General comments:

This a very important study that will have a significant impact on the field. The article is well structured and written. I have no major comments, just a few minor suggestions. I would like to propose that the article is accepted after the minor changes.

We appreciate the reviewer's acknowledgement on our work. The manuscript has been modified and much improved based on those valuable comments. A point-by-point response (text in blue) to the reviewer's comments can be found below.

My main suggestion is to use standard verification metrics when presenting the performance of the method, i.e. In Fig. 2. I would suggest that you show probability of detection, false alarm rate and critical success index as functions of Z, instead of CDF.

Response: We thank the review's suggestions. In Fig.2 we want to highlight the significant difference between the ML and the conventional drizzle detection method, thus we think it is more intuitive and appropriate to show the CDF of the detected drizzle signal as a function of reflectivity. We followed the review's suggestions by adding the standard verification metrics for the two-drizzle detection methods as shown in Fig S9 in the supplement. We also modified the text as follows:

"…A more detailed performance comparisons of the two drizzle detection methods are shown in Fig. S9, where the results are similar with Fig.2, the rise of the false detection rate for the ML-based method for reflectivity lower than -20dBZ is due to the exists of the extremely weak drizzle signals as will be discussed later…"

Minor comments:

Line 157: "...cloud/drizzle datasets is trained by a machine learning algorithm..."

Do you mean a machine learning algorithm is trained by the datasets?

Response: We thank the reviewer for pointing out this oversight, the modification has been made:

"…Finally, a machine learning algorithm is trained by the collection of well-defined cloud/drizzle datasets to resolve the drizzle identification function…."

Line 170: "...turbulence broadening is set as 0.2 m/s which is obtained from local observations..."

You use spectra width for cases where Z is less than -20 dBZ, do you do any other data filtering? As you show later, spectra shape might be modified by auto conversion even for such low reflectivity values.

Response: We didn't perform other filtering process as we believe reflectivity < -20 dBZ is an adequate constrain to isolate the effect of microphysics on Doppler spectrum width. For the vertical pointing radar, the observed spectral width is a measure of the Doppler spectrum broadening which is mainly contributed by three factors: turbulence (σ_t), microphysics (i.e., the falling velocity difference among hydrometers with different size) and the wind shear effects (usually is negligible compared to other two terms). In our study, we assume that for the non-drizzling (or weakly drizzling) clouds, Doppler spectrum broadening is mainly contributed by the turbulence factor, thus the observed second-moment of the Doppler spectrum, i.e. spectral width, can be used to estimate the turbulence broadening factor (σ_t).