Three Daily Radiosonde Datasets from NOAA Project on "Homogenization of global radiosonde humidity data"

Liz Zhang and June Wang
Nov. 2012

I. Background

The main goal of the project "Homogenization of global radiosonde humidity data" is to homogenize radiosonde humidity data from individual soundings with records over the globe and produce a reliable humidity dataset for community use.

The efforts were made to collect and homogenize global radiosonde data and to make radiation bias corrections to Vaisala RS92 radiosonde profiles in order to get better reference segments for homogenization. Three datasets were created as a result of the project.

II. Three Datasets

1) **Raw Dataset:** This dataset contains daily radiosonde data at its native vertical resolution from the 1960s to the end of 2011.

This data set used the NOAA/NCDC IGRA (Integrated Global Radiosonde Archive) data as the base data (http://www.ncdc.noaa.gov/oa/climate/igra/index.php, Durre et al. 2006), which is compared to the NCEP ADP (Automated Data Processing) Global Upper Air Observational Weather data (http://rda.ucar.edu/datasets/ds351.0) to add new stations and fill in data gaps at existing IGRA stations.

New stations have to have at least 120 months of record, at least 10 records per month. Totally 99 new stations were added from NCEP ADP data. Some of new stations have data only in earlier years, mostly from 1973 to 1990. 27 out of 99 new stations have observations in 2011. At five of them, the data records started as early as 1973.

Existing stations with fewer soundings in some years will be replaced by NCEP APD data for that year, if NCEP ADP data have at least 50% more soundings for that station in that year.

NECP ADP data (1973-2000) were provided by Joey Comeaux from NCAR/CISL. ADP data after 2001 were downloaded from NCAR/CISL data archive ds351.0 (http://rda.ucar.edu/datasets/ds351.0)

2) **homoRaw Dataset**: This dataset contains daily radiosonde data at standard pressure levels from January 1945 to the end of 2011.

A homogenization method was applied to a radiosonde raw data from January 1945 (but around 1973 for most stations) to February 2009 (Dai et al. 2011). The raw data used in Dai et al. (2011) is different from the "Raw Dataset" described above, which contains the IGRA data supplemented by the CARDS (Comprehensive Aerological Reference Dataset) data and a special radiosonde dataset in China.

Since 03/2009, the IGRA data were reformatted as homogenized data format and appended to the homogenized data.

The homogenized data have different data format (see example data in Appendix) from the "Raw dataset", which only contain data on surface and 15 standard pressure levels (1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10mb). It has PW values in three layers (surface-500mb, 500-300mb and 300-100mb) in the sounding header, and have two additional columns, RH (in unit of 0.01%) and specific humidity (in unit of 0.00001g/kg). Temperature and DPD data are in unit of 0.01deg.C, while in the "Raw dataset", they are in unit of 0.1deg.C.

3) **homoRS92 Dataset**: This dataset contains daily radiosonde at standard pressure levels from January 1945 to the end of 2010.

The solar radiation dry bias was widely known for Vaisala RS92 soundings. A correction method was developed and applied to the data at 63 stations for the periods when Vaisala RS92 was used (Wang et al. 2013).

Vaisala introduced radiation corrections to their RS92 soundings during 2011, with unknown information about when and where those corrections were utilized. Therefore, this data set is only available to the end of 2010. The raw data at these 63 stations are extracted from the "Raw dataset" and first corrected for the solar radiation dry bias. Then the corrected data were homogenized using the method in Dai et al. (2011).

Data at other stations were directly copied from the "homoRaw dataset".

III. Comparisons of the three datasets

Monthly mean values were calculated for temperature (T), dew point depression (DPD) and relative humidity (RH) at 500mb and 700mb levels for all three data set at station 10393, 12374, 47646 and 94998. The time series are shown below (Figure 1-8). Temperatures from all three datasets look similar, while humidity values are quite different between "Raw" data and homogenized data. Homogenized humidity data (DPD and RH) are smoother along the time line, while raw values show jumps at both 500mb and 700mb due to heterogeneity. RS92 solar radiation corrections could have significant impact on humidity values for the whole data period (Figure 1&2) (Wang et al. 2013).

Reference:

Dai Aiguo, Junhong Wang, Peter W. Thorne, David E. Parker, Leopold Haimberger, Xiaolan Wang, 2011: A New Approach to Homogenize Daily Radiosonde Humidity Data. *J. Climate*, **24**, 965–991.

Durre Imke, Russell S. Vose and David B. Wuertz. 2006: Overview of the Integrated Global Radiosonde Archive. *Journal of Climate*: Vol. 19, No. 1, pp. 53-68.

Wang Junhong, Liangying Zhang, Aiguo Dai, Franz Immler, Michael Sommer and Holger Vomel, 2013: Radiation Dry Bias Correction of Vaisala RS92 Humidity Data and its Impacts on Historical Radiosonde Data. *J. Atmos. Oceanic Tech., 30, 197-214.*

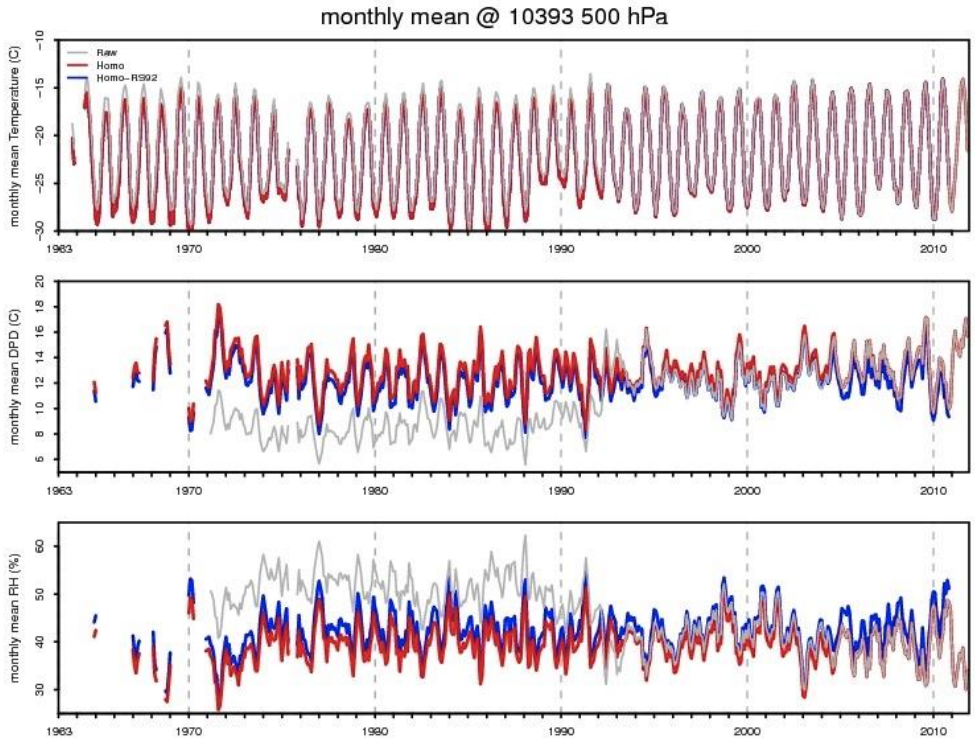

Figure 1. Time series for monthly mean values at 500hPa at Lindenberg, Germany (station #10393). Three panels are monthly mean temperature, dew point depression, and relative humidity from top to bottom. Data from "Raw Dataset" are in gray lines, data from "homoRaw Dataset" are in red lines, and data from "homoRS92 Dataset" are in blue lines.

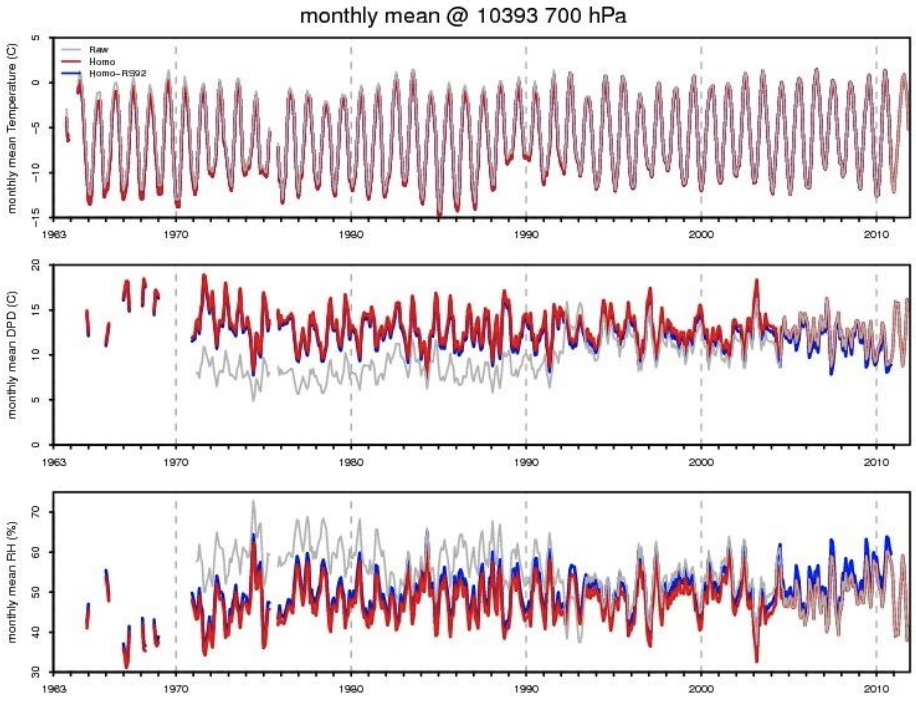

Figure 2. Same as Figure 1, but at 700hPa at Lindenberg, Germany (10393)

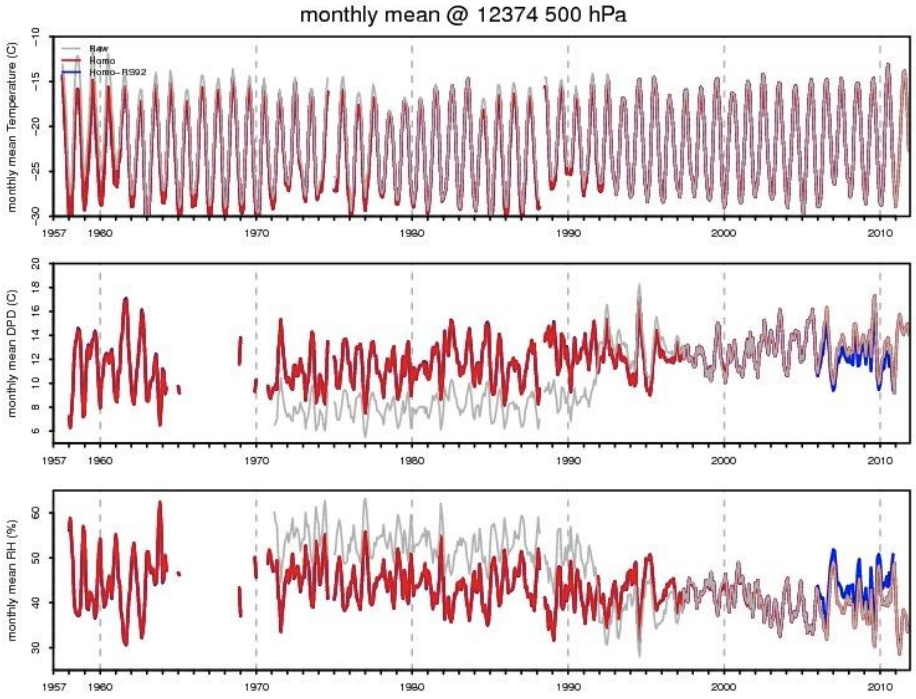

Figure 3. Same as Figure 1, but at 500hPa at Jozefoslaw, Poland (12374)

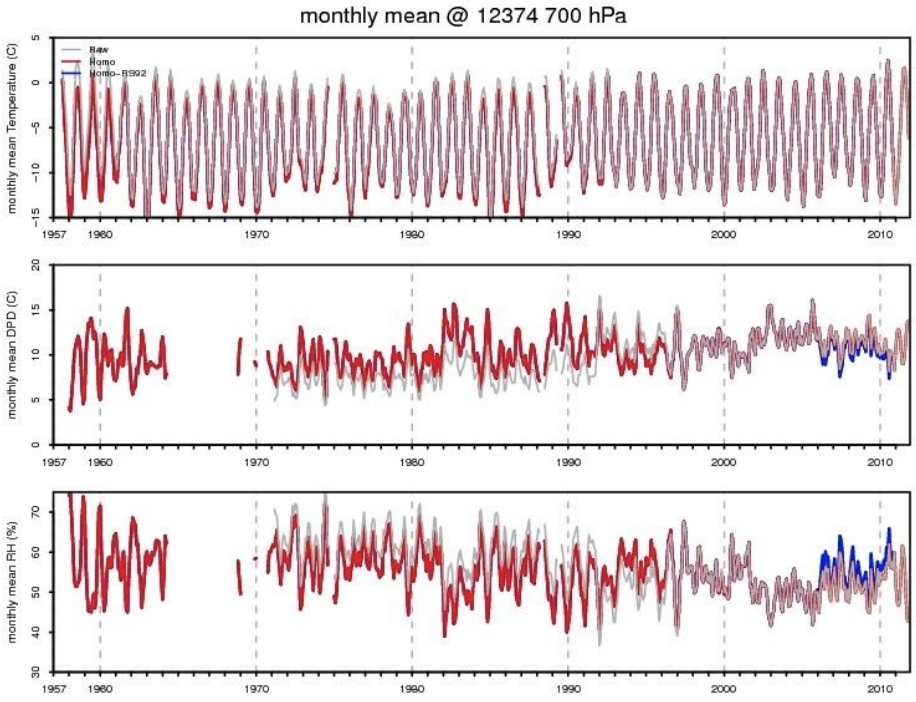

Figure 4. Same as Figure 1, but at 700hPa at Jozefoslaw, Poland (12374)

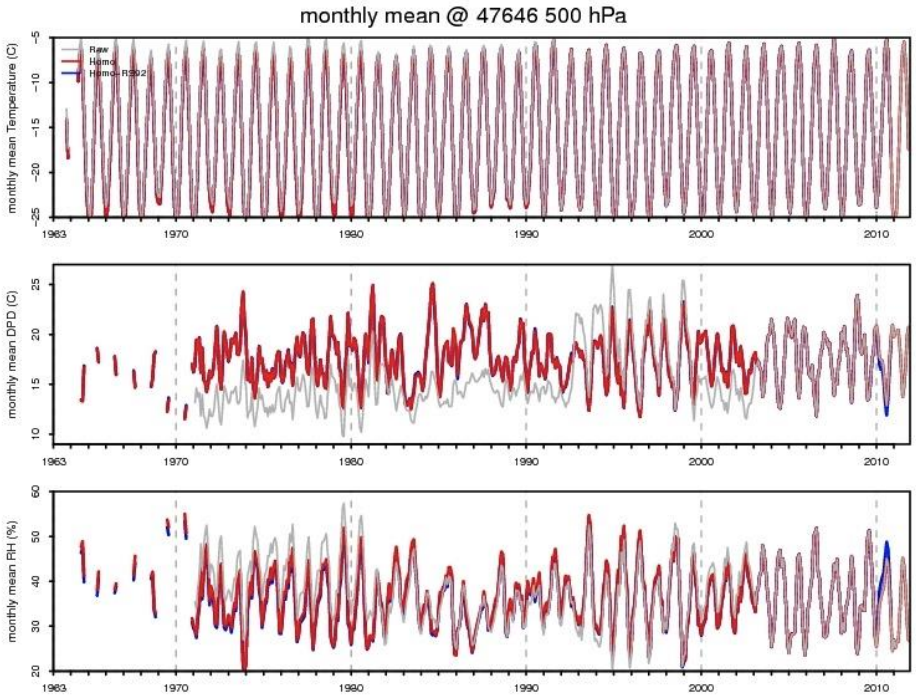

Figure 5. Same as Figure 1, but at 500hPa at Tateno, Japan (47646)

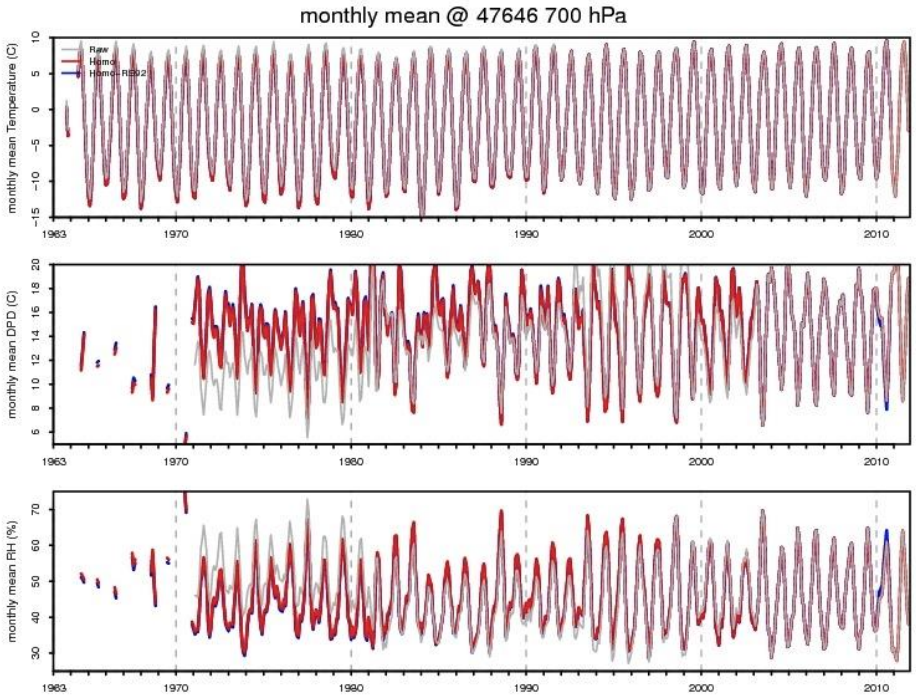

Figure 6. Same as Figure 1, but at 700hPa at Tateno, Japan (47646)

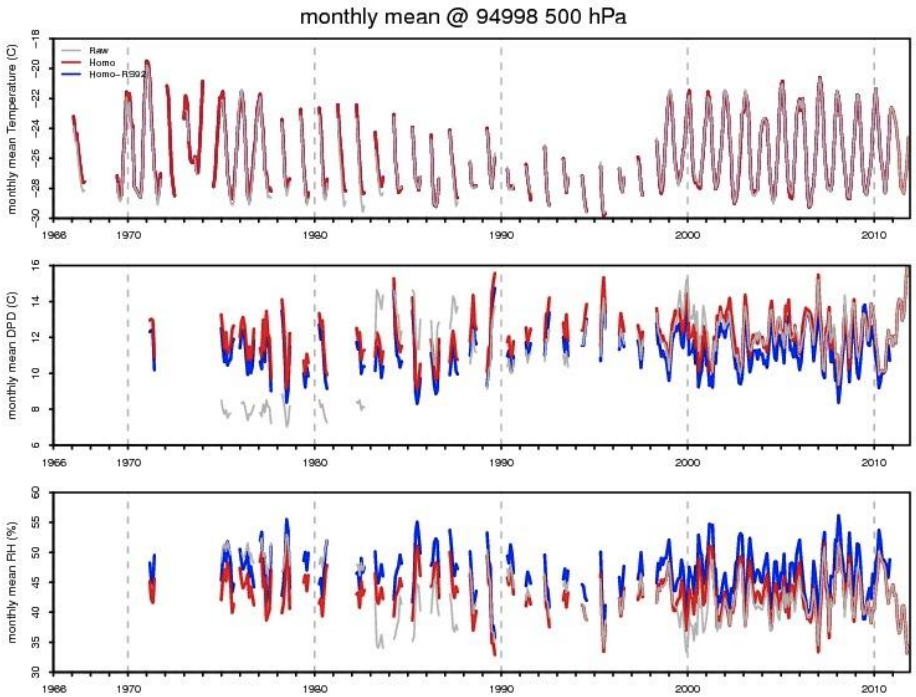

Figure 7. Same as Figure 1, but at 500hPa at Macquarie Island, Australia (94998)

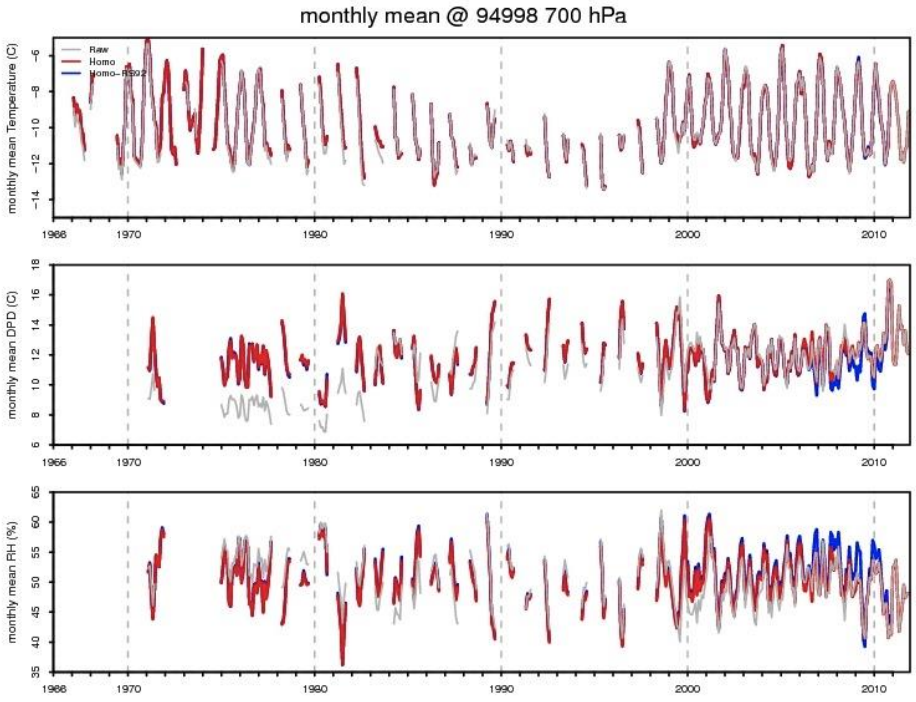

Figure 8. Same as Figure 1, but at 700hPa at Macquarie Island, Australia (94998)

Appendix:

Example data from "Raw Dataset"

```
#5451120111231239999  42
21103000B   55   -37B  140   70   30
10100000   265B  -43B  140   30   90
10 92500   877B  -65B  140  335  110
10 85000  1533B -101B  190  340  110
20 79200 -9999  -107B  200-9999-9999
10 70000  3018B -153B  210  340  150
```

Example data from "homoRaw Dataset" (same as "homoRS92")

```
#5451120111231239999  16 +2.1467E+00 +1.0126E-01 +4.3225E-02
21103000B   55   -370B  1400    70    30   92976  3292
10100000   265B  -430B  1400    30    90   91019  3273
10 85000  1533B -1010B  1900   340   110   40905  1955
10 70000  3018B -1530B  2100   340   150   24729  1482
```