Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., author comment AC1
https://doi.org/10.5194/amt-2022-149-AC1, 2022

Reply on RC1

Willi Schimmel et al.

Author comment on "Identifying cloud droplets beyond lidar attenuation from vertically pointing cloud radar observations using artificial neural networks" by Willi Schimmel et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-149-AC1, 2022

Dear Refer

Thank you very much for your feedback and for concidering the submission for AMT publication. Below are my answers and corrections to all of your questions and remarks.

Best regards,

Willi Schimmel

Major comments:

Figure1: Does this mean the CNN algorithm can only detect cloud droplets in the presence of bi-modal 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).

- Although we cannot say with certainty which morphological features are used by the CNN to identify cloud droplets, we can see that the calculated probability for the presence of cloud droplets in bi-modal spectra is very high (>0.8 see Fig. 7 right). In Fig. A4 on the right, it can be seen that the CNN shows increased cloud droplet probability values at the cloud top (>0.3) even though the radar Doppler spectrum is monomodal. Here the CNN seems to be oriented to other features (possible spectrum width, skewness, or similar). The same is true for Fig. 7 left above 2.8 km and Fig 7 right, between 3-3.3 km.

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?

- In short, yes. In the future, it will be possible for users to train the CNN model separately for the respective operational areas. Site specific training set will most likely enhance the performance. However, since the performance scores for the second site (Leipzig) are also high, we concluded that no retraining of the CNN for this dataset was required. Kalesse-Los et al., 2021 (https://doi.org/10.5194/amt-15-279-2022) had tested the applicability of a CNN pre-trained in the Arctic on mid-latitudinal clouds and also found it to perform sufficiently well.

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.

- The Punta Arenas data set is very strongly influenced by orographic gravity waves (see Radenz et al., 2021; https://doi.org/10.5194/acp-21-17969-2021). Longer updrafts and down-drafts cause the location of the peaks to change significantly. Nevertheless, we were able to obtain good results in the prediction of cloud droplets for Punta Arenas. I.e. the CNN seems to generalize to the extent that this phenomenon does not have a major impact on the prediction.

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.

- Added this information on Line 420-425: " Note that the method of \cite{Karstens1994} uses the adiabatic assumption to calculate the liquid water path in combination with the liquid water mask (time-height-mask for droplets presence), and ECMWF weather model data (temperature, air pressure, and specific humidity). Since the adiabatic assumption is not suitable in all cloud situations and the model data is subject to uncertainties, the idea was to compare only the correlation of both time series (MWR-LWP vs. LWP$_\mathrm{ad}$ of Cloudnet vs. LWP$_\mathrm{ad}$ of VOODOO). Therefore, we decided not to compare absolute values. "

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?

- We've noticed a drop in the recall score, when using more (e.g. 80% training vs. 20% validation) data for training. The distribution is 5 to 1 for noCD to CD samples. Arguably we assume, if the CNN is trained on more data, the CNN focuses also more on the noCD class.

Line 40: remove e.g. removed

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.

- reformulate: "Due to its very high sensitivity to cloud droplets, beta_att is used in the Cloudnet processing to identify cloud droplets."

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

- changed unit to kHz Line 150

Line 157: Do you mean "polarized signals"?

- yes, fixed typo

Line 162: Please check this sentence.

- fixed sentence: "The radar moments Z_e, $\bar{v}_D$, σ_w, and LDR are estimated from the spectra and stored as NetCDF files. Those files are used as input for the Cloudnet processing."

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

- precip rate from Vaisala WXT536 Compact Weather Station mounted to the cloud radar, I added this information to the manuscript Line 189

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

- No, by default Cloudnet extends a liquid layer to cloud top using the cloud top as it was observed by radar. If this radar cloud top is less than 500m above the liquid cloud base (detected by lidar), Cloudnet classifies those pixels between liquid cloud base and radar cloud top as liquid, even though no lidar signal is available.

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

- changed the sentence accordingly

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.

- Clarified that for the radar settings used, six spectra are concatenated per sample. Line 176

Line 361: Figure6 has no subplot.

- removed (A) and (B) in text.

Line 399: Please check this sentence: "… range are thatight than 0.8"

- rephrase sentence: "The first column of Fig. \ref{fig:s-metrix} shows that the precision values have a clustering close to 1"

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

- yes, change to "third"

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.

- This information is implicitly presented in Figure 9. B1 and B2 are PDFs of 2D variables, however the LWP is time-series data. We could add the LWC (liquid water content), which is a 2D variable, although this is likely not very meaningful because it relies on many assumptions made in the Frisch et al.1994 retrieval; mainly the adiabatic assumption, but also model temperature, pressure, humidity uncertainty, etc.)

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

- correct, changed TN to FP