General comments from the author on the revision:

We would like to thank the reviewers for their thoughtful and thorough responses and comments which have helped to strengthen the manuscript. In the process of revising and in light of some of the reviewers' points, we have made some additional edits to clarify and clean up the manuscript.

- In responding to the review comments, the organization and ordering of the material emerged as an issue. We have sought to address this by moving the analytic treatment of the error distributions (which appeared along with the results in Section 3 describing the underlying mechanisms) to a new subsection in Section 2 (Data and Methods). The restructuring required some large changes in the manuscript to accommodate the new ordering, combining repeated or dispersed statements into the new sections, and to help and with the flow. New or modified text has been added, but we have tried to avoid changing or introducing new technical material in the rearrangement.
    - Fig. 1-3 are the same
    - Fig. 4 moved (previously was Fig. 12)
    - Fig. 5 is a new diagram by reviewer request
    - Fig. 6 moved (previously was Fig. 4)
    - Fig. 7 is "new", summary of u and v moments from appendix by request
    - Fig. 8 combines the error decomposition from previous Fig. 9, 10, 11
    - Fig. 9 is new, compares RWF acting on mean profile to bias in full model (to address reviewer questions)
    - Fig. 10 is moved (previously was Fig. 13)
    - Fig. 11 is moved (previously was Fig. 5)
    - Fig. 12 is moved (previously was Fig. 14)
    - Fig. 13 is moved (previously was Fig. 15)
    - Fig. 14 is moved (previously was Fig. 8)
    - Fig. 15 replaces what was Fig. 16, highlights trends in time-averaging demonstrates error decay rates instead of showing reduction just in a 10-minute average (thank you for reviewer feedback)
    - Fig. 16 is new, part of the added analysis of comparison to cup measurements
    - Fig. 17 is new, also part of the added analysis
    - Fig. 18 combines plots of the vertical velocity reconstruction error (prev Fig. 6,17)
    - Fig. 19 was previously Fig. 19
    - (previous Fig. 7, showing the KDEs of a single time-series measurement to illustrate the selective sampling from the ensemble distribution was removed as a lower-priority figure for space and flow reasons)
- A community comment suggested additional analysis in comparing vector- and scalar-averaged lidar and point measurements. The bulk of the additions regarding this analysis have been made to the time-averaging theory (Section 2.3.5) and time-averaging results (Section 3.3, Fig. 16 and 17) to provide the background and present the results.
- In the derived wind speed error form, we have replaced our first-order expansion with an expansion using second order Taylor expansions in u and v to put it in the same form as mathematical analogs in existing literature (e.g. Courtney 2014 and Rosenbusch 2021). The change of form makes the correspondence with previous work clearer, especially when summarizing and discussing the comparison of the lidar and cup scalar averages of the wind speed. The analysis and conclusions from the original form hold equally for the alternative expansion.
- It came to our attention while clarifying some of the terminology based on reviewer comments that the mathematical notation could also be adjusted to support important distinctions. We have adjusted the notation to try to make each of the mathematical objects clear and sought to make the notational forms consistent.
    - We have changed the radial wind measured by the lidar at the beams from u_E to $v_{r,E}$ to make the relationship to v_r used for the radial velocity in the RWF model clearer and to avoid any confusion with the velocity perturbations at the beams (was u_E')
    - We have made a notational distinction between the perturbations from the volume-averaged wind (over the lidar scan volume) at each of the beam locations ($u_E'' = u(east\ beam) - U$) (now using u_E'' instead of u_E') turbulent fluctuations from a Reynolds average ($u' = u(x,t) - \overline{u}$). This supports our discussion of the ways the behavior can differ from Reynolds turbulent fluctuations when using a local volume-average, the distinction in the filtering of the turbulent velocity variances, and in the added material on the comparison of scalar wind speed averages.

- The decomposition of the vertical velocity error (w_err) was plotted against the wind-direction-corrected full model rather than the V1, equally weighted model. This has been corrected and changes the solid, ‘full model’ error lines in the panels for the vertical velocity in the decomposition in Fig. 8. (was Fig. 9,10,11)
- Correction of small typos, edits to improve clarity and flow, and bringing the text into compliance with journal standards (use of Fig., Eq., and hyphenation in particular)
- Figures have been regenerated as pdfs where possible. Panel labeling, legends, and use of mathematical terms in axis labels have been adjusted in some figures for consistency across the manuscript.

Response to Referee #1

Comment on amt-2022-73

General comments:

This is an interesting study that tackles a fundamental problem for the Doppler wind lidar observations which is the better understanding of the measurements' error. The authors used large-eddy simulations for different atmospheric stability conditions and analyzed which mechanisms drive the errors under each condition. This study can be very useful for researchers studying turbulence as observed by a Doppler wind lidar. It could help them understand possible error biases or the cause of large errors. It also shows the importance of multiple scans so that the observations can be averaged over time to decrease the error. I believe this study is worth publishing. Nevertheless, the authors should make several changes in the manuscript in order to improve it's structure. In the current state, it is difficult to read through. Introduction, methodology, results and conclusions are mixed, the authors should set apart these chapters to improve the flow of the manuscript. In the current version, parts of the introduction are presented in the methodology, whereas parts of the methodology are introduced in the results and some results are presented but discussed in a later section. In the conclusions, a summary of the work with the main points of the results should be presented, however the authors make a comparison of their results with previous studies and furthermore they introduce new references. Regarding the abstract, the results should be supported by numbers so it will be evident that these are the results of the current study. Some terms should be explained better throughout the study, so it is clear to the reader what is the magnitude of terms such as the "strong winds" or the "large structures". The figures and their captions need corrections as well. All the important information should be included in the plot with the description written in the caption, rather than important information, such as the colour of the lines, only explained in the caption. The language in the manuscript is fluent. Please find my comments below.

We would like to thank the reviewer for their time reviewing the manuscript and thoughtful comments and appreciate the acknowledgement of the applicability of our findings for researchers working with Doppler wind lidar. The insights into how the material can be better presented were particularly useful.

Scientific comments:

Page 2 Line 27-29: The authors mention that lidar data offer an indirect representation of the flow field. Although it is true that the wind lidar observations most likely need extra steps to extract the wind components compared to sonic anemometers, it is possible to directly observe the wind components if the beams are aligned with the wind direction e.g. horizontal beams with no elevation angle alongside the wind direction or vertical beams for the vertical wind.

Measuring the along-beam component is more direct than the reconstruction required by a profiling lidar, but we would still view it as less direct than e.g. a 'point' measurement made by sonic anemometers. As modeled by the RWF the probe length of the lidar sampling process is 10s of meters long vs decimeters and can interact with shear in the flow.

P2 L45-47: Similarly to my previous comment. In the phrase "Questions about measurements of vertical velocities"; do the authors mean vertical velocities variances? Because vertical velocity can be directly measured by the Doppler wind lidar and in fact even the vertical velocity variance is the "easiest" turbulence parameter that can be estimated using wind lidar observations, see Bonin et al. 2016: "Improvement of vertical velocity statistics measured by a Doppler lidar through comparison with sonic anemometer observations".

"Questions about measurements of vertical velocities and the ability of wind profiling lidar to measure turbulence remain areas of active research (Sathe et al., 2011; Sathe and Mann, 2012; Sathe et al., 2015; Newman et al., 2016)."

We referred to the way the volume-averaging can complicate even the 'direct' measurement of the vertical velocity and the variance, which is addressed in the Bonin et al. 2016 paper. We have amended the statement to reflect the particular challenges in the turbulence measurements and included the paper as an additional reference.

(LL49-50) "Questions about the ability of wind profiling lidar to measure turbulence remain an active area of research (Sathe et al., 2011; Sathe and Mann, 2012; Sathe et al., 2015; Newman et al., 2016; Bonin et al., 2016)."

P2-3 L59-62: The sentence "Compared to field studies of instrument accuracy, studies with virtual instruments in LES ...." should be supported by some relevant references for such studies.

We have amended this sentence to acknowledge studies that have leveraged the unique opportunities presented by a virtual instrument.

(LL62-66) "Compared to field studies of instrument accuracy, studies with virtual instruments in LES have unencumbered access to full knowledge of the flow field. This knowledge enables control over the case parameters (terrain, forcing, boundaries), and so users can "deploy" instruments in ways that may not be physically or financially possible in reality (e.g. re-sampling the same flow field or testing many locations in a domain) (Muschinski et al.,1999; Stawiarski et al., 2013; Wainwright et al., 2014; Gasch et al., 2019).

P3 L67: The PALM model was developed by Raasch and Schroter, 2001: "A large-eddy simulation model performing on massively parallel computers". The reference should be added in this sentence. Added here, thanks!

P4 L78-79: The Skamarock, 2008: "A description of the Advanced Research WRF version" reference should be included in this sentence. Added.

P4 L82: The Chow et al., 2005 is not relevant here as they do not mention the WRF model in their study.

We have corrected this citation to Kirkil (2012).

P10-11 L248-270: The authors state that they selected cases with "strong" and "weak" convective boundary layers. In line 262, these cases are characterized well-mixed layers. However, in Figure Table 1 we can see that the for the weak cbl the abl height is 525 m. This value seems to correspond to a developing boundary layer and not a well mixed. The authors should comment on that and whether these values can occur only in an ideal case of the simulation with a flat, homogeneous terrain etc. Moreover, the vertical range of the Windcube is portrayed in Figure 3 but it is not mentioned in the text and even in Table 2 it is not directly shown. This information should be included along with an explanation of this selection. The value is higher than the ABL height under stable conditions (170 m). Wouldn't this affect the comparison? See also my comment regarding Figure 4.

Regarding the "weak" convective boundary layer: The convective cases drawn from Rybchuk (2021) are emulate conditions of the Project Prairie Grass campaign (Barad, 1958) which looked at diffusion of tracer gasses. The ABL height in the weak CBL case can and did occur in real conditions. The depth is a function not just of the surface forcing but also of the strength of the inversion aloft which allows for the 525m ABL to occur in a well-developed boundary layer. We have noted in the text that these cases reflect real, observed conditions and added a reference to the campaign along with the existing reference to Rybchuk (2021).

Regarding the vertical range of the Windcube: We have added a line to Table 2 (now Table 1) to make the range explicit (40-240m AGL). The range arises from values fixed by the instrument and typical operating parameters, i.e. the number and vertical spacing of the range gates. We have added this explanation to the text (LL201-203). The fact that the range exceeds the ABL height under stable conditions does affect the comparison. We refer to the boundary layer structure as part of the discussion of the error behavior (e.g. the peaking in the perturbation variances mid-boundary layer and tapering aloft in the stable BL while the growing variances in the convective BLs only reach mid-boundary layer at the top of the range) but did not see fit to normalize by the boundary layer height since the geometry of the instrument is fixed.

P11 L266-267: The authors mention the limitation of the lidar range to include the entrainment zone. For the instrument, it is practically difficult to measure at this height due to the scarcity of aerosol in the zone. Do the authors refer to the instrument or the simulations? What are the limitations for a simulation?

Practical considerations like sufficient aerosols were outside the scope of our analysis, though important to actual operations. There is no reason the simulation / virtual lidar model couldn't be extended to higher ranges and the range limitation is introduced based on general Windcube instrument parameters (which do depend on considerations of aerosols and obtaining a usable signal) of the number and spacing of the range gates. Since the instrument geometry is more or less fixed for deployments across conditions, we have mimicked that geometry and range throughout the test cases (while ignoring, e.g. the possibility that a real instrument would have insufficient aerosol scattering to return a usable signal within that range).

P14: In Figure 4 panel (c), for the lower altitudes the median of the curve is not at zero as it also stated by the authors in P15 L356. The authors should explain at this point of the manuscript why this occur only at these particular levels of the SBL. It is also apparent from panel (c) that for the levels above the ABL height (170 m) the distribution becomes similar to the one near the surface. Any comments regarding this? Was this something the authors expected?

P15 L355: What could be the cause for overestimation during convective conditions?

For both these points, using the suggested re-ordering of the material to introduce the analytic error model in Section 2 allows the mechanisms behind these behaviors (RWF acting on the shear / vertical wind profile and the systematic positive bias term in the wind speed error) to be addressed as soon as they are shown.

The (moved-up) diagnosis of the causes of bias and variance in the error explains why the error above the stable ABL resembles that at the surface. The filtered turbulence under the scan-volume scale is similarly smaller at the surface and above the ABL (particularly the vertical velocity variances) which reduces the width of the distribution. Similar curvature in the background u velocity profile induces the negative biases in the shear layer in the surface but also aloft through the RWF.

A positive bias term (proportional the u,v variances / inversely proportional to the mean wind speed) gives rise to the overestimation in convective conditions when vertical velocity variances drive larger errors.

In the text, the discussion of these points and the connection of the error behavior to the vertical structure of the boundary layer turbulence (which were split between Sections 2 and 3 in the original manuscript) have been merged into a single, streamlined results section (Section 3).

P24 L506-507: The terms strong winds and strong shear are vague. Can you quantify these parameters? Is the underestimation expected above a specific threshold? Additionally, the argument that surface shear is one of the cause of the underestimation should be supported by some results in the form of Figures. Maybe the authors could add a secondary y axis with the wind speed and wind shear values at the different given heights in Figures 9, 10, 11. The estimation of the wind shear can be tricky but this claim should be supported by results.

"The most prominent influence of the RWF is near the surface layer (due to strong shear) and manifests as a negative shift in the mean error in the horizontal velocities. The weak CBL and stable BL, with strong winds and surface shear, exhibit the effect most strongly. "

(This material has been moved to Section 3.1 discussing the source of biases in the results. LL597-605)

The connection to the shear is mechanistically supported by our derivation of the bound (Eq. 16 in the edited manuscript) as well as previous studies (Clive, 2008; Courtney et al., 2014) that have derived the action of an RWF on idealized shear profiles. (These have been included in the discussion.)

The underestimation will occur for any concave function (like the semi-empirical log shear profile for U(z)); there is no threshold except what degree of bias is considered tolerable. The strong CBL, for example, does still have some shear near the surface but the resulting negative bias is <0.05 m/s, so would likely be considered negligible.

We have added Fig. 9 to help support the conclusion in the case-study results. Fig. 9 demonstrates that the bias due to the RWF in the full model and the bias resulting from applying the RWF to just the mean LES profile and shows the largest magnitudes in the bias in u (b) and v (d) occur in regions with the largest (nonlinear) shear.

P26 L540 & 548-550: Similarly to my previous comment, the term "large coherent structures" is vague. It is evident from Figure 12a that there are structures, upward motions followed by downward motions, of approximately 1 km size. On the contrary, for the stable boundary layer (Figure 12b) the size of the structures seems to be equal to few hundred meters. The lifecycle of such structures should be different. Do the authors categorize both structures' sizes as large? The authors also mention larger scale structures above the boundary layer in the SBL. Do they mean from 170 m up to 350 m that is shown in Figure 12b or in higher altitudes? Either way it should be clear to the reader what are the sizes of the structures and these should correspond to the figures presented in the manuscript.

"Without dissecting the mechanics more closely, we simply note that large coherent structure like turbulent plumes characteristic of the convective boundary layer can induce repeated, non-symmetric patterns in the relative perturbations of the beams, leading to non-zero means."

(This material has been moved to Section 3.1, LL590-596)

This is an important point about the size of the structures, which Fig 12 (now Fig. 4) was meant to help convey. The critical scale here is the size of the scan volume (which for the WindcubeV2 is ~42m across at the bottom range gate and ~255m across at the top, shown by the black lines in Fig. 4). "Large" coherent structures are meant to refer to features on the order of the scan volume or bigger so that their internal structure influences the sampling by opposing beams, which would include the 1km structures but potentially exclude some of the features in the lower stable BL.

We have explicitly called out "coherent structures large enough to span the scan volume" and denoted the scan radius in the text to help clarify this point.

Technical comments:

P1 L1: Lidars instead of lidar. Fixed.

P3: The full name for the abbreviation DBS should not be included in the caption of Figure 1, but rather in P4 L76.

The definition of the full name for the DBS abbreviation has been moved, along with checking the order and use of other abbreviations in the text.

In the first three instances Figure 1 is written with a capital F in the text, whereas in all the other instances figures in the text are written in parenthesis with a minor f. The citation of the figures should be consistent throughout the text.

Thanks for pointing this out. These instances along with all further references to figures and equations were changed to "Fig.", "Figures", and "Eq." to bring the text in line with journal style guidelines.

P3 L72: The full name for the abbreviation SOWFA should be given here. Fixed.

P3 L73: The full name for the abbreviation WRF-LES should be given here. Fixed.

P7 L156: The parameter "c" is defined after the equation 7, although it is also part of the equations 5 and 6. In order to avoid confusion, I suggest to move the definition before or after the equation 5.

Good catch. We have moved the definition in the text up to the first reference to the use of the speed of light before Eq. 4

P7 L161: The sentence "Using parameters .... can be made concrete" needs rephrasing. It is confusing in its' current state.

"Using parameters from the WindcubeV2 (Table 2), the shape and extent of the RWF (Eq. 6) can be made concrete (Fig. 2a)."

reworded to

(L170) "The RWF for the modeled WindcubeV2 (Fig. 2(a)) results from substituting the range gate and pulse parameters (Table 1) into the general, pulsed lidar equation (Eq. 6). The shape of the RWF peaks at the center and tapers symmetrically toward zero to either side."

P7 L167-170: The paragraph "The form of the pulsed lidar .... further distances by a pulsed lidar" is more suitable for the introductory section. The comparison of the RWF between pulsed and continuous lidar seems out of place in the methodology as only the pulsed lidar was used for this study.

We agree that the discussion here became more general than needed to serve the primary focus of the study. We've dropped the paragraph from the text.

P7 L172: Remove the word "found".

We have reworded the end of this sentence from

"…and the projection onto the beam direction found"

to

(L177) "…and then projected onto the beam direction".

Hopefully this makes it scan more cleanly for the reader.

P7 L173: The "Spe" in the parenthesis is a missing reference?

Yes; a reference used to split the scipy function (into triangulation and application of the interpolation weights to optimize repeated calls on the same grid) does not have a complete reference and therefore was mangled. We have moved that reference to the code and retained the complete technical scipy citation here.

P8 L185-188: The paragraph "Interpolation dominates ... subsequent developments" comments on the interpolation method and possible improvements in the data and methodology section. This paragraph is more suitable for a section like Conclusions/Discussion.

We have removed this paragraph and made a brief reference to the optimization of the interpolation as a target for future improvements in the conclusion (L948)

P8 L190-199: In the first paragraph of Section 2.1.2 the authors provide some general information regarding the lidar and the different scanning methods such as RHI and PPI. As these methods are not used in the particular study, they should not be mentioned in the data and method section. In my opinion, this paragraph should be removed entirely from the manuscript as it does not provide any valuable information for the study.

We had included this in a general approach to the virtual lidar model, but agree that the additional information in this paragraph is a distraction that doesn't serve the current study. We've cut out this portion of the paragraph and condensed it to simply refer the reader to Clifton (2015).

P10: In Figure 3 the explanation for the different lines (solid, dashed etc) representing the parameters is only included in the captions and not in the figures, hence it is not practical for the reader to study these figures, similarly for Figures 5, 6 etc.

We have reviewed each of the figures in the manuscript with this in mind and have included full legends reflecting each of the parameters used.

P13 L303: The claim that the components of equation 10 are commonly used should be supported by some examples-references.

We have included a reference to Cariou and Boquet (2010); the instrument itself returns the horizontal wind speed and direction and we believe they are fairly ubiquitous. (L414)

P13 L305-308: A figure showcasing the sign convention could be useful for the reader.

We've created a diagram (Fig. 5) to show the conventions used for the wind direction and sign of the direction error.

P14: The authors have introduced several parameters for the wind speed such as ulidar or horizontal wind speed, therefore it should be clear in Figure 4 which one is shown including its' name as defined by the authors.

We have re-labeled the figure axes with both the name and symbolic reference (|U_h|_err) which refers to the error in the magnitude of the horizontal winds (i.e. 'wind speed error'). Hopefully this addresses any possible confusion about the plotted quantity.

P14 L324: Remove "the". Fixed.

P15-P16: The caption of the Figures 5 and 6 mention dashed and dotted lines but only dashed and solid lines are depicted. The figures should be corrected and additionally there should be a legend in the figure with this information.

Thanks for catching this and we apologize for any confusion it may have caused in reviewing the article. The caption has been amended and the figure revised to include the line style distinction in the legend as well.

P15: Figures should be easily readable even when separated from the rest of the text. In Figure 8, the caption is linked to the parameters of Figure 5 which should be corrected. The parameters should be included in the caption and the legend of Figure 8 independently from Figure 5.

We have regenerated the figures to include comprehensive and consistent legends across these figures.

P17 L374: Converge instead of converges. Fixed.

P18 L401: The phrase "their respective height trends are similar to the previous section" should be accompanied by the respective values as a reminder for the reader. Not in the rearranged manuscript. We have tried to make sure to call out values and reiterate the recurring trends when we refer to them in the updated text.

P19-P33: In Chapter 4 the authors introduce several new equations. In my opinion, a manuscript flows better when all the equations and tools are presented in the data and methods chapter and subsequently the results are presented and discussed. So instead of presenting the figures in Chapter 3 and then using the equations to describe the results in Chapter 4, I believe it would be better if all the equations are already presented in Chapter 2 and the discussion of the results is moved to the corresponding figures. For example the explanation of the underestimation of the wind error although mentioned in Chapter 3 is explained much later in Chapter 4. By moving the equations to Chapter 2, the authors will also avoid repeating themselves.

Thank you for your thoughts on how best to organize the material. We had developed the analytic model for the error as a tool for understanding the output of the virtual tool and had thought it could be treated concurrently with the results in the paper. Seeing the degree to which the analytic development grew and how much we rely on it in explaining the results, we agree that the flow of the article is improved by re-ordering the presentation.

We have moved the development of the analytic model for the lidar error into a new subsection of the Data and Methods (Section 2). The results using the virtual instrument have been condensed into a single section (Section 3) and streamlined, referencing back to the analytic model in Section 2 when needed for the discussion. We hope this removes ambiguity and repetition and makes the paper stronger and more clear to the reader.

P20: The caption of Figure 9 seems more like a part of the manuscript than a caption. It should be rephrased in a way to resemble a caption.

P21-22: It would be easier to interpret the results from Figures 9, 10, 11 if these figure were merged in multiple panels and thus it would be possible to use the same caption for all instead of linking to the caption of Figure 9 which makes the Figures unreadable independently.

To both these points:

We have revised this set of figures into a single, combined decomposition figure (Fig. 8) which uses the same caption. The revised figure shows the covariances and obviates the need for some of the explanatory text in the caption. The specification that only the non-offset lidar are used was moved to the main text.

P29 L616: The word "term" is used two times. Fixed.

P30: The height of the wind speed is not mentioned in Figure 14. It should be included as part of the xaxis title. Figure 14 (now Fig. 12) aggregates data from all heights. Since the original manuscript left room for uncertainty, we have added an explicit statement to the caption to clarify the point.

P32: The legend showcasing the parameters that correspond to the colour lines is missing in Figure 16.

We have replaced this figure to better illustrate the trends over the time-averaging window. The new version (Fig. 15) does include the color-coded legend.

P32: The panels in Figure 16 are not numbered with (a), (b), (c) etc and the authors refer to the different plots as left and right panels. I believe it will be easier to use the numbering. Similarly the Figures 5, 6, 8, 9, 10, 11, 13, 17, 18 and the ones presented in Appendices A, D, E also miss the numbering.

Thank you for pointing out the inconsistency and confusion in (not) labeling the panels in these figures. We've worked through each of the figures to clean them up and add labels in a consistent style where they were missing and have updated the captions to reflect the changes.

P36-39: The comparison of the authors' results with previous studies should be included in the sections of the results and not in the conclusions. The authors should summarize the key points of their study and not include new information in this chapter. Although it is possible to include some previously mentioned references in the conclusions, it is not recommended to introduce new references such as Klaas and Emeis, 2021 and Teschke and Lehmann 2017. These references should be introduced earlier in the manuscript.

The comparisons of how our findings relate to previous studies have been mentioned as part of the results where appropriate (now Section 3) and moved into a discussion section (Section 4). We have correspondingly shortened the conclusion to summarize our key points.

P38 L840: Add the word of - "The form of our error...." Fixed.

P52 L1117: Remove typo " ". Fixed.

P52 L1119: Remove typo "&ndash". Fixed.