

Reviewer 1:

Thanks for the review and for recommending minor revisions. Below we respond point by point to the comments and suggestions.

I would like to see the authors address the following:

1. I think there needs to be a more careful discussion of uncertainty. A close look at the figures suggests that there is a distribution of retrievals at each height with a fairly strong peak. The distribution is particularly noticeable in the 20-24 plots on the top row of Figure 3. Is each of these points a reasonable retrieval or is the distribution caused by noise in the retrievals?

This is an important point. As explained in the second and third paragraphs of section 3, the main limitation of these comparisons with radiosondes is that an unknown part of the differences is the temporal and spatial variability of the wind components (including internal cloud dynamics at small scale), the other part being the error of the retrieval. One way to get a sense of this spatial and temporal variability is to plot the two successive radiosonde profiles, which span about four hours from the time of launch of the first one to the arrival time of the second in the upper troposphere. Then, for each of these four-hour time intervals, we have produced the distribution of wind retrievals. Our qualitative indication of a "good agreement" is essentially when the distribution of retrievals at each height is bounded by the two radiosonde measurements. When writing this paper, we initially thought that further disentangling those two sources of differences was not possible. However, as it turns out, radiosondes from this experiment usually did not get further away from the ship by more than about 10 km, and from the new quantitative analysis we have conducted (see new version of the manuscript) we have now been able to characterize the errors more quantitatively. To do so, we have binned all comparisons with the time and spatial difference between observations and retrievals. We believe this provides the best possible quantitative assessment of our retrievals given the independent measurements we have.

Is part of the error budget the precision in the Doppler velocity measurement itself, does the noise arise from the pointing, etc? The congestus highlights the instantaneous aspects of the retrievals whereas the stratiform cases as depicted represent many hours of data yet the same level of noise seems to be present in both.

What is referred to as "noise" by the reviewer is actually not noise, it is the variability of the retrieval over four hours, which ideally should be bounded by the two radiosonde profiles, noting that inside the four hours there may be more variability than what has been measured with only two radiosondes. Precision of Doppler measurements is indeed part of the error, but not expected to be the main one by far, as we do pulse pair over 2048 samples to describe a Nyquist velocity interval of about 10 ms^{-1} , resulting in a Doppler spectral accuracy of 0.6 cms^{-1} .

There does seem to be some rather odd spikes in Vy in the congestus example on the sides of the cloud and at the top that do not show up in the vertical or eastward components. Are these real or outliers?

Well spotted. Those areas on the edges of clouds are where the detection limits of the instrument are reached, so what we see could be a contamination of the retrieval at very low signal-to-noise ratio. However we cannot discard the possibility of local turbulent structures on the edges of the clouds responsible for entrainment and detrainment on these edges. Difficult to tell for sure.

A short discussion regarding these issues would demonstrate how accurate you expect the single retrievals to be and how much averaging is expected to be necessary to converge on a useful solution. It would be interesting to show an actual updraft if such an example is available.

The new quantitative analysis of errors presented in the new version of the manuscript fully addresses this question of individual retrieval accuracy. Our new results show that virtually unbiased wind components are produced (bias less than 0.2 ms^{-1}) with a standard deviation of about 2.5 ms^{-1} on the horizontal wind components.

2. It would be nice to include plots of the radar reflectivity in the figures.

As radar scientists we can only agree that reflectivity would provide a nice context to the Doppler observations and wind retrievals, however, we feel those plots are not fully needed to discuss the wind retrievals and would result in large four-panel figures in the paper. As a result, unless there is an important reason we have overlooked, we have not added them in the new version of the manuscript.

The only typographical issue I see is on line 28 where it should read Plan Position Indicator not Plane.

This has been corrected, thanks.