

## ***Interactive comment on “Development and validation of a supervised machine learning radar Doppler spectra peak finding algorithm” by Heike Kalesse et al.***

### **Anonymous Referee #2**

Received and published: 30 May 2019

This manuscript submission describes a new proposed algorithm to find Doppler velocity spectral peaks using an automated methodology. This algorithm purports to more consistently and accurately identify peaks due to different hydrometeor populations - a sometimes difficult task in mixed phase clouds. The authors, in my opinion, present a convincing case that the new machine learning algorithm outperforms previous Doppler spectral peak identifier algorithms. Example cases are shown in the manuscript, with further examples presented as supplementary material. I offer a few minor comments below that will hopefully improve the manuscript.

\*\*\*\*\*

C1

1. Abstract, lines 12-14: The authors state that “The new algorithm is found to perform well.” “Well” is a very subjective description. A more quantitative descriptor, or at least less subjective language, is preferred. Possible alternative wording that combines two sentences:

The new algorithm consistently identifies Doppler spectra peaks and outperforms other algorithms by reducing noise and increasing temporal and height consistency in detected features.

2. Lines 23-25: Suggest changing nominalized language for stronger writing:

The first step towards characterizing hydrometeor types is determining the number of different populations within a certain cloud volume.

3. Introduction, Lines 21-31: This section of the introduction is very fragmented. The authors inject various Doppler spectra analysis studies in a somewhat non-coherent manner. Maybe it's just a spatial issue (i.e., the authors indent the various studies as stand-alone paragraphs comprised of 1-3 sentences). One way to mitigate this issue to allow the science themes, rather than the referenced studies, to drive the content. I envision these lines recast in terms of scientific topic that would allow a more natural flow to the discussion. A suggestion:

Other studies have utilized Doppler spectra analyses to identify cloud microphysical composition and cloud processes operating in Arctic clouds. For instance, four Arctic cloud hydrometeor populations (background ice, cloud, drizzle and new ice) were successfully classified using continuity of spectral modes in time and height combined with high spectral resolution lidar (HSRL) and in-situ observations (Verlinde et al. 2013). BAECC {what does the BAECC acronym represent?!} field campaign analyses have also distinguished up to three noise-floor separated peaks in the recorded Doppler spectra for frontal snow falling through a supercooled water layer (SWL) that produced rimed snowflakes (Kalesse et al 2016). These respective peaks were then used to track microphysical processes along slanted fall streaks, although this documented case was

C2

special due to the separation of peaks by the noise-floor (merged peaks are usually observed, motivating the need to develop robust cloud radar Doppler spectrum peak separation techniques). Finally, KAZR observations of liquid-only and mixed-phase clouds at Oliktok Point, Alaska have been used to identify multiple Doppler peaks using the depth of the local minimum between the main peak and sub-peak as the main separation criteria (Williams et al. 2018).

All these efforts, using somewhat differing approaches, show that there is a need... [continue with the rest of the content from the last introductory paragraph].

I also suggest adding a final sentence to the introduction that briefly introduces what the current study will accomplish. For example, "This study describes a new algorithm that adopts machine learning tools to classify Doppler spectra peaks in complex mixed phase cloud scenarios" – or something similar to this statement that properly whets the readers' appetites.

4. Figure 1: Are these truly random spectra chosen from 16 February 2014? Or are they neighboring spectra, where neighboring can be defined as either spatial (height) or temporal?

5. Page 5, Line 2: 21 February 21 2014 -> 21 February 2014

6. Page 6, Lines 3-4: How did the chosen smoothing method produce the most promising results? Is there any quantitative measure to optimally select the smoothing method (like line fitting parameters)?

7. Page 6, Lines 5-9: Was there a compelling reason to choose the 16 s temporal and 90 m spatial smoothing parameters? This question is probably related to the previous comment. The obvious answer is that spatiotemporal smoothing needs to capture the multi-modal peaks shown in Fig. 2 without completely smearing out the features. I guess I'm having a difficult time being convinced that one could empirically derive the best smoothing parameters and method based only on an "eye test" without further

C3

quantitative support.

Post-hoc comment: The appendix content nicely lends further support for how the algorithm works with the adopted spatiotemporal constraints. I was initially going to suggest appendix material that shows how the algorithm would perform with different smoothing methods and parameters - maybe include a final brief appendix section illustrating the sensitivity of one or two cases to different smoothing schemes or spatiotemporal averaging parameters?

8. Figure 8 caption: I recommend adding what the black dashed line indicates. It is obviously the SLW layer that is again repeated in a later figure, but it should probably be mentioned here, too.

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-48, 2019.

C4