General comments

This paper addresses lidar improvements with a neat comparison of two new lidar prototypes with commercial systems. The finding that a reduced sampling rate is the best improvement is however poorly supported by the data.

The main drawback of the work is the lack of reference turbulent quantities to compared with. One of the systems was deployed close to a sonic anemometer but this valuable instrument is deliberately omitted. The basis on which improved turbulence estimates are claimed are mainly two and not convincing:

Increased variance with respect to the reference lidar is by itself not indicative of improvement. As also mentioned in the introduction, lidars can overestimate variances due to cross-contamination, so how do we know that the increased sampling rate is not indeed exacerbating a positive bias in the variance? Increased variance could also come from noise, and this is has not been ruled out either.

Exactly. Lidar can either overestimate or underestimate the variance of along-wind velocity (u) when the variance is computed from the reconstructed velocities provided by the lidar. This is known as the cross-contamination effect. To mitigate this, we focused our analysis on cases where the wind is aligned with one pair of opposite beams (beam 1/beam 3 or beam 2/beam 4) and computed the along-wind variance by combining the variance of the LOS velocities from these beams. In the specific case where the wind is aligned with the pair of beams 1 and 3, we have $\sigma_x^2 = \sigma_u^2$. Conversely, when the wind is aligned with the pair of beams 2 and 4, it holds that $\sigma_y^2 = \sigma_u^2$. Also, under these conditions, it can be reasonably hypothesized that the covariance term, σ_{uv} is negligible (e.g., Newman et al., 2016), where v represents the cross-wind velocity.

We have also corrected the variance for noise, which was quantified using two approaches: a spectral approach and the autocorrelation approach that you suggested.

The reduced noise estimated from the spectra of w is also not compelling. Increasing sampling rate extends the spectrum to higher frequency (Fig. 9), so the behavior of the fitting can change significantly. It is also mentioned that for the commercia lidar a noise plateau was not identified, so we cannot trust noise estimates from the reference lidar so what observed in Fig. 10a can be a numerical artifact

Also, the spectral analysis shows that spectra are very noisy and therefore the results should be interpreted more carefully. For instance, the laminar flow case in Fig. 11 is very questionable as laminar flow generally does not occur in the field and also because the supposedly laminar spectrum has more variance than the turbulent spectrum.

It is suggested to profoundly revise this work to make the most out of this useful dataset:

Thank you very much for your thorough review and valuable feedback. Initially, the study focused on evaluating the impact of two modifications to the lidar system: (1) reducing the probe length and (2) increasing the sampling rate of the WindCube v2.1 lidar profiler. These modifications were assessed separately for their influence on turbulence measurements. The updated version of the manuscript now focuses exclusively on the increased sampling rate, for which a 47-day dataset was collected. This dataset is accompanied by reference turbulence measurements provided by a sonic anemometer installed on a nearby met mast. The lack of reference data in the original manuscript was another concern raised by the reviewer. In this revised version, we incorporate the sonic anemometer dataset for more robust comparisons.

In the updated version of the manuscript, we narrowed the focus to examine the effect of the increased sampling rate on the variance and standard deviation measured by both lidars, as these values are used to compute turbulence intensity (TI), the most commonly used metric in the wind power industry to quantify turbulence. In the first version of the manuscript, comparing multiple turbulence metrics (such as dissipation rate and integral length scale) created confusion, so we chose to concentrate on variance and standard deviation. These estimates were corrected for the variance of instrumental noise, which was quantified using two methods: a spectral approach and an autocorrelation approach.

We estimate that the updated version of the manuscript is approximately 90% revised compared to the original version.

1. Calculate the turbulent statistics form the sonic (or even cups) as well and use it as reference

The updated version of the manuscript now focuses exclusively on the increased sampling rate, for which a 47-day dataset was collected. This dataset is accompanied by reference turbulence measurements provided by a sonic anemometer installed on a nearby met mast. In this revised version, we incorporate the sonic anemometer dataset for more robust comparisons and compute error metrics such as MAE, RMSE and relative to compare, along-wind variance and standard deviation derived the commercial and prototype configuration in comparison to the reference sonic anemometer measurement.

Moreover, the paper now addresses the impact of the increased sampling rate on data availability and key performance indicators (KPIs), such as the slope of the scatter plot between the mean wind speed measured by the lidar and that measured by the reference, the correlation coefficient (R^2) of the linear regression of the scatter plot, and the mean absolute difference in mean wind speed.

2. Do not provide overall biases only, but also RMS error on a 10-minute basis or, even better, scatter plot like the one in Fig. 5 for lidar vs sonic

We have included the RMS error along with additional metrics such as MAE, relative error, bias, and R^2. Additionally, we have added a scatter plot comparing the standard deviation obtained from both the commercial and prototype lidars against sonic anemometer measurements (see Fig. 9 in the updated version).

3. For the lidar with reduced probe volume where there is no met mast and very few data points, consider a smaller section with a lot of caution advised in the interpretation of the results

You and the first reviewer strongly recommended exercising caution when drawing conclusions related to the reduced probe length, given the very limited dataset (only 4 days) and the lack of reference measurements (e.g., from a sonic anemometer). In response to this feedback, we decided to remove the analysis on the reduced probe length. This is an ongoing work that is not ready to be published.

4. Evaluate lidar noise also using a non-spectral approach, like the autocorrelation method by Lenschow et al., 2000 (https://doi.org/10.1175/1520-0426(2000)017<1330:MSTFOM>2.0.CO;2)

Thank you for the reference. We have implemented the ACF method and included the results in the revised manuscript. Additionally, we have compared the spectral and ACF methods for estimating noise variance (Section 3.4.1, page 15). Our analysis shows that the spectral method yields a median variance 1.5 times higher than the ACF method for the commercial lidar and twice as high for the prototype lidar, highlighting differences in how each method characterizes noise. However, the spectral method also estimates a mean instrumental noise 30–40% lower than the ACF method, indicating variations in noise quantification. Moreover, the spectral method results in a significantly narrower spread of mean values, particularly for the commercial lidar, where the spread is reduced by half compared to the ACF method. This suggests a potential advantage in terms of consistency and stability. Based on these findings, we used the spectral method to correct the measured variance, as it provided more stable estimates of instrumental noise.

5. The introduction could mention the effect of pulses accumulation, which is different from the sampling rate. The accumulation acts as a low-pass filter in the time domain in an analogous way as the probe average does in the spatial one. The sampling rate refers more to how quickly the lidar moves through the scan cycle, regardless of how long it takes to measure a single LOS.

We have added this text to the introduction to address your recommendation:

"The intra-beam effect generates underestimation of turbulence metrics. It arises from two anisotropic filtering processes: (1) spatial filtering due to averaging over the probe volume and (2) temporal filtering caused by averaging over the beam's pulse accumulation time, Δt, at a given measurement position. These two effects give rise to a transfer function, H, applied by the instrument on the signal measured within the probe. The transfer function includes a part due to time-averaging (the sinc term) and a part due to space-averaging (the Gaussian term), such that (e.g., Kristensen et al., 2011):" Lines 43-47, page3.

"Pulsed lidar profilers require several seconds to complete a full scanning cycle resulting in a low sampling rate that causes discrepancies between turbulence measurements taken by anemometers and those by lidar profilers (Pena et al., 2009). While the sampling rate governs how quickly the lidar progresses through a scan cycle, it is directly influenced by pulse accumulation time". Lines 58-61, page 3.

These are some modifications that would bring the paper to the standards of the other publications in the topic.

Thank you for your feedback. We have implemented all the suggested modifications to ensure the paper meets the standards of other publications in the field. The revised manuscript includes the recommended analyses, additional metrics, and methodological comparisons to enhance its rigor and clarity.

Specific comments

L71: "mea" instead of "mean"

Corrected.

L77: is the increased sampling rate achieved through a faster accumulation or a higher pulse repetition frequency? In the second case, the maximum range may be reduced, and it should be explained.

The increased sampling rate is reached through a reduction of pulses sent into the atmosphere. Lines 101-104, page 4.

Please add Fig. 1 angles and axis clearly indicated for readers that are unfamiliar with this technique.

We have added the positions of beams and the axis, x, y and z.

L 94: please explain what the test requirements were to consider it as "passed".

We removed this part and presented some of the test requirements. The paper now addresses the impact of the increased sampling rate on data availability and key performance indicators (KPIs), such as the slope of the scatter plot between the mean wind speed measured by the lidar and that measured by the reference, the correlation coefficient (R^2) of the linear regression of the scatter plot, and the mean absolute difference in mean wind speed. (Section 2.6, 2.7 and 3.1).

Equations 1 and 2: b terms that should be the LOS velocities are not defined.

We defined the terms (Eq. 5-6).

Fig. 5: please add the colorbar of data density.

Done.

Not addressed anymore in the updated version of the manuscript.

L81: sampling rates of 0.25 Hz for wind speed may be misleading. The lidar uses a moving averaging window of 5 beams, so it does deliver a new wind speed estimate every second, but these estimates are not independent. This time overlapping effect should be made clear.

L130: the explanation of the rotation of velocity is unclear. In general, Vx and Vy are not 0, but after rotation v=0 (not Vy as indicated). Aligning the x axis to North is also not the common practice in atmospheric science, where x is W-E and y is S-N, and it may be worth mentioning this as well. Please add Fig. 1 angles and axis clearly indicated for readers that are unfamiliar with this technique.

L201: it is true that the inertial subrange is limited to the right by the viscous regime where dissipation reduces TKE, but it is also limited to the left by the integral scales that supply TKE, please add this detail.

Eq 9: the | symbol to indicate the range of frequencies may be mistaken for an integration. If a fit is instead performed in this region, it would be better to remove it and explain that it is a fitting operation in the inertial subrange.

L245: is the specification of 1% relative to the error over 10 minutes or the whole dataset? Please specify.

Fig. 6: please make the box and whisker format consistent between the two subplots.

L255: have you considered that the increased difference close to the ground may be due to the lidar with reduced probe length being able to resolve better nonlinear mean wind shear?

L272: the increase in interquartile range cannot be automatically ascribed to a better sensitivity since it could very much be noise (instrumental or statistical). The fact that larger increases in standard deviation are seen at high altitude is also suspect in this sense, since one could expect the reduced probe length to lead to more recovery of turbulence variance close to the ground where length scales are smaller. If it happens at larger range, it could be noise not sensitivity.

L283: "iterative" may not be the right word, "trial and error" maybe?

L330: it is confusing saying that β=5/3 was imposed for the dissipation energy, but then β<1 were excluded. Is this a two-step process where first we fit β to the whole spectrum, then if it passes the check it is used for the dissipation energy with a new fit in the inertial subrange and β=5/3? Please clarify.

L359: the integral length scale is not associated with a peak in the spectrum (not premultiplied), but it is by definition its value at 0 frequency, as shown in Pope 2020, Eq. 3.114. Please remove or rephrase.