07-02-2019

Dear David Carlson,

I am pleased to submit the revised version of essd-2018-95 ‘Completeness of radiosonde humidity observations based on the IGRA’ for peer-review completion. In revising the manuscript, I and my coauthors have carefully considered the reviewers’ comments. We must acknowledge that the revised version has benefited much from them.

Please note that the former authors’ reply to referee comments (AC1, AC3) already contained a point-by-point response along with a detailed description of the authors’ changes in manuscript. Therefore, the authors´ responses to the Referees are copied below without editing *save for the following amendments*:

→ regarding changes that were just summarized as “Planned changes in the manuscript”, we now refer their accomplishment, denoted by “Actual changes” and highlighted in grey.

→ additional changes to fully address the reviewers’ comments are denoted by “Further changes”, also highlighted in grey.

→ a few minor corrections to the previous responses are highlighted in yellow.

Indications of page and line numbers, as well as figure numbers, refer, of course, to the discussion version of the manuscript, unless stated otherwise. The marked-up revised manuscript is provided after the responses to Referees.

Sincerely,

António P. Ferreira

Authors' Response to Referee #1 Comments on

"Completeness of radiosonde humidity observations based on the IGRA" by António P. Ferreira, Raquel Nieto, Luis Gimeno

Referee comments (highlighted in blue) are copied before the authors' responses.

The manuscript by Ferreira et al. presents an analytical description and a rationalized statistical analysis of the radiosounding data archive available through the NOAA-IGRA initiative. The authors report an extensive and precious description of a huge number of information about the changes occurred in the number, geographical density and type of radiosoudings and payloads used around the globe since the beginning of last century to present for the measurement of atmospheric humidity. The authors focus they study on humidity measurements and they investigate the datasets in a critical way to show if, in term of continuity and coverage, the available data may support climate studies, though they do not assess this aspect on a scientific sound basis. The paper is well written and curated and the provided analysis and it update expected on a biannual basis may support many activities currently ongoing, not ultimately the Copernicus Climate Change Service (C3S). I am in favor of the manuscript publication. Nevertheless, I have to provide the authors with major revisions, because to my opinion a few concepts must be clarified and modified in the text of the manuscript, though this does not affect my positive view on the manuscript itself. Below I enclose my general comment and a bunch of minor issues to solve.

General comments

1. In the manuscript, starting form the abstract, it is mentioned that the IGRA V2 datasets is investigated according to various completeness criteria in order to show if the length of the available data records is sufficient for the purposed to estimate climate variability trends. The authors mainly refer to the temporal continuity of the observation and not the spatial gaps which are anyhow investigated in the statistics. Despite this purpose, the authors clarify in the conclusion that their work can only be supportive of climate studies or works related to the climate homogenization of the time series, and nor they do not discuss on a quantitative basis what is the length of data records required for climate studies neither they refer to past literature to asses this aspect. Assuming that the authors are not interested in the going ahead in the assessment of the effect of gaps in the data records on the estimation of climate signals (which is well beyond the scope of the paper), I'd ask them to change the tone of the manuscript, where needed, to better clarify the scope of the manuscript.

The abstract states the paper´s purpose: "The sounding data compiled in the Integrated Global Radiosonde (…) are examined here until the end of 2016, aiming to describe the completeness of humidity observations from radiosondes in different times and locations" [P10, L10-12], further expressing the expectation 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 according to various completeness criteria – even if differences in instrumentation and observing practices require extra attention" [P1, L25-27]. Similarly, the introductory section states: "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." [P3, L14]. In short, our main purpose is not exactly to show if the available data may support climate studies. Still, the interest of our work to climate studies is stressed in several parts of the manuscript. The next comment summarizes how climate studies are addressed in the body of the text.

The use of radiosonde data for climate studies is mentioned in the Introduction (P2, L25-29) because radiosonde archives provide the longest atmospheric humidity records available. While the notion of climatic variability may refer to different time scales, it usually involves a few decades (to define a reference mean state). Concerning climatic trends, the longer and stable the time series, the better. Therefore, Section 2.3.4 draws attention to the importance of accessing large interruptions in long term humidity time series, by identifying the time series without gap years and measuring the continuity of data within each year. The availability of multidecadal humidity time series of upper-air humidity is illustrated in Figure~~s 9 and~~ 10 of Section 3.4, using different completeness requirements for time series without gap years (concerning the fraction of observing days in a year, the maximum size of gap days and the vertical sampling in the lower troposphere). As explained in that section, "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." [P20, L4-6]. Fig~~s~~. 10~~-11~~ simply indicates how much the availability of long term humidity time series can depend on specific data completeness requirements, besides the length of the period of record of course. We think that the summary results in the concluding Sect. 6 (last bullet paragraph on P23, L1-8) are sufficiently objective in this respect. The paper ends with a suggestion on the selection of time series for studying the homogenization of humidity data.

Regarding the availability of data for climate studies, the abstract reads: "For illustration, the study presents a global picture of the completeness of radiosonde humidity observations over the years, including their latitudinal coverage. This overview shows that the number of radiosonde stations having a long enough record length for studies on the climatic variability and trends of humidity-related quantities depends critically on the temporal continuity, regularity and vertical sampling of the humidity time-series." [P1, L22-25]. We understand that the latter sentence can give the wrong impression that our paper quantifies the completeness *requirements* for performing climate studies on humidity-related quantities. This misperception can be briefly solved. We further understand that the Short Summary appearing on the ESSDD publication can be slightly misleading about the scope of the paper. This can be changed too.

Changes in the Abstract:

Paragraph on P1, L23-25:
This overview shows that the number of radiosonde stations having a long enough record length for studies on the climatic variability and trends of humidity-related quantities depends critically on the temporal continuity, regularity and vertical sampling of the humidity time-series.

It will be changed as follows:
This overview indicates that the number of radiosonde stations having a record length potentially long enough (multidecadal) for climate studies involving humidity-related quantities depends not only on the temporal range, but also on the continuity, regularity and vertical sampling of the humidity time-series.

Amendment to the Short Summary:

Where it reads:
The work shows that the potential use of radiosonde humidity data for climate studies depends on the continuity, regularity and vertical sampling of long time series.

It should read:

The work illustrates how humidity data potentially available for climate studies depend on the length, continuity, regularity and vertical sampling of time series.

Further changes:

Following the rearrange of Sect. 3 [see bellow response to point 3 of general comments], Sect. 3.4 discussing the availability of multidecadal humidity time series was retitled and numbered as Sect. 3.5. The related Fig. 10 was renumbered as Fig. 11.

2. The authors make use of IGRA data V2, which is the most updated version of IGRA which embed several improvements compared to the previous data version (V1). Nevertheless, in my personal researches I had a chance to find many bugs in the IGRA V2 where many data present in the V1 are missing and this is not dependent on the extended quality check applied within the IGRA V2. I had also a chance to report this bug to Imke Durré (PI of IGRA). I got similar feedbacks from other EU and US colleagues during discussion at various workshops. The station of Lindenberg in Germany (WMO index=10393), so accurately described in its past history in the manuscript, is one those affected by this issue (at least until a few weeks ago, my last access to IGRA). Therefore, I am wondering whether a comparison between IGRA V1 and V2 has been carried out and if gaps have been found and fixed somehow in the datasets investigated in the manuscript.

Such apparent data loss during the processing of data sources in IGRA 2 is concerning. We have used IGRA 2 as is, because of the extended data coverage and improvements of quality assurance on humidity data. On the positive side, any future corrections to sounding data in IGRA will translate into the updated versions of the meta-dataset introduced in the present paper.

3. The description of statistical analysis of IGRA V2 data is very extensive way, providing several details and a long description of each figure and table. When reading, I have been very interested by the content, though sometimes the reader may get tired by the way the manuscript is written. In a way similar to the conclusions, I'd suggest to the authors to change the style of their writing and privilege a description in "bullets" to describe the results whenever possible. this will help also to clarify the text itself. For example, for the relative humidity observations, At pag.5, it is reported that the average number of non-standard levels in weather-balloon sounding reports increased from about zero by 1945 to about 30 by 2000, but later in the text it is said that RH observation were already available since the 1930, and again later on (pag. 18) it is stated that RH are becomes more abundant since 1949. Though all of this information are exact the reader may get confused and their comparability and usefulness could be limited if the statical analysis is not described in a more schematic way.

We believe that the largest sections (Sect. 2 and 3) will become much easy to read once the tables and figures are placed besides the text. Please note that paragraphs were used extensively to organize ideas around themes. Also, number bullets were used four times before the concluding section. Data analysis is schematized in the following way: Sect. 2.1, 2.2 and 2.3 describe, respectively, the IGRA sounding data set used in our work, the identification of mostly-radiosonde data series, and the main data analysis needed to compute the parameters used to describe 'completeness of humidity observations'. Figures 1 and 2 are auxiliary to Sect. 2.1 and 2.3. Each subsection of Sect 2.3 ('Analysis of humidity data') introduces the corresponding completeness parameters and ends with a description of the secondary statistical analysis used later in Sect. 3 ('Overview on the completeness of radiosonde humidity observations' – in which the several completeness measures are graphically illustrated by Figs. 3-10. The introductory part of section 3 and several parts of subsections 3.1 to 3.4 recall the data and methods used to draw the related figures, each time they are first mentioned in the text.

On P8 (not P18), L2-4, what we have said is that *until 1969 humidity was reported only in the form of RH* (not that RH data became more abundant since 1949), as documented in IGRA's Readme file. This is unrelated with the fact that the earliest upper-air humidity observations are from 1930 [see P5, L28; P9, L16]; or that in 1945 the average number of significant levels in weather-balloon reports was nearly zero [see P5, L15] – meaning that standard pressure levels were virtually the only ones reported by then. None of this information has to do with the data analysis explained later in Sect. 2.3.

Further changes:

Concerning Sects. 2 and 3, we hope that the following changes bring more clarity to the statistical analysis. First, the material of Sect. 2.3 and Sect. 3 – which now both have five subsections instead of four – was slightly rearranged to separate issues of temporal completeness from other issues. Second, the introductory part of Sect. 2.3 and the numbering of figures in Sect. 3 were modified accordingly.

On P5 it is clarified that the number of vertical levels in Durre (2006) refers to IGRA v1

4. The use of the metric present in the Eq.1, "zonal coverage index" at pag 11 is not clear to me. Have been it used in the past and its adde value with respect to other metrics was shown? What's the added value with respect to a station density per 1000 km, for example? Has the ocean surface been excluded from the global surface, given that this can be calculate more clearly for fixed stations? There are many concerns to me on the use of this index, which requires clarification from the authors. Personally, I think that the user can make use of a much simpler index or statistics (a few of these are used later on in the manuscript by the authors themselves) to show if a zonal belt are under and over- represented. This is also depending on the different atmospheric circulation occurring in the different zonal belts, so I am wondering what's the usefulness of adopting this metric. I ask the authors to clarify in the manuscript of the added value due to the use of this metric.

The "zonal coverage index" index defined by Eq. 1 was intended to compare the fraction of stations among in different climatic zones, normalized by the surface area fraction. To our knowledge it was never used before. [Note: by mistake, the modulus term in Eq. (1) is not raised to -1, as it should be]. Although that index was used to construct Fig. 4b, note that it does not form part of the parameters represented in the dataset introduced in the paper.

We agree that the average station density in different latitude belts – or alternatively the average spacing of stations – would give a more direct information about the data coverage. Moreover, the plot of Fig. 4a representing the fraction of humidity-reporting stations at different latitude bands is not essential, since Fig. 3 (which uses yearly absolute numbers) gives a detailed picture of the same distribution. Therefore, the index in question can be replaced by a more familiar metrics, as suggested in the Referee comment, without affecting the dataset introduced in the paper.

Planned changes in the manuscript:

Section 2.3.1 will be fully revised, with the purpose of replacing Eq (1) by a parameter describing the average separation between adjacent stations in a given zonal band. Separate calculations will be presented for land and ocean areas. The **plots in Figure 4a and 4b** will be replaced to represent both regions. The **comments to Fig. 4 in Sect. 3.1** and **the first bullet paragraph of Sect. 6** will be modified accordingly (but retaining the essential about the relative amount of stations in each climatic zone as shown in Fig. 3).

Actual changes:

The full revision of Sect. 2.3.1, as sketched above, involved the addition of a new figure in Sect. 3.1 showing the locations of IGRA-RS stations: this is the newly Fig. 3 [see p. 43 of the marked-up manuscript]. Consequently, former Fig. 3 was renumbered as Fig. 4. In addition, the former Fig. 4 was redesigned according to the major change in Sect. 2.3.1 and the replacement of Eq. (1), being now Fig. 5. Figure 5 was renumbered as 6; Fig. 9 was moved and is now Fig. 7; Figures 6, 7, 8, 10 were thus renumbered as 7, 8, 9, 11.

NB – The rearrange of the subsections of Sect. 2.3 [see response to point 3 above] implied a similar rearrange in the subsections of Sect. 3 [idem]as well as in the bullet paragraph of Sect. 6.

5. The average vertical resolution could be useful information, but to my opinion, thinking about different applications it could be useful to have a statistic of the available level for the different regions of the atmosphere: Planetary Boundary Layer, Free Troposhere, Upper Troposphere/Lower Stratosphere (UT/LS), Upper Stratosphere. This classification may have a stronger impact to orienting users' application in the selection of the available data, e.g for trends calculation.

Figure 6 only shows the distribution of vertical resolution among humidity observations over time. The dataset introduced in the discussion paper, however, contains four related parameters for each IGRA-RS station:

i. Annual mean vertical resolution of yearly humidity obs. from each station
ii. Annual mean vertical resolution of yearly Sfc-to-500hPa humidity obs. from each station
iii. Vertical resolution of individual humidity obs. from each station
iv. Vertical resolution of individual Sfc-to-500hPa humidity obs. from each station

where 'resolution' stands for geometric mean resolution. Parameters (i) and (iii) refer to reported humidity data up to the highest measuring level. Parameters (ii) and (iv) refer to humidity data between the surface and the 500-hPa level, requiring humidity data at the surface and at a minimum of upper-air levels to estimate precipitable water vapor. Therefore, the dataset accompanying the paper provides specific information on the vertical resolution at the lower troposphere.

Moreover, the dataset also provides information on the highest measuring level:

v. Annual geometric mean pressure of the highest level with humidity data at each station
vi. Pressure of the highest level with humidity data in individual obs. at each station
vii. Altitude of the highest level with humidity data in individual obs. at each station

where (vii) is obviously only calculated if humidity is not missing at the surface level. Such information allows to select the soundings with humidity observations reaching the UT and the LS, as well as the stations and years that normally have observations in the UT/LS, even if the vertical resolution at those regions is not detailed.

Considering the parameters describing temporal completes of humidity observations [periods of record for humidity; fraction of observing days in a year; largest number of consecutive missing days], analysing vertical resolution for each separate atmospheric layer (PBL, lower/middle troposphere, UT, LS) – each one possessing its own temporal completeness – would imply extending the work beyond what might be expected from a first approach. Undoubtedly, a recommendation for future research.

6. Please check the use of the term "error" throughout the manuscript and replace it with "uncertainties" where more appropriated.

The term 'error' is used only twice in the manuscript: "Concerning humidity, radiosonde-based climatic studies are for now confined to the lower and middle troposphere, because of the large errors and insufficient data in the upper troposphere and lower stratosphere." [P2, L25-26] "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 amounts to less than 1 %." [P12, L28-29]

In the first case, we refer to measuring errors related to the poor accuracy of humidity sensors in the upper troposphere (UT) and lower stratosphere (LS). Given that different sensors have different accuracies under the same conditions, equal instruments behave differently in different conditions, and climate studies use data averaging from many observations, we must accept that the expression 'uncertainty' is more appropriate. Moreover, missing data arising from varying observation practices (low RH reporting, low-temperature cutoff) introduce uncertainties in the time series and their averages. This affect particularly the measurements above 500 hPa. (On passing, we should refer the issue of observation practices right away on P2, L22.) However, systematic biases in the UT are well known for certain radiosonde models. Thus, to include both situations, we would say 'uncertainty of measurements and biases'. In the LS, most measurements show a wet RH bias which is very large in relative terms. We think it's also adequate to add a few citations regarding accuracy and uncertainty of humidity measurements in the UT and LS.

In the second case, "error" is the error in calculating geopotential height; this will be clarified

Changes in the manuscript:

P2, L22
Where it reads

instrument changes and sampling differences

It will read

instrument changes, reporting practices and sampling differences

P2, L25-26

Where it reads:

the large errors and insufficient data in the upper troposphere and lower stratosphere

It will read:

the large uncertainty of measurements and biases in the upper troposphere and lower stratosphere (Elliot and Gaffen, 1991; Soden and Lazante, 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)

P12, L29

Where it reads:

the error amounts to

It will read:

the error in calculating geopotential height amounts to

Additions to the reference list:

Wang, J., D. J. Carlson, D. B. Parsons, T. F. Hock, D. Lauritsen, H. L. Cole, K. Beierle, and E. Chamberlain (2003), 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, doi:10.1029/2003GL016985, 16.

Soden, Brian & Lanzante, John. (1996). An Assessment of Satellite and Radiosonde Climatologies of Upper-Tropospheric Water Vapor. J. Climate. 9. 1235-1250. 10.1175/1520-0442(1996)009<1235:AAOSAR>2.0.CO;2.

Nash, J., T. Oakley, H. Vömel, and L. I. Wei, 2011: 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]

Specific comments

Pag.2, lines 10-12: the authors could mention that recent reanalysis products, for example ERA5 from ECWMF, will improve the 4 times daily frequency of the products up to hourly.

Perhaps, we should rather refer some ECMWF reanalysis as it is now, with a different period of record. This can be amended.

Actual change:

P2, L12

Where it reads:

up to 4 times daily from 1948 (as in NCEP/NCAR products) to present time

It should read:

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

Pag.2, line 13: replace “not to mention : : :..” with “with not negligible : : :..”
Agree. Changed to “with significant”.

Pag.2, lines 17-19: uncertainties due to balloon drifting and observation time are considered negligible citing the publication by Kitchen (1989) as well as the radiosonde profile accuracy. To my opinion these cannot be considered minor, and anyhow if minor or not this is depending on the considered application. First of all there more recent papers by Seidel et al. (2011) dealing with radiosonde balloon drifting. Estimates of elapsed time from balloon launch to various pressure levels, due to vertical balloon rise, have median values increasing from about 5 min at 850 hPa to about 1.7 h at 10 hPa, with ranges of about 20% of median values. Observed elapsed times always exceed those estimated using assumed 5 or 6 m/s rise rates. Regarding the data data accuracy, if we are referring to the ensemble of effect which may alter the sensors’ optimal measurement conditions, like the solar radiation effect to the effect of a sensor time-lag, these have been better quantified for the more recent radiosonde types and are quite relevant for any kind of climate application (see for example the GRUAN quantification by Dirksen et al. 2014). This is also the reason why many scientific groups have developed homogenized dataset of radiosonde data for climate application. For there reason above, I’d reformulate these line in the manuscript and I’d provide more updated references.

We understand that uncertainties related to the observation time and balloon drift are currently recognized to be relevant for data assimilation in models (both in numerical weather prediction and reanalysis), besides impacting forecast verification, satellite validation and climatic statistics, specially at high levels. Kitchen (1989) pointed out its importance for weather forecasting. We appreciate that the referee comment calls for more recent literature to substantiate the subject.

We remove “minor” and improve the whole sentence, adding several citations and new references to make a balance between the different problems mentioned. Dirksen et al. (2014) will be cited on P 6, L30.

Changes in the manuscript:

P2, L17-19
Where it reads:
despite sampling differences among stations and over time, minor uncertainties related to observation time and balloon drift (Kitchen, 1989), and differences in data accuracy – which depends on humidity sensors and varies with measured conditions (e.g., Nash, 2002).
It will read:
despite 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 thein) and differences in humidity data accuracy – which depends on humidity sensors and varies with measured conditions (e.g., Nash, 2002; Sappuci et al., 2005; Moradi et al., 2013; Dirksen et al., 2014).

P6, L30
Insert reference: (Dirksen et al., (2014) and references therein)

Additions to the list of references:

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

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

McGrath, R., T. Semmler, C. Sweeney, and S. Wang, 2006: Impact of Balloon Drift Errors in Radiosonde Data on Climate Statistics. J. Climate, 19, 3430–3442, https://doi.org/10.1175/JCLI3804.1

Sapucci, L.F., L.A. Machado, R.B. da Silveira, G. Fisch, and J.F. Monico, 2005: Analysis of Relative Humidity Sensors at the WMO Radiosonde Intercomparison Experiment in Brazil, J. Atmos. Oceanic Technol., 22, 664–678, doi:10.1175/JTECH1754.1

Wallis, T.W., 1998: A Subset of Core Stations from the Comprehensive Aerological Reference Dataset (CARDS). J. Climate, 11, 272–282, doi: 10.1175/1520-0442(1998)011<0272:ASOCSF>2.0.CO;2

~~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, https://doi.org/10.5194/amt-7-4463-2014, 2014.~~ ← IT WAS ALREADY CITED

Pag.2, line 26: please replace "errors" with uncertainties and then add also that the radiosonde sensors may have a limited sensitive to ppm water vapor concentrations in the UT/LS as one of the main reason because humidity data above 300hPa are unreliable.

Regarding the word replacement, please refer to our response to point 6 of the general comments. The details about limitations of humidity sensors were given in Sect. 1.3. However, we understand that the limited response to low water vapor pressures should be directly mentioned in Sect 1.3.

Change in the manuscript:

P5, L29
Where it reads:
humidity has been always difficult to measure in very cold or dry air.
It will read:
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).

Pag.2, line 32: please change "lag" time "time-lag".
Correct.

Pag.3, lines 19: "the remainder of this section".
Correct.

Pag.3, lines 21: put "available" in between of "levels" and "in radiosonde"
Fine.

Pag3, line 26: ": : :.relevant for the study". Study of what?

It means it is relevant to our study, referring to the subset of IGRA without pilot-balloon stations (as explained in lines 23-25 right above). See rephrasing.

Change in the manuscript:

P23, L26
Where it reads:
from each IGRA station relevant to the study
It will read:
from each relevant IGRA station

Pag3, line 26: " please change "indicated" with "reported"
Okay (referring to P3, L28).

Pag4, lines 7-10: Please put a reference related to the importance of data continuity for climate studies (trends, annual cycles).

Previous literature regarding temporal inhomogeneities in radiosonde time series have mainly focused in instrument changes (responsible for the largest time-varying biases), paying less attention to temporal sampling. Nevertheless, this issue has been assumed relevant to climate studies, as detailed next.

Concerning standards for long-term monitoring of the climate system: Karl et al. (1995) puts the *continuity* and *frequency* of in-situ and other observations – as well as the maintenance of local observations with a long and uninterrupted record – among the "critical issues for long-term climate monitoring".

Concerning temporal homogenization of radiosonde time series for determining long-term temperature and humidity trends: Lanzante et al. (2003), mentioned the interest of "counts of numbers of observations per month as a function of time and by level; these aid in finding sampling biases or less reliable time periods". Moreover, they have computed separate monthly means at 0000 and 1200 UT from 1947-98 CARDS data with the requirement of at least 16 valid values per month. [Since this paper focus on the detection of changing points in temperature time series related to instrument changes and reporting practices, it will also be cited in Sect.1; see P2, L20-23]. McCarthy et al. (2009) [already cited on P7, L8] addresses the problem of "important sampling biases in the raw humidity data, from missing dry observations and missing cold observations"; they also require at least 15 days of observations within a month to calculate monthly mean temperatures.

Concerning the temporal completeness requirements in climate studies on humidity or integrated water vapor: Gaffen et al. (1991), on studying spatial-temporal variability of global tropospheric moisture, discuss temporal sampling, arguing that a minimum of three observations per month is required to obtain an estimate of the monthly mean specific humidity that falls within the 0.1 confidence bands. Using a 1978-85 dataset from 119 stations, they noted that almost 5% of the station months did not meet that criterion. Gaffen and Elliot (1992) [already cited in the manuscript] selected radiosonde stations with at least one observation per day to accurately estimate the seasonal cycle of RH at different locations of the globe. Zhai and Eskridge (1997) used Gaffen´s criterion to study changes in PW over China, further rejecting stations with more than 3 years of data missing at the same level. Ross and Elliot (1996) [already cited in the manuscript], on studying water vapor trends over North America have required at least two months to estimate seasonal anomalies and 10 months to annual anomalies, a month being considering as missing if less than 10 observations during the month were missing.

The manuscript changes to the lines in question consists of rephrasing and inserting the proper references.

Changes in the manuscript:

P4, L9-10
Where it reads:
in addition, the regularity and continuity of humidity profiles are important to trend analysis and to detail annual cycles ~~of humidity or integrated water vapor~~ ← THIS WAS NOT ACTUALLY IN THE MS
It will read:
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)

P2, L23 | Add citation:
Lanzante et al. (2003)

Additions to the list of references:

Gaffen, D.J., T.P. Barnett, and W.P. Elliott, 1991: Space and Time Scales of Global Tropospheric Moisture. J. Climate, 4, 989–1008, https://doi.org/10.1175/1520-0442(1991)004<0989:SATSOG>2.0.CO;2

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

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

Lanzante, J.R., S.A. Klein, and D.J. Seidel, 2003: Temporal Homogenization of Monthly Radiosonde Temperature Data. Part I: Methodology. J. Climate, 16, 224–240, doi:10.1175/1520-0442(2003)016<0224:THOMRT>2.0.CO;2

Pag4, Line 10-12: I tend to disagree with the introductory sentence while I like the authors approach in the manuscript; therefore, I think in this sentence you must report which is needed to investigate the simultaneous spatial and time sub-sampling on the data whatever challenging this might be in the practice.

We think that advancing what is needed to investigate the simultaneous spatial and time sub-sampling on the data is beyond the scope of our work. Furthermore, it depends on the application of radiosonde data. Anyway, knowing the content of a comprehensive radiosonde data archive such as IGRA 2 represents a first step. The present paper might be useful in two ways: first, it gives a general picture of the spatial-temporal distribution of in-situ humidity observations worldwide. Second, it provides a tool (dataset) that helps to select either humidity time-series or individual observations based on their temporal and vertical completeness of humidity (*i.e. of simultaneous observations of pressure, temperature and humidity*). It is left to the user to find the data coverage and spatial continuity from the geographical coordinates of stations [please note the final recommendation on P23, L20-23].

However, we agree that the sentence in question can be improved in the following way. First, a station's period of record cannot be separated from its temporal completeness. Secondly, it is appropriate to cite again the paper of Wallis (1998)* since it represents a major effort towards a selection of radiosonde stations for trend studies on a regional or global scale.

(*) Wallis, T.W., 1998: A Subset of Core Stations from the Comprehensive Aerological Reference Dataset (CARDS). J. Climate, 11, 272–282, doi: 10.1175/1520-0442(1998)011<0272:ASOCSF>2.0.CO;2

Changes in the manuscript:

P4, L10-12
Where it reads:
Although the vertical and temporal completeness of station-based humidity time-series can be treated separately from the issues of spatial coverage and record length of stations, it seems to us that studying the completeness of observations in a global, historical data set of radiosonde observations should address these issues too.
It will read:
Although the vertical and temporal completeness of station-based humidity time-series can be treated separately from the issue of ~~spatial (horizontal)~~ geographical coverage, studying the completeness of observations in a global, historical data set of radiosonde observations should address that issue too. This is particularly true ~~regarding~~ concerning the subsampling of radiosonde stations for studies of atmospheric temperature or water vapor trends on a regional or global scale (Wallis, 1997).

Pag. line 24: pressure in not measured anymore in the most recent radiosonde types (e.g. RS-41 operation since a couple of years); please for completeness you may mention this.

Right, not only in the most recent GPS radiosondes but also in some Russian radiosonde systems which used ground radar and a radiosonde without pressure sensor. Being an exception to common radiosondes, this detail deserves a footnote. As to a reference to modern GPS radiosondes (with or without a pressure sensor), Nash et al. (2011) will be also cited in the former footnote 1 on page 30 [as well on page 2 – please see response to general comment 6]

Changes in the manuscript:

P4, L22-24
Where it reads:
In radiosonde soundings, temperature, relative humidity (eventually dewpoint depression too) and wind speed and direction are measured together with atmospheric pressure; geopotential height is indirectly measured from hypsometric calculations but may be missing in radiosonde reports.
It will read:

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).

---

(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)

P7, L30 [former footnote 1, now footnote 2 according to the above change]
Where it reads:
(Dabberdt et al. 2002)
It will read:
(Dabberdt et al. 2002; Nash et al., 2011)

Additions to the list of references:

Zaitseva, N. A., 1993: Historical developments in radiosonde systems in the former Soviet Union. Bull. Amer. Meteor. Soc., 74, 1893–1900, doi:10.1175/1520-0477(1993)074<1893:HDIRSI>2.0.CO;2.

Nash, J., T. Oakley, H. Vömel, and L. I. Wei, 2011: 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]

Pag.5, line 17-19: this sentence could be a good opportunity to claim for the importance of having high resolution measurements and, therefore, more levels available in the radiosounding report. This is in line with the high resolution BUFR files already flowing to the Met services from more than 100 station worldwide.

The current migration from TEMP to high-resolution BUFR reports will be referred in Sect. 2.1; the fact that BUFR data are not currently available in open archives (e.g. IGRA) should be noted.

Planned change to the manuscript:

P5, L19: Insert sentence about the migration to BUFR messages sent to the GTS, with a reference to Ingleby et al. (2016).

Actual change:

P5, L23 – Inserted text:
The current migration of RAOB reports from alphanumeric (TEMP) to the binary universal form for the representation of meteorological data (BUFR), together with the conversion of radiosondes to generate 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 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.

Addition to list of references:

Ingleby, B., P. Pauley, A. Kats, J. Ator, D. Keyser, A. Doerenbecher, E. Fucile, J. Hasegawa, E. Toyoda, T. Kleinert, W. Qu, J. St. James, W. Tennant, and R. Weedon, 2016: Progress toward High-Resolution, Real-Time Radiosonde Reports. Bull. Amer. Meteor. Soc., 97, 2149–2161, doi:10.1175/BAMS-D-15-00169.1

Pag. 7, line 13: "anything but uniform" can be modified "quite heterogeneous".

Correct. After all, there is substantial uniformity as to the launching time and STD levels.

Pag. 7, lines 22-24: these sentence is the "official" IGRA description, but going through the data the authors may realize that among the 2761 stations, many of them are "nearsurface" stations and not radiosounding station. All the reports are empty (-9999) for many station. These must be clarified and the reported number estiamted in a more precise way.

The NOAA's IGRA web page says, "over 2700", without giving a precise number, since it may increase as new sources are added. Maybe what is lacking here is the time when the data were downloaded (or the date of the IGRA stations list, since the data files were downloaded using a script based on that list). Regarding "near-surface stations", we have checked on this and, based on the dataset complementary to the paper, it seems there are about 120 stations (among the IGRA-RS subset given in Table S1) in which *the first one or two years of record* contains only surface data. This happens mostly during ~~1946-48~~ 1946-49. This can be mentioned in Sect. 2.2.

Planned changes in the manuscript:

P7, L22: Indication of the time when the data (or the stations list) was downloaded.
P10, L18: Mention to only-surface data at the beginning of record at some stations.

Actual changes:

P7, L16 – Inserted word:
IGRA 2 dataset → IGRA2 main dataset
P7, L16 – Deleted/moved text:
comprising over 45 million soundings

P7, L22-23
Where it reads:
taken at 2761 globally and temporally distributed stations (2662 fixed and 99 mobile)
It should read:
taken at over 2700 globally and temporally distributed stations

P7, L24 – Inserted text:
comprising over 45 million soundings from 2761 (2662 fixed and 99 mobile) stations [based on data accessed in September 2017].

P10, L18 – Inserted text:
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.

Pag. 7, line 16: Is the reported typical average accuracy in the troposphere only related to the most recent radiosonde types? Please clarify

(Referring to P8, L16) Recent radiosondes achieve an accuracy of 2% RH, while very old radiosondes have worse accuracy (10%). 5% is roughly the accuracy accepted for the troposphere, the layer that presents the most favourable conditions for measurements. The remainder of the sentence specifies that accuracy can in fact deviate significantly from this round number. At extremely low temperatures is not uncommon to see errors of 15%. Accuracy is compromised in high resolution profiles if the response time turns too long. And then there is precision for RH or DPD. In 1991, Elliot and Gaffen reported that by that time radiosondes presented a precision of about 3.5% RH. But this depends on environment conditions too. Moreover, TEMP reports may have different restrictions for precision over time. For all these reasons, maybe it is wise to withdraw the "typical accuracy", and shorten the whole sentence keeping the main idea (Note: the related references were already cited in other parts, mainly in Sect. 1.3)

Changes in the manuscript:

P8, L15-18

Where it reads:

Note that RH measurements from radiosondes have a typical absolute accuracy of ~ 5 % on average tropospheric conditions. However, accuracy varies 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 (Nash, 2002; Miloshevich et al., 2006; Nash, 2015).

It will read:

Note that the precision and accuracy of RH and DPD data varies 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.

Pag.11, line 1: please put a descriptive reference for GRUAN, I suggest Bodeker et al., 2016 BAMS.
Pag.11 line 3: It is not true that all of the GRUAN sites are transmitting data to the GTS. A few sites are still working to establish this data flow. This is also connected to the sentence at line 9, reporting the fact that not all GRUAN station are present within IGRA.
Pag.11, line 6: The added value of GRUAN is not only to provide data the quality of which should be "above the average", but to provide traceable uncertainties and a fully disclosed data processing described in peer-reviewed literature. Please add more details to show the real added value coming from GRUAN

We will add the suggested reference, clarify that by "All GRUAN sites" we mean the ones present in IGRA, and add something about the real value of GRUAN.

Changes in the manuscript:

P11, L1-3

Where it reads:

Moreover, the IGRA-RS contains 16 stations that form part of the GCOS Reference Upper-Air Network (GRUAN), (half certified and half to be certified according to current GRUAN status), of which eight (half certified too) are also GUAN stations. All GRUAN sites report

It will read:

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

P11, L4 – Insert text:

GRUAN aims to serve as reference to 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.

Addition to list of references:

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

Pag.11, eq.1: see my general comments above.

See our response to general comment 4.

Pag.12, line 18: what's the meaning of the "relevant" in this sentence?

It means of interest, i.e. a vertical level among the levels with non-missing humidity data. It is understood in the context, considering the preceding phrase in the whole sentence "M is the number of levels with humidity data above the lowest level", as well as the previous sentence which presents Eq. (1): "the mean vertical resolution of a single humidity sounding was defined by the geometric mean of $\{dzk\}$ for all levels with humidity data in the sounding profile"

Pag.12, line 19: How did you choose the value of the constant temperature T0, please clarify in the text.

The way Eq. (1) was written, the terms inside the product operator are non-dimensional. To prevent *underflow* in a calculation for high-resolution profiles and limited magnitude range for real numbers (because the logarithm terms are inferior to 1), T_0 should rather be chosen so that $RT_0/g \approx$ *average distance between vertical levels* (but this is precisely what we don't know).

Since we used double precision real numbers, and the number of levels with humidity data is not excessively large in IGRA, the result was insensible to the value used (250 K).

However, the term T_0 is not needed providing that Eq. (1) is reformulated in a way that (i) is mathematically equivalent and (ii) it can be safely implemented as is in a computer programme, without leading to underflow when the number of levels is too large. (Although we could alternatively define geometric mean by the anti-log of the arithmetic mean of log-transformed values, this is not needed.) In the revised and straightforward form, each individual term between curved brackets is raised to $1/M$ inside the product operator (instead of doing the product first and then the M^{th} root, as usually seen in the formal definition of geometric mean).

Changes in the manuscript:

P12, L15 | Reformulation of Eq(1):

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

P12, L19-20 | Delete text:

~~T_0 is a constant temperature in the range of the values found in the troposphere (used in the calculation to avoid overflow in case M is too large),~~

Section 2.3.3: this section refers to the quantification of the number of soundings which can be qualified to estimate precipitable water vapour. I am not sure to what extent this section may really confuse the reader. From one side, I think this is redundant with assessment of other indicators in the manuscript and would add value if then the selected radiosoundings according the criteria reported in this action may ready represent soundings for which an accurate estimation of the water vapor is feasible. i think the authors should clarify that thorn the radiosoundings selected according to the presented criteria allow to calculate an estimation of water vapor content which is the closest possible to the true one given the small number of vertical level available. The accuracy of the calculation of precipitable water vapour for these radiosoundings is anyhow affected by many other aspects: presence of clouds affecting the measurement sensors, homogeneity the water vapor field close to the surface, non-linearity of the water vapor variability along the vertical profile and so on.

We know that the calculation of precipitable water vapour (PWV) is affected by radiosonde data imperfections like defective RH measurements inside (or on the way in/out to) clouds, poor resolution near the surface (where specific humidity is highest and it rapidly varying with height), and possibly insufficient resolution to describe the vertical variations of humidity in the troposphere. Leaving aside wet data biases, our approach deals precisely with vertical resolution.

To our knowledge, the effect of the vertical resolution of in-situ observations on the accuracy of PWV estimation is poorly, if ever, quantified. However, all practical studies try to guarantee that at least the surface and standard levels, normally up to 500 hPa (infrequently, up to 400 hPa) are represented in humidity soundings. This is also the requirement used by NOAA/NCEI in their IGRA-derived data set.

In the Abstract and in other parts of the manuscript, including the concluding section, we clarify that our Sfc-to-500hPa soundings are those with adequate/suitable vertical sampling (coverage and resolution) to *estimate* PWV. Our approach is this: 1) To allow an additional near-surface level to fill the lack of the 925 hPa level before 1992, on the condition that such additional level provide a similar resolution near the surface (firs 2 km); 2) To include soundings in which some standard level is missing, providing that an additional level is given in its close vicinity.

In short, the proposed definition of Sfc-to-500hPa eligible to estimate PWV [P6, L5-13] is more inclusive than the usual requirement; furthermore, it assures that the near-surface resolution is similar to that of current standard levels, regardless of the absence of 925 hPa in a large portion of past radiosonde data. The definition is explained in detail in Sect. 2.3.3, with the aid of Fig.2, and is recalled in a straightforward manner in the beginning paragraph of Sect. 3.3.

We must agree that the wording “eligible to estimate PWV”, as used in the beginning line of Sect 2-3-3, is better than “qualified to estimate PWV”. This applies to the Abstract and to the title of Sect. 2.3.3.

Changes in the manuscript:

P1, L19 and P13, L1: Replacement of “qualified” by “eligible”

Pag.16, line 26: Please provide more details to explain the differences in the maximum range covered by the measurement of these parameters. For example, the way wind and humidity have been measured in the past compared to temperature?

The limited vertical range of humidity measurements was mentioned throughout Sect. 1, and its relationship with the working range of humidity sensors was explained in Sect. 1.3. Regarding wind, we had in mind the limitations of PIBAL (optical theodolite and radar) wind observations, but we cannot ascertain the same about wind profiles using radio-theodolites, as in most of the radiosonde systems of the past.

In the section where page 16 is we focus on humidity data, which must be accompanied by collocated temperature data. So, we think it is wise to withdraw wind and to clarify the difference between the top of humidity and temperature data in RAOB reports. (Of course, the burst altitude evolved with rubber balloons technology; we avoid such detail, since the increase of the vertical range of mandatory levels in the early years must have adapted to the vertical reach of balloons.) For the sake of accuracy, several minor corrections will be made in other parts of the manuscript.

Changes in the manuscript:

P7, footnote 2 regarding PILOT
Insert text:
The common single theodolite technique requires the approximate ascent rate to obtain position, while the double-theodolite method allows a pure trigonometric calculation.

P16, L20-21
Where it reads:
of temperature measurements in RAOB soundings
It will read:
of temperature measurements (i.e. of RAOB)

P16, L26-27
Where it reads:
Contrary to wind and humidity, temperature is usually measured up to the height achieved by the sounding balloon (or close to it), i.e., the burst altitude.
It will read:
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 ~~range~~ reach of humidity measurements depends on the working range of humidity sensors (and, to some degree, on reporting practices).

P16, L28
Where it reads:
at present situated almost 5 km below the burst altitude
It will read:
at present situated almost 5 km below the maximum altitude of RAOB (close to the burst altitude)

P17, L3-4
Where it reads:
faster than the burst altitude
It will read:
faster than the RAOB-top

P22, L18
Where it reads:
to the burst altitude (denoted by the maximum height of the temperature measurements)
It will read:
to the maximum height of the temperature measurements (close to the burst altitude)

P34, L4-5 | Caption of Figure 6
Where it reads:
in the annual soundings of temperature (TMP) and humidity (HUM) across the globe: mean and quartiles. (b) As in (a), but for the vertical extent of soundings.
It will read:
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.

Pag. 21, line 1: Did the metadata adhere to any international standards like ISO19115 or WIGOS? Please clarify this aspect.

The answer is no. In fact, such standards are of little use to our (meta) dataset for the following reasons. The present metadata do not substitute IGRA's own metadata (data sources, quality assurance, data format, stations list, periods of record, stations history). They are very specific in content, consisting on a set of values describing completeness of humidity data in a subset of the IGRA v2 main dataset – the sounding data from the stations reporting a minimum of RAOB data. Incidentally, as far as we understand, the metadata pertaining to IGRA do not adhere strictly to 'WIGOS metadata categories' or any other standard for geographic information. To avoid confusion with conventional metadata, we have named the dataset supplementary to the paper without using the word metadata: "A dataset of completeness of radiosonde humidity observations based on the IGRA".

As to metadata of the supplementary dataset, Sect. 4 and 5 of the manuscript provide primary information on the dataset, including a brief description of the represented parameters (see Tables 2 and 3), the spatial-temporal organization of data, data usability and access. The reader is referred to the Readme file for further details. Besides the short summary posted in the corresponding Zenodo's landing page, the Readme file provides the essential, mandatory elements*:* identification, purpose, data resources, sampling, spatial and temporal schema, data file description and format, authorship, contact. Ownership and data police are of course defined by Zenodo.

Pag.22, line 32: “exhibits” instead of the plural.

Correct.

Finally, a general comment about the plots in the different Figure. They are good and clear but the quality of the figures must be improved for the printing. Supplementary material is quite useful for the reader.

The original size of some figures is smaller than others because they were designed as in-text-figures, although observing the minimum recommended size (8 cm). If necessary, this issue can be corrected in due time.

Authors’ Response to Referee #2 Comments on

“Completeness of radiosonde humidity observations based on the IGRA” by António P. Ferreira, Raquel Nieto, Luis Gimeno

Referee comments (highlighted in blue) are copied before the authors’ responses.

The manuscript examines the completeness of newly released IGRA V2 humidity data. It is useful to have such a documentation to help users decide whether IGRA V2 has enough data for their own research before putting more efforts into downloading and analyzing the data. I would also like to appraise the authors for making their results (data) available and plan to update it on a two-year basis. I am little bit surprised on why the authors only look at humidity data, not including temperature and wind data. In “Introduction”, the authors did not provide the rationale for only studying humidity observations, such as less humidity data than temperature data and degraded performance for hygrometers. Based on my evaluation, I think that the manuscript in current version needs some revision. Some of specific comments are listed below.

Although the present work focuses on humidity, for comparison purposes, Sect. 2 and 3 provide some information about global wind and temperature data collected in IGRA [amount of stations and daily observation of the three parameters since 1905 (Fig. 1); vertical resolution/extent of temperature and humidity observations since 1945 (Fig. 6)].

The opening paragraphs of Sect. 1 recall the significance of radiosondes in accessing atmospheric humidity, as well as the role and limitations of historical radiosonde data to humidity studies. The problem of missing humidity observations due to limitations of operational radiosonde hygrometers and reporting practices, among other factors was pointed out in Sect. 1.2 and explained Sect. 1.3. Differences in vertical coverage related to reporting practices of humidity measurements will be also briefly mentioned on P2 [please refer to our reply to Referee #2].

Nevertheless, there is a particular reason why we have examined the completeness of radiosonde humidity observations, which in fact was not expressed on the manuscript. Humidity measurements (either as relative humidity or dew-point depression) require simultaneous temperature measurements, both quantities being measured at specified pressure levels. So, radiosonde humidity observations, in fact, represent simultaneous observations of pressure, temperature and humidity. Leaving aside horizontal wind – which is indirectly measured with the aid of a tracking device – humidity observations represent the most accomplished of the radiosonde observations. We think this is worth to be remarked in the revised version, not only in Sect. 1, but also in the Abstract and the concluding Sect. 6.

Changes in the manuscript:

P1, L11-12
Where it reads:
aiming to describe the completeness of humidity observations from radiosondes
It will read:
aiming to describe the completeness of humidity observations (i.e., of simultaneous observations of temperature and humidity) from radiosondes

P3, L11 (beginning of paragraph) – Insert text:

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' represent the most accomplished of the radiosonde observations.

P23, L9
Where it reads:
a dataset detailing the completeness of humidity observations
It will read:
a dataset detailing the completeness of humidity observations (RH and/or DPD together with pressure and temperature)

Specific comments:

1. P1 L16: spell out GUAN.

GUAN is now spelled right in the Abstract, except for the letter corresponding to GCOES which (like WMO) we believe is better known to most readers and is only translated the first time it appears in the main text.

P1, L16
Where it reads:
GUAN network
It will read
GCOES Upper-Air Network (GUAN)

2. P2 L30, add some of new references, such as Dai et al. (2011). Dai, A., J. Wang, P. W. Thorne, D. E. Parker, L. Haimberger, and X. L. Wang, 2011: A new approach to homogenize daily radiosonde humidity data J. Climate, 24, 965-991. This applies to other places in "Introduction".

Dai et al. (2011) was only cited in the concluding section; we agree it should be first mentioned in Sect. 1. [Please note that we add a few more updated refs regarding Referee #1 comments.]

P2, L29; P7, L8 – Insert citation:
Dai et al. (2011)

3. P3, L4: Durre et al. (2018, JTECH) should be used for the IGRA V2 reference.

In fact, that paper was not referred before because it was published after the submission date of our manuscript. A proper citation will be inserted in several places of the revised manuscript. We think we should also put a reference to CARDS, since it is the predecessor of IGRA v1. On P7, L24, Durre et al. (2016) refers specifically to the IGRA dataset used in our work.

Note: the information on the humidity quality checks on P8, L5-12 was kindly provided to the authors by Imke Durre; however, we have lately found that it became part of the documentation publicly available on the NOAA website for IGRA since April 18, 2018.

Changes in the manuscript:

P3, L3-4
Where it reads:
The first version contained practically data after 1945 (Durre et al., 2006). The IGRA Version 2 used in this paper (Durre et al., 2016), extends back in time as early as 1905
It will read:
The first version of IGRA (a successor of the 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, very recently documented in Durre et al. (2018), has enhanced data coverage and extends back in time as early as 1905

P8, L11-12
Where it reads:
(i) for RH (with the later introduction of this variable in the archive) and (iii)–(iv) were added in IGRA 2 (Imke Durre, personal communication, April 12, 2018)
It will read:
(ii) for RH (with the later introduction of this variable in the archive) and (iii)–(iv) were added in IGRA 2.

P8, L26 (before "NB") – Insert text:
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 submission date of this work.

Additions to the reference list:

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

Eskridge, R. E., O. A. Alduchov, I. V. Chernykh, Z. Panmao, A. C. Polansky, and S. Doty, 1995: A Comprehensive Aerological Reference Data Set (CARDS): Rough and systematic errors. Bull. Amer. Meteor. Soc., 76, 1759–1775, doi:10.1175/1520-0477(1995)076<1759:ACARDS>2.0.CO;2.

4. P4, L29: "although most of the soundings did not reach beyond 700 hPa", I think that this is outdated. Most of modern radiosonde soundings can reach above 700 hPa.

The phrase, inside parenthesis at the end of a sentence, reports to the earliest soundings in IGRA, coincident with the earliest upper-air temperature data at Lindenberg station. We will rephrase it for clarity.

P4, L29
Where it reads:
most of the soundings
It will read
most of those soundings

5. P7, L2: China doesn't use goldbeater skin anymore. Again, this info is outdated.

Nor Russia (!), according to the very recent Bruce Ingleby's report "An assessment of different radiosonde types 2015/2016, ECMWF Technical Memorandum 807 (2017). We believe that fast-response hygristors are not yet completely outdated (e.g., GTS1-1 radiosonde).
Changes in the manuscript:

Where it reads:

Although the capacitive thin-film sensors have been widespread (with Vaisala radiosondes), two older sensor types are still in use: the carbon hygristor (in VIZ radiosondes) and the goldbeater's skin sensor, introduced in 1950s and still used in radiosondes made in Russia and China; this peculiar sensor responds too slowly to be useful at temperatures lower than –20 °C and suffers from hysteresis following exposure to low humidity (Nash, 2015; Moradi et al., 2013).

It will read

Although the capacitive thin-film sensors have been widespread (with Vaisala radiosondes RS80 and RS92) two old humidity 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).

Addition to the reference list:

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

6. P13, L3: 500 hPa threshold might be too high for high elevation sites. I think that it would be 300 hPa.

Yes, most likely. However, according to usual practice, we take 500 hPa as minimal requirement to estimate PWV for all stations. A very few studies have extended this upper limit to 400 hPa and even 300 hPa (without making any distinction between low and elevated stations). Evidently that turns out to be too restrictive for past radiosonde observation. For a first approach, we think it's necessary to pay more attention to the vertical resolution up to the 850-hPa level, since the estimation of PWV is known to be particularly sensible to the humidity profile in that region.

7. Fig. 4: Is decreasing of radiosondes stations in 35-65N due to reducing number of radiosonde launches in Europe given the budget constrain?

We really don´t know. The related trends at other latitudes, after 1990, seem more significant on a global perspective. A detailed analysis by country would probably give more information. However, this is out of the scope of the current analysis.

Further author comment:

The geographical coverage of radiosonde stations is now discussed in the related Sect. 3.1
A new figure was introduced (Fig. 3 in the revised manuscript), while Fig. 4 was redesigned.

Summary of figure changes:

Fig. 3 ← NEW (page 46 of this document)
Fig. 4 ← former Fig. 3
Fig. 5 ← former Fig. 4, but redesigned according to the major changes in Sect. 2.3.1 and Sect. 3.1
Fig. 6 ← former Fig. 5
Fig. 7 ← former Fig. 9
Fig. 8 ← former Fig. 6
Fig. 9 ← former Fig. 7
Fig. 10 ← former Fig. 8
Fig. 11 ← former Fig. 10

Completeness of radiosonde humidity observations based on the IGRA

António P. Ferreira, Raquel Nieto, Luis Gimeno

Environmental Physics Laboratory (EPhysLab)
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 sounding data compiled in 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 humidity observations – i.e., of simultaneous observations of temperature and humidity – from radiosondes 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'), including 1300 WMO upper-air stations, of which 178 belong to the current GCOES Upper-Air Network (GUAN) ~~network~~. Completeness of humidity observations (either relative humidity or dewpoint-depression) for a radiosonde station and a full year is defined by: the number of humidity soundings; the fraction of days having humidity data; the mean vertical resolution of humidity data; the mean atmospheric pressure and altitude at the highest measuring level; and the maximum number of consecutive days without humidity data. The completeness of the observations ~~qualified~~ eligible for calculating precipitable water vapor – i.e., having adequate vertical sampling between the surface and 500 hPa – is 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, the study presents a global picture of the completeness of radiosonde humidity observations over the years, including their latitudinal coverage. This overview ~~shows~~ indicates that the number of radiosonde stations having a ~~long enough record length for studies on the climatic variability and trends of humidity-related quantities depends critically on the temporal continuity~~ record length potentially long enough (multidecadal) for climate studies involving humidity-related quantities depends not only on the temporal range, but also on the continuity, regularity and vertical sampling of the humidity time-series. 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 according to various completeness criteria – even if differences in instrumentation and observing practices require extra attention. A dataset is presented for that purpose, consisting of two main sub-sets: 1) humidity metadata for each of the IGRA-RS stations and year within the period of record (yearly metadata); and 2) humidity metadata for individual observations from the same stations (ascent metadata). These are complemented by 3) a list of the stations represented in the dataset, along with

the observing periods for humidity 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 ~~1948 (as in NCEP/NCAR products)~~ 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, ~~not to mention~~ 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-term and long-term distribution of moisture in the troposphere, despite ~~sampling differences among stations and over time, minor uncertainties related to observation time and balloon drift (Kitchen, 1989), and differences in data accuracy~~ 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 thein) and differences in humidity data accuracy – which depend~~s~~ on humidity sensors and ~~varies~~ vary with measured conditions (e.g., 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 ~~errors and insufficient data~~ uncertainty of measurements and biases in the upper troposphere and lower stratosphere (Elliot and Gaffen, 1991; Soden and Lazante, 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 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 ~~(Durre et al., 2016)~~ , very recently documented in Durre et al. (2018), has enhanced data coverage and extends back in time as early as 1905, although 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 also a combination of radiosonde and pilot-balloon observations) and the differences in the observing period, temporal regularity, 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’ represent 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 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 ~~relevant to the study~~ – both in terms of annual statistics and in 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 ~~indicated~~ 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;~~ in addition, the temporal regularity and continuity of humidity profiles are important to trend analysis and to detail annual cycles~~ 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~~ issues of spatial coverage and record length of stations, it seems to us that studying the completeness of observations in a global, historical data set of radiosonde observations should address these issues too.~~ issue of geographical coverage, studying the completeness of observations in a global, historical data set of radiosonde observations should address that issue too. 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 (~~eventually~~ and/or dewpoint depression~~ too~~) 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 ~~the~~ 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, until it was adopted worldwide by the end of 1991 (WMO, 1977, p. 15; WMO, 1987, pp. 57-58; Oakley, 1993, p. 23). 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

---

[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).

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 ~~the first IGRA data set~~ 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. 2.3.3).

The current migration of RAOB 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 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, radiosonde measurements reach the upper troposphere, much beyond the layers where most of the atmospheric water vapor resides. 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, 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 balloonsondes 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 on, 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 where 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 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. 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 newly 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), have improved the response time at low temperatures and the capability of measuring very low humidity. 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~~ are still in use~~ 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~~ still used in radiosondes made in Russia and China~~ used in some radiosonde types made in Russia and China until a few years ago; this peculiar sensor respond~~s~~ed too slowly to be useful at temperatures lower than –20 °C and suffer~~s~~ed 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 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. In sum, the vertical extent, vertical resolution, temporal regularity and continuity of humidity reports are ~~anything but uniform~~quite heterogeneous.

2 Input data and methods

We have examined the IGRA 2 main dataset until the end of 2016~~, comprising over 45 million soundings~~. 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] soundings taken ~~at 2761~~ over 2700 globally and temporally distributed stations~~ (2662 fixed and 99 mobile)~~, even though the coverage over oceans is limited to ships, buoys and remote

---

[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., 2011). 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

islands. IGRA 2 has also derived data for a selection of WMO stations, but 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 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 priory 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 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;
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); ~~(i)~~(ii) for RH (with the later introduction of this variable in the archive) and (iii)–(iv) were added in IGRA 2~~ (Imke Durre, personal communication, April 12, 2018)~~. 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 ~~RH measurements from radiosondes have a typical absolute accuracy of ~ 5 % on average tropospheric conditions. However, accuracy~~ the precision and accuracy of RH and DPD data ~~varies~~ 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~~ (Nash, 2002; Miloshevich et al., 2006; Nash, 2015)~~. 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).

---

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.

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 submission date of this work. NB – In the remaining of this paper, 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* – i.e., the stations with RAOB 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). 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 represent nowadays only 13 % of the total. The reason for the apparent discontinuity in the PIBAL stations between 1972 and 1973 is this: beginning in 1973, IGRA data largely come from the GTS ~~(Global Telecommunication System)~~ and include many more PIBAL data than prior data sources (Imke Durre, personal communication, April 12, 2018). The number of RAOB stations increased rapidly from the mid-1940s, staying in the range 850–900 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.' and 'RAOB' in Fig. 1a). The major relative change occurred between 1945 and 1946, coincident with the end of the second world war, when the global count of radiosonde stations triplicated. 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. After a major development until around 1970, the number of reporting radiosonde stations has been in the range 850–900. 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' and 'Higrom. > 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[4]. 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 temperature and humidity data, i.e. days without any RAOB data: 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).

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. 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, 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 missing years are considered, the full period of record of one station may be segmented into two or more periods for humidity. 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 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,

---

[4] 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.

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. ~~All~~ 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[5], personal communication, September 18, 2018). The GRUAN aims to serve as reference to 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. Despite the differences in data processing between GRUAN internal data and real-time data, the quality of routine observations carried out at GRUAN sites in recent years should be above average (WMO, 2011a; Dirksen et al., 2014). 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 to answer the following questions:

1. What is the ~~amount of humidity observations~~ spatial coverage of humidity-reporting stations in different years and latitudes~~ regions~~?

~~1.~~2. What is the fraction of days in a year with humidity data and the number of consecutive missing days on average?

~~2.~~3. What is the typical vertical resolution and vertical extent of humidity observations?

---

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

~~3.~~4. How many stations have enough data in the vertical to allow the estimation of precipitable water?

~~4. How long are the humidity time-series and how many consecutive missing days there is on average?~~

5. How does the temporal and vertical completeness affect the availability of long-term humidity time-series?

Each question was explored as detailed below in Sects. 2.3.1–2.3.4, with the results presented later in Sect. 3. The description of the related metadata parameters regarding ~~(*a*) RS station and year in record or (*b*)~~ each IGRA-RS station~~, date and nominal hour~~, is deferred to Sect. 4.

2.3.1 ~~Amount, frequency and latitudinal coverage~~Global coverage

~~The amount and frequency of humidity observations over time was studied in terms of *annual number of stations with any humidity observations* and *fraction of humidity-observing days in a year averaged across stations*. In both cases, the data were divided into subsets according to climatic zones, considering only the fixed stations.~~ 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 latitudinal distribution of humidity-reporting stations in different years was further studied by means of an index that accounts for the area of a latitude band between latitudes and :~~

$$\begin{aligned} zonal\text{–}coverage\ index &\stackrel{\text{def}}{=} \left(\frac{N_{band}}{A_{band}}\right) \div \left(\frac{N_{globe}}{A_{globe}}\right) \\ &= 2\left(\frac{N_{band}}{N_{globe}}\right)\left|\sin\theta_1 - \sin\theta_2\right| \end{aligned} \qquad \text{~~(1)~~}$$

~~where *N* denotes number of stations and *A* denotes surface area, the last expression using the spherical-Earth approximation. Of course, the mean station density over a latitude band, or the whole globe, may differ a lot from the local density because radiosonde stations are concentrated in continental, near urban areas of developed countries. The above-defined index is simply the mean stations density within a latitude band, such as a climate zone, divided by the global mean density. So, a value greater (/smaller) than 1 indicates that the mean stations' density for the latitude band is above (/below) global average, which is to say that the corresponding latitudes are overrepresented (/underrepresented).~~

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 φ and longitude λ 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 φ_1 and φ_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').

Eq. (1) was applied to the IGRA-RS 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 regularity

The frequency of humidity observations over time was studied in terms of the *fraction of humidity observing days in the year*. Although ~~the fraction of days with humidity observations during a year~~ 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. ~~Using the above definitions, first w~~We have ~~computed the size of missing days~~ averaged both quantities across all fixed stations reporting humidity within each major latitude region, year by year.

2.3.~~2~~ 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 defined by the geometric mean of $\{dz_k\}$ for all levels with humidity data in the sounding profile:

$$mean\ vertical\ resolution \overset{\text{def}}{=} \frac{R_d T_0}{g_0}\left\{\prod_{k=1}^{M}\left(\frac{\bar{T}_k}{T_0} ln \frac{p_{k-1}}{p_k}\right)\right\}^{1/M}$$

~~(2)~~

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

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 $k - 1$,~~T_0 is a constant temperature in the range of the values found in the troposphere (used in the calculation to avoid overflow in case M is too large)~~, 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 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, the statistical distributions of the vertical extent and the vertical resolution in humidity soundings 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 humidity soundings.

2.3.~~3~~ 4 Soundings ~~qualified~~ 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 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 show. 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 have been in widespread use since 2000. Secondly, the percentage of the surface-upper-air soundings having humidity data at the 925-Pa 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 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.~~4~~ 5 ~~Size of missing days and current~~ Current record length of time-series

~~Although the fraction of days with humidity observations during a year 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 '*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, 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 station may be divided into several sub-periods of record for humidity. Recall that the years with humidity observations at each RS station are indicated in Table S1.

~~. 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. Using the above definitions, first we have computed the size of missing days averaged across all fixed stations reporting humidity within each major latitude region, year by year. Then, w~~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,~~ including ocean stations with fixed location~~ i.e., over 1600 stations on continents and islands, 14 ocean fixed weather-ships and 2 environmental buoys. In the remaining Sects. 3.~~2~~3–3.~~4~~ 5 mobile stations (99 ships) are equally included~~, apart from one case at the beginning of Sect. 3.4~~ (see Table S1).

3.1 Amount and frequency of humidity observations by climate zones

Figure 3 shows the geographical distribution of the IGRA-RS fixed stations (see Table S1) 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. 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 number of stations measuring humidity is practically unchanged since then (see 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. Judging from IGRA data, between 1967 and 1972 there were 7 to 12 fixed ocean weather-ships transmitting radiosonde observations, almost all in the North Atlantic except for one 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 ship 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 10 to 20 ships of opportunity have launched radiosondes concurrently along their routes from the 1980s to present. The automated ice-drifting stations surveying the north polar region during 1950–91 are also worth noting. For statistics about the humidity observations from all floating stations included in IGRA (fixed and mobile), see end of Table S1.

Figure ~~3~~ 4 shows the count of fixed IGRA-RS stations with humidity observations in each year, by climatic zones on Earth. ~~Although~~ One can see that most of the few early humidity-reporting stations existing by 1940 were placed in the northern polar zone (Arctic)~~, from~~. From 1945 on they were located predominantly in the northern temperate zone, even though they raised in all regions, 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 on, compares with the Arctic in absolute number. Noticeably, the number of arctic stations decreased considerably between 1970 and 2000. The absence of mobile stations affects very little the curves of Fig. ~~3~~4, since mobile stations are short numbered and have a short period of record (a few years). E.g., considering the polar regions, according to IGRA, in 2016 there were only two ships reporting radiosonde observations in waters around the Arctic Circle – although the ~~cumulative~~ number of observations gathered over the years~~, gathered from different arctic missions,~~ may be relevant.

~~The latitudinal distribution of humidity-reporting stations is better expressed in terms of percentages for the same latitude bands, out of the total number of radiosonde stations over the globe, as shown in Fig. 4a; percentages are calculated at different times separated by an interval of 15 years. Clearly, the~~ The northern temperate zone ~~accounts~~ has accouted for about half of the world stations for decades, although its relative weight has been decreasing~~ since 1955~~over the years. The decrease of the ~~percent~~ number of stations situated in ~~mid- and high latitudes north of parallel 35° N~~ the northern extratropics after 1990 is counterbalanced by ~~a general increase in other regions, notably~~ an increase in the Tropics since then. ~~A similar histogram, but for the relative density of stations, i.e. the zonal-coverage index defined by Eq. (1), gives a different picture of the station coverage across latitude zones, as in Fig. 4b. The differences between the North and South Hemispheres are enlightened in the following way. From at least 1970 to 2015 the northern subtropical and temperate zones are equally overrepresented. In contrast,~~ 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 ~~this~~ these ~~is a~~ are regions with significantly more ocean and less land. Lastly, the Arctic has been much better covered than Antarctica. ~~Although it has decreased since the 1950s, the present-day mean artic stations density represents half of the global mean density.~~

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 from points on mostly land or ocean/sea regions within each climatic zone for every 15 years beginning in 1955, when the global radiosonde network was barely established. Fig. 5c shows a similar calculation but for the total surface area of each latitude band. Clearly, $\bar{s}$ has changed little over continents and large islands since at least 1970 and at all latitudes, with values ranging from about 200 km in northern temperate latitudes to 700 km 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, with values around 500 km over land, 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 Fig. 3. 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, except in the Southern Ocean when compared to Antarctica, but has also degraded slightly over time in the northern extratropical oceans and seas. 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 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 are accepted, in view of the relatively mild climatic conditions on oceans and the fulfillment from surface and satellite observations. 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 4b – which by including costal stations does not represent the actual station density in deep-ocean areas – indicates larger distances in most oceans, particularly in the southern midlatitudes where $\bar{s} >$ 1000 km.

~~Figure 5 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 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 corresponding standard deviations indicate that a non-negligible percentage of the stations have observations on every day in the year since the 1950s.~~

~~The fraction of days with humidity observations, detailed by station and year, is part of the first metadata set presented in this paper (see Sect. 4.1). Similar information is given for the Sfc-to-500hPa humidity soundings alone.~~

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

Figure ~~5~~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 corresponding standard deviations indicate that a non-negligible percentage of the stations have observations on every day in the year since the 1950s.

Figure ~~9~~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 with 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. ~~9~~7 gives a glint of the typical continuity of humidity time-series having any observations in the year. As a general picture, the average size of missing days decreased from 4–6 months to about 1 month between 1945 and 1960. 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 and the size of missing days, detailed by station and year, is part of the first metadata set presented in this paper (see Sect. 4.1). Similar information is given for the Sfc-to-500hPa humidity soundings alone~~.~~ (see Sect. 4.1).

3.~~2~~ 3 Global vertical extent and resolution of humidity observations

Figure ~~6a~~ 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 ~~6b~~ 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. 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, the soundings from mobile stations with missing geopotential height at the surface level had to be excluded. Curves are smoothed by a 5-years running mean, as in Fig. ~~6a~~8a. Note that the vertical extent of temperature measurements (i.e. of RAOB) ~~in RAOB soundings~~ 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 wind and humidity, temperature is usually measured up to the height achieved by the sounding balloon (or close to it), i.e., the burst altitude~~ 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 ~~6b~~ 8b we can observe that the top of the humidity observations is at present situated almost 5 km below the ~~burst altitude~~ 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 where 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. ~~6b~~ 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 ~~6b~~ 8b undoubtedly shows that humidity was mostly measured in the troposphere until the mid-1960s. In contrast, from the mid-1980s on, the vertical extent of humidity observations has increased consistently and faster than the ~~burst altitude~~ RAOB-top – roughly by 4 km per decade on global average –, denoting improvements of humidity sensors (see Sect. 1.3). In 2015, ¾ 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 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. ~~6b~~ 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. 2.3.3, 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. ~~6a~~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, 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.~~3~~ 4 Global relative amount of Sfc-to-500hPa humidity soundings

Recall that our definition of Sfc-to-500hPa humidity soundings (see Sect. 2.3.3) is intended to represent the sounding with a minimal amount of data, 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 $\{p: p \in \{1000, 925, 850, 700, 500\ \text{hPa}\} \cup \{p_{\mathrm{SFC}}\},\ p \le p_{\mathrm{SFC}}\}$;

B) Definition given in Sect. 2.3.3.

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.

Figure ~~7a~~ 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. ~~7a~~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 %. 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 ~~8a~~ 10a shows how the probability of finding a RS station with a percentage of Hum-A soundings exceeding a given value has reversed over about fifty years ago: only 12 % of the stations had at least 20 % of Hum-A soundings, whereas at present ≈ 91 % of the stations have at least 80 % of Hum-A soundings.

Figure ~~7b~~ 9b is the counterpart of Fig. ~~7a~~9a, when using the less stringent definition of Sect. 2.3.3 for “Sfc-to-500hPa humidity soundings” (Hum-B). 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. 3.2.3 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 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 ~~8b~~ 10b shows the probability of finding a 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 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.~~4~~ 5 ~~Average size of missing days and amount~~ Amount of long-term time-series

~~Figure 9 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 with 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. 9 gives a glint of the typical continuity of humidity time-series having any observations in the year. As a general picture, the average size of missing days decreased from 4–6 months to about 1 month between 1945 and 1960. 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.~~

Figure ~~10a~~ 11a (black lines) illustrate 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. In this case 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. 10a 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. ~~10a~~ 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. ~~10b~~ 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. 2.3.3) 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, ~~Figs. 9 and~~ Fig. ~~10~~ 11 ~~are~~ 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. ~~The size of missing days of humidity data at each station and year is given in the first metadata set provided in this paper; the same information is given for the sub-data consisting of Sfc-to-500hPa humidity soundings (see Sect. 4.1).~~ 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 same metadata set. In addition, the years with humidity observations can be found in Table S1 (as well as in a file accompanying the two main metadata sets): one or more periods per station, depending on whether there are gap years present 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 section 2.1 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 (stated in Sect. 2.3.3) 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 presented chronologically. However, to put humidity metadata into context, the yearly number of soundings and of RAOB soundings are also given besides HUMa. 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 1930, the pre-radiosonde is 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 2006, denoted herein as ‘IGRA-RS’. The selected set comprises 1723 stations, including 1300 WMO stations, of which 178 and 16 are, respectively, current GUAN and GRUAN sites. The early humidity reports 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, regarding the data later than 1945, are as follows.

- ~~The relative coverage of radiosonde stations measuring humidity in different climatic zones can be accessed in terms of the ratio between the stations’ mean density across the corresponding surface area and the mean density across the globe (‘zonal-coverage index’). The northern subtropical and temperate regions are clearly overrepresented (although not as much as in the past, regarding the extratropics). The zonal-coverage index of the tropical region has increased constantly since the mid-1950s, being presently above average. The worst coverage is found in the southern temperate region (largely oceanic), followed by Antarctica. The zonal-coverage index in the Artic is several times better than in Antarctica; however, it has decreased significantly over the years.~~ The radiosonde network providing humidity observations in the northern temperate latitudes and arctic region was essentially established in the mid-1950s. In the mid-1970 the globe was already 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 over time. Remarkably, the averaged distance to the nearest station over continents and large islands (between Greenland and Ireland in size), ranging from ≈ 200 km in the northern midlatitude countries to ≈ 700 km in Antarctica, has little changed. 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. Concerning ocean/sea areas, and disregarding mobile stations (‘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. Since the 1970s 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.
- 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.
- 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 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 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.
- ~~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 exhibit averages around 20 days in middle and high latitudes; in low latitudes it has increased over time from around 30 to 40 days.~~
- The amount of long-term humidity time-series with consecutive years of data until a given year depends on the year, the current record length and other completeness criteria. E.g.: 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 data that are available to perform climate studies, discounting discontinuities in accuracy, depends strongly on the strictness of the completeness criteria.

Finally, the paper presented 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 would like to thank to Dr. Imke Durre (Center for Weather and Climate, NOAA's National Centers for Environmental Information, Ashville, NC) for clarifying some points about the Integrated Global Radiosonde Archive - Version 2 (IGRA 2). We also thank to Orlando García (EPhysLab, Faculty of Science - University of Vigo, Ourense) for IT support. This work was funded by the Spanish government within the EVOCAR project (CGL2015-65141-R), co-funded by the European Regional Development Fund (ERDF), and it was also 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., 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.

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.

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., 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

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., 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: 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

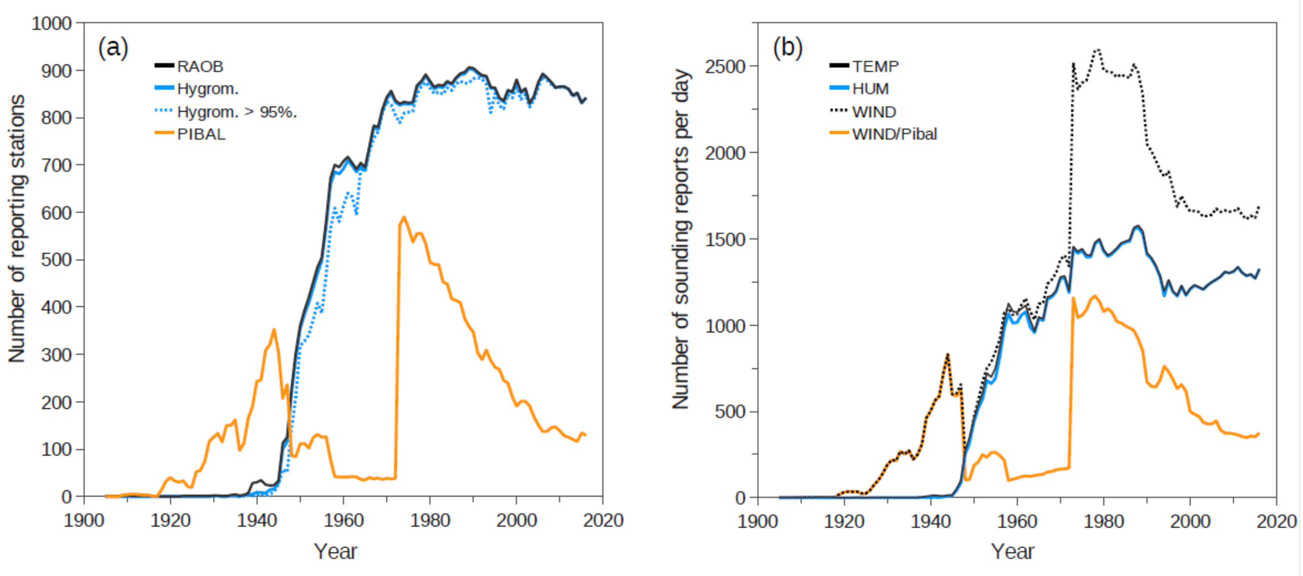
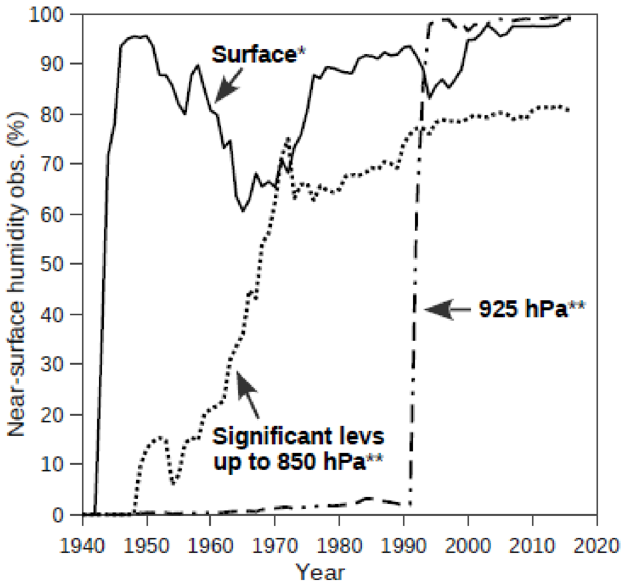


Figure 1

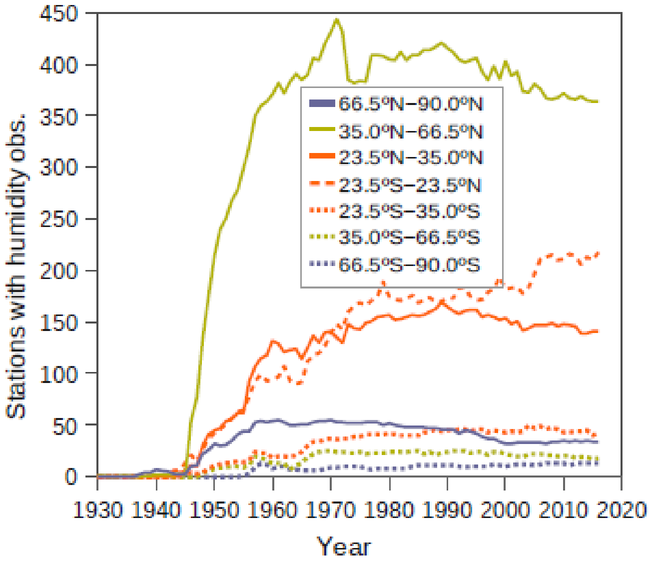
(a) Number of IGRA stations, for each year until 2016, reporting: any amount of radiosonde observations (RAOB); only wind observations (PIBAL); any humidity observations (Hygrom.); humidity data in at least 95 % of the RAOB soundings (Hygrom. > 95 %). (b) Number of IGRA sounding reports per day by atmospheric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



(*) surface-level observations out of all humidity soundings
(**) obs. out of reports with sfc data below the specified level

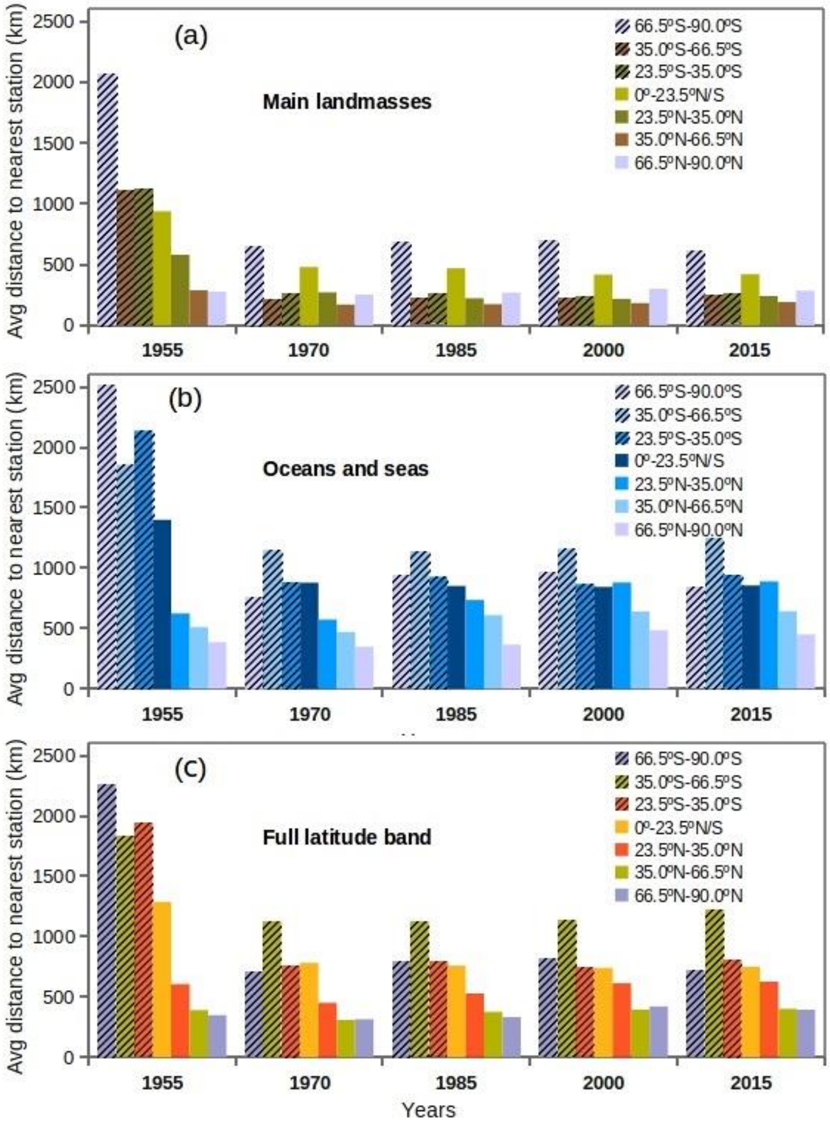
Figure 2

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~~ Figure 4

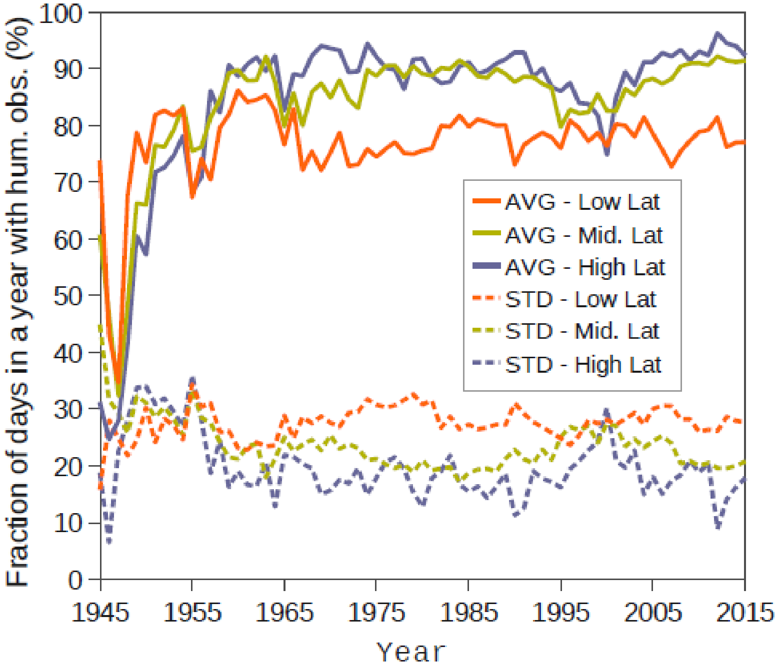
Yearly number of fixed stations reporting humidity observations in each climatic zone, since the time of the earliest humidity data in IGRA.



~~Figure 4~~ Figure 5

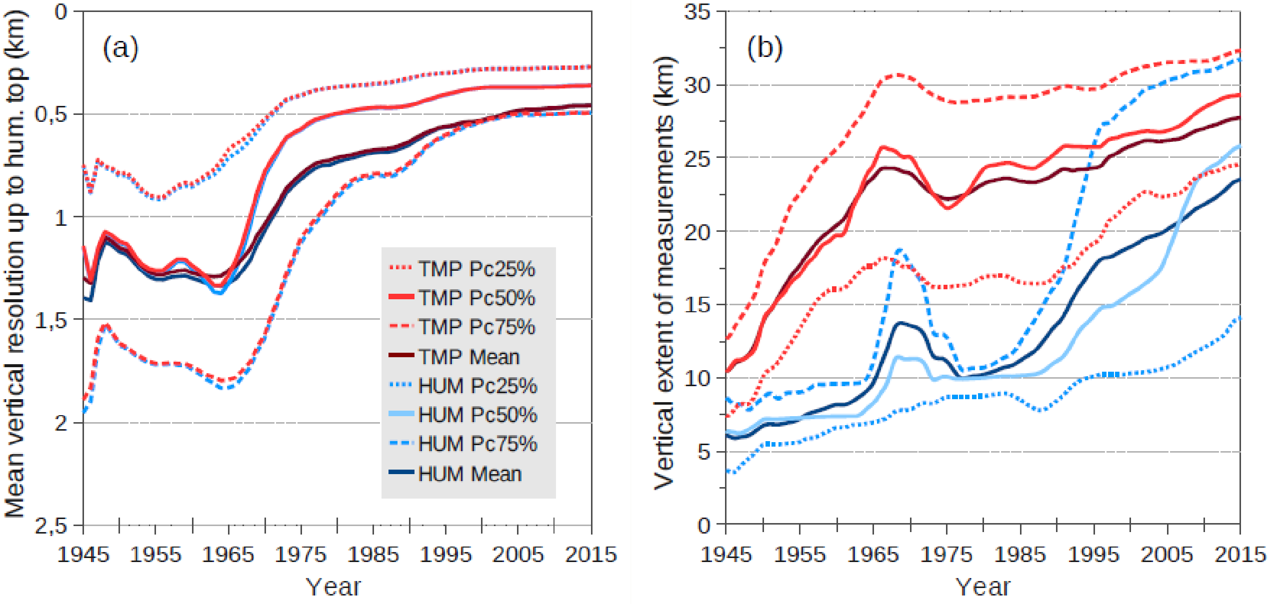
~~(a) Percentage of radiosonde stations measuring humidity, by climate zone, from 1955 to 2015 with 15 years interval. (b) Zonal-coverage index corresponding to (a), calculated by Eq. (1).~~

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)



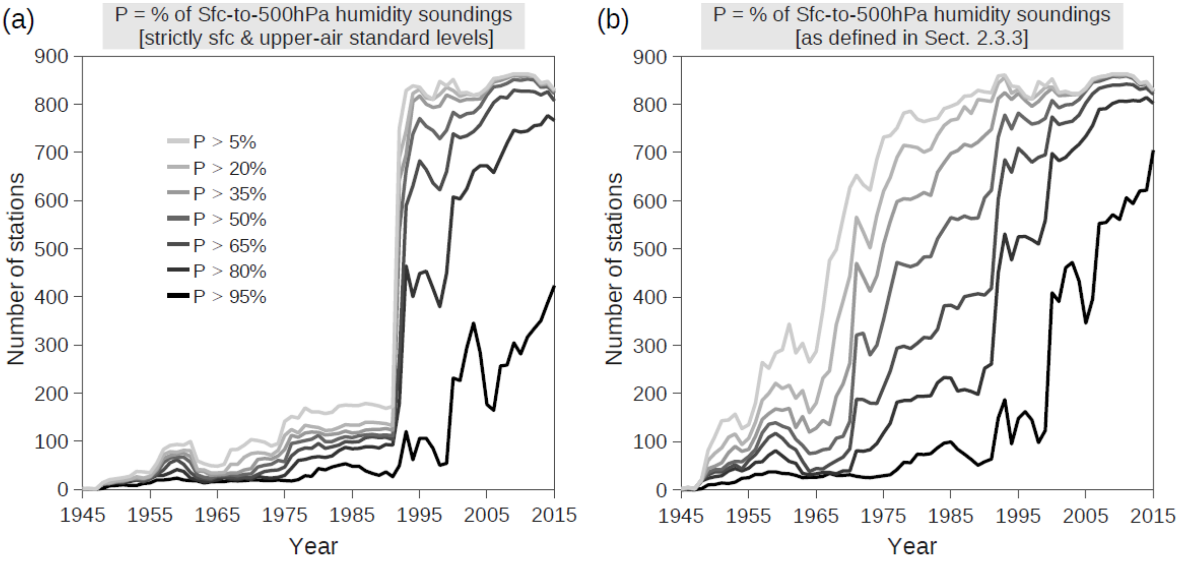
~~Figure 5~~ Figure 6

Statistics 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°.



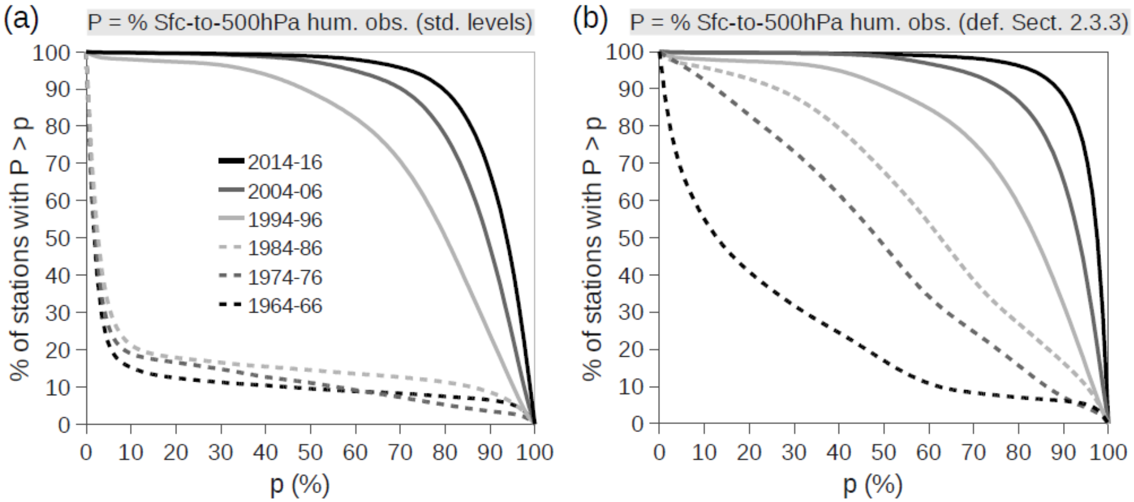
~~Figure 6~~ Figure 8

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



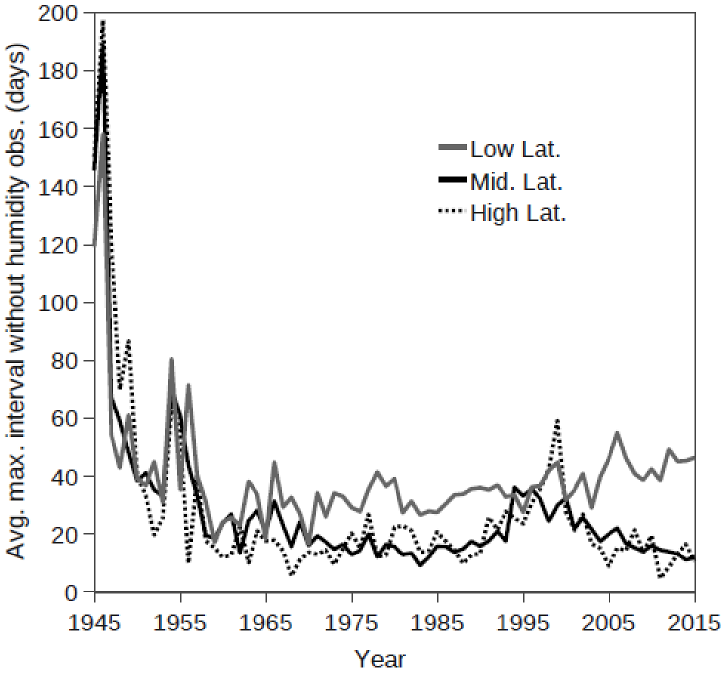
~~Figure 7~~ Figure 9

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. 2.3.3.



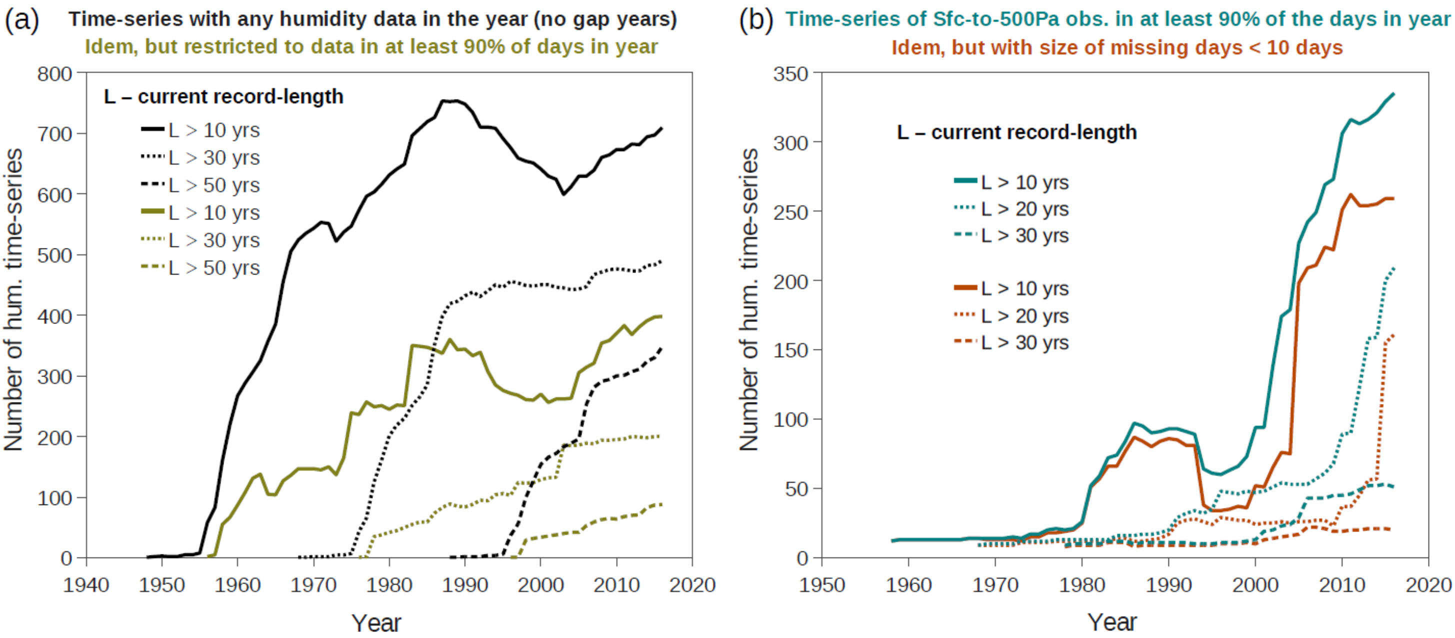
~~Figure 8~~ Figure 10

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. 2.3.3.



~~Figure 9~~ 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 10~~ 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).

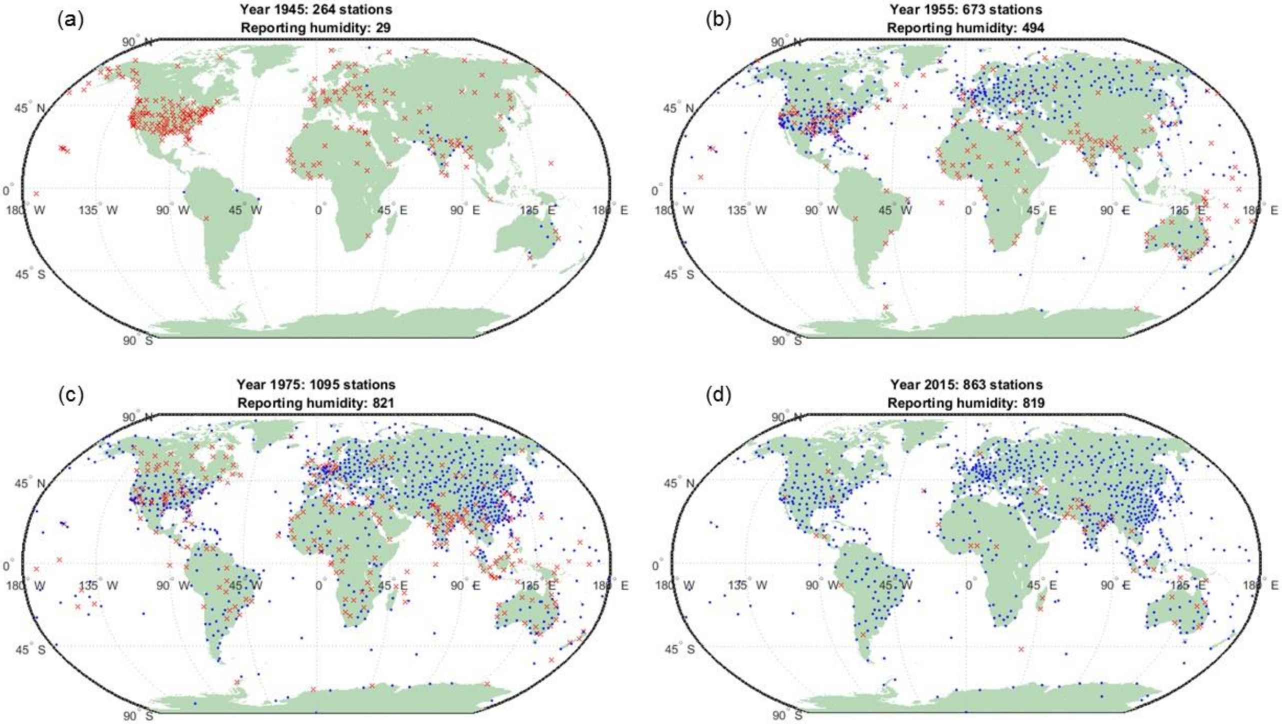


Figure 3

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.

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. 2.3.3. ([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 RAOB according to the thermodynamic data reported: 0 = none (only wind data); 1 = TEMP data, but missing HUM; 2 = TEMP and HUM data; 3 = TEMP and Sfc-to-500hPa HUM(1)
RESa	Mean vertical resolution of humidity data[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. 2.3.3. (2) Calculated by Eq. (2).

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