

The current version of the manuscript "Enhancing turbulent fluctuation measurement with tailored wind lidar profilers" has been fully revised and is substantially different from the version initially submitted. The most significant change is that now only one prototype lidar with faster sampling rate is being tested against the commercial version of the WindCube 2.1 and that a great effort has been put into quantifying the effect of noise on the measurements. The narrowed scope appears reasonable and has the potential to result in a better paper.

A strength of the manuscript is the work put into a quantification of noise in the LOS data from the WindCube. Though, it is unclear why the authors did not write about the CNR value that is provided by the instrument as a standard parameter and could have been used to compare the noise levels of the prototype lidar and the commercial version.

The main weakness of the manuscript lies in the assumption that a reduction in sampling frequency leads to a reduction in variance of the measurement data, which is not true. This false assumption that increasing the sampling rate could capture additional energy associated with smaller eddies leads the interpretation of the experimental data into a wrong direction. Instead, more focus should be put on the relationship of intra beam and temporal averaging and how it is influenced by the prevailing mean wind speeds.

In its current form, the manuscript is not ready for being accepted by WES and it should be reconsidered after major revisions. Please note that the following comments are not capturing all aspects that should be improved and that a revision should be done with care before submission.

Specific comments:

Response to the reviewer:

It is good practice to acknowledge the referee's effort put into reviewing the manuscript. The authors missed this opportunity which is discouraging. Further, the response to the reviewer is suffering from mistakes, e.g., "We are right." instead of "You are right." and statements that are not covered in the updated manuscript, e.g., "We have implemented logarithmically spaced [sic] frequency bins [...] See Fig. 7b".

1.:

The introduction gives some valuable insights into the history behind the topic, but the state of the art is insufficiently covered. Please add the most relevant and significant findings from the cited literature instead of just listing it in groups. The section should end with a guidance through the structure of the paper.

2.1:

Include a table with a comparison of the two lidar configurations showing parameters like sampling rate, accumulation time per LOS, number of samples per 30 min, range gate...

2.1:

The authors should reflect on the relationship between the industry demand for TI data (10 min.) and the variance of the u-component of the wind (30 min.) provided by the methods described in the paper.

2.2.1:

This subsection is the only content of subsection 2.2. This does not make sense.

2.2.1:

There are wind turbines only 210m away from the lidars, so there is no "undisturbed winds from almost all sectors". Please explain if only wind from the wind turbine's upstream direction was used in the study.

2.2.1:

Please describe the purpose of creating smaller subsets of data sampled at 0.25Hz and 1 Hz respectively. If the sonic was configured with higher sampling rate, the entire dataset could be used with 0.25Hz and 1Hz. This is unclear.

2.3:

It is wrong that Kelberlau and Mann (2020) recommend to not fit lidar-derived reconstructed velocity component data to turbulence models. They do it in their study, are satisfied with the approach and think it clarifies lidar-specific effects of turbulence sampling.

l. 245:

Provide information about the "alignment condition". What range in degrees is accepted to end with 17.1 % of the data? Is this including wind from beam 3 to beam 1, downstream of the wind turbine?

2.6:

The authors should not just claim "DNV-GL has defined acceptance criteria" but refer to the source explicitly.

2.7:

The verbal description of the quality parameters (RMSE, MAE, R2, rel. error) should be accompanied by equations that define them unambiguously.

3.2:

The description of the amount of variance included in different frequency ranges might be correct. But the conclusion that by a higher sampling rate could capture an additional percentage of the energy associated with smaller eddies is wrong. Sampling with too low frequency leads to aliasing and in a spectral display the energy from higher frequencies is folded into the lower frequency range. Instead, more focus should be put onto the relative influence of the temporal averaging caused by lower the accumulation time of the prototype lidar. Averaging does decrease the LOS variance.

3.4.1:

It is unclear why the CNR value as determined by the WindCube is not used as an indicator for the instrument noise. The median variance from spectral method for the prototype (0.0129) is also approx. 1.5 times higher than the corresponding value from the ACF method (0.0081). It is not twice as high as written in the manuscript.

Fig. 7:

The caption should be revised to explain the different purpose of subfigures (a) and (b). Also, describe which LOS direction has been used (5, vertical?)

3.5:

If the mean standard deviation is 2.9% higher, the corresponding variance must be 5.9% higher. It is unclear why the authors report 7.2%?

4:

The discussion refers to the impact of the prototype configuration on TI but it does not critically reflect on it. What happens to TI estimates if for example the v component of the turbulence wind field becomes significant, when the inflow is not aligned with one of the beams?

I suggest reducing the discussion of the potential of the prototype lidar for floating lidar systems to one sentence since floating lidar systems are not within the scope of this study.

Technical corrections:

l. 92: "True North" is wrong here because the lidar is rotated.

2.2.1: $450+1800=2250$? What happened to the remaining 6 intervals?

2.6 and other occurrences: DNV-GL does not exist anymore. They are named DNV now.

2.6: Refer to Table 1 and include availability thresholds.

l. 298: Replace "almost similar" by "similar"

l. 358: Replace "bin-averaged" by "wind speed-binned"

l. 408: Replace "relatively slight" by "slight"