Geoscientific
Model Development
Discussions

Interactive comment on "Model evaluation by a cloud classification based on multi-sensor observations" *by* Akio Hansen et al.

Anonymous Referee #2

This study presents a cloud classification product for numerical models, so that model output can be directly compared with observations via CloudNet. While I find this an interesting research question, I find that the current paper does not show sufficiently originality, either for the development of the technique or for the evaluation thereof. I unfortunately have to recommend rejection of the current paper, mainly for the following reasons:

1) While the authors state that this goes beyond a comparison of liquid/ice/etc. . . cloud fraction, in practice it remains close. I was expecting much more of a simulator approach here: Simulate the radar retrievals in the simulations, and classify clouds based on that. Such a project would have a more significant additional value on top of current apples to oranges comparisons.

2) The data set of the comparison with observations is very limited: Only two months over a single location. Not only results this in a much larger statistical uncertainty than necessary, it also means that the melting ice & cloud droplet category is not hit within the dataset.

3) The coarse graining algorithm ("take the most common category") does not seem all that accurate to me. Ideally, if CloudNet predicts some columns with (e.g.) liquid clouds and some separate columns with drizzle, the result on the model grid should be the mixture of the two.

4) Following up on point 3, it is unclear how the authors treat subgrid variability in the model. Are the subgrid parameterizations included at all in counting cloudiness? If so, does a pixel classify as liquid/ice/. . . if that phase is the dominant process, or is it possible to have multiple classifications within one model box?