

Major comments:

Figure1: Does this mean the CNN algorithm can only detect cloud droplets in the presence of bimodal spectra? I noticed the identified CD aera in Fig.7, Figure A4 are associated with this bimodality. If so, this may be related to the relative low “recall” number shown in Fig 9 and Fig 11 as the algorithm may tend to miss the scenario when cloud droplets exist but manifested as single peak spectrum (e.g., when the ice-contributed Doppler spectra is not distinguishable with the cloud Doppler spectra).

Line 243: I’m thinking the microphysics and the related process may be different in these two sites. Do you expect the proposed machine learning algorithm would generate a better performance if it was trained for these two sites separately?

Line 172-Line 175: Besides the “vertical generalization”, have you considered the normalization horizontally? Specifically, I am wondering the effect of vertical air motion on the model performance. If the trained model is very sensitive to the spectra peak location, then the spectra location offset caused by the air motion may confuse the ML model.

Line 413-Line 415: I’m thinking some other quantities, like root mean square error, are more preferable to indicate the LWP difference between observation and retrieval. Please elaborate more on the reasoning of using correlation coefficient.

Minor Comments:

Line 243: I would tend to think more datasets are preferable for model training if the training datasets are properly cleaned. Did you notice the prediction performance reduced when more training dataset are included?

Line 40: remove e.g.

Line 106-Line 107: The conclusion “... is used to distinguish between cloud droplets and aerosol in Cloudnet” is not supported by the reasoning discussed above.

Line 150: The frequency modulation unit here (MHz) is not consistent with the one shown in Table3 (kHz), please confirm.

Line 157: Do you mean “polarized signals”?

Line 162: Please check this sentence.

Line 189: I couldn’t find the precipitation rate products in the datasets, please add this information.

Line 202: Do you mean the “... if the detected cloud top by **lidar** was less than 500m above...”?

Line 214: This sentence should read like “Fig. 3 shows the architecture of the VOODOO retrieval algorithm”

For Figure 3: For the input, if I understand correctly, six Doppler spectra are constructed as a spectrogram for training. However, line 177 indicates that 30 time steps are used. Please confirm the number is consistent.

Line 361: Figure6 has no subplot.

Line 399: Please check this sentence: “... range are thatight than 0.8”

Line 400: Are you referring to the third column of Fig. 9?

Figure11: Please also add the probability density functions of LWP of the same datasets as shown in FigureB1 and B2, so the readers can see the relative frequency of the performance scores.

Line 426: Recall is not directly estimated by TN, please check the explanation here is reasonable.