Characterization of the Long-term Radiosonde Temperature Biases in the Upper Troposphere and Lower Stratosphere using COSMIC and Metop-A/GRAS Data from 2006 to 2014

Shu-peng Ho¹, Liang Peng¹, Holger Vömel²

¹ COSMIC Project Office, University Corporation for Atmospheric Research, Boulder, CO, USA
² National Center for Atmospheric Research, Boulder, CO, USA

Manuscript for Atmospheric Chemistry and Physics
September 2016

Shu-Peng Ho, COSMIC Project Office, Univ. Corp. for Atmospheric Research, P. O. Box 3000, Boulder CO. 80307-3000, USA (spho@ucar.edu)
Abstract

Radiosonde observations (RAOBs) have provided the only long-term global in situ temperature measurements in the troposphere and lower stratosphere since 1958. In this study, we use consistently reprocessed Global Positioning System (GPS) radio occultation (RO) temperature data derived from the COSMIC and Metop-A/GRAS missions from 2006 to 2014 to characterize the inter-seasonal and inter-annual variability of temperature biases in the upper troposphere and lower stratosphere for different radiosonde sensor types. The results show that the temperature biases for different sensor types are mainly owing to i) uncorrected solar zenith angle dependent errors, and ii) change of radiation correction. The mean RO-radiosonde global daytime temperature difference for Vaisala RS92 is equal to 0.22 K. The mean global daytime difference is equal to -0.12 K for Sippican, 0.87 K for VIZ-B2, 0.8 K for Russian AVK-MRZ, and 0.1 K for Shanghai. The global daytime trend of differences for Vaisala RS92 and RO temperature at 50 hPa is equal to 0.074 K/5yrs. Although there still exist uncertainties for Vaisala RS92 temperature measurement over different geographical locations, the global trend of temperature differences between Vaisala RS92 and RO from June 2006 to April 2014 is within +/-0.09 K/5yrs. Comparing with Vaisala RS80, Vaisala RS90 and sondes from other manufacturers, the Vaisala RS92 seems to provide the most accurate RAOB temperature measurements, and these can potentially be used to construct long term temperature Climate Data Records (CDRs). Results from this study also demonstrate the feasibility of using RO data to correct RAOB temperature biases for different sensor types.
1. Introduction

Stable, long-term atmospheric temperature climate data records (CDRs) with accurate uncertainty estimates are critical for understanding the impacts of global warming in both troposphere and stratosphere and their feedback mechanisms (Thorne et al., 2011; Seidel et al., 2011). Radiosonde observations (RAOBs) have provided the only long-term global in situ temperature, moisture, and wind measurements in the troposphere and lower stratosphere since 1958. Several groups have used multiple years of RAOB temperature measurements to construct long term CDRs (e.g., Durre et al., 2005; Free et al., 2004, 2005; Sherwood et al., 2008; Haimberger et al., 2008, 2011; Thorne et al., 2011; Seidel et al., 2009). However, it has long been recognized that the quality varies for different sensor types and height (e.g. Luers and Eskridge, 1995, Luers 1997, Luers and Eskridge 1998). Therefore, except for some sensor types where a relatively objective radiation correction had been applied (i.e., Vaisala RS90), it is difficult to objectively identify, trace, and remove most of the sensor-dependent biases for the historical sonde data and use the corrected RAOB temperatures to construct consistent temperature CDRs. The large uncertainties among temperature CDRs constructed from satellite and in situ measurements are still one of the most challenging issues for climate change researches (IPCC AR5).

The causes of temperature errors among RAOB sensor types include the changing of instruments and practices (Gaffen, 1994) and errors occurring due to the influence of solar and infrared radiation on the thermistor. In the past decade, many homogenization methods have been proposed to identify and correct errors due to changing of instruments and practice (Luers and Eskridge 1998; Lanzante et al., 2003;
Andrae et al., 2004; Free et al., 2004, 2005; Sherwood et al., 2008; Haimberger et al., 2008, 2011; Thorne et al., 2011; Seidel et al., 2009). Possible errors due to changes of instruments were identified by comparing with temperature measurements from adjacent weather stations. However, this approach is limited by the low number of co-located observations and large atmospheric variability. In addition, due to lack of absolute references, the remaining radiation temperature biases from adjacent stations may not be completely removed. As a result, only relative temperature differences of a possibly large uncertainty among stations are identified.

To correct possible RAOB temperature errors due to radiative effects, Andrae et al., (2004) and Haimberger et al., (2007, 2008, 2011) calculated temperature differences between observations and reanalyses data which were then used to minimize the differences between daytime and nighttime temperature differences. Nevertheless, because changes of reanalysis systems and possible incomplete calibration of satellite instruments may complicate the temperature bias correction, long-term stability of the derived temperature trends is still of great uncertainty. To correct the RAOB solar/infrared radiation errors, radiation correction tables (for example, RSN96, RSN2005 and RSN2010 tables from Vaisala) were introduced by manufactures to correct for radiation errors of particular sensors. However, when and how exactly different countries start to apply these corrections and whether there are remaining uncorrected radiative effects over different geographic regions are still unknown. It is important to use stable and accurate temperature references to characterize these errors from multiple sensors in different geographical regions over a long period of time.

The fundamental observable (time delay) for the Global Positioning System (GPS)
radio occultation (RO) satellite remote sensing technique can be traced to ultra-stable international standards (atomic clocks) on the ground. While time delay and bending angles are traceable to the international standard of units (SI traceability), the derived temperature profiles are not. To investigate the structural uncertainty of RO temperature profiles, Ho et al., (2009a and 2011) compared CHAMP (CHAllenging Minisatellite Payload) temperature profiles generated from multiple centers when different inversion procedures were implemented. Results showed that the mean RO temperature biases for one center relative to the all center mean is within ±0.1K from 8 km to 30 km, except for the South Pole above 25 km.

The mean temperature difference between the collocated soundings of COSMIC (Constellation Observing System for Meteorology, Ionosphere, and Climate) and CHAMP was within 0.1 K from 200 hPa to 20 hPa (Ho et al., 2009b; Anthes et al., 2008; Foelsche et al., 2009). At 20 hPa, the mean temperature difference between COSMIC and CHAMP was within 0.05K (Ho et al., 2009b). Schreiner et al. (2014) compared reprocessed COSMIC and Metop-A/GRAS (Meteorological Operational Polar Satellite–A/Global Navigation Satellite System (GNSS) receiver for Atmospheric Sounding) bending angles and temperatures produced at COSMIC Data Analysis and Archive Center (CDAAC). The mean layer temperature difference between 200 hPa to 10 hPa was within 0.05 K where the mean temperature difference at 20 hPa is equal to 0.03K. This demonstrates the consistency of COSMIC and Metop-A/GRAS temperatures.

The precision of RO temperature is ~ 0.1 K (Anthes et al., 2008; Ho et al., 2009a), and the precision of the trend of RO-derived temperature data is within ±0.06 K/5yrs (Ho et al., 2012). To estimate the uncertainty of RO temperature in the upper troposphere and
lower stratosphere, Ho et al., (2010) compared RO temperature from 200 hPa to 10 hPa to those from Vaisala-RS92 in 2007 where more than 10,000 pairs of coincident Vaisala-RS92 and COSMIC data were collected. The mean bias in this height range was equal to -0.01 K with a mean standard deviation of 2.09 K. At 20 hPa, the mean bias was equal to -0.02 K. These comparisons demonstrate the quality of RO temperature profiles in this height range.

RO derived atmospheric variables have been used as reference to identify RAOB sensor dependent biases. For example, Kuo et al., (2004) used RO data to identify sensor type dependent refractivity biases. Ho et al., (2010a) demonstrated that RO-derived water vapor profiles can be used to distinguish systematic biases among humidity sensors. He et al., (2009), hereafter He2009 and Sun et al. (2010, 2013) used RO temperature data in the lower stratosphere to quantify the temperature biases for several sensor types. While He2009 used the COSMIC post-processed temperature profiles from August 2006 to February 2007 to quantify the radiosonde radiation temperature biases for different sensor types, Sun et al., (2010; 2013) used COSMIC real-time processed temperature profiles to identify radiosonde temperature biases for numerical weather prediction analysis. Because complete GPS orbital information is not available in real-time, approximate GPS orbital information was used in the real-time inversion processing. The differences between real-time and post-processed RO temperatures in the lower stratosphere range from 0.3 K to 0.1 K depending on the comparison period. Although real-time COSMIC data, which are processed by using periodically revised inversion packages, may be suitable for weather analysis, they may not be suitable for climate studies. Both of these RAOB-RO comparisons are constructed from a relatively limited
A consistent validation of the variability of inter-seasonal and inter-annual RAOB temperature biases over a longer time period (close to ten years) for different temperature sensor types has not yet been done.

Recently, the UCAR CDAAC has developed an improved reprocessing package, which is used to consistently process RO data from multiple years of multiple RO missions including COSMIC (launched in April 2006) and Metop-A/GRAS (launched in October 2006). A sequence of processing steps is used to invert excess phase measurement to retrieve atmospheric variables including bending angle, refractivity, pressure, temperature, and geo-potential height.

The new inversion package uses improved precise orbit determination (POD) and excess phase processing algorithm, where a high-precision, multiple GNSS data processing software (i.e., Bernese Version 5.2, Dach et al., 2015) is applied for clock estimation and time transfer. In the reprocessing package, the POD for COSMIC and Metop-A/GRAS are implemented separately (Schreiner et al., 2011). The re-processed RO data produce more consistent and accurate RO variables than those from post-processed (periodically updated inversion packages were used) and real-time processed datasets.

The objectives of this study are to use consistently reprocessed GPS RO temperature data to characterize i) solar zenith angle (SZA) dependent temperature biases, ii) potential residual temperature errors due to incomplete radiation correction, iii) temperature biases due to change of radiation correction over different geographical regions, iv) the inter-seasonal and inter-annual variability of these temperature biases, and v) the trends of these biases and their uncertainty for different sensor types in the
upper troposphere and lower stratosphere. In contrast to previous studies (i.e., He2009 and Sun et al. 2010, 2013) that used shorter time periods, close to 8 years (from June 2006 to April 2014) of consistently reprocessed temperature profiles derived from COSMIC and Metop-A/GRAS are used. Because the quality of RO data does not change during the day or night and is not affected by clouds (Anthes et al., 2008), the RO temperature profiles co-located with RAOBs are useful to identify the variation of temperature biases over time of different temperature sensors.

In Section 2, we describe the RO and RAOB data and the comparison method. The global comparison of RO-RAOB pairs for different temperature sensor types for daytime and nighttime are summarized in Section 3. The global SZA dependent temperature biases for various sensor types at different geo-graphical regions are also compared in this section. The inter-seasonal variations of RAOB-RO temperature biases are assessed in Section 4. We conclude our study in Section 5.

2. Data and Comparison Method

2.1 RAOB data

The radiosonde data used in this study were downloaded from CDAAC (http://cosmic.cosmic.ucar.edu/cdaac/index.html). The data include the temperature, pressure and moisture profiles generated from the original radiosonde data in the NCAR data archive (http://rda.ucar.edu/datasets/ds351.0), which provides global radiosonde data with the detailed instrument type.

There are more than 1100 radiosonde stations globally. Figure 1 depicts the geophysical locations for all RAOB data from June 2006 to April 2014. These include
Vaisala RS80, RS90, RS92, AVK-MRZ (and other Russian sondes), VIZ-B2, Shippican MARK II A, Shanghai (from China), and Meisie (Japan). Table 1 summarizes the availability for different instrument types. In total, seventeen different types of radiosonde systems were used. The solar absorptivity ($\alpha$) and sensor infrared emissivity ($\varepsilon$) for the corresponding thermocap and thermistor for different instrument types are also summarized in Table 1. Most of the radiosonde data are collected twice per day.

Because the Vaisala RS80 sensor was never changed and should be the same across all RS80 models and the software uses the same radiation correction table which should not show any differences, we do not further separate Vaisala RS80 sensors (i.e., ID=37, 52, 61, and 67). For the same reason, all RS92 sensors (ID=79, 80, 81) are summarized together and all Russian sensors (ID=27, 75, 88, 89, 58) are summarized as AVK sonde (see Table 2 and Section 3.1).

### 2.2 GPS RO data

The re-processed COSMIC (Version 2013.3520) and Metop-A/GRAS (Version 2016.0120) dry temperature profiles downloaded from UCAR CDAAC (http://cosmic.cosmic.ucar.edu/cdaac/index.html) are used in this study. With six GPS receivers on board LEO satellites, COSMIC produced about 1000 to 2500 RO profiles per day since launch. With one receiver, Metop-A/GRAS produced about 600 RO profiles per day. The detail inversion procedures of COSMIC Version 2013.3520 and Metop-A Version 2016.0120 are summarized in http://cdaac-www.cosmic.ucar.edu/cdaac/doc/documents/Sokolovskiy_newroam.pdf. The general description of CDAAC inversion procedures is detailed in Kuo et al., (2004), and Ho et
In a neutral atmosphere, the refractivity \( N \) is related to pressure \( (P \text{ in hPa}) \), temperature \( (T \text{ in K}) \) and the water vapor pressure \( (e \text{ in hPa}) \) according to Smith and Weintraub (1953):

\[
N = 77.6 \frac{P}{T} + 3.73 \times 10^5 \frac{e}{T^2}.
\]

Because in the upper troposphere and stratosphere moisture is negligible, the dry temperature is nearly equal to the actual temperature (Ware et al., 1996). In this study, we use RO dry temperature from 200 hPa to 20 hPa to quantify the temperature biases for different sensor types.

### 2.3 Detection of RAOB Temperature Biases Using RO Data over Different Geographical Regions

The RO atmPrf data from COSMIC and Metop-A/GRAS were first interpolated into the mandatory pressure level of the radiosondes (i.e., 200, 150, 100, 50, and 20 hPa). To account for the possible temporal and spatial mismatches between RO data and RAOBs, the RO data within 2 hours and 300 km of the radiosonde data were collected for different ROAB instrument types. These matching criteria are similar to the criteria used by He2009. However, in contrast to He2009, positions of RO measurements at the corresponding heights are used in the RAOB-RO ensembles. We compute temperature differences between RO atmPrf and the corresponding RAOB pairs in the same pressure level \( i \) using the equation
\[ \Delta T(i,j) = \left( \frac{1}{n} \right) \times \sum_{s=1}^{s_{max}} \{ T_{RAOB}(i,j,s) - T_{RO}(i,j,s) \} , \]  

(2)

where \( j \) is the index for eighteen instrument type listed in Table 1, and \( s \) is the index for all the matched pairs for each of seventeen instrument types.

In addition, we compare the monthly mean temperature biases \( \Delta T^{\text{Time}} \) for the matched pairs at different geo-graphical regions from

\[ \Delta T^{\text{Time}}(l,m,k) = T_{RAOB}(l,m,k) - T_{RO}(l,m,k) , \]  

(3)

where \( l, m, \) and \( k \) are the indices of the month bin for each vertical grid \( (l) \), zone \( (m) \) and month for the whole time series \( (k = 1 \) to 95) from June 2006 to April 2014, respectively. The geographical zones \( (m) \) are from USA \( (m=1) \), Australia \( (m=2) \), Germany \( (m=3) \), Canada \( (m=4) \), United Kingdom \( (m=5) \), Brazil \( (m=6) \), Russia \( (m=7) \), China \( (m=8) \), and Japan \( (m=9) \), respectively. The standard deviation of the time series is also computed to indicate the variability of \( \Delta T^{\text{Time}} \). In this study, daytime data are from SZA from 0 to 90 degree and nighttime data are from SZA from 90 to 180 degrees. The SZA is computed from the synoptic launch time and location of sonde station since the time and location of the sonde at different heights are not available.

3. Global Mean RAOB Temperature Biases for all Sensor Types Identified by RO Data

RS92 (ID=79,80,81) data were used in this study. Since 1981, Vaisala RS80
(from 1981 to 2014), RS90 (from 1995 to 2014), and RS92 have been widely used for numerical weather prediction (NWP) and atmospheric studies. For many modern radiosondes (for example RS92) the structural uncertainties are +/- 0.2 K below 100 hPa and somewhat higher at higher levels. While the Vaisala data have been corrected for possible radiation errors (see RS92 Data Continuity link under the Vaisala website), some radiation corrections were also made for other sensor types, although they may not be clearly indicated in the Metadata files. We quantify the global mean residual radiation correction biases for all sensor types in this section.

3.1 The RAOB Temperature Biases during the Daytime and Nighttime for All Sensor Types

In total, we have more than 600,000 RAOB-RO pairs. Using Eq. (2), we compute the temperature biases of radiosonde measurements for each individual sensor type. The mean temperature bias for ensembles of the RAOB-RO pairs from June 2006 to April 2014 for the layer between 200 hPa and 20 hPa for different RAOB sensor types is summarized in Table 2. The standard deviations for each radiosonde type are also shown. The radiosonde temperature biases vary for different sensor types. All biases are less than 0.25 K, except for AVK and VIZ-B2, which reach 0.66 and 0.71 K respectively during the day.

The solar radiation effect on sensors is the dominant error source of RAOB temperature biases (Luers et al., 1998 and He2009). We assume that all operational data have a radiation correction already applied. The global temperature biases relative to the co-located RO temperature at 50 hPa for various radiosonde sensor types for daytime and
nighttime are shown in Figure 2. Only those stations containing more than 50 RO-ROAB pairs are plotted. Figure 2a shows biases for different sensor types, which vary with geographical region. Most of the sensor types contain positive temperature biases ranging from 0.1 to 0.5 K during the daytime. This bias during daytime may be a result of the residual error of the systematic radiation bias correction. Although we only include stations containing more than 50 RO-RAOB pairs, some level of heterogeneity (i.e., Fig. 2a over Brazil) may be due to low sample numbers. For example, stations with temperature biases larger than 0.5 K in eastern Brazil contain only about 60 RO-RAOB pairs. The cause of the heterogeneity in temperature bias between North and South China is not certain at this point.

The mean nighttime biases are very different from those in the daytime for the same sensors. Figure 2b shows that most of the sensor types show a cold bias at night except for Vaisala in South American, Australia, and Europe. The mean biases at night for the two sonde types with the largest warm bias at daytime (AVK and VIZ-B2) are equal to -0.06 K and -0.42 K, respectively (Table 2). The scatter of $\Delta T$ is similar for all sonde types during the day and night with standard deviations between 1.50 K and 1.71 K (Table 2).

The global mean $\Delta T$ for the Vaisala RS92 of 0.16 K during the comparison period is slightly larger than the temperature comparison between COSMIC and Vaisala RS92 in 2007 (Ho et al., 2010b) (~0.01 K) and in He2009 (~0.04 K from ~200 hPa to 50 hPa). This could be in part because more RS92-RO pairs from lower solar zenith angle regions (for example, from the southern Hemisphere and near Tropics, see Section 3.2) are included after 2007 (see section 4).
3.2 Solar Zenith Angle Dependent Temperature Biases for Vaisala Sondes

More than 50% of RAOB data are from Vaisala sondes, from a number of different countries. In total, 161,019 RS92 (ID=79, 80, 81) ensemble pairs are distributed in all latitudinal zones during the daytime. To quantify a possible residual radiation correction error for Vaisala RS92 measurements in the lower stratosphere, which may vary with SZA, we compare the mean temperature differences from 200 hPa to 20 hPa for daytime and nighttime over different regions in Figures 3 and 4, respectively.

Figure 3 indicates that RS92 in different regions demonstrate a similar quality in terms of mean differences from RO with a small warm bias above 100 hPa, as well as similar standard deviations relative to the mean biases of approximately 1.5K. Because some stations in the United States are only interested in the tropospheric profiles and use smaller balloons, fewer RO-RS92 samples are available above 70 hPa compared to those in other countries.

Figure 4 depicts the mean RS92-RO temperature differences from 200 hPa to 20 hPa for nighttime. The nighttime RS92 data over different regions show similar standard deviations of about 1.5 K compared to those at daytime. In most of the regions, the mean nighttime temperature biases are similar to those in the daytime results, with small (0.1-0.2 K) warm biases above 100 hPa. These residual nighttime warm biases are not seen in the ROAB-RO ensemble pairs for Sippican MARK, VIZ-B2, AVK, and Shanghai Sondes (see Section 3.3). This 0.1 K – 0.2 K warm bias for RS92 at night could be due to calibration of the RS92 temperature sensor (see Dirksen et al., 2014).

Because the quality of RO temperature is not affected by sunlight, the small but
obvious geographic-dependent biases are most likely due to the residual radiation correction for RS92 and when and how different countries apply the radiation correction (see Section 4.1).

To consider a possible SZA dependence of the temperature bias due to residual radiation errors for Vaisala RS92, we bin the computed temperature differences in 5-degree bins at each of the ROAB mandatory pressure levels above 200 hPa using all the RAOB-RO ensembles. Figure 5 depicts the temperature biases at 50 hPa as function of SZA in six regions. Only those bins that contain more than 50 RAOB-RO pairs are included. Zero SZA is at noon and 90 degrees SZA corresponds to sunrise or sunset. Figure 5 shows that the daily mean difference varies from 0.09 K (Canada) to 0.31 K (Brazil), with a slightly larger warm bias for low SZA (near noon) than that at higher SZA (late afternoon and in the night).

3.3 Temperature Biases for Sippican MARK, VIZ-B2, AVK-MRZ, and Shanghai Sondes

Unlike Vaisala sondes, which are distributed in almost all latitudinal zones, other sonde types are distributed mainly in the northern mid-latitudes. Fig. 6 depicts the mean temperature differences from 200 hPa to 20 hPa in the daytime for Sippican, VIZ-B2, AVK, and Shanghai. The biases for VIZ-B2 and AVK-MRZ are positive everywhere above 200 hPa, with means of about 0.7K. The biases are smaller for Sippican and Shanghai. These mean biases are similar to those from He2009. The small differences between these and He2009 results are likely due to the sampling differences between He2009 (August 2006 to February 2007, or 7 months) and this study (95 months).
Fig. 7 depicts the mean temperature differences from 200 hPa to 20 hPa in the nighttime also for Sippican, VIZ-B2, AVK-MRZ, and Shanghai. The nighttime biases are generally less than 0.1K except from VIZ-B2 above 100 hPA where they exceed 0.5K. The small positive values for VIZ-B2 and AVK-MRZ, which were present in the daytime (Fig. 6) are not present during the night (Fig. 7).

We also bin the temperature differences for these four sonde types in 5-degree SZA bins for each mandatory pressure levels above 200 hPa using all the RAOB-RO pairs from June 2006 to April 2014. Only those bins contains more than 50 RAOB-RO pairs are included. Figure 8 depicts the differences at 50 hPa as a function of SZA for Sippican MARK, VIZ-B2, AVK-MRZ, and Shanghai.

The VIZ-B2 sonde has a large warm bias (as high as 1.75 K) during daytime and a cold bias (as low as -0.8K) at night. AVK has a bias from about 0.7 K to 1.1 K in the daytime where its nighttime biases are close to zero. The mean biases for the Sippican and Shanghai sondes show less diurnal variation and are 0.08 and -0.17 K respectively.

4. Comparison of the Seasonal RAOB Temperature Biases in different Regions

Since there is some residual radiation error, we characterize the long-term stability of RAOB temperature measurements for different RAOB sensor types by quantifying their seasonal temperature biases relative to those of co-located RO data.

4.1 Identification of RS92 Temperature Biases due to Change of Radiation Correction

The Vaisala RS92 radiosonde was introduced in 2003 and is scheduled to be replaced by the Vaisala RS41 in 2017. Vaisala included a reinforcement of the RS92
sensor in 2007, which impacted the radiation error. To account for this sensor update, the radiation correction tables were updated in 2011 (RSN2010, software version 3.64), which is used to replace the original radiation correction table. Between 200 and 20 hPa, the correction in RSN2010 is about 0.1 K larger than in RSN2005 (see http://www.vaisala.com/en/products/soundingsystemsandradiosondes/soundingdatacontinuity/RS92DataContinuity/Pages/revisedsolarradiationcorrectiontableRSN2010.aspx). It is likely that each country updated the correction table for their entire network. However, when exactly each country implemented these updated tables is unknown.

To identify possible RS92 temperature biases due to changes of the radiation correction table (i.e., RSN2010), we compare the mean $\Delta T$ from January 2007 to December 2010 to those from January 2011 to April 2014 over the United States, Australia, Germany, Canada, United Kingdom, and Brazil (Figs. 9a-f). There is no consistent pattern of differences in these two periods over the six regions, with mean differences ranging from -0.122 K (Australia) to 0.047 K (United States). The small differences in profile shapes and magnitudes are an indication of the magnitude of the uncertainty in RS92 temperatures due to differences in implementing the radiation correction tables.

The Deutscher Wetterdienst (DWD), Germany’s Meteorological Service, implemented the updated radiation correction for the Vaisala RS92 in the spring of 2015 rather than 2011, to avoid inconsistencies with corrections already implemented in their data assimilation system. This may in part explain the greater consistency of $\Delta T$ over Germany for these two time periods than over other countries. This also indicates the importance of establishing traceability through careful documentation and metadata
tracking, which is especially important for using radiosonde data in climate studies. The relatively small temperature difference between these two periods over the United States is most likely a statistical artifact due to the very small number of coincidences in this period.

4.2 Time Series and trends of de-seasonalized radiosonde-RO differences

In this section we look at time series and trends in the de-seasonalized radiosonde-RO temperature differences from 2007 to 2014 in order to determine the long-term stability of these differences. Ideally, if both radiosondes and RO were free of biases, the time series would be stable and show small differences near zero with small standard deviations and no trends. We choose 50 hPa for showing these time series because the biases tend to be larger at this level than at lower levels. We also computed time series for 150 hPa, but except for lower biases, the results were similar to those at 50 hPa (not shown).

Figure 10 shows daytime and nighttime time series of monthly mean temperature biases at 50 hPa for Vaisala RS92 for the United States, Australia, Germany, Canada, United Kingdom, and Brazil. Table 3 summarized the mean and std of the monthly mean temperature differences for RS92 and RO at 50 hPa. Fig. 10 indicates that there is little variation over time in the monthly mean temperature differences at 50 hPa in all six regions, with little difference between day and night values. The magnitudes of the mean biases range from -0.01 K for Canada to over 0.2 K in Australia, Germany and Brazil. The standard deviations range from a low of 0.18 K (Australia, day) to a high of 0.46 K (United States, night). The small (less than 0.5 K) standard deviation for RS92 over
daytime and nighttime over these six regions demonstrates the long-term stability of RS92 data.

Figure 11 shows the daytime and nighttime time series of monthly mean temperature biases for each of the other sensor types at 50 hPa in North hemisphere mid-latitude (60°N-20°N) are also summarized in Table 4, respectively. All daytime biases are below 0.25 K in magnitude, except for Russia (0.8 K) and VIS-B2 (0.87 K). The magnitudes of the mean nighttime biases are all less 0.25 K except for VIS-B2, which is -0.56 K. The daytime biases for Russia and VIS-B2 contain obvious inter-seasonal variation.

Figure 12 shows daytime and nighttime time series of monthly mean de-seasonalized temperature biases at 50 hPa for Vaisala RS92 for the United States, Australia, Germany, Canada, United Kingdom, and Brazil. Table 3 summarizes the trends of the de-seasonalized temperature differences, and shows the de-seasonalized trends in RO temperatures for comparison. The root mean square (RMS) of the de-seasonalized time series (RMS of difference) in Table 3 indicates the trend uncertainty of the time series.

The de-seasonalized temperature differences are computed from

\[
\Delta T^{\text{Deseason}}(l,m,k) = T^{\text{RAOB}}(l,m,k) - T^{\text{Time}}(l,m,k'),
\]

where \(l\), \(m\), and \(k\) are the indices of the month bin for each layer \((l)\), zone \((m)\) and month for the whole time series \((k = 1\) to 95), respectively, and \(k'\) is the index of the month bin of the year \((k' = 1\) to 12). \(T^{\text{Time}}(l,m,k')\) is the mean RO temperature co-located for
different sensor types for each level \((l)\), zone \((m)\), and averaged over all available years for a particular month \((k')\). Note that because the period of available measurements for each of the sensor types is different, the months used to compute \(T^{\text{time}}(l,m,k')\) may vary for different sensor types.

Fig. 12 indicates the de-seasonalized trends in daytime temperature differences are within ±0.26 (K/5 yrs, see Table 3). The greatest magnitudes of the trends are 0.232 K/5 yrs and 0.264 K/5 yrs over Canada and United Kingdom respectively. These larger de-seasonalized trends may be a result of incomplete daytime radiation corrections applied in these regions in 2007-2010 and 2011-2014 (Fig. 9). The largest nighttime de-seasonalized trend is in the United States (-0.211 K/5 yrs).

The de-seasonalized trends in RO temperatures are generally larger than those for the radiosonde-RO differences. A maximum de-seasonalized trend of 1.143 K/5 yrs is found for nighttime temperatures over the United Kingdom. A minimum de-seasonalized trend of -0.69 K/5 yrs is found for daytime temperatures over Canada. Trends with magnitude greater than 0.5 K/5 yrs are found over the United States, Germany, Canada and the United Kingdom. The fact that these de-seasonalized trends in RO are significantly greater than the de-seasonalized trends in the differences suggests that they represent a physical signal in these regions. However, the time series is too short to represent a long-term climate signal; instead these likely represent real but short-term trends associated with natural variability. A long-term (de-seasonalized) trend in temperature at this level associated with global warming (stratospheric cooling) might be approximately -0.1 to -0.2 K/decade or -0.05 to -0.1 K/5 years (Randel et al., 2016).

Trends reported in this paper for the Vaisala RS92 radiosonde at 50 hPa (Table 3) range
from -0.211 K/5 years (U.S., night) to 0.264 K/5 years (United Kingdom, day), which are comparable to those reported by Randel et al., (2016).

We compare the global trend of radiosonde – RO temperature differences for the Vaisala and other radiosondes at 50 hPa in Table 4. The Vaisala RS92 biases are 0.22 K (day) and 0.12 K (night). The global de-seasonalized temperature differences for Vaisala RS92 for daytime and nighttime are equal to 0.074 K/5yrs and -0.094 K/5yrs, respectively. The 95% confidence intervals for slopes are shown in the parentheses in Table 4. This indicates that although there might be a small residual radiation error for RS92, the trend in RS92 and RO temperature differences from June 2006 to April 2014 is within +/-0.09 K/5yrs globally. These values are just above the 1-sigma calibration uncertainty estimated by Dirksen et al. (2014). This means that probably the stability of the calibration alone could explain most of this very small trend. It is also consistent with the change in radiation correction.

Figure 13 depicts the de-seasonalized temperature differences for Sippican MARK IIA, VIZ-B2, AVK-MRZ, and Shanghai in North hemisphere mid-latitude (60°N-20°N) at 50 hPa and the results are summarized in Table 4. The 95% confidence intervals for slopes are shown in the parentheses in Table 4. The de-seasonalized trend of the daytime differences varies from -0.137 K/5 years (Russia) to 0.468 K/5 years (VIZ-B2). The magnitudes of the daytime trends are less than 0.2 K/5 yrs for all sensor types except for VIZ-B2 and Sippican, both of which exceed 0.4 K/5 yrs. These are much larger than those of the Vaisala RS92 (0.074 K/5 yrs).

The corresponding nighttime de-seasonalized trends in the biases vary from -0.348 K/5 yrs (VIZ-B2) to 0.244 K/5 yrs (Sippican). Again, these are much larger than
those of Vaisala RS92 (-0.094 K/5 yrs). Thus the VIZ-B2 sensor stands out as having larger biases and trends than do the other sensors.

5. Conclusions and Future Work

In this study, we used consistently reprocessed GPS RO temperature data to characterize radiosonde temperature biases and the inter-seasonal and inter-annual variability of these biases in the upper troposphere and lower stratosphere for different radiosonde types. We reach the following conclusions.

1. Solar zenith angle dependent biases: The solar radiative effect on different sensors is the dominant error source of RAOB temperature biases during daytime. With the consistent precision of RO temperature data between COSMIC and Metop-A, we are able to identify the mean temperature biases from the 200 hPa to 20 hPa layer among older sensors (i.e., Vaisala RS80 sensors), and new sensors (i.e., RS92 sensors), and the daytime and nighttime biases for the same sensor types which are usually distributed in the same countries (i.e., Shanghai sensor in China, AVK in Russian, VIZ-B2 in United Stated). Because the quality of RO temperature is not affected by sunlight, those daytime/nighttime biases mainly originate from uncorrected radiation biases for each individual sensor types. Most of the sensor types contain positive temperature biases ranging from 0.1 to 0.5 K during the daytime. Among all the sensors, the Vaisila RS92 has the smallest temperature biases. The daytime mean RS92 bias is about 0.1 K to 0.3 K globally, which is statistically insignificant. The bias of the AVK (Russian) sonde is as large as 0.8 K, which is statistically significant. Most of the sensor types show a cold bias at night, where the AVK and VIZ-B2 biases are as large as -0.22 K and -0.54 K,
respectively.

2. Residual solar zenith angle dependent biases: After applying the solar radiation correction, most of the RS92 daytime biases are removed. However, a small residual radiation bias for RS92 remains, which varies with different geographical region or operating organization. Similar to the results of He2009 and Sun et al., (2010, 2013), we find that there exists a small SZA dependent biases among different sensor types. The global mean residual temperature biases for RS92 from SZA 0 to 45 degrees at 20 hPa, 50 hPa, and 150 hPa are approximately 0.3 K, 0.15 K, and 0.05 K, respectively. These biases are less than the uncertainty described in Dirksen et al., (2014).

3. Changes of the radiation correction and RAOB temperature uncertainty due to when and how the radiative correction was implemented: the correction for RSN2010 is about 0.1 K higher than those from RSN2005. To identify the possible RS92 temperature biases due to changes of radiation correction table, we compared the mean RS92 temperature differences from January 2007 to December 2010 to those from January 2011 to April 2014. Results show that there are no obvious changes between these two periods over the United States and Germany at 20 hPa. However, the daytime temperature differences over Australia, Canada, United Kingdom, and Brazil show changes close to 0.1 K to 0.15 K varying at different heights. Changing sensors independently of the appropriate radiation correction introduces extra uncertainties of the RS92 trends. The relatively small temperature difference between these two periods over the United States is most likely a statistical artifact due to the small number of coincidences in this period. The relatively small temperature difference between these two periods over the Germany may be because the DWD implemented the updated
radiation correction for the Vaisala RS92 in the spring of 2015 rather than 2011, to avoid inconsistencies with corrections already implemented in their data assimilation system. This also indicates the importance of establishing traceability through careful documentation and metadata tracking, which is especially crucial for radiosonde data used in climate studies.

4. We used time series of RAOB-RO differences to indicate the long-term stability for each sonde type. The uncertainties are from the combined effects of i) uncorrected solar zenith angle dependent biases, ii) change of radiation correction, iii) when and how the radiation correction was implemented, and iv) small samples used in the time series and trend analysis. Results show that the time series of the RS92 differences at all regions are, in general, stable in time with a small day-night difference in each region. Other sensors have much larger variation than those of Vaisala RS92.

5. We found that the variation of mean radiosonde-RO temperature differences in different regions is closely related to the corresponding variation of SZA, especially for VIZ-B2 and AVK-MRZ during the daytime. The Sippican MARK IIA over the United States and the Shanghai sondes do not show significant seasonal variation. The de-seasonalized trend in RS92 and RO differences from June 2006 to April 2014 is within +/-0.09 K/5yrs globally. The trend of de-seasonalized daytime temperature differences for Sippican, VIZ-B2, Russia AVK, and Shanghai are much larger than those of RS92. Overall, the Vaisala RS92 radiosondes show a quality and stability that make them suitable for use in long-term climate trend studies.

Note that the analyses we performed here do not include other error sources (i.e., cloud radiative effect, ventilation, and sensor orientation, meta data errors) mentioned by
Dirksen et al., (2014). Since it is not possible to investigate these errors, we assume these errors introduce more or less random errors when a relative large sample is used. In addition, although RO derived dry temperature data are not directly traceable to the international standard of units (SI traceability), it has been shown that the high precision nature of the basic RO observations of time delay and bending angle are preserved through the inversion procedures (Ho et al., 2009a, 2011). This makes RO-derived dry temperature uniquely useful for assessing the radiosonde temperature biases and their long-term stability including the seasonal and inter-annual variability in the lower stratosphere. Results from this study also demonstrate the potential usage of RO data to identify RAOB temperature biases for different sensor types.

Acknowledgments

This work is supported by the NSF CAS AGS-1033112. The authors acknowledge the contributions to this work from members of the COSMIC team at UCAR.

References


Schreiner, W., S. Sokolovskiy, D. Hunt, C. Rocken, and Y.-H. Kuo (2011), Analysis of


**Figure Captions**

Figure 1. Global distribution of radiosonde stations colored by radiosonde types. Radiosonde types updated from June 2006 to April 2014 are used. The percentage of each type of radiosonde used among all stations is listed. For those stations that radiosonde types are changed during this period, the latest updated radiosonde type is used in this plot. Vaisala RS92 ship observations contain less than 3% of the total RS92 profiles.
Figure 2. Mean RAOB-RO temperature biases at 50 hPa for the RAOB-RO ensembles from June 2006 to April 2014 for a) daytime, and b) nighttime. Only those stations containing more than 50 RO-RAOB pairs are plotted.

Figure 3. Comparisons of temperature between RS92 and RO for daytime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red line is the mean difference; the black line is the standard deviation of the mean difference; the dotted line is the sample number. The top X axis shows the sample number. The same symbols are also used for the following plots. We also plot the standard error of the mean (black dot) superimposed on the mean. The value of the standard error of the mean is less than 0.03 K depending on the sample numbers.

Figure 4. Comparisons of temperature between RS92 and RO for nighttime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.

Figure 5. The mean temperature biases (RS92 minus RO) at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for each 5 SZA bins; the red vertical line is the standard deviation of error defined as standard deviation divided by sample numbers; the vertical red lines superimposed on the mean are the standard error of the mean; the black line to indicate zero mean; the blue dash line is the sample number. The right Y axis shows the sample number. Only bins for more than 50 RAOB-RO pairs are plotted.

Figure 6. Comparisons of temperature between radiosonde and RO during the daytime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.

Figure 7. Comparisons of temperature between radiosonde and RO during the nighttime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.

Figure 8. The mean temperature biases at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO. Only bins for more than 50 RAOB-RO pairs are plotted.

Figure 9. The temperature differences between RS92 – RO from January 2007 to December 2010 (ΔT (RS92200701-201012)) and those from January 2011 to December 2015 (ΔT (RS92201101-201512)) over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.

Figure 10. The time series of monthly mean temperature differences to RO at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) England, and f) Brazil.
The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is for that during the nighttime; the vertical lines superimposed on the mean values are the standard error of the mean for daytime and nighttime, respectively; the back line indicates zero temperature bias; the pink/green dash line is the sample number for the daytime and nighttime, respectively. The right Y axis shows the sample number. The same symbols are also used for the following plots.

Figure 11. The time series of temperature difference in 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N).

Figure 12. The time series of de-seasonalized temperature differences at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is for that during the nighttime; the vertical lines superimposed on the mean values are the standard error of the mean for daytime and nighttime, respectively. The number of the monthly RAOB-RO pairs for daytime is indicated by the pink dashed line and that for nighttime by the green dashed line. The vertical lines superimposed on the monthly mean are the standard errors of the mean. Day and night trends are shown by solid red and blue lines respectively. The zero difference is indicated by the dashed black line. The 95% confidence intervals for slopes are shown in the parentheses. The right Y axis shows the sample number. The same symbols are also used in Fig. 13.

Figure 13. The time series of de-seasonalized temperature differences at 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N). The 95% confidence intervals for slopes are shown in the parentheses.
Figure 1. Global distribution of radiosonde stations colored by radiosonde types. Radiosonde types updated from June 2006 to April 2014 are used. The percentage of each type of radiosonde used among all stations is listed. For those stations that radiosonde types are changed during this period, the latest updated radiosonde type is used in this plot. Vaisala RS92 ship observations contain less than 3% of the total RS92 profiles.
Figure 2. Mean RAOB-RO temperature biases at 50 hPa for the RAOB-RO ensembles from June 2006 to April 2014 for a) daytime, and b) nighttime. Only those stations containing more than 50 RO-ROAB pairs are plotted.
Figure 3. Comparisons of temperature between RS92 and RO for daytime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red line is the mean difference; the black line is the standard deviation of the mean difference; the dotted line is the sample number. The top X axis shows the sample number. The same symbols are also used for the following plots. We also plot the standard error of the mean (black dot) superimposed on the mean. The value of the standard error of the mean is less than 0.03 K depending on the sample numbers.
Figure 4. Comparisons of temperature between RS92 and RO for nighttime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.
Figure 5. The mean temperature biases (RS92 minus RO) at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for each 5 SZA bins; the red vertical line is the standard deviation of error defined as standard deviation divided by sample numbers; the vertical red lines superimposed on the mean are the standard error of the mean; the black line to indicate zero mean; the blue dash line is the sample number. The right Y axis shows the sample number. Only bins for more than 50 RAOB-RO pairs are plotted.
Figure 6. Comparisons of temperature between radiosonde and RO during the daytime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.
Figure 7. Comparisons of temperature between radiosonde and RO during the nighttime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.
Figure 8. The mean temperature biases at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO. Only bins for more than 50 RAOB-RO pairs are plotted.
Figure 9. The temperature differences between RS92 – RO from January 2007 to December 2010 ($\Delta T_{\text{RS92 200701-201012}}$) and those from January 2011 to December 2015 ($\Delta T_{\text{RS92 201101-201512}}$) over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.
Figure 10. The time series of monthly mean temperature differences to RO at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is for that during the nighttime; the vertical lines superimposed on the mean values are the standard error of the mean for daytime and nighttime, respectively; the back line indicates zero temperature bias; the pink/green dash line is the sample number for the daytime and nighttime, respectively. The right Y axis shows the sample number. The same symbols are also used for the following plots.
Figure 11. The time series of temperature difference in 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N).
Figure 12. The time series of de-seasonalized temperature differences at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is for that during the nighttime; the vertical lines superimposed on the mean values are the standard error of the mean for daytime and nighttime, respectively. The number of the monthly RAOB-RO pairs for daytime is indicated by the pink dashed line and that for nighttime by the green dashed line. The vertical lines superimposed on the monthly mean are the standard errors of the mean. Day and night trends are shown by solid red and blue lines respectively. The zero difference is indicated by the dashed black line. The 95% confidence intervals for slopes are shown in the parentheses. The right Y axis shows the sample number. The same symbols are also used in Fig. 13.
Figure 13. The time series of de-seasonalized temperature differences at 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N). The 95% confidence intervals for slopes are listed in the parentheses.
Table 1. Summary of the availability for different instrument types and their solar absorptivity (α) and sensor infrared emissivity (ε) for the corresponding thermocap and thermistor and the sample number of RAOB-RO pairs used in this study from June 2006 to April 2014.

<table>
<thead>
<tr>
<th>ID</th>
<th>Sensor type</th>
<th>Availability</th>
<th>Solar absorptivity</th>
<th>Infrared emissivity</th>
<th>Number of RO-RAOB pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS80 37</td>
<td>Bead thermocap</td>
<td>1981–2014</td>
<td>0.15</td>
<td>[Luers and Eskridge, 1998]</td>
<td>0.02</td>
</tr>
<tr>
<td>Vaisala RS80-57H 52</td>
<td>Bead thermocap early 1990s [Redder et al., 2004] ~ Jul 2012</td>
<td>0.15</td>
<td>0.02</td>
<td>13192</td>
<td></td>
</tr>
<tr>
<td>Vaisala RS80/Loran 61</td>
<td>Bead thermocap ~ 2014</td>
<td>0.15</td>
<td>0.02</td>
<td>11591</td>
<td></td>
</tr>
<tr>
<td>Vaisala RS80/Digicora III 67</td>
<td>Bead thermocap ~ 2012</td>
<td>0.15</td>
<td>0.02</td>
<td>2864</td>
<td></td>
</tr>
<tr>
<td>Vaisala RS90/Digicorn I, II 71</td>
<td>Thin wire F-thermocap [Sun et al., 2010] 1995–2014</td>
<td>0.15</td>
<td>[Luers, 1997]</td>
<td>0.02</td>
<td>18082</td>
</tr>
<tr>
<td>Vaisala RS92/Digicora I/II 79</td>
<td>Thin wire F-thermocap [Sun et al., 2010] 2003–2014</td>
<td>0.15</td>
<td>0.02</td>
<td>40478</td>
<td></td>
</tr>
<tr>
<td>Vaisala RS92/Digicora III 80</td>
<td>Thin wire F-thermocap 2004–2014</td>
<td>0.15</td>
<td>0.02</td>
<td>184542</td>
<td></td>
</tr>
<tr>
<td>Vaisala RS92/Autosonde 81</td>
<td>Thin wire F-thermocap 2011–2014</td>
<td>0.15</td>
<td>0.02</td>
<td>42577</td>
<td></td>
</tr>
<tr>
<td>AVK-MRZ 27</td>
<td>Rod thermistor [Sun et al., 2010] ~ 2014</td>
<td>0.2</td>
<td>[He et al., 2009]</td>
<td>0.04</td>
<td>48954</td>
</tr>
<tr>
<td>AVK-BAR Russian 58</td>
<td>Rod thermistor 2007–2014</td>
<td>0.2</td>
<td>0.04</td>
<td>26020</td>
<td></td>
</tr>
<tr>
<td>AVK-MRZ (Russian) 75</td>
<td>Rod thermistor ~ 2013</td>
<td>0.2</td>
<td>0.04</td>
<td>9472</td>
<td></td>
</tr>
<tr>
<td>MARL-A or Vektor-M-MRZ (Russian) 88</td>
<td>Rod thermistor ~ 2014</td>
<td>0.2</td>
<td>0.04</td>
<td>23326</td>
<td></td>
</tr>
<tr>
<td>MARL-A or Vektor-M-BAR (Russian) 89</td>
<td>Rod thermistor ~ 2014</td>
<td>0.2</td>
<td>0.04</td>
<td>25715</td>
<td></td>
</tr>
<tr>
<td>VIZ-B2 51</td>
<td>Rod thermistor [Sun et al., 2010] 1997 [Elliott et al., 2002]–2014</td>
<td>0.15</td>
<td>[Luers and Eskridge, 1998]</td>
<td>0.86</td>
<td>16310</td>
</tr>
<tr>
<td>Sippican MARK II A Chip 87</td>
<td>Chip thermistor [Sun et al., 2010] 1998 [Elliott et al., 2002]–2014</td>
<td>0.07</td>
<td>[Luers and Eskridge, 1998]</td>
<td>0.85</td>
<td>59775</td>
</tr>
<tr>
<td>Shanghai 32</td>
<td>Rod thermistor 1998–2012 &lt;0.07 [Wei, 2011] &gt;0.90</td>
<td>71605</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meisei Japan 47</td>
<td>Thermistor [KOBAYASHI et al., 2012] 1994–2013</td>
<td>0.18</td>
<td>[Luers and Eskridge, 1998]</td>
<td>0.84</td>
<td>7888</td>
</tr>
</tbody>
</table>
Table 2. Mean and standard deviation of temperature differences (K) from the layer from 200 hPa to 20 hPa between RO and eight types of radiosonde\textsuperscript{a,b}. \textsuperscript{a}The values of standard deviations of temperature differences are shown in the parentheses. \textsuperscript{b}The sample number are for the ROAB-RO pairs available in the same time period.

<table>
<thead>
<tr>
<th>ID</th>
<th>All Day and night mean(std)/ sample numbers</th>
<th>Day mean(std)/ sample numbers</th>
<th>Night mean(std)/ sample numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaisala RS80</td>
<td>37, 52, 61, 67</td>
<td>0.10 (1.54)/29271</td>
<td>0.10 (1.53)/15947</td>
</tr>
<tr>
<td>Vaisala RS90</td>
<td>71</td>
<td>0.13 (1.54)/18082</td>
<td>0.16 (1.51)/8758</td>
</tr>
<tr>
<td>Vaisala RS92</td>
<td>79, 80, 81</td>
<td>0.16 (1.52)/267597</td>
<td>0.20 (1.50)/161019</td>
</tr>
<tr>
<td>AVK</td>
<td>27, 75, 88, 89, 58</td>
<td>0.33 (1.58)/133487</td>
<td>0.66 (1.51)/67679</td>
</tr>
<tr>
<td>VI-Z-B2</td>
<td>51</td>
<td>0.22 (1.67)/16310</td>
<td>0.71 (1.54)/9246</td>
</tr>
<tr>
<td>Sippican MARKIIA Chip</td>
<td>87</td>
<td>-0.08 (1.59)/59775</td>
<td>-0.06 (1.56)/31230</td>
</tr>
<tr>
<td>Shanghai</td>
<td>32</td>
<td>0.05 (1.68)/71605</td>
<td>0.18 (1.67)/33360</td>
</tr>
<tr>
<td>Meisei Japan</td>
<td>47</td>
<td>0.11 (1.69)/7888</td>
<td>0.03 (1.71)/3849</td>
</tr>
</tbody>
</table>
Table 3. Mean, standard deviation (std) of monthly temperature differences (K), de-seasonalized trend of temperature differences (K/5yrs), and root mean square (RMS) of de-seasonalized RS92-RO temperature difference time series at 50 hPa over United States, Australia, Germany, Canada, United Kingdom, and Brazil.

<table>
<thead>
<tr>
<th></th>
<th>United States</th>
<th>Australia</th>
<th>Germany</th>
<th>Canada</th>
<th>United Kingdom</th>
<th>Brazil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Bias</td>
<td>0.08</td>
<td>0.19</td>
<td>0.22</td>
<td>0.23</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>std of Mean Bias</td>
<td>0.4</td>
<td>0.46</td>
<td>0.18</td>
<td>0.3</td>
<td>0.2</td>
<td>0.24</td>
</tr>
<tr>
<td>De-seasonalized Trend of Differences (K/5 yrs)</td>
<td>0.001</td>
<td>-0.211</td>
<td>0.167</td>
<td>-0.083</td>
<td>-0.016</td>
<td>-0.135</td>
</tr>
<tr>
<td>De-seasonalized Trend of RO Temperature (K/5yrs)</td>
<td>0.941</td>
<td>0.506</td>
<td>-0.26</td>
<td>0.082</td>
<td>0.29</td>
<td>0.708</td>
</tr>
<tr>
<td>RMS of de-seasonalized difference</td>
<td>0.365</td>
<td>0.439</td>
<td>0.161</td>
<td>0.275</td>
<td>0.173</td>
<td>0.22</td>
</tr>
</tbody>
</table>
Table 4. Mean, standard deviation (std), de-seasonalized trend of temperature differences (K/5yrs), and root mean square (RMS) of de-seasonalized time series of temperature difference at 50 hPa for global Vaisala (RS80, RS90, and RS92), and other sensor types in the North hemisphere mid-latitude (60°N-20°N). The 95% confidence intervals for trend of differences are listed in the parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Mean Bias</th>
<th>STD Of MB</th>
<th>De-seasonalized Trend of Difference (k/5yrs)</th>
<th>RMS of Difference</th>
<th>Mean Bias</th>
<th>STD Of MB</th>
<th>De-seasonalized Trend of Difference (k/5yrs)</th>
<th>RMS of Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS80</td>
<td>37,52,61,67</td>
<td>0.18</td>
<td>0.187 (0.073,0.301)</td>
<td>0.268</td>
<td>0.13</td>
<td>0.33</td>
<td>0.114(-0.019,0.248)</td>
<td>0.301</td>
</tr>
<tr>
<td>RS90</td>
<td>71</td>
<td>0.16</td>
<td>-0.006 (-0.123,0.111)</td>
<td>0.26</td>
<td>0.17</td>
<td>0.38</td>
<td>0.043(-0.115,0.201)</td>
<td>0.352</td>
</tr>
<tr>
<td>RS92</td>
<td>79,80,81</td>
<td>0.22</td>
<td>0.074 (0.051,0.097)</td>
<td>0.062</td>
<td>0.12</td>
<td>0.12</td>
<td>-0.094(-0.131,-0.057)</td>
<td>0.093</td>
</tr>
<tr>
<td>Russia</td>
<td>27,75,88,89,88</td>
<td>0.8</td>
<td>0.137 (-0.205,-0.068)</td>
<td>0.164</td>
<td>-0.03</td>
<td>0.21</td>
<td>-0.143(-0.218,-0.067)</td>
<td>0.18</td>
</tr>
<tr>
<td>VIZ-B2</td>
<td>51</td>
<td>0.87</td>
<td>0.468 (0.358,0.579)</td>
<td>0.322</td>
<td>-0.56</td>
<td>0.43</td>
<td>-0.348(-0.518,-0.177)</td>
<td>0.386</td>
</tr>
<tr>
<td>Sippican MARKIIA Chip</td>
<td>87</td>
<td>-0.12</td>
<td>0.405 (0.304,0.507)</td>
<td>0.292</td>
<td>-0.12</td>
<td>0.21</td>
<td>0.244(0.168,0.320)</td>
<td>0.197</td>
</tr>
<tr>
<td>Shanghai</td>
<td>32</td>
<td>0.1</td>
<td>0.179 (0.081,0.276)</td>
<td>0.161</td>
<td>-0.2</td>
<td>0.17</td>
<td>-0.015(-0.120,0.090)</td>
<td>0.159</td>
</tr>
<tr>
<td>Meisei Japan</td>
<td>47</td>
<td>0.07</td>
<td>0.006 (-0.353,0.365)</td>
<td>0.619</td>
<td>0.05</td>
<td>0.51</td>
<td>-0.086(-0.369,0.197)</td>
<td>0.494</td>
</tr>
</tbody>
</table>