

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

This discussion paper is/has been under review for the journal Earth System Science Data (ESSD). Please refer to the corresponding final paper in ESSD if available.

# A global historical Radiosondes and Tracked Balloons Archive on standard pressure levels back to the 1920s

L. Ramella Pralungo<sup>1</sup>, L. Haimberger<sup>1</sup>, A. Stickler<sup>2,3</sup>, and S. Brönnimann<sup>2,3</sup>

<sup>1</sup>Department of Meteorology and Geophysics, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

<sup>2</sup>Oeschger Centre for Climate Change Research, Bern, Switzerland

<sup>3</sup>Institute of Geography, University of Bern, Bern, Switzerland

Received: 5 December 2013 – Accepted: 9 December 2013 – Published: 23 December 2013

Correspondence to: L. Ramella Pralungo (lorenzo.ramella-pralungo@univie.ac.at)  
and L. Haimberger (leopold.haimberger@univie.ac.at)

Published by Copernicus Publications.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Abstract

Long observed time series of climate state quantities are needed to extend our understanding of global weather, climate variability and monitoring not only at the Earth's surface but also the free atmosphere, since climate anomalies and climate change have a three-dimensional spatial structure.

Radiosonde data before 1958 and tracked balloon (PILOT) data are rarely available at standard times on standard pressure levels, which complicates their use for climate studies. This paper describes an upper air dataset on standard pressure levels at 00:00 GMT and 12:00 GMT for parameters temperature and wind. The input data sources of this dataset are the Comprehensive Historical Upper Air Network (CHUAN version 1.7), the Integrated Global Radiosonde Archive (IGRA) and upper air observations from the reanalysis datasets ERA-40 and ERA-Interim. As such it contains many data that have been digitized and collected in the EU FP7 project ERA-CLIM. Those data, which partly exist at asynoptic times and on altitude levels instead of pressure levels, have been brought to synoptic times and standard pressure levels using state of the art interpolation techniques, employing geopotential information National Oceanic and Atmosphere Administration (NOAA) 20th Century Reanalysis (NOAA 20CR) to aid the interpolation.

The standardized time series have been merged if necessary, taking care that the data are traceable back to their original sources. Only time series longer than 300 days have been saved in the merged archive, since the main purpose of this dataset is to aid climatological studies. If possible WMO numbers have been given to the station records. For some records which have never been identified by a WMO number, a local ID above 100 000 has been attached.

This paper describes the merging procedure, data count and data quality and how traceability of the data is ensured. It does not describe a homogenization procedure for both temperature and wind data. Homogeneity adjustments for both temperature and wind will be provided in a forthcoming paper.

ESSDD

6, 837–874, 2013

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

All the archives and metadata files are available in the *PANGAEA* archive with associated DOI <http://doi.pangaea.de/10.1594/PANGAEA.823617>. If users prefer netCDF files, they can be downloaded via <http://www.univie.ac.at/theoret-met/research/raobcore/>.

## 1 Introduction

The radiosonde network was practically the only upper air observing system up to the late 1970s and is still a valuable source of meteorological and climatological information, although there are now plenty of other observations such as satellite or aircraft data (Dee et al., 2011). While several global radiosonde archives exist and are publicly available, such as IGRA (Durre et al., 2006) or CHUAN (Stickler et al., 2010), they only partly fulfil the needs of climate scientists due to inhomogeneities in the data and since wind data from tracked balloons are often not available on standard pressure levels.

Almost all homogenized radiosonde datasets published so far, most notably Lanzante et al. (2003), McCarthy et al. (2008), Gruber and Haimberger (2008), Haimberger et al. (2012), Allen and Sherwood (2008), have been restricted to the period 1958 onwards, since before many radiosondes were not launched at standard synoptic times and data have been provided only on significant levels, not pressure levels. For tracked balloons, the situation is even worse, since those were tracked by theodolite or radar without any information on pressure. As such the observations have been collected on altitude levels, where these altitudes often were relative to the station height and the levels were not standardized either.

Even if millions of historical pilot balloon profiles were always digitally available on tape decks TD52 and TD53, a small fraction of these data (only post-1958) were used for input into NCEP/NCAR after a reworked process.

There are very few upper air wind climatologies so far that go back beyond 1958, and most of them work with monthly data although daily data are available back to the 1920s. Some studies have analysed the flow fields during special climatological

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


events such as the Dust Bowl, draught in the 1930s (Brönnimann and Luterbacher, 2004; Ewen et al., 2008). In a pioneering study, Brönnimann and Luterbacher (2004) used the data presented in “*A historical upper air data set for the 1939–1944 period*” (Brönnimann, 2003) to characterize the climate of troposphere and stratosphere in 1940–42 related to particularly strong El Niño event. Also Grant et al. (2009) did a first look on low frequency variability and trends of upper air temperature and geopotential but not winds.

The present study intends to improve the data availability by providing temperature and wind time series as far back as such data exist, but only on standard pressure levels. Data on altitude levels are interpolated to standard pressure levels using temperature information from the NOAA 20th century reanalysis (Compo et al., 2011). It is required that the time series are from ascending balloons (not kites or tethered balloons) and are at least 300 days long. As such the dataset is smaller than CHUAN (Wartenburger et al., 2014; Stickler et al., 2013) but is easier to use for time series analysis. The source data sets are described in the next section, details on the interpolation methods to standard time/pressure are given in Sect. 3, data counts and some results are presented in Sect. 4.

## 2 Input data

Creating a uniform radiosonde dataset is challenging since there are many different data sources and digitization efforts are still ongoing around the globe, producing valuable data but in different formats. The dataset presented here can, however, draw on the results of earlier integration efforts such as IGRA (Durre et al., 2006) and on input data preparation efforts for reanalyses (Uppala et al., 2005; Dee et al., 2011). During the assimilation process, altitude level data are supplemented with pressure level information and from those it is relatively easy to get standard pressure level data. These are available for a large fraction of post-1957 data since the ERA-40 (Uppala et al., 2005) and NCEP/DOE (Kistler et al., 2001) reanalyses that went back to 1957 and

1948, respectively. For altitude data that have not been assimilated yet, pressure information is constructed using geopotential information from the NOAA Twentieth Century Reanalysis (NOAA 20CR) (Compo et al., 2011) as described in the next section. The detailed list of input data used is as follows:

- The Comprehensive Historical Upper-Air Network (CHUAN) data set version 1.7, (Stickler et al., 2010; Wartenburger et al., 2014) and the ERA-CLIM Historical Upper-Air Data (Stickler et al., 2013). The ERA-CLIM historical upper-air dataset (acronym ECUD) contains upper air data collected and digitised within the EU 7th framework project ERA CLIM. These archives contain mainly historical upper-air data prior to the International Geophysical Year 1957. The data sets consist of 20 million balloon ascents written in around 5000 files that represent ca. 2000 stations with geopotential, temperature, wind and humidity data. The first record goes back to 1900. Those data, as well as some post 1957 data, have never been actively assimilated.
- The Integrated Global Radiosonde Archive (IGRA) (Durre et al., 2006), updated until 2012. IGRA contains data at standard and significant pressure levels and, sometimes, complementary altitude levels (not used in this work, since they do not add information to the pressure data). It is quite comprehensive and goes back to 1938. Being a dataset collected in the US it lacks a lot of data, however, over Europe prior to the mid-1960s.
- The ERA-40 observation input dataset. This dataset in BUFR format has lots of overlap with IGRA but contains several additional data over Europe, Japan and Antarctica that are missing in IGRA. The ERA-40 dataset starts in late 1957 and ends in 2002, however only data from 1958–1978 are used.
- The ERA-Interim observation input dataset. It is equivalent with ERA-40 observation input from 1979–2012 but is available in the far more convenient ODB format. The ERA-Interim input dataset goes up to present and is preferred to ERA-40 from 1979 onward.

Title Page

## Instruments

## Data Provenance & Structure

## Figures



[Back](#)

Close

Full Screen / Esc

[Printer-friendly Version](#)

## Interactive Discussion



## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


- The NOAA 20th century reanalysis (Compo et al., 2011). It does not contain radiosonde information but its geopotential field can be used to calculate pressure information for altitude levels. Since it is available back to 1872 even the oldest upper air data can be brought to pressure levels. In this work, it has been used as reference for data time/pressure interpolation and for quality control purposes.

Not only the observation input data are used from the reanalyses. ERA-40, ERA-Interim and the NOAA 20th century reanalysis all provide valuable reference fields for comparison with radiosonde data. Observation minus analysis departures (obs-an) from the NOAA-20CR have been calculated for all observations used. In addition Observation minus background departures (obs-bg) have been extracted for both ERA-40 and ERA-Interim. These are integral part of the dataset prepared here and they can greatly facilitate homogenization efforts (Haimberger, 2007; Haimberger et al., 2008, 2012; Gruber and Haimberger, 2008).

Next step is now to merge all those archives to get long timeseries, spanning the whole operative time of all the available stations.

### 3 Station identification

The station identification procedure is crucial in order to be able to join different records coming from the same station but stored in different archives.

For data assimilated in ERA-40 or ERA-Interim, this is relatively straightforward since the data must have WMO number and precise coordinates (latitude, longitude, altitude and time) in order to be assimilated. Also IGRA archives offers WMO ID numbers and coordinates for all the stations. The situation is much more difficult for CHUAN and ECUD archives: they have been delivered with metadata files with geographical coordinates (latitude, longitude, altitude, launch time), station name and, if available, WMO identification number and/or WBAN (Weather Bureau Army Navy) ID number. Only around 42 % of the stations have a WMO and 74 % have WMO and/or WBAN ID number. It has been recognized that many of the *unknown* stations can be marked with

WMO ID number. Automatic methods to assign the correct WMO numbers to these stations are complex, if not impossible, for the following reasons:

- the station names differ in different archives and, sometimes, they are reported in local language including nonstandard ASCII characters: an automatized unification appears to be very expensive;
- station relocations have often split records. In many cases it is possible to join them without introducing inhomogeneities. In other case, in the same area/city different stations were operative simultaneously (PILOT and radiosonde, for instance) and even if they are close by, they should be identified as independent stations and be merged only in a second phase.
- Cases of erroneous coordinates (typos, excenged signs) were detected.

For the aforementioned points, a first manual check has been performed to cross check the existing CHUAN and ECUD inventory files with the following metadata list (they can be downloaded from <ftp://srvx7.img.univie.ac.at>)

- WMO Observing Stations and WMO Catalogue of Radiosondes;
- Radiosonde comprehensive metadata catalogue (Texas University);
- ERA40 radiosonde list with metadata events;
- NOAA WBAN and WMO collection

Since no single reference file is complete, all of them are essential to assign and validate stations WMO ID numbers. Where the four listed station inventory files are incoherent themselves (lat/lon differs more than 2°) and after a manual check via Google-Maps, the most recent station location has been trusted. If WMO ID number and lat/lon were perfectly matching (most of the cases) the station identification was straightforward. When the station name was the same (considering the different languages and

## ESSDD

6, 837–874, 2013

### Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


typos) and lat/lon were close by ( $< 0.5^\circ$ ), but not perfectly matching, the station location was monitored with Google-Maps and, as expected, in many times it was simply a new collocation within the same area/city (e.g. from university to airport, from an airport to a new one). Stations with implausible lat/lon (wrong sign, inverted digits...) were highlighted and, after a double check, if possible their metadata have been corrected. Simultaneously operating stations in the same area/city (distance less than few km), where not more than one station has been reported as active<sup>1</sup>, have been treated as the same station. Where the localization was unsafe, no WMO ID number has been assigned.

As result of those efforts, we are able to identify roughly 95 % of the stations with a WMO ID number. The ones not included are, nevertheless, analyzed and stored. A table with station identifiers (file station\_list.txt) is presented in the supplements section. Where the latitude and/or the longitude are missing or are unrealistic in the station list inventory, the station has been rejected.

## 4 Interpolation from altitude to standard pressure levels

The PILOT balloons used for upper air wind measurements were tracked using theodolites or RADAR, both instruments report geometrical height as vertical coordinate. In both, CHUAN and ECUD, the wind observations are reported on altitude levels (meters above the sea level). An accurate interpolation from altitude to pressure (most likely not standard) and, in a second step, from not standard pressure to standard pressure levels requires either temperature plus humidity or geopotential information. Using standard atmosphere temperature values would introduce unnecessarily large errors. Geopotential information is available globally every 6 h on a  $2^\circ \times 2^\circ$  grid from the NOAA 20CR. These are interpolated bilinearly to the respective station locations (latitude/longitude).

<sup>1</sup>in the above cited metadata collection



## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


At the station location we can now find the interpolation weight  $a$  from the formula

$$\phi_x = a \cdot \phi_1 + (1 - a) \cdot \phi_2 \quad (1)$$

where  $\phi_1 < \phi_x < \phi_2$  where  $\phi_1$  and  $\phi_2$  are geopotential values at NOAA-20CR model levels at the station location and  $\phi_x$  is the reported altitude of the measurement multiplied by  $g$ . Now it is possible to determine the corresponding pressure at the station location  $p_x$ :

$$\ln p_x = a \cdot \ln p(\phi_2) + (1 - a) \cdot \ln p(\phi_1). \quad (2)$$

The pressure  $p_x$  is, most likely not yet a standard pressure level. In order to obtain values on standard pressure levels we perform again a linear interpolation from the available pressure levels to standard levels. This procedure was necessary also for assimilated PILOT (from ERA-40 and Interim input data) since those are only available on significant levels but not standard levels.

## 5 Time interpolation

Not all radiosonde and PILOT stations report at 00:00 UTC and 12:00 UTC. Particularly before 1958 the launch times were not standardized. In order not to lose too much data at asynoptic times, also a time interpolation has been implemented that allows backwards continuation of many records. To take into account the diurnal cycle we assume that the difference between observation and the reference NOAA 20CR is constant within  $\pm 6$  h of the observations. Thus, a simulated observation is generated by measuring the analysis departures obs-an at the time of the observation and assuming that the same departure exists at 00:00 UTC or 12:00 UTC.

The observations are divided in three time categories (Table 2).

These time categories are particularly important for the temperature measurements since the sensor may be affected by solar radiation bias. For temperature the time offset

is at most 6 h, which is crude. It could, in principle, be reduced to 3 h if 4 synoptic times per day were considered instead of just 2. A cubic interpolation is considered suitable to interpolate the 20CR to the observation time  $t_{\text{Obs}}$ .

Using the departure definition:

$$\text{departure}(t_{\text{Obs}}) = \text{Obs}(t_{\text{Obs}}) - 20\text{CR}(t_{\text{Obs}}) \quad (3)$$

we calculate the observation at the synoptic time 00:00 UTC and 12:00 UTC:

$$\text{Obs}(t_0) = 20\text{CR}(t_0) + \text{dep}(t_{\text{Obs}}) \quad (4)$$

$$\text{Obs}(t_{12}) = 20\text{CR}(t_{12}) + \text{dep}(t_{\text{Obs}})$$

Figure 1 summarizes the idea: the first two observations at synoptic time have been not manipulated. The third one has been reported at 03:00 UTC and we would like to shift it to 00:00 UTC. For this purpose we interpolate cubically, using the 4 closest analysis data, the NOAA 20CR (it could be temperature or  $U$  or  $V$  wind component) to the observation time and we calculate the departure observation minus reference. As second step, we add the departure to the NOAA 20CR value at 00:00 UTC (in this case,  $t_a$  in the picture), obtaining the reconstructed observation at the standard time 00:00 UTC. We take care that the same observation is not duplicated at 00:00 and 12:00 UTC.

In order to ensure the time interpolation consistency, we compare the results with the raw radiosonde data from ERA-40. ERA-40 used “first guess at appropriate time”, meaning that the background was compared to observation at the time of observation and not at the nearest synoptic time. Figures 2 and 3 evidence the good agreement of the CHUAN data after the two interpolations (altitude to pressure and time) with the ERA40 data, for the year 1957–1958 over USA for temperature and  $U$ ,  $V$  wind components. The mean difference CHUAN – ERA40 and the RMS are plotted against the standard pressure levels. The temperature shows excellent agreement, as expected, since the source data should be the same in CHUAN and ERA40. The constant difference of  $0.05^\circ$  is likely attributable to a different conversion from  $^\circ\text{C}$  to K. In this work, the

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



constant 273.15 has been adopted. The  $U$  and  $V$  wind components point out the good harmony between CHUAN and ERA-40, tiny differences are admissible and originate from the different adopted methods for the conversion altitude to standard pressure levels. The largest departures are located around 300 and 150 hPa, where there are large vertical wind gradients and the NOAA 20CR reanalysis has large temperature biases in some regions that may also lead to geopotential biases.

## 6 Merging the different archives

The good agreement and homogeneity between the time series coming from ECUD, CHUAN, IGRA, ERA-40 and ERA-Interim suggests that it is generally safe to merge these archives in a global one, in order to get longer, more complete and usable time series.

From the Figs. 7 and 8 is possible to follow the development of the upper air temperature and wind networks. While systematic wind observations begin in the 1920s (data stored in the ECUD and CHUAN archives), the systematic temperature observation starts only after 1945 (CHUAN and IGRA). Few pioneering temperature observation were already performed from 1900 (The Meteorologisches Observatorium Lindenberg/Richard Assmann Observatorium station (10393 WMO), in Germany, holds the longest record, with the first ascent dated 4 April 1900).

In order to merge all the stations and to ensure efficiency, the following rules have been adopted:

- station WMO ID number must be the same;
- station location (latitude, longitude and altitude) must be the same ( $\pm 0.5^\circ$ ) in the source archives, in order to preserve relocated station;
- a data priority has been set: (1) ERA-Interim, (2) ERA-40, (3) IGRA, (4) CHUAN and (5) ECUD;

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- only stations with more than 365 days have been considered;
- spike and consistency statistic tests have been performed in order to discard values have been performed erroneously.

For each observed value (temperature, wind) also *analysis departures* from the NOAA 20CR have been calculated:

$$\text{departure}(\text{day}, \text{pressure}, \text{time}) = \text{Obs}(\text{day}, \text{pressure}, \text{time}) - 20\text{CR}(\text{day}, \text{pressure}, \text{time}) \quad (5)$$

A simple quality control has been performed on the raw data:

- date and time limits must be plausible ( $0 =: 00 < \text{hour} < 23:59$ , we assume  $24:00 = 00:00$  of the next day, Gregorian calendar);
- temperature between  $-100$  and  $+60^\circ\text{C}$  or the equivalent in K;
- wind speed between  $0$  and  $200 \text{ ms}^{-1}$ ;
- wind direction between  $0$  and  $360^\circ$ ;

Inside those ranges, the observations may still contain very unlikely/wrong values due to many possible causes, the most likely being typos in the log books and digitization mistakes. The observation has been dropped during the merging procedure if its analysis departure is bigger than 4 times the standard deviation  $\sigma$  of the departures for the considered pressure level.

Figure 6 (Moscow station, temperature at 500 hPa, green CHUAN, red ERA40) highlights the presence of spikes in the raw time series. Those erroneous values are not propagated to the merged archive. A more comprehensive spike evaluation is shown in the Fig. 4, where all the archives illustrate spikes density below 0.5 % for all pressure levels. Exception is ECUD for temperature at 150 hPa: the spike rate is 1 %. The ECUD Temperature data are prior to 1957 and coming mostly from Siberia, region

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where the NOAA-20CR suffers by strong bias respect ERA-40 and ERA-Interim, as Brönnimann et al. (2012) report. In order to avoid implausible spike flag due the above mentioned bias, the NOAA-20CR has been adjusted with the montly difference respect ERA-Interim in the year 1979 calculated on the station location (bilinear interpolation),

5 assuming the gap significan and constant. In Fig. 5, the global temperature,  $U$  and  $V$  wind component difference ERA-Interim minus NOAA-20CR at 150 hPa, mean over 00:00 and 12:00 UTC, for the year 1979 has been plotted. Particular evident and strong is the warm bias present at high latitudes (beyond  $60^\circ$  N and S), up to 12K. Also a year cycle is visible with strong signal between October and June. Opposite situation for the

10  $U$  wind component where the strong difference (up to  $8 \text{ ms}^{-1}$ ) is concentrated in the tropical regions equally distributed during the year. The  $V$  wind component does not show strong bias.

## 7 The time series viewer

For simple time series visualization a Javascript-based time series viewer, available at the page [http://srvx7.img.univie.ac.at/~lorenzo/DEVL\\_rrvis\\_2.0/html/](http://srvx7.img.univie.ac.at/~lorenzo/DEVL_rrvis_2.0/html/), has been developed. It allows quick monitoring of the data archive which permits visual detection of outliers and shifts. One can choose between observed variables (temperature and wind speed, direction,  $U$  and  $V$  components) and departures from different background series (NOAA 20CR, ERA-Interim and ERA40). Observation time (00:00 UTC, 12:00 UTC

15 and 00:00–12:00 UTC difference) and the pressure level (from the 16 standard pressure levels) can be selected from the self explanatory menu.

20 Figure 10 shows the analysis departures (i.e. Observations minus NOAA 20CR) for Moscow (26 712, Russia). More details about the viewer can be found on <http://reanalyses.org/observations/raobcorerich-visualization>.

## 8 The MERGED ARCHIVE

The union of all data-sets gives a total of 3217 stations (land stations and anchored weather ships that use radiosondes, tracked balloons, with time series longer than 365 days), where 3020 have been recognized as WMO stations with valid WMO ID number.

5 1598 (1596 with WMO ID) stations contain temperature observations and 3152 (2957 with WMO ID) stations contain wind observations (as  $U$  and  $V$  components). The time series span from 1905 until today.

The Meteorologisches Observatorium Lindenberg/Richard Assmann Observatorium station (10393 WMO), in Germany, has the longest record, going back to 4 April 1900, but has several gaps due to war time disruptions. The longest continuous upper air temperature record comes from Moscow with data available from 1938.

Regarding the wind observations, the longest records have been collected in the USA, where in the 1920s the first upper air network was installed. The observations were performed with tracked balloons with all the difficulties and challenges of this practise by that time: manual measurements of speed and direction using theodolites did not allow to reach level higher than 400 hPa. Only with improvements in the instrumentation, it was progressively possible to reach higher levels (100 hPa were reached around 1950).

20 The global upper air network time coverage and distribution are visible in Figs. 7 and 8. In order to explore the developed of the global upper air network, it is interesting to examine, decade by decade, the number and the position of the operative stations. In Fig. 9 the development of the upper air observing network has been displayed<sup>2</sup>. The first systematic wind observations are dated 1920 over the United States and they become dense from 1935 onward. In the years 1945/1950 in Europe, Russia (even if in Moscow temperature observations has been maintained since 1938) and Japan a rudimentary upper air observation network is present, but also in Australia, New

<sup>2</sup>Only stations with more than 5 yr of Observations for the selected decade have been plotted

ESSDD

6, 837–874, 2013

### Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Zealand, the Hawaii, Polonesia and Africa few, but important, stations are working. Also stationary weatherships are operative in Atlantic and Pacific Oceans.

Already in the decade 1950–1960 in the North Hemisphere the global coverage is satisfactory.

From 1955 the Chinese and Australian radiosonde network become fully operational. While the already existent observations are extended and reinforced (more stations with both 00:00 UTC and 12:00 UTC launches), the South American network (mainly Chile and Argentina) is set up. Almost in the same years, permanent stations turn fully functional on the Antartica coast and new weather ships are operative.

In the most recent times and today, a homogeneous coverage over the Globe is reached, even if there are still not dense enough regions, as central Africa and South America. The maximum coverage was in the decade 1980–1990, afterwards there is a decrease, especially over former European colonies and the former Soviet Union. Anyway, a good spatial coverage has been maintained over whole globe, but observation scarcity is still particularly evident over central Africa and South America, key tropical regions (the stations in remote regions are particularly important especially before the satellite era, since these are the only anchors for reanalyses where no other data are available).

The global radiosonde network reaches its maximum extension, in terms of number of stations, in 1957/58, during the International Geographical Year where many new stations were set and measure campains were performed in remote regions (Siberia, polar regions, central Africa and South and Central Asia), in this biennium, more than 1600 stations were operative, rough 1200 reporting wind and 900 temperature, but unfortunately a large fraction was stopped after a few months and they don't contribute actively to the merged archives since their time series are too short. After IGY, which was a crucial step, in order to expand (new stations) and standardize (unified procedure for the observation method), for the global upper air network, the observing system remained quite stable for 30 yr, with around 800 stations reporting TEMP and 1000 reporting wind, before it declines, expectially for the PILOT, over former colonized

**Radiosondes and  
Tracked Balloons  
Archive**

L. Ramella Pralungo  
et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

regions. In December 2012, 825 stations were active, 713 reporting temperature and 804 wind.

The Table 1 summarizes how the single archives contribute to the merged archive. For temperature ECUD and CHUAN data, 66.3 % and 45.1 %, respectively, of the available observations have been ingested in the merged archive. The percentages are not higher because the most recent data stored in these archives are partly overlapping with IGRA and/or ERA-40, and those have higher data priority. For wind, more than 70 % of the ECUD and CHUAN data flow into the merged archive. The new digitized (ECUD and CHUAN) data contribute rough 4.8 % (Temperature) and 10.4 % (Wind) to the merged archive.

After the merging procedure, many stations now own more than 70yr of continuous observations, which makes them extremely interesting and valuable for further studies. One should note, however, that these data are not yet homogenized. The inhomogeneities are, however, relatively easily visible if one studies the NOAA-20CR or ERA-40/ERA Interim departure time series. One unequivocal example is reported in Fig. 12: the plot shows wind direction departures from NOAA-20CR for the station 072 764 (Municipal Bismark, North Dakota USA). As expected, the departures (running mean 200 days is applied) are constant and well balanced around the null line but, a strong bias (roughly 15°) is visible in the period 1938–1948 (only in data coming from the CHUAN archive).

## 9 Conclusions

The presented merged dataset contains upper air temperature and wind records on standard pressure levels back to the 1920s. It is specifically targeted for advanced quality control and bias adjustments, and, of course, climatological analysis. It complements existing upper air datasets (ECUD, CHUAN, IGRA, ERA-40 input, ERA-Interim input) that are in total perhaps more complete (they contain altitude and/or pressure levels and also short time series with less than 300 days with observations) but also

### Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


more difficult to use and not always aligned as time series. It contains not only the raw observations but also departures to the NOAA 20th century reanalysis (Merged, IGRA, CHUAN and ECUD archives) and ERA-40/ERA-Interim background forecasts. As such the dataset is particularly suitable as a basis for a homogenized temperature and wind dataset, that uses RAOBCORE technology for bias adjustments. The homogeneity adjustments and their effect on the time series and global mean trends are described in upcoming papers (Ramella-Pralungo and Haimberger, 2014; Haimberger et al., 2013).

The altitude to standard pressure level conversion involved the use of NOAA 20CR geopotential information. The time resolution is relatively coarse and future surface pressure only reanalyses, such as ERA-20C (Poli et al., 2013), will help to improve on that since they have passively assimilated the upper air data and thus measure the background departures at the right time, which allows to avoid the time interpolation step. Future surface data only reanalyses also may have smaller temperature and wind biases than do the NOAA-20CR. The archive is available in convenient NetCDF format and can be visualized with a simple online plotting tool. The archive will be updated once a year shortly after a full year has been completed in ERA-Interim.

## Appendix A

### NetCDF File structure

We provide, for the merged and, as well, the input data archives a reformatted version, stored in NetCDF files that contain the time series with daily resolution (00:00 and 12:00 UTC ascents) at the 16 standard pressure level (10, 20, 30, 50, 70, 100, 150, 200, 250, 300, 400, 500, 700, 850, 925 and 1000 hPa); one file for each variable (temperature, U-wind and V-wind components) and station.

The file name has been created for a easy and quick search, like:

123456\_V\_t.nc

(A1)

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


The first digit could be:

- 0 the station has been identified as a WMO station;
- 1 the station has been identified as a NO WMO station;

The next 5 digits are the WMO station identification number, if the first is a 0, otherwise they are the progressive number with which the station has been saved in the respective archive (CHUAN or ECUD, since only those two archives contain unknown stations). The *V* refers to the Variable reported in the file and it can be:

- *T* ⇒ temperature
- *U* ⇒ *U* wind component
- *V* ⇒ *V* wind component

The *\_t* refers to *time series*, as is the form in which the data have been stored.

In the file are defined 13 dimensions, 18 variables and the global attributes.

The variables list is composed by (type ⇒ name):

- integer ⇒ stations ⇒ Station ID, it works as the first 6 digits in the file name (according to A1)
- float ⇒ lat ⇒ Station Latitude, in degrees North, range  $[-90^{\circ}, 90^{\circ}]$
- float ⇒ lon ⇒ Station Longitude, in degrees East, range  $[-180^{\circ}, 180^{\circ}]$
- float ⇒ alt ⇒ Station Altitude, in m a.s.l, range  $[-400, 8000]$  m
- integer ⇒ pressure\_layers(16) ⇒ array with the 16 standard pressure levels;
- integer ⇒ obs\_time(2) ⇒ array with launch times (00:00 and 12:00) UTC;
- integer ⇒ date(45000) ⇒ progressive day index from 1900-01-01 with Gregorian calendar, range [19000101, 20230316]

- integer Varno  $\Rightarrow$  the variable number identifier following the ECMWF line guides (Upper Air Temperature 2 [K]; U-component of wind 3 [ms<sup>-1</sup>]; V-component of wind 4 [ms<sup>-1</sup>])
- integer  $\Rightarrow$  index\_days  $\Rightarrow$  progressive (from 1). Date(index\_days) returns the corresponding day that refers to index\_days for the selected station.
- float  $\Rightarrow$  obs  $\Rightarrow$  The observations array could be named, in agreement with the variable reported in the file name:
  - Temperature  $\Rightarrow$  Temperature
  - U\_Wind  $\Rightarrow$  *U* wind component
  - V\_Wind  $\Rightarrow$  *V* wind component

The arrays have dimensions

obs(obs\_time, pressure\_layers, index\_days)

In this way, we use the minimum number of days in order to map the time series.

After the observed time series, the departures (background departures from ERA-Interim, ERA-40 and analysis departures from NOAA 20CR) flags and sonde type information are stored as follows:

- float  $\Rightarrow$  biascorrect  $\Rightarrow$  biascorrect(obs\_time, pressure\_layers, index\_days), only available for ERA-Interim and ERA-40 archives, where biascorrections procedure has been performed by ECMWF (Haimberger and Andrae, 2011; Andrae et al., 2004);
- float  $\Rightarrow$  fg\_depar  $\Rightarrow$  fg\_depar(obs\_time, pressure\_layers, index\_days), only available for ERA-Interim and ERA-40 archives, where it has been performed as Observation-Forecast first guess (ECMWF forecast);

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



– float  $\Rightarrow$  an\_depar  $\Rightarrow$  an\_depar(obs\_time, pressure\_layers, index\_days), for ERA-Interim and ERA-40 archives it has been performed as Observation-Analysis (ERA-40 or ERA-Interim Reanalysis ECMWF); for IGRA, CHUAN, ECUD and Merged archive, the Analysis comes from NOAA 20CR fields (updated to 2010).

– integer  $\Rightarrow$  sonde\_type  $\Rightarrow$  sonde\_type(index\_days), contains informations for the sonde type utilized for each day with observations as suggested by WMO (see file WMO\_sondetype<sup>3</sup>, in ftp://srvx7.img.univie.ac.at

– integer  $\Rightarrow$  status  $\Rightarrow$  status(obs\_time, pressure\_layers, index\_days) it contains the data source archive for the current obs\_time, pressure\_layers, index\_days:

– status=1  $\Rightarrow$  ERA-Interim input;

– status=2  $\Rightarrow$  ERA-40 input;

– status=3  $\Rightarrow$  IGRA archive;

– status=4  $\Rightarrow$  CHUAN archive;

– status=5  $\Rightarrow$  ECUD;

– integer  $\Rightarrow$  anflag anflag(obs\_time, pressure\_layers, index\_days) flag, not used for this data type;

– integer  $\Rightarrow$  event1 event1(obs\_time, pressure\_layers, index\_days) flag, not used for this data type;

The file is equipped with global attributes:

– Conventions = “CF-1.4”  $\Rightarrow$  NetCDF files conventions;

– title  $\Rightarrow$  “”

<sup>3</sup>in the file:

<http://www.wmo.int/pages/prog/www/IMOP/meetings/Upper-Air/Radiosonde-netw/Doc5-3.pdf>

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


- institution ⇒ file owner “University of Vienna”;
- history ⇒ when the file has been produced;
- datatype ⇒ “RADIOSONDE INPUT DATA”
- source ⇒ source archive
- references: ⇒

Routines for reading the archived NetCDF files are available in FORTRAN and IDL and can be downloaded from <ftp://srvx7.img.univie.ac.at>.

*Acknowledgements.* This work has been funded by projects P21772-N22 and P25260-N29 of the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (FWF), as well as by the EU 7th Framework Programme collaborative project ERA-CLIM (Grant No. 265229) and by the Swiss National Science Foundation (project EVALUATE). The authors thank the collaborators within ERA-CLIM, most notably Hans Hersbach and Alexander Sterin for their support and constructive comments.

## References

- Allen, R. J. and Sherwood, S. C.: Warming maximum in the tropical upper troposphere deduced from thermal winds, *Nat. Geosci.*, 1, 399–403, 2008. 839
- Andrae, U., Sokka, N., and Onogi, K.: The radiosonde temperature bias correction in ERA-40, Volume 15 of ERA-40 Project Report Series, ECMWF, 2004. 855
- Bronnimann, S.: A historical upper air-data set for the 1939–44 period, *Int. J. Climatol.*, 23, 769–791, 2003. 840
- Brönnimann, S. and Luterbacher, J.: Reconstructing northern hemisphere upper-level fields during World War II, *Clim. Dynam.*, 22, 499–510, doi:10.1007/s00382-004-0391-3, 2004. 840
- Brönnimann, S., Grant, A. N., Compo, G. P., Ewen, T., Griesser, T., Fischer, A. M., Schraner, M., and Stickler, A.: A multi-data set comparison of the vertical structure of temperature variability and change over the arctic during the past 100 years, *Clim. Dynam.*, 39, 1577–1598, doi:10.1007/s00382-012-1291-6, 2012. 849

- Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Å., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, *Q. J. Roy. Meteor. Soc. A*, 137, 1–28, doi:10.1002/qj.776, 2011. 840, 841, 842
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011. 839, 840
- Durre, I., Vose, R., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, *J. Climate*, 19, 53–68, 2006. 839, 840, 841
- Ewen, T., Brönnimann, S., and Annis, J. L.: An extended pacific North American index from upper air historical data back to 1922, *J. Climate*, 21, 1295–1308, 2008. 840
- Grant, A., Brönnimann, S., Ewen, T., and Nagurny, A.: A new look at radiosonde data prior to 1958, *J. Climate*, 22, 3232–3247, 2009. 840
- Gruber, C. and Haimberger, L.: On the homogeneity of radiosonde wind time series, *Meteorol. Z.*, 17, 631–643, 2008. 839, 842
- Haimberger, L.: Bias correction of radiosonde observations, in: *Proceedings of the ECMWF/GEO Workshop on Atmospheric Re-analysis*, ECMWF, 19–22 June 2006, Shinfield Park, RG2 9AX Reading, UK, 123–126, 2007. 842
- Haimberger, L. and Andrae, U.: Radiosonde temperature bias correction in ERA-Interim, ERA report series, 8, available at: [http://www.ecmwf.int/publications/library/ecpublications/\\_pdf/era/era\\_report\\_series/RS\\_8.pdf](http://www.ecmwf.int/publications/library/ecpublications/_pdf/era/era_report_series/RS_8.pdf) (last access: 19 December 2013), 2011. 855
- Haimberger, L., Gruber, C., Sperka, S., and Tavolato, C.: Homogenization of the global radiosonde temperature and wind dataset using innovation statistics from reanalyses, in: *Proceedings of the Third WCRP International Conference on Reanalyses*, 5 pp., available at: [http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/G4-411\\_ea.pdf](http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/G4-411_ea.pdf) (last access: 19 December 2013), 2008. 842

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo et al.

Title Page

Abstract

Instruments

Data Provenance & Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


- Haimberger, L., Tavolato, C., and Sperka, S.: Homogenization of the Global Radiosonde Temperature Dataset through Combined Comparison with Reanalysis Background Series and Neighboring Stations, *J. Climate*, 25, 8108–8131, 2012. 839, 842
- Haimberger, L., Ramella-Pralungo, L., Hersbach, H., and Dee, D.: Temperature and wind bias adjustments used as input for ERA-SAT, ECMWF, in preparation, 2013. 853
- Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP/NCAR 50-year reanalysis: monthly mean CDROM and documentation, *B. Am. Meteorol. Soc.*, 82, 247–267, 2001. 840
- 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, 2003. 839
- McCarthy, M. P., Titchner, H., Thorne, P., Haimberger, L., and Parker, D. E.: Assessing bias and uncertainty in the HadAT adjusted radiosonde climate record, *J. Climate*, 21, 817–832, 2008. 839
- Poli, P., Hersbach, H., Tan, D., Dee, D., Thépaut, J. N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., Dragani, R., Trémolet, Y., Holm, E., Bonavita, M., Isaksen, L., and Fisher, M.: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), ECMWF, 2013. 853
- Ramella-Pralungo, L. and Haimberger, L.: Homogenization of the global historical upper air wind dataset, *J. Climate*, in preparation, 2014. 853
- 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. Meteorol. Soc.*, 91, 741–751, doi:10.1175/2009BAMS2852.1, 2010. 839, 841
- Stickler, A., Brönnimann, S., Jourdain, S., Roucaute, E., Sterin, A., Nikolaev, D., Valente, M. A., Wartenburger, R., Hersbach, H., Ramella-Pralungo, L., and Dee, D.: Description of the era-clim historical upper-air dataset, *Earth Syst. Sci. Data Discuss.*, submitted, 2013. 840, 841
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saari-nen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L.,

- Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 Re-analysis, Q. J. Roy. Meteor. Soc., 131, 2961–3012, 2005. 840
- 5 Wartenburger, R., Stickler, A., and Brönnimann, S.: A novel physics-based adjustment of upper-air temperature and geopotential height, J. Climate, in preparation, 2014. 840, 841

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)[Abstract](#)[Instruments](#)[Data Provenance & Structure](#)[Tables](#)[Figures](#)[I◀](#)[▶I](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)



**Radiosondes and  
Tracked Balloons  
Archive**L. Ramella Pralungo  
et al.**Table 1.** Archives contribution to the merged archive

	ECUD	CHUAN	IGRA	ERA-40	ERA-Int
Temperature					
% data used	66.30	45.1	13.8	48.5	99.9
% data respect total	0.1	4.7	10.0	32.8	52.4
Wind					
% data used	70.4	78.8	13.9	53.7	99.9
% data respect total	0.7	9.7	8.8	32.7	48.2

Title Page

Abstract

Instruments

Data Provenance &amp; Structure

Tables

Figures

I◀

▶I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Radiosondes and  
Tracked Balloons  
Archive**L. Ramella Pralungo  
et al.

Title Page

Abstract

Instruments

Data Provenance &amp; Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

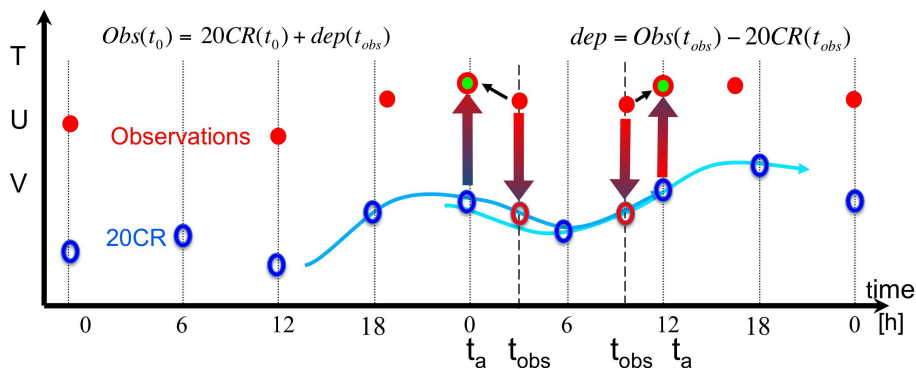
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Table 2.** Assignment strategy for observations for interpolation to nearest synoptic time.

Observation	Closest synoptic time
< 06:00 UTC	00:00 UTC
≥ 06:00 UTC and ≤ 18:00 UTC	12:00 UTC
> 18:00 UTC	00:00 UTC (day after)

Radiosondes and  
Tracked Balloons  
ArchiveL. Ramella Pralungo  
et al.

**Fig. 1.** Time interpolation. When there are no observations available at 00:00 or 12:00 UTC but only at other times, a reference value at the time of the asynoptic observation  $t_{obs}$  is calculated from the NOAA 20CR, employing a cubic interpolation with the 4 closest values. The difference  $Obs(t_{obs}) - 20CR(t_{obs})$  is assumed constant between  $t_{obs}$  and the closest synoptic time  $t_a$ . The observation at time  $t_a$  is, gained by adding the departure  $dep(t_{obs}) = CHUAN(t_{obs}) - 20CR(t_{obs})$  to  $20CR(t_a)$ .

Title Page

Abstract

Instruments

Data Provenance &amp; Structure

Tables

Figures

◀

▶

◀

▶

Back

Close

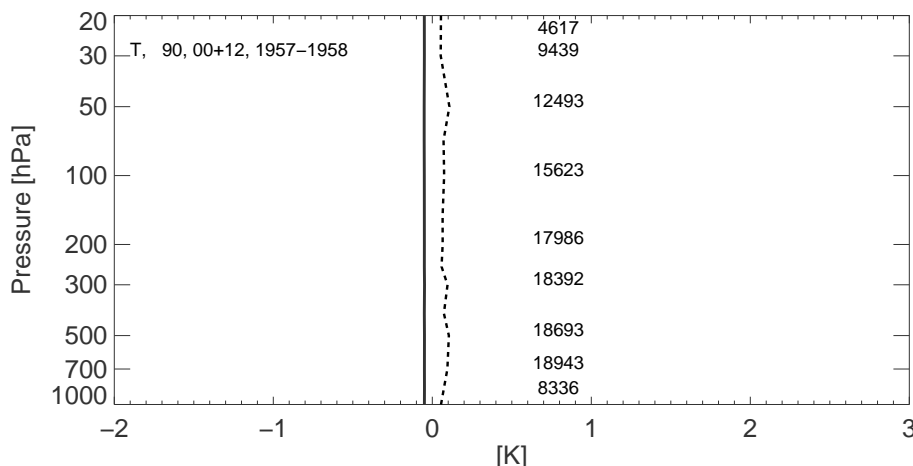
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

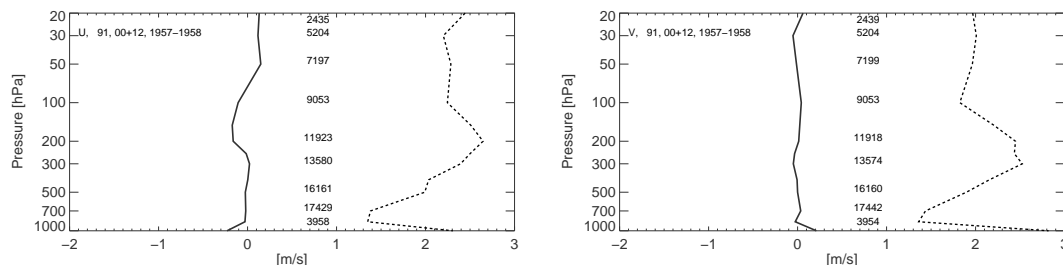


**Fig. 2.** Mean (solid) and RMS (dashed) temperature difference between observations from CHUAN and ERA-40 averaged over 90 stations at 00:00 and 12:00 UTC, for the years 1957–1958 in North America, on standard pressure levels. The total number of observations at each pressure level is also reported. Since those data are already on pressure levels, no altitude to pressure interpolation is needed. The constant vertical shift of 0.05 K is likely due to different conversion from Celsius to Kelvin. In this work, we have used  $0^{\circ}\text{C} = 273.15\text{K}$

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

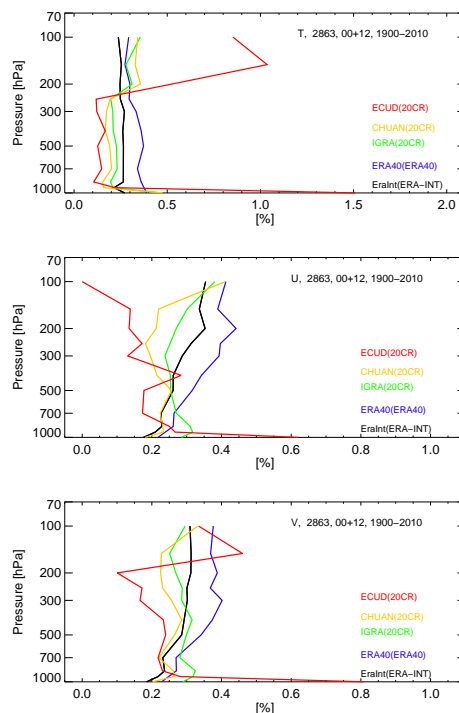


**Fig. 3.** Mean (solid) and RMS (dashed)  $U$  (left panel) and  $V$  (right panel) difference between observations in CHUAN and ERA-40, averaged over 91 stations at 00:00 and 12:00 UTC, for the period 1957–1958 in North America. The total number of observations at each pressure level is also reported. Differences come from different interpolation from altitude to standard pressure levels. Only WMO stations have been used.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

# Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

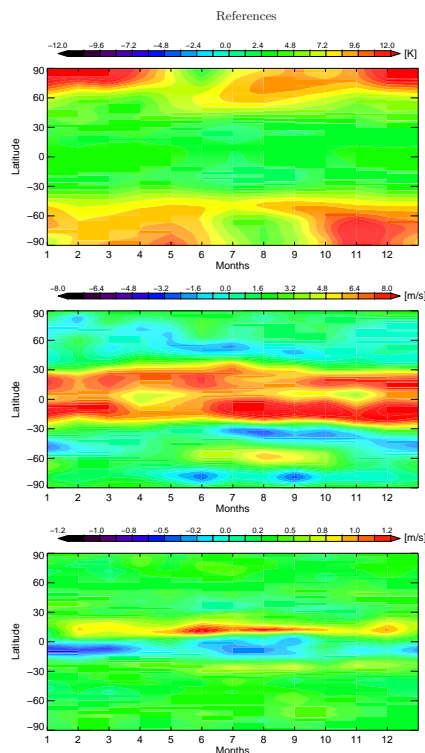


**Fig. 4.** Temperature (upper panels) and  $U$  (middle panels) and  $V$  (bottom panels) wind component spike frequency (%) as function of pressure, for the whole period 1900–2010. In the legend is reported the archive and, between brackets the reference used for the spike check (departures bigger than  $4\sigma$ ). Globally, the spikes individuate are always  $< 1\%$  except for the ECUD Temperature (only data prior to 1957) where at 150 hPa, for Temperature, the spikes density is roughly  $1\%$ . For all the others archives and variable, the density remains below  $0.5\%$ . Only WMO stations have been used.

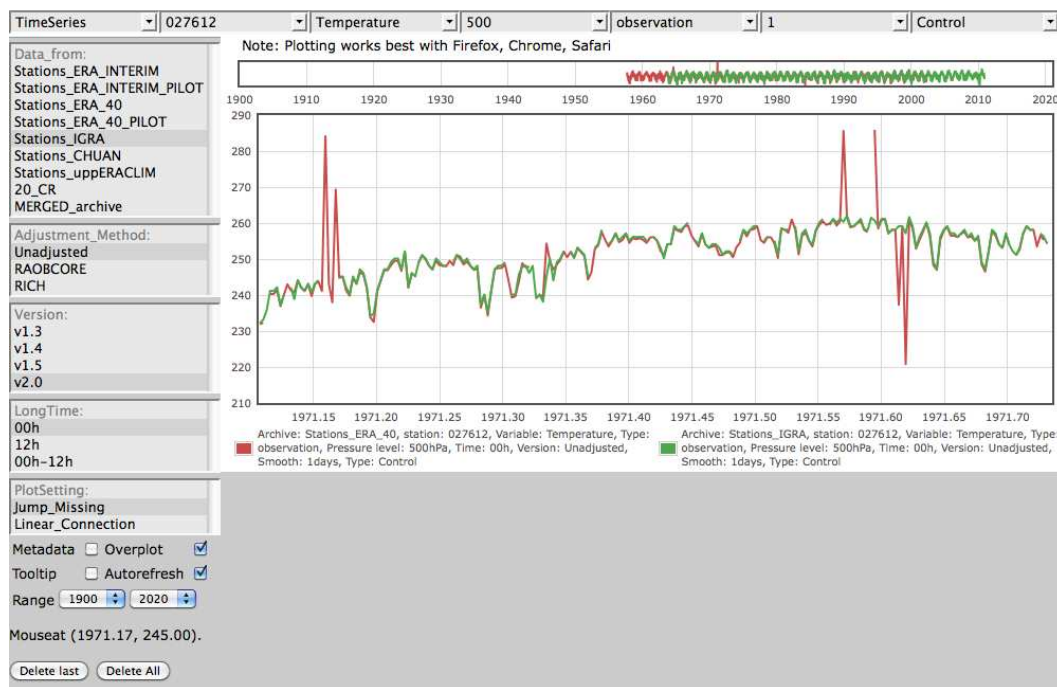
[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

# Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


**Fig. 5.** Temperature (top),  $U$  (middle) and  $V$  (bottom) Wind components: difference ERAInterim vs. NOAA-20CR, at 150 hPa mean over 00:00 and 12:00 UTC, for the year 1979. While  $V$  wind component shows only weak (less than  $1.5 \text{ ms}^{-1}$ ) bias while  $T$  and  $U$  wind components evidence strong differences. Remarkable are the discrepancy affecting temperature expectially in the polar regions and concentrate between October and May (up to 12 K). Opposite situation for the  $U$  wind component, where the difference is focused mainly in the tropical regions with amplitude in the range  $[-8, 8] \text{ ms}^{-1}$ .

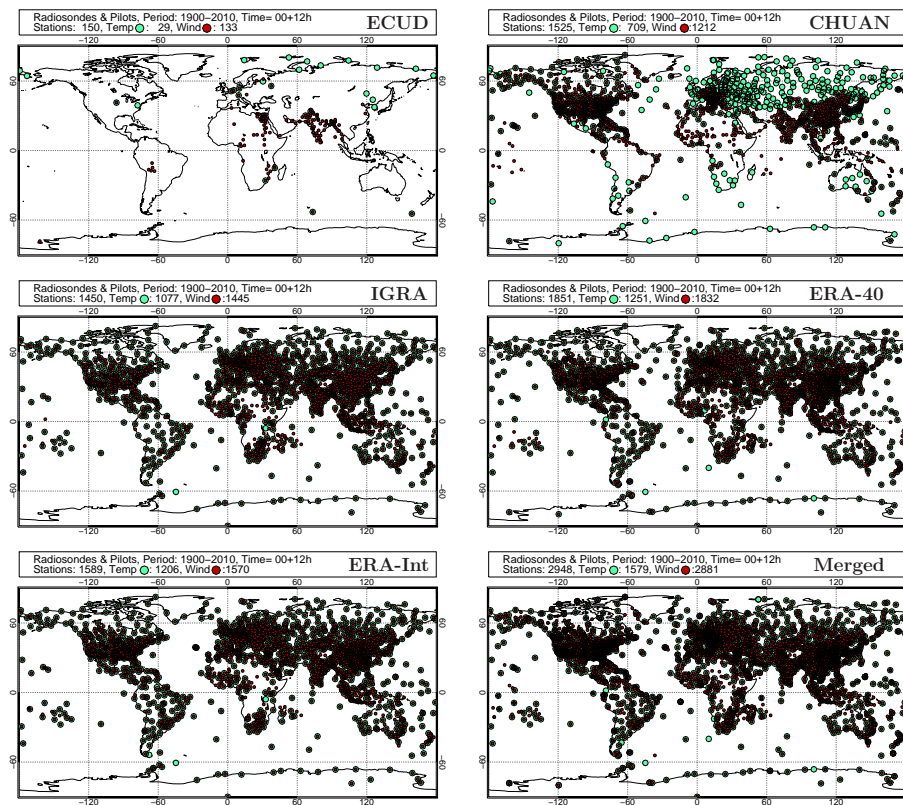


**Fig. 6.** Temperature observation time series of Moscow station (027612, Russia), at 500 hPa and at 00:00 h, green curve IGRA archive, red curve ERA-40 station archive. It is possible to see the two spikes on the left hand side of the plot: while no data are available in the IGRA archive, there are suspicious data in the ERA-40 archive. On the right hand side of the plot there are four suspicious values reported by the ERA-40 archive while, in the same days, the IGRA archive contains more plausible values.

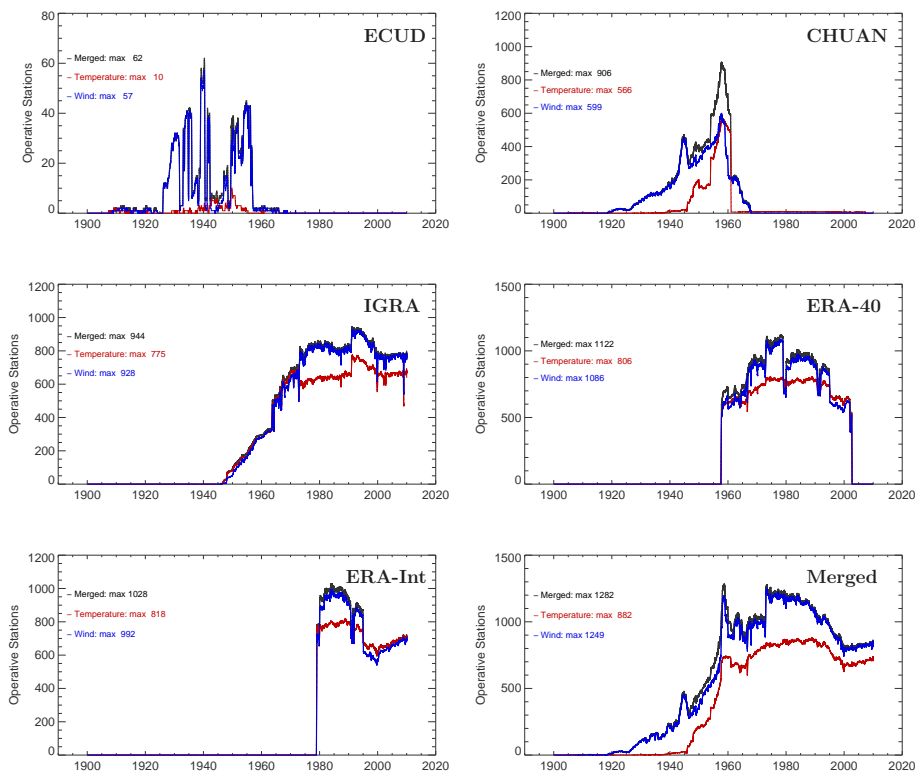


# Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


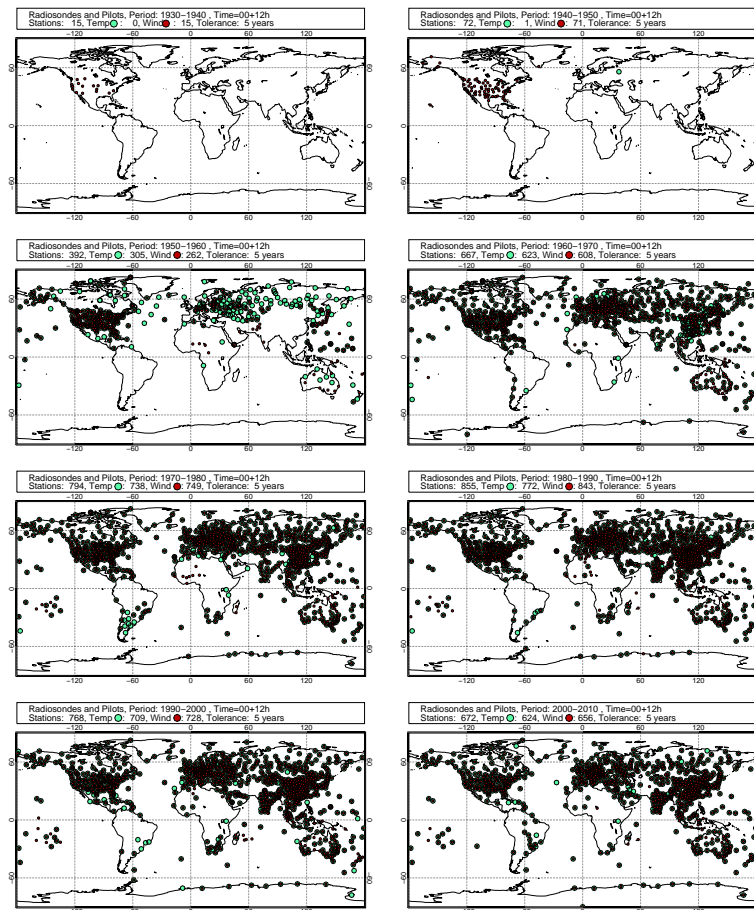
**Fig. 7.** Spatial distributions of stations with at least 365 days of data, at the pressure levels of 850, 700 and 500 hPa, in the different datasets used in this study. The bottom right panel shows the distribution in the merged archive. Only WMO stations have been used.



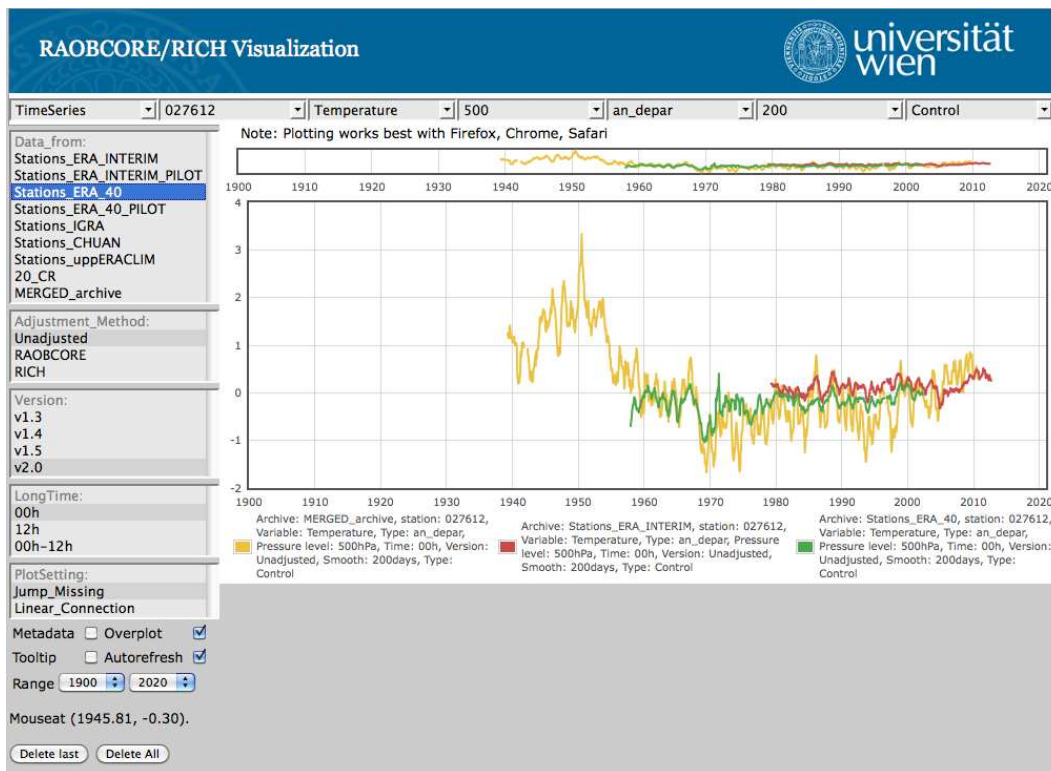
**Fig. 8.** Time series of number of active stations from the respective archives considered in this study. Only data from those stations are counted that have at least 365 ascents. The bottom right picture is the merged archive. Only WMO stations have been used.

# Radiosondes and Tracked Balloons Archive

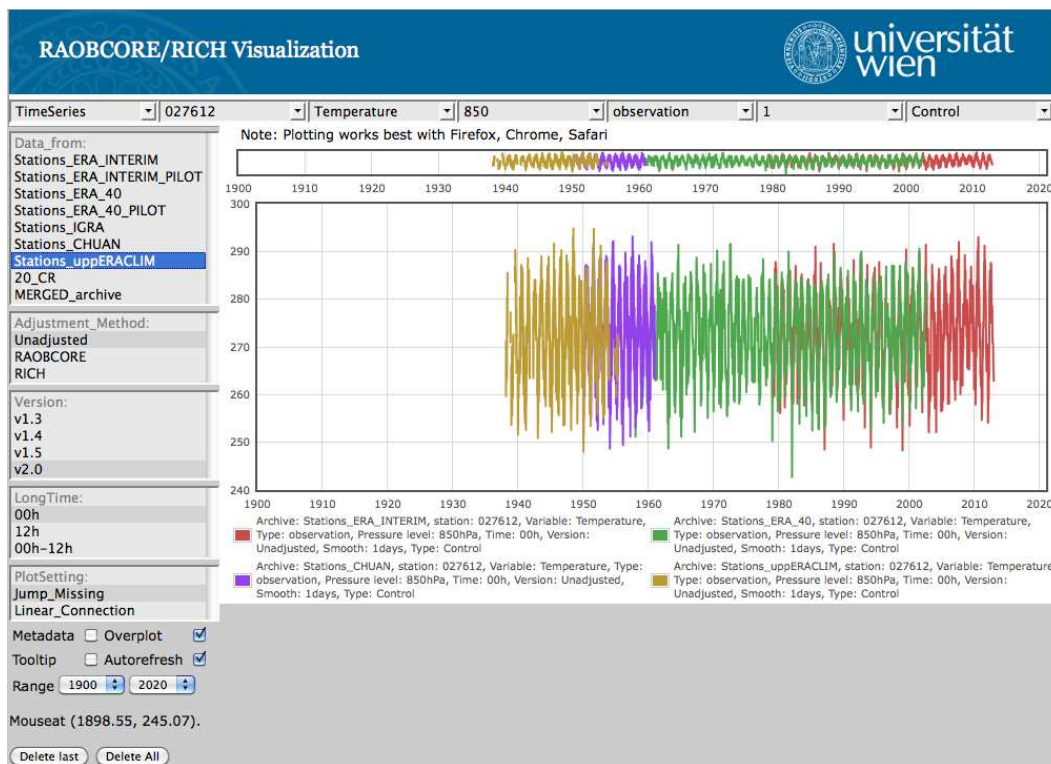
L. Ramella Pralungo  
et al.

[Title Page](#)[Abstract](#)[Instruments](#)[Data Provenance & Structure](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

**Fig. 9.** Decadal distribution of upper air stations in the merged archive with at least 5 yr of data in the range 850–500 hPa, starting 1930–1940 (upper left) to 2000–2010 (bottom right). Only WMO stations have been used.



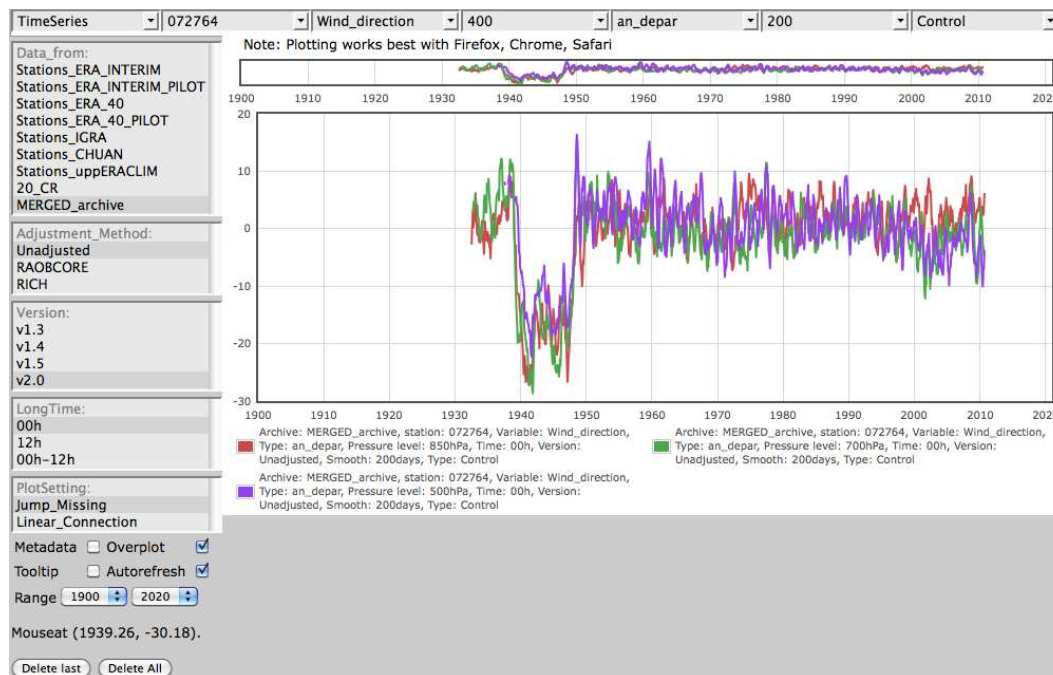
**Fig. 10.** Time series of departures between observations at Moscow station (027612, Russia) at 500 hPa and 00:00 UTC, and reference datasets derived from reanalysis efforts: obs-NOAA 20CR analysis (yellow); obs-ERA-40 (green) 6 h background forecasts, obs-ERA-Interim 12 h background forecasts (red). Even if the obs-NOAA 20CR show deeper wiggles compare to the other time series, it is still useful for detecting potential breaks in the observation time series, as can be seen from the jump in 1969 detectable in obs-NOAA20CR and obs-ERA40 departures, both.



**Fig. 11.** Observed temperature at Moscow station (027612, Russia) at 850 hPa, at 00:00 UTC. The plot shows the longest and complete temperature time series in the Merged archive, from 1938 up to now: red is ERA\_Interim, green ERA40, violet CHUAN and yellow ECUD. The not perfectly overlap between the common parts of the archives is a javascript artefact: it applies a smoothing function if the number of points exceed the plot resolution. Zoom in with the Time Serie Viewer the discrepancies vanish.

## Radiosondes and Tracked Balloons Archive

L. Ramella Pralungo  
et al.



**Fig. 12.** Station Municipal Bismarck (072764, North Dakota, USA), Wind Direction analysis departures (reference NOAA 20CR), at 850 (red), 700 (green) and 500 (violet) hPa. Well visible is the strong shift (rough 15°) in the period 1938/1948.

[Title Page](#)
[Abstract](#)
[Instruments](#)
[Data Provenance & Structure](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)