Review of Franklin et al. (2016, ACP) Revised Manuscript, 5/18/16

Thank you for considering many of my suggested revisions and making substantial changes to the manuscript. I believe that the manuscript is much improved over the initial submission, and I hope that you feel this way as well. I still have a few major comments (the most important being #4-5) and several minor comments that I would like addressed before the manuscript is published, as described in further detail below.

Major Comments

1. I still can't make sense of the argument presented as a response to major point 5a in my first review. If larger, faster-falling particles in the generic ice PSD simulation are removing more LWC via riming, I would expect to see more graupel, but Figure 13 show that graupel mass is not correlated with the RH aloft in Figure 4. I'd also expect supercooled cloud water to be lower if it is indeed being collected more efficiently by the larger, faster falling ice, but again, Figure 13 does not show that. Additionally, if supersaturation remains higher in a simulation, I would expect lesser latent heating since a larger supersaturation implies less overall condensation. Whether heating occurs via condensation+riming or via deposition, the overall heating is the same as long as the mass that has changed phase remains the same. If larger particles were being offloaded more efficiently to reduce water loading impacts on buoyancy, then I would expect to see less ice in the simulations with the highest RH aloft, but in fact those simulations have the most ice aloft in Figure 13. Lastly, if the updrafts were theoretically stronger in the simulations with higher RH aloft, then I would expect to see that in Figure 11, but there isn't a correlation between the strongest vertical velocities aloft in Figure 11 and RH aloft in Figure 4. Therefore, I don't see any evidence for the argument presented (Pg. 11, lines 16-20). This is not clear from Figure 5 either, where, for example, "eg" and "qcf2rainfreeze" have very similar updraft mass divergence profiles, but their relative humidities aloft are far different. Therefore, I suggest more evidence be shown to support the argument presented.

2. I don't follow the argument presented in response to major point 5b in my first review (presented on pg. 11 line 29 to pg. 12 line 21). Figure 5 b compares "3d" and "qcf2", which do clearly differ in Fig. 5c, but if "qcf2noqgr" is selected instead of "qcf2", its distribution of buoyancy in Figure 5c is the same as "3d", so how can the general conclusion that smaller ice leads to more entrainment and lower theta-e aloft be true? Additionally, "3d" with greater turbulent mixing is supposed to have more entrainment with lower theta-e than "nd", but the buoyancy (delta theta-d) at 6 km is greater in "3d" than "nd" and they only slightly differ in buoyancy at 14 km, so greater entrainment doesn't seem to lead to less buoyant updrafts that don't penetrate as high. The deleted figure that showed cloud top height and OLR also showed little difference between "3d" and "nd". I think more evidence needs to be presented to support the arguments presented in this section of the manuscript.

3. The reasoning on pg. 19, lines 1-3 that the 23-24Z maximum reflectivity profile differences for the simulations that use the generic ice PSD are a result of differing entrainment and water loading alone caused by differences in the model schemes doesn't seem to be the full explanation. If it was, then I'd expect to see larger differences for the 17-18Z period as well. My guess is that the location of the system and convection differs between the simulations by late in the simulation (23-24Z), and this may also be impacting the statistics, especially since the differences are so vast for the 23-24Z period in Figures 10b and 11b, but so small for the 17-18Z period in Figure 10a. Has the different evolution of the system and convection in different simulations in the domain considered for statistics in Figures 10-11 been examined? Is this a possible contributor to differences between simulations in Figures 10-11?

4. In Figure 11, the aircraft observed maximum vertical velocity profile is incorrect. A correct profile at 5-second (~750-m) resolution is attached as Figure 1. Therefore, the maximum 1-second profile in Figure 11 (solid black) should be a bit higher velocities and the dashed black profile average to 1 km resolution should be slightly less than the profile shown in the attached Figure 1. Therefore, I suggest reviewing the data again and correcting Figure 11.

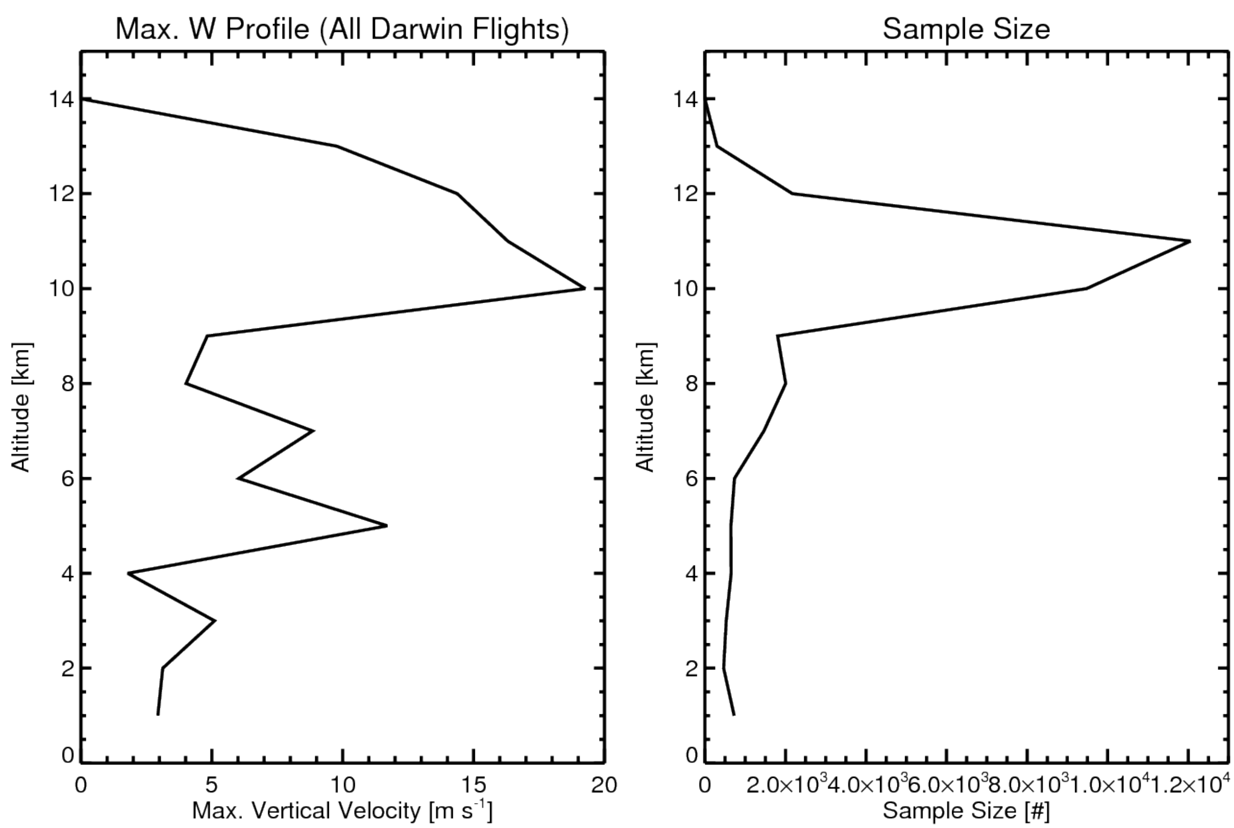

Figure 1. The maximum vertical velocity profile observed by the Falcon-20 aircraft during the HAIC-HIWC Darwin campaign including all flights at 5-second (~750-m) resolution with aircraft roll angles > 5° removed (left). The sample size is shown on the right.

5. In Figure 12, the aircraft mean TWC as a function of vertical velocity and temperature range is incorrect. The correct statistics are attached as Figure 2. There are very few samples between 0 and -10°C. There are a few more between -10 and -20°C, which show that TWC (~IWC) increases quickly to 2 g m^{-3} between negligible vertical velocities and a few m/s. It then changes slope and increase to something like 2.5 g m^{-3}, which is different than Figure 12. There aren't vertical velocities stronger than 10 m/s in this range even though Figure 12 shows that. This is similar to what the simulations do except that they have a bit higher IWCs. There are far more samples between -20 and -30°C, which show a similar relationship to -10 to -20°C for updrafts. For downdrafts, the TWC increase linearly to 2 g m^{-3} between 0 and -4 m/s, which is very much like some of the simulations in Figure 12. The slightly higher IWCs in the simulations in updrafts between -20 and -30°C could be because updrafts are more intense in the model, but it could also be because aircraft did not sample the most intense cells with lightning or high reflectivity. I suggest reviewing the data and correcting Figure 12. This may also change the conclusions that you drew from the figure.

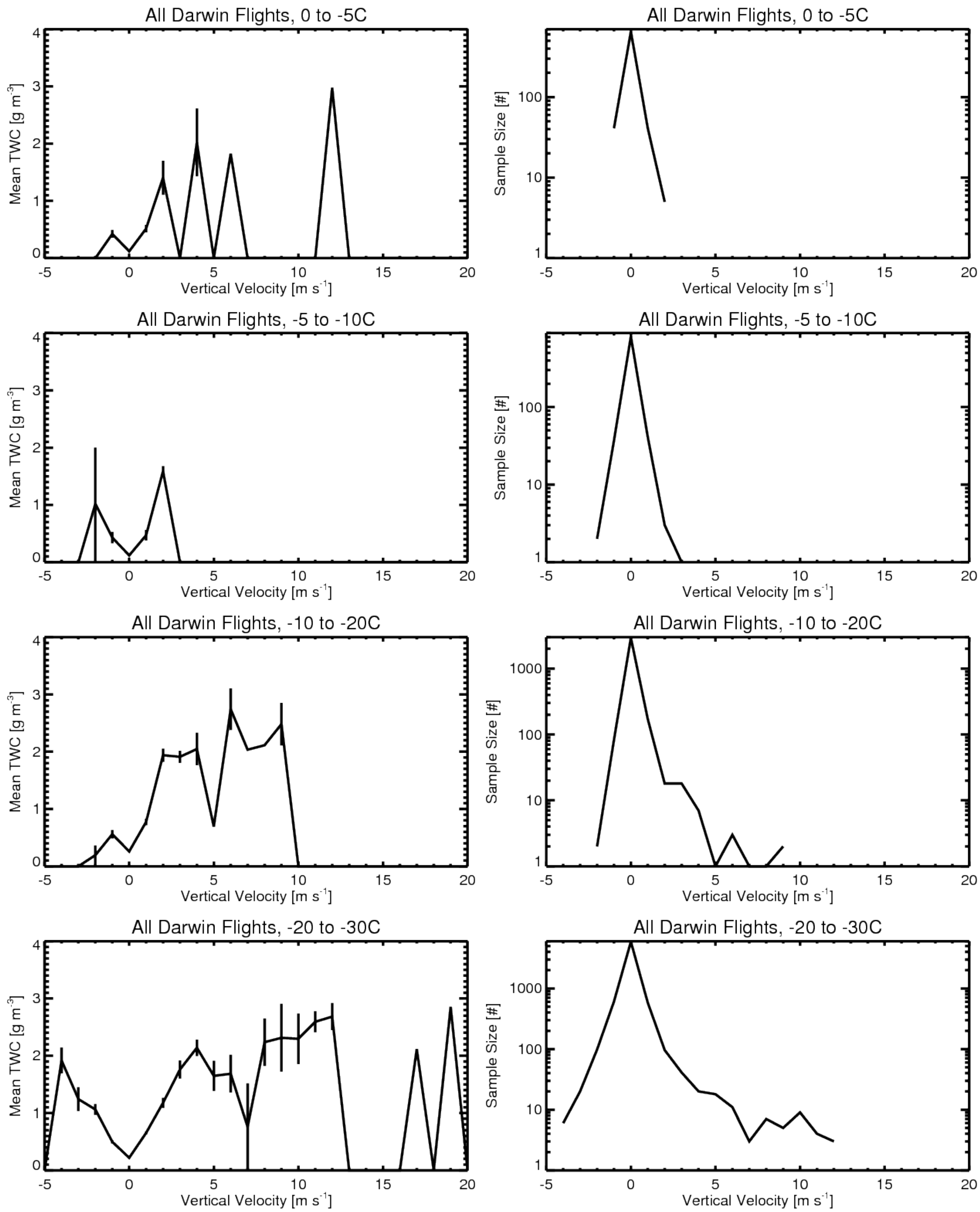

Figure 2. Mean IKP-2 total condensed water content as a function of vertical velocity for 4 temperature ranges (0 to -5°C, -5 to -10°C, -10 to -20°C, and -20 to -30°C) at 5-second (~750-m) resolution with data from all Falcon-20 aircraft flights during the HAIC-HIWC Darwin campaign (roll angles > 5° removed). Standard error bars are shown for vertical velocity bins where more than 1 sample exists. The sample size for each temperature range as a function of vertical velocity is shown on the right on a logarithmic scale.

Minor Comments

1. For the new Fig. 2, can a color bar scale for panels (a-d) be provided? Alternatively, those panels could be removed and the remaining panels could simply be presented or (e-h) could be gotten rid of and a brightness temperature from your model OLR could be computed and plotted on the same scale as observed brightness temperatures (which could be a different scale than (a-d) since the observed brightness temperatures could also be re-plotted).

2. At times, when different simulations from a figure are discussed in the text, I find it difficult to match them up with the line that refers to them in the figure since there are so many lines. I think that it would aid the reader if the simulation name (e.g., eg, qcg2, etc.) were put in parentheses when it is being discussed in the text.

3. On page 14, line 26, please insert "exist" between "errors" and "in" to make the sentence read more clearly.

4. I suggest putting standard error bars on the all flights observations line in Figure 14 so that uncertainty in the mean IWC resulting from sample size or variability can be seen.