Review of: "Coupling physics-informed neuronal networks with 3D scanning pulsed Doppler lidar"

This paper used a Physics-Informed Neural Network (PINN) to attempt to reconstruct a wind field using sparse measurements from a scanning doppler lidar. The authors present a method for filling in missing data using the polar Navier–Stokes equations as boundary conditions. The method is evaluated for two cases: 1) reconstruction of missing lidar data and 2) extension of lidar range using integration and reconstruction over sparse regions.

I am not convinced that this method accurately estimates the winds, and I am unsatisfied with the presentation of the statistics that assess the accuracy and precision. The authors need to take more care in how they evaluate the reconstructed winds vs. measured winds, how they assess separately the uncertainty, variability, and bias, as well as how sampling and range contribute to the final intercomparison. I suggest that a few of these major issues be addressed before publication.

Major Comments:

1a) Much more care needs to be taken when comparing the "true" measured winds to the reconstructed winds. The generic term "errors" is encompassing too much and makes reading this paper frustrating. Please be sure to use consistent and clearly defined terms to separate between:
"Measurement Uncertainty" (Poisson, instrument, and retrieval error),
"Residuals" (difference as a function of range and azimuth between the measurement and reconstruction)
"Bias" (systematic tendency in the residuals between measurement and reconstruction),
"Accuracy" (comments on the magnitude of the Bias or Residuals)
"Precision" (comments on the magnitude of the measurement uncertainty)
"Significance" (assessment of the bias with respect to the precision of the measurement)

1b) Use RMSE instead of MSE, as it will put your results in the same units as the measurement (m/s). It is more intuitively useful for the reader to know the bias rather than the square of the bias, which has no physical meaning. Figure 4 is extremely misleading! Must be range dependent or quoted at a few characteristic ranges.

1c) When presenting the final results (line 12 and conclusion) the median plus/minus the standard error should be given, not the 98th and 99th percentile. Currently, it presents a skewed picture to the reader and makes it hard to assess whether the comparison is actually any good.

1d) Line 45: You have reported a scalar 'accuracy' of less than 0.5 m/s. This tells me nothing as a reader. Given the higher data density closer to the lidar, a single value for bias will be range-weighted. Accuracy needs to be assessed as a function of range and azimuth.

2) Lines 73-75: I'm uncertain why these particular values for characteristic length and velocity are chosen. Please consider adding more text to explain your choice.

3) I'm unclear about your Loss Function in Eq. 3. Why have you chosen to minimize the MSE rather than a more standard technique of Least Squares. As the Loss function is written, it appears to always minimize the sum of two positive numbers based only on the boundary condition lidar profiles and the retrieval. This doesn't seem like it will work very well at larger lidar ranges where there is a significant gap between the data. It could lead to results where the data contributes very little.

4) Figure 5 looks very strange. The reconstructed wind field does not match the measurements on the left and has an "unphysical" polar coordinate rectangle shape for the yellow wind gust. This tells me that there is insufficient data to influence the retrieval grid. A second quality check that could be done is looking at the left panel at the beams near 257 and 309 degrees. At the outer ring, the wind speeds are low (teal-green, 0-5 m/s). In the PINN the outer ring on this segment starts at yellow and goes to green. Why does the PINN not match the boundary point values provided by the data?

5) Figure 6 (lower panel for 20h45) makes me very skeptical about what you are doing!!! The PINN reconstruction for 12 and 16 beams has a very large (> -15 m/s) radial wind speed around the 257 degrees mark that doesn't exist in the data!!!!

6) Multiple times while reading the paper, I wanted to ask how does this PINN method compare to simply interpolating between adjacent lidar profiles? Does PINN give a meaningfully better result than simple interpolation?

<u>Minor Comments:</u>

L13-14: Please use seconds to be consistent with later text (L12 and L111)

L26: Please cite a paper regarding a typical lidar to which this method could be applied.

L28: Please cite a paper and tell the reader what the safety requirements are. Do airports require wind information every second?, 10 seconds?, every minute? And at what resolution? This helps the reader evaluate the required resolution of the lidar scans. For example, the system is over-engineered if you produce a complete scan every 1s at 1.5 m range gating, but the airports only require 60 s at 100 m. Then it's better to drop the resolution and speed of your system and reduce the Poisson error. Same comment on L97 and L174: How fast does the system need to work to satisfy user requirements?

L51: "extent" → "extend"

All figures were a bit "granular" on the PDF. Please save figures at a higher resolution (300 dpi) and include larger fonts on axis labels

L112: What is a “lidar scheduler”?

Figure 3 (left) lacks relevance or motivation. Does this tell me about range limitations in the hardware for signal acquisition, local geographical gaps, or about the algorithm?

Figure 3 (right) is OK, but why does it need a figure? It's good enough to describe as text

L142-143: This sentence is a bit unclear. Please consider rephrasing.

Figures 5 and 6 need clearer range information included either in the figure or caption.

L160-161: You could try to mitigate this by offsetting the pointing direction on each rotation (ie. non-harmonic of π azimuthal step size). Similar to a satellite with a precessing orbit.

Figure 1 is not useful. You have tried to fit too much, and it is difficult to interpret. I had to spend time to understand the figure, rather than the figure aiding my understanding of the text. Please consider revising it to something simpler. I only need to know: Measurement (Θ, r, t) → “black box” → Retrieved (v_r, v_Θ, P) using Navier-Stokes (NS) and a loss function. The NS equations are better presented in Eq. 1 and Eq.2 than in the figure.

Table 1 doesn’t add to the paper. The information is in the text. Please move to an Appendix or remove

L175: In essence, you have a reliable data product at short-range, then you use some sparse scans to attempt to extrapolate further. There is nothing wrong with this, but perhaps adding a red line in the polar plot to mark a “reliability” limit might be helpful. Calculating and showing your median bias and standard error will allow you to assess over what range your reconstruction is significantly equal to the measurement.

L183: Is this future work? You can already test if parallelization gives a large performance boost on your laptop. I generally test/develop lidar code locally on 4 to 6 cores using my work laptop, then allow many more cores on the server for batch processing years of data. It is very straightforward to try.

L195: Section 1.3?

L197-200: A bit unclear. Please consider rephrasing. Compressed sensing wasn’t mentioned in the introduction, so it doesn’t really fit in the discussion.

L200: Section 1.5?

L230: I have no idea what the “Cyclops dilemma” is. Either explain it in the introduction or remove it.

L231-232: The last line of your conclusion is rather weak. Comparison with (Zhang and Zhao, 2021) should be in the discussion section.