General comments:

In this manuscript, the authors combine two algorithms (PEAKO and peakTree) for detecting and characterizing peaks in cloud-radar Doppler spectra. Based on two test cases, they investigate the capabilities and limitations of their newly implemented PEAKO-peakTree toolkit. Their results indicate that the presented analysis method can be used to identify supercooled liquid water and ice peaks in mixed-layer clouds as long as turbulence is not too high and radar settings guarantee high-quality Doppler data.

Overall, the manuscript is well written and the results are presented clearly. For easier readability, long compound nouns that are used occasionally (particularly in the abstract) could still be simplified (s. comments below).

I particularly like that the authors make an effort to assess under which conditions the best results can be expected for analyzing multimodal cloud-radar Doppler spectra with the PEAKO-peakTree toolkit. Nonetheless, I would still suggest to include more information about how different factors impact the results and conclusions, as listed in the specific comments below. The description of how turbulence is modeled and included in the analysis should be expanded, in particular, so the readers can better understand this crucial part of the analysis, even if they are not familiar with the cited paper(s) of how turbulence can be modeled and relates to radar Doppler spectra.

Considering the high quality of the study and the novelty of the results, I would therefore suggest to publish the manuscript after these points are addressed.

Specific comments:

Abstract: Some sentences rely heavily on nominal style with long compound nouns, which makes them difficult to read and understand promptly.

In l. 8, for example, '…use of two cloud radar Doppler spectra peak analysis algorithms' could be replaced by '… use of two algorithms for analyzing the peaks in cloud-radar Doppler spectra' or '… use of two algorithms for analyzing the peaks in Doppler spectra recorded by cloud radar measurements'. Similarly, in l. 10, 'The learned Doppler spectrum peak detection parameters …' could be replaced by 'The optimal parameters derived for detecting all relevant peaks in the Doppler spectra …'.

Analysis: What is the impact of clutter (= strong non-meteorological signal, e.g. reflections from the (wet) ground) in the radar data on the analysis results? Do the cloud-radar data have to be free of clutter? Is it not necessary to filter the data before detecting peaks, because the clutter is filtered automatically by the PEAKO peak finding algorithm? Can the analysis also be applied to a different radar with different clutter characteristics?

These points should be addressed briefly somewhere in the manuscript, e.g,. in Section 2.2 and in the Conclusions.

Section 2.4: I do not fully understand what is done here. Particularly, how does MDV enter the calculations? From the text, I assume v_h is obtained from the atmospheric models at each radar range bin, but it is not clear how this is then translated into a power spectrum and what the role of MDV is.

Considering the importance of turbulence in the radar analysis later on, this section should be expanded to make it easier to understand how turbulence is estimated for the analysis.

Section 2.5: Why do you use a constant fill value for empty size bins and why is a value of 10 m-3 chosen, in particular? Intuitively, I would assume filling the missing values with another value or by spline interpolation could also be a valid approach. Would using different approaches modify the peak characteristics of the forward-modeled Doppler spectra and thus change the results significantly?

Section 3: How would the results and conclusions change, if a different prominence threshold and minimum peak width (instead of 1.5 dB, 0.03 ms-1) were chosen? Or, why did the authors select these specific values?

For example, if the peak prominence threshold and/or the peak width threshold for this evaluation were increased, it seems plausible to expect that a lower number of spectral averages will already lead to a reasonable number of identified peaks, i.e. the curves in Fig. 2 reach more realistic values more quickly.

Section 4.1: During training, is it important to already include Doppler spectra from the type of precipitation that is to be studied later on, e.g., if we want to identify liquid peaks do we need to include (many) Doppler spectra that have liquid peaks in the training dataset? This would be important to mention, because, then, the training dataset cannot be chosen at random but would require a pre-selection process to obtain a suitable subset of all available data.

Genenerally, how 'realistic' are the peaks identified by a human user and what effect on the results would be observed if the same data were used by different humans to create the training dataset?

l. 450 ff: Just by visually comparing Figs. 6c and e, the correlation seems rather vague. Is this correlation evident (more clearly) from plotting log(EDR) vs. the peakTree no. of peaks, and maybe even quantifiable through the Pearson correlation coefficient?

l. 511 ff: In Fig. 6h, time 3 is characterized by a higher ice particle concentration (including larger ice particles according to Figs. 7d-f) than times 1 and 2. Could this be the main reason why an additional ice peak can be observed in the Doppler spectrum at this time, and the changes in turbulence only play a minor role here?

l. 573 f: This statement seems rather aggressive, considering that no additional information is given on how these applications could actually be achieved from the methods and results presented in this study and no citations are included where radar Doppler spectra are already used for those applications.

Technical corrections:

l. 32f: No line break between minus sign and 37 °C. Use minus instead of hyphen.

l. 141: Should the units of the sampling volume be cm^3 instead of cm^-3, i.e. no minus sign?

l. 211: Peak prominence in dB instead of dBZ?

Figure 2: Add the units of velocity and spectral reflectivity to panel c?

l. 359: maybe better rephrase as '… reflectivities of undetected and falsely detected peaks…'

l. 440f: …., with cloud top height around 2 km and cloud top temperature of approx. ….

l. 507 f: Replace compound noun 'hydrometeor mode fall velocities' with 'the fall velocities of different hydrometeor modes' or something similar

l. 551: 'cloud radar Doppler spectrum peak detection …' could be replaced by 'Detecting peaks in cloud-radar Doppler spectra …'

l. 553: maybe use present tense 'point to' instead of past tense 'pointed to' (similar to 'the results suggest that …' in l. 558)

l. 557: … up to 2500 spectra as training data, …

l. 561: co-located instead of closely located

l. 577: 'capability' instead of 'capacity'