Referee's comments to gmd-2022-150

General comments

This paper discusses the results of a comparison of wind profiles between the profiling lidar measurements and numerical weather predictions. The topic is very relevant, and the information provided by remote sensing is extremely valuable to improve the state-of-the-art forecast models.

The literature is fairly reviewed, although there is no mention of the limitations of profiling lidars, i.e. the along-beam average and the assumption of horizontal homogeneity. It is recommended to add a paragraph in the introduction discussing existing studies of errors introduced by complex terrain or front passages on the retrieval of mean quantities from profiling lidars (e.g [1,2]).

The experimental dataset and the model are described thoroughly, and the effort devoted to the development and discussion of the data quality check is commendable. The description of the error metrics would require additional details (see below).

The analysis of the results is interesting, but the section regarding the daily and seasonal statistics of the error can be improved as follows:

- By adding the error on the mean or stricter statistical error check to ensure that the random noise is identified or removed before discussing physical trends
- By expanding the discussion on possible causes of the seasonal variation of the error observed at SGP and Graciosa

Provided that the previous points are addressed (which are deemed as minor reviews for such a comprehensive study), the paper is suitable for publication in this journal, in the Reviewer's opinion.

Specific comments

Line 50: Please replace "lack temporal resolution" with "are affected by limited data availability". In fact, the radiosondes can have a good temporal resolution within one single launch.

Lines 132-133, "Therefore…Doppler lidar": please discuss possible errors introduced by the vertical upsampling of simulation data.

Line 135: Please justify the choice of the SNR threshold.

Line 136: Please provide a reference for the speckle filter.

Lines 180-181, "This…surface": It is unclear why the uncertainty in the balloons' position may lead to higher error near the ground, as the sondes generally drift further away at higher altitude due to the stronger winds. If the balloons' launch sites are not collocated with the lidar locations, then the error can be due to the larger horizontal variability of wind speed close to the surface due to the effect of the local terrain. Please clarify.

Line 195: Please provide an explicit formula for the mean absolute wind vector error.

Figure 6: Please add error bar equal to the monthly standard deviation to give a sense of the inter-annual variability.

Figures 7, 8: In some cases (Kumpula, Cape Cod) the daily patterns do not emerge significantly on top of the statistical noise. Please add a colormap with the error on the mean of the error metrics (see [3]), or reject data affected by an error on the mean higher than a reasonable threshold.

Figure 8, 10: Please use a colormap suitable for data containing positive and negative values (e.g. BlueWhiteRed in Matlab, ReBu in python) for easier readability.

Lines 237-238, “Kumpula…land”: this statement could be substantiated by plotting the dependence of the error on the wind direction.

Line 288: Please correct the typo “an LLJ”.

References:

[1] Bingöl, F. (2009). Complex terrain and wind lidars.

[2] Klaas-Witt, T., & Emeis, S. (2022). The five main influencing factors for lidar errors in complex terrain. Wind Energy Science, 7(1), 413-431.

[3] Zięba, A., Ramza, P. (2011). Standard deviation of the mean of autocorrelated observations estimated with the use of the autocorrelation function estimated from the data. Metrology and Measurement Systems, 18(4), 529-542.