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 Ho[1], Liang Peng[1], Holger Vömel[2]

[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;

77 Andrae et al., 2004; Free et al., 2004, 2005; Sherwood et al., 2008; Haimberger et al.,
78 2008, 2011; Thorne et al., 2011; Seidel et al., 2009). Possible errors due to changes of
79 instruments were identified by comparing with temperature measurements from adjacent
80 weather stations. However, this approach is limited by the low number of co-located
81 observations and large atmospheric variability. In addition, due to lack of absolute
82 references, the remaining radiation temperature biases from adjacent stations may not be
83 completely removed. As a result, only relative temperature differences of a possibly large
84 uncertainty among stations are identified.

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

99 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 re-processed 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 -0.02K. 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

period of time. 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 senor 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

192 Vaisala RS80, RS90, RS92, AVK-MRZ (and other Russian sondes), VIZ-B2, Shippican
193 MARK II A, Shanghai (from China), and Meisie (Japan). Table 1 summarizes the
194 availability for different instrument types. In total, seventeen different types of
195 radiosonde systems were used. The solar absorptivity (α) and sensor infrared emissivity
196 (ε) for the corresponding thermocap and thermistor for different instrument types are also
197 summarized in Table 1. Most of the radiosonde data are collected twice per day.

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

204

205 **2.2 GPS RO data**

206 The re-processed COSMIC (Version 2013.3520) and Metop-A/GRAS (Version
207 2016.0120) dry temperature profiles downloaded from UCAR CDAAC
208 (http://cosmic.cosmic.ucar.edu/cdaac/index.html) are used in this study. With six GPS
209 receivers on board LEO satellites, COSMIC produced about 1000 to 2500 RO profiles
210 per day since launch. With one receiver, Metop-A/GRAS produced about 600 RO
211 profiles per day. The detail inversion procedures of COSMIC Version 2013.3520 and
212 Metop-A Version 2016.0120 are summarized in http://cdaac-
213 www.cosmic.ucar.edu/cdaac/doc/documents/Sokolovskiy_newroam.pdf. The general
214 description of CDAAC inversion procedures is detailed in Kuo et al., (2004), and Ho et

al., (2009a, 2012). In a neutral atmosphere, the refractivity (N) is related to pressure (P in hPa), temperature (T in K) and the water vapor pressure (e in hPa) according to Smith and Weintraub (1953):

$$N = 77.6\frac{P}{T} + 3.73*10^5\frac{e}{T^2} \quad . \tag{1}$$

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

$$\Delta T(i,j) = (1/n) \times \sum_{s=1}^{s=n} \{T_{RAOB}(i,j,s) - T_{RO}(i,j,s)\}, \quad (2)$$

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 ΔT^{Time} for the matched pairs at different geo-graphical regions from

$$\Delta T^{Time}(l,m,k) = T_{RAOB}(l,m,k) - T_{RO}(l,m,k), \quad (3)$$

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 ΔT^{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 Senor 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

283 nighttime are shown in Figure 2. Only those stations containing more than 50 RO-ROAB
284 pairs are plotted. Figure 2a shows biases for different sensor types, which vary with
285 geographical region. Most of the sensor types contain positive temperature biases ranging
286 from 0.1 to 0.5 K during the daytime. This bias during daytime may be a result of the
287 residual error of the systematic radiation bias correction. Although we only include
288 stations containing more than 50 RO-RAOB pairs, some level of heterogeneity (i.e., Fig.
289 2a over Brazil) may be due to low sample numbers. For example, stations with
290 temperature biases larger than 0.5 K in eastern Brazil contain only about 60 RO-RAOB
291 pairs. The cause of the heterogeneity in temperature bias between North and South China
292 is not certain at this point.

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

300 The global mean ΔT for the Vaisala RS92 of 0.16 K during the comparison
301 period is slightly larger than the temperature comparison between COSMIC and Vaisala
302 RS92 in 2007 (Ho et al., 2010b) (~0.01 K) and in He2009 (~0.04 K from ~200 hPa to 50
303 hPa). This could be in part because more RS92-RO pairs from lower solar zenith angle
304 regions (for example, from the southern Hemisphere and near Tropics, see Section 3.2)
305 are included after 2007 (see section 4).

306

307 **3.2 Solar Zenith Angle Dependent Temperature Biases for Vaisala Sondes**

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

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

320 Figure 4 depicts the mean RS92-RO temperature differences from 200 hPa to 20
321 hPa for nighttime. The nighttime RS92 data over different regions show similar standard
322 deviations of about 1.5 K compared to those at daytime. In most of the regions, the mean
323 nighttime temperature biases are similar to those in the daytime results, with small (0.1-
324 0.2 K) warm biases above 100 hPa. These residual nighttime warm biases are not seen in
325 the ROAB-RO ensemble pairs for Sippican MARK, VIZ-B2, AVK, and Shanghai Sondes
326 (see Section 3.3). This 0.1 K – 0.2 K warm bias for RS92 at night could be due to
327 calibration of the RS92 temperature sensor (see Dirksen et al., 2014).

328 Because the quality of RO temperature is not affected by sunlight, the small but

329 obvious geographic-dependent biases are most likely due to the residual radiation
330 correction for RS92 and when and how different countries apply the radiation correction
331 (see Section 4.1).

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

341

342 **3.3 Temperature Biases for Sippican MARK, VIZ-B2, AVK-MRZ, and Shanghai**
343 **Sondes**

344 Unlike Vaisala sondes, which are distributed in almost all latitudinal zones, other
345 sonde types are distributed mainly in the northern mid-latitudes. Fig. 6 depicts the mean
346 temperature differences from 200 hPa to 20 hPa in the daytime for Sippican, VIZ-B2,
347 AVK, and Shanghai. The biases for VIZ-B2 and AVK-MRZ are positive everywhere
348 above 200 hPa, with means of about 0.7K. The biases are smaller for Sippican and
349 Shanghai. These mean biases are similar to those from He2009. The small differences
350 between these and He2009 results are likely due to the sampling differences between
351 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

376 sensor in 2007, which impacted the radiation error. To account for this sensor update, the
377 radiation correction tables were updated in 2011 (RSN2010, software version 3.64),
378 which is used to replace the original radiation correction table. Between 200 and 20 hPa,
379 the correction in RSN2010 is about 0.1 K larger than in RSN2005 (see
380 http://www.vaisala.com/en/products/soundingsystemsandradiosondes/soundingdatacontin
381 uity/RS92DataContinuity/Pages/revisedsolarradiationcorrectiontableRSN2010.aspx). It is
382 likely that each country updated the correction table for their entire network. However,
383 when exactly each country implemented these updated tables is unknown.

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

393 The Deutscher Wetterdienst (DWD), Germany's Meteorological Service,
394 implemented the updated radiation correction for the Vaisala RS92 in the spring of 2015
395 rather than 2011, to avoid inconsistencies with corrections already implemented in their
396 data assimilation system. This may in part explain the greater consistency of ΔT over
397 Germany for these two time periods than over other countries. This also indicates the
398 importance of establishing traceability through careful documentation and metadata

399 tracking, which is especially important for using radiosonde data in climate studies. The
400 relatively small temperature difference between these two periods over the United States
401 is most likely a statistical artifact due to the very small number of coincidences in this
402 period.
403
404 **4.2 Time Series and trends of de-seasonalized radiosonde-RO differences**
405 In this section we look at time series and trends in the de-seasonalized radiosonde-
406 RO temperature differences from 2007 to 2014 in order to determine the long-term
407 stability of these differences. Ideally, if both radiosondes and RO were free of biases, the
408 time series would be stable and show small differences near zero with small standard
409 deviations and no trends. We choose 50 hPa for showing these time series because the
410 biases tend to be larger at this level than at lower levels. We also computed time series
411 for 150 hPa, but except for lower biases, the results were similar to those at 50 hPa (not
412 shown).
413 Figure 10 shows daytime and nighttime time series of monthly mean temperature
414 biases at 50 hPa for Vaisala RS92 for the United States, Australia, Germany, Canada,
415 United Kingdom, and Brazil. Table 3 summarized the mean and std of the monthly mean
416 temperature differences for RS92 and RO at 50 hPa. Fig. 10 indicates that there is little
417 variation over time in the monthly mean temperature differences at 50 hPa in all six
418 regions, with little difference between day and night values. The magnitudes of the mean
419 biases range from -0.01 K for Canada to over 0.2 K in Australia, Germany and Brazil.
420 The standard deviations range from a low of 0.18 K (Australia, day) to a high of 0.46 K
421 (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

$$\Delta T^{Deseason}(l,m,k) = T_{RAOB}(l,m,k) - \overline{T^{Time}}(l,m,k'), \quad (4)$$

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). $\overline{T^{Time}}(l,m,k')$ is the mean RO temperature co-located for

445 different sensor types for each level (l), zone (m), and averaged over all available years
446 for a particular month (k'). Note that because the period of available measurements for
447 each of the sensor types is different, the months used to compute $\overline{T^{Time}}(l,m,k')$ may vary
448 for different sensor types.

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

455 The de-seasonalized trends in RO temperatures are generally larger than those for
456 the radiosonde-RO differences. A maximum de-seasonalized trend of 1.143 K/5 yrs is
457 found for nighttime temperatures over the United Kingdom. A minimum de-seasonalized
458 trend of -0.69 K/5 yrs is found for daytime temperatures over Canada. Trends with
459 magnitude greater than 0.5 K/5 yrs are found over the United States, Germany, Canada
460 and the United Kingdom. The fact that these de-seasonalized trends in RO are
461 significantly greater than the de-seasonalized trends in the differences suggests that they
462 represent a physical signal in these regions. However, the time series is too short to
463 represent a long-term climate signal; instead these likely represent real but short-term
464 trends associated with natural variability. A long-term (de-seasonalized) trend in
465 temperature at this level associated with global warming (stratospheric cooling) might be
466 approximately -0.1 to -0.2 K/decade or -0.05 to -0.1 K/5 years (Randel et al., 2016).
467 Trends reported in this paper for the Vaisala RS92 radiosonde at 50 hPa (Table 3) range

468 from -0.211 K/5 years (U.S., night) to 0.264 K/5 years (United Kingdom, day), which are
469 comparable to those reported by Randel et al., (2016).

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

481 Figure 13 depicts the de-seasonalized temperature differences for Sippican
482 MARK IIA, VIZ-B2, AVK-MRZ, and Shanghai in North hemisphere mid-latitude
483 (60°N-20°N) at 50 hPa and the results are summarized in Table 4. The 95% confidence
484 intervals for slopes are shown in the parentheses in Table 4. The de-seasonalized trend of
485 the daytime differences varies from -0.137 K/5 years (Russia) to 0.468 K/5 years (VIZ-
486 B2). The magnitudes of the daytime trends are less than 0.2 K/5 yrs for all sensor types
487 except for VIZ-B2 and Sippican, both of which exceed 0.4 K/5 yrs. These are much
488 larger than those of the Vaisala RS92 (0.074 K/5 yrs).

489 The corresponding nighttime de-seasonalized trends in the biases vary from -
490 0.348 K/5 yrs (VIZ-B2) to 0.244 K/5 yrs (Sippican). Again, these are much larger than

491 those of Vaisala RS92 (-0.094 K/5 yrs). Thus the VIZ-B2 sensor stands out as having
492 larger biases and trends than do the other sensors.
493
494 **5. Conclusions and Future Work**
495 In this study, we used consistently reprocessed GPS RO temperature data to
496 characterize radiosonde temperature biases and the inter-seasonal and inter-annual
497 variability of these biases in the upper troposphere and lower stratosphere for different
498 radiosonde types. We reach the following conclusions.
499 1. Solar zenith angle dependent biases: The solar radiative effect on different sensors
500 is the dominant error source of RAOB temperature biases during daytime. With the
501 consistent precision of RO temperature data between COSMIC and Metop-A, we are able
502 to identify the mean temperature biases from the 200 hPa to 20 hPa layer among older
503 sensors (i.e., Vaisala RS80 sensors), and new sensors (i.e., RS92 sensors), and the
504 daytime and nighttime biases for the same sensor types which are usually distributed in
505 the same countries (i.e., Shanghai sensor in China, AVK in Russian, VIZ-B2 in in United
506 Stated). Because the quality of RO temperature is not affected by sunlight, those
507 daytime/nighttime biases mainly originate from uncorrected radiation biases for each
508 individual sensor types. Most of the sensor types contain positive temperature biases
509 ranging from 0.1 to 0.5 K during the daytime. Among all the sensors, the Vaisila RS92
510 has the smallest temperature biases. The daytime mean RS92 bias is about 0.1 K to 0.3 K
511 globally, which is statistically insignificant. The bias of the AVK (Russian) sonde is as
512 large as 0.8 K, which is statistically significant. Most of the sensor types show a cold bias
513 at night, where the AVK and VIZ-B2 biases are as large as -0.22 K and -0.54 K,

514 respectively.

515 2. Residual solar zenith angle dependent biases: After applying the solar radiation
516 correction, most of the RS92 daytime biases are removed. However, a small residual
517 radiation bias for RS92 remains, which varies with different geographical region or
518 operating organization. Similar to the results of He2009 and Sun et al., (2010, 2013), we
519 find that there exists a small SZA dependent biases among different sensor types. The
520 global mean residual temperature biases for RS92 from SZA 0 to 45 degrees at 20 hPa,
521 50 hPa, and 150 hPa are approximately 0.3 K, 0.15 K, and 0.05 K, respectively. These
522 biases are less than the uncertainty described in Dirksen et al., (2014).

523 3. Changes of the radiation correction and RAOB temperature uncertainty due to
524 when and how the radiative correction was implemented: the correction for RSN2010 is
525 about 0.1 K higher than those from RSN2005. To identify the possible RS92 temperature
526 biases due to changes of radiation correction table, we compared the mean RS92
527 temperature differences from January 2007 to December 2010 to those from January
528 2011 to April 2014. Results show that there are no obvious changes between these two
529 periods over the United States and Germany at 20 hPa. However, the daytime
530 temperature differences over Australia, Canada, United Kingdom, and Brazil show
531 changes close to 0.1 K to 0.15 K varying at different heights. Changing sensors
532 independently of the appropriate radiation correction introduces extra uncertainties of the
533 RS92 trends. The relatively small temperature difference between these two periods over
534 the United States is most likely a statistical artifact due to the small number of
535 coincidences in this period. The relatively small temperature difference between these
536 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

Anthes, R. A., P. Bernhardt, Y. Chen, L. Cucurull, K. Dymond, D. Ector, S. Healy, S.-P. Ho, D. Hunt, Y.-H. Kuo, H. Liu, K. Manning, C. McCormick, T. Meehan, W. Randel, C.R. Rocken, W. Schreiner, S. Sokolovskiy, S. Syndergaard, D.Thompson, K. Trenberth, T.- K. Wee, Z. Zeng (2008), The COSMIC/FORMOSAT-3 Mission: Early Results, *Bul. Amer. Meteor. Sci*. 89, No.3, 313-333.

Andrae, U., N. Sokka, and K. Onogi (2004), The radiosonde temperature bias correction in ERA-40. ERA-40 Project Report Series, Vol. 15, 34 pp.

Dach, R., S. Lutz, P. Walser, P. Fridez (Eds) (2015), Bernese GNSS Software Version 5.2.User manual, Astronomical Institute, Universtiy of Bern, Bern Open Publishing. DOI: 10.7892/boris.72297; ISBN: 978-3-906813-05-9.

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

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

Free, M., J. K. Angle, I. Durre, J. Lanzante, T. C. Peterson, and D. J. Seidel (2004), Using first differences to reduce inhomogeneity in radiosonde temperature datasets, *J. Clim.,* **17**, 4171–4179.

Free, M., D. J. Seidel, J. K. Angell, J. Lanzante, I. Durre, and T. C. Peterson (2005), Radiosonde atmospheric temperature products for assessing climate (RATPAC): A new data set of large-area anomaly time series, *J. Geophys. Res.*, 110, D22101, doi:10.1029/2005JD006169.

Foelsche, U., B. Pirscher, M. Borsche, G. Kirchengast, and J. Wickert (2009), Assessing the climate monitoring utility of radio occultation data: From CHAMP to FORMOSAT-3/COSMIC, *Terr. Atmos. and Oceanic Sci.*, Vol. 20, doi: 10.3319/TAO.2008.01.14.01(F3C).

Haimberger, L., (2007), Homogenization of radiosonde temperature time series using innovation statistics. *J. Climate*, 20, 1377– 1403.

——, C. Tavolato, and S. Sperka (2008), Toward elimination of the warm bias in historic radiosonde temperature records—Some new results from a comprehensive intercomparison of upper- air data. *J. Climate*, 21, 4587–4606.

——, and U. Andrae (2011) Radiosonde temperature bias correction in ERA-Interim. ERA Rep. Series, Vol. 8, 17 pp.

He, W., S.-P. Ho, H. Chen, X. Zhou, D. Hunt, and Y. Kuo (2009), Assessment of radiosonde temperature measurements in the upper troposphere and lower stratosphere using COSMIC radio occultation data, *Geophys. Res. Lett.,* 36, L17807, doi:10.1029/2009GL038712.

Ho, S.-P., Y. H. Kuo, Zhen Zeng, and Thomas Peterson (2007), A Comparison of Lower Stratosphere Temperature from Microwave Measurements with CHAMP GPS RO Data, *Geophy. Research Letters,* 34, L15701, doi:10.1029/2007GL030202.

Ho, S.-P., and co-authors (2009a), Estimating the Uncertainty of using GPS Radio Occultation Data for Climate Monitoring: Inter-comparison of CHAMP Refractivity Climate Records 2002-2006 from Different Data Centers, *J. Geophys. Res.*, doi:10.1029/2009JD011969.

637 Ho, S.-P., M. Goldberg, Y.-H. Kuo, C.-Z Zou, W. Schreiner (2009b), Calibration of
638 Temperature in the Lower Stratosphere from Microwave Measurements using
639 COSMIC Radio Occultation Data: Preliminary Results, *Terr. Atmos. Oceanic Sci*.,
640 Vol. 20, doi: 10.3319/TAO.2007.12.06.01(F3C).
641
642 Ho, S.-P., Zhou X., Kuo Y.-H., Hunt D., Wang J.-H., (2010a), Global Evaluation of
643 Radiosonde Water Vapor Systematic Biases using GPS Radio Occultation from
644 COSMIC and ECMWF Analysis. *Remote Sensing*. 2010; 2(5):1320-1330.
645
646 Ho, S.-P., Ying-Hwa Kuo, William Schreiner, Xinjia Zhou (2010b), Using SI-traceable
647 Global Positioning System Radio Occultation Measurements for Climate
648 Monitoring [In "States of the Climate in 2009]. *Bul. Amer. Meteor. Sci*., 91 (7),
649 S36-S37.
650
651 Ho, S.-P., and co-authors (2012), Reproducibility of GPS Radio Occultation Data for
652 Climate Monitoring: Profile-to-Profile Inter-comparison of CHAMP Climate
653 Records 2002 to 2008 from Six Data Centers, *J. Geophy. Research.* VOL. 117,
654 D18111, doi:10.1029/2012JD017665.
655
656 Ho, S.-P., L. Peng, and D. Hunt (2016), Construction of consistent temperature climate
657 data records in the lower Troposphere using COSMIC, Metop-A/GRAS, and
658 Metop-B/GRAS from 2006 to 2015, *Atmos. Meas. Tech.,* submitted.
659
660 Kuo, Y. H., T. K. Wee, S. Sokolovskiy, C. Rocken, W. Schreiner, and D. Hunt (2004),
661 Inversion and error estimation of GPS radio occultation data, *J. Meteorol. Soc. Jpn.*,
662 **82**, 507– 531.
663
664 Lanzante, J. R., S. A. Klein, and D. J. Seidel (2003), Temporal homogeni- sation of
665 monthly radiosonde temperature data, Part I: Methodology, *J. Clim.,* 16, 224–240,
666 doi:10.1175/1520-0442(2003)016.
667
668 Luers, J. K., and R. E. Eskridge (1995), Temperature corrections for the VIZ and Vaisala
669 radiosondes, *J. Appl. Meteor.*, 34, 1241–1253.
670
671 Luers, J. K., (1997), Temperature error of the Vaisala RS90 radiosonde, *J. Atmos.*
672 *Oceanic Technol.*, 14, 1520–1532.
673
674 Luers, J. K., and R. E. Eskridge (1998), Use of radiosonde temperature data in climate
675 studies, *J. Clim.,* 11, 1002–1019.
676
677 Randel, W. J., A. K. Smith, F. Wu, C.-Z. Zhou, and H. Qian (2016), Stratospheric
678 temperature trends over 1979–2015 derived from combined SSU, MLS, and
679 SABER satellite observations, *J. Clim.*, *29*, 4843–4859, doi:10.1175/JCLI-D-15-
680 0629.1.
681
682 Schreiner, W., S. Sokolovskiy, D. Hunt, C. Rocken, and Y.-H. Kuo (2011), Analysis of

GPS radio occultation data from the FORMOSAT-3/COSMIC and Metop/GRAS missions at CDAAC, *Atmos. Meas. Tech.*, 4, 2255–2272.

Seidel, D. J., and Coauthors (2009), Reference upper-air observations for climate: Rationale, progress, and plans. *Bull. Amer. Meteor. Soc.,* 90, 361–369.

Seidel D.J., N. P. Gillett , J. R. Lanzante, K. P. Shine, P. W. Thorne (2011), Stratospheric tem- perature trends: Our evolving understanding. Wiley Interdisciplinary Reviews 2: 592–616.

Sherwood, S. C., C. L. Meyer, R. J. Allen, and H. A. Titchner (2008), Robust tropospheric warming as revealed by iteratively homogenized radiosonde data. *J. Climate,* 21, 5336–5352.

Smith, E. K., and S. Weintraub (1953), The constants in the equation for atmospheric refractive index at radio frequencies, *J. Res. Natl. Bur. Stand. U.S.,* 50, 39–41.

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

Sun, B., A. Reale, S. Schroeder, D. J. Seidel, and B. Ballish (2013), Toward improved corrections for radiation-induced biases in radiosonde temperature observations, *J. Geophys. Res.,* 118, 4231–4243, doi:10.1002/jgrd.50369.

Thorne, P. W., D. E. Parker, J. R. Christy, and C. A. Mears (2005), Uncertain- ties in climate trends: Lessons from upper-air temperature records, *Bull. Am. Meteorol. Soc.*, 86, 1437–1442, doi:10.1175/BAMS-86-10-1437.

Thorne, P. W., and Coauthors, 2011a: A quantification of uncertainties in historical tropical tropospheric temperature trends from radiosondes. *J. Geophys. Res.*, 116, D12116, doi:10.1029/2010JD015487.

Ware, R., et al. (1996), GPS sounding of the atmosphere from low Earth orbit: Preliminary results, *Bull. Am. Meteorol. Soc.*, 77, 19–40, doi:10.1175/1520-0477.

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 ($RS92_{200701\text{-}201012}$) and those from January 2011 to December 2015 (ΔT ($RS92_{201101\text{-}201512}$) 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.

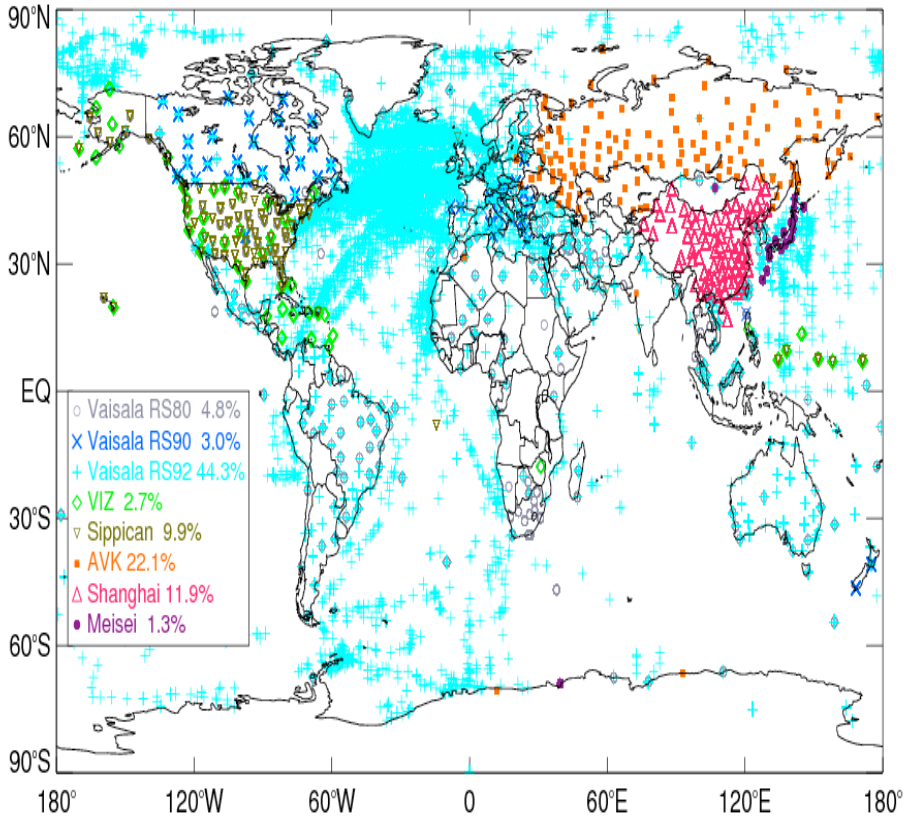

Figure 1. Global distribution of radiosonde stations colored by radiosonde types. Radiosonde types updated from June 2006 to April 2014 are used. The percentage of each type of radiosonde used among all stations is listed. For those stations that radiosonde types are changed during this period, the latest updated radiosonde type is used in this plot. Vaisala RS92 ship observations contain less than 3% of the total RS92 profiles.

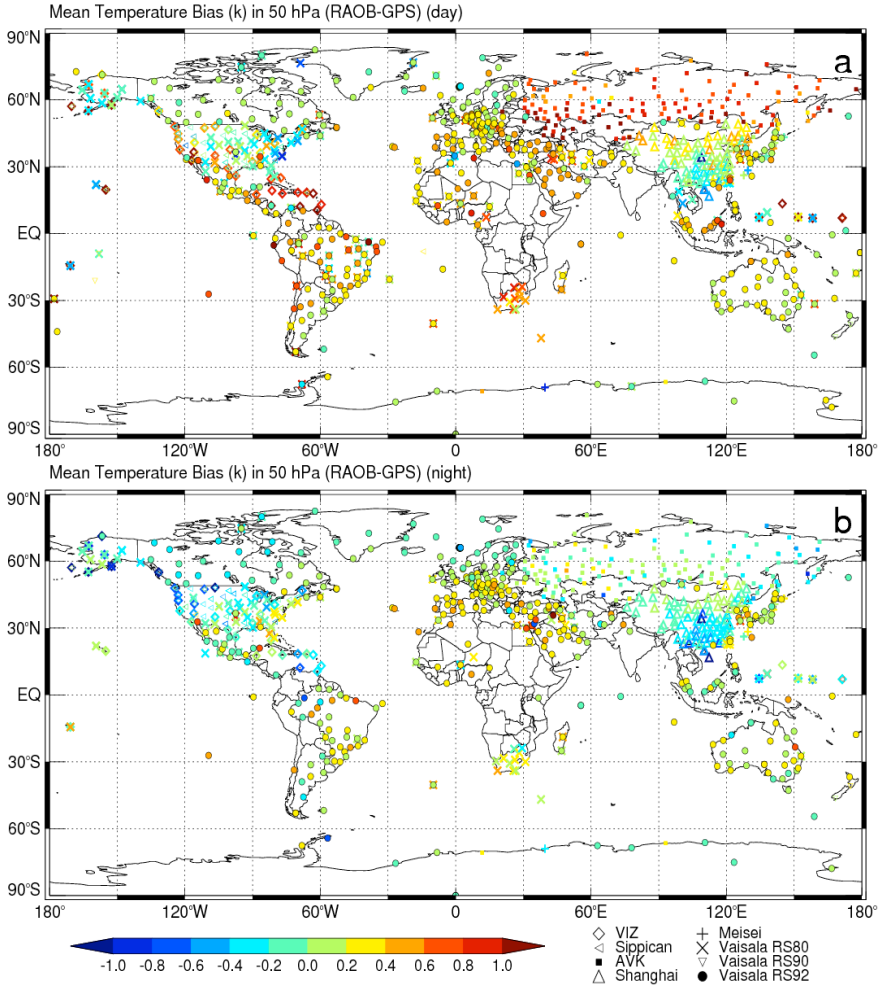

Figure 2. Mean RAOB-RO temperature biases at 50 hPa for the RAOB-RO ensembles from June 2006 to April 2014 for a) daytime, and b) nighttime. Only those stations containing more than 50 RO-ROAB pairs are plotted.

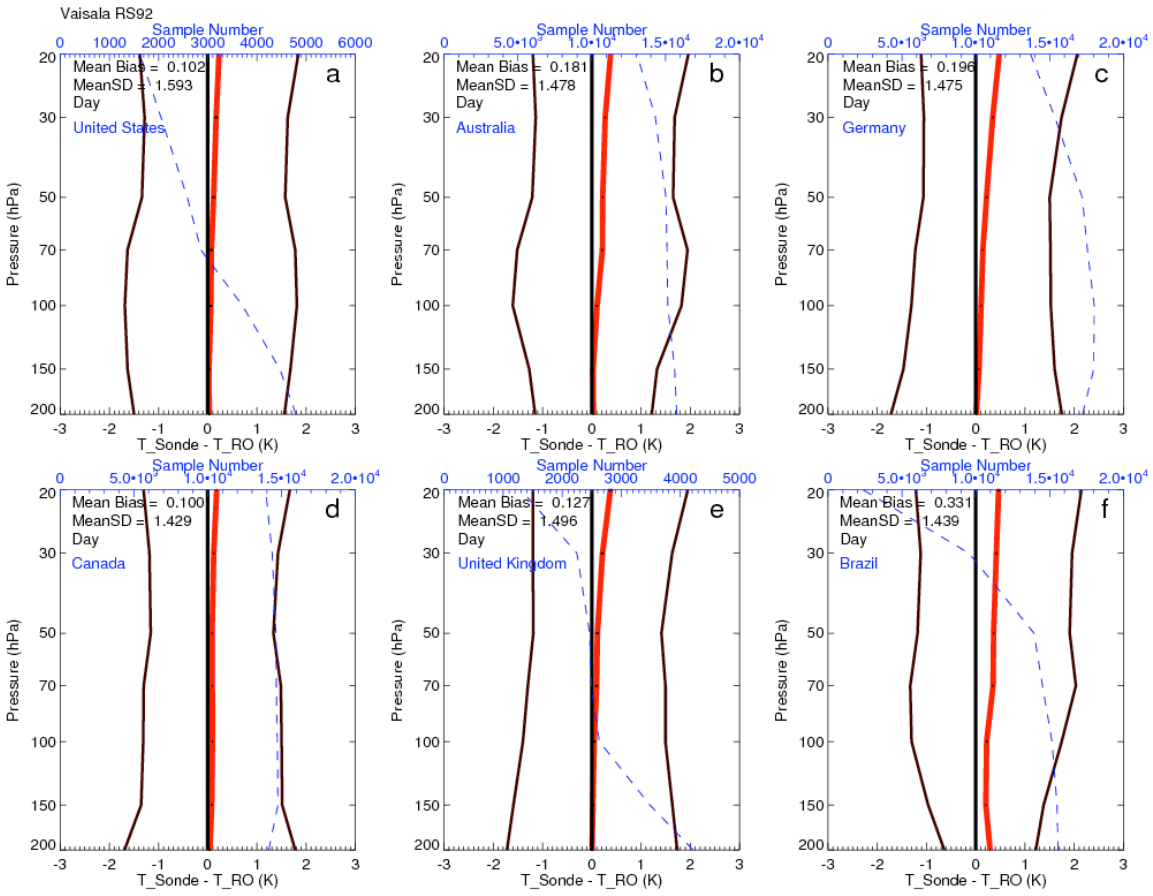

Figure 3. Comparisons of temperature between RS92 and RO for daytime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red line is the mean difference; the black line is the standard deviation of the mean difference; the dotted line is the sample number. 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.

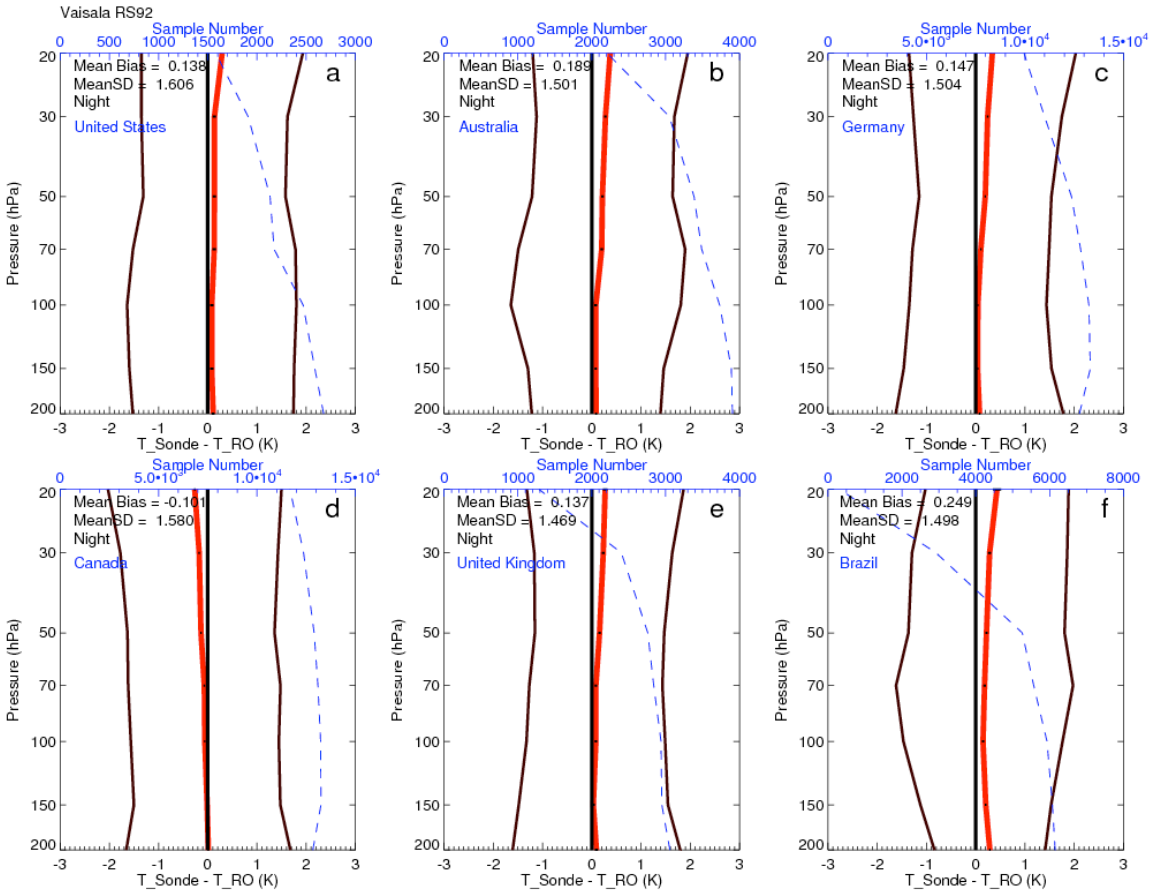

Figure 4. Comparisons of temperature between RS92 and RO for nighttime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.

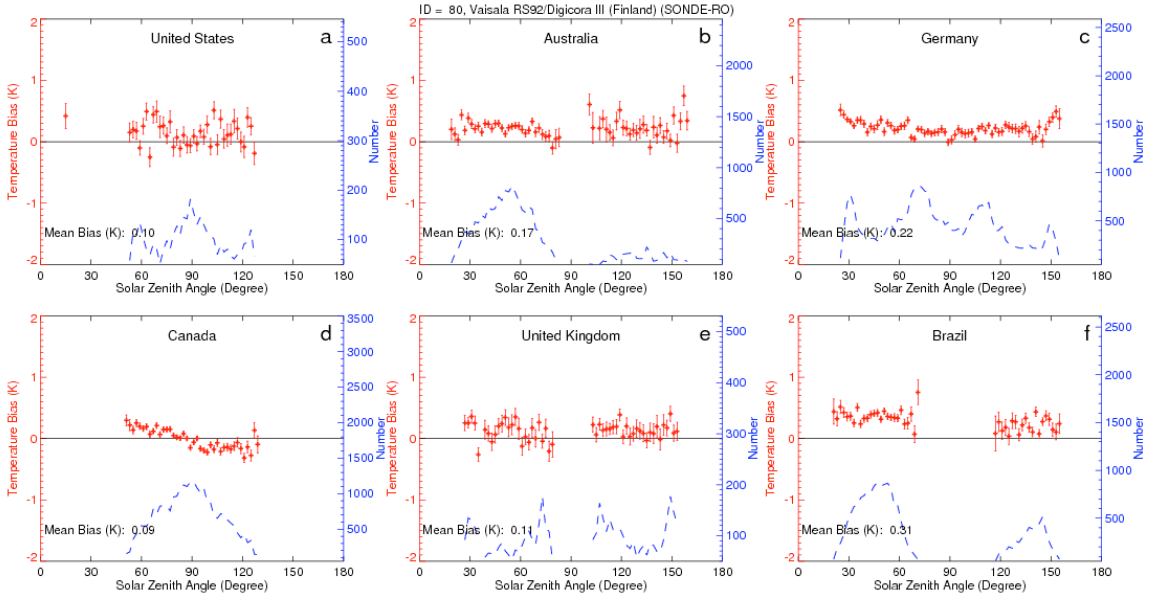

Figure 5. The mean temperature biases (RS92 minus RO) at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for each 5 SZA bins; the red vertical line is the standard deviation 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.

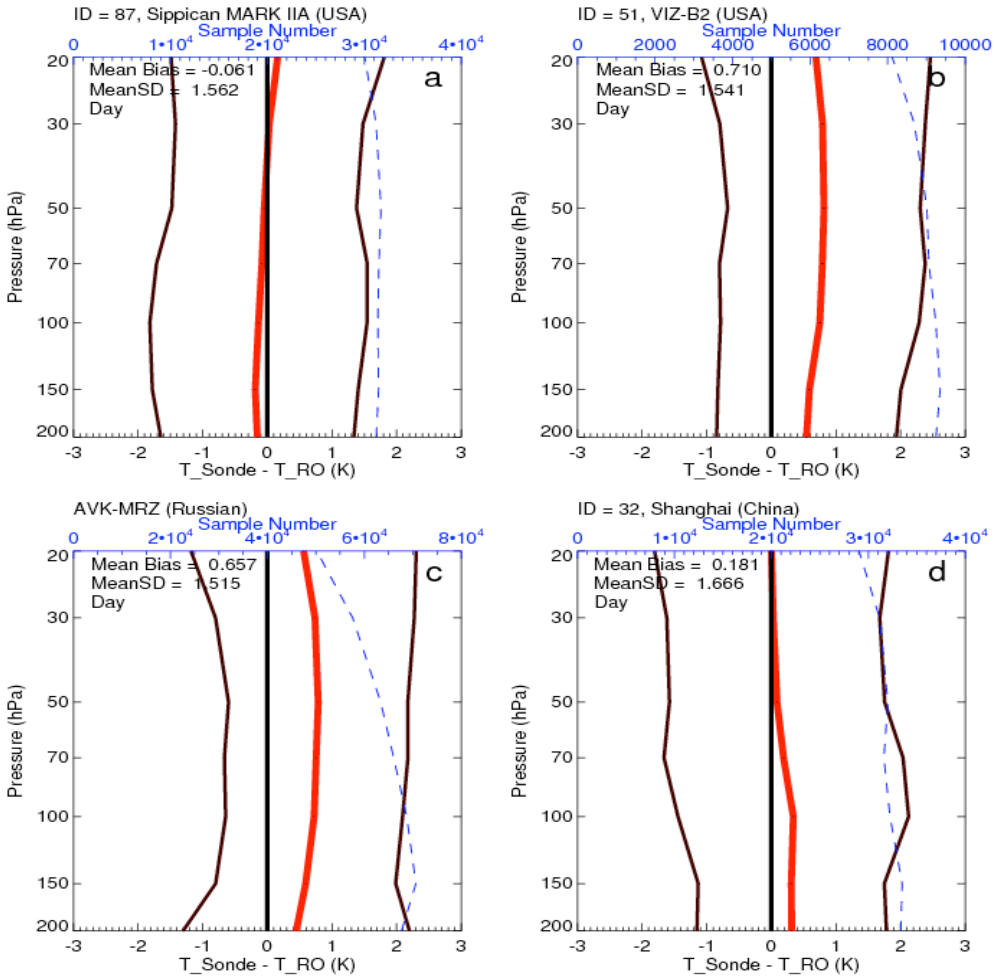

Figure 6. Comparisons of temperature between radiosonde and RO during the daytime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.

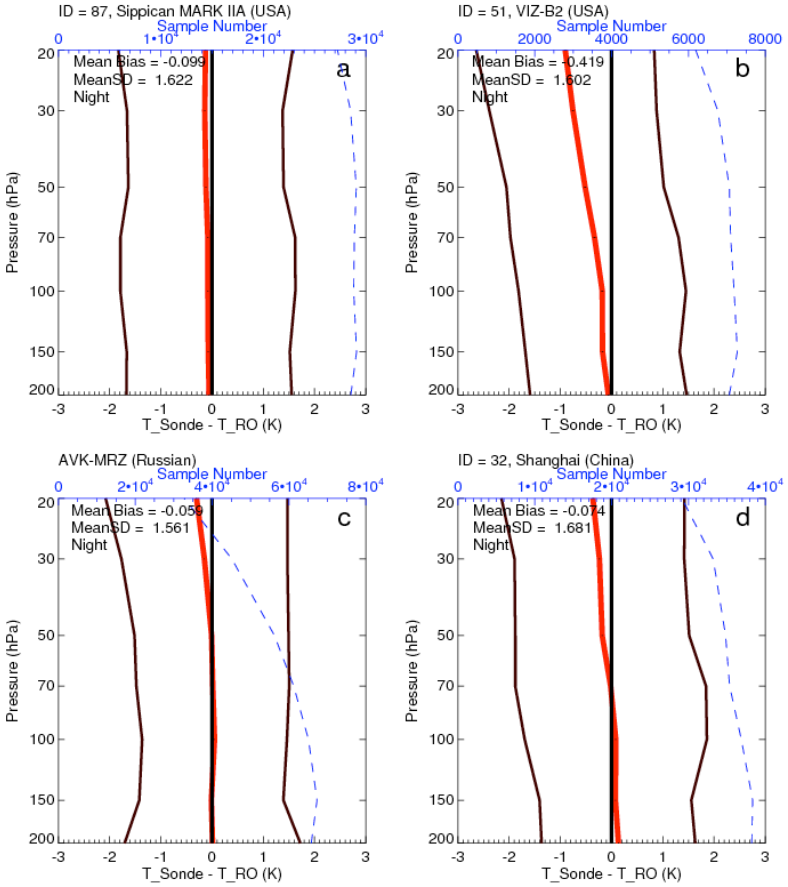

Figure 7. Comparisons of temperature between radiosonde and RO during the nighttime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.

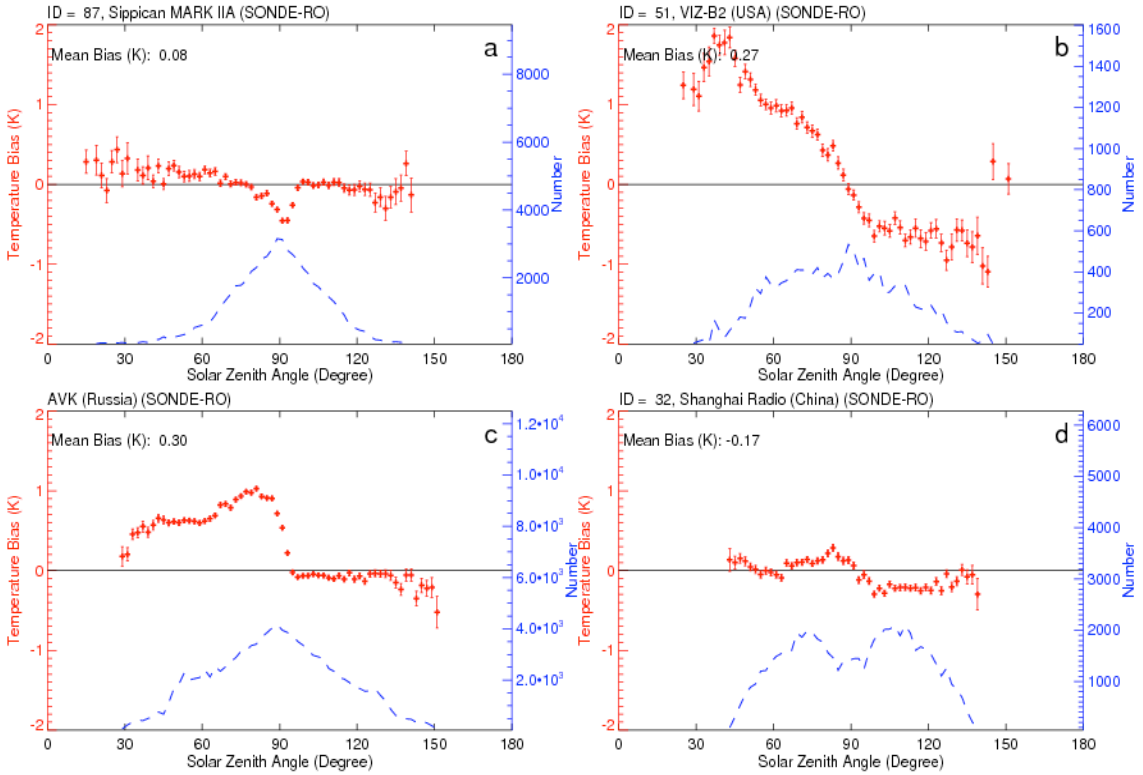

Figure 8. The mean temperature biases at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO. Only bins for more than 50 RAOB-RO pairs are plotted.

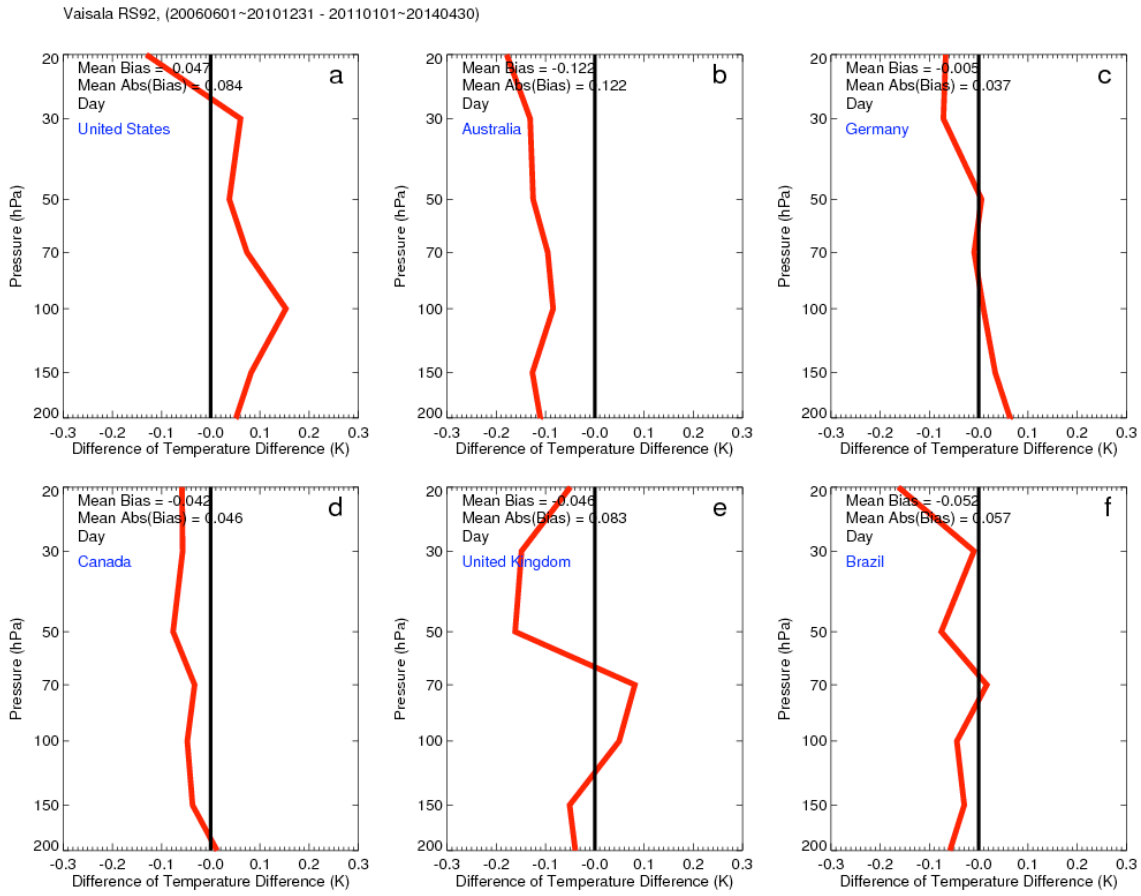

Figure 9. The temperature differences between RS92 – RO from January 2007 to December 2010 (ΔT ($RS92_{200701\text{-}201012}$) and those from January 2011 to December 2015 (ΔT ($RS92_{201101\text{-}201512}$) over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.

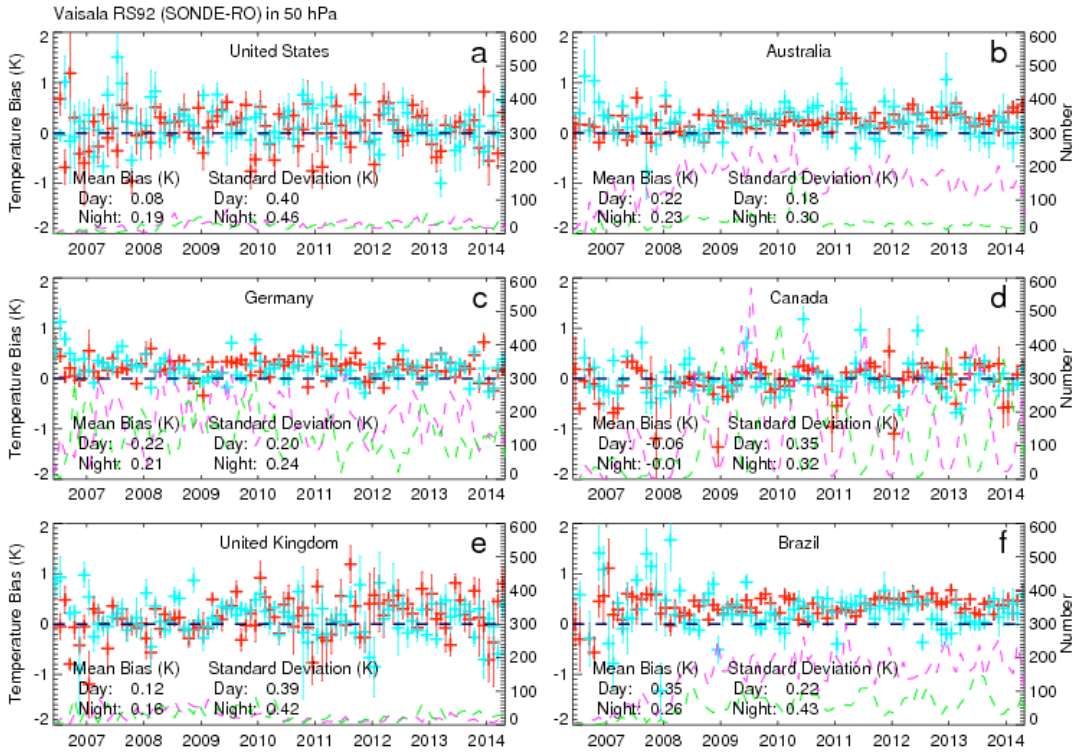

Figure 10. The time series of monthly mean temperature differences to RO at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is 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.

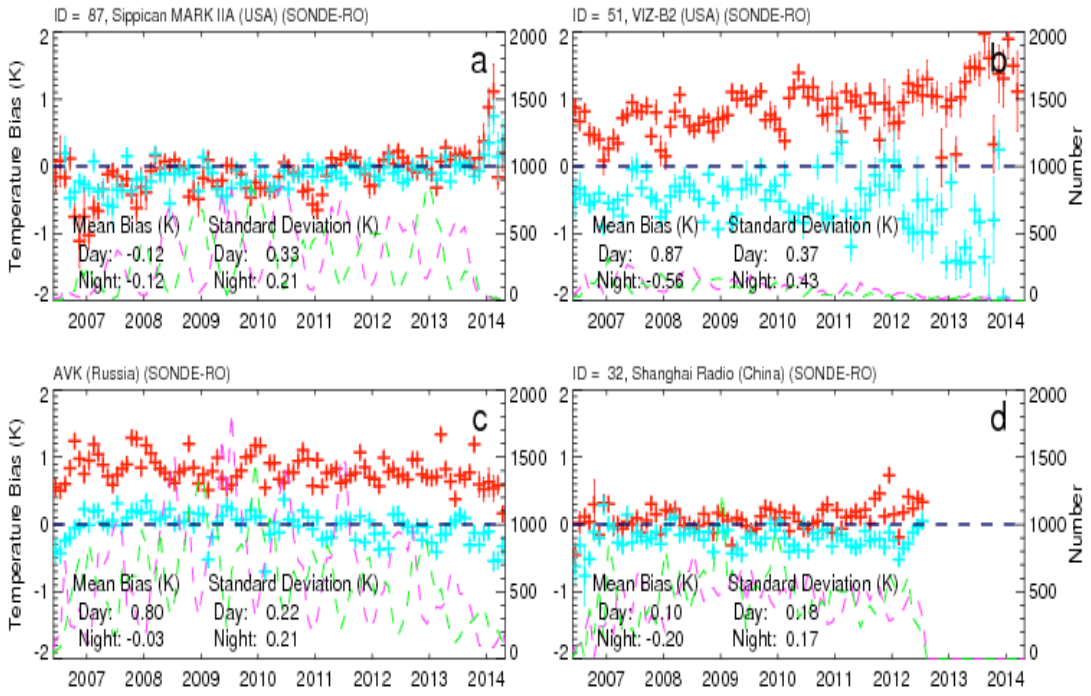

Figure 11. The time series of temperature difference in 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N).

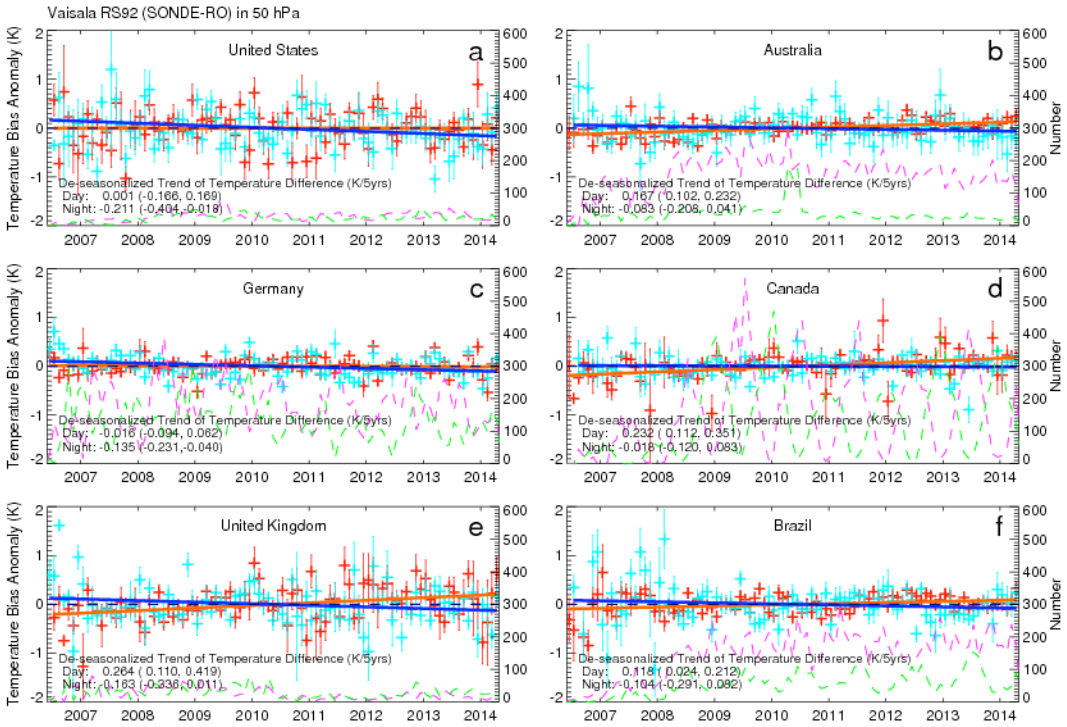

Figure 12. The time series of de-seasonalized temperature differences at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is for 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.

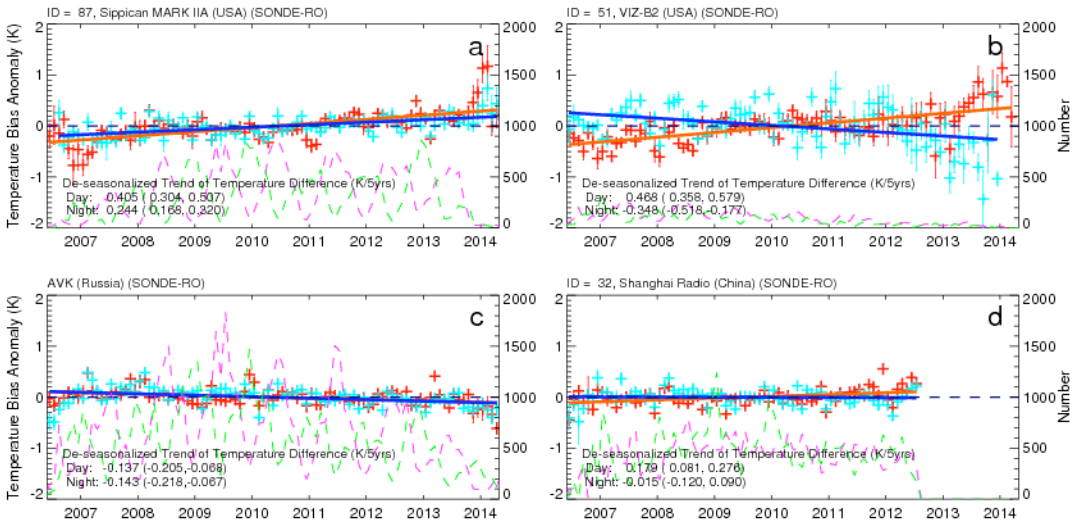

Figure 13. The time series of de-seasonalized temperature differences at 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N). The 95% confidence intervals for slopes are listed 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.

	ID	Sensor type	Availability	Solar absorptivity	Infrared emissivity	Number of RO-RAOB pairs
RS80	37	Bead thermocap	1981~ 2014	0.15[*Luers and Eskridge*, 1998]	0.02	1624
Vaisala RS80-57H	52	Bead thermocap	early 1990s [*Redder et al.*, 2004] ~ Jul 2012	0.15	0.02	13192
Vaisala RS80/Loran	61	Bead thermocap	~ 2014	0.15	0.02	11591
Vaisala RS80/Digicora III	67	Bead thermocap	~ 2012	0.15	0.02	2864
Vaisala RS90/Digicorn I, II	71	Thin wire F-thermocap [*Sun et al.*, 2010]	1995 ~ 2014	0.15[*Luers*, 1997]	0.02	18082
Vaisala RS92/Digicora I/II	79	Thin wire F-thermocap [*Sun et al.*, 2010]	2003 ~ 2014	0.15	0.02	40478
Vaisala RS92/Digicora III	80	Thin wire F-thermocap	2004~2014	0.15	0.02	184542
Vaisala RS92/Autosonde	81	Thin wire F-thermocap	2011~2014	0.15	0.02	42577
AVK-MRZ	27	Rod thermistor [*Sun et al.*, 2010]	~ 2014	0.2[*He et al.*, 2009]	0.04	48954
AVK-BAR Russian	58	Rod thermistor	2007 ~ 2014	0.2	0.04	26020
AVK-MRZ (Russian)	75	Rod thermistor	~ 2013	0.2	0.04	9472
MARL-A or Vektor-M-MRZ (Russian)	88	Rod thermistor	~ 2014	0.2	0.04	23326
MARL-A or Veltor-M-BAR (Russian)	89	Rod thermistor	~ 2014	0.2	0.04	25715
VIZ-B2	51	Rod thermistor[*Sun et al.*, 2010]	1997[*Elliott et al.*, 2002]~ 2014	0.15[*Luers and Eskridge*, 1998]	0.86	16310
Sippican MARK II A Chip	87	Chip thermistor[*Sun et al.*, 2010]	1998[*Elliott et al.*, 2002]~ 2014	0.07[*Luers and Eskridge*, 1998]	0.85	59775
Shanghai	32	Rod thermistor	1998 ~ 2012	<0.07 [*Wei*, 2011]	>0.90	71605
Meisei Japan	47	Thermistor [*KOBAYASHI et al.*, 2012]	1994 ~ 2013	0.18[*Luers and Eskridge*, 1998]	0.84	7888

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[a,b]. [a]The values of standard deviations of temperature differences are shown in the parentheses. [b]The sample number are for the ROAB-RO pairs available in the same time period.

	ID	All Day and night mean(std)/ sample numbers	Day mean(std)/ sample numbers	Night mean(std)/ sample numbers
Vaisala RS80	37, 52, 61, 67	0.10 (1.54)/29271	0.10 (1.53)/15947	0.09 (1.55)/13324
Vaisala RS90	71	0.13 (1.54)/18082	0.16 (1.51)/8758	0.11 (1.57)/9324
Vaisala RS92	79, 80, 81	0.16 (1.52)/267597	0.20 (1.50)/161019	0.09 (1.55)/106578
AVK	27, 75, 88, 89, 58	0.33 (1.58)/133487	0.66 (1.51)/67679	-0.06 (1.56)/65808
VIZ-B2	51	0.22 (1.67)/16310	0.71 (1.54)/9246	-0.42 (1.60)/7064
Sippican MARKIIA Chip	87	-0.08 (1.59)/59775	-0.06 (1.56)/31230	-0.10 (1.62)/28545
Shanghai	32	0.05 (1.68)/71605	0.18 (1.67)/33360	-0.07 (1.68)/38245
Meisei Japan	47	0.11 (1.69)/7888	0.03 (1.71)/3849	0.19 (1.66)/4039

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.

	United States		Australia		Germany		Canada		United Kingdom		Brazil	
	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
Mean Bias	0.08	0.19	0.22	0.23	0.22	0.21	-0.06	-0.01	0.12	0.16	0.35	0.26
std of Mean Bias	0.4	0.46	0.18	0.3	0.2	0.24	0.35	0.32	0.39	0.42	0.22	0.43
De-seasonalized Trend of Differences (K/ 5 yrs)	0.001	-0.211	0.167	-0.083	-0.016	-0.135	0.232	-0.018	0.264	-0.163	0.118	-0.104
De-seasonalized Trend of RO Temperature (K/5yrs)	0.941	0.506	-0.26	0.082	0.29	0.708	-0.69	-0.534	0.509	1.143	-0.076	-0.354
RMS of de-seasonalized difference	0.365	0.439	0.161	0.275	0.173	0.22	0.276	0.215	0.358	0.392	0.212	0.398

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.

		Mean Bias	STD Of MB	De-seasonalized Trend of Difference (k/5yrs)	RMS of Difference	Mean Bias	STD Of MB	De-seasonalized Trend of Difference (k/5yrs)	RMS of Difference
RS80	37,52, 61,67	0.18	0.29	0.187 (0.073,0.301)	0.268	0.13	0.33	0.114(-0.019,0.248)	0.301
RS90	71	0.16	0.29	-0.006 (-0.123,0.111)	0.26	0.17	0.38	0.043(-0.115,0.201)	0.352
RS92	79,80, 81	0.22	0.07	0.074 (0.051,0.097)	0.062	0.12	0.12	-0.094(-0.131,-0.057)	0.093
Russia	27,75, 88,89 58	0.8	0.22	-0.137 (-0.205,-0.068)	0.164	-0.03	0.21	-0.143(-0.218,-0.067)	0.18
VIZ-B2	51	0.87	0.37	0.468 (0.358,0.579)	0.322	-0.56	0.43	-0.348(-0.518,-0.177)	0.386
Sippican MARKIIA Chip	87	-0.12	0.33	0.405 (0.304,0.507)	0.292	-0.12	0.21	0.244(0.168,0.320)	0.197
Shanghai	32	0.1	0.18	0.179 (0.081,0.276)	0.161	-0.2	0.17	-0.015(-0.120,0.090)	0.159
Meisei Japan	47	0.07	0.69	0.006 (-0.353,0.365)	0.619	0.05	0.51	-0.086(-0.369,0.197)	0.494