10/24/16 6:00 AM
Comments on *Characterization of the Long-term Radiosondes Temperature Biases….* by Ho et al. (ACP-2016-801)
Rick Anthes

Line numbers refer to ACP-2016-801.pdf

Overall comments

The authors have done a lot of careful work and there are interesting results in their paper on the different accuracies, stabilities and trends of various types of radiosonde data. The authors use radio occultation (RO) data in the upper troposphere and lower stratosphere as a reference data set for the period June 2006 to April 2014. Since radiosonde data are often used in climate studies, it is important to document the accuracies and uncertainties of different types of radiosondes in different countries.

1. However, the paper is too long (58 pages) and the sections on results (Sections 3-5) are tedious and difficult to read because of too much detail in the text that merely duplicates what is in the tables and figures as well as too many symbols in the text (e.g. ΔT ($RS92_{200701\text{-}201012}$)). Furthermore, much of the detail is describing statistics that are small and probably not statistically significant or of general interest. The reader is overwhelmed with the reporting of many numbers without a focus on what is really important and what is of little or no interest. The paper would be greatly improved and have more impact if it were shortened significantly and only the major results included in the text.

- ⇨ As suggested by the reviewer, we shorten this paper significantly. We rewrote Section 3-5, combining section 4.2, section 4.3, and section 5 into a new section (new section 4.2). Many symbols (e.g. ΔT ($RS92_{200701\text{-}201012}$)) are removed in the text. In addition, we removed results for 150 hPa (i.e., Tables 4 and 6) because the results were similar to those at 50 hPa. Appendix A is also removed, yet a part of the Appendix A is inserted in the Introduction section. To demonstrate whether the results are statistically significant or not, we performed statistical significance tests for RAOB – RO trend difference. We only mention that the statistical significance results in the revised paper. The revised manuscript is now reduced from 58 pages to 47 pages. All changes are tracked and the tracked manuscript is also submitted. In addition, since the heights from 150 hPa to 50 hPa are in part of upper troposphere, we added " "Upper Troposphere and" to the title.

2. There are many statistics of radiosonde minus RO temperatures from various types of radiosondes at different levels of the atmosphere between 200 and 20 hPa over six different regions of the world. It is not clear which of these results are statistically significant and which ones we should be concerned about. This makes interpretation of the results difficult as we could be looking at small, quasi-random differences that have no physical meaning, nor even meaning relative to the specific types of radiosonde data. Differences are often 0.1K or less, which are well below the accuracy of radiosonde

sensors. When the different atmospheric sampling volumes of the radiosondes and RO are considered, sampling errors alone can be much larger than 0.1K. It would be helpful if the authors could do statistical significance tests and describe in the text only the results that are significant at the 95% or higher level.

- ⇨ To demonstrate if the computed de-seasonalized trend differences (RAOB – RO) are statistically significant or not, we performed statistical significance tests for the trend difference. In Figs. 12 and 13 we list the 95% confidence intervals for trend difference (ROAB – RO) in the parentheses.

- ⇨ The 95% confidence intervals for trend differences for global Vaisala (RS80, RS90, and RS92) and other sensor types during the daytime and night time in the North hemisphere mid-latitude (60°N-20°N) are summarized in the Table 4. We also discuss what the trends in radiosonde minus RO temperatures and RO temperatures means is the text (see the reply for comment 5).

3. The authors compute trends of the differences between individual types of radiosonde and RO over a 7-year period. Most of the trends are small (of order 0.2 K per five years) and quite different, with some being positive and some being negative. It is not clear what these trends mean, except as an indication of the uncertainty of the radiosonde minus RO temperatures over this short period. Again an estimate of the statistical significance of these trends would be useful. What would the magnitude of trends computed from a similar time series of random data with the same standard deviation as these differences be? A comparison of what sort of trends in temperature at these levels due to long-term climate change would be useful as well. For example, from climate models we might expect a temperature trend in the lower stratosphere to be something like 5 K per 100 years or 0.25 K per five years. 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 (England, day), so they are comparable or slightly smaller than what one would expect for a long-term climate trend signal.

- ⇨ To compare trends in temperature at these levels due to long-term climate change with RO trends in this paper, we refer to stratospheric temperature trends over 1979–2015 computed by Randel et al., (2016) in line 466. Randel et al., (2016) is also added in the references. Randel et al., (2016) indicated that the linear trends over 1979–2015 show that cooling in the lower stratosphere is about -0.1 K to -0.2 K/decade. In line 464 we added, "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)."

4. It would also be interesting to compare these trends in radiosonde-RO temperature differences to the corresponding trends in the RO temperatures over this period. Indeed, Tables 3 and 4 give the RO trends, but they are never mentioned in the text!

- ⇨ As suggested by the reviewer, we rewrote sections 4 and 5 and compared de-seasonalized trends of radiosonde-RO temperature differences to the corresponding trends in RO temperatures over this period.

- ⇨ In line 455 we added "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."

- ⇨ To shorten the paper, we also removed results for RAOB and RO temperature comparisons at 150 hPa (old Tables 4 and 6). Now the new Tables 3 and 4 are specifically mentioned in the new text.

5. The RO temperature trends at 50 hPa (Table 3) range from -0.69 (Canada, day) to 1.143 (England, night). Quite different values are found at 150 hPa (Table 4), with the 5-year trends ranging from -0.797 (Canada, day) to 1.508 (U.S. day). In general, the magnitudes of the trends of radiosonde-RO temperature differences are smaller than the trends in RO temperatures, which is an indication of the consistency between the radiosonde and RO temperatures. The large differences in RO temperature trends between regions (much larger than expected for a long-term climate change signal) probably indicates natural variability in the six different regions. The fact that they are larger than the trends in the differences indicates to me that they are a real signal in the different regions over this 7-year time period. Presumably, since the radiosonde-RO trends are smaller, the radiosondes (at least the good ones) would pick up similar trends to the RO trends. A discussion of what the trends in radiosonde minus RO temperatures and RO temperatures means is needed.

- ⇨ We specifically added several discussions of what the trends in radiosonde minus RO temperatures and RO temperatures means in the new Section 4.2.

  As mentioned in the reply for comment 4, we added a discussion of what the trends in radiosonde minus RO temperatures and RO temperatures means in Line 455. In line 460, we specifically discuss what the trends in radiosonde minus RO temperatures and RO temperatures means: "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. "

- ⇨ In lines 470-480, we discuss what trends in radiosonde minus RO temperatures and RO temperatures means by stating "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."

- ⇨ We discuss the mean bias in the last two paragraphs of Section 4.2. In line 481 we state "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 VIZ-B2 and Sippican, which both exceed 0.4 K/5 yrs. These are much larger than those of the Vaisala RS92 (0.074 K/5 yrs)."

- ⇨ In line 489, we state, "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."

In summary, the paper contains some interesting and important results and should be published, but it requires significant rewriting, editing, and shortening with greater emphasis on what the important results are and less detail on all the individual numbers.

<u>Detailed comments</u>

1. The papers use three terms to describe the radiosonde-RO temperature differences: differences, biases, and anomalies. I suggest using only differences and biases, and eliminate all references to anomalies.

⇨ We replace all "anomalies" with "differences" in this paper.

2. Do you mean United Kingdom rather than England?

   ⇨ Yes, it shall be "United Kingdom". We replace all "England" with "United Kingdom" in this paper and Figures. The revised Figs. 3, 4, 5, 9, 10, and 12 are inserted.

3. An example of how a difficult to read paragraph containing a repetition of data in a table can be simplified, shortened, and made more readable is lines 264-267:

"In general, the radiosonde temperature biases vary for different sensor types. The mean *ΔT* for RS92 (0.16 K), RS80 (0.10 K), RS90 (0.13 K), Sippican MarkIIA (-0.08 K), Shangai (0.05 K) and Meisei (0.11 K) are smaller than those for AVK (0.33 K) and VIZ-B2 (0.22 K) (see Table 2)" (50 words) may be replaced with

"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." (31 words).

   ⇨ The sentence in lines 264-267 was revised as suggested by the reviewer. Similar sentences in old Sections 3-5 were also revised and are not specifically mentioned. The tracked manuscript is submitted.

4. An example of unnecessary use of symbols in a sentence which makes reading difficult is: "The mean temperature biases in this region for $\Delta^{Time}_{Sippican}$, Δ^{Time}_{VIZ-B2}, Δ^{Time}_{AVK}, and $\Delta^{Time}_{Shanghai}$ for 50 hPa are summarized in Table 5." This can be written in much more readable form as "The mean temperature biases in this region for the Sippican, VIZ-B2, AVK, and Shanghai radiosondes at 50 hPa are summarized in Table 5."

   ⇨ We removed many symbols (for example, $\Delta^{Time}_{Sippican}$) in this paper and many sentences are revised as suggested. For example, in line 426, we now state "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." No symbols are used in the sentence.

5. Lines 439-442. I don't understand this sentence. If the U.S. did not use RS92 radiosondes before 2012, there would be no data for comparison with RO before 2012 (i.e. none in the period 2007-2010). However, this section talks about RS92 from Jan 2007 to Dec 2010 for the U.S. and Fig. 10 shows RS92 vs. RO going back to 2007. Also, a small number of pairs in the comparison does not necessarily imply small differences—in fact, a small number of pairs could lead to large differences due to an inadequate sample size.

- ⇨ The US National Weather Service (NWS) did not use Vaisala RS92 radiosondes before 2012. The Vaisala RS92 radiosondes before 2012 were mainly launched by research groups (for example, at the ARM site and during individual field experiments from universities, etc).

- ⇨ To avoid confusion by the readers, we deleted the statement "since the US National Weather Service (NWS) did not use Vaisala RS92 radiosondes before 2012."

6. It seems strange that Table 3 is not mentioned until line 563, long after Tables 4, 5 and 6 are mentioned and discussed.

   - ⇨ In the revised paper, we refer to Table 3 before Table 4. The Table 3 is now referred in line 415 whereas Table 4 is first referred in line 426.

***Interactive comment on* "Characterization of the Long-term Radiosonde Temperature Biases in the Lower Stratosphere using COSMIC and Metop-A/GRAS Data from 2006 to 2014" *by* Shu-Peng Ho et al.**
Anonymous Referee #2

Received and published: 15 December 2016

This manuscript compares COSMIC and MetOP/GRAS GPS RO data with radiosonde temperatures in the interval 2006-2014. While this is not the first comparison of these data sets, the one presented here is an improvement due to the long time interval considered and because reprocessed COSMIC data have been used. The comparison is comprehensive and detailed. Accurate estimation of differences between radiosondes and GPS-RO is important since both are potentially used as "anchors" in reanalysis efforts. Other, less accurate data such as satellite radiances or aircraft temperatures are often bias-adjusted adaptively (Dee and Uppala 2009), whereas "anchors" are not.

Therefore I recommend eventual publication of the manuscript after undergoing the following major revisions:

1) Before the actual intercomparison, it should be specified what the structural uncertainties of the two measurement technologies are. These are mentioned for RO retrievals only in the supplement (+/-0.1 K in the 20-200 hPa range). For many modern radiosondes (RS92 in particular) they are specified as +/- 0.2 K below 100 hPa and somewhat higher at higher levels. RS-RO differences that fall within this range, especially if they are different in different regions of the world, should not be considered as "bias", as they may have other causes than systematic measurement errors. Small sample sizes or the different volumes sampled may be the reasons for the differences.

- ⇨ As suggested by the reviewer#1, we shortened this paper significantly. In the revised paper line 261, we added "For many modern radiosondes (for example RS92) the structural uncertainties are +/- 0.2 K below 100 hPa and somewhat higher at higher levels."

2) Modern radiosondes measure up to the 5hPa level, whereas this comparison stops at 20 hPa. Presumably this conservative choice is related to uncertainties in the inversion of the Abel integral necessary for the conversion of bending angles to refractivities. They lead to larger strucural uncertainties of the RO method. Could you elaborate on this, and also if the +/- 0.1K uncertainty specified for RO profiles applies to the 20 hPa level.

- ⇨ In the revised paper, we moved parts of Appendix A (The Quality of GPS RO Data as Benchmark References and the New Reprocessed Package) to the introduction section. In line 112, we specifically stated "At 20 hPa, the mean temperature difference between COSMIC and CHAMP was within 0.05K (Ho et al., 2009b)."
- ⇨ In line 113 we specially stated the mean layer temperature difference between 200 hPa to 10 hPa is within 0.05 K, and at 20 hPa, the mean temperature difference is

equal to 0.03 K:” 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 current results are for 200 hPa and 20 hPa, where the ionospheric effect is minimal.

- ⇨ To estimate the uncertainty of RO temperatures in the upper troposphere and lower stratosphere, particularly between 200 hPa and 10 hPa, we stated in line 122: “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.”

- ⇨ Based on above studies, we are confident that an uncertainty of +/- 0.1K for RO profiles does apply to the 20 hPa level.

The other review points are minor:

3) The trend comparisons are difficult to interpret since the time interval is so short. Also the regional trend variability is much larger than the trend differences between RS and RO, at least for the more accurate radiosonde types.

- ⇨ As suggested by the reviewer#1, we shortened this paper significantly. We rewrote sections 3-5, combining section 4.2, section 4.3, and section 5 into a new section (new section 4.2). In the new section 4.2, we add several paragraphs to discuss what the trends in radiosonde minus RO temperatures and RO temperatures means.

  As mentioned in the reply for comment 4, we added a discussion of what the trends in radiosonde minus RO temperatures and RO temperatures means in Line 455. In line 460, we discuss what the trends in radiosonde minus RO temperatures and RO temperatures mean: “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.”

- ⇨ In lines 470, we discuss what the trends in radiosonde minus RO temperatures and RO temperatures mean by stating "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."

- ⇨ We discuss the mean bias in the last two paragraphs of Section 4.2. In line 481 we stated "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)."

- ⇨ In line 489, we stated "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."

4) When looking at the maps in Figure 2, it seems there is quite some heterogeneity even in countries with the same sensor, particularly at daytime, e.g. over China and Brasil. Can you give an explanation? It appears that the radiosonde type is not the only factor that determines the temperature biases. Do you think it is possible to estimate the biases also for each station individually? This has been done by several authors when homogenizing radiosonde time series.

- ⇨ We suspect that the heterogeneity over China may be due to inconsistent corrections applied in northern and southern provinces of China. In general, the Chinese sondes contain their corrections, which are not documented in public literature.

- ⇨ The heterogeneity over Brazil may be due to a smaller sample size at certain stations. For example, we found that stations with temperature biases larger than 0.5 K in the east Brazil contain only about 60 RO-RAOB pairs.

- ⇨ Two more sentences were added in section 3.1 “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.”

5) The thresholds for daytime/nighttime (SZA < or > 90 deg) may not be optimal. Fig. 8c clearly shows positive biases at 90 deg, only at >95deg they are negative. Also the VIZ B2 and Shanghai sondes seem to reach their nighttime value at SZA clearly larger than 90 deg. I am also asking for which times the SZA have been calculated? Nominal launch time of the radiosonde or time of collocation? Please clarify.

- ⇨ We tested several criteria and decided to use thresholds of SZA < 90 as daytime and SZA > 90 as nighttime. It is possible that this could contain scattering effects during dawn and sunset, but we given the uncertainties of the actual time of observations, this threshold appeared most appropriate.

- ⇨ The SZA is computed from the launch time and location of sonde station since the information of specific time and location of sonde at different height is not available.

- ⇨ We added: “The SZA is computed from the synoptic launch time and location of sonde station since the time and location of the sonde at different height is not available.” in the end of section 2.3.

6) What is the reason for the strong decrease in measurement numbers already at 70 hPa over the US (Fig. 3a)? Did the reports from higher levels go missing? At most other places RS92 sondes consistently reach 20 hPa.

- ⇨ The height is determined by the balloon used at various sites. There are about 15 stations launching RS92 during the study period. Our best guess is that these US stations are only interested in the tropospheric profiles and use smaller balloons. Meteorological services usually try to get to 50 or 30 hPa for all soundings and use slightly larger balloons. GRUAN stations are required to reach 5 hPa and should use larger balloons. The sondes launched ARM site also reach to 5 hPa.

- ⇨ To provide the possible reason for the strong decrease in measurement numbers already at 70 hPa over the US (Fig. 3a), in Line 314 we added “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."

7) Fig. 3 onwards: You plot means and standard deviations. Instead you could plot means and the standard deviations of the MEAN (sigma^2/sqrt(N)) or 95% confidence intervals. This would allow a smaller scale for the x-axes.

- ⇨ I think the review is talking about "standard error of the mean". We did plot the standard error of the mean. Since there are a lots of RO-RAOB pairs, the standard error of the mean is too small to see.

- ⇨ The standard deviation of the mean are plotted in current Figs. 3, 4, 6, 7. We also add the standard error of the mean in Figs. 3, 4, 6, and 7. To make it clear, we re-state "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 " in the caption of Fig. 3.

8) Figs 5,8: Please triple number scale so that there is less intersection between number line and departures.

- ⇨ The number scale in Figs. 5 and 8 is revised. The new Figs. 5 and 8 are used in the paper.

9) Fig. 9: Are these differences significant? The samples are smaller here. If std is 1.5K and number is 1000 for both samples, then the std of the means is roughly +/-0.05. For a 95% confidence interval you have to multiply with 1.96. Thus a large fraction of the differences shown in Fig. 9 would be insignificant.

- ⇨ The purpose of Fig. 9 is to use RO temperature as references to identify the RS92 temperature biases due to change of radiation correction. With the uncertainty of RO data (+/- 0.1K uncertainty) and RAOB data ((+/- 0.2K uncertainty below 100 hPa and larger uncertainty above that), it is hard to say the results are significant.

- ⇨ Therefore, we added a new paragraph in line 387 "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."

10) Figs 11-13: Are the trends or trend differences significant? Please give confidence intervals for slopes.

- ⇨ Confidence intervals for slopes are added in each panel in Figs. 12 and 13. The confidence intervals for slopes are shown in the parentheses in each panels of Figs. 12 and 13.

11) Layer mean 20-200 hPa bias values in tables 1,2 are of limited use, since the biases changes a lot over this range of pressure.

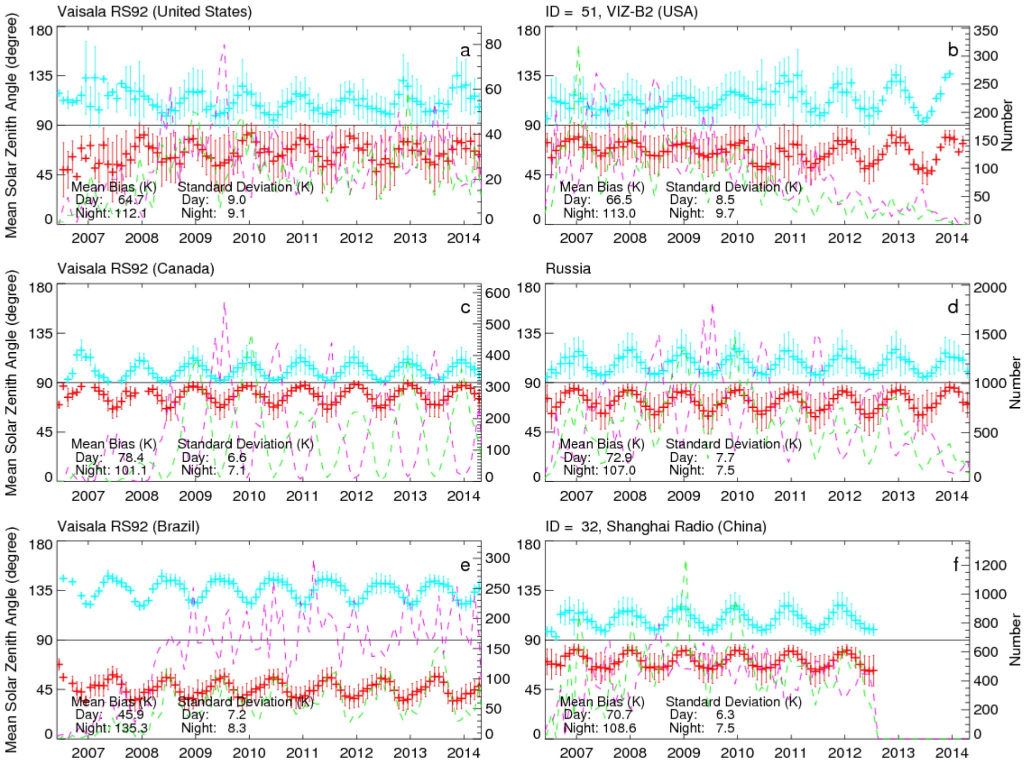
- ⇨ Table 2 summarizes the change of the mean and standard deviation of temperature differences (K) between 200 hPa and 20 hPa between RO and eight types of radiosonde. This is to demonstrate that RO temperature can be used as references to distinguish the temperature biases among sonde types and their biases at daytime and nighttime of the comparison in the rest of the paper. We think this is important and we will keep Tables 1 and 2.

12) Tables 3,4: What do you think is the reason for the very different biases over Brasil at 150 hPa for RS92 sondes? This level is well below the tropopause. Is it possible that water vapor or cloud content could adversely affect the RO estimates there. These effects have been neglected in Formula 1.

- ⇨ In the revised paper we remove results for those for 150 hPa (i.e., old Tables 4 and 6) because the results were similar to those at 50 hPa.

- ⇨ The reason for the larger biases over Brazil for RS92 may be due to the incomplete bias correction. Figure 5 shows that the mean ΔT (RS92) has a slightly larger warm bias for low SZA (near noon) than that at higher SZA (late afternoon and in the night). The mean SZA for the RO-RS92 pairs over USA, Canada, and Brazil are 64.7 degree, 78.4 degree, and 45.9 degree, respectively. Because daytime SZA over Brazil is in general smaller (close to the noon) than other regions, the Brazil temperature biases relative to the collocated RO data are higher than other regions.

- ⇨ The attached figure (this will not be shown in the paper) depicts the seasonal variation of SZA over different regions.



13) l718: traceability, not tractability.

⇨ In line 718 and 491, the “tractability“ is replaced by “traceability “

14) l1385: RAOB instead of ROAB

⇨ In line 11385, the “ROAB” is replaced by “RAOB”

15) l653-655: Some words are missing, the sentence does not seem to be complete.

⇨ The sentence is completed now.

16) l558: non-trivial

⇨ In line 1558, “non-trivial” is added.

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.

Ben Ho 12/5/2016 2:20 PM
Deleted: RAOB

Ben Ho 12/5/2016 2:22 PM
Deleted: RAOB

Ben Ho 12/5/2016 2:23 PM
Deleted: (ΔT)

Ben Ho 12/8/2016 10:21 AM
Deleted:

Ben Ho 12/5/2016 2:23 PM
Deleted: 0.18 K in Australia, 0.20 K in Germany, 0.10 K in Canada, 0.13 K in England, and 0.33 K in Brazil

Ben Ho 12/5/2016 2:24 PM
Deleted: ΔT

Ben Ho 12/5/2016 2:24 PM
Deleted: 06

Ben Ho 12/5/2016 2:24 PM
Deleted: 71

Ben Ho 12/5/2016 2:24 PM
Deleted: 66

Ben Ho 12/5/2016 2:25 PM
Deleted: 0.18

Ben Ho 12/5/2016 2:25 PM
Deleted: anomalies

Ben Ho 12/5/2016 2:25 PM
Deleted: 0

Ben Ho 12/5/2016 2:25 PM
Deleted: yrs over United States, -0.02 K/5 yrs over Germany, 0.17 K/5yrs over Australia, 0.23 K/5yrs over Canada, 0.26 K/5yrs over England, and 0.12 K/5yrs over Brazil, respectively.

Ben Ho 12/5/2016 2:26 PM
Deleted: measurements

Ben Ho 12/5/2016 2:28 PM
Deleted: anomaly

Ben Ho 12/5/2016 2:28 PM
Deleted: globally

Ben Ho 12/5/2016 2:28 PM
Deleted: best

Ben Ho 12/5/2016 2:29 PM
Deleted: which

Ben Ho 12/5/2016 2:29 PM
Deleted: to use

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;

Ben Ho 12/5/2016 2:46 PM
Deleted: Siedel

Ben Ho 12/5/2016 2:46 PM
Deleted: been using

Ben Ho 12/5/2016 2:46 PM
Deleted: Siedel

Ben Ho 12/5/2016 2:47 PM
Deleted: also

Ben Ho 12/5/2016 3:12 PM
Deleted: measurement

Ben Ho 12/5/2016 3:12 PM
Deleted: beside

Ben Ho 12/5/2016 3:13 PM
Deleted: to some sensor types

Ben Ho 12/5/2016 3:13 PM
Deleted: very

Ben Ho 12/5/2016 3:13 PM
Deleted: type

Ben Ho 12/5/2016 3:14 PM
Deleted: global

Ben Ho 12/5/2016 3:14 PM
Deleted:

Ben Ho 12/5/2016 3:14 PM
Deleted: ing

Ben Ho 12/5/2016 3:14 PM
Deleted: measurement

Ben Ho 12/5/2016 3:14 PM
Deleted: were p

Ben Ho 12/5/2016 3:15 PM
Deleted: sonde

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)

Ben Ho 12/5/2016 3:15 PM
Deleted: i

Ben Ho 12/5/2016 3:15 PM
Deleted: temperature

Ben Ho 12/5/2016 3:15 PM
Deleted: ing

Ben Ho 12/5/2016 3:16 PM
Deleted: those

Ben Ho 12/5/2016 3:17 PM
Deleted: used

Ben Ho 12/5/2016 3:18 PM
Deleted: to identify temperature anomalies between observations and backgrounds, which

Ben Ho 12/5/2016 3:19 PM
Deleted: are

Ben Ho 12/5/2016 3:19 PM
Deleted: anomalies

Ben Ho 12/5/2016 3:19 PM
Deleted: anomaly

Ben Ho 12/5/2016 3:20 PM
Deleted: critically

Ben Ho 12/5/2016 3:20 PM
Deleted: Unlike RAOBs, t

Ben Ho 12/8/2016 1:46 PM
Formatted: Indent: First line: 0.5", Line spacing: double

Ben Ho 12/5/2016 3:20 PM
Deleted: i.e.,

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

Ben Ho 12/5/2016 3:26 PM
Deleted: The

Ben Ho 12/5/2016 3:26 PM
Deleted:

Ben Ho 12/5/2016 3:27 PM
Deleted: GPS

Ben Ho 12/8/2016 1:48 PM
Deleted: FORMOSAT-3/Constellation Observing System for Meteorology, Ionosphere, and Climate (

Ben Ho 12/8/2016 1:48 PM
Deleted:)

Ben Ho 12/5/2016 3:28 PM
Deleted: from April 2008 to August 2009 and from May 2008 to August 2011, respectively,

Ben Ho 12/5/2016 3:28 PM
Deleted: the

Ben Ho 12/5/2016 3:29 PM
Deleted: a

Ben Ho 12/5/2016 3:29 PM
Deleted: mode

Ben Ho 12/5/2016 3:29 PM
Deleted: s (not shown)

Ben Ho 12/5/2016 3:29 PM
Deleted: the

Ben Ho 12/5/2016 3:30 PM
Deleted: it ma

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

Ben Ho 12/5/2016 3:37 PM
Deleted: in

Ben Ho 12/8/2016 1:50 PM
Deleted: Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) Data Analysis and Archive Center (

Ben Ho 12/8/2016 1:50 PM
Deleted:)

Ben Ho 12/8/2016 1:53 PM
Deleted: Meteorological Operational Polar Satellite–A (

Ben Ho 12/8/2016 1:53 PM
Deleted:)

Ben Ho 12/8/2016 1:53 PM
Deleted: (GNSS Receiver for Atmospheric Sounding

Ben Ho 12/5/2016 3:38 PM
Deleted: Brief description of the improved reprocessing package and the general quality of GPS RO data for climate as benchmark references for climate studies are described in Appendix A.

Ben Ho 12/5/2016 3:38 PM
Formatted: Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Ben Ho 12/5/2016 3:39 PM
Deleted: analysis

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

Ben Ho 12/5/2016 3:40 PM
Deleted: C

Ben Ho 12/5/2016 3:40 PM
Deleted: ing

Ben Ho 12/5/2016 3:40 PM
Deleted: (i.e., COSMIC2013 covering from June 2006 to April 2014)

Ben Ho 12/5/2016 3:41 PM
Deleted: (i.e., Metop-A2016 covering from to September 2007 to December 2015)

Ben Ho 12/5/2016 3:41 PM
Deleted: Because COSMIC contains dominate sample numbers (six receivers) than those of Metop-A (one receiver), we limited this study from June 2006 to April 2014 (see Section 2.2).

Ben Ho 12/5/2016 3:41 PM
Deleted: in the lower stratosphere

Ben Ho 12/5/2016 3:42 PM
Deleted: The trend uncertainties for specific RAOB senor types are also specified. Systematic RAOB temperature biases and their uncertainties relative to RO data are documented for different sensor types, which in turn is useful to quantify uncertainties of temperature CDRs constructed from different RAOB sensor types.

Ben Ho 12/5/2016 3:43 PM
Deleted: We perform the trend analysis for temperature biases among sensors in Section 5.

Ben Ho 12/5/2016 3:43 PM
Deleted: 6

Ben Ho 12/5/2016 3:50 PM
Deleted: (sonPrf) from June 2006 to April 2014

Ben Ho 12/5/2016 3:50 PM
Deleted: are

Ben Ho 12/5/2016 3:50 PM
Deleted: sonPrf

Ben Ho 12/5/2016 3:50 PM
Deleted: over both lands and islands

Ben Ho 12/5/2016 3:51 PM
Deleted: , which are used in this study. T

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 (α) and sensor infrared emissivity (ε) 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

Ben Ho 12/5/2016 3:51 PM
Deleted: These radiosonde data are transmitted through the Global Telecommunication System (GTS).

Ben Ho 12/5/2016 3:52 PM
Deleted: during the comparison time period

Ben Ho 12/5/2016 3:53 PM
Deleted: also

Ben Ho 12/5/2016 3:53 PM
Deleted: , available from June 2006 to April 2014)

Ben Ho 12/5/2016 3:54 PM
Deleted: , available from April 2008 to Dec. 2015)

Ben Ho 12/5/2016 3:54 PM
Deleted: GPS

Ben Ho 12/5/2016 3:54 PM
Deleted: To maintain similar sample numbers in each month, we use RO data from June 2006 to April 2014 in this study.

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

Ben Ho 12/5/2016 3:55 PM
Deleted: are

Ben Ho 12/5/2016 3:55 PM
Deleted: hPa

Ben Ho 12/5/2016 3:56 PM
Deleted: hPa

Ben Ho 12/5/2016 3:56 PM
Deleted: hPa

Ben Ho 12/5/2016 3:56 PM
Deleted: hPa

Ben Ho 12/5/2016 3:56 PM
Deleted: atmPrf

Ben Ho 12/5/2016 3:56 PM
Deleted: with those

Ben Ho 12/5/2016 3:56 PM
Deleted: are

Ben Ho 12/5/2016 3:56 PM
Deleted: , which are similar to

Ben Ho 12/5/2016 3:57 PM
Deleted: the

Ben Ho 12/5/2016 3:57 PM
Deleted: in

Ben Ho 12/5/2016 3:58 PM
Deleted: Different

Ben Ho 12/5/2016 3:58 PM
Deleted: from

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

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

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

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

Ben Ho 12/5/2016 3:59 PM
Deleted: further

Ben Ho 12/5/2016 3:59 PM
Deleted:

Ben Ho 12/5/2016 3:59 PM
Deleted: . The equation we used to compute the monthly mean temperature biases between RAOB and RO data at the mandatory height is

Ben Ho 12/8/2016 2:15 PM
Deleted: December

Ben Ho 12/8/2016 2:16 PM
Deleted: 5

Ben Ho 12/8/2016 11:31 AM
Deleted: England

Ben Ho 12/8/2016 2:17 PM
Deleted: and

Ben Ho 12/8/2016 2:17 PM
Deleted:

Ben Ho 12/5/2016 4:00 PM
Deleted: ($Std(\Delta T^{Time})$)

Ben Ho 12/5/2016 4:00 PM
Deleted: tion

Ben Ho 12/5/2016 4:01 PM
Deleted:

Ben Ho 12/5/2016 4:01 PM
Deleted: we define

Ben Ho 12/5/2016 4:01 PM
Deleted: degree

Ben Ho 12/8/2016 2:18 PM
Deleted: a

Ben Ho 12/5/2016 4:01 PM
Deleted: degree

Ben Ho 1/22/2017 9:51 AM
Formatted: Font:12 pt

Ben Ho 12/5/2016 4:02 PM
Deleted: is

Ben Ho 1/12/2017 12:49 PM
Deleted: Introduced since 2003,

Ben Ho 1/12/2017 12:50 PM
Formatted ... [1]

Ben Ho 12/5/2016 4:03 PM
Deleted: from June 2006 to April 2014 was

Ben Ho 12/5/2016 4:04 PM
Deleted: Over oceans, Vaisala RS92 ... [2]

Ben Ho 12/5/2016 4:04 PM
Deleted: to the current,

(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

Ben Ho 1/17/2017 2:02 PM
Formatted: Font:12 pt

Ben Ho 1/17/2017 2:02 PM
Formatted: Font:12 pt, No underline

Ben Ho 1/12/2017 12:50 PM
Formatted: Font:12 pt

Ben Ho 12/5/2016 4:06 PM
Deleted: that

Ben Ho 12/5/2016 4:07 PM
Deleted: of the temperature difference ($Std(\Delta T)$) for the layer between 200 hPa and 20 hPa

Ben Ho 12/5/2016 4:08 PM
Deleted: individual

Ben Ho 12/5/2016 4:08 PM
Deleted: s

Ben Ho 1/22/2017 9:48 AM
Formatted: Line spacing: double

Ben Ho 12/5/2016 4:08 PM
Deleted: calculated.
<#>Although we only include those stations stations containing more than 50 RO-RAOB pairs, some levels of heterogeneity (i.e., Fig. 2a over Brazil) may be, in part, due to lower sampling numbers. For example, those stations with temperature biases larger than 0.5 K in the east Brazil contain only about 60 RO-RAOB pairs. The causes of the temperature biases heterogeneity over North and South China are not certain in this point. ... [3]

Ben Ho 12/5/2016 4:09 PM
Deleted: different

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 ΔT 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 ΔT 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).

Ben Ho 12/5/2016 4:10 PM
Deleted: s

Ben Ho 12/5/2016 4:10 PM
Deleted: a and b, respectively

Ben Ho 12/5/2016 4:10 PM
Deleted: depicts

Ben Ho 12/5/2016 4:11 PM
Deleted: that there exists obvious different mean ΔT

Ben Ho 1/22/2017 9:48 AM
Formatted: Font:12 pt

Ben Ho 1/22/2017 9:48 AM
Deleted: <#>Although we only include those stations containing more than 50 RO-RAOB pairs, some levels of heterogeneity (i.e., Fig. 2a over Brazil) may be, in part, due to lower sampling numbers. For example, those stations with temperature biases larger than 0.5 K in the east Brazil contain only about 60 RO-RAOB pairs. The causes of the temperature biases heterogeneity over North and South China are not certain in this point. ... [4]

Ben Ho 12/5/2016 4:13 PM
Deleted: ΔT is

Ben Ho 12/5/2016 4:14 PM
Deleted: ΔT

Ben Ho 12/5/2016 4:14 PM
Deleted: is

Ben Ho 12/5/2016 4:14 PM
Deleted: at night

Ben Ho 12/5/2016 4:15 PM
Deleted: *Std(ΔT)*

Ben Ho 12/5/2016 4:15 PM
Deleted: 5

Ben Ho 12/5/2016 4:15 PM
Deleted: 68

Ben Ho 12/5/2016 4:15 PM
Deleted: see

Ben Ho 12/8/2016 2:29 PM
Deleted: b

Ben Ho 12/5/2016 4:17 PM
Deleted: 5

Ben Ho 12/5/2016 4:17 PM
Deleted: The Stds of ΔT for daytime and nighttime combined are 1.52 K for Vaisala RS92, 1.58 K for AVK, 1.67 K for VIZ-B2, 1.59 K for Sippican, 1.68 K for Shanghai, and 1.69 K for Meisei.

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

Ben Ho 12/5/2016 4:18 PM
Deleted: <#>**a. Regional Comparison Results**

Ben Ho 12/5/2016 4:18 PM
Deleted: and over different regions,

Ben Ho 12/5/2016 4:18 PM
Deleted: further

Ben Ho 12/5/2016 4:19 PM
Deleted: Figure 3a is for United States, and Figs. 3b-f are for the Australia, Germany, Canada, England, and Brazil, respectively.

Ben Ho 12/5/2016 4:21 PM
Formatted: Indent: First line: 0.56", No bullets or numbering

Ben Ho 12/5/2016 4:19 PM
Deleted: a

Ben Ho 12/5/2016 4:19 PM
Deleted: depicts

Ben Ho 12/5/2016 4:20 PM
Deleted: when comparing to RO data.

Ben Ho 12/5/2016 4:21 PM
Deleted: less

Ben Ho 12/5/2016 4:21 PM
Deleted: ing th

Ben Ho 12/5/2016 4:21 PM
Deleted: o

Ben Ho 12/5/2016 4:21 PM
Deleted: The *Std(ΔT)* (RS92) from 200 hPa to 20 hPa for different countries are United States (1.59 K), Australia (1.48 K), Ger ... [5]

Ben Ho 12/5/2016 4:22 PM
Deleted: ,

Ben Ho 12/5/2016 4:22 PM
Deleted: where Figure 4a is for Unite ... [6]

Ben Ho 12/5/2016 4:22 PM
Deleted: a

Ben Ho 12/5/2016 4:23 PM
Deleted: scatter

Ben Ho 12/5/2016 4:23 PM
Deleted: The *Stds(ΔT)* (RS92) are 1. ... [7]

Ben Ho 12/5/2016 4:24 PM
Deleted: (RAOB minus RO) are 0.1 K ... [8]

Ben Ho 12/5/2016 4:24 PM
Deleted: from

Ben Ho 12/5/2016 4:25 PM
Deleted: except for those in United S ... [9]

Ben Ho 12/5/2016 4:26 PM
Deleted: The mean nighttime *ΔT* (R ... [10]

Ben Ho 12/5/2016 4:27 PM
Deleted: in the lower stratosphere,

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).

Ben Ho 12/5/2016 4:28 PM
Deleted: al

Ben Ho 12/5/2016 4:28 PM
Deleted: Because all sondes are launched same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section. ... [11]

Ben Ho 12/5/2016 4:33 PM
Deleted: Solar Zenith Angle Dependent

Ben Ho 12/5/2016 4:35 PM
Deleted: s… 6a-d …depicts the mea ... [12]

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

Ben Ho 12/5/2016 4:38 PM
Deleted: This may owe, in part, to the sampling differences between He2009 (from 2006 August to 2007 Feb., 7 months) and this study (95 months).

Ben Ho 12/5/2016 4:38 PM
Deleted: s

Ben Ho 12/5/2016 4:38 PM
Deleted: a-d

Ben Ho 12/5/2016 4:38 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:39 PM
Deleted:)

Ben Ho 12/5/2016 4:39 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:39 PM
Deleted:)

Ben Ho 12/5/2016 4:39 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:39 PM
Deleted:)

Ben Ho 12/5/2016 4:39 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:39 PM
Deleted:), respectively.

Ben Ho 12/5/2016 4:39 PM
Deleted: The mean ΔT is -0.10 K for Sippican, -0.42 K for VIZ-B2, -0.06 K for AVK, and -0.07 K for Shanghai.

Ben Ho 12/5/2016 4:40 PM
Deleted: Their *Stds*(ΔT) are 1.62 K for Sippican, 1.60 K for VIZ-B2, 1.56 K for AVK, and 1.68 K for Shanghai, respectively. ... [13]

Ben Ho 12/5/2016 4:40 PM
Deleted: computed ΔT

Ben Ho 12/5/2016 4:41 PM
Deleted: for different sensor types

Ben Ho 12/5/2016 4:42 PM
Deleted: s

Ben Ho 12/5/2016 4:42 PM
Deleted: a-d

Ben Ho 12/5/2016 4:42 PM
Deleted: ΔT at 50 hPa varying for SZA ranging from 0 degrees to 180 degrees for Sippican MARK, VIZ-B2, AVK-MRZ, and Shanghai, respectively.

Ben Ho 12/5/2016 4:42 PM
Deleted: <#> The SZA for Sippican, VIZ-B2, Russian, and Shanghai sonde ... [14]

Ben Ho 12/5/2016 4:43 PM
Deleted: The RAOB-RO monthly m ... [15]

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 ΔT 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 ΔT over Germany for these two time periods than over other countries. This also indicates the importance of establishing traceability through careful documentation and metadata

Ben Ho 12/5/2016 4:44 PM
Deleted: hPa

Ben Ho 12/5/2016 4:44 PM
Deleted: stronger

Ben Ho 12/5/2016 4:44 PM
Deleted: a

Ben Ho 12/5/2016 4:45 PM
Deleted: Vaisala radiation correct

Ben Ho 12/5/2016 4:45 PM
Deleted: (i.e., ΔT ($RS92_{200701-201012}$)

Ben Ho 12/5/2016 4:46 PM
Deleted: (i.e., ΔT ($RS92_{201101-201404}$)

Ben Ho 12/8/2016 11:32 AM
Deleted: England

Ben Ho 12/5/2016 4:47 PM
Deleted: RO temperature is used as references for these two periods. Results show that there is no obvious ΔT change between these two periods for the RS92 sondes over United States and Germany (the mean daytime temperature difference between these two periods are about -0.05 K and -0.01 K in 50 hPa for United States and Germany, respectively, see Figs. 9a and c). However, the daytime temperature difference between ΔT ($RS92_{200701-201012}$) and ΔT ($RS92_{201101-201404}$) over Australia, Canada, England, and Brazil show obvious close to 0.1 K to 0.15 K difference varying at different heights (see Figures 9b, c, e, and f, respectively). Note that over Australia, the temperature difference between these two periods at 20 hPa is also as large as -0.2 K, which may also be resulted in the incomplete radiation correction. The incomplete radiation correction likely leads to small but not negligible anomaly in the time series. In this case, the trend anomalies in Australia, Canada, England, and Brazil at 50 hPa is larger than those over the United States and over Germany (see Section 5.2).

Ben Ho 12/5/2016 4:50 PM
Deleted: not in 2011, but

Ben Ho 12/5/2016 4:49 PM
Deleted: better

Ben Ho 12/5/2016 4:49 PM
Deleted: ($RS92_{200701-201012}$) and ΔT ($RS92_{201101-201404}$)

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

Ben Ho 12/5/2016 4:51 PM
Deleted: crucial

Ben Ho 12/5/2016 4:51 PM
Deleted: used

Ben Ho 12/8/2016 3:02 PM
Deleted: , since the US National Weather Service (NWS) did not use Vaisala RS92 radiosondes before 2012.

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

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

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 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 $\overline{T^{Time}}(l,m,k^{'})$ may vary for different sensor types.

Fig. 12 indicates the de-seasonalized trends in daytime temperature differences are within ±0.26 (K/ 5yrs, 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 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,

Ben Ho 12/5/2016 4:53 PM
Deleted: **4.2 Time Series Anomaly for RS92** ... [16]

Ben Ho 1/17/2017 10:47 AM
Formatted ... [17]

Ben Ho 12/5/2016 4:52 PM
Deleted: 6...

Ben Ho 12/5/2016 4:55 PM
Deleted: with ID=37, 52, 61, and 67

Ben Ho 12/5/2016 4:55 PM
Deleted: with ID=79, 80, 81

Ben Ho 12/5/2016 4:56 PM
Deleted: obvious

Ben Ho 12/5/2016 4:56 PM
Deleted: is

Ben Ho 12/5/2016 4:56 PM
Deleted: in the lower stratosphere

Ben Ho 12/5/2016 4:56 PM
Deleted: shall

Ben Ho 12/5/2016 4:57 PM
Deleted: in the lower stratosphere co ... [18]

Ben Ho 12/5/2016 4:57 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:57 PM
Deleted:)

Ben Ho 12/5/2016 4:58 PM
Deleted: .

Ben Ho 12/5/2016 4:58 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:58 PM
Deleted:)

Ben Ho 12/5/2016 4:59 PM
Deleted: mainly distributed over Russian

Ben Ho 12/5/2016 4:59 PM
Deleted: .

Ben Ho 12/5/2016 4:59 PM
Deleted: contain

Ben Ho 12/5/2016 4:59 PM
Deleted: in the

Ben Ho 12/5/2016 4:59 PM
Deleted: mean ΔT (

Ben Ho 12/5/2016 4:59 PM
Deleted:)

Ben Ho 12/5/2016 4:59 PM
Deleted: ΔT (

Ben Ho 12/5/2016 4:59 PM
Deleted:)

Ben Ho 12/5/2016 5:00 PM
Deleted: in the night time

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

Ben Ho 12/5/2016 5:01 PM
Deleted:

Ben Ho 12/5/2016 5:01 PM
Deleted: (i.e., ΔT (RS92)

Ben Ho 12/5/2016 5:01 PM
Deleted: in

Ben Ho 12/5/2016 5:02 PM
Deleted: close to

Ben Ho 12/5/2016 5:02 PM
Deleted: In the daytime, the mean ΔT (Sippican), ΔT (VIZ-B2), ΔT (AVK-MRZ), and ΔT (Shanghai) are (-0.06, 0.71, 0.66, 0.18) K.

Ben Ho 12/5/2016 5:02 PM
Deleted: ing

Ben Ho 12/5/2016 5:02 PM
Deleted: warmer

Ben Ho 12/5/2016 5:03 PM
Deleted: (i.e., RSN2010)

Ben Ho 12/5/2016 5:03 PM
Deleted: ΔT (

Ben Ho 12/5/2016 5:03 PM
Deleted:)

Ben Ho 12/5/2016 5:03 PM
Deleted: (i.e., ΔT ($RS92_{200701-201012}$)

Ben Ho 12/5/2016 5:04 PM
Deleted: ΔT (RS92)

Ben Ho 12/5/2016 5:04 PM
Deleted: for the RS92 sondes

Ben Ho 12/5/2016 5:04 PM
Deleted: in

Ben Ho 12/5/2016 5:04 PM
Deleted: between ΔT ($RS92_{200701-201012}$) and ΔT ($RS92_{201101-201404}$)

Ben Ho 12/8/2016 11:33 AM
Deleted: England,

Ben Ho 12/5/2016 5:04 PM
Deleted: obvious clo

Ben Ho 12/5/2016 5:06 PM
Deleted: difference

Ben Ho 12/5/2016 5:06 PM
Deleted: very

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

Ben Ho 12/5/2016 5:07 PM
Deleted: not in 2011, but

Ben Ho 12/5/2016 5:08 PM
Deleted: anomalies

Ben Ho 12/5/2016 5:08 PM
Deleted:

Ben Ho 12/5/2016 5:09 PM
Deleted: individual

Ben Ho 12/5/2016 5:09 PM
Deleted: s

Ben Ho 12/5/2016 5:09 PM
Deleted: ΔT^{Time} (

Ben Ho 12/5/2016 5:09 PM
Deleted:)

Ben Ho 12/5/2016 5:10 PM
Deleted: persistent

Ben Ho 12/5/2016 5:10 PM
Deleted: difference during the day

Ben Ho 12/5/2016 5:10 PM
Deleted: time and nighttime

Ben Ho 12/5/2016 5:10 PM
Deleted: individual r

Ben Ho 12/5/2016 5:10 PM
Deleted: s

Ben Ho 12/5/2016 5:11 PM
Deleted: While $\Delta^{Time}_{Sippican}$ and $\Delta^{Time}_{Shanghai}$ are largely constant in time, Δ^{Time}_{VIZ-B2} has obvious seasonal variations with a negative trend and night and positive trend during daytime. Δ^{Time}_{AVK} has an irregular seasonal variation particularly during daytime, with a large warm bias.

Ben Ho 12/5/2016 5:11 PM
Deleted: ΔT^{Time} at

Ben Ho 12/5/2016 5:11 PM
Deleted: highly

Ben Ho 12/5/2016 5:12 PM
Deleted: where the

Ben Ho 12/5/2016 5:12 PM
Deleted: USA

Ben Ho 12/5/2016 5:12 PM
Deleted: ΔT^{Time}

Ben Ho 12/5/2016 5:13 PM
Deleted: The daytime trend of anomalies for RAOB and RO temperature at 50 hPa for United States and Germany are equal to (0.00, -0.02) K/5yrs, the trend of anomalies are equal to (0.18, 0.24, 0.26, 0.12) K/5yrs over Australia, Canada, England, and Brazi ... [19]

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.

Ben Ho 12/5/2016 5:15 PM
Deleted: does

Ben Ho 12/5/2016 5:15 PM
Deleted:

Ben Ho 12/5/2016 5:15 PM
Deleted:

Ben Ho 1/17/2017 11:59 AM
Deleted:

Ben Ho 1/17/2017 11:59 AM
Formatted: Indent: First line: 0"

Ben Ho 12/8/2016 3:03 PM
Deleted: ... [20]

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.

Ho, S.-P., M. Goldberg, Y.-H. Kuo, C.-Z Zou, W. Schreiner (2009b), Calibration of Temperature in the Lower Stratosphere from Microwave Measurements using COSMIC Radio Occultation Data: Preliminary Results, *Terr. Atmos. Oceanic Sci*., Vol. 20, doi: 10.3319/TAO.2007.12.06.01(F3C).

Ho, S.-P., Zhou X., Kuo Y.-H., Hunt D., Wang J.-H., (2010a), Global Evaluation of Radiosonde Water Vapor Systematic Biases using GPS Radio Occultation from COSMIC and ECMWF Analysis. *Remote Sensing*. 2010; 2(5):1320-1330.

Ho, S.-P., Ying-Hwa Kuo, William Schreiner, Xinjia Zhou (2010b), Using SI-traceable Global Positioning System Radio Occultation Measurements for Climate Monitoring [In "States of the Climate in 2009]. *Bul. Amer. Meteor. Sci*., 91 (7), S36-S37.

Ho, S.-P., and co-authors (2012), Reproducibility of GPS Radio Occultation Data for Climate Monitoring: Profile-to-Profile Inter-comparison of CHAMP Climate Records 2002 to 2008 from Six Data Centers, *J. Geophy. Research.* VOL. 117, D18111, doi:10.1029/2012JD017665.

Ho, S.-P., L. Peng, and D. Hunt (2016), Construction of consistent temperature climate data records in the lower Troposphere using COSMIC, Metop-A/GRAS, and Metop-B/GRAS from 2006 to 2015, *Atmos. Meas. Tech.,* submitted.

Kuo, Y. H., T. K. Wee, S. Sokolovskiy, C. Rocken, W. Schreiner, and D. Hunt (2004), Inversion and error estimation of GPS radio occultation data, *J. Meteorol. Soc. Jpn.*, **82**, 507– 531.

Lanzante, J. R., S. A. Klein, and D. J. Seidel (2003), Temporal homogeni- sation of monthly radiosonde temperature data, Part I: Methodology, *J. Clim.,* 16, 224–240, doi:10.1175/1520-0442(2003)016.

Luers, J. K., and R. E. Eskridge (1995), Temperature corrections for the VIZ and Vaisala radiosondes, *J. Appl. Meteor.*, 34, 1241–1253.

Luers, J. K., (1997), Temperature error of the Vaisala RS90 radiosonde, *J. Atmos. Oceanic Technol.*, 14, 1520–1532.

Luers, J. K., and R. E. Eskridge (1998), Use of radiosonde temperature data in climate studies, *J. Clim.,* 11, 1002–1019.

Randel, W. J., A. K. Smith, F. Wu, C.-Z. Zhou, and H. Qian (2016), Stratospheric temperature trends over 1979–2015 derived from combined SSU, MLS, and SABER satellite observations, *J. Clim.*, *29*, 4843–4859, doi:10.1175/JCLI-D-15-0629.1.

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.

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 1/16/2017 10:01 AM
Deleted: December

Ben Ho 1/16/2017 10:01 AM
Deleted: 5

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

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.

Ben Ho 1/16/2017 10:01 AM
Deleted: December

Ben Ho 1/16/2017 10:01 AM
Deleted: 5

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 1/17/2017 2:54 PM
Deleted: OAB

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:33 AM
Deleted: England

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Deleted: England

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Deleted: England

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Deleted: England

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

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.

Ben Ho 1/17/2017 11:36 AM
Deleted: anomaly

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 1/17/2017 11:39 AM
Deleted: .

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 9:24 AM
Deleted: Figure 12. The time series of de-seasonalized temperature anomaly in 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) England, and f) Brazil. The 95% confidence intervals for slopes are shown in the parentheses. ... [21]

Ben Ho 12/8/2016 9:24 AM
Deleted: 3

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 9:24 AM
Deleted: anomaly in

Ben Ho 1/16/2017 2:05 PM
Deleted: .

Ben Ho 12/8/2016 11:34 AM
Formatted: Font:12 pt

Ben Ho 1/16/2017 4:09 PM
Deleted: ... [22]

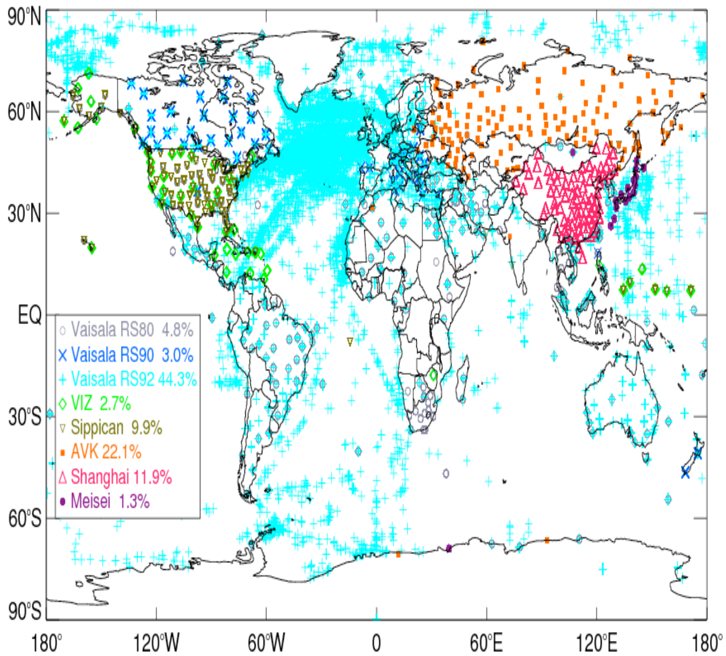


Figure 1. Global distribution of radiosonde stations colored by radiosonde types. Radiosonde types updated from June 2006 to April 2014 are used. The percentage of each type of radiosonde used among all stations is listed. For those stations that radiosonde types are changed during this period, the latest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

Ben Ho 1/16/2017 10:02 AM
Deleted: ... [23]

Ben Ho 1/16/2017 10:02 AM
Deleted: December

Ben Ho 1/16/2017 10:02 AM
Deleted: 5

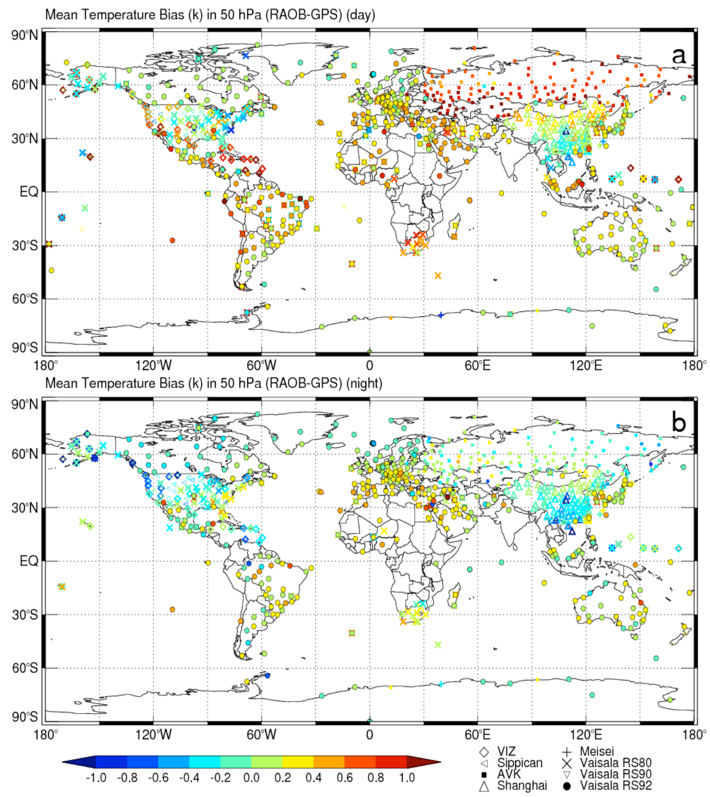


Figure 2. Mean RAOB-RO temperature biases at 50 hPa for the RAOB-RO ensembles from June 2006 to April 2014 for a) daytime, and b) nighttime. Only those stations containing more than 50 RO-ROAB pairs are plotted.

Ben Ho 1/16/2017 10:02 AM
Deleted: December

Ben Ho 1/16/2017 10:02 AM
Deleted: 5

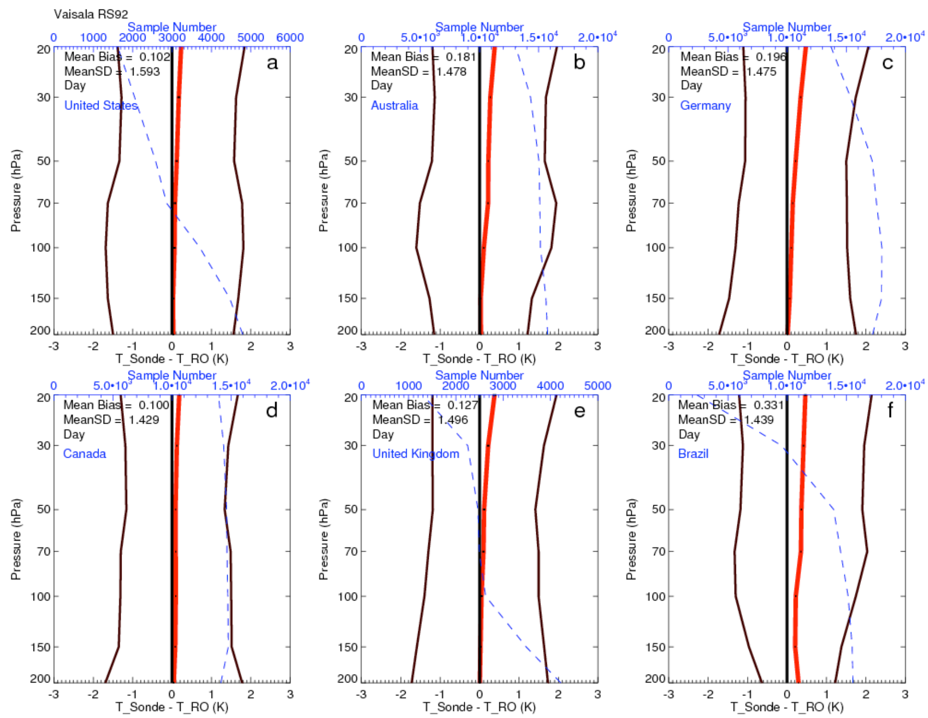
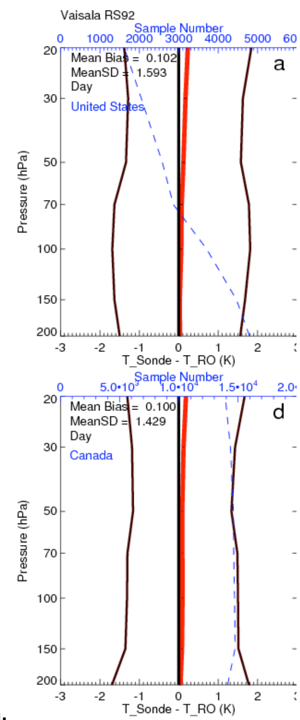


Figure 3. Comparisons of temperature between RS92 and RO for daytime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red line is the mean difference; the black line is the standard deviation of the mean difference; the dotted line is the sample number.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

Ben Ho 12/11/2016 8:10 AM



Deleted:

Ben Ho 12/8/2016 11:35 AM

Deleted: England

Ben Ho 1/22/2017 10:03 AM

Deleted: the horizontal black lines superimposed on the mean are the standard error of the mean;

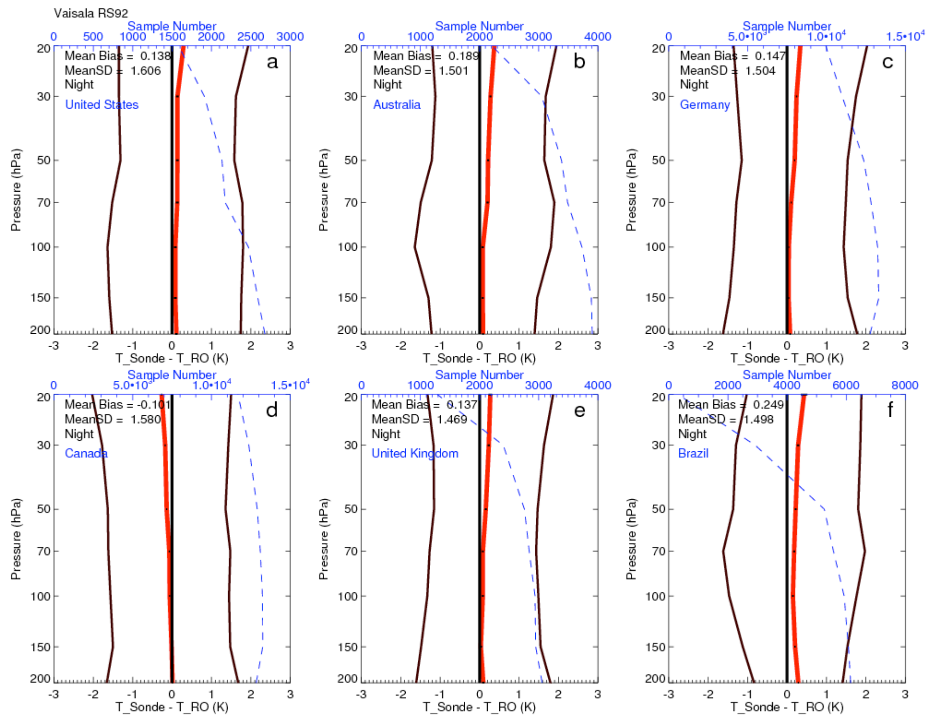
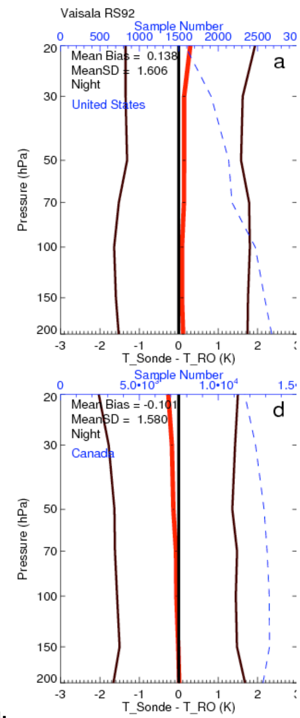


Figure 4. Comparisons of temperature between RS92 and RO for nighttime over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.

Ben Ho 12/11/2016 8:10 AM



Deleted:

Ben Ho 12/8/2016 11:35 AM

Formatted: Font:12 pt

Ben Ho 12/8/2016 11:35 AM

Deleted: England

Ben Ho 12/8/2016 11:35 AM

Formatted: Font:12 pt

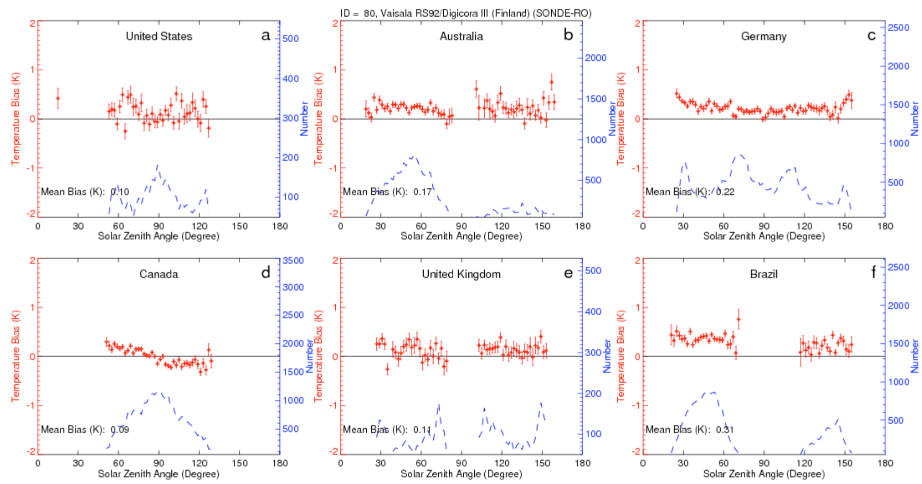
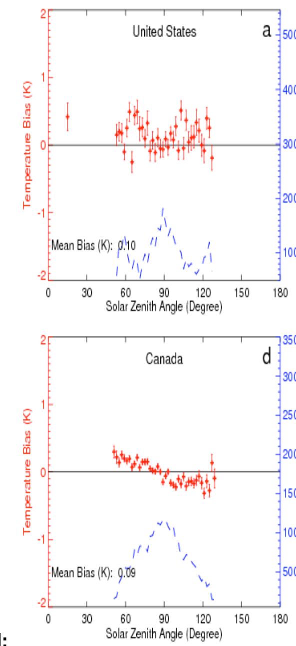


Figure 5. The mean temperature biases (RS92 minus RO) at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for each 5 SZA bins; the red vertical line is the standard deviation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

Ben Ho 12/11/2016 8:10 AM



Deleted:

Ben Ho 12/8/2016 11:36 AM

Formatted: Font:12 pt

Ben Ho 12/8/2016 11:37 AM

Deleted: England

Ben Ho 12/8/2016 11:36 AM

Formatted: Font:12 pt

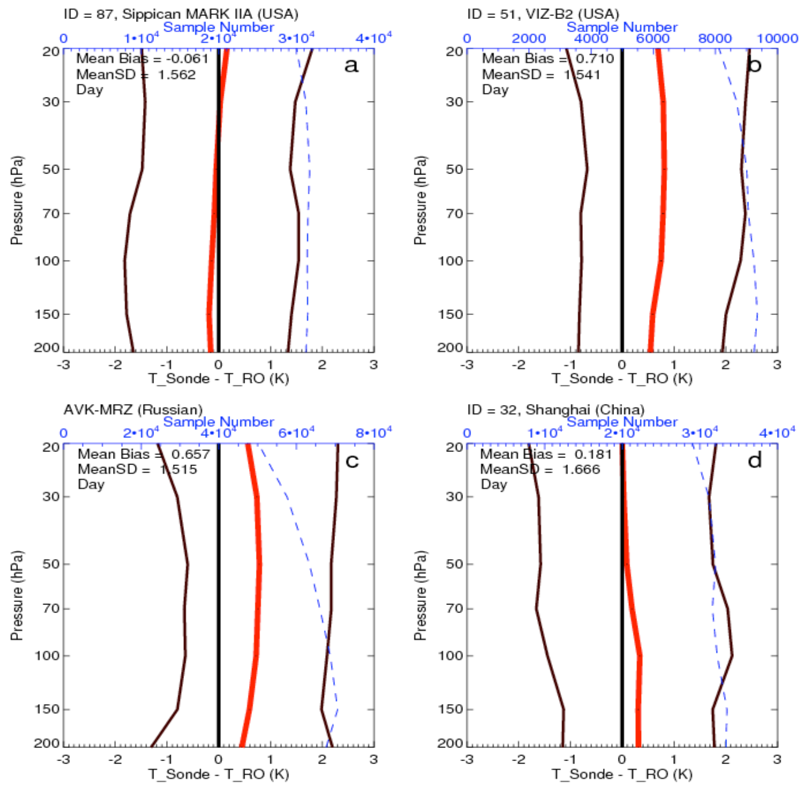


Figure 6. Comparisons of temperature between radiosonde and RO during the daytime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

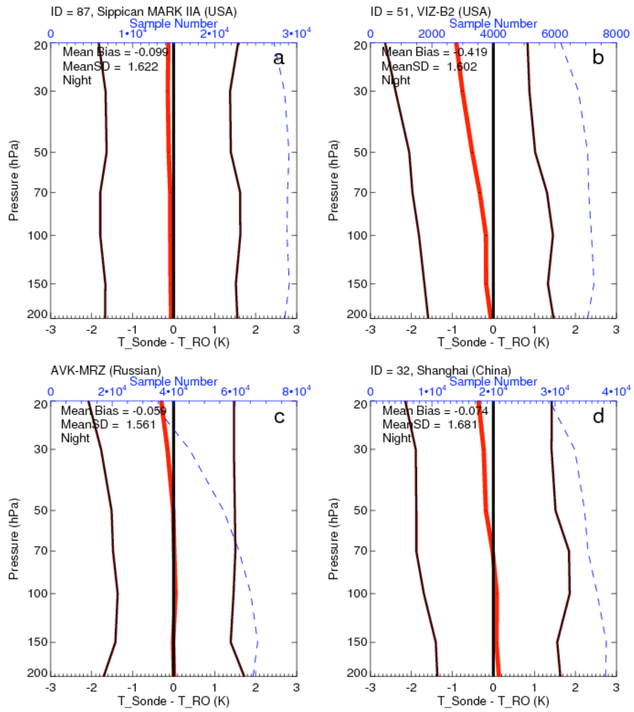


Figure 7. Comparisons of temperature between radiosonde and RO during the nighttime for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO.

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

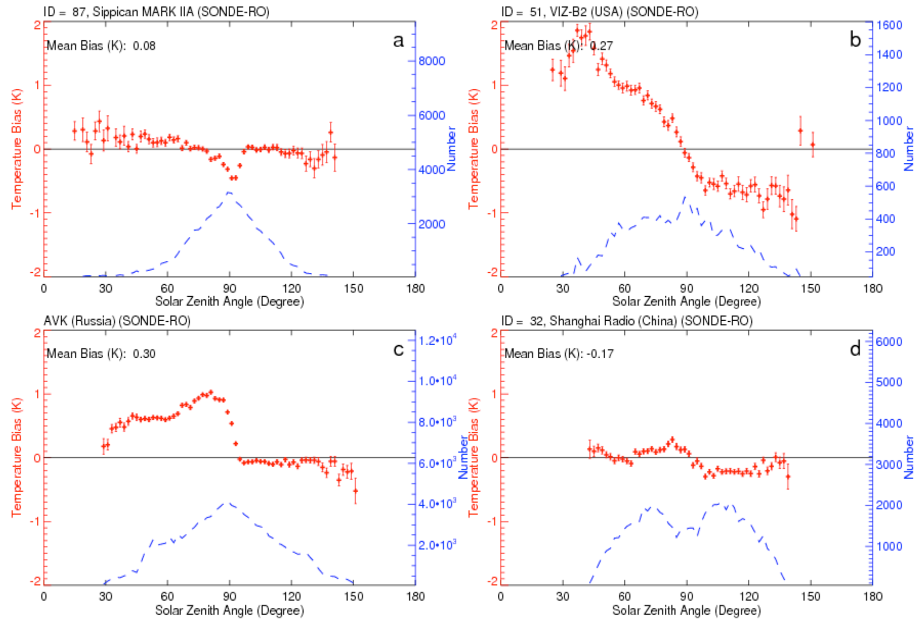


Figure 8. The mean temperature biases at 50 hPa varying for SZA from 0 degrees to 180 degrees for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO. Only bins for more than 50 RAOB-RO pairs are plotted.

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

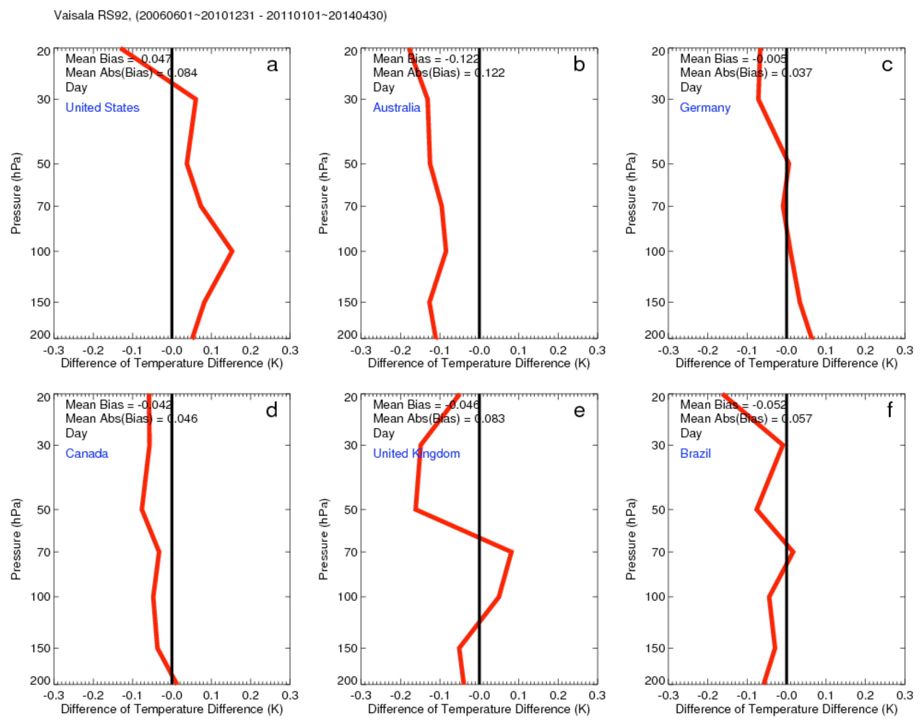
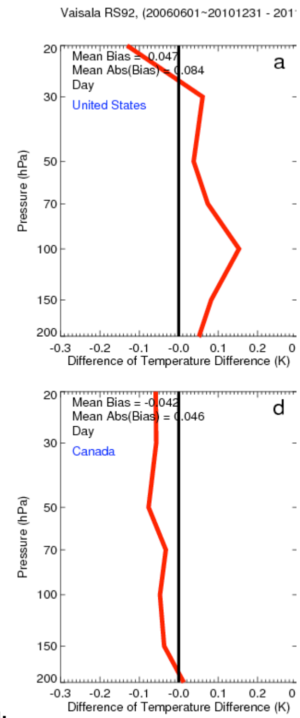


Figure 9. The temperature differences between RS92 – RO from January 2007 to December 2010 (ΔT (RS92$_{200701\text{-}201012}$) and those from January 2011 to December 2015 (ΔT (RS92$_{201101\text{-}201512}$) over a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil.

Ben Ho 12/11/2016 8:10 AM



Deleted:

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:36 AM
Deleted: England

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

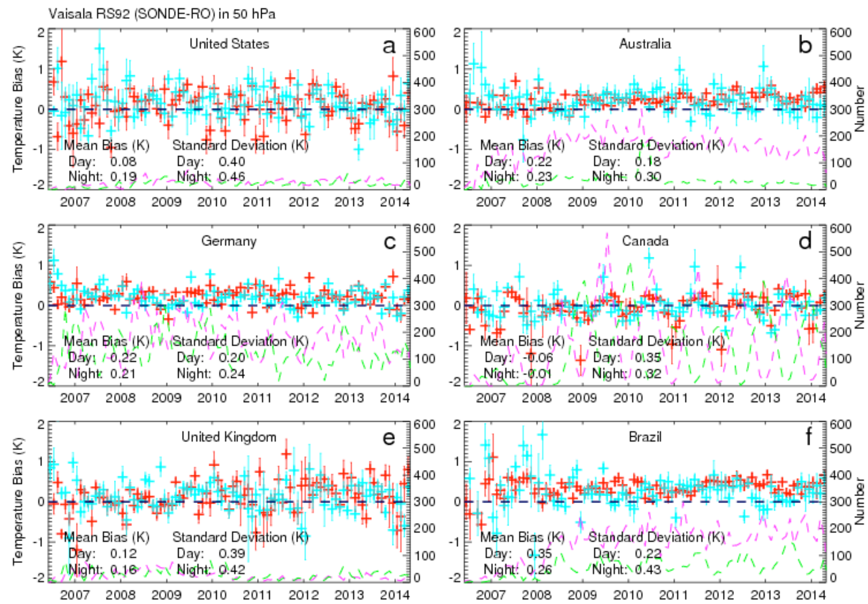
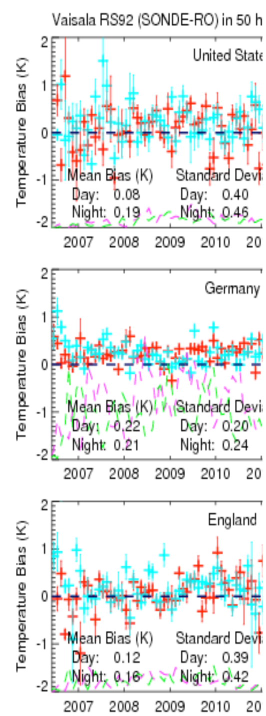


Figure 10. The time series of monthly mean temperature differences to RO at 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) United Kingdom, and f) Brazil. The red cross is the mean difference for RS92 minus RO temperature at 50 hPa during the daytime and the blue cross is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

Ben Ho 1/17/2017 12:50 PM



Deleted:

Ben Ho 1/17/2017 11:39 AM

Formatted: Font:12 pt

Ben Ho 1/17/2017 12:49 PM

Deleted: England

Ben Ho 1/17/2017 11:39 AM

Formatted: Font:12 pt

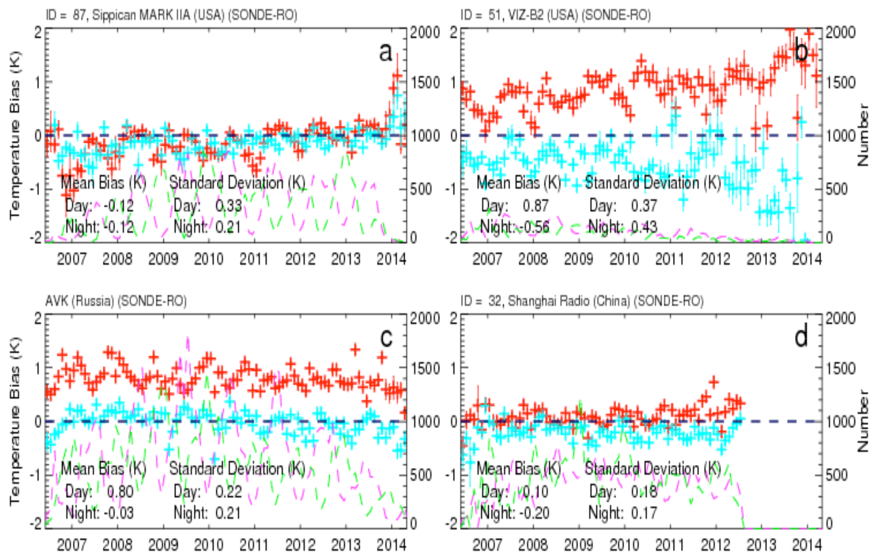


Figure 11. The time series of temperature difference in 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian Sonde minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N).

Ben Ho 1/17/2017 11:39 AM
Formatted: Font:12 pt

Ben Ho 1/17/2017 11:37 AM
Deleted: anomaly

Ben Ho 1/17/2017 11:38 AM
Deleted: .

Ben Ho 1/17/2017 11:39 AM
Formatted: Font:12 pt

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.

Ben Ho 1/17/2017 10:50 AM
Deleted: ... [24]

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 11:39 AM
Formatted: Font:12 pt

Ben Ho 12/8/2016 9:28 AM
Deleted: ... [25]

Ben Ho 12/8/2016 9:28 AM
Deleted: ... [26]

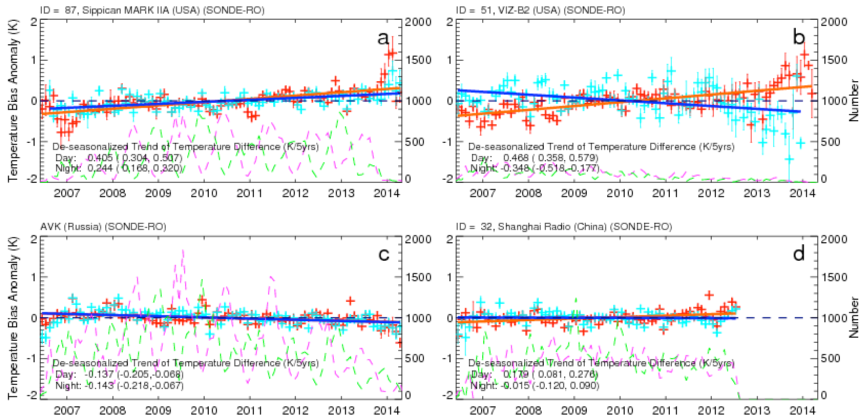
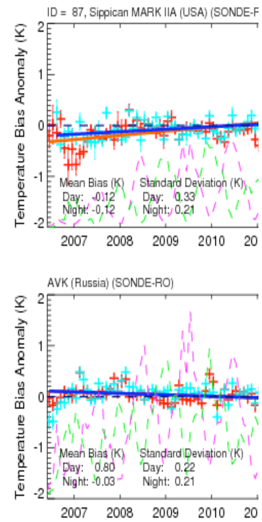


Figure 13. The time series of de-seasonalized temperature differences at 50 hPa for a) Sippican over United States minus RO, b) VIZ-B2 over United States minus RO, c) Russian minus RO, d) Shanghai minus RO in the North hemisphere mid-latitude (60°N-20°N). The 95% confidence intervals for slopes are listed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

Ben Ho 1/17/2017 11:28 AM



Deleted:

Ben Ho 12/8/2016 11:39 AM

Formatted: Font:12 pt

Ben Ho 12/8/2016 11:39 AM

Formatted: Font:12 pt

Ben Ho 12/8/2016 9:28 AM

Deleted: Figure 13. The time series of de-seasonalized temperature anomaly 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. 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

Ben Ho 12/8/2016 11:39 AM
Deleted: ... [27]

Ben Ho 1/17/2017 10:51 AM
Deleted: December

Ben Ho 1/17/2017 10:51 AM
Deleted: 5

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.

Ben Ho 12/8/2016 9:29 AM
Deleted: anomaly

Ben Ho 12/8/2016 11:36 AM
Deleted: England

	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

Ben Ho 1/16/2017 4:27 PM
Deleted: trend (K/5yrs), and
Ben Ho 12/8/2016 9:30 AM
Deleted: anomaly
Ben Ho 12/8/2016 11:40 AM
Formatted: Font:12 pt
Ben Ho 1/17/2017 7:08 AM
Deleted: anomaly
Ben Ho 1/16/2017 4:16 PM
Formatted: Font:Calibri, 10 pt
Ben Ho 1/16/2017 12:23 PM
Deleted: ANOM
Ben Ho 1/16/2017 12:23 PM
Deleted: ANOM
Ben Ho 1/16/2017 3:36 PM
Deleted: ANOM
Ben Ho 1/16/2017 3:36 PM
Deleted: ANOM
Ben Ho 12/8/2016 11:08 AM
Deleted:
Day ... [28]

Page 11: [1] Formatted	Ben Ho	1/12/17 12:50 PM

Indent: First line: 0.5", No bullets or numbering, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Page 11: [2] Deleted	Ben Ho	12/5/16 4:04 PM

Over oceans, Vaisala RS92 launched from ships were also used (see Figure 1).

Page 12: [3] Deleted	Ben Ho	12/5/16 4:08 PM

calculated.

In general, the radiosonde temperature biases vary for different sensor types. The mean *ΔT* for RS92 (0.16 K), RS80 (0.10 K), RS90 (0.13 K), Sippican MarkIIA (-0.08 K), Shangai (0.05 K) and Meisei (0.11 K) are smaller than those for AVK (0.33 K) and VIZ-B2 (0.22 K) (see Table 2).

Page 13: [4] Deleted	Ben Ho	1/22/17 9:48 AM

Although we only include those stations containing more than 50 RO-RAOB pairs, some levels of heterogeneity (i.e., Fig. 2a over Brazil) may be, in part, due to lower sampling numbers. For example, those stations with temperature biases larger than 0.5 K in the east Brazil contain only about 60 RO-RAOB pairs. The causes of the temperature biases heterogeneity over North and South China are not certain in this point.

The mean *ΔT* between 200 hPa and 20 hPa for different sensor types in the daytime is summarized in Table 2. The daytime mean *ΔT* (RS92) is about 0.1 K to 0.3 K globally, and the *ΔT* (AVK) is as large as 0.8 K. The relatively larger *Std(ΔT)* for Shanghai (~1.67 K) may be mainly due to large *ΔT* difference in north and south China under different solar zenith angle especially during the daytime. In the daytime *ΔT* (Shanghai) can be as large as 0.2 K to 0.4 K in east and north China and range from -0.2 K to -0.4 K in the south China.

Page 14: [5] Deleted	Ben Ho	12/5/16 4:21 PM

The *Std(ΔT)* (RS92) from 200 hPa to 20 hPa for different countries are United

States (1.59 K), Australia (1.48 K), Germany (1.48 K), Canada (1.43 K), England (1.5 K), and Brazil (1.44 K).

However, there still exist small but not negligible *ΔT* (RS92) between 200 hPa and 20 hPa in different regions. The mean *ΔT* (RS92) in the United States is close to zero near the 200 hPa then slightly increases with height. The mean *ΔT* (RS92) in United States from 200 hPa to 20 hPa is equal to 0.10 K. The mean *ΔT* (RS92) are 0.18 K for Australia, 0.20 K for Germany, 0.10 K for Canada, 0.13 K for England, and 0.33 K for Brazil (Figs. 3b-e).

Page 14: [6] Deleted	Ben Ho	12/5/16 4:22 PM

where Figure 4a is for United States, and Figs. 4b-f are for the Australia, Germany, Canada, England, and Brazil, respectively.

Page 14: [7] Deleted	Ben Ho	12/5/16 4:23 PM

The *Stds(ΔT)* (RS92) are 1.61 K for United States, 1.50 K for Australia, 1.50 K for Germany, 1.58 K for Canada, 1.47 K for England, and 1.50 K for Brazil.

Page 14: [8] Deleted	Ben Ho	12/5/16 4:24 PM

(RAOB minus RO) are 0.1 K to 0.2 K smaller (colder) than

Page 14: [9] Deleted	Ben Ho	12/5/16 4:25 PM

except for those in United States and England. The nighttime mean temperature bias for USA is 0.14 K whereas the daytime mean bias is 0.10 K. The nighttime mean temperature bias for England 0.14 K whereas the daytime mean bias is 0.13 K.

Page 14: [10] Deleted	Ben Ho	12/5/16 4:26 PM

The mean nighttime *ΔT* (RS92) are 0.14 K for United States, 0.19 K for Australia, 0.15 K for Germany, -0.10 K for Canada, 0.14 K for England, and 0.25 K for

Brazil, respectively (Figs. 4a-e).

The small but not negligible mean Vaisala RS92 temperature biases in different regions may indicate a small residual error after applying the radiation correction tables (i.e., RSN96, RSN2005, and RSN2010 tables) for the respective sonde type.

Page 15: [11] Deleted **Ben Ho** **12/5/16 4:28 PM**

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted **Ben Ho** **12/5/16 4:28 PM**

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted **Ben Ho** **12/5/16 4:28 PM**

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted **Ben Ho** **12/5/16 4:28 PM**

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted	Ben Ho	12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted	Ben Ho	12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted	Ben Ho	12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted	Ben Ho	12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted	Ben Ho	12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted Ben Ho 12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted Ben Ho 12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [11] Deleted Ben Ho 12/5/16 4:28 PM

Because all sondes are launched close to the same UTC time, RS92 in different regions are launched at different local times, i.e. different SZA. The analyses for the SZA dependent temperature biases are further discussed in next section.

b. Solar Zenith Angle Dependent Temperature Biases

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 15: [12] Deleted Ben Ho 12/5/16 4:35 PM

s

Page 16: [13] Deleted Ben Ho 12/5/16 4:40 PM

Their *Stds*(ΔT) are 1.62 K for Sippican, 1.60 K for VIZ-B2, 1.56 K for AVK, and 1.68 K for Shanghai, respectively.

Similar to Vaisala, we

Page 16: [14] Deleted Ben Ho 12/5/16 4:42 PM

The SZA for Sippican, VIZ-B2, Russian, and Shanghai sondes is mainly ranging between 30 degree and 150 degree. The VIZ-B2 has an obvious warm bias during daytime and a cold bias at night relative to those of RO temperature profiles (Figure 8b). At 50 hPa, the VIZ-B2 warm bias can be as large as 1.75 K near the noon, and it decreases to -0.8 K during the night. AVK has a temperature bias of from about 0.7 K to 1.1 K in the daytime where its nighttime biases are close to zero (Figure 8c). The mean temperature biases for the Shanghai sondes is about 0.16 K and -0.07 K for daytime and nighttime (Figure 6d), respectively.

Page 16: [15] Deleted	Ben Ho	12/5/16 4:43 PM

The RAOB-RO monthly mean temperature biases in the lower stratosphere at different geographical regions are highly dependent on the seasonal variation of the SZA.

Page 22: [16] Deleted	Ben Ho	12/5/16 4:53 PM

4.2 Time Series Anomaly for RS92

SZA-dependent biases may result in seasonally and regionally dependent temperature biases for different sensor types, which may result in unexpected trend uncertainty. With a residual RS92 radiation error identified in the Section 3.2, the time series of the RS92-RO temperature bias behave slightly different for different regions. Figures 10a-f show daytime and nighttime time series of monthly mean temperature biases computed using Eq. (3) at 50 hPa for ΔT^{Time} (RS92) at United States (ΔT^{Time} ($RS92_{USA}$)), Australia (ΔT^{Time} ($RS92_{Australia}$)), German (ΔT^{Time} ($RS92_{German}$)), Canada (ΔT^{Time} ($RS92_{Canada}$)), England (ΔT^{Time} ($RS92_{England}$)), and Brazil (ΔT^{Time} ($RS92_{Brazil}$)), respectively. The number for the monthly RAOB-RO pairs for daytime is in pink dash line and that for the nighttime is in green dash line. The vertical lines superimposed on the monthly mean are the standard error of the mean.

Figures 10a-f indicate that the time series of ΔT^{Time} (RS92) at all regions are largely constant in time with a small difference during the daytime and nighttime in each individual regions. The consistency of RAOB and RO time series data is best represented by their standard deviation. The *Stds*(ΔT^{Time}) are 0.4 K for United States, 0.18 K for Australia, 0.20 K for Germany, 0.35 K for Canada, 0.39 K for England, and 0.22 K for Brazil, respectively. The relatively larger *Std*(ΔT^{Time}) for United States and England may be owing to smaller samples (less than 40 RAOB-RO pairs in most of the months from 2006 to 2014). The relative larger *Std*(ΔT^{Time}) in Canada is mainly caused by the seasonal sampling difference. During summer, daytime RAOB-RO pairs are as many as 400 and drop to less than 10 pairs during winter (Figure 10d). The mean daytime temperature biases are 0.08 K for United States, 0.22 K for Australia, 0.22 K for Germany, -0.06 K for Canada, 0.12 K for England, and 0.35 K for Brazil.

The *Std*(ΔT^{Time}) for RS92 at night are larger than those during daytime, except for those in Canada, which may be due to a relative smaller RAOB-RO ensemble pairs in the nighttime over those regions. The *Stds*(ΔT^{Time}) for RS92 are 0.46 K for United States, 0.30 K for Australia, 0.24 K for Germany, 0.32 K for Canada, 0.42 K for England, and 0.43 K for Brazil in the nighttime. Their mean nighttime temperature biases are 0.19 K for United States, 0.23 K for Australia, 0.21 K for Germany, -0.01 K for Canada, 0.16 K for England, and 0.26 K for Brazil. The less than 0.5 K *Std*(ΔT^{Time}) for RS92 over daytime and nighttime over these six regions actually demonstrate the long-term stability of RS92 data.

The variation of mean ΔT^{Time} at different regions is highly related to the corresponding variation of SZA. The largest mean ΔT^{Time} (RS92) is over Brazil (i.e., ΔT^{Time} ($RS92_{Brazil}$) see Figure 10f), where the mean ΔT^{Time} ($RS92_{Brazil}$) is equal to 0.35 K and 0.26 K for the daytime and nighttime, respectively.

A seasonal variation of ΔT^{Time} (RS92) is not apparent expect over Canada. Although the mean temperature biases are very small (less than +/- 0.06 K) over Canada (i.e., northern high-altitudes), there still exist some seasonal-dependent temperature bias, which could be a result of the very few RAOB-RO ensemble pairs for night time in summer, and for daytime night time in winter (Fig. 10d). Over Canada, daytime SZA is as high as 50 degree in summer which becomes 88 degree during the winter. Therefore, the daytime ΔT^{Time} (RS92) can be as large as 0.3 K during the summer and as low as -0.3 K during the winter.

With less radiative effect on sondes, the magnitude of RAOB-RO temperature bias at150 hPa is in general smaller than those in 50 hPa (see Table 4). The mean ΔT^{Time} (RS92) at 150 hPa daytime temperature differences are 0.00 K for United States, 0.03 K for Australia, 0.07 K for Germany, 0.03 K for Canada, 0.04 K for England, and 0.21 K for Brazil. The corresponding ΔT^{Time} (RS92) for nighttime for these countries are 0.09 K for United States, 0.08 K for Australia, 0.06 K for Germany, 0.12 K for Canada, 0.05 K for England, and 0.23 K for Brazil. The $Std(\Delta T^{Time})$ for RS92 at these six regions are all less than 0.37 K during the day and less than 0.52 K during the night (Table 4).

4.3 Time Series Anomaly for Sippican MARK, VIZ-B2, AVK-MRZ, and Shanghai Sondes

To demonstrate the inter-seasonal and inter-annual variation of the RAOB-RO temperature biases, the time series of the monthly mean temperature bias for Sippican MARK IIA (ID=87), VIZ-B2, AVK, and Shanghai (i.e., $\Delta^{Time}_{Sippican}$, Δ^{Time}_{VIZ-B2}, Δ^{Time}_{AVK}, and

$\Delta^{Time}_{Shanghai}$) in the northern mid-latitudes (from 60°N to 20°N) are shown in Figure 11. The mean temperature biases in this region for $\Delta^{Time}_{Sippican}$, Δ^{Time}_{VIZ-B2}, Δ^{Time}_{AVK}, and $\Delta^{Time}_{Shanghai}$ for 50 hPa are summarized in Table 5.

Figure 11 shows the time series of the monthly mean temperature bias at 50 hPa. During daytime $\Delta^{Time}_{Sippican}$ (-0.12 K), Δ^{Time}_{VIZ-B2} (0.87 K), Δ^{Time}_{AVK} (0.80 K), and $\Delta^{Time}_{Shanghai}$ (0.10 K) are warmer than those in the nighttime. The monthly mean temperature biases at 50 hPa for nighttime are -0.12 K for $\Delta^{Time}_{Sippican}$, -0.56 K for Δ^{Time}_{VIZ-B2}, -0.03 K for Δ^{Time}_{AVK}, and -0.20 K for $\Delta^{Time}_{Shanghai}$. While $\Delta^{Time}_{Sippican}$ and $\Delta^{Time}_{Shanghai}$ are largely constant in time, Δ^{Time}_{VIZ-B2} has obvious seasonal variations with a negative trend during nighttime and positive trend during daytime. The number of VIZ-B2 observations drops off after 2012 (see Figure 11b), which contributes to the larger variation of the Δ^{Time}_{ViZ-B2} after then.

Δ^{Time}_{AVK} has an irregular seasonal variation particularly during daytime, with a large warm bias. A part of this irregular bias may be due to an unidentified change of instrumentation and large seasonal variations in sample numbers. The standard deviation of the temperature differences for these four sensors (i.e., $Std(\Delta^{Time}_{Sippican})$, $Std(\Delta^{Time}_{VIZ-B2})$, $Std(\Delta^{Time}_{AVK})$, and $Std(\Delta^{Time}_{Shanghai})$) at daytime are 0.33K for $\Delta^{Time}_{Sippican}$, 0.37 K for Δ^{Time}_{VIZ-B2}, 0.22 K for Δ^{Time}_{AVK}, and 0.18 K for $\Delta^{Time}_{Shanghai}$. The corresponding nighttime variations are 0.21 K for $\Delta^{Time}_{Sippican}$, 0.43 K for Δ^{Time}_{VIZ-B2}, 0.21 K for Δ^{Time}_{AVK}, and 0.17 K for $\Delta^{Time}_{Shanghai}$. The mean time series temperature bias at 50 hPa for these sensor types for the northern mid-latitude is summarized in Table 5 and the corresponding mean time series temperature differences at 150 hPa is summarized in Table 6.

5. Trend Analysis and Potential Causes of RAOB Temperature Trend Uncertainty

5.1 Comparison Method

To further quantify inter-annual variation of RAOB temperature biases for different sensor types, we conduct the trend analysis for the time series of RAOB-RO temperature anomaly. The anomaly of trend for each of individual sensor types relative to those of co-located RO temperature are computed and compared. We focus on the trend analysis for individual sensor types over specific regions similar to previous sections. The de-seasonalized temperature anomalies are computed by:

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

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 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 $\overline{T^{Time}}(l,m,k^{'})$ may vary for different sensor types. The mean trend of temperature difference anomalies for each of the sensor types at 50 hPa and 150 Pa are summarized in Tables 5 and 6, respectively.

5.2 Trend of Temperature Anomalies for Vaisala Sondes

The trend uncertainty for RAOB over different regions are mainly due to i) uncorrected solar zenith angle dependent biases, ii) changing of radiation correction, iii) and iv) small samples used in the trend analysis. While it is not possible to identify the bias for each of the individual causes, we can only quantify the combined statistical biases using RAOB-RO ensembles.

Figure 12 depicts the de-seasonalized temperature anomalies for Vaisala RS92 over United States ($\Delta T_{RS92\_USA}^{Deseason}$), Australia ($\Delta T_{RS92\_Australia}^{Deseason}$), German ($\Delta T_{RS92\_Germany}^{Deseason}$), Canada ($\Delta T_{RS92\_Canada}^{Deseason}$), England ($\Delta T_{RS92\_England}^{Deseason}$), and Brazil ($\Delta T_{RS92\_Brazil}^{Deseason}$), respectively. In general, daytime trend differences at 50 hPa in all six regions are within ±0.26 (K/ 5yrs, see Table 3). While the daytime trend of anomalies for RAOB and RO temperature at 50 hPa for United States and Germany are 0.00 and -0.02 K/5yrs, the trend of anomalies are equal to 0.18 K/5yrs over Australia, 0.24 K/5yrs over Canada, 0.26 K/5yrs over England, and 0.12 K/5 yrs over Brazil, respectively. This non-trivia trend anomaly in the later regions may be owing to the incomplete daytime radiation correction applied in these regions between ΔT ($RS92_{200701-201012}$) and ΔT ($RS92_{201101-201404}$) (see Figure 9). The corresponding nighttime trend differences in these six regions are -0.21 K/ 5yrs for United States, -0.08 K/ 5yrs for Australia, -0.14 K/ 5yrs for Germany, -0.02 K/ 5yrs for Canada, -0.16 K/ 5yrs for England, and -0.10 K/ 5yrs for Brazil (see Table 3).

To further examine the temperature trend uncertainty for global Vaisala sensors, we compare the global trend of anomaly for RS80, RS90, and RS92 at 50 hPa and 150 hPa in Tables 5 and 6, respectively. The global de-seasonalized temperature anomalies for Vaisala RS92 for daytime and nighttime are equal to 0.07 K/5yrs and -0.09 K/5yrs, respective (Table 5). The 95% confidence intervals for slopes are shown in the parentheses. This indicates that although there might be a small residual radiation error for RS92, the trend anomaly between RS92 and RO 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.

The trend anomaly between RS80 and RO is about 0.19 K/5yrs during the day and 0.11 K/5yrs during the night (Table 5). Those between RS90 and RO temperature in the lower stratosphere are equal to -0.01 K/5yrs and 0.04 K/5yrs for daytime and nighttime, respectively (Table 5).

To compute the degree of deviation between RAOB temperature and RO temperature, we also calculate the root mean square (RMS) temperature difference of the derived de-seasonalized anomalies. The global RMS for $\Delta T_{RS92}^{Deseason}$ in the daytime is equal to 0.06 K. This indicates the consistency of RS92 temperature measurements relative to the RO temperature. Both of the global RMS for RS80 and RS90 in daytime are (0.27, 0.26) K (Table 5).

Because the RAOB temperatures in 150 hPa are less biased compared to those at 50 hPa, the de-seasonalized temperature anomalies for Vaisala Sondes at 150 hPa are even smaller than those at 50 hPa. The trend differences for $\Delta T_{RS92}^{Deseason}$ at 150 hPa for RS92 are -0.13 K/5yrs for United States, 0.12 K/5yrs for Australia, -0.02 K/5yrs for Germany, 0.23 K/5yrs for Canada, 0.06 K/5yrs for England, and 0.00 K/5yrs for Brazil during the daytime (see Table 4). The corresponding trend differences during the nighttime are -0.23 K/5yrs for United States, -0.07 K/5yrs for Australia, -0.19 K/5yrs for Germany, -0.21 K/5yrs for Canada, -0.08 K/5yrs for England, and -0.01 K/5yrs for Brazil.

The global RMS of RAOB-RO anomalies for RS92 at 150 hPa are 0.04K for daytime and 0.07 K for nighttime (Table 6).

5.3 Trends of Temperature Anomalies for Sippican MARK, VIZ-B2, AVK-MRZ, and Shanghai Sondes

Figure 13 depicts the de-seasonalized temperature anomalies for Sippican MARK IIA (ID=87), VIZ-B2 (ID=51), AVK-MRZ (ID=27), and Shanghai (ID=32) (i.e., $\Delta T^{Deseason}_{MARK-IIA}$, $\Delta T^{Deseason}_{VIZ-B2}$, $\Delta T^{Deseason}_{AVK}$, and $\Delta T^{Deseason}_{Shanghai}$) at 50 hPa, respectively. The trends of temperature anomalies for these sensor types are listed in Table 5. The 95% confidence intervals for slopes are shown in the parentheses. The daytime temperature trend anomalies are 0.41 K/5yrs for $\Delta T^{Deseason}_{MARK-IIA}$, 0.47 K/5yrs for $\Delta T^{Deseason}_{VIZ-B2}$, -0.14 K/5yrs for $\Delta T^{Deseason}_{AVK}$, and 0.18 K/5yrs for $\Delta T^{Deseason}_{Shanghai}$, which are much larger than those of the Vaisala RS92. The corresponding nighttime trend anomalies are 0.24 K/5yrs ($\Delta T^{Deseason}_{MARK-IIA}$), -0.35 K/5yrs ($\Delta T^{Deseason}_{VIZ-B2}$), -0.14 K/5yrs ($\Delta T^{Deseason}_{AVK}$), -0.02 K/5yrs ($\Delta T^{Deseason}_{Shanghai}$). Since the number of AVK - RO pairs decrease significantly after 2012, the trend anomaly for AVK-RO pairs before and after 2012 vary.

The root mean square (RMS) of the de-seasonalized time series *Std(ΔT*^{Time}*)* is used to indicate the trend uncertainty of the time series. The trend differences and RMS for all the sonde types at 50 hPa and 150 hPa are summarized in Tables 5 and 6, respectively.

Page 22: [17] Formatted	Ben Ho	1/17/17 10:47 AM

Indent: First line: 0"

Page 22: [18] Deleted	Ben Ho	12/5/16 4:57 PM

in the lower stratosphere comparing to the co-located RO temperatures.

1.

Page 24: [19] Deleted	Ben Ho	12/5/16 5:13 PM

The daytime trend of anomalies for RAOB and RO temperature at 50 hPa for United States and Germany are equal to (0.00, -0.02) K/5yrs, the trend of anomalies are equal to (0.18, 0.24, 0.26, 0.12) K/5yrs over Australia, Canada, England, and Brazil, respectively. The trend anomaly between RS92 and RO from June 2006 to April 2014 is within +/- 0.09 K/5yrs globally. The trend anomaly between RS80 and RO is about 0.19 K/5yrs during the day and 0.11 K/5yrs during the night. The daytime temperature trend anomalies for $\Delta T_{MARK-IIA}^{Deseason}$, $\Delta T_{VIZ-B2}^{Deseason}$, $\Delta T_{AVK}^{Deseason}$, and $\Delta T_{Shanghai}^{Deseason}$ are (0.40, 0.47, -0.14, 0.18) K/5yrs, which are much larger than those of RS92.

Page 25: [20] Deleted	Ben Ho	12/8/16 3:03 PM

Appendix A: The Quality of GPS RO Data as Benchmark References and the New Reprocessed Package

A1. The Quality of GPS RO Data as Benchmark References for Climate Studies

While the position and speed of the GPS and low earth orbit (LEO) satellites are known, we can inverse the time delay to bending angles, refractivity, and temperature vertical distribution with high precision and accuracy (Ho et al., 2009a,b, 2012). While time delay and bending angle 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 shown that the mean RO temperature biases for one center (for example from UCAR) relative to the all center mean is within ±0.1K from 8 km to 30 km except for south pole above 25 km (see Fig. 6d in Ho et al., 2011). Ho et al., (2007, 2009b) demonstrated that the RO derived temperature profiles in the lower stratosphere are extremely useful to identify and calibrate the inter-satellite microwave brightness temperature differences form Advanced Microwave Sounding Units (AMSU) and Microwave Sounding Units (MSU) on board different satellite missions. In this study, UCAR RO temperature profiles will be used in this study.

GPS RO observations are of high vertical resolution (from ~60 m near the surface to ~1.5 km at 40 km). The mean temperature difference between the collocated soundings of COSMIC and CHAMP is within 0.1 K from 200 hPa to 20 hPa (Ho et al., 2009b;

Anthes et al., 2008; Foelsche et al., 2009). Schreiner et al., (2014) compared re-processed COSMIC and Metop-A/GRAS bending angles produced at CDAAC. The mean COSMIC and Metop-A/GRAS bending angle differences are about 0.02–0.03 μrad which demonstrates the reproducibility of COSMIC and Metop-A/GRAS. The mean layer temperature difference between 200 hPa to 10 hPa is within 0.05 K (not shown). This is consistent with those between COSMIC and CHAMP at the same height (Ho et al., 2009a). The precision of RO temperature is ~ 0.1 K (e.g., 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 accuracy 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 Vaisala-RS92 and COSMIC coincident data are collected. The mean bias in this height range is equal to -0.01 K with a mean standard deviation of 2.09K. Although the quality of Vaisala-RS92 may vary in different regions (see Section 4.1), this comparison demonstrates the quality of RO temperature profiles in this height range.

A2. Brief Description of the New Inversion Package from CDAAC

Comparing with the previous version, the new inversion package used improved precise orbit determination (POD) and excess phase processing algorithm, where a high-precision, multiple Global Navigation Satellite System (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 (Schrein et al., 2011). Compared to the real-time processed RO data, much improved and more completed satellite POD data are used in the reprocessed package. The re-processed COSMIC and Metop-A/GRAS data would produce more consistent and accurate RO variables than those from post-processed (periodically updated inversion packages were used) and real-time processed datasets.

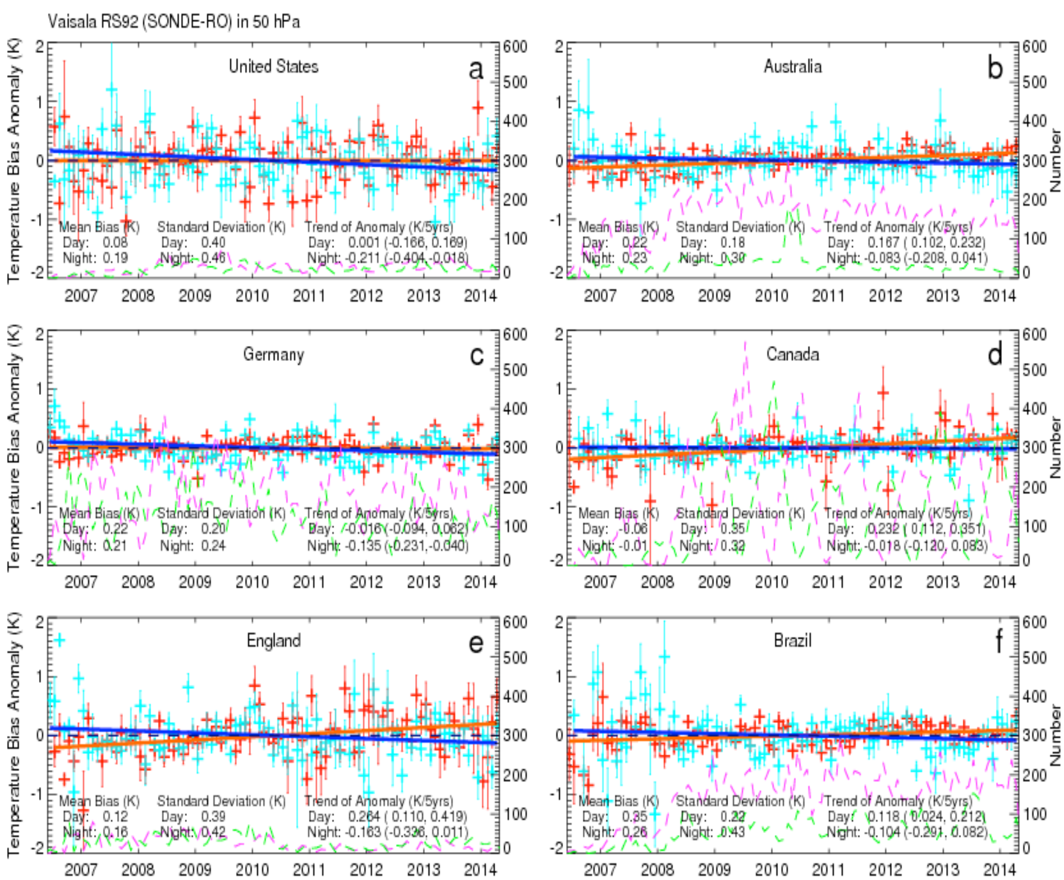
Page 30: [21] Deleted	Ben Ho	12/8/16 9:24 AM

Figure 12. The time series of de-seasonalized temperature anomaly in 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) England, and f) Brazil. The 95% confidence intervals for slopes are shown in the parentheses.

Page 31: [22] Deleted	Ben Ho	1/16/17 4:09 PM

Page 31: [23] Deleted	Ben Ho	1/16/17 10:02 AM

Page 42: [24] Deleted	Ben Ho	1/17/17 10:50 AM



Page 42: [25] Deleted	Ben Ho	12/8/16 9:28 AM

Figure 12. The time series of de-seasonalized temperature anomaly in 50 hPa for RS92 for a) United States, b) Australia, c) Germany, d) Canada, e) England, and f) Brazil. The 95% confidence intervals for slopes are listed in the parentheses.

Page 42: [26] Deleted	Ben Ho	12/8/16 9:28 AM

Page 44: [27] Deleted	Ben Ho	12/8/16 11:39 AM

Page 47: [28] Deleted	Ben Ho	12/8/16 11:08 AM

Table 6. Mean, standard deviation (std), trend (K/5yrs), and root mean square (RMS) of time series of temperature anomaly at 150 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 anomaly are shown in the parentheses.

		Day				Night			
		Mean Bias	STD Of MB	Trend of ANOM (k/5yrs)	RMS of ANOM	Mean Bias	STD Of MB	Trend of ANOM (k/5yrs)	RMS of ANOM

RS80	37,52, 61,67	0.18	0.19	0.045 (-0.036,0.126)	0.18	0.21	0.29	0.063(-0.055,0.181)	0.263
RS90	71	0.1	0.31	-0.058 (-0.181,0.065)	0.275	0.11	0.33	-0.065(-0.203,0.072)	0.307
RS92	79,80, 81	0.08	0.05	0.013 (-0.005,0.031)	0.041	0.05	0.08	-0.068(-0.094,-0.042)	0.066
Russia	27,75, 88,89 58	0.54	0.19	-0.194 (-0.254,-0.134)	0.16	0	0.21	-0.147(-0.199,-0.094)	0.135
VIZ-B2	51	0.65	0.38	0.370 (0.227,0.514)	0.362	-0.15	0.3	-0.051(-0.182,0.079)	0.272
Sippican MARKIIA Chip	87	-0.23	0.19	0.217 (0.148,0.285)	0.181	-0.1	0.18	0.160(0.095,0.226)	0.158
Shanghai	32	0.32	0.12	-0.086 (-0.149,-0.023)	0.1	0.08	0.19	-0.302(-0.394,-0.210)	0.176
Meisei Japan	47	0.06	0.53	-0.102 (-0.371,0.168)	0.458	0.19	0.5	0.001(-0.271,0.272)	0.455