Summary

The authors present an evaluation study of wind profile forecasts in the boundary layer, from the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF), against winds retrieved by Doppler Lidars. Lidar data from 6 locations with different characteristics are used, while the retrievals are primarily validated using collocated radiosonde profiles. Traditional methods, consisting of various statistics, such as the calculation of mean absolute errors and wind speed bias, are used to compare observations with forecast. The authors show that IFS winds are more accurately predicted in marine and coastal locations (Graciosa, Cape Cod and Kumpula), while errors increase at stations with more complex topography. Moreover, the authors present statistics on seasonality and discussion on the effect of local characteristics (Low Level Jet at SGP, Katabatic winds at Granada).

The study is very relevant to the journal and the results are valuable to improve the state-of-the-art forecast models. The manuscript is well-written, carefully organized and provides the right amount of detail for the reader. The authors conducted a study with novelty and clarity; thus, I recommend for its publication, with only minor comments and revision (see below).

General comments

The manuscript mainly discusses the difference between the wind lidar measurements and the modeled ones. In the relevant plots, only the biases are plotted and discussed, without a mention on the mean wind values, which could be used as a reference of the importance of the absolute bias. Can the authors explain why this is not included? The authors could consider to include in the manuscript one of the following: (a) the mean wind speed values (by the model or the lidar), when they discuss their absolute errors/biases; (b) a scatterplot of the lidar vs the modeled wind speeds for the different stations and altitudes (similar as Figure 3); or (c) a diurnal colorplot of the mean wind speed for the different stations (similar as Figure 7).

A math appendix with the equations used for the calculation of the different parameters and their errors would be useful to the readers (e.g. for the wind speed and direction calculated from u and v components, the B, MAE_S, MAE_D and MAE_V).

It is good to mention somewhere in the manuscript that (a) all elevations in the plots are above surface level and (b) the authors used the IFS outputs, which are above surface elevation, to compare with the relevant lidar measurements. If this is the case.

Specific comments

At Lines 68-88, and in table 1, the 6 stations are described. It would be of interest to the readers to include the elevation above sea level of the Lidars (e.g. in table 1, next to the coordinates). The terrain altitude variability around the stations, in the IFS resolution, would be an interesting information also.

Lines 96-104: Could you comment on the reason why in different locations the VAD scans were scheduled in different angels? Is it related with the vertical resolution and max altitude of interest in each site, or something else?

Lines 109-110: The syntax can be improved to read the sentence better. Consider revising to this end.

Lines 110-113: Please include a reference on the ECMWF IFS HRES model used. Furthermore, some additional information for the model could be interesting for the reader. Namely: (a) how many (approx.) of the 137 vertical levels are below 1km, since PBL winds are discussed in this study, (b) what is the range of the vertical grid spacing in the PBL heights, (c) the information that IFS HRES runs on a Gaussian grid or (if available) the horizontal resolution in degrees.

Lines 115-117: Consider including a reference on these assimilations.

Line 136: Speckle filter: add a reference on this filter used.

Lines 143- 145: Rain filter: in case this filtered is used in the past for another study (known to the authors), they could include a reference of prior use when describing this filter.

Figure 2: I suggest to present the different sub-plots with the same order as discussed in the manuscript (so the speckle filter (c) to be before the residual filter (b)). Additionally, consider including titles at the colorbars (and at y-axis in d) or subtitles with the information of each filter in the subplots.

L 175-177: It would be good if the resolution of Wind Lidar is also mentioned here, to highlight the difference between Radiosonde (10m) and Lidar (30m) resolution.

Figure 3: It would be of interest to add the station names above the 4 columns of the doppler lidar wind direction plot.

Section 3.2: It would be interesting to include the information of the distance between the Radiosonde Launch points and the Wind Lidars for the 4 sites used.

Line 180: “This was expected as the radiosonde launch may still be impacting the balloon track”. This sentence is not clear. Please rephrase to include why this is more relevant in the lower altitudes.

Lines 186-189: At 117m, in Darwin, all wind speeds bellow 15m/s (not only the ones bellow 3m/s) have a significant different between the lidar and the radiosonde (in comparison to the differences in the other stations/altitudes). So please include a quantification of the agreement discussed in the sentence: “Fig. 3 shows that if low wind speeds < 3 m s−1 were ignored, the agreement would also be much better at Darwin”.

Line 211–212: Could you comment on the reason why the errors are larger during summer and lower during winter for SGP? At Line 217, you mention a weaker seasonal cycle for Graciosa. It is also an opposite cycle, with larger errors during winter and lower during summer. Could you comment on the reason for this contradiction? Is this connected with relatively higher/lower wind speed values in the periods of higher/lower MAEv in the two sites? Did you investigated also the Mean Absolute Percentage Error (MAPE) for these two sites, and if so, do you observe the same seasonal variability?

Figure 6: Please consider adding error bars in this plot.

Figures 7 and 8: It would be helpful for the reader if you can include sub-legends with each station name/location in the sub-plots. Also, a symbol indicating the local solar time in each site discussed in the manuscript) would be useful.

Line 236: “...indicating the influence of a growing boundary Layer”. Do you imply that the growing boundary layer is not correctly captured from the model, or that the PBL winds are modeled with higher biases? Please specify if possible.

Lines 265-266: Please improve this sentence to read better. Maybe break it to two smaller ones.

L 288-289: Can the authors briefly describe the method according to which they detected LLJ? The reference provided, shows the clarity of the data processing, but it would be easy for the reader to have a short briefing of the method.

Technical comments

Line 119: “surface Hersbach (2010)”: correct the reference syntax.

Line 158: “filter is more active”: consider revising to "more effective”.

In Figure 8: A diverging colorbar separating the negative from the positive values of the wind speed bias (for example red=negative, white=0, blue=positive) would be really helpful for the reader. This way, it would be easier to point out the diurnal evolution characteristics of PBL mentioned at 231-238.

Lines 113-114: Please improve the syntax of this sentence.

Curtain plots: What does the white space at the bottom of each curtain plot represents? Is it the height that Wind Lidar starts to measure? If so, it would be interesting to include this information in the general description of the systems.

Figure 12: Please add labels: “Wind lidar” for (a), (c) and “IFS” for (b), (d). Here, a diverging colorbar would also be very helpful.

Line 323: “... to observed”: I think “to” need to be omitted.

Figures A1, A2: Please add labels with the locations.