Review of the manuscript wes-2014-108 entitled "On the lidar-turbulence paradox and possible countermeasures" by A. Peña, G. G. Yankova, and V. Mallini

This article addresses LiDAR-turbulence paradox, i.e. the underestimation of the LiDAR-based turbulence intensity due to the probe volume averaging. Based on the data collected during a months-long campaign by two continuous-wave profiling LiDARs and four sonic anemometers, the authors adopt a neural network (NN) algorithm to correct the LiDAR's along-velocity variance. The problem is addressed at three different levels: first, utilizing a physics-based turbulence model to efficiently reproduce the LiDAR-to-sonic turbulence ratio; second, the physics-based model is used to train a physics-based NN (PBNN) on LiDAR turbulence data to correct the probe-volume averaging effect and positively assess them against the sonic anemometry; finally, the data from one LiDAR are used to train a data-driven NN (DDNN) to correct the turbulence prediction from a second co-located LiDAR and compare the latter against sonic anemometry. The article is well-written, the results are encouraging, and the limitations of the current analysis are thoroughly discussed. Thus, I recommend this article for publication after few major revisions I have detailed in the remainder of this review.

General comments

- Eq. 1: If you assume a right-handed reference frame (i.e. with the z-axis pointing upward), the radial wind speed is: $v_r = u\cos\phi + w\sin\phi$. Please cross check this equation.
- Eq. 3 (Line 115): Should the argument of the complex exponential have a negative sign?
- Line 124: I would explicitly write the Mann's model of the spectral tensor.
- Lines 133-134: Which values of L and dissipation rate did you use?
- Line 164: To improve the overall clarity of the manuscript, I suggest moving the "Methods" section after the site description section, since the content of Section 4 is mentioned several times in Section 3.
- Line 197: Please provide a reference for the estimation of the dissipation rate.
- Line 261: I believe a schematic of the experiment (or pictures) is useful to the reader to better understand the remainder of this study.
- Line 265: A maximum time lag of 10 s is acceptable when second-order statistics are cross-validated among different instruments. However, I would not mention "time synchronization" for this study as, to my understanding, the different instruments were not sampling based on a common GPS time stamp.
- Line 306: The agreement between mean velocity measured by sonic and Lidar looks excellent. Please add the slope, intercept and coefficient of determination to quantify the agreement.
- Lines 308-312: In Fig. 4, the information may be better stated quantifying the linear regression parameters between lidar- and sonic-based variance. You can just provide one global value for each height (regardless of the stability and length scale), or specific values

for each data subset. I leave this up to the authors' decision, but I believe that the initial comparison between sonic and lidar variance estimates requires a thorough quantification.

- Line 347: Previously (Line 320) you mentioned that a common value of $\Gamma = 3$ is used across all stability and length scale cases. Considering that the thermal stability influences the isotropy/anisotropy of the energy-carrying turbulent structures, would you observe any difference in Fig. 7 letting $\Gamma = 2$ or $\Gamma = 4$?
- Lines 388-389: Do you have an explanation of why the case with the lowest number of inputs (P4) returns the best agreement? During the cross-validation of the PBNN the result was opposite (case P8 in Table 1).
- Lines 394-398: In Fig. 12 (and following) I would rather just plot the error bars of the along-velocity variance against the wind speed, instead of the TI (which is not independent from the mean wind speed). The error on the TI inevitably carries the error on the mean velocity (tough small), whereas here you want to highlight only the error on the second-order turbulence statistic.
- Line 406: I agree that having a large number of samples within the $5.25 - 5.50$ m s^{-1} bin improves the agreement with the sonic-based values. However, I see that you have a good agreement also for the highest bins, where the number of occurrences is much smaller. Thus, I don't think that the occurrences are the only reason why the agreement is poorer at low wind speed. My suggestion is to separate the along-velocity variance from the wind speed and see how the results change (see my previous comment).

Technical comments

- Table 3: The header of the last column should report $\frac{z}{L_O} \leq -0.05$.
- Figure 9: I do not understand the unit on the ordinate axis. The figure's caption states that the histogram is normalized, but the sum of all the bar values is clearly greater than 1 (or 100%). Also, please change the upper limit of the *y*-axis to show the exact values of the left-most bar.
- Line 374: I found this sub-subsection a little bit dispersive in terms of results visualization. If the authors think it is possible, I suggest concentrating the top panels of Fig. 12, 14, 16 and 18 in a single figure, the bottom panels of these figures in another one and do the same for Figs. 11, 13, 15 and 17. The current back-and-forth between the text and the figures makes this portion of the manuscript a little bit hard to follow.
- Line 410: Please replace “that of two” with “the ones of the two [...]”.
- Line 471: Please correct the equation: $U \leq 5$ m s^{-1}.
- Line 526: Before summarizing the results of this study, it is useful to start the Conclusion with a summary of the experimental setup and data analysis methods used in this work.