Atmos. Meas. Tech. Discuss., 4, C1644–C1653, 2011
www.atmos-meas-tech-discuss.net/4/C1644/2011/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Comparisons of temperature, pressure and humidity measurements by balloon-borne radiosondes and frost point hygrometers during MOHAVE 2009" *by* D. F. Hurst et al.

D. F. Hurst et al.

dale.hurst@noaa.gov

Received and published: 26 September 2011

Replies to Reviewer #1:

{Reviewer Comment} The authors include anomalous profiles in their determination of uncertainties; when trying to obtain information about the level of instrument quality and uncertainty shouldn't the best available data be used. Is it the basic premise of uncertainty analysis to examine homogeneous objects? The paper assumes the objects

analyzed are alike. Only the RS92 instruments are the same, i.e., same configuration, same sensors, and same software. IMET sensors are very different than the RS92, the instrument is a different size and configuration, and the processing used two different software packages, not the manufacturers software. The authors should consider justifying why the uncertainty analysis between RS92 and IMET radiosondes is valid in view of the extremely different character of these radiosondes. The discussion concentrates on differences, but RS92 temperature measurements are corrected while IMET does not apply corrections. Wouldn't results using corrected and uncorrected temperature differences be bogus.

{Author Reply} This paper strives to provide critical information for those who launch radiosondes, and especially for those who assume that one sonde is as good as the next. The results clearly demonstrate this is not the case. Not only are there important differences between the temperature and pressure measured by the RS92 and IMET during the majority of flights, there are flights where one sonde produces anomalous data relative to the majority of flight data. We consider it vital to communicate this finding to the radiosonde users, raising their awareness that some sondes simply do not perform within their manufacturer's specifications. This may seem obvious to someone with extensive radiosonde experience, but there are probably many potential readers who are not aware of the intermittent, significant problems with radiosondes.

Point of clarification: yes, the nighttime RS92 temperature measurements are "corrected" for solar radiation effects by the DigiCORA software, by a maximum of 0.01°C (according to Vaisala Table RSN2005), but this is insignificantly small. Temperature data for the two daytime RS92 flights are more substantially corrected by the Vaisala software. Since this manuscript was submitted back in April, solar radiation corrections for the IMET temperature measurements have become available. We have now applied these corrections to the IMET temperature data for the two daytime flights (TF027, TF040) and recomputed the RS92-IMET differences and statistics for T and RH. I will make it perfectly clear in the manuscript which temperature data were corrected and which were not.

{RC} Eight of 26 profiles stood out that the authors considered being anomalous. What purpose is achieved by including poor data? Justification for including anomalous data is not found in the Introduction of elsewhere in the text. The authors should justify why including the anomalous profiles is important. Eight profiles are 30 percent of all profiles, a relatively large portion.

{AR} The purpose of including the anomalous data is clearly explained in our reply to your previous comment. A brief justification of why it is important to retain and show the anomalous data will be added to the manuscript.

{RC} The unfortunate lack of exact timing regarding the release of RS92 radiosondes (2- second time hacks) and the IMET radiosonde (1-3 second non-uniform time hacks) makes the fidelity of the match up procedure suspect. The authors depended on the software to obtained release times that provided ill-defined precision. Were he actual times of instrument release visually noted? These would be the basis for using elapsed times. Simultaneous comparison between two and more instruments depends on precise time match up which is not apparent from reading section 2.3. From the discussion given, time stamping as described does not seem to have been very successful and may have led to relatively poor profile synchronization. The paper notes the RS92-SGP and IMET radiosondes have GPS capability. GPS times would have been a more precise method with which to align the profiles. Although the RS92-K radiosonde did not have GPS capability better match ups probably could have been accomplished. A short explanation why the available GPS times were not utilized would be useful, otherwise why mention the GPS capability. Neither were the number of K-type instruments not mentioned in the text or tables.

{AR} There weren't enough RS92-SGP sondes flown with the IMET sondes to synchronize the data for every flight using GPS time. Since the sondes were physically affixed to the same payload, at the same height, they were launched at the same exact

time. The launch time mismatch occurs because the two different receiving systems employ slightly different algorithms to detect the moment of launch (t=0), at which point the system begins recording elapsed times (since launch) and flight data. The method employed to synchronize the two data streams is very robust and greatly improves the temporal correlations of temperature structures observed by the 2+ sondes on a given balloon. Shifting the RS92 elapsed timestamps by 2 seconds at a time relative to the IMET timestamps, until arriving at the best correlation between temperatures measured by the two sondes, yields a potential maximum temporal mismatch of ± 1 second between the two data streams. This temporal matching is as close as can be obtained given the 2-s RS92 data timestamps.

{RC} Mistakes and typos should have been noted and removed by the authors before sub- mitting the paper. One example is found on line 18 of section 1, Introduction. The date given for MOHAVE 09 should have been 2009, not 2011. Another significant mistake is found on line 13 of section 2.1 Radiosonde. Vaisala's pressure sensor is not piezo resistive, it is a silicone capacitive sensor.

{AR} Corrected.

{RC} Page 4362 Line 23. The authors noted that the IMET radiosonde did not have temperature corrections but should also have mentioned the RS92 radiosonde has temperature corrections, although its nighttime corrections are small. The authors might mention whether they took into account the possible influence of these errors on the RS92-IMET differences?

{AR} It will be made perfectly clear in the manuscript which RS92 and IMET temperature measurements were or were not corrected, now that the two IMET daytime flights have been corrected. The RS92 nighttime corrections (maximum 0.01 °C) are insignificant and have no impact on the comparison results.

{RC} Page 4363 Data Matching section. A clearer description of the time stamp method is needed. Data profile matching is the key ingredient for a valid analysis and should be clearly described, including positive and negative issues. The authors claim the time stamps to be' most reliable' (line 23) but continue the discussion with an explanation of the offsets encountered. After carrying out the match up process, average time differences of 2.5±6.0 seconds remain. This much difference compromises the match up fidelity and constitutes a flaw in the analysis. The authors should justify how were these differences actually were handled.

{AR} The 2.5±6.0 seconds is the average of time shifts necessary to synchronize the RS92 and IMET data streams (that were not naturally synchronized because of the launch time "mismatch" described above), not the average time stamp difference that remains after the synchronization process. The potential maximum time mismatch between the two data streams is within ±1 second.

{RC} Matching profiles using correlation seems to be a good idea but did not recover 100 percent of the profiles. A secondary method (page 4364 lines 11-17) was used to match profiles, but use of this method contradicts the statement made earlier (page 4363 lines 23-24). There is a lack of match up consistency that points to arbitrary profile alignment. The authors need to mention why GPS was not used?

{AR} The two matching methods generally indicate the same required shifts in the elapsed time stamps of the RS92 data. For 15% of the flights the correlation method does not provide a clear and definitive maximum correlation coefficient, meaning it cannot choose between two sequential timestamp shifts (that differ by 2 seconds) that yield the same maximum correlation coefficient. The correct choice of time stamp shift for each flight is crucial to the quality of the data comparison, so visual inspection of the two temperature measurement time series was used as a supplemental method of determining the correct time shift for these 15% of cases. The "contradiction" you report stems from our poor choice of the word "match" in Line 22 Page 4363 and the confusing statement that follows. This will be clarified. As for GPS, please see our reply above about the low number of RS92 sondes with GPS capabilities.

{RC} Page 4365 Line 19 to end of paragraph. The method used to identify some of the profiles as being anomalous seems arbitrary. If some (number needed) profiles are indeed anomalous the authors should explain why they were not totally excluded from the analysis or for that matter, included? Furthermore, eight anomalous profiles is a fairly large portion of the total measurements, thus a further reason to remove them.

{AR} Again, this paper aims to demonstrate that all radiosondes are not created equal, and that sonde quality can vary substantially. By clearly identifying the anomalous profiles (as defined in Line 25, Page 4365) in graphs, and by providing separate statistics of the sonde measurement differences for all profiles, including those deemed anomalous, and for only the non-anomalous profiles, we accomplish several objectives. The first has already been stated in the first line of this paragraph. Second is the "reality" of including statistics from all sonde data in the comparison, i.e., this is what you get if no sondes are rejected. Third is that the "majority" statistics, obtained when the poorly performing sondes are excluded, provide the best numbers to evaluate measurement biases between "typical" RS92 and IMET sounding data.

{RC} The authors use median differences, which reduces the data ensemble to one point per kilometer per radiosonde pair, while arguing the median is better than using averages because it removes random differences. Were the median points tested for randomness? It is true using the medians reduce calculations to a minimum while eliminating the effect of extreme differences but the question is whether or not extreme differences should be removed or allowed to remain. How was this decision reached? Were the data normally distributed and were outliers removed before selecting the median?

{AR} The use of median differences instead of average differences prevents a few large differences, typically the result of sporadic sensor noise, from disproportionately influencing the difference statistics. Figure 1, with visual examples of this sporadic noise, shows the potential for average differences in some altitude bins to be significantly skewed by this noise. Another important reason to employ median instead average differences is that averages assume a normal distribution of the data and would necessarily require the potentially subjective identification and removal of statistical outliers.

{RC} Page 4366 Lines 16-25. The authors need to be clear when discussing uncertainty, reproducibility, and accuracy. In fact, definitions would help. Furthermore, RS92-RS92 uncertainty differences are quoted for pressure regimes 1080-100 hPa and 100-3 hPa and reproducibility is given for three layers: 1080-100 hPa; 100-20 hPa; and 20-3 hPa. The authors have mistakenly transformed the manufacturer-quoted reproducibility pressure boundaries to 400 hPa and 100 hPa. Why 400 hPa when it is not a boundary given earlier in the text or by Vaisala?

{AR} You are correct that the RS92 reproducibility values for temperature measurements should be different in three different pressure regimes: 3-20, 20-100 and 100-1080 hPa, while the pressure reproducibility values differ in the 3-100 and 100-1080 hPa layers. The 400 hPa is only a boundary for IMET pressure uncertainties and should therefore only be considered when combining the IMET and RS92 pressure uncertainties. A table presenting the manufacturers' measurement uncertainties and the combined uncertainties of measurement differences will be added to the manuscript to improve the clarity of this discussion and to remove some "wordiness" of the text. In addition, a table presenting the percentages of measurement differences that exceed the combined uncertainties (i.e., "excessive differences") will be added, further reducing "wordiness" of the text.

{RC} Page 4367 Line 3. If the TF028 and TB028b profiles are well inside the uncertainty limits the authors should explain why they are considered anomalous? In fact, agreement between these profiles is the best of the ensemble.

{AR} As defined in the manuscript (Line 25, Page 4365), "anomalous" describes profiles that do not conform to the majority of profiles. Though these two difference profiles are inside the combined uncertainty limits and show the best RS92-IMET temperature agreement, they do not conform to the majority of profiles.

{RC} Page 4368 Line 18. If it is clear from Figure 3 that the two anomalous profiles will skew the statistics negatively, why were they included in the analysis?

{AR} The two anomalous do expand the standard deviations and negative skew the difference statistics when all the profiles are included, but have no impact on the difference statistics when the anomalous profiles are excluded. Both sets of statistics are provided in the CDFs and Tables for this very reason.

{RC} Page 4369 Paragraph beginning with line 6. The anomalous differences near the surface suggest time differences between the instrument profiles may be the reason for such large separations (See previous comment page 4363). The radiosonde operator should have noted pre-release discrepancies and followed the manufacturers' instructions by rejecting the radiosondes.

{AR} Though not intentional, we launched some radiosondes that should have been rejected based on pre-launch checkout criteria. The anomalous difference profiles that were produced, including significant differences at the lower altitudes, are a result of poor sensor calibration and/or performance, not of poor data synchronization with the other radiosonde(s). This strongly reinforces the importance of performing pre-launch checks of sonde sensor quality and, if need be, the outright rejection of certain sondes.

{RC} Page 4369 Line 15. I may have missed a definition for 'conforming' and 'non conforming' earlier in the paper. These are subjective terms, please define or include limits, if any.

{AR} Conform ("to be similar in form or character") is a widely used word that we don't think requires a definition. There are no quantitative limits involved in selecting "anomalous" profiles – this task was performed through close examination of the median difference profiles. Though subjective (i.e., not strictly quantitative), this procedure was very straightforward and unambiguous.

{RC} Page 4370 Line 15. At 16 km ($\sim$100 hPa) $\pm$0.78 hPa is an uncertainty of approximately 0.8 percent while at 32 km (∼10 hPa) this uncertainty limit is 8.0 percent, in the first instance a calculated height error is ∼50 meters, the second instance the height error is about ∼530 meters.

{AR} It is clear from this sentence that the RS92-IMET pressure differences exceed their combined uncertainty limits more often than not at the highest altitudes because the limits shrink from ±1.12 hPa (7-16 km , ∼100-400 hPa,) to ±0.781 hPa above 16 km (<100 hPa). If your comment is only about the use of the word "smallest" being incongruent with our consistent use of absolute uncertainty limits, the word will be replaced with one that is more suitable.

{RC} Page 4370 Line 27. This seems to be another instance when the radiosonde operator should have noted large differences prior to balloon release. If acceptance/release criteria were not met prior to balloon release should those profiles that include large surface anomalous data records be removed from the data set?

{AR} Again, some sondes that should have been rejected were inadvertently launched. This is the realistic outcome of assuming that all sonde sensors will perform well and that pre-launch checks are unnecessary.

{RC} Page 4373 Lines 22 through 27. The second sentence contradicts the first. Sentence beginning with 'A big temperature difference . . . is ok, but the following sentence does not follow. Unless the atmospheric temperatures were unusual wouldn't the temperature at 4-7 km be higher (warmer) than temperature near 10-12 km?

{AR} These few lines of text will be removed because they only detract from what should be a very simple argument – that RH values above 20 km are so small that the absolute differences between RH measurements by different sensors can never be large.

{RC} Page 4377 Paragraph at line 19. I disagree with this statement and the next. Considering each level has a different pressure and/or temperature bias it is possible

the total error would be small or none at all. However, a biased pressure measurement at the surface of 1 hPa (i.e., an offset) results in approximately 670-meter error at 10 hPa given the temperature measurement is accurate. Therefore, two instruments with different pressure offsets at the surface would not have a constant difference with altitude.

{AR} The geopotential height calculation algorithm we use for IMET sondes, and we believe is used by Vaisala for their RS92 sondes, anchors the launch altitude as its actual value (based on GPS or survey elevation), and is unaffected by a bias in the surface pressure measured by the radiosonde. Once the launch altitude is established, all subsequent geopotential heights are calculated incrementally from the small pressure, temperature and RH changes that occur between one time stamp and the next. If one radiosonde is uniformly biased by 1 hPa throughout the total sounding and one radiosonde has no bias, the geopotential heights calculated from these sondes will be exactly the same if they both have the same response function for changes in pressure (i.e., same slope but one has y-intercept of 1 hPa). On a standard altitude vs. pressure chart you are correct, a 1hPa bias in pressure at 10hPa represents a nearly 0.7 km difference in altitude. But our geopotential height calculation algorithm does not employ such a "lookup" chart, only measured incremental changes, hence if two radiosondes record the exact same changes they will report the exact same geopotential altitudes, even if one sonde has uniformly biased pressure and the other has no pressure bias.

Interactive comment on Atmos. Meas. Tech. Discuss., 4, 4357, 2011.