Date: February 22, 2018

Manuscript #: amt-2017-486
Manuscript title: ***Evaluating tropospheric humidity from GPS radio occultation, radiosonde, and AIRS from high-resolution time series***

Brief Summary of the Manuscript
This manuscript studies the statistical differences of specific humidity among numerous data sets, including GPS radio occultations from various processing centers, AIRS, ERA-Interim, and radiosondes. The analysis is performed over specific geographic locations, which are characterized by different environmental conditions. The authors use a year-worth of data, in 2007, so that they have an increased number of GPS radio occultation sampling. The advantage of this investigation over previous research is that it quantifies the statistical differences of specific humidity among all the different data sets in short temporal scales and not in an ensemble climatological mean. The manuscript contains useful results that are worth publishing in the journal of Atmospheric Measurement Techniques, only after major revisions, according to the guidelines provided below.

Comments:

1) **Page 3, Line 24.** The objectives are too general and slightly confusing. Please, re-write this section. For example, I am not sure what "... quantify RO humidity retrievals..." or "...quantify how these RO humidity..." mean. I recommend be specific on the objectives, as you describe them in the next paragraph.

2) **Section 2.1.** I recommend including the specific humidity accuracy for each center, including appropriate references.

3) **Section 3.1, Figure 2.** The authors should provide more discussion on the observed differences among the data sets, with respect to ERA-Interim, and try to provide an explanation of where these differences may be coming from. E.g., discuss the 1000-400 hPa dry bias seen in UCAR 1D-Var, JPL direct, and GFS. Also, at 200-100 hPa, the 1D-Var approach (UCAR and WEGC) show an excellent agreement with ERA-Interim, whereas the rest of the data sets show much larger deviations. The authors should describe that and try to explain the observed behavior. Additionally, the standard deviations and RMS of all the regressions could provide a quantitative insight on the degree of agreement with respect to ERA-Interim.

4) **Section 3.2, Page 12.** Similar to the previous comment, it would be nice if the authors could provide a more detailed discussion on the differences of each data set with respect to ERA-Interim at the GUAM location and explain these differences within the context of the environmental conditions over this region and within the context of how each profile has been retrieved by each center.

5) **Page 12, Lines 14-15.** I find this statement a bit strong, because it may exclude other possible sources that could cause the observed discrepancies. Could it also be that ERA-Interim may be overestimating the degree of entrainment above the PBL, thus introducing more water vapor in the free troposphere aloft? Have the authors considered convection as another possible source of such discrepancies? Also, it should read: "Figure 5 shows that all data sets are dry-biased with respect to..." In Figure 5, the title of the color bar is missing.

6) **Figures 6 and 7.** The title is missing from the x-axis. Also, the definition of ND assumes than ND is unitless, yet the differences here are given as percentage. Either modify the definition of the ND, or redo the figure accordingly. Perhaps, add 2-3 sentences to explain to the reader what the smaller RMS values physically mean in the radio occultations (e.g., smaller scatter, steadier daily variability, better long-term robustness, better accuracy... something along these lines, so that the reader attach a physical interpretation to the results). This way, I believe the analysis becomes clearer.

7) **Page 16, Line 2.** It has not become clear from the discussion so far how different atmospheric conditions could influence the data biases. It would be great if there were a transitioning paragraph before Section 4 that summarizes in 3 sentences the conclusions of Section 3. I believe this would be a smooth transition.

8) **Figure 8.** At 800-400 hPa over Guam, at 1000-800 hPa over Mina, and at 800-400 hPa over Mina, the three orange asterisks that indicate values about -21%, -18%, and -35%, respectively, could they be outliers? Also, it would be good to include the latitude and the longitude of each station at the beginning.

9) **Page 17, Line 5.** Could the observed AIRS dry bias be due to cloud-contaminated radiances? And therefore, the AIRS statistical differences might be statistically insignificant? I think it would be good if this were mentioned in the interpretation of the results for completeness, unless only AIRS cloud-cleared pixels are used in the analysis.

10) **Page 17, Line 29.** Any physical explanation as to why the signal is strongest in the layers above 600 hPa?

11) **Page 18, Lines 4-6.** Although this may be true, there have been no results showing refractivity variations. Therefore, aren't these lines out of context? What purpose do they serve? On another note, I would think that due to deep convection within the eye and eyewall regions of a typhoon, together with the water vapor entrainment and vertical overshooting from the well-mixed moist layer that sits at the bottom of the typhoon, there would be an increase in the water vapor concentration in the free troposphere that could lead to refractivity increase high up based on Eq. (1). Perhaps, this could be an explanation to my previous comment above?

12) **Section 6.** Why the inter-center mean and not the GUAM sub-data set? Is it because there is no "true" RO data set, and thus an inter-center mean is regarded more realistic? But then again, wouldn't the inter-center mean smooth out differences? Why not use the GUAM radiosonde data set as the "true" and redo Figure 10? Would this change the conclusions of the analysis?

13) **Page 19, Line 3.** You mean the absolute value of the mean deviation?

14) **Figure 10.** More discussion is required in the refractivity analysis. E.g., over GUAM, the JPL and WEGC refractivity differences are systematically larger than the inter-center mean above about 800 hPa and the JPL refractivity difference is larger even down to 950 hPa. Over the Japanese stations, the JPL refractivity difference is systematically larger than the inter-center mean. Any explanation as to why these may be? Could these be associated with the different environmental conditions over GUAM and the Japanese stations? Additionally, the figure caption in Figure 10 needs fixing at the part where the authors describe what line represents each data set.

15) **Section 7.** I feel that the conclusion section needs re-writing, in order to focus on the objectives of this investigation. For example, these eight concluding remarks could be summarized into one single paragraph and then a second paragraph should describe the findings of this investigation regarding the behavior of the radio occultations within the context of the: a) wet vs dry conditions, b) typhoon passages, and c) different geographic locations.