 2009; Durre et al., 2009; Dai et al., 2011). On a rather different scale, radiosonde measurements with high vertical resolution near the ground are suitable for studying water evaporation over land and the structure of the planetary boundary layer, provided that the time-lag of humidity sensors as they move through a rapidly changing environment is accounted for (Sugita and Brutsaert, 1991; Connell and Miller, 1995; Seidel et al., 2010).

Since its creation in 2004, the Integrated Global Radiosonde Archive (IGRA) was meant to be the largest data set of up-to-date weather-balloon observations freely available, by collecting quality-controlled data from upper-air stations across all continents. The first version of IGRA (a successor of the Comprehensive Aerological Reference Data Set (CARDS; Eskridge et al., 1995) contained practically data after 1945 (Durre et al., 2006). The IGRA Version 2 used in this paper, released by the NOAA’s National Centers for Environmental Information (Durre et al., 2016) and ~~very~~ recently described in Durre et al. (2018), has enhanced data coverage and extends back in time as early as 1905, although (for historical reasons) humidity data begin in 1930 with a sole location in Europe. The extension to observations prior to 1946 resulted mainly from the addition of data from the Comprehensive Historical Upper-Air Network (CHUAN), which is the most important collection of upper-air observations taken before 1958 (Stickler et al., 2009). In view of the huge amount of data collected in IGRA (which is a combination of radiosonde and pilot-balloon observations) and the differences in the observing period, temporal regularity and continuity, vertical resolution and vertical extension of humidity data among different stations, finding the most suitable humidity-reporting stations (or humidity soundings from different stations) for a specific purpose can be difficult to put into practice.

Radiosonde humidity measurements involve the simultaneous measurements of pressure, temperature and relative humidity or dew-point depression. Therefore, except for horizontal wind, which is indirectly measured with the aid of a remote tracking device, ~~the ‘~~humidity~~observations’~~ represents the most accomplished of the radiosonde observations. The purpose of this paper is to study the completeness of humidity observations collected in IGRA according to various needs – number and latitudinal distribution of observing stations, fraction of observing days in a year, resolution and range of vertical levels, length and continuity of the time-series, minimal sampling between the surface and the 500-hPa level – aiming to facilitate the use of radiosonde humidity data by atmospheric and environmental scientists. The task is two-fold: first, to elucidate the completeness of the humidity observations from IGRA for each year in global terms, including the latitudinal coverage of stations and the length of regular time-series; second, to provide metadata describing the completeness of humidity observations from each station. The observing periods without missing years in humidity data must be clarified. Latitudinal and regional differences should be easily derived from the geographic coordinates of stations.

The remainder of this section is intended to clarify the term “completeness of observations” concerning the use of radiosonde data and to present an historical account of the main factors that limit the completeness of humidity observations from radiosondes: vertical levels available in radiosonde reports; missing observations associated with humidity sensor limitations. The next sections are organized as follows. Section 2 indicates the IGRA data set used in the study; selects the IGRA stations reporting a minimum of radiosonde data (coined as ‘IGRA-RS’) by discharging stations with practically wind-only data in their period of record; and explains the data analysis. Section 3 presents a global picture of the completeness of humidity observations over the years, as derived from the IGRA-RS stations~~ reporting a minimum of radiosonde data, i.e., discharging stations with practically wind-only data in their period of record~~. Section 4 provides the definition of the metadata parameters describing the completeness of humidity observations from each ~~relevant~~ IGRA-RS station – ~~both in terms of~~ either as annual statistics ~~and~~ or ~~in~~for individual soundings – and the format description of the corresponding data sets. The availability of the resulting dataset supplied by the study (Ferreira et al., 2018) is reported in Sect. 5. A summary of results and some suggestions for future application are given in Sect. 6.

### 1.1 Completeness of observations for radiosonde humidity studies

Data completeness in a data set refers to the extent to which the data set collects the expected elements: quality-assured data are not left-out; missing or invalid values are properly indicated. This is a basic requirement for data quality, and it is assured in IGRA. In a different way, and uncommon, data completeness may refer to *whether the required data for a specific purpose are available or not.* That meaning is not new in the field of meteorology. For instance, the WMO recommendations on “data completeness” (sic) required for calculating monthly means and climate normals from meteorological surface data refer to the temporal continuity and regularity of observations for different climate elements (WMO, 1989). Broadly, Bellamy (1970) discussed the acceptability of meteorological observations in terms of their degree of completeness, considering that the goal of meteorological observations is to “depict the space-time distributions of everything-atmospheric everywhere always, ever more completely in ever-increasing detail”; appropriately, he used the expression *completeness of observations*. This is the terminology used in the present paper.

Concerning the completeness of radiosonde humidity observations, the vertical coverage and vertical resolution of sounding data is of first concern, chiefly between the surface and the middle troposphere (~500 hPa) regarding the precipitable water vapor content; furthermore, the period of record and the regularity and continuity of radiosonde data is a relevant issue for long-term monitoring of the climate system (Karl et al., 1995), as exemplified by temporal sampling requirements used in trend and seasonal analysis of temperature, humidity and integrated water vapor (Gaffen et al., 1991; Gaffen and Elliot, 1992; Karl et al., 1995; Ross and Elliot, 1996; Zhai and Eskridge 1997, Lanzante et al. 2003; McCarthy et al., 2009). Although the vertical and temporal completeness of station-based humidity time-series can be treated separately from the ~~issue of~~ geographical coverage of stations, studying the completeness of observations in a global, historical data set of radiosonde observations should address ~~that~~ both issues ~~too~~simultaneously. This is particularly true concerning the subsampling of radiosonde stations for studies of atmospheric temperature or water vapor trends on a regional or global scale (Wallis, 1997).

Several factors contribute to differences in the completeness of humidity observations among radiosonde stations and individual soundings: i) The geographical coverage of radiosonde stations evolved over time, and so, the period of usage varies among stations; ii) A lack of equipment maintenance may result in interruptions of observations; iii) The number of vertical levels and vertical extent in radiosonde reports depend on the standard pressure levels in use, as well as on the reported significant levels (assuming that the balloon bursts at the proper altitude); iv) Missing humidity observations arise from difficulties associated with the performance of humidity sensors and the observing practices related to their working range. While (i) and (ii) are of a random nature, points (iii) and (iv) deserve an explanation because of historical changes with implications in the vertical coverage and resolution of radiosonde humidity profiles.

### 1.2 Vertical levels in radiosonde observations

In radiosonde soundings, temperature, relative humidity (and/or dewpoint depression) and wind speed and direction are measured together with atmospheric pressure, while geopotential height is indirectly measured from hypsometric calculations[1] (but may be missing in radiosonde reports). As a common practice, only standard pressure levels and significant levels are stored and reported. Currently, the *standard levels* are 1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, and 10 hPa (WMO, 1996). But historical changes deserve due attention. An inspection of the earliest soundings collected in IGRA – made in 1905 at Lindenberg, Germany, a quarter of a century before radiosondes were available – reveals temperature data reaching sometimes 100 hPa, with the reported levels being 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150 and 100 hPa (although most of those soundings did not reach beyond 700 hPa). Radiosonde humidity measurements at the same station, as collected in IGRA, began in 1950. Nevertheless, the 150- and 100-hPa levels were first recommended by the WMO in 1953, while the levels 70, 50, 30, 20, and 10 hPa were proposed in 1957, the International Geophysical Year. Even so, the levels above 200 hPa were still referred as non-standard by the WMO in 1958, until the 100-hPa level was finally adopted that year (WMO, 1957; WMO 1958). In the years that followed, the pressure levels ≤ 150 hPa (representing roughly the stratosphere) became common worldwide. As to the lower levels ≥ 200 hPa (representing roughly the troposphere), they were in general use since the early 1940s, with two exceptions: first, the 250-hPa level was only adopted in 1970, to satisfy aviation demands (WMO, 1970); second, the 925-hPa (within the planetary-boundary layer above low-altitude stations), although planned since 1977, was first required in WMO Antarctic stations in 1987, given the low surface pressure over the Antarctic Plateau, until it was adopted worldwide by the end of 1991 (WMO, 1977, p. 15; WMO, 1987, pp. 57-58; Oakley, 1993, p. 23).

[1] Except in some Soviet/Russian radiosonde-radar systems and the last generation of GPS radiosondes – in which pressure is deduced from the (radar or GPS, respectively) profile of geometric height and the radiosonde profiles of temperature and humidity (Zaitseva, 1993; Nash et al., 2011).

Note, however, that these two levels were in use to in some stations before international agreement, and for a long time as exemplified by the Lindenberg station. Besides the standard levels, some intermediate fixed levels within the troposphere (e.g., 800, 750, 650, 600 hPa) are regularly used in some stations (Shea et al., 1994). The Lindenberg station also indicates the early use of 600 hPa. The additional (high-stratospheric) levels 7, 5, 3, 2 and 1 hPa have been used in agreement with WMO recommendations (WMO, 1970), depending on regional and national practices. E.g., they form part of 'upper-level' observations in the U.S. National Weather Service (OFCM, 1997).

The number of s*ignificant levels* – non-standard levels needed to reproduce the vertical temperature and dew-point temperature profiles, capturing turning points or abrupt changes (such as thermal inversions and the tropopause) – depend on atmospheric conditions, manual rules and, before automation, on the observers' skills. By the late 1950s, the rules for choosing significative levels were still under discussion (WMO, 1957), being established over time by WMO regulations (WMO, 1988). Interestingly, the almost linear increase in the average number of non-standard levels in weather-balloon sounding reports (radiosondes + pilot balloons) from about zero by 1945 to about 30 by 2000 – as revealed from IGRA v1, inferring from Fig. 7 in Durre et al. (2006) –, can hardly be attributed to an increased attention to significant levels alone. It suggests that a significant number of stations have reported additional levels apart from the standard and significant levels (both "mandatory" in WMO's nomenclature).

The *surface level*, which is treated separately in upper-air sounding reports, was reported in most of the radiosonde stations since the mid-1940s. However, it has been reported systematically only since around 2000 (as shown later in Sect. 2.3).

The current migration of ~~RAOB~~ radiosonde reports from alphanumeric (TEMP) to the binary universal form for the representation of meteorological data (BUFR), together with the conversion of radiosondes to generate native BUFR messages, allows the transmission of high-resolution data (2 to 10 s sampling rate, i.e., ~ 5 to 50 m resolution in a typical balloon ascent) along with the balloon drift position, the observation time for each level and other metadata (Ingleby et al., 2016). Currently, 20% of the radiosonde stations send high-resolution BUFR reports through the Global Telecommunication System (GTS), many coming from Europe; however, such data are not yet available in an open archive.

### 1.3 Missing humidity observations

Combining adequate spatial and temporal resolution with enough accuracy for synoptic use, modern radiosonde measurements reach the upper troposphere and lower stratosphere, much beyond the layers where most of the atmospheric water vapor resides. That ~~was no~~has not always been so. While the vertical sampling of temperature soundings is limited by the burst altitude and the mandatory levels (standard and significant), the maximum height and the vertical resolution of humidity soundings are further restricted by sensor limitations. Upper-air humidity measurements began in the 1930s but became substantial only in the 1940s. Despite radiosonde hygrometers (measuring relative humidity (RH)) have improved over time, humidity has been always difficult to measure in very cold or dry air due to the poor response of many instruments at very small vapor concentrations (by lowering saturation vapor pressure, cold temperatures are associated with low water vapor pressures). As it was once pointed out, “humidity measurements in the free atmosphere are probably the least satisfactory of the regular aerological observations” (Hawson, 1970). Balloon-based chilled mirror hygrometers, designed to measure water-vapor mixing ratios in the stratosphere (an extremely cold and dry environment), has been used for more than half a century but are exclusive to scientific research or comparison with humidity measurements from operational radiosondes (Mastenbrook and Daniels, 1980; Vömel et al. 2007; Hurst et al., 2011; Hall et al., 2016). Since a long time ago, whether services need to rely on meteorological radiosondes consisting of~~meteorological radiosondes are~~ expandable balloons carrying relatively low-cost and light instrument packages (Brettle and Galvin, 2003).

Here is a brief review of the main humidity sensor types and their limitations, since the time when registering balloons ~~sondes~~ were abandoned by national weather services and electric hygrometers began to be incorporated in radiosondes (circa 1940; DuBois (2002)). The lithium chloride humidity sensors, which were widely used in radiosondes between the mid-1940s and the mid1960s, did not respond to temperatures below around −40º C. From the early 1960s onwards, the new carbon hygristor allowed measurements at lower temperatures – down to −65°C in the early 1990s, however with a time lag in the sensor´s response as large as 10 minutes (Garand et al., 1992). In practice, humidity measurements at temperatures below −40°C w~~h~~ere discontinued in many countries before the 1990s, limiting the vertical extent of routine humidity observations to about 400 hPa (≈ 7 km altitude) (Gutnick, 1962; Gaffen, 1993). Besides, the radiosondes using lithium chloride hygrometers suffered from a low-frequency limitation in the transmission of RH less than 15–20 %, known as motorboating (Wade, 1994). The radiosondes using the carbon hygristor enabled, in principle, measurements in that low RH region – however, the accuracy and reproducibility of low-RH values was little known and suspected to be poor for many years, giving the wrong impression that relative humidity lower than about 20 % did not occur in the lower troposphere (Wade, 1994; Nash, 2015). Therefore, values of RH below 20 % were usually cut off in humidity reports; in the radiosonde network of the U.S.A. this happened between 1973 and 1992 (Elliott and Gaffen, 1991). Note that changes in instrument and reporting practices in different countries took place at different times~~. –~~: the threshold value of RH varied in the range 10–20%; the lowest temperature of −40°C for reporting humidity, and the shift to lower temperatures, was applied in different periods depending on country; humidity could be reported up to a specified pressure level. Moreover, mechanical sensors are not exclusive to pre-1940s radiosondes: hair hygrometers were only abandoned in the mid-1950s and rolled hair hygrometers were used in a few places until about 1980; the goldbeater's skin sensors introduced in the 1950s became particularly important in the Soviet Union. [For historical details on these changes, see Gaffen (1993).]~~E.g.: hair hygrometers were only abandoned in the mid-1950s; rolled hair hygrometers were used in a few places until about 1980; the threshold value of RH varied in the range 10–20 %; the lowest temperature of −40ºC for reporting humidity, and the shift to lower temperatures, was also applied in different periods; humidity could be reported up to a specified pressure level (Gaffen, 1993).~~ The new~~ly~~ capacitive thin-film sensors introduced in 1981, with ~~different versions emerging in the 2000s to avoid contamination from rain and thick clouds (using a protective cap, or rather using two sensors alternately heated and used for measurement)~~the RS80 radiosonde, have improved the response time at low temperatures and the capability of measuring very low humidity. Two important enhancements occurred in the late 1990s. First, the  protection from chemical contamination arising from outgassing of RS80 radiosonde packages, thus making the dielectric polymer more selective to water vapor molecules and reducing dry bias (Wang et al., 2002). Second, the dual sensors introduced in the RS90 radiosonde, in which two sensors were alternately heated to remove condensation from the measuring sensor, thus preventing wet biases after measurements in saturated conditions. In the RS92 radiosonde, in use since 2004, the lowest temperature of the heating cycle extended down from – 40 $^0$C to – 60 $^0$C. The smaller size of and the better ventilation of the RS90 and RS92 sensors compared to RS80 improved the response time. However, RS80 sondes were less affected by dry biases in daytime measurements because of the protective rain cap which also prevented direct sunlight (Smit et al., 2013). However, RH reports at temperatures lower than –40 °C did not develop significantly until about 2000; in recent years, for temperatures of –50 °C to –70 ° C only the newest humidity sensors respond quickly enough to make useful measurements; moreover, the best ones had an uncertainty of around 16 % RH at temperatures as low –70 ºC (occurring over Antarctica and around the tropical tropopause), which is barely acceptable for numerical weather prevision but not suitable for climatic studies (Nash, 2015). Improvements over time were not restricted to sensor type but also to data reduction and calibration. E.g.: measurements from the carbon hygristor in VIZ radiosondes were improved in the 1990s by correcting the low-humidity algorithm; some modern radiosonde systems apply corrections for slow time constant of response and for daytime heating of the humidity sensor; calibration at low temperatures was perfectioned (Dirksen et al., (2014) and references therein). While radiosonde humidity measurements are now generally reliable in the troposphere, uncertainties remain concerning the upper-stratosphere, with temperatures below –50ºC, in addition to dry conditions found above the lower troposphere and wet conditions that occur in thick clouds (Miloshevich et al., 2006). Although the capacitive thin-film sensors have been widespread (with Vaisala radiosondes RS80 and RS92), two older sensor types continued in use for many years: the carbon hygristor (in VIZ/Sippican radiosondes, currently in disuse, and in the GTS1 radiosonde, in use in China) and the goldbeater's skin sensor ~~introduced in 1950s and~~ used in some radiosonde types made in Russia and China until a few years ago; this peculiar sensor responded too slowly to be useful at temperatures lower than –20 °C and suffered from hysteresis following exposure to low humidity (Nash, 2015; Moradi et al., 2013). For the current radiosonde types, see Ingleby (2017).

The trouble in measuring upper-air humidity affects the completeness of observations in several ways: the vertical extent of humidity soundings varies much among radiosonde stations and over time owing to sensor limitations in very cold air; likewise, vertical gaps in low humidity regions are expected, due to cut-off of RH below sensors' measuring capability; lastly, missing days in radiosonde humidity records may originate from adverse conditions (dry days, wet days, cold days) at individual stations (Garand et al., 1992; Ross and Elliott, 1996; McCarthy et al., 2009; Dai et al., 2011). As explained above, the actual extent of missing data depends on the observing practices ~~intricated~~ combined with sensor limitations. In addition, failures in some part of the radiosonde system can compromise soundings. Faulty ground-equipment used for control checks (sensors' calibration before balloon release), data reduction and data recording or telecommunication of coded reports may cause long inoperative periods; poor signal reception from the radiosonde make sometimes data processing impossible.

Radiosonde operations in remote environments, particularly performed from ships, present their own challenges; Hartten et al. (2018) give a vivid illustration. In sum, the vertical extent, vertical resolution, temporal regularity and continuity of humidity reports are quite heterogeneous.

## 2 Input data and methods

We have examined the IGRA 2 main dataset until the end of 2016. Section 2.1 presents briefly that dataset, including the quality assurance of humidity data. Section 2.2 provides a first look on the data, to find out how many and which of the IGRA stations have a non-negligible amount of radiosonde observations (RAOB), and at the same time to give a hint of the amount of humidity data. Section 2.3 describes the data analysis, aiming to explore the completeness of humidity observations in the sense introduced in Sect. 1.

### 2.1 IGRA 2 – sounding data

The IGRA 2 consists primarily of radiosonde[2] and pilot-balloon[3] observations from over 2700 globally and temporally distributed stations, even though the coverage over oceans is limited to ships, buoys and remote islands. ~~IGRA 2 has also derived data for a selection of WMO stations, but t~~This paper concerns with ~~the~~ sounding data (Durre et al. 2016), comprising over 45 million soundings from 2761 (2662 fixed and 99 mobile) stations [based on data accessed in September 2017].

The main difference between IGRA 2 and ~~the former version~~IGRA 1 is the amount of sounding data: 33 data sources instead of the initial 11, implying about 80 % more stations; new data from hundreds of stations before 1946 and the addition of floating stations (fixed weather ships and buoys, mobile ships, and Russian ice islands); furthermore, humidity data prior~~y~~ to 1969 were added. The latter change is related to how humidity values were stored in radiosonde reports. Until 1969, humidity observations were given only as RH; from then on, RH measurements have normally been converted to dewpoint depression (DPD) and reported ~~too~~ mostly in that form. Different assumptions in the conversion code can lead to inconsistencies of data (Garand el al., 1992). The former IGRA contained only DPD, while IGRA 2 contains humidity data in either form, as available in original reports, provided they pass the following conditions:

i) Data completeness: valid temperature accompanies humidity data;
ii) Valid range: 0-100 % for RH; 0 to 70 ºC for DPD;
iii) Internal consistency: DPD-derived RH differs from reported RH by 10 % at most;

---

[2] In modern usage, the term radiosonde refers not only to the early radiosondes but also to the rawindsondes (in use since the 1950s), which, besides measuring thermodynamic parameters, provide wind information with the aid of a radio-theodolite, a radar device, a radio navigation system or, more recently, GPS (Dabberdt et al. 2002; Nash et al., 2007). Observations from either radiosonde type are often abbreviated as 'raob' in meteorological jargon.

[3] Free balloon tracked by optical theodolites or radar to measure upper-air winds (Wenstrom, 1937, Hickman, 2015). Often abbreviated as 'pibal'. The common single theodolite technique requires the approximate ascent rate to obtain position, while the double-theodolite method allows a pure trigonometric calculation. In visual tracking, rarely used today but still important where radar tracking or wind measurements from a rawindsonde are not possible, a flashlight is used during night or twilight hours.

iv) Plausibility: (derived water-vapor pressure) ≤ 0.1 × (atmospheric pressure).

Quality checks (i)–(ii), save for the RH range, are integral to IGRA from its creation (Durre et al., 2006); (ii) for RH (with the later introduction of this variable in the archive) and (iii)–(iv) were added in IGRA 2 (Durre, 2016; Durre et al., 2018). Note that a RAOB message must have at least temperature data at several pressure levels, while humidity or wind data may be missing, and geopotential height is not always given. The recording of pressure levels, and consistence between pressure and geopotential height whenever the latter is reported in source data, has been assured in IGRA since its first version. IGRA uses a consistent data format, irrespective of the provenience of the data (PIBAL or RAOB). Therefore, RAOBs in IGRA can be simply identified by the presence of temperature data. Wind observations from pilot-balloons (PIBAL) have only wind data at several geopotential heights (adjusted from geometrical height measurements and the gravitational field).~~In radiosonde data, consistence between pressure and geopotential height (whenever the latter is reported in source data) is assured in IGRA since its first version. In wind data from pilot-balloons, the vertical coordinate is geopotential height (presumably adjusted from geometrical height measurements and the gravitational field)~~. ~~Note that~~Concerning humidity data, the precision and accuracy of RH and DPD data vary substantially as a function of RH and temperature, degrading in dry or cold conditions to a greater or lesser extent depending on the radiosonde type (for a review on the subject, see Smit et al. (2013)). The information about instrument changes (stations' history), whenever available, is provided in a separate metadata file in IGRA 2 [update of the metadata given in the first version of IGRA, which were mostly taken from Gaffen (1996)].

The most frequent nominal observation times are 0300 and 1500 UT until 1957 and 0000 and 1200 UT afterwards, which reflects the shift in observing time that occurred in 1957 in major WMO radiosonde networks. In the beginning of 1958 the primary standard hours of WMO upper-air observations were already 0000 and 1200 GMT (WMO, 1958). However, in some countries, different synoptic hours were practiced over the years; sometimes stations have performed up to four soundings per day for certain demands (see Gaffen, 1993).

For a description of data coverage and data sources of IGRA 2, a full description of quality assurance of data, and further detail on the differences between IGRA 1 and IGRA 2, the reader should see Durre et al. (2018) published after the initial draft of this paper was written. NB – Hereafter, IGRA 2 is simply referred to as IGRA, unless stated otherwise.

## 2.2 Identification of radiosonde stations

The examination of the IGRA reveals that 958 stations have wind-only observations in their full period of record, i.e., 34.7 % of the stations represented in the entire archive. These stations form part of the global pilot-balloon (PIBAL) station network, which evolved over time. As to the rest of the stations, some of them changed from a PIBAL launching site to a radiosonde launching site at some point in their period of record, – meaning that they are not, strictly, PIBAL stations nor radiosonde stations. Obviously, the number of PIBAL stations or radiosonde stations at a certain time depends on the stations opened and closed before that time, of either type. In the following, terms like "observations", "soundings" or "reports" refer to the upper-air stations and balloon data of IGRA, which retains most of the source data.

Figure 1a shows the yearly number of ~~the *RAOB stations*~~stations reporting RAOB any time of the year – ~~i.e., the stations with radiosonde~~meaning they have at least observations ~~at any time of the year, at least~~ of temperature, regardless of simultaneous humidity/wind observations – and of stations reporting PIBAL observations alone~~the *PIBAL stations*~~ – i.e., ~~the stations~~ reporting ~~measuring~~ only wind throughout the year ~~(assuming data are not lost)~~. For comparison, the number of stations reporting any humidity data – indicating that radiosondes are equipped with a hygrometer – and with humidity observations in more than 95 % of the radio soundings, is also shown. Constituting the bulk of the IGRA stations until the early 1940s, the ~~PIBAL~~ stations with only PIBAL data represent nowadays only 13 % of the total. The reason for the apparent discontinuity in the amount of stations performing only PIBAL ~~stations~~observations between 1972 and 1973 is this: beginning in 1973, IGRA data largely come from the GTS and include many more PIBAL data than prior data sources (Imke Durre[4], personal communication, April 12, 2018). The number of ~~RAOB stations~~stations reporting RAOB increased rapidly since the mid-1940s, staying in the range 800–900 from around 1970 to ~~virtually~~ present (2016). Note that, before the advent of the radiosonde, upper-air measurements of temperature and relative humidity were made using kites, registering balloons and aircrafts; these platforms were gradually abandoned until the radiosonde era was established in the mid-1940s (DuBois, 2002; Stickler et al., 2010). Since the first radiosonde prototypes were developed between 1929 and 1930, the early observations of temperature collected in IGRA, beginning with one station in 1905 (Lindenberg, Germany), were made by meteographs without radio-telemetry. The first upper-air humidity observations are from 1930, coming from a single station: Kjeller, Norway. According to IGRA, until 1942 there were less than 10 stations reporting humidity, with that number growing rapidly in the following years together with the total number of radiosonde stations, reflecting the widespread use of radiosondes (see '~~Higrom.~~HUM' and '~~RAOB'~~ TEMP' in Fig. 1a). The major relative change occurred between 1945 and 1946, coincident with the end of ~~the second world war~~World War II, when the global count of radiosonde stations tripl~~icat~~ed. Note that the replacement of hair hygrometers by the lithium chloride humidity element began in some radiosonde networks shortly before; in the U.S.A., apparently that change took place between 1940 and 1943 (Elliott and Gaffen, 1991; DuBois, 2002), although IGRA does not contain data prior to 1946. Fig 1a shows also that the fraction of radiosonde stations measuring humidity in more than 95 % of the soundings increased over the years, getting very close to 100% in the last decade (see relative difference between the curves '~~RAOB'~~ TEMP' and '~~Higrom.~~HUM > 95 %').

Figure 1b shows the evolution of the global, annual mean number of soundings performed per day, for the different atmospheric parameters apart from pressure: temperature, humidity and wind. Recall that pressure is always measured in RAOB soundings, while in PIBAL soundings wind is measured as a function of altitude[5]. For clarity, the PIBAL wind soundings are depicted separately from all wind measurements, which also comes from RAOB soundings since the mid-1940s. Although, as a rule, radiosonde launches are carried out twice a day, in fact there is a significant number of missing days in

[4] Center for Weather and Climate, NOAA's National Centers for Environmental Information, Ashville, NC.

[5] Pressure levels are present in the wind-only data coming from 28 IGRA stations (standard levels with missing height), indicating that an on-board pressure sensor with a radio transmitter was used along with the wind-finding system.

temperature and humidity data, i.e. days without any RAOB data: roughly 1 in 5 days during the year, on average for the years after the mid-1950s~~, as it can be~~ [as concluded by comparing the yearly number of observing stations (TEMP, HUM in Fig. 1a) ~~(RAOB)~~ with half the global number of daily observations ~~Fig. 1b~~ (TEMP, HUM in Fig. 1b)].

Aiming to study humidity completeness, the IGRA stations having a negligible amount of temperature data in every year of their period of record were excluded, because temperature is required to measure RH or DPD and so all humidity data in IGRA are accompanied by temperature data. Specifically, we have selected the *stations with RAOB soundings in 5 % or more of the annual soundings in at least one calendar year within their full period of record* until the end of 2016. These will be hereafter referred to as *IGRA-RS stations* (RS stands for radiosonde), even if some of them contribute with relatively few RAOB data. The above selection reduces the number of IGRA stations by 38 %, whereas the number of soundings is only reduced by 13 %, amounting to 39.5 million, out of which there are 30.2 million radio soundings including 29.8 million humidity soundings (see Table 1). Wind-only soundings are still present in 23.6 % of the soundings from the selected stations. Note, however, that 92.2 % of the removed stations are strictly PIBAL stations and the remaining 7.8 % have RAOB data in less than 0.6 % of the corresponding soundings, apart from two cases with a period of record shorter than one year. In sum, the IGRA-RS subset retains practically all the RAOB soundings (99.999%), particularly the humidity soundings, as shown in Table 1.

The IGRA-RS stations and their locations are listed in supplementary Table S1, along with the full periods of record (Full POR), the periods of record for humidity (Hum POR) and the corresponding numbers of humidity observations (Hum Obs), i.e., the number of individual soundings reporting either DPD or RH data. Since humidity time series can be interrupted for long periods of time~~missing years are considered~~, the full period of record of one station may be segmented into two or more periods for humidity (both are rounded to years). Table S1 comprises 1723 stations, of which 1300 are WMO stations (denoted by the letter 'M' following the 2-character country code of IGRA identifier codes). Note: in the data from around 120 land stations the early years of record for humidity (normally 2 to 3 years), contains only surface or near-surface data; this happens in about 100 stations of the former Soviet Union, mostly during the years 1946-49.

Focusing on the usefulness for climatic studies, the subset of WMO upper-air stations ~~integrating~~ integrated into the Global Climate Observing System (GCOS), i.e., the GCOS Upper-Air Network (GUAN), deserves attention. Formally established in the 1990s~~, together with the surface stations of the GCOS Surface Network~~, the GUAN is aimed to provide long-term, consistent, homogeneous and reliable observations needed to monitor the atmospheric component of the global climate system (WMO, 2002; McCarthy, 2008). At present the GUAN comprises 178 stations, all of which are represented in IGRA-RS. The IGRA-derived statistics of humidity observations from the GUAN stations for the period 2001/10/01 to 2016/12/31 is shown in supplementary Table S2, as explained next. 'Hum POR' indicates the years with any humidity data in the year, as found in IGRA, beginning at the time when each station was included in GUAN, or, at least, at the earliest time for which performance indicators for the GUAN stations are available through the NOAA/National Centers for Environmental Information website; this is the first day of the month of 'Begin Date' indicated in Table S2. '# Days' is the number of days in

Hum POR, excluding the months before Begin Date. The last three columns give the corresponding count of humidity observations around the principal nominal hours, 0000 UT and 1200 UT (± 1 h), and at any other times (0200 UT through 1000 UT and 1400 UT through 2200 UT). Stations are identified by the WMO region and WMO number, followed by the station name and country. (To find out the corresponding IGRA ID codes in Table S1 it suffices to observe that the last nine characters must be 'M000' followed by the WMO number.) Note that most of the GUAN stations have humidity data at or around 0000 UT and 1200 UT almost every day; however, the exceptions to the rule, and even gap years, are not negligible.

Moreover, the IGRA-RS contains 16 stations that form part of the GCOS Reference Upper-Air Network (GRUAN; Bodeker et al., 2016): half certified and half to be certified according to current GRUAN status, of which eight (half certified too) are also GUAN stations. Those specific GRUAN sites report default data (from radiosonde manufacturers) to the GTS; at present, most of them already send BUFR messages with high resolution (Michael Sommer[6], personal communication, September 18, 2018). The GRUAN aims to serve as reference network for climate applications, satellite validation and in support of other radiosonde networks, by providing long-term high-quality records of vertical profiles of selected essential climate variables, accompanied by traceable estimates of measurement uncertainties (WMO, 2011a; Dirksen et al., 2014). Naturally, real-time meteorological data transmitted from GRUAN sites to the GTS may differ from GRUAN internal data regarding raw data processing. For reference, the IGRA ID codes of the GRUAN sites appearing in the IGRA-RS station list are underlined on Table S1; of course, other GRUAN sites performing only research measurements are not part of IGRA. Likewise, the WMO numbers of the GUAN stations coincident with GRUAN sites are underlined on Table S2.

### 2.3 Analysis of humidity data

Overall, the analysis of data from IGRA-RS stations, selected as described in the previous section, ~~aimed~~ aims to answer the following questions:

1. What is the spatial coverage of humidity-reporting stations in different years and latitudes?
2. What is the fraction of days in a year with humidity data and the number of consecutive missing days on average?
3. What is the typical vertical resolution and vertical extent of humidity observations?
4. How many stations have enough data in the vertical to allow the estimation of precipitable water?
5. How does the temporal and vertical completeness affect the availability of long-term humidity time-series?

Each question ~~was~~is explored as detailed below in Sects. 2.3.1–2.3.5, with the results presented later in Sect. 3. The description of the related metadata parameters regarding each IGRA-RS station, is deferred to Sect. 4.

[6] GRUAN Lead Centre, Lindenberg Meteorological Observatory - Richard Aßmann Observatory, Germany.

### 2.3.1 Global coverage

The geographical distribution of the radiosonde network measuring humidity evolved over time. Its size and distribution were studied in terms of the annual number of stations with any humidity observations during the year in different climatic zones, considering only the fixed IGRA-RS stations. The spatial coverage of observations was further studied by a parameter that is closely related to the average spacing of stations, but it represents better the data coverage if stations are unevenly distributed.

The average separation between adjacent stations ($L$) over a region of the Earth's surface can be estimated by $\sqrt{A/n}$, where $n$ is the number of stations lying on a surface of area $A$, $n/A$ representing the average station density. This measure is, however, insensible to the spatial distribution of stations. The global radiosonde network has highly variable density since the observation stations are concentrated in continental regions, mostly in populated areas of developed countries. Sparse-data areas occur on oceans and seas, near the poles and in certain parts of land continents. $L$ can be alternatively defined as the mean distance between each station and its nearest neighbor; but this definition ignores data-void areas. The average distance from a point on the surface to the nearest station ($\approx L/2$ for a uniform network) is more informative because it depends on the distribution of concentrated- and sparse-data areas. Therefore, to study the global coverage of observations it is convenient to use the *average distance to the nearest station, as measured from every point over the main landmasses or ocean/sea areas within a given latitude band*. Let $s\,(x)$ be the geodesic distance from a given point $x = (\varphi, \lambda)$ of latitude $\varphi$ and longitude $\lambda$ to the position of the nearest station: $s = \min\{dist(x, x_i);\ i = 1, 2, \ldots, N\}$, where $x_i$ denotes the positions of individual stations, say $N$ in total. Averaging $s$ over a zonal band bounded by latitudes $\varphi_1$ and $\varphi_2$, under the spherical-Earth approximation,

$$\bar{s}(\varphi_1, \varphi_2) = \int_0^{2\pi} \int_{\varphi_1}^{\varphi_2} \sigma\, s \cos\varphi\, d\varphi\, d\lambda \Big/ \int_0^{2\pi} \int_{\varphi_1}^{\varphi_2} \sigma \cos\varphi\, d\varphi\, d\lambda \qquad (1)$$

where the overbar denotes area-weighted average, and $\sigma(\varphi, \lambda)$ is a mask value that can be used to restrict the calculation to mostly land or water regions by switching the values $\sigma = 0$ and $\sigma = 1$ appropriately. The following method was applied. First, the calculation for main landmasses excludes points on landmasses smaller than Ireland, since they give irrelevant information about the spacing of stations over land; however, continental archipelagos are treated as part of continents. Any regions outside the above defined main landmasses are treated as belonging to ocean/sea, excluding lakes which are included in continents. Finally, the determination of the nearest station from points on ocean/sea areas involves not only stations surrounded by sea water (stations on oceanic islands plus a few fixed weather-ships, since we focus on fixed stations) but also stations located on the coastline of continents and large islands, as well as on the shores of seas enclosed by continents. This scheme assumes that upper-air observations at such locations are partly representative of atmospheric conditions above the nearby waters, because the physical frontier between land and sea is blurred in the atmosphere (incidentally, the island and coastal *surface* stations are classified by the WMO as 'sea stations').

We have applied Eq. (1) ~~was applied~~ to the IGRA-RS fixed stations reporting humidity in specific years to examine the global coverage of upper-air humidity observations in different climatic zones over time, regardless of the temporal and vertical completeness of time series. Such information is not part of the dataset introduced in this paper, which focus precisely on the time series at each station. Nevertheless, Eq. (1), with possible adaptions for the latitude and longitude intervals, may be used to study the spatial coverage of any subset of stations selected according to a given range for the metadata parameters presented in Sect. 4.

### 2.3.2 Annual frequency and temporal continuity

The frequency of humidity observations over time ~~was~~ is studied in terms of the *fraction of humidity observing days in the year*. Although this gives a sense of the regularity of observations, it says little about the continuity of data over the year. In this respect, it is of interest to know the size of the *maximum interval of consecutive days without humidity data in a year* – denoted hereafter as *'size of missing days'*.

The above defined measures of temporal completeness are critical to study climatic trends (long-term changes in the annual mean or in the seasonal cycle of humidity-related quantities) on specific locations or areas of the globe, which otherwise requires merging procedures using radiosonde data from nearby locations to circumvent large data gaps. We have averaged both quantities across all fixed stations reporting humidity within each major latitude region, year by year.

### 2.3.3 Vertical resolution and vertical extent

Since the vertical resolution varies with height – according to the height of the reported pressure levels (standard and significant) and depending on the number of levels with non-missing data for humidity –, the vertical resolution of an individual sounding must be defined by a vertical average. Since the vertical distance between consecutive levels, say $dz_k$, generally increases with height, with the lower levels being more populated than the upper layers, a geometric mean is more suitable than an arithmetic mean. So, the *mean vertical resolution of a single humidity sounding* ~~was~~ is here defined by the geometric mean of $\{dz_k\}$ for all levels with humidity data in the sounding profile:

$$mean\ vertical\ resolution = \frac{R_d}{g_0} \prod_{K=1}^{M} \left( \bar{T}_k \ln \frac{p_{k-1}}{p_k} \right)^{1/M} \qquad (2)$$

where $p_k$ is the atmospheric pressure at level $k$ ($k = 0$ denoting the lowest level with humidity data), $M$ is the number of levels with humidity data above the lowest level, $\bar{T}_k$ is the estimated mean temperature between level $k$ and its immediate, relevant lower level $k-1$, $R_d$ is the specific gas constant for dry air and $g_0$ is the standard gravity. (Note: IGRA's data-quality checks assures that vertical levels with valid humidity data have also valid temperature and pressure data.)

Since geopotential altitude is only given in part of the RAOB reports, the *vertical extent of an individual humidity sounding*, i.e. its *maximum height above mean sea level reached by the humidity measurements*, was estimated by adding the station's elevation to the height from the surface calculated upon pressure and temperature data from the surface level up to the top of the humidity sounding (highest level with a non-missing value for DPD or RH), whenever values of temperature and pressure at the surface are given; otherwise the height from the surface cannot be calculated. For mobile stations (ships and buoys), the elevation of the stations can be approximated to zero, unless the vertical extent of the sounding is too small, requiring data for the balloon release height.~~For moving stations, the elevation of station is taken equal to the geopotential height at the surface level, if given (otherwise, the vertical extent was not calculated)~~. For the purpose, it suffices to neglect moisture in the hypsometric equation; given that the virtual temperature is typically within 4 K above the actual temperature, the error in calculating geopotential height amounts to less than 1 %.

Following the above definitions, we have studied the statistical distributions of the vertical extent and the vertical resolution in humidity soundings from all IGRA-RS stations (including mobile) ~~were studied~~ over time by grouping individual values of both parameters in annual bins. To assess the shortness of humidity observations in RAOB, we have also calculated the vertical extent of temperature observations and their vertical resolution up to the top of the co-located humidity observations.

#### 2.3.4 Soundings eligible to estimate precipitable water vapor

Usually, the precipitable water vapor (column integrated water vapor mass per unit surface area) is estimated from the profile of water vapor mixing ratio between the surface to the 500-hPa level – i.e. the layer where ~ 95 % of the columnar mass of water vapor is and where humidity data from radiosondes are more often available and generally more accurate (Elliot et al., 1991; Gaffen et al., 1992; Ross and Elliott, 1996; Durre et al., 2009). In this paper, a humidity profile is considered eligible to estimate precipitable water vapor under the following conditions:

i. *Humidity data are given at the station's surface and at all standard levels laying between the surface and the 500-hPa level, except for the 925-hPa level.*
ii. *If humidity data is missing at a standard level apart from 925-hPa, a nearby significant level is acceptable if its height from the surface differs from the height of the missing standard level by less than 5 %*.
iii. *The distance between any consecutive levels with humidity data between the surface and the lowest level located more than 1 km away from the surface should not exceed 1 km, unless the station elevation is larger than 500 m.*

The 925-hPa level is not required here because this level was not standard until 1991. However, condition (iii) assures a minimal resolution in the planetary boundary layer, by including near-surface significant levels as well as the 925-hPa level when it is given; this is required because water vapor is highly variable and abundant in this region; the exception for very elevated stations contemplates the case when the first upper-air humidity record is at 850-hPa, i.e. ≈ 1.5 km above the mean sea level, but the height from the surface is less than 1 km. The IGRA-RS soundings fulfilling the above conditions will be hereafter referred to as *Sfc-to-500hPa humidity soundings*.

Typically, the first standard level higher than 1 km from the ground is 850-hPa. Thus, by including enough data at significant levels below the 850-hPa level, instead of requiring data at the current standard levels in the same layer – sorted out of 925-hPa and 1000-hPa, depending on the surface pressure – the definition given above accommodates much more soundings, particularly before 1992 when the 925-hPa level was not mandatory. To be sure, the relative amount of humidity data at near-surface levels is now examined, excluding 1000-hPa (around 0.1 km altitude) since this is frequently placed below the stations' elevation (339 m on average for the IGRA-RS fixed stations).

Figure 2 shows the evolution of the global percentage of humidity observations at the 925-hPa level and at any significant level between the surface and 850-hPa, out of the IGRA-RS soundings having humidity data at the surface on condition that the surface level pressure is higher than 925 hPa and 850 hPa, respectively (referred in the following as surface-upper-air soundings). The percentage of surface observations in all humidity soundings is also shown. First, note that the surface observations began in 1943 (in fact, not only of humidity but also of temperature), rising rapidly in the next 5 years to ≈ 95 % of the humidity soundings, decreasing then to a minimum of 60% in 1965 and broadly increasing again until 2000, staying above 95 % since then. In short, the humidity measurements at the surface level are mostly available since 1945 ~~and~~but ~~have been~~were in widespread use only since 2000. Secondly, the percentage of the surface-upper-air soundings having humidity data at the 925-Pa increased almost as a step-function~~is almost step-wised, with a discontinuity~~ around 1992. It increases from only 2 % in 1991 to 60, 86 and 98 % in the following three years. This change is coincident with the introduction of the 925-hPa level as an additional standard level in radiosonde messages in November 1991, and consistent with the fraction of stations already reporting that level in mid-1993 (Oakley, 1993, p. 24). Lastly, the percentage of the surface-upper-air soundings that have humidity data at any significant level below the 850-hPa level, beginning in 1948, generally increased with time, mainly in the 1960s, with a value larger than 80 % in recent years.

Using the definition given at the beginning of this section, we have studied the number and the percentage of stations (fixed and mobile) whose Sfc-to-500hPa humidity soundings exceed a given percentage out of the humidity soundings made in each year. The distance between missing standard levels and nearby significant levels was calculated from pressure and temperature data, neglecting moisture. A stricter definition of Sfc-to-500hPa humidity soundings, specifically, having humidity data at the surface and all upper-air, current standard levels up to 500-hPa was also studied for comparison.

**2.3.5 Current record length of time-series**

The '*current record length' of a humidity time-series in a given station and year*, is herein defined as the *number of elapsed years in the time-series with no gap years in the interim*. ~~For readiness~~To simplify, a calendar year with any amount of humidity data is counted as one. For past years, the current record length generally differs from the span of the entire time-series, which may continue after the year under consideration. Also, note that one station can have more than one time-series for humidity, as the humidity observations can be interrupted by one or more gap years. So, the full 'period of record' of a RS radiosonde station may be divided into several sub-periods of record for humidity. Recall that the years with humidity observations at each IGRA-RS station are indicated in Table S1.

We have studied the evolution of the average current record length of the humidity time-series, from all stations, year by year. The same was done for the time-series with humidity soundings 90 % of the days in the year, at least, and particularly consisting of Sfc-to-500hPa soundings, with their own record lengths. The calculation of the distribution of the current record length of the time-series with Sfc-to-500hPa soundings 90 % of the days in the year or more was repeated, by restricting the corresponding size of missing days to less than 10 days.

## 3 Overview on the completeness of radiosonde humidity observations

This section gives a general picture of the completeness of humidity observations over the years, using the data from the IGRA-RS stations defined in Sect. 2.2 and following the data analysis described in Sect. 2.3. The results of Sects. 3.1 and 3.2 refer to fixed stations, i.e., over 1600 stations on continents and islands, 14 ocean fixed weather-ships and 2 environmental buoys. In the remaining Sects. 3.3–3.5 mobile stations (99 ships) are equally included (see Table S1; moving stations are denoted by unspecified geographical coordinates).

### 3.1 Geographical coverage of humidity observations

Figure 3 shows the geographical distribution of the IGRA-RS fixed stations at specific years illustrating the growth of the global radiosonde network (cf. Fig. 1a); stations reporting humidity observations are highlighted and counted. Recall that a single station can change from PIBAL observations to RAOB, or even the reverse, during its period of activity. The IGRA-RS retains practically all RAOB data of IGRA; however, some stations have years with only PIBAL observations. Since almost all of the IGRA stations measuring temperature in a given year do also measure humidity at least part of the time (as seen by comparing black and solid blue lines in Fig. 1a), it is clear that most of the IGRA-RS station-years without humidity data (red crosses in Fig. 3) correspond to years of PIBAL observations alone (no RAOB).Although IGRA-RS excludes the IGRA stations without any RAOB at all, a comparison between Fig. 3 and Fig. 1a indicates that most of the IGRA-RS station-years without humidity data correspond to periods of wind-only observations. Concerning humidity observations, Fig. 3 shows that, by 1945, 2/3 of the stations were set in South Asia (British India) and Australia. Most of the data coverage over North-America, Greenland, Europe and North Asia, including the Artic region, as well as over the surrounding oceans [island stations and up to four fixed weather-ships operating simultaneously] took place between 1945 and 1955. By 1975, Central and South-East Asia, Africa, South-America, Antarctica (along its coastline, except for the Amundsen–Scott South Pole Station) and the surrounding oceans [islands, except for one or two fixed weather-ships operating in the North Pacific in the meantime] were already covered – although not as well as farther north regarding the continental regions, with the noteworthy exception of China territory. While the ~~total~~ global number of fixed stations measuring humidity ~~is~~ has remained practically unchanged since then (see Fig. 4 ~~Fig. 1a for the intermediate time~~), their geographical distribution changed significantly. At present (2015) there are more observation sites in South America, much less in Central and East Africa, much more in Western Asia, and a more even distribution in the rest of the world.

Although oceans have been mainly covered by island stations, ocean weather-ships were important for more than 30 years, between the late 1940s and the early 1970s. Judging from IGRA data, coverage from these ships was optimal between 1967 and 1972, ~~there were~~with 7 to 12 fixed ocean weather-ships transmitting radiosonde observations simultaneously, almost all located in the North Atlantic except for one ~~station~~ ship in the Norwegian Sea and another one in the North Pacific. Apparently, upper-air observations from these platforms ended a few years later (save for the station "M" in the Norwegian Sea, which continued until 1990), coincident with the growing use of satellite retrievals in weather forecasting. However, in-situ data coverage on oceans was improved by using balloon sounding systems on board of merchant ships~~,~~ – which obviously are not shown in Fig. 3~~;~~. ~~about~~ About 10 to 20 ships of opportunity have launched radiosondes concurrently along their routes from the 1980s to present. Other floating, mobile stations are also worth mention.~~The polar missions between 1985 and 2014 of the ice-breaker and research vessel Polarstern (Driemel et. al, 2017) and the~~ The automated ice-drifting stations surveying the north polar region during 1950–91 ~~are also worth noting~~were important considering their location and the amount of data gathered over the years. The missions to the Arctic and Antarctica performed by the ice-breaker and research vessel *Polarstern* (Driemel et al., 2016), covering also Atlantic Ocean regions during transit, provided substantial radiosonde humidity data in the periods 1985–1993 and 2000–2014. The weather ship *Polarfront* accounts for the largest amount of moving radiosonde data (after manning station "M" in the 1980's) but it operated only in the Norwegian Sea. For statistics about the humidity observations from all floating stations included in IGRA (fixed and mobile), see end of Table S1.

Figure 4 shows the count of fixed IGRA-RS stations with humidity observations in each year since 1930, by ~~climatic zones on Earth~~latitude bands of equal area, representing approximately tropical ($0^0$–$30^0$) and extratropical ($30^0$–$90^0$) latitudes in each hemisphere. One can see that before 1937 there was only one observing station and three years without data (1933-34 and 1936). ~~One can see that most of t~~The few ~~early~~ humidity-reporting stations existing by 1940 were placed in the ~~n~~Northern ~~polar zone (Arctic)~~ Hemisphere (in fact, most of them were in the Artic). From 1945 on they were located predominantly in the northern ~~temperate zone~~extratropics, even though they raised ~~in all regions~~other latitudes, at the fastest rate in the subsequent two decades. ~~In the climatic zones to the south of the Southern Tropic the number of stations is much smaller than in other zones, except for the southern subtropical zone which, from the 1970s onwards, compares with the Arctic in absolute number. Noticeably, the number of arctic stations decreased considerably between 1970 and 2000.~~ From 1970 onwards, the number of stations in the southern extratropics did not follow the observed growth observed elsewhere. This is not surprising since regions south of parallel $30^0$ S have significantly more ocean and less land. Note that ~~T~~the absence of mobile stations affects very little ~~the curves of~~ Fig. 4, since mobile stations are short numbered and have a short period of record (a few years), despite being important in covering the oceans. ~~E.g., considering the polar regions, according to IGRA, in 2016 there were two ships reporting radiosonde observations in waters around the Arctic Circle – although the number of observations gathered over the years is relevant.~~ The northern ~~temperate zone~~extratropics ha~~s~~ve accounted for about half of the world stations for decades, although its relative weight has been decreasing over the years~~. The~~; the decrease ~~of~~ in ~~the~~ number of stations situated in ~~the northern extratropics~~that region after 1990 is counterbalanced by an increase in the Tropics since then. ~~Considering the relative area of climatic zones, the southern temperate latitudes are always the most poorly represented, followed by the southern subtropics, which is not surprising since these are regions with significantly more ocean and less land. Lastly, the Arctic has been much better covered than Antarctica.~~

Using Eq. (1) and the method outlined in Sec. 2.3.1, Fig. 5a-b shows the average distance to the nearest station ($\bar{s}$) as ~~measured~~ calculated from points on mostly land or ocean/sea regions within ~~each climatic zone~~several latitude bands for every 15 years beginning in 1955, when the global radiosonde network was barely established. Fig~~.~~ure 5c shows a similar calculation but for the total surface area of each latitude band. Clearly, since at least 1970, $\bar{s}$ has changed little over continents and large islands ~~since at least 1970 and~~ at all latitudes, with values ranging from ∼ ~~about~~ 200 km in land areas of the northern temperate latitudes (35.0$^0$–66.5$^0$ N) to ∼700 km in land areas of the southern polar region (66.5$^0$–90.0$^0$ S), i.e., in Antarctica. ~~Considering the decreasing number of stations in the northern extratropics over time (cf. Fig. 4), this is an indication that stations became more evenly distributed.~~ In the Tropics (23.5$^0$ S – 23.5$^0$ N), ~~with~~ where $\bar{s}$∼ ~~≈~~ 500 km over land during the same period, the little change observed despite considerable changes in the distribution of continental stations can be explained as follows: the better spatial coverage over South America at present is offset by the poorer coverage over a large part of ~~the~~ sub-Saharan Africa~~, as seen in~~, save for the country of South Africa (see Fig. 3, panels c and d). ~~As to the continental lands in the southern extratropics, the number and distribution of stations has little changed in for several decades (cf. Fig. 3).~~ In contrast, in ocean/sea regions $\bar{s}$ not only is two to three times larger than over land regions~~,~~ – except in the southern polar region, i.e. in the Southern Ocean when compared to Antarctica~~,~~ – but has also degraded slightly over time in the ~~n~~Northern ~~extratropical~~ Hemisphere oceans and seas from the subtropics (23.5$^0$–35$^0$ N) to the Arctic (66.5$^0$–90$^0$ N). From a global perspective, the hemispheric differences in $\bar{s}$ due to the distribution of oceans and continents can be appreciated in Fig. 5c. Note that $2\bar{s}$ gives an estimate of the average separation between adjacent stations in regions where the radiosonde network is relatively regular. For example, Fig. 5a indicates a typical separation of about 400 km in the northern temperate continental regions~~, including~~ (coincident with the wealthier countries of the North Hemisphere – which is acceptable for synoptic weather forecasting. 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 need to be filled by satellite-based data and supplemented by surface observations, which are generally much denser than radiosonde stations.~~are accepted, in view of the relatively mild climatic conditions on oceans and the fulfillment from surface and satellite observations.~~ Nevertheless, ~~On~~ on a scale suitable for climate monitoring, the WMO recommends that upper-air stations should have a maximum average separation of 1000 km (WMO, 2011b). This would require $\bar{s} \leq 500$ km. Figure 5b

– which by including costal stations does not represent the actual station density in deep-ocean areas – indicates larger distancesthat fixed stations are too far apart in most oceans, particularly in the southern midlatitudes where $\bar{s} > 1000$ km.

### 3.2 Average fraction of days in a year with humidity observations and size of missing days

Figure 6 represents the fraction of days in a year having humidity observations, averaged across all humidity-reporting fixed stations in each of the three major latitude regions of both hemispheres, along with the standard deviation. The plot begins in 1945, about when upper-air humidity measurements became routine in radiosonde soundings (cf. Fig. 1). In all regions, the average fraction of the days in a year with humidity observations increased rapidly until around 1960, stabilizing since 1965 to values in the range 70–80 % in low latitudes and 80–95 % in mid- and high latitudes. Besides, the The corresponding standard deviations indicate that a non-negligible percentage of the stations have observations on every day in the year since the 1950s.

Figure 7 represents the typical ‘size of missing days’ in humidity observations in each year, as averaged across the fixed stations located within each major latitude zone. Recall that we only care withfocus only on sub-year missing days, since gaps of one or more years are exceptions related to interruptions of station’s operation or maybe to the lack of a functioning hygrometer. So, Fig. 7 gives a glint ofsummarizes the typical continuity of humidity time-series having any observations in the year. As a general picture, the average size of missing days decreased dropped from 4–6 months to about 1 month between 1945 and 1960; much of this change occurred before 1950, indicating that radiosonde measurements became rapidly regular in the early years. From 1960 to 2015, the stations at low latitudes present a trend in the size of missing days, from ≈ 30 days to 40 days on average; while stations at mid- to high latitudes present typical values of ≈ 20 days, except during the mid-1990s (≈ 30 days). Nonetheless, the dispersion of values away from the mean (not shown) indicates that some stations have time-series much more continuous than others, e.g. with daily data throughout the whole year.

The fraction of days with humidity observations, the size of missing days and the count of observations, detailed by station and year, is part of the first metadata set presented in this paper (see Sect. 4.1). Note that gap years are also evaluated in the metadata. Similar information is given for the Sfc-to-500hPa humidity soundings alone.

### 3.3 Global vertical extent and resolution of humidity observations

Figure 8a shows the distribution of the vertical-mean resolution in the annual humidity soundings since 1945, along with the homologous distribution in the simultaneous temperature soundings, limited to the highest level with humidity measurements for comparison. The vertical-mean resolution of each soundings was calculated by Eq. (2), rounded to the nearest decameter. For clarity, the curves displaying the mean and the quartiles are smoothed by a 5-year running mean. The differences between the distributions for temperature and humidity indicate missing humidity data in a few percent of the vertical levels with temperature data; although statistically irrelevant, such vertical gaps may be quite significant in individual soundings. The vertical resolution was relatively poor on average and highly variable in the early decades: until around 1965 the mean and median were coincident and varied in the range 1.1–1.4 km, with a midspread (interquartile range) of almost 1 km. Both the average resolution and the midspread improved consistently from 1965 to between 2000 and 2005. Since 2005, ¾ of the soundings have a vertical-mean resolution better than 0.5 km, with half of the values ranging from 0.3 to 0.5 km.

Figure 8b shows how the maximum height above the mean sea level reached by either temperature or humidity measurements is distributed among the corresponding soundings on each year. Curves are smoothed by a 5-year running mean, as in Fig. 8a. However, the distribution is restricted to temperature and humidity soundings with surface data, these representing, respectively, 86.7 % and 86.0 % of the total RAOB soundings in the period 1945–2016. ~~Also~~In addition, the soundings from mobile stations with missing geopotential height at the surface level ~~had to be~~are excluded, for consistency with the hypsometric calculations used in the dataset presented in this paper. According to Fig. 8b, by 2003, 75% of the temperature soundings with surface data reached an altitude of 22 km, i.e. ~ 50 hPa. Note that Durre et al. (2006) reported that, by the same year, 74% of all IGRA 1 soundings reached at least the 100-hPa level, i.e. ~ 16 km; this lower height is due to the inclusion of PIBAL data in their analysis of IGRA 1, whereas our analysis is restricted to RAOB in IGRA 2.~~Note that the vertical extent of temperature measurements (i.e. of RAOB) is normally much larger than the vertical extent of balloon-born observations of any kind: e.g., ¾ of all soundings reach 100 hPa (≈ 10 km) in 2003, from the data of the first version of IGRA, as reported in Durre et al. (2006); whereas ¾ of the temperature soundings with surface data reach 22 km in the same year, according to the present analyses of IGRA-RS data selected from IGRA v2. This difference is essentially due to the removal of PIBAL stations in our analysis.~~

Contrary to temperature, which can be measured up to the maximum height achieved by the sounding balloon (burst altitude, although the highest reported level is usually limited by the standard levels in use), the vertical reach of humidity measurements depends on the working range of humidity sensors (and, to some degree, on reporting practices). In Fig 8b we can observe that the top of the humidity observations is at present situated almost 5 km below the maximum altitude of the temperature measurements (close to the burst altitude) on global average (either mean or median). But the difference in vertical extent between temperature and humidity observations has changed greatly over the years, with a maximum mean value of 12.5 km by 1980. Until the late 1960s the burst altitude increased at a much faster rate that the top of humidity observations, except for a few years after 1965 when the reverse happened. This last feature ~~and~~ is coincident with the introduction of the carbon hygristor. Likewise, the isolated peak of the global mean maximum height of humidity observations around 1970, as seen in Fig. 8b ('HUM'), seems to indicate an exploratory period with the new instrument. Noteworthy, in 1970 the WMO stated that "no routine observations of humidity are made in the stratosphere and no practical use is envisaged for such current observations" (Hawson, 1970). Figure 8b undoubtedly shows that humidity was mostly measured in the troposphere until the mid-1960s. In contrast, from the mid-1980s onwards, the vertical extent of humidity observations has increased consistently and faster than the RAOB-top – roughly by 4 km per decade on global average –, denoting improvements of humidity sensors (see Sect. 1.3). In 2015, ~~¾~~ three quarters of the humidity-top values extend to over 13 km, and half to over 26 km, indicating that many humidity reports extend into the lower stratosphere. The accuracy of radiosonde humidity measurements in the stratosphere is beyond the scope of this paper~~; however, the accuracy of fairly recent measurements in the upper-troposphere, and certainly more above too, was considered inadequate for climate studies~~. However, only 15 years ago, international experts pointed out that the accuracy of current operational measurements of humidity in the upper-troposphere was inadequate for addressing climate variability and change (despite the usefulness of some sensors) and a challenge for future operational radiosondes (Durre et al., 2005). It is interesting to note that the interquartile range of the top height of humidity soundings has increased much in the period 1985–2000, i.e., individual values became widely dispersed around the median. This is likely related to the proliferation of humidity sensors of different kinds, thus increasing the instrument variations among stations as new instruments coexist with older ones. For instance, by 1989 the WMO had identified 20 major radiosonde types in use worldwide (Kitchen, 1989a). While this number has fallen to 13 by 2002 (Elms, 2003), Fig. 8b shows that the dispersion in the vertical range of different humidity sensors has increased in the meantime; changes in instrument-dependent observing practices may have a role.

The average vertical resolution of the humidity observations, by year and station, as well as the individual values by station, date and time, are given in the metadata sets provided in this paper (see Sect. 4). Similar metadata is provided for the Sfc-to-500hPa humidity soundings, as defined in Sect. 2.3.4, in which case the vertical resolution is calculated for the levels between the surface and the 500-hPa level and is normally finer than depicted in Fig. 8a.

Since IGRA-RS observations do not have surface humidity data prior to 1945, and surface data is not always given after that year, to account for all soundings the vertical extent of humidity observations must be alternatively represented by the lowest pressure corresponding to humidity data. The same applies to soundings from mobile stations~~,~~ if geopotential height is not given at the surface level (it is missing in 30 % of the corresponding RAOB). Although the local sea level is normally within ± 10 m from mean sea level, taking zero as the baseline height can lead to large relative errors if humidity is only measured very close to the radiosonde station elevation (balloon release height)~~since the station's elevation is variable, and the geopotential height of the surface level is missing in 30 % of the corresponding RAOB~~. Therefore, the average pressure of the top of humidity soundings, by station and year, are given in the first metadata set presented in this paper; average values are represented by a geometric mean (see Sect. 4.1). As to the second set for individual observations –metadata by station, date and time – both the top pressure and the corresponding altitude, whenever this can be calculated, are provided (see Sect. 4.2).

### 3.4 Global relative amount of Sfc-to-500hPa humidity soundings

Recall that our definition of Sfc-to-500hPa humidity soundings (Sect. 2.3.4) is intended to represent the soundings with ~~a minimal amount of data~~enough vertical level, and almost evenly distributed near the surface, such that the water vapor profiles can be properly described and the precipitable water can be estimated. In this respect, the completeness of such humidity observations based on current standard pressure levels alone is unsatisfactory for two reasons: first, the level 925-hPa was barely used before 1992 (see Fig. 2); second, the sounding data at significant levels, often related to features of temperature rather than RH, or at other additional levels, are equally good provided that the vertical resolution (habitually increasing towards the surface) is not too different. Next, we will compare the distribution of Sfc-to-500hPa soundings among IGRA-RS stations in each year using alternative definitions:

A) Humidity data at the surface and all current standard levels above the surface up to 500 hPa, i.e. at pressure levels $\{p\colon p \in \{1000, 925, 850, 700, 500\ \mathrm{hPa}\} \cup \{p_{\mathrm{SFC}}\},\ p \le p_{\mathrm{SFC}}\}$;

B) Definition given in Sect. 2.3.4.

The soundings meeting either of the above two definitions are coined as *Hum-A* or *Hum-B* in the following analysis. Note, however, that only definition B was used to prepare the metadata sets supplied in this paper.

Let P be the percentage of Sfc-to-500hPa soundings (HUM-A or HUM-B, at our choice), out of all humidity soundings from an arbitrary station in a given year. Figure 9a shows the evolution of the number of stations with ~~the percentage of~~P for Hum-A soundings exceeding given values, ~~out of all soundings with humidity in each year –~~ since nearly the time radiosonde humidity data at the surface level are first available. Comparing Fig. ~~1a~~ 9a with Fig. ~~9a~~1a, we can see that between 1945 and 1991 only a small fraction of the ~~RS~~ radiosonde stations carried out Hum-A observations in most of the soundings, say, with P within ~~a percentage~~the range P > 95 %. Things changed drastically between 1992 and 1994, when the number of stations with a significant percentage of Hum-A soundings, e.g. 20 % < P < 80 %, increased by an order of magnitude. This change reflects the change in the observing practice shortly after 925-hPa was internationally adopted as a standard pressure level (see Fig. 2). In recent years there are over 400 stations with P > 95 % and almost 800 with P > 80 %. Figure 10a shows how the probability of finding a IGRA-RS station with a percentage of Hum-A soundings exceeding a given value was reversed over the last five decades: by 1965, only 12 % of the stations had at least 20 % of Hum-A soundings, whereas by 2015 ≈ 91 % of the stations had at least 80 % of Hum-A soundings.

Figure 9b is the counterpart of Fig. 9a, for the less stringent Hum-B soundings. Recall that 925-hPa is now treated in the same way as any near-surface non-standard level, and, for the rest, significant levels close to standard levels are allowed. (Note: the filling of missing data according to condition (ii) of Sect. 2.3.4 affected 22 % of the total Hum-B soundings.) One can observe that, before 1992, the number of stations having moderate to high percentages P of Hum-B observations is much larger than the number of stations having the same percentages of Hum-A observations; from 1992 on, that difference is moderate and only significant when P is very high. Besides the sudden increase of the stations with many Hum-B observations in 1992, there other two noteworthy change points: a sudden increase around 1970 and another one 2000. These changes are related to the increase of the global percentage of humidity soundings with surface data happening at about the same time (cf. Fig. 2). For example, the number of stations providing Hum-B vertical profiles in more than 80 % of the soundings doubled from ≈ 40 to 80 between 1969 and 1971, it doubled from 250 to 500 between 1990 and 1993, it increased from 520 to 700 between 1999 and 2001; a much more constant value of around 800 is observed in the period 2005–2015. The inverse cumulative distribution function in Fig 10b shows the probability of finding a IGRA-RS station with a percentage of Hum-B soundings exceeding a given value in different trienniums. About fifty years ago, the fraction of stations having at least 20 % of Hum-B soundings was as large as 40 %, increasing gradually over time to nearly 100% at present. Furthermore, 97 % of the presently active stations have at least 80 % of Hum-A soundings.

The number of Hum-B Sfc-to-500hPa humidity soundings in each station and year, as well as the fraction of days in a year having such soundings, are both given in the first metadata set presented in this paper (see Sect. 4.1). Information on other parameters describing humidity completeness but focusing on these soundings is also provided. The metadata set regarding individual observations identifies the Sfc-to-500hPa humidity soundings and provides information on their vertical resolution between the surface and 500 hPa (see Sect. 4.2).

### 3.5 Amount of long-term time-series

Figure 11a (black lines) illustrates the number of humidity time-series, one per station, fixed or mobile, with a 'current record length' (elapsed years until the year in abscissa) exceeding a given number of decades: 1, 3 and 5. The time-series refer to periods of consecutive years with any observations in the year. For comparison (color lines) the same is shown for the time-series with data 90 % of the days in the year or more. It should be noted that the concurrent series with more than 10 years of back data begin in 1948, even though the first upper-air humidity measurements began in 1930 and by 1949 there were already around 300 stations measuring humidity (cf. Fig. 1a). The initial slope of the curves in Fig. 11a denotes the rapid growth of the global radiosonde network after the second world war, with many stations measuring humidity regularly. Nevertheless, the curves are not monotonous, due to the closing of stations in the past and to the existence of gap years in many stations. E.g., if we want to collect the largest amount of parallel humidity time-series extending back in time to over 10 years, Fig. 11a tells us that we should pick IGRA data until 1987–90, corresponding to about 750 radiosonde stations. However, as of 1976, the number of parallel time-series with humidity observations 90 % of the time or more and extending back in time to over 30 years has been increasing over the years, representing 200 stations in 2016. By restricting the series to those having sufficient vertical sampling between the surface and 500 hPa to estimate precipitable water vapor, the number drops to about 15 between 1978 and 2000, only then increasing steadily to about 50 by 2015. If we consider a much shorter duration, e.g. 10 years as a minimum, the time-series until the same year becomes several times more numerous and start much earlier.

Fig. 11b represents the distribution of the current record length under more restrictive conditions. First, it considers only the time-series consisting of Sfc-to-500hPa soundings (definition of Sect. 2.3.4) covering 90 % of the days in the year or more; secondly, it further limits the time-series to those with less than 10 consecutive missing days in every year, meaning that all months are evenly represented. We can see that the number of parallel time-series with regular (fraction of days in a year ≥ 90 %) Sfc-to-500hPa soundings and extending back in time to more than 10 (30) years has increased rapidly since 2000, after four (two) decades with a value of less than 100 (20), only then increasing steadily to about 330 (50) by 2016. However, the Sfc-to-500hPa time-series with the same length but being almost continuous (fraction of days in a year ≥ 90 %; size of missing days < 10 days) are much less numerous.

~~Evidently,~~ Fig~~.~~ure 11 is only illustrative. There is no simple answer to the question of ~~since~~ when we have enough data to (in theory) perform climate studies from radiosonde humidity data: it depends on the strictness of completeness criteria. (In practice, it depends also on the accuracy and homogeneity of measurements by different instruments.) Note that the length of the time-series under user-specified conditions for any of the metadata parameters defined in the next section, either backward or forward in time, can be derived from the information contained in the related metadata set (Sect. 4.1). In addition, the years with humidity observations can be found in Table S1 (as well as in a data file accompanying the two main metadata sets): one or more periods per station, depending on whether there are gap years or not in the station's full period of record.

## 4 Metadata sets of completeness of radiosonde humidity observations based on the IGRA

The metadata sets describing the completeness of radiosonde humidity records collected in the IGRA are outlined next. They are constructed upon the metadata parameters introduced in Sect. 2.3 and examined in Sect. 3, except for minor adaptions described in the subsections below. The combination of selection criteria, by simply specifying value ranges for the metadata parameters, offers a plethora of choices to the user. Both sets refer to sounding data from the IGRA-RS sub-set of IGRA 2 stations, as defined in Sect. 2.2 and listed in Table S1, which essentially excludes pilot-balloon stations.

### 4.1 Metadata by station and year

The metadata parameters regarding each IGRA-RS station and year (annual statistics) are defined in Table 2. Since pressure is the vertical coordinate that is always present in humidity radiosonde data, and the calculation of height above mean sea level is impossible when the surface level is missing, the average vertical extent of humidity data at a given station and year is represented by the geometric mean value of the lowest pressure with humidity data. (Given the huge number of values involved, the arithmetic mean of the logarithm of pressure was calculated first and then exponentiated.) Note that the geometric mean pressure of the highest level with humidity data provides a natural measure of the corresponding arithmetic mean altitude: insomuch as pressure decrease almost exponentially with height, the arithmetic mean height above sea level is roughly proportional to the logarithm of the geometric mean pressure. The statistics for the vertical resolution of humidity data consists in the annual mean of the 'mean vertical resolution' of individual soundings at each station, calculated by Eq. (2). The IGRA data meeting the definition of Sfc-to-500hPa humidity soundings as stated in Sect. 2.3.4 have their own metadata: number of observations, average vertical resolution, size of missing days, and fraction of days in a year having humidity data.

The metadata set is as a plain-text file, each data record containing the metadata values, station-by-station and year-by-year. For each station, the yearly variables defined in Table 2 are displayed chronologically. To put humidity metadata into context, the yearly number of soundings and of RAOB soundings are also given in the dataset. The years within a station's period of record with observations of any meteorological variable (wind, temperature, humidity) are kept, using appropriate missing values when humidity data is missing throughout a whole year. While humidity data begin in 1930, pre-radiosonde years are included to preserve the stations history.

NB – The metadata set is complemented by a list of the IGRA-RS stations with information on geographic coordinates (if fixed), name, and country, along with original metadata describing the periods of record for humidity and the corresponding amounts of observations. This is an ASCII version of Table S1, for computing purposes.

### 4.2 Metadata by station, date and time

The metadata parameters regarding individual soundings at each IGRA-RS station are defined in Table 3. The utility of the corresponding metadata set is to allow a fine selection of humidity data, as a complement to a first selection based on the statistical parameters of Table 2. This metadata set is organized into one plain-text file per station, and one data record per sounding. Below the headline, each record contains: the date, nominal hour, latitude and longitude of the balloon launch, followed by the related humidity metadata.

## 5 Availability of metadata sets

A dataset combining the two metadata sets presented in Sect. 4 is available on Zenodo, DOI: 10.5281/zenodo.1332686. The accompanying 'Readme' file gives the necessary information about the data format and file contents. The update is planned to take place each time two full years have been completed in IGRA, starting on January 2019 for the period 2017–2018.

## 6 Summary and recommendations

~~The completeness of radiosonde humidity observations from IGRA (Version 2) was studied, upon selecting the stations with a minimal amount of radiosonde observations in their period of record until the end of 2016, denoted herein as 'IGRA-RS'.~~ This work has studied the completeness of radiosonde humidity observations compiled in the IGRA Version 2 (Durre et al., 2016; Durre et al., 2018) upon setting aside the IGRA stations with more than 95% pilot-balloon data in every year of their period of record until the end of 2016. The selected set (denoted IGRA-RS) retains virtually all RAOBs distributed by~~comprises~~ 1723 stations, including 1300 WMO stations, of which 178 and 16 are, respectively, current GUAN and GRUAN sites. The earliest~~y~~ humidity ~~reports~~ observations are from the 1930s, when the radiosonde era had begun, but the data amount is only significant after around 1945. Several parameters describing the completeness of humidity observations were defined and then examined in statistical terms~~ based upon IGRA-RS sounding data~~, providing a global picture of humidity completeness in radiosonde observations over time. The main conclusions, for the years beginning in 1945, are as follows.

- The radiosonde network providing humidity observations in the northern temperate and polar latitudes ~~and arctic region~~ was essentially established in the mid-1950s. In the mid-1970 the globe ~~was already~~became covered by practically the same number of humidity-reporting stations (nearly the same as radiosonde stations) as it is today, although the distribution of stations has changed considerably over time. ~~Remarkably, t~~The averaged distance to the nearest station measured from points over continents and large islands (from Mindanao up to Greenland in size), ranging from ~~~~≈ 200 km in the northern midlatitude countries to ~~~~≈ 700 km in Antarctica, has little changed too. ~~This can be explained as follows: in the northern extratropical land regions, the general decrease in the number of observation stations with time was accompanied by a more even spatial distribution; in the tropical land regions, the overall increase in the number of stations with time was accompanied by a much worst coverage of sub-Saharan Africa; in the southern extratropical land regions, both the number and the distribution of stations have been almost unchanged.~~ However, the spatial distribution of continental stations continue to show important regional disparities, such as the poor coverage of most of sub-Sharan Africa which has worsened from 1970 to present. Concerning ocean/sea areas, and disregarding mobile stations (mostly 'ships of opportunity'), the oceans of the southern temperate latitudes always exhibit the poorest coverage, with an average distance to the nearest station exceeding 1000 km. Remarkably, ~~S~~since ~~the 1970s~~during the last half century the average distance to the nearest station has increased considerably in the North Hemisphere oceans and seas extending from the subtropics to the Arctic.
- The fraction of days in a year with humidity observations has greatly increased until 1960, on average calculated for each year across the stations on each of the main latitude regions. Since 1965, the mean values are in the range 70–80 % in low latitudes and 80–95 % in mid- and high latitudes annually, although individual values vary widely among stations. Between 1945 and 1955, the 'size of missing days' in humidity observations (largest gap in days) in each year having any humidity data decreased from about five months to one month on global average. After 1960, the size of missing days exhibits averages around 20 days in middle and high latitudes; in low latitudes it has increased over time from around 30 to 40 days.
- Humidity was measured mostly in the lower to middle troposphere until the mid-1960s. Since the mid-1980s, the mean height achieved by humidity measurements has increased by ≈ 4 km per decade on global average; the gap distance to the burst altitude (denoted by the maximum height of the temperature measurements) more than halved. At present, that gap is ≈ 5km and ~~¾~~ three quarters of the humidity soundings reach to at least 13 km altitude. However, the dispersion of the maximum height of humidity data around the mean has increased too, likely due to the coexistence of older humidity sensors with new sensors. The vertical-mean resolution of humidity observations (geometric mean of the distance between consecutive levels with data), has a median in the range of 1.3–1.4 km until 1965, among all humidity soundings in each year, with an interquartile range of ≈ 1 km. Thereafter, the median and midspread have improved constantly before stabilizing in the early 2000s; since then, half of the annual soundings present a vertical-mean resolution between 0.3 and 0.5 km.

- The number of stations providing Sfc-to-500hPa profiles with adequate resolution for calculating precipitable water vapor in more than 80 % of the reported humidity soundings in a year has changed from a few dozen in the period 1945–1970 to around 800 in recent years (97 % of the active radiosonde stations by 2014/16). In general, the amount of stations having a significant percentage of Sfc-to-500hPa observations shows a sudden increase around the years 1970, 1991/92 and 2000. These change points are associated with the availability of data at the surface level and at non-standard near-surface levels; the latter provide important information in the planetary boundary layer before 1992, when the 925-hPa level was not standard and ~~it~~ was rarely used.
- The amount of ~~long-term~~ humidity time-series with a given number of consecutive years of data until a given year depends not only on the year~~, the time span and~~ and the length of record, but also on other completeness criteria. ~~E.g.:~~For example, the station-based time-series extending back in time for at least 30 years and being 90 % complete (in terms of fraction of days in a year) begin with a few units in 1977 and represent around 200 stations in recent years; by further requiring Sfc-to-500hPa completeness, the other criteria being the same, the number of time-series drops to 50 by 2016. Evidently, the equivalent time-series until the same year but with a much shorter duration, e.g. 10 years as a minimum, are several times more numerous and start much earlier. In short, the amount of humidity time-series that are *potentially* available to perform climate studies, i.e., discounting accuracy requirements and biases due to instrument changes, depends strongly on the strictness of the completeness criteria.

~~Finally~~Furthermore, th~~e~~is paper ~~presented~~ presents a dataset detailing the completeness of humidity observations (RH or DPD together with pressure and temperature) based on the data from the IGRA-RS stations. The dataset (Ferreira et al., 2018) consists of: 1) statistical metadata for each station and year, in a single file for the 1723 stations and their full period of record until 2016; 2) metadata specific to individual soundings, organized in one file per station and covering 39.5 million soundings for the same stations; and (3) list of stations along with the observing periods for humidity and the corresponding number of observations. The metadata parameters were designed to facilitate the selection of upper-air humidity data from IGRA according to a plethora of choices, therefore being able to meet specific research needs in the areas of atmospheric and environmental sciences.

It is widely known that the usefulness of historical radiosonde data depends crucially on metadata information about instrumentation and observing practices (Schwartz and Doswell, 1991; Elliot and Gaffen, 1991; Gaffen et al., 1996; Parker and Cox, 1995). However, sampling differences among stations associated with geographical coverage, observing periods and missing data, are no less important than differences in data precision and accuracy. Reporting practices related to limitations of humidity sensors affect particularly the humidity records (Garand et al., 1992; McCarthy et al. 2009). In this respect, the metadata presented in this paper, if used as a tool to find out the more complete humidity time series – relatively long, regular and continuous; vertically extensive and well resolved –, should be crossed with the coordinates of stations and the station history metadata available in IGRA. On the other hand, the present metadata might accelerate progress in the current research on the homogenization of radiosonde humidity data (McCarthy et al., 2009; Dai et al., 2011), by sampling stations with coincident observing periods and satisfying reasonable completeness criteria.

## Acknowledgments

We thank Imke Durre ,for clarifying some points about the Integrated Global Radiosonde ArchiveIGRA Version 2 during the preparation of this work, and David Carlson for editorial advice on the first draft of the paper, and two anonymous reviewers for their helpful comments. We also thank to Orlando García for IT support at EPhysLab (Environmental Physics Laboratory at the Faculty of Science of the University of Vigo, Ourense Campus). This work was funded by the Spanish government within the EVOCAR project (CGL2015-65141-R), co-funded by the European Regional Development Fund, and partially supported by the Galician regional government through the program Consolidation and Structuring of Competitive Research Units – Competitive Reference Groups (ED431C 2017/64-GRC).

## References

Andersson, E., Hólm, E., Bauer, P., Beljaars, A., Kelly, G. A., McNally, A. P., Simmons, A. J., Thépaut, J., and Tompkins, A. M.: Analysis and forecast impact of the main humidity observing systems, Q. J. Roy. Meteor. Soc., 133, 1473–1485, doi:10.1002/qj.112, 2007.

Bellamy, J. C.: Some basic characteristics of observational data, in Meteorological Observations and Instrumentation – 1970, Proceedings of the American Meteorological Society Symposium on Meteorological Observations and Instrumentation, Washington, D. C., 10–14 February 1969, doi:10.1007/978-1-935704-35-5, 1970.

Bodeker, G. E., Bojinski, S., Cimini, D., Dirksen, R. J., Haeffelin, M., Hannigan, J. W., Hurst, D. F., Leblanc, T., Madonna, F., Maturilli, M., Mikalsen, A. C., Philipona, R., Reale, T., Seidel, D. J., Tan, D. G. H., Thorne, P. W., Vömel, H., and Wang, J.: Reference Upper-Air Observations for Climate: From Concept to Reality, B. Am. Meteor. Soc., 97, 123–135, doi:10.1175/BAMS-D-14-00072.1, 2016.

Brettle, M. J. and Galvin, J. F. P. : Back to basics: Radiosondes: Part 1 – The instrument, Weather, 58, 336–341, doi:10.1256/wea.126.02A, 2003.

Connell, B. H. and Miller, D. R.: An interpretation of radiosonde errors in the atmospheric boundary layer, J. Appl. Meteorol., 34(5), 1070–1081, doi:10.1175/1520-0450(1995)034<1070:AIOREI>2.0.CO;2, 1995.

Dabberdt, W. F., Cole, H., Paukkunen, A., Horhammer, J., Antikainen, V., and Shellhorn, R.: Radiosondes, in Encyclopedia of Atmospheric Sciences, Vol. 6, ed. By Holton, J.R., Pyle, J., and Curry, J. A., Elsevier Science/Academic Press, Amsterdam, The Netherlands, 2002, pp. 1900–1913, 2002.

Dai, A., Wang, J, Thorne, P. W., Parker, D. E., Haimberger, L., and Wang, X. L.: A new approach to homogenize radiosonde humidity data, J. Climate, 24, 965–991, doi:10.1175/2010JCLI3816.1, 2011.

Dirksen, R. J., Sommer, M., Immler, F. J., Hurst, D. F., Kivi, R., and Vömel, H.: Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde, Atmos. Meas. Tech., 7, 4463-4490, doi:10.5194/amt-7-4463-2014, 2014.

DuBois, J. L., Multhauf, R. P., and C. A. Ziegler, C. A.: The Invention and Development of the Radiosonde, Smithsonian Institution Press, Washington, DC, 78 pp, 2002.

Durre, I.: Integrated Global Radiosonde Archive V2, Dataset Description, Version 1.0, 15 pp, [Word document available at https://www1.ncdc.noaa.gov/pub/data/igra], 2016

Durre, I., Reale, T., Carlson, D., Christy, J., Uddstrom, M., Gelman, M., and Thorne, P.: Improving the usefulness of operational radiosonde data, B. Am. Meteor. Soc., 86, 411–418, doi:10.1175/BAMS-86-3-411, 2005.

Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, J. Climate, 19, 53–68, doi:10.1175/JCLI3594.1, 2006.

Durre, I., Williams Jr., C. N., Yin, X., and Vose, R. S.: Radiosonde-based trends in precipitable water over the Northern Hemisphere: An update, J. Geophys. Res., doi:10.1029/2008JD010989, 2009.

Durre, I., Vose, R.  S., Yin, X., Applequist, S., and Arnfield, J: Integrated Global Radiosonde Archive (IGRA) Version 2, full period-of-record (POR) data, NOAA National Centers for Environmental Information, doi:10.7289/v5x63k0q, 2016.

Durre, I., Yin, X., Vose, R. S., Applequist, S., and Arnfield, J.: Enhancing the data coverage in the Integrated Global Radiosonde Archive, J. Atmos. Ocean. Tech., 35, 1753–1770, doi:10.1175/JTECH-D-17-0223.1, 2018.

Driemel, A., Loose, B., Grobe, H., Sieger, R., and KönigLanglo, G.: 30 years of upper air soundings on board of R/V POLARSTERN, Earth Syst. Sci. Data, 8, 213–220, doi:10.5194/essd-8-213-2016, 2016.

~~Driemel, A., Fahrbach, E., Rohardt, G., Beszczynska-Möller, A., Boetius, A., Budéus, G., Cisewski, B., Engbrodt, R., Gauger, S., Geibert, W., Geprägs, P., Gerdes, D., Gersonde, R., Gordon, A. L., Grobe, H., Hellmer, H. H., Isla, E., Jacobs, S. S., Janout, M., Jokat, W., Klages, M., Kuhn, G., Meincke, J., Ober, S., Østerhus, S., Peterson, R. G., Rabe, B., Rudels, B., Schauer, U., Schröder, M., Schumacher, S., Sieger, R., Sildam, J., Soltwedel, T., Stangeew, E., Stein, M., Strass, V. H.,~~

~~Thiede, J., Tippenhauer, S., Veth, C., von Appen, W.-J., Weirig, M.-F., Wisotzki, A., Wolf-Gladrow, D. A., and Kanzow, T.: From pole to pole: 33 years of physical oceanography onboard R/V Polarstern, Earth Syst. Sci. Data, 9, 211-220, doi:10.5194/essd-9-211-2017, 2017.~~

Elliott, W. P. and Gaffen, D. J.: On the utility of radiosonde humidity archives for climate studies, B. Am. Meteor. Soc., 72, 1507–1520, doi:10.1175/1520-0477(1991)072<1507:OTUORH>2.0.CO;2, 1991.

Elliott, W. P., Smith, M. E, Angell, J. K.: Monitoring tropospheric water vapor changes using radiosonde data, ed. M. E. Schlesinger, Developments in Atmospheric Science, Elsevier, Amsterdam, Vol. 19, pp. 311-327, doi:10.1016/B978-0-444-88351-3.50027-1, 1991.

Elms, J.: WMO catalogue of radiosondes and upper-air wind systems in use by members in 2002, Report No. 80, WMO/TD No. 1197, Part A [available at https://library.wmo.int], 2003.

Eskridge, R. E., Alduchov, O. A., Chernykh, I. V., Panmao, Z., Polansky, A. C., and Doty, S.: A Comprehensive Aerological Reference Data Set (CARDS): Rough and systematic errors, B. Am. Meteor. Soc., 76, 1759–1775, doi:10.1175/1520-0477(1995)076%3C1759:ACARDS%3E2.0.CO;2, 1995.

Ferreira, A. P., Nieto, R., and Gimeno, L.: A dataset of completeness of radiosonde humidity observations based on the IGRA [Data set], Zenodo, doi:10.5281/zenodo.1332686, 2018.

Free, M. and Seidel, D. J.: Causes of differing temperature trends in radiosonde upper air data sets, J. Geophys. Res., 110, D07101, doi:10.1029/2004JD005481, 2005.

Gaffen, D. J., Barnett, T. P. and Elliott, W. P.: Space and time scales of global tropospheric moisture, J. Climate, 4, 989–1008, doi:10.1175/1520-0442(1991)004<0989:SATSOG>2.0.CO;2, 1991.

Gaffen, D. J.: Historical Changes in Radiosonde Instruments and Practices, IOM Report No. 50, WMO/TD-No. 541, 128 pp. [available at https://library.wmo.int], 1993.

Gaffen, D. J.: A Digitized Metadata Set of Global Upper-Air Station Histories, NOAA Technical Memorandum ERL ARL-211, Silver Spring, MD, 38 pp. [ftp://ftp.ncdc.noaa.gov/pub/data/images/gaffen1996.pdf], 1996.

Gaffen, D. J., Robock, A., and Elliott, W. P.: Annual cycles of tropospheric water vapor, J. Geophys. Res., 97(D16), 18185–18193, doi:10.1029/92JD01999, 1992.

Garand, L., Grassotti, C., Hall, J., and Klein, G. L.: On differences in radiosonde humidity-reporting practices and their implications for numerical weather prediction and remote sensing, B. Am. Meteor. Soc., 73, 1417–1423, doi:10.1175/1520-0477(1992)073%3C1417:ODIRHR%3E2.0.CO;2, 1992.

Gutnick, M: Mean Annual Mid-Latitude Moisture Profiles to 31 km, Air Force Research Report, AFCRL-62-681, July 1962 [http://www.dtic.mil/dtic/tr/fulltext/u2/286147.pdf], 1962.

Haimberger, L., Tavolato, C, and Sperka, S: Toward elimination of the warm bias in historic radiosonde temperature records–Some new results from a comprehensive intercomparison of upper-air data, J. Climate, 21, 4587–4606, doi:10.1175/2008JCLI1929.1, 2008.

Hall, E. G., Jordan, A. F., Hurst, D. F., Oltmans, S. J., Vömel, H., Kühnreich, B., and Ebert, V.: Advancements, measurement uncertainties, and recent comparisons of the NOAA frost point hygrometer, Atmos. Meas. Tech., 9, 4295-4310, doi:10.5194/amt-9-4295-2016, 2016.

Hartten, L. M., Cox, C. J., Johnston, P. E., Wolfe, D. E., Abbott, S., McColl, H. A., Quan, X.-W., and Winterkorn, M. G.: Ship- and island-based soundings from the 2016 El Niño Rapid Response (ENRR) field campaign, Earth Syst. Sci. Data, 10, 1165-1183, doi:10.5194/essd-10-1165-2018, 2018.

Hawson, C. L.: Performance requirements of aerological instruments, WMO Tech. Note 112, WMO 267, 49 pp. [available at https://library.wmo.int], 1970.

Hickman, A.: History of pilot ballooning, Weather, 70, S21-S23, doi:10.1002/wea.2526, 2015.

Hurst, D. F., Hall, E. G., Jordan, A. F., Miloshevich, L. M., Whiteman, D. N., Leblanc, T., Walsh, D., Vömel, H., and Oltmans, S. J.: Comparisons of temperature, pressure and humidity measurements by balloon-borne radiosondes and frost point hygrometers during MOHAVE-2009, Atmos. Meas. Tech., 4, 2777-2793, doi:10.5194/amt-4-2777-2011, 2011.

Ingleby, B., Pauley, P., Kats, A., Ator, J., Keyser, D., Doerenbecher, A., Fucile, E., Hasegawa, J., Toyoda, E., Kleinert, T., Qu, W., St. James, J., Tennant, W., and Weedon, R.: Progress toward high-resolution, real-time radiosonde reports, B. Am. Meteor. Soc., 97, 2149–2161, doi:10.1175/BAMS-D-15-00169.1, 2016.

Ingleby, B.: An Assessment of Different Radiosonde Types 2015/2016, ECMWF Technical Memorandum No. 807, European Centre for Medium Range Weather Forecasts, Reading, UK, [http://www.gaia-clim.eu/system/files/publications/17551-assessment-different-radiosonde-types-20152016.pdf], 2017.

John, V. O. and Buehler, S. A.: Comparison of microwave satellite humidity data and radiosonde profiles: A survey of European stations, Atmos. Chem. Phys., 5, 1843-1853, doi:10.5194/acp-5-1843-2005, 2005.

Karl, T. F., Derr, V. E., Easterling, D. R., Folland, C. K., Hofmann, D. J., Levitus, S., Nicholls, N., Parker, D. E., and Withee, G. W.: Critical issues for long-term climate monitoring, Climatic Change, 31(2/4), 185-221, doi:10.1007/BF01095146, 1995.

Kishore, P., Venkat Ratnam, M., Namboothiri, S. P., Velicogna, I., Basha, G., Jiang, J. H., Igarashi, K., Rao, S. V. B., and Sivakumar, V.: 2011. Global (50°S–50°N) distribution of water vapor observed by COSMIC GPS RO: Comparison with GPS radiosonde, NCEP, ERA-Interim, and JRA-25 reanalysis data sets, J. Amos. Sol-Terr. Phy. 73, 1849–1860, doi:10.1016/j.jastp, 2011.04.017, 2011.

Kitchen, M.: Representativeness errors for radiosonde observations, Q. J. Roy. Meteor. Soc., 115, 673-700, doi:10.1002/qj.49711548713, 1989.

Kitchen, M.: Compatibility of radiosonde geopotential measurements, IOM Report No. 36, WMO/TD No. 344 [available at https://library.wmo.int], 1989a.

Kley, D., Russell, J. M., and Phillips, C.: SPARC assessment of upper tropospheric and stratospheric water vapour, WCRP-No. 113, WMO/TD – No. 1043, SPARC Report No. 2 [available at https://library.wmo.int], 2000.

Kuo, Y.-H., Schreiner, W. S., Wang, J., Rossiter, D. L., and Zhang, Y.: Comparison of GPS radio occultation soundings with radiosondes, Geophys. Res. Lett., 32, L05817, doi:10.1029/2004GL021443, 2005.

Lanzante, J. R., Klein, S. A. and Seidel, D. J.: Temporal homogenization of monthly radiosonde temperature data. Part I: Methodology, J. Climate, 16, 224–240, doi:10.1175/1520-0442(2003)016%3C0224:THOMRT%3E2.0.CO;2, 2003.

Laroche, S. and Sarrazin, R.: Impact of Radiosonde Balloon Drift on Numerical Weather Prediction and Verification, Weather Forecast., 28, 772–782, doi:10.1175/WAF-D-12-00114.1, 2013.

Luers, J. K. and Eskridge, R. E.: Use of radiosonde temperature data in climate studies. J. Climate, 11, 1002–1019, doi:10.1175/1520-0442(1998)011<1002:UORTDI>2.0.CO;2, 1998.

Mastenbrook, H. J. and Daniels, R. E.: Measurements of stratospheric water vapor using a frost-point hygrometer, in Atmospheric water vapor, ed. By Deepak, A., Wilkerson, T. D., and Ruhnke, L. H., Academic Press, pp. 329-342, 1980.

McCarthy, M. P.: Spatial sampling requirements for monitoring upper-air climate change with radiosondes, Int. J. Climatol., 28: 985-993, doi:10.1002/joc.1611, 2008.

McCarthy, M. P., Thorne, P. W., and Titchner, H. A.: An analysis of tropospheric humidity trends from radiosondes. J. Climate, 22, 5820–5838, doi:10.1175/2009JCLI2879.1, 2009.

McGrath, R., Semmler, T., Sweeney, C., and Wang, S.: Impact of balloon drift errors in radiosonde data on climate statistics, J. Climate, 19, 3430–3442, doi:10.1175/JCLI3804.1, 2006.

Miloshevich, L. M., Vömel, H., Whiteman, D. N., Lesht, B. M., Schmidlin, F. J., and Russo, F.: Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation, J. Geophys. Res., 111, D09S10, doi:10.1029/2005JD006083, 2006.

Moradi, I., Soden, B., Ferraro, R., Arkin, P., and Vömel, H.: Assessing the quality of humidity measurements from global operational radiosonde sensors, J. Geophys. Res-Atmos., 118, 8040–8053, doi:10.1002/jgrd.50589, 2013.

Nash J.: Review of test results on the accuracy of radiosonde relative humidity sensors, in Proc. ECMWF/GEWEX Workshop on Humidity analysis, Reading, UK, 8–11 July 2002, pp. 117–123, 2002

Nash, J., Gaffard, C., Smout, R., and Smees, M.: Introduction to upper-air measurements with radiosondes and other in-situ observing systems, Observation Development, Met Office, Exeter, 03-07 September 2007, L'Aquila, Italy, [https://www.wmo.int/pages/prog/www/IMOP/meetings/Upper-Air/ET-IOC-3/Doc3.1(1).pdf], 2007.

Nash, J., Oakley, T., Vömel, H., and Wei, L. I.: WMO intercomparison of high quality radiosonde systems (Yangjiang, China 12 June–3 August 2010), WMO Instruments and Observing Methods Rep. 107, 238 pp., [Available at www.wmo.int/], 2011.

Nash, J.: Measurement of upper-air pressure, temperature and humidity, IOM Report No. 121, World Meteorological Organization [available at https://library.wmo.int], 2015.

Noh, Y.-C., Sohn, B.-J., Kim, Y., Joo, S., and Bell, W: Evaluation of temperature and humidity profiles of Unified Model and ECMWF analyses using GRUAN radiosonde observations, Atmosphere, 7, 94, doi:10.3390/atmos7070094, 2016.

Oakley, T: Report by the Rapporteur on Radiosonde Compatibility Monitoring. Part B – Compatibility of Radiosonde Geopotential Measurements 1990, 1991 and 1992. IOM Report No. 56, WMO/TD No. 587 [available at https://library.wmo.int], 1993.

OFCM – Office of the Federal Coordinator for Meteorology: Federal Meteorological Handbook No. 3 – Rawinsonde and pibal observations, National Oceanic and Atmospheric Administration, Washington, DC, available at https://www.ofcm.gov/publications/fmh/allfmh2.htm, 1997.

Parker, D. E. and Cox, D. I.: Towards a consistent global climatological rawinsonde data-base, Int. J. Climatol., 15, 473–496. Doi:10.1002/joc.3370150502, 1995.

Rieckh, T., Anthes, R., Randel, W., Ho, S.-P., and Foelsche, U.: Evaluating tropospheric humidity from GPS radio occultation, radiosonde, and AIRS from high-resolution time series, Atmos. Meas. Tech., 11, 3091–3109, doi:10.5194/amt-11-3091-2018, 2018.

Ross, R. J. and Elliott, W. P.: Tropospheric water vapor climatology and trends over North America: 1973–93. J. Climate, 9, 3561–3574, doi:10.1175/1520-0442(1996)009<3561:TWVCAT>2.0.CO;2, 1996.

Ross, R. J. and Elliott, W. P.: Radiosonde-based Northern Hemisphere tropospheric water vapor trends, J. Climate, 14, 1602–1611, doi:10.1175/1520-0442(2001)014<1602:RBNHTW>2.0.CO;2, 2001.

Sapucci, L.F., Machado, L. A., da Silveira, R. B., Fisch, G., and Monico, J. F.: Analysis of relative humidity sensors at the WMO Radiosonde Intercomparison Experiment in Brazil, J. Atmos. Ocean. Tech., 22, 664–678, doi:10.1175/JTECH1754.1, 2005.

Seidel, D. J., Sun, B., Pettey, M., and Reale, A.: Global radiosonde balloon drift statistics, J. Geophys. Res., 116, D07102, doi:10.1029/2010JD014891, 2011.

Soden, B. and Lanzante, J.: An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor, J. Climate, 9, 1235-1250, doi:10.1175/1520-0442(1996)009%3C1235:AAOSAR%3E2.0.CO;2, 1996.

Schwartz, B. E. and Doswell, C. A.: North American rawinsonde observations: Problems, concerns, and a call to action, B. Am. Meteor. Soc., 72, 1885–1896, doi:10.1175/1520-0477(1991)072<1885:NAROPC>2.0.CO;2, 1991.

Seidel, D. J., Ao, C. O. and Li, K.: Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis, J. Geophys. Res., 115, D16113, doi:10.1029/2009JD013680, 2010.

Shangguan, M., Matthes, K., Wang, W., and Wee, T.-K.: Validation of COSMIC water vapor data in the upper troposphere and lower stratosphere using MLS, MERRA and ERA-Interim, Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-248, in review, 2016.

Shea, D. J., Worley, S. J, Stern, I. R. and Hoar, T. J.: An introduction to atmospheric and oceanographic data, National Center for Atmospheric Research, Tech. Note 404, 136 pp. [available at https://library.ucar.edu/], 1994

Smit, H, Kivi, R, Vömel, H, and Paukkunen, A: Thin film capacitive sensors, in Monitoring Atmospheric Water Vapour Ground-Based – Remote Sensing and In-situ Methods, ISSI Scientific Report No 10, Niklaus Kämpfer (Ed.), Springer, doi:10.1007/978-1-4614-3909-7, 2013

Stickler, A., Grant, A. N., Ewen, T., Ross, T. F., Vose, R. S., Comeaux, J., Bessemoulin, P., Jylhä, K., Adam, W. K., Jeannet, P. Nagurny, A. Sterin, A. M. Allan, R., Compo, G. P. Griesser, T., and Brönnimann, S.: The Comprehensive Historical Upper Air Network (CHUAN), B. Am. Meteor. Soc., 91, 741-751. doi:10.1175/2009BAMS2852.1, 2010.

Sugita, M. and Brutsaert, W.: Daily evaporation over a region from lower boundary layer profiles measured with radiosondes, Water Resour. Res., 27(5), 747–752, doi:10.1029/90WR02706, 1991.

Sun, B., Reale, A., Seidel, D. J., and Hunt, D. C.: Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics, J. Geophys. Res., 115, D23104, doi:10.1029/2010JD014457, 2010.

Thorne, P. W., Parker, D. E., Tett, S. F. B., Jones, P. D., McCarthy, M. Coleman, H. and Brohan P.: Revisiting radiosonde upper air temperatures from 1958 to 2002, J. Geophys. Res., 110, D18105, doi:10.1029/2004JD005753, 2005.

Vergados, P., Mannucci, A. J., Ao, C. O., Verkhoglyadova, O., and Iijima, B.: Comparisons of the tropospheric specific humidity from GPS radio occultations with ERA-Interim, NASA MERRA, and AIRS data, Atmos. Meas. Tech., 11, 1193–1206, doi:10.5194/amt-11-1193-2018, 2018.

Vömel, H., David, D. E., and Smith, K.: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer (CFH): Instrumental details and observations, J. Geophys. Res., 112, D08305, doi:10.1029/2006JD007224, 2007.

Wade, C. G.: An evaluation of problems affecting the measurement of low relative humidity on the United States radiosonde, J. Atmos. Ocean. Tech., 11, 687–700, doi:10.1175/1520-0426(1994)011%3C0687:AEOPAT%3E2.0.CO;2, 1994.

Wallis, T. W.: A subset of core stations from the Comprehensive Aerological Reference Dataset (CARDS), J. Climate, 11, 272–282, doi:10.1175/1520-0442(1998)011%3C0272:ASOCSF%3E2.0.CO;2, 1998.

Wang, J., Cole, H. L., Carlson, D. J., Miller, E. R., Beierle, K., Paukkunen, A. and Laine, T. K.: Corrections of Humidity Measurement Errors from the Vaisala RS80 Radiosonde – Application to TOGA COARE Data, J. Atmos. Ocean. Tech., 19, 981–1002, doi:10.1175/1520-0426(2002)019<0981:COHMEF>2.0.CO;2, 2002.

Wang, J., Carlson, D. J., Parsons, D. B., Hock, T. F., Lauritsen, D., Cole, H. L., Beierle, K., and E. Chamberlain, E.: Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication, Geophys. Res. Lett., 30, 1860, 16, doi:10.1029/2003GL016985, 2003.

Wenstrom, W. H., 1937: On pilot balloons and sources of light for high altitude upper-wind observations. Mon. Weather. Rev., 65, 326–331, doi:10.1175/1520-0493(1937)65%3C326:OPBASO%3E2.0.CO;2

WMO: Meteorology – A three-dimensional science: Second session of the commission for aerology, in WMO Bull., IV, (4), 134-138 [available at https://library.wmo.int], 1957.

WMO: Commission for Synoptic Meteorology, Abridged Final Report of the Second Session (New-Delhi, 21 January – 15 February 1958), WMO–No. 74, RP. 30 [available at https://library.wmo.int], 1958.

WMO: Commission for Synoptic Meteorology, Abridged Final Report of the Second Session. (Geneva, 15 June – 3 July 1970), WMO–No. 269, RP. 86 [available at https://library.wmo.int], 1970.

WMO: Twenty-Ninth Session of the Executive Committee, Abridged Report with Resolutions (Geneva, 26 May – 15 June 1977), WMO–No. 483 [available at https://library.wmo.int], 1977.

WMO: Thirty-Ninth Session of the Executive Council, Abridged Report with Resolutions (Geneva, 1-5 June 1987), WMO–No. 682 [available at https://library.wmo.int], 1987.

WMO: WMO Manual on Codes, vol. I. , World Meteorological Organization, Geneva, 1988.

WMO: Calculation of Monthly and Annual 30-Year Standard Normals, WCDP-No. 10, WMO-TD/No. 341 [available at https://library.wmo.int], 1989.

WMO: Radiosonde humidity sensor intercomparison, Phase II – Field Test, Laboratory for Hydrospheric Processes, Goddard Space Flight Center, Wallops Flight Facility, Wallops island, Virginia, USA, 8–26 September 1995, [Available at https://wmo.int/], 1995.

WMO: Guide to Meteorological Instruments and Methods of Observation, 6th ed, WMO Rep. 8, World Meteorological Organization, Geneva, 1996.

WMO: Guide to the GCOS Surface and Upper-Air Networks: GSN and GUAN, WMO/TD No. 1106 [available at https://library.wmo.int], 2002.

WMO: Report of the Third GCOS Reference Upper Air Network Implementation and Coordination Meeting (GRUAN ICM-3), WMO/TD No. 1575 [available at https://www.gruan.org/articles], 2011a.

WMO: Guide to climatological practices, WMO No. 100., 2011 edition, World Meteorological Organization, Geneva, [Available at www.wmo.int/], 2011b.

Zaitseva, N. A.: Historical developments in radiosonde systems in the former Soviet Union. B. Am. Meteor. Soc., 74, 1893–1900, doi:10.1175/1520-0477(1993)074%3C1893:HDIRSI%3E2.0.CO;2, 1993.

Zhai, P. and Eskridge, R. E.: Atmospheric water vapor over China. J. Climate, 10, 2643–2652, doi:10.1175/1520-0442(1997)010<2643:AWVOC>2.0.CO;2, 1997.

## LIST OF FIGURES 1–11

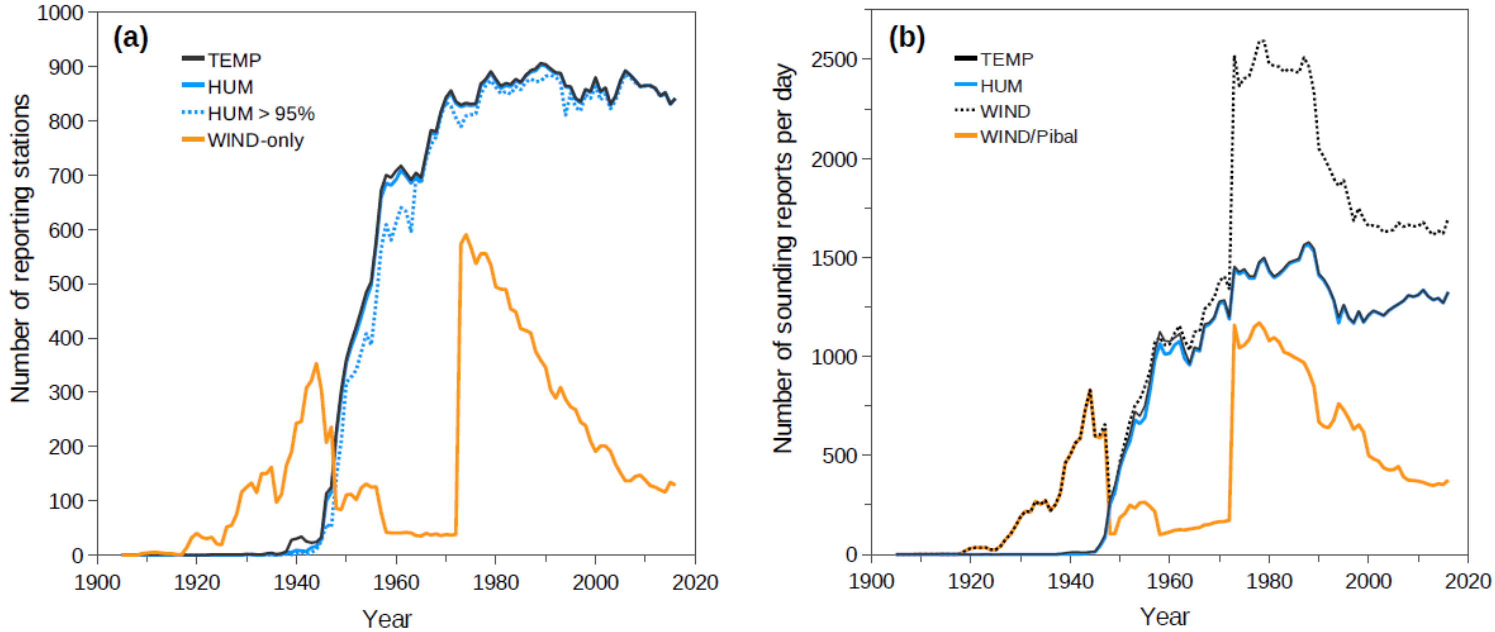


**Figure 1**

**(a) Number of IGRA stations, for each year until 2016, reporting: ~~any amount of radiosonde~~ temperature observations, regardless of humidity and wind (~~RAOB~~TEMP);~~ only wind observations (PIBAL);~~ ~~any~~ temperature and humidity observations, regardless of wind (~~Hygrom.~~HUM); humidity ~~data~~ in at least 95 % of the RAOB ~~soundings~~ (~~Hygrom.~~HUM > 95 %); only wind observations (WIND-only). (b) Number of ~~IGRA~~ sounding reports per day compiled in IGRA, by atmospheric parameter: temperature (TEMP); humidity (HUM); wind from radiosonde or pilot-balloon measurements (WIND); wind ~~only~~ from pilot balloons ~~data~~ (WIND/Pibal). Note: in panel (a), the sum ~~PIBAL + RAOB~~TEMP + WIND-only gives the global number of stations with any data in IGRA~~; considering data gaps due to interruptions of operation or data recording at individual stations, the yearly number of active stations is somewhat larger~~.**

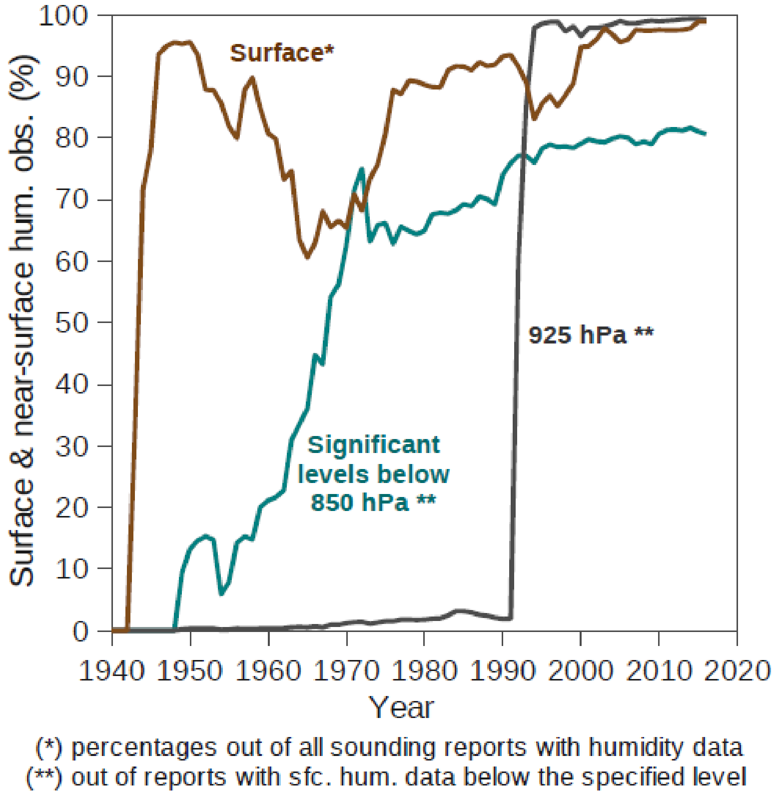


**Figure 2**

**Percentage of humidity observations: at the surface level out of all humidity soundings in each year; at 925-hPa, out of the soundings with surface data below that level; and at any non-standard level – excluding the 925-hPa, irrespective of the year – between the surface and the 850-hPa level, whenever this is above the surface.**

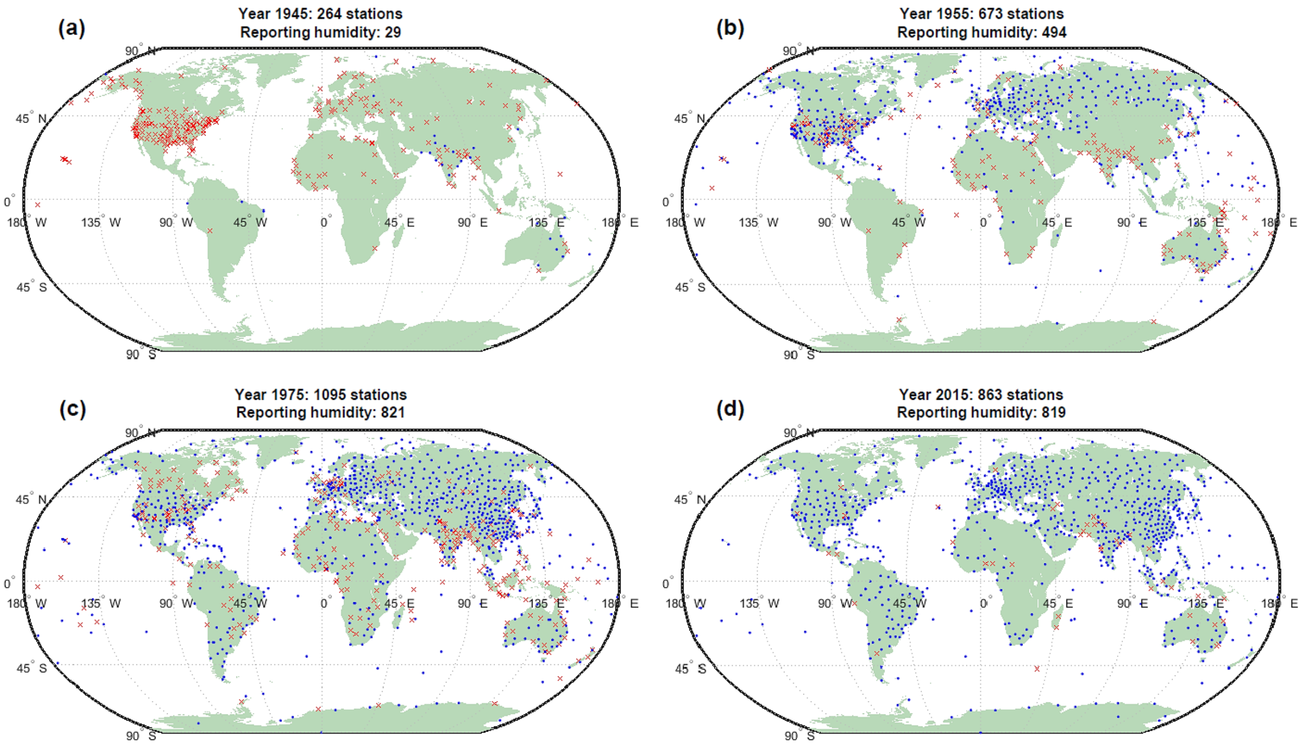


**Figure 3**

**Locations of IGRA-RS fixed stations active in the years of (a) 1945, (b) 1955, (c) 1975 and (d) 2015. Note that the interval of time doubles between consecutive panels. Blue dots denote the stations reporting humidity (RH/DPD + temperature) anytime in the year; red crosses denote stations reporting only temperature/wind, or else presenting a data record gap, during the same year. The total number of stations (crosses + dots) refers to the active stations in each year regardless of any data record gaps.**

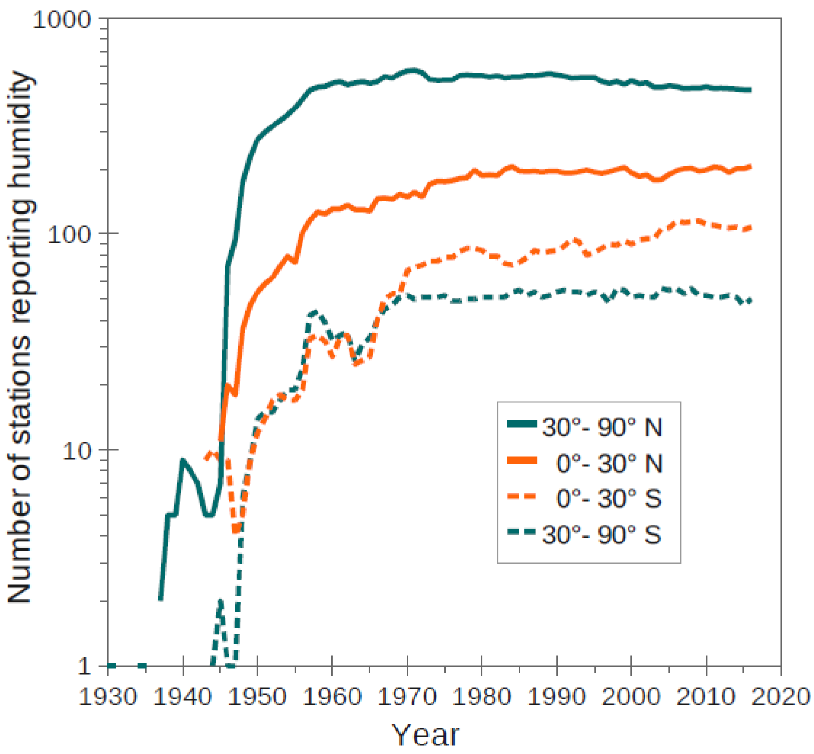


**Figure 4**

~~Yearly number of fixed stations reporting humidity observations in each climatic zone, since the time of the earliest humidity data in IGRA.~~

Yearly number of IGRA-RS fixed stations reporting humidity in northern and southern tropical and extratropical latitudes, since the time of the earliest radiosonde humidity observations.

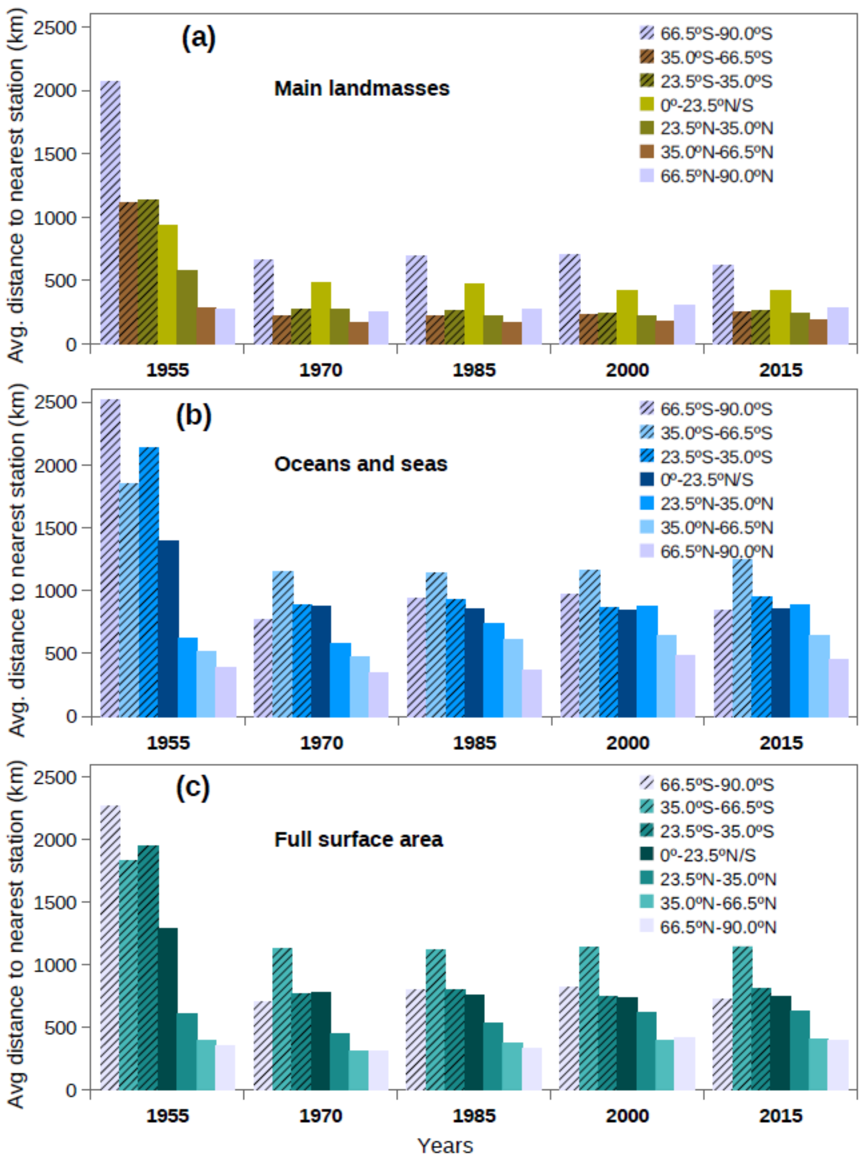


**Figure 5**

**Average distance to the nearest station, among the humidity-reporting stations from 1955 to 2015 in 15-year intervals, as calculated on specific areas of each climatic zone: (a) Main landmasses; (b) Oceans and seas; (c) Full surface area. See text for calculation details (Sec. 2.3.1).**

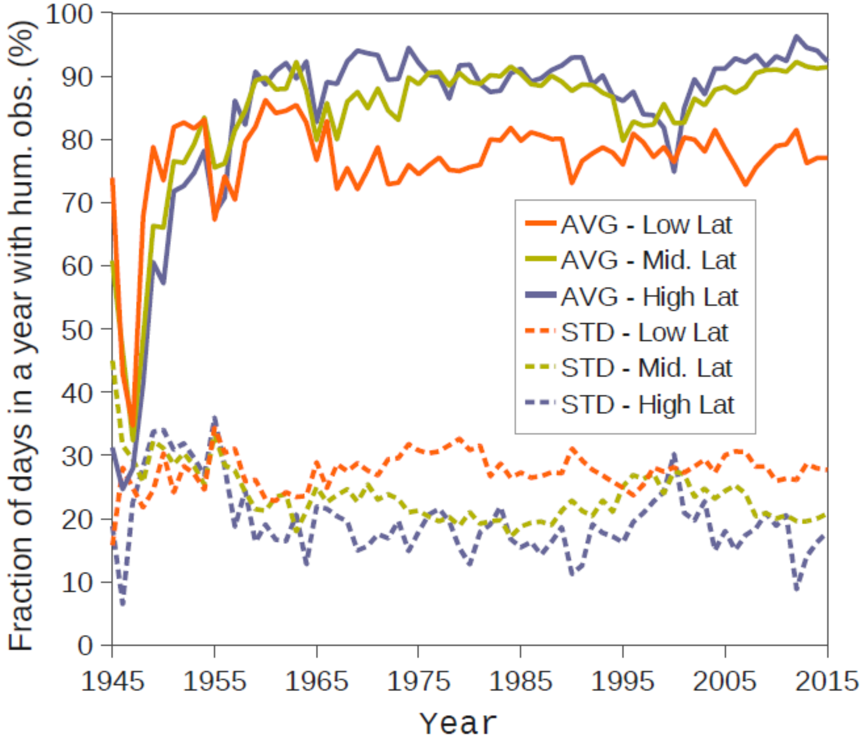


**Figure 6**

**Statistical measures of the fraction of days in a year with humidity observations from 1945 to 2015, representing the fixed radiosonde stations over the major latitude regions for both hemispheres: average among stations with humidity data, and standard deviation. Low Lat – abs. latitudes < 30°; Mid. Lat – abs. latitudes between 30° and 60°; High Lat – abs. latitudes > 60°.**

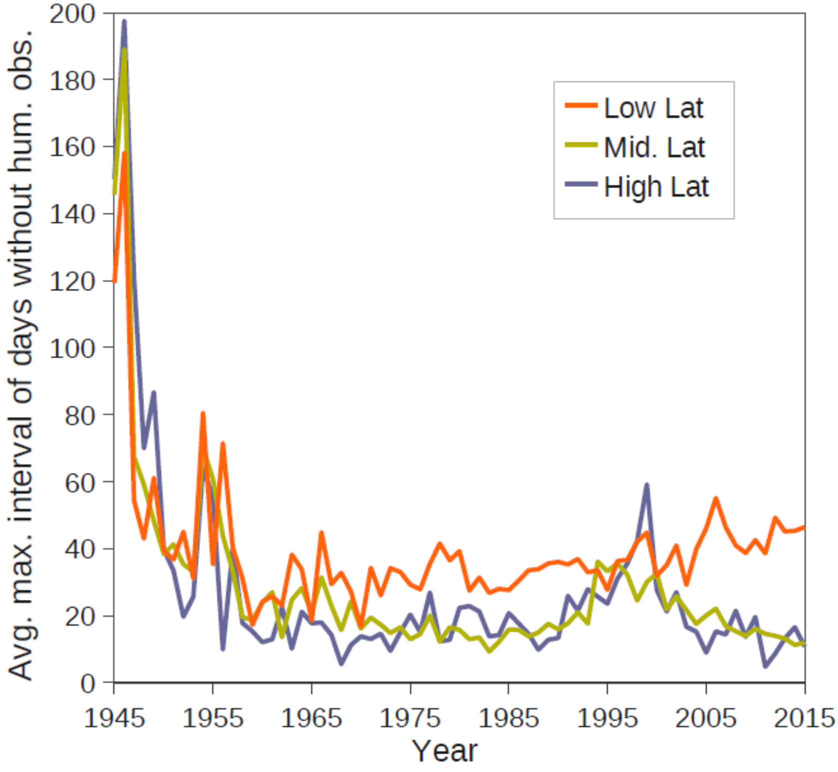


**Figure 7**

**Maximum number of consecutive days without humidity data in each year with some humidity observations ('size of missing days'), averaged among the fixed stations within each of the major latitude regions defined as in Fig. 6.**

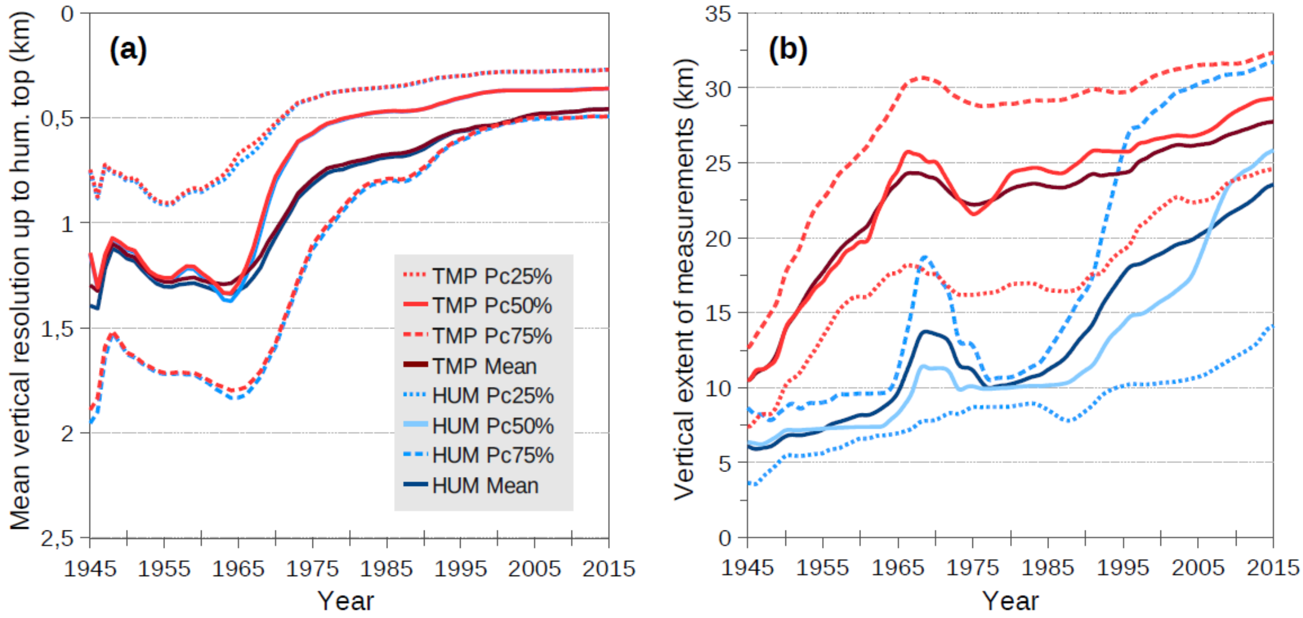


**Figure 8**

**(a) Distribution of the 'mean vertical resolution' (definition of Eq. (2)) in the temperature (TMP) and humidity (HUM) observations across the globe, calculated up to the highest measuring level for humidity: Mean and quartiles. (b) As in (a), but for the vertical extent of TMP and HUM observations, restricting to observations with surface temperature data. All the curves are smoothed by a 5-year running mean.**

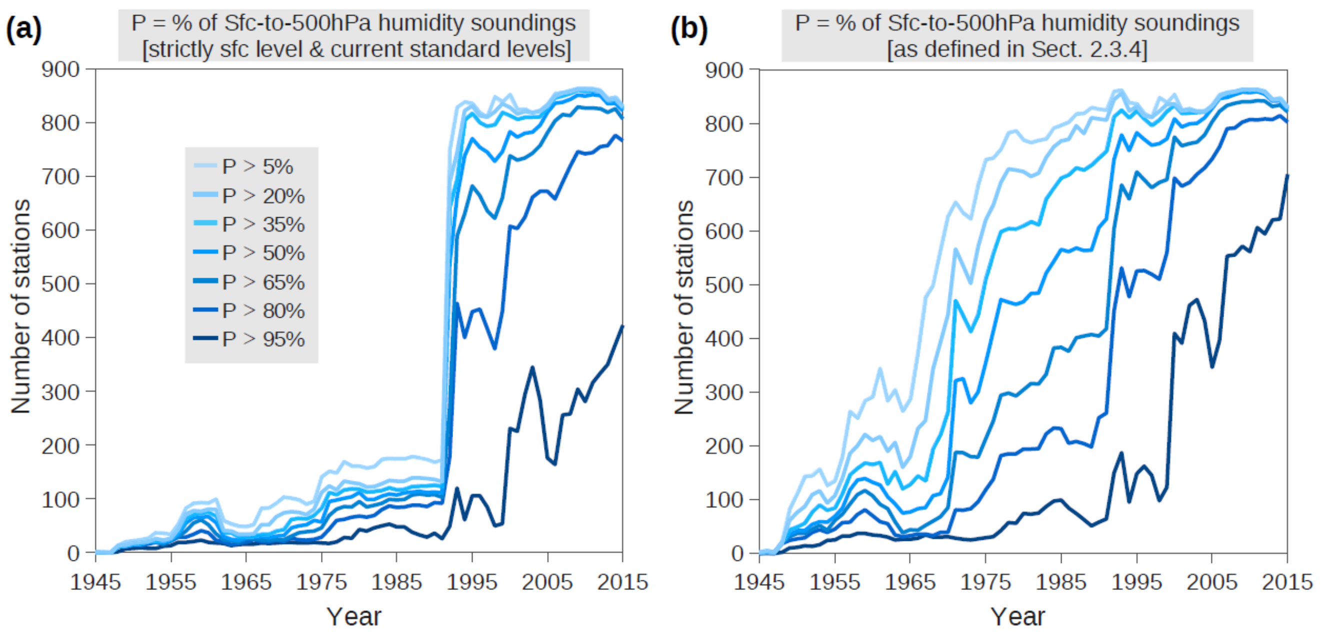


**Figure 9**

**Number of stations for which the percentage of Sfc-to-500hPa humidity soundings, out of all soundings with humidity in each year, exceeds a given value (see inset of panel a): (a) Requiring humidity data at the surface level and all current standard levels above the surface up to 500 hPa; (b) Using the definition of 'Sfc-to-500hPa humidity soundings' given in Sect. 2.3.4.**

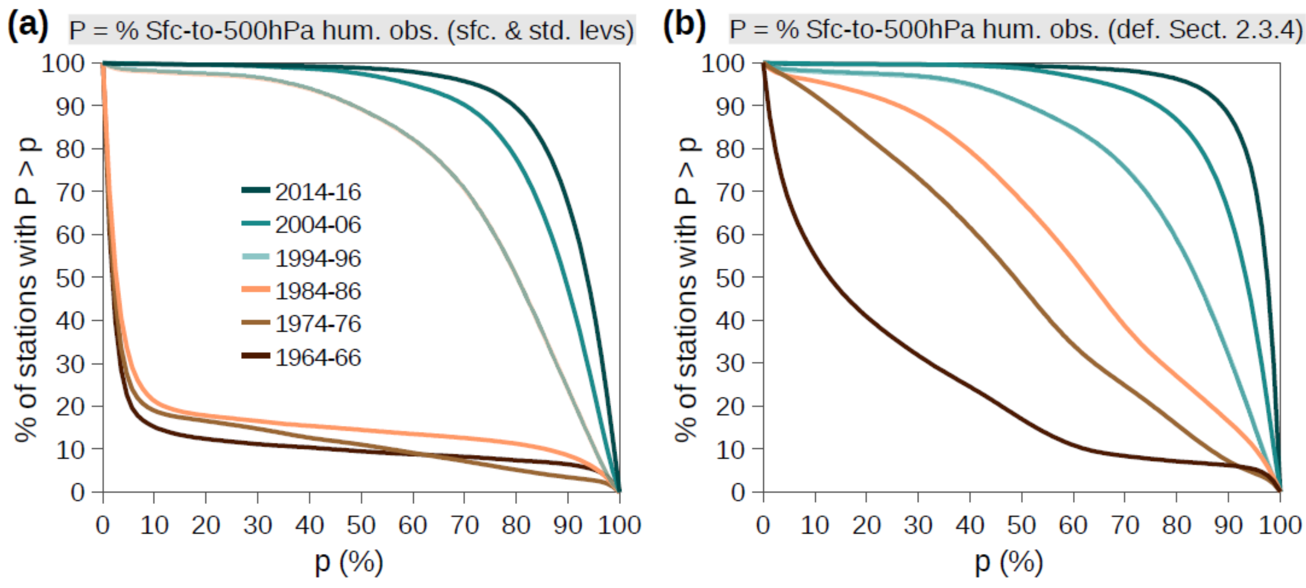


**Figure 10**

**Percentage of stations, out of all IGRA-RS stations, with a percentage of Sfc-to-500hPa humidity soundings exceeding a given value (shown in abscissa), out of the humidity soundings in several trienniums with 10 years interval (see inset of panel a): (a) Requiring humidity data at the surface level and all current standard levels above the surface up to 500 hPa; (b) Using the definition of 'Sfc-to-500hPa humidity soundings' given in Sect. 2.3.4.**

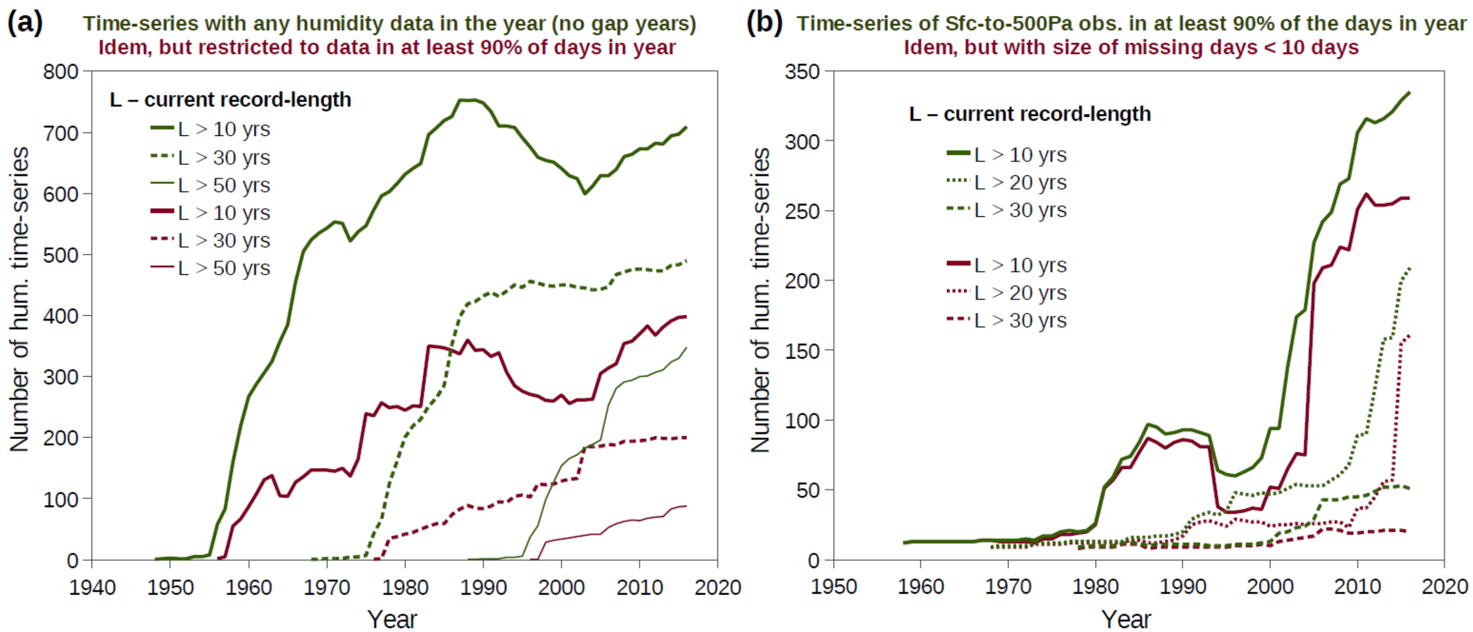


**Figure 11**

**Number of humidity time-series with a record length exceeding a given number of years until the year in abscissa (see insets), under different completeness criteria (indicated on top of the graphs).**

## LIST OF TABLES 1–3

**Table 1. Number of soundings in IGRA and IGRA-RS subset until the end of 2016**

| | **IGRA** (2761 stations) | **IGRA-RS** (1723 stations) |
|---|---|---|
| Total (PIBAL + RAOB) | 45,677,409 | 39,526,638 |
| PIBAL (*wind-only data*) | 15,463,235 | 9,312,891 |
| RAOB | 30,214,174 | 30,213,747 |
| RAOB with humidity data | 29,801,708 | 29,801,324 |

**Table 2 Parameters of humidity completeness for each IGRA-RS station and year**

| **Parameter** | **Description** (*Statistics refer to a year within the period of record of the station*) |
|---|---|
| HUMa | # of soundings with humidity observations |
| HUMb | # of soundings with Sfc-to-500hPa humidity observations[1] |
| RESa | Annual mean vertical resolution of humidity observations[2] |
| RESb | Annual mean vertical resolution of Sfc-to500hPa humidity observations[3] |
| TOPP | Annual geometric mean pressure of the highest level with humidity data |
| GAPa | Largest interval of days in the year for which humidity observations are missing |
| GAPb | Largest interval of days in the year for which Sfc-to-500hPa hum. obs. Are missing |
| FDYa | Fraction of days in a year with humidity observations |
| FDYb | Fraction of days in a year with Sfc-to-500hPa humidity observations |

(1) As defined in Sect. 2.3.4. (2) Arithmetic mean of the ‘mean vertical resolution’ given by Eq. (2).
(3) Calculated as RESa, but between the surface and 500 hPa.

**Table 3 Parameters of humidity completeness for each sounding**

| Parameter | Description |
|---|---|
| RAOB | Classifies the sounding according to RAOB data content:<br>0 = none (only wind data); 1 = TEMP data, but missing HUM;<br>2 = TEMP and HUM data; 3 = TEMP and Sfc-to-500hPa HUM(1) |
| RESa | Mean vertical resolution of humidity data values[2] |
| RESb | Mean vertical resolution of Sfc-to-500hPa humidity data[3] |
| TOPP | Atmospheric pressure of the highest level with humidity data |
| TOPZ | Altitude above mean sea level of the highest level with humidity data |

([1]) As defined in Sect. 2.3.4. ([2]) Calculated by Eq. (2).

([3]) Calculated as RESa, but between the surface and 500 hPa.