Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-801-AC2, 2017

Atmospheric Chemistry and Physics
Discussions

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.

Shu-Peng Ho et al.

spho@ucar.edu

Received and published: 24 January 2017

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 (sigmaËĘ2/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.

This figure is inserted in the supplement. This figure is also submitted as Figure 1.

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.

Please also note the supplement to this comment:
http://www.atmos-chem-phys-discuss.net/acp-2016-801/acp-2016-801-AC2-supplement.pdf

---

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-801, 2016.

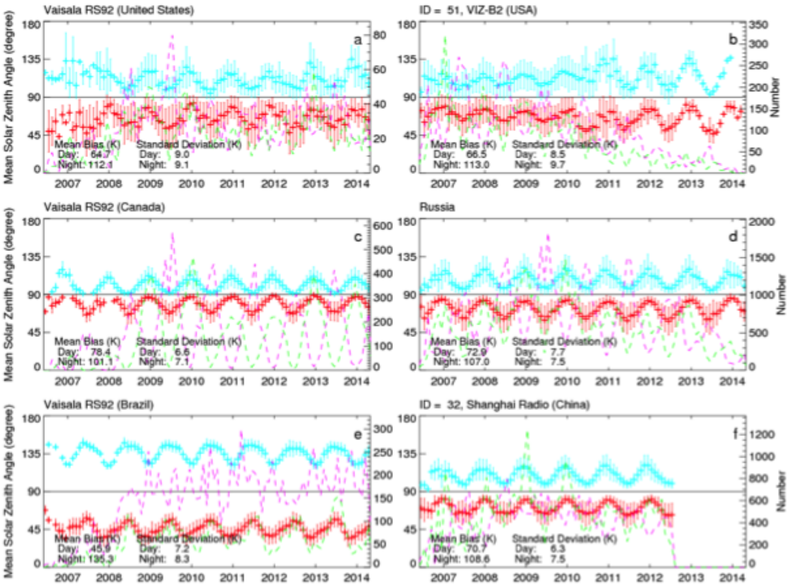

Fig. 1.