Atmos. Meas. Tech. Discuss., 8, C5728–C5760, 2016
www.atmos-meas-tech-discuss.net/8/C5728/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Atmospheric Measurement Techniques Discussions

Interactive comment on "Comparison of Vaisala radiosondes RS41 and RS92 at the ARM Southern Great Plains Site" *by* M. P. Jensen et al.

M. P. Jensen et al.

mjensen@bnl.gov

Received and published: 18 March 2016

We thank two anonymous referees and Dr. Fujiwara for their thoughtful comments and suggestions on the manuscript. Following these suggestions we have made many changes to the manuscript, outlined in detail below with specific responses to each comment. Some of the major changes to the manuscript involve: 1) The clarification that the conclusions regarding the impact of the differences in the radiosonde observations, while minimal for many applications, could be significant for other applications such as long-term trend analysis of atmospheric thermodynamics, and 2) The reworking of the analysis of the impact of clouds on the radiosonde measurements, both for "wet-bulbing" effects and radiative heating effects, and statistical representativeness.

Response to comments from Anonymous Referee #2 on "Comparison of Vaisala radiosondes RS41 and RS91 at the ARM Southern Great Plains Site" by M. P. Jensen et al.

Major comments:

[1] Referee #2 writes, "There seems to be a problem in the pressure measurements of the Vaisala RS92 and I suspect that ARM is not following proper procedure in preparing the pressure sensor of the RS92. Proper Vaisala RS92 ground check corrections require a correction of the radiosonde pressure using an independent surface pressure measurement. If this has been done correctly the pressure measurements at launch will be nearly identical to the reference pressure. This reference pressure is also used for the initialization of the Vaisala RS41 pressure and therefore the pressure measurements between the two radiosondes must almost by definition be in near agreement near the surface. This is not the case here. This shortcoming has already been recognized by the GRUAN processing of the ARM data, but apparently not been considered here. All RS92 soundings should be reprocessed using the proper ground check correction for pressure."

Reply: We carefully reviewed the ground check procedures performed during the field phase of the intercomparison study. These procedures followed those outlined by Vaisala and were under the observation of a Vaisala technician for the first half of the intercomparison launches. During this review we identified a small difference in the height of the radiosonde ground check station and the independent surface pressure measurement that was not accounted for during operations. We have applied a correction to the reported pressures to account for this difference and produced new plots using this pressure value. The difference was small enough that there is no noticeable change in the figures.

The reported mean pressure difference of the RS41-RS92 radiosondes is 0.5 hPa close to the ground and the 10/90 percentiles are within +/- 1.0 hPa. This result is to be expected as it is achieved with two instruments using different technology and both having a measurement uncertainty of 1.0 hPa. It is true that at the ground level the two radiosondes are set to output the same pressure readings “by definition.” However, the GPS-based pressure measurement is subjected to disturbances caused by multi-path propagaton of the GPS signal, especially at altitudes ranging from 0-3 km above ground level. Therefore, it is typical that one sees small differences between GPS-based and sensor-based pressure measurements close to the ground. The fact that the mean pressure differences are so small (< 0.15 hPa) above 3 km provides confidence that there was no significant error in the initializations of the measurements. It may be possible to minimize uncertainties, in GPS-based pressure measurements, due to multi-path propagation by using an antenna model that is more robust to these disturbances.

Changes: Fig. 8, former Fig. 13 (current Fig. 11), and former Fig. 16 (current Fig. 15) were recreated applying the small pressure correction. As noted in the reply, the difference in the plots is not noticeable. We have also added a sentence in section 4, within the discussion of Fig. 8a of the possible influence of GPS signal propagation effects on the RS41 pressure measurement near the surface.

[2] Referee #2 writes, “I am somewhat concerned about the statistics and the significance of the cloud related results. The authors define 8 could types and end up having mostly just two soundings per cloud type. One type only has one sounding. Furthermore, there are daytime and nighttime soundings, which will further reduce the number of soundings per cloud type and time of day. Only for daytime soundings would I expect a significant cloud influence on the comparison, thus, only daytime soundings should be used here. Since the differences are generally small, basing statements on a low number of soundings is not ideal. For example, Figure 12 shows very large differences in cc6 (two sondes total, one daytime, one nighttime) compared to the other cloud groups. I would speculate that a single comparison strongly influences this statistics. Likewise, the nighttime comparison of pressure seems to be significantly different than

the daytime comparison, which is unexpected. I can speculate that here as well a single comparison influences the statistics. The statistics of the entire cloud discussion should either be improved or the entire discussion should be removed.”

Response: We agree that our original analysis based on the cloud categories was not well constrained, and particularly, the conclusions based on cloud categories with only one or two soundings in them were not robust. Because we believe that the availability of the cloud profiling observations is a unique aspect of this study, and can provide some additional insights we have chosen to re-formulate this analysis, rather than remove it, within the construct of the limited dataset that we have. In the best case scenario, we would have many more twin soundings with which to build robust statistics. We have reformulated our cloud categories considering the cloud occurrence in three separate layers: Low < 3 km, 3 km < Middle, < 8 km, and High > 8 km. This provides 3 categories with three or more daytime sounding flights as shown in Table 7. We then compare only those categories that include three or more sounding profiles.

Changes: Table 7 has been completely reformulated using these predefined cloud layer descriptions. We have also indicated (*) which soundings were taken during nighttime. Former Fig. 12 (current Fig. 13) has been redone comparing only cloud categories 2, 3, 4, which have three or more daytime sounding flights. The discussion of current Fig. 13, in Section 4 has been completely rewritten.

[3] Referee #2 writes, “The conclusion and abstract state, that there will be no significant impacts in the switch from the RS92 to the RS41. This statement only applies to the currently used MW31 version 3.66, but not for the MW31 version 3.62, which has been used until 2011. The changes between these two Vaisala RS92 versions strongly impact long-term climate data series, and therefore, this change must be evaluated as well. The authors must point out that there already is an inconsistency in the RS92 record, which has not been properly accounted for. The ARM radiosoundings are also part of the GRUAN network and these data have been processed by GRUAN using an independent processing system, which considers these inconsistencies. All

ARM RS92 data have been or are scheduled to be reprocessed using this processing system. Therefore, the comparison between RS41 and RS92 should also refer to the GRUAN processed ARM data, which do not suffer from changes in processing version (3.62 to 3.66) or changes in sensors (coating of humidity sensors). A comparison to the GRUAN processing of the same data is required if the authors truly want to make a statement that there will not be significant impacts on the long term series in switching from the Vaisala RS92 to the Vaisala RS41. As is, their statement is somewhat misleading and only refers to the changes between Vaisala RS92 processing MW31 3.66 to Vaisala RS41 processing MW41 2.1, which is valid only for the last four years, but not for the entire ARM RS92 data record.

Reply: For the majority of science applications, particularly those of users of the ARM Climate Research Facility, the differences between the RS92 and RS41 measurements (which were generally smaller than the manufacture defined instrument combined uncertainties) will have little to no impact. However, for applications such as determining long-term trends in atmospheric thermodynamics these small differences may have significant impacts. Therefore, the conclusion was admittedly overstated.

Based on the intercomparison that was done any conclusions are only applicable to the RS92 and RS41 radiosondes using the current MW31 version 3.66 for the SGP locale and the conditions encountered during the intercomparison. We are not considering the changes from MW31 version 3.62 to 3.66.

The purpose of the manuscript is the comparison of the operationally available measurements from the RS92 and RS41 radiosondes and the documentation of those differences. These are the observations that many operational and scientific users work with and therefore it is important to characterize these measurements without the additional post-processing that is done as part of the GRUAN project. For this reason we have chosen not to include the GRUAN processed data in this manuscript. We do see the need for additional work to characterize the RS41 observations within the GRUAN framework particularly for the purpose of homogenization of the long-term radiosonde

record.

For completeness, we have obtained the GRUAN processed RS92 radiosondes and compared to both the Vaisala processed RS92 and the RS41 radiosonde observations. For pressure measurements, the GRUAN corrections are smaller than 1 mb for nearly 97% of the measurements in this intercomparison. For most measurements, especially below 10 km the GRUAN processed pressure measurement is less than the original measurement. For temperature measurements, the GRUAN processed data decreases the temperature measurement less than 0.5oC for nearly 95% of the observations in this intercomparison. For relative humidity the GRUAN processing produces differences less than 5(2)% for more than 96(91)% of the measurements in the intercomparison. Above 15 km the GRUAN corrected value is generally greater than the original RS92. The GRUAN processed soundings result in larger differences from the RS92 than that between RS92 and RS41. Here we show the comparison of the GRUAN processed RS92 and the RS41 sounding observations in Fig. A. Compared to the differences without the GRUAN processing, the pressure comparison indicated that the RS41 are generally lower than the GRUAN calues in the lower troposphere, the temperature measurements show the RS41 measurements warmer than the GRUAN processed values at all levels, the RH differences are similar with an without GRUAN processing and the wind differences show larger differences following the GRUAN processing at all heights.

Changes: We have changed the last sentence of the abstract to more specifically state. “For a majority of science applications, particularly those of users of ARM data, the differences between the RS92 and RS41 measurements should have little impact.” We have also edited the conclusions to state, “For many science applications, particularly those of users of ARM data, the described differences between RS92 and RS41 will have little impact, however, for long-term trend analyses of atmospheric thermodynamic quantities additional characterization of the RS41 measurements and their relation to the long-term observational records will be required. “

Figure A Vertical profiles of the median (black), 25/75th percentile (green) and 10/90th percentile (grey) differences between GRUAN- processed RS92 and original RS41 observations (RS92-RS41) for (a) pressure, (b) dry bulb temperature, (c) relative humidity, (d) zonal wind and (e) meridional wind.

Minor comments:

[4] Referee #2 writes, "Page 11326, lines 21-23: Please better describe the difference between the RS41 and RS92 humidity sensor. The expression 'heating functionality' is not clear, since both sondes use sensor heaters."

Reply: We have expanded upon the description of the differences between the RS41 and RS92 humidity sensors.

Changes: The text has been changed to, "For humidity observations the RS41 uses a thin-film capacitor with an integrated temperature sensor and heating functionality while the RS92 uses a thin-film capacitor with a heated twin sensor. In both radiosonde models heating is used as a means for deicing the humidity sensor when a radiosonde traverses through cloud layers with freezing conditions. In the case of the RS41, a controlled heating is applied for the purpose, whereas in the RS92, the two sensors are pulse-heated sequentially. In general, the RS41 humidity sensor has improved resolution, response time and accuracy compared with the RS92 (Table 3, Vaisala 2014)."

[5] Referee #2 writes, "Page 11327, line 1: The authors should also note that the Vaisala ground software for the RS92 can be configured such that the GPS measurements are used for pressure calculations of the RS92. Although this may not be practices widely, it could lead to a similar performance of pressure determinations in the RS92 as the RS41."

Reply: The RS41 model (RS41-SG) used in this study uses the GPS observations as stated to determine the pressure. A different model, RS41-SGP, has a pressure sensor

similar to the RS92-SGPD model used in the study. When using the RS41-SGP and the RS92-SGPD it is possible to configure the sounding system to utilize either the pressure sensor or GPS sensor.

Change: We have changed the text to make this distinction clear, "The RS41 model used in this trial, RS41-SG, makes use of GPS observation of vertical displacement along with the temperature and humidity measurements to derive the atmospheric pressure, while the RS92 model, RS92-SGPD, uses a direct measurement of pressure with a silicon, capacitive sensor. Note that there is also a model RS41-SGP with a pressure sensor, similar to the RS92-SGPD, and, with both models, it is possible to configure the sounding system to utilize either sensor or GPS based pressure for the sounding profile."

[6] Referee #2 writes, "Page 11327, lines 27-28: The ground check device of the RS41 only reconditions the RS41 humidity sensor comparable to the RS92 reconditioning, but no 0% ground check is performed, since no 0% RH reference is involved. This sentence should be deleted or modified accordingly."

Reply: The RS41 performs the equivalent of a zero humidity check through heating of the sensor. In order to clarify this, we have explained the ground check procedure more thoroughly.

Change: The sentence describing the zero humidity check was changed to, "Also, the RS41 ground check device enables an accurate zero humidity check without the use of a desiccant, as in the GC25 ground check device used with the RS92. For the RS41, the dry reference condition of the zero humidity check is generated in open air by heating the sensor using the integrated heating element on the sensor chip. The procedure utilizes the commonly known behavior of relative humidity: relative humidity decreases towards zero humidity level as the temperature rises high enough."

[7] Referee #2 writes, "Page 11328, lines 1-3: This statement is highly speculative. The operations have certainly been simplified and remove some chance for operator error,

but recommending less training may lead to new additional operator errors compensating the gains of the new system. Time savings is probably minimal. These statements should be removed or substantiated with real data."

Reply: Perhaps this statement is speculative. Our experience during the intercomparison was that there was an obvious time savings and simplicity to the launch procedures associated with the RS41 radiosonde compared to the RS92 radiosonde. We did not however quantitatively time the preparation procedure.

Changes: Since we did not quantify the time savings, and have already outlined the procedural changes that make the preparation of the RS41 simpler, we have removed the sentence that indicated the preparation time was shorter and that less training would be required.

[8] Referee #2 writes, "Page 11328, line 10: In 2011 Digicora version 3.64 introduced a new correction for time lag and solar radiation correction for humidity measurements. These corrections can be turned off; therefore it is important to specifically mention that they were used. Furthermore, this version has an updated solar radiation correction table and algorithm for the temperature measurement, which should be pointed out."

Reply: The new corrections for time lag and solar radiation were used.

Changes: The text has been modified to clearly state that the latest corrections were applied. We have added a sentence to this paragraph that states, "All correction algorithms were enabled in the sounding systems, and, specifically, the solar radiation corrections for the temperature and humidity measurements, updated since version 3.64, were applied in MW31 calculations."

[9] Referee #2 writes, "Page 11328, lines 23-24: The main ventilation term is due to the ascent rate. This does not change in the larger payload. This statement does not hold and should be deleted."

Reply: Higher inertia and drag of the payload, result in a more stable flight which will

generally result in less ventilation.

Changes: We have edited the text to describe this better, "In the twin sounding, due to higher inertia and drag of the payload, and thus more stable flight, the sensors generally have slightly less ventilation. A larger payload may also magnify the effects of some error sources, for example, temperature sensor orientation error caused by solar radiation."

[10] Referee #2 writes, "Page 11329, line 1: The abstract states a sounding period from June 3-8 not June 3-7."

Reply: The first sounding of the intercomparison was launched at 12:55 LT on 3 June. The final sounding was launched at 23:55 LT on 7 June and terminated at 1:07 LT on June 8.

Change: The dates have been changed to June 3-8 in the text.

[11] Referee #2 writes, "Page 11331, line 27: This statement is somewhat generic and based only on two profile comparisons shown. Since a larger number of soundings passed through clouds can the authors substantiate this statement with a statistics of all their comparisons showing that the RS41 is 'less prone' to wet-bulbing? Can the authors provide information that the hydrophobic coating has changed to explain this improvement? I am somewhat concerned that there appears to be a 250 m separation between the cloud top and the when the radiosonde sensors see the cloud top. Is that an indication that the RS41 may also suffer from wet-bulbing for something like 50 s? Or can the ARM Raman Lidar confirm that the moist layer extended significantly above the cloud?"

Reply: Based on Figs. 9 and 10, the statement of the RS41 being less prone to "wet bulbing" effects was admittedly overstated. However, it does appear that for these two flights the RS41 measurements were not impacted by "wet bulbing" effects as much as the RS92. We have edited the text such that the statement only applies to these two sounding flights. The sequential pulse heating method of the RS92 with substantially long sensor element heating and cooling (non-heating) periods may not be sufficient to eliminate sensor ice/wetting in some in-cloud conditions. A review of the height coordinates for the radar image (Fig. 10) and the sounding image (Fig. 9) showed that the Fig. 10 heights are above ground level (AGL) and the sounding heights are above mean sea level (AMSL). The SGP site is 315 meters AMSL. This explains the difference in the cloud-top height from the radar and the radiosonde.

Change: We have changed the sentence to read, ". For these two sounding flights, the RS41 measurements seem to have less impact from "wet bulbing" effects compared to the RS92, consistent with the results of Edwards et al. (2014). " The heights in Fig. 10 have been changed to AMSL. This height difference also impacted Fig. 6 and for Fig. 14 (current Fig. 12). The heights in these figures have also been changed to AMSL.

[12] Referee #2 writes, "Page 11332, lines 6-9: The wind measurements appear to be basically in agreement and the differences appear not to be statistically significant. The differences may be a result of filtering and the authors should make a statement, whether the (small) differences are statistically significant or not. If they are statistically insignificant, which I believe they are, then they shouldn't be called differences."

Response: We have determined that the differences in the zonal wind measurements are not statistically significant, while the differences in the meridional winds are statistically significant (though still small). Some of this difference in statistical significance is likely the result of the fact that differences in wind direction measurements will propagate to larger differences in the meridional wind than the zonal wind due to the prevailing wind direction being near 270o. (see response to comment [48].

Changes: We have indicated that the observed differences in the zonal wind are not statistically significant, while those for the meridional winds are statistically significant. We have also added some text describing why these differences are expected to be larger for the meridional component for the wind for the observed conditions.

[13] Referee #2 writes, "Page 11332, lines 8-10: These statements are speculative and may not hold up to scrutiny. Despite increase radiative heating, the processing for both radiosondes uses a correction for radiative heating. Difference in the correction algorithm though may contribute to the (very small) differences in temperature. Wet-bulbing at 10 km can be excluded, since cirrus clouds don't tend to lead to sensor icing. Sensor response time may be a reason, but again, the processing does account for time response issues, which therefore should not be an issue. The differences in humidity are significant and may also be caused by sensor calibration, which should be mentioned."

Response: We believe that the reviewer was referring to lines 13-16. We agree that the statements made conclusions beyond what could be surmised from the analysis. We have rewritten this to indicate where there is speculation from just a couple cases. We agree that "wet-bulbing" at 10 km (well above the freezing level) does not contribute to differences at those heights. We also agree that sensor calibration may be providing some contribution to the observed differences.

Changes: We have rewritten this sentence to indicate speculation about possible causes for the differences, including sensor calibration. Since we removed Fig. 11, we have removed much of this paragraph. However, the sentence referred to in the comment has been used as a concluding sentence to the preceding paragraph and now reads, "The relative peaks in the temperature and relative humidity differences near a height of 10 km may be related to a combination of sensor calibration, differences in radiative heating impacts (measurements plus correction algorithms) of sensors due to contributions from cloud albedo and sensor response time in regions of strong gradients as the sondes traverse cloud layers."

[14] Referee #2 writes, "Page 11332, line 20: I cannot see a 'broader peak' between 8 and 10 km. Maybe a broad peak between 5 and 13, but the correlation between general cloud occurrence and the features shown in Figure 8 is weak at best. This should either be deleted or strengthened. As is, I would not argue that the behavior in

and around clouds represent the largest differences between the sondes."

Response: In retrospect, the original Fig. 11 did not illustrate the cloud frequency of occurrence very well, particularly important day vs. night differences. We have opted to remove this figure, and reorganize the following discussion to better link the occurrence of clouds and the observed differences both as a function of day vs. night and cloud category differences.

Changes: Former Fig. 11 has been removed. We have moved former Fig. 13 to current Fig. 11, former Fig. 14 to Fig. 12, and former Fig. 12 to Fig. 13 (the following figures former Fig. 15-19 then become Fig. 14-18 respectively). We have also rewritten the discussions of current Figures 11-13.

[15] Referee #2 writes, "Page 11333, line 19: Better write '. . . noisier as the daytime differences.' The comparison of temperature measurements is not a true comparison of atmospheric temperature measurements, but also a comparison of the radiation correction schemes used for the RS92 and RS41. It would be very helpful, if the authors could discuss the difference in radiation correction schemes and how they relate to the observed difference in temperature."

Response: The greater amount of noise in the nighttime differences are the result of the smaller sample size (6 nighttime vs. 14 daytime) and so the levels of noisiness in the profiles are not directly comparable. The temperature comparisons do indeed include both the measurement differences and any differences in the radiation corrections. That being said, the two radiosonde types use the same principles to correct for solar radiation effects.

Changes: We have edited the text in two places, first to read "Note that there were only 6 nighttime and 14 daytime soundings during the intercomparison and due to the notable difference in sample sizes, the levels of noisiness in the nighttime/daytime median difference profiles are not directly comparable." And we have also added the following sentences, "The heating effect of daytime solar radiation is corrected in both

radiosonde models using the same principles. However, the radiation effect is still seen as an increased measurement noise during daytime soundings and, mainly due to differences in sensor design, also as small deviations between the two measurements." In order to clarify the role of both the measurement and the radiation corrections in the differences.

[16] Referee #2 writes, "Page 11336, line 8: The authors should not only make use of the other measurements at the site, but in particular of the different RS92 processing provided by GRUAN for ARM. Since ARM SGP is formally a site within GRUAN, all data are also processed by GRUAN and these data are available to the authors. It would be most useful to make use of this processing, since this processing is more self-consistent than the operational Vaisala processing."

See response and changes for comment [3].

[17] Referee #2 writes, "Page 11336: The agreement between the two radiosondes is remarkably good; however, there appears to be a systematic dry bias of the microwave observations compared to the radiosondes. This is somewhat surprising, since the Yu et al., (2015) showed much better agreement between the radiosondes and the microwave at the ARM site at Manus. The authors should use the GRUAN data at SGP, which already provide uncertainties for RH, to estimate the uncertainty for PWV of their own data instead of referencing data from a tropical site with much higher PWV. Can they comment on the apparent dry bias of the microwave?"

Reply: We have revisited both the calculation of the PWV from the radiosonde measurements, and in consultation with Dr. David Turner (NSSL) have considered a 2-channel microwave radiometer retrieval with improved representation of the water vapor continuum along with a 3-channel MWR retrieval. We did identify a small bug in our calculation of PWV from the radiosondes that resulted in an approx. 0.4 mm over-estimate for all soundings. However, this is a small change relative to the reported differences. After this review, we have additional confidence that the estimates provided in former Fig. 19 (current Fig. 18) are correct for both the sonde and MWR PWV. With these corrections, nearly half of the launches show the sonde-derived PWV greater than the range of PWV observed by the MWR over the half hour following launch. This is not entirely unusual and has been observed previously at the SGP (Jensen et al. 2015b) and at the ARM site at Manus, PNG (Ciesielski et al. 2014). These cases do not correlate with observed cloud cover, surface wind speed/direction, humidity, or PWV. We believe that non-local variability in soil moisture, and low-level humidity is contributing significantly to the sonde PWV estimates. The Oklahoma Climatological Survey report for June 2014 shows the SGP site near the edge of a strong gradient in soil moisture, with much larger values to the northeast of the SGP site. Many, but not all of the radiosonde flights travelled to the NE of the site over the lowest 2 km of their flight and likely experienced higher humidity values than over the SGP site.

Changes: We have updated former Fig. 19 (current Fig. 18) correcting for a small bug (0.4 mm overestimate) in the estimation of the PWV from the radiosondes. We have added a statement suggesting the importance of horizontal inhomogeneity of the low-level humidity and we have adjusted the numbers regarding agreement between the PWV estimates based on the new values after the bug fix.

[18] Referee #2 writes, “Figure 8: Pressure difference: The pressure difference near the surface seems to be off by about 0.5 hPa on average. This should not happen, if both the Vaisala RS92 and Vaisala RS41 were initialized using the same reference pressure (see major comment above).”

See response and changes for comment [1].

[19] Referee #2 writes, “Figure 11: The X-axis label should probably be ‘Number of soundings’, where the legend explains what is meant by this.”

Reply: We agree that the x-axis label is not clear.

Changes: As suggested, we have changed the X-axis label.

[20] Referee #2 writes, “Figure 18: Since RH over liquid at ice saturation decreases with temperature, it would be better to show the RH difference as relative difference, not as absolute difference. It would be most enlightening, if the authors would also include the GRUAN processed data in this figure.”

Reply: We duplicated this figure using relative differences rather than absolute differences (Fig. B). This version of the figure is more difficult to explain, and the language can quickly get confusing (e.g. relative differences of relative humidity). The new figure also reduces the dynamic range of the resulting differences and does not tell a different story, so we have chosen to leave the original figure.

Figure B Relative difference in relative humidity between the RS92 and RS41 sondes as a function of temperature for four difference relative humidity ranges.

Changes: None.

[21] Referee #2 writes, “Figure 19: It would be better to show a PWV difference as function of PWV, rather than PWV as function of sonde number. The plot should also indicate daytime and nighttime comparisons.”

Reply: We made the plot as a function of PWV rather than sonde launch number. The resulting figure is more difficult to read and did not shed any further light on the reasons for the observed differences. Five of the nighttime soundings do show the sonde estimates greater than the MWR range, so we have highlighted the nighttime cases. However, it should be noted that some of the daytime cases also show the sonde estimates greater than the MWR range.

Changes: We have designated the nighttime soundings in former Fig. 19 (current Fig. 18).

Technical comments:

[22] Referee #2 writes, “Page 11324, line 25: Some important references seem to be missing here.”

Response: It would be too much to list all of the references that have ever presented information on the historical changes in radiosonde instrumentation, practices, processing and other issues. That being said, our original reference list was overly short and did not include many important and recent references.

Changes: We have expanded the list of references here to include: Milosevich et al. 2004, Haimberger 2007, Vomel et al. 2007, Haimberger et al. 2008, Sherwood et al. 2008, Seidel et al. 2009, Dai et al. 2010, Immler et al. 2010, Thorne et al. 2001, Moradi et al. 2013, Wang et al. 2013, Yu et al. 2015, Bodeker et al. 2016.

[23] Referee #2 writes, "Page 11325, line 19: Delete 'first.'"

Change: Deleted.

[24] Referee #2 writes, "Page 11325, lines 20-21: Delete 'with new technical solution"'

Change: Deleted.

[25] Referee #2 writes, "Reference Yu et al. appeared 2015, not 2014."

Change: This has been corrected in the text and the list of references.

[26] Referee #2 writes, "Table 3: The resolution of the humidity observations from the RS92 can also be set to 0.1%."

Reply: The RS92 can be set to a resolution (P[hPa], T[oC], U[m/s]) of (0.1, 0.1, 1) for EDT output and (0.01, 0.01, 0.01) for FLEDT output.

Change: Since we used FLEDT output, we have changed tables 2,3,4 to represent the FLEDT output resolutions, and we have included a note at the bottom of the table.

[27] Referee #2 writes, "Figure 3: The unwinder string maximum (not minimum) is 30 m."

Reply: It should indeed be minimum because with shorter string lengths the balloon wake becomes more of an issue.

Changes: We have added a note in the text where the twin-sounding rig is described.

[28] Referee #2 writes, "Figure 7: This is not a Skew-T plot, but rather a simple T log-P plot. Please explain the dotted lines. The figure legend refers to a 'parcel path', which doesn't make sense here and shows two dew point lines. I believe the figure tries to overlay both the RS41 and RS92 data. Since the differences are very small, the authors may just show one of the two. The explanation in the text for this figure is sufficient. CAPE and CIN need not be shown as part of the figure, since they are not discussed in the text and add little to the comparison."

Reply: The reviewer is correct, this plot is a T-log P plot rather than a skew-T. The dotted lines were moist adiabats, but we have removed them for simplicity. The figure does indeed overlay the temperature and dew points for both the RS92 and RS41 sondes. We have opted to keep both on the plot because part of the message is how well these observations agree. We have removed the calculated values for CAPE and CIN since they were not relevant.

Changes: We have produced a new Fig. 7 and changed the caption to correctly reflect what is being show, The new figure removes the moist adiabats, corrects the legend and removes the unnecessary information on CAPE/CIN. The new caption reads, "Profiles of dry bulb and dew point temperature from balloon flight #3 which was launched on 03 June 2014 at 17:46 LT. Dry bulb temperature for RS92 (cyan) and RS41 (magenta). Dew point temperature for RS92 (blue) and RS41 (red)."

[29] Referee #2 writes, "Figure 15: Legend refers twice to panel c."

Changes: Corrected.

[30] Referee #2 writes, "Figure 15: The median RH value cannot be 65.2%, if only integer values are used as shown. The median RH of the data shown in table 6 is 65.0%. The median temperature of the data shown is 26.2 C (not 26.5 C)."

Reply: The reviewer is correct. In calculating the medians, more significant digits were used than are presented in the table. These extra significant digits are not needed for the criteria in former Figs 15, 16, 17 (current Figs. 14, 15, 16). We have re-calculated the medians using the numbers in Table 6. Based on the values in the Table, the medians should be: T=26.2oC, RH=65%, SC=41.45%, PWV=3.63.

Changes: The values of the medians have been changed in the captions of former Figs. 15, 16, 17 (current Figs. 14, 15, 16) and in the text.

[31] Referee #2 writes, "Figure 18: Lower left panel should have units in % instead of mb. Pressure units should be written as hPa throughout, not as mb."

Reply: This error was actually in former Fig. 17 (current Fig. 16).

Changes: The units in the lower panel of former Fig. 17 (current Fig. 16) have been changed to hPa. Pressure units have been changed to hPa throughout the document.

Response to comments from Anonymous Referee #3 on "Comparison of Vaisala radiosondes RS41 and RS91 at the ARM Southern Great Plains Site" by M. P. Jensen et al.

[32] Referee #3 writes, "This paper would benefit from discussion about the process for merging an historical RS92 radiosonde measurements series with a future RS41 radiosonde measurement series. For example, if a site is switching from RS92 to RS41, what corrections will need to be made to the RS92 data to homogenise them with the newer RS41 data? I am quite sure that the results from the study are relevant to addressing that question but have not actually been used to do so. Of particular interest would be the need to apply different bias corrections under different cloud cover conditions. This is both particularly challenging but also a strength of this study since it is able to address how different cloud conditions affect the biases between the two sonde types."

Response: While we agree that an exercise to determine: 1) further corrections that need to be done to the RS41 observations, 2) how to homogenize the RS92 data with

RS41 data and 3) how to apply different bias corrections under different cloud cover conditions is important, this is beyond the scope of what we can accomplish with the short dataset we have collected here. There is a need to follow up this short exercise with a study comparing RS41 observations to an independent standard, such as the frost point hygrometer as has been done for the RS92 in Milosevich et al. (2009). There is also a need to collect a much longer dataset of RS41 radiosonde observations at a site like the ARM site in order to statistically characterize the impacts of different cloud conditions on the radisonde observations. We hope that this more detailed study can be part of our future activities.

Changes: None.

[33] Referee #3 writes, "Page 11324, line 10: What exactly is meant by "manufacturer specified tolerances"? Do you mean the manufacturer specified random uncertainties on the measurements?"

Response: By "manufacturer specified tolerances" we mean "manufacturer specified combined uncertainties".

Changes: We have changed the sentence in the abstract to read, "The results suggest that the RS92 and RS41 measurements generally agree within manufacturer specified combined uncertainties of the measurements with notable exceptions when exiting liquid cloud layers where the "wet bulbing" effect appears to be mitigated for several cases in the RS41 observations."

[34] Referee #3 writes, "Page 11324, line 13: With regards to "a smaller impact from solar heating", is that both for temperature and humidity measurements or just temperature measurements?"

Response: The smaller impact is on both temperature and humidity.

Changes: We have changed this sentence in the abstract to read, "The RS41 measurements also appear to show a smaller impact on temperature and humidity measurements from solar heating."

[35] Referee #3 writes, "Page 11324, lines 16-17: The conclusion that "and so a switch to RS41 radiosondes will have little impact on long-term observational records" does not follow from the previous statements. It is possible that while the measurements agree within their random uncertainties, that a switch from RS92 to RS41 could introduce a small systematic bias into the measurement time series that would compromise the long-term record, particularly in regards to trend analyses. As far as I can see, having read the abstract, you have not evaluated the extent to which such biases may compromise a long-term trend. Therefore, I suspect that such a conclusion is not well-founded."

Response: We agree. This conclusion was overstated. In this study, we have only evaluated the differences in RS92 and RS41 radiosondes using the MW31 and MW41 sounding systems with the latest (as of June 2014) software upgrades. Our comparison is only relevant to these operational components for the conditions observed at the SGP site during June 2016. Our results do indicate that under these conditions the differences between these two sonde types is smaller than the manufacturer defined uncertainties in the measurements. For many applications by ARM data users, these differences will have no impact. However, for long-term trend analysis of atmospheric thermodynamics, these differences will be important and more work will need to be done to further characterize and account for these differences. For further discussion of the reconsideration of this conclusion please see the response to comment [3].

Changes: We have changed the text in the abstract and conclusions to better reflect the conclusions that can be made. We note that the differences between the radiosonde types observed during the campaign were smaller than the manufacturer specified uncertainties and so for many applications a switch from RS92 to RS41 will have no impact, but for studies such as long-term trend analysis this difference will be important and need to be further quantified.

[36] Referee #3 writes, "Page 11325, line 19-20: What do the "SGP" and "SG" on the RS92 and RS41 stand for? Are they acronyms for something? I also see in Table 1 "SGPD" instead of "SGP". This is a little confusing. Are the SGP and SGPD sondes different?"

Response: The SG, SGP, and SGPD are product codes. We do understand how they could cause confusion.

Changes: We have limited the use of these product codes to places where the exact definition of the sonde product is necessary.

[37] Referee #3 writes, "Page 11326, line 2: Jensen 2015 is not a peer-reviewed publication and so I would recommend that you do not cite this."

Change: This citation has been removed.

[38] Referee #3 writes, "Page 11326, line 10: The "RS41/MW41" nomenclature is confusing here. What does the "MW41" refer to? Likewise for "RS92/MW31" 2 lines later. I am guessing that the MW stands for Marwin but maybe not all readers will be able to guess at this. "

Reply: We agree that this could be confusing. We have rewritten this sentence to be clear that the MW41 and MW31 refer to the DigiCORA® Sounding System model names.

Change: The sentences have been changed to, "Figure 1 shows a picture of the two radiosonde types and figure 2 the complete system set-up used in the trial. When comparing the radiosonde RS41 and Vaisala DigiCORA® Sounding System MW41 with the older generation RS92 and Vaisala DigiCORA® Sounding System MW31, the new set-up includes improved sensor technologies and easier operational sounding preparations, aimed at higher accuracy and better data consistency in operational radiosoundings."

[39] Referee #3 writes, “Table 4: I think that it is misleading to refer to the "radiosonde pressure sensor manufacturer specifications" for the RS41 sonde given that the sonde does not have a pressure sensor. This needs to be made clearer.”

Reply: We agree that this could be misleading. We have changed the wording to be more specific.

Change: The caption for Table 4 has been changed to, “Table 4. Radiosonde pressure measurement specifications by the manufacturer (based on Table 5 from Jauhiainen et al., 2014).” And within the table we have changed “Sensor Type” to “Measurement Principle.”

[40] Referee #3 writes, “Page 11326, line 24: Replace "The RS41 sensor" with "The RS41 humidity sensor" to avoid confusion.”

Change: Replaced.

[41] Referee #3 writes, “Page 11327, line 2: Replace "The GPS-derived values of the RS41" with "The GPSderived pressure values for the RS41".”

Change: Replaced.

[42] Referee #3 writes, “Page 11327, line 4: Either use mb or hPa for your units of pressure but please don’t switch. I would suggest hPa.”

Change: Have changed all pressure units to hPa. This includes Fig. 7, Fig. 8, former Fig. 13 (current Fig. 11), and former Fig. 16 (current Fig. 15).

[43] Referee #3 writes, “Page 11327, line 6: I would have thought that the directional uncertainty would be a function of the velocity i.e. under strong wind conditions the uncertainty on the wind direction should be smaller than under light wind conditions.”

Reply: The referee is correct. The stated uncertainty in the wind direction applies for wind speed greater than 3 m/s.

Change: We have added text to indicate that this uncertainty is derived for wind speeds in excess of 3 m/s.

[44] Referee #3 writes, “Page 11328, line 24: And this is primarily because the ascent rate is slower with the twin flights right? Though when I look in Table 5, the ascent rates seem pretty close to what you would expect from a single sonde flight.”

Reply: Higher inertia and drag of the payload, result in a more stable flight, which will generally result in less ventilation.

Changes: We have edited the text to describe this better, “In the twin sounding, due to higher inertia and drag of the payload, and thus more stable flight, the sensors generally have slightly less ventilation. A larger payload may also magnify the effects of some error sources, for example, temperature sensor orientation error caused by solar radiation.”

[45] Referee #3 writes, “Page 11329, lines 21-22: This is not a very wide range of surface temperature conditions. I am particularly interested in the RS41 performance when surface temperatures are below freezing but of course you cannot report on this. I will find some other intercomparison papers that hopefully report on flights done at very low temperatures.”

Reply: While this admittedly is not a very wide range of surface temperatures, it was representative of the average range of temperatures generally experienced in north-Central OK during the month of June. According to the National Weather Service Climatology, the average high (low) temperature during the month of June in north-Central, OK is 31(18.3)oC. We refer the referee to the conference paper by Jauhiainen et al. (2014) who present some comparisons done in Finland, Feb. 2014.

Changes: None

[46] Page 11329, lines 25-29: Much of what is said in these five lines is repeated almost verbatim in the caption for Figure 6. I don’t think that this duplication is necessary and

the information should appear either only in the manuscript or in the caption.

Response: We agree that this duplication is unnecessary.

Changes: We removed the description of how occurrence statistics are calculated and the meaning of the vertical black lines from the text. We have removed the description of the ARSCL data product from the figure caption.

[47] Referee #3 writes, “Figure 7: I don’t understand why the dry bulb temperature is labeled as "parcel path" in the legend. The figure caption also needs to explain what CAPE and CIN are.”

Response: The “parcel path” label was a mistake from an edited plotting code. This has been fixed. We have also removed the CAPE/CIN numbers, as they were unnecessary for the point being made.

Changes: This figure has been redone. Not only has the legend been fixed, and CAPE/CIN numbers removed, but the moist adiabats were also removed.

[48] Referee #3 writes, “Figure 8: When looking at this figure I can’t immediately understand why the zonal wind differences are smaller than the meridional wind differences.”

Response: The larger (but still rather small) differences in the meridonal wind differences, particularly in the 5-10 km height range, are due to the wind direction being near 270o (i.e. westerlies) for all of the soundings. At these angles, a small difference in the wind direction will result in a larger difference in the meridional wind component due to the cosine dependence.

Changes: We have added the sentence, “The larger (but still rather small) differences in the meridional wind speeds compared to the zonal wind speeds, particularly in the 5-10 km height range, is the result of the prevailing winds being westerly (near 270o) at these heights, where the cosine dependence of the meridional wind has the largest rate of change and so a small difference in wind direction will propagate to a larger difference in the wind speed.”

[49] Referee #3 writes, “Page 11330, line 15: I think that you have to be careful in your terminology. When you say "accuracy" I interpret this as "the systematic error" whereas what is being reported in Table 4 is the uncertainty. Shouldn’t you be referring to "precision" rather than "accuracy"?”

Response: We believe the referee meant line 25. The referee is correct, that accuracy is not the correct term here, but precision is not applicable either.

Changes: We have changed the sentence to read “These differences are well within the given combined uncertainties of the radiosonde models (see Table 4) and are consistent with the results of Motl (2014)

[50] Referee #3 writes, “Page 11331, line 2: Coming back to the point that I raised earlier: If there was a systematic 0.13_C temperature difference between RS92 and RS41 radiosondes this would certainly compromise the homogeneity of the long-term temperature climate data record for trend analyses.”

See response and changes for comment [3]

[51] Referee #3 writes, “Page 11333, line 22: Yes but a 1 hPa pressure difference at 25 km is far more concerning than a 1 hPa pressure difference at 2 km altitude.”

Reply: Agreed. The observed differences are much smaller than 1 mb, and are much, much smaller (very close to zero) than that above 13 km, where the average pressure is in the 150-200 hPa range.

Changes: None.

[52] Referee #3 writes, “Figure 17: The bottom left X axis is labelled incorrectly.”

Reply: The reviewer is correct. The units on the bottom left x-axis should be hPa.

Changes: We have changed the units to hPa.

[53] Referee #3 writes, “Figure 19: Why are there no uncertainty bars on the RS92 and

RS41 PWV measurements?

Response: The “error bars” on the MWR PWV do not represent a traditional uncertainty, but instead represent the range of PWV observed by the MWR over the half hour following the launch of the radiosonde. We chose the first half hour because this should be a reasonable representation of the variability that the radiosonde may experience as it traverses the lower atmosphere and drifts away from the MWR site. We have not included uncertainty bars on the radiosonde PWV, because we cannot easily quantify the variability as the balloon drifts away from the site, which is likely a larger contribution to the uncertainty than the instrument uncertainties. See also response to comment [17].

Changes: See changes for comment [17].

[54] Page 11337, lines 16-17: Again I don’t think that this is a robust conclusion based on the results that you have presented. You have not shown that a small e.g. 0.1_C systematic bias between the RS41 and RS92 radiosondes would not compromise a 20 year trend analysis where the first 10 years of measurements were made using an RS92 system and the last 10 years with an RS41 system.

See response and changes for comment [3]

GRAMMAR AND TYPOGRAPHICAL ERRORS

[55] Page 11327, line 17: Replace "On the contrary" with "In contrast".

Change: Replaced.

[56] Page 11328, line 17: Should this be "UW1-30 ozonesonde unwinder"?

Reply: Correct, this should be the UW1-30 ozonesonde unwinder.

Changes: “ozone” was changed to “ozonesonde.”

[57] Page 11331, line 11: Replace "again noticeable" with "again a noticeable".

Change: Replaced.

[58] Page 11335, line 14: Replace "There is some different behaviors" with "There are some different behaviors".

Change: Replaced.

Response to comments from M. Fujiwara on “Comparison of Vaisala radiosondes RS41 and RS91 at the ARM Southern Great Plains Site” by M. P. Jensen et al.

[59] M. Fujiwara writes, “The authors conclude in Section 5, "... but under most observational conditions, the RS41 and RS92 measurements agree to within the manufacturer specified limits and so a switch to RS41 radiosondes will have little impact on long-term observational records." However, Figure 8 (and other figures) shows characteristic profiles of temperature and relative humidity (RH) differences: RS92 temperature is _0.05 C warmer than RS41 temperature around 3 km and is _0.1 C colder around 9 km; and RS92 RH is always smaller than RS41 RH in the troposphere, with its peak of _2%RH around 10km. I am not sure whether 20 comparisons give statistically significant (and robust) results, but it seems to me that these results are rather robust at least at this site during June 2014 (and for the production batches for these radiosondes). “

See response and changes for comment [3]

[60] M. Fujiwara writes, “ If these results are robust for all RS92 and RS41 data regardless of location, time/season, and batches, the climate community needs to take these into account when analyzing long-term variability (or when homogenizing the time series). Furthermore, the pressure differences are very important, if they exist, because the climate community usually uses pressure as the vertical coordinate; even if there is no difference in temperature/RH measurements, the difference in pressure measurements would create temperature/RH differences where temperature/RH has vertical gradients. Therefore, it would be very useful to provide a summary table for

the "profile" differences, not only for "individual sensor" differences. See, for example, p. 928 of Kobayashi et al. (2012), who discussed "simultaneous sensor comparison" versus "comparison on pressure levels."

Responses: The comparison protocol we have followed is similar to the “simultaneous sensor comparison” used by Kobayashi et al. (2012). This technique gives the most independent comparison of the measurements that are taken. We also did comparisons with respect to height, but the convolution of the radiosonde height estimation with the temperature/humidity/wind observations made the results more difficult to interpret. The same would be true using pressure as the vertical coordinate and the convolution of the pressure differences. This could have the result of masking differences due to compensating errors, something we want to avoid in this study.

Changes: None.

[61] M. Fujiwara writes, “My questions are related to the reasons for the above mentioned temperature/RH differences. The authors discuss (p.11332, lines 13-16), "The relative peaks in the temperature and relative humidity differences near a height of 10km are likely related to a combination of increased radiative heating of sensors due to contributions from cloud albedo, "wet-bulbing" effects and sensor response time in regions of strong gradients as the sondes traverse cloud layers." Is there also a possibility that the factory calibration procedure could play a role? The choice of calibration temperature/humidity points and the choice of fitting curve function might play a role. “

Response: In general the observed median temperature difference around 10 km is about 0.1oC only, thus, for such a small difference it is possible to name several potential reasons. For RS41 the calibration uncertainty of the temperature measurement is less than 0.05oC over the practical measurement range, and the RS92 is not significantly less accurate. Uncertainty in humidity calibration increases at cold conditions, thus a fraction of the observed average difference of < 2.5 % RH may be related to calibration. Yet, the effects of varying solar radiation on RS92 humidity measurements

are more likely to induce more significant errors than the calibration uncertainties.

Changes: We have added sensor calibration as a possible contributor to observed differences.

[62] M. Fujiwara writes, “If this is not the case (i.e., these factors would not create the above mentioned magnitudes of differences), then how about the possibility that the sensor supporting structure might cause heat contamination to the temperature and RH sensors differently for RS92 and RS41? “

Response: Ventilation at the altitude of around 10 km {pressure 300-250 hPa) is still pretty good, and thus, the heat contamination is not likely to be a major factor. At low pressures (10 hPa or so) the RS92 temperature measurement may suffer fro spikes, to our understanding caused by the frame structure around the sensor.

Changes: None

[63] M. Fujiwara writes, “Also, RS92 has two RH sensors heated alternatively to reduce the sensor icing and thus to reduce wet-bulbing effects, while RS41 only has one RH sensor. Why does the RS41 is less prone to wet-bulbing effects (p. 11331, lines 27-28)? “

Response: The sequential pulse heating method of the RS92 with substantially long sensor element heating and cooling (non-heating) periods may not be sufficient to eliminate sensor ice/wetting in some in-cloud conditions.

Changes: We have added this description of the role of the RS92 sequential pulse heating to the text.

[64] M. Fujiwara writes, “But, Anonymous Referee #2 pointed out that RS41 also has a heating mechanism. Please give more detailed descriptions about both RS92 and RS41 RH sensors and their heating mechanisms.”

See response and changes for comment [4]

[65} M. Fujiwara writes, “Also, let me point out that for RS41 which uses GPS to measure pressure, temperature and RH errors, if they exist, would propagate to the pressure data. Does this explain (at least part of) the RS92-RS41 pressure differences? “

Response: According to Vaisala uncertainty analysis, the uncertainty in GPS height measurements is the dominant source of error in GPS-based pressure measurements in the lower troposphere (where the largest deviations were observed in this intercomparison). The temperature measurement uncertainty may also have a small effect, while we estimate the relative humidity effect to be negligible.

Changes: None.

[66] M. Fujiwara writes, “But, it looks from Anonymous Referee #2’s review that the authors should check the pressure ground check procedure (or the treatment of surface pressure data in the data processing) first.”

See response and changes for comment [1]

[67] Finally, I agree with Anonymous Referee #2 that the authors should also analyze GRUAN RS92 data product so that we can clarify whether the differences they found come from the correction algorithms for RS92 or not.

See response and changes for comment [3].

Interactive comment on Atmos. Meas. Tech. Discuss., 8, 11323, 2015.

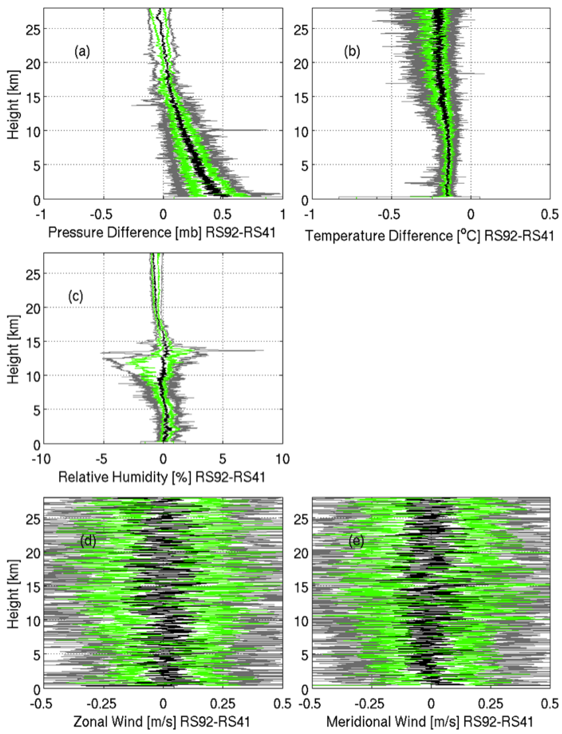

Fig. 1. Figure A Vertical profiles of the median (black), 25/75th percentile (green) and 10/90th percentile (grey) differences between GRUAN- processed RS92 and original RS41 observations (RS92-RS41)

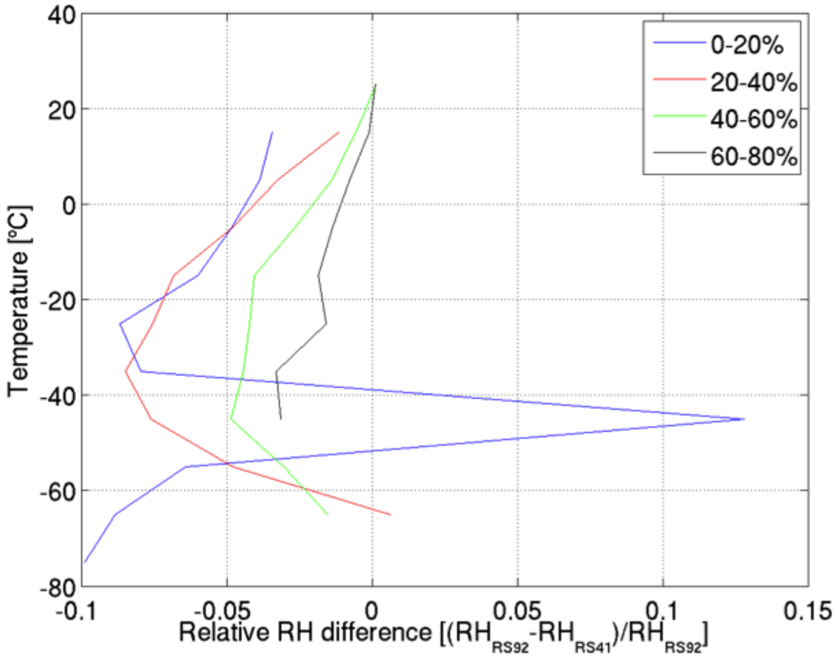

Fig. 2. Figure B Relative difference in relative humidity between the RS92 and RS41 sondes as a function of temperature for four difference relative humidity ranges.