Dear Reviewer,

First of all, we thank the reviewer very much for agreeing to act as a reviewer and for investing their time to review this manuscript. The reviewer’s constructive critique is much appreciated and their comments are very valuable to us. Some excellent points have been raised that we did consider in the revision process.

Response to general comments:

1. The implementation of the virtual lidar in the LES seems to be able to change the azimuth and elevation angle of the lidar beam instantaneously. Depending on the lidar type, the scanner head can have a significant travel time resulting in either shorter dwell times or the measurements covering an azimuth range instead of a single point. I believe this behavior could be added to the virtual lidar with either pauses between the sampling of the lidar beams or by averaging a distribution of lidar beams within an azimuth range. At the least, this aspect should be touched upon in the discussion.

Yes, it is correct that the scanner head of a lidar either moves continuously, resulting in a measurement over a range of azimuth instead of a measurement along a single azimuth direction, or spends time changing positions, resulting in a shorter dwell time. We added a remark that this way of operating the virtual lidar in the LES does not fully represent the scanning regime of a Doppler lidar in reality. However, one should be aware of the fact, that a lidar works with pulse repetition rates of several kHz, a sampling we will not be able to simulate. In the end the lidar provides one radial velocity measurement every 5 seconds. We think that averaging the LES output over 25 values within 5 seconds is a fair representation. We added this argumentation to the description of the virtual lidar scanning regime.

2. The methods sections suffers from fragmentation of information and some inconsistencies in the usage of terms (see specific comments for examples). It might benefit from further streamlining.
This is a good point. We considered this in our revision by consistently using the term ‘reference value’ instead of ‘truth value’.

**Response to specific comments:**

*Line 11: Instrument orientation could refer to both north alignment or horizontal leveling.*

We added “with respect to the mean flow”.

*Line 13: The authors could consider to include the scan duration and time averaging aspect of the results into the abstract.*

We added the following sentence to the abstract: “Furthermore, we find that extending the averaging interval length of lidar measurements reduces the error.”

*Line 65: The abbreviation RMSD was only introduced in the abstract, but not the text.*

Corrected.

*Line 150 – 152: From the description I gather that the virtual LiDAR measures 5 seconds at the first position (alpha=0), then travels instantaneously to the second position (alpha=15) and measures the for the next 5 seconds. Wouldn't it be more realistic that the virtual LiDAR measures e.g. for 1 second at first position and then four seconds later the next measurement for 1 second is made at alpha=15? This way the travel time of scanner is accounted for, which results in less temporal averaging in reality.*

See answer to general comment no. 1.

*Line 168: The same for averaging effects from the pulse length.*

We fully agree that pulse-length averaging effects might be an additional source of error or uncertainty. However, it was not the aim of the present study to provide a full error analysis of Doppler-lidar wind measurement. Our focus was clearly on the errors introduced by different scan configurations, this was already explained in the text (now lines 182-184 in the revised manuscript). In fact, the pulse-averaging effect was investigated in a previous study by Gehrke (2019), and it was found to have a negligible effect when analyzing the errors of different scan regimes at a given height or averaged across the bulk of the atmospheric boundary layer.

*Line 178: The height interval is always from 50 m to the top of the atmospheric boundary layer. To avoid fragmentation of information, this could be stated directly here.*

Adapted.
Section 3 (and maybe throughout the manuscript in general): The reference values from the LES are referred to as “predicted values”, “truth value”, or “reference”. The output of the virtual lidar are referred to as “observed values”, “measured values”, “virtual measurements”, or “lidar values”. I believe that settling on one specific term for each would increase clarity of the manuscript.

See answer to general comment no. 1.

Line 182: The description of the processing here is different to the description in line 154. The previous page states that the data is first accumulated in 120 s steps and later averaged again to 10/30/60 minutes. Here the description states that the individual beams are accumulated and then directly averaged 10/30/60 minutes.

We realize that the description of the processing might be contradictory to earlier descriptions of the scan procedure. We therefore rephrased this passage as well as the part in line 154. Note that the averaging process is described in line 160 – 164 (now line 173 to 177 in the revised manuscript).

Line 186: Insert “the” before lidar.

We assume the reviewer was referring to line 187 “…measurement of the horizontal wind with lidar scan technique …”. We decided to leave this part unchanged, as we believe this is the grammatically correct form.

Line 190-192: Sometimes the wind profiles from a VAD scan are also interpreted to be representative for a spatial average across the scanning cone (in opposition to being representative for a column above the instrument). I believe this would be another aspect that could be investigated with the setup used here.

Yes, the reviewer made a very good point here. We did actually also investigate the spatial representativity of the measurements by comparing the measured profiles to the spatial average across the scanning cone, as well as to the average across the entire simulation domain. For this publication we decided to restrict ourselves to comparing the measurements to the vertical column for two reasons: Fist of all, as we stated in line 187 ff., one is usually interested in a wind profile above a certain location, for example when it comes to data assimilation into NWP models. Second, we feel that extending the discussion to additional reference wind profiles would overload this publication, which is already quite long.

Line 277: insert “the” before wind speed.

We left this in the original form since we feel that both versions are possible, and two lines later we use "wind direction" also without an article and the reviewer agreed on that.
Line 279: Did the authors use an arithmetic mean or an angular mean for the wind directions?

Here, the arithmetic mean was used.

Line 299-304: Was there a specific motivation why the north alignment of the Doppler lidar was investigated? To me this would fall into the same category as scanning the VAD in a counter-clockwise direction: I would expect some small random variations of the numbers, but no systematic difference. There is nothing wrong with this part, but it strikes me as an odd thing to investigate and was not well motivated in the introduction.

The motivation for including this investigation was that the orientation of the lidar could have an impact on the measurement result if quasi-stationary non-symmetric coherent structures are present in the boundary layer. We did not find this hypothesis confirmed, nevertheless, we believe it is a relevant result that should be included. To make our motivation for this more clear, we added a sentence in section 3.1.

Fig. 4 and Fig. 5: The panels for the 10-min average and 30-min average are unclear to me, because I only see one line for the reference / observation, but I would expect that there are 6 and 2 observations, respectively. Are those panel showing the mean RMSD?

In line 275 (in the revised version now line 289 ff) we note that all RMSD values are mean values over a 60-min measurement period. Thus, there is only one single value for the 60-min interval, while for the 30 (10) -min interval, the value represents an average over two (six) values. We added a note to the figure caption to make this more clear for the reader. For the profile figures, such as Fig. 5, for reasons of clarity, we display only one exemplary measurement period. We changed the text of the figure caption to “...profiles obtained over one exemplary 10-min period...”.

Line 331-332: To me it seems that identical values exist for top left, top center, and bottom left and remarkably close values for top right?

Corrected.

Line 498: Some of the limitations of this study like not accounting for range gate effects were already brought up in the methods section. Maybe the same could be done with the effect of surface heterogeneity on the error to put the readers mind at ease that it has not been forgotten.

We added a sentence to section 4, mentioning that all simulations are carried out over flat and homogeneous surface (line 209 f. in the revised manuscript).