Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-58-RC5, 2018

Atmospheric Measurement Techniques Discussions

Interactive comment on "Unraveling hydrometeor mixtures in polarimetric radar measurements" *by* Nikola Besic et al.

Anonymous Referee #5

Received and published: 10 May 2018

Summary: Nikola Besic et al. This paper presents a technique for performing hydrometeor identification (HID) on polarimetric radar measurements. In particular, it demonstrates a technique to retrieve a mixture of hydrometeors within a single bin, rather than only returning the dominant hydrometeor type. Only being able to return the dominant hydrometeor type has been an Achilles heel of HID. I believe the paper does have a high level of significance. Additionally the figures, other than a few places mentioned below, are of high quality.

The writing is adequate, though some portions would benefit from some editing for clarity. The paper is split into two effective sections. The mathematics in the paper can be confusing to follow as notations and terminology seems to change as the paper progresses. This is caused in part by trying to pull in techniques from several different

fields. Overall while I believe the concept has merit, I have some issues with some of the base assumptions. My largest objection is that after identifying the dominant hydrometeor, that the second closest centroid represents the most likely mixture component. I think the radar variables forming correlated frame vectors rather than a true basis make this assumption require quite a bit of justification.

The comparison of the data with the MASC is a good first step down this path, but as it reports fewer number of hydrometeors (3 ice classified in this case) than the classifier, it makes it hard to know for sure if there is any type of interference that would cause misclassification. I'm not convinced the authors have truly demonstrated that their secondary mode classification works. I would like to see potentially another dataset brought in to show some generality to the method.

While my review is somewhat critical, I do think the technique has a fair bit of merit and I would like to see the authors continue it through. In particular, I would love to see a way to address the issue of proportions of secondary components. The authors do equal mixtures for the development of the thresholding of the probability scores, but I believe one could do some simulation to give actual error bars on the retrieval by varying the components of the mixture at different amounts and then performing the retrieval. While real observations always need to be used as a final step, I would have liked to see more simulation data used to characterize the behavior of this retrieval.

Major Comments: From the abstract: "Intrinsically related to the concept of entropy, introduced in the context of the radar hydrometeor classification in Besic et al. (2016), the proposed method, based on the hypothesis of coherency in backscattering, estimates the proportions of different 10 hydrometeor types in a given radar sampling volume " As far as I can tell, entropy is not actually used in any of the proposed method, but rather as a followup parameter to further support the conclusions of mixed hydrometeor types in the bins.

This paper starts with an analogy to the Cloud and Pottier work on classifying po-

larimetric SAR, and then attemps to expand this to weather radar using reflectivity, differential reflectivity, Kdp and Rhohv as well as a phase indicator variable. I would like to echo the other commenters that the coherent and incoherent terminology should be explained in text due to its similarity to other common uses of that term in radar.

The exposition on page 5 related to POLSAR, while interesting, obscures the discussion of the technique as it applies to weather radar. While entropy is used, the Cloude and Pottier work uses much more than just entropy for their classification, namely adding anisotropy and alpha angle which provide useful information for the POLSAR classification. Those do not necessarily apply here (or at least are not explored in this paper). The switching between different notation and vocabulary can make the work difficult to follow.

I think the approach and workflow could be made much more clear. The classifier listed in this paper is essentially comparing the distance of a measured observation vector to each of the examplars and the minimum distance is assigned to that bin (equation 1). My biggest problem with the work is because the vectors of the space you are embedding into are not orthogonal, you can add 2 hydrometeor types to find a third type. So while I buy that the classifier predicts the dominant type, assuming that the second most likely hydrometeor is present in the mixture feels like a big leap and is not supported in the development of this work. Also what does proportion mean here? I understand it is somewhat of a mathematical construct here, but do the authors feel it is most akin to weight, number, or possibly relative amplitude of the signal returns?

On page 14 the statement : “As suggested in Subsection 4.1, the first uncorrelated components is supposed to represent a "pure" component, freed up of the effect of incoherence in the data acquisition “

Page 6.5-20: There is a lot of switching of terminology here between standard mechanism, end member, centroid, etc.

Page 6.11-12: “Nevertheless, basing the estimates of proportions on the distance in the

space populated by the standard mechanisms and measurements is a reasonable way to sum up the standard mechanisms.” I don’t feel like this has actually been justified and is a key premise of this work.

“Now that we have identified our centroids as standard mechanisms (analogy with PolSAR) and have found a mode for their addition via the distances in the Euclidean space of the classification “ Again, I feel like this is overstated. The reduction in equation 5 and 6 of this to a linear mixing problem does not feel justified. Perhaps this is because m_i is not actually defined here. I assume this is the observations matrix (k) associated with that pure material, but maybe I misread this?

“Boxes of hydrometeor mixtures contain different versions 5 of mostly plausible mixtures obtained by linearly combining in equal shares polarimetric parameters of two hydrometeor types involved, following assumptions of coherence and linearity “ I would like to see more text as to how this mixing was accomplished. Figure 5 implies you only used equal mixtures of the two hydrometeors (50/50). Why limit it only to equal mixtures? I assume it has to do with finding the threshold value between two centroids that exists at exactly the 50% crossover point, but this should be made more clear.

Figure 1: I don’t understand what the authors are attempting to demonstrate with this figure.

Figure 2: The definition of hydrometeor types in this paper is only in this figure caption. I would add a table or descriptive text with this information.

Equation 7: This c_i is essentially the same as m_i above correct? Maybe some clarifying statements here, or normalizing the notation would help. In general many of the equations seem to use one set of notation to introduce an equation (I assume the notation from that field), then apply it using a entirely different set of notation. You should unify the different representations.

Page 10: In the comparisons with the MASC, where are the centroids for each HID

class drawn from? In particular, are the centroids trained on this dataset, or on another dataset? I'm wondering how much overfitting may be involved in this process that would limit it's generalizability.

Minor Comments: Page 3.25 – leptokurticity, while a valid term, will cause many readers to have to stop reading and go look this term up as it is not commonly used in this field. Perhaps add a subclause to describe the term.

Figure 5: It took a few readings to understand that the greyscale level of the boxes and the text in (a) are related. Maybe make this more clear.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-58, 2018.