Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2022-35-RC2, 2022

Comment on amt-2022-35

Anonymous Referee #2

Referee comment on "High-fidelity retrieval from instantaneous line-of-sight returns of nacelle-mounted lidar including supervised machine learning" by Kenneth A. Brown and Thomas G. Herges, Atmos. Meas. Tech. Discuss.,
https://doi.org/10.5194/amt-2022-35-RC2, 2022

Overall a well written paper on solid scientific grounds that fits well to the scope of AMT. I suggest publication after a few points have been clarified.

general comments

- In contrast to the two other techniques the Machine Learning approach not only filters spectra, but processes the whole way down to u_LOS. It is not really a QA/QC technique anymore but rather a full retrieval, hence I consider statements like "Our work compares three QA/QC techniques, including conventional thresholding, advanced filtering, and a novel application of supervised machine learning ..." (l. 10) as inaccurate.
Why did you not choose to just using ML for QA/QC to produce filtered spectra and then run the peak detection on these spectra as you do with the other approaches? This would make the approach more general and give more insight on what the model actually does, hence allow for more targetted improvements

- The paper could make more clear right from the abstract and the introduction that the main aim of the current work is to get rid of solid interferences while keeping the data availability high. Indeed Fig 11 and Tab. 3 show that for the other cases already the very basic thresholding approach is enough to achieve comparable results than for the two more adavanced techniques

specific comments

---

l. 15: please re-word "overlapped meteorological tower" to something like "on-site meteorological tower" or "meteorological tower within the sampling volume"

Sect 2.2: This description is quite long for that it actually principally follows Herges and Kayantuo, 2019. Please consider a more concise formulation focussing on potential differences to what is presented in the above-mentioned reference. Nevertheless, the illustrations Fig. 2 and 3 are rather valuable, I wouldn't skip these.

Fig. 2: I found it a bit confusing to have the most "raw" data in subfigure c) while excerpt spectra of it are in b) and processed median LOS winds in a). If I didn't miss anything fundamental, I would prefer to change c) <-> a)

Fig. 3: Please leave a reference to the original spectra in Fig. 2b) in the caption of Fig 2)

l. 301: you deliberately exclude double-peaked spectra for trainging what is to be expected present for wake situations. Why this choice? How to explain that the Machine Learning results fit the anemometer data that well in Fig 10d)?

I would prefer to have the experiment site description (Sect. 3) before the data processsing description (Sect. 2) as I think it might ease the reading process. Especially it will also make the section about the ML easier to read.

l. 528: "This source stems from the difference..." Your explanation to this shows plausibility for this interpretation but is not sufficient to exclude all other possible sources which could cause such a bias. For this you would need to deliver a more quantitative estimate of your interpretation. Anyways, it is maybe beyond the scope of this manuscript so you might simply go for "This source PROBABLY stems from the difference..."

Fig: 10: The green "x" can be misinterpreted as a fully trusted data point (instead of an outlier well handled by your mehtod) if not carefully reading the legend. Another notation would be preferred.

l. 709: Please add a reference to the out-of-distribution-detection techniques

l. 734: In what you show rather than providing "higher accuracy" ML appears to provide "very similar accuracy" to the advanced filtering method. Please re-word

technical corrections
---

l. 331: please include space between "3." and "Experimental Techniques"