

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2022-35-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2022-35

Anonymous Referee #1

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-RC1>, 2022

General Comments

The authors present an investigation which focusses on using machine learning to improve the processing of Doppler spectra acquired by scanning Doppler lidars. The manuscript is both well structured and written, which enables the understanding of the processing steps used in the study as well as the rationale for following them. Their results show that through the use of machine learning it is possible to increase the data availability especially in the case when a scanning Doppler lidar is measuring complex flows, like a wake of a wind turbine. In those cases, the differentiation between wind signals and noise is a challenging task and the conclusions of this study will contribute in addressing these challenges.

Please find below my comments on the manuscript, as well as suggestions for improvements that I think will enhance the understanding of the presented study.

- The authors throughout the manuscript discuss about the interferences that the optical window of the scanning wind lidar is creating in the measured Doppler Spectra. I think that rather a general problem for wind lidars, this issue is a characteristic of the scanning wind lidar used in this study. It is probably attributed to alignment of the top window to the line-of-sight of the lidar. This can be dealt if the top window is not installed normal to the line-of-sight of the lidar, as it happens for example in the commercial ZX300 wind profiler. I suggest that the authors state that this is relevant for this instrument or provide a reference that this is a general problem. Furthermore, the near zero velocity contributions could be suppressed using an electrical high-pass filter.
- In Figures 11 and 13 are presented scatter plots of the line-of-sight measurements from the lidar and sonic anemometers. The lidar data used correspond to quasi-instantaneous measurements, acquired at 400 Hz, and the sonic anemometer data that

was at 100 Hz “has been temporally interpolated at the same instance as the lidar passing”. In those high sampling rates, it is very important to ensure a very good synchronization between the two data acquisition systems. Can the authors add information about the expected synchronization accuracy?

- The authors in the lines 197 – 198 mention the wind turbines WTGa1 and WTGa2, but the presentation of the experimental setup is first presented in Figure 7. I suggest that the authors should consider moving the presentation of the Experimental Techniques section after the Introduction, so it is easier for a reader to understand the discussion about the processing techniques.
- The authors in Section 2.2 give an overview of the “Advanced Filtering” technique. I find the overview quite extended. Maybe the authors could consider highlighting the differences in relation to the method presented in Herges and Keyanto (2019) and reduce a bit this part.

Specific Comments

- Line 46. Can the authors elaborate more what they mean with “typical setups”?
- Line 56. Here the authors talk about the “time series of u_{los} ”, thus, if I understand this correctly, referring to the line-of-sight velocity that is derived by the estimated characteristic frequency shift in each Doppler spectrum. However, in the caption of Figure 1 they use “ u_{los} ” to also denote the various line-of-sight contributions over a frequency bandwidth of the region of interest. I find this a bit confusing, thus I suggest that the authors clarify this part. For example, they could use the same symbol as the one they use for the median line-of-sight.
- Lines 68 – 72. Here the authors describe noise contributions in a laser Doppler spectrum. First, I suggest that they should mention here that this is relevant to a cw Doppler lidar. Furthermore, I think that this description is not accurate: a. The presence of a flat noise floor doesn’t result to a loss of precision of the line-of-sight velocity estimation, since it can be effectively removed by thresholding. The mean noise floor due to shot noise is dependent on the intensity of the local oscillator laser and not on the backscatter. b. The SpinnerLidar uses a coherent detection scheme which is based on a balanced photodetector, therefore I don’t see how is relevant the reference of Liu et. al 2006, c. The variance of the noise floor, which is dependent on the number of FFTs averaged to produce a Doppler spectrum will introduce a noise in the measured Doppler spectrum, the impact of which will be dependent on the intensity of the backscattered light, the frequency bandwidth of the wind speed fluctuations, and the method used to estimate the characteristic Doppler shift.
- Line 86. Rather than an increase in the signal strength, the accumulation of a number of spectra reduces the noise variance.
- Lines 154 – 156. Can the authors explain this part a bit better? I am not that I understand the statement “...this magnitude is the maximum among the bins...”
- Lines 168-169. Can the authors explain what they mean when they write that they “don’t apply any correction for the alteration of the skew distribution due to thresholding”?
- Line 184. What kind of scan-head motor speed rates did the authors test?
- Line 199. In Figure 2c the authors present a time series of normalized Doppler spectra. For the normalization they subtract the mean noise floor from the measured spectra and subsequently divide with a fixed value. I don’t understand the point of this division, it is more a rescaling of the spectrum rather a normalization.
- Line 205. The authors state that “the mask regions were then increased horizontally by 2 pixels”. What is it meant here with both “horizontally” and “2 pixels”?

- Line 229. How large is the area of the sliding neighbourhood and how did the authors reach to that value?
- Line 234. The authors state that in Figure 4b the “peak returns signals from the operational rotor are clear”. However, I am not sure if I can visually detect them. Can you please explain a bit more in the text what are the observations that one should notice in that figure?
- Lines 277 and Lines 285. Since, I am not a machine learning expert I don't know how typical are the values of B and of the percentages of the data split. Can the authors add a short comment or a reference in the manuscript to support or explain the selection of these values?
- Lines 329 – 330. Here the authors the authors write that the selected threshold of the standard error provides an acceptable balance between data availability and variance error. However, they also state that the trade off has not yet been studied exhaustively. On which basis it is concluded then that it provides an acceptable balance?
- Line 371 I suggest replacing the reference of Mikkelsen et al. 2013 with the

Sjöholm, M., Pedersen, A. T., Angelou, N., Abari, F. F., Mikkelsen, T., Harris, M., Slinger, C., and Kapp, S.: Full two-dimensional rotor plane inflow measurements by a spinner-integrated wind lidar, in: European Wind Energy Conference & Exhibition 2013, 2013.

Since, in the Mikkelsen et al. 2013 a spinner lidar with a single prism scanner head was used.

- Line 414. Table 2. The mean wind direction that is reported here does not seem correct. For that wind direction the wake of the wind turbine would not be in the location of the mast. How is the mean direction estimated here?
- Line 441. Please define δ also in the text, not only in the caption of Figure 8.
- Page 21. In the legend of the Figure 9 it is written that the spectra presented in Figures c – g are scaled with a fixed number ($2^{16}-1$). However, just from a visual inspection of the figures, all presented spectra appear to have the same maximum value, which maybe is 1? Are the authors sure that they scale the spectra with a fixed number and not with the maximum value of the measured spectra? it is otherwise strange that all measured spectra have the same maximum value.
- Page 24. Figure 11. In this figure are presented scatter plots between lidar line-of-sight measurements and the corresponding sonic measurements projected on the lidar's line-of-sight. Here it looks like the measurements at 10 m (presented in orange) have overall higher values than the ones at 58 m (presented in red). This is inconsistent with the profile presented in Figure 9. Can the authors explain why is this difference observed?
- Lines 538 – 543. What is the direction of the booms in relation to the geographic North?
- Line 553. Why do the authors use the letter T to denote standard deviation and not the Greek letter sigma?
- Line 557. What is the “parameter estimation problem”?
- Line 584. Table 3. Would not it more relevant here to estimate the mean and standard deviation of the absolute difference between lidar and sonic line-of-sight measurements? And where it is attributed the drop in the data availability that is observed in the machine learning technique? The same comments apply to Tables 4, 5

and 6.

- Page 28. Table 4. Why there are no measurements at 45 m and 58 m in the case of the "thresholding" and "machine learning" techniques? Furthermore, how do the authors explain that even though the machine learning technique can eliminate biases that are attributed to the truncation of the low energy part of the Doppler spectrum the statistical results are still similar with the other techniques?
- Line 635. The reverse process of Eq. 3 is not straightforward using line-of-sight measurements from one Doppler lidar and it requires assumption of spatial homogeneity of the flow or the use of a model of the spatial variations of the flow. Please explain in which context the correction by the inverse cosine of delta is discussed.
- Line 710 – 712. It is stated that the machine learning technique requires 1 second to evaluate an estimate on personal computer. What is the duration of the data set that this processing time corresponds to?

Technical Corrections

- Line 7. I suggest replacing "error" with "errors"
- Line 15. I suggest replacing "an overlapped" with "an adjacent". I understand that the authors mean that the meteorological tower was at the same location as the scanning pattern, but the adjective "overlapped" without additional information is a bit confusing here.
- Line 33. Can you add here what type of models you are referring to?
- Line 36 – 38: Please rephrase this sentence. The uncertainties in lidar measurements cannot stem from the modelling approaches for the reconstruction of the velocity vector. It is rather that uncertainties in the estimated radial wind speeds that propagate to uncertainties in the reconstructed wind vector.
- Line 41. The spectrum presented in Figure 1 corresponds to a Doppler spectrum that is acquired by a continuous-wave (cw) Doppler lidar. I suggest stating here that this is an example of a cw lidar.
- Line 43. Add "the" to "...are related to the line-of-sight velocity...".
- Figure 1, label. I suggest replacing the "Example power spectral density.." with "Example of a power spectral density.."
- Line 64. Add "the" to "...sources in the measured lidar..."
- Line 66. "...instrument errors ..."
- Line 67. "... nacelle-mounted lidars"
- Line 72. "... is a fast scanning ..."
- Line 72. "... CW lidar as have been mounted..." please edit this sentence, it looks that something is missing
- Line 73. The reference Mikkelsen et al. 2013 refers to a field study with a lidar installed on the spinner of a wind turbine, which did not sample at 500 Hz.
- Lines 192 – 194. The authors write "Line colors correspond between each of the three subfigures..." . Please clarify this sentence.
- Line 248. Replace "is" with "has".
- Line 262. Add "using the Levenberg-Marquardt method".
- Line 283. Add "the model refinement"
- Line 377. What is meant with the "full weight"?
- Line 409. More than how many orders of magnitude?
- Line 411. The units of the values presented in Table 1 and 2 should not be written in italic fonts.
- Line 448. Remove one of the two section references.

- Line 451. Add "...on the ensemble..."
- Line 650. Add "than the inflow cases".
- Line 721. Add "the three lidar processing techniques"