Dear Referee #2,

We thank you for the review of our manuscript and your detailed remarks. We have revised the manuscript, considering your remarks and those from the other referees. The revision led to a substantially changed and shortened version of the manuscript.

In order to allow a focus on the most important results we now reduced the content to the introduction of the LES-based airborne Doppler lidar simulator. Following this, illustrating the benefits of this new tool, the error in wind profiling due to violation of AVAD retrieval assumptions by boundary layer turbulence is investigated for a standard system setup and retrieval strategy. The result sections on system setup and retrieval strategy characteristics are not part of the manuscript anymore due to their lengthy nature as well as comments by other referees.

Please find our answers to your major and specific comments below.

Many thanks for your work so far and kind regards,

Philipp Gasch and Co-authors

Initial, short replies are marked in this color

The final response addition is marked in this color

Major comments:

> (1) The writing style, English and grammar needs work and many comments are listed below. I stopped after several pages because it was taking too much time. Significant re-writing and organizational changes are needed in the manuscript. I also think the paper is much too long. It appears that this journal does not have a page limit, but it would help readers to shorten the discussion in several places and remove sections that are not needed (some examples given below). The paper reads a bit like a dissertation with too much background and drawn out detail. A published paper should be more concise without sacrificing understanding of the problem. Please shorten the results section, it looks like too much information is presented and it might not be worthy to publish all of it.
>
> We are sorry to hear that the style, English and grammar did not suit your tastes. We are surprised about the comments about the English, as one of our co-authors is a native speaker and carefully proofread the manuscript. We are therefore very sorry for the inconsistencies that have slipped our attention.
>
> Regarding the shortening: The manuscript was designed for publication and not as a dissertation. Due to the many aspects considered we felt that some background information might be helpful for some readers.

We have considerably shortened the manuscript now in order to be more concise. Our native speaker co-author has carefully proofread the revised manuscript version again.

(2) Throughout the manuscript the word "homogeneous" and "inhomogeneities" are used and this represents a critical aspect of the study and results. For example, "...mismatch between assumed homogeneous wind field models and the wind field inhomogeneities during the measurement process". These are ambiguous terms and I don't understand how they are being used in this context. Since they represent critical points of the paper, it is hard for me to assess the method and results. The authors need to lay out in detail what they are referring to here and clarify this throughout the manuscript. Are you talking about wind variability below the scale of the instrument footprint, grid spacing of the wind retrievals, something else?

In this study we are concerned about errors in wind profiling that arise from the fact that the analysis method assumes that the flow is homogeneous (i.e. constant mean wind profiles), whereas an actual boundary layer experiences flow inhomogeneities driven by turbulence, such as discussed in Shapiro and Fedorovich (2007), Kiemle et al. (2011) and Lundquist et al. (2015).

By the term "wind flow inhomogeneities" we refer to deviations of the wind speed from the mean state due to boundary layer turbulence. In order to consider only turbulent conditions, we limit the extent of our analysis domain to 0-800 m vertically (inside the boundary layer) in Sec. 4. The structure of the turbulence present in the LES is further detailed in our answer to remark (3). The range of the wind variability is presented by the spectra of the turbulence (fig. 5.8 in Stawiarski, 2014). The dominant range of the turbulence spectra is larger than the instrument footprint (10 cm beam diameter, 300 ns Gaussian pulse width, 72 m range gate length), enabling an accurate representation of the turbulence in the simulated measurements. The dominant range of the turbulence spectra is smaller than the horizontal grid spacing of the retrieval (1.3x1.9 km) but larger than the vertical spacing (60 m). This spatial mismatch clearly influences our results in Sec. 4 and is discussed there.

The term homogeneous flow (a synonym is homogeneous wind field conditions) means a uniform flow and the absence of any deviations from the mean flow state, both spatially and temporally (Stull, 2000). Homogeneous flow is assumed in the most simple form of the Velocity Azimuth Display (VAD) and Volume Velocity Processing (VVP) retrievals (Koscielny et al. 1984, Boccippio, 1994, Banakh et. al, 1995, Leon and Vali, 1996). Homogeneous flow is rarely present in the real atmosphere and certainly not inside the boundary layer. The violation of the homogeneity assumption in the VAD/VVP retrieval due to boundary layer turbulence causes an error in the retrieved wind profile. This error is the focus of our study.

We have clarified the meaning of the term 'inhomogeneity/inhomogeneous flow' as well as 'homogeneous flow' in the revised version of the manuscript. This includes rewording the manuscript in many places in order to clarify and stress that we are concerned about wind profiling errors due to boundary layer turbulence violating the AVAD assumptions. We have included an illustrative

figure, in which the presence of flow inhomogeneities is clear (fig. 4), and another one, in which the resulting error is visible (fig. 5)

(3) The LES domain size of 5 km X 5 km X 1.8 km is extremely small and I have doubts that this domain will represent a realistic environment to test the lidar wind sampling. The authors state that a single flight through the LES does not yield sufficient statistics. However, making 25 different aircraft trajectories through a very small box, probably does not generate any real independent statistics since the retrieved winds are sampling almost the same flow (the decorrelation spatial scale is probably larger than the box itself). It appears the grid spacing of the wind retrievals might be 1.3 – 1.9 km for along/across track. Given this spacing, I don't think the authors can generate independent flight tracks and statistics through a 5 km X 5 km domain. The authors should try their simulations with a flow in a larger domain (with coarser resolution) in order to study a more realistic environment and allow for independent statistics.

While we appreciate the reviewer's request for a larger domain, applying coarser resolution would actually undermine the attempt to represent fully developed turbulence. We are confident of the decorrelation spatial scale, as the integral length scale (Tennekes and Lumley, 1972, Lenschow and Stankov, 1986) of the turbulence present in the LES is $L_i < 500$ m, as stated in section 2.1. Further documentation of the decorrelation scale of the turbulence is provided in the attached fig. 6.3 and fig. 6.6 from Stawiarski (2014) below.

As we constrain the analysis to heights < 800 m (above which gravity waves are present, increasing the integral length scale), the assumption of statistical independence is valid for our analysis volumes of 1.3x1.9 km, which are much larger than the integral length scale (keeping in mind that turbulence is the driver of profiling error which we are interested in).

At the given spatial and temporal integral length scales, we argue that the 16 flight directions, repeated at 1-minute spacing and with a random initial profiling offset, do sample independent air masses. The spatial distance of the wind profiles is larger than the integral length scale. The temporal spacing is on the order of the integral time scale, but aided by decorrelation through advection between subsequent wind profiles.

Unfortunately, we think a coarser model would be problematic to investigate the errors due to turbulent structures in the boundary layer. In our study the grid resolution is 10 m (corresponding to a resolution > 50 m). Thereby, Doppler lidar measurements at a measurement frequency of 1 Hz, with a range gate length of 72 m and a flight speed of 65 m s^{-1} can be realistically represented, as the resolution corresponds to the sampling rate and spatial extent of the range gates. In a coarser model, the high sampling rate would not produce independent measurements. Further, the structure of the turbulence, causing the wind profiling error we are interested in, would be less accurately represented.

Please note that our approach follows established methods. These LES wind fields have been used for Doppler lidar studies in a similar way before

(Stawiarski, 2014, Stawiarski et al., 2015) and that other authors choose similar approaches for ground-based studies (Scipion et al., 2009, Scipion, 2011).

We will make sure to discuss the independence of the acquired data in a more clear way in the revised version of the manuscript.

Sufficient independence of the acquired data is an important point. Therefore, we are addressing the concerns about independence of the sample data in-depth in the revised manuscript version in-depth (Sec. 3.2). To this end, we have changed the sampling strategy to only include profiles which are sampled with spatial or temporal independence. This is done by using a checkerboard approach, which is discussed and illustrated in the manuscript now. Further, the independence of the sampled wind profiles is now checked using statistical methods including autocorrelation.

Specific comments (some major, some minor):

Page 1

Line 2: should say "...additional insights **relative** to ground-based systems..." Included.

Line 2: what does "spatially resolved" mean here? This term is too ambiguous. Changed to spatially distributed.

Line 3: "...prepares the ground...", is poor English and needs a change. Deleted.

Line 4: spell out LES for the first time used; what is meant by "first"? Note that other studies have used large eddy simulations to study remote sensing instruments. LES is now spelled out. We mean the first LES-based airborne Doppler lidar simulator, which to our knowledge doesn't exist so far.

Line 6: I believe it should be "...wind profiles **in** inhomogeneous flow...". Changed.

Line 7: Need to clarify with numbers what is meant by "acceptable error margins". Acceptable is ambiguous and could mean very different things to different people. Changed.

Line 7/8: sentence that starts with "Results allow for determination..." should be removed. This is an obvious outcome of the simulations. Changed.

Line 8: What is meant by "flow inhomogeneities"? Seems like this is key since much of the manuscript mentions this, but again, this term seems ambiguous to me. As stated above, we tried to clarify what we mean by flow inhomogeneities.

Line 16: What is meant by "short horizontal averaging distances"? Changed.

Page 2

Line 3: Need a comma after “benefits”. Included.

Line 11/12: “considering both wind profiling and nadir measurements of the wind field”; I don’t understand the difference between these two things as it is written. Changed.

Line 13: What is meant by “mean” horizontal flow? Average over space/time and what scales? Changed.

Line 18: I don’t understand how winds are retrieved through “inversion of the beam matrix”, this sounds like an incorrect statement or writing error. Winds are retrieved by inverting the least squares fit between the model and observations. It is crucial for reader to understands how the inversion process works as this is the basis of the AVAD retrieval. Therefore, this statement is now moved to the section explaining the retrieval process in order to provide more context. We follow the notation of Leon and Vali (1998). What is inverted (in our case using a SVD decomposition) is the beam matrix G, yielding a least-square solution to the problem. The LSQ-fit is the result of the inversion process (to our knowledge, the LSQ-fit cannot be inverted). The LSQ-fit is obtained by multiplying the beam matrix with the estimated wind vector, obtained by the inversion. We have included fig. 4, illustrating the input truth wind field as well as the retrieval procedure and LSQ-fit. We quote Leon and Vali (1998, p. 865):

“The matrix of beamvectors (for the selected form of the velocity field) is then inverted using a singular value decomposition (Bevington 1969; Menke 1989), such that

$$\mathbf{V}_{est} = \mathbf{G}^{-g}\,\mathbf{V}_{Dopp}\,. \qquad (19)$$

The estimated velocity field parameters can then be used, together with the matrix of beamvectors, to produce an array of predicted Doppler velocities:

$$\mathbf{V}_{Dopppre} = \mathbf{G}\,\mathbf{V}_{est}. \qquad (20)$$

Comparison between the predicted and actual data values can reveal how well the wind fields have been fit and the appropriateness of the velocity field form.“

Line 25: “...assume homogeneous conditions throughout the sample volume.”, what is meant by this statement? One can’t measure things that are sub-grid-scale, but I don’t understand what this is referring to. Please see our answer to your major comment 2 for an in-depth answer on this. We assume that you are referring to the retrieval volume when using the term sub-grid-scale. Sub-retrieval volume flow inhomogeneities can introduce error into the retrieved wind vector when using AVAD (as AVAD assumes homogeneous flow throughout the retrieval volume). As you state, the inhomogeneities cannot be retrieved by the mean wind vector obtained as a solution. Therefore, we are not trying to measure sub-retrieval volume processes. However, deviations from the assumed mean flow state on scales smaller than the retrieval volume will introduce error into the AVAD retrieval. This is the error which we are concerned about. The advantage of the ADLS is that the sub-retrieval volume inhomogeneity is exactly known, because the LES is used as a known input wind field, which is impossible in reality. Therefore, we can quantify the error due to sub-retrieval volume flow inhomogeneities.

We have tried to give more context by referring to boundary layer turbulence directly, as well as rewording.

Line 28: Statement about how high elevation angles are used to constrain the footprint. Tilt angles closer to nadir will provide a shorter slant path and thus smaller footprint, but there are other reasons for choosing this steep tilt. Some things could be hardware limitations, range limitations and attenuation. We agree and have included your suggestion.

Line 29: I think it should be "measured radial velocity" and not "retrieved radial velocity". The process of obtaining a radial velocity also involves a process termed retrieval by some researches (e.g. spectral peak estimation), which is why we chose this wording. We now reworded according to your suggestion in order to avoid confusion.

Line 34: I am still confused on what is meant by "homogeneity assumption". As stated above we have tried to clarify and explain.

Page 3

Lines 15 – 17: Sentence is too long, need to break up for clarity. Done.

Line 24: Need a comma after "capabilities". Done.

Line 25: Need a comma after "systems". Done.

Line 29: "challenged" should be "challenges". Done.

Page 4

First paragraph: I don't understand what "assume homogeneous wind field and inhomogeneities during the measurement process" means in this context. Need to make significant changes to clarify this and possibly make a diagram to illustrate what this is referring to. We reworded in order to clarify in combination with the above changes. In addition, as suggested by you, we have included fig. 4. Fig. 4 illustrates the input truth wind field as well as the retrieval procedure. It also includes an illustration of the flow inhomogeneities due to turbulence, which we are concerned about.

Line 16: Need a period after "follows" instead of a colon. Done.

Page 5

General question: what is this a large eddy simulation of, homogeneous, isotropic turbulence? The domain size of 5 km X 5 km X 1.8 km is extremely small and I have doubts that this domain will represent a realistic environment to test the lidar wind sampling. The authors state that a single flight through the LES does not yield sufficient statistics. However, making 25 different aircraft trajectories through a very small box, probably does not generate any real independent statistics since the retrieved winds are sampling almost the same flow (the decorrelation spatial scale is probably larger than the box itself). The authors should try to find a simulation with a larger domain (with coarser resolution) to test the sampling and/or mention that the results of this study are limited to very idealized flow conditions. Please see our answer to your major comment number 3.

Page 6

Line 19: 65 m/s seems like a low speed to me; what type of aircraft is this instrument targeted for? This is a low speed indeed, the full sentence reads 'The aircraft speed relative to air (IAS) is set to 65 m s^{-1}, representative for a medium-range turboprop aircraft at measurement speed'. The system is intended for use aboard a medium-range turboprop aircraft, the Dornier 128-6 (D-IBUF), which measures at such a speed (Corsmeier et al., 2001).

Lines 20 – 25: regarding the difference between aircraft heading and ground track...I assume you are talking about drift here. Note that Guimond et al. (2018) found an error in the Lee et al. (1994) mapping equations for Earth-relative coordinates, which don't contain a correction for drift. If you are incorporating drift into your mapping coordinates, this correction should be applied. Guimond, S.R., J.A. Zhang, J. Sapp and S.J. Frasier, 2018: Coherent turbulence in the boundary layer of Hurricane Rita (2005) during an eyewall replacement cycle. *J. Atmos. Sci.*, **75,** 3071- 3093. We are talking about drift indeed. The error reportedly found by Guimond et al. (2018) does not affect our simulations, as the unbiased wind direction retrieval results show. We base our equations on the approach by Leon and Vali (1998). We have removed an erroneous reference to Lee et al. (1994) and replaced it with the correct one to Lenschow (1972). We have also adjusted our notation by switching the signs in eq. (1-3) in order to be fully consistent with them.

Regardless of the above, in this paragraph, we are concerned about the change in sampling distance in the LES due to drift. As the change in sampling distance due to drift is important, we have reworded and explained with an example in order to make this clear.

Line 25: I don't understand this sentence and the bold claim that this is the "first presentation of a correct airborne sampling simulation", please explain more clearly. Due to the general shortening, we have removed this statement. Nevertheless, we are not aware of any discussion of the effect of drift on the sampling spacing when simulating airborne measurement systems (if you are, please let us know). This is an important effect, and we believe this is the first description on how to take it into account in simulations.

Page 7

Top half of page: I am confused with this section and Appendix A1. The aircraft position (lat,lon,height) are provided by the GPS on any aircraft and the mapping equations provide locations of the pulse volume centers relative to these positions. Can't the authors just generate a realistic aircraft position vector (possibly from real data) and sample the model winds with that? This extra stuff seems irrelevant. If there is a large aircraft head wind, then the plane might only go forward very slowly and the wind retrievals would only cover a small region. In reality head winds are usually very small relative to the aircraft speed so I don't understand the motivation to get into all this detail. Just use a realistic aircraft position vector because ultimately this will be applied to real situations. We have significantly shortened this paragraph in order to make it more clear. Unfortunately, due to the effect of wind on aircraft track development, real GPS aircraft data should not be used in the ADLS for the time varying wind field (unless it would stem from a flight with the exact same wind field conditions and flight direction).

Further, due to the slow measurement speed of the aircraft, the effect of wind speed on track development is not small. We now illustrate the magnitude of the effect with a small calculation example in the revised manuscript.

Line 26: Is the 20-second full circle scan time (3 revolutions-per-minute) the lidar scan rate used in the remainder of the study? So with the aircraft speed of 65 m/s, the along-track spacing is 1.12 km? Make these parameters more clearly stated in the paper. Also, what is the grid spacing of the wind retrievals? In the revised, simplified study we are now only using the 20-second full circle scan time in order to avoid confusion. This gives an along-track spacing of 1300 m (20 s*65 m s^{-1} = 1300 m). We are now illustrating the system setup and retrieval procedure, as well as the new checkerboard approach in fig. 6.

Section 2.4 Retrieval – nadir as an example application: I suggest removing this section. The paper is getting much too long and this method has little practical use. Also, the authors say that "wind profiling" is the focus of this study rather than the nadir method. We have significantly shortened the manuscript. In our opinion, this section provides an illustrative example of the ADLS sampling, as well as the turbulence present in the LES. Due to the significant shortening in other parts of the manuscript we would therefore like to keep the nadir section as part of the manuscript as it is.

Page 12

Line 15/16: "The model is given by the beam geometry...". As stated, this is incorrect, the model is the radial velocity equation, which includes the beam geometry. We have reworded this, please also see our answer regarding the LSQ-fit procedure and the terminology followed above.

Line 22: "...the wind field is usually assumed to be homogeneous...". Again, I don't understand what you mean by homogeneous here. See major comments. Please see our answer to your major comment 2, as well as our answer to your previous comment on this above.

Page 13

Equations (8) and (9): The matrices U, S and W are not defined so I have no idea how they are used. It is hard for me to evaluate this paper without proper identification of variables. We are sorry that introduction of these variables slipped our attention, we are now identifying these variables (they are based on the standard SVD procedure).

References

Banakh, V. A., Smalikho, I. N., Köpp, F., & Werner, C., Representativeness of wind measurements with a CW Doppler lidar in the atmospheric boundary layer. Applied optics, 34(12), 2055-2067, 1995.

Boccippio, D. J., A diagnostic analysis of the VVP single-Doppler retrieval technique. Journal of Atmospheric and Oceanic technology, 12(2), 230-248, 1995.

Corsmeier, U., Hankers, R., & Wieser, A., Airborne turbulence measurements in the lower troposphere onboard the research aircraft Dornier 128-6, D-IBUF. Meteorologische Zeitschrift, 10(4), 315-329, 2001.

Guimond, S. R., Zhang, J. A., Sapp, J. W., & Frasier, S. J., Coherent turbulence in the boundary layer of Hurricane Rita (2005) during an eyewall replacement cycle. Journal of the Atmospheric Sciences, 75(9), 3071-3093, 2018.

Kiemle, C., Wirth, M., Fix, A., Rahm, S., Corsmeier, U., & Di Girolamo, P., Latent heat flux measurements over complex terrain by airborne water vapour and wind lidars. Quarterly Journal of the Royal Meteorological Society, 137(S1), 190-203, 2011.

Koscielny, A. J., Doviak, R. J., & Zrnic, D. S., An evaluation of the accuracy of some radar wind profiling techniques. Journal of Atmospheric and Oceanic Technology, 1(4), 309-320, 1984.

Lee, W. C., Dodge, P., Marks Jr, F. D., & Hildebrand, P. H., Mapping of airborne Doppler radar data. Journal of Atmospheric and Oceanic Technology, 11(2), 572-578, 1994.

Lenschow, D. H. The measurement of air velocity and temperature using the NCAR Buffalo aircraft measuring system. Boulder: National Center for Atmospheric Research, 1972.

Lenschow, D. H., & Stankov, B. B.: Length scales in the convective boundary layer. Journal of the Atmospheric Sciences, 43(12), 1198-1209, 1986.

Leon, D., & Vali, G., Retrieval of three-dimensional particle velocity from airborne Doppler radar data. Journal of Atmospheric and Oceanic Technology, 15(4), 860-870, 1998.

Lundquist, J. K., Churchfield, M. J., Lee, S., & Clifton, A., Quantifying error of lidar and sodar Doppler beam swinging measurements of wind turbine wakes using computational fluid dynamics. Atmospheric Measurement Techniques, 8, 2015.

Scipion, D.: Characterization of the convective boundary layer through a combination of large-eddy simulations and a radar simulator, Phd, University of Oklahoma, 2011.

Scipión, D., Palmer, R., Chilson, P., Fedorovich, E., and Botnick, A.: Retrieval of convective boundary layer wind field statistics from radar profiler measurements in conjunction with large eddy simulation, Meteorol. Zeitschrift, 18, 175–187, https://doi.org/10.1127/0941-2948/2009/0371, 2009.

Shapiro, A., & Fedorovich, E., Coriolis effects in homogeneous and inhomogeneous katabatic flows. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134(631), 353-370, 2008.

Stawiarski, C.: Optimizing Dual-Doppler Lidar Measurements of Surface Layer Coherent Structures with Large-Eddy Simulations, KIT Scientific Publishing, Karlsruhe, 2014.

Stawiarski, C., Traumner, K., Knigge, C., and Calhoun, R.: Scopes and challenges of dual-doppler lidar wind measurements-an error analysis, J. Atmos. Ocean. Technol., 30, 2044–2062, https://doi.org/10.1175/JTECH-D-12-00244.1, 2013.

Stull, R., *Meteorology for scientists and engineers*. Brooks/Cole Thomson Learning, Pacific Grove, USA, 1-528, 2000.

Tennekes, H., Lumley, J. L., & Lumley, J. L., *A first course in turbulence*. MIT press, Cambridge, USA, 1-310, 1972.

Figures from Stawiarski (2014)

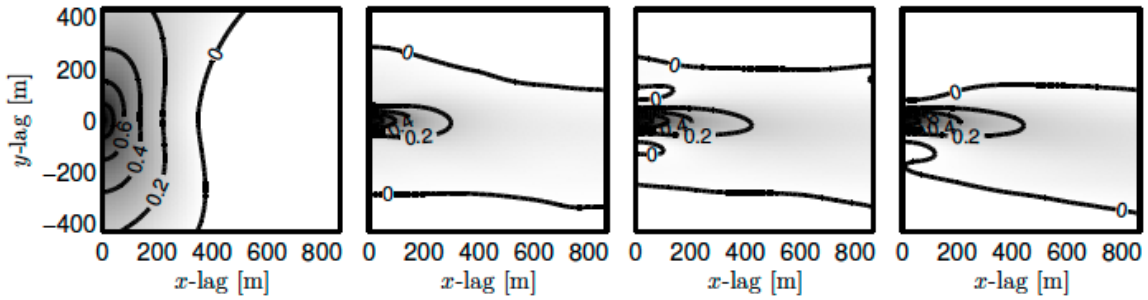

Figure 6.3.: Development of the spatial autocorrelation in the LES u wind fields with the background wind, $u_G = \{0, 5, 10, 15\}$ m/s from left to right.

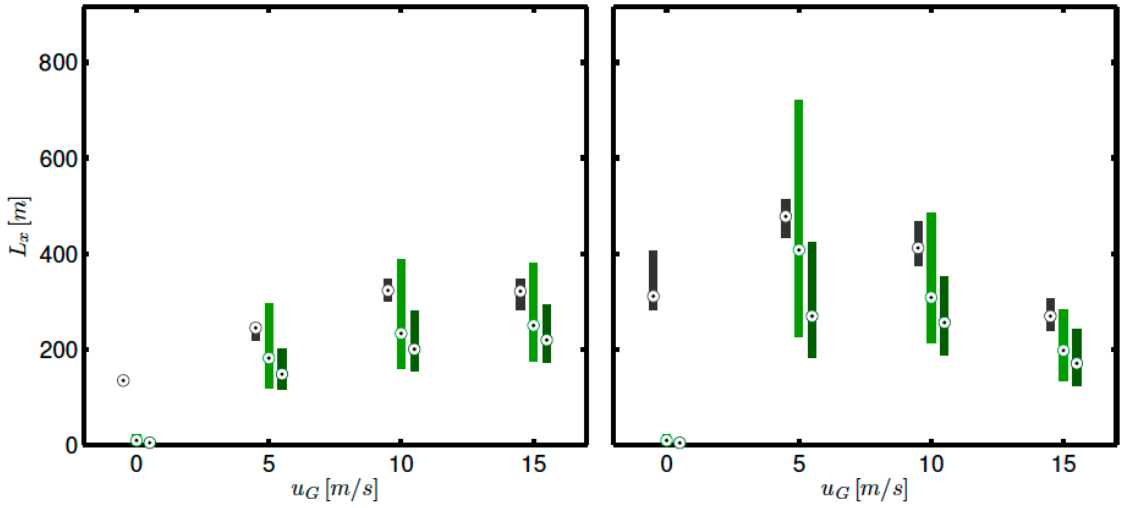

Figure 6.6.: *Integral length scales*: Scales L_x of the u (left) and v (right) wind fields in mean wind direction for the time series, computed with the full mean and variance of the data set (light green) and the mean and variance of the respective series (dark green). The black bars show the comparative LES results. The bars cover the range between the 25th and 75th percentile, the circles mark the median.

Dear Referee #3,

We thank you for the review of our manuscript and your detailed remarks. We have revised the manuscript, considering your remarks and those from the other referees. The revision led to a substantially changed and shortened version of the manuscript.

In order to allow a focus on the most important results we now reduced the content to the introduction of the LES-based airborne Doppler lidar simulator. Following this, illustrating the benefits of this new tool, the error in wind profiling due to violation of AVAD retrieval assumptions by boundary layer turbulence is investigated for a standard system setup and retrieval strategy. The result sections on system setup and retrieval strategy characteristics are not part of the manuscript anymore due to their lengthy nature as well as comments by other referees.

Please find our answers to your major and specific comments below. Many thanks for your work so far and kind regards,

Philipp Gasch and Co-authors

Initial, short replies are marked in this color

The final response addition is marked in this color

Major comments:

1. The LES domain is too short for this type of study. In the 5 km x 5 km x 1.8 km domain, the authors only use 2 points within to do the analysis for a given trajectory. And approximately 20 trajectories are used in their analysis. The authors also mention that they do not have enough statistics, which is a big concern in such a study. The authors base their methods based on Stawiarski 2014 2015, which is for a ground-based Doppler Lidar and who were looking at detection of coherent structures and the ability of Lidars to resolve those structures. But here, the authors are trying to assess the accuracy of the in-homogeneity, and more simulations and larger domains are a necessity. I can understand that the authors want to have high-resolution LES model data to validate their simulator, but statistically these results cannot be taken seriously. Again, it is very well known that flow in-homogeneity causes errors in Doppler Lidar retrievals, but quantifying the amount for airborne Doppler Lidar measurements is key. And the reviewer feels, that the quantification part is missing here, which is key.

It seems that a misunderstanding occurred. Because the error due to flow inhomogeneity is a well-known issue, we do use the simulator to quantitatively assess this error. (We are not trying to validate the simulator and/or assess the accuracy of the inhomogeneity, further we don't state that our statistics are insufficient). To our knowledge, the LES-based analysis is new, which is why we present the first LES-based airborne Doppler lidar simulator.

In the study we are using 16 different flight directions (called trajectories by you) through the LES domain. Each flight direction yields 2 retrieved wind profiles. The 16 flight directions are repeated 25 times at 1-minute temporal spacing (the wind field is

not assumed to be frozen during each passage but evolving instead). Thereby, this procedure yields 800 independent wind profiles for each of the four background wind cases (3200 wind profiles overall). Additionally, each of the wind profiles consists of 12 wind speed retrievals at different vertical levels (these 12 wind profile *points* of each wind profile are indeed correlated, as they are in reality), giving a total number of 9600 wind profile points for each background wind case (38400 wind profile points overall). If you are concerned about the spatial and temporal independence of the wind profiles we would like to refer you to our short reply to referee #1.
The 3200 independent wind profiles (38400 retrieved wind profile points) are much more than what is typically used in real world measurement comparisons (e.g. 1612 wind profiles in Weissmann 2005, approx. 10 wind profiles in DeWekker 2012, a single profile in Kavaya 2014, 2056 wind profile points in Bucci 2018).

We would thereby like to know what is insufficient about our statistical database and what you would consider a sufficient statistical database for a quantitative error analysis.

We have significantly altered the sampling procedure used to obtain the wind profiles as well as reworded its description. A newly developed checkerboard approach is used and illustrated (fig. 6). Using this approach a sufficient number of profiles are sampled, while maintaining statistical independence through clear temporal and spatial separation. We are also investigating the independence of the sample wind profiles in-depth in the revised manuscript version (Sec. 3.2). Even with the reduced checkerboard sampling approach, our statistical basis is larger than that of other studies reported in literature (who were using real world measurements which are inherently more expensive and complicated to conduct). We are emphasizing the comparably large statistical sample size more clearly in the revised version of the manuscript (we have to correct a small error in our initial answer: Weissmann et al. (2005) used 33 wind profiles for comparison, not 1612).

2. The optical effects of the laser beam due to motion has not been considered in the results. It is just demonstrated that the simulator can do motion correction. It would complicate the analysis and make it more realistic and interesting to study such realistic effects (Kavaya 2014 Hill 2008) rather than only flow in-homogeneity (again its a well known issue). Maybe the authors can take motion data from a real aircraft and simulate the effects of wind retrieval accuracy.

As you noted, we do consider the geometric transformations applicable to real world airborne Doppler lidar measurements. Further, we also investigate the effect of a measurement system imperfection. We do this through an assumed scanner pointing inaccuracy (called beam pointing inaccuracy by us), which is likely to occur in real world measurements. We are thereby able to analyze the effect of an imperfect scanner and/or aircraft inertial navigation system, which is done in Sec. 4.1. Further, we also analyze the effect of an imperfect laser system by adding a random radial velocity fluctuation to the measured radial velocities.

We do not see an advantage in using real world aircraft motion data, as the actual motion of the aircraft does not influence the VVP (not VAD) wind speed retrieval quality (only the accuracy with which the motion is known, which we take into account through the beam pointing inaccuracy).

We would thereby like to know what optical effect of the laser beam due to motion you are referring to.

Following referee comments by Mr. Brewer, we have removed the noisy system simulation from the manuscript. We now focus purely on an assumed ideal measurement system in order to highlight the wind profiling error introduced due to violation of the AVAD assumptions in the turbulent boundary layer.

3. The article has several repetitive statements all over the document and is extremely long. I feel a large portion of this article can be moved to the supplement section or deleted. The sentence formations are very abrupt, which sometimes I feel are unnecessary and not adding any additional information. I have noted a couple below, but I would urge the authors to take a deeper look at the entire article.

We will make sure to have another critical look at the wording of the manuscript. Further, we will also shorten its length.

As stated above, we have significantly revised and shortened the manuscript.

4. The Figures are not very helpful in understanding the results. They all are dumped into one figure with results from several simulations and different type of statistics. Figures 6-16 are very confusing and do not make it obvious without reading the winded text. I can understand the authors sentiment in showing a lot of information, but they need to use smarter approaches such as Taylor's diagram or similar.

We are sorry for the inconvenience. It is very difficult to convey the amount of information we present without a detailed description in the text. We will make sure to have a look at alternative forms of displaying the information as well as a more compact wording.

We have tried to reduce the amount of information we try to convey. The problematic figures mentioned by you are not part of the manuscript anymore.

Specific comments:

Page 5 Line 10 – The flight temporal sampling is not clear. 1-minute temporal sampling, does that mean the aircraft takes 1 minute to travel the 5 km domain? Isn't that a function of the scan trajectory? Its not very clear. Is the wind assumed to be stationary over that 1 minute? We have changed the sampling procedure to a checkerboard approach. We are now illustrating the used checkerboard sampling procedure using figure 6 in order to make the sampling procedure clear. The time it takes to transect the LES domain depends on the distance, the aircraft speed and the wind speed (as drift is taken into account). The wind field is changing while the aircraft is flying through the box, as we are using the time-varying wind field approach.

Page 8 Line 7: Instead of saying "their" weighting function . . . please mention the type of weighting function. Was is a Gaussian or rectangular or? Similarly, other things in the paper. It is a Gaussian weighting function and we now mention this.

Page 8 Line 8 – What is deltaX? Please use lidar technical terms or variables to be consistent with the literature. We have changed the notation to be consistent with Stawiarski et al. (2013) for lidar technical terms.

Page 8: I feel a short table of these characteristics would make it much more legible and easier for the readers. We hope that adapting the notation and rewording is sufficient.

Page 8 line 19: “can be added to the simulated wind field”? Please refine, as it has been added in the flow field, is my understanding. The sentence read: „Therefore, in the ADLS, a Gaussian noise with standard deviation σ_{vD} can be added to the simulated measured radial velocities.“ As stated, the noise was added to the simulated radial velocities. This statement has now disappeared due to the focus on an ideal measurement system.

Page 8 line 9 – What is FME? The acronym comes doesn’t make sense with the sentence before. We are not using the acronyms anymore following the shortening of the manuscript. FME was measurement frequency.

Section 2.3 – A motion correction algorithm is introduced, but the author mentions that they have not used any of that in their paper. I feel this section is not useful for this paper, if not used. Please cite a reference, such as Kavaya et al., 2014 or Hill et al., 2009, which has similar motion correction algorithms implemented to real lidar data. If its different that these algorithms, then please state only the difference. Please move this the next section to supplement. We are using the motion correction algorithm in section 2.4 (nadir retrieval), as well as figure 1. Therefore, we would like to keep the statement, but as before we are not presenting the algorithm itself as it is simple.

Page 3, Line 26. . . Remove “These include” and just list the references. Done.

Page 5 Line 6... Gravity waves being present above what? – Sentence structure is wrong! The sentence structures sometimes are too abrupt and unnecessary. So maybe the authors have to give it another read and remove some of the abrupt sentence structures. Such as Page 4, Line 26, it is not necessary to have the first statement “The simulator is tested with a set of underlying wind fields”. There are other instances such as this and needs to be addressed. I understand that it’s a personal choice of the reviewer. We have significantly rewritten the manuscript and hope to provide a smoother reading now. As you say, this is a matter of personal choice as well.

Page 27 line 11: why bold number of retrieved parameters (VAR)? And similar other locations. Please change. We are not using the acronyms anymore following the shortening of the manuscript.

Since there were too many issues with the paper, I feel like the authors need to revise the paper carefully and provide a better and shorter version for the reviewers. This was done.

References

Bucci, L. R., O'Handley, C., Emmitt, G. D., Zhang, J. A., Ryan, K., & Atlas, R. (2018). Validation of an Airborne Doppler Wind Lidar in Tropical Cyclones. *Sensors*, *18*(12), 4288.

De Wekker, S. F. J., Godwin, K. S., Emmitt, G. D., & Greco, S. (2012). Airborne Doppler lidar measurements of valley flows in complex coastal terrain. *Journal of Applied Meteorology and Climatology*, *51*(8), 1558-1574.

Kavaya, M. J., Beyon, J. Y., Koch, G. J., Petros, M., Petzar, P. J., Singh, U. N., ... & Yu, J. (2014). The Doppler aerosol wind (DAWN) airborne, wind-profiling coherent-detection Lidar system: overview and preliminary flight results. *Journal of Atmospheric and Oceanic Technology*, *31*(4), 826-842.

Stawiarski, C., Traumner, K., Knigge, C., and Calhoun, R.: Scopes and challenges of dual-doppler lidar wind measurements-an error analysis, J. Atmos. Ocean. Technol., 30, 2044–2062, https://doi.org/10.1175/JTECH-D-12-00244.1, 2013.

Weissmann, M., Busen, R., Dörnbrack, A., Rahm, S., & Reitebuch, O. (2005). Targeted observations with an airborne wind lidar. *Journal of Atmospheric and Oceanic Technology*, *22*(11), 1706-1719.

Dear Alan Brewer,

We thank you for the careful review of our manuscript and your detailed remarks. We have revised the manuscript, considering your remarks and those from the other referees. The revision led to a substantially changed and shortened version of the manuscript.

As stated previously, we reduced the content of this study to the introduction of the simulator. As you suggested, we highlight the ability of the LES-based simulator tool to investigate wind profiling error due to flow inhomogeneity with a basic example for a standard system setup and retrieval strategy. The simulator is then available for use by the community and specific setups, as suggested by you. We will explore the trade space of an airborne system and the effect of measurement system inaccuracies in detail in a future study.

Please find our answers to your major and specific comments below. Many thanks for your work so far and kind regards,

Philipp Gasch and Co-authors

Initial, short replies are marked in this color

The final response addition is marked in this color

An LES-based airborne Doppler lidar simulator for investigation of wind profiling in in-homogeneous flow conditions This is an excellent tool for evaluating airborne Doppler lidar measurement design and will be an invaluable part of experimental design. The authors have done an excellent job of addressing the key considerations of airborne Doppler lidar sampling strategy.

General comments: It was not clear to me if the paper was designed to illustrate the capabilities of the simulator using a single set of LES and lidar configurations or if the authors are attempting to explore the operational trade space of an airborne system with sufficient detail so that the community could use this paper as a design tool rather than the simulator itself. If the goal of the paper is to illustrate the application of the tool, then I think the authors need not explore as much of phase space in every degree of freedom. A single variable plot accompanied by discussion of the underlying mechanism for the relationship for each of the major impacts would suffice. Concentrate on those plots/relationships that show significant structure and focus there.

It is true that we tried to achieve both, introducing the simulator as well as exploring the trade space of an airborne system, at the same time. We now realize that this is too complex for one study. Therefore, we will reduce the content of this study to the introduction of the simulator. As you suggested, we will highlight the ability of the LES-based simulator tool to investigate wind profiling error due to flow inhomogeneity with a basic example. The simulator is then available for use by the community and specific setups, as suggested by you. We will explore the trade space of an airborne system and the effect of measurement system inaccuracies in detail in a future study.

We have proceeded as outlined in our initial answer. We now focus on the ideal system simulation only, as the noisy system simulation was lacking some aspects which can be important in reality (see answers below). Further, we are not exploring the trade space of an airborne Doppler lidar system anymore, as suggested by you. This is subject to a further study (including an improved noise simulation). Specific questions on system setups and retrieval strategies can be addressed directly by using the simulator after publication.

For example Fig 10 would benefit from another approach that highlights the differences between the two plots (highlighting the effect/impact of noise) and changing structure within a plot (slow increase in RMSE/REL). The actual values of the static variables in the plots are only meaningful for one particular set of conditions (both LES and lidar characteristics) and may not be of interest to a wide range of folks. How certain variables change wrt the chosen independent variable is of interest as it will generally do so for all configurations. These relationships should be highlighted and the underlying mechanism explained if possible. If the goal of the paper is to probe the trade space of all systems/conditions, then I think some effort should go into finding fewer "normalized" variables that are independent of specific choices for system values or LES characteristics. For example – my understanding is that the underlying bias sensitivity of the horizontal wind fit to coherent structures in the vertical wind field comes down to how many "pairs" of these features you average/accumulate over prior to doing a fit. As you increase the number of pairs of up/down motions in the volume – the residual of the fit may have a higher RMS, but the bias will decrease. This condition will depend on the dominant spatial scale of the turbulence, the size of the sample volume, the density of the sample points. . . A potential "normalized" variable might be the number of independent turbulent scale lengths per sample volume. For a given turbulence profile in the LES and a given scan pattern and beam PRF, one could imagine a profile of this "normalized" variable and corresponding uncertainty in the fit. . . If the authors are able to combine variables and break the analysis into the underlying mechanisms, it may serve to widen the impact of the paper and reduce the number of variables that have to be studied. I'm happy to discuss directly with the authors if need be – I'll ask the editor to share my contact information. For our application, we have a fixed scan geometry (wedge scanner) so the only system variables we have to adjust are scan rate, beam PRF & pulse width. The adjustable fit parameters are height resolution and number of sweeps to integrate over. The hope would be to use a vertical transect to characterize the strength and dominant spatial scale of the turbulence as a function of height then combine that with the system/scan parameters to come up with a normalized variable "number of turbulent scale lengths per sweep" as a function of height. The hope would be that we could use that profile and your results to determine the uncertainty in the horizontal wind fit as a function of number of sweeps integrated. . . Another set of variables that could be combined – SNR, beam PRF, and LOS vel uncertainty (using the CRLB discussed below). Once the results are expressed in these terms, they are no longer only applicable to the system defined in the study. At 50 pages, this is a long manuscript and the length may limit its impact and applicability. Much of the wind profiling theory section, evaluation of errors section (definitions) could be moved to appendices.

With the limitation of this study to the basic simulator and method description we cannot include more general findings, using normalized variables as suggested, in this study. Nevertheless, your remarks make total sense and we are happy to discuss these aspects directly with you in the near future. As you proposed, based on the simulator,

we have already developed a method to estimate the uncertainty of the retrieved wind profiles in-situ without the need for additional data except for some basic boundary layer parameters. Unfortunately, including this method in the study would be beyond the scope of this manuscript as it is too long already, but we are happy to share it with you immediately.

We are now focusing on a basic example as stated above. Exploration of the system trade space as well as generalization and uncertainty estimation will be investigated in a further study. According to your suggestion, we have moved many sections to the appendices.

Some concerns/questions on the approach taken in the study (some of these may be redundant with first section)

1. Vertically averaging/combining results when you have a height dependent footprint combined with a height dependent turbulence profile / integral scale. Are each of the 9600 profile points from different distances below the AC? If so – how does the dominant spatial scale of the turbulence compare to the spatial scale of the sample volume/arc? (ie “number of independent samples” within a scan arc).

The 9600 profile points are from different altitudes below the aircraft. As you say, they were determined using height dependent footprints as well as turbulence profiles in order to achieve a spread among these important input variables.

With our proposed new uncertainty estimation method the average ratio of the dominant turbulence scale compared to the scale of the sample volume is taken into account by calculating an effective sample size. Therefrom, the uncertainty is determined. Unfortunately, this is too complicated to include in this study, but as stated before we are happy to share it with you immediately.

We have now simplified the retrieval approach to the checkerboard sampling procedure. We hope that this make the sampling and retrieval more understandable. The uncertainty forecast method is prepared and will be published at a later point.

2. Sensitivity to static errors in pointing/orientation – does this just fall within the RMS of the assumed uncertainty in pointing or does a static offset impact the fit differently?

As you state below, our error emulation does not capture all aspects of errors encountered in real world systems. The effect of a static error in pointing/orientation depends on its origin, whether it is a static offset in the scanner pointing or the INS orientation. In general, it should lead to increased error levels and potentially a bias depending on the flight and wind direction. Due to the complexity of the situation and our incomplete emulation we will remove this part from the manuscript and investigate more in-depth in the discussed further study.

As suggested in our preliminary answer, we have removed the noisy system simulation from the manuscript due to the over-simplification as well as complexity. Possible system errors require further study, as there is a multitude of possible error sources and the associated influences are complex.

3. Dependence on one set of operational parameters (one lidar design) coupled with non-physical simplifications (constant uncertainty in vel as a function of height) is problematic. Need to find a set of independent variables that the user can calculate for their system / scan design.

As stated above, this will be the subject of a future study. Nevertheless, we are willing to share the method developed with you immediately.

As stated above, we have removed the noisy system simulation from the manuscript. Any interested user can use the simulator directly to answer specific questions related to their system / scan design.

4. Break analysis into single dimensions where possible and describe the underlying process if you have been able to glean that from the analysis so the user can project the result into their operational space.

The system trade space analysis will be removed from the manuscript and subjected to further study.

As the exploration of system trade space was removed from the manuscript we are happy to investigate this for specific user setups if needed.

5. Describe mechanism behind bias mentioned in paper in more detail.

The mentioned possible bias at low wind speeds will remain in the revised manuscript, as this is a new and important finding. We will explain in more depth.

The mentioned potential bias at low wind speeds is an important new finding of this study. We are now illustrating an AVAD LSQ-fit for the no-background wind speed case (fig. 4). This figure serves to illustrate the mechanism responsible for erroneous mapping of vertical wind into horizontal wind.

We will respond to the specific comments below in our final response together with the revised manuscript. Thank you once again for all the effort invested.

Specific Comments Pg 1 Line 10: “laser system noise” - detection uncertainty (see Cramer Rao Lower bound- CRLB) “beam pointing inaccuracy due to system vibrations” – We have found dominant “pointing” errors come from Inertial Navigation Unit (INU) orientation uncertainties. Most concerning are bias errors rather than RMS from vibration. As important error aspects were missing for the noisy system simulation we have removed this part of the study in order to extend research into these topics and publish them at a later point.

Pg 1 Line 11: “system setups” – define first use. This was reworded.

Pg 1 Line 16: “short horizontal averaging distances” Along track vs cross track? This was removed, we are now using equal along- vs. across-track averaging distances.

Pg 2 Line 9: “the vertical wind through nadir ” - the vertical wind with nadir. This was removed due to the shortening of the manuscript.

Pg 2 Line 20: “while neglecting the non-standard beam geometry” poorly worded / not clear. This was moved to Sec. 3.1 and ‘non-standard’ replaced with ‘altered’.

Pg 2 Line 22: anelastic -> an elastic. They actually use the anelastic mass continuity equation.

Pg 4 Line 13: “collinearity in model geometry” you mention this on Pg 2 in terms of the matrix inversion, but given that this is a primary research question, you should describe more fully. This will allow the reader to better interpret the phrase at the end of this sentence. We are not focusing on collinearity filtering using the condition number CN anymore (pseudo difficult conditions through sector blanking), but rather on filtering by the correlation coefficient R^2 in low wind speed situations. We have reworded accordingly.

Pg 4 Line 30: “5 x 5 x 1.8 km” I am concerned whether this will allow for sufficient independence in the multiple sampling configuration described at Pg 6 Line 10 (see comment below). By using one small set of data and varying the sampling heading angle of the plane through the domain, all paths share a common volume in the center and are not completely independent. This will impact each height differently due to sampling footprint as a function of distance below plane. I assume that the dominant turbulent spatial scales at each height are also different – so this further complicates the interpretation. We have substantially changed the sampling procedure in order to ensure spatial and temporal disjunct sampling. Further, we are now investigating the independence of the acquired wind profiles (Sec. 3.2).

Pg 5 Line 5: “The boundary layer height is approximately. . .” At first consideration, it would seem that an important quantity would be the relationship between the integral scale and the scan volume/sample arc length - how many "integral scale" lengths are averaged over in one scan/retrieval. Do you vary this relationship and, if so, is it done by changing the input wind field or only the scan configuration? After my initial reading, it seems you are averaging your results in height, so you may be diluting the effect by averaging over multiple conditions in height. As mentioned in the prior comment, you have both height dependent sample volumes and turbulent length scales. Indeed, the relationship between the integral scale and the scan volume/sample arc length is important. In our study it is inherently varied over a wide range by using retrievals from different heights below the aircraft, as well as the changing boundary layer structure and the different wind cases (fig. B4). With the newly developed uncertainty forecast method this parameter is accounted for and estimated.

Pg 6 Line 11: “As the different flight tracks lead to different air mass volumes being sampled, the different flight directions are independent and can be used to increase the sample size of the statistical analysis” This is the phrase I was basing the Pg4 line 30 comment about independence. We refer to our answer there.

Pg 6 Line 27: “For the frozen-in-time wind field, the LES coordinate” Are you considering atmospheric features that might be correlated to the ground - ie land usage or topographic effects? Making measurements in complex terrain might lead to making measurements in where coherent features in the vertical wind field might be present.

We are not considering features correlated to the ground at the moment, as the LES is using a homogeneous, constant surface heat flux. We agree, this effect is very important to consider when making measurements in complex terrain and/or different land usage scenarios. It definitely requires further study but unfortunately is beyond the scope of this study.

Pg 7 Line 29: “direction noise” Several issues to consider here: 1) While it may end up being a similar effect in the end our experience is that the estimation of orientation angles p,r,y and their angular rates are more prone to error then reading an encoder on the scanner. 2) The uncertainty in pointing angle also feeds into the LOS plat- form motion correction algorithm which then feeds back into the wind profile. 3) The static/low frequency component of the pointing offset is more problematic than a zero mean Gaussian noise source on the pointing. Depending on the inertial navigation unit (INU), errors in the drift correction of the sensor can lead to low frequency/static errors. 4) Latency in the communication with INU sensor and fast scanning/beam rates can lead to static offsets in orientation/angular rates estimation. Because of this comment, as written above, we have removed the noisy system simulation from the manuscript due to the over-simplification as well as complexity. We agree, that all other important effects mentioned by you can occur and need to be investigated. Due to the multitude of possible system errors this requires further study, as the associated influences are complex.

While you can use the ground to “calibrate” the static angle offsets – time varying, low frequency effects can still be present in the data. Upward looking scans, cloud cover, operation over water could all lead to periods where ground strikes are not available. Being able to quantify the sensitivity of these errors propagated into the wind profiles would be great. We will investigate the effects of system noise in a further study.

Pg 8 Line 15: “detector noise, speckle effects and turbulence within the measure- ment volume” In areas where there are adequate aerosol, the dominant mechanism for uncertainty in the Doppler measurement comes from uncertainty in estimating the spectral peak (take a look at Frehlich, Coherent Doppler lidar measurements of winds in the weak signal regime). CRLB depends on SNR, pulse width, and averaging time – (beam rate, vertical resolution, distance from AC) these are quantities that are part of the trade space when designing a scan/sample strategy.

Within a beam you will get a range of SNR and hence uncertainty in the LOS Vel measurement (as a function of range). Setting up your runs with a single velocity error for all ranges is not representative of a true measurement. This effect is only exacerbated when you add a variable aerosol field as a function height as well.

Here, the same statement as above is valid. Due to the over-simplification in our existing approach we have removed this section from the manuscript.

Pg 8 Line 27 “System components” (and used elsewhere) this is an ambiguous term. In this case it would seem that you are referring to pointing vectors – but elsewhere you seem to have different meaning. System components refer to aircraft, scanner and lidar. We have reworded this section to clarify and be consistent throughout.

Pg 9 Line 6 “aircraft center of gravity...” The moment arm should be between the location of the INU and the center of the final turning mirror in the scanner. Correct, we changed it accordingly.

Pg 10 Line 2 “restore” do you mean isolate? Isolate is a better wording, we changed it accordingly.

Pg 10 Line 4 “system noise” No matter how well you apply the alignment calibration, there will always be a non-zero static offset. You should consider a sensitivity analysis of the wind profiles to static errors in the orientation. (rather than always assuming the errors are zero mean.) Here, the same statement as above is valid. Due to the over-simplification in our existing approach we have removed the noisy system simulation from the manuscript.

Pg 10 Line 34 “horizontal wind profile” It can remove the effect of a static profile, but not the natural variability. If you are trying to measure w’^2 – the variability in the horizontal wind will still manifest in the w’^2 profile. Correct, we changed the wording accordingly.

An LES-based airborne Doppler lidar simulator ~~for investigation of~~ and its application to wind profiling in inhomogeneous flow conditions

Philipp Gasch[1], Andreas Wieser[1], Julie K. Lundquist[2,3], and Norbert Kalthoff[1]

[1]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, 80303, USA
[3]National Renewable Energy Laboratory, Golden, 80401, USA

Correspondence: Philipp Gasch (philipp.gasch@kit.edu)

Abstract. Wind profiling by Doppler lidar is common practice and highly useful in a wide range of applications. Airborne ~~observations~~ Doppler lidar can provide additional insights relative to ground-based systems by allowing for spatially ~~resolved~~ distributed and targeted measurements. ~~This study prepares the ground for an upcoming airborne Doppler lidar system by investigating the measurement process theoretically. To evaluate the future system characteristics and measurementaccuracy~~ Providing a link between theory and measurement, a first ~~LES-based~~ LES (large eddy simulation)-based airborne Doppler lidar simulator (ADLS) has been developed. Simulated measurements are conducted based on LES wind fields, considering the coordinate and geometric transformations applicable to real world measurements. The ADLS provides added value as the input truth used to create the measurements is known exactly, which is nearly impossible in real world situations. Thus, valuable insight can be gained into measurement system characteristics as well as retrieval strategies.

~~The accuracy of retrieved wind profiles under inhomogeneous flow conditions in the boundary layer is investigated. In general, when using reasonable system setups,~~ As an example application, airborne Doppler lidar wind profiling is ~~possible with acceptable error margins. Results allow for determination of preferential system setups and wind profiling strategies. Under the conditions considered, flow inhomogeneities exert the dominant influence on wind profiling error. In comparison, both the errors caused by random radial velocity fluctuations due to laser system noise and beam pointing inaccuracy due to system vibrations are of smaller magnitude.~~

~~Airborne~~ investigated using the ADLS. For commonly used airborne velocity azimuth display (AVAD) techniques, flow homogeneity is assumed throughout the retrieval volume, a condition which is violated in turbulent boundary layer flow. Assuming an ideal measurement system, the ADLS allows to isolate and evaluate the error in wind profiling which occurs due to the violation of the flow homogeneity assumption. Overall, the ADLS demonstrates that wind profiling is possible in turbulent wind field conditions with reasonable errors (RMSE $= 0.36$ m s^{-1} for wind speed when using a commonly used system setup and retrieval strategy for the conditions investigated). Nevertheless, flow inhomogeneity, e.g. due to boundary layer turbulence, can cause an important contribution to wind profiling error and is non-negligible. Results suggest that airborne Doppler lidar wind profiling at low wind speeds (< 5 m s^{-1}) ~~requires adequate system setups, retrieval strategies and quality filtering as the retrieved wind speeds~~ can be biased~~otherwise. The utility of quality filtering criteria for wind profile reliability~~

~~(coefficient of determination and condition number) is examined. While the filtering by the condition number is useful for all circumstances, an inadequate coefficient of determination threshold can bias the retrieved wind speeds. Even with strict quality filtering criteria applied, considerable wind profile retrieval error can exist, especially for steep scan elevation angles of more than 70° from the horizontal or short horizontal averaging distances.~~ , if conducted in regions of inhomogeneous flow conditions.

1 Introduction

Doppler lidar has experienced rapidly growing importance and usage in remote sensing of atmospheric winds over the past decades (Weitkamp et al., 2005). Sectors with widespread usage include boundary layer meteorology, wind energy and airport management. Compared to ground-based Doppler lidar systems, airborne systems can provide advantages and are seen as a promising tool for future research (Baker et al., 1995, 2014; Davis et al., 2018). Airborne Doppler lidar extends the spatial coverage of flow phenomena, enabling stream-wise and span-wise investigation of flow phenomena (Kiemle et al., 2011; De Wekker et al., 2012; Chouza et al., 2016b) as well as targeted observations with rapid deployment (Weissmann et al., 2005; Zhang et al., 2018). Further, airborne Doppler lidar can serve as a testbed and validation tool for upcoming and existing space-based Doppler lidar systems (Paffrath et al., 2009; Lux et al., 2018; ESA, 2018; Baidar et al., 2018; Tucker et al., 2018).

Due to their benefits, various airborne Doppler lidar systems have been developed in the past. Most of the systems are based on long-range aircraft flying in the upper troposphere at high speeds (O(250m s^{-1})), thereby they are destined for sensing the free troposphere due to the coarser spatial resolution (O(5 km) for wind profiling) (Weissmann et al., 2005; Kavaya et al., 2014; Guimond et al., 2014; Tian et al., 2015; Baidar et al., 2018). Some deployments of airborne Doppler lidar have been reported based on medium-range aircraft flying in the lower troposphere and at slower speeds (O(50m s^{-1})). These systems yield higher spatial resolution (O(1 km) for wind profiling) and can enable boundary-layer studies ~~. Thus far, only results for wind profiling of the mean horizontal wind have been reported (De Wekker et al., 2012; Godwin et al., 2012; Koch et al., 2014). As of yet, a study analyzing the vertical wind through nadir or other measurement geometries such as Dual-Doppler have not been attempted for the medium-range aircraft. This study provides a critical component for an airborne Doppler lidar system for boundary-layer research by investigating flexible scanning geometries theoretically, considering both wind profiling and nadir measurements of the wind field.~~ (De Wekker et al., 2012; Godwin et al., 2012; Koch et al., 2014).

~~Airborne Doppler systems are often used to assess wind profiles of the mean horizontal flow (Weissmann et al., 2005; De Wekker et al., 2012; Kavaya et al., 2014; Guimond et al., 2014; Tian et al., 2015; Baidar et al., 2018). The theory and problems associated with airborne Doppler wind profiling are very similar between lidar and radar systems. Therefore, no distinction is made between the studies using either of the two instruments in the following, unless necessary due to explicit differences.~~

~~Similar to ground-based wind profiling, airborne wind profiling is usually conducted by scanning in conical scans along the flight path. Retrieval of the mean wind vector can be achieved through inversion of the beam matrix, yielding a least-squares~~

solution to the problem (Leon and Vali, 1998). De Wekker et al. (2012) and Tian et al. (2015) apply the original Velocity-Azimuth-Display (VAD) analysis directly, while neglecting the non-standard beam geometry due to aircraft movement. Other methods at higher computational cost exist as well. Guimond et al. (2014) discuss a variational approach which can improve the traditional solution by adding anelastic mass continuity constraints on the estimated solution. Accumulation of the Doppler spectra can be conducted for Doppler lidar and has been shown to extend the retrieval limits in clear air conditions with little return signal (Weissmann et al., 2005; Witschas et al., 2017).

Most currently used airborne wind Most currently used airborne wind profiling approaches use the airborne velocity azimuth display technique (AVAD) or closely related retrieval strategies to estimate wind profiles. These approaches assume homogeneous flow conditions throughout the sample volume . retrieval volume (e.g. no deviations from the mean flow condition). Especially in turbulent environments such as the atmospheric boundary layer, this condition homogeneity assumption is rarely fulfilled and wind profiling at high spatial resolution remains challenging (Leon and Vali, 1998; Guimond et al., 2014; Tian et al., 2015). The problem is further intensified by the fact that many airborne profiling systems use high scan elevation angles (closer to nadir)in order to constrain the footprint . Reasons for doing so include constraint of the measurement footprint, hardware and range limitations or signal attenuation. As a result, the retrieved measured radial velocity is strongly influenced by the vertical wind variation along the scan circle. Thereby, deviations from the mean homogeneous conditions violations of the homogeneous flow assumption can lead to non-negligible errors in the retrieved mean wind vector (Tian et al., 2015; Bucci et al., 2018). This problem also applies to ground-based wind profiling in complex terrain, at short sampling durations, e.g. when analyzing single scans, or for short sector scans (Cheong et al., 2008; Bingöl et al., 2009; Wang et al., 2015).

The coefficient of determination is often used to detect a violation of the homogeneity assumption (Päschke et al., 2015; Wang et al., 2015). When using this approach, it is assumed that deviations from the homogeneous state show up as deviations of the measurements from the LSQ-fit. If the matrix inversion is performed based on a singular value decomposition, additional quality criteria such as the condition number, describing the degree of collinearity among the model geometry, become available and are frequently utilized (Boccippio, 1995; Holleman, 2005; Cheong et al., 2008; Shenghui et al., 2014; Päschke et al., 2015; Wang et al., 2015). However, as shown already by Koscielny (1984), a linear change in the vertical wind biases the retrieved horizontal components and a linear change in the horizontal components biases the retrieved vertical component. Both deviations are not detectable as deviations from the LSQ-fit.

A common method to assess the reliability of retrieved Doppler lidar wind profiles is by comparison to wind profiles from other measurement systems. For airborne systems, additional problems exist for in-situ comparisons: instrumented towers are of limited use due to their small vertical extent and fixed location (as well as lidar measurement problems due to chirp close to the ground detailed by Godwin et al. (2012)), simultaneous aircraft measurement are challenging and expensive to execute (and still suffer from co-location problems) and systems with similar remote sensing characteristics also suffer from co-location problems and results can show large differences (De Wekker et al., 2012; Tian et al., 2015; Bucci et al., 2018). Therefore, the most prominent approach is the comparison of retrieved wind profiles to radiosondes and/or dropsondes, which can provide verification if conducted for a sufficiently large dataset (Weissmann et al., 2005; Chouza et al., 2016a; Bucci et al., 2018).

Nevertheless, both systems still exhibit very different sampling characteristics and volumes~~,~~. This difference often making a direct comparison of the results challenging as it is difficult to determine if the observed deviations occur due to the differing sampling volumes, ~~spatial inhomogeneity~~ violation of the homogeneity assumption or actual instrument error.

Due to these challenges, idealized simulations of Doppler measurement systems can provide detailed insight into the ~~limitations and capabilities~~ capabilities and limitations of these systems. Early studies determined the representativeness of Doppler lidar measurements based on a statistical description of turbulence and for idealized system set-ups. Many of the first studies emphasized the effects of measurement geometry and turbulence on system characteristics and performance (Waldteufel and Corbin, 1978; Boccippio, 1995; Banakh et al., 1995; Baker et al., 1995; Frehlich, 2001; Banakh and Werner, 2005). As a result, the reliability of measured radial velocities under different turbulent intensity conditions and its impact on retrieval quality are well described.

With increasing computational capabilities, a numerical approach based on simulated atmospheric wind fields became feasible. For ground-based systems, a number of investigations detailing the error characteristics associated with wind profiling exist based on ~~LES-simulated wind fields . These include Muschinski et al. (1999); Scipión et al. (2009); Wainwright et al. (2014); Lundquist et al. (2015) and Klaas et al. (2015).~~ LES (large eddy simulation)-simulated wind fields (Muschinski et al., 1999; Scipión et al., 2009; Wainwright et al., 2014; Lundquist et al., 2015; Klaas et al., 2015).

For airborne (or satellite-based) systems, the moving platform alters the measurement process and viewing geometry, thereby introducing new problems. These ~~challenged~~ challenges have been investigated with statistical models (Baker et al., 1995; Gamache et al., 1995; Frehlich, 2001) or real measurement data (Leon and Vali, 1998; Weissmann et al., 2005; Tian et al., 2015; Chouza et al., 2016a). For airborne systems, Lorsolo et al. (2013) and Guimond et al. (2014) show the importance of model-based simulator studies, while relying on coarser resolution model output and focused on errors introduced due to the measurement system inaccuracies. So far, to our knowledge, a simulation of an airborne wind profiling system with complex scanning geometries and high resolution atmospheric wind fields (O(100 m)) is missing.

Extending previous studies, this work ~~aims to investigate~~ investigates the impact of wind field inhomogeneities on airborne wind profiling at ~~high resolution (evaluation of single scans,~~ highest resolution (O(1 km)~~for wind profiling). The focus of this study is on the error introduced by the wind retrieval algorithm due to the mismatch between assumed homogeneous wind field models and the wind field inhomogeneities present during the measurement process.~~), Towards this goal, an LES-based airborne Doppler lidar simulator (ADLS) ~~allowing for simulation of various systems setups and retrieval settings~~ is presented. Using the ADLS the measurement and retrieval process can be replicated in great detail, however, with the advantage of knowing the atmospheric input data in the sampling volume exactly. In an example application of the ADLS, this study investigates airborne wind profiling at high resolution (evaluation of single scans, O(1 km) for wind profiling) in a turbulent, inhomogeneous wind field. The error observed in the ADLS between input and retrieved wind profile is ~~then~~ directly traceable to ~~wind field inhomogeneities and its magnitude can be evaluated against other sources of error~~ (the violation of the wind field homogeneity assumed in AVAD, e.g. ~~random radial velocity fluctuations and beam pointing direction inaccuracy).~~ due to

boundary layer turbulence. This direct approach is not possible in other measurement system comparisons or simulations so far.

Consequently, the ~~research questions addressed are as follows: Is~~ question addressed is whether the violation of the homogeneity assumption~~relevant for~~ , due to turbulent boundary layer flow, is a relevant driver of error in airborne Doppler wind profiling ~~? Is it possible to determine optimal system setups and retrieval settings under inhomogeneous conditions? Can the violation of the homogeneity assumption and collinearity in model geometry be reliably detected from the measurements? In other words, are the coefficient of determination and condition number valid quality filtering measures for wind profiling quality?~~ at high spatial and temporal resolution.

~~The outline of this paper is as follows:~~ The structure of the study is as follows. In the first section, the ADLS is introduced. This section contains a description of the underlying wind fields, the system simulation characteristics, as well as an outline of the measurement procedure. Further, a simulated nadir transect is analyzed as an illustrative example. In the second section~~the ADLS framework is introduced, in the third section the~~ , the error associated with airborne Doppler wind profiling ~~theory is provided and in~~ in inhomogeneous flow conditions is investigated using the ADLS. To this end, the wind profiling retrieval and associated error metrics are explained. Then, the ~~fourth section airborne wind profiling error is evaluated for various system setups and retrieval settings.~~ sampling procedure and independence of the sampled wind profiles are discussed. After the evaluation of wind profiling error in inhomogeneous flow, the section concludes with an investigation of commonly used quality filtering criteria. At last, the conclusions are given.

2 Airborne Doppler lidar simulator

~~Extending previous studies, this work investigates the impact of wind field inhomogeneities on airborne wind profiling at highest resolution (O(1 km)) using an LES-based airborne Doppler lidar simulator.~~ The description of the ADLS consists of four sections outlining and mimicking the ADLS structure and operation. First the underlying wind field options are specified, then the simulation of the airborne Doppler lidar system components is discussed and the measurement procedure is explained. Last, the nadir or wind profile retrieval can be performed on the simulated measurement data. ~~The wind profiles are then evaluated against the original wind field data supplied as input.~~

~~Simulator operation scheme and settings for the wind profiling quality analysis. System characteristics and retrieval options are varied according to the specifications given in Sec. 2.2 and Sec. 3. The corresponding abbreviations can be found in Tab. ??. Standard values used for system setup and wind profile retrieval are marked in red. Parameters which are always varied to generate statistical dataset are marked in blue.~~

2.1 Atmosphere - Wind field

~~The simulator is tested with a set of underlying wind fields.~~

In order to be as close as possible to a realistic measurement environment this study utilizes time-varying LES generated wind fields~~(WND)~~. The LES fields are simulated using the parallelized LES model (PALM) and provided by the University of Hanover (Raasch and Schröter, 2001). ~~The LES~~

When using LES, a trade-off between the resolution of the simulation and the domain extent has to be realized, as computational power is limited. On the one hand, the resolution has to be sufficient to ensure a realistic simulation of turbulence and associated Doppler lidar wind measurements. Doppler lidar measurements are assumed to be represented realistically when the lidar range gate length Δp is much larger than the grid spacing Δx of the LES simulation (Stawiarski et al., 2013). On the other hand, the domain has to be large enough to ensure a sufficient number of independently sampled wind profiles for statistical analysis, given the investigated turbulence characteristics, system setup, retrieval strategy and sampling procedure.

The LES used in this study employs a grid spacing of ~~10 m~~ $\Delta x = 10$ m (corresponding to a resolution finer than O(100 m)) and a domain size of ~~5 x 5 x 1.8 km, data output started with fully developed turbulence after a spin-up time of 1 h. The LES~~ 5 km x 5 km x 1.8 km. The grid spacing fulfills the condition $\Delta p >> \Delta x$ at the assumed range gate length $\Delta p = 72$ m (Sec. 2.2.3). Further, the domain size is sufficiently large to sample a sufficient number of independent wind profiles for statistical analysis, given the turbulence characteristics present, an adequate system setup, retrieval strategy and sampling procedure, as is shown in Sec. 3.1. The ADLS can be easily adopted for use with other LES wind fields (e.g. a larger or longer simulation domain), should this be necessary for future studies.

The present LES is driven with a geostrophic background wind of $u_G = 0, 5, 10$ and 15 m s^{-1} and a constant kinematic surface heat flux $\overline{w'\Theta'} = 0.03$ K m s^{-1} ($\overline{w'\Theta'} = 0.23$ K m s^{-1} for the 0 m s^{-1} background wind case). Data output started with fully developed turbulence at a temporal resolution of 1 s after a spin-up time of 1 h. The LES wind fields are the same as used by Stawiarski et al. (2013), and so a detailed description is available in Stawiarski (2014)[chap. 5.1] and Stawiarski et al. (2015). The convective situation is classified as unstable stratification with the stability criteria pointing to the development of organized convective structures for $u_G > 0$ m s^{-1}. The boundary layer height is approximately 600 m (1200 m for the 0 m s^{-1} background wind case), with the entrainment zone extending from $600 - 1000$ m~~and gravity waves being present~~. Gravity waves are present in the entrainment zone and above. The stream-wise integral length scales for u and v are in the range of $200 - 500$ m, whereas the span-wise integral length scales are much smaller in the range $50 - 200$ m (Stawiarski et al., 2015). Due to the in-depth description in Stawiarski et al. (2015) the used LES data ~~is~~ are not detailed further here.

~~The virtual aircraft transects the LES 25 times at 1-minute temporal spacing, as a singular flight through the LES does not yield sufficient statistics. At the given turbulence characteristics, flight directions and advection speed, the temporal and horizontal spacing is sufficient for the sampled air volume and turbulent structures to be different from each other during each transect.~~

~~The ADLS processes time-varying LES wind fields and the input can be adapted to other LES models easily. As an alternative to the time-varying LES wind fields, a fixed LES time step could be utilized because the virtual aircraft and scanner move through the LES data. This corresponds to the assumption of frozen turbulence which is reasonable for an airborne system. Both approaches were tested as a part of this study, yielding very similar results. In order to be as close as possible to reality, albeit requiring higher computational cost, only results obtained from the time-varying wind field are presented in this study.~~

2.2 System - Airborne Doppler lidar components

In the second section all main system components consisting of aircraft, scanner and lidar are simulated geometrically. An illustrative overview of the results obtained after system simulation is shown in fig. 1.

2.2.1 Aircraft

The aircraft trajectory is specified by the coordinates of the desired start and end points together with a prescribed aircraft speed relative to air. Curvilinear trajectories are also possible with a specified turn time.

To evaluate the effect of the aircraft motion relative to the wind, the aircraft transects the LES volume in 16 directions, with a heading difference of 22.5° between each (DIR). As the different flight tracks lead to different air mass volumes being sampled, the different flight directions are independent and can be used to increase the sample size of the statistical analysis. Consequently, the flight directions are varied for all investigations to generate a greater number of wind profiles for statistical analysis.

Another parameter varied Another parameter that can be varied is the aircraft flight altitude(AAL). Here, a standard value of 1700 m is chosen, representative for an unpressurized turboprop aircraft flying above the boundary layer under visual flight rules. The effect of lowering the AAL to 1450 m and 1200 m is analyzed .

For a time-varying wind field, as used in this study. Lower flight levels yield a smaller footprint of the measurement, however, the number of retrieved wind profile points is decreased as less atmosphere is covered vertically.

The aircraft speed relative to air (IAS) is set to 65 m s^{-1}, representative for a medium-range turboprop aircraft at measurement speed.

Depending on whether a frozen-in-time or time varying wind field is chosen, the aircraft , the aircraft trajectory and sampling positions inside the LES must be calculated differently than for ground based measurement systems because the aircraft moves relative to the air mass and not to the ground. As the ground, which results in a difference between aircraft heading and ground track (Lee et al., 1994). Therefore, depending on whether the LES coordinate system coincides with the groundor the air mass, different sampling distances have to be , the sampling distances applied in the LES . To our knowledge, this manuscript is the first presentation of a correct airborne sampling simulation for LES studies, therefore it is explained in-depth in the following.

For the frozen-in-time wind field, the LES coordinate system coincides with that of the air mass. The air mass and turbulent elements contained within are not advected through the domain during the measurement process. Thereby, the sampling is done at equidistant intervals in LES space along the flight trajectory. The spacing of the sampling points is calculated through the simple formula $s = IAS \cdot t$. Consequently, for a given sampling time and aircraft speed an equal volume of air mass (and turbulence) is sampled as is done by a real aircraft. However, in this case the aircraft motion due to the wind speed needs to be factored into the retrieval at a later point using a triangle of velocities calculation (Appendix A1), because the aircraft track have to be adjusted for the actual distance covered by the aircraft with respect to the groundis influenced by the wind speed. To illustrate the concept, albeit being unrealistic, consider an aircraft flying at 65 m s^{-1}, either up- or downwind, with a LES

wind speed of 65 m s^{-1}. The frozen-in-time wind field sampling distance in LES space will be the same, but compared to the ground, the aircraft will have moved a large distance in the first case and not at all in the second case.

For the time-varying wind field, the aircraft still moves relative to the air mass but now the LEScoordinate system coincides with the ground coordinate system. In this case, the , taking into account the movement of the air mass present in the LES. Adjustment of the sampling positions is achieved by taking into account the motion of the aircraft due to the air mass motion has to be considered during each time step to yield the correct actual measurement positions in the LES. Therefore To this end, the aircraft heading and ground speed are determined iteratively for subsequent time steps, also based on the triangle of velocities (Appendix A1). The effect can again of this procedure can be imagined with an aircraft flying at 65 m s^{-1} into 65 m s^{-1} headwind. Relative to the LES coordinate system this aircraft will stay in the same place. It will thereby sample at the same geographic coordinate in the LES at all times. Summarizing, the

The correct simulation of the aircraft motion is important as it changes the sampling of the wind field data by the lidar system . In the is altered. In this study, wind speeds reach up to 25% of the aircraft speed relative to air (IAS). Thereby, the sampling distance between measurements is changed by a factor of up to 0.4 between flying up- or downwind (50 m s^{-1} vs. 80 m s^{-1} ground speed), presenting a non-negligible effect. In the ADLS, the aircraft track development is dependent on the wind field as described above. Additionally, pitch, yaw and roll moments can be added to the aircraft position to simulate the effect of aircraft accelerations on the measurement process. These are termed aircraft movement (see Appendix Tab. **??** for an overview of general terminology) in this study and superimposed artificially and thereby independent of the track development. This independence is not realistic but deemed sufficient, as aircraft position correction maneuvers should generally not alter the track development (and thereby sampling) significantly, as they will cancel out over short periods of time. Thereby, the effect of aircraft movement due to flight maneuvers can be emulated as well.

2.2.2 Scanner

The scanner movement is simulated with freely selectable scanning geometries and includes an option to correct for the aircraft movement. Subsequent scan positions are calculated in an aircraft relative coordinate system by specifying the position which should be reached by the scanner, a time needed for the scan movement and a scan mode. Five scan modes are available. The scanner can exhibit stare-mode, scan with constant azimuth, constant elevation, along the shortest possible distance on a sphere between two positions or focus on specific positions in the ground reference system (thereby also correcting for aircraft motion). When the aircraft movement correction is enabled, scanner positions are calculated in an earth relative coordinate system and then transferred back to the respective aircraft relative coordinate system positions corrected for aircraft movement. The standard scan pattern (ELE, ROT) used for comparison is based on the scan geometry of existing systems (Weissmann et al., 2005; De Wekker et al., 2012). It consist of a 20-second full circle scan time (ROT, scan speed 18° s^{-1}) at 60° elevation (ELE, 30° off-nadir). For optimization, both elevation and speed of the scanner are varied (Sec. **??**).

Two scanner system simulations are performed as a part of In this study, an 'ideal' system without noise and a 'noisy' system with emulated scanner pointing direction noise. For the first setup, in order to purely focus on the effect of atmospheric wind profile retrieval error due to atmospheric flow inhomogeneities, an ideal scanner system is assumed without any beam pointing

inaccuracy (so-called 'ideal' system in the following). ~~For the second setup, in order to simulate an imperfect measurement system, a random Gaussian inaccuracy can be added to the scanner azimuth and elevation angles before being used for wind profile retrieval (details are explained in Sec. ??).~~

2.2.3 Lidar

~~In this study, the~~ The simulated lidar is based on the specifications of a Lockheed Martin Coherent Technologies WTX WindTracer system. Lidar systems with similar characteristics are often used in airborne Doppler lidar studies (Weissmann et al., 2005; De Wekker et al., 2012; Chouza et al., 2015; Witschas et al., 2017; Zhang et al., 2018). The lidar beam is emulated in accordance with Stawiarski et al. (2013) and their Gaussian range gate weighting function based on a pulse width of $\sigma_\tau = 3 \cdot 10^{-7}$ s is applied. The same cut-off value, 20 % of the maximum value of the weighting function, is chosen for calculating the effective length of the range gates for the averaging process. Variable range gate lengths and spacings can be specified, in accordance with the standard WTX settings a range gate length of ~~$l_{RG} = 72$ m with a spacing of $\Delta x_{RGC} = 72$ m~~$\Delta p = 72$ m with a non-overlapping spacing of $\Delta R_0 = 72$ m (the distance between subsequent range gate centers), starting at a distance of 400 m from the lidar is used for this study. ~~For wind profiling, the standard~~ The lidar measurement frequency ~~(FME) used in this study is 1 Hz~~can be varied in the range $1 - 10$ Hz. The details of how the volume scanned by the laser beam is constructed for averaging during the measurement process are explained in the next section.

Again, ~~an 'ideal' and a noisy system are defined but this time for the laser. Two laser system simulations are performed as a part of this study, an 'ideal' system without noise and a 'noisy' system with emulated measurement noise. For the first setup,~~ in order to purely focus on the ~~effect of atmospheric~~ wind profile retrieval error due to atmospheric flow inhomogeneities, an ideal laser system is assumed without any random radial velocity fluctuations or other system errors.

~~However, for coherent systems, random radial velocity fluctuations (uncorrelated noise) due to detector noise, speckle effects and turbulence within the measurement volume can occur as part of the measurement process (Frehlich, 1997, 2001). Random radial velocity fluctuations in the high-signal return regime can be approximated by a Gaussian distribution with zero mean and standard deviation σv_D. Therefore, in the ADLS, a Gaussian noise with standard deviation σv_D can be added to the simulated measured radial velocities. This inclusion allows for investigation of the effect of random radial velocity fluctuations on the quality of wind profiles (see Sec. ??). For the lidar system simulated here, based on previous studies, the random noise is expected to be below 0.25 m s^{-1} for 1 Hz measurement frequency, 0.5 m s^{-1} for 5 Hz measurement frequency and 1 m s^{-1} for 10 Hz measurement frequency (all at 750 Hz pulse repetition frequency) (Frehlich, 2001; Stawiarski et al., 2013).~~

2.3 Measurement procedure - motion combination and correction

For the measurement, the aircraft and atmospheric motion vectors need to be combined to give the measured radial Doppler ~~velocity~~velocities. Additionally, in order to obtain the atmospheric motion vector from the LES, the ~~range gate positions~~ locations of the lidar range gates need to be calculated~~from~~. Both operations are conducted based on the state of the system components, meaning the current aircraft orientation and motion, as well as scanner position and lidar setting. Conveniently, this calculation is achieved by defining two separate coordinate systems (fig. 2), following the theory outlined in

~~Lee et al. (1994)~~ Lenschow (1972) and Leon and Vali (1998). The LES wind fields are based in earthbound coordinate system (superscript E) oriented east-north-up (ENU). The aircraft coordinate system (superscript A) is oriented along aircraft-right wing-down (ARD). The scanner is emulated and the measurement is performed in the aircraft coordinate system. To transfer between the two systems, coordinate transformations are required, the details are given in Appendix A2.

~~Schematic depiction of a possible airborne Doppler lidar system. The sketch is based on an upcoming system for the Dornier 128-6 aircraft of the TU Brunswick (D-IBUF, Corsmeier et al. (2001)). The lidar is inside the aircraft pointing outward, with a scanner mounted on the side of the fuselage directing the beam in the atmosphere. Displayed are the aircraft, scan cone surfaces as well as the coordinate systems and vectors used in eq. 3, for additional information see Sec. 2.3.~~

The measured Doppler velocity v_D can be calculated through projecting the velocity vectors onto the lidar beam vector. The radial Doppler velocity measured by the lidar has a contribution from the lidar motion vector $\boldsymbol{v}_L^A$ and the atmospheric motion vector $\boldsymbol{v}_p^A$~~, both~~. Both are projected onto the beam direction vector $\boldsymbol{b}$, which is defined as a unit vector:

$$v_D = \boldsymbol{b}^A \cdot \left(-\boldsymbol{v}_p^A + \boldsymbol{v}_L^A \right). \tag{1}$$

The lidar motion can be split into two contributions, the aircraft motion $\boldsymbol{v}_a^A$ and an aircraft movement moment arm due to aircraft rotation ω with the moment arm $\boldsymbol{r}$ (specifying the distance between the ~~aircraft center of gravity and the lidar position~~), position of the inertial navigation unit in the aircraft and the position of the final mirror turning the lidar beam),

$$v_D = \boldsymbol{b}^A \cdot \left(-\boldsymbol{v}_p^A \pm -\boldsymbol{v}_a^A \pm -\boldsymbol{\omega}^A \times \boldsymbol{r}^A \right). \tag{2}$$

The above assumes that motion towards the lidar is negative, whereas motion away from the lidar is positive. Depending on the lidar beam direction, an airborne Doppler lidar system measures foremost the aircraft speed as it presents a vector with very large magnitude compared to the atmospheric motion vector and the aircraft movement moment arm. The atmospheric motion vector needs to be transformed into the aircraft coordinate system where the measurement is performed using the rotation matrix $\mathbb{T}'^{AE}$ (see App. A2). The same is true for the aircraft motion vector which is originally calculated with respect to the LES coordinate system. Consequently, the measurement is achieved according to the following formula:

$$v_D = \boldsymbol{b}^A \cdot \left(-\mathbb{T}'^{AE} \boldsymbol{v}_p^E \pm -\mathbb{T}'^{AE} \boldsymbol{v}_a^E \pm -\boldsymbol{\omega}^A \times \boldsymbol{r}^A \right). \tag{3}$$

Appendix A3 details how the weighted and averaged particle velocity $\boldsymbol{v}_p^E$ is obtained from the LES considering the range gate position and motion during the measurement process.

Before the application of the retrieval, the contribution of the aircraft motion and aircraft movement moment arm to the measured radial velocity is then removed again~~using Eq. 3 to restore the original atmospheric contribution~~.

$$v_{COR} = v_D + \boldsymbol{b}^A \cdot \left(\mathbb{T}'^{AE} \boldsymbol{v}_a^E + \boldsymbol{\omega}^A \times \boldsymbol{r}^A \right). \tag{4}$$

This calculation isolates the motion corrected velocity v_{COR}, which is the atmospheric wind contribution to the measured radial velocity (the projection of the atmospheric wind vector on the beam vector). The ADLS makes it possible to add inaccuracies

in any of the components relevant in the measurement or motion correction process. However, please note that throughout this study no system inaccuracies are introduced in the measurement or motion correction process ~~for the ideal system simulation~~ in order to focus on problems in wind profiling due to atmospheric inhomogeneity. In addition, it is assumed that for the real system the optical alignment and beam direction can be reliably quality controlled using a beam calibration procedure based on ground returns. The theory and procedure are outlined in Haimov and Rodi (2013). Consequently, for the ideal ADLS system, the motion-corrected Doppler velocity due to the particle velocity is equal to the particle velocity projection itself~~again~~. Therefore, compared to a ground-based system, only the measurement geometry is altered due to the moving system. The above discussed rotations and transformations do not influence the accuracy of the motion correction. As a result, ~~for the ideal system simulation,~~ all wind profiling errors discussed in the next sections stem purely from the inhomogeneous atmospheric ~~conditions~~flow conditions due to boundary layer turbulence. In a homogeneous wind field, the retrieved wind profile is exact.

~~In real measurements, system noise is expected to occur. Therefore a worst-case scenario simulation with emulated system noise is conducted besides the ideal system simulation to estimate the impact of system noise (see Sec. 2.2.2 and 2.2.3). For the noisy system simulation, two sources of system noise are added, both on the measured radial velocity v_D as well as the beam pointing direction $\boldsymbol{b}^A$. Emulating Doppler lidar noise, a Gaussian noise contribution with measurement frequency dependent σv_D is added to the measured radial velocity. This addition is done before removal of the lidar motion contribution. Emulating an imperfectly known measurement system state, a Gaussian noise contribution is also added to the beam pointing direction, both on azimuth and elevation angles, thereby introducing wind profiling errors due to imperfect motion correction.~~

2.4 Retrieval - nadir as an example application

~~Nadir transects through LES for 10 m s^{-1} background wind case with superimposed aircraft oscillation movement at 10 Hz measurement frequency. a) LES vertical wind along a crosswind transect. b) ADLS motion corrected velocity measurement along a crosswind transect. c) LES vertical wind along an upwind transect. d) ADLS motion corrected velocity measurement along an upwind transect. Scanner movement correction is disabled for the first half of the transects, whereas it is enabled for the second half.~~

An illustrative example of the ADLS capability is given by simulated nadir retrievals. In fig. 3 the simulated vertical wind measurement is compared to the LES input along two transects for an ideal system with a lidar measurement frequency of ~~FME = 10 Hz.~~ 10 Hz. The aircraft is flying at an altitude of 1700 m with an IAS of 65 m s^{-1} and a super-imposed 8 s, 5° roll oscillation, combined with a 15 s, 3° pitch oscillation.

The first transect, a crosswind flight (fig. 3 a,b), reveals the spanwise boundary turbulence structure for the 10 m s^{-1} background wind case. Along stream organization of turbulence into rolls occurs and gravity waves are present above the boundary layer. The ADLS results show that the simulated lidar is able to capture turbulent structures in the boundary layer, although smoothing occurs for the finest scales. Vertically, the resolution of the measurement is defined by the range gate length, whereas horizontally it is defined through the combination of aircraft speed and measurement frequency. Another noticeable effect in the measurement is the movement correction capability of the scanner, which is disabled for the first half of the transect and enabled for the second half. During the first half, the lidar beam is not pointing exactly nadir due to the superimposed, artificial

aircraft ~~movement~~roll and pitch oscillation. This deviation causes a portion of the horizontal wind to be projected into the measurement. For the crosswind case, this effect is caused by the (rather strong) roll oscillations of the aircraft. Consequently, the measured vertical wind shows some additional superimposed structures compared with the LES for the first half of the transect. This contribution ~~from the horizontal wind could~~ of the mean horizontal wind can be removed, as a second step, if the horizontal wind profile is known (Chouza et al., 2016b). However, variability in the mean horizontal wind profile will still manifest in the vertical wind measurement. The roll and pitch oscillations also result in a distorted curtain location, which is not directly below the aircraft anymore (illustrated in fig. 1). The winding curtain results in a non-equidistant sampling of turbulence, complicating analysis further. For the second half the scanner movement compensation is enabled. Consequently, the beam is pointing nadir at all times and no horizontal wind contribution is visible in the measurements.

The second transect, an upwind flight (fig. 3 c,d), reveals the streamwise boundary turbulence structure for the $10\ \mathrm{m\ s^{-1}}$ background wind case. The flight is conducted in an updraft region of the along-stream organized boundary layer convection, therefore positive vertical wind speeds dominate. In this case, the superimposed pitch movement contaminates the retrieved vertical wind measurement with a contribution from the horizontal wind if the scanner movement correction is disabled. ~~In real measurements, the degree of this beam offset due to pitch oscillations is detectable in the ground return velocity which is non-zero in this case.~~ Thereby, nadir-pointing measurements provide a good opportunity to check the accuracy of the scanner movement compensation and beam direction accuracy using the ~~vertical~~ motion corrected wind measurement and ground return velocity. Should the movement compensation or pointing accuracy not be satisfactory, it is still possible to calibrate them using the ground return velocity (Haimov and Rodi, 2013). Using this information, the amount of horizontal wind mapped into the vertical wind can ~~again~~ be removed in a further post-processing step if the vertical profile of the horizontal wind is known. In order for this method to yield reliable results, an accurate horizontal wind speed estimation is necessary. For this reason, the reliability of wind profiling measurements is the focus of this study.

3 Airborne Doppler lidar wind profiling in inhomogeneous flow conditions

~~Doppler wind profiling presents an inverse problem. The system state which produced the observations is inferred from a set of observations and an underlying model. The observations are the measured Doppler velocities. The model is given by the beam geometry, however, limitations arise because only a limited number of measurements , amount of beam directions and viewing geometries are possible. Choosing the beam geometry is an important step that can be altered depending on the desired measurement. On the one hand, it is desirable to maximize the spread among the beam pointing directions to reduce collinearity between the observations. On the other hand, the measurement volume should be minimized in order to have a small and localized footprint of the retrieved wind profile.~~ In this section, the ADLS is used to investigate the error associated with airborne Doppler lidar wind profiling in inhomogeneous flow conditions. Compared to in-situ measurements commonly used for wind profiling accuracy evaluation, the ADLS offers more insight, as the representation error due to the different sampling volumes between lidar and in-situ measurements does not exist. From real measurements, it is difficult to determine whether the observed wind profile differences are due to the difference in location and/or sampling volume (lidar scanned

volume vs. instrumented tower) or due to violated model assumptions (homogeneous wind field) used for retrieval. The ADLS can overcome the co-location problem, as the wind field contributions used to create the measurements are known exactly based on the LES input (fig. 4).

~~The system state is the original wind field on which the observations are performed, which is unknown in real measurements but known for the ADLS. To enable a robust estimation the wind field is usually assumed to be homogeneous, while knowing that this assumption is rarely fulfilled, especially in the boundary layer. Consequently, errors are introduced in the retrieved wind profile compared to the original wind field. This section describes the theory of the inversion process as wellas errors associated with it (for a description of the terminology used for different errors see Tab. ??).~~ To this end, in this section the wind profiling retrieval and associated error metrics are explained. Then, the sampling strategy and independence of the sampled wind profiles are discussed. After the evaluation of wind profiling error in inhomogeneous flow, the section concludes with an investigation of commonly used quality filtering criteria.

3.1 ~~Wind profiling theory~~

3.1 Wind profiling theory and associated error metrics

~~Throughout~~ The theory and problems associated with airborne Doppler wind profiling are very similar between lidar and radar systems. Therefore, no distinction is made between the studies using either of the two instruments in the following, unless necessary due to explicit differences. Similar to ground-based wind profiling, airborne wind profiling is usually conducted by scanning in conical scans along the flight path. Retrieval of the mean wind vector can be achieved through inversion of the beam matrix, yielding a least-squares solution to the problem (Leon and Vali, 1998). De Wekker et al. (2012) and Tian et al. (2015) apply the original Velocity-Azimuth-Display (VAD) analysis directly, while neglecting the altered beam geometry due to aircraft movement. Other methods at higher computational cost exist as well. Guimond et al. (2014) discuss a variational approach which can improve the traditional solution by adding anelastic mass continuity constraints on the estimated solution. Accumulation of the Doppler spectra can be conducted for Doppler lidar and has been shown to extend the retrieval limits in clear air conditions with little return signal (Weissmann et al., 2005; Witschas et al., 2017).

In this study, ~~the~~ a standard inversion based approach outlined by Leon and Vali (1998) for airborne wind profile retrieval ~~(termed 'AVAD')~~ is followed. In this method, multiple radial velocity measurements ~~v_{D_n}~~ under different beam pointing directions ~~$\boldsymbol{b}_n$~~ are sampled from an atmospheric wind field~~with mean wind vector $\boldsymbol{v}_p$:~~

$$\begin{bmatrix} v_{D_1} \\ v_{D_2} \\ v_{D_3} \\ \vdots \\ v_{D_n} \end{bmatrix} = \begin{bmatrix} b_{x_1} & b_{y_1} & b_{z_1} \\ b_{x_2} & b_{y_1} & b_{z_1} \\ b_{x_3} & b_{y_1} & b_{z_1} \\ \vdots & \vdots & \vdots \\ b_{x_n} & b_{y_n} & b_{z_n} \end{bmatrix} \begin{bmatrix} v_{p_x} \\ v_{p_y} \\ v_{p_z} \end{bmatrix}.$$

~~The beam directions from multiple measurements make up the model matrix $\mathbb{G}$. Consequently, the relation can be expressed in the following way:~~

$v_D = \mathbb{G}\boldsymbol{v}_p.$

~~Knowing v_D and $\mathbb{G}$, the~~ , e.g. by scanning (see fig. 1 for an illustration of the scan trajectory). The inverse problem is then solved by calculating the inverse $\mathbb{G}^{-g}$ of the beam pointing ~~model $\mathbb{G}$, in order to obtain an~~ matrix (for the associated formulas and details see Appendix A4). Thereby, a least-squares (LSQ) estimate of the wind vector ~~$\boldsymbol{v}_{ret}$~~ , which was responsible for the observations~~:~~

$\boldsymbol{v}_{ret} = \mathbb{G}^{-g} v_D.$

~~The general inverse $\mathbb{G}^{-g}$ of the overdetermined least-squares problem can be calculated as (Menke, 2012):~~

$\mathbb{G}^{-g} = [\mathbb{G}^T\mathbb{G}]^{-1}\mathbb{G}^T.$

~~Equation (A17) presents a least-squares solution to the problem. Instead of calculating the general inverse in the above way, in this study~~, is obtained (see fig. 4 for an illustration of the LSQ-fit procedure). In this study, a singular value decomposition (SVD) is performed ~~, which~~ to calculate the inverse. The SVD-solution yields benefits compared to the ~~above direct solution(Boccippio, 1995):~~

$\mathbb{G} = \mathbb{U}\mathbb{S}\mathbb{W}^T,$

$\mathbb{G}^{-g} = \mathbb{W}_p\mathbb{S}_p^{-1}\mathbb{U}_p^T.$

~~Assessing the reliability of the parameters retrieved through the inversion process is a common problem in inverse theory~~direct solution, as a number of reliability control parameters become available (Boccippio, 1995).

~~The first often used metric deals with the stability of the inversion process. For the SVD the stability is expressed through the condition number (CN). The CN provides a measure of the spread of the model space, thereby diagnosing collinearity (Boccippio, 1995; Leon and Vali, 1998; Shenghui et al., 2014; Päschke et al., 2015; Wang et al., 2015). The condition number is~~ The retrieval error is defined as the ~~ratio of the largest to the smallest diagonal entry of $\mathbb{S}$, which are the singular values of $\mathbb{G}$:~~

$$CN = \frac{max(\lambda_S)}{min(\lambda_S)}.$$

~~A high CN indicates high collinearity in the model geometry. Consequently, the real system state is not well explored in at least one direction and as a result the inferred system state is prone to a large error amplification (Boccippio, 1995). The CN is often used as a quality filtering criteria to exclude measurements where the spread of the beam pointing directions is not sufficient to explore the wind field adequately (Boccippio, 1995; Päschke et al., 2015; Wang et al., 2015).~~

Another measure commonly used to detect violations of the homogeneity assumption is the coefficient of determination (R^2) which can be obtained from the LSQ-solution. Using the estimated wind vector, an average radial velocity (LSQ-fit) is constructed by projecting it into radial velocities using the model geometry:

$$v_D^{ret} = \mathbb{G}\boldsymbol{v}^{ret}.$$

This average radial velocity differs from the measured radial velocities due to wind field inhomogeneities smaller than the measurement volume as well as model misspecification and measurement system noise (only in the noisy system case).

Using the average radial velocity, difference between the true wind speed V_m^T, calculated from the input LES wind field as discussed below, and the R^2 is defined as

$$R^2 = 1 - \frac{\sum_i (v_{Di} - v_{Di}^{ret})^2}{\sum_i (v_{Di} - \sum_i v_{Di})^2}.$$

Wind field inhomogeneities smaller than the scan volume size decrease the R^2. Therefore, it is commonly assumed that the R^2 captures the degree of violation of the homogeneity assumption. Often used filtering criteria are $R^2 > 0.8$ or $R^2 > 0.95$ for ground-based wind profiling (Wang et al., 2015; Päschke et al., 2015).

The reliability of both measures for ADLS retrieved wind speed V_m^R, obtained from wind profiling (see fig. 5 for an example on the obtained wind profiles and associated errors):

$$\Delta V_m = V_m^R - V_m^T. \tag{5}$$

In this study, the wind profile quality assessment in airborne Doppler lidar scanning is investigated as a part of this study(Sec. 3.3).

4 Evaluation of error in wind profiling

Using the ADLS, the accuracy of the retrieved wind profiles can be analyzed under inhomogeneous wind field conditionsspeed V_m is calculated from all three components, $V_m = \sqrt{u^2 + v^2 + w^2}$. For the quantitative analysis , four metrics are employed, consisting of the root-mean-squared error (RMSE), the relative root-mean-squared error (REL), the number of available wind profile points (N) and the retrieval bias. The RMSE is given as:

$$\mathrm{RMSE} = \sqrt{\left[\frac{\Sigma_i^N (V_{m\,i}^T - V_{m\,i}^R)^2}{N}\right]}.$$

Here, V_m^T is the true wind speed based on the input LES wind speeds, whereas V_m^R is the ADLS retrieved wind speed from wind profiling. The retrieval error is $\Delta V_m = V_m^R - V_m^T$. In this study the wind speed V_m is calculated from all three components,

$$V_m = \sqrt{u^2 + v^2 + w^2}.$$

N is the number of wind profile points fulfilling the quality filtering criteria. The relative root-mean-squared error (REL) is used in accordance with Guimond et al. (2014):

$$\mathrm{REL} = \sqrt{\left[\frac{\Sigma_i^N (V_{m\,i}^T - V_{m\,i}^R)^2}{\Sigma_i^N (V_{m\,i}^T)^2}\right]}.$$

The REL can provide additional information to the RMSE as its magnitude is independent of the mean wind speed, thereby enabling comparisons between the different background wind cases, especially for higher wind speeds. The RMSE is driven by two factors, the variance and bias of the retrieval errors. Therefore we also report the bias of the retrieval as its average mean deviation,

$$\mathrm{bias} = \frac{1}{N}\sum_{i=1}^{N}(V_{m\,i}^R - V_{m\,i}^T).$$

, defined in App. A5.

To compare the simulated measurement, a retrieved wind speed, an LES based model truth has to be defined. Here, the simulator offers more analysis insight than in-situ measurements, where a representation error exists due to the different sampling volumes between lidar and in-situ measurements, e.g. instrumented towers. From real measurements, it is difficult to determine whether the observed wind profile differences are due to the difference in location and/or sampling volume (lidar scanned volume vs. instrumented tower) or due to violated model assumptions (homogeneous wind field). The ADLS can overcome the co-location problem, as the wind field contributions used to create the measurements are known exactly. Therefore, this This study follows the approach described in Wainwright et al. (2014) to define the model truth. In this, the model truth is the average over the points in the sampling volume touched by the lidar beam after weighting by the lidar weighting function (Sec. 2.2.3). In this study, the vector average is used as the wind speed averaging method. Vector averaging is more representative of the lidar measured wind speed than scalar averaging, as the lidar averages the wind field over a large area rather than summing up individual contributions without respect for directional change. The method used makes a difference especially for lower wind speed cases (see illustrated in fig. 5, discussed in-depth in Sec. 3.3). The advantage of the direct, equal-volume-based LES-lidar comparison, as noted by Wainwright et al. (2014), is that differences which occur between the simulated measurement and model truth are exclusively traceable to the wind field inhomogeneities. Therefore, optimization of the measurement system setup with respect to the impact of wind field inhomogeneities is possible. When an ideal system is assumed, with no measurement system inaccuracy and no co-location problem, the results are also useful as they present a lower bound to the observable Doppler lidar error. Under the conditions investigated in this study, a A lower error cannot be observed from real in-situ comparisons (assuming equal statistical basis and comparable conditionsatmospheric conditions, measurement system setup and retrieval strategy), as the error given in this study specifies the error inherent inherent error due to the Doppler lidar viewing geometry and retrieval settings. Please note that despite investigating an overall random error, the error between vertically adjacent wind profile points is not necessarily random. This lack of statistical independence is because turbulent structures are correlated between adjacent wind profile points for an individual wind profile, making the profile appear smooth despite overall random error being present.

The simulation results presented in the following are not directly transferable to ~~the validation of Doppler lidar wind profile measurements through other measurement systems, as to-date all available measurement systems are unable~~ real-world wind profiling comparisons. Today, no measurement system is able to trace the exact volume touched by the lidar beam ~~. An in-situ~~ and thereby measure the input truth, as is done in the ADLS. Therefore, a real world comparison should result in larger ~~deviations~~ error levels than what is presented here, due to additional co-locations problems, and due to the different wind vector measurement principles (Bradley et al., 2012). It seems worthwhile to extend the analysis in this direction at a later point, as the underlying statistical foundation exists already (Frehlich, 2001). The topic is not addressed here as it is beyond the scope of this work~~and the inherent problems caused by the measurement geometry and viewing geometry are voluminous by themselves.~~ .

~~To avoid correlation among the analyzed data only one standard setting for system setup and retrieval setting is used for general analysis (see also Sec. 2 and fig. B1). Further, only wind profiles within the boundary layer and entrainment zone are compared as this study focuses on retrieval error under inhomogeneous conditions. Therefore, the maximum wind profile altitude is set to 800 m (see also fig. B7). Only the two wind profilesclosest to the center, contained fully within the LES, are analyzed to ensure full measurement coverage and avoid an artificially increased CN due to the lidar beam exiting the LES volume. The wind profiling start point is chosen randomly for each transect within one profiling length. The random start point prevents a systematic influence of the transect begin location on measurement quality for the different flight directions. It also contributes to a further de-correlation of analyzed wind profiles.~~

3.1 System setup, retrieval strategy and sampling procedure

The retrieved wind profiles are sampled from the LES domain in temporal and spatial proximity. Independence of the retrieved wind profiles is important to ensure a sufficiently large and independent dataset for robust analysis with reliable values. Therefore, correlation is minimized through an adequate measurement system setup, retrieval strategy and sampling procedure explained in the following. However, it should be noted that any potential correlation in wind profiling error does not affect the reported errors levels systematically (RMSE, REL, bias). Correlation reduces the number of effectively available wind profiles, and thereby influences the uncertainty with which the error levels are estimated, but it does not impact the error levels themselves. Nevertheless, in order to achieve a sufficiently large statistical sample size, correlation is minimized through an adequate measurement system setup, retrieval strategy and sampling procedure. To this end, all profiles are sampled from non-overlapping measurement domains in time and space (termed checkerboard approach in the following), illustrated in figure 6 and explained in the following.

The measurement system setup is simulated based on an intended upcoming airborne Doppler lidar system for boundary layer research. The simulator settings are adjusted accordingly, an overview of parameter settings is given in fig. B1. The aircraft characteristics are based on an unpressurized, medium-range turboprop aircraft, flying above the boundary layer at measurement speed under visual flight rules. Therefore, the aircraft is flying at an IAS of 65 m s^{-1} and the aircraft altitude is set to 1100 m. Combined with the chosen scan elevation angle, this setting gives equal along- and across-track retrieval volume averaging distances.

The ~~standard system setup (STP) used for analysis~~ scan pattern used in this study is based on the ~~four different wind speed cases and 16 flight directions through the LES. For this setting, only data from the 20 s scan pattern~~ scan geometry of existing systems (Weissmann et al., 2005; De Wekker et al., 2012; Bucci et al., 2018). It consists of a 20-second full circle scan time (scan speed $18° \, s^{-1}$) at 60° elevation (~~ROT 20 s, ELE 60°), 1 Hz measurement frequency (FME~~ 30° off-nadir). The scan trajectory starting point is chosen randomly at each time step in order to ensure sampling of different locations by the lidar beam. Additionally, the scan rotation direction is reversed for subsequent transects through the LES.

The lidar measurement frequency is set to 1 Hz~~) and the 1700 m flight altitude (AAL 1700 m)is analyzed. The standard~~, with pulse width and range gate settings according to Sec. 2.2.3. Changing the measurement frequency to 10 Hz does not alter the results significantly, as the additional measurements are strongly autocorrelated (see fig. 4). Consequently, the greater number of measurements does not provide new information to the retrieval, nor does it result in a better fulfilled homogeneity assumption.

The wind profile retrieval ~~setting (RET)~~ consists of u,v,w component retrieval ~~(VAR u,v,w)~~ using a volume-based profile separation~~(SEP Volume)~~. Retrievals are obtained from along-track averaging over $X_a = 1300$ m ~~(AVG 1300 m) at an~~ and the same across-track averaging distance of ~~$X_c = 1963$ m (corresponding~~ $X_c = 1300$ m. The along-track averaging distance is based on the distance covered by the aircraft during one full scan rotation. The across-track averaging distance corresponds to the maximum across-track distance covered by the lidar beam at 60° elevation and a flight altitude of ~~1700 m)~~1100 m above ground. The vertical wind profile resolution is chosen as 62 m~~(LEV 62 m)~~, yielding one range gate within every vertical layer. ~~In Sec. ?? each of~~ Only wind profiles within the boundary layer and entrainment zone are compared as this study focuses on retrieval error under inhomogeneous, turbulent conditions. Therefore, the maximum wind profile altitude is set to 800 m, preventing the impact of larger scale features such as gravity waves with longer de-correlation scales.

Three flight trajectories traverse the LES domain in parallel from south to north at a horizontal distance of 1750 m to each other, in a crosswind flight direction. Each of the three trajectories is repeated 28 times at 1-minute temporal spacing, giving a total number of 84 actual transects. To avoid correlation among the analyzed profiles from different transects, all profiles are sampled from non-overlapping measurement domains in time and space, illustrated in figure 6.

To this end, only one wind profile is retrieved for each transect. The retrieval volumes for the parallel transects at each time step are arranged in a checkerboard pattern to ensure maximum spatial independence. The distance between spatially neighboring transects and measurements is larger than the integral scale of turbulence in the LES, generating statistical independence.

The checkerboard pattern is shifted by X_a along-flight track for subsequent transects, thereby also ensuring temporal independence of the retrieved profiles. After three time steps, the retrieval volume shift procedure is repeated. Thereby, the same location is sampled only every three minutes. The decorrelation of the retrieved profiles is aided by crosswind advection and the ~~parameters are then varied individually to evaluate the effect of different system setups and profile retrieval settings.~~ pencil-beam nature of the lidar beam. While large scale structures in the flow, e.g. due to coherent turbulence, may persist over some time and distance, small scale variation is more rapid. The small scale variation is expressed by the short integral length scales of the LES-flow at $50 - 500$ m, denoting the spatial distance over which turbulence is statistically independent from

the previous sample. A random offset within one retrieval volume length is added to all retrieval volume start points within one retrieval volume shift procedure. The random start point prevents a systematic influence of the transect begin location on measurement quality for the different flight directions.

~~The described system setupand retrieval setting allow~~ An analysis of the spatio-temporal correlation of the obtained wind profile retrieval errors ΔV_w shows that the obtained errors are independent to a high degree for all background wind speed cases, time steps and retrieval altitudes (fig. B2, B3, B4, B5). The autocorrelation functions allow for evaluation of both, the spatial as well as the temporal correlation of profiling error, at the same time for the chosen checkerboard approach. On the one hand, the autocorrelation values at lags $1, 2; 4, 5; 7, 8; ...$ reveal the spatial correlation of wind profiling error between neighboring retrieval volumes. On the other hand, the autocorrelation values at lag $3, 6, 9, ...$ reveal the temporal correlation of retrieval error for the first retrieval volume (as the retrieval volume shift procedure is repeated after a cycle of three volumes, the same volume is sample again at these lags). Taking into account the small number of points used to calculate the autocorrelation functions, they do not display concerning systematic structure (e.g. repetitive with cycle duration 2 or 4, or similar between trajectory or altitudes). The marginal structure present beyond the noise resulting from the small sample size can be explained by large scale organization of turbulence on scales similar to the spacing of the retrieval volumes. Examples of weak structures are the alternating correlation coefficients for the first lags of trajectory 1 in the 0 m s^{-1} background wind case at low altitudes (fig. B2), as well as the slightly negative correlation for lag 1 in the 15 m s^{-1} background wind case at low altitudes (fig. B5. As the associated correlation coefficients remain small (< 0.5 in almost all cases), the number of effectively available wind profiles is not affected strongly by these effects.

Please note that despite investigating an overall independent error (between spatially and temporal different profiles), the error of vertically adjacent wind profile points is not random for an individual profile (see fig. 5). The error between neighboring vertical profile layers is not random because turbulent structures are correlated between vertically adjacent layers. This makes an individual wind profile appear smooth, despite overall random error being present. The magnitude of the random error only becomes noticeable when looking at profiles sampled at different locations or times (this disguise is likely one of the reasons why the error due wind field inhomogeneities has received little attention in literature so far). The vertical correlation effect is present in both the ADLS and real measurements with equal magnitude.

3.2 Evaluation of error in wind profiling

The previous section showed that a statistical analysis of error in wind profiling is possible using the given system setup, retrieval strategy and sampling procedure.

The ADLS simulation allows for retrieval of ~~38400 wind profile points~~336 individual wind profiles giving 4032 wind profile points for all altitudes (12 wind profile points per wind profile at the different retrieval altitudes). These consist of ~~9600 wind profile points~~ 84 wind profiles (1008 wind profile points) for each background wind case, yielding a sufficient statistical basis for evaluation. ~~For quality~~ The 336 wind profiles (4032 wind profile points) are more than what is typically available for comparison in real world measurements, as co-located validation measurements are very difficult and costly to conduct. For example, 33 wind profiles (740 wind profile points) are compared to dropsonde data in Weissmann et al. (2005),

approx. 10 wind profiles to a ground-based wind profiler in De Wekker et al. (2012), a single wind profile to dropsonde data in Kavaya et al. (2014) and approx. 49 wind profiles (2056 wind profile points) to dropsonde data in Bucci et al. (2018), with each of them mentioning the importance of the spatially differing sampling volumes and associated problems.

The wind profiling quality of the standard system setup and retrieval settings in the investigated inhomogeneous flow conditions can be evaluated from fig. 7. As discussed above, the wind profile retrieval is made assuming an ideal measurement system, thereby all deviations are directly traceable to AVAD assumption violation due to wind field inhomogeneities. For quality filtering purposes a minimum coefficient of determination of $R^2_{min} = 0.90$ and a maximum condition number $\mathrm{CN}_{max} < 9$ are applied. Similar values have been recommended and used in the past by Boccippio (1995); Päschke et al. (2015); Wang et al. (2015). The effect of quality filtering on retrieval quality is investigated in detail in Sec. 3.3.

~~Comparison of LES truth and ADLS retrieved wind speed for an ideal measurement system using STP, RET with quality filtering criteria applied. Color-coded are all measurements which pass quality filtering, gray crosses are the ones which do not.~~

~~The quality of the standard system setup and retrieval settings can be evaluated from fig. 7. In this case the wind profile retrieval is made assuming an ideal measurement system, thereby all deviations are directly traceable to wind field inhomogeneities.~~

A total number of ~~28778~~ 3006 (of the 4032 retrieved) wind profile points pass quality filtering and the results provide a number of interesting observations. Firstly, the wind speed retrieval is unbiased provided that an appropriate ~~system setup~~ retrieval strategy is used and appropriate quality filtering criteria are applied (discussed in Sec. 3.3). Overall retrieval quality is high with $R^2 = 0.99$ and ~~RMSE = 0.26 m s^{-1}~~ RMSE $= 0.36$ m s^{-1}. Yet, even with strict quality filtering applied, deviations ~~up to~~ above 1 m s^{-1} can occur. The LES cases with higher wind speeds show higher wind variability and increased absolute wind speed retrieval error. With decreasing wind speed (separately for each of the cases), associated with measurements in the boundary layer under turbulent conditions, deviations increase but remain unbiased. Interestingly, the ~~9600~~ 1008 retrieved wind profile points for the 0 m s^{-1} background wind case are completely excluded by the quality filtering criteria (as well as ~~22~~ 18 wind profile points from the 5 m s^{-1} background wind case). They also show much higher wind speed retrieval scatter and can introduce a positive bias if inappropriate quality filtering criteria are applied, although the individual retrieved components themselves are unbiased (fig. B6). The reasons behind this are discussed in Sec. 3.3, where the effect of quality filtering ~~on retrieval quality~~ is thoroughly investigated.

The vertical distribution of wind profile error mirrors that of the responsible boundary layer turbulence (fig. B7). Errors are largest in the middle of the boundary layer where up- and downdrafts have maximum intensity. Towards the ground, a slight reduction in wind profiling error is observable for all wind speed cases. There, the size of the turbulent elements becomes smaller and consequently the lidar scan averages over more eddies, thereby better fulfilling the homogeneity assumption. This is a theoretical result, in real measurements other effects such as ground chirp will degrade near-ground retrieval of wind profiles (Godwin et al., 2012). Towards the top of the boundary layer, wind profile error decreases as the homogeneity assumption is better fulfilled. Nevertheless, entrainment and detrainment processes can still cause noticeable retrieval error, especially for higher wind speed cases. In the free, homogeneous atmosphere wind profile error vanishes (not shown).

The wind direction retrieval (fig. B8) is of similarly good quality compared to the wind speed retrieval. The wind profile points which pass quality filtering cluster in a small area around 270°, representing the westerly wind direction. The wind direction retrievals for the 0 m s^{-1} background wind case scatter throughout the full range as the wind direction is not meaningful without a background wind speed. As before, these values are completely eliminated by quality filtering. ~~Overall, the~~ The retrieval exhibits slightly degraded quality criteria (lower R^2, small y-axis intercept) ~~, but~~ compared to the wind speed retrieval. However, this behavior is due to the data not being distributed over a wide range as is the case for wind speed.

3.3 ~~Influence of system noise on wind profiling quality~~Quality filtering criteria

~~RMSE, REL, N and bias for four background wind cases at varying random radial velocity fluctuations and beam pointing inaccuracies. a), b), c), d) 0 m s^{-1} background wind case. e), f), g), h) 5 m s^{-1} background wind case. i), j), k), l) 10 m s^{-1} background wind case. m), n), o), p) 15 m s^{-1} background wind case.~~

~~Using the ADLS the importance of different error sources for wind profiling accuracy can be evaluated. In this study, three error sources are considered. As a standard, the violation of the inhomogeneity assumption due to boundary layer turbulence is always present in the retrievals. Additionally, two error sources due to measurement system noise can be emulated:~~

1. ~~The Doppler lidar radial velocity noise is emulated through addition of a Gaussian distributed random fluctuation to the measured radial velocity. The standard deviation of the Gaussian distribution is varied in a range of $\sigma v_D = 0 - 1$ m s^{-1}. This range of variation presents a worst case scenario, as stated before the WTX-system is expected to show a random radial velocity fluctuation below 0.25 m s^{-1} at 1 Hz measurement frequency.~~

2. ~~A beam pointing inaccuracy is emulated through addition of Gaussian noise to the beam pointing direction. The emulation is achieved through adding Gaussian distributed random numbers to the beam azimuth and elevation before retrieval. The standard deviation of the Gaussian distribution is varied in a range of $\sigma \boldsymbol{b} = 0 - 1°$. This magnitude also presents a worst-case scenario assumption, the achievable beam pointing direction accuracy is assumed to be much better and can be improved by using beam pointing calibration (Haimov and Rodi, 2013).~~

~~The results obtained for the three error sources are presented separately for each background wind case in fig. **??**, using the standard system setup and retrieval strategy as above. As before, a clear influence of the background wind speed on retrieval error is detectable, therefore the error characteristics are discussed individually for the 0, 5, 10 and 15 m s^{-1} case. For each background wind case, a maximum number of 9600 wind profile points is retrievable.~~

~~For the 0 m s^{-1} background wind case all wind profile points are eliminated by the quality filtering criteria (fig. **??** c). The responsible criterion is the R^2, as all volumes are sufficiently covered by the lidar beam, thereby making CN filtering not applicable. The raw, not quality filtered values exhibit an RMSE and bias in excess of 1 m s^{-1}, highlighting the need for application of quality filtering criteria (fig. **??** a, d).~~

~~For the 5 m s^{-1} background wind case, retrieval error strongly increases if system noise levels exceed $\sigma v_D > 0.25$ m s^{-1} or $\sigma \boldsymbol{b} > 0.25°$ (fig. **??** c,f) . Alongside this result, the number of retrieved wind profile points decreases above these noise levels, as quality filtering eliminates more retrievals (fig. **??** g). At higher noise levels a bias develops due to quality filtering (fig. **??**~~

h). This bias is an undesirable feature of the applied quality filtering, although only being present for a very small number of retrieved wind profile points. The unfiltered, raw wind profile retrieval is unbiased and actually shows slightly lower RMSE levels for the highest noise cases due to Assessing the reliability of the non-existent bias (fig. **??** e,h). parameters retrieved through the inversion process is a common problem in inverse theory and two metrics are investigated as a part of this study.

For the 10 m s^{-1} background wind case, the increase of retrieval error with increasing system noise is less pronounced. Turbulence is the main driver of wind profiling error for noise levels up to 0.75 m s^{-1} for radial velocity noise or 0.75° beam pointing inaccuracy, if considered individually, or up to 0.5 m s^{-1} and 0.5°, if both system noise sources are combined (fig. **??** i). Again, quality filtering eliminates wind profile points with high error levels, but also introduces a positive bias.

For the 15 m s^{-1} background wind case, turbulence remains the main driver of wind profiling error for all system noise levels considered, if varied individually (fig. **??** m). If both noise sources are combined, the highest noise level of 1 m s^{-1} and 1° shows a slightly degraded profiling quality and number of retrieved wind profile points.

Overall, the system noise analysis shows that for the expected system accuracy the largest error in wind profiling is caused by flow inhomogeneities. An exception are low wind speeds, when system accuracy can become important if high system noise levels are present. Further, quality filtering is necessary to eliminate unreliable wind speed retrievals. However, quality filtering has to be conducted with adequate thresholds as it can introduce a retrieval bias and thereby even increase retrieval error for low wind speeds (see also Sec. 3.3).

3.4 Influence of system setup on wind profiling quality

The results from the previous section emphasize the importance of turbulence as a driver of wind profiling error. Therefore, in order to show the influence of the system setup on wind profiling quality, two cases are considered in the following: The first case focuses purely on the wind profiling error due to the violation of the inhomogeneity assumption. This case is termed 'ideal system' as no system noise is included. The second case presents results for the combined system noise case due to both radial velocity noise and beam pointing direction inaccuracy, termed 'noisy system'. Lowest level noise (0.25 m s^{-1} for radial velocity noise and 0.25° for beam pointing inaccuracy) is introduced in order to show the tendencies of wind speed retrieval quality with respect to the different system noise sources. Under normal operating conditions a real system is expected to perform better than even the lowest system noise levels (De Wekker et al., 2012).

Wind profiling error for four different background wind cases WND $0, 5, 10, 15$ m s^{-1} using STP, RET. a) RMSE, REL, N and bias for an ideal measurement system. b) RMSE, REL, N and bias for a noisy measurement system with radial velocity noise $\sigma v_D = 0.25$ m s^{-1} and beam pointing inaccuracy $\sigma \boldsymbol{b} = 0.25°$.

The variation in retrieval quality for the different **background wind cases (WND)** is shown in fig. **??**. As discussed before, the absolute RMSE increases with increasing background wind speed as the deviations become larger. The opposite behavior is observed for the REL. The reason for the decreasing REL lies in the ratio of the vertical wind magnitude to the horizontal wind magnitude. For lower wind speeds, relatively strong up- and downdrafts exist which cause error in the horizontal wind speed retrieval. As the horizontal wind is of smaller magnitude, any deviations due to the vertical wind have a stronger impact if measured in a relative way. For higher wind speeds, the vertical wind magnitude does not increase linearly with

the horizontal wind magnitudeThe coefficient of determination R^2 (App. A4, eq. A22) is often used to detect a violation of the homogeneity assumption (Päschke et al., 2015; Wang et al., 2015). When using this approach, it is assumed that deviations from the homogeneous state show up as deviations of the measurements from the LSQ-fit. Wind field inhomogeneities smaller than the scan volume size decrease the R^2. Therefore, higher wind speeds show a decreased REL. Consequently, although the RMSE increases further for higher wind speeds, the results suggest that the REL approaches a threshold level. The different background wind cases thereby also represent varying turbulence intensity conditions, which have been shown to be of importance for Doppler lidar retrieval quality before (Banakh et al., 1995; Wang et al., 2016).

The underlying gray symbols indicate the respective quantities before application of quality filtering. Results show that for the 0 m s^{-1} background wind case all 9600 theoretically available wind profile points are removed by it is commonly assumed that the R^2 filtering (gray vs. black diamond). For the 5 m s^{-1} background wind case just 22 wind profile points are filtered for the ideal system (gray vs. red diamond). For the 10 m s^{-1} and 15 m s^{-1} background wind cases no wind profile points are eliminated by quality filtering for the standard system setup and retrieval strategy as they exhibit a magnitude of turbulence that does not result in R^2 filtering. The adequately explored volumes make CN filtering not applicable for any wind speed case for the standard system setup and retrieval strategy. Because no wind profile points pass quality filtering for the 0 m s^{-1} background wind case, no RMSE, REL and bias are specified. However, the unfiltered values show that the RMSE and bias are much larger than for the higher background wind cases. An RMSE above 1 m s^{-1}, a REL above 0.1 and a bias above 1 m s^{-1} are treated as unreliable wind profile retrievals in the following, therefore values above this error level are not included in the analysis. A reliable wind profile retrieval under lowest wind speed conditions is impossible to achieve using the standard system setup and retrieval setting as the radial velocity contribution of the vertical wind is dominant in this case. Consequently, an erroneous mapping of vertical wind into horizontal wind results in a biased retrieval (further discussed in Sec. 3.3).

The noisy system results show that the assumed system inaccuracy contributes to the retrieval error, but turbulence still remains the main driver (fig. **??** b). A noticeable increase in RMSE and REL is observed for the 5 m s^{-1} background wind case, while at higher wind speeds the increase is less. The retrieval remains bias-free. Further, a reduced number of retrieved wind profile points is evident under noisy conditions as a larger number of wind profile points are excluded by the R^2 filtering criteria. This quality filtering, however, only leads to a marginal reduction of error levels.

RMSE, REL, N and bias for 16 different flight directions using STP, RET. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

Flight direction (DIR) also has an influence on wind profiling quality for two reasons: first, due to the aircraft flying relative to air (see Sec. 2.2), upwind flights contain more samples per volume than downwind flights. Second, due to the alignment of turbulence with mean flow, different flight directions can show a more or less severe captures the degree of violation of the homogeneity assumption. Both effects can be observed in the ADLS results (fig. **??**). The results discussed here have to be interpreted with care, because the sample size is reduced for the directional analysis (a maximum of 600 wind profile points is retrieved for each setup). In all other analysis the individual flight directions are combined to generate a reliable number of samples. Overall, and especially for the higher background wind cases, upwind profiling yields the lowest RMSE (DIR WSW, colored asterisks) as more points fall into the ground-based volume used for the retrieval. Down-wind profiling on the other

hand shows slightly degraded results in comparison (DIR ESE, colored asterisks). The influence of alignment of turbulence with flow is also noticeable. In particular, for the 15 m s^{-1} background wind case, the flight direction W does not yield the best retrieval quality (lowest RMSE, blue asterisk), as would be expected based on pure sampling grounds. At the W direction, a stronger horizontal shear of the vertical velocity is encountered by the Doppler lidar, possibly due to along-flow alignment of rolls, leading to a higher error of the retrieved wind profiles. Consequently, the directions next to the straight upwind flight exhibit the best retrieval quality. Some directions show a small tendency for retrieval bias, which could be caused by the varying turbulence structure and alignment. However, the reliability and magnitude of this bias cannot be determined with good confidence due to small sample sizes used in this analysis. As discussed before, the addition of noise decreases retrieval quality especially for the 5 m s^{-1} background wind case. Nevertheless, the above described influence of flight direction on wind profiling error remains valid.

In the following, the results for the different flight directions are combined and then averaged to create a larger statistical dataset, as they show small systematic differences and do sample different air mass volumes.

RMSE, REL, N and bias for three aircraft flight altitudes using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**. In the colored areas the across-track averaging distance X_c is changed to the values specified above in order to adapt to the distance covered by the lidar beam.

The **aircraft flight altitude (AAL)** is a system parameter which can be easily varied in real measurements. At a lower flight altitude the across-track distance covered by the lidar beam is smaller. Thereby, Often used filtering criteria are $R^2 > 0.8$ or $R^2 > 0.95$ for ground-based wind profiling (Wang et al., 2015; Päschke et al., 2015). However, as shown already by Koscielny (1984), a linear change in the across-track averaging distance can be decreased, resulting in a smaller measurement footprint. However, a lower flight altitude also results in a smaller number of available wind profile points above the boundary layer, due to the decreased vertical coverage. The impact of aircraft flight altitude on retrieval quality is shown in fig. **??** for three flight altitudes which allow for full vertical boundary layer coverage. For the conditions investigated here, lowered flight altitudes result in a slightly increased RMSE for the 5, 10 and 15 m s^{-1} background wind cases (colored asterisks). The flight altitude can change error levels as the ratio between the lidar scan volume and turbulent eddy size changes. Most likely, for the situation investigated here, at the lowest flight altitude the scan volume inside the boundary is closer to the turbulence dominating eddy size. Thereby, the average is performed over a smaller number of eddies, resulting in a less fulfilled homogeneity assumption due to the short averaging distance. The increase in retrieval error for lower flight altitudes is not severe and has the benefit of having a decreased measurement footprint. The retrieval also remains unbiased for all flight altitudes. As before, the introduction of system noise increases the wind profiling error. This increase is of similar magnitude for all settings.

RMSE, REL, N and bias for six scan elevation angles using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**. In the colored areas the across-track averaging distance X_c is changed to the values specified above in order to adapt to the distance covered by the lidar beam.

The **scan elevation angle (ELE)** has a strong influence on wind profiling quality but also measurement footprint. To demonstrate the influence of scan elevation on wind profiling quality, the scan elevation angles are varied between 30° (shallow) and 80° (steep), with 90° being nadir (fig. **??**). For more shallow scan elevation angles the lidar beam covers a larger

across-track distance. Consequently, the across-track averaging distance has to be adjusted, resulting in a larger measurement footprint. For steeper scan elevation angles the lidar beam covers a smaller across-track distance, resulting in a more confined measurement footprint. The along-track averaging distance is kept constant at the standard value of 1300 m for all setups.

At steep scan elevation angles the vertical wind can exhibit greater influence. Consequently, at elevations steeper than 60°, the wind profiling quality is degraded (higher RMSE, colored asterisks) and can only be partially improved through quality filtering. The degradation is not just due to a larger retrieval scatter, in addition the retrieval also becomes biased (colored plus signs) due to erroneous mapping of vertical wind into horizontal wind, enabled by its large radial velocity contribution. Quality filtering can improve the wind profiling quality for steep scan elevations (gray vs. colored asterisks). However, it also worsens a preexisting wind profiling bias for the 5 m s^{-1} background wind case (gray plus sign vs. red plus sign). Further, quality filtering introduces a previously non-existent profiling bias for the 10 m s^{-1} and 15 m s^{-1} background wind cases (gray plus signs vs. green, blue plus sign). Additionally, for steeper elevation angles the number of usable wind profile points decreases (colored diamonds). The reason is twofold: First, due to the increased projection of the range gate length onto the vertical, the maximum number of available wind profile points is lower. Second, quality filtering also decreases the number of usable wind profile points for scan elevation angles greater than 70°. The degradation is more severe for the lower background wind cases, as fewer points fulfill the quality filtering criteria (in this case the filtering occurs due to R^2). In summary, for the standard retrieval strategy, scans with more than 70° elevation do not allow for wind profile retrieval with reasonable error margins under turbulent conditions for all wind speeds, despite having smaller measurement footprints.

When using scans with elevation angles more shallow than the standard 60° elevation, an improved retrieval quality can be achieved at the cost of having a larger across-track averaging distance. On the one side, ADLS results show decreased error levels (colored asterisks) and an increased number of wind profile points being available (colored diamonds). The decrease in error level is due to a smaller influence of the vertical wind on retrieval error. The increased number of retrieved wind profile points is due to the more shallow scan elevation, which leads to a smaller vertical projection of the range gate length. The additional points do not necessarily contain additional, independent information due to the vertical autocorrelation of turbulence and thereby wind profile error (discussed in Sec. **??**), limiting the increase in vertical resolution. On the other side, for more shallow scans, the footprint of the measurement starts to increase as a larger ground distance is covered by the lidar beam and the across-track averaging distance has to be increased. ADLS results have to interpreted with care for elevation angles smaller than 40° due to the limited along- and across-track extent of the LES domain. Once the across-track distance covered by the lidar beam approaches the LES volume, the beam exits the LES volume frequently for some flight directions and estimation of wind profiling quality becomes unreliable. Nevertheless, for scan elevations angles more shallow than 50°, the unfiltered 0 m s^{-1} background wind case retrieval appears within the reliable retrieval limits (gray asterisks). Here, refraining from R^2 quality filtering would allow for retrieval of the lowest wind speeds, although with a biased retrieval.

The discussed wind profiling error characteristics are also present when introducing the system noise. The increase of profiling error and bias for steep elevation angles is more pronounced, and the number of measurements eliminated by quality filtering increases further. As before, the profiling error decreases for scans more shallow than 60° elevation. Further, the

slightly reduced number of retrieved wind profile points for the 5 m s^{-1} background wind case is not present for scans with $< 50°$ elevation.

RMSE, REL, N and bias for five scan speeds using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

Scan speed (ROT) also has an impact on wind profiling quality (fig. **??**). At slower scan speeds (longer rotation duration) the number of independent measurement points in the volume used to create the wind speed estimate is reduced due to autocorrelation among the measurements (the number of measurements is the same as measurement frequency is unchanged). Faster scan speeds allow for greater azimuth diversity and measurements which are less correlated, thereby enabling an improved wind profiling quality, despite having an overall equal number of measurement points available in each volume. All scan speeds show bias-free results and achieve the maximum number of retrievable wind profile points. For fast scan speeds with rotation duration of 16 s and below, the raw RMSE of the 0 m s^{-1} background wind case becomes smaller than 1 m s^{-1}. Similarly, the bias is also reduced, but the applied quality filter still eliminates all retrievals. One important aspect which is not captured by the ADLS is the influence of increasing scan speed on return signal quality and thereby Doppler speed measurement quality. A rapid scan movement leads to a strongly changing IAS contribution during the scan and sampling of more turbulent elements. If these rapid scans cause problems in Doppler velocity estimation due to spectral broadening or other effects, slower scan speeds become preferable again.

When system noise is considered, the wind profiling error reduction for fast scan speeds is lowered, as autocorrelation among the individual measurements is reduced due to the noise. Thereby, increasing the scan speed does not yield the same magnitude of wind profiling quality improvement as it is the case without system noise. However, for the 5 m s^{-1} background wind case, the slightly reduced number of retrieved wind profile points is corrected for by faster scan speeds. Further, raw retrievals for the 0 m s^{-1} background wind case again show a decreasing RMSE and bias.

RMSE, REL, N and bias for three lidar measurement frequencies using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

ADLS results show that an increased **lidar measurement frequency (FME)** does not improve wind profiling quality (fig. **??**). This insensitivity exists because higher measurement frequencies produce measurements in between the lower frequency measurements which are highly autocorrelated. Due to the autocorrelation (resolving the turbulent structures in greater detail but not fulfilling the homogeneity assumption better) the additional measurements do not provide new information that can be used to improve the LSQ-fit. In fact, vertical wind biases the retrieved horizontal components and a linear change in the R^2 is lowered for some wind profile points from the 5 m s^{-1} background wind case, as turbulence is resolved in more detail. This R^2 reduction results in a filtering of some wind profile points, thereby the number of retrieved wind profile points is slightly degraded for the ideal system case. The results of the measurement frequency analysis could be altered for a real system if the signal return quality is influenced by the signal accumulation time. On the one hand, if a longer pulse accumulation leads to a more reliably estimated Doppler velocity, lower measurement frequencies become preferable. On the other hand, if the change in radial velocity due to the scanner movement leads to problems in Doppler velocity estimation, higher measurement frequencies become preferable.

Adding artificial noise to the measured radial velocities in the ADLS leads to a loss of retrieved wind profile points for the lower background wind cases at high measurement frequencies. This behavior is caused by the severely increased radial velocity noise (1 m s^{-1} at 10 Hz) and the frequency constant beam pointing inaccuracy, both leading to more filtering of wind profile points by the R^2 criterion. Due to the non-linear increase in random radial velocity noise (see Sec. 2.2.3) the 10 Hz measurement frequency case even exhibits slightly elevated profiling error levels. The results discussed for the FME analysis are only valid for the specific noise setup used. They can be altered for a real system with differing radial velocity and beam pointing inaccuracy. If these are known, a further system optimization can be conducted using the ADLS.

3.4 Influence of retrieval settings on wind profiling quality

As a volume velocity processing (VVP) type of algorithm (Waldteufel and Corbin, 1978; Boccippio, 1995) is applied to infer the wind field parameters, a number of options exist to vary the retrieval and inversion settings. In this study, the impact of the number of retrieved parameters (VAR, only horizontal components vs. including vertical component) and the way in which the volume is defined (SEP, volume-based, time-based) is analyzed. The horizontal along-track averaging distance is varied (AVG 650, 1300, 2600 m) as well as the vertical extent of the volume used for processing (LEV 62, 124, 248 m). Further, stability of the wind retrieval with sectors being omitted is investigated, to assess effects such as data omissions due to the presence of clouds in real measurements (BLA $0-90°, 0-180°, 0-270°$).

RMSE, REL, N and bias for u,v,w and u,v component estimation approaches using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

The **number of retrieved parameters (VAR)** is an important setting that can be varied in the inversion process (Boccippio, 1995). This study focuses on the wind profile retrieval under inhomogeneous conditions at highest resolution with a minimum number of data points. Therefore, the two most simple options are analyzed:

1. Retrieval of the full wind vector, consisting of the u,v,w component (Leon and Vali, 1998; De Wekker et al., 2012).
2. Retrieval of the horizontal wind vector only, consisting of the u,v component (Koch et al., 2014).

If the w component is neglected in the inversion, any persistent vertical wind speed causes a bias in the horizontal components biases the retrieved vertical component. Both deviations are not detectable as deviations from the LSQ-fit. On the other hand, the stability of the solution could be increased for cases in which short scans/few points are available.

The number of retrieved components does not exhibit a discernible influence on retrieval quality (fig. **??**). Both retrieval methods show equal error levels and an equal number of retrieved wind profile points. This insensitivity goes hand-in-hand with the observation that the average vertical wind speed can be retrieved with reasonable quality (fig. B6), and so retrieving all three wind components is the preferable choice. Inclusion of system noise in the retrieval process does not change the discussed behavior. The results could be altered in more challenging measurement conditions (e.g. due to the presence of sector blanking), as for sector blanking higher CNs are reached.

RMSE, REL, N and bias for volume and time-based profile separation approaches using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

For the AVAD analysis, all measurements within a defined volume are evaluated for the wind profile retrieval (essentially making it a VVP approach). The vertical and horizontal extent of the volume need to be specified by the user depending on the intended observational task. In the ADLS two options for the **separation of wind profiles (SEP)** used for processing exist:

1. Retrieval of the wind profile in fixed, ground-based volumes with a given size that does not depend on the flight direction (Guimond et al., 2014; Tian et al., 2015). In this case, upwind flights contain more samples per given volume than downwind flights.

2. A time-based approach where all measurements within a defined time period (e.g. individual scans) are analyzed independent of their pointing direction and sampling volume (Leon and Vali, 1998; De Wekker et al., 2012). In this case, downwind flights yield slightly degraded and upwind flights slightly improved horizontal resolution. In addition, the scanned volume and thereby measurement footprint can be very large as aircraft motion and scanner movement are added together.

ADLS results show that the separation method does not exhibit a strong influence on the retrieval quality (fig. **??**). This insensitivity is expected, as both methods use a similar amount of data from the same setup to retrieve the wind vector. Due to the fixed and known retrieval footprint, the volume-based approach is therefore the preferable choice. The described results remain valid when artificial system noise is considered in the measurement process.

RMSE, REL, N and bias for three along-track averaging distances using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

For any of the strategies, the extent of the averaging period of data used for the VVP processing has to be specified. Consequently, the effect of varying the extent of the horizontal averaging domain is analyzed using the ADLS. For the standard system setup and retrieval setting the **along-track averaging distance (AVG)** is varied between 650, 1300 and 2600 m (determining the along-track wind profile resolution). The along-track distances correspond to the approximate distance covered by the aircraft during 0.5, 1 and 2 scanner rotations. However, the moving aircraft allows for directing the laser beam ahead and back and therefore measures the same spot at different times under varying azimuth angles. The across-track averaging distance is set to 1963 m for all volumes, corresponding to the across-track ground distance covered by the lidar beam at a flight altitude of $h = 1700$ m and the standard scan elevation of $\phi = 60°$.

As expected, varying the along-track averaging distance used for profile retrieval has a strong influence on wind profile error (fig. **??**). High-resolution wind profiling at 650 m is associated with a larger RMSE which can be lowered to some extent by quality filtering. For higher background wind cases, the strongest increase is caused by downwind profiling directions which exhibit elevated error levels (not shown). Interestingly, the application of the quality filtering criteria does not completely filter the 0 m s^{-1} background wind case anymore, leading to a severely biased retrieval. This is further discussed in Sec. 3.3. The error decreases for all wind speeds if a longer averaging distance of 1300 m is used, and in this case differences between up- and downwind wind profiling become smaller as discussed before. The decrease, however, is not linear and a further extension of the averaging distance to 2600 m does not yield an improvement of similar magnitude. It does, however, lead to the 0 m s^{-1} background wind case appearing within the reliable retrieval limits for both RMSE and bias. This analysis can also be inverted,

that is, if an acceptable error threshold is given, the required along-track averaging distance can be determined (including usage of the quality filtering analysis, for which the 650 m case is used as a challenging setting). When system noise is included in the analysis, the above described behavior holds at overall higher error levels. As expected, the increase in error due to system noise is more pronounced for the shorter along-track averaging distances. Further, the quality filtering eliminates more points for shorter distances, resulting in slightly decreased error levels.

RMSE, REL, N and bias for three vertical averaging distances using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

Another parameter which can be varied is the **vertical averaging distance (LEV)**, corresponding to the vertical wind profile resolution. Depending on the vertical analysis volume extent, the number of neighboring range gates used to create one wind profile point differs and can influence the LSQ-fit quality. The vertical averaging distance is varied between 62 mIf the matrix inversion is performed based on a singular value decomposition, additional quality criteria such as the condition number CN (App. A4, eq. A20), 124 m describing the degree of collinearity among the model geometry, become available and are frequently utilized (Boccippio, 1995; Holleman, 2005; Cheong et al., 2008; Shenghui et al., 2014; Päschke et al., 2015; Wang et al., 2015). The CN provides a measure of the spread of the model space, thereby diagnosing collinearity (Boccippio, 1995; Leon and Vali, 1998; Shenghui et al., 2014; Päschke et al., 2015; Wang et al., 2015). A high CN indicates high collinearity in the model geometry. Consequently, the real system state is not well explored in at least one direction and 248 m, which corresponds to one, two or four range gates falling into the retrieval volume at 60° elevation. The influence of varying the vertical averaging distance is hardly discernible in terms of wind profile error (fig. **??**). However, the number of retrieved wind profile points varies strongly, as expected, because the vertical profile resolution is changed. The reason for vertical averaging having no influence is founded in the turbulence structure. Neighboring range gates probe mostly similar turbulent elements and thereby suffer from similar errors. Therefore, including more range gates in the LSQ-fit does not improve the retrieval quality as the additional data is highly correlated. Only a marginal reduction is observed for the 248 m vertical averaging distance, but at the cost of having far fewer samples available, which can be problematic in regions of strong vertical shear. Another conclusion from the error characteristics is that wind profile error is usually correlated over multiple vertical levels. Consequently, Doppler lidar retrieved wind profiles appear smooth, but can still suffer from non-negligible error throughout the full profile height (resembling a bias, but as this error averages to zero over many profiles it is more accurately described as a random error). This correlation is the reason why the error discussed here is difficult to identify visually in real Doppler lidar wind profiles: No small-scale variations occur, but rather offsets over large sections of the profile. When artificial measurement noise is introduced, vertical averaging over multiple range gates improves the retrieval quality and the number of retrieved wind profile points slightly, with the improvement being stronger for the 5 m s^{-1} background wind case. With more vertical averaging, more measurements fall into the retrieval volume, and so consequently the radial velocity noise is reduced due to averaging (beam pointing inaccuracy is not , as it is the same for all vertically neighboring range gates).

RMSE, REL, N and bias for four blanked sectors using STP, RET otherwise. a) Ideal system. b) Noisy system. Color-coded points as in fig **??**.

Last, the robustness of the retrieved wind profile to **partial data unavailability (BLA)**, here termed 'blanking', is investigated (fig. **??**). In real measurements in the boundary layer, sectors of the scan can become unavailable due to the presence of clouds, rain, obstacles or other disturbances. The effects are simulated by discarding a specific sector of the scan geometry for wind profile retrieval (in the ground-based reference system ENU, in which the disturbance would be present). For the analysis, three cases are considered: sector blanking from $0-90°, 0-180°$ and $0-270°$. A sector blanking of $90°$ leads to a moderate increase in the average error levels for all background wind cases without significant loss in the number of retrieved wind profile points. With a further increase to $180°$ sector blanking the wind profiling quality is strongly degraded, which is more prominent for the downwind flight directions (not shown). More of the 5 m s^{-1} background wind speed retrievals are quality filtered, which again introduces a slight bias. Further, the exclusion of the 0 m s^{-1} background wind case is not completely successful anymore, which again results in a severely biased retrieval. With $270°$ sector blanking, the CN exceeds values of 25 and no reliable wind profile retrieval is achievable anymore. Once again, the introduction of noise does not change the discussed behavior significantly but results in elevated error levels especially for the 5 m s^{-1} background wind case. Further, a reduced number of wind profile points is retrievable and the retrieval becomes biased beyond $90°$ sector blanking.

3.4 Quality filtering criteria

as a result the inferred system state is prone to a large error amplification (Boccippio, 1995). The CN is often used as a quality filtering criteria to exclude measurements where the spread of the beam pointing directions is not sufficient to explore the wind field adequately (Boccippio, 1995; Päschke et al., 2015; Wang et al., 2015). In this section, the performance of the most commonly used quality criteria R^2 and CN is evaluated, specifically their relation to wind profile quality and their adequacy in detecting violations of the homogeneity assumption and collinearity in the model geometry. A first performance check is conducted for the standard system setup and retrieval setting for all parameters. In a second performance check, the high resolution retrieval setting (AVG 650 m) is used as in this case the analyzed volumes show varying coverages by the airborne Doppler lidar, making reliable estimation of the wind vector more challenging.

Quality filtering criteria for the standard system setup and retrieval strategy (STP, RET). a) Color-coded histogram of number of occurrence for retrieval error and R^2. Gray crosses show values eliminated by quality filtering. b) Retrieval error and R^2, color-coded is CN. c) LES wind speed and ADLS wind speed, color coded is R^2. Profile points which pass quality filtering are displayed as color-coded circles, profile points which are eliminated as color-coded crosses. d) Same as a), but for CN. e) Retrieval error and CN, color coded is R^2. f) Same as e), but for CN.

The analysis of the simulator results for the ideal measurement system setup and standard retrieval setting reveals a number of interesting findings (fig. B98 and fig. B9 for the full range of non-quality filtered values). Deviations from the geostrophic background wind towards lower wind speeds, associated with stronger turbulence in the boundary layer, generally decrease the value of the R^2 but do not influence the CN, as it is expected (fig. B9 8 c, f). Even at high $R^2 > 0.90$, significant wind profile errors up to 1 m s^{-1} do exist. The CN values cluster in a narrow range with values between $3-4$, as the sampling volume is generally well-explored by the lidar for all retrievals due to the checkerboard approach.

A noteworthy feature occurs for the 0 m s^{-1} background wind case. Here, a clear dependence of wind speed error on the R^2 value is observable (fig. B9, a, b, c). The estimated wind speeds are only slightly biased for coefficients of determination in the range close to 0. However, with increasing R^2, the bias of the retrieved wind speed increases linearly to values in the range $1-2$ m s^{-1}. This problem exists despite the fact that the individual estimated components themselves are unbiased (fig. B6). Coefficients of determination above 0.9 are non-existent for the 0 m s^{-1} background wind speed, therefore the biased values can be filtered ~~without problems.~~

completely. The observed behavior is caused by the lidar measurement method: at low wind speeds, the retrieval is strongly influenced by the vertical wind and less by the horizontal wind, especially at steep elevation angles ~~.~~(see also fig. 4). For high coefficients of determination, the w variations are more in-phase (with an expected horizontal wind contribution from the simplified horizontal wind field model) and thereby mapped more into horizontal wind, causing a stronger positive bias. The values of the horizontal wind speed in the range of $0-3$ m s^{-1} correspond with a mean vertical wind of $0-1.5$ m s^{-1} amplitude being mapped into the horizontal wind at $60°$ elevation. At lower coefficients of determination, the noisy (sub-scan-volume) w variations overwhelm the smaller horizontal wind speed signal. However, they do not lead to a mean wind speed being estimated as they cancel out due to averaging, thereby causing no bias. If the scalar averaged LES wind speed is used for lidar comparison instead of the vector averaged LES wind speed discussed here, the described linear trend is equally present, but with an intercept offset of -1 m s^{-1} along the y-axis. However, for shallow elevation angles, giving more accurate retrievals, the estimated lidar wind speed tends towards the vector average, making it the correct choice for comparison.

In summary, for low wind speeds, the application of a higher R^2 mainly filters for smoother in-phase variations of the vertical wind~~(and/or system noise, if present)~~. Consequently, selecting a higher threshold for R^2 can cause a higher bias, an important finding that is not expected and, to our knowledge, undocumented so far. For low wind speed cases, the R^2 is not an appropriate quality filtering criteria, unless an appropriately high threshold is chosen which then successfully filters the low wind speed cases. In the investigated case, the 0 m s^{-1} background wind cases are completely eliminated by setting strict quality filtering criteria. ~~However, if more challenging conditions are used, even the few points that remain can cause a non-negligible retrieval bias due to their large estimation error (fig. **??**).~~

~~Quality filtering criteria for along track averaging distance AVG 650 m and STP, RET otherwise. Panel labeling same as fig. B9.~~

~~In the challenging analysis setup using a shorter averaging distance AVG 650 m (fig. **??**), a higher spread of the retrieved wind profile points compared to the longer averaging distance is apparent. This increased spread arises from the only partial exploration of some retrieval volumes by the lidar. Therefore, a greater range of R^2 and CN are explored by the data, and relatively more wind profile points are excluded (see also fig. **??**). The previously discussed retrieval bias for low wind speeds with increasing R^2 is prominent again. As some of the biased points with a R^2 higher than 0.9 are determined from a good sampling distribution (small CN), they are not completely excluded by the quality filteringcriteria. Consequently, the vertical wind mapping causes a strong positive wind speed retrieval bias for low wind speeds~~ For the retrievals investigated here, a clear separation between biased retrievals from the 0 m s^{-1} background wind case and the non-biased retrievals from the 5 m s^{-1} background wind case is possible, based on $R^2 > 0.9$ threshold filtering.

~~Nevertheless, in general, outliers are connected to either a high CN or a decreased R^2, stressing the importance of both quality filtering criteria. For this retrieval setting, the wind speed retrieval error starts to show outliers beyond 3 m s^{-1} for CN$>$ 9. A reduction of profile error spread is observable for CN$<$ 5 with the loss of relatively few windprofile points. Additionally, relatively few points lie in the region $R^2 = 0.9 - 0.95$, therefore the results justify the usage of stricter quality filtering criteria in this case to obtain a reduced profiling error. In this way, the ADLS allows for determining a meaningful quality threshold for each setting depending on the accuracy desired.~~

However, in reality, this separation is not necessarily possible, as intermediate background wind speeds in the range $0 - 5$ m s^{-1} might be present. These cases may also result in biased retrievals, as long as the radial velocity contribution of the vertical wind is of comparable magnitude to that of the horizontal wind, giving the possibility for non-negligible erroneous mapping. A possible bias for unknown reasons, varying by day, has been reported by Weissmann et al. (2005) already. They stress the need for further investigation of this phenomena using simulated lidar data. The bias at low wind speeds is also noticeable in studies comparing lidar-measured wind speeds to in-situ measurements, however without discussion or explanation so far. Some recent examples of the observed overestimation at low wind speeds are, besides others, visible in Holleman (2005) (ground-based; fig. 4, fig. 5), Chouza et al. (2016a) (airborne; fig. 2, fig. 3b under turbulent conditions inside the boundary layer) and Pauscher et al. (2016) (ground-based; fig. 5e).

Therefore, based on the above discussion, we recommend the following procedure in wind profile analysis: In doubtful situations, where the approximate magnitude of the measured wind speed is unknown, one should always analyze a long spatial average over multiple scan rotations first (approx. 10 times the expected maximum eddy-size). For longer averaging, the mapping of vertical wind does not influence the retrieval significantly as the structure of the vertical wind is not in-phase with the measurement geometry over long distances ~~.~~ (except for extended regions of horizontally sheared vertical wind, e.g. in complex terrain). Thereby, the approximate magnitude of the wind speed can be analyzed reliably, although at coarse spatial resolution. If the outcome of this analysis is a wind speed magnitude below 5 m s^{-1}, ~~a~~ and a non-negligible vertical wind with similar magnitude is present, a further analysis at higher spatial resolution is not advisable due to the possibility of erroneous mapping. If, on the other hand, the wind speed magnitude is above this threshold, ~~the~~ or the vertical wind magnitude is negligible, the spatial resolution of the retrieval can be refined further. In the future, if advanced and flexible measurement system are available, it also appears possible to adjust the measurement system setup based on the encountered wind field conditions. For this procedure, an analysis of preferable measurement system setups and retrieval strategies which are less influenced by erroneous mapping appears worthwhile to conduct in future studies.

~~When leaving~~ Leaving the low wind speed problem aside, even when using strict quality filtering criteria, considerable wind speed retrieval error can still exist due to inhomogeneities. Or, stated in another way, even relatively small deviations from the homogeneity assumption that do not lead to significant degradation of the quality filtering criteria can cause noticeable retrieval error. For these cases, the possibility of an in-situ uncertainty forecast based on the measured ~~radial velocities appears worthwhile~~ data appears promising to investigate.

Overall, it can be concluded that wind profiling error due to violation of the homogeneity assumption can be of non-negligible magnitude. The errors reported in this study, solely due to flow inhomogeneity, are of comparable magnitude, or even above,

what has been reported as overall wind profiling error in other studies (Weissmann et al., 2005; De Wekker et al., 2012; Bucci et al., 2018). Wind profiling error due to flow inhomogeneity should therefore not be neglected as an important source of error when evaluating the measurement accuracy of airborne Doppler lidar systems.

4 Conclusions

In this study, an LES-based airborne Doppler lidar simulator ~~, the ADLS, and its capabilities are~~ is presented to provide a ~~valuable~~ link between Doppler lidar theory and measurements~~in order to optimize instrument deployment and to quantify wind profiling error~~. As an example application of the ADLS, the error associated with wind profiling in turbulent flow conditions is investigated. This error results from a violation of the homogeneity assumption used for wind profile retrieval and can be isolated using the ADLS, as the input truth used to create the measurements is known exactly.

The ADLS utilizes an LES wind field at 10-m grid spacing to simulate the turbulent convective boundary layer with geostrophic background wind speeds of $0, 5, 10$ and 15 m s^{-1}. The emulated measurement system consists of an aircraft, scanner and lidar, ~~for which varying setups are available,~~ considering many geometric transformations applicable in real-world measurements. After system emulation, a simulated measurement is conducted based on the LES wind fields. ~~Here, both an ideal system and a noisy system with an added Gaussian distributed random radial velocity fluctuation and beam pointing inaccuracy are simulated. Finally, the wind profile retrieval is performed on the simulated measurement data using a variety of retrieval strategies. The wind profiles are then evaluated against the original wind field data supplied as input, thereby revealing errors due to the lidar measurement process. The~~ The ADLS does neither include a simulation of the signal return (aerosol scattering) process nor of the lidar instrument physics. These processes have been investigated adequately elsewhere and problems associated with them are expected to be of manageable magnitude in practice. Instead, a direct measurement approach as chosen by other recent studies is further extended here (Stawiarski et al., 2013; Guimond et al., 2014). ~~Nevertheless, the effect of system noise in the measurement process can be emulated through superimposed random radial velocity fluctuations in the simulated measurements and are considered in this study.~~

As an example application of the ADLS, the accuracy of high-resolution airborne wind profiling under inhomogeneous~~flow conditions~~ , turbulent flow conditions inside the boundary layer is investigated. ~~The~~ For this application, the ADLS demonstrates its usefulness as representation problems present in studies comparing lidar to in-situ measurements can be avoided. ~~Thereby~~ Thereby, we are able to provide ~~an analysis of wind profile retrieval quality for various system setups and retrieval settings.~~

~~The results show that wind profiling retrieval errors due to inhomogeneous flow conditions can be of considerable magnitude and even render the retrieved wind profiles useless in~~ a quantitative analysis of airborne Doppler lidar retrieval error due to the violation of the ~~case of low wind speeds or inadequate~~ homogeneity assumption used in AVAD.

Results show that the standard system setup and retrieval ~~settings. Under the conditions investigated , errors due to flow inhomogeneities are larger than those due to the worst-case random radial velocity fluctuations and beam pointing inaccuracy.~~

For low wind speeds the system accuracy can become important if the assumed worst-case system noise levels are exceeded. Nevertheless, strategy investigated (based on the ADLS, wind profiling error margins can be constrained to an acceptable level, by choosing appropriate system setup and retrieval settings depending on the accuracy desired.

ADLS results allow for determination of preferential wind profiling strategies to achieve a small measurement footprint. In the comparison of flight altitude decrease versus scan elevation increase, this study shows that a decrease in flight altitude is the preferable option. Although at the cost of decreased vertical coverage, for a decreased flight altitude the retrieval quality remains almost unaffected. A possible remedy for the decreased vertical coverage could be a system which can scan down- as well as upward iteratively, thereby providing wind information inside the boundary layer as well as above. When using a fast-switching 10-Hz scanner and lidar system, wind profiling could be achieved without loss of along-track resolution characteristics of existing systems) allows for wind profiling with acceptable accuracy in inhomogeneous flow (RMSE $= 0.36$ m s^{-1} for wind speed). Nevertheless, retrieval errors due to the autocorrelation present in high-frequency measurements.

Scan elevation angles steeper than 70° are problematic under turbulent conditions in the boundary layer due to the strong influence of vertical wind on retrieval quality. With the specific retrieval setting used in this study, for the scan elevation angle an acceptable error vs. measurement footprint trade-off is achieved at 60° elevation. For elevations greater than 70°, the wind profiling error increases rapidly, the retrieval can become biased and a smaller number of wind profile points is retrieved. A faster scan speed can have a positive effect on retrieval quality if random system noise remains low, as in this case autocorrelation between adjacent measurements is decreased. The same principle is valid for increasing measurement frequencies: for low noise systems, the benefits of higher measurement frequencies are marginal as the additional measurements provide little more information due to strong autocorrelation. For higher noise systems the autocorrelation between measurements is decreased and higher measurement frequencies can provide better quality.

Retrieval settings can also impact wind profiling quality. For the standard system setup, our results show only marginal differences between retrieving only the horizontal wind components or including the vertical wind. Comparing volume and time-based approaches, both approaches yield comparable errorlevels but the volume-based approach has the advantage of having a defined measurement footprint. As expected, the along-track averaging distance has a strong influence on wind profiling quality and longer horizontal averaging increases retrieval quality. In comparison, the vertical averaging distance does not exhibit a strong impact on retrieval quality. Vertical averaging improves retrieval quality only slightly due to the similar structure of turbulence over multiple range gates. Under the standard conditions investigated here, the wind profile retrieval is robust to up to 90° of the azimuth scan sector being unavailable (due to clouds or other blocking), at slightly increased error levels.

Our results show that the careful usage of R^2 and CN as quality filtering criteria is necessary and adequate, as they can improve wind profiling quality by flagging unreliable measurements (e.g. 0 m s^{-1} background wind case). However, one has to bear in mind that quality filtering by R^2 can also introduce a retrieval bias for problematic system setups or retrieval strategies. Quality filtering with CN is always useful and does not introduce any systematic errors. violation of the homogeneity assumption used for retrieval are non-negligible and can present a major contribution to overall wind profiling error. An important new finding of this study is that Doppler lidar wind profiling, at low wind speeds under turbulent conditions, can

be unreliable and even biased despite or even because of applying quality filtering criteria. For low wind speeds, erroneous mapping and filtering of the vertical wind into horizontal wind can occur, thereby biasing the retrieved wind speed. ~~The problematic situations cannot be reliably detected by the usage of quality filtering criteria such as the R^2 and CN in the least-squares estimation. On the contrary, inadequate usage of the R^2 can lead to a higher bias of lidar wind speed retrievals in situations with low background wind speeds. System setups which reduce the magnitude of this problem are those using shallow scan elevation angles or fast scan speeds. However, these come at the cost of having a wider measurement footprint or possibly degraded signal quality. As a remedy, if these system setups are unavailable or undesirable, we propose to always begin the analysis of airborne Doppler lidar wind profiling at coarse resolution, by averaging over long distances. In this way, the wind speed regime can be reliably identified. If, on the one hand, this analysis yields a wind speed magnitude below 5 m s^{-1}, a further analysis at higher spatial resolution is not advisable. If, on the other hand, the wind speed magnitude is above this threshold, the spatial resolution of the retrieval can be refined further.~~ In order to resolve this issue, an analysis of preferable measurement system setups and retrieval strategies, which are less prone to erroneous mapping, appears worthwhile to investigate in future studies.

Overall, this study highlights the benefit of an LES-based airborne Doppler lidar simulator. ~~It~~ The ADLS offers a promising opportunity to investigate the lidar measurement process more closely, not only with respect to wind profiling, but for example also with respect to estimating turbulent quantities from ~~airborne Doppler lidar measurements.~~ nadir measurements. Further, the ADLS can also serve as a algorithm testbed for more complex retrieval approaches such as airborne Dual-Doppler. Regarding AVAD wind profiling, it is necessary to investigate the possibility of an in-situ uncertainty estimation of retrieved wind profiles based on the measured data. Here, the ADLS can provide a valuable contribution by allowing for method validation under close-to realistic conditions, while knowing the input truth, which is nearly impossible in real measurement setups.

Code and data availability. The underlying MatLab code and data are available from the author upon request.

Author contributions. PG developed the simulator, performed the simulations, conducted their evaluation and prepared the manuscript which was improved by all co-authors who acquired the project funding and also contributed in supervising the work.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Christoph Knigge and Siegfried Raasch from the University of Hanover are acknowledged for granting the right to use the LES data. The first author is thankful for PhD supervision by Christoph Kottmeier.

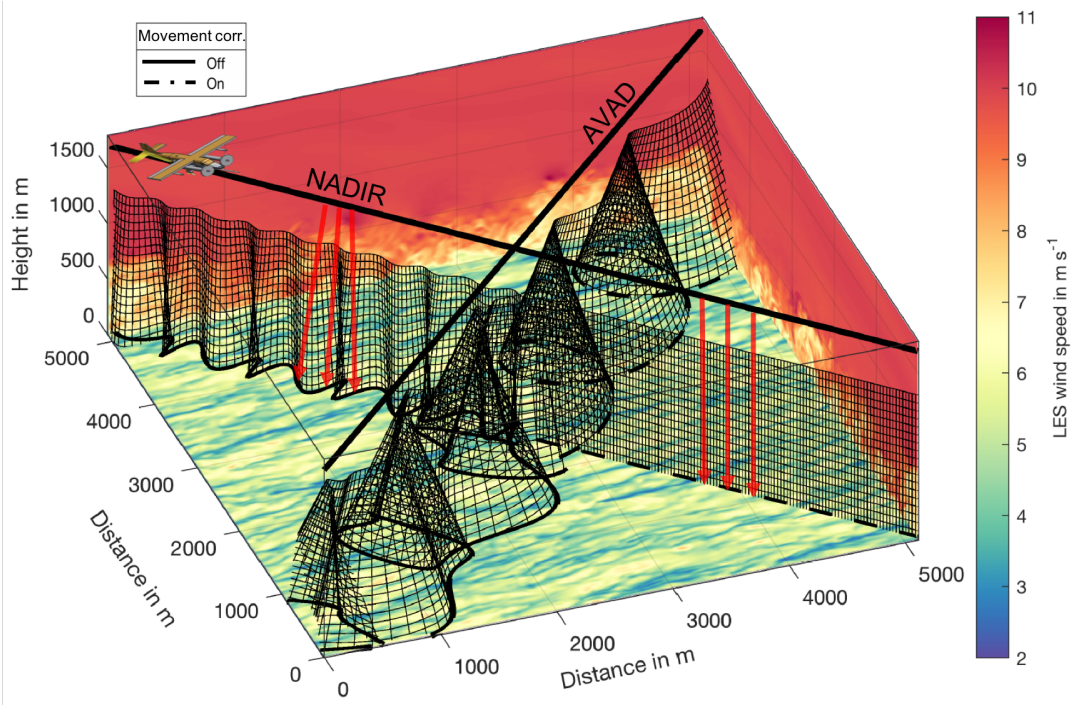

Figure 1. Illustration of system and measurement position simulation for two transects through the LES domain. The LES wind speed is color-coded. The black lines represent aircraft trajectories. The black curtains show the range gate positions used to conduct the measurement which are calculated from aircraft position, scanner and lidar system simulation. For the first half of each transect the scanner movement correction is disabled, whereas for the second half it is enabled. The nadir transect is used to retrieve the vertical wind, whereas the AVAD pattern is used to retrieve the horizontal wind.

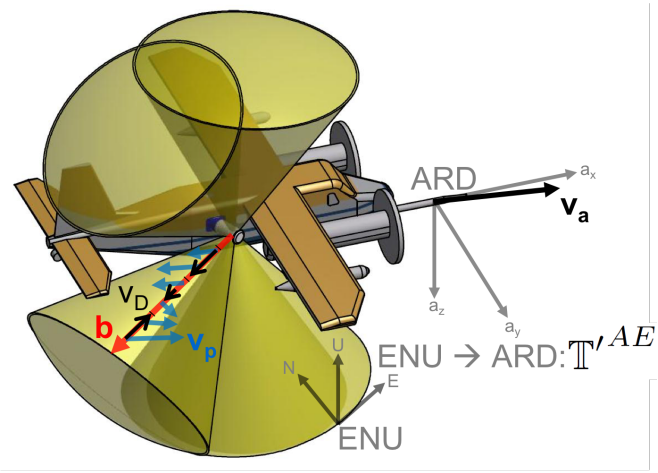

Figure 2. Schematic depiction of a possible airborne Doppler lidar system. The sketch is based on an upcoming system for the Dornier 128-6 aircraft of the TU Brunswick (D-IBUF, Corsmeier et al. (2001)). The lidar is inside the aircraft pointing outward, with a scanner mounted on the side of the fuselage directing the beam in the atmosphere. Displayed are the aircraft, scan cone surfaces as well as the coordinate systems and vectors used in eq. 3, for additional information see Sec. 2.3.

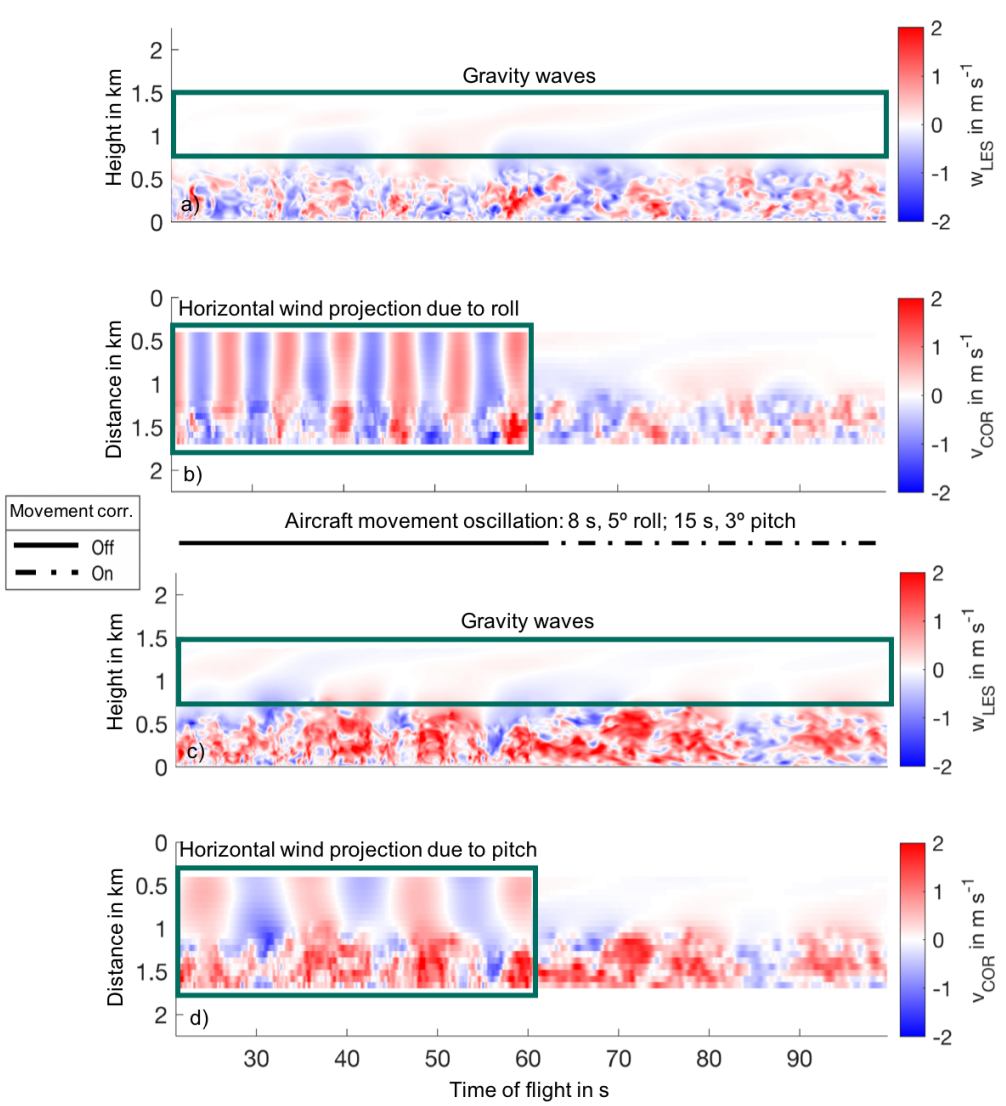

Figure 3. ADLS simulated nadir transects for the 10 m s^{-1} background wind case. The scanner movement correction is disabled for the first half of the transects, whereas it is enabled for the second half. a) LES vertical wind along a crosswind transect. b) ADLS motion corrected velocity measurement along a crosswind transect. c) LES vertical wind along an upwind transect. d) ADLS motion corrected velocity measurement along an upwind transect.

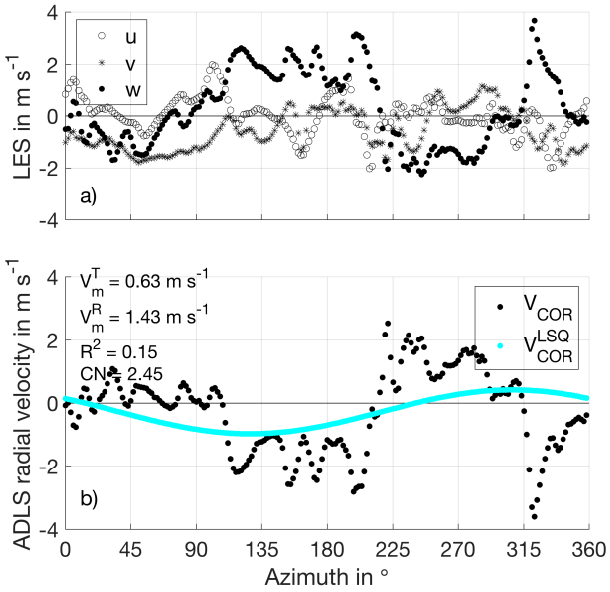

Figure 4. Illustration of the wind profile retrieval procedure example for the 0 m s^{-1} background wind case and 10 Hz lidar measurement frequency for illustrative purpose. a) Input u,v,w values from LES showing the model truth used to produce the simulated radial velocities. The homogeneity assumption used for AVAD retrieval is clearly violated as the u,v,w components show deviations throughout the scan. b) Simulated, motion corrected radial velocities as well as LSQ-fit obtained from SVD inversion used for retrieval. The retrieval results in a positively biased retrieved wind speed as the vertical wind is erroneously mapped into horizontal wind. Please note: As movement towards the lidar is negative by convention, an updraft in ENU (positive vertical velocity) results in a negative radial velocity, as the lidar is looking down and air is moving towards the lidar.

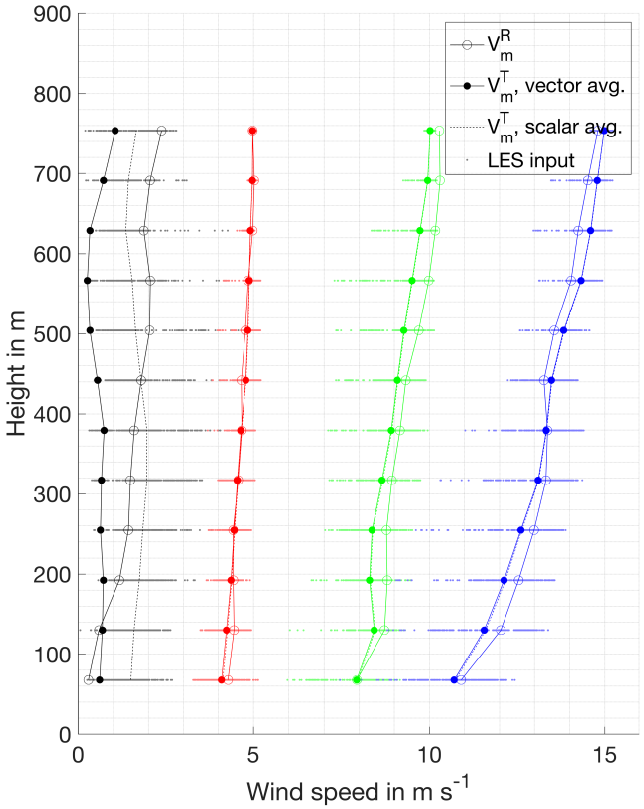

Figure 5. Examples of retrieved wind profiles and associated LES input truth for all four background wind cases. Retrieval errors are due to the homogeneity assumption used for retrieval being violated. For the 0 m s^{-1} background wind case the effect of scalar vs. vector averaging is detectable in the difference between the individual LES wind speed input values and the vector average value (discussed in Sec. 3.3). The retrieval for the 10 m s^{-1} background wind case shows that the retrieved value can lie outside the range of input values. Retrieval errors are correlated over multiple altitudes (resembling a bias, but being random for different profiles) as turbulence is correlated vertically.

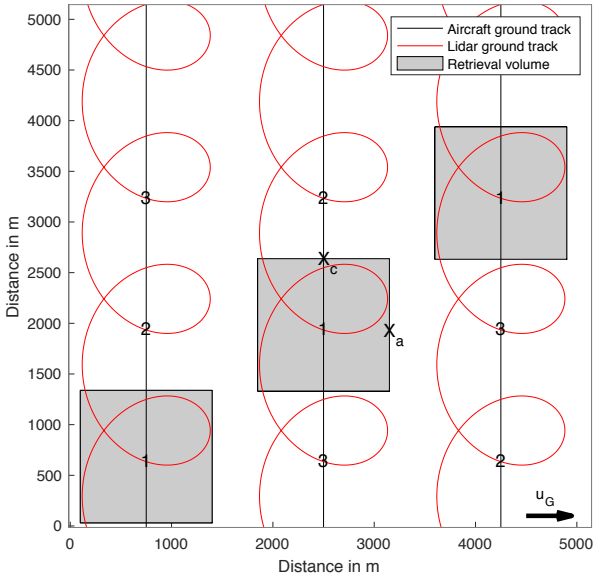

Figure 6. Illustration of checkerboard approach used for wind profile retrieval. Shown are the locations of the three transects and associated retrieval volumes for the first time step. Retrieval volumes are shifted to the enumerated positions for subsequent time steps, while flight trajectories are repeated. Additionally shown is the lidar ground track.

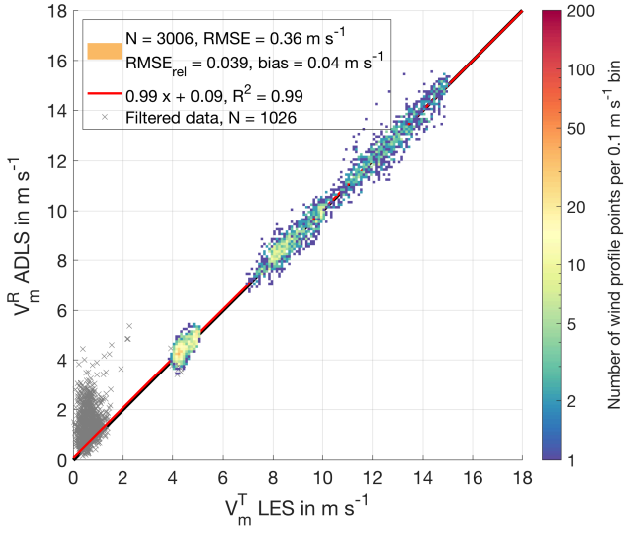

Figure 7. Comparison of LES truth and ADLS retrieved wind speed for an ideal measurement system with quality filtering criteria applied. Color-coded are all measurements which pass quality filtering, gray crosses are the ones which do not. The observable deviations from the 1:1 line reveal retrieval errors due to the violation of the homogeneity assumption used in the AVAD.

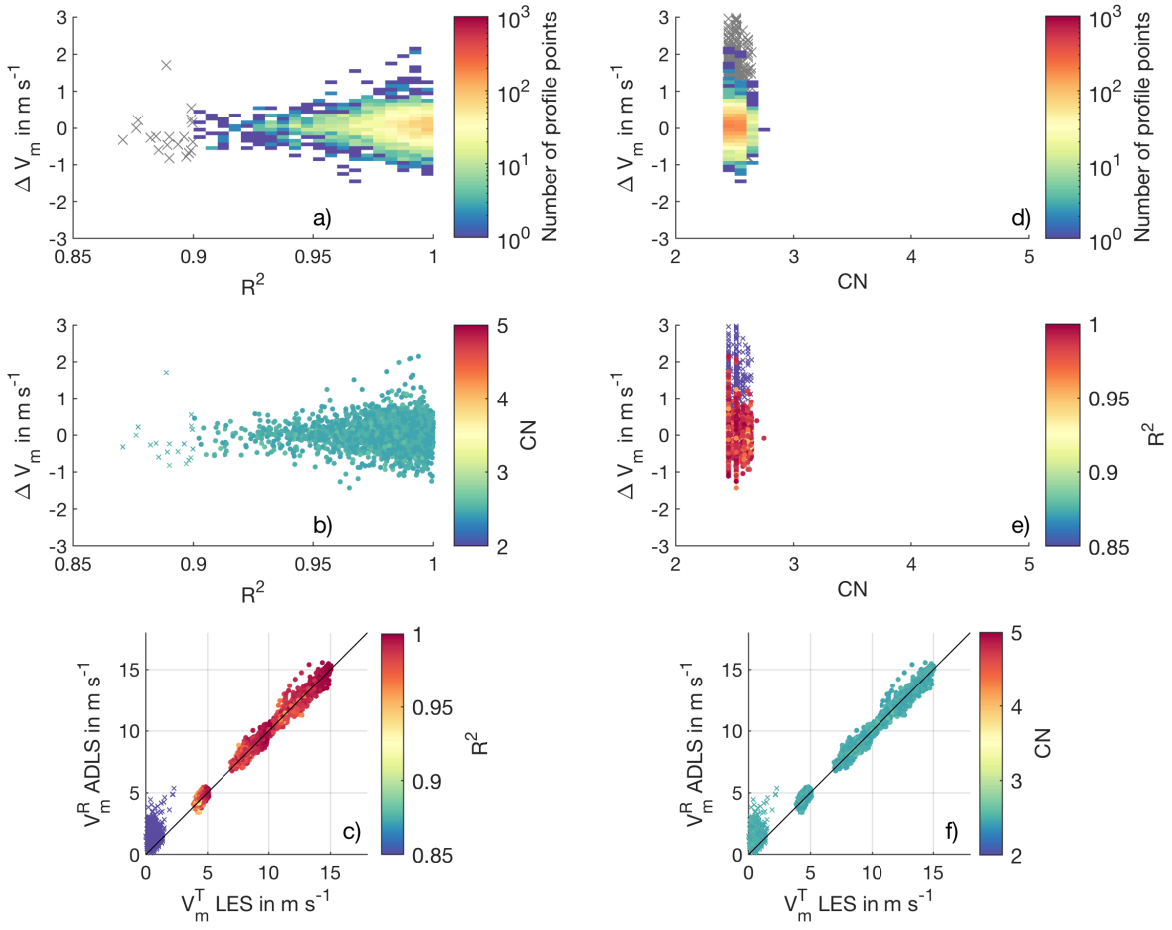

Figure 8. Quality filtering criteria for the standard system setup and retrieval strategy. a) Color-coded histogram of number of occurrence for retrieval error and R^2. Gray crosses show values eliminated by quality filtering. b) Retrieval error and R^2, color-coded is CN. c) LES wind speed and ADLS wind speed, color coded is R^2. Profile points which pass quality filtering are displayed as color-coded circles, profile points which are eliminated as color-coded crosses. d) Same as a), but for CN. e) Retrieval error and CN, color coded is R^2. f) Same as c), but for CN.

Appendix A: ~~Geometric calculations for~~ Mathematical formulation of concepts implemented in the ADLS

A1 Triangle of velocities

Given for this example are the in-air speed IAS, the aircraft ground track direction TR, the wind speed WS and the wind direction WD. Directions are to be given in degree from north and speeds in the same units. Needed are the aircraft heading

HDG and the ground speed GS. We can calculate them according to the following formulas:

$$HDG = TR + \arcsin\left(\frac{WS \cdot \sin(TR - WD)}{IAS}\right), \tag{A1}$$

$$GS = IAS \cdot \cos(HDG - TR) + WS \cdot \cos(TR - WD). \tag{A2}$$

The calculations can be performed for other combinations of given and needed variables as well.

A2 Coordinate transforms

As outlined in Lenschow (1972) ~~, Lee et al. (1994) and Leon and Vali (1998)~~and Leon and Vali (1998), transformations between the LES earthbound (E) coordinate system, oriented east-north-up (ENU), and the aircraft (A) coordinate system, oriented along aircraft-right wing-down (ARD), are achieved by using the standard heading-pitch-roll procedure using the transformation matrix $\mathbb{T}$. The coordinate transform matrix from the aircraft ARD reference frame to ground ENU reference frame is T^{AE} (A to E). Here, the transformation is ~~exemplary~~ conducted for the beam direction vector $\boldsymbol{b} = [b_x, b_y, b_z]$ as an example. It can be transferred between the two systems with

$$\boldsymbol{b}^E = \mathbb{T}^{AE}\boldsymbol{b}^A \tag{A3}$$

and

$$\boldsymbol{b}^A = \mathbb{T}'^{AE}\boldsymbol{b}^E. \tag{A4}$$

The transformation matrix can be split into separate rotations around the individual aircraft axes.

$$\mathbb{T}^{AE} = \mathbb{HPR}. \tag{A5}$$

Hereby, $\mathbb{H}$ denotes a heading rotation at angle ψ (yaw around the z-axis), $\mathbb{P}$ a pitch rotation at angle θ (pitch around the y-axis) and $\mathbb{R}$ a roll rotation at angle ϕ (roll around the x-axis). Individually, they are given as:

$$\mathbb{H} = \begin{pmatrix} \sin(\psi) & \cos(\psi) & 0 \\ \cos(\psi) & -\sin(\psi) & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{A6}$$

$$\mathbb{P} = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}, \tag{A7}$$

$$\mathbb{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi). \end{pmatrix}. \tag{A8}$$

Combined, this results in

$$\mathbb{T}^{AE} = \begin{pmatrix} \sin(\psi)\cos(\theta) & \cos(\psi)\cos(\phi) + \sin(\psi)\sin(\theta)\sin(\phi) & -\cos(\psi)\sin(\phi) + \sin(\psi)\sin(\theta)\cos(\phi) \\ \cos(\psi)\cos(\theta) & -\sin(\psi)\cos(\phi) + \cos(\psi)\sin(\theta)\sin(\phi) & \sin(\psi)\sin(\phi) + \cos(\psi)\sin(\theta)\cos(\phi) \\ \sin(\theta) & -\cos(\theta)\sin(\phi) & -\cos(\theta)\cos(\phi) \end{pmatrix}. \tag{A9}$$

A3 Lidar beam position, lidar beam motion and averaging of the LES wind field

In order to determine the LES wind velocity v_p^E which is projected onto the beam, the range gate position has to be calculated. As the position is needed for the LES, it must be calculated in the ground reference system. Therefore, after transferring the beam direction into the ground reference system (App. A2), the range gate center ~~distance x_{RGC}~~ distances R_0 (spaced $\Delta R_0 = \Delta p = 72$ m apart) in beam orientation to the aircraft position,

$$\boldsymbol{p}_{RGC R_0}{}^{E} = \boldsymbol{p}_{AC}^{E} + x_{RGC} R_0 \boldsymbol{b}^{E}. \tag{A10}$$

This calculation is repeated for the range gate begin and end positions by subtracting or adding half the range gate length to ~~x_{RGC}: R_0:~~

$$x_{RGB} R_{0B} = x_{RGC} R_0 - l_{RG} \Delta p/2, \tag{A11}$$

$$x_{RGE} R_{0E} = x_{RGC} R_0 + l_{RG} \Delta p/2. \tag{A12}$$

Thereby, the range gate position at the average measurement time is fully characterized, however, the range gate motion due to aircraft motion and scanner movement still need to be accounted for.

The motion of the range gate during one measurement is accounted for by defining a volume between the range gate position at the beginning and the end of the measurement process. The two positions are calculated using a range gate motion vector. The range gate motion vector is constructed by using the difference of the range gate center position compared to the end range gate position of the previous measurement (this assumption is valid as only continuous aircraft and scanner movements are investigated). Using these positions, difference vectors are constructed for the range gate begin, center and end positions. The first and last range gate positions during one measurement are then obtained by subtracting and adding the range gate motion vector from the range gate center position.

The real lidar beam only has a beam diameter of approx. 10 cm. This diameter is not enough to ensure an adequate sampling of LES data (grid spacing is ~~$\Delta x = 5$ m~~ $\Delta x = 10$ m). Therefore, the beam volume is artificially enlarged in the direction orthogonal to the motion. The factor is set in relation to the grid spacing, a minimum distance of half the grid point distance is employed to ensure points inside the volume,

$$d_{\text{inflate}} = \Delta x/2. \tag{A13}$$

After all points that fall inside the volume covered by the lidar beam are determined, they are weighted according to the range gate weighting function (Sec. 2.2) based on their orthogonal distance from the beam center. Last, a linear averaging is

applied to obtain the average velocity of all points in the volume. This averaged velocity v_p^E is then projected onto the beam direction according to Eq. 3.

A4 ~~Overview of terminology and error concepts~~Wind profile retrieval theory

~~Overview of used acronyms and their meaning. ADLS Airborne Doppler lidar Simulator AVAD Airborne Velocity Azimuth Display ADL Airborne Doppler lidar CN Condition Number LES Large-Eddy Simulation LSQ Least-square R^2 Coefficient of determination RET Retrieval strategy standard RMSE Root-mean-square error REL Relative root-mean-square error STP System setup standard VAD Velocity-azimuth display VVP Volume-velocity processing Name Description Standard Options AAL Aircraft flight altitude 1700 m 1700, 1450, 1200 m AVG Along-track averaging distance 1300 m 650, 1300, 2600 m BLA Sector blanking 0° 0°, 90°, 180°, 270° DIR Flight direction 16 directions 16 directions ELE Scan elevation angle 60° elev. 30° – 80° elev. FME Lidar measurement frequency 1 Hz 1, 5, 10 HzLEV Vertical averaging distance 62 m 62, 124, 248 mROT Scan rotation duration 20 s rot. 4 – 20 s rot. SEP Profile separation methodVolume Volume, time SYS Measurement system type Ideal, noisy Ideal, noisy TIME LES time steps 25 time steps 25 time steps VAR Retrieval variant u,v,w u,v,w; u,v WND Geostrophic background wind case 0, 5, 10, 15 m s^{-1} 0, 5, 10, 15 m s^{-1} σv_D Radial velocity noise 0, 0.25 m s^{-1} 0 – 1 m s^{-1} σb Beam pointing noise 0, 0.25° 0° – 1°~~

~~Definition of general terminology as used in this study. Term General definition ADLS implementation Measurement Obtaining a radial velocity estimate from the~~ In the AVAD method, multiple radial velocity measurements v_{COR_n} under different beam pointing directions $\boldsymbol{b}_n$ are sampled from an atmospheric wind field ~~. Averaging and weighting LES-wind speeds in lidar beam volume to obtain radial velocity estimate. Motion Change in position. Calculated using a fixed IAS and~~ with mean wind vector $\boldsymbol{v}_p$:

$$\begin{bmatrix} v_{COR_1} \\ v_{COR_2} \\ v_{COR_3} \\ \vdots \\ v_{COR_n} \end{bmatrix} = \begin{bmatrix} b_{x_1} & b_{y_1} & b_{z_1} \\ b_{x_2} & b_{y_1} & b_{z_1} \\ b_{x_3} & b_{y_1} & b_{z_1} \\ \vdots & \vdots & \vdots \\ b_{x_n} & b_{y_n} & b_{z_n} \end{bmatrix} \begin{bmatrix} v_{p_x} \\ v_{p_y} \\ v_{p_z} \end{bmatrix}. \tag{A14}$$

The beam directions from multiple measurements make up the beam pointing matrix $\mathbb{G}$. Consequently, the relation can be expressed in the following way:

$$v_{COR} = \mathbb{G}\boldsymbol{v}_p. \tag{A15}$$

Knowing v_{COR} and $\mathbb{G}$, the inverse problem is then solved by calculating the inverse $\mathbb{G}^{-g}$ of the ~~triangle of velocities for the aircraft, the range gate motion vector for the lidar beam. Motion correction Obtaining a corrected radial velocity from the measured radial velocity by accounting for aircraft motion and aircraft movement moment arm. Removes the exact aircraft contribution from the measured radial velocity, no error is introduced. (Wind profile) Retrieval Inferring an estimate of the~~

~~underlying wind field from a set of radial velocity observations based on an underlying wind field assumption. The beam matrix inversion process utilizing~~ beam pointing matrix $\mathbb{G}$, in order to obtain an estimate of the wind vector $\boldsymbol{v}_{ret}$ responsible for the observations:

$$\boldsymbol{v}_{ret} = \mathbb{G}^{-g} v_{COR}. \tag{A16}$$

The general inverse $\mathbb{G}^{-g}$ of the overdetermined least-squares problem can be calculated as (Menke, 2012):

$$\mathbb{G}^{-g} = [\mathbb{G}^T \mathbb{G}]^{-1} \mathbb{G}^T. \tag{A17}$$

Equation (A17) presents a least-squares solution to the problem. Instead of calculating the general inverse in the above way, in this study a singular value decomposition ~~. Range gate motion vector Change in range gate position from one measurement to the next due to aircraft motion as well as aircraft and scanner movement Calculated as the difference between range gate begin and end position Movement Change in orientation. Pitch, roll, yaw for the aircraft; elevation, azimuth for the scanner. Emulated by artificially prescribed values, not connected to aircraft motion in the ADLS. Ideal measurement system No measurement error in the radial velocities and model geometry is perfectly known. No artificial noise is added to the obtained radial velocities and no error introduced into the beam pointing direction. Noisy measurement system Random fluctuations present in the~~ (SVD) is performed, which yields benefits compared to the above direct solution (Boccippio, 1995):

$$\mathbb{G} = \mathbb{U}\mathbb{S}\mathbb{W}^T, \tag{A18}$$

$$\mathbb{G}^{-g} = \mathbb{W}_p \mathbb{S}_p^{-1} \mathbb{U}_p^T. \tag{A19}$$

SVD results in an orthogonal decomposition, where $\mathbb{U}^T\mathbb{U} = \mathbb{W}^T\mathbb{W} = \mathbb{I}$, with $\mathbb{I}$ being the identity matrix. Further, $\mathbb{S}$ consists of the singular values of $\mathbb{G}$ as its diagonal entries. The condition number is defined as the ratio of the largest to the smallest singular value λ_S:

$$CN = \frac{max(\lambda_S)}{min(\lambda_S)}. \tag{A20}$$

Another quality control measure which can be obtained from the LSQ-solution is the coefficient of determination (R^2). Using the estimated wind vector, an average radial velocity (LSQ-fit) is constructed by projecting it into radial velocities using the beam pointing geometry:

$$v_{COR}^{ret} = \mathbb{G}\boldsymbol{v}^{ret}. \tag{A21}$$

Using the average radial velocity~~estimates (due to lidar instrument noise, turbulence in the range gate volume etc.) and random fluctuations present in the beam pointing direction (due to imperfect aircraft navigation system, scanner pointing etc.). The noise is emulated by adding random numbers from a Gaussian distribution with specified standard deviation to the measured radial velocities and measured beam azimuth and elevation.~~, the R^2 is defined as

$$R^2 = 1 - \frac{\sum_n (v_{COR_n} - v_{COR_n}^{ret})^2}{\sum_n (v_{COR_n} - \sum_n v_{COR_n})^2}. \tag{A22}$$

~~Definition of error terminology as used in this study. Term General definition ADLS implementation Retrieval error The difference between the atmospheric input wind vector in the sampling volume and the retrieved wind vector after the measurement and inversion process. Directly obtained by subtracting the (vector) average of the wind field components used to create the radial velocitymeasurements from the wind vector retrieved using the measurements through inversion. Total random error, Baker et al. (1995); Frehlich (2001) All random differences present between multiple horizontal wind profile retrievals. Divided into random error and representation (or sampling) error. Consist only of random profile error as representation error can be avoided and no other measurement system is simulated. Random (profile) error, Baker et al. (1995); Frehlich (2001) Random fluctuations in the wind profile retrieval due to instrument noise (especially at low-signal quality), turbulence in the measurement volume and retrieval algorithm stability. This error contribution is~~

A5 Error metrics

The RMSE is given as:

$$\mathrm{RMSE} = \sqrt{\left[\frac{\Sigma_i^N (V_{m\,i}^T - V_{m\,i}^R)^2}{N}\right]}. \tag{A23}$$

Here, V_m^T is the true wind speed based on the input LES wind speeds, whereas V_m^R is the ADLS retrieved wind speed from wind profiling. The retrieval error is $\Delta V_m = V_m^R - V_m^T$. N is the number of wind profile points fulfilling the quality filtering criteria. The relative root-mean-squared error (REL) is used in accordance with Guimond et al. (2014):

$$\mathrm{REL} = \sqrt{\left[\frac{\Sigma_i^N (V_{m\,i}^T - V_{m\,i}^R)^2}{\Sigma_i^N (V_{m\,i}^T)^2}\right]}. \tag{A24}$$

The REL can provide additional information to the ~~focus of this study. Adjacent profile points for an individual profile appear smooth despite random error being present as the turbulence between them is similar. Representation or sampling error , Baker et al. (1995); Frehlich (2001) Observed differences due to sampling of different parts of the atmosphere and/or only partial coverage of the atmosphere by different measurement systems. E.g. between Doppler lidar wind profile retrieval and radiosonde wind profile. Can be avoided in the ADLS as the atmospheric input data is known exactly and the same volumes are used, therefore not present in this study. Random radial velocity fluctuation, Frehlich et al. (1994) Random fluctuations in the radial velocity measurement due to instrument noise (especially at low-signal quality), turbulence in the range gate volume and radial velocity estimation stability (spectral peak finding). Not modeled directly, but emulated by adding random Gaussian distributed values to the measured radial velocities for the noisy system simulation. Beam pointing inaccuracy Random fluctuations in the beam pointing orientation due to aircraft and scanner vibration or insufficient knowledge of mirror orientation. Not modeled directly, but emulated by adding random Gaussian distributed values to the scanner azimuth and elevation angles.~~ RMSE as its magnitude is independent of the mean wind speed, thereby enabling comparisons between the different background wind cases, especially for higher wind speeds. The RMSE is driven by two factors, the variance and bias

of the retrieval errors. Therefore we also report the bias of the retrieval as its average mean deviation,

$$\text{bias} = \frac{1}{N}\sum_{i=1}^{N}(V_{m\,i}^{R} - V_{m\,i}^{T}). \tag{A25}$$

Appendix B: Additional figures

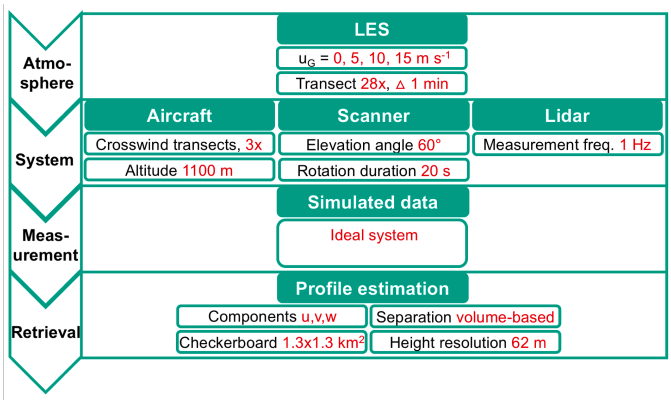

Figure B1. Simulator operation scheme and settings for the wind profiling quality analysis. Standard values used for wind profile retrieval are marked in red.

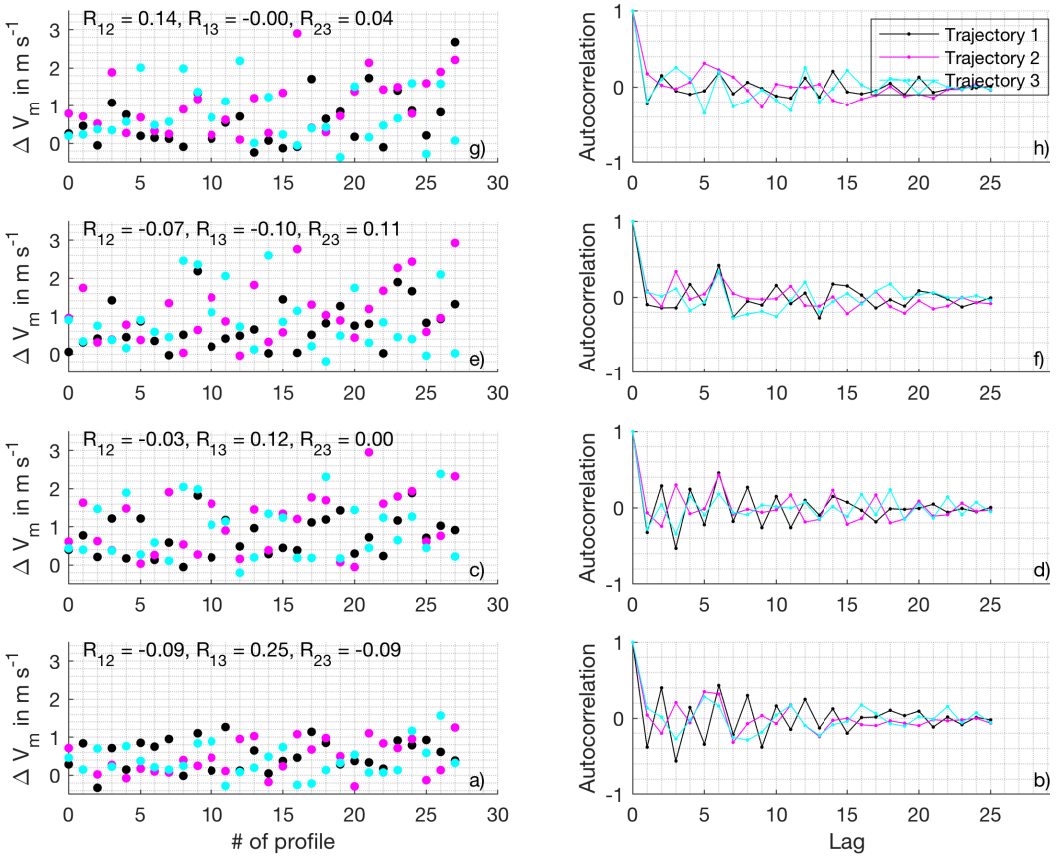

Figure B2. Error correlation analysis for the 0 m s^{-1} background wind case. Panels a), c), e), g) show the retrieval errors for every time step at 60, 320, 500, 760 m altitude respectively. Color-coded is the number of the parallel trajectory. Additionally given as numbers is the correlation between neighboring trajectories. Panels b), d), f), h) show the autocorrelation of the retrieval errors for each transect.

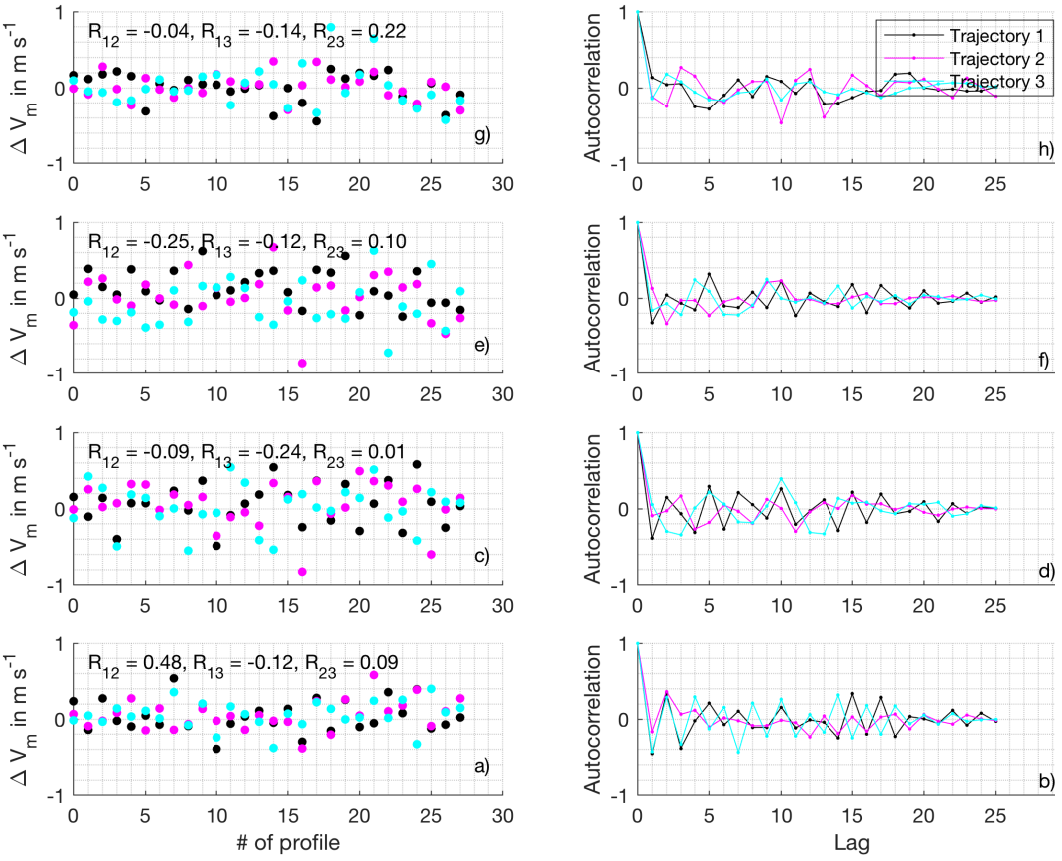

Figure B3. ~~Illustration of aircraft trajectory and scanner position simulation~~ Same as fig. B2 but for ~~two transects through~~ the ~~LES domain. The LES~~ 5 m s^{-1} background wind ~~speed is color-coded~~case. ~~The black vectors represent aircraft trajectories~~

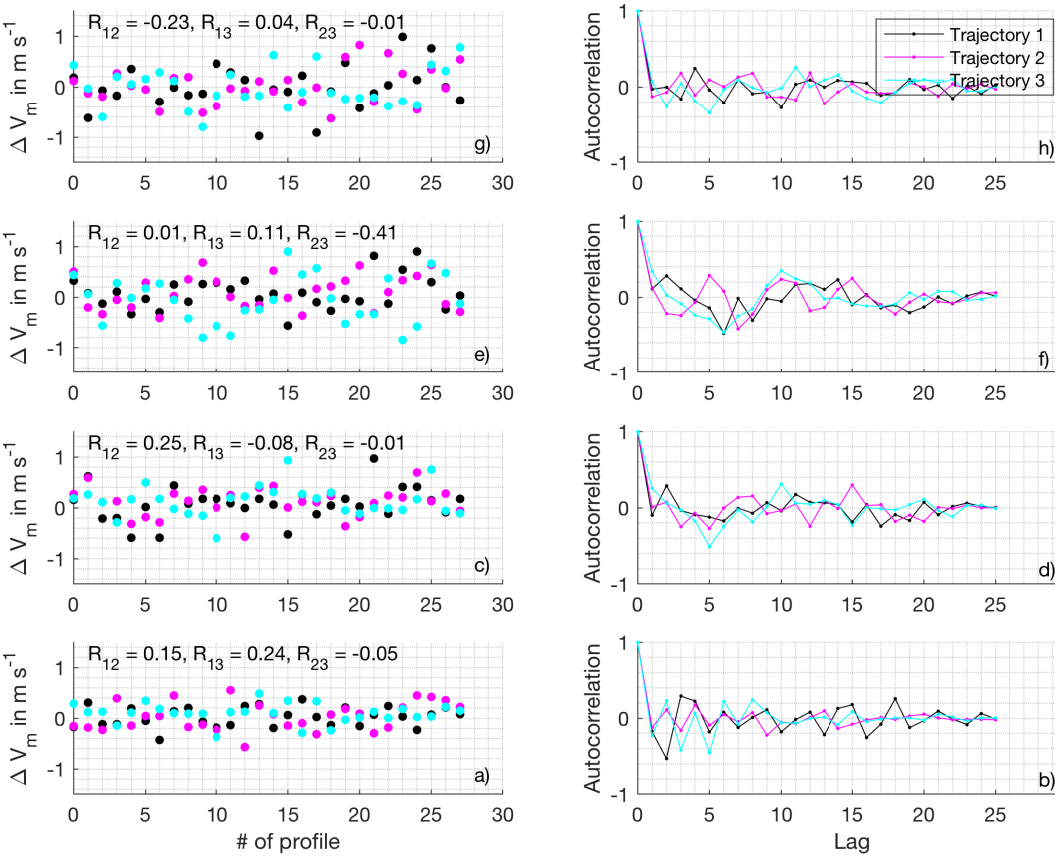

Figure B4. Same as fig. ~~The black curtains show the range gate positions used to conduct the measurement which are calculated from lidar and scanner setting. For the first half of each transect the scanner movement correction is disabled, whereas~~ B2 but for the ~~second half it is enabled. The nadir transect is used to retrieve the vertical~~ 10 m s^{-1} background wind ~~, whereas the AVAD pattern is used to retrieve the horizontal wind~~case.

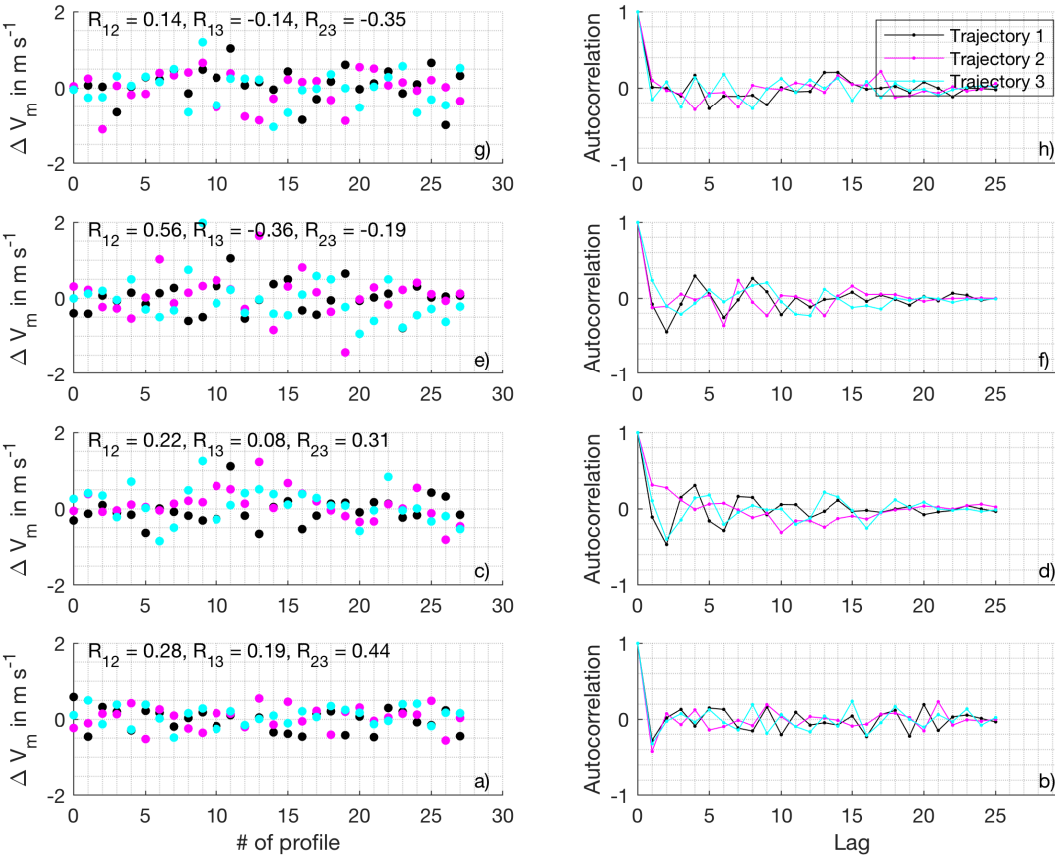

Figure B5. Same as fig. B2 but for the 15 m s^{-1} background wind case.

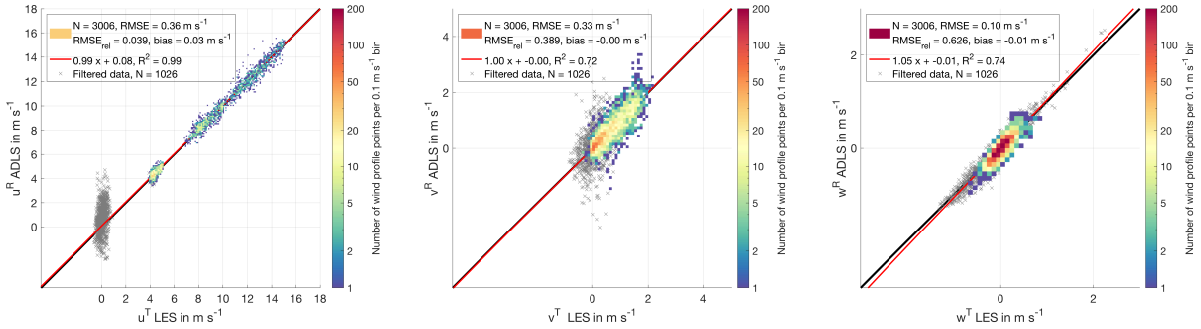

Figure B6. LES truth and ADLS retrieved wind ~~direction~~ speed components for an ideal measurement system with ~~STP, RET~~quality filtering criteria applied. Color-coded are all measurements which pass quality filtering, grayed out crosses are the ones which do not.~~Values which pass quality filtering cluster around 270° wind direction. The 0 m s^{-1} background wind case is filtered completely by applying quality filtering criteria explained in Sec. 3.3.~~

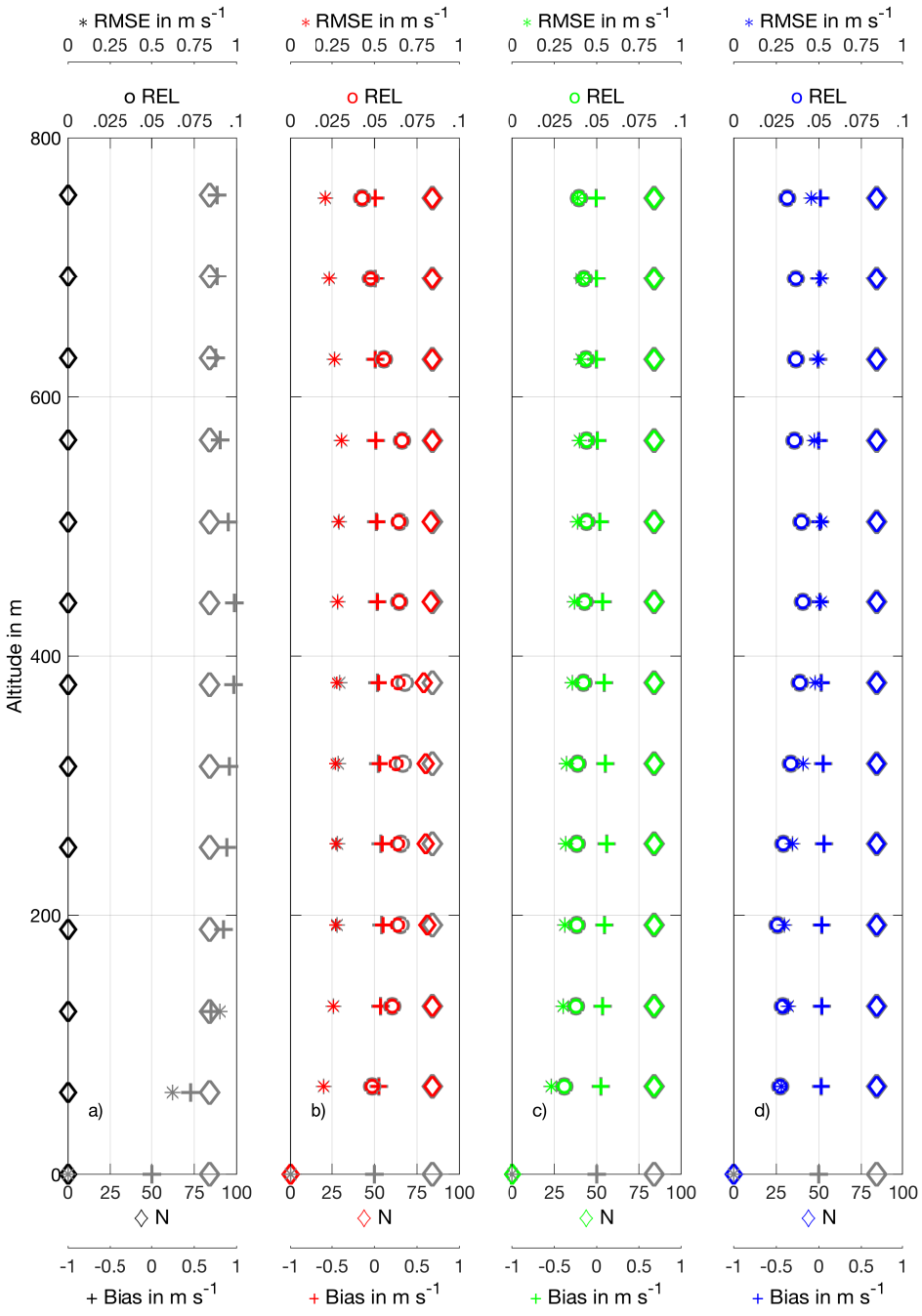

Figure B7. ~~LES truth~~ Profile of wind speed retrieval quality parameters RMSE, REL, bias and ~~ADLS~~ number of retrieved wind ~~speed components~~ profile points as a function of height for an ideal ~~measurement~~ system ~~with STP, RET. Color-coded are all measurements which pass quality filtering, grayed out crosses are~~ using the ~~ones which do not~~standard system setup and retrieval strategy. ~~Quality filtering criteria explained in Sec~~a) 0 m s^{-1} background wind case. ~~3.3 are applied~~b) 5 m s^{-1} background wind case. c) 10 m s^{-1} background wind case. d) 15 m s^{-1} background wind case.

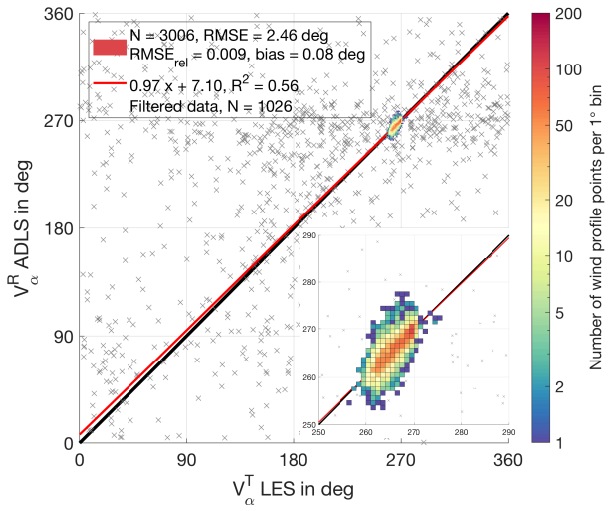

Figure B8. ~~Histogram of~~ LES truth and ADLS retrieved wind ~~speed~~ direction for an ideal measurement system with ~~AVG 650 m and STP, RET otherwise~~quality filtering criteria applied. Color-coded are all measurements which pass quality filtering, grayed out crosses are the ones which do not. ~~Quality~~ Values which pass quality filtering cluster around 270° wind direction, which is magnified in the inset. The 0 m s^{-1} background wind case is filtered completely by applying quality filtering criteria explained in Sec. 3.3~~are applied~~.

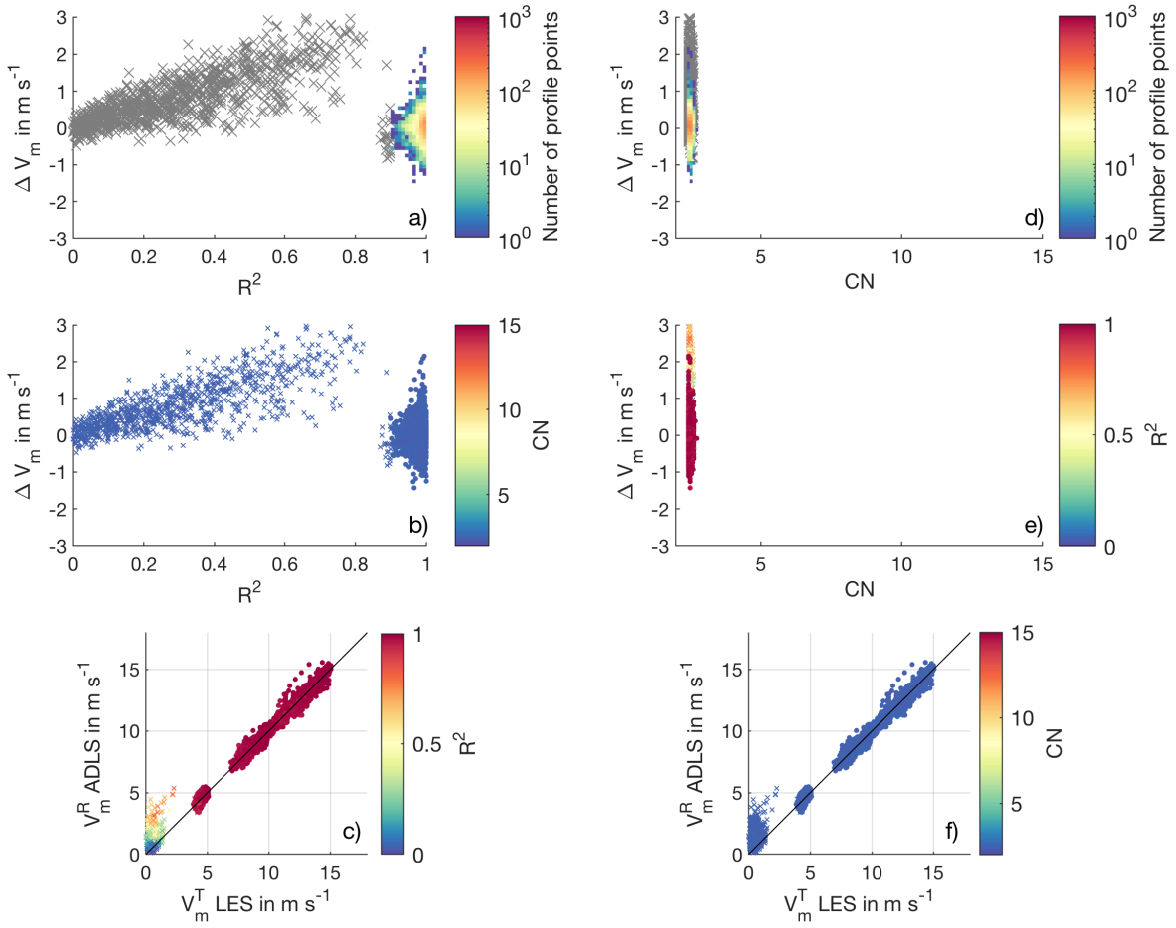

Figure B9. ~~RMSE, REL, N and bias~~ Quality filtering criteria for ~~four background wind cases at varying random radial velocity fluctuations~~ the standard system setup and ~~beam pointing inaccuracies without application~~ retrieval strategy. a) Color-coded histogram of number of occurrence for retrieval error and R^2. Gray crosses show values eliminated by quality filtering. ~~a),~~ b) Retrieval error and R^2, color-coded is CN. c) ~~, d) 0 m s^{-1} background~~ LES wind ~~case. e), f), g), h) 5 m s^{-1} background~~ speed and ADLS wind ~~case. i)~~speed, ~~j), k), l) 10 m s^{-1} background wind case~~color coded is R^2. ~~m), n), o), p) 15 m s^{-1} background wind case.~~Profile ~~of wind speed retrieval parameters~~ points which pass quality filtering are displayed as ~~a function of height for an ideal system using STP~~color-coded circles, ~~RET~~profile points which are eliminated as color-coded crosses. d) Same as a)~~0 m s^{-1} background wind case~~, but for CN. ~~b~~e) ~~5 m s^{-1} background wind case~~Retrieval error and CN, color coded is R^2. f) Same as c)~~10 m s^{-1} background wind case~~, but for CN.~~d) 15 m s^{-1} background wind case.~~

References

Baidar, S., Tucker, S. C., Beaubien, M., and Hardesty, R. M.: The Optical Autocovariance Wind Lidar. Part II: Green OAWL (GrOAWL) Airborne Performance and Validation, J. Atmos. Ocean. Technol., 35, 2099–2116, https://doi.org/10.1175/jtech-d-18-0025.1, 2018.

Baker, W. E., Emmitt, G. D., Robertson, F., Atlas, R. M., Molinari, J. E., Bowdle, D. A., Paegle, J., Hardesty, R. M., Post, M. J., Menzies, R. T., Krishnamurti, T. N., Brown, R. A., Anderson, J. R., Lorenc, A. C., and McElroy, J.: Lidar-Measured Winds from Space: A Key Component for Weather and Climate Prediction, Bull. Amer. Meteor. Soc., 76, 869–888, 1995.

Baker, W. E., Atlas, R., Cardinali, C., Clement, A., Emmitt, G. D., Gentry, B. M., Hardesty, R. M., Källén, E., Kavaya, M. J., Langland, R., Ma, Z., Masutani, M., McCarty, W., Pierce, R. B., Pu, Z., Riishojgaard, L. P., Ryan, J., Tucker, S., Weissmann, M., and Yoe, J. G.: Lidar-measured wind profiles: The missing link in the global observing system, Bull. Am. Meteorol. Soc., 95, 543–564, https://doi.org/10.1175/BAMS-D-12-00164.1, 2014.

Banakh, V. A. and Werner, C.: Computer simulation of coherent Doppler lidar measurement of wind velocity and retrieval of turbulent wind statistics, Opt. Eng., 44, 71 205, https://doi.org/10.1117/1.1955167, 2005.

Banakh, V. a., Smalikho, I. N., Köpp, F., and Werner, C.: Representativeness of wind measurements with a cw Doppler lidar in the atmospheric boundary layer., Appl. Opt., 34, 2055–2067, https://doi.org/10.1364/AO.34.002055, 1995.

Bingöl, F., Mann, J., and Foussekis, D.: Conically scanning lidar error in complex terrain, Meteorol. Zeitschrift, 18, 189–195, https://doi.org/10.1127/0941-2948/2009/0368, 2009.

Boccippio, D. J.: A diagnostic analysis of the VVP single-doppler retrieval technique, J. Atmos. Ocean. Technol., 12, 230–248, https://doi.org/10.1175/1520-0426(1995)012<0230:ADAOTV>2.0.CO;2, 1995.

Bradley, S., Perrott, Y., Behrens, P., and Oldroyd, A.: Corrections for Wind-Speed Errors from Sodar and Lidar in Complex Terrain, Boundary-Layer Meteorol., 143, 37–48, https://doi.org/10.1007/s10546-012-9702-0, 2012.

Bucci, L. R., O'Handley, C., Emmitt, G. D., Zhang, J. A., Ryan, K., and Atlas, R.: Validation of an Airborne Doppler Wind Lidar in Tropical Cyclones, Sensors, 18, 4288, https://doi.org/10.3390/s18124288, 2018.

Cheong, B. L., Yu, T. Y., Palmer, R. D., Yang, K. F., Hoffman, M. W., Frasier, S. J., and Lopez-Dekker, F. J.: Effects of wind field inhomogeneities on Doppler beam swinging revealed by an imaging radar, J. Atmos. Ocean. Technol., 25, 1414–1422, https://doi.org/10.1175/2007JTECHA969.1, 2008.

Chouza, F., Reitebuch, O., Groß, S., Rahm, S., Freudenthaler, V., Toledano, C., and Weinzierl, B.: Retrieval of aerosol backscatter and extinction from airborne coherent Doppler wind lidar measurements, Atmos. Meas. Tech., 8, 2909–2926, https://doi.org/10.5194/amt-8-2909-2015, 2015.

Chouza, F., Reitebuch, O., Benedetti, A., and Weinzierl, B.: Saharan dust long-range transport across the Atlantic studied by an airborne Doppler wind lidar and the MACC model, Atmos. Chem. Phys., 16, 11 581–11 600, https://doi.org/10.5194/acp-16-11581-2016, 2016a.

Chouza, F., Reitebuch, O., Jähn, M., Rahm, S., and Weinzierl, B.: Vertical wind retrieved by airborne lidar and analysis of island induced gravity waves in combination with numerical models and in situ particle measurements, Atmos. Chem. Phys., 16, 4675–4692, https://doi.org/10.5194/acp-16-4675-2016, 2016b.

Corsmeier, U., Hankers, R., and Wieser, A.: Airborne turbulence measurements in the lower troposphere onboard the research aircraft Dornier 128-6, D-IBUF, Meteorol. Zeitschrift, 10, 315–329, https://doi.org/10.1127/0941-2948/2001/0010-0315, 2001.

Davis, C. A., Markowski, P. M., Moore, J. A., Barth, M. C., Grubišić, V., Klein, P. M., Raymond, D. J., Detwiler, A., Geerts, B., Mullendore, G. L., and Lee, W.-C.: Recommendations for In Situ and Remote Sensing Capabilities in Atmospheric Convection and Turbulence, Bull. Am. Meteorol. Soc., 99, 2463–2470, https://doi.org/10.1175/bams-d-17-0310.1, 2018.

De Wekker, S. F. J., Godwin, K. S., Emmitt, G. D., and Greco, S.: Airborne Doppler lidar measurements of valley flows in complex coastal terrain, J. Appl. Meteorol. Clim., 51, 1558–1574, https://doi.org/10.1175/JAMC-D-10-05034.1, 2012.

ESA: ESA's Aeolus wind satellite launched, https://www.esa.int/For_Media/Press_Releases/ESA_s_Aeolus_wind_satellite_launched, 2018.

Frehlich, R.: Effects of wind turbulence on coherent Doppler lidar performance, J. Atmos. Ocean. Technol., 14, 54–75, https://doi.org/10.1175/1520-0426(1997)014<0054:EOWTOC>2.0.CO;2, 1997.

Frehlich, R.: Errors for space-based doppler lidar wind measurements: Definition, performance, and verification, J. Atmos. Ocean. Technol., 18, 1749–1772, https://doi.org/10.1175/1520-0426(2001)018<1749:EFSBDL>2.0.CO;2, 2001.

Frehlich, R., Hannon, S. M., and Henderson, S. W.: Performance of a 2-μm coherent Doppler lidar for wind measurements, J. Atmos. Ocean. Technol., 11, 1517–1528, https://doi.org/10.1175/1520-0426(1994)011<1517:POACDL>2.0.CO;2, 1994.

Gamache, J. F., Marks, F. D., and Roux, F.: Comparison of three airborne Doppler sampling techniques with airborne in situ wind observations in Hurricane Gustav (1990), J. Atmos. Ocean. Technol., 12, 171–181, https://doi.org/10.1175/1520-0426(1995)012<0171:COTADS>2.0.CO;2, 1995.

Godwin, K. S., De Wekker, S. F. J., and Emmitt, G. D.: Retrieving Winds in the Surface Layer over Land Using an Airborne Doppler Lidar, J. Atmos. Ocean. Technol., 29, 487–499, https://doi.org/10.1175/JTECH-D-11-00139.1, 2012.

Guimond, S. R., Tian, L., Heymsfield, G. M., and Frasier, S. J.: Wind retrieval algorithms for the IWRAP and HIWRAP airborne doppler radars with applications to hurricanes, J. Atmos. Ocean. Technol., 31, 1189–1215, https://doi.org/10.1175/JTECH-D-13-00140.1, 2014.

Haimov, S. and Rodi, A.: Fixed-antenna pointing-angle calibration of airborne doppler cloud radar, J. Atmos. Ocean. Technol., 30, 2320–2335, https://doi.org/10.1175/JTECH-D-12-00262.1, 2013.

Holleman, I.: Quality control and verification of weather radar wind profiles, J. Atmos. Ocean. Technol., 22, 1541–1550, https://doi.org/10.1175/JTECH1781.1, 2005.

Kavaya, M. J., Beyon, J. Y., Koch, G. J., Petros, M., Petzar, P. J., Singh, U. N., Trieu, B. C., and Yu, J.: The doppler aerosol wind (DAWN) airborne, wind-profiling coherent-detection lidar system: Overview and preliminary flight results, J. Atmos. Ocean. Technol., 31, 826–842, https://doi.org/10.1175/JTECH-D-12-00274.1, 2014.

Kiemle, C., Wirth, M., Fix, A., Rahm, S., Corsmeier, U., and Di Girolamo, P.: Latent heat flux measurements over complex terrain by airborne water vapour and wind lidars, Q. J. R. Meteorol. Soc., 137, 190–203, https://doi.org/10.1002/qj.757, 2011.

Klaas, T., Pauscher, L., and Callies, D.: LiDAR-mast deviations in complex terrain and their simulation using CFD, Meteorol. Zeitschrift, 24, 591–603, https://doi.org/10.1127/metz/2015/0637, 2015.

Koch, G. J., Beyon, J. Y., Cowen, L. J., Kavaya, M. J., and Grant, M. S.: Three-dimensional wind profiling of offshore wind energy areas with airborne Doppler lidar, J. Appl. Remote Sens., 8, 1–11, https://doi.org/10.1117/1.JRS.8.083662, 2014.

Koscielny, A.: An Evaluation of the Accuracy of Some Radar Wind Profiling Techniques, J. Atmos. Ocean. Technol., 1, 309–320, 1984.

Lee, W.-C., Dodge, P., Marks, F. D., and Hildebrand, P. H.: Mapping of Airborne Doppler Radar Data, J. Atmos. Ocean. Technol., 11, 572–578, https://doi.org/10.1175/1520-0426(1994)011<0572:MOADRD>2.0.CO;2, 1994.

Lenschow: The Measurement of Air Velocity and Temperature Using the NCAR Buffalo Aircraft Measuring System, Tech. Rep. June, NCAR, Boulder, 1972.

Leon, D. and Vali, G.: Retrieval of three-dimensional particle velocity from airborne doppler radar data, J. Atmos. Ocean. Technol., 15, 860–870, https://doi.org/10.1175/1520-0426(1998)015<0860:ROTDPV>2.0.CO;2, 1998.

Lorsolo, S., Gamache, J., and Aksoy, A.: Evaluation of the hurricane research division doppler radar analysis software using synthetic data, J. Atmos. Ocean. Technol., 30, 1055–1071, https://doi.org/10.1175/JTECH-D-12-00161.1, 2013.

Lundquist, J. K., Churchfield, M. J., Lee, S., and Clifton, A.: Quantifying error of lidar and sodar doppler beam swinging measurements of wind turbine wakes using computational fluid dynamics, Atmos. Meas. Tech., 8, 907–920, https://doi.org/10.5194/amt-8-907-2015, 2015.

Lux, O., Lemmerz, C., Weiler, F., Marksteiner, U., Witschas, B., Rahm, S., Schäfler, A., and Reitebuch, O.: Airborne wind lidar observations over the North Atlantic in 2016 for the pre-launch validation of the satellite mission Aeolus, Atmos. Meas. Tech., 1, 3297–3322, https://doi.org/https://doi.org/10.5194/amt-11-3297-2018, 2018.

Menke, W.: Describing Inverse Problems, Elsevier/Academic Press, Oxford UK, https://doi.org/10.1016/B978-0-12-397160-9.00001-1, 2012.

Muschinski, A., Sullivan, P. P., Wuertz, D. B., Hill, R. J., Cohn, S. A., Lenschow, D. H., and Doviak, R. J.: First synthesis of wind-profiler signals on the basis of large-eddy simulation data, Radio Sci., 34, 1437–1459, https://doi.org/10.1029/1999RS900090, 1999.

Paffrath, U., Lemmerz, C., Reitebuch, O., Witschas, B., Nikolaus, I., and Freudenthaler, V.: The airborne demonstrator for the direct-detection doppler wind lidar aladin on adm-aeolus. Part II: Simulations and rayleigh receiver radiometric performance, J. Atmos. Ocean. Technol., 26, 2516–2530, https://doi.org/10.1175/2009JTECHA1314.1, 2009.

Päschke, E., Leinweber, R., and Lehmann, V.: An assessment of the performance of a 1.5 μm Doppler lidar for operational vertical wind profiling based on a 1-year trial, Atmos. Meas. Tech., 8, 2251–2266, https://doi.org/10.5194/amt-8-2251-2015, 2015.

Pauscher, L., Vasiljevic, N., Callies, D., Lea, G., Mann, J., Klaas, T., Hieronimus, J., Gottschall, J., Schwesig, A., Kühn, M., and Courtney, M.: An inter-comparison study of multi- and DBS lidar measurements in complex terrain, Remote Sens., 8, https://doi.org/10.3390/rs8090782, 2016.

Raasch, S. and Schröter, M.: PALM - A large-eddy simulation model performing on massively parallel computers, Meteorol. Zeitschrift, 10, 363–372, https://doi.org/10.1127/0941-2948/2001/0010-0363, 2001.

Scipion, D.: Characterization of the convective boundary layer through a combination of large-eddy simulations and a radar simulator, Phd, University of Oklahoma, 2011.

Scipión, D., Palmer, R., Chilson, P., Fedorovich, E., and Botnick, A.: Retrieval of convective boundary layer wind field statistics from radar profiler measurements in conjunction with large eddy simulation, Meteorol. Zeitschrift, 18, 175–187, https://doi.org/10.1127/0941-2948/2009/0371, 2009.

Shenghui, Z., Ming, W., Lijun, W., Chang, Z., and Mingxu, Z.: Sensitivity analysis of the VVP wind retrieval method for single-doppler weather radars, J. Atmos. Ocean. Technol., 31, 1289–1300, https://doi.org/10.1175/JTECH-D-13-00190.1, 2014.

Stawiarski, C.: Optimizing Dual-Doppler Lidar Measurements of Surface Layer Coherent Structures with Large-Eddy Simulations, KIT Scientific Publishing, Karlsruhe, 2014.

Stawiarski, C., Traumner, K., Knigge, C., and Calhoun, R.: Scopes and challenges of dual-doppler lidar wind measurements-an error analysis, J. Atmos. Ocean. Technol., 30, 2044–2062, https://doi.org/10.1175/JTECH-D-12-00244.1, 2013.

Stawiarski, C., Träumner, K., Kottmeier, C., Knigge, C., and Raasch, S.: Assessment of Surface-Layer Coherent Structure Detection in Dual-Doppler Lidar Data Based on Virtual Measurements, Boundary-Layer Meteorol., 156, 371–393, https://doi.org/10.1007/s10546-015-0039-3, 2015.

Tian, L., Heymsfield, G. M., Didlake, A. C., Guimond, S., and Li, L.: Velocity-Azimuth display analysis of doppler velocity for HIWRAP, J. Appl. Meteorol. Clim., 54, 1792–1808, https://doi.org/10.1175/JAMC-D-14-0054.1, 2015.

Tucker, S. C., Weimer, C. S., Baidar, S., and Hardesty, R. M.: The optical autocovariance wind lidar. Part I: OAWL instrument development and demonstration, J. Atmos. Ocean. Technol., 35, 2079–2097, https://doi.org/10.1175/JTECH-D-18-0024.1, 2018.

Wainwright, C. E., Stepanian, P. M., Chilson, P. B., Palmer, R. D., Fedorovich, E., and Gibbs, J. A.: A time series sodar simulator based on large-eddy simulation, J. Atmos. Ocean. Technol., 31, 876–889, https://doi.org/10.1175/JTECH-D-13-00161.1, 2014.

Waldteufel, P. and Corbin, H.: On the Analysis of Single-Doppler Radar Data, J. Appl. Meteorol., 18, 532–542, https://doi.org/10.1175/1520-0450(1979)018<0532:OTAOSD>2.0.CO;2, 1978.

Wang, H., Barthelmie, R. J., Clifton, A., and Pryor, S. C.: Wind measurements from arc scans with Doppler wind lidar, J. Atmos. Ocean. Technol., 32, 2024–2040, https://doi.org/10.1175/JTECH-D-14-00059.1, 2015.

Wang, H., Barthelmie, R. J., Pryor, S. C., and Brown, G.: Lidar arc scan uncertainty reduction through scanning geometry optimization, Atmos. Meas. Tech., 9, 1653–1669, https://doi.org/10.5194/amt-9-1653-2016, 2016.

Weissmann, M., Busen, R., Dörnbrack, A., Rahm, S., and Reitebuch, O.: Targeted observations with an airborne wind lidar, J. Atmos. Ocean. Technol., 22, 1706–1719, https://doi.org/10.1175/JTECH1801.1, 2005.

Weitkamp, C., Asakura, T., Ha, T. W., Kamiya, T., Krausz, F., Monemar, B., Venghaus, H., Weber, H., and Weinfurter, H.: Lidar - Range-Resolved Optical Remote Sensing of the Atmosphere, Springer, New York, NY, 2005.

Witschas, B., Rahm, S., Dörnbrack, A., Wagner, J., and Rapp, M.: Airborne wind lidar measurements of vertical and horizontal winds for the investigation of orographically induced gravity waves, J. Atmos. Ocean. Technol., 34, 1371–1386, https://doi.org/10.1175/JTECH-D-17-0021.1, 2017.

Zhang, J., Atlas, R., Emmitt, G., Bucci, L., and Ryan, K.: Airborne Doppler Wind Lidar Observations of the Tropical Cyclone Boundary Layer, Remote Sens., 10, 825, https://doi.org/10.3390/rs10060825, 2018.